

AIS 2022

**17th International Symposium on
Applied Informatics and Related Areas**

PROCEEDINGS

Óbuda University

PROCEEDINGS of
17th International Symposium on
Applied Informatics and Related Areas

Editors:

Ildikó Petőné Csuka

Gyula Simon

Óbuda University
November 17, 2022
Székesfehérvár, Hungary
ISBN 978-963-449-302-0

CONTENTS

Oral Session I. - Applied Informatics I.....	1
László Sajó-Bohus and Monika Pogátsnik: Solar flare effect on microelectronic circuits assessed for Wigner's disease.....	2
Malgorzata Veróné Wojtaszek, Erick Noboa and György Eigner: Drought monitoring mapping by using satellite remote sensing: Study Case.....	8
Éva Hajnal, Gergely Vakulya and Péter Udvardy: Long Term Rumen Temperature Analysis of Dairy Cattle with Help of Bolus Sensor.....	12
Gyula Simon, Gergely Vakulya, Sándor Tarsoly and Péter Galambos: Semi-Automatic Detection and Tracking of Growing Mushrooms on Image Sequences.....	17
Oral Session II. - Applied Informatics II.	22
Erika Peukert, Anikó Mahler and József Halász: Computerized latency data correspond with Iowa Gambling Task responses in healthy individuals.....	23
Tamás Jancsó, Andrea Pődör, Éva Nagyné Hajnal and Rezső Nagy: Image Classification Methods to Detect the Growth of Industrial Areas.....	28
Péter Udvardy: LEGO robotics at service of education.....	33
János Katona and Andrea Pődör: Comparative Analysis of Land Value Capture Tools in Europe.....	38
Poster Session I. – Industrial and Embedded Systems.....	43
Zoltán Szilágyi and Csaba Hajdu: Design and Requirement Analysis of an Outdoor Autonomous Mobile Manipulator.....	44
Bertalan Beszédes: Artificial experimental environment for indoor plant cultivation.....	50
Bertalan Beszédes: Design of a Reliable Portable Power Supply Chain for an Embedded System.....	55
György Györök, Margit Makó and Borbála Györök, Olga Shvets: Capacitive sensor for gaseous substances, from old to new.....	60
Aslima Alimhanova, Svetlana Grigoryeva and György Györök: VLC technology concept in life support parameter management systems.....	64
Bolatova Anar and Svetlana Grigoryeva: Indoor LED lighting control based on embedded systems....	68
Bertalan Beszédes and Bence Márta: Development and Prototype Production of a Full-spectrum LED Lighting Driver Circuit.....	72
Bertalan Beszédes, Bence Márta and Péter Udvardy: Artificial Lighting Experimental Environment in Agriculture for Seed Germination.....	77
Poster Session II. - Computer Systems.....	82
Rozália Lakner: Simulation and Analysis of Route Plans for Autonomous Guided Vehicle Systems.....	83
Gergely Vakulya: Genetic algorithm based image approximation with geometrical shapes.....	89

Albina Kadyroldina, Assel Kussaiyn-Murat, Alexander Krasavin and Darya Alontseva: A new laser stripe detection algorithm for 3D scanning applications.....	93
Olga Shvets, Bauyrzhan Smakanov and György Györök: Stabilization of environmental conditions to improve the performance of a mobile application for the state of the driver monitoring.....	98
Olga Shvets, Assel Naizabayeva and Márta Seebauer: Autonomous life support using intelligent analysis in Kazakhstan.....	102
Daniyar Kaliyev, Olga Shvets and György Györök: Robust Control Design for Quadcopter for Path Following in Monitoring Tasks.....	106
József Borbély: Solving Olympiad Problems Via Lagrange Multipliers.....	110
Vanessza Tőke and Rozália Lakner: Boardgame Recommendation System with Collaborative Filtering.....	115
Erik Dobra and Éva Hajnal: Virtual Reality Game Development with External Hand Motion Sensor..	121
Poster Session III. – Geodesy and Geoinformatics.....	126
Gergely László, Gábor Péter Molnár and Gabriella Bor: Lever arm calculating algorithm for efficient and precise vector determinations.....	127
Ágnes Krausz and Gergely László: Comparison of lever arm determination methods in terms of their effect of the accuracy of aerial triangulation.....	131
Regina Hanna Balaton and Andrea Pődör: Design of Cemetery Cadastre of Saint Vince Home of the Sisters of Charity in Csákvár.....	139
Gibárt Gilányi and Gábor Molnár: Conversion of Laplace equation from the spherical coordinate description to Cartesian coordinate system.....	143
Gábor Nagy and Zsuzsanna Ungvári: Comparative Analysis of several contour line generation methods and softwares.....	147
Poster Session IV. - Education, Enterprise and Economics.....	151
Ildikó Petőné Csuka: Entrepreneurial activity and ecosystem in Hungary.....	152
Tímea Bekk and Monika Pogátsnik: Motivation potential in higher education after covid-19 based on personality traits.....	157
Csaba Bráda and Monika Pogátsnik: The changes in the concept and content of the mechanical engineer training from the 3rd industrial revolution to industry 4.0.....	163
Nikoletta Tolner, Judit Módné Takács, András Dávid and Monika Pogatsnik: Online examination from the point of view of teachers.....	169
Zsuzsanna Felícia Pretz and Éva Takács: Promotion of dual training form and its' tools at a large company.....	174

ORAL SESSION I.
Applied Informatics I.

Solar flare effect on microelectronic circuits assessed for Wigner's disease

Monika Pogatsnik
Alba Regia Technical Faculty
Obuda University
Szekesfehervar, Hungary
pogatsnik.monika@amk.uni-obuda.hu

László Sajó-Bohus
Alba Regia Technical Faculty
Obuda University
Szekesfehervar, Hungary,
Universidad Simón Bolívar,
Valle de Sartenejas Caracas, Venezuela
lsajo@usb.ve

Abstract— Heavy energetic charged (HZE) or neutral particles induce atomic and molecular changes in materials. Depending on the absorbed radiation dose, degradation and failure of the electronic components occur in electrical systems installed e.g., for space vehicle control or satellite communication. Galactic cosmic ray (GCR) increases abruptly during solar flares due to the emission of solar energetic particles (SEPs). These are event-related waves of energy ejected from the sun's surface. High energy photons and protons, alphas, neutrons, electrons, and muons among others are released in a time interval sequence of short bursts lasting a few minutes. Ionizing particles lose their energy in matter originating high charge density in a thick layer of material. Electron-hole pairs are produced in semiconductors. Most transistors are made from very pure silicon and have one or two kinds of charge carriers (e.g., field-effect transistors, and bipolar junction transistors). Existing a high-density ion, molecular structures are disrupted by Coulomb repulsion with the consequence of changing the transistor parameters; a phenomenon that alters the functionality of microelectronics circuitry ending in data corruption. We report consequences in microelectronic circuitry of the effect called Wigner disease. Atomic alteration due to solar flare neutron scattering induces dislocations per atom (dpa). For this purpose, the mathematical formalism as an approximate model of Kinchin and Pease is applied. The study is a contribution toward the *Artemis* NASA space program related to the Red Planet exploration.

Keywords— solar flare, radiation damage, Wigner disease, transistors, electronic circuitry.

I SOLAR PARTICLES AND GALACTIC COSMIC RAYS

The intergalactic space as well as the solar system is submerged in a medium where we observe plasma, photons micro/macro meteorites, and endless particles of a wide range of mass and energy. Particles can be emitted by the sun as a continuous process or by solar flares. These produced sporadically release matter with fluencies of approximately 3×10^{10} p.cm². Solar particles are superposed to the background radiation of galactic origin (galactic cosmic rays). During space extra-terrestrial exploration the most frequently detected charged particles are primary protons; their energy is around 1 MeV, with an almost constant flux in the energy range between 1 and 200 MeV (~ 3 p. cm⁻². s⁻¹) [1]. The next group to be reckoned with are neutrons due to their disruptive effect on atomic and molecular structure; charge-less particles penetrate the materials producing primary knock-on atom (PKA) i.e., displacements caused by recoil. That occurs by an indirect mechanism following a given interaction in a matter such as elastic/inelastic scattering, capture process (n, γ);

nuclear reactions e.g. (n, p), (n, α), (n, 2n), (n, f). Fast neutron scattering and nuclear reactions are dominating processes when their energy spectrum is above 1MeV. On the other hand, for slower energy regions, the neutron capture mechanism has a higher probability to take place [2]. Given the importance of neutron interaction with matter, a global neutron monitor network has been established [3].

High-energy protons or other high-energy charged particles (HEZ) from solar origin penetrate e.g., Si-wafers producing high-density ionization with characteristic Linear Energy Transfer (LET). The consequence is a Coulomb explosion taking place followed by large atomic arrays disruption. During the mechanism, low Z-nuclei such as, O, Al, and Si often are relocated to a different place from their original position, an effect that alters the matter property often referred to as radiation-induced damage [4].

That is a phenomenon of great importance for materials rich in silicon and silicon oxide-based electronic components conceived for interplanetary exploration. Specifically, we refer to the NASA (USA) Artemis ongoing mission. That is a program having the goal to establish a human settlement on the *Red Planet* Mars. The program starts setting up a sort of base camp, on the Moon, at the lunar south pole region by 2024. The plan is to facilitate manned space missions and at the same time will provide new experiences that could be convenient for the next step i.e., the colonization of Mars.

II RADIATION DOSE DURING AN INTERPLANETARY MISSION

The Artemis project is a continuation of the program named Apollo (Brother of Artemis in Greek mythology); to achieve this goal, several private space-related companies and international partners such as European Space Agency which includes the Hungarian Space Research Office, actively contribute to the NASA (USA) initiative.

Crewed spaceflight programs carried out in the past put in evidence that astronauts are exposed to a relatively high space radiation dose to the extent to be considered Worker's Occupationally Exposed to Ionizing radiation. During long-term spaceflight, doses in the range of 50-2,000 mSv have been reported. Worth to be noted that above 100 mSv radiation sickness and biological anomaly (cancer) are frequently observed.

The space radiation dose is considered so far one of the barriers against manned planetary missions. Before some practical solution can be applied to shield radiation, a robot

system will be employed for the *Artemis* mission. The three-stage plan starts with phase 1 going back to the moon, with a command center heavily relying on massive electro-mechanical, communication, and computing systems.

During space flight and the permanence on the planet, electronic equipment must operate in a rather harsh radiation environment and cumulative adverse effects are expected. Reliability will be reduced by radiation damage of SiC or GaN power electronics and generally space electronics. Their active lifetime will shorten and under particular conditions for instance during solar flare higher electron-hole pairs will be formed [5] and [6]. The space radiation generated damage, in vital electronic components relying on Si-lattice, silicon dioxide, and other elements employed in insulators and casing, is one major obstacle with other, for future human safety, [7] and [8].

III RADIATION EFFECTS AND NUCLEAR DISPLACEMENT DAMAGE

The most vulnerable components are based on semiconductor elements. The estimated radiation dose for electronic devices and insulation used in magnet coil is given by Liu et al [9] and matter simulating space suit as well as human elemental body compound [10].

The main issue is the total energy absorbed by matter that deteriorates electronic systems' performance. Since a mission time interval may be of several years, often it is convenient to carry out a test on electronics intended for space use by supplying high dose employing accelerators. The advantage of the method in a short time high dose rate values can be achieved for electronic components in the range of values between $0.5\text{-}3.0\text{ Gy}\cdot\text{s}^{-1}$. As an example, the technique could be applied to e.g., an electronic device based on SiO_2 with 6mm of Al radiation shield to obtain values corresponding to an average space mission (around $1\text{MGy}\cdot\text{a}^{-1}$).

The dose rate may increase considerably during solar events such as i.- proton rich Coronal Mass Ejections and ii.- heavy ion rich Impulsive Solar Flares. In both cases particles or heavy nuclei penetrating a transistor structure induce ions along their path; for illustration, we give a schematic representation of the phenomenon for a semiconductor substrate in fig.1.

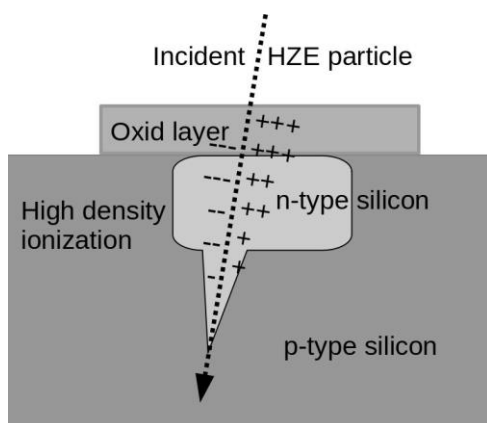


Fig.1 HEZ-particle track, indicated by the arrow, is where high-density Coulomb ionization is originating in a semiconductor layer.

Solar activities are related to a sporadic phenomenon where neutral particles originate by nuclear reaction with solar

corona elements namely: protons, 3He , 4He , C, N, O, Ne, Mg, Si, and Fe. The first on the list, are precursors of solar flare neutrons are those that increase the probability to induce permanent failure in space electronics.

The targeted matter suffers both elemental and molecular modifications, that depend strongly on the neutron spectral energy. The latter is characterized by three energy regions namely thermal, fast, and between them an interval corresponding to resonances. This thought of a relatively small energy interval is the most damaging.

Thermal neutrons absorbed by an electronic device under power, induce often the so-called single event effect or SEE; a phenomenon produced mainly through two nuclear processes namely elastic and inelastic reactions.

In the latter cases, the n-type substrate suffers severe disruption in its atomic structure; reaction fragments release a large amount of energy in a short range ($1\text{-}10^5\text{ keV}\cdot\mu\text{m}^{-1}$) as illustrated in fig. 2.

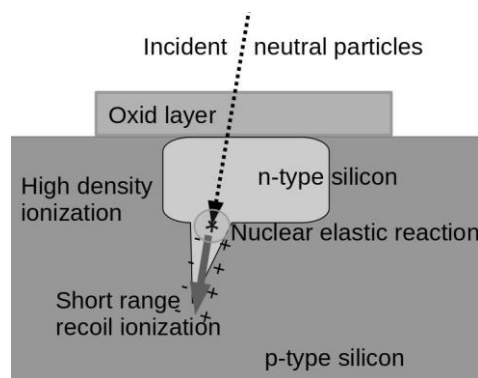
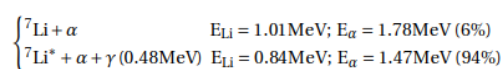
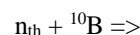


Fig.2. Schematic drawing showing atomic displacement by incoming neutrons; high LET particles are induced by a nuclear reaction (involving neutrons) or elastic scattering occurring in a transistor layer.

Another example is related to the element Boron (with isotopic composition of 80% ^{11}B +20% ^{10}B) often employed in the insulating layer. Thermal neutrons may induce a ^{10}B binary reaction having a high transversal cross-section (3838[b]). The nuclear neutron capture mechanism follows two processes one of which is a gamma-less reaction:



In these reactions, ${}^7\text{Li}$, ${}^4\text{He}$ fragments leave high ionization tracks (LET in silicon up to $1\text{keV}\cdot\mu\text{m}^{-1}$) moving in the opposite direction so that one of them probably alters the electronic device performance e.g., through SEE.

The other process that has a high impact is neutron scattering on Si atoms, SiO_2 molecule, or carbon nanotubes for space applications as given by Bohus et al. These are found in transistors, integrated circuits (IC), and generally in space electric devices.

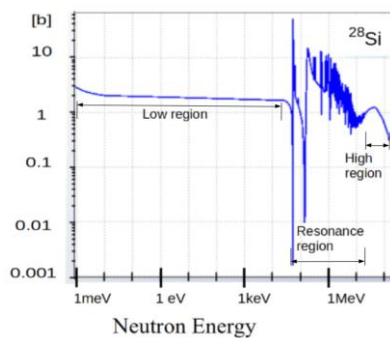


Fig.3. Scattering cross section for ^{28}Si selected for three neutron energy regions (data taken from: <https://www.nndc.bnl.gov/sigma>).

Neutrons depending on their translational energy, interact with different probability with ^{28}Si (92.23%), ^{29}Si (4.67%), and ^{30}Si (3.10%) isotopes. The elastic transversal cross section is given in Fig. 3, to show the large variability in the function of the energy.

The most significant aspect is related to the resonance region where scattering cross section may change almost five orders of magnitude.

The displacement per atom (DPA) due to neutrons at the mentioned energy region may change in the same order of magnitude. In the case of ^{28}Si , the number of displacements per neutron collision could be high near a sensitive region to induce a single event upset (SEU). The accumulated effect related to neutron scattering correlates to the absorbed dose rate. A faulty circuit performance can be observed for a dose rate of $0.1 \text{ [mGy}\cdot\text{s}^{-1}]$; Vibishna et al. [11], reported a type of SEU i.e., single-event effects impact on Avionic system and radiation-induced alteration on Si-rich electronic components e.g., FPGA in that a single HZE particle induces not at all or highly destructive and nondestructive damage.

Silicon technology today forms the basis of a worldwide, component industry. The reason for this expansion can be found not only in the physical properties of silicon but also in the unique properties of the silicon-silicon dioxide interface. However, silicon devices show a kind of instabilities that is related to crystal lattice structure (voids and clusters). In the case of exposure to an intense neutron or protons field the mentioned instabilities increases due to the occurring atomic displacement and the formation of new recombination centers. These through trapping of holes provide current leakage, loss of amplifier gains, and in deteriorate image quality camera sensors, (e.g., CCD and CMOS devices among others).

The major impact of data corruption, noise on images, and system shutdowns are on the reference stars' localization. Even a limited failure could be catastrophic from the crew safety point of view.

Depending on the Coulomb field intensities originating inside any electronic component the following cases are possible: Single Event Upset, Single Event Latchup, Single Event Gate rupture, and Single Event Out; the most disruptive of the mentioned, are the last two. These induce faulty responses leading to hardware or transitory failure even to the so-called "no fault found". The latter could be of gravity since under normal control that may go undetected.

IV NEUTRON DAMAGE BY WIGNER DISEASE

The neutron energy range during space missions is a relevant issue due to the radiation risk involved; the range is quite large since covers an interval of values between 0.1 and 200 MeV; relevant information on this subject is given in ref Palfalvi et al [12]. Any solid interacting neutron radiation can display atomic structure alteration known otherwise as the Wigner effect or Wigner's disease.

Silicon crystal and silicon oxide ionization energy are respectively 3.6 and 18.0 eV/ion pair. The corresponding number of ions produced by a charged particle along the penetration length of a few microns is between $2,8 \times 10^4$ and 11.1×10^6 ion pair. As mentioned, a high-intensity Coulomb field originates followed by rapid recombination that often results in a detriment of matter.

Watanabe et al. [24], have extensively studied neutron production (Solar Neutron Event) in connection to the Solar sudden outbreak. Lockwood [22], outlined space neutron existence and Lario [13] computed values for the total fluence of solar neutrons at energies: 1, 10, 100 and 1000 MeV and primary neutron fluence respectively 0.87; 0.72; 0.12 and 0.0013×10^6 neutrons cm^{-2} ; The measured values are considerably higher due to secondary neutron-producing reactions; we mention the most important: (p; n), (p; pn), and (p; 2pn). As the solar flare protons escape, they have a non-zero probability to interact with nuclei such as ^{13}C , ^{15}N , ^{18}O , ^{22}Ne , $^{25,26}\text{Mg}$, ^{29}Si , and ^{54}Fe .

Si atoms and SiO_2 structure under neutron exposure can be displaced from their original position as described by Eugene Wigner [14]; here we refer to it as Wigner's Disease (WD). The phenomenon originated from a neutral particle carrying kinetic energy above the threshold value of 25eV following an elastic scattering with Si-nuclei.

In a crystal structure, atoms are positioned a tetrahedral lattice from which they are displaced by the nuclear scattering mechanism. That could be a singular event or a cascade displacement in a silicon wafer depending on e.g. Solar flare neutron's energy.

The importance of the WD effect in semiconductors, microchip components, or highly packed transistors is highlighted through primary knock-on atoms PKA. If energy is available, then higher-order knock-ons; this process originates atom displacement (DPA) giving a measure on the radiation damage e.g., in silicon crystals by non-ionized field, [15]. The energy required to carry out the process values are given for some elements by [16] and [17]

To determine the DPA value an approximation given by Kinchin and Pease referred to as the K-P model is most conveniently applied [18]. Some assumptions make the basis of the model namely: pure elastic collisions between neutron and nuclei; there is a threshold energy E_d below that no atomic displacement is possible; annihilation process or radioactive decay does not intervene during lattice displacement. A schematic drawing of a DPA is given in Fig.4.

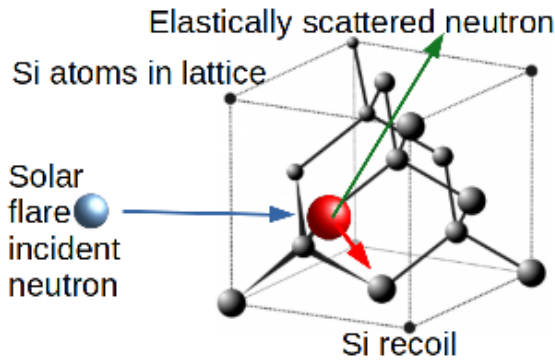


Fig.4. An artistic view of a neutron impinging onto a Si atom lattice. The small arrow indicates the atom recoil to occupy a new permanent site. The other two arrows show the impinging and the elastically scattered particle directions.

The damage rate equation R_d provides a value to quantify radiation damage in the unit time or DPA rate, by the following equation [17]:

$$R_d = N \int_l \varphi(E_n) dE_n \int_E^{\Lambda E_n} \sigma_d(E_n, E) \nu(E) dE$$

Where:

N = number of lattice atom density exposed to neutrons

$\sigma_d(E)$ = microscopic scattering cross section [b]

$\phi(E_n)$ = particle fluence [n/cc-s]

$\nu(T)$ is the number of displaced atoms resulting from neutron scattering given by Kinchin and Pease, (or K-P model) as $E_n/2E_d$

E_n = incident neutron energy [MeV].

E_d = threshold Energy [ev], i.e., the threshold displacement energy that is the minimum kinetic energy necessary to displace an atom in a solid from its lattice site to the defect position.

To remove an atom or induce a displacement (DPA) in a lattice, a given amount of energy E_d is required to break electronic bonds. At the target surface, its value is lower than the Lattice Binding Energy since less atom surrounds each other; consequently, it is expected to have a larger number of DPA. Electronic components made of thinner layers (lower number of targeted atoms) would experience less damage.

From the theory of elastic scattering [17], we find that

$$\sigma_n(E_n, E) = \frac{\sigma_n^{el}(E_n)}{\Lambda E_n}$$

and as mentioned the elastic collision cross-section $\sigma_n^{el}(E_n)$ is neutron energy dependent, for light elements is almost constant for incident neutron energies up to 1 MeV, above that value, several resonance peaks dominate then again, the spectrum shows a constant value. Detailed information is given with specific values at international data libraries available at nuclear data centers [19].

The atomic mass parameter Λ is determined by the expression: $\Lambda=4A/(1+A)^2$. Since the isotopic effect has little

influence on the parameter Λ , in this instance, we use the value of the naturally most abundant isotope mass.

Following arguments given in ref. [4 Ch.17], the coefficient, that provides the average number of displacements produced by a neutron of energy E , is determined by the compact expression: $\Lambda \cdot E_n / 4E_d$. The E_d defined above depends on the chemical elements. Its values, for some of those found in space electronic devices, are reported in Tab. 1

TABLE 1. MINIMUM ENERGY VALUES I.E., DISPLACEMENT THRESHOLD ENERGIES, E_d FOR SOME CHEMICAL ELEMENTS OF CONCERN [20].

Element	C	O	Al	Si	Fe	Cu
E_d [MeV]	32	24	25	25	24	30

We understand that the E_d magnitude depends on several parameters such as the crystallographic direction [hkl], temperature, applied strain, and other factors, however, for convenience in Table 1 the average or effective value, is reported i.e. $E_d = \text{ave}(E_d(\theta, \phi))$, averaged over all lattice directions.

The neutron average energy and the solar flare-induced neutron flux are given respectively by the expressions:

$$\bar{E}_n = \frac{\int_0^{\infty} E_n \phi(E_n) dE_n}{\phi} \quad \Phi = \phi t = t \int_0^{\infty} \phi(E_n) dE_n$$

where all variables have been already defined, except t , which is time [s]. Neutron spectra given by Hua et al [21], is used to determine the average neutron energy $E_{n,av}$ for three energy intervals < 5 ; 5-20; 100 MeV respectively as $E_n=1; 5; 10; 50$ MeV.

Watanabe [24] indicates a neutron flux on the Sun's surface is 690×10^{28} n.sr. This is a value that during a solar flare may increase considerably at least 35 times [22]. Considering the Moon's surface and the inverse square law the expected number of neutrons is around 1×10^{10} n.sr. Gorchacheva et al. [6] employing on board Resurs-DK1 satellite, reported on a single solar flare a neutron count rate of 660 n.sec^{-1} . This rather low counting rate is observed probably due to the existing radiation shielding.

V RESULTS AND DISCUSSION

Displaced atoms per neutron collision for the concerned elements are given in Table 2.

TABLE 2. DISPLACED ATOMS BY NEUTRON COLLISION FOR ENERGIES ABOVE E_d / Λ [20]

Element	C	O	Al	Si	Fe	Cu
E_d / Λ [eV]	83	145	116	187	348	363
DPA/n 10^3	3.0	1.7	2.2	1.3	0.7	0.69

The specific displacement rate R_d per targeted element is determined by the expression (9) in its simplified form:

$$R_d = \frac{N \wedge \sigma^{el}}{4E_d} \bar{E}_n \phi$$

Where: N is the number of atoms/cm³; σ_{el} is the neutron elastic cross section on average for a given energy average \bar{E}_n ; Φ is the number of n.cm⁻².s⁻¹.

The threshold energy values change with the atomic reduced mass value as given in Fig. 5.

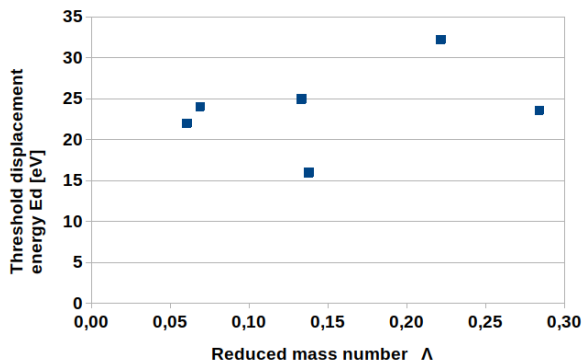


Fig. 5. Threshold energy variation in function of the coefficient Λ .

The Rd-values in units of [disp. cm⁻³. s⁻¹] are given in Table 3 considering values already reported in the previous tables.

TABLE 3. COEFFICIENTS USED TO DETERMINE SPECIFIC DPA FOR NEUTRON AVERAGE OF 1 MeV AND A FLUENCE

	C	O	Al	Si	Fe	Cu
N at.cc x10 ²³	1.05	0.27	0,6	0.5	0.84	0.84
σ^{el} [b]	6	5	3	2	3	2
Φ [n.cm ⁻²]	10 ¹⁵	10 ¹⁵	10 ¹⁵	10 ¹⁵	10 ¹⁵	10 ¹⁵

Elastic reaction cross section values are reported in the literature, e.g. Konobeyev et al. [19], reported some for structural materials below 800 MeV.

TABLE 4. AVERAGE NEUTRON ENERGY AND NEUTRON FLUX Φ (EN) [N.CM⁻².S⁻¹] VALUES WERE USED IN THIS STUDY.

Parameter	Case 1	Case 2	Case 3	Case 4
E_n (MeV)	1	10	100	150
Φ (En)	10 ⁷	10 ⁸	10 ⁹	10 ¹⁰

For case 1, employing the equation and data provided by mentioned tables, the number of displaced atoms per cm⁻³ per unit time in [s] and the resulting DPA for a selected neutron flux is given in Table 5:

TABLE 5. NUMBER OF ATOMS DISPLACED FOR CASE 1 OF TAB.4 (EN=1MeV AND Φ (EN) = 10⁷ [N.CM⁻². S⁻¹])

Displaced atoms	C	O	Al	Si	Fe	Cu
Rd (cm ⁻³ .s ⁻¹) x10 ⁸	18	1.4	3.8	0.69	1.8	2.1
DPA (s ⁻¹) x10 ⁻¹⁵	18	8.6	6.5	2.7	2.2	1.4

In Fig. 5 The number of displacements per atom for the case when only solar flare neutrons are of concern, energy, and fluence values are respectively 1MeV and 10⁹ [n .cm⁻².s⁻¹].

We may observe that the expected contribution to the total radiation damage is negligibly low.

On the other hand, a recent study [23] did show that solar flare protons interacting with solar matter produce a considerable number of neutrons. Solar protons observed during the solar event from August 4, 1972, had a fluence, in the energy range 10-100 [MeV], between 10⁺¹⁰ and 10⁺⁸ [p.cm⁻²] observed during an average solar flare time interval of 2 days (172800 [s]). To illustrate the impact on space electronics we report in Fig. 6 that the DPA displaced atoms due to solar flare neutrons value for an extreme case when the flux is 10⁹ n.cm⁻². s⁻¹ and that due to proton interaction with solar matter (1,7 x 10¹⁰ n.cm⁻². s⁻¹).

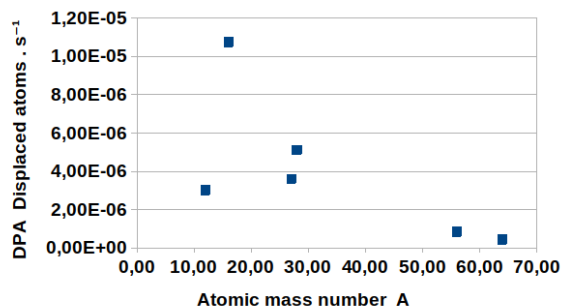


Fig.6 Neutron-induced displacement of some structural components employed in space electronics.

Figure 6, shows the general tendency for targeted matter; it is worthwhile to note that Si-atom ($A=27$) during solar flare will experiment approximately 10⁻⁶ displacements per second meaning that during exposure each atom suffers dislocation every 10⁶[s] from its normal lattice site.

VI CONCLUSION

Space electronics during any mission can be exposed to the radiation dose of several kGy.a⁻¹. Primary and secondary neutrons originating in solar flare and those produced by primary protons (p,n) nuclear reaction taking place, could induce atomic displacement of high frequency defined by the DPA-value. The related damage induces SEE in space electronic components ending in permanent damage. In the case of components employed in IC with two or more processors for parallel computing, the solar flare neutron flux could be of such an intensity to displace every Si-atom in a few weeks compromising the device performance. Currently, the *Wigner effect* is one of the obstacles that burden long-term astronaut missions and space habitats

VII SUMMARY AND FURTHER RESEARCH

Space electronic devices are exposed to radiation-induced damage due displacement of lattice atoms. The impact is related to the single-event effect process. Ionizing dose absorbed enhances the mechanism disrupting the electronic component response. An interplanetary journey could take several years during that time the space electronic devices could be exposed to solar flare neutrons. In the case of exposure, this study pointed out that in the extreme case the impact is negligible. That is in opposition to results when solar flare protons induce (p, xn) reactions with solar matter-producing reactions are included. In fact, the secondary

neutron flux can be considerably high to the extent that almost weekly every single Si-atom suffers site displacement referred to as *Wigner disease*. The annealing provides a recovery process often unsuccessful to re-establish the original structure with an outcome of deteriorating effect on the space electronic performance. To reduce radiation damage in the matter, shielding is devised. Some new mitigation strategies such as manufacturing smaller components and 3D circuits could be the solution for long-term space human exploration.

REFERENCES

- [1] Pálfalvi, J., & Sajó-Bohus, L. (2015). Cosmic radiation detection by solid state nuclear track detector technique. In *Solid State Phenomena* (Vol. 238, pp. 16-54). Trans Tech Publications Ltd.
- [2] Paillet, P., & Leray, J. L. (1999). Defects and radiation-induced charge-trapping phenomena in SiO₂. *Instabilities in Silicon Devices*, 3, 723-780.
- [3] Mishev, A., & Usoskin, I. (2020). Current status and possible extension of the global neutron monitor network. *Journal of Space Weather and Space Climate*, 10, 17.
- [4] Nordlund, K., Zinkle, S. J., Sand, A. E., Granberg, F., Averback, R. S., Stoller, R. E., ... & Simeone, D. (2018). Primary radiation damage: A review of current understanding and models. *Journal of Nuclear Materials*, 512, 450-479.
- [5] Lingenfelter, R. E., Flamm, E. J., Canfield, E. H., & Kellman, S. (1965). High - energy solar neutrons: 2. Flux at the Earth. *Journal of Geophysical Research*, 70(17), 4087-4095.
- [6] Goryacheva, V. S., Voronov, S. A., & Mikhailov, V. V. (2017). Method Of Solar Neutron Search with PAMELA Neutron Detector. In *Journal of Physics: Conference Series* (Vol. 798, No. 1, p. 012039). IOP Publishing.
- [7] Poizat, M. (2017). Radiation handbook for electronics TID Total Ionising Dose Presentation, Radiation environment and its effects in EEE components and hardness assurance for space applications, CERN-ESA-SSC workshop.
- [8] Bohus, L. S., Martínez, O. S. A., & Palacios, D. (2018). Cosmic neutrons radiation damage in carbon nanotubes for space applications. *Phys Astron Int J*, 2(2), 94-96.
- [9] Liu, J. C., Mao, S., Nelson, W. R., Kase, K. R., & Ipe, N. E. (1995). Calculations of the radiation doses to the electronic equipment and magnet coil insulation in the PEP-II tunnel. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 359(3), 463-473.
- [10] Sajó-Bohus, L., Pálfalvi, J. K., Akatov, Y., Arevalo, O., Greaves, E. D., Németh, P., ... & Eördögh, I. (2005). Neutron-induced complex reaction analysis with 3D nuclear track simulation. *Radiation measurements*, 40(2-6), 442-447.
- [11] Vibishna, B., Beenamol, K. S., & Singh, A. K. (2013). Understanding single-Event effects in FPGA for Avionic system design. *IETE Technical Review*, 30(6), 497-505.
- [12] Pálfalvi, J. K., Akatov, Y., Szabó, J., Sajó-Bohus, L., & Eördögh, I. (2006). Detection of primary and secondary cosmic ray particles aboard the ISS using SSNTD stacks. *Radiation protection dosimetry*, 120(1-4), 427-432.
- [13] Lario, D. (2012). Estimation of the solar flare neutron worst - case fluxes and fluences for missions traveling close to the Sun. *Space Weather*, 10(3).
- [14] Wigner, E. P. (1946). Theoretical physics in the metallurgical laboratory of Chicago. *Journal of Applied Physics*, 17(11), 857-863.
- [15] Moll, M. (2018). Displacement damage in silicon detectors for high energy physics. *IEEE Transactions on Nuclear Science*, 65(8), 1561-1582.
- [16] Trung, N. T. H., Phuong, H. S. M., Starostenkov, M. D., Romanenko, V. V., & Popov, V. A. (2018, November). Threshold displacement energy in Ni, Al and B₂NiAl. In *IOP Conference Series: Materials Science and Engineering* (Vol. 447, No. 1, p. 012004). IOP Publishing.
- [17] Olander, D. R. (1976). *Fundamental aspects of nuclear reactor fuel elements: solutions to problems* (No. TID-26711-P2). California Univ., Berkeley (USA). Dept. of Nuclear Engineering.
- [18] Pease, R. L., Johnston, A. H., & Azarewicz, J. L. (1988). Radiation testing of semiconductor devices for space electronics. *Proceedings of the IEEE*, 76(11), 1510-1526.
- [19] Konobeyev, A. Y., Korovin, Y. A., & Sosnin, V. N. (1992). Neutron displacement cross-sections for structural materials below 800 MeV. *Journal of nuclear materials*, 186(2), 117-130.
- [20] McLane, V. (1997). Nuclear data services of the nuclear data centers network available at the national nuclear data center. *Fusion engineering and design*, 37(1), 217-221.
- [21] Hua, X. M., Kozlovsky, B., Lingenfelter, R. E., Ramaty, R., & Stupp, A. (2002). Angular and energy-dependent neutron emission from solar flare magnetic loops. *The Astrophysical Journal Supplement Series*, 140(2), 563.
- [22] Lockwood, J. A., Yingst, H. E., Calawa, A. R., & Sarmaniote, G. (1956). Cosmic-ray neutron intensity increase associated with solar flare of February 23, 1956. *Physical Review*, 103(1), 247.
- [23] Sajó-Bohus, L., López, J. A., & Castro-Colin, M. (2021). Simulation of Dose Estimations from Solar Protons: A PMMA-Bi₂O₃ Shielding Model for Space Exploration.
- [24] Watanabe, K., Muraki, Y., Matsubara, Y., Murakami, K., Sako, T., Tsuchiya, H., ... & Shirasaki, Y. (2003). Solar neutron event in association with a large solar flare on 2000 November 24. *The Astrophysical Journal*, 592(1), 590.

Drought monitoring mapping by using satellite remote sensing: Study Case

Malgorzata Veróné Wojtaszek
Institute of Geoinformatics - Alba
Regia Faculty
Precision Farming Research Center
University Research and Innovation
Center
Óbuda University
Székesfehérvár, Hungary
wojtaszek.malgorzata@amk.uni-
obuda.hu

Erick Noboa
Doctoral School of Applied Informatics
and Applied Mathematics
Precision Farming Research Center
University Research and Innovation
Center
Óbuda University
Budapest, Hungary
erick.noboa@irob.uni-obuda.hu

György Eigner
Biomatics and Applied Artificial
Intelligence Institution, John von
Neumann Faculty of Informatics
Physiological Controls Research Center
Precision Farming Research Center
Óbuda University
Budapest, Hungary
eigner.gyorgy@nik.uni-obuda.hu

Abstract— Drought is the consequence of the absence of water, and the many variables that produce the meteorological drought (as the reduction of precipitation, climate changes, and surface temperature between others) are well understood. When this Meteorological drought is persistent, it leads to a reduction of soil moisture or agricultural drought which consequently reduces the stream of rivers, lake levels, and underground flows creating a hydrological drought. In growing crops, the water deficit reduces the production efficiency of the enzymes in charge of the photosynthesis process of the plants which harms the farm’s production inducing a socioeconomic drought. European countries went through one of the most droughted periods of the last centuries during the summer of the year 2022 increased by the frequent heatwaves and Hungary was not exempted from its consequences, where the higher drought index is commonly observed in the northern, western, and central section of the country. In the study, satellite images were used to map the state of vegetation. We monitored the effect of drought on crop-growing soils based on the statistical analysis of vegetation indices calculated from images, where the Normalized Difference Vegetation Index (NDVI) values are notably increasing until the 21st of Jun of the year 2022 reaching a maximal value of 0.6012 in one of the analyzed fields before it starts decreasing to a minimum observed value of 0.3209 in the same field after a period of thirty-seven days due to lack of water.

Keywords—Crop Growing, Drought, Remote Sensing, Image Index Analysis

I. INTRODUCTION

The dynamics of the soil-atmosphere moisture interaction start when the land content of water naturally evaporates, rising to the atmosphere, where these molecules lose energy, giving place to condensation while the temperature decreases, creating local cumulus of water, and finally precipitate back to the soil closing this dynamic cycle. When the field of study presents an anomalous leak of this precipitation during an extended time (known as a meteorological drought), other sources of water, such as lakes, rivers, or underground flows decrease their levels or caudal [1]. This hydrological drought eventually influences the plant metabolism, the stress produced in the crop impacts different levels of the vegetation modifying physiological and biochemical features, i.e., rooting, growing, and breeding processes considered as an agricultural drought [2], [3]. The effects of soil drought have shown consequences all around the Earth’s surface as example, the drought that California – United States of America – suffered between the years 2012 – 2015 registered an economical loss of \$810M only in the farming sector [4].

On the other hand, the European continent passed through the most drought summer of the last centuries according to the Global Drought Observatory (GDO) analytic report of August 2022 due to the frequent heatwaves that crossed the continent.

The availability of different Remote Sensing (RE) techniques provides a tool to support the decision-making activity within the involved agricultural processes, from the land preparation prior to seeding until the harvest time of the crop [5]. This RS trend presents a huge rising pivot in the last 20 years with all the unmanned aerial applications of this variety of systems i.e. UAV’s or dedicated satellites implemented to compile enough information to give a clear idea of the current health state of the analyzed field, allowing farmers to react on time and optimize the crop production while increasing the usage efficiency of resources as water.

Multi-spectral remote sensed imagery has become one of the most common practices to obtain new images from the combination of two or more frequencies for information extraction [6] to indicate canopy changes due to water stress. One of the earliest RS analytical products is the NDVI used to simplify the analysis of the multi-frequency images resulting from the usage of the near-IR band which is highly effective for expressing vegetation status and quantified vegetation attributes [6].

II. SATELLITE IMAGES AND DATA ANALYSIS

Multispectral ESA Sentinel-2 satellite images acquired from the ESA Sentinel Scientific Data Hub were used in this study. The investigation period was between June and August 2022. A total of 13 cloud-free images were available, but after preliminary interpretation, 7 of them were selected for further processing. Meteorological (the amount and the intensity of precipitation) data were considered in the selection. The characteristics of the data used in the research are included in Table 1.

TABLE I. REMOTE SENSING AND METEOROLOGICAL DATA

Dataset	Parameter	
	Data type	Acquisition date
Sentinel-2	Spectral resolution VIS, NIR, SWIR (bands: 3, 4, 8, 12)	03, 11, 21. June 2022 06, 18, 21, 28 July 2022
Meteorological station	Precipitation accum. total	May (13 mm) June (49 mm) July (17 mm) August (67 mm)

NDVI values range from -1 to 1. NDVI values close to -1 occur in areas covered by water. Values ranging from -0.1 to 0.0 occur in areas of exposed soil without vegetation cover. Values from 0.0 to 0.33 occur in areas covered with poor vegetation or dry vegetation. NDVI index values ranging from 0.33 to 0.66 are characteristic of areas covered with vegetation of average condition. Values NDVI index above 0.66, are characteristic of plants in good condition being at the highest stage of development. The NDVI thresholds describing the different quality of vegetation were selected based on the literature and on field surveys. In this work we considered the results of image analyses of previous years with adequate water supply and maize growing stages. The field work was carried out with the participation of the crop production manager.

During the study, the analysis of the values of NDVI indicators was done for maize cultivation in the period from June to the end of August 2022. The calculation of NDVI index values was done for all images and the minimum, maximum, and average values were determined within the studied fields. In order to determine the impact of drought on the vegetation condition, the NDVI change analysis was done within examine period (Figure 3), as well as a comparison with each value of the NDVI indices for the corresponding months in 2019 (Figure 4).

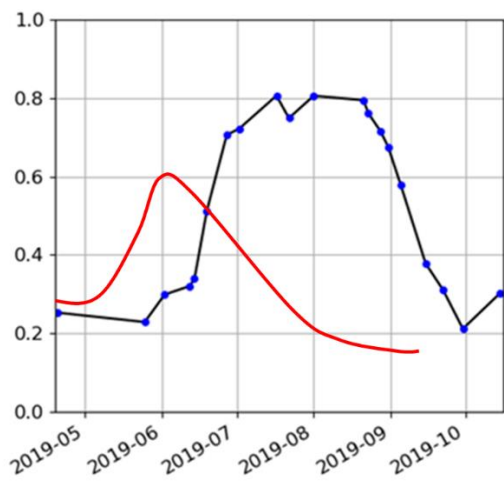


Figure 4. NDVI values change for the years 2022 (red line) and 2019 (blue line)

The calculated NDVI values exposed in the Figure 3 shows a decreasing tendency after the 21.06.2022 compared with the NDVI index values obtained the previous year (2019) in the corresponding months, when there was sufficient water supply for plant growth, with notable increasing numbers measured until the end of August reaching values above 0.8 as can be observed in blue dots on the Figure 4, in contrast with a maximum of 0.6 calculated in the 2022 (Figure 4 red line) which represents a reduction of 25% in three years. The calculated values of the NDVI index allow us to conclude that for all analyzed months, lower NDVI index values occurred in 2022, influenced by the drought.

IV. OBJECT-BASED CLASSIFICATION

In the object-based classification process we used segmentation to determine objects and it is included the following steps:

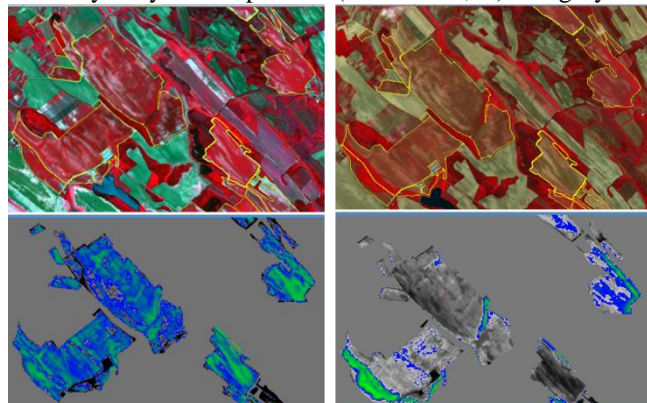
A. Segmentation

- Chessboard segmentation using thematic layer (Level 1, fields identification)
- Chessboard segmentation 20 m units (Main Level, analysis level)

B. Object classification

- Spectral values of 4 bands
- NDVI

In the classification process the objects were assigned into tree classes: maize with poor, average and high condition. Due to the draught only a few segments (cells) of area were assigned into the ‘high quality’ category. After 21 June basically only the ‘poor’ (NDVI ≤ 0.35) category was



detectable (Figure 5).

Figure 5 Sentinel2 images taken on 21.06.2022 and 07.28.2022 and NDVI based classification results. Pixels with the NDVI values of bigger than 0.35 are marked in color.

A view of the analyzed fields by using NDVI-based classification for two dates are presented in the Figure 5. The green-blue color indicates areas with NDVI values greater than 0.35.

V. CONCLUSION

The severe drought affecting many regions of Hungary since the beginning of the year has been further expanding and worsening as of early August. Dry conditions are related to a wide and persistent lack of precipitation combined with a sequence of heatwaves from May. One of the crops most affected by the negative precipitation anomalies in the three months ending on the end of August 2022 is maize. The NDVI index was applied to analyze the vegetation condition in this study area. According to the results, after 21 June the NDVI values show a sharp decline, while in the end of July the vegetation of most of the study area is completely dry.

On the other hand, the satellite Sentinel 2 provide data with relatively high spatial resolution (10-20 m) and high temporal resolution (3-5 days) observations of the Earth. So that remote sensing data has been increasingly used not only for large-area but also for within-field drought monitoring.

Acknowledgements: The authors gratefully acknowledge the support provided for this research by the GINOP_PLUSZ-2.1.1.-21-2022-00040 Project. The project is supported by Széchenyi Plan Plus, co-funded by the Government of Hungary and the European Union. We thank Zsolt Keresztes

plant production Director of the Gyermelyi Holding Zrt. for the cooperation.

REFERENCES

- [1] A. Berg and J. Sheffield, "Climate Change and Drought: the Soil Moisture Perspective," *Clim. Chang. DROUGHT*, vol. 4, pp. 180–797, 2018, doi: 10.1007/s40641-018-0095-0.
- [2] D. Kapoor, S. Bhardwaj, M. Landi, A. Sharma, M. Ramakrishnan, and A. Sharma, "The Impact of Drought in Plant Metabolism: How to Exploit Tolerance Mechanisms to Increase Crop Production," *Appl. Sci. 2020, Vol. 10, Page 5692*, vol. 10, no. 16, p. 5692, Aug. 2020, doi: 10.3390/APP10165692.
- [3] M. S. Iqbal, A. K. Singh, and M. I. Ansari, "Effect of drought stress on crop production," *New Front. Stress Manag. Durable Agric.*, pp. 35–47, Mar. 2020, doi: 10.1007/978-981-15-1322-0_3/COVER.
- [4] "Economic Analysis of the 2014 Drought for California Agriculture | Center for Watershed Sciences." <https://watershed.ucdavis.edu/2014-drought-report> (accessed Oct. 03, 2022).
- [5] S. Khanal, K. Kc, J. P. Fulton, S. Shearer, and E. Ozkan, "remote sensing Remote Sensing in Agriculture-Accomplishments, Limitations, and Opportunities", doi: 10.3390/rs12223783.
- [6] S. Huang, L. Tang, J. P. Hupy, Y. Wang, and G. Shao, "A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing," *J. For. Res.*, vol. 32, no. 1, pp. 1–6, Feb. 2021, doi: 10.1007/S11676-020-01155-1/FIGURES/2.
- [7] M.V Wojtaszek, I Abdurahmanov: Crop Water Condition Mapping by Optical Remote Sensing, International Journal of Geoinformtics Vol. 17 No. 1 (2021): Volume 17, No. 1 January - February 2021 <https://journals.sfu.ca/ijg/index.php/journal/issue/view/22>

Long Term Rumen Temperature Analysis of Dairy Cattle with Help of Bolus Sensor

Éva Hajnal
Alba Regia Technical Faculty
Óbuda University
Székesfehérvár, Hungary
hajnal.eva@amk.uni-obuda.hu

Gergely Vakulya
Alba Regia Technical Faculty
Óbuda University
Székesfehérvár, Hungary
vakulya.gergely@amk.uni-obuda.hu

Péter Udvardy
Alba Regia Technical Faculty
Óbuda University
Székesfehérvár, Hungary
udvardy.peter@amk.uni-obuda.hu

Abstract—The article contains the analysis of data from a dairy cattle rumen bolus experiment. The bolus contained a 3D accelerometer and a thermometer and was used to monitor a cow for 9 days. The frequency of drinking could be determined from the measured temperature data, checked on the basis of the control video recording. The monitoring of changes in the cow's body temperature was also analyzed based on the measured data series.

Keywords—dairy cattle, rumen bolus

I. INTRODUCTION

In the modern economy, there is an increasing demand for economical production, which typically manifests itself in the automation of processes. This phenomenon has also reached precision agriculture and dairy farms. As a result, the continuous increase in the size of herds, the creation of farms often containing several thousand animals, and the continuous decrease in the amount of human labor can be observed. At the same time, increasing the milk yield per cow is also a requirement. This strict set of requirements can only be fulfilled by using the cutting edge technical possibilities. Thus, the need for continuous monitoring of farm animals also arose [1]. In addition to the goals mentioned above, the basic requirements of animal welfare were already formulated in the 1970s, and farms are trying to pay more and more attention to complying with them. What are these principles? In 1979, the FAWC (the UK's Farm Animal Welfare Council) formulated the 5 basic animal rights, freedom from hunger and thirst, from discomfort, from pain, injury and disease, from fear and distress and freedom to express most normal behaviors. At the same time, it became clear that the observance of fundamental rights cannot be evaluated objectively [2]. The EU FP6 research program eventually developed semi-objective protocols for animal welfare assessment. The definition of good health was absence of disease, injuries and pain, with physical, social and mental well-being assessed through four measures of appropriate behavior [2]. Along these principles, it is advisable to develop monitoring systems that can indicate phenomena and problems related to animals, and can provide information and alerts to farmers in relation to animals. Such a possible monitoring system can be created with the help of a complex sensor placed in the cattle's rumen [3], which measures the conditions of the cow's rumen with several

sensor modalities, processes the information and sends it to the integrated animal welfare monitoring system. Such a system can detect problems related to feeding the cow [4], the animal's illness [5], lameness, heat stress during warm episodes in summer, but it can also indicate to the farmer the cow's estrus or the onset of calving [2].

In this publication, we would like to present such a sensor experiment and, during the processing of the temperature data measured by the sensor, we are looking for the answer of what questions the continuous measurement of the rumen temperature helps to answer, what physiological information it provides us with. There were similar attempts in the literature earlier [6]. Chapter II of the article presents the architecture of the sensor and the measurement system. Chapter III. discusses the evaluation of the received data and the interpretation of the results. The final version of the bolus is planned for working at least 2 or 3 years long, which is reachable if the communication events are radically restricted and the sent data amount also limited. In this preliminary experiment the main aim was to get real sensor data from the system for data processing and data compression algorithm development.

II. MEASUREMENT METHOD

This measurement was intended to be a preliminary experiment. A bolus sensor meant to be used throughout the entire life of a cattle must be optimized in terms of energy consumption with constrained sampling and processing strategies with as few radio communications as possible. As the opposite, an experimental bolus has to serve raw, unprocessed data with continuous sampling, with the same sensors and physical layout of the final design, if possible. The bolus has two integrated sensors: a triaxial accelerometer and a temperature sensor. Based on previous experience the sample rate was set to 25 Hz in case of the accelerometer and to 1 Hz for the temperature sensor. The collected raw data was sent to the gateways through a custom error resilient radio link, where the data is collected for off-line processing.

The sealed, water-tight bolus sensor was injected to one test animal at Agárdi Farm Kft. Seregélyes (see Fig. 1). During the time of the application the animal was taken into a cage to prevent injuries and to provide easier access. As a control,

the behavior of the animal was regularly inspected and video recordings were also made.



(a)



(b)

Fig. 1: Injecting the bolus sensor into the rumen of the cattle with the applicator. (a) The applicator is inserted in its total length. (b) The applicator is pulled out after injecting the bolus.

Every collected data was stored with timestamps. Temporal positioning of the recorded video data was achieved by regularly showing a clock to the camera, to create reference points.

During the first phase of the data processing only the temperature measurements were analyzed (see Chapter III.).

A. Hardware architecture

The experimental rumen bolus sensor is built into a durable water-tight enclosure. The cylindrical device has a diameter of 40 mm and length of 140 mm, and made from POM-C plastic. The enclosure has a cavity inside formed as a blind hole with diameter of 36 mm and length of 135 mm. The power source is a 3 V D type battery, which is able to store 14 Ah of capacity.

The main component of the electronic part is an ST STM32L496RET6 microcontroller with a maximum clock of 80MHz PLL. The bolus is equipped with 2 different sensors:

- A Sensirion STS35 thermometer
- An ST LSM6DSO triaxial accelerometer

The radio communication chip is a Würth AMB3626-M, operating on 169MHz. The software running on the microcontroller measured 1 temperature and 25 acceleration samples each second and sent this data through the radio link to the gateways. Each packet was transmitted two times to make the system more robust.

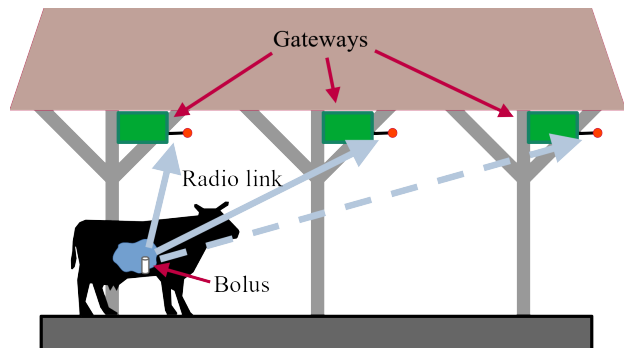


Fig. 2: The architecture of the experimental measurement setup

The test area (approx. 100 m × 15 m) was covered with 3 gateways. The hardware of the gateways is based on the Raspberry Pi 3B single board computer, which runs the data collection software. Measurement records are collected into CSV files. Time synchronization is done with Network Time Protocol (NTP) through 4G cellular connection. The architecture of the measurement system is shown in Fig. 2.

III. RESULTS AND DISCUSSION

In this paper the temperature measurements and data analysis is presented. First the data collected by the 3 gateways were cleaned and merged. Then the time series were split into 12-hour intervals and each segment was visualized. Fig. 3 shows two example 12-hour intervals.

TABLE I: Drinking events and temperature data during the experiments

Date	Drinking events	Drinking events in rarefaction	Maximum daily temperature
28/07	6	5	33.41
29/07	7-8	7-8	35.6
30/07	5	5	25.08
31/07	5	6	26.67
01/08	8	8	31.34
02/08	8	8	34.15
03/08	7	7	34.12
04/08	3	2	35.68

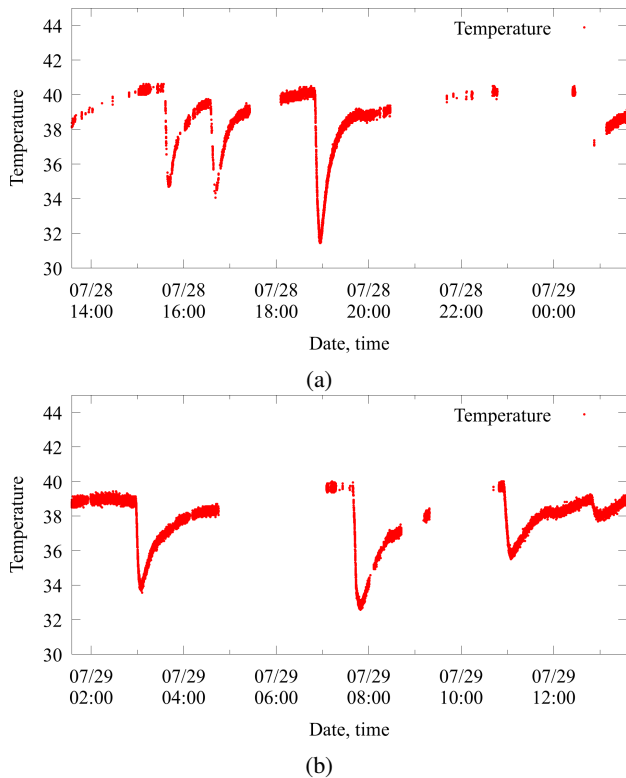


Fig. 3: The temperature of the rumen measured during a 12-hour period from noon to midnight (a) and from midnight to noon (b).

A. Data integrity

During the data cleaning and merging phase the nature of data integrity was well observable. Since the bolus had to be powered on approx. 2 days before the experiment to let the silicon sealing dry, the first received real measurement data packet has a sequence number of 64301. Then the sequence numbers were increasing for approx. 8 days. When the battery started to run low the bolus started to keep resetting at every few minutes, resulting in unstable measurements.

In the experiment 3 gateways were used to provide a broader coverage. As opposed to the expectations most of the data packets were received either by all the 3 gateways, or none of them at all. Probably in some positions (like lying) or when the direct line of sight between the bolus and the receiver antenna was covered by another cattle the attenuation was too high. Smaller and larger gaps are clearly visible in Fig. 3. Note that the cattle are regularly moved from the barn to another locations e.g. for milking.

B. Data quality

Assuming, that the mass of the rumen of the cattle is large (note that its volume is approx. 0.12 m^3) the temperature can be considered constant in the minute-scale. However in Fig. 3 a thick line is visible, which suggests a high noise level. To estimate measurement error a continuous sample series for 30 minutes was selected from a calm time interval (see Fig. 4),

then the most important statistical properties were calculated. The mean of the sample is $38.99 \text{ }^\circ\text{C}$ and the standard deviation is $\sigma = 0.12 \text{ }^\circ\text{C}$. This is in a good concordance with the factory description of this temperature meter sensor. The maximum and minimum values were $38.5 \text{ }^\circ\text{C}$ and $39.44 \text{ }^\circ\text{C}$, being $-0.49 \text{ }^\circ\text{C}$ and $+0.45 \text{ }^\circ\text{C}$ from the mean.

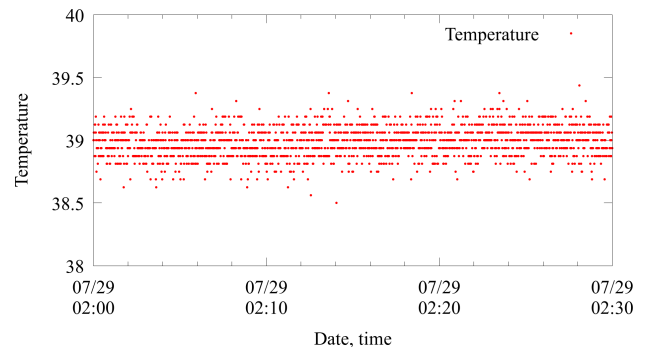


Fig. 4: The selected data slice.

The quantization of the measurements is clearly visible. The resolution of the thermometer is $\frac{1}{16} \text{ }^\circ\text{C}$.

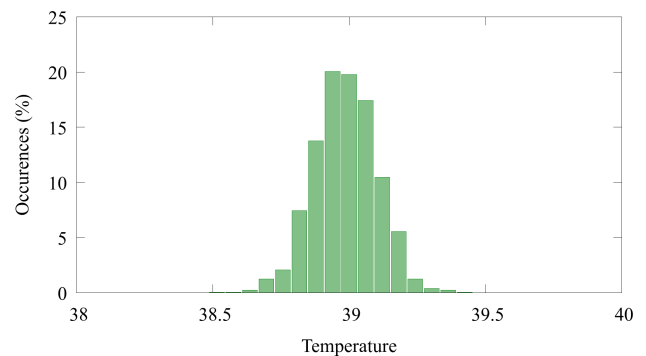


Fig. 5: The histogram of the selected slice.

The histogram of the slice of data is shown in Fig. 5, showing a Gaussian distribution.

C. Data analysis

In both Fig. 3(a) and (b) 3 large negative peaks are visible. In the (a) graph another similar peak is probable, but it is invisible, due to the missing data. In the (b) graph another, much smaller negative peak is visible with slightly different nature. The larger negative peaks belong to drinking events in good concordance with the literature [7]. The water intake happens nearly instantly at this time scale. The temperature of the water available in the barn is approx. $14 \text{ }^\circ\text{C}$. It mixes with the content of the rumen and the resulting temperature gives the minimum of the negative peak. Note that some time is required for the mixing and for the bolus to follow the temperature change. After this point an approximately exponential warming up follows, when the rumen content reaches the original temperature.

During the 8-day period a total of 51 drinking events were observed based on the temperature data, which were approximately uniform in each day. Drinking behavior was not significantly different between daytime and night. Daily drinking events were determined from the measured temperature curve (Table I.). The measurements were done during hot summer days, so it was interesting the correlation between drinking events and daily maximum temperature. The Pearson correlation was 0.052 that means no correlation between temperature and number of drinking events. It is possible that the amount of water drunk is correlates with the temperature so later it should be investigated with appropriate mathematical model and measured controls.

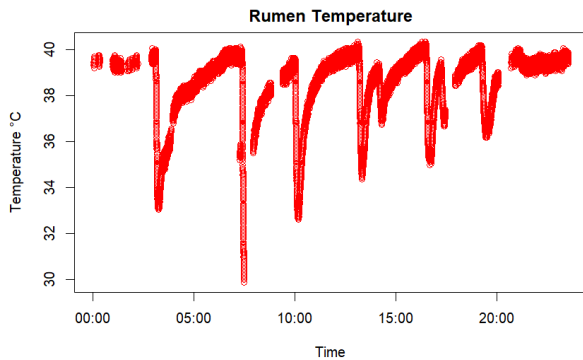


Fig. 6: Drinking events on native temperature data

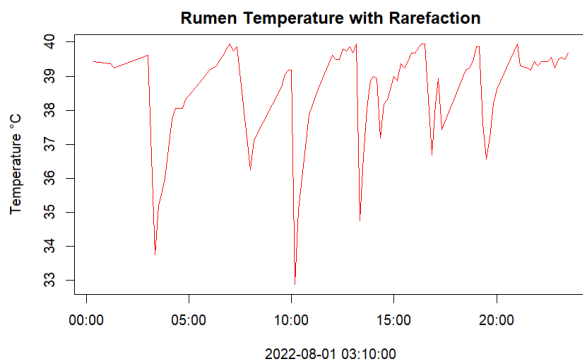


Fig. 7: Drinking events on resampled temperature data series (measurement period 10min)

In the next experiments the aim is to measure longer interval, which is achievable by decrease the frequency of measurements. So the original dataset was resampled to select measurements only every 10 minute (sampling frequency 1/600 Hz). This resampled data is shown on Fig.7. The line is thin because the noise of the sensor is not visible due to resampling. The large negative peaks connected to the water intake are also visible on this chart. The drinking events were determined and counted from the graph (Fig 7.) and Table I. shows the results which are relevant comparing with the higher frequency measurement. Mean error is 0.38. However, the

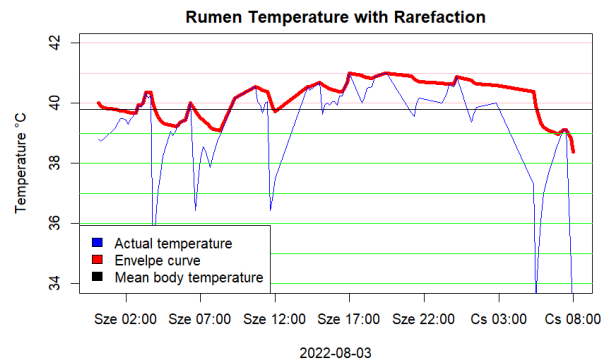


Fig. 8: Detection of estrus based on the temperature meter measurements

minimum values of the resampled graphs sometimes largely differ from the original chart, because the temperature fall after drinking is relatively quick, its duration between 20 and 30 minutes. Temperature measurements are useful in detection of drinking events and detection of illnesses or other physiological events connected with temperature changes (etc. estrus). To detect the base body temperature an envelope curve was created to compensate the effect of drinking on the body temperature. The standard deviation of the temperature is 0.12 and the temperature changes during the drinking is at least 5-6 °C. So the interval maximums were calculated (at first in every 6 hours), and the points of the curve were detected with a weighted average between maximum and actual temperature, in which the positive deviation is weighted with 20-fold weight. So the drinking threshold is 2-4 °C. The Figure 8 shows an example of the envelope curve. This envelope curve possibly reflects the real body temperature – however, it should be strengthened with further experiments. Figure 8. shows the body temperature in a two days' period. In this period the estrus is detectable (smaller temperature increase wave and after starts the estrus near 1 °C temperature increase for 12-18 hours. [8]) This detection should be support yet with activity data.

IV. SUMMARY

In this paper a long-term measurement was presented with a rumen bolus sensor. A sensor device was applied into the rumen of a dairy cattle, which measured acceleration and temperature data continuously for 10 days. The temperature data was processed and the typical pattern of drinking was identified. These patterns can easily be detected even with a much sparse sampling, even within the future bolus. The base rumen temperature also was detected, which can be applicable to monitor the health status of the animal or detect important physiological events as estrus.

ACKNOWLEDGMENT

”This research was funded by NKFI Hungary, grant number 2020-1.1.2-PIACI-KFI-2020-00109”. The authors would like

to thank to Levente Kovács MATE Hungary for his advice and Agárdi Farm Kft. Seregélyes to help in the investigation.

REFERENCES

- [1] C. H. Knight, "Review: Sensor techniques in ruminants: More than fitness trackers," *Animal*, vol. 14, no. S1, pp. S187–S195, 2020. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1751731119003276>
- [2] G. Caja, A. Castro-Costa, and C. H. Knight, "Engineering to support wellbeing of dairy animals," *Journal of Dairy Research*, vol. 83, no. 2, pp. 136–147, 2016.
- [3] É. Hajnal, L. Kovács, and G. Vakulya, "Dairy cattle rumen bolus developments with special regard to the applicable artificial intelligence (ai) methods," *Sensors*, vol. 22, no. 18, p. 6812, 2022.
- [4] L. Kovács, L. Rózsa, M. Pálffy, P. Hejmel, W. Baumgartner, and O. Szenci, "Subacute ruminal acidosis in dairy cows-physiological background, risk factors and diagnostic methods= subakutna acidoza buraga u mliječnih krava-fiziološko značenje, čimbenici rizika i dijagnostičke metode," *Veterinarska stanica*, vol. 51, no. 1, pp. 5–17, 2020.
- [5] L. Kovács, F. L. Kézér, F. Ruff, and O. Szenci, "Rumination time and reticuloruminal temperature as possible predictors of dystocia in dairy cows," *Journal of Dairy Science*, vol. 100, no. 2, pp. 1568–1579, 2017.
- [6] J. A. Vázquez-Diosdado, G. G. Miguel-Pacheco, B. Plant, T. Dottorini, M. Green, and J. Kaler, "Developing and evaluating threshold-based algorithms to detect drinking behavior in dairy cows using reticuloruminal temperature," *Journal of Dairy Science*, vol. 102, no. 11, pp. 10471–10482, 2019. [Online]. Available: <http://dx.doi.org/10.3168/jds.2019-16442>
- [7] J. M. Bewley, M. W. Grott, M. E. Einstein, and M. M. Schutz, "Impact of intake water temperatures on reticular temperatures of lactating dairy cows," *Journal of Dairy Science*, vol. 91, no. 10, pp. 3880–3887, 2008. [Online]. Available: <http://dx.doi.org/10.3168/jds.2008-1159>
- [8] T. A. Burnett, M. Kaur, L. Polsky, and R. L. Cerri, "Rumen-Reticular Temperature During Estrus and Ovulation Using Automated Activity Monitors in Dairy Cows," *Frontiers in Veterinary Science*, vol. 7, no. December, pp. 1–11, 2020.

Semi-Automatic Detection and Tracking of Growing Mushrooms on Image Sequences

Gyula Simon
Alba Regia Technical Faculty
University of Óbuda
Székesfehérvár, Hungary
simon.gyula@amk.uni-obuda.hu

Gergely Vakulya
Alba Regia Technical Faculty
University of Óbuda
Székesfehérvár, Hungary
vakulya.gergely@amk.uni-obuda.hu

Sándor Tarsoly
Antal Beczy Center for Intelligent Robotics
University of Óbuda
Budapest, Hungary
sador.tarsoly@irob.uni-obuda.hu

Péter Galambos
Antal Beczy Center for Intelligent Robotics
University of Óbuda
Budapest, Hungary
peter.galambos@irob.uni-obuda.hu

Abstract—The efficient management of autonomous mushroom production plants requires the model of growth rate of mushrooms. Photos of the plants are used as input for the growth models, which then predict the development of individual mushrooms. Recently machine learning techniques have been successfully applied to create such models. For the machine learning systems, however, large number of training samples are required. The training samples include photos of the plant and also ground truth markers indicating the position and size of the mushrooms on the photo. In this paper an image processing system is introduced, which is able to create good quality ground truth from sequences of images of the plant. The proposed system can automatically detect the mushroom positions and sizes on each of the pictures, but also allows user intervention to minimize the number of detection errors.

Keywords—image processing, object detection, computer vision, mushroom cultivation

I. INTRODUCTION

In precision agriculture machine learning techniques are widely used to improve crop quality, e.g. using weather prediction, providing pest prevention, or maintaining optimal conditions in greenhouses [1]-[8]. Machine learning techniques have been successful to forecast various events based on time series, e.g. in weather and climate modelling [9], business [10], finance [11], and lately crop growth modelling [5]-[8] as well.

In mushroom production, technological greenhouses are used, where optimal parameters can be precisely adjusted [5]. Recently machine learning systems have been proposed to detect and measure mushrooms, and estimate their growth rate [5], [6]. To allow the training of machine learning systems, however, a large number of training samples are required, e.g. in the form of a series of photos showing the evaluation of the crop, with the synchronized timeline of various measured environmental parameters possibly affecting the growth rate (e.g. temperature or humidity). The training samples must include the ground truth, e.g. for the photos the mushrooms must be identified for the learning system, e.g. in the form of parameters center and radius. The generation of ground truth is burdensome, especially for large number of samples. In this paper an image processing system is proposed, which aids the creation of good quality ground truth data from sequences of images of the plant. The proposed system models mushroom as a circle, and can automatically detect the position and radius of each mushroom on each of the pictures. The system also allows user intervention to minimize the number of detection errors.

The outline of the paper is the following: Section II will provide the outline of the system and explains the mode of operation. In Section III the main components will be introduced. Section IV illustrates the operation of the system. Section V concludes the paper.

II. SYSTEM OPERATION

The input of the system is a series of photos taken on the crop, as illustrated in Fig. 1. The photos are taken so that the growth of the mushrooms is observable, but the difference is not too large. In practice one photo in each hour gives good results. The operation of the system is based on the following trivial observation: large mushrooms are easier to detect than small ones. Thus, the processing of images is started from the last one, where the mushrooms are the largest. In the first phase all mushrooms with considerable size are detected on the latest image. In the second operation phase, stepping backwards, the mushrooms are tracked on each image. In the second phase information, found in the previous image, is used to locate the mushrooms.

The concept is illustrated in Fig. 1, where the input images are denoted by I_1, I_2, \dots, I_N . In phase #1 the last image I_N is processed. Notice that in this phase the only input is the image. The output of the processing is a vector P_N of estimated circle positions, where each position includes the center and radius of one circle, corresponding to a mushroom. In Fig. 1, four mushrooms are detected, the detections are shown as red circles.

In phase #2, the remaining of the images are processed in reverse order, i.e. $I_{N-1}, I_{N-2}, \dots, I_1$. The result of each step is a

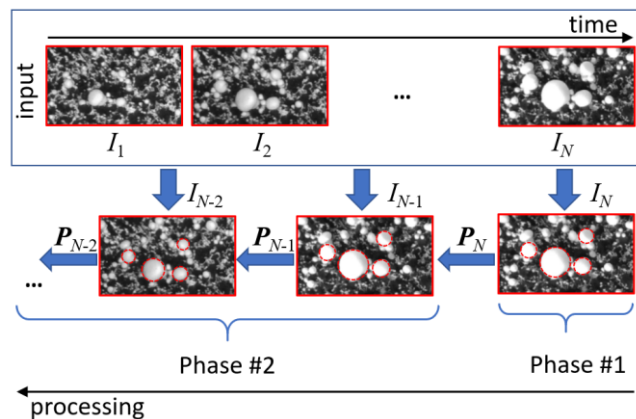


Fig. 1. The concept of image processing

vector of positions, similarly to phase #1. When image I_n is processed, the output is P_n , and the input also includes the result vector P_{n+1} of the previous processing step.

Notices:

- As Fig. 1 illustrates, the processing is performed in reverse order.
- The lengths of the result vectors are the same, determined in phase #1. If the size of a mushroom becomes too small to detect, its position is still maintained in the next processing steps.

III. SYSTEM COMPONENTS

In this Section the image processing tools will be introduced.

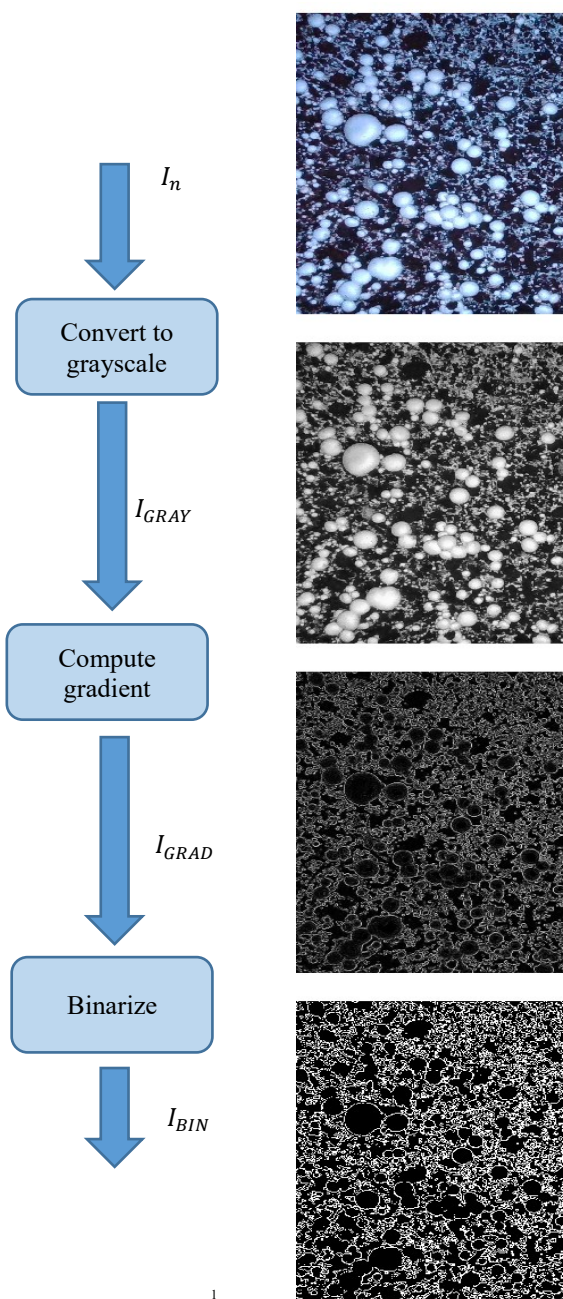


Fig. 2. The image preprocessing steps

A. Image Preprocessing

The input of phase #1 is the last image I_N . In the preprocessing step the image is converted to a grayscale image (I_{GRAY}). To highlight the edges on the image, the image gradient is computed (I_{GRAD}). Finally, the gradient image is binarized (I_{BIN}). The steps of the preprocessing are illustrated in Fig. 2.

B. Phase One

In phase #1 circles are searched for in the binary image I_{BIN} . For circle finding, the classical method is the Circle-Hough-transformation (CHT) [12], which is able to find circles with a given radius on binary images. Extensions of the Hough-transform allow the search of various shapes [13]. Also, circles with different radii can be searched, with iterative usage of the CHT. The Phase Coding Method (PCM) uses a scale-invariant kernel operator, thus it provides more efficient implementation in case the radius is not known [14]. The MATLAB Image Processing Toolbox contains implementations for both CHT and PCM [15]. The performance of the methods is illustrated in Fig. 3.

The input of the circle search was the binary image I_{BIN} , but the figure also contains the grayscale image I_{GRAY} to ease visual evaluation. Both methods have a sensitivity parameter, where higher values close to 1 allow detection of more circles. The examples shown in Fig. 3 allow the following conclusions:

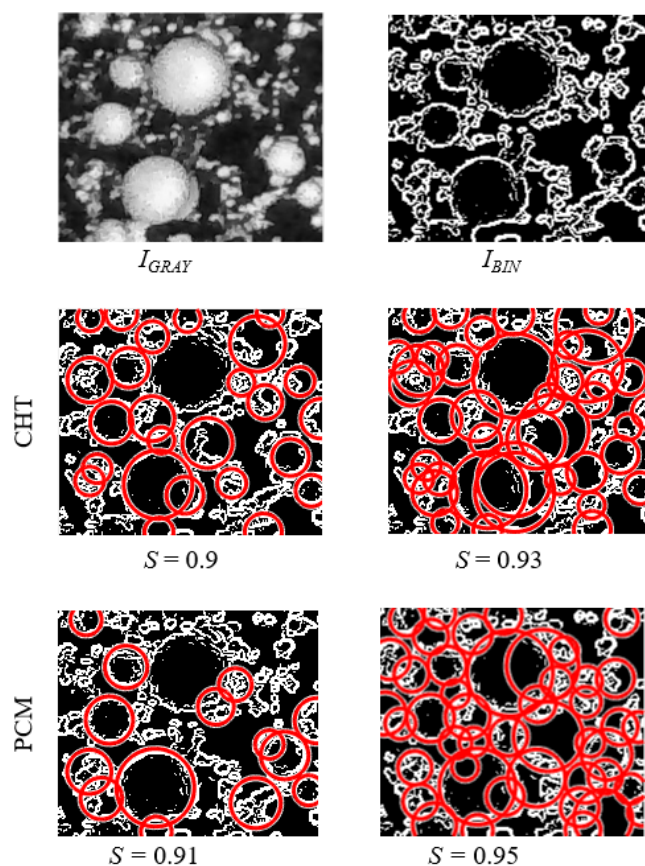


Fig. 3. The performance of the circle finder algoritms CHT and PCM, with various sensitivity values S .

- Small sensitivity values provide fewer false detections but also may miss real targets.
- Large sensitivity values provide a great amount of false detections but the chance of missing a target is smaller.

Since the general circle search algorithms cannot provide robust detection by themselves, they are used to provide initial estimates, which are further evaluated. The sensitivity S is set high enough to provide a rich set of detections, with potentially no misses. The initial estimates filtered and pruned as follows:

The quality of the circle edge is measured by three quantities Q_{per} , Q_{plate} , and Q_{color} . The perimeter quality index Q_{per} of a circle with center C and radius R is calculated as follows:

$$Q_{per} = \frac{N_{white-per}}{2R\pi}, \quad (1)$$

where $N_{white-per}$ is the number of white pixels in I_{BIN} along the circle drawn at center C and radius R . Obviously, Q_{per} measures the ratio of white pixels of the current circle and the ideal circle (C,R) and shows the “fullness” of the circle. A full circle has $Q_{per} = 1$, while a circle with fewer white pixels along the ideal perimeter has $Q_{per} < 1$.

The plate quality of the circle measures the ratio of the black pixels and the total number of pixels inside (C,R) in I_{BIN} , as follows:

$$Q_{plate} = \frac{N_{black-plate}}{2R\pi}, \quad (2)$$

where $N_{black-plate}$ is the number of black pixels. High plate quality value close to 1 indicates that almost all of the pixels inside the circle are black, i.e. there is no significant change of color inside the circle. An ideal round mushroom would have $N_{black-plate} = 1$, while a false detection on the mycelium has small plate quality.

The color index measures the average brightness I_{mean} of the circle in image I_{GRAY} . The color quality is defined as follows:

$$Q_{color} = I_{plate}/I_{max}, \quad (3)$$

where I_{max} is the maximum possible value of the brightness (e.g. in an 8-bit picture $I_{max} = 255$). White mushrooms have high, while the ground and the mycelium have low brightness values.

Using quantities Q_{per} , Q_{plate} , and Q_{color} , each circle is tested. If all three quantities are higher than experimentally defined values L_{per} , L_{plate} , and L_{color} , respectively, then the circle is accepted, otherwise rejected. To each of the accepted circles a combined quality index Q is assigned:

$$Q_{comb} = \mu_{per}Q_{per} + \mu_{plate}Q_{plate} + \mu_{color}Q_{color}, \quad (4)$$

where weighting factors μ_{per} , μ_{plate} , and μ_{color} are set experimentally.

After the quality check the pruning is performed. Initially, the candidate set A contains all of the accepted circles with

their quality indices, while the output set B is empty. The core of the pruning algorithm is the following:

Step 1. Select circle C with the highest $*$ in A .

Step 2. Move C from A to B .

Step 3. Remove circles from A , which are covered by C .

Repeat Steps 1-3 until A is empty.

The pruning is performed in two rounds. In the first round the pruning algorithm is run with $*$ = Q_{comb} . Then the output of the first round is moved to the candidate set, and the pruning algorithm is repeated with $*$ = R . The output set of the second round constitute the final output P_N of phase #1.

Notice that user intervention is possible at the end of phase #1. Here the found circles may be adjusted, deleted, or new circles may be added to P_N .

C. Phase Two

The output of phase #1 provided an initial set of circle estimates P_N on image I_N , as shown in Fig. 1. In phase #2 images are processed in backward direction, and the search on image I_n is performed using the results P_{n+1} . Since no new mushrooms can appear backwards in time, and the size and position of existing mushrooms do not change significantly from one image to another, the search in image I_n is concentrated to positions and sizes P_{n+1} found in image I_{n+1} .

The search process for one particular mushroom is illustrated in Fig. 4. The position of this mushroom on image I_{n+1} is $P_{n+1} = (C_{n+1}, R_{n+1})$. We assume that for position $P_n = (C_n, R_n)$ the following constraints are true:

$$|C_{n+1} - C_n| \leq \Delta C, \quad (5)$$

and

$$0 \leq R_{n+1} - R_n \leq \Delta R. \quad (6)$$

Constraint (5) expresses the fact that the change of the center position of a mushroom is limited, the limit being ΔC . Notice that this change may happen because of the uneven growth of a mushroom, or a mushroom may be pushed by another touching mushroom, and estimation inaccuracies also can result a small shift.

Constraint (6) expresses the fact the radius of a mushroom increases in time. The limit of increase between two images is limited by parameter ΔR .

The core circle finding algorithm is the following:

Input: a central point C in the vicinity of C_{n+1} .

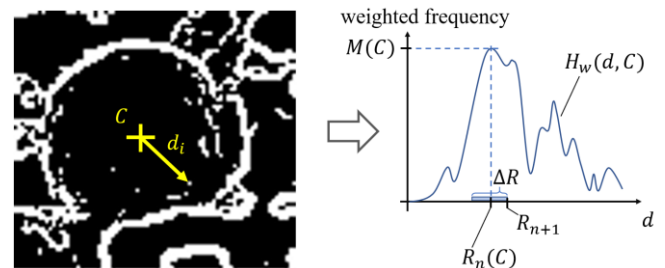


Fig. 4. Estimation of a circle's radius and quality at a given center (x,y) .

Step 1. On image I_n measure the distance of every white pixels from C . The distance of pixel i from C is d_i , as shown on the left-hand side of Fig. 4.

Step 2. Create the $H(C, d)$ histogram of distances, with bin size b equal to the required resolution (e.g. one pixel), so that $H(C, d)$ counts the number of white pixels on a circle with center C and radius between $d - b$ and $d + b$.

Step 3. Create a weighted histogram $H_w(C, d)$, as follows:

$$H_w(C, d) = \frac{H(C, d)}{2\pi d}, \quad (7)$$

which defines the ratio of the actual number and the maximum number of white pixels on the circle. $H_w(C, d)$ is illustrated on the right-hand side of Fig. 4. The closer H_w is to 1 the “more perfect” the circle is.

Step 4. Search for the maximum of H_w in the region of $R_{n+1} - \Delta R \leq d \leq R_{n+1}$. Thus the radius of the best circle with center C is estimated as follows:

$$R_n(C) = \operatorname{argmax}_{R_{n+1} - \Delta R \leq d \leq R_{n+1}} H_w(C, d), \quad (7)$$

The quality Q of circle is defined as follows:

$$M(C) = \max_{R_{n+1} - \Delta R \leq d \leq R_{n+1}} H_w(d), \quad (8)$$

The algorithm defined by steps 1-4 provides the best radius, given the center C of the circle. During the estimation process these steps are repeated starting from different values of $C = C^k$, such that $|C_{n+1} - C^k|$, so that the region around C_{n+1} is searched with the required resolution (e.g. 1 pixel). The best estimate will define the new estimated circle:

$$C_n = \operatorname{argmax}_{|C_{n+1} - C^k| \leq \Delta C} M(C^k), \quad (9)$$

and the radius is $R_n(C_n)$, using (7).

Notice that the circle finder solution used in phase #2 is not suitable for general circle finding, since it accepts any circular portion of a white are as a solution. Since the previous steps already provides a good estimate, it is used to fine-tune the solution.

IV. EXPERIMENTS

A. Image collection

The data collection device contained a Raspberry Pi 4 mini-PC with 4GB RAM and two BME280 temperature, humidity, and barometer sensor modules. The images were taken with an Intel RealSense SR305 RGBD camera and saved to a flash drive. Every day a backup was created in a private cloud. The camera provided a VGA size depth map (not used in the current processing) and a Full HD color image.

The data collection device was placed approx. 700mm above the mushroom growing bed. The illumination was provided by neutral white LEDs. Two meters of LED stripe was used with 20W total power. A 50W power supply provided the energy supply, and an IP67-protected servomotor opened and closed the camera protector cover. The data collection equipment had to withstand several environmental

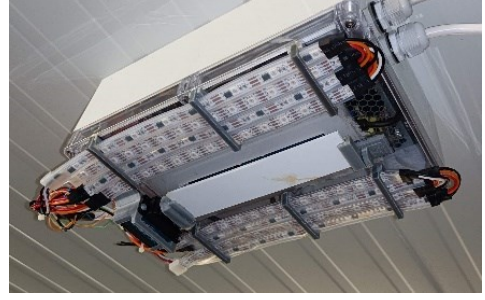


Fig. 5. Automatic image collection device mounted over the mushroom growing bed

challenges, e.g. 100% relative humidity, 70 degrees Celsius temperature, and a corrosive environment during the 48-hour-long sterilization between cultivation cycles. Fig. 5 shows the image collection device installed in the mushroom house.

B. Image Processing

An example processing is shown in Fig. 6. The pictures were taken in a mushroom cultivation facility over 60 hours. One picture was taken in each hour, producing input $I_1 \dots I_{60}$. The last picture I_{60} is shown in Figs. 6(a) and 6(b), while I_{30} and I_1 are shown in Figs. 6(c) and 6(d), respectively.

Phase #1 results are shown in Fig. 6(a) as red circles.. A few misdetections and inaccurate detections can be observed, which were manually corrected, as shown in Fig. 6(b). This was the input of phase #2. Two further frames are shown in the illustration: I_{30} was taken in the middle of the growing process, while I_1 is the first stage. The detections are shown with red circles.

The advantages of the proposed approach are clearly visible in Fig. 6(d). Here the small mushrooms are hard to detect and distinguished from the background pattern. Using detection results from later frames makes the process much more reliable.

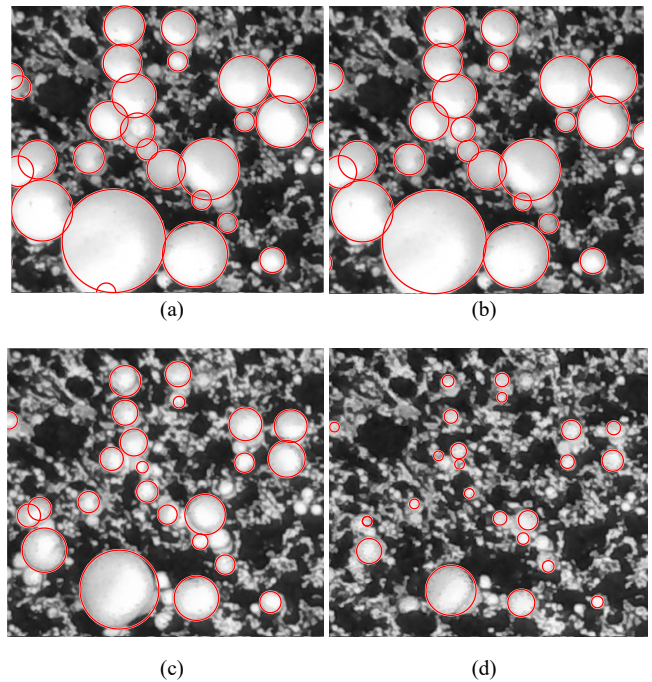


Fig. 6. Example processing. (a) input image I_{60} and phase #1 detection results for $10 \leq R \leq 60$, (b) manually corrected results, (c)-(d) phase #2 results for images I_{30} and I_1

CONCLUSIONS

In this paper a semi-automatic image processing system was proposed to detect mushrooms in a sequence of pictures taken on the plant. The proposed method contains two phases: in phase #1 the last image with the largest mushrooms is used and sufficiently large items are detected. In phase #1 classical circle-finding methods were used with specially tailored scoring and pruning mechanism. In phase #2, proceeding in backward direction in time, detection results of later images are utilized to find smaller items in the image, presumably nearly at the same position with almost the same size. In phase #2 a histogram-based detection was utilized to find items in the near vicinity of their presumed location.

Future work includes the improvement of the detection of non-circular mushrooms, e.g. using ellipsoids instead of circles, as the core shapes to search for.

The proposed method is intended to be utilized in the ground-truth generation phase of machine learning algorithms, which can provide growth models for mushroom crops.

ACKNOWLEDGMENT

Peter Galambos is a Bolyai Fellow of the Hungarian Academy of Sciences. Peter Galambos is supported by the UNKP-22-5 (Bolyai+) New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research Development and Innovation Fund.

REFERENCES

- [1] C.-H. Chen, H.-Y. Kung, and F.-J. Hwang, "Deep Learning Techniques for Agronomy Applications," *Agronomy*, vol. 9, no. 3, paper 142, Mar. 2019, doi: 10.3390/agronomy9030142.
- [2] M.D. Bah, A. Hafiane, R. Canals, "Deep learning with unsupervised data labeling for weed detection in line crops in UAV images," *Remote Sens.* vol. 10, no. 11, paper. 1690, Oct. 2018.
- [3] M. Mehra, S. Saxena, S. Sankaranarayanan, R.J. Tom, M. Veeramaniandan, "IoT based hydroponics system using deep neural networks," *Comput. Electron. Agric.* vol. 155, pp. 473–486, Dec. 2018, doi: 10.1016/j.compag.2018.10.015.
- [4] L. Zhong, L. Hu, H. Zhou, "Deep learning based multi-temporal crop classification," *Remote Sensing of Environment*, vol. 221, pp. 430-443, Feb. 2019, doi: 10.1016/j.rse.2018.11.032.
- [5] C.-P. Lu, J.-J. Liaw, T.-C. Wu, and T.-F. Hung, "Development of a Mushroom Growth Measurement System Applying Deep Learning for Image Recognition," *Agronomy*, vol. 9, no. 1, paper 32, Jan. 2019, doi: 10.3390/agronomy9010032.
- [6] T.M. Gundoshmian, S. Ardabili, C. Mako, A. Mosavi, "Modeling and optimization of the oyster mushroom growth using artificial neural network: Economic and environmental impacts," *Mathematical Biosciences and Engineering*, vol. 19, no. 10, pp. 9749-9768, July 2022, doi: 10.3934/mbe.2022453
- [7] C. Folberth, A. Baklanov, J. Balkovič, R. Skalský, N. Khabarov, M. Obersteiner, "Spatio-temporal downscaling of gridded crop model yield estimates based on machine learning," *Agricultural and Forest Meteorology*, vol. 264, pp. 1-15, Jan. 2019, doi: 10.1016/j.agrformet.2018.09.021.
- [8] D. Paudel, et al ., "Machine learning for large-scale crop yield forecasting," *Agricultural Systems*, vol. 187, paper 103016, Feb. 2021, doi: 10.1016/j.agsy.2020.103016.
- [9] M.G. Schultz et al., "Can deep learning beat numerical weather prediction," *Phil. Trans. R. Soc. A.*, vol. 379, paper 20200097, Feb. 2021, doi: 10.1098/rsta.2020.0097
- [10] A. Kolková, M. Navrátil, "Demand Forecasting in Python: Deep Learning Model Based on LSTM Architecture versus Statistical Models," *Acta Polytechnica Hungarica* vol. 18, no. 8, pp., 123/141, 2021, doi: 10.12700/APH.18.8.2021.8.7.
- [11] P. Gogas, T. Papadimitriou, "Machine Learning in Economics and Finance," *Comput Econ*, vol. 57, 1–4 (2021). <https://doi.org/10.1007/s10614-021-10094-w>
- [12] E.R. Davies, "A Modified Hough Scheme for General Circle Location," *Pattern Recognition Letters*, vol. 7, no. 1. pp. 37–43, Jan. 1988, doi: 10.1016/0167-8655(88)90042-6.
- [13] Priyanka Mukhopadhyay, Bidyut B. Chaudhuri, "A survey of Hough Transform," *Pattern Recognition*, vol. 48, no. 3, pp. 993-1010, 2015, doi: 10.1016/j.patcog.2014.08.027.
- [14] T.J. Atherton, D.J. Kerbyson, "Size invariant circle detection," *Image and Vision Computing*, vol. 17, no. 11, pp. 795-803, Sep. 1999, doi: 10.1016/S0262-8856(98)00160-7.
- [15] MathWorks: "imfindcircles. Find circles using circular Hough transform." online: <https://www.mathworks.com/help/images/ref/imfindcircles.html>. Accessed: Sept 1., 2022.

ORAL SESSION II.
Applied Informatics II.

Computerized latency data correspond with Iowa Gambling Task responses in healthy individuals

Erika Peukert
 Dept. of Psychiatry,
 St. George Univ. Teaching Hospital,
 Szekesfehervar, Hungary
peukert.erika@gmail.com

Aniko Mahler
 Pro Rekreatioe TOC,
 Gardony, Hungary
mahlerne.kofner.aniko@toparti.hu

Jozsef Halasz
 Obuda University AMK,
 Szekesfehervar, Hungary;
 Pro Rekreatioe TOC,
 Gardony, Hungary
halasz.jozsef@amk.uni-obuda.hu

Abstract— In recent years, the most important decision making task was the so-called Iowa Gambling Task, to interpret reward sensitivity and the ability to realize long-term consequences in a card selection game. Surprisingly, the importance in basic and clinical research in relation with the working of the orbitofrontal cortex did not result latency data in this very important, clinically used context. In the present paper, a computerized version of the Iowa Gambling Task was used focusing on latency issues. The aim of the present study was to describe deck-selection-latency and deck-switching data within the Iowa Gambling Task after selecting different blow-like cards throughout the test. General linear model was used to establish statistical differences within 500 responses. High and low blow card selections resulted in higher deck switching responses compared to non-blow card selections. Interestingly, increased latency was only observed in the case of high blow card selections. While the age was accompanied by significantly higher latency data, card number had the opposite effect. Our data indicate the importance of studying latency data in the interpretation of the working of prefrontal cortex via the Iowa Gambling Task.

Keywords — *beneficial strategy, computerized latency data, decision making, Iowa Gambling Task, prefrontal cortex.*

I. INTRODUCTION

The prefrontal cortex is an associative neocortical part, what makes us human, and alterations in the working of the prefrontal cortex might result in dysfunction of emotion regulation and violent behavior [1]. The prefrontal cortex is developing slowly in comparison with even other neocortical structures, the full structure might be acquired in individuals within their late twenties. Three different substructures can be described within the prefrontal cortex, the so called dorsolateral part (mainly involved in cold executive functions, like working memory), the anterior cingular part (medially, mainly involved in motivation), and the orbitofrontal prefrontal cortex (mainly involved in social behavior, warm

decision making). The famous story of Phineas Gage created famous revival in the mid nineties, when contemporary reconstruction technics were used to describe precisely injured prefrontal regions in the background of later impaired decision making and disrupted social behavior [2]. In 1994, Bechara and coworkers described the first results of the so-called Iowa Gambling Task [3], where a card selection task was used to describe the alteration of decision making in subjects with orbitofrontal damage. The task became famous, as no other task was able to map warm (emotionally involved) decision making in humans, and moreover because of its sensitivity to orbitofrontal damage (Figure 1).

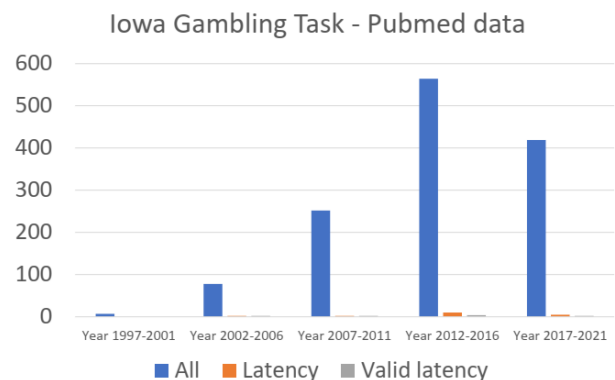


Figure 1. Sum of 5-year Iowa Gambling Task number of publications in the PubMed database (NIH). Albeit a substantial number of publications were present, the analysis of latency data was practically missing from the literature.

Early publications not only described the applicability of the task, but also served as basis for the so-called somatic marker hypothesis. Control individuals began to choose advantageously before they realized which strategy worked best, while patients with prefrontal damage continued to choose disadvantageously even after they knew the correct strategy. Patients never developed anticipatory skin conductance responses [4]. In a very interesting paper describing two adult subjects where the prefrontal lesion was

acquired before 16 months, normal basic cognitive abilities were accompanied by defective moral and social reasoning and impaired functioning in the Iowa Gambling Task resembled psychopathic traits [5]. Later the workgroup published a review paper about the concept of the somatic marker hypothesis, to address the problems of decision making encountered in patients with certain kinds of prefrontal damage and with compromised emotions [6]. In their paper, the authors describe the functional alteration of the working of the prefrontal cortex, thus it is not necessary, that the alteration of the prefrontal cortex is caused by organic lesions, but also might be the consequence of maltreatment or social segregation.

Later on the construct validity of the Iowa Gambling Task was examined, also in comparison to other tests examining other domains of the functions of the prefrontal cortex [7-8]. Mainly, executive domains were studied and it seemed that results acquired from the Iowa Gambling Task are largely independent from cold measures (variables without major emotionality, e.g. working memory). Alternative measures of deck choices were also deeply analyzed [9], but latency data were not used as target measures. Executive functions are depending on age and environmental context, and highly connected to certain psychopathologies [10-12], but affected individuals seems to have a different kind of vulnerability than in individuals with functional orbitofrontal prefrontal damage [13]. Biochemical manipulations might subtly modify the performance in the Iowa Gambling Task, but the therapeutic potential of the above operations is highly questionable at present in subjects with functional alteration of the working of the orbitofrontal prefrontal cortex [14-15].

The aim of the present study was to describe deck-selection-latency and deck-switching data within the Iowa Gambling Task after different blow-like responses throughout the test. The analysis of these additional strategical data might reveal subtle differences in populations with altered functioning of the prefrontal cortex. E.g. in children with conduct disorder, the long-term decision making is impaired, thus children with violent behavior do not only have problems with regulation of emotions, but they have serious difficulties to understand the long-term consequences of simpler and more complicated social decisions, like hitting or bullying someone [16]. Thus, new and more sensitive variables might help in the interpretation and the therapy of the children with the mentioned condition.

II. METHODS

The computerized version of the original Iowa Gambling Task [2] was used, provided by Inquisit (updated in 2013 by K. Borchert for Millisecond), translated and adapted by JH. Instead of only using classical measures, the latency and deck-switching data were analyzed. Inquisit provides highly reliable latency data, thus this Iowa Gambling Task version seemed the best solution for analyzing latency data. Additional to individual click latency data, relative latency data were collected, the so called blow-induced data. Relative post-blow

latency data (difference between the click latency of posterior and prior to the click) were also collected. Post-blow latency data were not interpreted for the first and the last click. Five Caucasian healthy adult females aged 26-36 years provided the data, with education of after-high school at least two years. The screenshot of the test start can be seen in Figure 2. 500 latency data were analyzed, and the clicks were grouped according to the blow effect. Non-blow meant zero or less than 100 USD loss by the click (Figure 3), low-blow was considered when the loss was above 100 USD but below 400 USD (Figure 4), and high-blow was considered when the loss was above 1000 USD (Figure 5). In the original task, the gain by click can be 50 USD or 100 USD. In the long run, the smaller gain will be beneficial, thus among the four available decks, the low gain decks will be beneficial (two from the four decks), as the high-gain decks (with 100 USD) will be accompanied by high losses (again, two from the four decks). Interestingly, to our best knowledge, not only the latency data were out of the scope in previous studies, but the deck-switching strategies as well.



Figure 2. Screenshot from the Inquisit version of the Iowa Gambling Task. Each decks represent different strategies, two beneficial and two non-adaptive decks. The players have 2000USD at the beginning, and have 100 draws (clicks).

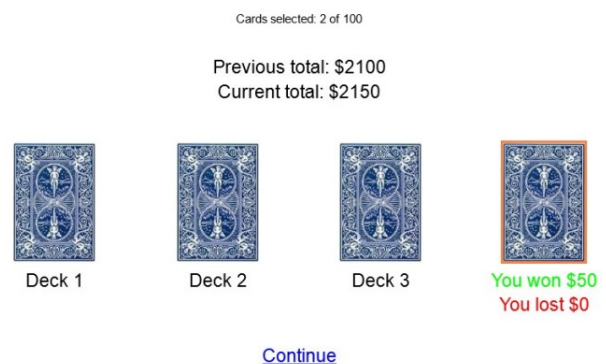


Figure 3. Screenshot from the Inquisit version of the Iowa Gambling Task, non-blow (the loss is 0 USD, only winning was present, 50 USD).

General Linear Model (GLM) was used to analyze between group differences (non-blow, low-blow, high-blow clicks) in terms of relative latencies and deck-switching strategies, while age was used as a covariant. Newman-Keuls post hoc comparisons were also run where appropriate, $p < 0.05$ was considered as significant difference. Spearman correlations were used to establish correlation between age and latency data, moreover card number and latency data.

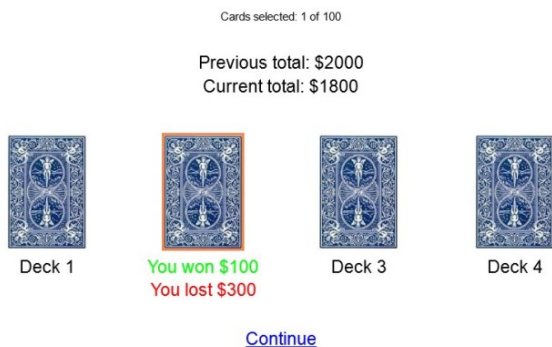


Figure 4. Screenshot from the Inquisit version of the Iowa Gambling Task, low-blow (the loss is 300 USD).

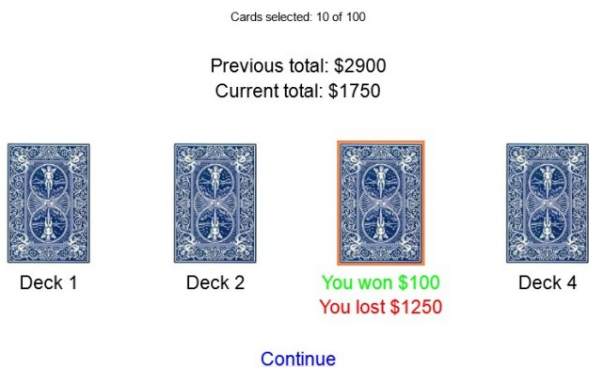


Figure 5. Screenshot from the Inquisit version of the Iowa Gambling Task, high-blow (the loss is 1250 USD).

III. RESULTS

In the deck-switching parameter, non-blow responses resulted a change in 55.7%, while a significant increase was present in low-blow and high-blow deck-switching responses ($F_{(2,497)}=10.953$, $p < 0.001$), Figure 6. Newman-Keuls post hoc comparison did not reveal significant differences between low-blow and high-blow responses.

In the relative latency parameter, there was a significant difference between groups ($F_{(2,492)}=4.401$, $p < 0.05$). While low-blow responses did not differ from non-blow responses, high-blow responses showed a significant increase to both non-blow and low-blow choices, Figure 7.

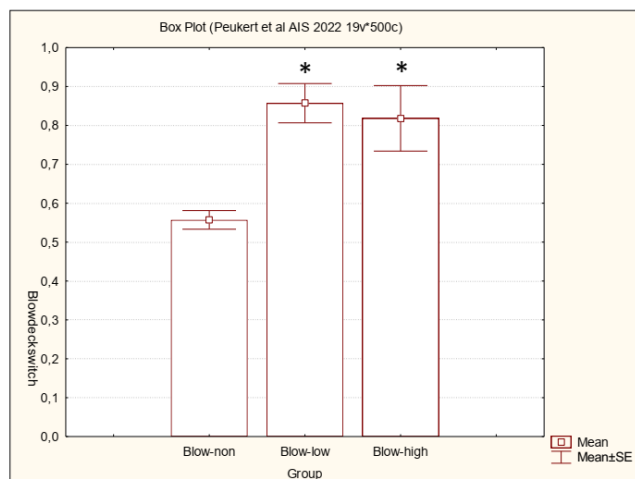


Figure 6. The probability of switching a deck after non-blow, low-blow or high-blow choice. *, significantly different ($p < 0.05$) from non-blow card-selection.

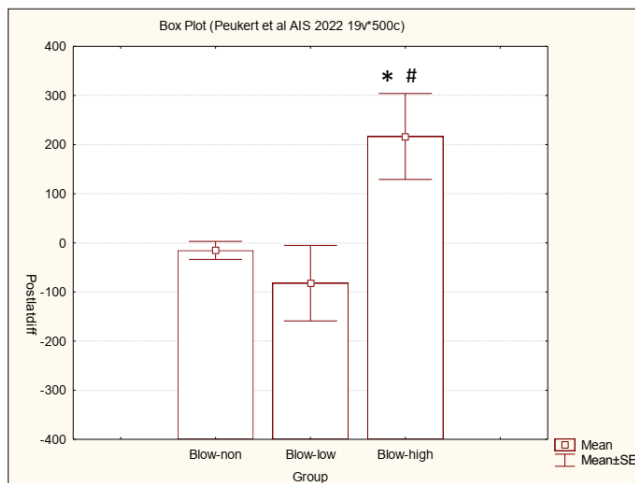


Figure 7. The difference in latencies (milliseconds) after non-blow, low-blow or high-blow choice. *, significantly different ($p < 0.05$) from non-blow card-selection; #, significantly different ($p < 0.05$) from low-blow card-selection.

Absolute post-blow latency data showed similar trend (average values for the three groups, non-blow: 969.07 ms; low-blow: 961.98 ms; high-blow: 1125.64 ms; ($F_{(2,492)}=2.293$, $p < 0.10$), Figure 8.

Spearman correlation revealed small, but significant correlation between the age and general latency data ($R=0.125$, $p < 0.01$), and a small, but significant inverse correlation was present between the selection number (number of the card in the task) and general latency data ($R=-0.126$, $p < 0.01$).

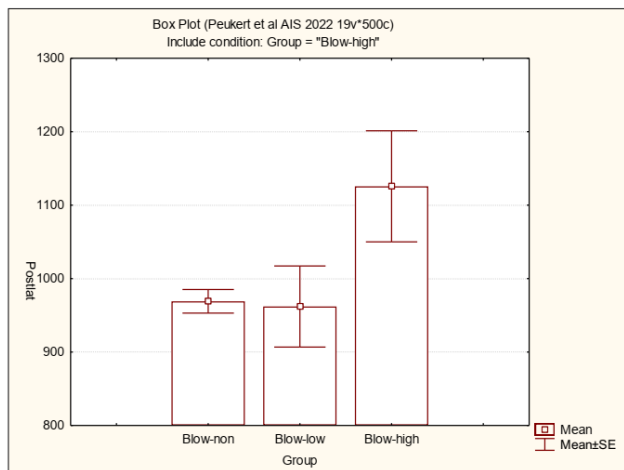


Figure 8. Latencies (milliseconds) after non-blow, low-blow or high-blow choice.

IV. DISCUSSION

The main results of the present study were the following. First, increased deck-switching strategies were observed after high-blow and even low-blow cards compared to non-blow responses. Second, increased latencies were observed compared to low-blow and non-blow selections during high-blow choices. Third, age was significantly accompanied by longer general latency data, and by the end of the test, shorter general latencies were observed.

Technically, early publications on the Iowa Gambling Task [2-6] were using physical card selections, thus measuring latencies could have been difficult to accomplish. Difficult, but not impossible, still, it did not happen. In several tasks related to executive functions, like non-classical go/no-go tests, one can be easily understand the importance of latencies [10-12]. On the stimulus level, executive functions largely depend on work-load related to certain time parameters. Classically, these tests measure performance, scoring not related to latencies. Analyzing latency data in relation of executive functions is out of the scope of the present paper, but could have high relevance in conditions like attention-deficit/hyperactivity disorder (ADHD). On the other hand, ADHD is highly comorbid condition with conduct disorder, and conduct disorder is the most important childhood condition with altered prefrontal functioning [16]. Thus, it is really surprising that no earlier paper analyzed latency data during Iowa Gambling Task in subjects with ADHD or conduct disorder, or even in control subjects. The other surprisingly overlooked Iowa Gambling Task parameter is the description of deck-switching strategies. Understanding the reason and the timing of the deck-switching might help to understand those factors what might modify preformed decisions.

Surprisingly, our defined two parameters together were able to dissect non-blow, low-blow, and high-blow responses. Interestingly, only one parameter was not enough to outline the three kind of responses. In this respect, non-blow choice is inducing a non-switcher low latency response, low-blow choice is inducing a switcher low latency response, while a

high-blow choice is accompanied by a switcher high latency response. According to interpreting high latency data, in general, high blow choices induce a longer, more thoughtful cortical response. These kind of data can only be collected in highly precise systems, and the Inquisit is excellent choice for data collection. Inquisit might be used to deliver latency-sensitive responses to the subjects, what might have therapeutic meaning in children with conduct disorder. Of course, the original Iowa Gambling Task can not be directly used for teaching sensitive choice selections. Additional sensorial measures, like heart rate variability, skin conductance responses and even the analysis of prefrontal scalp responses might help in giving choice feedback in real-life input scenarios, thus Iowa Gambling Task can only be used as model system.

The limitations of our study that only healthy adult subjects were involved in the study. As no literature data could be acquired on the measured alternative responses, this study can only be used as preliminary data collection on a valuable and important alternative measure. Definitely, skin conductance measurement would have been a valuable additive variable to the present context.

In a further step, we would like to extend the present concept to control children and children with externalization problems. Decision making ability is a sensitive variable in individuals with conduct disorder [16], and interpreting latency data in their case might add extremely important aspect in evaluating certain scenarios. This therapeutic asset only could work with computerized data collection and feedback, where arteficial barriers could enhance their abilities to interpret non-adaptive decisions, like, visual or auditory enhancement of risky pre-selection scenarios.

ACKNOWLEDGMENT

The authors declare no conflict of interest. The authors would like to thank for the participants of this preliminary study (Eszter Kalapacs, Edina Lukacs, Zsuzsanna Miklos, Gabriella Papp, and Imola Szollosi-Konsza from Pro Rekreatio TOC) for their valuable help in the data collection.

REFERENCES

- [1] R.J. Davidson, K.M. Putnam, and C.L. Larson, "Dysfunction in the neural circuitry of emotion regulation--a possible prelude to violence", *Science*, vol. 289, pp. 591–594, 2000.
- [2] H. Damasio, T. Grabowski, R. Frank, A.M. Galaburda, and A.R. Damasio, "The Return of Phineas Gage: clues about the skull of a famous patient.", *Science*, vol. 264, pp. 1102–1105, 1994.
- [3] A. Bechara, A.R. Damasio, H. Damasio, and S.W. Anderson, "Insensitivity to future consequences following damage to human prefrontal cortex.", *Cognition*, vol. 50, pp. 7–15, 1994.
- [4] A. Bechara, H. Damasio, D. Tranel, and A.R. Damasio, "Deciding advantageously before knowing the advantageous strategy.", *Science*, vol. 275, pp. 1293–1295, 1997.

- [5] S.W. Anderson, A. Bechara, H. Damasio, D. Tranel, and A.R. Damasio, "Impairment of social and moral behavior related to early damage in human prefrontal cortex.", *Nature Neuroscience*, vol. 2, pp. 1032–1037, 1999.
- [6] A. Bechara, H. Damasio, D. Tranel, and A.R. Damasio, "The Iowa Gambling Task and the somatic marker hypothesis: some questions and answers.", *TRENDS in Cognitive Sciences*, vol. 9, pp. 159–162, 2005.
- [7] M.T. Buelow and J.A. Suhr, "Construct validity of the Iowa Gambling Task.", *Neuropsychology Reviews*, vol. 19, pp. 102–114, 2009.
- [8] D.A. Gansler, M.W. Jerram, T.D. Vannorsdall, and D. Schretlen, "Does the Iowa Gambling Task measure executive function?", *Archives of Clinical Neuropsychology*, vol. 26, pp. 706–717, 2011.
- [9] D.A. Gansler, M.W. Jerram, T.D. Vannorsdall, and D. Schretlen, "Comparing alternative metrics to assess performance on the Iowa Gambling Task.", *Journal of Clinical and Experimental Neuropsychology*, vol. 33, pp. 1040–1048, 2011.
- [10] M.E. Toplak, G.B. Sorge, A. Benoit, R.F. West, and K.E. Stanovich, "Decision-making and cognitive abilities: A review of associations between Iowa Gambling Task performance, executive functions, and intelligence.", *Clinical Psychology Review*, vol. 30, pp. 562–581, 2010.
- [11] C. Lamm, P.D. Zelazo, and M.D. Lewis, "Neural correlates of cognitive control in childhood and adolescence: Disentangling the contributions of age and executive function.", *Neuropsychologia*, vol. 44, pp. 2139–2148, 2006.
- [12] M. Huizinga, C.V. Dolan, and M.W. van der Molen, "Age-related change in executive function: Developmental trends and a latent variable analysis.", *Neuropsychologia*, vol. 44, pp. 2017–2036, 2006.
- [13] S.A.K. Oberg, G.J. Christie, and M.S. Tata, "Problem gamblers exhibit reward hypersensitivity in medial prefrontal cortex.", *Neuropsychologia*, vol. 49, pp. 3768–3775, 2011.
- [14] S. Vrshek-Schallhorn, D. Wahlstrom, T. White, and M. Luciana, "The effect of acute tyrosine phenylalanine depletion on emotion-based decision-making in healthy adults.", *Pharmacology, Biochemistry and Behavior*, vol. 105, pp. 31–57, 2013.
- [15] A. Galbiati, S. Marelli, E. Giora, M. Zucconi, A. Oldani, and L. Ferini-Strambi, "Neurocognitive function in patients with idiopathic Restless Legs Syndrome before and after treatment with dopamine-agonist.", *International Journal of Psychophysiology*, vol. 95, pp. 304–309, 2015.
- [16] National Institute for Health and Care Excellence, "Antisocial behaviour and conduct disorder in children and young people. Recognition, intervention and management.", British Psychological Society and The Royal College of Psychiatrists, *National Clinical Guideline Number 158*, pp. 1–468, 2013.

Image Classification Methods to Detect the Growth of Industrial Areas

Tamas Jancso
Alba Regia Technical Faculty
Óbuda University
Szekesfehervar, Hungary
jancso.tamas@amk.uni-obuda.hu

Andrea Podor
Alba Regia Technical Faculty
Óbuda University
Szekesfehervar, Hungary
podor.andrea@amk.uni-obuda.hu

Éva Hajnal
Alba Regia Technical Faculty
Óbuda University
Szekesfehervar, Hungary
hajnal.eva@amk.uni-obuda.hu

Rezso Nagy
Eötvös Loránd University
Faculty of Science
Budapest, Hungary
nagy.rezso17@gmail.com

Abstract—The incorporation of agricultural areas is an increasing problem. During the construction, we can typically observe the expansion of industrial parks, logistics centers or the construction of leisure centers, sports facilities or buildings and halls for other purposes realized as part of state investments. In the article, we try to show the changes in the agricultural area using several image classification methods. The purpose of the presented three types of classification is to compare the values of the calculated areas and to present the possibilities of the applied software.

Keywords— remote sensing, image classification, environmental analysis, GIS

I. INTRODUCTION

The size of industrial areas within a city can be increased if the industrial area is located on the outskirts of the city. In such cases, industrial areas grow mostly at the expense of agricultural areas. In addition, municipal or state investments can also reduce the size of areas suitable for cultivation.

The sample area on the outskirts of Székesfehervár refers to such a case, where a sporting complex is being built and the expansion of the aluminum processing plant is observed at the same time. The decrease in agricultural land is clearly visible if we compare the space images taken in 2011 and 2022 (See Fig. 1 and 2).



Fig. 1. WorldView-2 RGB composite image (2011)



Fig. 2. Sentinel2 RGB composite image (2022)

If we want to find out how much the size of the industrial area has been increased at the expense of agricultural areas, we have to face several challenges [4], [5]. The WorldView-2 image has a resolution of 2 m, while the Sentinel 2 image has a resolution of 10 m. If we want to compare the two images, we have to reduce the resolution of the higher resolution WorldView-2 image to resolution of 10 m.

Another problem is that buildings in industrial areas have very diverse spectral properties, so for image classification based on training areas, samples would have to be taken in many places. Therefore, it is more appropriate to choose another solution for the separation of thematic classes.

In this article, we try to determine the increase in the size of the industrial area in three ways for three types of software. The obtained results are discussed in the conclusion chapter.

II. IMAGE CLASSIFICATION

A. Image classification in Idrisi by NDVI

The normalized difference vegetation index (NDVI) index is suitable for distinguishing the vegetation within an area.

The NDVI is calculated by the formula (1) [1].

$$NDVI = \frac{NIR-R}{NIR+R} \quad (1)$$

Since the WorldView-2 and Sentinel 2 images have red (R) and near infrared (NIR) channels, we can obtain the NDVI index image for both years. Fig. 3 and 4 show the NDVI maps.

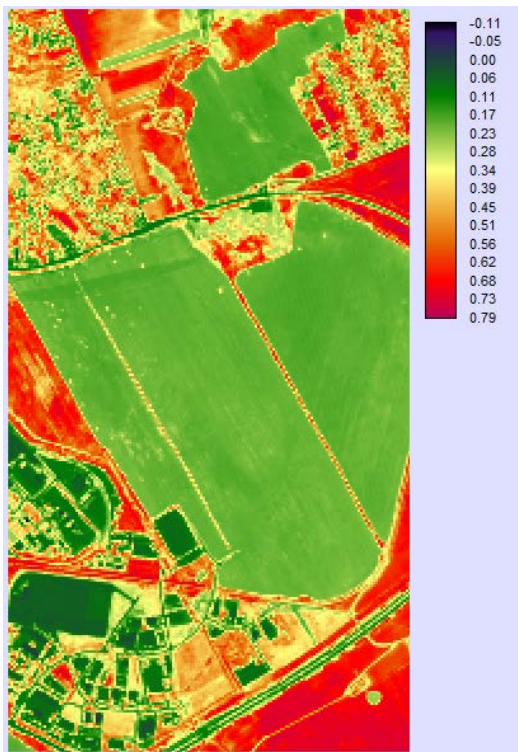


Fig. 3. WorldView-2 NDVI map (2011)

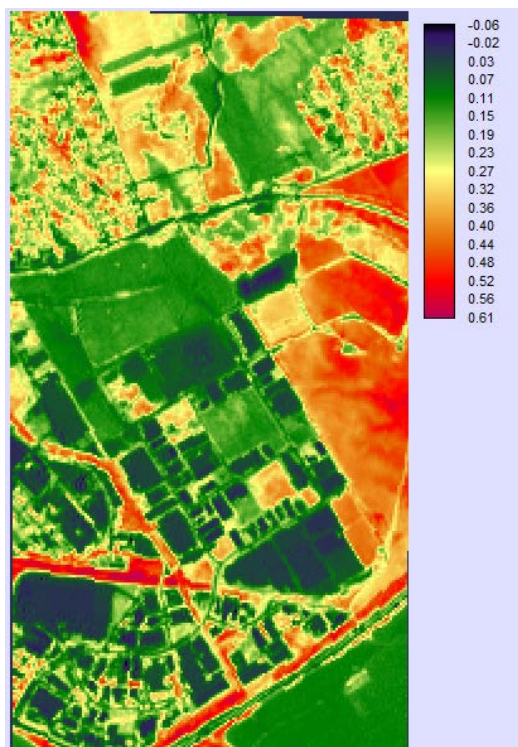


Fig. 4. Sentinel2 NDVI map (2022)

After the visual interpretation, the NDVI map can be divided into three categories. Table I shows the NDVI value limits of each category, where Min is greater than or equal to, and Max means values less than the given value.

TABLE I. NDVI LIMITS

Categories	WorldView-2		Sentinel 2	
	Min	Max	Min	Max
Built-in area	-0.11	0.14	-0.06	0.08
Soil and bare land	0.14	0.39	0.08	0.29
Vegetation	0.39	0.8	0.29	0.62

Applying the RECLASS function by Table I, we can produce thematic maps with three categories in Idirisi. Fig. 5 and 6 show the resulting maps.



Fig. 5. WorldView-2 thematic map (2011)

In Fig. 6, we can notice that red stripes have appeared at the edges of the image. The reason of this that empty areas appeared at the edges of the Sentinel 2 image field during the image transformation to the Hungarian HD72 projection. We can cut off these empty areas later.



Fig. 6. Sentinel 2 thematic map (2022)

In Fig. 5 and 6, we can see that the vegetation and soil categories show large differences. The reason for this is that the recordings were taken at the beginning and at the end of summer during different vegetation periods. If these two categories are combined, the built-up area becomes clearly visible. The result of this can be seen in Fig. 7 and 8.



Fig. 7. WorldView-2 classified image with 2 categories (2011)



Fig. 8. Sentinel 2 classified image with 2 categories (2022)

Using the AREA function, we can calculate the size of built-up areas in hectares. By subtracting the two areas, we get that the size of the built-up area increased by 39.71 hectares in 10 years.

The Idrisi software provides the possibility to fit the individual processing steps into a model. this model is shown in figure 9.

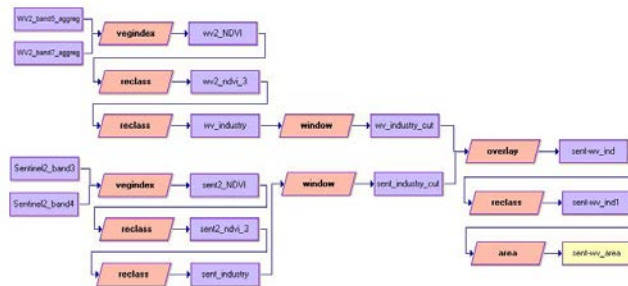


Fig. 9. Classification process built in the Idrisi macromodeler

B. Image classification in ArcGIS Pro

In ArcGIS Pro the first step of the image classification is the preparation of segmentation. Image segmentation in ArcGIS Pro is using the Mean Shift method (Comanicu- Meer, 2002). The characteristics of the image segments is based on spectral detail, spatial detail, and minimum segment size. The approach is applying the moving window technique, which calculates an average pixel value within the window, and as the window is moving it recalculates the value to make sure that all segments are correct. The outcome is a gathering of pixels in the picture into a segment of average color.



Fig. 10. The results of the segmentation process

Based on the segmented image we can implement the supervised image classification. During the classification we used the support vector machine method. In this case we also acquired 3 types of class, bare land, built-up area and vegetation.



Fig. 11. The results of classification process in ArcGIS Pro (grey-built-up area, green-vegetation, orange-bare land)

Following the classification, the images were transformed into HD72 projection, and after vectorization the areas of the different classes were calculated.

TABLE II. THE AREAL DIFFERENCES BETWEEN THE TWO IMAGES

Image	Bare land (ha)	Built-up area (ha)	Vegetation (ha)
WV 2	221,702	99,279	149,906
Sentinel 2	155,9186	116,8466	200,8336
Difference	-65,7836	17,5676	50,9296

The comparison of the result shows that the built-up area increased during the given period. The difference in the vegetation shows that the two images were taken at different vegetation period.

C. Image classification in QGIS

Image classification was made with SCP Python plugin [2] of QGIS software. For the sake of comparability, the resolution of the WorldView image was reduced to the

resolution of the Sentinel recording using the QGIS raster calculator. During the rescaling, the resolution had to be reduced for each band, and then the recording had to be stacked again. The Worldview recording was originally recorded on 8 bands, but only the red, green, blue and near-infrared bands were used. At the end of the operation, we received a version of the original image with the same resolution as the Sentinel-2 recording. In the false color method, in the last phase of preparing the images, the order of allocation of the bands was determined in such a way that the categories targeted by the study could be better distinguished: Industry and inhabited area, Vegetation and Soil and bareland. In the case of the new band allocation, we tried to make the two images as uniform as possible, i.e., the near-infrared band was placed first, the green band was placed second, and the red band was placed third. This is how we achieved that the vegetation became a bright red color, the buildings and roads were white or blue, while the soil became gray, brown or pink (Fig 12).



Fig. 12. QGIS false color image (left WorldView 2 image 2011, right Sentinel 2 image 2022)

The first task was the classification of each area into the three categories mentioned above. For both images, sample areas for all three categories were specified, based on which the artificial intelligence algorithm classified the pixels into one of the classes. When trying the algorithms of the SCP plugin for the classification, the use of the minimum distance algorithm gave the most accurate result. Increasing the number of sample areas per category could also increase the accuracy, but above four or five sample areas, the accuracy no longer increased. Classification errors can be checked visually (Fig 13), and their value can be inferred from the standard deviation obtained by different methods. After the classifications, the two classified images were compared in the SCP Land cover change function. The result was that the built-up area increased by 40.23 hectares.

The second method used was classification based on the normalized vegetation index (NDVI). Using the formula (1) presented in the previous chapter, the NDVI maps were created, and then classified pixels based on the classification table. The limit values in Table III were used for the classification, so that the lower values were greater than or equal to, while the upper values worked with a smaller condition. Value limits were determined based on the histogram of the individual rasters.

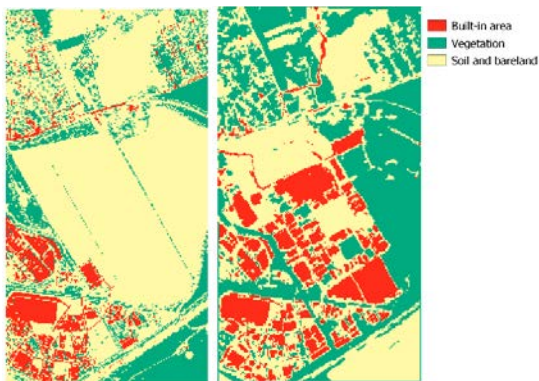


Fig. 13. Images after classification

After the classification was completed, the raster “surface area” function were ran on both classified images (Fig 14). Based on the calculations, the size of the built-up area increased by 38.06 hectares.

TABLE III. NDVI LIMITS IN QGIS SOFTWARE

Categories	WorldView-2		Sentinel 2	
	Min	Max	Min	Max
Built-in area	-0.63	0.03	-0.1	0.22
Soil and bareland	0.03	0.18	0.22	0.38
Vegetation	0.18	0.8	0.38	0.8

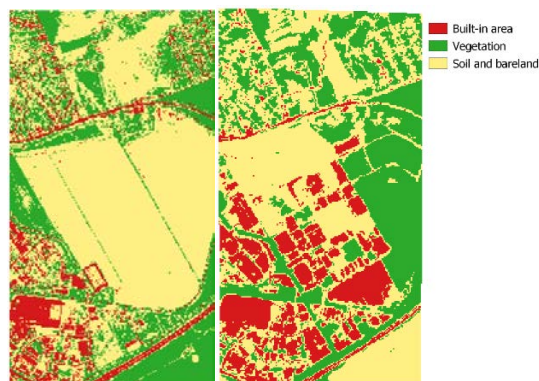


Fig. 14. Classification by NDVI. Left WorldView-2 image(2011) Right Sentinel-2 (2022).

III. CONCLUSION

The change in the size of the built-up area was determined in a given area is located on the outskirts of the Székesfehérvár city between 2011 and 2022. The task was complicated by the fact that the resolutions of the available satellite images were different. Another difficulty was that the recordings were made in different vegetation periods. On one, the agricultural area is completely bare, on the other, the same area is covered with green vegetation. In the course of the solution, several methods and software (Idrisi, ArcGIS, QGIS) were used, the results obtained are more or less consistent with one exception. Taking into account all the results obtained, the change in the built-in area was 33.9 ± 10.9 ha. If the outlier was skipped, the increase of the built-in area 39.3 ± 1.1 ha. The most accurate results were provided by NDVI-based methods. However, an interesting question is why the NDVI index limits, which can be determined based on the histogram of values, differed in the case of Idrisi and QGIS software.

ACKNOWLEDGMENT

The paper was supported by the Bilateral Chinese-Hungarian Project No. 2019-2.1.11-TÉT-2020-00171 with the project title of “Investigation of the characteristics of surface shapes in rural environment based on point clouds and remote sensing data”.

REFERENCES

- [1] pJ. R. Jensen, “Introductory Digital Image Processing: A Remote Sensing Perspective”, Pearson, 4th Edition, pp. 656, 2015
- [2] Congedo, L., “Semi-Automatic Classification Plugin: A Python tool for the download and processing of remote sensing images in QGIS”, Journal of Open Source Software, 6(64), 3172, 2021
- [3] Comanicu D., Meer P., “Mean shift: A robust approach toward feature space analysis”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, May 2002.
- [4] Verőné Wojtaszek M., Balázsik V., Jancsó T., Horoszné Gulyás M., Meng Q., “Comparison of Three Image Classification Methods in Urban Environment”, Journal od Geoscience and Environment Protection, 2015 : 3 pp. 54-59. , 6 p., 2015
- [5] Zhou X. , Verőné Wojtaszek M., Chen Ch., Jancsó T., “Urban Land Cover Mapping Based on Object Oriented Classification Using WorldView 2 Satellite Remote Sensing Images”, In: Neményi M ; Heil B (ed.), The Impact of Urbanization, Industrial and Agricultural Technologies on the Natural Environment : International Scientific Conference on Sustainable Development and Ecological Footprint, Sopron, Hungary, 2012.03.26. - 2012.03.27., Sopron: Nyugat-magyarországi Egyetem Kiadó, Paper ZVWChJ. 7 p., 2012

LEGO robotics at service of education

Péter Udvardy
 Alba Regia Technical Faculty
 Institute of Science and Software Technology
 Obuda University
 Szekesfehervar, Hungary
 udvardy.peter@amk.uni-obuda.hu

Abstract— Learning by doing seems to be the best method in education as students interact with their environment by forming it and by adapting knowledge. Interactive LEGO sets can play a great role in engineering and informatics' education from the beginning of the primary school up to the university level. This paper introduces the LEGO sets and their possible role in education, the levels, methods and tasks applied in teaching and the steps taken toward programming. The programming solutions are shown in LEGO's native programming environment and in Matlab and Simulink environment, too.

Keywords—LEGO Mindstorms EV3, robotics, education

I. THE INTERACTIVE LEGO SETS

For nearly twenty years, LEGO has been producing building sets that include not only hardware but also software, in which the built model can be programmed and controlled. The first such set appeared in the late nineties and was called LEGO Mindstorms RCX (Robotic Command eXplorers). The brain of the set was the brick, which connected to the computers via an IR port. The brick contained 3 output and 3 input ports to which sensors and motors were connected.

The next generation of LEGO Mindstorms has been available since 2006 as Mindstorms NXT. Two years later, the NXT 2.0 version was released. In the original 1.0 kit, 3 motors and 4 sensors were available (ultrasound, sound, touch and light sensors), the 2.0 version already contained more parts, and a new sensor was added to the system, a color sensor. A USB interface was added to the brick and the program included the construction of 4 robots. The brick contained 3 output and 4 input ports to which the sensors and motors were connected. A special version of Mindstorms NXT was also released, the 'Educational version', which was primarily intended for educational purposes.



Fig. 1. LEGO Mindstorms EV3 [1]

In 2013, Mindstorms EV3 was released, which has changed significantly compared to its predecessors. 4 motors can already be connected to the brick, the robot is available

via WIFI and Bluetooth, the gyro sensor is included in the kit and a new type of motor is also available. This motor is special because its axis is parallel to the longitudinal axis of the motor. The robot can also be controlled with an infrared remote control. There are descriptions of 17 robots in the set. [1]

II. THE PROGRAMMING ENVIRONMENT

Mindstorms EV3 can be programmed with its own free software, the modules required for the sensors can also be downloaded from the LEGO website. The application is available for tablets/smartphones (<https://www.lego.com/hu-hu/mindstorms/apps/ev3-programmer-app>), and the software for computers (<https://www.lego.com/hu-hu/mindstorms/downloads/download-software>). Programming is icon-based and does not require any prior training. The icons must be aligned one after the other in the correct order, and then the operating parameters must be entered within each icon. The program is built by connecting the icons. In the EV3 help menu, there is a detailed description of setting the icon parameters. You can open a new project or access previous projects on the opening screen of the software. The usual elements are also available here, as well as five 'factory' projects, complete with description, construction guide, programming aid and demonstration.

When programming Mindstorms EV3, you can place the desired instructions by placing icons one after the other. The instructions will be executed from left to right in the order of location. In all cases, the start icon must be placed from the process blocks as a starting point, and the program is closed with the stop program icon from the special blocks. By clicking on the connection point between the icons, the connection between the icons will be visible and each icon can be placed anywhere in the project. The order of execution does not change because of this, it is determined by the connections. The program can be single or multi-threaded. Most of the icons have a parameter list, which can be used to set the command mode associated with the icon. For example, the angle, time, number of revolutions, speed and power of the engine, or the detection and measurement parameters of the sensors. [2]

The Matlab (matrix laboratory) software is a mathematical system developed by MathWorks, which was developed to perform numerical calculations in the 1970s in the United States. The software, which was previously developed as an educational aid and later applied to engineering tasks, is widely used in education and in support of other, not strictly speaking, engineering work (environmental sciences, solving economic and social problems). [3]

Simulink is a new generation graphical programming and development environment related to Matlab, which can be used for modeling, simulation, and implementation. It is built with graphical block diagrams and customizable block libraries, and includes development tools such as the graphical

debugger, the model editor and analyzer, and the signal monitor. The created models can be broken down hierarchically, so that more complicated models are clearer and easier to edit. The different hierarchy levels can be broken down further within the models, so they can be broken down almost to the basic units and edited in parallel.

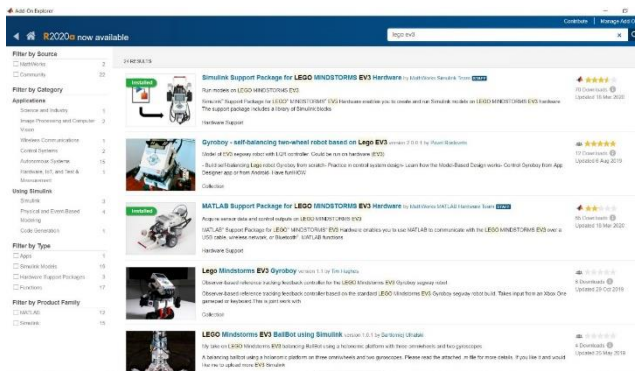


Fig. 2. LEGO Mindstorms EV3 add-ons in Matlab [3]

Hundreds of additional toolkits are available for the Matlab program system, by downloading and installing them you can solve many mathematical, engineering and other tasks. During the task solution two of these add-ons in Matlab were downloaded and installed, both from the 'official' MatWorks source. With the Matlab support package for Lego Mindstorms EV3 hardware and the Simulink support package for Lego Mindstorms EV3 hardware add-ons, Matlab can communicate with the EV3 hardware via USB, Wi-Fi and Bluetooth connections and Simulink models can be created and run with it. By installing the Simulink Support Package for LEGO MINDSTORMS EV3 Hardware additional package, the same sensors and actuators that can be found in the LEGO software become available in Matlab.

A LEGO Mindstorms EV3 (31313) programmable robotics kit was available to perform the tasks. The purpose of the construction was to build a stable, simple basic robot that can be easily assembled at any time during education, even with minimal 'Lego' practice. Among the construction instructions on the official LEGO website, the basic structure of the TRACK3R robot seemed to be the right choice for this, since the massive basic frame can be built in 10 steps, by placing the bricks on it and only minimal additions are needed to ensure mobility.

Through the Simulink interface, the computer can control the EV3 robot and instruct it to perform movements. Once we have connected our computer to the network and the Mindstorms EV3 robot is also connected to the same network, communication can be established between them. We only use the bricks and the two large engines from the LEGO Mindstorms EV3 set. The presented simple model helps to understand the potential of Matlab and Simulink in robot control during education. [4]

First, a Simulink model of the computer was created, the buttons and switches necessary to create the model from the Simulink Library browser were chosen, and then the Mindstorms EV3 control panel for modeling was created.

III. LEGO MINDSTORMS EV3 IN EDUCATION

The LEGO Mindstorms EV3 mobile robot set contains all the elements that can be used to demonstrate the structure, main parts and programming possibilities of robots at all

levels of education. When compiling the material, we progress from simple tasks ('move the robot') to more complex tasks requiring sensors and even multiple motors. The tasks on the own graphical programming interface of the LEGO Mindstorms EV3 robot set and on the Matlab/Simulink interface are both presented. Experience shows that early technical and science education can result in greater engagement among students in the long term. [5] [6]

The first task for those new to the LEGO Mindstorms EV3 robot is to move the robot. First the robot was connected to the computer via a Bluetooth connection and the linkage can be seen in the lower right corner of the graphical programming interface. The status of the motors and sensors can also be seen here (port view) and the program can also be started from here. The port view shows that in this case three sensors are connected to the Mindstorms EV3 brick, infrared sensor for port one, color and light sensor for port two, and gyro sensor for port four.

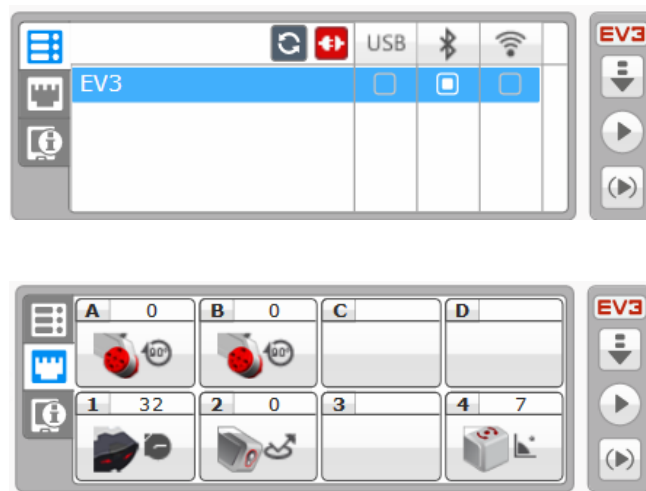


Fig. 3. Connection and port view of the EV3 set

The second task is based on what was learned during the first task and a sensor, the infrared sensor, will be used. The infrared sensor can be used to determine the distance of objects in the environment. There are two ways to use the sensor, in one the approach or distance is compared with a predetermined value, while in the other case a given distance is measured. The infrared sensor can also be used as the receiver part of the infrared transmitter (remote control) included in the kit, in which case the robot receives instructions or searches for the source of the infrared signal (beacon mode) and the degree of angular deviation from the signal source can also be set. The latter can be useful when performing more complex tasks, when, for example, the infrared remote control is part of a pallet and the forklift uses the infrared sensor to locate the pallet and turn completely in front of it to lift it.

When avoiding obstacles, the infrared sensor placed at the front of the robot detects obstacles in its range and the actuator intervenes when the preset value is reached. There are several options for avoiding obstacles, even the joint work of several sensors can help Mindstorms EV3 progress. During programming, a switch is incorporated into the program, where the state of the infrared sensor determines the true or false branch. The sensor setting can be less than, less than or equal to, equal to, greater than or equal to and greater than. In

general, we use the smaller or equal value for object avoidance, according to the movement speed of the Mindstorms EV3 robot, we specify a minimum distance of at least 20 centimeters for obstacle avoidance.

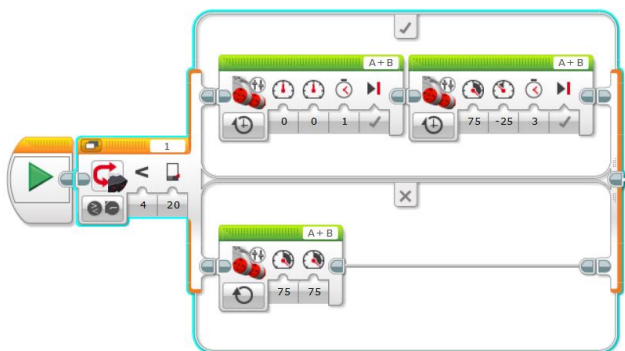


Fig. 4. Object avoidance programme

For the robot to continuously move and constantly avoid obstacles, the entire program is in a loop, so the program runs endlessly and only stops when it is switched off.

In the true and false branches, tank drive was put, in the false branch we program continuous straight forward movement, and in the true branch stop and turn movements. The time of these movements and the angle of rotation determine the dynamics of the Mindstorms EV3 robot, which can be important, for example, in the movement of the interacting robots of a swarm of humanoid mobile.

When running the program, the Mindstorms EV3 robot moves straight ahead with 75 percent engine power for both engines, and when it gets closer than 20 centimeters to an obstacle, it stops for 1 second and then turns through 3 seconds (with 75 and -25 percent engine power). This movement lasts until there is an obstacle closer than 20 centimeters in front of the robot when it continues straight, indefinitely due to the inclusion of the loop.

By completing the task, students can understand the role of the tank drive, the switch and the loop, their application possibilities, as well as the general role of sensors in switches and the application possibilities of the infrared sensor.

By building a Simulink model of the Mindstorms EV3 robot in a Simulink environment, you can program a robot that avoids objects or obstacles with the help of its infrared sensor, too. This can be achieved with several settings, in one example, the Mindstorms EV3 robot moves straight ahead, then changes direction upon detecting an obstacle within a given distance. The opposite of this can be the case when the robot is continuously moving laterally, but the direction of movement is determined by the value of the predetermined distance, above it moves to the right and below it to the left.

For the object search, constant values must be entered, based on which the constant speed of the Mindstorms EV3 robot is determined, and the distance value is set on the infrared sensor. The speed was determined as 0.25 m/s.

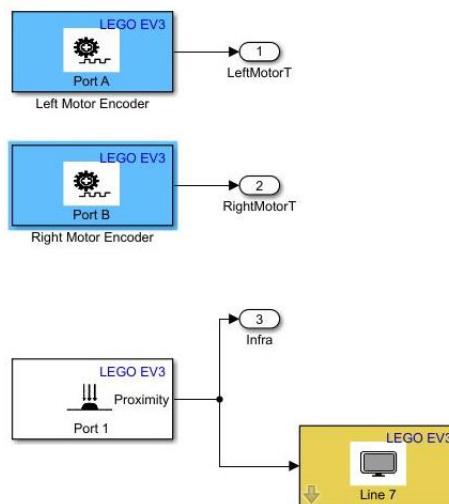


Fig. 5. Motors, sensor and display in Simulink

If the distance detected by the infrared sensor is greater than 30 centimeters, it moves to the right, otherwise to the left, so if an object gets too close, the Mindstorms EV3 robot changes direction and avoids a collision. The current distance is continuously displayed between 0 and 100 on the robot's display.

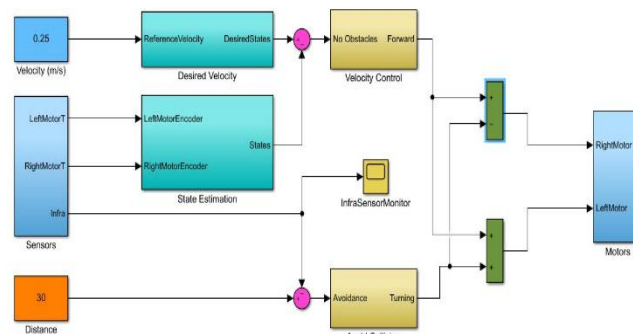


Fig. 6. Object avoidance model in Simulink

Distance measurement was also performed with the infrared sensor in a Simulink model as Mindstorms EV3 infrared sensor and its output was the result of distance measurement. During the run of the Simulink model, the robot displayed the distance in the seventh line of the brick display, and the same value appeared graphically on the scope. For the correct setting, the sampling time of the infrared sensor had to be set from the 0.1 second in the basic setting to 0.001 second, as previously the values displayed on the scope were unintelligible.

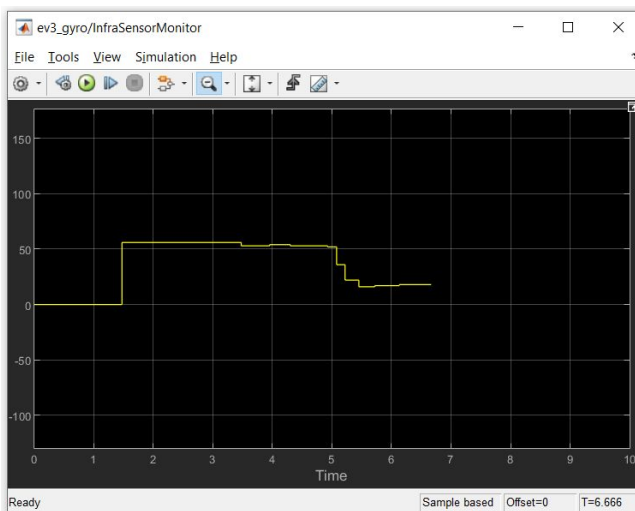


Fig. 7. Distance measurement scope

In the third task, a similar model was built to the first task, but with different sensors (color and light sensor).

Line tracking is a simple task for a human being. Looking in front of us, we follow the line with our eyes at a distance of a few meters, using our other senses as well, so we can walk along the designated path in a roughly straight line.

In comparison, the robot does not see ahead, the color and light sensor on the Mindstorms EV3 robot looks down, it only sees the one-centimeter part in front of it. It must be noted that modern robots already use other tools for spatial orientation, machine vision, cameras, distance sensors, laser scanners help with their navigation (this is called natural navigation, when the robot makes a map of the room for itself and uses this to find its way around [7]).

Industrial robots can use several methods during line tracking: they follow a fixed magnetic line or travel on a line operating on the induction principle, although the cost of building them is high and they are not mobile systems, and their change is difficult and expensive.

The color and light sensor of the Mindstorms EV3 robot can choose between eight different colors and can distinguish between dark and light. By modeling its industrial environment, Mindstorms EV3 color and light sensors enable accurate color-based object sorting.

In practice a black line is used (a path made of insulating tape, as its flexibility and width are suitable for the task) as a line (path), and line tracking can be done in two ways: the robot's program based on the shade difference between the dark and light parts, or black colored parts differ other surfaces. Generally, the latter method is more reliable.

At the start, the robot is started from the black line, the Mindstorms EV3 robot 'scans' the path forward and if the detected area below is black it moves to the right, as soon as it leaves the black area and detects a light part of the space it turns back to the left. An important aspect for accurate line tracking is the optimal choice of travel speed because if the robot turns too quickly it will permanently leave its path and will not find its way back. Due to the method described above, the robot moves forward by 'jerk' left and right.

For programming, the switch and loop used in the second task are used again, the switch is set here by the color and light

sensor. If the condition is true, the robot moves to the right, and if it is false, the robot moves to the left, the set speed is 50 percent for both motors: movement to the right is 50 and 50, while movement to the left occurs at 50 and -50 percent motor power in the steered drive setting.

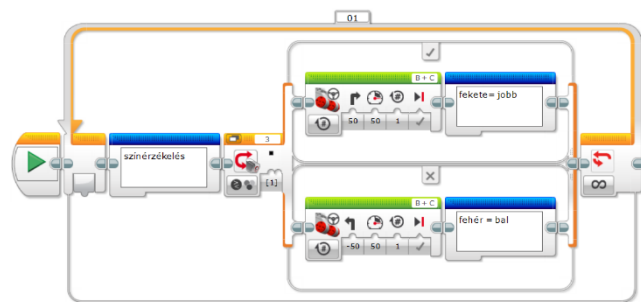
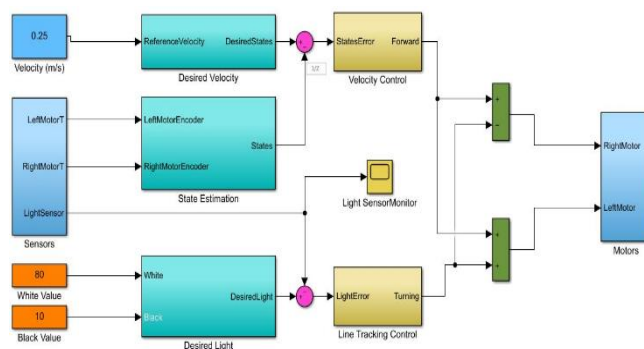


Fig. 8. Object avoidance model in Simulink

The same task is solved in the Matlab/Simulink environment as follows. The path tracking is carried out in the same environment as when using the 'own' software, so black insulating tape on a white/light basis represents the route to be followed, but in the Simulink model the dark and light parts get their own value and the program compares this during its progress. By the way, this solution would also be available on the 'own' graphical programming interface, there I chose the other method for the sake of simplicity and the presentation of the two options. For line tracking, constant speed and color values must be entered. The value of the speed was determined at 0.25 m/s, this value is still adequate for the correction in case of leaving the track.

If, based on the intensity of the reflected light, the sensor moves on the black color, the robot continues its path straight, and if the light moves more towards the light color (white color), it corrects its path by changing the two motors' travel speed and the direction of rotation of the right big motor, while the direction of rotation of the left big motor does not change. The word 'Lite' appears in the fifth line of the Mindstorms EV3 robot display and continuously shows the light intensity value during the program's run (typically between 0 = black and 75-80 = white). The Mindstorms EV3 robot does not 'jerk' with this control mode, its progress feels much more continuous than with trajectory tracking based on 'black or white' perception.



With the color and light sensor, a light intensity measurement (reflected light intensity) was also performed in a Simulink model environment. In the simple model the Mindstorms EV3 color and light sensor was included, and its output was the result of the intensity value of the reflected

light. During the run of the Simulink model, the robot displayed the light intensity in the seventh line of the brick display, and the same value appeared graphically on the scope.

Both methods are suitable for track tracking, in the case of one the progress on the line is more 'sophisticated', but the appropriately chosen speed is perhaps the most important factor in both cases. If the robot moves too fast, there is a greater chance that the Mindstorms EV3 robot will leave its track and not find its way back to the designated black line.

IV. SUMMARY

LEGO Mindstorms EV3 sets are available for schools and for universities for educational purposes. From the primary school level up to the higher educational level these sets seem to be a good tool for programming education. In this paper different programming tasks were introduced, and different tools were shown. Matlab and Simulink programming environment could be a good solution beside the LEGO's 'native' programming environment.

At the Óbuda University Alba Regia Technical Faculty these versatile LEGO sets are used at different educational levels. LEGO sets are in education in Bejczy Antal junior programming competition for the primary school students, Bejczy Antal LEGO programming contest for the secondary schools and Multiagent robotic systems on university level.

REFERENCES

- [1] www.lego.com
- [2] R. Higashi, R. Shoop: Introduction to programming LEGO Mindstorms EV3 Teacher's Guide. Carnegie Mellon Robotics Academy, 2014. <https://www.cmu.edu/roboticsacademy/PDFs/Curriculum/Intro-to-EV3/EV3-teachers-guideWEB.pdf>
- [3] www.mathworks.com
- [4] A. Behrens et al.: First Steps into Practical Engineering for Freshman Students Using MATLAB and LEGO Mindstorms Robots. : Acta Polytechnica Journal of Advanced Engineering, Volume 48, 2020, Pages 44-49, ISSN 1210-2709
- [5] V. Chaudhary et al.: An Experience Report on Teaching Programming and Computational Thinking to Elementary Level Children Using Lego Robotics Education Kit 2016 IEEE Eighth International Conference on Technology for Education (T4E), Mumbai, 2016, pp. 38-41, doi: 10.1109/T4E.2016.016.
- [6] N. Montes et al.: A novel educational platform based on matlab/simulink/lego ev3 for teaching with robots, 2018. INTED2018 proceedings, pp. 975-980. ISSN: 2340-1079
- [7]] L. Pérez, i. Rodríguez, N. Rodríguez, R. Usamentiaga, and D. F. García, "Robot Guidance Using Machine Vision Techniques in Industrial Environments: A Comparative Review," Sensors, vol. 16, p. 335, Mar. 2016.

Comparative Analysis of Land Value Capture Tools in Europe

János Katona
Óbuda University Alba Regia Technical Faculty
Székesfehérvár, Hungary
dr.katona.janos@uni-obuda.hu

Andrea Pődör
Óbuda University Alba Regia Technical Faculty
Székesfehérvár, Hungary
andrea.podor@uni-obuda.hu

Abstract— Scarcity of financial resources is a problem across Europe. As countries and local authorities emerge from the economic and financial crisis, they have fewer resources to meet all their public obligations. This is why the COST 17125 project (Public Value Capture of Increasing Property Values) was set up, with the main objective of developing a common framework for value capture. Although a substantial database is available, it is not comparable due to differences. This paper attempts to compare Land Value Capture practices in 29 countries and draw conclusions.

Keywords— value capture, urban development, land management

I. INTRODUCTION

The main aim and objective of the project nr. CA17125 is the development of a common framework for value capturing and the provision of innovative tools for public value capture based on comparative analysis to optimize the allocation of development costs and benefits as well as to disencumber the public budget [1] [2] [3].

The fulfillment of the main goal is served by three tasks:

1. Develops a common understanding of public value capture through a thorough review of existing data and definitions from different research areas and countries.
2. Identifying research gaps and ensuring conditions for future research and cooperation.
3. Establish a permanent, Europe-wide network of land management experts to work with political and administrative stakeholders to exchange knowledge on public value capture and propose policy measures to improve public value capture [4] [5] [6] [7].

Currently 36 countries joint this COST Action (33 MC Members and 3 MC Observers). Hungary was among the Inclusiveness Target Countries (ITC), so Hungarian researchers also joined the Action. A prerequisite for research is that all participants understand each term in the same way. Therefore, the research adopted modified version of the Hendricks' system as a common basis. This classification distinguishes basically recurring and non-recurring forms of public value capture. Recurring forms are further differentiated in annual payments and payments in case of sale/purchase. Non-recurring forms are further differentiated in tools focussing on one factor of value increase and tools focusing on more than one factor [8] [9].

The countries were asked to describe the current PVC situation according to a template. The research methodology is discussed in the next chapter.

II. RESEARCH METHODOLOGY

Each country participating in the research has produced a national report, which explains the following topics:

- Local authorities and planning system
- Recurring forms of public value capture (annual payments): Real estate tax for buildings; Real estate tax for lands; Municipal/Communal tax
- Recurring forms (in case of sale/purchase): Real estate transfer tax; Capital gain tax
- Non-recurring forms (focussing on one factor of value increase): Fee for construction of infrastructure
- Non-recurring forms (focussing on more than one factor of value increase): Development according General Urban Planning Legislation; Development according Special Urban Planning; Cooperative development by urban contracts; Contract models

After explaining the above themes, the percentage of Public value capture observed in the different stages of development process had to be given. The numbers of the Hungarian report are shown in Table I.

TABLE I. PERCENTAGE OF PCV DEVELOPMENT SAGES

	Public	Private	Depending
Planning	0	50	50
Reallocation	0	50	50
Land for internal infrastructure	0	20	80
Construction of internal infrastructure	0	90	10
External infrastructure	0	100	0
Economic Development	20	80	0

As PVC in Hungary is limited to recurring forms, the Public column often contains a value of 0. Unfortunately, in many cases, development agreements are not publicly available, so the percentages are only approximate.

III. RESULTS

Principally, the situation about public value capture is subject to regional variations for two main reasons: the federal structure of administrative competences and responsibilities, and the economic differences. Public value capture is particularly current where there are high property values. Different categories were therefore introduced in the tables to take account of this problem. When creating the statistics for the national chapters, the problem arose that the term public value capture is not defined in many countries and does not play a role in the public discussion. Therefore, the different instruments have not yet been assigned to the factor of the value increase. The term public value capture is not defined in many countries (e.g. Hungary), therefore, the different instruments have not yet been assigned to the factor of the value increase.

TABLE II. PUBLIC VALUE CAPTURE INSTRUMENTS IN THE COUNTRIES

	Recurring forms		Non-recurring forms	
	Annual payments	In case of sale	Focusing on one factor	Focusing on more than one factor
Austria	4	4	4	2
Belgium	1; 2; 4	1; 4	1; 2	3
Bosnia	4; 1	1; 2	4	3; 1
Bulgaria	4	4	2	2
Croatia	1; 2; 4	1; 4	1; 2	0
Czechia	4	0	3	1
Estonia	4	4	3	3
Finland	4	4	1; 4	1; 4
France	1	1; 4	4	4; 2
Germany	4	4	4	3; 4
Greece	4	1; 4	1	3; 4
Hungary	3	4	1; 4	0
Israel	1	4	4	4
Italy	3	3; 4	0	4
Latvia	4	4	4; 2; 3	0
Lithuania	2	3	2	3
Malta	0	4	4	2
North-Macedonia	4	4	4	2
Norway	3; 4	4	3; 4	0
Poland	4	4	2	1
Portugal	4	4	2	1
Serbia	4	4	4	1
Slovakia	4	4	3; 4	1; 3
Slovenia	4	4	4	2
Sweden	4	4	1; 4	3; 4
Switzerland	4	3; 4	4	3; 4
Turkey	4	4	3; 4; 4; 2	3
Ukraine	4	4	0	0
United Kingdom	4	4	4	3; 4

Explanation: 0 – no instrument, 1 – instrument is used limited, 2- instrument is used rarely, 3 – instrument is used regular one part of the country, 4- instrument used regularly throughout the country.

The recurrent LVC distinguishes from non-recurrent LVC in who the payers are. Recurring LVC is mainly a tax for individuals, but the non-recurrent LVC is foremost paid by professionals. This difference is most likely the main reason why the recurring tax has a social profile in so many of the sample countries. As Table II. shows, the use of these tools varies widely between countries.

A. *Recurring forms of public value capture*

The non-recurrent land value instrument usually rests on increase in land value affected by specific interventions (e.g. investments, rezoning, delivery of planning permissions). Recurrent capture of value in land is contrary to this: it refers to systems that capture value partly indifferent to any increase in value and always independent on any explicit interventions that affect the value of the land. The recurrent instruments rest on different motivations.

Recurrent tools usually do not distinguish between the value of buildings from the value of the land. Recurrent value capture has significant similarities to taxes. Except from Croatia, Malta, all countries apply property tax based on annual payments in by any means possible. A few countries make a distinction between land and the buildings, although most countries tax the property including land and building.

Recurrent tools usually do not distinguish between the value of buildings from the value of the land. Recurrent value capture has significant similarities to taxes. Except from Croatia, Malta, all countries apply property tax based on annual payments in by any means possible. A few countries make a distinction between land and the buildings, although most countries tax the property including land and building.

The property tax rate varies significantly between the countries. The Baltic countries charge the highest rate (Estonia 2.5%, Latvia 3%; Lithuania 4%). The rate of recurring property tax is dependent on annual payments according to the use of the property, the value, and its location.

Recurring income from property is revenue to the owner from capital that the owner has invested in property. The lower tax rate, 10% is in Bosnia Herzegovina. Countries with higher tax rates include Czechia (15%), Norway (22%), Italy (23-26%) and Israel (25%).

All countries in the survey except Croatia have a tax on the sale of property. That type of tax is the lowest in Malta (8%) and the highest in North Macedonia (70%). In North Macedonia, the tax applies only to households that have lived in the house for less than 3 years prior to the sale; in Austria, no tax is levied if the seller has lived in the house for more than 8 years. Germany, on the other hand, imposes a higher tax rate if the seller has sold more than three properties within the last five years. Obviously, countries design taxes to protect individual one-time sellers and tax professional dealers.

Transaction tax is also a commonly used type of tax. This tax has abolished in Czechia during the Covid-19 crisis, apparently to stimulate the markets. There are countries where transaction taxation is uniform. This group includes Bosnia Herzegovina, Croatia, Estonia, Finland, France, Greece, Norway, Poland, Slovenia, Turkey, and Hungary. Median rate is just above 2% of the purchase price. At the

lowest end is Ukraine with 1%, and at the higher end is Bosnia Herzegovina with 5%. The other group includes countries with differentiated rates of transaction tax. Austria, the United Kingdom, and Italy have a progressive tax rate. The tax rate is 1-2% for average properties, and 3-9% for luxury properties. Some countries also distinguish between kinds of buyers and the location of the property. Serbia relieves both first-time buyers and buyers of modest houses from transaction tax. Sweden differentiates between natural persons and companies, charging 1.5% of the sales sum to buyers who are natural persons and 4.25% for business buyers.

B. Non-recurring forms of public value capture

The recently introduced new urban agenda (UN Habitat, 2022), which emphasises the link between sustainable urbanisation and a better quality of life, will require higher revenues for public spending. There is therefore a strong need for knowledge exchange on the most effective PVC tools that can be applied in European countries.

It is a common non-recurring form that local authorities request the developer/owner to provide in-kind road and service infrastructure within the development project boundary. However, for smaller scale developments, local authorities may ask for in-kind contributions, for example to build a playground or small public spaces outside the development boundary.

PCV is already present in elements of the planning process, such as: land use change, general/spatial/detail plan, planning permission. The development will include the construction of roads and social infrastructure (e.g. schools, public open spaces etc.) with the support of PVC tools. The local authority may require the developer/owner to make an in-kind contribution to the provision of roads and services within and outside the development project. Fig. 1 shows PVC's role in the developing process.

The PVC increases due to the internal infrastructure factor (Bosnia Herzegovina, Bulgaria, Croatia, Germany, Italy, Latvia, Lithuania, North Macedonia, Serbia, Slovakia, Slovenia, Sweden, Switzerland, Turkey and the UK) in the majority of countries. In some countries, in particular Finland, France, Germany, Israel, Italy, Finland, France, Germany, Israel, Serbia, Sweden, Switzerland, Israel, and the UK, developers/owners may be asked to contribute to external infrastructure.

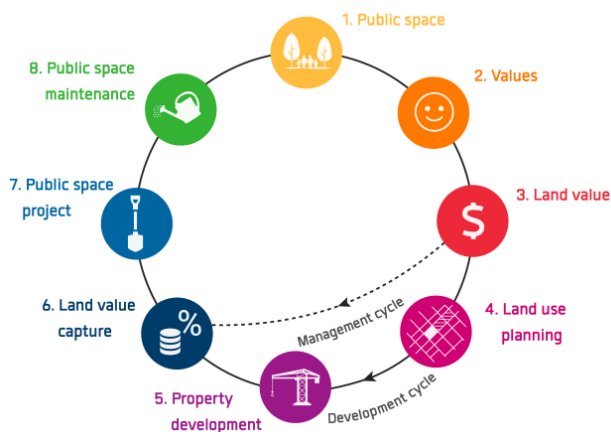


Fig. 1. The role of PVC in developing process [10]

It is important to note that the most significant PVC instrument is the developer obligation, which has been adopted by the majority of the countries analysed, in particular Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Estonia, Finland, Germany, Hungary, Israel, Italy, Latvia, Lithuania, Malta, the Netherlands, Poland, Hungary, Malta, Slovakia, Slovenia, the Netherlands, Sweden, Switzerland, Norway, Northern Macedonia, and the United Kingdom. The Fig. 2 shows the practice of PVC in Israel.

Developer contributions in kind have a long tradition in many countries. Today, this instrument has been extended to the payment of contributions in kind or levies. This is how owners/developers contribute to social infrastructure, affordable housing, and community spaces. Except for United Kingdom, all the above-mentioned countries use some form of developer contribution. Developer contributions in the UK can cover contributions to affordable housing and other infrastructure (such as education, health, transport, flood and water management, green and digital infrastructure).

However, there are countries where the developer/landowner benefits fully from the increase in land value due to the planning process and there is no PVC to finance community functions based on this factor. These countries include: as Austria, Bulgaria, Croatia, Czechia and Slovakia. In most of the above countries, the local authority is responsible for providing the technical infrastructure inside and outside the development project, and the developer/landowner is obliged to pay a fee for this. In Austria, there have been extensive discussions on the need to adopt a policy of planning profit set-aside in the future in order to support the provision of affordable housing through the regional land fund. In the Czech Republic, the local authority charges the owner of a developable plot a fee for connection to an existing water or sewerage network. The expectation is that roads within the development project and connections to the existing road network are the responsibility of the developer/landowner.

The land readjustment can be an effective instrument for PVC. In Germany, Spain and Israel, where the compulsory land readjustment procedure has a long tradition, this land management tool is successfully used to ensure efficient PVC. Finland, Sweden, France, Belgium, however, do not necessarily apply the land readjustment instrument for PVC, but in fact they apply the land management instrument well in their land use planning policies.

Most countries have created non-recurring forms of PVC that can be considered as developer obligations. The developer contributes in kind or by paying a fee to public development in exchange for a decision that enables the development process. Developer obligations for developer contributions are usually based on negotiations between the local authority and the developer. The agreed obligations are usually agreed before planning permission is granted.

The developer obligations and the conditions when and how they are agreed vary from county to county, but some generalities can be made. According to the analysis, development obligations can serve up to three purposes:

1) to capture the value of land, as the benefits derived from land are "unearned" (a link to the concept of land rent);

2) to finance public costs (provision of technical and social infrastructure, including playgrounds, open spaces, affordable housing, etc.);

3) to finance public budgets in a more general sense. It should be noted that the same objectives can also be linked to recurrent forms.

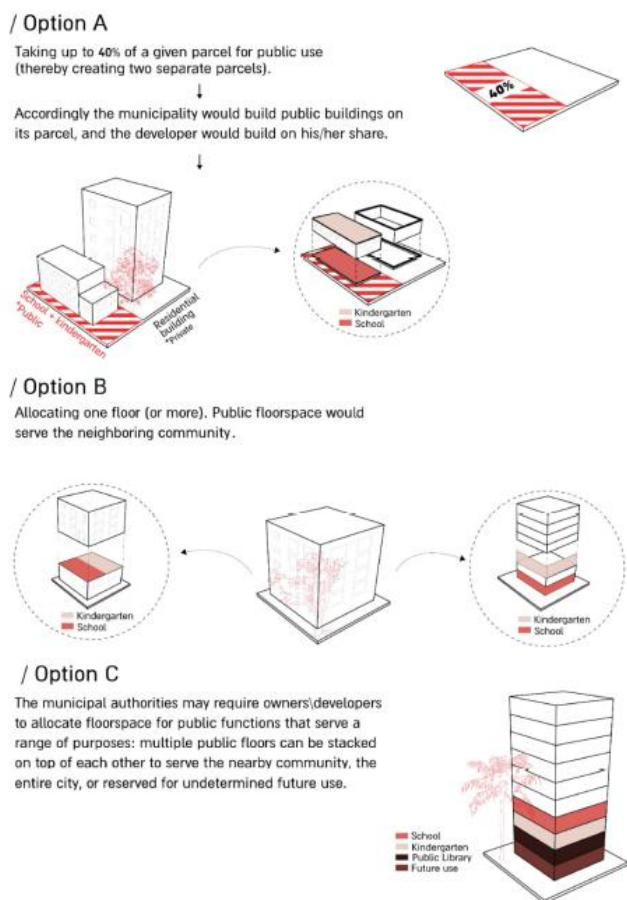


Fig. 2. Pathways for the vertical allocation of public floor space in Israel [11]

In Italy, building contributions are proportional to construction costs and are set at regional level. Urbanisation charges are levied to finance the costs incurred by the municipality for new urbanisation. Austria, the Czech Republic, Germany, Israel, Slovakia, Sweden, and Switzerland apply a developer/owner levy (fee) for the construction of technical infrastructure within the project in addition to the agreed developer obligations for technical and/or social infrastructure outside the development project. In United Kingdom, developers make in-kind contributions to the technical and social infrastructure within the development project, in addition to the agreed developer obligations, which are either in-kind contributions or a levy (charge) on the technical and/or social infrastructure outside the project. Malta applies the infrastructure service contribution and developer obligations associated with the development project, which would contribute to large private sector projects. France applies a development tax/m² to finance the infrastructure of the project, in addition to the social mix obligations. Belgium applies planning conditions for the developer to contribute in kind to the infrastructure to be built in the project, and planning fees to compensate for the impact of the development. Poland applies an

infrastructure charge to the development project and an obligation on the developer to upgrade roads in order to obtain planning permission. Slovenia only applies a charge for utilities and Latvia applies a developer obligation for infrastructure outside the development project.

There are basically two approaches to developer obligations. One group includes value capture assets directly related to the development project, and the other includes major revenues that can be used to finance the public budget. Developer obligations based on a value-based approach are used in Germany and the UK, where a development value assessment by an independent property valuer forms the basis for negotiations between the developer and the local authority. The value established for the development valuation is calculated using the comparative valuation method in Germany and the residual valuation method in the UK. The negotiated approach to builders' obligations is very much applicable in Switzerland, but the negotiations are facilitated by lawyers rather than planners and surveyors.

There is an alternative instrument to obligations - the builder's levy, which is applicable in most of the countries analysed. The levy is usually calculated using a method adopted by the local authority, based on a fee per m², per unit and/or per overall size of the development.

The countries analysed can be divided into two groups according to the responsibility for financing technical infrastructure within a development project.

Group 1: Austria, Belgium, Czech Republic, Estonia, Finland, Germany, Hungary, Israel, Norway, Sweden, Switzerland, Latvia, Switzerland. In these countries under the approved local plan, the landowner is responsible for financing the technical infrastructure.

Group 2: Bosnia and Herzegovina, Bulgaria, Croatia, France, Greece, Italy, Lithuania, Malta, Northern Macedonia, Poland, Serbia, Slovakia, Slovenia, Turkey, United Kingdom. In these countries the developer becomes the landowner and is responsible for providing the technical infrastructure in kind or paying a fee.

IV. SUMMARY

The tools of non-recurring forms varies from country to country, but generally similar variations of forms can be found across participating countries. This occurrence heavily depends on existing political-legal-Institutional framework, which includes property rights, planning and development policies and practice in each country.

It can be generally stated that both the recurrent and the non-recurrent value capture tools contribute significantly to the social welfare. Quite a lot of the countries in this sample expect developers to finance new and upgraded infrastructure and most of the countries also require the developers to contribute to the financing of technical, green, and grey infrastructure outside of the development project.

The idea of unearned value increase is not accepted in all countries, because few countries use non-recurrent instrument that can extract developer contributions to more than one factor. In addition, relatively few countries apply mandatory land readjustment instruments or temporarily public landownership. Both public land readjustment and temporarily public landownership are instruments well suited for extracting land for very different kinds of goods –

like land for greens and kindergartens, and affordable houses, and technical infrastructure. This appears as an area to be developed in these countries.

The sample shows that, value extracted from increased land value is very limited degree target social aspects of urban life. Analysis of recurrent land value capture is more nuanced, but the property tax and the transaction tax types might be further differentiated for fair and equitable contribution.

REFERENCES

- [1] Hendricks, A. et. al., “Public value capture of increasing property values – What are unearned increments?” In: Erwin Hepperle, Rober Dixon-Gough, Reinfried Mansberger, Jenny Paulsson, Jozef Hernik and Thomas Kalbro (eds.), *Land Ownership and Land Use*
- [2] János Katona - Andrea Pődör, Increasing urban development efficiency through land value capture mechanisms In: Csuka Ildikó; Simon Gyula (szerk.) *AIS 2021-16th International Symposium on Applied Informatics and Related Areas – Proceedings Konferencia helye, ideje: Székesfehérvár, Magyarország 2021.11.11. Székesfehérvár: Óbudai Egyetem, pp 97-101 (2021)*
- [3] Botticini, F.; Auzins, A.; Lacoere, P.; Lewis, O.; Tiboni, M. *Land Take and Value Capture: Towards More Efficient Land Use. Sustainability* 2022, 14, 778. <https://doi.org/10.3390/su14020778>
- [4] Marquard, E.; Bartke, S.; Font, J.G.I.; Humer, A.; Jonkman, A.; Jürgenson, E.; Marot, N.; Poelmans, L.; Repe, B.; Rybski, R.; et al. *Land Consumption and Land Take: Enhancing Conceptual Clarity for Evaluating Spatial Governance in the EU Context. Sustainability* 2020, 12, 8269.
- [5] Hendricks, A. *Public Value Capture—An Opportunity to Improve the Economic Situation of African Municipalities. In Responsible and Smart Land Management Interventions: An African Context; de Vries, W.T., Bugri, J.T., Mandhu, F., Eds.; CRC Press–Taylor&Francis Group: Boca Raton, FL, USA, 2020; pp. 251–262*
- [6] Hendricks, A.; Kalbro, T.; Lorente, M.; Vilmin, T.; Weitkamp, A. *Public Value Capture of Increasing Property Values—What are the “Unearned Increments”?* *Eur. J. Sustain. Dev. Res.* 2017, 2, 34.
- [7] Alterman, Rachelle, *Land Use Regulations and Property Values: The 'Windfalls Capture' Idea Revisited (2012). Chapter in: 'The Oxford Handbook of Urban Economics and Planning' (Nancy Brooks, Kieran Donaghy, and Gerrit-Jan Knaap, eds.) pp. 755-786, Oxford University Press, 2012, Available at SSRN: <https://ssrn.com/abstract=2309571>*
- [8] Balázsik V., Tóth Z., Abdurahmanov I. *Analysis of Data Acquisition Accuracy with UAV INTERNATIONAL JOURNAL OF GEOINFORMATICS (1686-6576)*: 17 1 pp 1-10 (2021)
- [9] Kugler, Zsófia ; Tóth, Zoltán ; Szalay, Zsuzsa ; Szagri, Dóra ; Barsi, Árpád, *Supporting microclimate modelling with 3D UAS data acquisition QUARTERLY JOURNAL OF THE HUNGARIAN METEOROLOGICAL SERVICE* 123 : 3 pp. 279-294. , 16 p. (2019)
- [10] Alexander Stähle: *Public-Space-Making, Searching for Best Practice from Planning to Management 191018, Spacescape, Östgötagan 100, 116 92 Stockholm, 47 p.*
- [11] Nir Mualam, Andreas Hendricks, Vida Maliene, Eyal Salinger, *Value Capture and Vertical Allocations of Public Amenities, Academic Editor: Miguel Amado, Sustainability* 2021, 13(7), 3952

POSTER SESSION I.
Industrial and Embedded Systems

Design and Requirement Analysis of an Outdoor Autonomous Mobile Manipulator

Zoltán Szilágyi
Óbuda University
Alba Regia Technical Faculty

Székesfehérvár, Hungary
szilagyi.zoltan@amk.uni-obuda.hu

Csaba Hajdu
Széchenyi István University
Department of Informatics

Győr, Hungary
hajdu.csaba@ga.sze.hu

Abstract— Industrial Mobile Robots (IMRs) gained prominence in research and technical applications. Besides their original interior application environment, the expanding demand for such devices has recently led to their growing application in exterior environments. In general, the external setting challenges robots with higher complexity and variety in the sense of tasks and the target environment.

This article presents the proposal of a mobile manipulator robot suitable for external applications like industrial and urban settings. The robot achieves flexible manipulation freedom with two manipulator arms and a revolving base beneficial in solving a wide variety of complex tasks. The related research process included additional characteristic requirements for similar external devices. These requirements include IP protection, rough ground mobility, acceptable movement range, and autonomous self-monitoring and control capabilities. The research further focuses on analyzing system-wide capabilities with the help of computer simulations and virtual test environments. Besides developing a novel robotic device, the further goal is to survey the potentials and challenges of external robotic applications and to summarize the gained technical knowledge and experience.

Keywords—mobile robotics, IMR, mobile manipulator

I. INTRODUCTION

The theme of industrial robots has been a paramount research field and market since the 60s. Traditionally, arm-like kinematic structures – robotic manipulators - were sufficient in automatizing many industrial applications. However, parallel with the optimization of manipulator solutions, autonomous mobile robots gained individual attention in the industry to automatize material handling in industrial departments.

The field of Industrial Mobile Robots (IMR) is a currently prominent subfield regarding its market size and research opportunities. Mobile robots are perfect candidates to automatize tasks in an industrial setting, like material handling and other collaboration tasks with human actors. For example, the growing Automatic Ground Vehicles (AGV) market focuses on material handling between factory departments, with only a minimal extension of existing infrastructure. Furthermore, new regulations and standards are applied to the robots to overcome possible safety and reliability issues.

Overall, most interior industrial tasks can be addressed with currently available solutions. On the other hand, some tasks that require precise manipulation capability or fine movement still need to be addressed by most market-ready mobile robots. An example is when a robot moves in a workshop setting handling equipment.

The market typically omits exterior industrial tasks for a good reason. The internal industrial environment is usually more controllable, requiring minimal preparations inside the departments to install the robot. External settings are much more stochastic, with the minimal capability to change and equip the infrastructure of the operating environment. Also, exterior robots require a more rugged structure to resist weather and terrain effects (e.g., vibration, rough ground, rain, and electrical disturbances). The exterior setting usually requires high manipulability (e.g., with a robotic arm) as specific tasks can be complex and tedious pick-and-place scenarios. Furthermore, subtasks can vary even during the execution of the task, such as in an agricultural setting. Many industrial fields await flexible automatization of monotonous yet complex tasks. The development of a multi-purpose highly maneuverable with refined manipulator capabilities could be beneficial in many application areas such as agricultural, urban, mining, and the traditional military applications of mobile robots.

This paper presents the design of an autonomous mobile robot equipped with dual robot manipulators to overcome the actuation challenges of exterior industrial use cases. In the presented attempt, the primary goal was to combine the precise capabilities of a robotic manipulator with the endurance and mobility of a mobile robot platform suited for exterior use. The robot is currently in the design phase, with organized related requirements and concepts to provide rugged, highly maneuverable, and swift manipulation capabilities. As the design is unconventional, the current scope of the paper is exploratory research organizing the conceptual design and posed requirements. In addition, the current paper also attempts to provide a flexible software architecture inspired by currently prominent software frameworks for autonomous vehicles and distributed robotic system paradigms.

II. RELATED WORK

Traditionally, scholars tend to differentiate between mobile and industrial robots, categorizing their usage and capabilities. Industrial robot manipulators are regulated under different standards for example, in ANSI/RIA R15.06 and ISO 15066:2016 (more safety-focused standard). Apparently, due to the increasing popularity of mobile robots in an industrial context, a similar standardization process has been recently started to define proper and standard requirements for mobile robots. As different approaches and solutions appeared, manufacturers started to designate robots under varying terminology, with no precise term scope (popular terms are listed in TABLE I.). Regarding mobile robots in the industry, manufacturers typically differentiate between the following two categories:

- **Autonomous mobile robots (AMRs):** are expected to handle complex and unexpected situations in a low infrastructure setup, with vast capabilities of human interactions and a relatively high degree of autonomous operation. Although, their maximum payload and speed are typically limited for safety reasons.
- **Automatic Ground Vehicles (AGVs):** these robots are expected to handle only a limited set of tasks (e.g., material handling) in a built-up support infrastructure setting. Their human interaction capabilities are limited and not expected to resolve complex interferences autonomously.
- **Industrial Mobile Robot (IMRs):** a mobile robot targeted explicitly for usage in an industrial environment.

TABLE I. OVERVIEW OF POPULAR ABBREVIATIONS IN ROBOTIC INDUSTRY

Abbreviation	Term
AGV	Automatic Guided Vehicles
AMP	Autonomous Mobile Platform
AMR	Autonomous Mobile Robot
IMR	Industrial Mobile Robot
UGV	Unmanned Guided Vehicle
UAV	Unmanned Aerial Vehicle

Mobile robots equipped with a robot manipulator are typically referred to as mobile manipulators. These robots are expected to carry out simple pick-and-place tasks and similar manipulation tasks. The requirement for manipulation is prominent in the traditional use of mobile robots in military and surveillance applications.

Due to the growing market of autonomous mobile robots in industrial applications, new standards targeting IMRs appeared. A relatively new standard, ANSI/RIA R15.08-2020 defines a broad terminology for IMRs:

- **IMR Type A:** autonomous mobile robot, with only an autonomous mobile platform (essential capability to move and navigate through free space).
- **IMR Type B:** a base platform with any (passive or active) attachment that is not a manipulator, beyond the capability to navigate autonomously.
- **IMR Type C:** a mobile manipulator qualifying to R15.06 as an attachment on an IMR Type A mobile platform or an AGV as the base.

Exterior applications have been under massive development in recent years. Traditionally, mobile robots are applied in military and academic context with special tasks and highly qualified support. However, as mobile robots became widespread, numerous industrial applications applied robots for automatization. In this article, the agricultural context is further investigated. Some authors envisioned that mobile robots and unmanned aerial vehicles (UAVs) could aid in agricultural tasks [1], which are demanding for agricultural workers and monotonous. Promising results have been achieved in the cultivation of crops, vineyards, and crop

monitoring [2]. As a result, many manufacturers started to provide robotized solutions for agricultural tasks.

The software system of the robot must be flexible in regards of extendibility and adjustability, implying the capabilities can be adjusted as components and the refining of components can be changed during a mission. A general control loop model for autonomous mobile robots popularized by Siegwart et al. [3], indicates mandatory tasks for autonomous operation. The model is derived from the widespread sense-think-act cognitive model [4]. In this model, the robot software is organized into a set of components each associated with sensing (i.e., proprioceptive/exteroceptive perception, processing), thinking (i.e., planning, inferring) and actuation (i.e., control of motors, manipulators) each bound to a valid input/output interface closing the control the loop. This model does not rigidly define each subtask but rather provides a starting framework for robot software design. The control loop, with some addition to mobile robot manipulators in detail, is depicted in Fig. 1, with additional elements inspired by the field of Cognitive Infocommunications (CogInfoCom) [5].

Nowadays, new software systems tend to be organized into cooperating components (i.e., microservices) besides robotic systems. In robotic systems, this approach is beneficial and evident regarding the development and verification subtasks. For example, a trajectory planner can be developed and tested as a separate component and then integrated and refined in the target robot's context. This approach is currently favored in the design of cyber-physical systems (CPS), providing numerous communication frameworks (e.g., Data Distribution Service- DDS [6]). In robotic systems, Robot Operating System 2 (ROS2) [7] [8] started to gain high prominence and support from academic and industrial actors alike. It should be noted that ROS 2 uses DDS as an intermediate framework to achieve automotive- and military-grade stability and reliability [9].

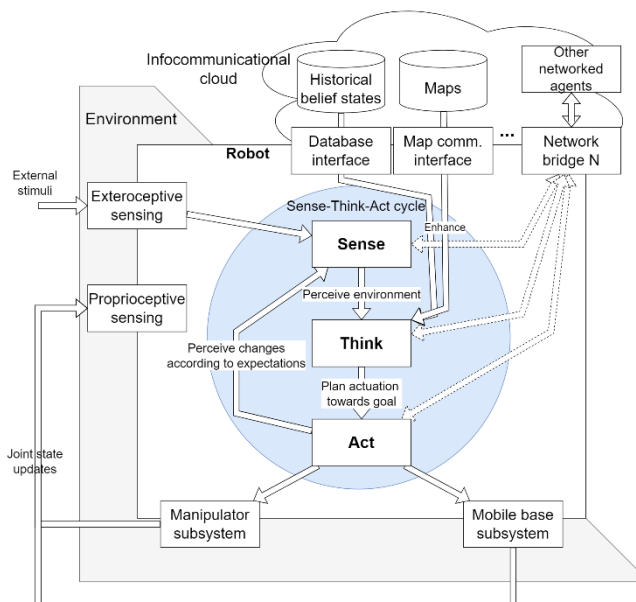


Fig. 1. Modified sense-think-act cognitive model for mobile robots

III. CONCPETUAL DESIGN

In the following chapter, the conceptual design is discussed in detail, including the typical use cases, requirements based on industry-related standards, and

physical requirements related to the electronic equipment, structural integrity, and safety of the robot. Additionally, the specific requirements for each possible environment setting are briefly discussed.

A. High-level description

In the following subsection, the primary modules of the robot are identified according to its primary goal. The discussed robot conceptually is a composition of a mobile robot platform (with a low-level platform controller) equipped with two industrial manipulators (each with its independent controller system). While this robot design is relatively widespread in academic research [10] and disaster management [11] in industrial applications, this kinematic setup remains quite unconventional. At a bare minimum, the robot provides proprioceptive information of its joint states (motor states), thus its whole kinematic structure by the controller systems. Proprioceptive information can be amended with inertial data (IMU sensor) in the frame of the platform and arms joints. The mobile robot platform must provide a way to the extension with additional sensor equipment (e.g., cameras, LiDARS, further IMUs, GNSS) to process external sensory information in a modular way. The perception extension point on the robot should be placed distant from the working space and payload of the robot (*perception module*). *Power management* is treated as a separate component, as all other components must be powered. Notably, a possible significant difference between the electric supply of the mobile base and both manipulators. *High-level control* (i.e., cognitive) is composed of computers with extensive computational capabilities (e.g., perception, mapping, and planning tasks). *Low-level control* (i.e., actuator-close) indicates the computational units that presumably have real-time capabilities and high reliability with direct access to the actuators of the robot. The overview of the robot with the highlight of mandatory components is depicted in Fig. 2.

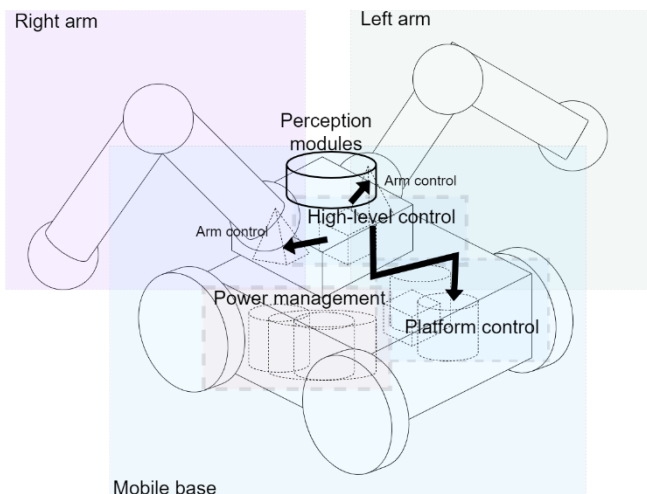


Fig. 2. Overview of the robot with the indication of important components

B. Requirements analysis

The robot is designed to operate in an exterior environment, categorized as:

- **External industrial environment:** the ground is typically paved, and the infrastructure is well-developed. Typical environments include urban and industrial environments (e.g., large factories, oil refineries).

- **External natural environment:** the ground is unpaved (rough) with scarce infrastructural development. Typical environments include agricultural areas, natural reserves, raw material extraction, and military training grounds.

Compared to mobile autonomous robots operating in an industrial environment, exterior mobile robots are posed with numerous additional challenges (summarized in Fig. 3). In the following subsections, these main challenges are discussed in detail.

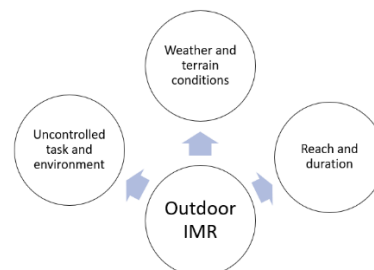


Fig. 3. Challenges for outdoor IMR

1) Weather and terrain conditions

In contrast to a controlled industrial environment, outdoor applications must cope with unpredictable weather and varied terrain conditions. Even in a built environment, the device faces many challenges: rough road conditions, potholes, inclines, slopes, and curbs. The road surface is built from materials of different mechanical attributes, such as asphalt, concrete, gravel, and unpaved dirt. Moreover, the weather condition is a stochastically and rapidly changing factor. Rain or strong wind can be just as problematic as intense sunshine. Defense against these conditions requires appropriate countermeasures, all in such a way that the device itself has to apply for the appropriate protection since usually there is no possibility for operator intervention during autonomous work.

2) Range and runtime duration

Another set of challenges is the question of operating range and runtime. Compared to typical industrial applications, the device is intended to operate in a much higher distance range from the mission starting point and usually is idle in this position for much longer. These use cases place a double requirement on the device. The high distance range requires a high degree of autonomy, problem-solving, and adaptation for various missions. On the other hand, this requirement also sets higher performance and movement speed standards than widespread applications. The movement speed of traditional industrial mobile robots is typically unfeasible to cover several kilometers in a typical exterior mission time or is only possible to a very limited extent. The environment requires a traveling speed of at least 15-20 km/h or faster, which corresponds to the average speed of a cyclist. The runtime duration must be also adjusted to cover the operating range with electronic equipment.

3) Uncontrolled task and environment

In the controlled industrial environment, the tasks of the mobile robots are optimized and precalculated. Nowadays, the entire operation of robots is managed under constant monitoring in a relatively static environment. However, in the case of outdoor applications, the environment is much more

variable and nondeterministic. As a result, the tasks that arise are not controlled, and their solution in many cases requires autonomous problem solving and swift improvisation from the mobile robot. In summary, the nature of the tasks and the environment pose the primary challenge to the robot's systems (especially to its software). Therefore, the robot components must be flexible, modular, and standardized as possible.

C. Conceptual design of the robot

1) Drive system

Mobility is a primary requirement for the presented robot. The requirement for IMRs varies widely. In our case, the drive system requirements are high-precision positioning at low speed, starting with high torque, and traveling at high speed are also among the requirements. Low speed and high accuracy are usually characteristics of some servo drive systems, so their application is essential in the current case. High starting torque is also a characteristic of synchronous motors, which is also satisfied by the previously mentioned servo systems. On the other hand, moving at high speed and momentum requires a different speed and torque range, which is more characteristic of asynchronous drives [12]. Based on these, using a hybrid, synchronous-asynchronous drive system would be most favorable in outdoor robots.

2) Power management

Ensuring the power supply is critical for an outdoor autonomous mobile robot, especially if the primary requirement of increased distance range is considered. Furthermore, breakdown or discharge events can also be a problem. Therefore, a power supply system that includes redundancy and self-management capability is required. The self-management capability means that the robot can control the power of individual subsystems. The associated tasks include, for example, determining the required energy needed for return to the base location and preventing an unexpected event by shutting down the least critical individual subsystems. Continuous energy management is fundamental to the operation of the robot. The power supply module must manage the charging process, the state of the accumulator, and individual subsystem energy consumption. Providing the possibility of remote monitoring and intervention is also a mandatory capability. Power management must operate in the most energy-efficient way possible. The issue of on-the-go charging and energy recovery may also arise, which can be implemented with solar panel or energy feedback-generator methods.

IV. DETAILED DESIGN

In the following chapter, the robot's design in more detail is discussed. This includes some insights to mechanical design, software architecture, and simulation.

A. Mechanical design

The mechanical design of the robot considered the challenges listed in the previous chapter. Optimally low weight and low center of gravity generally benefit long-range and higher speed. The loading area located in the structure of the robot platform (similarly as depicted in Fig. 4) allows the main geometric parameters of the robot to be displayed unchanged even when transporting cargo. This mechanical design contrasts with the solution for indoor industrial applications, where the robot carries the payload on top. The design with two manipulators is a means of high flexibility

and adaptability, but proper energy management is necessary to avoid unwanted leakage.

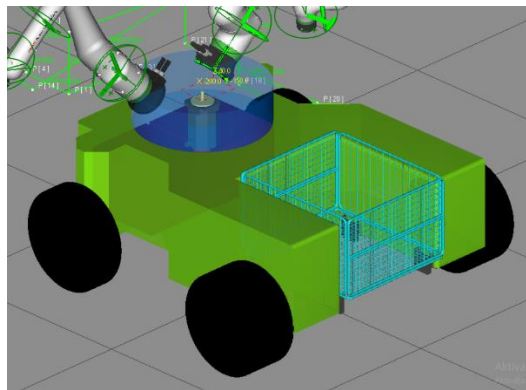


Fig. 4. Payload area located in the structure of the robot.

B. Software system design

The software system should fulfill current expectations of modern robotic software paradigms, such as modularity and distributed organization. Recent development in robotic software frameworks provides a flexible toolset to develop and integrate different software functionalities into a working robot. Widespread frameworks (e.g., ROS2) model (remote and memory-resident) component communication with a communicational graph providing a realization of popular machine-to-machine communicational patterns (e.g., Publisher/Subscriber, Remote Procedure Call), aiding the development of specialized but interconnected components can be created that can be each verified separately or replaced according to the task. This approach corresponds with modern software design and provides a solid framework for iterative testing and development but is also congruent with the connectionist approach in cognitive psychology.

Using the extended sense-think-act model and inspired by the advances in advanced-driving assistance systems (ADAS) [13], the minimal subset of tasks can be identified to perform autonomous tasks:

- **Sensing:** the interface to the robot's sensors (proprioceptive and exteroceptive) and minimal processing, e.g., a camera driver, GNSS equipment driver.
- **Perception:** the processing of sensory data, detection of features, or sensor fusion, e.g., object detection & tracking, state estimation [14].
- **Mapping:** is the estimation of the robot's position on a wide global range (i.e., estimating its geocoordinate state) by using historically processed sensory data and external sources (e.g., geographical maps, landmark databases). In short, placing the robot on a constructed map. Example components include SLAM methods and geographical mapping.
- **Inference representation:** based on instantaneous and historical sensory data, the robot infers a formally representable state of its belief state, with different features weighted according to the current scenario. The most straightforward solution is to represent the belief state as a grid, but more complex and computationally efficient solutions are preferred, especially in an outdoor setting. It should be noted that many widespread solutions treat the representation of the environment and belief state as a monolithic part of planning. New research focuses

on separating representation as sets of independent components [15].

- **Planning:** based on historical state estimates, current environment representation, and detected features, the robot navigates towards its goal and optionally provides manipulator working trajectories. This plan could be a set of abstract actions (using a planning language) or minimally a trajectory towards a state. Most researchers approach the planning problem with a hierarchical solution, separating *global* planning (the robot's environment in a wide range) and refining the intermediate result by *local planning* (close surroundings of the robot). Widespread planners are sampling-based (e.g., RRT derivatives [16]), graph-based (e.g., A*, Hybrid A* [17]), and rely on some semantic data (e.g., kinematic structure, dynamic parameters) related to the robot or its surrounding environment elements. Separately for arm manipulators, specialized software suites are available with an interface for other software components (e.g., MoveIt2 [18]). Ultimately, the planner provides data for the control components, thus connecting other components with the actuators.
- **Control:** control components continuously provide reliable control signals for the actuation drivers based on the output of the planner components. If the planner provides a trajectory, the control components attempt to track this trajectory based on the instantaneous state estimate and an optimum strategy, therefore solving an optimization problem in this regard. Some methods can integrate local trajectory planning with control tasks with direct application to manipulators [19, p.].
- **Actuation:** the direct access to the actuator components (i.e., motors) through the provided control signals, e.g., motor drivers, gripper control.
- **Interaction:** the interface that provides interaction and monitoring capabilities for external actors (e.g., human users, other robots, other software), e.g., runtime monitors, user interfaces.

These cognitive components, their instantiation on one of the computational units of the robot, and their intermediate relationships are summarized in Fig. 5.

C. Simulation and verification

Simulation of robots typically aids the development of robots in every phase. In addition to the apparent boost to continuous software development, it also provides a way to verify different subsystems of the robot continuously. However, in the case of the presented robot, each manipulator and the platform could operate independently. Consequently, the simulation of each part must be checked and revised alongside the integrated system.

For general robotic systems, widespread rigid-body simulators (e.g., Gazebo [20]) are available. Such simulators enable the proper simulation of multiple robots and provide a mockup of their actuator-sensor interface for programs. With this methodology, it is possible to simultaneously develop programs for the simulated and the actual physical robot.

In the case of the presented dual-manipulator mobile robot, the construction and the low-level interfacing of components are challenging, making the physical robot available only in a later phase. A conceptual mockup can be created to start software development.

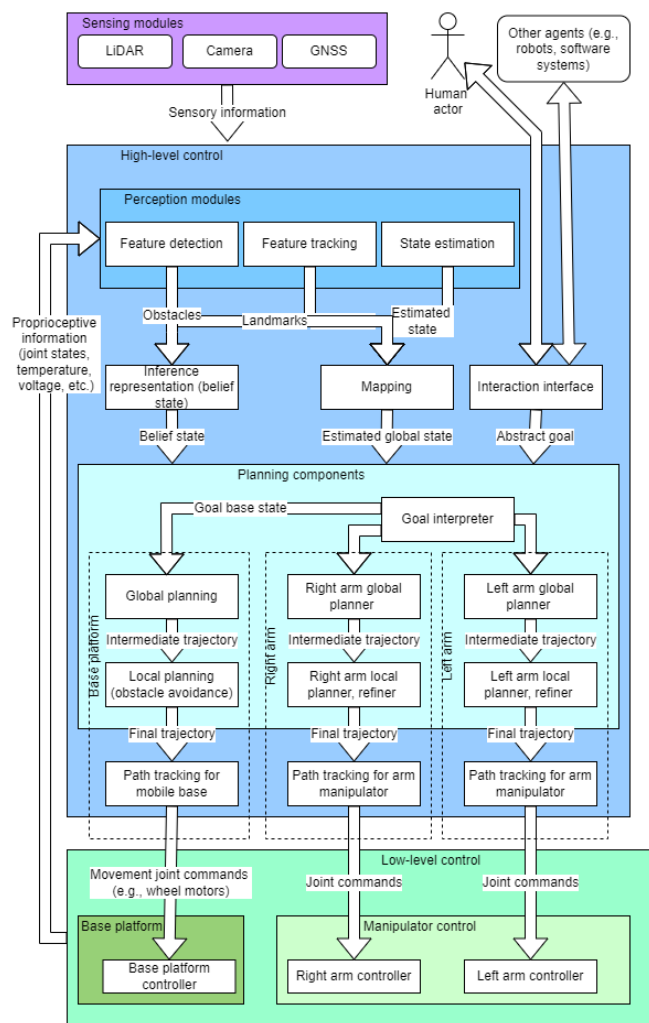


Fig. 5. Proposed software architecture of dual arm mobile robot

V. FUTURE WORK

The construction of the presented robot is a challenging and promising project from research and industrial aspects. This section discusses some further development steps ahead.

Creating the simulation mockup is mandatory in terms of software and mechanical verification. Therefore, the next step is to create a conceptual simulation of the robot with a set of different test scenes, each testing the capabilities of the robot. On the other hand, parallel with the simulation, a physical scaled mockup could be created to test the kinematic and mechanical attributes of the robot.

A scaled physical mockup would help integrate electronic components such as power supply, motor drivers, and low-level control. The electronics would be customized to some degree in the case of low-level control requiring substantial verification before use. An important step is the electronic integration of all low-level control components with related software bridge components. The sensory and actuation equipment requires further integration steps and additional bridging component. A considerable future milestone is the electronic integration of each independent component, followed by the composed setup.

Even a coarse simulation could be used to develop the essential identified software components. For example, the development of planner and perception components requires only a conceptual model of the robot. Therefore, creating

working test scenarios and environment setups that are meaningful and challenging enough is the focus of implementing simulation tasks.

VI. CONCLUSION

This paper presented a dual-arm mobile manipulator for intended industrial use. The presented setup is somewhat widespread in academic and military applications but rather unconventional in industrial applications, particularly not in outdoor environments. This paper presented the challenges and basic requirements posed to the robot. The paper also conceptualizes the software architecture of the robot enhanced for dual-arm application. Finally, the paper discusses the development steps, including the simulation. In summary, many tests and developments are required to achieve the outlined results with promising measurements. The number of solution options is high, and the challenge includes a variable and unpredictable element, possibly leading to pitfalls and breakthroughs in robotic development.

REFERENCES

- [1] C. Zhang and J. M. Kovacs, "The application of small unmanned aerial systems for precision agriculture: a review," *Precision Agriculture*, vol. 13, no. 6, pp. 693–712, 0 2012, doi: 10.1007/s11119-012-9274-5.
- [2] B. Ambrus, G. Teschner, A. J. Kovács, and M. Neményi, "DEVELOPMENT OF SMALL SMART DATA LOGGER ROBOTS EMBEDDED IN IOT SYSTEM FOR CROP PRODUCTION," in *International Scientific Conference „AGRICULTURAL MECHANIZATION AND TECHNOLOGY IN EUROPE AND PERSPECTIVES“ PROCEEDING, 2022*, pp. 50–55. [Online]. Available: <https://m2.mtmt.hu/api/publication/33079729>
- [3] R. Siegwart and I. R. Nourbakhsh, *Introduction to Autonomous Mobile Robots*. Scituate, MA, USA: Bradford Company, 2004.
- [4] M. Siegel, "The sense-think-act paradigm revisited," Jul. 2003, p. 5 pp. doi: 10.1109/ROSE.2003.1218700.
- [5] P. Baranyi, Á. Csapó, and G. Sallai, *Cognitive Infocommunications (CogInfoCom)*, 1st ed. Springer Publishing Company, Incorporated, 2015.
- [6] J. M. Schlesselman, G. Pardo-Castellote, and B. Farabaugh, "OMG data-distribution service (DDS): architectural update," *IEEE MILCOM 2004. Military Communications Conference, 2004.*, vol. 2, pp. 961-967 Vol. 2, 2004.
- [7] M. Quigley *et al.*, "ROS: an open-source Robot Operating System," 2009.
- [8] Y. Maruyama, S. Kato, and T. Azumi, "Exploring the performance of ROS2," in *Proceedings of the 13th ACM SIGBED International Conference on Embedded Software (EMSOFT)*, 2016, pp. 1–10.
- [9] D. Thomas, W. Woodall, and E. Fernandez, "Next-generation ROS: Building on DDS," Mountain View, CA, 0 2014. doi: 10.36288/ROSCon2014-900183.
- [10] G. Gillini *et al.*, "A dual-arm mobile robot system performing assistive tasks operated via P300-based brain computer interface," *IR*, vol. 49, no. 1, Art. no. 1, Jan. 2022, doi: 10.1108/IR-07-2020-0137.
- [11] A. Sharp, K. Kruusamae, B. Ebersole, and M. Pryor, "Semiautonomous dual-arm mobile manipulator system with intuitive supervisory user interfaces," Mar. 2017, pp. 1–6. doi: 10.1109/ARSO.2017.8025195.
- [12] M. Ehsani, Y. Gao, S. Gay, and A. Emadi, *Modern Electric, Hybrid Electric, and Fuel Cell Vehicles: Fundamentals, Theory, and Design*. 2004. doi: 10.1201/9781420037739.
- [13] S. Kato *et al.*, "Autoware on Board: Enabling Autonomous Vehicles with Embedded Systems," in *2018 ACM/IEEE 9th International Conference on Cyber-Physical Systems (ICCPS)*, Apr. 2018, pp. 287–296. doi: 10.1109/ICCPS.2018.00035.
- [14] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press, 2005.
- [15] C. Hajdu and Á. Ballagi, "Proposal of a Graph-based Motion Planner Architecture," in *2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*, 2020, pp. 000393–000398. doi: 10.1109/CogInfoCom50765.2020.9237891.
- [16] S. M. Lavalle, "Rapidly-Exploring Random Trees: A New Tool for Path Planning," 1998.
- [17] D. Dolgov, S. Thrun, M. Montemerlo, and J. Diebel, "Practical Search Techniques in Path Planning for Autonomous Driving," *AAAI Workshop - Technical Report*, 0 2008.
- [18] M. Görner, R. Haschke, H. Ritter, and J. Zhang, *MoveIt! Task Constructor for Task-Level Motion Planning*. 2019, p. 196. doi: 10.1109/ICRA.2019.8793898.
- [19] B. Magyar, N. Tsiogkas, J. Deray, S. Pfeiffer, and D. Lane, "Timed-Elastic Bands for Manipulation Motion Planning," *IEEE Robotics and Automation Letters*, vol. PP, pp. 1–1, 2019, doi: 10.1109/LRA.2019.2927956.
- [20] N. Koenig and A. Howard, "Design and Use Paradigms for Gazebo, An Open-Source Multi-Robot Simulator," in *In IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2004, pp. 2149–2154.

Artificial Experimental Environment for Indoor Plant Cultivation

Bertalan Beszedes
 Óbuda University
 Alba Regia Technical Faculty
 Székesfehérvár, Hungary
 beszedes.bertalan@uni-obuda.hu
<https://orcid.org/0000-0002-9350-1802>

Abstract—This paper explains the basic needs of indoor plants. The developed environment uses additional lighting for supplement plant production. The article deals with the control of power light emitting diodes, the optimization and diagnostic possibilities of the spectrum are part of the presented experimental environment. Eventually conclusions were drawn about the feasibility and requirements of such a system to make it possible to implement.

Keywords—greenhouse lighting, artificial seed germination, grow LED driver, light spectrum control, indoor plant monitoring

I. INTRODUCTION

Different plants need specific humidity, temperature and amount of light in order to develop properly. In order to create the right environment and conditions, for plants, it need to know, the plant's natural habitat and its properties. This includes the fact that plants also have a daily cycle, there is a period when the sun does not shine, when it is a little cooler, and when the air around them is drier or wetter. This concept can be extended by adding different plant profiles, monitoring different parameters and implementing schemes.

The proposed concept is based on modular subsystems, due to easier modifiability [1-5]. The optimization of the structure is considered according to the following papers [6-8]. Real time data monitoring and logging is also a significant part of the experimental environment [9, 10] [11, 12]. One effective ways control the operation is closed-loops and compensation of disturbances [13-15]. The electronic circuit implementation enables maintainability and robust operation [16-19].

A. Active period

To ensure daily cyclicity, the system links intervention limits to the local time, such as how long the system is allowed to automatically switch on lights and water. In the evening cycle, the temperature in the plant's natural habitat drops, the humidity increases and the amount of light decreases, for ideal growth, the control system must also follow this scheme.

As a result, a so-called "active period" was defined in the system, i.e. the actuators of the system can intervene during this period. The rest of the day is the "passive period" when only measurements are taken automatically, but no interventions.

B. Software background

In order to store the measurement results in an orderly, transparent and long-term way, it is worth storing the data in a database. With the help of queries, the system can produce easily interpretable, informative graphs. In addition to the database, a graphical display application is also required, which is compatible with the other elements of the system.

The two utilities must also work together with the user program written in python. The time-based database management program InfluxDB and the graphical display program Grafana were used in the experimental environment.

C. Database

InfluxDB is an open source time-based database designed to store data with time stamps, see on Fig 1. The management of time stamps is already basically built-in, so querying by time interval is much more transparent at the user level than in traditional databases. The control system is based on a Raspberry Pi3 base, the storage capacity of which can be expanded to a limited extent, it is necessary to properly compress and average the data at certain intervals, i.e. to implement archiving.

The database keeps high-precision and high-sampling frequency data in the real-time time period, and after a specified time, the data is reduced and averaged, and continues to exist as archival data occupying a smaller storage space. These "aged" data can also be queried, but their accuracy is reduced compared to real-time sampled data.

```
datapoints =
[
  {
    "measurement": "brushEvents",
    "user": "Carol",
    "time": "idoBelyeg",
    "fields":
    {
      "duration": 127, "intensity": 500
    }
  }
]
```

Figure 1. Inserting data to DB

D. Graphic display

The data must be displayed to the user in such a way that it provides valuable information both about the current values and, upon request, about the values of a

given time interval. After starting the Grafana client from the terminal, the user can connect to the appropriate port of localhost, where the graphical interface can be edited and viewed.

After writing the query, you can select a specific form of display (diagrams, graphs or indicators), and these can be given any color, unit of measurement, data accuracy and rounding. On Fig. 2. an example measurement can be seen.



Figure 2. Graphical user interface

II. DESIGN OF THE ENVIRONMENT

The environment was designed to provide the plants with different amount of light, water, air, and humidity. [20, 21]. The test equipment consists of several hardware elements sensors, actuators and a microcontroller. [22, 23].

The main control unit is a Raspberry Pi microcontroller. The output of the DHT22 digital temperature and humidity sensor can be connected directly to the digital input of the control unit.

The TSL25911FN chip is capable of detecting electromagnetic waves in the infrared and visible range, as well as measuring Lux. With the two built-in AD converters it can transmit the voltage measured by the photodiode via I2C communication. The module works with 16-bit resolution, so it can provide accurate data on the lighting of the environment.

The output of the soil moisture sensor is an analog signal, and a TLC1543 type 10-bit, 11 channel, high-speed CMOS A/D converter was used to digitize it. Communication with the module is done serially.

DS1302 is a real time clock with an I2C interface.

As an intervention module, a fan is needed for ventilation, a pump for irrigation and an LED for lighting. [24] The operation of these units can be controlled using the Raspberry Pi based on the processed sensor data. [25] A drive circuit was created to power the interventionists. [26] The system design can be seen on Fig. 3. The experimental environment can be seen on Fig. 4.

III. SOFTWARE SOLUTIONS

The program needs the ideal parameter list, with the help of which it can set the maximum and minimum temperature and humidity values, the time of artificial light illumination and the frequency of sprinkling. [27].

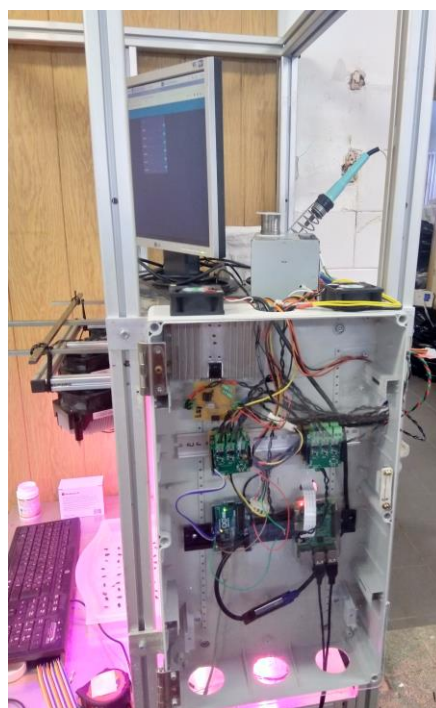


Figure 3. Picture of the test setup

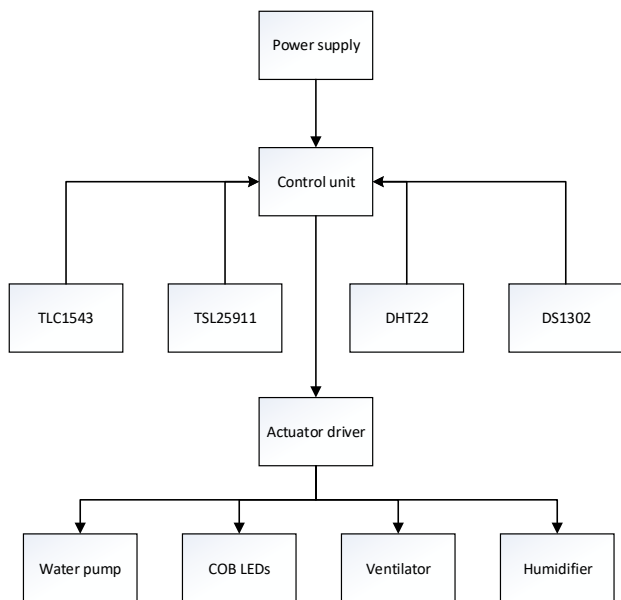


Figure 4. Block diagram of the hardware

A. User interface

The user must select the name of the appropriate plant from a list when the program starts, then the values of the variables are set after a short setup cycle.

The system is basically designed for automatic operation and intervenes in response to the values processed by the sensors, if necessary. However, as an experimental system, user interactivity still plays an important role, so the program was expanded with such functions. All functions, which are basically automatically executed cyclically, can be started at any time as a user. It is possible to have a full measurement performed on the system, to switch actuators on and off, and to modify the daily intervention counters.

User commands are sent using hotkeys, but user intervention cannot interrupt the cyclical running of the program. A software monitoring channel monitors the set of incoming keyboard interrupts. When a given keyboard combination is pressed, it is transferred to a frozen set of virtual keys, to which a function is individually assigned. The given function will be executed on a secondary thread, the main thread will deal with the automatic execution completely independently of this and will continue to run unharmed.

With this method, pressed button combinations can be observed for the entire Linux operating system, i.e. they take effect even after opening the graphical interface.

B. Daily cyclicality

The part of the software responsible for intervention operates according to the conditions described in point I/A. With the help of the built-in DateTime library, it is possible to query the system time and decide whether it is permissible to intervene in the given time interval.

In order to facilitate monitoring, new variables were also introduced, which track how long the given actuators were switched on a given day. These variables can also play an important role in deciding the necessity and limitation of the intervention. During the "active period" defined with DateTime, it is possible to determine, for

example, how many hours of artificial light the system should provide. These counters can be deleted every experimental period.

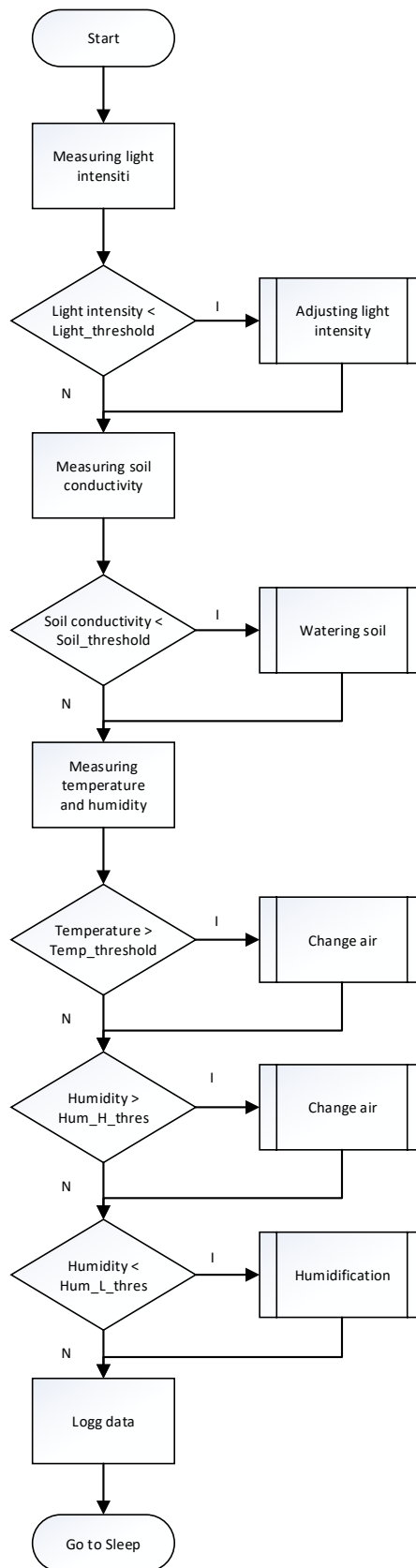


Figure 5. Flowchart of the software

C. Operation of the system

Taking into account the needs of the plant, a specific operating diagram can be defined:

- If the system temperature is too high, the fan will turn on.
- If the humidity is too high, the fan will turn on.
- If the humidity is too low, humidification is necessary.
- If the soil is too dry, it should be watered.
- If there is not enough light, the lamp turns on.

In summary, a flowchart shows the system operation in Fig. 5. The specific comparison levels were determined empirically.

IV. CONCLUSION

This paper aimed to show an optimized indoor plant production environment. The design contains power LEDs and driving circuits, a control system and a data logging application. The shown modular design can also be used for other indoor plant production applications. The proposed architecture is easily scalable, user-friendly, and robust. The article describes the relationship between the system components and includes an algorithm that demonstrates the control and measuring options.

ACKNOWLEDGMENT

The authors would like to thank all the faculty staff and member that provide help and assistance throughout the project completion.

REFERENCES

- [1] Alexander Baklanov, Svetlana Grigoryeva, György Györök, Control of LED Lighting Equipment with Robustness Elements, Acta Polytechnica Hungarica Acta Polytechnica Hungarica Vol. 13, No. 5, 2016. pp. 105-119.
- [2] Ramazan Şenol, Kubilay Taşdelen, A New Approach for LED Plant Growth Units, Acta Polytechnica Hungarica Vol. 11, No. 6, 2014. pp. 57-71.
- [3] A. Baklanov, S. Grigoryeva, Gy. Györök. Intelligent control of LED luminaries. 9th International Symposium on Applied Informatics and Related Areas - AIS 2014, Székesfehérvár, 2014. pp. 87-91 (ISBN:978-615-5460-21-0)
- [4] Attila, Sáfár ; Bertalan, Beszédes. Educational Aspects of a Modular Power Management System. In: Orosz, Gábor Tamás (szerk.) AIS 2019 : 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University. Székesfehérvár, Magyarország : Óbudai Egyetem, (2019) pp. 163-166. , 4 p.
- [5] György, Györök ; Bertalan, Beszédes Concept of a Reliable Redundant Off-grid Power Supply Chain In: Szakál, Anikó (szerk.) SACI 2019 : IEEE 13th International Symposium on Applied Computational Intelligence and Informatics : PROCEEDINGS Temesvár, Románia : IEEE (2019) 383 p. pp. 205-10. , 6 p.
- [6] Beszédes, Bertalan ; Széll, Károly ; Györök, György Redundant Photo-Voltaic Power Cell in a Highly Reliable System ELECTRONICS 10 : 11 p. 1253 , 20 p. (2021)
- [7] Bertalan, Beszédes ; Károly, Széll ; György, Györök A Highly Reliable, Modular, Redundant and Self-Monitoring PSU Architecture ACTA POLYTECHNICA HUNGARICA 17 : 7 pp. 233-249. , 14 p. (2020)
- [8] Györök, György ; Bertalan, Beszedes Fault tolerant power supply systems In: Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016) Székesfehérvár, Magyarország : Óbudai Egyetem (2016) pp. 68-73. , 6 p.
- [9] Bartos, Gaye Ediboğlu, et al. "A Multilingual Handwritten Character Dataset: THE Dataset." Acta Polytechnica Hungarica 17.9 (2020).
- [10] Meng Qingyan ; Chen Xu ; Zhang Jiahui ; Sun Yunxiao ; Li Jiaguo ; Jancsó Tamás; Sun Zhenhui Canopy Structure Attributes Extraction from LiDAR Data Based on Tree Morphology and Crown Height Proportion PHOTONIRVACHAK / JOURNAL OF THE INDIAN SOCIETY OF REMOTE SENSING (0255-660X 0974-3006): 46 9 pp 1433-1444 (2018)
- [11] Xiaojiang, Li ; Weindong, Li ; Quingyan, Meng ; Chuanrong, Zhang ; Tamas, Jancso ; Kangli, Wu Modelling building proximity to greenery in a three-dimensional perspective using multi-source remotely sensed data JOURNAL OF SPATIAL SCIENCE , 16 p. (2016)
- [12] G. Simon and L. Sujbert, "Special issue on "Recent advances in indoor localization systems and technologies", " Applied Sciences, Vol. 11, No. 9, paper 4191, May 2021.
- [13] L. Sujbert, G. Simon and G. Peceli, "An Observer-Based Adaptive Fourier Analysis," IEEE Signal Processing Magazine, Vol. 39 No. 4, pp.134-143, Oct. 2020.
- [14] G. Simon, G. Zachár, and G. Vakulya, "Lookup: Robust and Accurate Indoor Localization Using Visible Light Communication," IEEE Trans. on Instrumentation and Measurement, Vol. 66, No. 9, pp.2337 - 2348, Sept 2017.
- [15] Tóth, Peter; Pogatsnik, Monika. Advancement of inductive reasoning of engineering students. HUNGARIAN EDUCATIONAL RESEARCH JOURNAL (HERJ) pp. 1-21., 21 p. (2022)
- [16] Módné Takács, J., Pogátsnik, M., Kersánszki, T. (2022). Improving Soft Skills and Motivation with Gamification in Engineering Education. In: Auer, M.E., Hortsch, H., Michler, O., Köhler, T. (eds) Mobility for Smart Cities and Regional Development - Challenges for Higher Education. ICL 2021. Lecture Notes in Networks and Systems, vol 389. Springer, Cham. https://doi.org/10.1007/978-3-030-93904-5_81
- [17] G. Györök and B. Beszédes, "Highly reliable data logging in embedded systems," 2018 IEEE 16th World Symposium on Applied Machine Intelligence and Informatics (SAMi), 2018, pp. 000049-000054, doi: 10.1109/SAMI.2018.8323985.
- [18] S. Nosratabadi, K. Szell, B. Beszedes, F. Imre, S. Ardabili and A. Mosavi, "Comparative Analysis of ANN-ICA and ANN-GWO for Crop Yield Prediction," 2020 RIVF International Conference on Computing and Communication Technologies (RIVF), 2020, pp. 1-5, doi: 10.1109/RIVF48685.2020.9140786.
- [19] Baklanov, A., Baklanova, O., Grigoryeva, S., (...), Vais, Y., Györök, G. The development of hybrid IP architecture for solving the problems of heating networks (Using pipeline-parallel data processing technology). 2020. Acta Polytechnica Hungarica 17(1), pp. 123-140
- [20] Györök, G. Continuous Operation Monitoring of Electronic Circuits with Embedded Microcontroller. 2018 18th IEEE International Symposium on Computational Intelligence and Informatics, CINTI 2018 – Proceedings 8928147, pp. 155-160
- [21] Zhaparova, A., Titov, D., Baklanov, A.Y., Györök, G. Study of the effectiveness of switching-on LED illumination devices and the use of low voltage system in lighting. 2015 Acta Polytechnica Hungarica. 12(5), pp. 71-80
- [22] Györök, G. Crossbar network for automatic analog circuit synthesis 2014 SAMI 2014 - IEEE 12th International Symposium on Applied Machine. Intelligence and Informatics, Proceedings 6822419, pp. 263-267
- [23] Ladislav Fozo ; Rudolf Andoga ; Radovan Kovacs. Thermodynamic cycle computation of a micro turbojet engine. 2016.In: CINTI 2016. Danvers : IEEE, 2016 P. 000075-000079. ISBN 978-1-5090-3909-8
- [24] FŐZŐ, Ladislav ; ANDOGA, Rudolf ; KOVÁCS, Radovan. Experimental identification of a small turbojet engine with variable exhaust nozzle. 2015. In: CINTI 2015. Danvers : IEEE, 2015 P. 65-69. ISBN 978-1-4673-8519-0
- [25] KOMJÁTY, Maroš ; FŐZŐ, Ladislav ; ANDOGA, Rudolf. A Digital Diagnostic System for a Small Turbojet Engine 2013. In: Acta Polytechnica Hungarica. Vol. 10, no. 4 (2013), p. 45-58. ISSN 1785-8860

- [26] ANDOGA, Rudolf ; FŐZŐ, Ladislav ; MADARÁSZ, Ladislav - KAROL, Tomáš. Innovative approaches in modeling, control and diagnostics of small turbojet engines. 2013. In: Acta Polytechnica Hungarica. Vol. 10, no. 5 (2013), p. 81-99. ISSN 1785-8860
- [27] György, Györök ; Bertalan, Beszédes. Artificial Education Process Environment for Embedded Systems. In: Orosz, Gábor Tamás (szerk.) 9th International Symposium on Applied Informatics and

Related Areas - AIS2014. Székesfehérvár, Magyarország : Óbudai Egyetem, (2014) pp. 37-42. , 6 p.

Design of a Reliable Portable Power Supply Chain for an Embedded System

Bertalan Beszédés
 Óbuda University
 Alba Regia Technical Faculty
 Székesfehérvár, Hungary
 beszedes.bertalan@uni-obuda.hu
<https://orcid.org/0000-0002-9350-1802>

Abstract— Microcontrollers and electrical circuits are fail in case of low supply voltage. They can switch off, oscillate or do unpredictable things because they complexity. To prevent such scenario additional protection circuits and firmware elements should be added. In this article shows a reliable energy supply chain and proposes solutions to prevent this problem. The research aimed to show various PCB designs for diverse environments and supply voltages. The article deals with battery cells, charging and protection circuits, in-circuit current measurement and USPs.

Keywords— voltage supervisor, digital current measurement, in-circuit current measurement, lithium iron phosphate, LiFePo4 protection, reliable microcontroller supply, reliable Raspberry PI supply

I. INTRODUCTION

Low power supply voltage can cause instabilities and vulnerabilities. While powering-up a device, the voltage starts at 0V and ends around the supply voltage. During this surge it passes the hazardous voltage range. In battery powered operation, if the battery discharges, the output cell voltage is decreasing, without proper under voltage protection, the power supply also can fall to the hazardous voltage range, especially with high temporary type load current spikes.

High value capacitors in the supply lines can slow down the startup voltage rise time, and corrupt microcontroller operation. The fail of the system can handle by a restart button, but it is inconvenient for the user, especially in a remote location. Till the stable supply voltage is reached, it is advisable to keep the reset pin in active state or the enable pin in inactive logic level. In practice the task can be done with an RC element connected to the reset pin, it will delay the supply voltage on the pin, but there are further solutions.

Numerous solutions exist to power MCUs and computers, like Raspberry Pi [1-3], within this portable [4] and data acquisition hardware and software realizations have got a wide range of solutions [5, 6]. MCUs can be hardened for external disturbances and also well controllable [7, 8]. Many of supply chains can be realized in the field of the embedded systems [9-11].

The experimental system is based on modular sub-circuits, due to easier modifiability [12, 13]. The optimization of the layout is considered according to the

following papers [14-16]. Real time data monitoring and logging is also a significant part of the development [17-20].

One of the effective ways of controlling the system is closed-loops and compensation of disturbances [21-23]. Various optimized solutions and compensating for uncontrolled disturbances are discussed in a series of papers [24-26]. Numerous schemes of control systems architecture are considered [27, 28].

II. SYSTEM ELEMENTS

A. Battery cells

To power mobile applications several energy sources can be considered. Alkaline cells are starts at 1.5V and ends around 1V. The capacity of an AAA battery is about 1200mAh, and an AA battery is about 3400mAh. Alkaline batteries are not rechargeable in default. For a 3.3V power supply with three cells, a low dropout voltage regulator is needed. For a 5V power supply with one, two or three cells, a boost converter is needed. Regulators and converters are shortening the life capacity due to the voltage conversion losses. Alkaline cells do not need external protection circuit, and they do not explode under non-operating conditions.

Nickel cadmium (NiCd) cells are starts under 1.4V and ends above 1V. The capacity of an AAA battery is around 300mAh, and an AA battery is about 800mAh. Nickel cadmium batteries are rechargeable. In the discharge cycle the cell voltage mostly between 1.2V and 1.3V. Three fully charged cells voltage (4.2V) is too high for 3.3V components, and LDO also needed for secure operation.

Nickel metal hydride (NiMH) battery discharge voltage curve is similar to the nickel cadmium batteries. The capacity of an AAA battery is about 900mAh, and an AA battery is around 2200mAh, and they are also rechargeable.

Lithium polymer (LiPo) cells are starts at 4.2V and ends around 2.8V. For a 3.3V power supply an LDO, and for a 5V power supply, a boost converter is needed. Lithium polymer batteries are made with different packages, with different capacities, and they are rechargeable. Lithium polymer cells are good for indoor usage, but performs poorly at low temperature field applications.

Lithium iron phosphate (LiFePo₄) cells voltage starts at 3.6V and ends just under 3V (2V in the edge of deep discharge). Most of the discharge curve the cell voltage stays above 3V, so no voltage regulator is necessary. Lithium polymer batteries are made in cylindrical and rectangular packages. The capacity of an AAA (14500) battery is about 700mAh, and an 18650 cell is around 1500mAh. The main advantages of LiFePo₄ are its excellent low temperature performance (LiPo cells are losing capacity under 0°C).

For a 3.3V power source, in case of powering a mobile device, three alkaline cell is equal to one lithium iron phosphate cell. Lithium iron phosphate batteries are suitable for higher current applications, maximum discharge current is less than 1.5C, and they are rechargeable. LiFePo₄ cells are also safer than LiPo cells, and have got lower self-discharge rate. A 100% depth of discharge (DoD) is also less detrimental to battery life. The capacity of LiFePo₄ batteries are smaller and the price is approximately the same than LiPo batteries.

Lithium iron phosphate cells need to charge with 3.6V, the charging current is 0.4C, the fast charging current is close to 0.9C.

B. Battery charging and protecting

For a 5V power source, in case of powering a mobile device two lithium cell and a voltage regulator or one lithium cell and a boost converter is needed. Most of mobile applications (except the high power applications) energy density and charging time is not primary, but safety is quite important.

Lithium polymer cells are protected against overcharge and deep discharge with (for example) TP4056 integrated circuit. Lithium iron phosphate cells can be protected with TP5000 and CN3058 ICs.

A rechargeable battery with a boost converter can supply embedded microcontrollers. For protection against over-discharge, an additional - for example - a TP4056 board is advisable between the battery cell and the boost converter. It also prevents the battery from short circuits and also can charge it from 5V.

The TP5410 IC based modules can charge a LiPo cell, it also has got a boost converter, with a low quiescent input current (under 15uA), but it has no over-discharge and short circuit protection.

To cover the needs, the TPS6109x synchronous boost converter is suitable till 2A output current, it has a built-in low battery comparator and it disconnects the load during startup and shutdown, to prevent fail controller operation.

For higher current applications the TPS6123x is a possible choice till 5A output current, with all the previously mentioned features, it also has programmable soft start operation, and output over voltage protection.

C. Portable PV extension

For portable operation, it is a good practice to add a small solar panel to the system. Solar panels output power is depending from the incoming solar radiation and from the output load. Both parameter can alter during operation. An MPPT charge controller can optimize the PV cell output for maximum power. CN3791 is a photovoltaic cell maximum power point tracker and a complete charge controller for single cell lithium batteries, it works up to 28V input voltage and to 4A charge current. The charging

current is programmed with a current sense resistor, the IC has battery overvoltage protection, the precharge phase (16% of the charge current at CN3791) threshold voltage is 2.8V. It is low enough not to compromise the operation, the battery cell protection circuit is disconnecting the load before the cell would discharge till 2.8V.

D. Current measurement

In-circuit current measurement requires a serially placed shunt resistor. With current sensors the solar power, the load current and the battery charge level can be measured. The INA219 is a bi-directional DC current, voltage sensor and power monitor, with I2C interface. The INA3221 is also capable for the task, it is a three channel high-side voltage and current sensor device, with I2C interface. Fig. 1 is showing the measuring concept. With the help of an 0.1Ohm shunt resistor, the measurable current range is $\pm 1.6A$ and the resolution is 0.4mA. The shunt resistor should be high precisions and with a power of at least 2W.

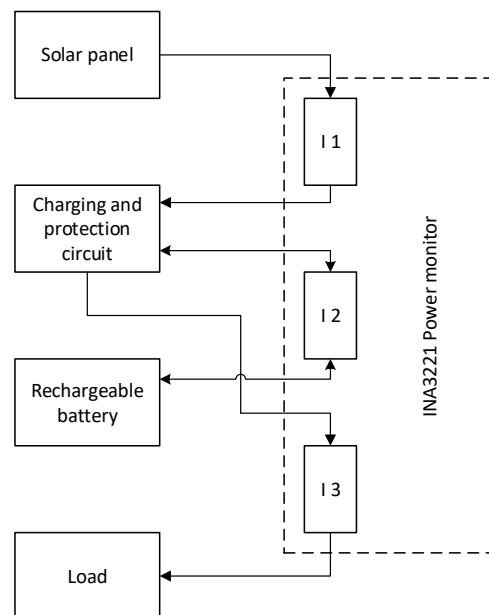


Figure 1. In-circuit current measurement

In the shunt resistor there is always some small, but potentially significant voltage drops, which is called burden voltage. Hall sensor based current sensors can measure charge movement with electrically isolated from the power leads, they can be used both in the low and the high side of the power supply loop, and they will not create a burden voltage. Most of the hall sensor based current measuring modules are designed for high currents, so they are not accurate at low currents (for example the ACS712 is designed $\pm 5A$, $\pm 20A$ and $\pm 30A$). Hall sensors can be disturbed with electromagnetic fields.

LTC4150 also a useful tool, a coulomb counter, with two voltage sensor and one current sensor. It can measure the input and output charge before a rechargeable battery, and can monitor the amount of stored charge. These solutions can be useful for low power microcontroller based portable devices, a small solar panel also can be integrated into the system.

III. REALIZATION

A. Energy supply chain

For a portable power supply, mains or sun could be the power source. Microcontrollers and sensors are run on 3.3V or 5V, or sometimes both. To get 5V from mains, a USB charger or a power supply can be used. For 3.3V supply, the use of a low dropout voltage regulator (LDO) is recommended. Some of the microcontroller boards have got a built-in 3.3V voltage regulator, if they accept 5V supply.

For portable devices, where the device should be independent from mains voltage from time to time an energy storage element is required. From the solar cell, a battery charger charges the battery, it is proposed to implement a LiFePo4, a LiIon or a LiPo battery. The batteries must also be able to be charged from mains voltage, a charging circuit is required for this task. The protector circuit should cut off the battery before over charge or over discharge. An LDO or a boost converter provides the power supply for the microcontroller and the additional elements. Fig. 2 shows the described architecture.

B. UPS for Raspberry Pi

The energy chain is able to charge the battery and supply the load at the same time, most of the external uninterruptible power supplies (UPS) do not have this feature. The used UPS module can intelligently switch between main power and standby power, even in case of small load current. The operation enables to use multiple power supplies and energy storages in a portable device. In the prototype it switches between a wall adapter DC output and the PiFePo4 battery's boost converter output. The first way is the main power source, and the second one is the battery cell.

The UPS module also can be used in redundant portable applications, where are duplicated batteries. When the first battery become empty, the module automatically switches to the second battery. The user also could parameterize the switching algorithm, and decide a priority list – the function assumes continuous power supply to the microcontroller

Because of the battery charging IC charging curve, the load should be disconnected for a few minutes if the battery cell voltage is under 3V. The charging process is starting with a trickle current, which is 1/10 of the charging current. Most of the microcontrollers are not consume more power in normal operation than the trickle current, in case of an 18650 cell. For a Raspberry Pi, an additional voltage supervisor system can be added, to shut down the Raspberry Pi if the cell voltage goes under 3V, and starts up if the cell voltage is higher than 3V.

In a Raspberry Pi the software is stored in SD cards most of the time. SD cards can become corrupted if the Raspberry Pi loses power without a proper shutdown. Reliability with SD cards can be achieved by decreasing writing cycles. SSDs are a good solution to swap SD cards, but they are relative expensive, they take up a lot of space, for connection they need an external cable and they block one of the USB3 connectors.

Voltage supervisors are also a good opportunity to solve the problem, they increase the reliability of the

device. They have got a voltage reference, a comparator and an output driver transistor. The component keeps the reset pin low, till the power supply voltage level is under the reference voltage. With a MOSFET it is feasible to handle the whole project.

Voltage supervisors are working on lower voltages than the supply voltage and with hysteresis to prevent oscillation. KA75xxx series is made by TO-92 packages, KA75330 is for 3.3V, KA75290 is for 2.9V, from the TPS383x series TPS3839 is switches at around 3.14V. The voltage supervisor pin is also connected to a Raspberry Pi GPIO pin to start a shutdown sequence. An auto-launch application during the booting process, a Python script is checking the GPIO pins, and automatically shut down the Raspberry Pi in case of low supply voltage.

With a 5.5V supercapacitor bank, the MCU can safely shut down. Capacitors store energy in electrostatic fields, can handle heavy charging, discharging currents and they can be used in harsh environments. Supercapacitors energy density is much higher than ordinary capacitors, but they store less energy than batteries.

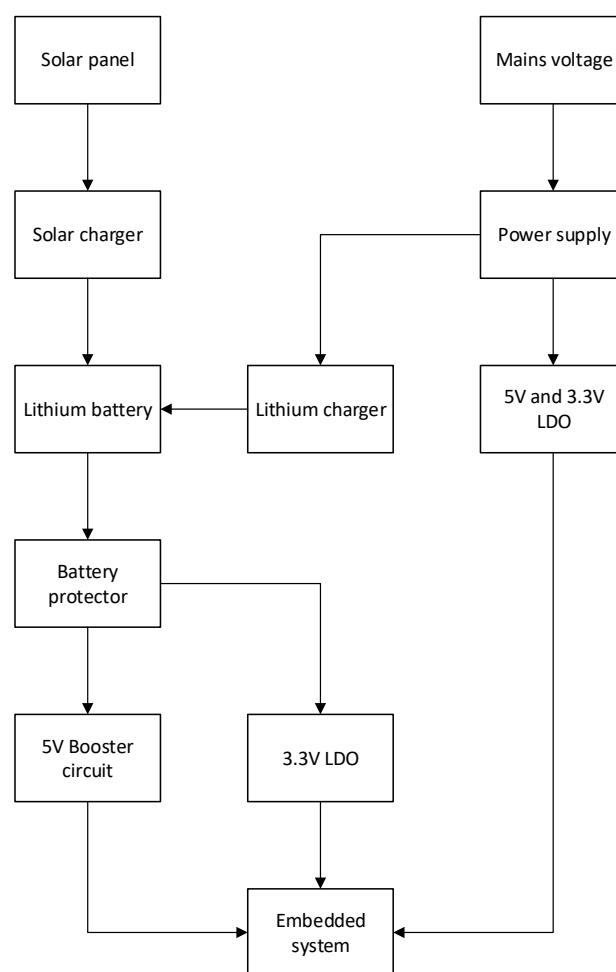


Figure 2. Portable supply chain architecture

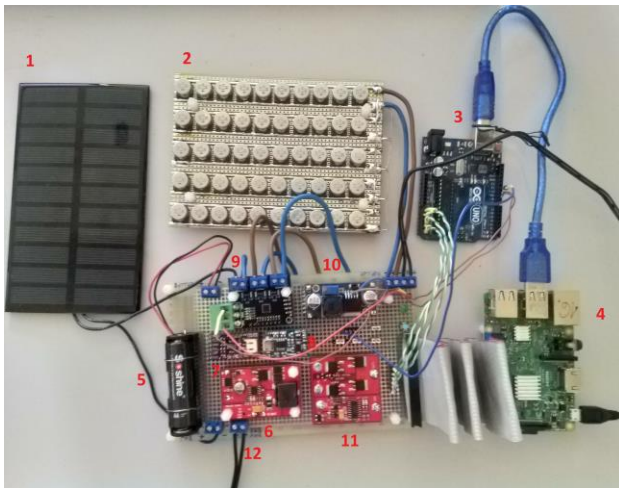


Figure 3. Experimental environment (1 – Solar panel, 2 – Supercapacitor bank, 3 – microcontroller with ADC, 4 – Raspberry Pi, 5 – PiFePo4 cell, 6 – Solar charger, 7 – Coulomb counter, 8 – Battery charger and protector, 9 – Power monitor, 10 – Boost converter, 11 – UPS, 12 – PSU input)

IV. CONCLUSION

The paper proposed reliable solutions to prevent MCUs and computers to fail due to power supply imperfections. The realizations shown voltage indication possibilities detailed the types of electrochemical energy storage and in-circuit current measurement techniques, for portable embedded systems in a practical manner. The results of the development showed that the use of voltage supervisors are capable to increase reliability. The shown modular design can also be used for other supply chain applications. The proposed architecture is easily scalable and user-friendly. The article describes the relationship between the system components. Eventually conclusions were drawn about the feasibility and requirements of such a system to make it possible to implement.

ACKNOWLEDGMENT

The authors would like to thank all the faculty staff and member that provide help and assistance throughout the project completion.

REFERENCES

- [1] Sati, Salem & El-bareg, Ahmed. (2018). MANET Testbed using Raspberry Pis. *International Journal of Wireless and Microwave Technologies*. 8. 52-63. 10.5815/ijwmt.2018.02.05.
- [2] Zoltán Balogh, Martin Magdin, György Molnár. Motion Detection and Face Recognition using Raspberry Pi, as a Part of, the Internet of Things. *Acta Polytechnica Hungarica* Vol. 16, No. 3, 2019
- [3] PANKOV, Pavel & NIKIFOROV, Igor & DROBINTSEV, Dmitry. (2020). Hardware and software data processing system for research and scientific purposes based on Raspberry Pi 3 microcomputer. *Proceedings of the Institute for System Programming of the RAS*. 32. 57-69. 10.15514/ISPRAS-2020-32(3)-5.
- [4] Ademi Ospanova, Aizhan Zharkimbekova, Lazzat Kussepova1, Aizhan Tokkuliyeva, Makhabbat Kokkoz. Cloud Service for Protecting Computer Networks of Enterprises Using Intelligent Hardware and Software Devices, Based on Raspberry Pi Microcomputers. *Acta Polytechnica Hungarica* Vol. 19, No. 4, 2022
- [5] Lazgheb, Skander & Bayar, Bacem & Belouda, Malek & Oueslati, Hatem & Mabrouk, S.. (2019). Raspberry Pi-based smart platform for data acquisition, supervision and management of a hybrid PV/WT/Batteries system. 10.1109/MMS48040.2019.9157307.
- [6] Ádám Csapó, György Wersényi, and Myounghoon Jeon. A Survey on Hardware and Software Solutions for Multimodal Wearable Assistive Devices Targeting the Visually Impaired. *Acta Polytechnica Hungarica* Vol. 13, No. 5, 2016
- [7] Alexander Baklanov, Svetlana Grigoryeva, György Györök, Control of LED Lighting Equipment with Robustness Elements, *Acta Polytechnica Hungarica*. Vol. 13, No. 5, 2016. pp. 105-119.
- [8] Bertalan, Beszédes ; Károly, Széll ; György, Györök A Highly Reliable, Modular, Redundant and Self-Monitoring PSU Architecture *ACTA POLYTECHNICA HUNGARICA* 17 : 7 pp. 233-249. , 14 p. (2020)
- [9] Attila, Sáfár ; Bertalan, Beszédes. Educational Aspects of a Modular Power Management System. In: Orosz, Gábor Tamás (szerk.) AIS 2019 : 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University. Székesfehérvár, Magyarország : Óbudai Egyetem, (2019) pp. 163-166. , 4 p.
- [10] György, Györök ; Bertalan, Beszédes Concept of a Reliable Redundant Off-grid Power Supply Chain In: Szakál, Anikó (szerk.) SACI 2019 : IEEE 13th International Symposium on Applied Computational Intelligence and Informatics : PROCEEDINGS Temesvár, Románia : IEEE (2019) 383 p. pp. 205-10. , 6 p.
- [11] Györök, György ; Bertalan, Beszedes Fault tolerant power supply systems In: Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016) Székesfehérvár, Magyarország : Óbudai Egyetem (2016) pp. 68-73. , 6 p.
- [12] Beszédes, Bertalan ; Széll, Károly ; Györök, György Redundant Photo-Voltaic Power Cell in a Highly Reliable System *ELECTRONICS* 10 : 11 p. 1253 , 20 p. (2021)
- [13] Györök, György ; Bertalan, Beszédes. Using Thermal Imaging Cameras to Test Electrical Systems. In: IEEE - Kádár, Péter; Lamacchia, Francesco P.; IEEE (szerk.) IEEE CANDO EPE 2018 : 2018 International IEEE Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE). New York (NY), Amerikai Egyesült Államok, Piscataway (NJ), Amerikai Egyesült Államok : IEEE (2018) pp. 147-152. Paper: 8 , 6 p.
- [14] Bartos, Gaye Ediboğlu, et al. "A Multilingual Handwritten Character Dataset: THE Dataset." *Acta Polytechnica Hungarica* 17.9 (2020).
- [15] Meng Qingyan ; Chen Xu ; Zhang Jiahui ; Sun Yunxiao ; Li Jiaguo ; Jancsó Tamás ; Sun Zhenhui Canopy Structure Attributes Extraction from LiDAR Data Based on Tree Morphology and Crown Height Proportion *PHOTONIRVACHAK / JOURNAL OF THE INDIAN SOCIETY OF REMOTE SENSING* (0255-660X 0974-3006): 46 9 pp 1433-1444 (2018)
- [16] A. Baklanov, S. Grigoryeva, Gy. Györök. Intelligent control of LED luminaries. 9th International Symposium on Applied Informatics and Related Areas - AIS 2014, Székesfehérvár, 2014. pp. 87-91 (ISBN:978-615-5460-21-0)
- [17] G. Györök and B. Beszédes, "Highly reliable data logging in embedded systems," *2018 IEEE 16th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2018, pp. 000049-000054, doi: 10.1109/SAMI.2018.8323985.
- [18] ANDOGA, Rudolf ; FÖZŐ, Ladislav ; MADARÁSZ, Ladislav - KAROL, Tomáš. Innovative approaches in modeling, control and diagnostics of small turbojet engines. 2013. In: *Acta Polytechnica Hungarica*. Vol. 10, no. 5 (2013), p. 81-99. ISSN 1785-8860
- [19] Xiaojiang, Li ; Weidong, Li ; Qingyan, Meng ; Chuanrong, Zhang ; Tamas, Jancso ; Kangli, Wu Modelling building proximity to greenery in a three-dimensional perspective using multi-source remotely sensed data *JOURNAL OF SPATIAL SCIENCE* , 16 p. (2016)
- [20] Tóth, Peter; Pogatsnik, Monika. Advancement of inductive reasoning of engineering students. *HUNGARIAN EDUCATIONAL RESEARCH JOURNAL (HERJ)* pp. 1-21., 21 p. (2022)
- [21] S. Nosratabadi, K. Szell, B. Beszedes, F. Imre, S. Ardabili and A. Mosavi, "Comparative Analysis of ANN-ICA and ANN-GWO for Crop Yield Prediction," *2020 RIVF International Conference on*

- Computing and Communication Technologies (RIVF)*, 2020, pp. 1-5, doi: 10.1109/RIVF48685.2020.9140786.
- [22] KOMJÁTY, Maroš ; FŐZŐ, Ladislav ; ANDOGA, Rudolf. A Digital Diagnostic System for a Small Turbojet Engine 2013. In: *Acta Polytechnica Hungarica*. Vol. 10, no. 4 (2013), p. 45-58. ISSN 1785-8860
- [23] G. Simon and L. Sujbert, "Special issue on "Recent advances in indoor localization systems and technologies"," *Applied Sciences*, Vol. 11, No. 9, paper 4191, May 2021.
- [24] L. Sujbert, G. Simon and G. Peceli, "An Observer-Based Adaptive Fourier Analysis," *IEEE Signal Processing Magazine*, Vol. 39 No. 4, pp.134-143, Oct. 2020.
- [25] G. Simon, G. Zachár, and G. Vakulya, "Lookup: Robust and Accurate Indoor Localization Using Visible Light Communication," *IEEE Trans. on Instrumentation and Measurement*, Vol. 66, No. 9, pp.2337 - 2348, Sept 2017.
- [26] Ladislav Fozo ; Rudolf Andoga ; Radovan Kovacs. Thermodynamic cycle computation of a micro turbojet engine. 2016.In: *CINTI 2016*. Danvers : IEEE, 2016 P. 000075-000079. ISBN 978-1-5090-3909-8
- [27] FŐZŐ, Ladislav ; ANDOGA, Rudolf ; KOVÁCS, Radovan. Experimental identification of a small turbojet engine with variable exhaust nozzle. 2015. In: *CINTI 2015*. Danvers : IEEE, 2015 P. 65-69. ISBN 978-1-4673-8519-0
- [28] Módné Takács, J., Pogátsnik, M., Kersánszki, T. (2022). Improving Soft Skills and Motivation with Gamification in Engineering Education. In: Auer, M.E., Hortsch, H., Michler, O., Köhler, T. (eds) *Mobility for Smart Cities and Regional Development - Challenges for Higher Education*. ICL 2021. *Lecture Notes in Networks and Systems*, vol 389. Springer, Cham. https://doi.org/10.1007/978-3-030-93904-5_81

Capacitive Sensor for Gaseous Substances, from Old to New

György Györök¹, Margit Makó¹, Borbála Györök², Olga Shvets³

¹ Óbuda University, Alba Regia Technical Faculty,
Budai Str. 45, 8000 Székesfehérvár, Hungary
{gyorok.gyorgy, margit.mako}@amk.uni-obuda.hu

² Óbuda University, Rejtő Sándor Faculty of
Light Industry and Environmental Engineering,
Doberdó st. 6, 1034 Budapest, Hungary
borbala.gyorok@cl.uni-obuda.hu

³ D. Serikbayev East Kazakhstan Technical University
Protozanov Street 68, Ust-Kamenogorsk, Kazakhstan
oshvets@ektu.kz

Abstract—In a previous development, it became important to be able to detect changes in the concentration of certain gases in the air. To this end, we want to develop a capacitive sensor that can effectively detect very small changes in composition. To design such a sensor, we wanted to use an old capacitive component, the application of which is described in the following article.

The capacitive sensor-processor circuit described in this paper, as a useful solution for operating a planar capacitor, can be advantageously used in the context of certain interactive suits, active wearable.

In this paper, the sensor itself, the analog-to-digital solution that interrogates the sensor, is presented.

I. INTRODUCTION

A few decades ago, rotary capacitors were widely used in electronic equipment, be it measuring instruments or even radio receivers. These were electronic components that consisted of a stator and a rotor disk package. By rotating a shaft, the position of the rotor could be changed, thus changing the effective surface area between the plates. Using this method, such air-insulated rotary capacitors could produce capacitance variations from a few pF to a few hundred pF [1][2].

The capacitance variation in a variable capacitor through an oscillator or filter in a circuit was an important means of determining frequency [3]. Nowadays, these components are much less important, but they can still be found in very large numbers and can be easily removed from older equipment [4][5].

The photo of the rotating capacitor shown in Fig.1 is a three-cell, $3 \times 330 pF$ capacitor. Other insulated rotary capacitors have been used, but are irrelevant for the present application.

II. CAPACITOR AS GAS SENSOR

We know from the theory of electricity that the capacitance of a planar capacitor depends on distant of plate (d), the surface of the plates facing each other (A) and the dielectric between the capacitor plates ($\epsilon_0 \epsilon_r$), where ϵ_0 is permittivity of vacuum, ϵ_r relative value specific to the material; so $C = \epsilon_0 \epsilon_r \frac{A}{d}$ [6][7][8].

When using such an air-insulated rotary condenser, changes in the air composition of this device are naturally associated with changes in capacitance. In this case, the relative change in the epsilon derived from the relationship will give a different value from the air or vacuum permittivity. Since such a capacitor is in contact with air or gas over a fairly large surface area, its capacitance can be greatly influenced for a

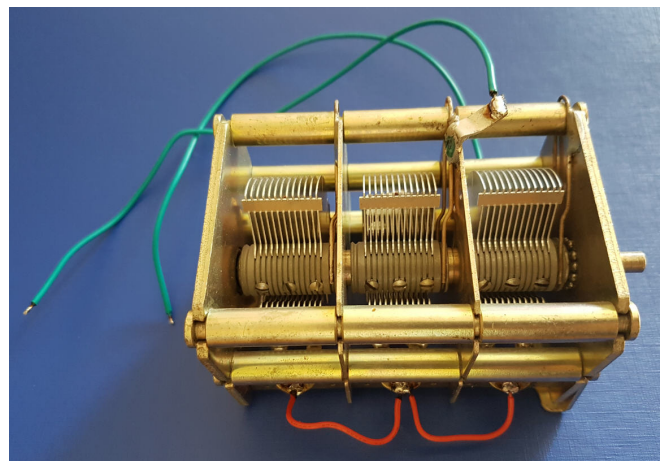


Fig. 1. Three-cell variable capacitor. The capacitance per cell is 330pF.

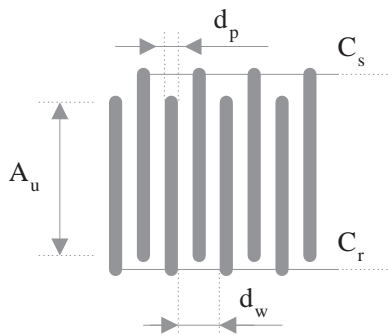


Fig. 2. Interpretation of rotary capacitor stator and rotor plates and derivation of useful surface area.

given rotor position (best when fully rotated).

Fig. 2 shows a simplified model of the capacitor. The stator (C_s) and rotor (C_r) plates are shown separately. The effective surface area of the plates is marked with A_u [9]. The distance between the stator plates is denoted by d_p , while the thickness of the rotor plates is denoted by d_w . Thus, the capacitance of the plate capacitor can be calculated using Equation 1.

$$C_v = n_c \epsilon_0 \epsilon_r \frac{2A_u}{d_w - d_d}, \quad (1)$$

where; n_c the number of cells, $\epsilon_0 = 8.854187817 \cdot 10^{-12} \frac{As}{Vm}$. With the approximate most accurate measurements, the capacitance is close to 1nF [10].

From the geometry ($A_u \simeq 5,6cm^2$, $d_p = 0,8mm$, $d_w = 0,6mm$) it is easy to calculate that the volume (V_d) of the dielectric (air, gas) in the capacitor is about $\simeq 40cm^3$. The volume of the dielectric between the arms of a capacitor is synonymous with the sensitivity and speed of the sensor [11][12].

III. ANALOGUE SENSOR CIRCUIT FOR CAPACITANCE CHANGE

We need a solution where the output signal of the sensor changes the most with the change in capacitance, the change in permittivity.

Frequency mixing, or heterodyne, is usually a good answer. In the commonly used solution shown in Fig.3, the signal

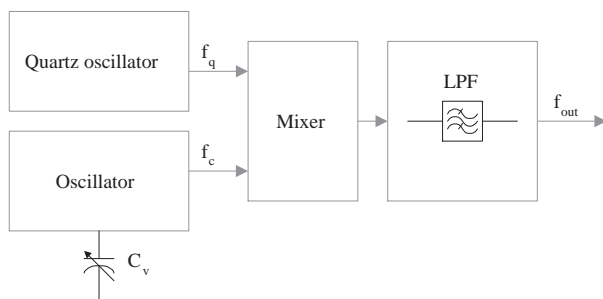


Fig. 3. Detecting the change in capacitance by changing the difference frequency between an oscillator and a reference oscillator.

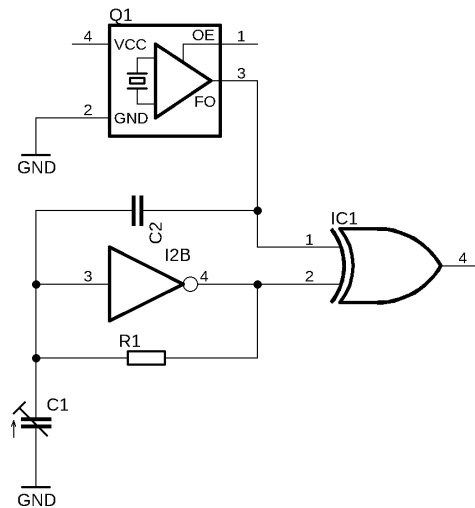


Fig. 4. Digital circuit for frequency or phase difference based capacitance difference detection.

from a fixed-frequency oscillator (f_q) is mixed with the signal from the capacitance-dependent (C_v) oscillator (f_c) and the difference frequency is selected using a low-pass filter (f_{out}). So we can formally write; $f_{out} = f_q - f_c$ [13].

As described in the equation 1, the equation follows 2;

$$\Delta f_c = \mathbb{F}(\Delta C_v), \quad (2)$$

that is, Δf_c is some function of ΔC_v), where formally write (3);

$$\Delta C_v = \mathbb{G}(\Delta \epsilon_r). \quad (3)$$

To achieve a high sensitivity, even with small changes in capacitance, a significant frequency difference must be achieved using relatively high frequency values. The experimental example operated at 10MHz. Also important is the stability of the reference oscillator, which in practice is quartz-precision [14][15].

In this case, it is the change in output frequency and its magnitude that provides information about the change in capacity [16] [17].

Our "gas detection" device was tested for CO_2 and O_3 gases. The change in capacitance was detectable even for not very exact measurements, the value was not greater than 0,3% (a few pF) [18].

IV. DIGITAL SENSOR CIRCUIT FOR CAPACITANCE CHANGE

On Fig. 4 is a circuit solution to detect the change in capacitance.

Capacitor (C_1) at the input of inverter (I_2B) is charged with resistor (R_1). When the comparator level is reached, the output of the inverter goes low and C_1 discharges. The frequency can be determined approximately; $f = \frac{1}{2\pi R_1 C_1}$. The $Q1$ is a quartz oscillator. Its output ($Q1_3$) and the output of the oscillator (I_2B_4) are connected to the inputs of an exclusion gate (IC_1). An important part of the circuit is C_2 , which synchronizes the

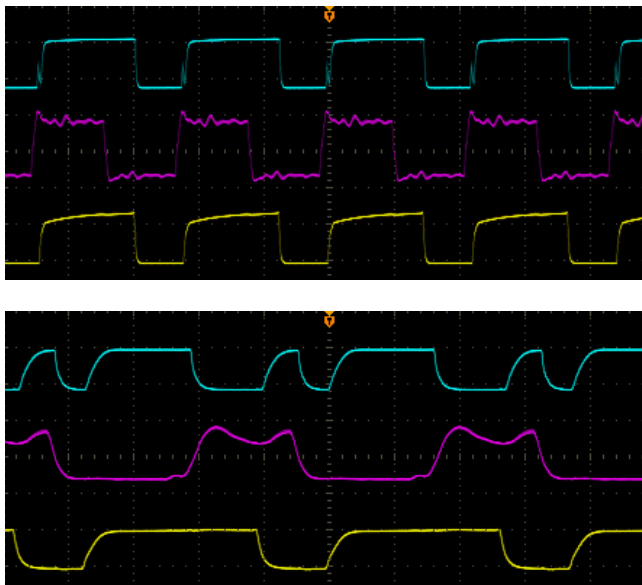


Fig. 5. Output signals of the phase change detection circuit (yellow for the output of the reference oscillator, purple for the output of the RC oscillator, blue for the output of the EXOR). C_1 1nF and $f_{Q1} = 1\text{MHz}$ in the upper figure, C_1 700pF and $f_{Q1} = 6\text{MHz}$ in the lower figure.

$I2B$ oscillator to the quartz oscillator signal [19].

A signal appears at the output of IC1 when a signal change occurs if there is a phase difference or change in phase between the output signal of Q_1 and the output signal of $I2B$. Since the frequency of $I2B_4$ is a function of C_1 , the output of IC_1 will also be a function of C_1 [20].

Fig. 5 shows the typical circuit in Fig. 4 at two capacitance values. For the experiment, two reference oscillators with different frequencies and the RC oscillator synchronized to them were also tested.

V. SPECIFIC APPLICATION, SPECIFIC FIELD

The fashion industry (Fig. 6) can also use the capacitive principle sensor and its special application as a threshold switch. The digital circuit shown in the figure can be used for this purpose. In this case, the output of the sensor is not the magnitude of the change in capacitance, but its state relative to a threshold value. In this case, this could mean a change in the position of a bodily surface (palm movement) [21]. In this case, a certain response can be given within the limits of the designer's imagination. Different lighting effects are shown in the figure [22].

The "closing in" of the environment in active clothing can be perspective. In this case, the wearer can expect reactions that depend on the social environment. The correlation between carbon dioxide-rich air and room occupancy warrants further investigation [23].

This application - by nature - requires a different capacitor design [24]. A sandwich structure capacitor (Fig.2) strapped with a porous material (special polymer) could be a perspective [25].

Also in perspective is the capacitance retrieval technique,

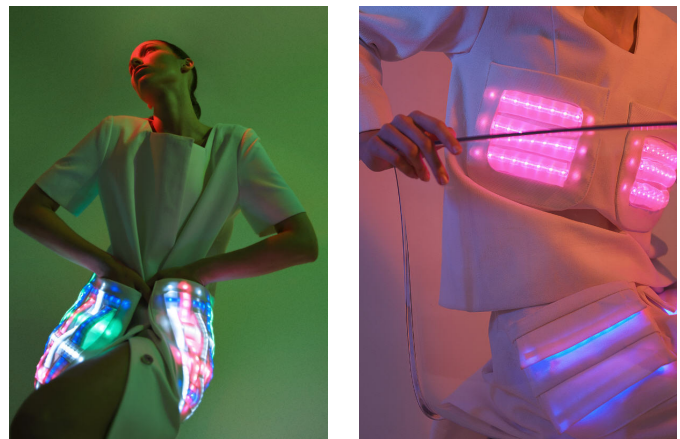


Fig. 6. Interactive fashion products on the actuator side. Color LEDs as indicators of the wearer's "emotion".

where the digital input circuit of a microcontroller is measured with an internal TIMER peripheral to measure the time-delay variation [26][27].

VI. CONCLUSIONS

With this article, we wanted to present a re-use of an existing device that can be used with some ingenuity and care. To this end, we have proposed two processing circuit techniques, with the proviso that a conventional technique using cap-sense can also be used if sufficiently sensitive. The capacitance-change detection circuit as a method predicts a new field of application in the context of interactive clothing.

All the areas highlighted in this paper are the beginning of a specific field that will require further research, often under laboratory conditions.

REFERENCES

- [1] Gy. Györök. Pwm tesztgenerátor. In *Alkotás/MTMT*, 2018.
- [2] L. O. Chua. *Computer aided nalysis of electronic circuit*. Prentice-Hall INC., New Jersey, USA., 1975.
- [3] J. F. Wakerly. *Digital design*. Prentice-Hall, Newjersey, USA., 2001.
- [4] B. Beszédes. Reliable presence and intrusion detection with collaborative sensor modules in electronic property protection systems. In *SAMI 2019, IEEE 17th World Symposium on Applied Machine Intelligence and Informatics*, pages 354–360, Feb 2019.
- [5] Tolner N.; Beszédes B. Cseh D. Györök, Gy.; Dávid. *Supervision of the operation of digital circuits by Embedded Microcontroller*. 13th International Symposium on Applied Informatics and Related Areas: AIS 2018, Óbuda University: Budapest, Hungary., 2018.
- [6] Gy. Györök. *Continuous Operation Monitoring of Electronic Circuits with Embedded Microcontroller*. IEEE 18th International Symposium on Computational Intelligence and Informatics (CINTI 2018), Budapest, Hungary, 2018.
- [7] A. Y. Baklanov; A.E.Baidildina; S. Grigoryeva; Gy. Györök. *Optimization of the work of LED lighting systems*. The Public fund editorial publishing center.
- [8] Gy. Györök. *Interactive Monitoring of Serial Electronic Circuit with Embedded Microcontroller*. IEEE 23rd International Conference on Intelligent Engineering Systems: proceedings (INES 2019), Budapest, Hungary, 2019.
- [9] C. K. Sedra, A. S.; Smith. *Microelectronic circuit*. Oxford University Press, New Yor, Oxford, 1998.

- [10] B. Beszédes; K. Széll; Gy. Györök. *A Highly Reliable, Modular, Redundant and Self-Monitoring PSU Architecture*. ACTA POLYTECHNICA HUNGARICA 17: 7, Budapest, Hungary, 2020.
- [11] Gy. Györök; B. Beszédes. *Concept of a Reliable Redundant Off-grid Power Supply Chain*. IEEE 13th International Symposium on Applied Computational Intelligence and Informatics: SACI 2018, Timisoara, Romania, 2019.
- [12] Grigoryeva Y.; Koshubaev Z. Baklanov, A. Y.; Györök. *Embedded LED lamp control system based on the ESP8266 microcontroller*. 14th International Symposium on Applied Informatics and Related Areas: AIS 2019, Óbuda University: Budapest, Hungary, 2019.
- [13] A. Zhaparova; D. Titov; A. Y. Baklanov; Gy. Györök. *Study of the Effectiveness of Switching-on LED Illumination Devices and the Use of Low Voltage System in Lighting*. The ACTA POLYTECHNICA HUNGARICA 12: 3, Budapest, Hungary, 2015.
- [14] Gy. Györök; B. Beszédes. *Duplicated Control Unit Based Embedded Fault-masking Systems*. IEEE 15th International Symposium on Intelligent Systems and Informatics: SISY 2017, New York, United States of America, 2017.
- [15] Gy. Györök; B. Beszédes. *Fault-tolerant Software Solutions in Microcontroller Based Systems*. 12th International Symposium on Applied Informatics and Related Areas: AIS 2017, Óbuda University, Budapest, Hungary, 2017.
- [16] Kaliyev Daniyar, Shvets Olga, and Gyrgy Gyrok. Uav based navigation system for forest fires monitoring. In *AIS 2021-16th International Symposium on Applied Informatics and Related Areas - Proceedings*, pages 47–50, 2021.
- [17] SHVETS OLGA, SMAKANOV BAUYRZHAN, Levente Kovcs, and Gyrgy Gyrok. Intelligent system for driver support using two classifiers for simulation. *JOURNAL OF THEORETICAL AND APPLIED INFORMATION TECHNOLOGY*, 100:4767–4782, 2022.
- [18] Gyrgy Gyrok and Shvets Olga. Adaptive ozone generator with embedded microcontroller. In *2022 IEEE 26th International Conference on Intelligent Engineering Systems (INES 2022)*, pages 155–158, 2022.
- [19] Bertalan Beszdes and Gyrgy Gyrok. Nagy megbzhatsg, egy tengely mentn forgathat fotovoltaiikus feszltsgforrs. *BIZTONSGTUDOMNYI SZEMLE*, 3:73–83, 2021.
- [20] Dr. Gy. Györök. Mikrokontrollerek és programozható analóg áramkörök egyedi alkalmazásai. In *Óbudai Egyetem, Disszertáció benyújtásának éve: 2015, Védés éve: 2015 Megjelenés/Fokozatszerzés éve: 2015*, pages 1–40, 2015.
- [21] Bertalan Beszdes, Kroly Szll, and Gyrgy Gyrok. Redundant photo-voltaic power cell in a highly reliable system. *ELECTRONICS*, 10:1253–1273, 2021.
- [22] Gy. Györök. *The Veterinary Horse Circuit for a Microcontroller Supervised System*. IEEE 16th International Symposium on Intelligent Systems and Informatics (SISY 2018), Budapest, Hungary, 2019.
- [23] Gyrgy Gyrok. Interactive monitoring of electronic circuits with embedded microcontroller. In *2021 IEEE 19th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, pages 000223–000228, 2021.
- [24] Dr. Gy. Györök. Számítógép perifériák ii. (háttértárak). In *Budapest, Magyarország: Óbudai Egyetem, Alba Regia Műszaki Kar (2015)*, pages 1–60, 2015.
- [25] Gy. Györök; B. Beszédes. *Highly Adaptive optocoupler degradation compensation in isolated feedback loops*. IEEE 12th International Symposium on Applied Computational Intelligence and Informatics: SACI 2018, Timisoara, Romania, 2018.
- [26] Gy. Györök; A. Y. Baklanov; B. Beszédes. *Extension of Nodal Voltage Method with the Thermosensing*. 12th International Symposium on Applied Informatics and Related Areas: AIS 2017, Óbuda University: Budapest, Hungary, 2017.
- [27] Gy. Györök; B. Beszédes. *Highly reliable data logging in embedded systems*. IEEE 16th World Symposium on Applied Machine Intelligence and Informatics: SAMI 2018, Seattle (WA), United States of America, 2018.

VLC technology concept in life support parameter management systems

A. Alimkhanova*, S.Grigoryeva* and G.Gyorok**

* D.Serikbayev East Kazakhstan Technical University/ Faculty of Information Technology, Ust-Kamenogorsk, Kazakhstan

** Obuda University/ Alba Regia Technical Faculty, Hungary

AAlimkhanova@ektu.kz, SGrigorieva@ektu.kz, gyorok.gyorgy@arek.uni-obuda.hu

Abstract— The article describes the concept of using LED lighting devices for data transmission in home automation systems. The scheme of the organization of the control system of life support parameters through the luminous flux using VLC technology is proposed. The capabilities of VLC technology allow you to implement intelligent control systems. In our work, it is proposed to take into account the influence of indirect indicators when regulating life support parameters. For the effective operation of this system, we suggest using existing indoor lighting sources. The automated life support management system is considered by the example of temperature data transmission.

Keywords : VLC, LED lamps, data transmission, control system.

I. INTRODUCTION

The home automation system is becoming more and more popular. People want to live in intelligent "smart" living spaces equipped with home automation systems. Such systems not only provide them with convenience, comfort, safety, but also reduce their daily expenses due to energy-saving solutions. Traditional control systems use solutions for wired connection of devices. However, the introduction of such systems requires laying cables and it is most rational to do this simultaneously with the construction of a house. Currently, in home automation, wireless technologies Wi-Fi, Bluetooth, Zigbee are mainly used for receiving and transmitting data, which use the radio frequency spectrum [1-3].

For the process of data exchange and the operation of applications of "smart" systems, it is necessary to form a special environment that will provide high data transfer rates, low signal latency, and maintain a high density of subscriber devices. Wireless technologies operating in the radio frequency range are overloaded and cannot fully meet the requirements of modern life.

The home automation system requires low cost, low power consumption and does not require high data transfer rates. In our work we propose to use optical wireless communication technology Visible Light Communication. VLC is a new wireless communication technology based on data transmission through LED lighting systems.

LED is currently replacing universal incandescent and fluorescent lamps from our lives. The advantages of LED lighting devices are long service life, low energy consumption and high light output, reliability, environmental friendliness. In addition, the LED has a high response sensitivity with the ability to support high-speed communication. In VLC, the LED performs the

functions of both communication and lighting, that is, with dual functionality.

Despite the fact that VLC technology has been developed and widely researched only for the last ten years, transmission systems with data transfer rates from 100 Mbit/s to 100 Gbit/s have been demonstrated, depending on the design of the LED, which is confirmed by research results [4-7].

In 2018, the International Telecommunication Union initiated the issue of using visible light waves for short-range communication, and the spectrum management working group presented a report on the possibility of reducing the congestion of the radio frequency spectrum. The report stated that "research on the development of new technologies in the combination of communication based on electromagnetic waves of visible light could create an interesting combination for solving problems of effective use of the radio frequency spectrum" [8].

The advancement of LED-based semiconductor lighting technology into lighting systems and the ability of powerful white LEDs to switch quickly and be controlled have influenced the development of wireless data transmission systems in the optical range integrated into lighting systems. Data transmission by light stream has become possible since LED lighting is mainly used in buildings, offices and industrial premises, which reduces operating costs.

The relevance of the topic of work is due to the need to meet the needs of a rapidly growing number of subscribers with wireless traffic with support for the quality of the required services. The research is aimed at developing applications using VLC technology for indoor communication, where it complements Wi-Fi and cellular wireless communication. The authors have previously conducted studies for possible applications [9-11].

II. MAIN PART OF THE RESEARCH

In order to properly and fully understand the use of data transmission technology with the help of powerful white LEDs for the organization of a life support parameter management system, it is necessary to form a concept. It is also worth considering the technical features of building the automation system itself using data transmission using the visible spectrum.

The system should integrate into existing control systems and complement them with its strengths, leveling the weaknesses of currently popular systems (increased radio noise, the possibility of signal interception, the final data transfer rate, etc.). An excellent analogy is the possibility of parallel use of mobile traffic via cellular

networks and Wi-Fi, which in everyday life allows the user to stay "online".

VLC technology implies data transmission when lighting devices are switched on. The question arises about the energy efficiency of the system. In the residential sector of modern cities, the peak load of wireless systems falls in the evening, i.e. when the lighting system is active in the premises. The proposed approach makes it possible to reduce the noise and load of the radio frequency range at this time.

Thus, the introduction of the proposed data transmission technology does not imply complete dominance and replacement of existing technologies. It is necessary to have a clear understanding that the proposed system should work in symbiosis with existing wireless data transmission technologies.

For effective integration of the system, it is advisable to use existing lighting sources in the room. Since at the moment in the Republic of Kazakhstan it is mandatory to use LED lighting with GOST-specified parameters in a number of state institutions, such integration is quite real. Despite the fact that every citizen can use any type of permitted lighting in his personal home, there has been a steady trend in the use of LED lighting in residential buildings in the Republic over the past years. Thus, the introduction of the proposed technology is possible everywhere in Kazakhstan on the basis of ready-made infrastructure.

It is proposed to implement the construction of a life support parameter management system using data transmission technology by means of LED lighting on the principle of Master – Slave operation (master device – slave device). Since the system is limited by the possibility of sending a data transmission signal, the Master (master device) must broadcast data to all devices at once, and the Slave (slave devices) must "listen" to the master device in real time. If the Slave understands that it is being accessed, it must accept the transmitted information and execute the command without forming a response message.

The proposed control system is based on the principle of data transmission over the visible spectrum in one direction. In simple words, in a room, LED lamps on the ceiling will be part of the master device, and the slave device is not equipped with lighting elements, but only a photodiode. Therefore, the slave device will not form a response signal in the direction of the Master, but only execute its commands. This approach corresponds to the concept of building unidirectional data transmission inherent in VLC technology.

When building a stable automation system for life support, it is necessary to take into account the presence of inter-circuit connections. Modern technologies allow us to take a fresh look at the principles of regulation and management and apply them to solve the problems of emerging problems.

To organize a life support management system, it is required to form control circuits for the following parameters – room temperature, air humidity, gas (air) analysis, window and door positions, light level.

The following organization of the control system is proposed.

The master controller is equipped with a touch panel and is located indoors. With the help of this controller, the entire control system is programmatically controlled. It is required to provide for the possibility of choosing manual control of the system and automatic operation mode.

The control unit consists of LED lighting devices and a controller. The controller is connected to the lighting devices. It is possible to carry out communication both with the help of cable products (wires) and with the help of wireless communication (Wi-Fi, Bluetooth). It is worth noting that when using wireless systems, there is both a positive side and a negative side. A positive feature is the ability to quickly integrate the system into an existing one. Additionally, such a construction will allow you to control the system remotely, but as a result, the unique increased security of the system from penetration based on data transmission technology using the visible spectrum is lost.

Interface controller. It is required to organize a data transfer interface between the receiver and the transmitter. An interface controller is used for this purpose. In an LED device, the interface controller performs a more complex function, since it is also part of the LED driver. In the receiver, the interface controller "decrypts" the analog signal received from the photodiode, brings it into the required sending format and executes the command according to the protocol.

The measuring unit includes a set of the following sensors: temperature sensor, humidity sensor, gas analysis sensor, end position sensors, light sensor. The sensors are connected to the controller and are located in accordance with existing standards and GOST standards. There are no specific requirements for installing sensors using VLC technology. The connection can also be made both wirelessly and wirelessly.

The block of the executive mechanism is considered as a set of actuators for this system: intelligent batteries / heating element / air conditioner, humidifier, forced exhaust, servos for windows and doors. The actuators are additionally equipped with photodetectors for the implementation of data transmission technology over the visible spectrum. The actuators are not interconnected, and do not have information and communication channels to communicate with each other. All intellectual work is performed by the master device, and the slave devices only obey commands.

Figure 1 shows a diagram of the described approach on the example of the organization of temperature control in the room.

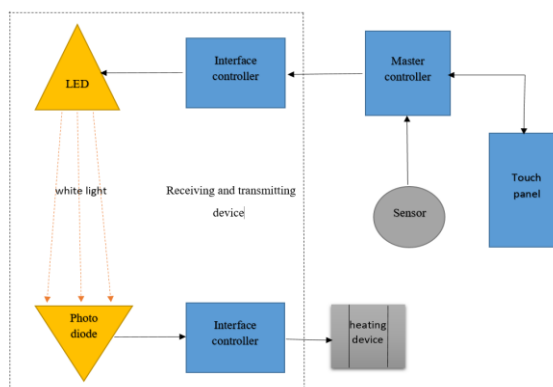


Fig. 1 – Diagram of an automated indoor temperature control system using VLC technology

There are several control circuits in the life support management system. All the data obtained can be divided into direct and indirect indicators for each contour. In each, highlight the possible relationship that affects the adjustable parameter. VLC technology provides for unidirectional communication in the data transmission system, so it makes sense to pay attention to indirect parameters in addition to taking into account direct control parameters during the control of each specific circuit. Despite the fact that a number of parameters do not have a clear linear relationship between each other, modern technologies allow you to process a large array of data, learn and adapt. This will eventually allow the formation of control signals by indirect actuators, ensure the stability of a particular circuit and increase the stability of the entire system.

Figure 2 shows a block diagram of the room temperature control circuit.

The following designations are introduced in the figure:

$X_{temp}(t)$, $X'_{temp}(t)$ – the signal at the input and output of the temperature sensor, respectively;

$X_h(t)$, $X'_h(t)$ is the signal at the input and output of the humidity sensor, respectively;

$X_{gaz}(t)$, $X'_{gaz}(t)$ is the signal at the input and output of the gas sensor, respectively;

$X_{1...n\ position}(t)$, $X'_{1...n\ position}(t)$ is the signal at the input and output of the position sensors, respectively;

$Y_{temp}(t)$ – control action.

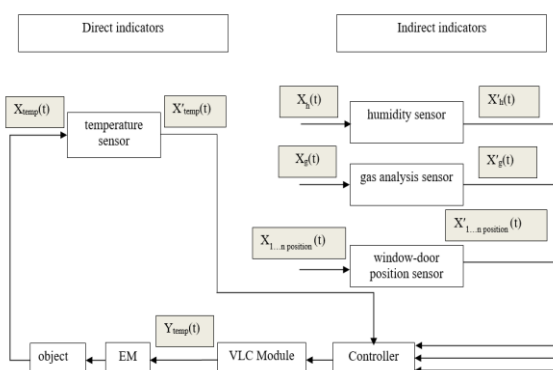


Fig. 2 – Block diagram of the temperature control circuit

The room temperature according to GOST is considered to be 18-24°C, and the minimum relative humidity for human comfort and health is 30% -60%.

There is a linear relationship between air temperature and relative humidity. In winter, the outdoor air on a foggy day at 0°C has a humidity of about 100%, and the heated air into the room up to 22 °C gives about 30% humidity.

By standards, every person who is in the room for an hour requires from 30 to 60 cubic meters of clean air per hour. With the help of the sensor of the air analyzer, it is possible to obtain data on the state of the air in the room. The inflow of fresh air into the room can be provided in various ways, for example, with the help of a forced exhaust or an air conditioner. In the proposed system, ventilation is carried out by opening / closing windows. With the help of a servo, it is possible to open the window according to the specified degree of opening (percentage

of opening). The end sensor is designed to check the position of the window, and as a protection against overshoot, the percentage of opening is set programmatically to the servo, and is tracked only in the zero position. It is logical that when windows and /or doors are opened for ventilation, both the temperature and humidity of the room will change.

Having made a number of assumptions, and using known linear dependencies, it is possible to calculate the behavior of the control object.

Figure 3 shows a block diagram of the humidity control circuit of the room.

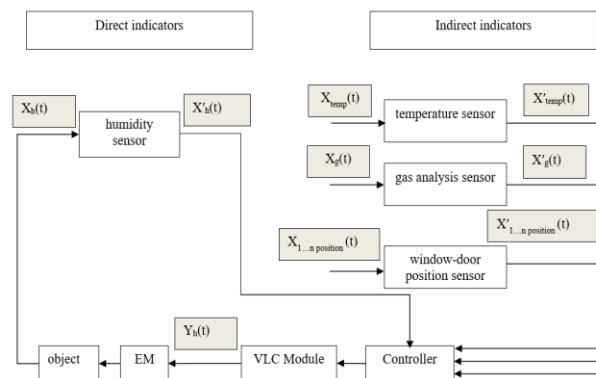


Fig. 3 – Block diagram of the humidity control circuit

According to state standards, indoor air quality is an important parameter for life support. The humidity in the room should be maintained from 30-60%.

Figure 4 shows a block diagram of the control circuit for the composition of indoor air.

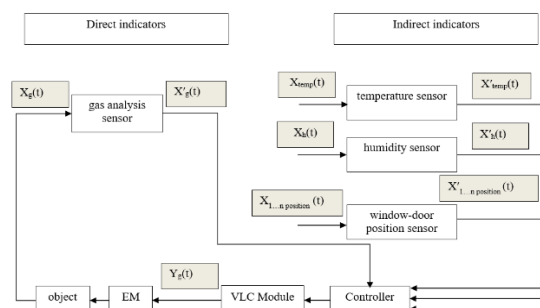


Fig. 4 – Block diagram of the air composition control circuit

Figure 5 shows a block diagram of the window/door position control circuit.

With this approach to building an automation system, an important step will be writing an algorithm for the operation of the system, taking into account the specific features of the control object. This task falls on the engineering staff, who will directly carry out the design, installation and commissioning works. The purpose of the work is not to provide a specific engineering solution, the scientific component, the concept and description of the process will allow the engineer to make calculations and configure the system for specific requirements. Depending on the parameters of the room and the functional purpose, both the dependencies in the formulas and the state standards for life support will differ.

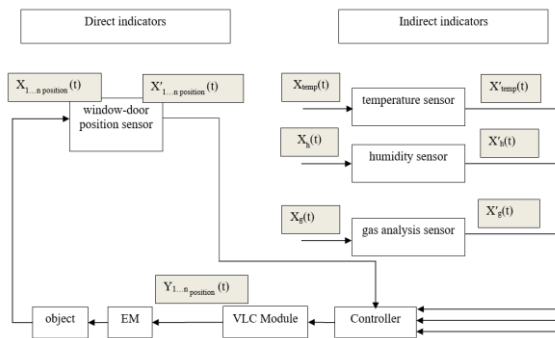


Fig.5 – Block diagram of the window/door position control circuit

Modern automated control systems can be implemented as intelligent. In our control system, taking into account in addition to direct control parameters and indirect indicators, it is possible to implement an intelligent system, since the proposed system has the following properties.

- the system has the ability to interact with the outside world using information and information-communication channels, in our case through visible light of lighting sources;
- the system is open, which allows you to replenish and acquire the necessary knowledge;
- the ability to predict the behavior of the control object (parameters) due to indirect data;
- the system has the property of compensating for inaccurate information about the state of the control object by training and implementing a new control algorithm;
- the system does not reduce the quality of functioning in case of violation of inter-circuit connections in the management process.

The advantage of intelligent automated control systems is the ability to train the system and adapt to impacts based on the analysis of the information received about the state of the control object, as well as information about the behavior of the surrounding world with which the object is in interaction. To implement such a system, it is necessary to have not only a database, but also a knowledge base, an algorithm for machine logical data entry, etc. Taking into account all the requirements for the construction of intelligent automated control systems, it becomes possible to organize the management of poorly formalized or complex technical objects by “training” the system as a whole in real time.

CONCLUSION

The article proposes and describes a new approach to the organization of an automated control system for the parameters of life support in a room using data transmission technology by means of LED lighting. The control system is based on the principle of data transmission over the visible spectrum in one direction, which corresponds to the concept of building VLC technology. Taking into account the unidirectionality in data transmission, a block diagram has been developed. In addition to direct regulation parameters, it is proposed to take into account the influence of indirect indicators. With this in mind, the control circuits have been developed.

REFERENCES

- [1] Frenzel L. Understanding solutions for the crowded electromagnetic frequency spectrum Available at: <https://www.electronicdesign.com>.
- [2] Swain F. Our hunger for wireless data is threatening to crash our communication networks - but there are ways to avoid disaster Available at: <https://www.bbc.com/future/article>.
- [3] Cisco Annual Internet Report (2018–2023). White Paper. 2020, P.35.
- [4] M. Khalid, G. Cossu, R. Corsini, P. Choudhury, E. Ciaramella: 1-Gb/s transmission over a phosphorescent white LED by using rate adaptive discrete multitone modulation, *IEEE Photon. J.*, Vol. 4, Issue 5, 2012, pp. 1465–1473.
- [5] Azhar, T. Tran, D. O'Brien: A Gigabit/s indoor wireless transmission using MIMO-OFDM visible-light communications, *IEEE Photon. Technol. Lett.*, Vol. 25, No. 2, 2013, pp. 171–174.
- [6] Ricardo X. G. Ferreira, EnyuanXie, Jonathan J. D. McKendry, et al.: High Bandwidth GaN-Based Micro-LEDs for Multi-Gb/s Visible Light Communications, *IEEE Photonics Technology Letters* Vol.28(19), 2016. DOI:10.1109/LPT.2016.2581318
- [7] D.O'Brien, S. Rajbhandari, H. Chun: Transmitter and receiver technologies for optical wireless, *Royal Society*, Vol. 378 Issue 2169, 2020. doi.org/10.1098/rsta.2019.0182
- [8] The use of visible light waves for broadband communication. ITU-R report. Geneva, 2019. – URL: https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-SM.2422-2018-PDF-R.pdf.
- [9] S.Grigoryeva, A.Baklanov, A.Alimkhanova, A.Dmitriev, Gy.Györök. Usage of Light-Emitting Diode Lighting and Visible Light Communication Technology for Temperature Control // *ActaPolytechnicaHungarica*. Budapest.2021. – vol.18, Issue Number 4 – pp.7-24. doi:10.12700/APH.18.4.2021.4.1
- [10] S. Grigoryeva, A. Baklanov, A. Alimkhanova, “Organization of an Automated Heat Modes Control System through LED Lighting,” 2020 International Scientific Conference “FarEastCon”/ Far Eastern Federal University (FEFU), October 2020.
- [11] S. Grigoryeva; A. Alimkhanova; E. Grigoryev; G. Gyorok Integration of the Data Transmission Function into Light-Emitting Diode Lighting Systems // 2021 IEEE Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT), 15-16 May 2021, Yekaterinburg, Russia

Indoor LED lighting control based on embedded systems

Bolatova Anar

Faculty of Information Technology
D.Serikbayev East Kazakhstan Technical University
Ust-Kamenogorsk, Kazakhstan
bolatova.anar@mail.ru

Svetlana Grigoryeva

Faculty of Information Technology
D.Serikbayev East Kazakhstan State Technical University
Ust-Kamenogorsk, Kazakhstan
SGrigorieva@mail.ru

Abstract—This article describes a specific approach in designing an office lighting system. An intelligent lighting control system based on LED lighting devices with built-in microcontrollers is considered. A portable small-sized sensor is proposed to use in order to control the light level and color spectrum emitted by the LED lighting device in real time. The system for monitoring and controlling lighting parameters is built using Wi-Fi wireless data transmission technology. The proposed organization of the lighting system allows for detailed and individual adjustment, taking into account the specified operating parameters and user requirements.

Keywords— *smart system, LED lighting, wireless technology, embedded systems*

I. INTRODUCTION

Interest in the design and promotion of intelligent lighting systems is based on the development of electronic information processing devices, info-communication technologies, and is also stimulated by energy conservation programs at the international and national levels [1,2].

Statistical data show an increase in the production and consumption of electricity in Kazakhstan. Almost 10 billion kW are spent annually on lighting needs, which is 12-14% of all electricity generated in the country [3]. It should be noted that in Kazakhstan this consumption indicator is less than the world one, which is 19% [4]. However, this is not due to the use of energy-efficient lighting devices and the introduction of modern lighting control systems, but to the high consumption of energy resources in the industry and the energy sector itself.

The total potential for reducing energy consumption for Kazakhstan is estimated at 30% [5]. The lighting industry is one of the areas with great energy-saving potential. The share of electricity consumption by lighting fixtures in public and residential buildings in Kazakhstan is 40% - 60%, which is much higher than in European countries. The costs of these key resources are constantly growing and an understanding of how these resources are spent is needed. This requires managing them and optimizing their use. At the same time, one of the possibilities to "take" resources under control is associated with lighting systems, which can be made not only much more cost-effective but also can be turned into an intelligent infrastructure. Therefore, one of the priority areas for reducing energy consumption is to reduce energy consumption for lighting in this sector.

Saving electricity consumption for lighting can be achieved through the use of energy-saving lighting lamps, as well as through the introduction of intelligent control systems.

The development of semiconductor lighting technology has made it possible to use LEDs in lighting devices for lighting systems. Powerful white LEDs used as light sources are distinguished by a long service life, high luminous flux

with low power consumption, high-quality lighting, the ability to control the luminous flux without compromising technical and operational characteristics. The high cost of LED lighting, which was initially a negative indicator and strongly influenced user demand, is steadily declining. In combination with improved reliability and uninterrupted operation indicators, solution of operational problems in connection and installation, as well as real financial savings due to reduced energy consumption, the price factor fades into the background. Energy-saving LED luminaires are used to solve problems of lighting a workspace in an office. They provide the required level of illumination, the necessary color rendition, and contrast for comfortable work throughout the entire declared service life.

Up to 40% of the electricity spent on lighting can be saved by taking into account natural light, disconnecting the load without the direct presence of people in the room, and planning various scenarios of the lighting environment [6]. Thanks to their technical and operational characteristics, LED light sources provide the opportunity to design intelligent lighting control systems for the workspace.

The main design criteria for modern lighting systems are energy savings and satisfaction of consumer demands. Open office layouts are the most used type of office space, where the illumination of each work area is the result of co-lighting from multiple ceiling lights. Existing commercial projects of lighting systems for office premises are aimed at reducing the cost of consumed energy and show good results through the use of different approaches in control systems [6-9]. It should be noted that the creation of a comfortable light environment, taking into account the specifics of work and user preferences, is most often overlooked or not considered at all. In addition, the installation of a control system in existing buildings with an established engineering infrastructure requires large financial investments due to the need for a global modernization.

The intelligent lighting optimization algorithm formulates lighting control as a linear programming problem, i.e. minimize energy consumption and satisfy user preferences. Currently, much less attention is paid to individual visual comfort than energy conservation. Studies conducted in standard office spaces show a positive correlation between satisfaction with the quality of the lighting environment and employee productivity. The data on the differences in the requirements for lighting for the implementation of tasks that differ in function are given [10]. More often than not, satisfying personal visual comfort is seen as a counterbalance to energy conservation. However, if the user is provided with the ability to independently adjust the lighting environment settings in accordance with his preferences in addition to the existing lighting scenario management system, then user satisfaction can be increased without increasing energy costs.

The organization of an individual comfortable light environment is relevant in the design of lighting control systems in an office since lighting devices are usually connected and grouped into zones in which the same uniform illumination is provided. The widespread development and implementation of embedded systems based on microcontrollers allows the introduction of modern methods of automation and control to solve this problem. By leveraging the versatility of wireless networking technologies embedded in today's sensors and actuators, an intelligent lighting control system can be created that simultaneously maximizes user satisfaction and conserves energy in offices.

In this paper, the authors describe their own approach to designing an intelligent LED lighting control system in an office premise with common space. It is proposed to use wireless data transmission technology to control the level of illumination and color spectrum based on embedded systems to solve the problem of efficient management of electricity and space at the same time.

II. LIGHTING CONTROL SYSTEM USING EMBEDDED SYSTEMS

The basis for research and implementation of our own approach in the design of an intelligent lighting system for an office with a common space was the educational laboratory "Intelligent life support systems" developed on the basis of D. Serikbayev VKTU. The appearance is shown in Figure 1.



Fig. 1. Teaching laboratory: 1 - Siemens controller; 2 - Touch panel; 3 - LED lamp

All modules of the system are controlled by the Siemens Simatic S7-400 controller. The software allows the control and monitoring of parameters in the room. The resources of the program are located on the webserver and allow connection to the system remotely. Administration levels are provided. For visualization and ease of control, the system is equipped with a touch panel. When switching to manual control, it is possible to change the scenario, set new settings.

This system has the following main modules (Figure 2):

- water supply control module - controls the absence of leaks in the water supply and sewerage systems;
- access control module - controls access to the room, and also monitors the movement and presence of people in the room to execute the developed scenarios;

- temperature control module - controls the temperature in the room and controls the valves of the heating system;
- lighting control module - controls the level of lighting in the room depending on the time of day, controls the level of the luminous flux of LED lamps, taking into account natural light, and also, turns on / off all lamps in the room depending on the information received from the access control module.

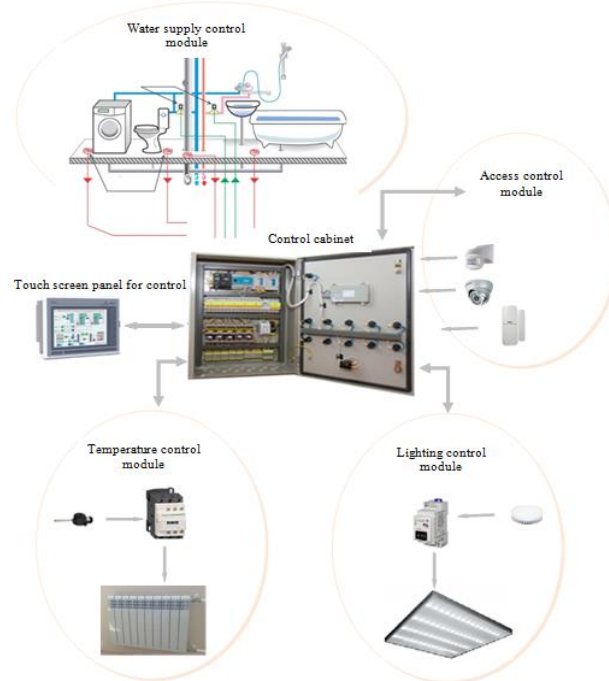


Fig. 2. Modules of the "Smart House" complex

Taking a closer look at the lighting control subsystem, it works on the basis of data received from motion sensors installed on doorways and light sensors installed in the middle of the ceiling of the room. The control algorithm takes into account the following parameters:

- motion sensors read the number of people present in the room. Upon entering the room, lighting devices are automatically turned on. In the absence of people in a given period of time, the lighting is turned off.
- light sensors take into account natural light and adjust the lighting level in the room in accordance with the specified one.

This approach is classic when designing a lighting control system.

Most of the existing projects in the "Smart Home" system implement control modules focused on the use of powerful controllers. The advantages of such systems are large computing capabilities and the integration of connecting a large number of devices to organize a comfortable environment and control life support parameters. However, the detriments are the high price, the incompletely used functionality and power of such a system, the complexity of integration into the existing engineering and informational infrastructure. To create an energy-efficient system with high functionality, this investigation proposes to integrate a lighting control module into an existing system, which includes LED lamps with built-in microcontrollers and a portable sensor.

III. PRACTICAL RESULTS

Description of a general approach to system design.

The main computing power is taken over by the master controller. It is suggested to use a single-board computer Raspberry Pi, which is much cheaper than industrial controllers. For the convenience of user control and visualization of the system operation, a touch screen monitor can be installed additionally.

The slave device will be LED ceiling lights. In this case, it is advisable to use two or more lamps.

For monitoring and control, its own portable small-sized light and color spectrum sensor with low power consumption is used.

Wi-Fi wireless technology is used to transfer data and control signals between elements of the lighting system. With this approach, both internal (individual) system configuration and connection to a common network with the possibility of remote control are possible.

The main elements of an intelligent lighting system based on embedded systems are shown in Figure 3.

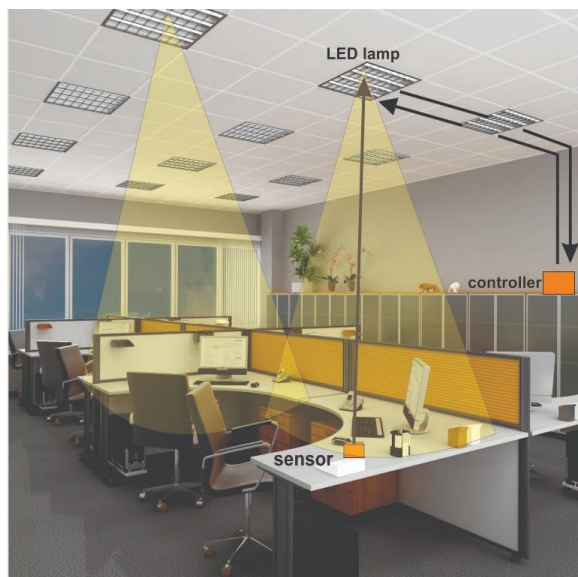


Fig. 3. General scheme

Modern microprocessor electronics allow the development of a low-cost smart lighting device with the ability to control and transmit data wirelessly. For implementation, it is proposed to use a microcontroller on the Espressif ESP8266 chip. This microcontroller has established itself at the present time due to its low cost, as well as the opportunity provided for the development of devices supporting the concept of the "Internet of Things" based on data transmission over a Wi-Fi network. An additional advantage when choosing a microcontroller is the presence of a NodeMCU development board.

Espressif offers a new line of ESP32 microcontrollers with support for wireless data transmission over two communication channels - Wi-Fi and Bluetooth, with significant computing power and data transmission security. [11].

Both chips have become the market leaders of the "Internet of Things" concept in their respective segments,

providing connection to a huge number of devices around the world. However, in this project, there is no need to overpay for functionality that will not be used.

As analogy with the lighting device, the developing portable sensor will be equipped with the same ESP8266 microcontroller. The advantage of the selected microcontroller in the sensor design is its low power consumption, as well as the ability to operate in various energy-saving modes. This parameter is very important in terms of portable sensor design. In the device under development, it is suggested to use two sensitive photocells. One of them will react to the ambient light level in the room, and the second will analyze the visible light spectrum. Both parameters will be read in real-time.

Description the operation algorithm of the control system.

The user installs the handheld transducer on the work surface. To get started, it should be activated. The activation process is confirmed by a light indication on the handheld sensor and on the touchscreen of the master controller. The next step is setting the parameters. This is possible both for a group of sensors and for each sensor separately. The system can operate in two modes - user and automated. In the user mode, a person in the room parameterizes the system independently in accordance with his requirements and preferences. Using the touch screen connected to the master controller, the user can set the required parameters of the illumination in the room. To reduce energy costs and set the optimal characteristics of the working surface, an automated version of room lighting is provided.

After activating the portable sensor, the system needs to determine it in the space of the room and bind it to the LED lamp. For this purpose, the system transmits the binary sequence of each lighting fixture using visible light communication technology. The sensor photocell reads the signal and transmits it to the master controller. There is a binding of the lighting fixture and the measuring device. The next step is to compare the measured data with the setpoints for regulation.

CONCLUSION

The proposed lighting control system will allow real-time adjustment of the optimal or user-defined level of illumination of the working surface. Leveraging the ease of integrating wireless networking with existing lighting equipment, the fixtures are connected to a portable, small-sized sensor via a Wi-Fi link to provide individual switching and dimming. The proposed approach to the organization of the LED lighting control system will reduce the computing power in the system, avoid installing a system of stationary sensors, and provide point and individual settings, taking into account the specified operating parameters.

REFERENCES

- [1] Resolution of the Government of the Republic of Kazakhstan. 29.08.2009. № 904. On approval of the Program «Energy Saving – 2020».
- [2] The concept of development of the electric power industry of the Republic of Kazakhstan until 2035 dated 26.09.2022.
- [3] A.Babko, S.Inyutin. Energy and light audits in buildings and street lighting. Astana: Publishing ..., 2014. 174 p.
- [4] International Energy Agency. Light's Labour's lost. Policies for Energy-efficient Lighting, OECD/IEA, 2006, 516 p.

- [5] S. Grigoryeva, A. Baklanov, D. Titov, V. Sayun, E. Grigoryev, "Analysis energy efficiency of automated control system of LED lighting," International Siberian Conference on Control and Communications, 2017, Astana.
- [6] Y. Aussata, A. Rosmanisa, S. Keshav, "A power-efficient self-calibrating smart lighting system," Energy and Buildings, Volume 259, 15 March 2022. <https://doi.org/10.1016/j.enbuild.2022.111874>
- [7] S. Bunjongjit, S. Ananwattanaporn, A. Ngaopitakkul, Ch. Jettanasen, T. Patcharoen, "Design and Application of Daylight-Based Lighting Controller on LED Luminaire," Appl. Sci. 2020, doi:10.3390/app10103415
- [8] K.R. Wagiman, M.N. Abdullah, M.Y. Hassan, N.H.M. Radzi "A new metric for optimal visual comfort and energy efficiency of building lighting system considering daylight using multi-objective particle swarm optimization," Journal of Building Engineering, Vol. 43, 2021, doi:10.1016/j.jobee.2021.102525
- [9] F. Suna, J. Yua, "Indoor intelligent lighting control method based on distributed multi-agent framework," Optik, Vol. 213, 2020. <https://doi.org/10.1016/j.ijleo.2020.164816>
- [10] L.T. Doulos, A. Tsangrassoulis, et. all, "Examining the Impact of Daylighting and the Corresponding Lighting Controls to the Users of Office Buildings," Energies, Vol. 13(15), 2020, doi:10.3390/en13154024
- [11] Announcing ESP32-C6: a Wi-Fi 6 + Bluetooth 5 (LE) SoC, https://www.espressif.com/en/news/ESP32_C6, 9.04.2021

Development and Prototype Production of a Full-spectrum LED Lighting Driver Circuit

Bertalan Beszédes
Óbuda University
Alba Regia Technical Faculty
Székesfehérvár, Hungary
beszedes.bertalan@uni-obuda.hu
<https://orcid.org/0000-0002-9350-1802>

Bence Márta
Óbuda University
Alba Regia Technical Faculty
Székesfehérvár, Hungary
martabence2000@gmail.com

Abstract—This paper presents an optimised DC/DC control solutions for full spectrum LED lighting equipment. It also deals with the proposed electronic structure and the key parameters of a lighting system. The research aimed to show various PCB designs for diverse environments, and different converter topologies for dissimilar supply voltages. The article deals with the control and diagnostic possibilities of power light-emitting diodes, with wide spectrum range of an experimental environment.

Keywords— COB LED driver, light spectrum control, high power LED, full-spectrum LED driver, DC-DC converter topologies, LED DC/DC converter, PCB prototype production

I. INTRODUCTION

Numerous solutions exist to enhance LED driving [1-3], within this the DC to DC conversion based current sources are have got a wide range of solutions [4, 5]. Robust solutions are hardened for external disturbances and also well controllable [6, 7]. Many of these solutions can be realized with different type of converters [8].

The experimental PCBs are based on modular sub-circuits, due to easier modifiability [9, 10], a prototype production line also helped to assemble the diver circuits [11-13]. The optimization of the layout is considered according to the following papers [14-16]. Real time data monitoring and logging is also a significant part of the development [17-20].

One of the effective ways of controlling the system is closed-loops and compensation of disturbance [21-23]. Various optimized solutions and compensating for uncontrolled disturbances is discussed in a series of papers [24-26]. Numerous schemes of control systems architecture are considered [27-31].

II. LIGHT SOURCES

Light Emitting Diodes (LEDs) emit light in a narrow band of the spectrum (Fig. 1), opposed to more traditional methods of lighting such as incandescent light bulbs or fluorescent lamps. LEDs also generally have well defined directional radiation characteristics. such characteristics make it possible to achieve almost any spectral distributions with various areal distributions.

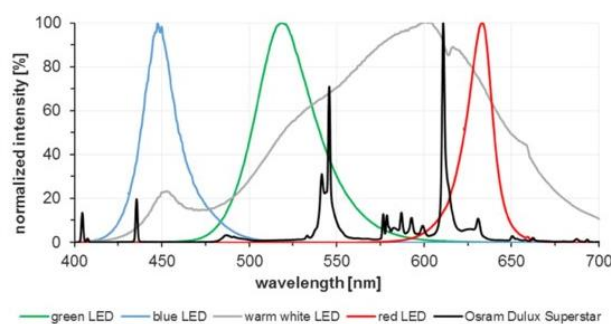


Figure 1. Emission spectra of different LEDs [31]

III. PROTOTYPE PRODUCTION

The purpose of the prototype PCB production is to create a module that can be used to perform various tests during product development. Rapid prototype manufacturing processes significantly accelerate the product development, however, human intervention is required to operate it.

The fast PCB prototyping production line provide a good in-house research and development solution. It contains a PCB milling machine, a stencil printer, an UV exposure box, a pick and place station and a reflow oven, each module is an individual workstation.

EasyEDA is used as the ECAD software, and as an additional software Gerbv is also used for checking gerber files.

The boards are designed with SMD components. The PCB and the stainless steel stencil manufactured by an external company. First step in this type of prototype production process is to paste the PCB with a manual stencil printer (Fig. 2). The stencil should be in perfect alignment with solder pads on the PCB and the squeegee should be at a 45°-60° to the stencil. With a semi-automatic SMT pick and place machine, the electrical SMD components can be placed directly onto the pasted circuit board. After the component placement (see on Fig. 3), the PCB can go to the reflow oven for soldering. Through hole components like connectors are placed and soldered manually to the PCB.

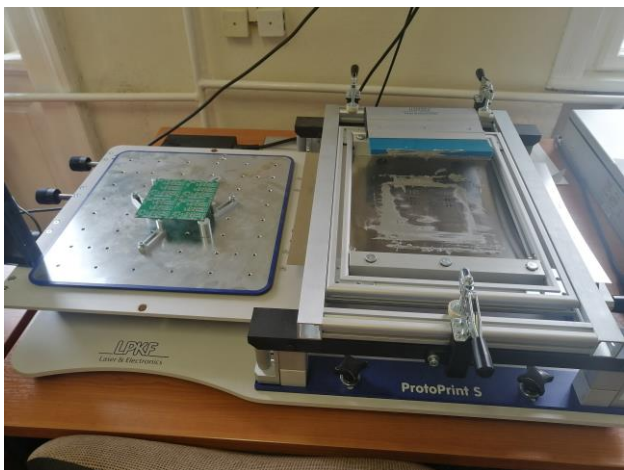


Figure 2. Manual stencil printer station

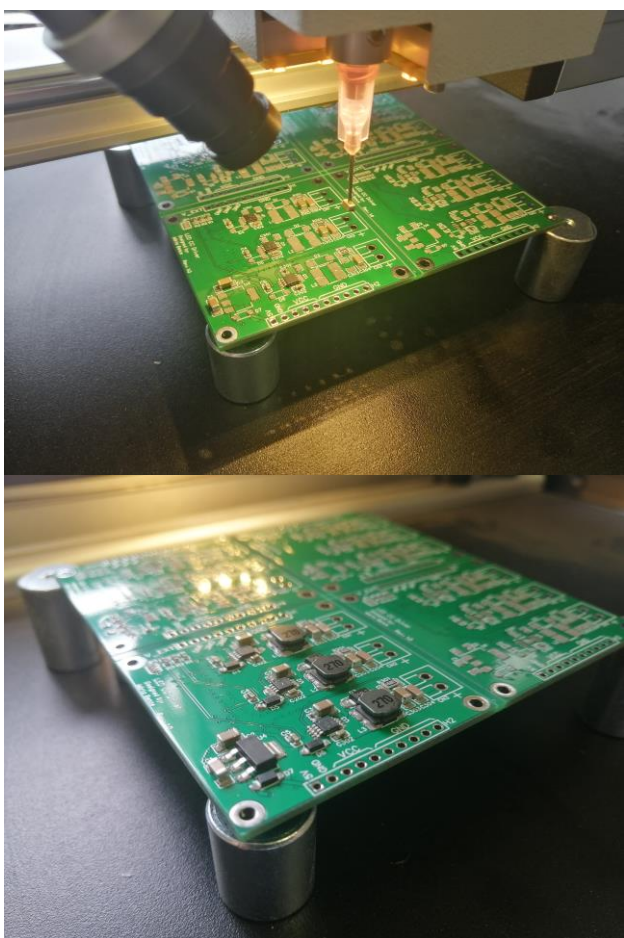


Figure 3. Semi-automated PCB manufacturing

IV. REALIZATION

A. Local lighting

Local lighting solutions are useful in many situations like soldering, handling tiny components, professional photography etc.

The realized circular PCB (see on Fig. 4 and Fig. 5) was made for use with a desktop magnifier. It uses the same driver circuitry as the one presented on Figure 6,

with three 1w LEDs connected to its output. for dimming control, it uses an ATTiny85 microcontroller which translates a potentiometer's analog output into a PWM signal. By using 3 of the PCBs, a full circle can be made.

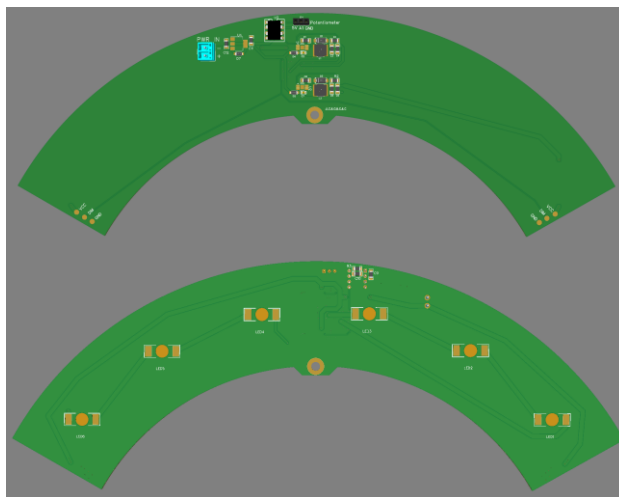


Figure 4. Front side of circular PCBs



Figure 5. Circular PCB

B. Grow power LEDs

As an experimental solution, 900mA buck type driver PCBs were created, see on Fig. 6. These boards are stackable in arrays of two or three. Due to the amount of heat produced by the connectors contact resistance and by the switching losses, a stack of three boards is only recommended when sufficient airflow and/or temperature sensing is provided.

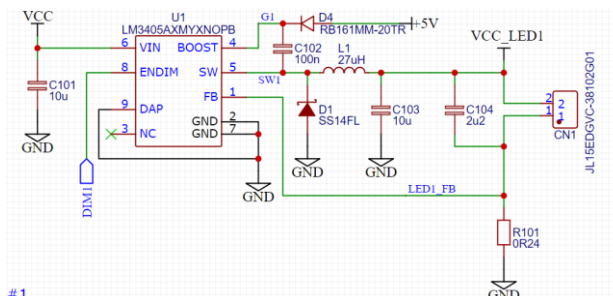


Figure 6. LM3405 LED driver circuit

The power supply of the application area is a 12V high current PSU. The driver board experimental environment can be seen on Fig. 7. The power consumption of the circuit in standby mode is 114mW. The result of the efficiency measurement depending on the control signal duty cycle can be traced in Fig. 8.



Figure 7. LED driver circuit testing environment

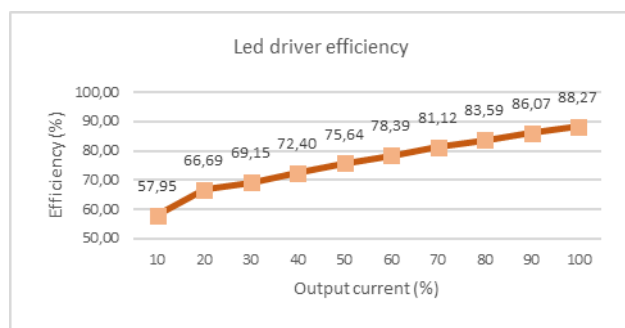


Figure 8. Led driver efficiency graph

C. Control circuit

A boost type 100mA driver board was also realized, this time with 6 channels and a microcontroller socket (ATmega328P or compatible models) for per-channel control capability. The channels can be individually controlled through analog potentiometers or an I2C bus. In I2C mode, all devices must be configured through a UART connection before use. After the successful configuration, any of the devices can be marked as master and control the array through its UART connection. Additional I2C devices, such as light or temperature sensors can also be added with an appropriate firmware. The schematic diagram can be seen on Fig. 9 and the 3D model is shown in Fig. 10.

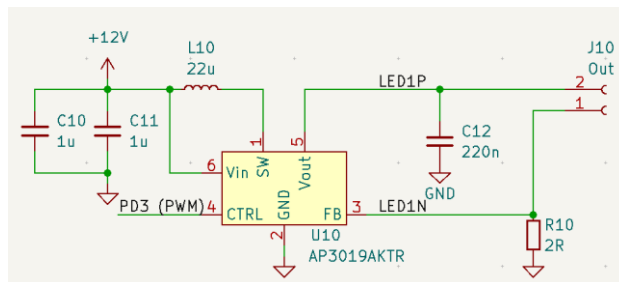


Figure 9. AP3019AKTR LED driver circuit

The array can be configured to work independently from any external devices or subordinate to an external control system shown on Fig. 11. These capabilities can also be combined for fail-safe operation.

In both modes, sensors can be attached to the system, creating a feedback loop for any desired parameter of the system, such as light intensity, spectral distribution or even input/output currents.

The physical realization of the intervening elements of the complex spectrum spatial illumination can be seen in Figure 12.

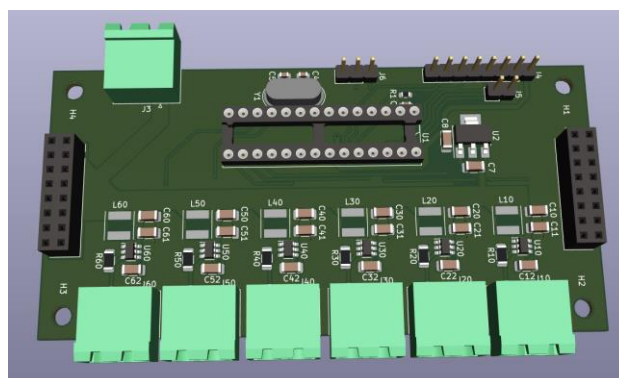


Figure 10. AP3019AKTR LED driver circuit 3D model

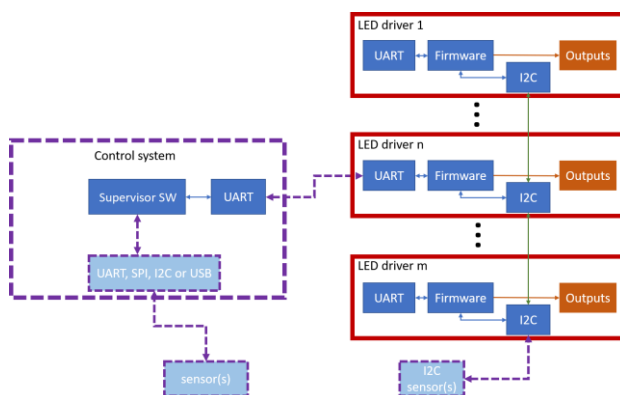


Figure 11. LED driver circuit control system architecture



Figure 12. Grow LED matrix

V. CONCLUSION

The received results showing the effectiveness of the developed power LED driving circuit concept and the realizations also confirms the functionality of the proposed solution. The results of the development showed that the use of current driver circuits are capable to use for full-spectrum LED lighting, and they can be controlled properly. The shown modular design can also be used for other applications. The proposed architecture is easily scalable, user-friendly, and robust. The article describes the relationship between the system components, also demonstrates the control and measuring options. Eventually conclusions were drawn about the feasibility and requirements of such a system to make it possible to implement.

ACKNOWLEDGMENT

The authors would like to thank all the faculty staff and member that provide help and assistance throughout the project completion.

REFERENCES

- [1] Barwar, Manish & Sahu, Lalit & Tripathi, Prabhat & Bhatnagar, Pallavee & Chander, Hema & Gupta, Krishna & Guerrero, Josep. (2022). Demystifying the Devices Behind the LED Light: LED Driver Circuits. *IEEE Industrial Electronics Magazine*. 2-13. 10.1109/MIE.2022.3164526.
- [2] Musumeci, Salvatore. (2021). Passive and Active Topologies Investigation for LED Driver Circuits. 10.5772/intechopen.97098.
- [3] Ramazan Şenol, Kubilay Taşdelen, A New Approach for LED Plant Growth Units, *Acta Polytechnica Hungarica* Vol. 11, No. 6, 2014. pp. 57-71.
- [4] Verma, Deepak. (2020). DC-DC converter topologies for LED driver circuit: A review. *Science Journal of Circuits Systems and Signal Processing*. 14. 542-547. 10.46300/9106.2020.14.70.
- [5] Attila, Sáfár ; Bertalan, Beszédes. Educational Aspects of a Modular Power Management System. In: Orosz, Gábor Tamás (szerk.) AIS 2019 : 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University. Székesfehérvár, Magyarország : Óbudai Egyetem, (2019) pp. 163-166. , 4 p.
- [6] Baklanov, A. & Grigoryeva, Svetlana & Titov, D.. (2016). The practical realization of robustness for LED lighting control systems. 52-56. 10.1109/IFOST.2016.7884310.
- [7] Alexander Baklanov, Svetlana Grigoryeva, György Györök, Control of LED Lighting Equipment with Robustness Elements, *Acta Polytechnica Hungarica* Acta Polytechnica Hungarica Vol. 13, No. 5, 2016. pp. 105-119.
- [8] Xu, Renbo & Li, Yongzhi & Zhong, Lixin & Liu, Jiaming. (2014). Research of an Efficient LED Lighting Driver Based on Boost-Buck Converter. *Circuits and Systems*. 05. 153-159. 10.4236/cs.2014.56017.
- [9] Bertalan, Beszédes ; Károly, Széll ; György, Györök A Highly Reliable, Modular, Redundant and Self-Monitoring PSU Architecture *ACTA POLYTECHNICA HUNGARICA* 17 : 7 pp. 233-249. , 14 p. (2020)
- [10] György, Györök ; Bertalan, Beszédes Concept of a Reliable Redundant Off-grid Power Supply Chain In: Szakál, Anikó (szerk.) SACI 2019 : IEEE 13th International Symposium on Applied Computational Intelligence and Informatics : PROCEEDINGS Temesvár, Románia : IEEE (2019) 383 p. pp. 205-10. , 6 p.
- [11] Beszédes, Bertalan. Fast Printed Circuit Board Prototype Production Line in Educationan Environment. In: *EEITE2019*. (2019) Paper: 1 , 10 p. ISSN: 2654-2099
- [12] Beszédes, Bertalan. Electronic Design and Fast Prototipe Production - ECAD/ECAM. Budapest, Magyarország : Óbudai Egyetem, Alba Regia Műszaki Kar (2019) , 56 p. ISBN: 9789634491507
- [13] Györök, György ; Bertalan, Beszédes. Using Thermal Imaging Cameras to Test Electronical Systems. In: *IEEE - Kádár, Péter; Lamacchia, Francesco P.; IEEE (szerk.) IEEE CANDO EPE 2018 : 2018 International IEEE Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE)*. New York (NY), Amerikai Egyesült Államok, Piscataway (NJ), Amerikai Egyesült Államok : IEEE (2018) pp. 147-152. Paper: 8 , 6 p.
- [14] Beszédes, Bertalan ; Széll, Károly ; Györök, György Redundant Photo-Voltaic Power Cell in a Highly Reliable System *ELECTRONICS* 10 : 11 p. 1253 , 20 p. (2021)
- [15] Györök, György ; Bertalan, Beszedes Fault tolerant power supply systems In: Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016) Székesfehérvár, Magyarország : Óbudai Egyetem (2016) pp. 68-73. , 6 p.
- [16] A. Baklanov, S. Grigoryeva, Gy. Györök. Intelligent control of LED luminaries. 9th International Symposium on Applied Informatics and Related Areas - AIS 2014, Székesfehérvár, 2014. pp. 87-91 (ISBN:978-615-5460-21-0)
- [17] Bartos, Gaye Ediboglu, et al. "A Multilingual Handwritten Character Dataset: THE Dataset." *Acta Polytechnica Hungarica* 17.9 (2020).
- [18] Meng Qingyan ; Chen Xu ; Zhang Jiahui ; Sun Yunxiao ; Li Jiaguo ; Jancsó Tamás; Sun Zhenhui Canopy Structure Attributes Extraction from LiDAR Data Based on Tree Morphology and Crown Height Proportion *PHOTONIRVACHAK / JOURNAL OF THE INDIAN SOCIETY OF REMOTE SENSING* (0255-660X 0974-3006): 46 9 pp 1433-1444 (2018)
- [19] G. Györök and B. Beszédes, "Highly reliable data logging in embedded systems," *2018 IEEE 16th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2018, pp. 000049-000054, doi: 10.1109/SAMI.2018.8323985.
- [20] Xiaojiang, Li ; Weindong, Li ; Quingyan, Meng ; Chuanrong, Zhang ; Tamas, Jancso ; Kangli, Wu Modelling building proximity to greenery in a three-dimensional perspective using multi-source remotely sensed data *JOURNAL OF SPATIAL SCIENCE* , 16 p. (2016)
- [21] G. Simon and L. Sujbert, "Special issue on "Recent advances in indoor localization systems and technologies", *Applied Sciences*, Vol. 11, No. 9, paper 4191, May 2021.
- [22] L. Sujbert, G. Simon and G. Peceli, "An Observer-Based Adaptive Fourier Analysis," *IEEE Signal Processing Magazine*, Vol. 39 No. 4, pp.134-143, Oct. 2020.
- [23] G. Simon, G. Zachár, and G. Vakulya, "Lookup: Robust and Accurate Indoor Localization Using Visible Light Communication," *IEEE Trans. on Instrumentation and Measurement*, Vol. 66, No. 9, pp.2337 - 2348, Sept 2017.
- [24] Pogátsnik, M. (2021): Dual education: connecting education and the labor market, *OPUS ET EDUCATIO* 3 pp. 304-313.
- [25] S. Nosratabadi, K. Szell, B. Beszedes, F. Imre, S. Ardabili and A. Mosaavi, "Comparative Analysis of ANN-ICA and ANN-GWO for Crop Yield Prediction," *2020 RIVF International Conference on Computing and Communication Technologies (RIVF)*, 2020, pp. 1-5, doi: 10.1109/RIVF48685.2020.9140786.

- [26] Pogatsnik, M., Kendrovics, R. "Communication and Reading Comprehension among Informatics and Engineering Students," 2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI), Herlany, Slovakia, 2020, pp. 235-240
- [27] Ladislav Fozo ; Rudolf Andoga ; Radovan Kovacs. Thermodynamic cycle computation of a micro turbojet engine. 2016. In: CINTI 2016. Danvers : IEEE, 2016 P. 000075-000079. ISBN 978-1-5090-3909-8
- [28] FÖZŐ, Ladislav ; ANDOGA, Rudolf ; KOVÁCS, Radovan. Experimental identification of a small turbojet engine with variable exhaust nozzle. 2015. In: CINTI 2015. Danvers : IEEE, 2015 P. 65-69. ISBN 978-1-4673-8519-0
- [29] KOMJÁTY, Maroš ; FÖZŐ, Ladislav ; ANDOGA, Rudolf. A Digital Diagnostic System for a Small Turbojet Engine 2013. In: Acta Polytechnica Hungarica. Vol. 10, no. 4 (2013), p. 45-58. ISSN 1785-8860
- [30] ANDOGA, Rudolf ; FÖZŐ, Ladislav ; MADARÁSZ, Ladislav - KAROL, Tomáš. Innovative approaches in modeling, control and diagnostics of small turbojet engines. 2013. In: Acta Polytechnica Hungarica. Vol. 10, no. 5 (2013), p. 81-99. ISSN 1785-8860
- [31] Schachtner, Josef & Bayer, Patrick & Jacobi von Wangelin, Axel. (2016). A flow reactor setup for photochemistry of biphasic gas/liquid reactions. Beilstein Journal of Organic Chemistry. 12. 1798-1811. 10.3762/bjoc.12.170.

Artificial Lighting Experimental Environment in Agriculture for Seed Germination

Bertalan Beszédés
Óbuda University
Alba Regia Technical Faculty
Székesfehérvár, Hungary
beszedes.bertalan@uni-obuda.hu
<https://orcid.org/0000-0002-9350-1802>

Bence Márta
Óbuda University
Alba Regia Technical Faculty
Székesfehérvár, Hungary
martabence2000@gmail.com

Péter Udvardy
Óbuda University
Alba Regia Technical Faculty
Székesfehérvár, Hungary
udvardy.peter@uni-obuda.hu
<https://orcid.org/0000-0002-8897-9326>

Abstract—This paper explains the basic biology behind the idea of using artificial lighting for acceleration of a plant's development. The article deals with the control of power light-emitting diodes, the optimization and diagnostic possibilities of the spectrum are part of the presented experimental environment. Eventually conclusions were drawn about the feasibility and requirements of such a system to make it possible to implement.

Keywords— photoreceptors in plants, greenhouse lighting, chlorophyll a, chlorophyll b, artificial seed germination, grow LED driver, light spectrum control

I. INTRODUCTION

Numerous solutions exist to enhance the yield of soil, ranging from chemical compounds to systems with full environmental regulation. Many of these solutions use some form of complementary or complete illumination, usually using LED technology.

The experimental environment is based on modular subsystems, due to easier modifiability [1-5]. The optimization of architectural structure is considered according to the following papers [6-8]. Real time data monitoring and logging is also a significant part of the research [9-12]. One of the effective ways of controlling the operation is closed-loops and compensation of disturbance [13-15].

II. BIOLOGICAL BACKGROUND

A previous study has shown that chlorophyll a, chlorophyll b, cry2 and Phytochromes play a significant role in the development of plants. [16]

A. Chlorophyll a and chlorophyll b

Plants absorb the light spectrum in a similar but broader range as the human eye, but unlike humans, they absorb best red and blue light. Chlorophyll is one of the main molecules that enable plants to absorb and use the energy of light thus making photosynthesis work. In the higher plants, two main types of chlorophyll can be found with slightly different absorption curves. The small difference in the absorption characteristics allows them to capture different wavelengths, catching more of the sunlight spectrum. All oxygenic, photosynthetic organisms use chlorophyll a, but are differentiated by additional pigments such as chlorophyll b. Chlorophyll a

absorbs light at wavelengths of purple, blue and red range of the spectrum, but reflects the majority of light in the green range. Additional photosynthetic pigments expand the spectrum of absorbed light, increase the range of wavelengths used in photosynthesis. The addition of chlorophyll b a to chlorophyll a expands the absorption spectrum as shown on Fig. 1.

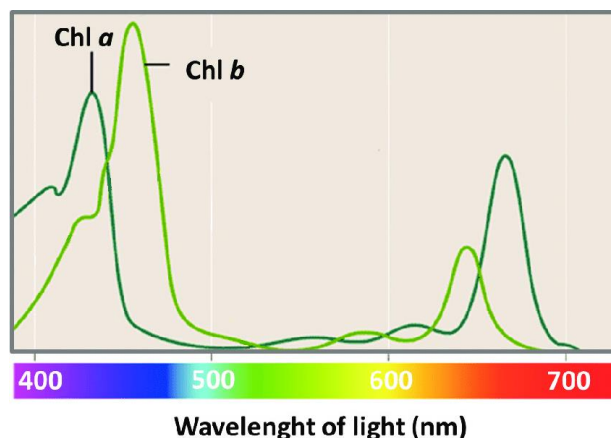


Figure 1. Absorption spectrum of chlorophyll a and chlorophyll b [17]

B. CRY2

Cryptochromes are a class of flavoproteins found in animals and plants. They capture light-related external stimuli and control the internal clock of plants. They are also associated with morphological reactions such as inhibition of stem elongation, dilation of the cotyledon, anthocyanin production, and photoperiodic flowering. Cryptochromes absorb UVA (ultraviolet), blue and green wavelengths. In plants, cryptochromes mediate phototropism in response to blue light. This response is now known to have its own set of photoreceptors, the phototropins.

Cry2 is responsible for the leaf and cotyledon expansion mediated by blue light. In transgenic plants, excessive expression of Cry2 increases the expansion of cotyledons stimulated by blue light, resulting in numerous broad leaves and no flowers instead of primary leaves and flowers. Cry2 genes have been shown [17] to delay flowering in continuous light and accelerate it in

long and short days, indicating that CRY2 may play a role in accelerating flowering time in continuous light.

C. Phytochromes

Phytochromes are responsible for induction of flowering and seed development, they regulate stem and leaf elongation, seed germination, chlorophyll synthesis, seedling elongation, size, shape, number and movement of leaves, timing of flowering in adult plants and "shade avoidance syndrome." Phytochrome-regulated responses are mediated by the ratio of red and far-red light, which affects the photostationary state of the phytochrome molecule. They can be classified as either type I, which are activated by far-red light, or type II activated by red light. Phytochromes also sense light, which causes the plant to grow towards it (phototropism). The above is illustrated in Fig. 2.

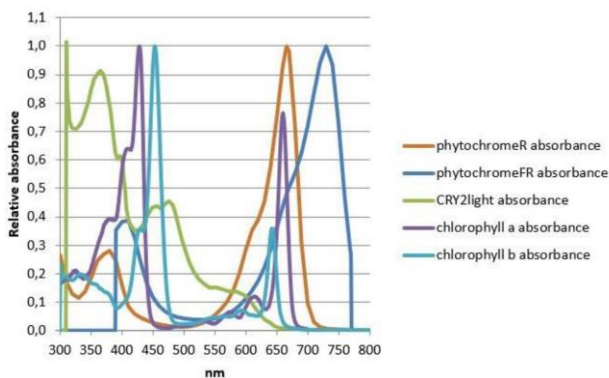


Figure 2. Relative absorbance of different photoreceptors in plants [16]

III. EXPECTATIONS FROM A SYSTEM

The main goal of an artificial lighting system is to provide the plants with light that has an optimal spectral distribution, while adopting to the circadian cycle of the plant. To achieve this, such a system needs to be able to keep track of the time of day and control the lighting of the environment accordingly.

The energy-efficiency of such a system can be greatly improved if the sun's own radiation is used and the artificial lights only provide supplementary illumination at the peak wavelengths of the absorption spectrum. This way only red and blue lights need to be used to achieve notable results as the rest of the spectrum is covered by the natural light. When using this technique, a system could easily adopt to a plant's natural circadian cycle just by monitoring the natural light, therefore it could be used in geographic areas of widely varying sunny hours.

IV. LIGHT SOURCE

A. Incandescent light

Incandescent light bulbs have a small initial cost compared to the cost of energy it is using over its lifetime. [20] They typically have a lifetime of around 1000 hours. [21]

The luminous efficiency of a typical 120V incandescent light bulb is 16 lm/W, compared to the 60 lm/W for

compact fluorescent lights and 150 lm/W of some LED sources [22].

B. Fluorescent lamps

While fluorescent lamp fixtures have a higher initial cost, compared to incandescent bulbs, because of the required external circuitry, it has lower energy consumption, which offsets for the initial costs. [23]

C. LEDs

Peak efficiency can be reached by using Light Emitting Diodes (LEDs) as they can emit light in a narrow band of the spectrum, as opposed to more traditional methods of lighting such as incandescent light bulbs or fluorescent lamps. LEDs also generally have directional radiation characteristics compared to the other mentioned light sources. These characteristics make LEDs a perfect candidate for this application as the relevant molecules' absorption curves have their peaks close to each other, therefore 2 well-chosen LED types can cover the most sensitive parts of the absorption spectrum. [24]

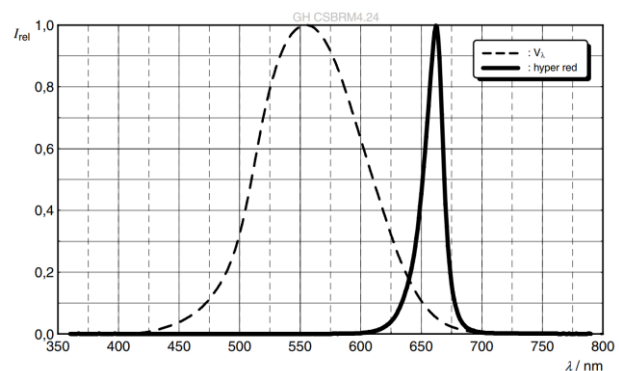


Figure 3. Relative Spectral Emission of OSRAM GH CSBRM4.24 LED [18]

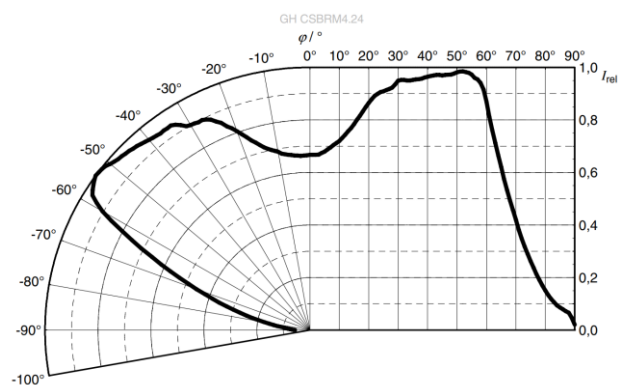


Figure 4. Radiation Characteristics of OSRAM GH CSBRM4.24 LED [18]

The LEDs' characteristics also allow it to be easily dimmed by using PWM controls or constant current sources (Fig. 5). [25] As different plants may require different ratios of the supplementary red and blue light for optimal development, this enables for more universal solutions. [26]

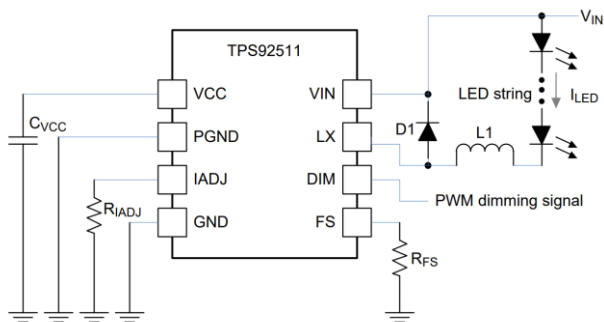


Figure 5. Example circuit for TPS92511 constant current LED driver IC [19]

D. Solution used for the experiment

For the experiment, 10w LEDs were chosen, together with a custom-made, LM3405 based constant-current driver circuit to provide them with a maximum of 900 mA (Fig. 6). The 10W IR COB LED needs more than 2A driving current, an individual current source circuit is added for the experimental setup for this purpose.

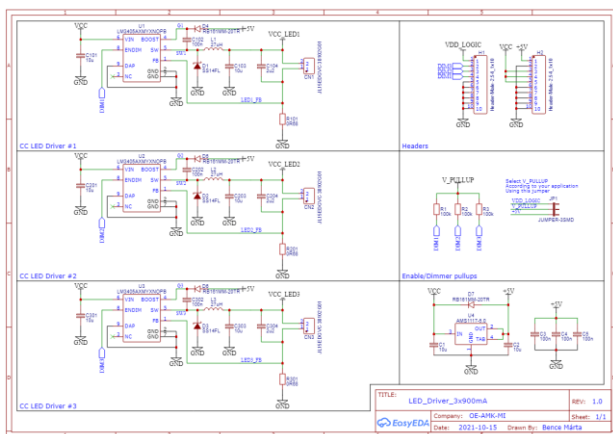


Figure 6. Schematic of the LED driver board

For practical reasons, the boards were designed in a way that they can be stacked on top of each other. The power and dimming signals can be fed to the PCBs through a base board, see on Fig. 7.

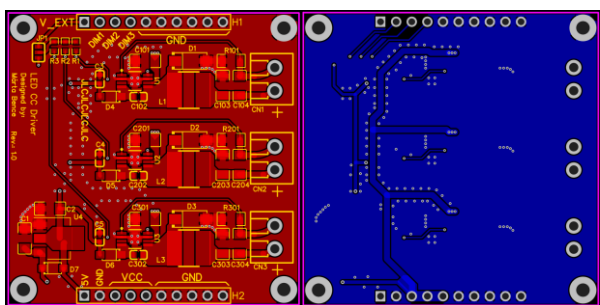


Figure 7. Top (left) and bottom (right) copper layer of the LED driver board

V. DESIGN OF THE EXPERIMENTAL ENVIRONMENT

The environment was designed to provide the plants with different wavelengths of light from above in a controllable manner (see Fig. 8). The test equipment consists of several hardware elements:

- Tray for the plants
- 10W LEDs with cooling
 - 4x red
 - 2x blue
 - 2x warm white
 - 1x infrared
- Constant current LED drivers
- Arduino UNO
- Raspberry Pi 3 with camera



Figure 8. Picture of the test setup

The Raspberry Pi is used to set the brightness of the LEDs and also take a photo of the plants every hour with a camera module (Fig. 9).

The Arduino UNO is used to generate the PWM dimming signals for the LED driver circuits.

The software running on the Raspberry Pi is created with the help of Node-RED (Fig. 10). The flow consists of only 4 nodes, excluding the UI elements.

The Arduino UNO's task is simply to interpret a packet of bytes sent by the Raspberry Pi and set its PWM outputs' duty cycles accordingly.

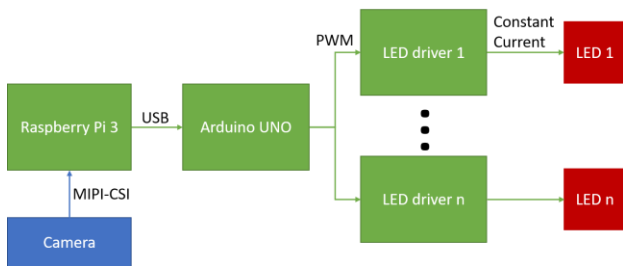


Figure 9. Block diagram of the hardware

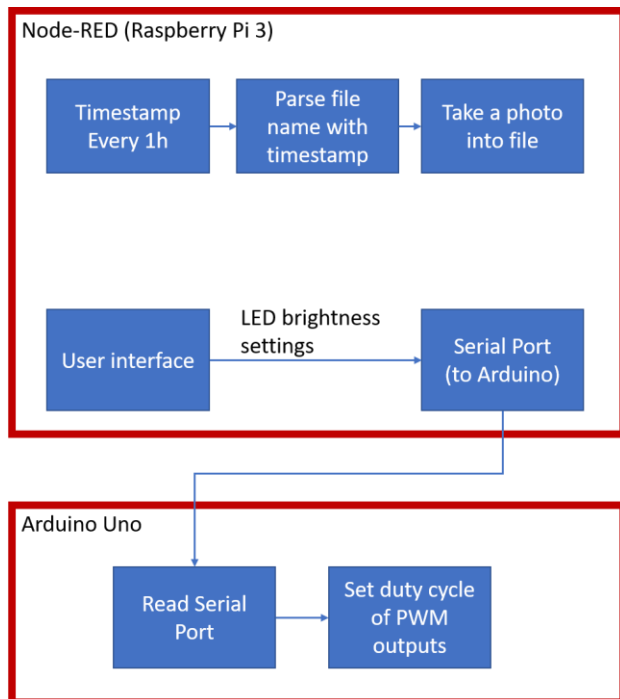


Figure 10. Block diagram of the software

VI. CONCLUSION

This paper aimed to show an optimized experimental indoor plant production environment. The design contains power LEDs and LED drivers, a light spectrum control system and a data logging application. The shown modular design can also be used for other indoor plant production applications. The proposed architecture is easily scalable, user-friendly, and robust. The article describes the relationship between the system components and includes an algorithm that demonstrates the control and measuring options.

ACKNOWLEDGMENT

The authors would like to thank all the faculty staff and member that provide help and assistance throughout the project completion.

REFERENCES

- [1] Attila, Sáfár ; Bertalan, Beszedes. Educational Aspects of a Modular Power Management System. In: Orosz, Gábor Tamás (szerk.) AIS 2019 : 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University. Székesfehérvár, Magyarország : Óbudai Egyetem, (2019) pp. 163-166. , 4 p.
- [2] György, Györök ; Bertalan, Beszedes Concept of a Reliable Redundant Off-grid Power Supply Chain In: Szakál, Anikó (szerk.) SACI 2019 : IEEE 13th International Symposium on Applied Computational Intelligence and Informatics : PROCEEDINGS Temesvár, Románia : IEEE (2019) 383 p. pp. 205-10. , 6 p.
- [3] Beszedes, Bertalan ; Széll, Károly ; Györök, György Redundant Photo-Voltaic Power Cell in a Highly Reliable System ELECTRONICS 10 : 11 p. 1253 , 20 p. (2021)
- [4] Bertalan, Beszedes ; Károly, Széll ; György, Györök A Highly Reliable, Modular, Redundant and Self-Monitoring PSU Architecture ACTA POLYTECHNICA HUNGARICA 17 : 7 pp. 233-249. , 14 p. (2020)
- [5] Györök, György ; Bertalan, Beszedes Fault tolerant power supply systems In: Orosz, Gábor Tamás (szerk.) 11th International Symposium on Applied Informatics and Related Areas (AIS 2016) Székesfehérvár, Magyarország : Óbudai Egyetem (2016) pp. 68-73. , 6 p.
- [6] Bartos, Gaye Ediboğlu, et al. "A Multilingual Handwritten Character Dataset: THE Dataset." Acta Polytechnica Hungarica 17.9 (2020).
- [7] Meng Qingyan ; Chen Xu ; Zhang Jiahui ; Sun Yunxiao ; Li Jiaguo ; Jancsó Tamás ; Sun Zhenhui Canopy Structure Attributes Extraction from LiDAR Data Based on Tree Morphology and Crown Height Proportion PHOTONIRVACHAK / JOURNAL OF THE INDIAN SOCIETY OF REMOTE SENSING (0255-660X 0974-3006): 46 9 pp 1433-1444 (2018)
- [8] G. Györök and B. Beszedes, "Highly reliable data logging in embedded systems," 2018 IEEE 16th World Symposium on Applied Machine Intelligence and Informatics (SAMII), 2018, pp. 000049-000054, doi: 10.1109/SAMI.2018.8323985.
- [9] Xiaojiang, Li ; Weidong, Li ; Qingyan, Meng ; Chuanrong, Zhang ; Tamas, Jancso ; Kangli, Wu Modelling building proximity to greenery in a three-dimensional perspective using multi-source remotely sensed data JOURNAL OF SPATIAL SCIENCE , 16 p. (2016)
- [10] G. Simon and L. Sujbert, "Special issue on "Recent advances in indoor localization systems and technologies", " Applied Sciences, Vol. 11, No. 9, paper 4191, May 2021.
- [11] L. Sujbert, G. Simon and G. Peceli, "An Observer-Based Adaptive Fourier Analysis," IEEE Signal Processing Magazine, Vol. 39 No. 4, pp.134-143, Oct. 2020.
- [12] G. Simon, G. Zachár, and G. Vakulya, "Lookup: Robust and Accurate Indoor Localization Using Visible Light Communication," IEEE Trans. on Instrumentation and Measurement, Vol. 66, No. 9, pp.2337 - 2348, Sept 2017.
- [13] Pogátsnik, M. (2021): Dual education: connecting education and the labor market, OPUS ET EDUCATIO 3 pp. 304-313.
- [14] S. Nosratabadi, K. Szell, B. Beszedes, F. Imre, S. Ardabili and A. Mosavi, "Comparative Analysis of ANN-ICA and ANN-GWO for Crop Yield Prediction," 2020 RIVF International Conference on Computing and Communication Technologies (RIVF), 2020, pp. 1-5, doi: 10.1109/RIVF48685.2020.9140786.
- [15] Pogátsnik, M., Kendrovics, R. "Communication and Reading Comprehension among Informatics and Engineering Students," 2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMII), Herlany, Slovakia, 2020, pp. 235-240
- [16] Bures, Silvia; Urrestarazu Gavilán, Miguel; Kotiranta, Stiina; Artificial lighting in agriculture [online]. January 2018. Horticulture Library.
- [17] Guidi, Lucia & Tattini, Massimiliano & Landi, Marco. (2017). How Does Chloroplast Protect Chlorophyll Against Excessive Light?. 10.5772/67887.
- [18] OSRAM GH CSBRM4.24 datasheet
- [19] Texas Instruments TPS92511 datasheet
- [20] Incandescent Lamps, Publication Number TP-110, General Electric Company, Nela Park, Cleveland, OH (1964) pg. 3
- [21] Krajewski, Markus (24 September 2014). "The Great Lightbulb Conspiracy". IEEE Spectrum. IEEE. Archived from the original on 6 November 2014.

- [22] Vincenzo Balzani, Giacomo Bergamini, Paola Ceroni, Light: A Very Peculiar Reactant and Product. In: *Angewandte Chemie International Edition* 54, Issue 39, (2015), 11320–11337, doi:10.1002/anie.201502325
- [23] National Research Council (U.S.). Building Research Institute. Building illumination: the effect of new lighting levels Publisher National Academies, 1959. Page 81
- [24] Alexander Baklanov, Svetlana Grigoryeva, György Györök, Control of LED Lighting Equipment with Robustness Elements, *Acta Polytechnica Hungarica Acta Polytechnica Hungarica* Vol. 13, No. 5, 2016. pp. 105-119.
- [25] Ramazan Şenol, Kubilay Taşdelen, A New Approach for LED Plant Growth Units, *Acta Polytechnica Hungarica* Vol. 11, No. 6, 2014. pp. 57-71.
- [26] A. Baklanov, S. Grigoryeva, Gy. Györök. Intelligent control of LED luminaries. 9th International Symposium on Applied Informatics and Related Areas - AIS 2014, Székesfehérvár, 2014. pp. 87-91 (ISBN:978-615-5460-21-0)

POSTER SESSION II.

Computer Systems

AIS 2022 – 17th International Symposium on Applied Informatics and Related Areas

Simulation and Analysis of Route Plans for Autonomous Guided Vehicle Systems

Rozália Lakner
Alba Regia Technical Faculty
Óbuda University
Székesfehérvár, Hungary
lakner.rozalia@amk.uni-obuda.hu

Abstract – Automated Guided Vehicle Systems (AGVS) are designed to transport objects automatically, efficiently, and safely between locations. In AGVS, vehicles follow a defined route, which is determined taking into account, for example, collisions and deadlocks. To do so, an important part of the operation of AGVS is the examination of the planned routes in order to detect potential problems. The paper demonstrates the application of Colored Petri nets for simulation and analysis of AGVS and for detecting and avoiding deadlocks and collisions.

Keywords – Automated Guided Vehicles System (AGVS), Colored Petri net, simulation and analysis, deadlock.

I. INTRODUCTION

Autonomous Guided Vehicles (AGVs) are used to automate the transport of parts and materials within a production facility, either indoors or outdoors. The study of an AGV System (AGVS) involves several activities such as layout planning, synthesis of transport activities and feasible pathways, determining the number of AGVs, deadlock prevention and scheduling transport activities, so the design and further performance analysis of AGVS can be quite complex. In general, discrete event simulation is used as the primary tool for modelling and analyzing AGV systems [1, 2].

Petri nets [3] are a modelling formalism that allows a natural representation of discrete event systems. A Petri net has both graphical and mathematical formalism for describing concurrent, asynchronous, distributed systems [4]. Over the past half century, various extensions to Petri nets have been developed, e.g., colored, timed, hierarchical, and many successful applications have been made in a wide variety of domains. These methods support both the modeling and the analysis of systems in various application areas, including process engineering, manufacturing systems and computer science.

The current work presents the construction and application of Colored Petri net models for simulation and formal analysis of AGV systems.

II. COLORED PETRI NETS

Colored Petri net (CPN) is a Petri net extension that allows tokens to be distinguished by assigning a data value to the tokens [5]. Places in a CPN can be assigned types, and tokens have values taken from the types associated with the places. These values, which may represent complex types, are called colors. The colors assigned to places can be specified using the place color set.

A. Formal definition of Colored Petri nets

Colored Petri net formally can be described by an octuple in the form $CPN = \langle P, T, A, \Sigma, C, E, G, I \rangle$, where P is the finite set of the places, and T is the finite set of the transitions, with $P \cup T \neq \emptyset$ and $P \cap T = \emptyset$. A defines the arcs from places to transitions and contrariwise such as $A \subseteq (P \times T) \cup (T \times P)$. A

Petri net is also known as a place/transition net, where places play a role as passive elements, e.g., states, conditions, and transitions correspond to active elements, e.g., events, actions, movements. Σ defines a finite set, called color set, and C is a color function. The color set is defined by specifying a data type, which can also be a multiset. The color function specifies the type of tokens it can accept for each place, i.e., it assigns to the place the color set of possible tokens there, $C: P \rightarrow \Sigma$. The arc expression function, E , is used to specify the condition for allowing transitions and the consequence of firing transitions, $E: A \rightarrow e$ (where e is an expression). The value of the arc expression function must correspond to the color sets assigned to the corresponding vertices. G defines the guard function, $G: T \rightarrow g$. The guard functions (or guard conditions) allow additional constraints to be specified as eligibility conditions for transitions, which are expressions interpreted on multi-sets. The value of a guard functions is of Boolean type (true, false), and the transition is allowed if the evaluation value is "true". The initial state of the CPN defines the initial marking (distribution of colored tokens), which assigns to each place the multi-set of its associated color set, $I: P \rightarrow i$.

Petri nets can be described as bipartite graphs, where the places are represented by circles and the transitions by rectangles. The places and transitions are connected by weighted directed arcs, denoted by the arc expression functions.

B. Simulation and analysis of Colored Petri nets

In the state M of a Petri net, the transition t is enabled, i.e., firing is possible, if all of its input places contain at least the number of tokens defined by the weight function, i.e., $\forall p \in \bullet t: M(p) \geq w(p, t)$. When a transition t fires, a defined number of tokens are consumed from its input places and produced at the output places, so that the state of the net changes, producing a new marking M' as follows: $\forall p \in \bullet t: \forall p \in t \bullet: M'(p) = M(p) - w(p, t) + w(t, p)$. The simulation of these two basic rules of enabling and firing transitions presents the dynamic behavior of a Petri net. Note that the execution of a Petri net is nondeterministic, i.e., if several transitions are enabled at the same time, any one of them can fire, and different sequences of transitions can lead to different states from the same initial state.

Petri nets are also used for system analysis, where various system properties can be investigated, e.g., reachability, boundedness, reversibility, liveness, deadlock. Most of these properties can be investigated using the reachability graph, which contains all possible states of the Petri net and is constructed from a sequence of all firing transitions from the initial state M_0 .

The reachability graph provides information about system properties, such as

- **Reachability** - The state M is reachable, if there is a sequence of firing transitions from the initial state to M . If M is found as a node of the reachability graph, then M is reachable, otherwise is not.
- **L_1 liveness** - The Petri net is L_1 live if each transition fires at least once in some firing sequence from the initial state. If all transitions can be found as arc labels in the reachability graph, then the Petri net is L_1 live, otherwise it is dead.
- **Liveness** - The Petri net is live if all transitions can fire from any reachable states.
- **Deadlock** - A reachable marking for a Petri net where no transition can fire.

The properties of Petri nets can be marking-dependent, in other words, behavioral, which depend on both the structure and the initial state of the Petri net, and structural, which are independent of the initial marking. Behavioral properties, e.g., reachability, liveness, deadlock, can be investigated in terms of a reachability graph containing the state space of the Petri net constructed from the initial state M_0 . The main drawback of these studies is their exhaustive nature, as the reachability graph can be very large, and the time and space requirement increase exponentially with the number of places. This type of analysis is therefore computationally hard.

III. SIMULATION AND ANALYSIS OF AGVS WITH CPN

The chapter presents the CPN model developed for modelling of the route plans for AGVS, as well as the simulation and analysis of the model. The paths of the AGVs were determined as a synthesis task using a P-graph framework [6], and the possible paths were analyzed using CPN. The Petri net model of the AGVS implemented by CPN Tools [7].

A. The CPN model of the AGVS

To determine the AGV routes, first, it is necessary to know the environment. Fig. 1. illustrates the AGV layout under study, where three rooms (A, B and C) are defined, with gateways (G1, G2 and G3) and with three special stations (S charging station in room A, T1 and T2 pick-up and delivery stations in room B and C). AGVs can travel between different locations. For the safe movement of AGVs, safety zones must be established, which must be defined considering the size of the vehicle. These zones are shown in green on the left side of Fig. 1, together with their specific locations. This environment, implemented using the CPN-Tools CPN simulator, is shown on the right side of Fig. 1., where the specific locations and the possible movements of AGVs between two neighboring locations correspond to the places and transitions in Petri net, respectively.

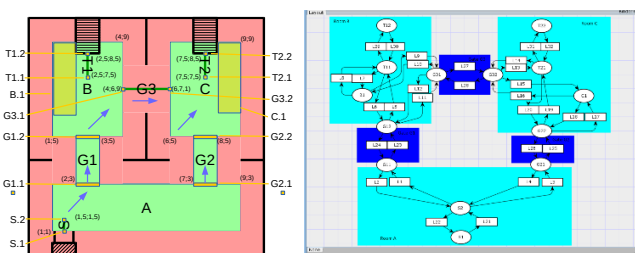


Fig. 1: The layout of the AGVS and its CPN representation.

After specifying the environment, the Petri net must be completed with the properties of the AGVs and their associated routes, and conditions describing the movement of the AGVs between the specific locations in the layout. The color set declarations used to identify AGVs and describe the routes to be taken are shown in Fig. 2.

```

▼ colset PLACE = with S1|S2|G11|G12|G21|G22|G31|G32|B1|T11|T12|
C1|T21|T22|Wh|Delivery|none;
▼ colset ID = INT;
▼ colset PACK_ID = STRING;
▼ colset PACK = product PACK_ID * PLACE;
▼ colset PACK_FROM_TO = product PACK_ID * PLACE * PLACE;
▼ colset PACKLIST = list PACK_FROM_TO;
▼ colset PATH = list PLACE;
▼ colset FORK = with no|ahead|behind;
▼ colset AGV = product ID * PACK * PATH * PACKLIST * FORK * PLACE * PLACE;
    
```

Fig. 2: The color set declarations of the AGVS

Accordingly, an AGV can be specified by a combination of an identifier (ID), the object on the AGV ($PACK$), a route to be completed ($PATH$), the objects to be transported ($PACKLIST$), the position of the fork if exists ($FORK$), the current and previous location of the AGV ($PLACE$). $PACK$ is described by specifying its ID and the destination of the objects, $PATH$ contains the list of locations in the order of the route to be taken, and $PACKLIST$ contains the ID , pick-up station, and delivery station of the objects to be transported by the AGV. An example of specifying the properties associated with an AGV based on the above declarations is: $(2, ("", none), [S2, G11, G12, T11, T12, T11, B1, G31, G32, T21, T22], [{"pack1", Wh, Delivery}], behind, S1, none)$. This means that the AGV with ID 2 is empty, its route to be completed is via the points $S2, G11, \dots, T22$ in the order given in the list, and during the tour it must transport the object with ID "pack1" from the pick-up point Wh to the delivery point $Delivery$. The AGV has fork and it is in a behind state (it is positioned at the back according to the direction), the current position of the AGV is $S1$ and its previous position is irrelevant.

If the AGV with ID 2 is located at the marked point $S1$ (charging station), i.e., the above value is assigned to the location $S1$ in CPN, the AGV is assigned the route as shown in Fig. 3. The AGV is travelling empty on the red and purple sections of the route and is delivering an object on the blue and purple sections of the route (the AGV is travelling both empty and delivering an object at the purple points of interest). The green sections represent the pick-up and drop-off of the object.

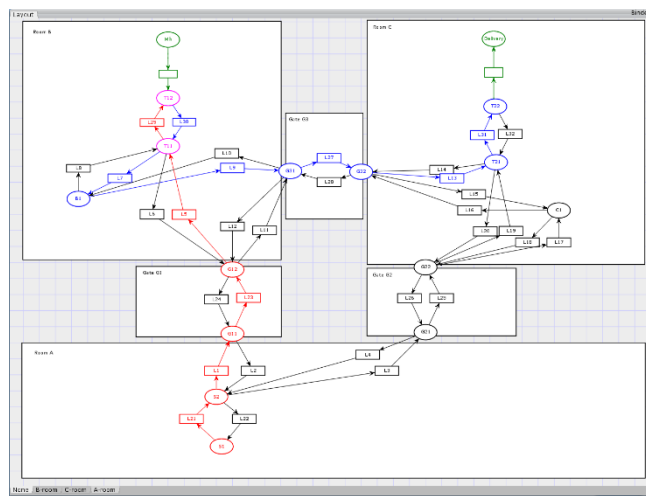


Fig. 3: Access route assigned to the AGV.

To specify the movement of an AGV, the transitions of the CPN require the definition of condition and consequence functions. To perform a transition T representing a step of an AGV, i.e., to move the AGV from a given place $P1$ to a given place $P2$, it is necessary that

1. the AGV is at place $P1$ and
2. the first element of the AGV's route list must be exactly place $P2$.

Of the firing conditions for a transition T , condition 1 is given by arc expression function associated with the input arc of the transition and condition 2 is given by guard function associated with the transition.

The consequence of firing transition T is that

1. the AGV is positioned at place $P2$ and
2. place $P2$ is removed from the AGV's route list.

The description of the above conditions and consequences using CPN functions is shown in Fig. 4.

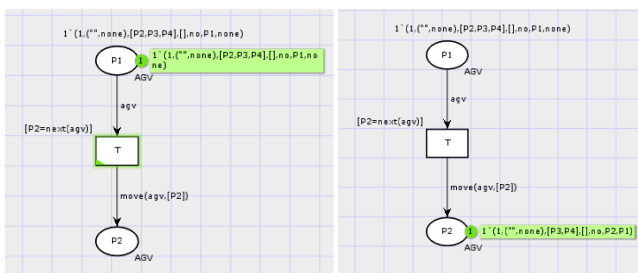


Fig. 4: Transition describing the movement of the AGV with pre- and post-firing states.

In addition to the transitions describing the movement of AGVs, transitions describing the handling of transported objects (pick-up, drop-off) are also required.

To implement the *Pick_up* transition in CPN, i.e., the loading of an object onto an AGV, requires

1. the AGV is at the pick-up point (P),
2. the AGV is empty and
3. the storage associated with P contains the object that is included in the list of objects to be transported by the AGV.

Of the firing conditions for the *Pick_up* transition, condition 1 can be specified by arc expression function associated with the input arc of the transition, and conditions 2 and 3 can be specified by the guard function associated with the transition.

The consequence of firing the *Pick_up* transition is

1. the object to be transported is deleted from the container associated with the transition and
2. the object to be transported is placed on the AGV and
3. the object to be transported is deleted from the list of further packages to be transported of the AGV.

The conditions and consequences for picking up an object by means of functions assigned to the arcs of CPN are shown in Fig. 5.

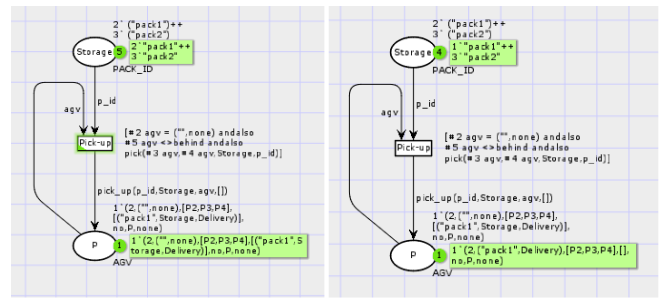


Fig. 5: Transition describing pick-up of object with pre- and post-firing states.

To implement an activity representing the delivery of an object (*Drop* transition) with CPN, the preconditions are as follows:

1. the AGV is at the delivery point (P) and
2. the place of arrival of the object transported by the AGV is the container of P .

Both firing conditions of the *Drop* transition can be specified by arc expression function associated with the input arcs of the transition.

The consequence of firing the *Drop* transition is that

1. the object to be delivered is removed from the *PACKLIST* of AGV and
2. the object to be transported appears on the container of P .

The preconditions and consequences of a *Drop* transition are described by functions assigned to the arcs as shown in Fig. 6.

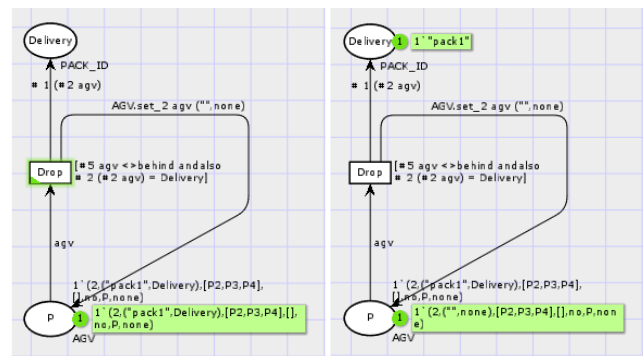


Fig. 6: Transition describing delivery of object with pre- and post-firing states.

Examining the AGVS, it is important to describe the orientation of the AGV and how it changes direction, which also allows questions to be asked about when and how a change of direction is possible without moving the AGV's center of gravity. When managing the orientation of AGV, there are basically two types. One type of AGV does not have a fork, in which case the lifting and unloading of objects is done by moving the AGV under the pallet and then raising and lowering the AGV. This is the operation of the SEW AGV shown on the left in Fig. 7. The other type of AGV is equipped with a fork (lifting arm) for transporting objects. Example is Linde AGV shown on the right in Fig. 7. When modelling these AGVs, it is also necessary to specify whether the AGV travels with a fork in front or behind it on the given route section.



Fig. 7: AGV without fork (left side) and with fork (right side).

In relation to the possible change in the direction of movement of the AGV, it is necessary to identify the places where it is possible to move in the opposite direction. In addition, the points of interest in the layout where the change of orientation of the AGV is mandatory in the event of a further movement shall be identified. Such places are pick-up points, delivery points, charging points and walls. These movements are shown by pairs of steps in different colors in Fig. 8 and 9.

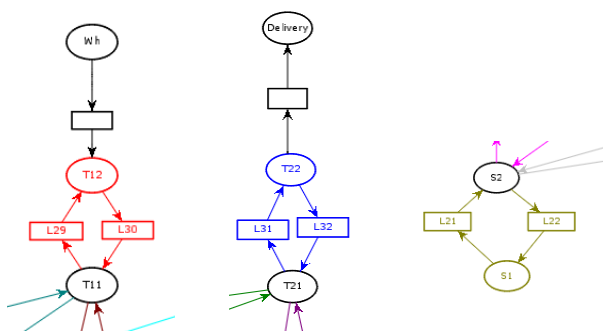


Fig. 8: Orientation change: at the pick-up point ($T12$), at the delivery point ($T22$), at the charging point ($S1$).

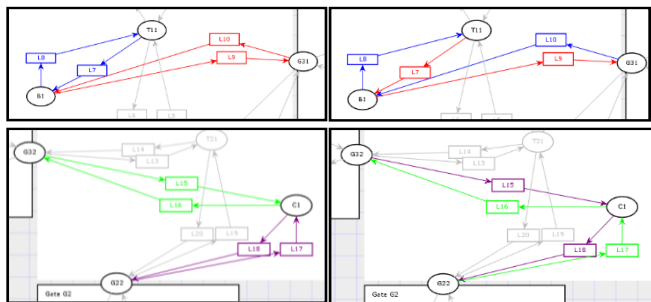


Fig. 9: Orientation change at the walls ($B1$ and $C1$).

In addition, depending on the route section to be traversed, mandatory orientation change may also occur due to the direction of the trajectory arc associated with the route section. These locations in the layout are shown in Fig. 10., where at the gateways associated with locations B and C (marked points $G12$, $G31$, $G32$ and $G22$), opposite directional movement is possible according to the route sections to be accessed. These movements are shown by the pairs of steps in different colors in the figure.

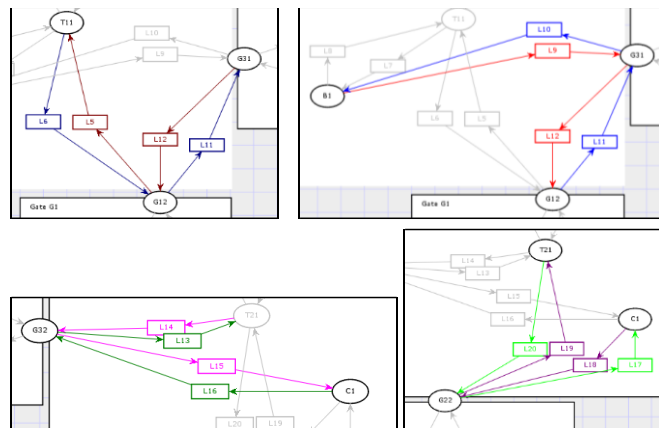


Fig. 10: Orientation change at the gateways ($G12$, $G31$, $G32$ and $G22$).

Based on the above, the change in orientation of AGVs can be summarized as follows:

- Orientation change required (mandatory, regardless of the road section to be covered)
 - proceeding from a delivery point,
 - proceeding from a pick-up point,
 - proceeding from a charging point
 - passing from a wall.

The required change of orientation depends on the type of the actual location.

- Change of orientation possible (optional, determined by the route to be followed)
 - previous location and next location are the same (change of vehicle orientation due to "turning"),
 - other orientation change, due to the direction of the path of the route section.

The possible change of orientation is influenced by the previous position of the route travelled and the current step. The change in orientation of the AGVs during each step (transition) is implemented in the arc expression functions and guard functions associated with the transitions. (Comment: the capacity limits of the places represented by anti-places of the CPN.)

B. The simulation of AGVS with CPN

The following case study shows the simulation of the colored Petri net model. Fig. 11 shows the initial state of the system: one AGV at location $S1$ (charging point of AGV): $(1, ("", none), [S2, G11, G12, T11, T12, T11, B1, G31, G32, T21, T22], ("pack1", Wh, Delivery)), behind, S1, none)$. That is, the AGV with ID 1 is empty, the route to be taken is $[S2, G11, G12, T11, T12, T11, B1, G31, G32, T21, T22]$, the list of the object to be delivered is $[("pack1", Wh, Delivery)]$, the fork is in the *behind* state (backwards according to the direction of the alignment to the charging point), the current location of the AGV is $S1$, and the previous location is irrelevant.

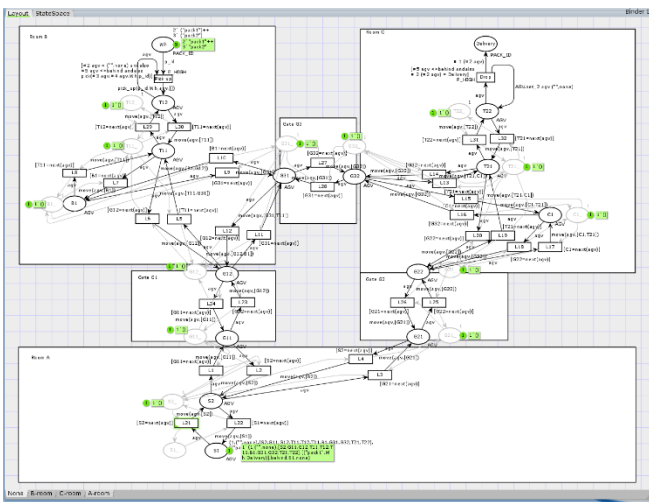


Fig. 11: The initial state of the AGVS.

In the initial state shown in Fig. 11., transition *L21* can be performed, after which the system is in the state shown in Fig. 12. (AGV with ID 1 moves to the state *S2*, its orientation changes to *ahead*, i.e., the AGV is positioned in the front according to the direction of movement).

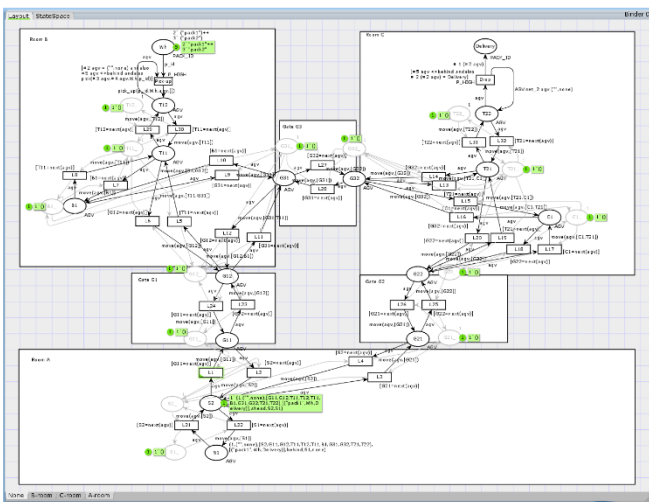


Fig. 12: The state of the AGVS after firing *L21 - L1* transition is fireable.

In the state shown in Fig. 12, transition *L1* is firing, the AGV is moved from the marked position *S2* to the marked position *G11* after the transition is executed. Subsequent firing transitions are *L23, L5, L29*, which, after successive transitions, move the AGV to *T12* marked point in room *B*. The orientation of the forklift remains unchanged (*ahead*) throughout. The AGV at the *T12* location picks up the "*pack1*" object, which must be delivered to the *Delivery* location. The state after the execution of the *Pick-up* transition is shown in Fig. 13.

In the state shown in Fig. 13, transition *L30* is firing, i.e., the AGV can move to the marked position *T11*, during which the AGV orientation changes (becomes *behind*). After executing transition *L30*, transitions *L7* and *L9* are firing. Upon execution of the latter, the orientation of the AGV will change again (become *ahead*). Then, transitions *L27, L13* and *L31* will be executed, which will bring the AGV to the marked point *T22* of room *C*, i.e., the delivery point, were, after execution of the *Drop* transition, the system will be in the state shown in Fig. 14.

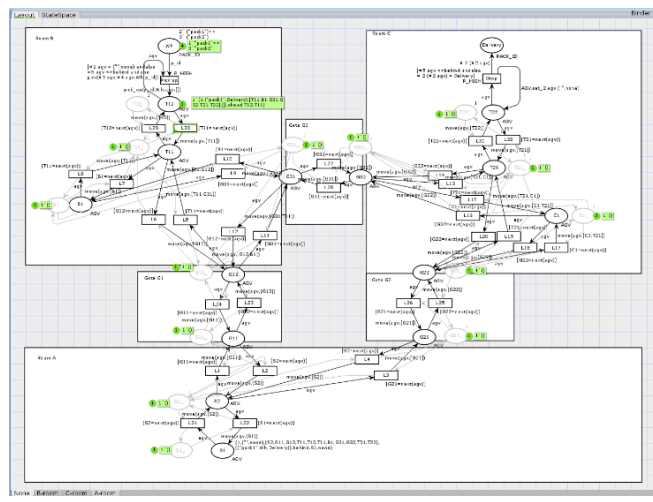


Fig. 13: The state of the AGVS after firing *Pick-up* transition.

The package with the ID "*pack1*" has been dropped from the AGV and at the same time has appeared in the *Delivery* container. In the state shown in Fig. 14., the AGV is in the empty state at the marked position *T22* (delivery point). There is no transition to be executed, therefore the current state is deadlock. The list representing the current route to be taken by the AGV and the list of objects to be delivered are both empty (*[]*), indicating that the AGV has completed its assigned route and delivered the assigned package(s). Thus, the deadlock reached is also the expected end state of the system.

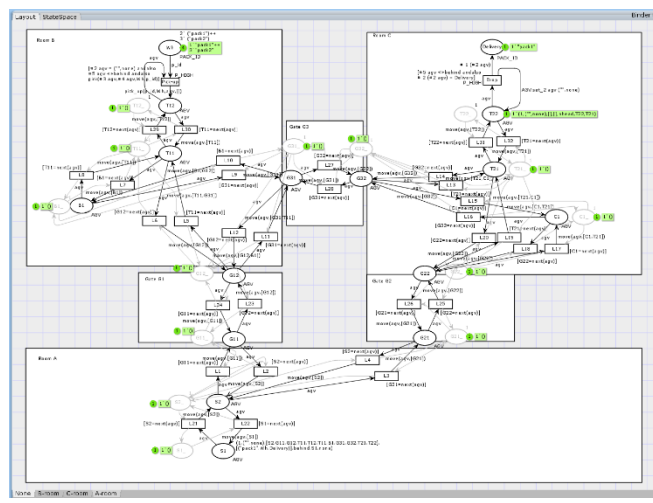


Fig. 14: The deadlock state of the AGVS – expected state.

C. The analysis of AGVS with CPN

In addition to simulating the AGVS, the relevant dynamic properties of the system, such as reachability and deadlock, are investigated using the reachability graph [8] with the State Space Tool integrated in CPN Tools.

In the initial state of the case study, there are two AGVs in the environment, as shown in Fig. 15. One of them is at location *T22*, with characteristics $(1, ("", none), [T21, G22, G21, S2, S1], [], ahead, T22, none)$. The other AGV is at location *T12*, with characteristics $(2, ("", none), [T11, B1, G31, G32, T21, T22], ("pack1", Wh, Delivery), ahead, T12, none)$.

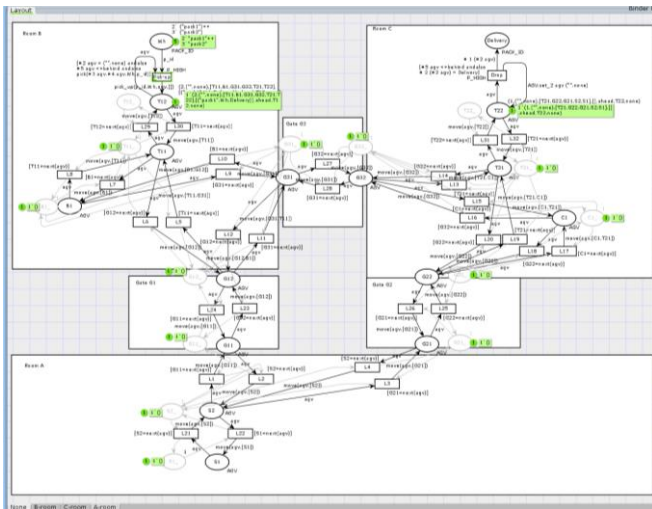


Fig. 15: The initial state of the AGVS with two AGVs.

The reachability graph of the system is shown in Fig. 16. State 1 of the graph represents the initial state, states 21 and 44 represent deadlocks, and their corresponding states values are given in the rectangles.

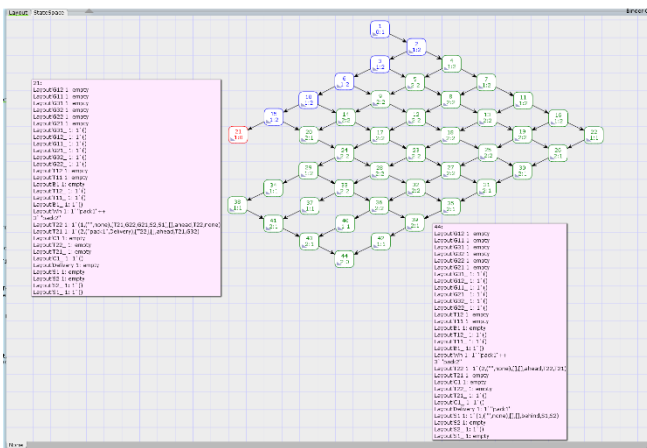


Fig. 16: The reachability graph of the AGVS with two AGVs.

In state 44, AGV 1 is located at S1 (charging station) and AGV 2 at T22 (delivery station). Both AGVs have an empty ([]) remaining route to follow and package list to deliver, i.e., they have traversed their assigned route and delivered their assigned objects. The deadlock is the intended end state of the system. In state 21, AGV 1 is at T22 and AGV 2 is at T21. The properties of AGV 1 are the same as in its initial state, i.e., the AGV is 'stationary'. AGV 2 has travelled all but one place along the designated route but can no longer enter place T22. The state is a deadlock, which is the unexpected final state of the system.

The partitions of the reachability graph are shown in different colors in Fig. 16.: the states in the red partition whose result in unexpected final state(s) regardless of the execution order of subsequent transitions. The states in the green partition show the states that lead to an expected final state. The states in the blue partition are the ones that can lead to both types of final states.

If the orientation of the AGV with ID 2 in the initial state is changed while leaving the other properties unchanged (i.e., the color of the token representing the AGV at location T12 is (2, ("",none), [T11,B1,G31,G32,T21,T22], [{"pack1",Wh,

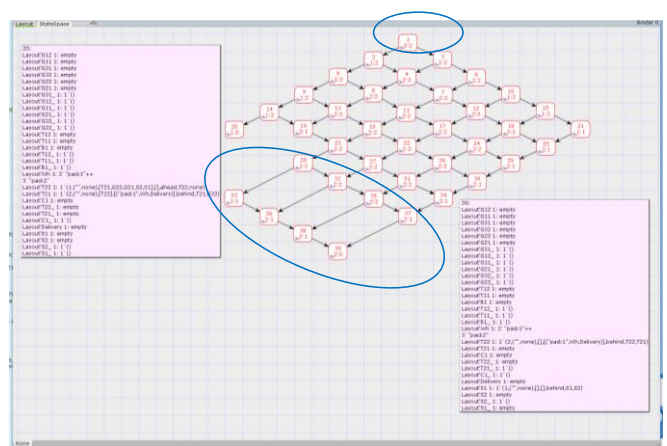


Fig. 17: The reachability graph of the AGVS with two AGVs (modified orientation of AGV 2 in the initial state).

Delivery)],behind,T12,none)), the reachability graph shown in Fig. 17. is obtained.

Comparing this with the reachability graph in Fig. 16., it can be seen, that the number of vertices and edges of the graph has decreased (i.e., previously feasible transitions have become infeasible), and therefore the structure of the reachability graph has naturally changed (the location of the changes is shown by blue ovals). Both deadlocks of the system are not expected system states, and therefore there is no transition sequence leading to an expected system state (each vertex belongs to a "red partition").

IV. CONCLUSIONS

This paper presents the application of colored Petri nets to the description and analysis of automated guided vehicle (AGV) systems, showing that the colored Petri net model is suitable for system simulation and formal analysis. The analysis of the Petri net reachability graph has shown that for the design and testing of AGVS, it is critical to identify deadlocks and to avoid them by detecting unexpected system states as soon as possible to determine alternative step sequences.

V. REFERENCES

- [1] H. Kaid, A. M. El-Tamimi, E. A. Nasr, A. Al-Ahmari, "Applications of Petri nets based models in manufacturing systems: A review" Proceedings of the International Conference on Operations Excellence & Service Engineering, Orlando, FL, pp. 516-28, 2015.
- [2] W. Wu, Z. Xing, H. Yue, H.. Su, H., S. Pang, "Petri-net-based deadlock detection and recovery for control of interacting equipment in automated container terminals" IET Intelligent Transport Systems, 2022.
- [3] C. A. Petri, "Kommunikation mit Automaten", Rheinisch-Westfälisches Institut für Instrumentelle Mathematik an der Universität Bonn, Bonn, Germany, 1963.
- [4] J. L. Peterson, "Petri Net Theory and the Modeling of Systems", Prentice Hall, New Jersey, 1981.
- [5] K. Jensen and L.M. Kristensen, "Coloured Petri Nets. Modelling and Validation of Concurrent Systems", Springer-Verlag 2009.
- [6] R. Lakner and B. Bertók, "Synthesis and analysis of AGV systems", AIS 2020, 15th International Symposium on Applied Informatics and Related Areas, Óbuda University, pp. 63-67., ISBN 978-963-449-209-2, 2020.
- [7] CPN Tools: A tool for editing, simulating, and analyzing Colored Petri nets, cpntools.org, Last accessed 01/10/2022.
- [8] T. Murata, "Petri Nets: properties, analysis and applications", In: Proc IEEE, vol. 77. pp. 541-80. 1989.

Genetic algorithm based image approximation with geometrical shapes

Gergely Vakulya

Alba Regia Technical Faculty

Óbuda University

Székesfehérvár, Hungary

vakulya.gergely@amk.uni-obuda.hu

Abstract—In this paper the design parameters of a genetic algorithm (GA) based image approximation / compression algorithm are analyzed. The efficiency using different approximation shapes and with different GA parameters (population size, mutation rate) are evaluated using test runs. Recommended parameter sets for different scenarios are provided.

Keywords—genetic algorithm, image approximation

I. INTRODUCTION

The genetic algorithm [1] concept can be efficiently used in many optimization problems to provide an alternative solution. It can be used in several areas from data mining [2] to neural networks and deep learning [3]. GA does not guarantee the optimum, and the quality of the approximation highly depends on the implementation and on the carefully chosen design parameters.

One large area, where GA can efficiently be used, is image processing [4], e.g. image approximation.

In [5] an image approximation algorithm was presented using arbitrary triangles to approximate the black and white areas of images. Using different shapes, however, can lead to different results. In this paper additional shapes (rectangles and circles) will be used for the approximation, among the triangles.

The efficiency of the algorithm is highly affected by the design parameters. In this paper the effect of the used shapes, the maximum number of shapes, the mutation and crossover ratio and the distribution and the parameters of the random numbers used in generation of the new shapes and during the mutations are evaluated.

II. RELATED WORK

Genetic algorithms [1] are widely used [3], primarily for solving problems, where the number of parameters is too high and no quick solution is known. For most of those problems finding the optimal solution is not critical, and an approximation (or maybe any possible solution) can be acceptable.

An approximate solution can be given to classic NP-complete problems, Vehicle Routing Problem (VRP) [6] and Travelling Salesman Problem (TSP) [7], with genetic algorithms. Another two typical applications are transporting [8] and scheduling [9].

In mechanical engineering genetic algorithms can be used to optimize mechanical components, e.g. to make support elements with minimum weight (and from minimum amount of material) and with maximal strength [10].

Different areas, like fuzzy logic [11], data mining [2] or scheduling [12] can efficiently be combined with genetic algorithms.

Genetic algorithms can be effectively used for economic applications [13], e.g. for automatized trading [14] or portfolio analysis [15].

A promising application is using genetic algorithm to optimize the weights of a neural network instead of using back propagation [3], [16]. Two interesting applications of neuro-evolution are an automated player for the classic Nintendo Super Mario Brothers video game (MarIO) [17] and an application, where a virtual robot battle player is controlled by a neuro-evolutionary algorithm [18].

A good application of genetic algorithms is image processing [19], e.g. image segmentation [4] or image enhancement [20]. This paper will focus on a similar application: image approximation [5].

III. THE GA-BASED IMAGE COMPRESSION

A. Problem statement

The algorithm works with 8-bit grayscale images with, and represent them as an $(n \times n)$ matrix:

$$IMG = (a_{i,j} \in 0, 1, \dots, 255^{n \times n}), \quad (1)$$

where $a_{i,j}$ is the pixel in the (i, j) coordinates of the input image.

An approximation B of the image consists of the series of k shapes:

$$B = (s_1, s_2, \dots, s_k) \quad (2)$$

where each shape is given with a series of parameters. A shape s is defined as follows:

$$s = (t, p_1, p_2, \dots, p_{pn(t)}, c), \quad (3)$$

where t is the type of the shape and $p_1, \dots, p_{pn(t)}$ are the parameters of the shape, given that $pn(t)$ gives the number

of parameters for the t^{th} shape. The $c \in -255, \dots, 255$ parameter gives the color of the shape.

Each approximation has a graphical representation R , similarly to the input image:

$$R = (b_{i,j} \in 0, 1, \dots, 255^{n \times n}), \quad (4)$$

where $b_{i,j}$ is the pixel in the (i, j) coordinates of the approximation.

B. Graphical representation

The genetic algorithm handles the following shapes:

- $t = 1$ represents a triangle with $np(1) = 6$ parameters, where (p_1, p_2) , (p_3, p_4) and (p_5, p_6) belong to the 3 vertices.
- $t = 2$ defines a rectangle with $np(2) = 4$ and (p_1, p_2) and (p_3, p_4) give the opposite vertices.
- Finally, $t = 3$ determines a circle with $np(3) = 3$ parameters, where (p_1, p_2) is the center of the circle and p_3 is the radius.

During calculation of R the shapes are rendered overlaying each other with adding their color to the actual image, starting from a white image. After adding a shape, the value of each pixel is kept between 0 and 255.

The aim of the algorithm is to find an approximation B with minimum error $e(B)$, i.e. with minimum difference between the original image IMG and the graphical representation R of the approximation B :

$$e(B) = \left| \sum_{i,j} (IMG_{i,j} - R_{i,j}) \right| \quad (5)$$

The conventional term related to genetic algorithms is *fitness function*, which can be defined as the opposite of the error:

$$f(B) = -e(B). \quad (6)$$

C. The genetic algorithm

The genetic algorithm uses a P population of p approximations:

$$P = B_1, \dots, B_p. \quad (7)$$

At first each approximation of the population is initiated with an approximation contains a number of randomly generated shapes. Then every population is generated as follows.

- The graphical representation is rendered for each approximation.
- The fitness function is calculated using the graphical representations and the original image.
- A fixed number of approximations are selected from the population with the highest fitness, which are left intact.
- A fixed number of approximations are overwritten by new ones generated using crossover.
- The rest of the approximations are randomly modified in terms of the coordinates and the color with predefined probabilities.

For sake of simplicity the details of the genetic operators are not discussed here in detail. They can be found in [5].

D. Software architecture

The image approximation algorithm is implemented in C language with as few external dependencies as possible to be portable to different operating systems. The different parameters (i.g. population size, resolution, maximum number of shapes, shape types) can be set at compile time.

The software expects the input image in BMP format in grayscale with a fixed resolution. The graphical representation of the approximation with the smallest error is saved in every tenth iteration in raw format and converted to JPEG using an external program.

The implementation uses a single thread only, but several instances with different settings can be run in parallel.

IV. TEST RESULTS

To test the performance of the method with different settings and to analyze the effect of different parameters the genetic algorithm was run with several different settings. The parameters can be summarized as follows:

- Maximum number of shapes: $20 \leq k_{max} \leq 100$
- Population size: $p = 4096$
- Intact approximations: 256
- Overwritten approximations: 1024
- Shapes: triangles only, rectangles only, circles only and all shapes mixed.
- Random shape generation and mutation with random numbers with uniform and Gaussian distribution.

During generating and modifying the shapes two different strategies are used. The *uniform* strategy uses uniform distribution for all random numbers, i.g. the vertices of the triangles and the rectangles, and the center and the radius of the circles are generated with uniform random distribution. The same is true for the modification of the vertices and the radii as well.

The *Gaussian* strategy uses random numbers with Gaussian distribution. During generation of triangles and rectangles first a center point is generated with uniform distribution and the vertices are shifted with Gaussian offsets. The circles are generated with uniform centers and Gaussian radii.

The test image can be seen in Fig. 1(a). It is a moderately detailed, well distinguishable silhouette of Sherlock Holmes. Fig. 1(b) contains the best approximation during the experiments with the number of shapes limited to 80 ($k_{max} = 80$).

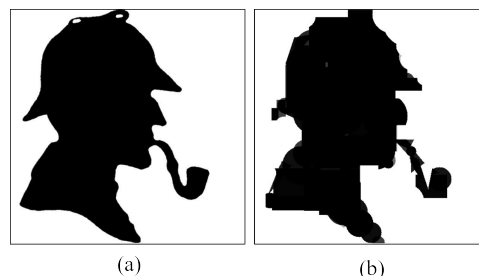


Fig. 1. The test image (a) and the best matching image after 5000 generations with the best performing settings.

Fig. 2 shows the best approximation of the test image after 50, 200 and 1000 generations using mixed shapes (upper row) and circles (lower row). The maximum number of shapes were $k_{max} = 80$.

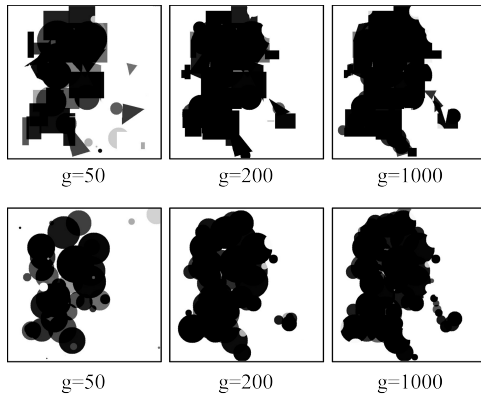


Fig. 2. The best approximation of the test image after 50, 200 and 1000 generations using mixed shapes (upper row) and circles only (lower row).

On Figs. 3 and 5 the effect of using different shapes and different random number strategies are visualized. For this experiment $k_{max} = 80$ was set. Fig. 3 shows the results of the uniform strategy, while Fig. 4 belongs to the Gaussian random number strategy. Gaussian distribution leads to slightly lower errors in each case. During the experiments triangles gave the worst performance, while the other two shapes and the mixed mode performed almost equally, being the mixed mode the best setting.

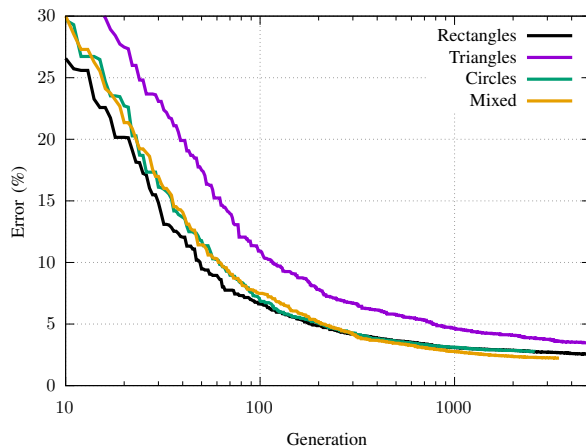


Fig. 3.

To evaluate the effect of k_{max} and the random distribution the same test image was approximated with 9 different k_{max} values between 20 and 100, and the genetic algorithm was run with the Gaussian random number strategy and with mixed shapes. Fig. 5 shows the error of the best approximation in each generation.

According to the tests the achievable minimum error decreases with more and more shapes. Using less than 40 or 50

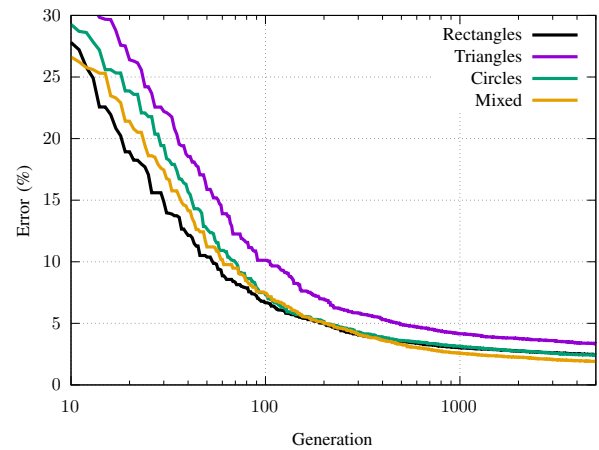


Fig. 4.

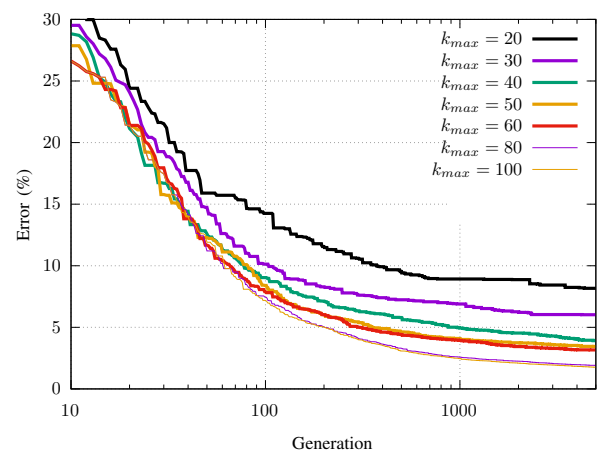


Fig. 5.

shapes does not give a recognizable approximation. On the other hand, using more than 80 shapes does not significantly reduce the achievable error.

V. SUMMARY

In this paper a genetic algorithm based image approximation method was presented, which extends a similar, existing solution. The approximation is based on a genetic algorithm and uses a limited number of triangles, rectangles, circles, or all of them. The paper discussed several different design parameters of the algorithm: different limits for the number of shapes, different random number generation strategies and different shapes. In all tests the triangle based approximation was outperformed by the others, and the mixed strategy provided the best results. With using Gaussian distribution during some part of the algorithm the results can slightly be improved.

REFERENCES

- [1] S. N. Sivanandam and S. N. Deepa, *Introduction to Genetic Algorithms*. Berlin Heidelberg: Springer-Verlag, 2008.

- [2] T. Barman and N. Deb, "State of the art review of speech recognition using genetic algorithm," in *2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI)*. IEEE, 2017, pp. 2944–2946.
- [3] F. P. Such, V. Madhavan, E. Conti, J. Lehman, K. O. Stanley, and J. Clune, "Deep neuroevolution: Genetic algorithms are a competitive alternative for training deep neural networks for reinforcement learning," *arXiv preprint arXiv:1712.06567*, 2017.
- [4] S. Pare, A. K. Bhandari, A. Kumar, G. K. Singh, and S. Khare, "Satellite image segmentation based on different objective functions using genetic algorithm: a comparative study," in *2015 IEEE international conference on digital signal processing (DSP)*. IEEE, pp. 730–734.
- [5] G. Vakulya, "Evaluation of a genetic algorithm based image approximation method," in *AIS 2021-16th International Symposium on Applied Informatics and Related Areas - Proceedings*, 2021, pp. 12–15.
- [6] K. Q. Zhu, "A new genetic algorithm for vrptw," in *Proceedings of the international conference on artificial intelligence*, 2000.
- [7] Z. H. Ahmed, "Genetic algorithm for the traveling salesman problem using sequential constructive crossover operator," *International Journal of Biometrics & Bioinformatics (IJBB)*, vol. 3.6, p. 96, 2010.
- [8] C. Lin, J. Yu, J. Liu, and C. Lee, "Genetic algorithm for shortest driving time in intelligent transportation systems," in *International Conference on Multimedia and Ubiquitous Engineering 2008*, 2008, pp. 402–406.
- [9] C. Lim and E. S., "Production planning in manufacturing/remanufacturing environment using genetic algorithm," in *Proceedings of the 7th annual conference on Genetic and evolutionary computation*, 2005.
- [10] M. T. Bhoskar, M. O. K. Kulkarni, M. N. K. Kulkarni, M. S. L. Patekar, G. Kakandikar, and V. Nandedkar, "Genetic algorithm and its applications to mechanical engineering: A review," *Materials Today: Proceedings*, vol. 2, no. 4-5, pp. 2624–2630, 2015.
- [11] J. J. Buckley and Y. Hayashi, "Fuzzy genetic algorithm and applications," *Fuzzy sets and systems*, vol. 61, no. 2, pp. 129–136, 1994.
- [12] J. F. Gonçalves, J. J. de Magalhães Mendes, and M. G. Resende, "A hybrid genetic algorithm for the job shop scheduling problem," *European journal of operational research*, vol. 167, no. 1, pp. 77–95, 2005.
- [13] H. Dawid and M. Kopel, "On economic applications of the genetic algorithm: a model of the cobweb type," *Journal of Evolutionary Economics*, vol. 8, no. 3, pp. 297–315, 1998.
- [14] R. Kuo, C. Chen, and Y. Hwang, "An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network," *Fuzzy Sets and Systems*, vol. 118, no. 1, pp. 21–45, 2001.
- [15] S. Slimane and M. Benbouziane, "Portfolio selection using genetic algorithm," *Journal of Applied Finance & Banking*, vol. 2, no. 4, pp. 143–154, 2012.
- [16] R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Neural networks for perception*. Elsevier, 1992, pp. 65–93.
- [17] B. A., S. Y., R. G., and C. J., "Learning levels of mario ai using genetic algorithms," *Advances in Artificial Intelligence. CAEPIA 2015. Lecture Notes in Computer Science*, vol. 9422.
- [18] S. Y., Z. E., and S. M., "Using genetic programming to evolve robocode players," vol. 3447.
- [19] S. Jayaraman, S. Esakkirajan, , and T. Veerakumar, *Digital image processing*. India: Mc Graw Hill, 2009.
- [20] H. Deborah and A. M. Arymurthy, "Image enhancement and image restoration for old document image using genetic algorithm," in *2010 Second International Conference on Advances in Computing, Control, and Telecommunication Technologies*. IEEE, pp. 108–112.

A new laser stripe detection algorithm for 3D scanning applications

Albina Kadyroldina
*School of Information Technologies
 and Intelligent Systems*
*D Serikbayev East Kazakhstan
 Technical University*
 Ust-Kamenogorsk, Kazakhstan
akadyroldina@gmail.com

Assel Kussaiyn-Murat
*School of Information Technologies
 and Intelligent Systems*
*D Serikbayev East Kazakhstan
 Technical University*
 Ust-Kamenogorsk, Kazakhstan
akussaynmurat@edu.ektu.kz

Alexander Krasavin
*School of Information Technologies
 and Intelligent Systems*
*D Serikbayev East Kazakhstan
 Technical University*
 Ust-Kamenogorsk, Kazakhstan
akrassavin@ektu.kz

Darya Alontseva
*School of Information Technologies
 and Intelligent Systems*
*D Serikbayev East Kazakhstan
 Technical University*
 Ust-Kamenogorsk, Kazakhstan
dalontseva@ektu.kz

Abstract—The paper proposes a new and simple laser stripe detection algorithm for 3D scanning with surface illumination by a laser stripe. The algorithm for selecting areas of a black-and-white image containing images of a laser stripe is based on the calculation of the gradient and Laplacian of the image intensity using standard digital image processing procedures. The proposed detection algorithm is characterized by low computational cost and does not imply any a priori knowledge about the object being detected and the image, including the knowledge of the laser stripe width. The paper also describes a new algorithm for converting a laser stripe image into a broken line, suitable for use in 3D scanning with laser illumination.

Keywords—3D scanning, image processing, machine vision)

I. INTRODUCTION

This Nowadays, new 3D scanning methods for various purposes are developed and applied around the world, including for the development of additive technologies of robotic microplasma spraying of coatings. In particular, the authors of this study used an industrial robotic arm for microplasma layer-by-layer spraying of tantalum coatings on a titanium implant, ensuring the movement of the robot along a 3D model obtained by 3D scanning the implant surface [1].

For 3D scanning and quality control methods for industrial products based on surface illumination with a laser stripe, as well as for machine vision methods that use structured light, the task of detecting a laser stripe is a key task [2], [3], [4]. These areas of application dictate a number of requirements and limitations for laser stripe detection algorithms, firstly, subpixel precision is often required, and secondly, these are restrictions on the computing power required to implement these algorithms, which makes it possible to implement detection algorithms in real-time embedded systems.

At present, a number of laser stripe detection algorithms are known and widely used [4], [5], [6], and the question of the reliability and accuracy of these algorithms has also been studied quite well. It should be noted that algorithms

that provide high accuracy and reliability usually require relatively large amounts of calculations, that is, the algorithms are characterized by high computational cost and duration of the computation process. Another feature of this kind of algorithms is the need for some data about the detected object and image, for example, data about the width of the laser stripe.

The aim of this study was to develop a simple and reliable laser stripe detection algorithm as well as to develop algorithm for converting a laser stripe image into a broken line, suitable for 3D scanning with surface illumination by a laser stripe.

II. METHODS

A. Formalization of the task of detecting the laser strip in the digital image.

The discrete intensity function of the black-and-white image I_{ij} can be interpreted as a discretization of a continuous function of two variables $I(x,y)$ on a rectangular grid. Such a representation makes it possible to formalize the problem of detecting a laser stripe in a digital image. The intersection of the plane of the laser beam with the surface is a continuous curve in space. The projection of this curve onto the image plane forms a continuous two-dimensional curve in the image plane. This curve is piecewise continuous, and for the points of the curve that belong to the intervals of continuity, the vector of the tangent to the curve is defined. Of course, in reality, images of laser-illuminated surface points do not form a curve, but a two-dimensional region in the image plane, called a laser stripe. Ideally, the curve is a level isoline where the maximum intensity is within the region of the laser stripe. Moreover, in such an ideal case, the isolines of the function are curves obtained by parallel transfer of the curve to the image plane. For a given point M on the curve γ , consider a line p , perpendicular to the curve, i.e., a line with a direction vector perpendicular to the tangent vector to the curve. The restriction of a function of two variables $\Phi(x,y)$

to a line with a unit direction vector $\vec{e} = (e_x, e_y)$ passing through the point $M(x_0, y_0)$ is the function defined by formula (1):

$$\varphi(s) = \Phi(x_0 + s \cdot e_x, y_0 + s \cdot e_y) \quad (1)$$

The restriction of $i(x)$ of the intensity of $I(x, y)$ on the line p (in the ideal case) is called the laser strip profile. For real digital images, the restriction $i^*(x) I(x, y)$ to the line p is a linear superposition of the form (2):

$$i^*(x) = i(x) + B + N(x) \quad (2)$$

Using digital filtering methods, it is possible to suppress the constant B and noise $N(x)$ components $i^*(x)$, thus highlighting the profile of the laser strip $i(x)$ [6]. The width of the laser stripe is the length of the interval where the function $i(x)$ is nonzero

B. Laser stripe detection

Laser stripe detection is a curve reconstruction or polyline fitting process (approximation by broken line). In a number of methods of laser strip detection, the algorithm of detection reconstructs the position of individual points of the curve, and then linking the found points into the sequence, which can be interpreted as the sequence of the vertices of the broken line, which approximate curve γ . Of course, if necessary, based on the constructed sequence of points belonging to the curve, it is possible to perform an approximation more accurate than the simplest approximation by a broken line. In this study, the task of a laser stripe detection as a task of a curve reconstruction is different from the task of detecting an image area containing an image of a laser stripe. It should be noted that if during image processing it is possible to reliably detect the image area of the laser stripe, then the task of the curve reconstruction is greatly simplified. In this paper, the authors focus mainly on the task of detecting the image area corresponding to the laser stripe. In the classical work of Steger [7], when classifying methods for detector of curvilinear structures in digital images, a class of methods based on considering a curve image as an "object with edges" is singled out, for example, when recognizing a line of a handwritten image of a symbol or a riverbed in aerial photographs. The method proposed in this study uses gradient image processing specifically to highlight the "edges" of the laser strip in the image and can be attributed to this group of methods with a certain stretch. The next section provides the necessary information about the gradient processing of the images used in this study, and the advantages and features of the described algorithm are discussed.

Details of the experiment on obtaining digital images of objects with a laser stripe can be found in a previous paper [8].

III. RESULTS AND DISCUSSIONS

A. Applying an Image Gradient for Laser Stripe Detection

The notion of a continuous function of image intensity allows us to consider the gradient of the image function $\vec{g}(x, y)$ (3), often referred to simply as the *image gradient*:

$$\vec{g} = \text{grad } I = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right) \quad (3)$$

Of course, real digital images of discretely derivative images (functions $\frac{\partial I}{\partial x}$ and $\frac{\partial I}{\partial y}$) in the classical sense (as private derivatives of analytically given functions) are not defined. In the technique of digital image processing, methods have been developed for calculating partial derivatives and differential operators of the image intensity function, usually based on the application of the operation of convolution of the discrete function of the image intensity and the corresponding kernel of the discrete function defined in a small (in comparison with the image dimensions) rectangle of fixed sizes [6]. In the technique of image recognition, the image gradient is used to highlight the image features (the boundaries of the regions). Figuratively speaking, images containing features are characterized by high values of the gradient module function $G(x, y)$ (4) (see Fig. 1).

$$G(x, y) = \left| \vec{g} \right| = \sqrt{g_x^2 + g_y^2} \quad (4)$$

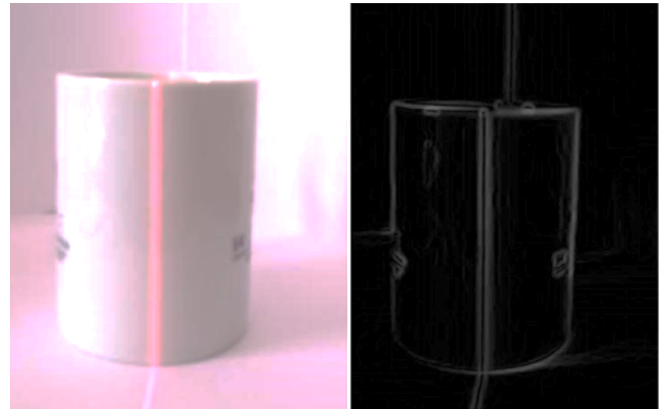


Fig. 1. Digital image of an object with laser backlight (left) and the distribution of the image gradient module (right).

The image gradient is a vector field. A two-dimensional vector field is defined by two vector fields, i.e. scalar functions of two variables $g_x(x, y)$ и $g_y(x, y)$. During software image processing, the components e_x and e_y of the gradient direction vector \vec{e} were calculated by formula (5) for each image pixel and the magnitude of the image gradient modulus $G_{ij} = G(i, j)$ at this point was calculated for each grid point (i, j)

$$\vec{e}(x, y) = \frac{1}{G(x, y)} \cdot \vec{g}(x, y) \quad (5)$$

Thus, the gradient field was set by three two - dimensional arrays. The direction of the image gradient vector at a point is perpendicular to the tangent to the isoline of the image intensity function passing through this point.

As can be seen from Fig. 1, the laser strip on the visualization of the distribution of the gradient module corresponds to a specific pattern, namely two parallel stripes corresponding to the high values of the image gradient module. The first impulse of the authors of this study was to apply a line scan of the gradient modulus function to detect this pattern. Scanning along the lines of one or another function of the $N \times M$ dimension, the nodes of which correspond to the image pixels will be called the process of the repeated application of the sequence processing procedure. $\{F_{i,1}, F_{i,2}, \dots, F_{i,M}\}$ for all lines of the image, i.e., for all index values $j=1,2, \dots, N$. Scanning along columns and diagonals is defined similarly. However, experiments quickly showed the low reliability of this approach to the problem of the laser stripe region detection. The problem was the presence of a large number of high-level gradient modulus regions occurring on image features other than the edges of the laser stripe. This led to an unacceptably high number of false positives of the detector. However, the authors managed to develop a reliable algorithm of the laser stripe area detection, which was also based on the recognition of a certain pattern during progressive scanning, using the calculation of another scalar field - the Laplacian field of the image.

B. Algorithm of the laser stripe area detection in a digital image.

The first stage of the laser stripe detection method proposed here is the pre-processing of the image with a Gaussian Blurring Filter. After such processing, the profile of the laser strip with a sufficiently high accuracy is described by the Gaussian type (6):

$$f(x) = A \cdot e^{-\frac{x^2}{2\sigma^2}} \quad (6)$$

The first and second derivatives of the function f are given by formulas (7) and (8), respectively:

$$\frac{df}{dx} = -\frac{A}{\sigma^2} \cdot x \cdot e^{-\frac{x^2}{2\sigma^2}} \quad (7)$$

$$\frac{d^2f}{dx^2} = \frac{A}{\sigma^2} \cdot \left(\frac{x^2}{\sigma^2} - 1 \right) \cdot e^{-\frac{x^2}{2\sigma^2}} \quad (8)$$

Gradient and Laplacian operators are invariant to coordinate system rotations. Let us consider a Cartesian coordinate system on the image plane, the origin of which is located at point O of the curve γ , the Y axis is tangent to the curve γ at the point O , and the X axis is perpendicular to γ . Then the limitation of the gradient modulus function on the X axis will look like (9):

$$g_x(x) = \frac{A}{\sigma^2} \cdot |x| \cdot e^{-\frac{x^2}{2\sigma^2}} \quad (9)$$

Laplacian image L is determined at the arbitrary point of the plane of image M as (10):

$$L(M) = \left(\frac{\partial^2 I}{\partial x^2} \Big|_M + \frac{\partial^2 I}{\partial y^2} \Big|_M \right) \quad (10)$$

In the selected coordinate system $\frac{\partial^2 I}{\partial y^2} = 0$ and the Laplacian limitation l_x on the X axis will have the form (11):

$$l_x(x) = \frac{d^2 f}{dx^2} = \frac{A}{\sigma^2} \cdot \left(\frac{x^2}{\sigma^2} - 1 \right) \cdot e^{-\frac{x^2}{2\sigma^2}} \quad (11)$$

Graphs of the distribution functions of the gradient modulus and Laplacian in the cross section of the laser strip are shown in Fig.2.

Let's introduce two threshold levels: $G_m > 0$ for the values of the modulus of the gradient in the image pixel and $L_m < 0$ for the magnitude of the image Laplacian in the image pixel. Let's define three predicates (12), (13), (14) for testing image pixels:

$$P(M) = \begin{cases} \text{True, if } L_m < 0 \\ \text{False, otherwise} \end{cases} \quad (12)$$

$$Q(M) = \begin{cases} \text{True, if } G(M) > G_m \text{ and } L(M) = \text{False} \\ \text{False, otherwise} \end{cases} \quad (13)$$

$$N(M) = \bar{P}(M) \text{ and } \bar{Q}(M) \quad (14)$$

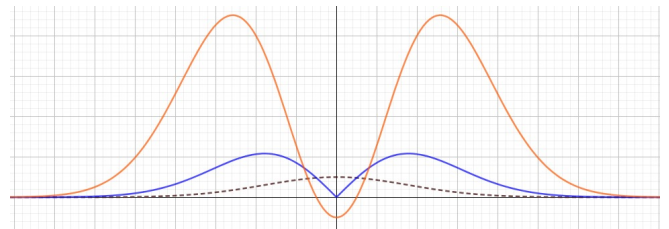


Fig. 2. Graphs of the distribution functions of the modulus of the gradient (blue) and Laplacian (red) in the cross section of the laser strip, the intensity of which is given by the Gaussian function (black dashed line)

Obviously, an image pixel can satisfy either one of these three predicates and not satisfy the other two, or not satisfy any of these predicates. Consequently, the set of image pixels D is divided into three pairwise disjoint sets $S_1 = \{M \in D \mid P(M) = \text{True}\}$, $S_2 = \{M \in D \mid Q(M) = \text{True}\}$, $S_3 = \{M \in D \mid N(M) = \text{True}\}$ moreover, an arbitrary image pixel belongs to one of these sets ($S_1 \cup S_2 \cup S_3 = D$). Let us construct a discrete function $F(M)$ (15) given on the pixels of the image:

$$F(M) = \begin{cases} 0, N(M) = \text{True} \\ 1, Q(M) = \text{True} \\ 2, P(M) = \text{True} \end{cases} \quad (15)$$

The laser strip corresponds to a characteristic pattern - a band within which $F = 2$ (the value of the Laplacian is less than the threshold value), framed by areas in which $F = 1$ (the value of the gradient modulus is greater than the threshold value and the Laplacian value is greater than the level of the negative threshold L_m) (see Fig. 3).

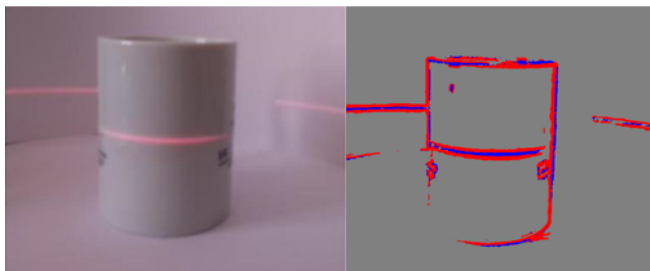


Fig. 3. Image of the scanned object (left). Visualization of the distribution of the F function (right). Regions of the image in which the modulus of the image gradient is higher than the specified threshold G_m are highlighted in red. Regions of the image in which the value of the Laplacian of the image is less than the level of the negative threshold L_m are highlighted in blue

To detect a laser stripe, the proposed method uses scanning of a discrete function in one of the directions (in rows, columns, or diagonals). The scanning direction is selected either according to the previously known orientation of the laser strip on the image, or scanning is carried out in three directions, followed by the procedure for selecting the scanning results.

The scanning procedure was implemented using a Finite State Machine (FSM), the alphabet of input symbols of which was the set of values of the function F , and the set of states of which corresponded to the stages of the pattern $(F = 0) \rightarrow (F = 1) \rightarrow (F = 2) \rightarrow (F = 1) \rightarrow (F = 0)$. The strip boundaries were determined during the transition from the interval $F = 0$ to the interval $F = 1$ and the reverse transition from the interval $F = 1$ to the interval $F = 0$. In addition, during the scanning process, intervals with a high level of the modulus gradient value ($F = 1$) were determined, on which the direction of the gradient vector was subsequently calculated.

C. Laser Line Detection Algorithm.

As noted above, preliminary detection of the laser stripe region in the image greatly simplifies the procedure for approximating the curve γ . The following is a brief description of the detection method, which is only one of the possible methods for detecting a laser stripe. Since the laser stripe image area detection algorithm allows estimation of the stripe width and its position, it becomes possible, in particular, to apply the methods described by Sun et al. in paper [4]. So, in the case of a positive result of scanning the line, in the region of a high value of the gradient modulus, the point $P(x_p, y_p)$ was chosen, which is the middle of the corresponding interval. Based on the scanning results, the interval $[a, b]$ was selected on the straight line l , passing through the point P along the direction vector of the gradient (e_{xp}, e_{yp}) , calculated for the point P . The constraint $f(x)$ of the gradient intensity function I on the

straight line l on the segment $[a, b]$ was constructed by the method of linear interpolation.

To determine the coordinate of the point of intersection $Q(x_q, y_q)$ of the straight line l with the curve γ , the position of the center of mass x_c of a flat figure bounded by the graph $f(x)$ was calculated using formula (16):

$$x_c = \frac{\int_a^b x \cdot f(x) dx}{\int_a^b f(x) dx} \quad (16)$$

After that, the coordinates of the point $Q(x_q, y_q)$, were calculated using formulas (17) and (18):

$$x_q = x_p + x_c \cdot e_{xp} \quad (17)$$

$$y_q = y_p + y_c \cdot e_{yp} \quad (18)$$

Thus, for each laser stripe interval detected during scanning of the line, the position of the corresponding point of the imaginary "ideal" curve γ , that is, the image of the line of intersection of the laser beam plane with the surface was calculated. Further, the constructed points of the curve were connected into a sequence of vertices of a broken line approximating the curve γ .

D. Software implementation of the algorithm and discussion of the results

The algorithm was implemented and tested as software [9] written in C# for detecting a laser stripe on the image of a scanned object and representing an elongated area of the laser stripe as a broken line without self-intersections (Fig. 4).

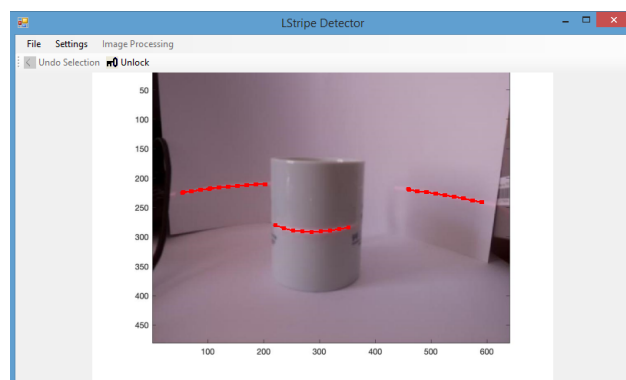


Fig. 4. View of the main window of the software [9] after performing the laser stripe detection procedure. The constructed broken lines are highlighted in red in the image

When a color image was input to a computer program, it was first converted into a black and white image using a standard algorithm. As emphasized above, the described general detection algorithm consists of two sequentially applied algorithms: the detection algorithm for the laser stripe area in the image, and the detection algorithm itself, which reconstructs the image of the intersection curve of the laser beam plane and the surface. To assess the impact of the

parameters of the detection procedure on its quality, the graphical interface of the software [9] allows to visualize the map of the distribution of the values of the Laplacian of the image and the modulus of the image gradient, as shown in Fig.5.

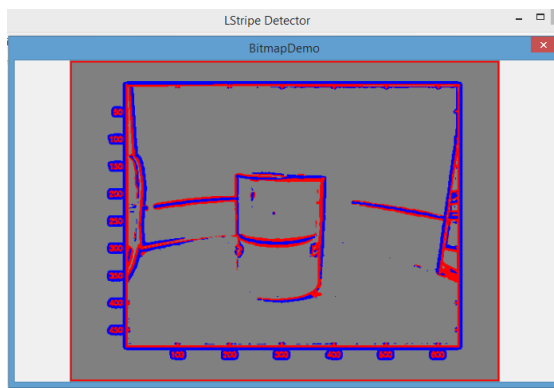


Fig. 5. Pop-up window of the computer program [9] with a map of the distribution of the values of the Laplacian of the image and the module of the image gradient

When testing, the reliability of the original algorithm for detecting the laser strip in the image was important primarily. Testing on the images at the disposal of the authors showed the high reliability of the proposed algorithm with the expected high speed. Since a set of images corresponding to such different initial conditions as various lasers used, different types of scanned surfaces, various types of video cameras used in testing was very limited, this work can be considered as an announcement of the method, designed to draw attention to the proposed detection method for assessment of its capabilities and restrictions.

IV. CONCLUSIONS

A simple algorithm for extracting areas of a black-and-white image containing images of a laser stripe is developed. The algorithm is based on the calculation of the gradient and Laplacian of the image intensity using standard digital image processing procedures. The proposed algorithm is characterized by low computational cost and does not imply any a priori knowledge about the detected object and image, in particular, no knowledge of the strip width is required.

Preliminary application of the described algorithm to the image, in order to subsequently apply one of the known methods for detecting a laser stripe in the regions of stripe localization already detected at the previous stage, can

significantly increase the overall performance and reliability of the detection algorithm

A new algorithm for converting a laser stripe image into a broken line is developed, suitable for use in 3D scanning tasks using laser illumination.

The results of the research are of significance for a wide range of researchers developing image processing algorithms for 3D scanning with laser stripe illumination and machine vision methods using structured light.

ACKNOWLEDGMENT

This research was funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP13068317).

REFERENCES

- [1] A. Kussaiyn-Murat, A. Krasavin, D. Alontseva, A. Kadyroldina, A. Khozhanov, Iu. Krak, P. Muñoz de Escalona, and I. Dyomina, "Development of an intelligent robotic system for plasma processing of industrial products with complex shape", 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), pp. 572-579, 2021, DOI: 10.1109/IDAACS53288.2021.9660960
- [2] X. Shi, Y. Sun, H. Liu, et al., "Research on laser stripe characteristics and center extraction algorithm for desktop laser scanner", SN Appl. Sci. 3, vol. 298 (2021), pp. 1-12, 2021, <https://doi.org/10.1007/s42452-021-04309-w>
- [3] J. Wang, W. Cheng, W. Li, X. Ma, and W. Yan, "Research on extraction method of centerline of large format laser stripe", Proc. SPIE 12079, Second IYSF Academic Symposium on Artificial Intelligence and Computer Engineering, vol. 1207927, 1 December 2021, <https://doi.org/10.1117/12.2622741>
- [4] Q. Sun, J. Chen, and C. Li, "A robust method to extract a laser stripe centre based on grey level moment", Optics and Lasers in Engineering, vol. 67, pp. 122-127, 2015.
- [5] J. Forest, J. Salvi, E. Cabruja, and C. Pous, "Laser stripe peak detector for 3D scanners. A FIR filter approach", Proceedings - International Conference on Pattern Recognition, vol. 3, pp. 1-5, 2004, DOI: 10.1109/ICPR.2004.528.
- [6] R. Gonzalez, R. and Woods, Digital Image Processing, 2nd Edition, Ad. Prentice Hall, 2002.
- [7] C. Steger, "An unbiased detector of curvilinear structures," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 20, no. 2, pp. 113-125, 1998.
- [8] A. Kadyroldina, A. Kussaiyn-Murat, B. Bertalan, A. Krasavin, D. Alontseva, "Image Acquisition and Processing on Raspberry Pi in Matlab for 3D-Scanning", PROCEEDINGS of 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University, Székesfehérvár, Hungary, pp. 39-42, November 2019
- [9] A.L. Krasavin, D.L. Alontseva, A.T. Kadyroldina, The computer program for laser strip detection in digital images of 3D scanning objects. Certificate of the Republic of Kazakhstan for the copyright object (software) No. 5870 of October 17, 2019

Stabilization of environmental conditions to improve the performance of a mobile application for the state of the driver monitoring

Olga Shvets
IT and intelligent system
sub-department
East-Kazakhstan
Technical University
Ust-Kamenogorsk, Kazakhstan
olga.shvets75@gmail.kz

Bauyrzhan Smakanov
IT and intelligent system
sub-department
East-Kazakhstan
Technical University
Ust-Kamenogorsk, Kazakhstan
bauka.10.10@mail.ru

György Györök
The Alba Regia
Technical Faculty
Obuda University
Székesfehérvár, Hungary
gyorok.gyorgy@amk.uni-obuda.hu

Abstract—The use of modern digital technologies in the field of security enhancement is a promising area of research. The results of a study in the field of increasing the level of human security in the case of monotonous, routine work using neural networks were reported at the AIS2021 conference. It has been developed a model and an algorithm for detecting a dangerous state of the driver (wakeful or sleeping). This work has been continued. The conditions of the environment in which the developed software is intended to be used depend on the smartphone used and how the smartphone is mounted in the vehicle. It should be noted that in a real environment, smartphones work with not always complete and accurate data. These factors directly affect the performance of the mobile app by changing the speed, accuracy, and functionality of the mobile app. This paper presents algorithms for manual and automatic calibration of the system based on data from the camera, sensors and settings of the driver's smartphone. It is proposed to use a mandatory calibration method that adapts to the current context of the driver and vehicle in order to reduce errors in various calculations of physical quantities numerical values and ensure the uniformity of measurements during the operation of the system. The calibration takes into account the input information about the driver, vehicle and smartphone. The processing of the critical area and the state of the driver in the image were implemented in the C ++ programming language using the computer vision software libraries OpenCV and Dlib, and the JNI interfaces were realized in the Java language.

Keywords—neural networks, emotions recognition, software development

I. INTRODUCTION

Currently, there are enough works aimed at improving human safety with the help of driver monitoring systems. Some scientists use methods and algorithms for tracking the closed eyes of the driver in their developments. For example, in the article [1] such models are used. Attention is paid to the health of the driver, for example, pulse, pressure, temperature are measured [2]. The paper [3] uses equipment to control the biometric indicators of the driver. In a study [4], published by us in the journal JATIT, the method of video analytics was used. The article [4] provides a detailed analysis of the literature on the research topic over the past 5 years. The results of a study in the field of increasing the level of human security in the case of monotonous, routine work using neural networks were reported at the AIS2021 conference [5]. At the moment, we have paid attention to the problem of environmental conditions in the cabin of the vehicle. How the smartphone is fixed in the car, as well as the phone model, directly affect the performance of the

mobile application, its speed, accuracy and functionality. To solve this problem, it is supposed to use a mandatory calibration method that adapts to the current context of the driver and vehicle. The calibration takes into account the input information about the driver, vehicle and smartphone.

II. MANUAL CALIBRATION OF THE SYSTEM

From a technical point of view, the method of calibrating a certain parameter is the determination of the characteristics of the deviation of a particular physical quantity associated with the measurement conditions, namely, the comparison of the numerical value of a physical quantity measured using the driver's smartphone with a separate value or falling into the measurement range, previously measured and established on the basis of analysed scientific studies and test data from drivers of a driver condition monitoring system. By analysing and taking into account the profile and contextual information about the driver and the characteristics of the hardware and software of the smartphone, two calibration modes were identified by the nature of the implementation, which are the initial manual calibration carried out directly by the driver of the vehicle and automatic calibration - already in the process of using the mobile application.

Manual calibration of the system is mandatory at the first start after installation and can be recommended to the driver if there is a need to re-configure and adapt the mobile application to the current conditions of use.

This type of calibration is intended to ensure the initial correct operation of the monitoring system. When using manual calibration, the parameters (Figure 1) to be calibrated are the position and size of the driver's face and head (the face is directed to the smartphone's front camera, there are no foreign objects obstructing the view in the visibility zone, and the face itself does not go beyond the boundaries of the smartphone's camera viewfinder), the volume level of audible warning signals, etc.



Fig. 1. Normal positions

The parameters characterizing the position of the driver's head are the angle of inclination and the angle of rotation of the head (see Figure 2).

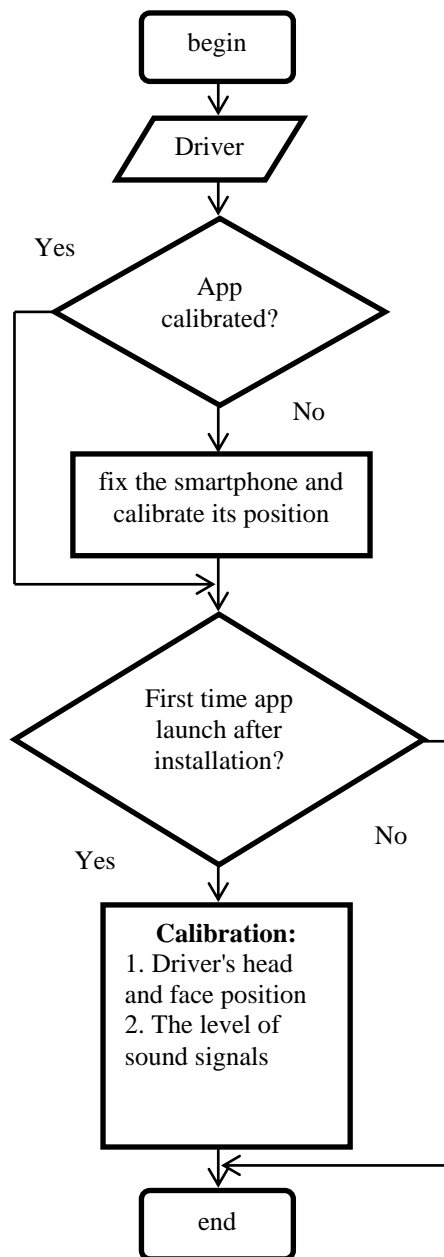


Fig. 2. Manual calibration algorithm

In this calibration mode, the driver's head deviation angle is calculated with respect to the plane of the front camera of the smartphone. This information about the position of the head allows to detect a dangerous deviation of the angle of inclination / rotation of the head from the one specified in the application settings. Determining the angle of rotation / inclination of the driver's head:

Begin

1. $diff \leftarrow dAngle - offsetAnglePref$
2. $cAngle \leftarrow diff * (if\ diff < 0\ do\ (-1)\ else\ 1)$
3. $isDangerAngle \leftarrow cAngle > anglePref\ or\ cAngle < -anglePref$

End

where

- $dAngle$ is the recognized angle of head rotation/tilt relative to the camera plane,
- $offsetAnglePref$ is the angle of deviation of the head rotation/tilt from zero degrees,
- $anglePref$ is the allowable head rotation/tilt angle (non-negative value),
- $cAngle$ is the driver's head position angle, taking into account the specified deviation

III. AUTOMATIC CALIBRATION OF THE SYSTEM

In addition to the manual mode, the automatic mode is another option for using calibration, which consists in program setting and adaptation of the parameters of the mobile software complex of the monitoring system without the participation of the driver during the operation of the application. The automatic calibration mode is mainly used in the case when during the operation of the monitoring system a large number of consecutive false alarms about dangerous conditions occur in the absence of a reaction from the driver, confirming or rejecting this or that event. In this case, as a result of applying this calibration mode, the dangerous states detected by the monitoring system are transferred to the "non-dangerous" category and are taken into account by the system when further adapting the mobile software package to the driver's driving style.

The automatic calibration mode is described in more detail as follows. Images of the face of one driver, obtained as a result of the operation of the front camera of a smartphone at various positions of his head, including the angle of rotation to the left or right, or its angle of inclination forward or backward, are one of the main sources of information used by the monitoring system when determining a dangerous condition. Thus, the characteristics of the head position have a direct impact on the operation of the facial characteristics recognition module from the frontal image of the driver's face.

The automatic calibration mode algorithm (see Figure 3) consists in iterative and regularly repeated adjustment of the driver's facial characteristics by assessing and taking into account deviations in the angles of inclination and rotation of the head in the plane of the front camera of the smartphone.

A predetermined number of measurements (observations) of the driver's head position angles, collected over a certain time series in the process of performing automatic calibration, make it possible to describe each input image from the video sequence of the smartphone's front camera by a set of certain real numbers denoting the degree of head rotation and inclination in the range from 0° up to 360° . A threshold operation is applied to the resulting interval set of numbers calculated for each image, which makes it possible to filter out the values that make the set noisy with inconsistent values. For example, as a threshold operation, you can choose the lower (or upper) threshold so that the number of frames from the video sequence in which the tilt (or rotation) angle exceeds the threshold is less (or more) than the predicted proportion of all observed images from the front camera of the smartphone.

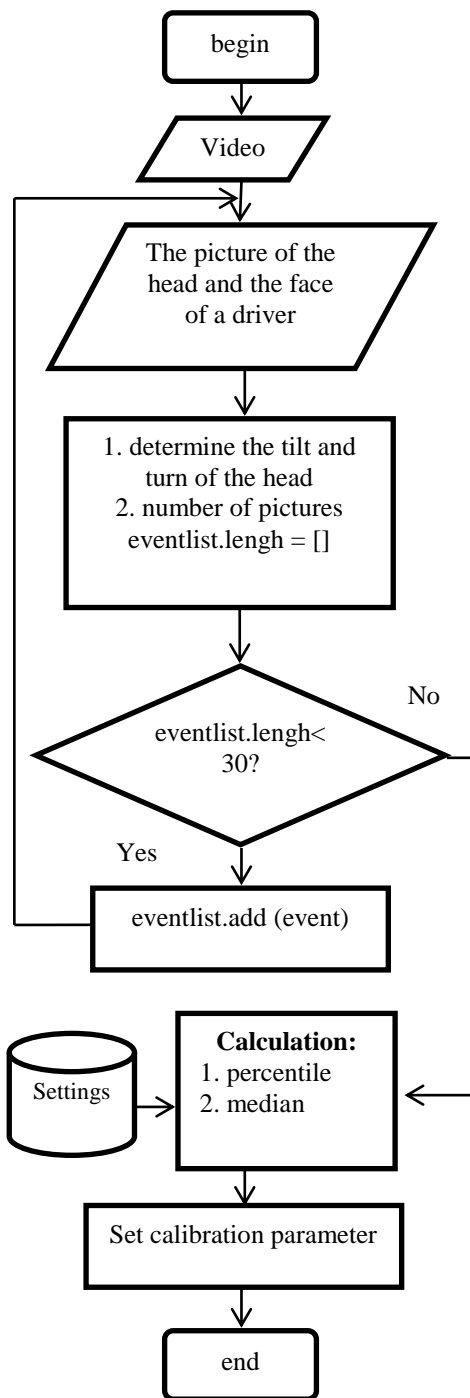


Fig. 3. Automatic calibration algorithm

The described method is based on the use of a percentile represented by such a number that the given random value does not exceed it only with a fixed probability given in percent. The final step in the automatic calibration mode is the calculation of the median of the filtered range of values. It will determine the calibrated observed parameter (head tilt or turn) and save it to the driver settings stored on the smartphone.

It is known, that the calibration process is based on information about the objects that are captured by the camera: their proportions, orientation, and spatial position (see Figure 4).

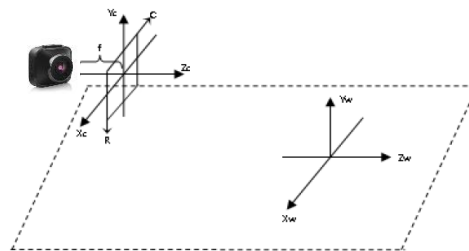


Fig. 4. The model of the visual system of the surveillance camera [4]

The model of the visual system of the surveillance camera includes:

- Rectangular image coordinate system with integer R and C axes, where R and C coordinates are the numbers of rows and columns of pixels in the image.
- Rectangular camera coordinate system with real axes Xc, Yc, Zc.
- Rectangular world coordinate system with real axes Xw, Yw, Zw.

The optical center in the camera system has coordinates (0, -f, 0). The image plane RC, modeling the camera matrix, lies in the XcYc plane, and the Zc axis is the optical axis of the camera [4]. This model is described in details in our research paper [4].

IV. SOFTWARE

It was found as a result of the experiments that the resolution of the processed image in width and length, not exceeding 300 pixels on each side, is sufficient for most modern smartphone models.

The processing of the critical area, the state of the driver's behavior in the image, was implemented in the C++ programming language using the computer vision software libraries OpenCV and Dlib, and the JNI interfaces were implemented in the Java language. The OpenCV open cross-platform computer vision library developed by Intel, as well as the Dlib machine learning library, greatly simplifies computer vision programming by providing a convenient interface for face detection, tracking and recognition. An important component in OpenCV is the mathematical apparatus and functionality for image processing.

In the general case, the process of image processing and face search consists of the following successive steps: face detection and localization in the image, face image alignment (geometric and brightness), feature calculation and direct recognition - comparison of the calculated features with the standards stored in the database.

An approach is used for each frame coming from the front camera of a smartphone to analyze facial characteristics based on combining image processing algorithms from the Dlib and OpenCV libraries, implemented in C++ and focused on selecting faces, key points on it and working with these points. The following algorithms are used to search and highlight facial characteristics in images, which are included in the package from Dlib:

- Calculation of descriptors by the method of histogram of directed gradients (HOG, Histogram of Oriented Gradients) and search for faces in the image;

- Use of the support vector machine to classify descriptors;
- Application of a set of decision trees (Random Forest) to select 68 key points on a human face (see Figure 5).



Fig. 5. Program settings window

The mobile application sets a time period equal to 10 seconds, to ensure the completeness and accuracy of the calculated parameters, during which the results of the analysis of the driver's facial characteristics are accumulated from the front camera of the smartphone.

This task is performed using Android's built-in CountdownTimer class, which allows you to implement a programmatic countdown timer. During this calibration, the driver's head must be directed straight ahead in the direction of the vehicle. The CircularFifoQueue class is a standard implementation of this type of queue from the Apache Commons Collections library. The return parameter of the findAverage function is an object of the Pair system class, containing as the first argument the SMA (Simple Moving Average) for the head left/right angle, and the second argument SMA for the forward/backward tilt of the driver's head. Calculation of the calibrated parameters of the driver's head on the interval private fun findAverage (list: CircularFifoQueue<PureDistraction>):

```
Pair<Float, Float> {
    val iter = list.iterator()
    var sumYawAngle = 0f
    var sumPitchAngle = 0f
    while (iter.hasNext()) {
        val item = iter.next()
        sumYawAngle +=
        item.yawHead.angleHead.detectedAngleValue
        sumPitchAngle +=
        item.pitchHead.angleHead.detectedAngleValue }
    val size = list.size.toFloat()
    val yaw = sumYawAngle / size
    val pitch = sumPitchAngle / size
```

```
return Pair(yaw, pitch) }
```

The mobile application installed on the driver's smartphone is designed for devices based on the Android operating system with the minimum supported version of Android 4.2. It is a result of our work.

Our system has a number of advantages despite of many very good existing developments in the field of driver monitoring:

- 1) Portability, because the system is installed on the driver's smartphone and can be used in different cars by the same person.
- 2) The cost for local Kazakh drivers is lower than the cost of proposed existing systems by about 30%.
- 3) The system is adapted to local road conditions (roughness of the road surface in some places, off-road, sharply continental climate, causing extreme fluctuations in temperature, etc.).
- 4) Our software can be customized for a specific driver to improve the accuracy of situation recognition.

We are going to continue the research and realize special platform for smartphone holder.

V. CONCLUSION

Algorithms have been developed for automatic and manual calibration of the driver's condition monitoring system based on data from the camera, sensors and settings of the driver's smartphone in order to minimize the influence of external conditions on the implementation of the mobile application and to maximize the adaptation of the system for the driver. All the considered methods and algorithms are implemented in a software package in the C # programming language and Java.

ACKNOWLEDGMENT

Part of the research was carried out with the support of the East-Kazakhstan technical university.

REFERENCES

- [1] J. Orlovska, F. Novakazi, B. Lars-Ola, M. Karlsson, C. Wickman, R. Söderberg, Effects of the driving context on the usage of automated driver assistance systems (ADAS) - naturalistic driving study for ADAS evaluation, *Transportation Research Interdisciplinary Perspectives*, 4, 2020, p. 100093, DOI: 10.1016/j.trip.2020.100093 <http://www.sciencedirect.com/science/article/pii/S259019822030004X>.
 - [2] Q. Chao, H. Bi, W. Li, T. Mao, Z. Wang, M.C. Lin, Z. Deng, "A survey on visual traffic simulation: Models, evaluations, and applications in autonomous driving", *Computer Graphics Forum*, 2020, <https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.13803>, DOI:10.1111/cgf.13803. arXiv: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/cgf.13803>.
 - [3] M. Johns, G. Strack, W. Ju, "Driver assistance after handover of control from automation", 21st International Conference on Intelligent Transportation Systems (ITSC), 2018, pp. 2104-2110.
 - [4] O. Shvets, B. Smakanov, L. Kovacs, G. Gyorok. Intelligent system for driver support using two classifiers for simulation. *JATIT*, 15th August 2022. Vol.100. No 15, pp 4767-4782. <http://www.jatit.org/volumes/Vol100No15/20Vol100No15.pdf>
- O. Shvets, B. Smakanov and Zh. Bakatbayeva. Monitoring Of The State Of A Person At Hazardous Work. *Proceedings of 16th International Symposium on Applied Informatics and Related Areas*, 2021, p. 88-92.

Autonomous life support using intelligent analysis in Kazakhstan

Olga Shvets
*IT and intelligent system
 sub-department
 East-Kazakhstan
 Technical University*

Ust-Kamenogorsk, Kazakhstan
 olga.shvets75@gmail.kz

Assel Naizabayeva
*IT and intelligent system
 sub-department
 East-Kazakhstan
 Technical University*

Ust-Kamenogorsk, Kazakhstan
 aselnaizabaeva@mail.ru

Márta Seebauer
*The Alba Regia
 Technical Faculty
 Obuda University*

Székesfehérvár, Hungary
 seebauer.marta@uni-obuda.hu

Abstract—The Republic of Kazakhstan has chosen a fundamentally new path for the development of society with the adoption by of the “Strategy “Kazakhstan 2050” and the Concept of transition to a “green” economy. One of the central points in the gradual transition to a green economy is energy efficiency. The use of modern neural network technologies in this area is a promising and relevant area of research. In this paper, we propose a model consisting of two similar perceptrons, one of which is applicable for predicting the hourly load profile of a working day, the other for a weekend or holiday. It was developed the software "Smart House Energy Management System". The software product includes four main blocks: Measurement data; Power indicators; Generation and consumption of electricity; and predictive model. The demo example predicted the current drawn by the load for two connected devices. It is possible to play different scenarios, disconnect and connect devices. We can adjust the work of alternative energy sources (we have solar panels) depending on the forecast.

Keywords— *smart house, intelligent analysis, energy efficiency*

I. INTRODUCTION

Now much attention all over the world and in Kazakhstan in particular is paid to the issues of energy saving and energy efficiency. This topic has been relevant for many years. This shows the presence of publications devoted to this issue, for example [1-11]. Particular attention is paid to solving the problems of uninterrupted power supply of residential buildings. The Republic of Kazakhstan has chosen a fundamentally new path for the development of society after the adoption of the “Strategy “Kazakhstan 2050” and the Concept of transition to a “green” economy. According to the Concept, the focus of state policy on reducing environmental impact, resource conservation and achieving a high level of quality of life for the population will play a key role. Kazakhstan has significant opportunities for improving energy efficiency in industry, energy, housing and communal services and transport [12]. Energy efficiency is one of the central points of a phased transition to a green economy.

The aim of the study is to develop software for regulating the energy consumption of a residential building using, along with the traditional source of an alternative energy source - solar panels.

We have discussed this problem in detail during Hungarian Science Festival 2020 by Óbuda University, 12.11.2020 [13]. After the study, we came to the conclusion that the topic is practically significant and very relevant.

II. MODELING

We propose a model consisting of two similar perceptrons, one of which is applicable for predicting the hourly load profile of a working day, the other for a weekend or holiday. In the proposed model, the learning processes for workdays and weekends (or holidays) were separated, i.e., four matrices of synaptic weights were calculated (two for each perceptron). The architecture of the neural network includes three layers: Input, Hidden and Output (see Figure 1).

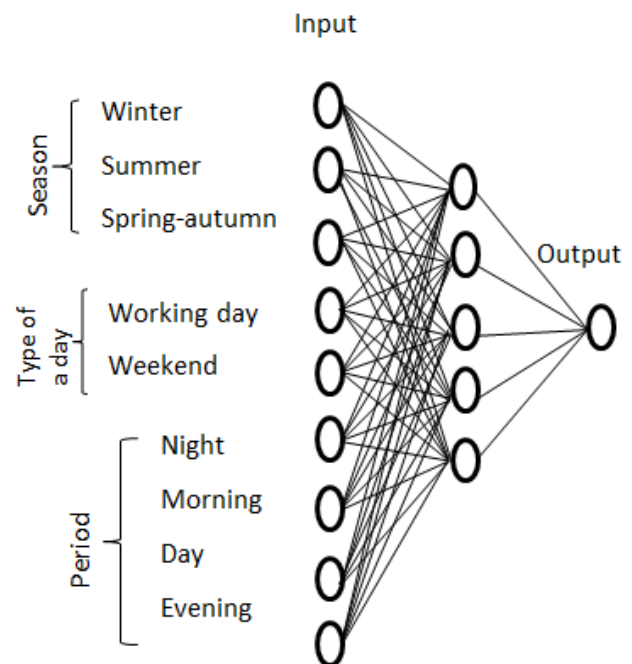


Fig. 1. The architecture of the neural network

Input layer consists from 3 groups:

- 1) Season (winter, summer, spring-autumn),
- 2) Type of a day (weekend or working day) and
- 3) Period (night, morning, daytime, evening).

Input data of the neural network include information about seasons, type of a day and time of day (period). Output data present a profile of the power consumption: enough or not the energy from solar panel or we need additional energy.

18 daily load curves were selected, of which 16 formed a training sample, and 2 - a control one to train a neural network that predicts the power consumption of individual residential buildings. The initial synaptic weights were

chosen in such a way that before the training of the neural network, the relative load of a residential building with electric stoves during the morning peak in autumn will be 50%.

The backpropagation algorithm was used for training. Due to the limited training sample, it was necessary to repeat the training cycle. The cyclic loading of the training sample into the neural network in both cases was repeated until the root-mean-square change in the weights of neurons during the training cycle fell below 3%. For this, 12 cycles were enough.

III. THE EXPERIMENTAL INSTALLATION

The choice of generators for the experimental setup was dictated by the available equipment. Monocrystalline and polycrystalline photovoltaic modules were used in this installation. Once installed on the roof, they can be considered as built-in renewable energy sources.

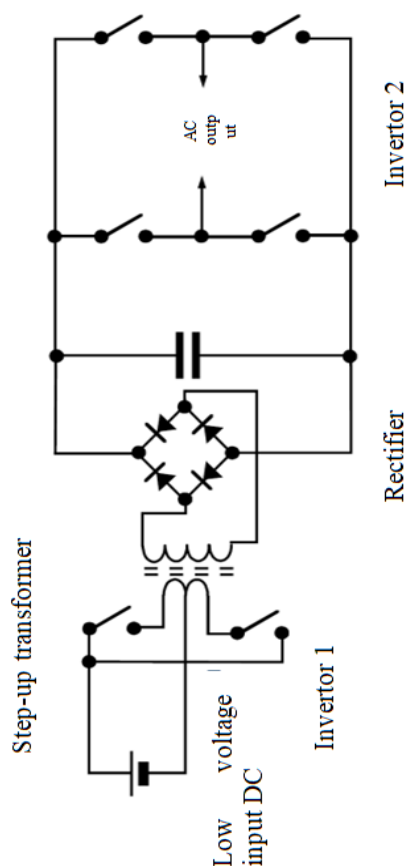


Fig. 2. The topology of the inverter

Equipment is needed to ensure interaction with the grid that converts direct current to alternating current and changes the voltage to the desired level, since batteries and photovoltaic modules operate on direct current, while the electrical network is 230 volts AC, 50 Hz. In other words, the use of an inverter is required. In the system under consideration, the inverter acts as a bridge between the part of the system containing generators and storage and the part containing loads.

As part of this work, a 550 W network inverter was designed and built. Let's define the basic connection diagram. The system will have two buses: a DC bus and an AC bus. All DC installations (photovoltaic, DWT, ballast

load, batteries) are connected to the DC bus, and all AC installations (loads, mains) are connected to the AC bus. The inverter connects to both buses and links them together.

It is necessary to know the energy flow going from and to each installation object to connect measuring equipment. This flow, depending on the type of device, must either be measured or controlled to the desired value. The topology of the inverter used in the experimental setup is shown in Figure 2.

This part of the experimental setup consists of an inverter (in this case a breaker) followed by a transformer, which in turn is followed by a rectifier. Although it is more difficult, it has certain advantages. When the inverter output passes through a transformer, the transformer must be sized to operate with the inverter's output frequency.

The proposed system required some way to make electrical loads "smart" using fuzzy logic elements. The practical part of designing a Smart Plug device is discussed in detail and published in the materials of the AIS2020 symposium [13].

IV. SOFTWARE

The software consists of four main blocks: Measurement data; Power indicators; Generation and consumption of electricity; Predictive model.

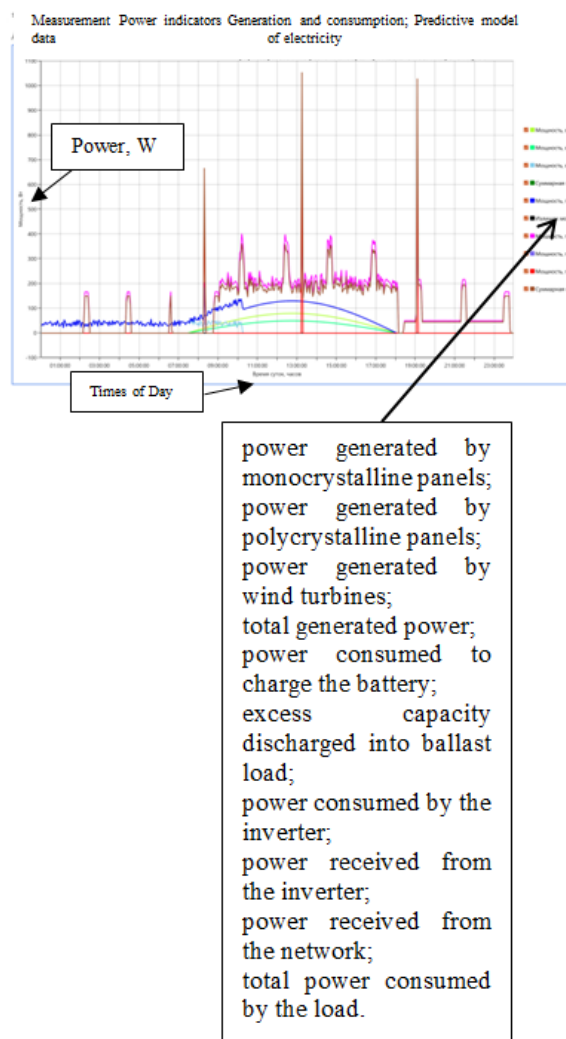


Fig. 3. Tab "Power indicators"

Figure 4 shows graphs for three indicators depending on the time of day:

- the power received from the inverter;
- power received from the network;
- the total power consumed by the load.

When we click on the button "Calculate the balance of electricity for the day" the calculation of the following indicators is calculate:

- generated by single-crystal panels, $W \cdot h$;
- produced by polycrystalline panels;
- produced by wind turbines;
- total electricity generated;
- power consumed to charge the battery;
- consumed from the inverter;
- consumed from the network;
- total electricity consumption.

The figure shows an example of forecasting for two devices turned on simultaneously.

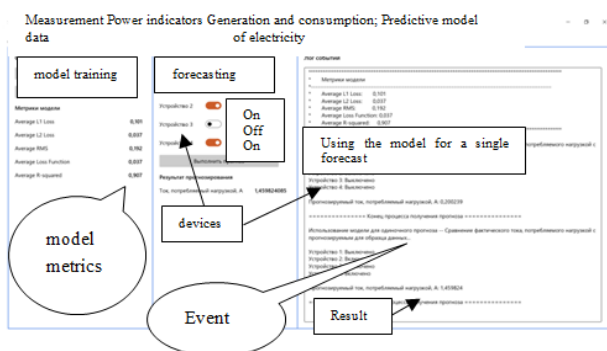


Fig. 4. An example of forecasting for two switched on devices and certificate of authorship for software

The software is implemented in the C # environment in accordance with the developed schemes and requirements. (see Figure 5).



Fig. 5. Certificate of authorship for software № 16772 from 14.04.2021 "Residential building intelligent energy management system "Smarthouse"

The interface of neural connections is presented below; each connection stores a weight and a pointer to a neuron (see Figure 6).

```

template <typename T>
class Neuron;

template <typename T>
class NeuralLink
{
public:
    NeuralLink( ) : mWeightToNeuron( 0.0 ),
        mNeuronLinkedTo( 0 ),
        mWeightCorrectionTerm( 0 ),
        mErrorInformationTerm( 0 ),
        mLastTranslatedSignal( 0 ){};

    NeuralLink( Neuron<T> * inNeuronLinkedTo,
        double inWeightToNeuron = 0.0 ) :
        mWeightToNeuron(
            inWeightToNeuron ),
        mNeuronLinkedTo(
            inNeuronLinkedTo ),
        mWeightCorrectionTerm( 0 ),
        mErrorInformationTerm( 0 ),
        mLastTranslatedSignal( 0 ){};

    void SetWeight( const double& inWeight ){ mWeightToNeuron = inWeight; };
    const double& GetWeight( ) { return mWeightToNeuron; };

    void SetNeuronLinkedTo( Neuron<T> * inNeuronLinkedTo ){ mNeuronLinkedTo =
        inNeuronLinkedTo; };
    Neuron<T> * GetNeuronLinkedTo( ) { return mNeuronLinkedTo; };

    void SetWeightCorrectionTerm( double inWeightCorrectionTerm ){
        mWeightCorrectionTerm = inWeightCorrectionTerm; };
    double GetWeightCorrectionTerm( ) { return mWeightCorrectionTerm; };

    void UpdateWeight( ){ mWeightToNeuron = mWeightToNeuron +
        mWeightCorrectionTerm; };

    double GetErrorInformationTerm( ) { return mErrorInformationTerm; };
    void SetErrorInformationTerm( double inEITerm ){ mErrorInformationTerm =
        inEITerm; };

    void SetLastTranslatedSignal( double inLastTranslatedSignal ){
        mLastTranslatedSignal = inLastTranslatedSignal; };
    double GetLastTranslatedSignal( ) { return mLastTranslatedSignal; };
protected:
    double mWeightToNeuron;
    Neuron<T> * mNeuronLinkedTo;
    double mWeightCorrectionTerm;
    double mErrorInformationTerm;
    double mLastTranslatedSignal;
};
    
```

Fig. 6. Listing of the interface of neural connections

The neural network itself stores pointers to neurons organized in layers (in general, pointers to neurons are stored in vectors that need to be replaced with layer objects), includes an abstract factory of neurons, as well as a network learning algorithm.

It is not necessary to demonstrate all algorithms realized in the work.

Figure 7 shows the realization of neural network training.


```

template <typename T>
class NeuralNetwork;

template <typename T>
class TrainAlgorithm
{
public:
    virtual ~TrainAlgorithm(){};
    virtual double Train(const std::vector<T>& inData, const std::vector<T>&
inTarget) = 0;
    virtual void WeightsInitialization() = 0;
protected:
};

template <typename T>
class Hebb : public TrainAlgorithm<T>
{
public:
    Hebb(NeuralNetwork<T> * inNeuralNetwork) :
mNeuralNetwork(inNeuralNetwork){};
    virtual ~Hebb(){};
    virtual double Train(const std::vector<T>& inData, const std::vector<T>&
inTarget);
    virtual void WeightsInitialization();
protected:
    NeuralNetwork<T> * mNeuralNetwork;
};

template <typename T>
class Backpropagation : public TrainAlgorithm<T>
{
public:
    Backpropagation(NeuralNetwork<T> * inNeuralNetwork);
    virtual ~Backpropagation(){};
    virtual double Train(const std::vector<T>& inData, const std::vector<T>&
inTarget);
    virtual void WeightsInitialization();
protected:
    void NguyenWidrowWeightsInitialization();
    void CommonInitialization();
    NeuralNetwork<T> * mNeuralNetwork;
};

```

Fig. 7. The algorithm of neural network training

V. CONCLUSION

An intelligent system has been developed that allows to regulate uninterrupted power consumption from solar panels, supplementing this process with power supply from the AC network, to reduce the power consumption of a residential building. Moreover, the system gives preference to power from an alternative energy source - solar panels. The intelligence of the system lies in the selection of the ratio of alternative and traditional energy supply in favor of the alternative, i.e. more economical option. Uninterrupted power supply is also maintained during peak hours in case of insufficient power of solar panels.

All the considered methods and algorithms are implemented in a software package in the C # programming language.

We are going to continue the research and try to test our system in real conditions.

ACKNOWLEDGMENT

Part of the research was carried out with the support of the East-Kazakhstan technical university.

REFERENCES

- [1] BP Global Statistical Review of World Energy, 2021, <http://www.bp.com/>
- [2] Today in energy, Energy Information Agency, USA, Sept. 2021, <https://www.eia.gov/todayinenergy/>
- [3] Baklanov A., Naizabayeva A., Seebauer M., Soltan A. Efficient Monitoring and Control of Solar Energy Consumption by LED Lighting Devices with Embedded Microcontroller. AIS 2019, 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University, ISBN 978-963-449-156-9 November 14, 2019 Székesfehérvár, Hungary, p 13-16.
- [4] Jakkula, V., Youngblood, G. & Cook, D.: Identification of lifestyle behavior patterns with prediction of the happiness of an inhabitant in a smart home. Approaches to Beauty and Happiness. Available at: <http://www.aai.org/Papers/Workshops/2006/WS-06-04/WS06-04-005.pdf> [Accessed March 12, 2021].
- [5] Oliveira V., Jaschke J., Skogenstad S. Optimal operation of energy storage in buildings: Use of hot water system // Journal of energy storage. 2016. V.5. P. 102-112.
- [6] DNV's Energy Transition Outlook is an independent, model-based forecast of the world's most likely energy future through to 2050, 2021, <https://eto.dnv.com/2021#ETO2021-top>
- [7] He Y. Energy saving of central air-conditioning and control system: Case study: Nanchang Hongkelong Supermarket. Available at: <http://theseus17-kk.lib.helsinki.fi/handle/10024/21077> [Accessed March 4, 2021].
- [8] European policies in energy saving, 2019, <http://europa.eu/!Tj97Qn>.
- [9] Asaithambi S., Venkatraman S., Venkatraman R. Big Data and Personalisation for Non-Intrusive Smart Home Automation, Big Data and Cognitive Computing, 2021, 5. 6. 10.3390/bdcc5010006.
- [10] Beszédes B., Széll K., Györök G. Redundant photo-voltaic power cell in a highly reliable system, Electronics (Switzerland), V. 10, Issue 11, 1 June 2021, No 1253.
- [11] Björkskog C. Human Computer Interaction in Smart Homes. Helsinki, Finland. Available at: <http://www.hiit.fi/~oulasvir/58307110/smarthomes.pdf> [Accessed March 20, 2021].
- [12] Message from the President of the Republic of Kazakhstan Kassym-Jomart Tokayev, Nur-Sultan, March 16, 2022.
- [13] Shvets O., Seebauer M., Naizabayeva A., Toleugazin A.: Autonomous power supply systems optimization for energy efficiency increasing. 15th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2020 by Óbuda University, 12.11.2020, pp. 128-132.

Robust Control Design for Quadcopter for Path Following in Monitoring Tasks

Daniyar Kaliyev

School of Information Technology and
Intellectual Systems

D.Serikbayev East Kazakhstan
Technical University)

Ust-Kamenogorsk, Kazakhstan
daniyar.1996.samarka@gmail.com

Olga Shvets

School of Information Technology and
Intellectual Systems

D.Serikbayev East Kazakhstan
Technical University)

Ust-Kamenogorsk, Kazakhstan
olga.shvets75@gmail.com

György Györök

Alba Regia Technical Faculty
Óbuda University

Budapest, Hungary

gyorok.gyorgy@amk.uni-obuda.hu

Abstract—The control system design plays a critical role in unmanned aerial vehicles to perform pre-planned tasks such as fire exploration and monitoring operations. The UAV encounters various interferences when carrying out outdoor missions. Therefore, it is very important to work on the development of a robust control algorithm that is resistant to external disturbances and changes in parameters. This article proposes the structure of the robust UAV control design by setting their parameters using the example of a quadcopter, a robotic device that is perfect for monitoring tasks. The results of modeling and experiments confirm the good performance of the system in tracking trajectories even in the presence of disturbances.

Keywords—UAV, quadcopter, robust controller, PID controller, path tracking.

I. INTRODUCTION

In recent years, there has been a steady increase in interest in autonomous air vehicles, especially the most common type, the quadcopter, a four-engine UAV. The advantages of quadcopters are low cost, simplicity, few movement restrictions, and the ability to carry large payloads. Their use has increased due to their easy access to remote areas for wide applications such as monitoring, agricultural services, mapping and photography, combat damage assessment, border interception prevention, and others [1]. The drones' promising applications are exploration and fire monitoring [2, 3]. Environmental factors, such as gusts of wind, affect the quality of data collection (photos) during reconnaissance flights. From this point of view, it is essential to obtain reliable data. Therefore, a reliable controller with trajectory control algorithms is required for high-quality flight performance. For autonomous UAV navigation, the controller must include trajectory planning and path tracking [4]. There are several control laws to stabilize the quadcopter during flight (nonlinear, linear, robust, etc.). Many control laws have been developed using a simplified nonlinear system or linear models [5], [6]. In [7], a linear and nonlinear model predictive control scheme was presented for controlling a quadcopter to track various reference trajectories. The authors also presented an analysis of the stability of linear and nonlinear feedback control schemes. In [8], the authors proposed a unified motion control scheme for quadcopters, which solves the problems of stabilization, trajectory tracking, and the problem of path following. In a later paper [9], researchers proposed a trajectory generation algorithm, which allows you to quickly calculate high-performance flight trajectories for moving a quadcopter from a large class of initial states to a given target location. Thus, the actual task is to develop a controller for a quadcopter control system that allows autonomous flight along a given route with minor deviations

and is resistant to external influences. This work proposes a simple quadcopter controller structure and a trajectory control algorithm.

II. SYSTEM DESCRIPTION

This section describes the dynamic quadcopter model and controller circuit used in the experiments. The dynamic model of a quadcopter is obtained by representing the vehicle as a 3D rigid body driven by forces and torques generated by the propellers, see Fig. 1.

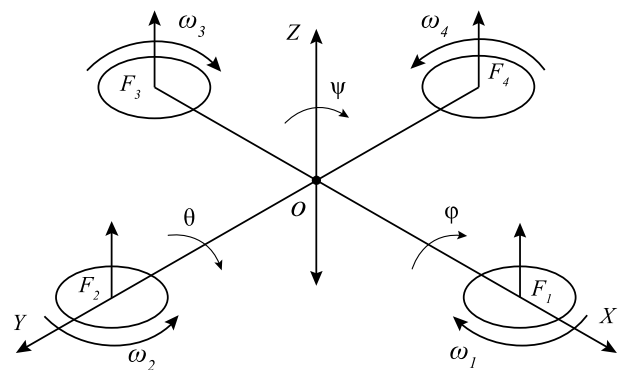


Fig. 1 Dynamic model of a quadcopter

A widely used simplified quadcopter model is shown below. [10]:

$$\begin{aligned} m\ddot{x} &= -u \sin \theta, & I_\psi \ddot{\psi} &= \tilde{\tau}_\psi \\ m\ddot{y} &= u \cos \theta \sin \phi, & I_\theta \ddot{\theta} &= \tilde{\tau}_\theta \\ m\ddot{z} &= u \cos \theta \cos \phi - mg, & I_\phi \ddot{\phi} &= \tilde{\tau}_\phi \end{aligned} \quad (1)$$

where x and y — horizontal coordinates, z — vertical position, g — acceleration of gravity, ψ , θ and ϕ — yaw, pitch and roll angles respectively, m — device weight, I_i — moment of inertia about the i - axis, and $\tilde{\tau}_i$ — torque acting on the i - axis.

The quadcopter control system can be divided into sub-controllers such as position controller (to control x and y), height controller (to control z), and attitude controller (to control the three Euler angles ϕ , θ , and ψ) (see Fig. 2).

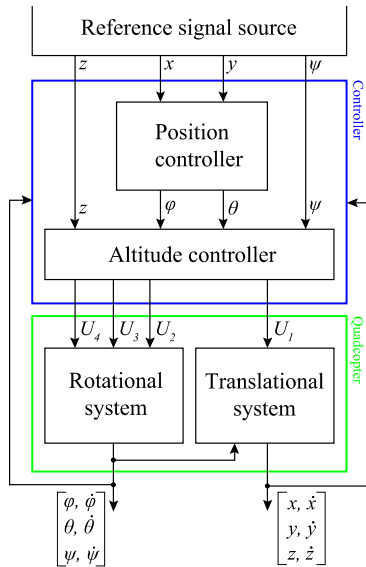


Fig. 2 Basic controller system structure for quadcopter

It is worth noting that in the system described above, the angles and their time derivatives are independent of the translation components. Ideally, we can imagine the structure described above consisting of two subsystems: angular rotations and linear displacements, as shown in Fig. 3.

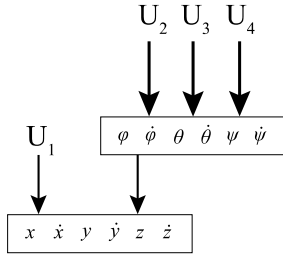


Fig. 3 Relationship between the rotational and translational subsystems of the quadcopter

III. METHODOLOGY

A. Control system for quadcopter

The mathematical equations of the quadcopter model without perturbations can be written as follows:

$$\ddot{x} = (\cos \varphi \sin \theta \cos \psi + \sin \varphi \sin \psi) \frac{U_1}{m} \quad (2.1)$$

$$\ddot{y} = (\cos \varphi \sin \theta \sin \psi - \sin \varphi \cos \psi) \frac{U_1}{m} \quad (2.2)$$

$$\ddot{z} = (\cos \varphi \cos \theta) \frac{U_1}{m} - g \quad (2.3)$$

$$\ddot{\varphi} = \frac{I_y - I_z}{I_x} \dot{\theta} \dot{\psi} - \frac{J_r}{I_x} \Omega_r \dot{\theta} + \frac{1}{I_x} U_2 \quad (2.4)$$

$$\ddot{\theta} = \frac{I_z - I_x}{I_y} \dot{\varphi} \dot{\psi} + \frac{J_r}{I_y} \Omega_r \dot{\varphi} + \frac{1}{I_y} U_3 \quad (2.5)$$

$$\ddot{\psi} = \frac{I_x - I_y}{I_z} \dot{\varphi} \dot{\theta} + \frac{1}{I_z} U_4 \quad (2.6)$$

The system can be rewritten in state space form:

$$\dot{X} = f(X, U) \quad (3)$$

where U – input vector, and X - state vector chosen as follows:

$$X = [x \ y \ z \ \varphi \ \theta \ \psi \ \dot{x} \ \dot{y} \ \dot{z} \ \dot{\varphi} \ \dot{\theta} \ \dot{\psi}]^T \quad (4.1)$$

$$U = [U_1 \ U_2 \ U_3 \ U_4]^T \quad (4.2)$$

Let us present the equations of connection of control channels U_1 , U_2 , U_3 , and U_4 with the rotation speeds of the quadcopter propellers. The quadcopter altitude changes with U_1 , steer angle changes with U_2 , pitch angle depends on U_3 and U_4 controls yaw angle. These input forces will help the operator move the quadcopter in all directions. Each input force is calculated by the equation shown below. Inputs are displayed as follows:

$$U_1 = k_f (\omega_1^2 + \omega_2^2 + \omega_3^2 + \omega_4^2) \quad (5.1)$$

$$U_2 = k_f (-\omega_2^2 + \omega_4^2) \quad (5.2)$$

$$U_3 = k_f (\omega_1^2 - \omega_3^2) \quad (5.3)$$

$$U_4 = k_z (-\omega_1^2 + \omega_2^2 - \omega_3^2 + \omega_4^2) \quad (5.4)$$

and

$$\Omega_r = \omega_1 - \omega_2 + \omega_3 - \omega_4 \quad (6)$$

where ω_i for $i = 1, 2, 3, 4$ denotes the speed of the i -th rotor, and T_i for $i = 1, 2, 3, 4$ - thrust generated by the i -th rotor, and the thrust $T_i(t)$ is equal to the function of the rotor speed, determined by the expression:

$$T_i(t) = k_f \omega_i^2 \quad (7)$$

where k_f and k_z - constant coefficients, and

$$f(X, U) = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{\varphi} \\ \dot{\theta} \\ \dot{\psi} \\ u_x \frac{U_1}{m} \\ u_y \frac{U_1}{m} \\ (\cos \varphi \cos \theta) \frac{U_1}{m} - g \\ \frac{I_y - I_z}{I_x} \dot{\theta} \dot{\psi} - \frac{J_r \Omega_r}{I_x} \dot{\theta} + \frac{1}{I_x} U_2 \\ \frac{I_y - I_x}{I_y} \dot{\varphi} \dot{\psi} + \frac{J_r \Omega_r}{I_y} \dot{\varphi} + \frac{1}{I_y} U_3 \\ \frac{I_x - I_y}{I_z} \dot{\varphi} \dot{\theta} + \frac{1}{I_z} U_4 \end{bmatrix} \quad (8)$$

where

$$u_x = \cos \varphi \sin \theta \cos \psi + \sin \varphi \sin \psi \quad (9.1)$$

$$u_y = \cos \varphi \sin \theta \sin \psi - \sin \varphi \cos \psi \quad (9.2)$$

B. Trajectory control algorithm

The critical element of the quadcopter control system (as an unstable system) is the stabilizing PID controller, which is a component capable of fixing the course of the control deviation, its cumulative value, and change over time. The PID controller calculates the "error" value as the difference between the measured and desired values. The deviation value is multiplied by the P factor (proportional term), the accumulated value is multiplied by the I factor (integral term), and the difference is multiplied by the D factor (differential term). The sum of the individual products gives us the amount of regulation [11]:

$$u_t = P \cdot E_t + I \cdot \sum_{i=0}^t E_i + D \cdot (E_t - E_{t-1}) \quad (10)$$

where u - amount of regulation, P - proportional coefficient, E - deviation size, I - integral coefficient, D - derivative coefficient.

The projections of the horizontal force, without resistance of the non-load-bearing part of the device, according to the dynamics equations have the form below:

$$U_x = P(-\cos \gamma \cos \psi \sin \nu + \sin \gamma \sin \psi) \quad (11.1)$$

$$U_z = P(\cos \gamma \sin \psi \sin \nu + \sin \gamma \cos \psi) \quad (11.2)$$

from where it is possible to determine the roll and pitch angles at which the required effects are created with a known total thrust P :

$$\gamma_d = \arcsin \frac{U_{zd} \cos \psi_d + U_{xd} \sin \psi_d}{P} \quad (12.1)$$

$$\vartheta_d = \arccos \frac{U_{zd} \sin \psi_d - U_{xd} \cos \psi_d}{P} \quad (12.2)$$

Controlling influences U_{xd} and U_{zd} , and also $U_{yd} = U_1$ for the height control channel can be obtained by considering the trajectory control subsystem as a control system that tracks the required coordinates of the center of mass, in particular, as the output signals of the PID controller according to deviations of the center of mass coordinates from the required ones:

$$\begin{aligned} U_{xd} &= K_{px}(x_d - x) + K_{ix} \int (x_d - x) dt + K_{dx}(\dot{x}_d - \dot{x}) \\ U_{yd} &= \frac{m}{\cos \gamma \cos \vartheta} K_{py}(y_d - y) + K_{iy} \int (y_d - y) dt + K_{dy}(\dot{y}_d - \dot{y}) + mg \\ U_{zd} &= K_{pz}(z_d - z) + K_{iz} \int (z_d - z) dt + K_{dz}(\dot{z}_d - \dot{z}) \end{aligned} \quad (13)$$

The values of the coefficients of the regulators for the relevant variables, selected by the Ziegler-Nichols method [12], are shown in Table 1.

TABLE I. VALUES OF THE PID CONTROLLER COEFFICIENTS

	K_p	K_i	K_d
x	8.5	0.02	8.1
y	24	10	28
z	4.9	0.015	5.1

IV. EXPERIMENTAL RESULTS

In this section, we show the effectiveness of the proposed UAV trajectory tracking control for various trajectory curvatures. We tested the control algorithm, which stabilizes the flight coordinates of a quadcopter and tracks a given trajectory. The simulation for an arbitrary curvature path to show that the proposed design provides satisfactory performance under multiple scenarios is shown in Fig. 4. The quadcopter flew at a fixed altitude of 30 meters and pass through ten waypoints, where at the top points the UAV stops to take some photos. And Fig. 5 shows changes in x , y and z coordinates.

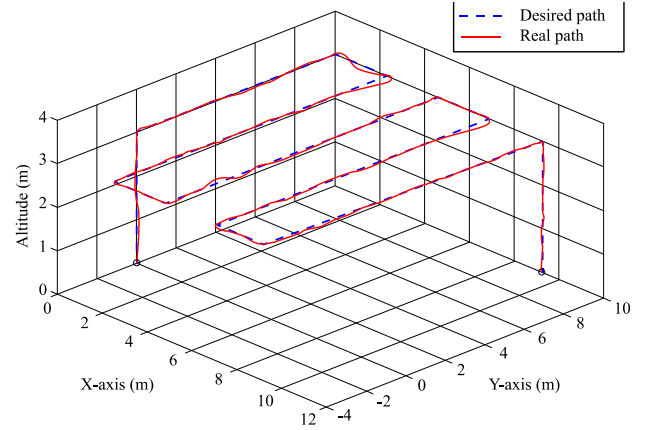


Fig. 4 Trajectory tracking test when the quadcopter follows the path in presence of wind and manual disturbances

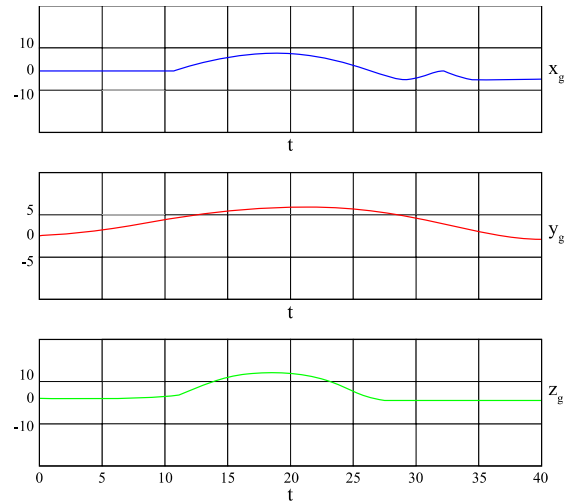


Fig. 5 Coordinate changes (x_g , y_g , z_g) of the PID controller

V. DISCUSSION

During experimental studies on a quadcopter, the following effect was noticed. When adjusting the parameters of the stabilizing PID controller directly during the flight according to the above algorithm based on telemetry data, the most difficult process is to adjust the differential component of the controller, a slight change in this parameter leads to overshoot. At the same time, dosed overshoot, which does not lead to loss of control and is poorly distinguishable visually, has a positive effect on the stability of the quadcopter to external influences. The stabilization of the quadcopter occurs

much faster, and with an increase in the payload of the apparatus, the growth of the component D generally has a positive effect on the actual controllability of the quadcopter.

CONCLUSION

This article proposes a robust controller scheme for path tracking of a quadcopter for automatic monitoring of a target area. The control structure is based on a simple non-linear controller with sub-controllers. The control scheme has been verified in flight tests to track the trajectory, considering external disturbances. The simulation results demonstrate the proposed method's good performance and show the algorithm's efficiency in its implementation.

REFERENCES

- [1] H. N. Lashari, H. Mansoor Ali, S.-U.-R. Massan, "Applications of unmanned aerial vehicles: a review", 3C Technology. Special Edition, 2019. doi: 10.17993/3ctecno.2019.specialissue3.85-105.
- [2] D. Kaliyev, O. Shvets. "Application of robotic systems for fire exploration". AIS 2019: 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University. Székesfehérvár, pp. 91-92.
- [3] D. Kaliyev, O. Shvets, G. Györök. "UAV based navigation system for forest fires monitoring". AIS 2021: 16th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2021 by Óbuda University, Székesfehérvár, 2021, pp. 47-50.
- [4] I. D. Cowling, O. A. Yakimenko, J. F. Whidborne, A. K. Cooke. "A prototype of an autonomous controller for a quadrotor UAV," 2007 European Control Conference (ECC), 2007, pp. 4001-4008, doi: 10.23919/ECC.2007.7068316.
- [5] H. Mo, G. Farid. "Nonlinear and adaptive intelligent control techniques for quadrotor uav—a survey". Asian Journal of Control, 2019. Vol. 21, №. 2, pp. 989-1008.
- [6] S. H. Derrouaoui, Y. Bouzid, M. Guiatni. "Nonlinear robust control of a new reconfigurable unmanned aerial vehicle". Robotics, 2021, Vol 10. № 2, p. 76.
- [7] P. Ru, K. Subbarao. "Nonlinear Model Predictive Control for Unmanned Aerial Vehicles," MDPI aerospace, 2017, pp. 1–26. doi: 10.3390/aerospace4020031.
- [8] V. H. Andaluz, H. Victor, Gallardo, M. Cristian, Chicaiza, A. Fernando, Carvajal, P. Christian. "Robot nonlinear control for Unmanned Aerial Vehicles' multitasking". Assembly Automation. 2018. doi: 10.1108/AA-02-2018-036.
- [9] Y. Al Younes, M. Barczyk. "Nonlinear model predictive horizon for optimal trajectory generation". Robotics, 2021. Vol. 10, № 3, p. 90.
- [10] P. Castillo, L. E. Munoz, P. Garcia. "Indoor Navigation Strategies for Aerial Autonomous Systems". Elsevier, Feb. 2017 ISBN:9780128051894.
- [11] A. A. Najm, I. K. Ibraheem. "Nonlinear PID controller design for a 6-DOF UAV quadrotor system". Engineering Science and Technology, an International Journal, 2019, Vol. 22, № 4, pp. 1087-1097.
- [12] V. V. Patel. Ziegler-Nichols tuning method. Resonance, 2020, Vol. 25, № 10, pp. 1385-1397.

Solving Olympiad Problems Via Lagrange Multipliers

Borbély József
 Óbudai Egyetem, Székesfehérvár, Hungary
 borbely.jozsef@uni-obuda.hu

Abstract— There are a lot of hard problems concerning inequalities in mathematical olympiads which have an elegant elementary solution. In this paper we will do the contrary of the traditional approach: we will try to give a method which solves a large number of these problems using the same basic idea (but the solution will not always be less complicated).

This is the basic idea of Lagrange multiplier method, which can be combined with Weierstrass theorem, if S is a closed and bounded set.

In the following section we demonstrate this method by proving olympiad level inequalities.

I. INTRODUCTION

In this paper we will prove four olympiad level inequalities using Lagrange multipliers. These solutions were not published anywhere else, and use the same basic idea.

II. ABOUT LAGRANGE MULTIPLIERS

We will need the following theorem of Weierstrass (see [2]):

Theorem (Weierstrass): Let n be a positive integer, and let D be a closed, bounded subset of the real n -dimensional space. If the function $f: D \rightarrow \mathbb{R}$ is continuous, then it has a minimum and a maximum.

A closed set is a set which contains all its limit points, which means that every convergent sequence of points in a closed set converges to a point of this given set.

Let us consider the following problem: let us take a real valued function g which is defined on a subset S of the real n -dimensional space. Furthermore let us assume that the functions g_1, g_2, \dots, g_m all vanish on the set S (these are functions with n real variables as well). Using the conditions concerning the g_i 's we want to determine the extreme values of the function g .

Here the so called method of Lagrange multipliers comes in handy: if these functions are differentiable by all of the n variables, then we define the modified function

$$G = g + \sum_{i=1}^m \lambda_i \cdot g_i.$$

Here the λ_i 's are called the Lagrange multipliers which will be calculated later.

But why even this function? Because this function G is identical to g on the set S , and in this function we encoded our conditions in some way.

We will determine the partial derivatives of G , which must equal zero at any extremal points. By solving this system of equalities we will obtain a set of potential extremal points.

III. OLYMPIAD LEVEL INEQUALITIES

We begin with an inequality which was set in the Hungarian journal KÖMAL in 2007 (problem A.433., see [5]):

Inequality 1: If the sum of the squares of the real numbers

a, b, c equals 1 then

$$a+b+c \leq 2abc + \sqrt{2}.$$

Proof:

According to the condition we have

$a^2 + b^2 + c^2 = 1$. The real triples of numbers (a, b, c) satisfying

this condition define a closed, bounded set, so by Weierstrass theorem the continuous function

$$f(a, b, c) = a + b + c - 2abc$$

must have a maximal and a minimal value on this set. Let us introduce the function

$F(a, b, c) = a + b + c - 2abc + \lambda \cdot (a^2 + b^2 + c^2 - 1)$, where λ is the Lagrange multiplier.

The partial derivatives of $F(a, b, c)$ are the following functions:

$$F_a' = 1 - 2bc + 2\lambda a$$

$$F_b' = 1 - 2ac + 2\lambda b$$

$$F_c' = 1 - 2ab + 2\lambda c$$

Thus one has to solve the system of equations

$$0 = 1 - 2bc + 2\lambda a$$

$$0 = 1 - 2ac + 2\lambda b$$

$$0 = 1 - 2ab + 2\lambda c$$

in order to determine the potential extreme points.

By subtracting the second equation from the first one we get

$$0=(a-b) \cdot (2c+2\lambda). \text{ Similarly we get}$$

$$0=(b-c) \cdot (2a+2\lambda) \text{ and}$$

$$0=(c-a) \cdot (2b+2\lambda).$$

If there are two different ones among the numbers a,b,c, then at least two of the differences (a-b), (b-c), (c-a) is nonzero. We can assume that a-b≠0 and b-c≠0. This implies that 2c+2λ=2a+2λ=0 which means that a=c=-λ. By substitution we get the system of equations

$$0 = 1-2bc+2\lambda a=1+2\lambda b-2\lambda^2,$$

$$2\lambda^2+b^2=1.$$

$$\text{So } 2\lambda b = 2\lambda^2 - 1 \text{ and}$$

$$4\lambda^2=4\lambda^2 \cdot (2\lambda^2 + b^2)=8\lambda^4+4\lambda^2 b^2=8\lambda^4+(2\lambda^2 - 1)^2, \text{ thus we get}$$

$$8\lambda^4+(2\lambda^2 - 1)^2-4\lambda^2 = 0, \text{ which is equivalently}$$

$$12\lambda^4-8\lambda^2+1=0. \text{ We get the solutions}$$

$$\lambda = \sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{6}}, -\sqrt{\frac{1}{6}}.$$

So if there are two different ones among the numbers a, b, c then the potential extreme points will be

$$\left(\sqrt{\frac{1}{2}}, \sqrt{\frac{1}{2}}, 0\right), \left(-\sqrt{\frac{1}{2}}, -\sqrt{\frac{1}{2}}, 0\right), \left(\sqrt{\frac{1}{6}}, \sqrt{\frac{1}{6}}, \sqrt{\frac{4}{6}}\right), \left(-\sqrt{\frac{1}{6}}, -\sqrt{\frac{1}{6}}, \sqrt{\frac{4}{6}}\right),$$

$\left(-\sqrt{\frac{1}{6}}, -\sqrt{\frac{1}{6}}, -\sqrt{\frac{4}{6}}\right)$, and the points we get by permutation of the coordinates of these vectors. If the coordinates are the same, then the possible extreme points are

$$(\sqrt{3}, \sqrt{3}, \sqrt{3}), (-\sqrt{3}, -\sqrt{3}, -\sqrt{3}).$$

By substitution we can check that the maximal value of the expression f(a,b,c)=a+b+c-2abc with respect to the given condition equals $\sqrt{2}$.

The following inequality is a shortlisted problem for the International Mathematical Olympiad from the year 2010. This problem was posed at the Surányi János Emlékverseny in 2011 (see [3]):

Inequality 2: If a,b,c,d are real numbers satisfying the relations a+b+c+d=6 and a²+b²+c²+d²=12 then

$$36 \leq 4(a^3 + b^3 + c^3 + d^3) - (a^4 + b^4 + c^4 + d^4) \leq 48.$$

Proof:

The vectors (a, b, c, d) satisfying the two given conditions clearly determine a closed bounded set, thus by the theorem of Weierstrass the expression

$$4(a^3 + b^3 + c^3 + d^3) - (a^4 + b^4 + c^4 + d^4)$$

has a minimal and a maximal value with respect to the given conditions.

Let us define F(a,b,c,d) in the following way:

$$F(a,b,c,d) = 4(a^3 + b^3 + c^3 + d^3) - (a^4 + b^4 + c^4 + d^4) + \lambda_1(a + b + c + d - 6) + \lambda_2(a^2 + b^2 + c^2 + d^2 - 12),$$

where λ_1 and λ_2 are the Lagrange multipliers.

The partial derivatives of F(a,b,c,d) are the following functions:

$$F_a' = -4a^3 + 12a^2 + 2\lambda_2 a + \lambda_1$$

$$F_b' = -4b^3 + 12b^2 + 2\lambda_2 b + \lambda_1$$

$$F_c' = -4c^3 + 12c^2 + 2\lambda_2 c + \lambda_1$$

$$F_d' = -4d^3 + 12d^2 + 2\lambda_2 d + \lambda_1$$

In order to determine the potential extremal points we have to solve the system of equalities

$$F_a' = F_b' = F_c' = F_d' = 0.$$

First we will prove that for these solutions at most two of the numbers a, b, c, d can be different. We prove this statement indirectly. Let us assume to the contrary that there are three different ones among the numbers a, b, c, d satisfying the system of equalities. Without loss of generalization we can assume that a, b, c are different numbers.

$$0 = F_a' - F_b' = 2(b - a)(2a^2 + 2ab + 2b^2 - 6a - 6b - \lambda_2).$$

Due to the assumption a≠b we have

$$2a^2 + 2ab + 2b^2 - 6a - 6b - \lambda_2 = 0.$$

Similarly we get

$$2b^2 + 2bc + 2c^2 - 6b - 6c - \lambda_2 = 0.$$

By taking the difference of these equations we get

$$0 = (2a^2 + 2ab + 2b^2 - 6a - 6b - \lambda_2) -$$

$$-(2b^2 + 2bc + 2c^2 - 6b - 6c - \lambda_2) =$$

$$= 2(a - c)(a + b + c - 3).$$

Due to the assumption a≠c we have

$$a + b + c = 3, \text{ which implies } d = 3.$$

In this case a²+b²+c²=3.

It is easy to see that for real numbers a, b, c we have

$(a - b)^2 + (b - c)^2 + (c - a)^2 \geq 0$, and we have equality if

and only if $a=b=c$. This inequality is equivalent to the following one:

$$(a + b + c)^2 \leq 3(a^2 + b^2 + c^2).$$

Each side of this inequality equals 9, which means the only solution for real numbers in the case $d=3$ is $a=b=c=1$. But this contradicts the assumption that there are three different ones among the numbers a, b, c, d.

Thus we know that among the numbers a, b, c, d there are at most two different ones. If $a=b=c=d$ then all of them must equal $3/2$, but in this case the sum of their squares is different from 12.

Thus there are exactly two different ones among the numbers a,b,c,d. If three of them are identical and the fourth one is different, then we get the point (0, 2, 2, 2), (3, 1, 1, 1), and the points we get by permutation of the coordinates of these vectors.

If the two groups consist of two-two identical numbers then the system of equation does not have a solution in real numbers.

By substitution we can check that the minimal value is 36, the maximal value is 48.

Inequality 3 was a problem at the International Mathematical Olympiad in 1984:

Inequality 3: *If the sum of the nonnegative numbers x, y, z equals 1 then*

$$0 \leq xy+yz+zx-2xyz \leq \frac{7}{27}.$$

Proof:

For some real numbers a, b, c we have

$$x=a^2, y=b^2, z=c^2 \text{ and } a^2+b^2+c^2=1.$$

Clearly the vectors (a, b, c) in question determine a closed, bounded set, so by Weierstrass theorem the function

$f(a,b,c)=a^2b^2+b^2c^2+c^2a^2-2a^2b^2c^2$ has a minimal and a maximal value with respect to the condition $a^2+b^2+c^2=1$.

Let us introduce the function

$$F(a,b,c)=a^2b^2+b^2c^2+c^2a^2-2a^2b^2c^2+\lambda \cdot (a^2 + b^2 + c^2 - 1),$$

where λ is the Lagrange multiplier.

The partial derivatives of $F(a,b,c)$ are the following functions:

$$F_a' = 2a \cdot (b^2 + c^2 - 2b^2c^2 + \lambda)$$

$$F_b' = 2b \cdot (a^2 + c^2 - 2a^2c^2 + \lambda)$$

$$F_c' = 2c \cdot (b^2 + a^2 - 2b^2a^2 + \lambda)$$

Thus one has to solve the system of equations

$$0 = 2a \cdot (b^2 + c^2 - 2b^2c^2 + \lambda)$$

$$0 = 2b \cdot (a^2 + c^2 - 2a^2c^2 + \lambda)$$

$$0 = 2c \cdot (b^2 + a^2 - 2b^2a^2 + \lambda)$$

If $abc \neq 0$ then

$$-\lambda = b^2 + c^2 - 2b^2c^2 = a^2 + c^2 - 2a^2c^2, \text{ so}$$

$$b^2 \cdot a^2 + 2c^2(a^2 - b^2) = (a^2 - b^2)(2c^2 - 1) = 0.$$

In the case $abc \neq 0$ similarly we get the equations

$$(b^2 - c^2)(2a^2 - 1) = 0 \text{ and}$$

$$(c^2 - a^2)(2b^2 - 1) = 0.$$

If at least two of the numbers $(2a^2 - 1)$, $(2b^2 - 1)$, $(2c^2 - 1)$ equal zero then one of the numbers a, b, c must equal zero which contradicts our assumption $abc \neq 0$. Thus at least two of the numbers $(a^2 - b^2)$, $(b^2 - c^2)$, $(c^2 - a^2)$ must equal zero, which means that $a^2=b^2=c^2$.

Thus the three dimensional points having the coordinates $\sqrt{3}$ or $(-\sqrt{3})$ are potential extreme points. By substitution we can check that the desired inequalities hold.

In the case $abc=0$ one of the numbers a, b, c must equal 0. Let us assume that $c=0$. In this case $a^2+b^2=1$ and so

$$xy+yz+zx-2xyz = a^2b^2 = a^2(1 - a^2).$$

Of course a^2b^2 is nonnegative and $a^2(1 - a^2)$ cannot be greater than $\frac{1}{4}$ which is less than $\frac{7}{27}$.

Inequality 4 is an own problem of the author of this paper which was set at the Nemzetközi Magyar Matematikaverseny in 2010 (class 11). Here we give a solution with Lagrange multipliers (which is more complicated than the original elementary solution):

Inequality 4: *Let n be a positive integer number, and let c be a positive constant. If the nonnegative numbers x_1, x_2, \dots, x_n satisfy the condition*

$$\sum_{i=1}^n (x_i^2 + ix_i) = c, \text{ then the inequalities}$$

$$\sqrt{c + \frac{n^2}{4}} - \frac{n}{2} \leq \sum_{i=1}^n x_i \leq \sqrt{n} \cdot \sqrt{c + \sum_{j=1}^n \frac{j^2}{4}} - \sum_{j=1}^n \frac{j}{2}.$$

hold.

Proof:

For some real numbers y_1, y_2, \dots, y_n we have $y_i^2 = x_i$ for every $1 \leq i \leq n$. Thus we have to give the minimal and maximal value of $\sum_{i=1}^n y_i^2$ with respect to the condition

$$c = \sum_{i=1}^n (y_i^4 + iy_i^2).$$

The existence of the extreme values is guaranteed by the theorem of Weierstrass.

Using the method of Lagrange multipliers we define the function

$$F(y_1, y_2, \dots, y_n) = \sum_{i=1}^n y_i^2 + \lambda \cdot (\sum_{i=1}^n (y_i^4 + iy_i^2) - c),$$

where λ is the Lagrange multiplier.

Taking the derivative of F by the i-th variable we get

$$4\lambda y_i^3 + (2\lambda i + 2)y_i = 2y_i(2\lambda y_i^2 + \lambda i + 1).$$

In order to determine the potential extremal points we have to solve the system of equations

$$0 = 2y_i(2\lambda y_i^2 + \lambda i + 1), \text{ where } i=1, 2, \dots, n.$$

For every possible index i we have either $y_i=0$ or

$$2\lambda y_i^2 + \lambda i + 1 = 0.$$

Thus if for some index i the number y_i is different from zero then we have

$$2\lambda y_i^2 + \lambda i + 1 = 0, \text{ which is equivalent to the equation}$$

$$-1 = \lambda(2y_i^2 + i).$$

Here λ must be different from zero, thus

$$y_i^2 + \frac{i}{2} = \frac{-1}{2\lambda}, \text{ and } y_i^4 + iy_i^2 + \frac{i^2}{4} = \frac{1}{4\lambda^2}.$$

Let $1 \leq i_1 < i_2 < \dots < i_t \leq n$ be the indices of the numbers in the set $\{y_1, y_2, \dots, y_n\}$ which are different from zero.

Using this notation we have

$$\sum_{j=1}^t \left(y_{i_j}^2 + \frac{i_j}{2} \right) = \frac{-t}{2\lambda},$$

which is equivalent to the equation

$$\sum_{j=1}^t y_{i_j}^2 = \frac{-t}{2\lambda} - \sum_{j=1}^t \frac{i_j}{2}.$$

Similarly we get

$$\begin{aligned} c &= \sum_{i=1}^n (y_i^4 + iy_i^2) = \sum_{j=1}^t (y_{i_j}^4 + i_j y_{i_j}^2) = \\ &= t \cdot \frac{1}{4\lambda^2} - \sum_{j=1}^t \frac{i_j^2}{4}. \end{aligned}$$

Thus $c + \sum_{j=1}^t \frac{i_j^2}{4} = t \cdot \frac{1}{4\lambda^2}$, and

$$tc + t \sum_{j=1}^t \frac{i_j^2}{4} = \frac{t^2}{4\lambda^2}.$$

The square root of $\frac{t^2}{4\lambda^2}$ is $\frac{-t}{2\lambda}$, because λ must be positive if the

numbers y_1, y_2, \dots, y_n are real (namely in the equation

$$y_i^2 + \frac{i}{2} = \frac{-1}{2\lambda} \text{ the left hand side is positive}).$$

Thus we have

$$\sqrt{t} \cdot \sqrt{c + \sum_{j=1}^t \frac{i_j^2}{4} - \sum_{j=1}^t \frac{i_j}{2}} = \sum_{j=1}^t y_{i_j}^2 = \sum_{i=1}^n y_i^2.$$

We have to determine the minimal and maximal value of this expression.

We will prove that for every positive integer k the inequality

$$\begin{aligned} &\sqrt{t} \cdot \sqrt{c + \sum_{j=1}^t \frac{i_j^2}{4} - \sum_{j=1}^t \frac{i_j}{2}} < \\ &< \sqrt{t+1} \cdot \sqrt{c + \frac{k^2}{4} + \sum_{j=1}^t \frac{i_j^2}{4} - \sum_{j=1}^t \frac{i_j}{2} - \frac{k}{2}} \text{ holds.} \end{aligned}$$

This inequality is equivalent to the following one:

$$\begin{aligned} &\sqrt{tc + t \sum_{j=1}^t \frac{i_j^2}{4} - \sum_{j=1}^t \frac{i_j}{2}} < \\ &< \sqrt{tc + \frac{tk^2}{4} + t \sum_{j=1}^t \frac{i_j^2}{4} + c + \frac{k^2}{4} + \sum_{j=1}^t \frac{i_j^2}{4} - \sum_{j=1}^t \frac{i_j}{2} - \frac{k}{2}}, \end{aligned}$$

which is equivalently

$$\begin{aligned} &\sqrt{tc + t \sum_{j=1}^t \frac{i_j^2}{4} + \frac{k}{2}} < \\ &< \sqrt{tc + \frac{tk^2}{4} + t \sum_{j=1}^t \frac{i_j^2}{4} + c + \frac{k^2}{4} + \sum_{j=1}^t \frac{i_j^2}{4}}. \end{aligned}$$

By taking the squares on both sides we have to prove that

$$\begin{aligned} &tc + t \sum_{j=1}^t \frac{i_j^2}{4} + \frac{k^2}{4} + k \sqrt{tc + t \sum_{j=1}^t \frac{i_j^2}{4}} < \\ &< tc + \frac{tk^2}{4} + t \sum_{j=1}^t \frac{i_j^2}{4} + c + \frac{k^2}{4} + \sum_{j=1}^t \frac{i_j^2}{4}, \text{ which inequality} \end{aligned}$$

is equivalent to the following one:

$$0 < c + \sum_{j=1}^t \frac{i_j^2}{4} - k\sqrt{t} \sqrt{c + \sum_{j=1}^t \frac{i_j^2}{4} + \frac{tk^2}{4}} + c$$

On the right hand side we have

$$\left(c + \sum_{j=1}^t \frac{i_j^2}{4} - \frac{k\sqrt{t}}{2} \right)^2 + c, \text{ which is clearly positive, because } c \text{ is a positive number. Thus we proved that}$$

$$\begin{aligned} & \sqrt{t} \cdot \sqrt{c + \sum_{j=1}^t \frac{i_j^2}{4} - \sum_{j=1}^t \frac{i_j}{2}} < \\ & < \sqrt{t+1} \cdot \sqrt{c + \frac{k^2}{4} + \sum_{j=1}^t \frac{i_j^2}{4} - \sum_{j=1}^t \frac{i_j}{2} - \frac{k}{2}}. \end{aligned}$$

This (using induction) implies that $\sum_{i=1}^n y_i^2$ is maximal if all the numbers y_1, y_2, \dots, y_n are nonzero. Solving the equalities we get that the maximal value is

$$\sqrt{n} \cdot \sqrt{c + \sum_{j=1}^n \frac{j^2}{4} - \sum_{j=1}^n \frac{j}{2}}.$$

Similarly, the minimal value will be attained if exactly one of the numbers y_1, y_2, \dots, y_n is nonzero. It is easy to check that in this case $y_n \neq 0, y_1 = y_2 = \dots = y_{n-1} = 0$, and the minimal possible value is

$$\sqrt{c + \frac{n^2}{4} - \frac{n}{2}}.$$

REFERENCES

- [1] Dušan Djukić, Vladimir Janković, Ivan Matić, Nikola Petrović, The IMO Compendium: A Collection of Problems Suggested for The International Mathematical Olympiads: 1959-2009 Second Edition, Problem Books in Mathematics, 2nd ed., 2011
- [2] Leindler László: Analízis, Polygon, 2012

Online sources:

- [3] <https://www.imo-official.org/problems/IMO2010SL.pdf>
- [4] <https://nmmv.berzsenyi.hu/feladatok/2010>
- [5] <https://www.komal.hu/feladat?a=feladat&f=A433&l=hu>

Board Game Recommendation System with Collaborative Filtering

Vanessza Tőke
Alba Regia Technical Faculty
Óbuda University
Székesfehérvár, Hungary
tvanessza@stud.uni-obuda.hu

Rozália Lakner
Alba Regia Technical Faculty
Óbuda University
Székesfehérvár, Hungary
lakner.rozalia@amk.uni-obuda.hu

Abstract—This paper presents the building of a recommendation system, based on collaborative filtering, using real data about board games and their ratings from users. It reviews the notion of web scraping, data cleaning and recommendation systems, and the used technology for development. Describes the building steps of a python-based application. Shows how to collect data using web scraping technique. Also mentions how to prepare, analyze, and visualize the gathered data with popular python libraries.

Keywords—python, recommendation system, collaborative filtering, web scraping, data cleaning

I. INTRODUCTION

Nowadays, playing board games is popular than ever. The board games already existed a millennium ago, used in the ancient time, but their main purpose was more strategical, than fun. In the last few years, the number of game releases has increased, so from year-to-year it is harder to decide which games should we purchase, which ones are fitting best to our appetite. In Fig. 1. the number of released games can be seen in an annual breakdown, for which the data about boardgames collected from BoardGameGeeks.com [1].

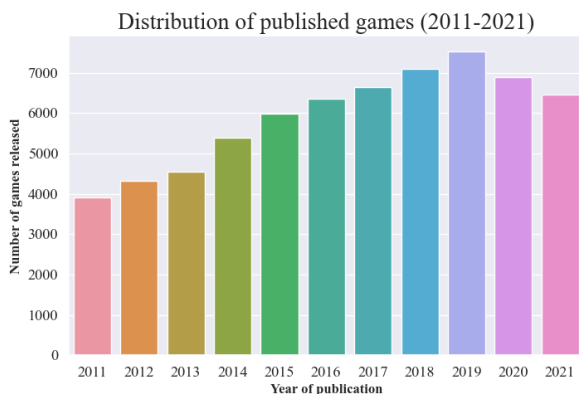


Fig. 1. Distribution of published board games in last 10 years

Thanks to the proliferation of high-performance computers and the developing information technology, the Big Data projects - working with huge amount of data, with the shortest possible processing time -, thus, the recommendation systems spread and used in a more widely scale.

For building recommendation systems, the most popular programming language is Python, and for making the development phase much easier, many python-based libraries are available to develop with more easily and faster. This article describes the process of building board game recommendation system, guiding through the basic concepts involved, using popular recommendation technique, the collaborative filtering and through the earlier steps like collecting and preparing the data with popular data science libraries.

II. RECOMMENDATION SYSTEM

Recommendation systems have become a defining part of today's world. Many companies, like Netflix, Amazon, and Facebook use recommendation systems to get more customers to buy their products/services, or to made them to purchase something directed contents and advertisements.

According to the definition recommendation systems' main function is to suggest items for users according to their preferences, knowing their history for the recommendation [2]. These suggestions can be used in any kind of decision-making processes. The word *item* is a general term used to express what the system recommends to the users.

A. Types of recommending systems

For creating powerful and accurate systems we need as much data as possible, and preferably containing ratings from existing users and properties about the recommended item(s). That which kind of data will be more important from a recommendation point of view will be dependent on the type of used recommendation technique. The recommendation systems have different types, according to what preferences are mainly used for the recommendations. The mainly spread types are the collaborative filtering, the content-based filtering, and the hybrid processing.

1) Collaborative filtering

For collaborative filtering process the data about the users' rating habits are needed. If the data is enough for discover correlation between the similar users or the similar items, then the explored similarity can be used to recommend items to a given user according to what other users liked. The system's suggestion completely based on the ratings [3].

Within collaborative filtering the memory-based and model-based filtering are differentiated. The memory-based filtering has two types, which are be presented in this paper.

2) User-based collaborative filtering:

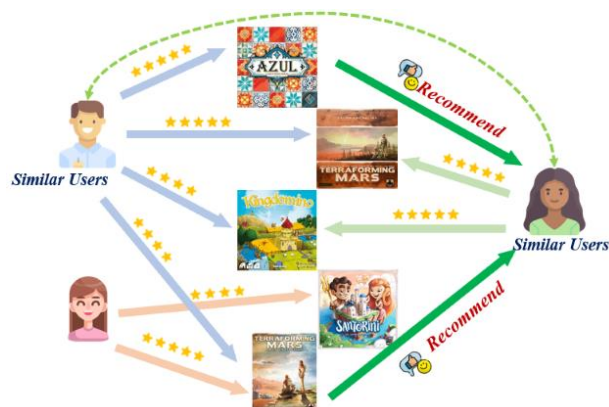


Fig. 2. User-based collaborative filtering method example, based on board game items. The recommendations based on the similar users' ratings.

The user-based collaborative filtering algorithm produces recommendation list for given user according to the view of other users, as it can be seen in Fig. 2. The basic concept is that if the ratings of rated items are similar between users, then it can be said that these users would rate other items similarly as well [4].

3) Item-based collaborative filtering:



Fig. 3. Item-based collaborative filtering method example based on board game items. The recommendation based on similarity between rated games.

The item-based collaborative filtering algorithm based on similarity, just like user-based filtering, but while there the similarity between the users is searched, using ratings, here among the rated items the similarity is researched. If two users like the same games, the system try to find the similar items according to this and then recommend the appropriate items, as shown in Fig. 3.

4) Content-based filtering

Content-based recommendation technique makes its suggestions relies on descriptions and attributes about the items and uses the user's interests and preferences (the users predefined profile) to make the recommendations personal [3].

5) Hybrid techniques

As it is in its name, hybrid recommendation system uses multiple recommendation techniques to create more accurate recommendations and personalized suggestions to the users.

B. Similarity measurement techniques

Talking about recommendation systems, measuring the similarity is a key point in the systems. Even if we talk about collaborative filtering or content-based systems, somehow the similarity should be determined to find the patterns between the users or items.

For measuring similarity there are many techniques and equation to define it between items. The two most popular similarity measures used for recommendation systems are the cosine similarity and the Pearson similarity.

1) *Cosine similarity*: Items are corresponding for vectors of an n-dimensional space and their similarity is the cosine of their angle [2].

$$\cos(x, y) = \frac{(x \cdot y)}{\|x\| \|y\|} \quad (1)$$

where \cdot indicates vector dot product and $\|x\|$ is the norm of vector x .

2) *Pearson similarity*: Use the correlation between the items to give the similarity between two objects [2].

$$Pearson(x, y) = \frac{\sum(x, y)}{\sigma_x \times \sigma_y} \quad (2)$$

where the covariance of data points x and y \sum and their standard deviation σ .

III. DEVELOPMENT AND IMPLEMENTATION

For developing the board game recommendation system three main steps were determined. Talking about a Big Data project, the first step is collecting data about board games. After the data is stored to files, the next step is the data cleaning and preparation phase. The last step demonstrate in this article is the building of the collaborative filtering recommendation systems.

Python programming language was used to write the code, and a few useful libraries to the given steps. These libraries will be discussed and introduced later in the relevant chapters. For backup and version controlling Git and GitHub were used to commit the changes in the program and store the code in GitHub. For writing the relevant code the Visual Studio Code open-source code editor was used.

A. Collect the data

As it can be seen in the previous chapter creating a project under the scope of Big Data need a huge amount of data. However, owning the required quantity and quality of data is not always an easy task.

Data can be collected from the following places:

- from databases of different systems (e.g., database of enterprise software)
- from survey forms filled by users
- from Internet, using a script to scrape interesting data from a given website – web scraping technique

In the present work the web scraping technique was used. For this purpose, the playwright and the beautiful soup python-based libraries were used [5].

Web scraping is a method to collect data without any connection with an API or without a user's interaction with the collectible data. The basis of web scraping is creating an automated program which will ask the web server, using queries to the needed data.

The collection of the data was scraping goes 100% from BoardGameGeeks.com [1].

Three types of datasets were created from the gathered values:

1) *Board games general data*: Gathered data as a table format from collection pages, where all the games – existing in the website's database – are displayed. Here the most important attributes were like game's title, the ranking order, the general informations about ratings (number of ratings and the average ratings) and the unique link to the game which later will be the identifier of games in all datasets, as can be seen in Fig. 4.

link_to_game	rank	title	avg_rating	number_voters
/boardgame/174430/gloomhaven	1.0	Gloomhaven	8.71	51212.0
/boardgame/161936/pandemic-legacy-season-1	2.0	Pandemic Legacy: Season 1	8.58	47092.0
/boardgame/224517/brass-birmingham	3.0	Brass: Birmingham	8.66	29491.0
/boardgame/291457/gloomhaven-jaws-of-the-lion	4.0	Gloomhaven: Jaws of the Lion	8.61	21080.0
/boardgame/167791/terraforming-mars	5.0	Terraforming Mars	8.40	80129.0

Fig. 4. Part of a general dataset about board games' main properties.

2) *Board games detailed data*: Using the unique identifier link for the games, it is possible to reach more information about board games to reference this. For collaborative filtering this dataset will not be used in the later steps.

3) *Ratings data*: The user's ratings to the games can be gathered also with the help of the game's unique identifier, which is a part of the whole hyperlink leads to the board games. For collecting ratings a simple schema was built for the file, where the data is stored. As can be seen in Fig. 5, the schema contains the user's username in the website, the rating which they gave to the particular games and the identifier to the games.

	user	rating	game_id
0	annecoleens	8.0	/boardgame/174430/gloomhaven
1	gvaldizan	8.0	/boardgame/174430/gloomhaven
2	Patryk_Purczynski	8.0	/boardgame/174430/gloomhaven
3	hankAA	8.7	/boardgame/174430/gloomhaven
4	Crash G	6.0	/boardgame/174430/gloomhaven

Fig. 5. Ratings dataset's structure, and the collected properties.

Playwright library was used in web scraping process for requesting HTML pages from web server and for browser automation. The other important extension library is BeautifulSoup, this one was responsible for detecting the HTML tags storing the data and scraping this information into python data structures. For collecting data from the website another library Pandas was used as well for piling up data, stored in table format on the site. Finally, the gathered data was saved into CSV files.

B. Cleaning the data

After the data is collected, it is obvious that for using them for the main purposes, some transformation and preparation is essential before start operating with them.

Data cleaning is an integral part of data preprocessing. Data preprocessing is a well-defined process which give back from given data the same- or less amount of data. The essence of data cleansing to detect and remove the errors and inconsistencies occur in data, in the interest of increasing the quality of data [6].

In this step knowledge about our own data and make some analysis on them is important. The better the data is prepared; the more accurate results can be expected from the system in the future.

In this project the used libraries for the data preprocessing were the Pandas and Seaborn python-based libraries. The transformation and cleansing steps were made with Pandas, and the Seaborn was reliable for making charts and plots for further analysis.

Preparing boardgame and ratings datasets for recommendation system, the below steps were taken:

- remove unimportant attributes from datasets
- rename the attributes to bring them to the same naming convention
- detect and drop duplicated rows
- set the index columns for the data frames
- transform inconsistent data to missing values or drop them
- delete board games which would not be recommended by the recommendation system (remove board game extensions from dataset)
- drop rows from ratings dataset where the username or the rating are missing (in that case the row is irrelevant from the point of view of recommendation system's purpose)
- remove the whitespaces (trim) from all values and set the data types to the given columns

Analysis mainly created for ratings dataset to determine some interesting facts on the data, as illustrated in Fig. 6. and Fig. 7.

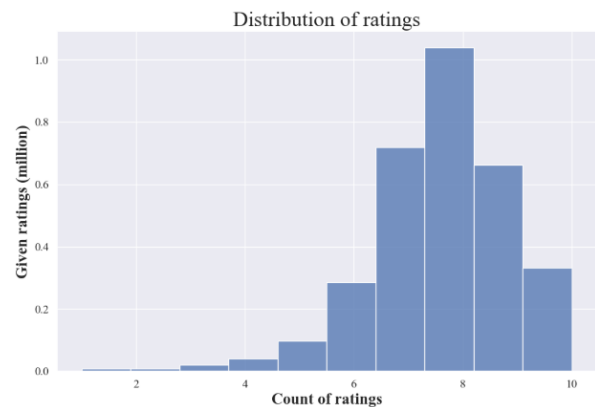


Fig. 6. The distribution of ratings. The most given ratings to the games was between the range of [6.5;9] intervals.

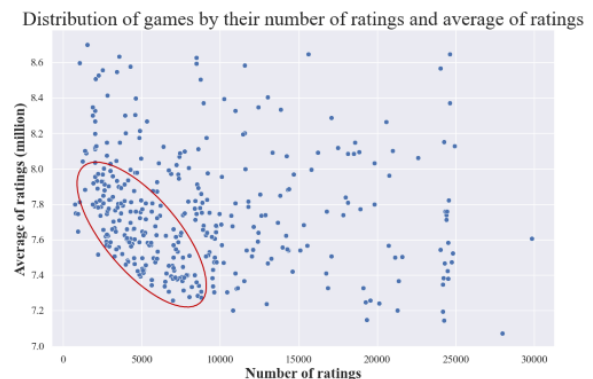


Fig. 7. Distribution of games by their number of ratings and average of ratings. Within the red oval there are the most games appearing.

User-item matrices were generated from ratings dataset to the user-based and item-based recommendation systems as well. Because of the size of the ratings data is so enormous, some processes were executed on the created matrices.

The following attempts were fulfilled to make the appropriate matrices:

- drop users whose number of ratings are below a given threshold number
- drop ratings which are below a given rating threshold, it was determined by using the scale of the ratings and chose an appropriate value which above the user rather likes the game than not
- reduce ratings per game up to a threshold value

The above transformations are required for optimizing the dataset to be able to process on the computer with lower capacity, used for developing this project. At the end of the preparation and transformation of the data for creating, the dataset is counting 370 games with 24570 users and their ratings.

C. Build the recommendation systems

For creating a system, resulting satisfying results, more recommendation engines need to be developed with different types of parameterizations and implemented logic, to determine which variations are the best.

The Pandas, Scikit-learn and the SciPy python libraries were used for creating the user-item matrices, to use and execute the Pearson similarity and cosine similarity.

Developing the following recommendation engines, the logic from [7] and [8] were implemented, optimized, and customized to board game ratings dataset.

The general steps in user-based and item-based recommendation algorithms are slightly differ from each other, as can be seen below.

General Steps of user-based recommendation algorithm:

1. Find similar users based on how they interacted with the common items.
2. Find the items which are rated high by both the similar users and the user to whom the recommendation is made. Filtering the cases the rated movies are same between both sides.
3. Calculate the weighted average score for all of the items.
4. Calculate the rank of the items, using the score of the games. Finally pick the top n recommendations [7].

General Steps of item-based recommendation algorithm:

1. Using the user ratings dataset calculate the item similarity scores.
2. Find the top n items which have the highest similarities for the rated and liked items by the user.
3. Calculate the weighted average score for the most similar items by the users.
4. Rank items based on the scores and pick top n items to recommend [8].

For the above algorithms the data needs to be standardized to bring the values to the same interval. The created engines for identifying similarity used the cosine similarity or the Pearson similarity. In the case of cosine similarity, the missing values were replaced with 0 values. The matrices new values will be numbers between [-1;1] intervals, where the closer the number to one the bigger the similarity is between elements. Fig. 8. shows an example for similarity matrix.

user	-mIDE-	...Hammer	0010	0492372665
-mIDE-	1.000000	0.866025	NaN	-0.612372
...Hammer	0.866025	1.000000	0.000000	0.394719
0010	NaN	0.000000	1.000000	NaN
0492372665	-0.612372	0.394719	NaN	1.000000

Fig. 8. User similarity matrix with Pearson correlation for the first 4 users.

After this process, in the case of user-based collaborative filtering, the test user's ratings were removed from the similarity matrix's index. Setting the threshold to filter out the positive similarity between the users, then sort the similarity values from highest to lowest. Next step to create a matrix already does not contain the boardgames rated by the test user and only contains items rated by the similar users. The following process is the recommendation, where the recommended items are determined by the weighted average of user similarity score and the boardgame rating.

The final similarity weighted rank for the boardgames was calculated by (3):

$$game\ rank = \frac{\sum user\ similarity \times game\ rating}{all\ ratings} \quad (3)$$

where *user similarity* is the similarity measure between the target user and the similar user, *game rating* is the rating given by the similar user and *all ratings* is the number of rated games by all the similar users. Fig. 9. contains the code snippet for implementation of user-based algorithm.

```
ratings_UCF = ratings_UCF \
    .pivot_table(index='user', columns='Title', values='rating')
ratings_UCF = ratings_UCF \
    .subtract(ratings_UCF.mean(axis=1), axis = 'rows')
user_similarity_p = ratings_UCF.T.corr()
user_similarity_p.drop(index=test_user_id, inplace=True)
similar_users = \
    user_similarity_p[user_similarity_p \
        [test_user_id]>similarity_threshold] \
    [test_user_id].sort_values(ascending=False)[:topN]
rated_games_by_test_user = \
    ratings_UCF[ratings_UCF.index==test_user_id] \
    .dropna(axis=1,how='all')
rated_by_similar_users = \
    ratings_UCF[ratings_UCF.index.isin(similar_users.index)] \
    .dropna(axis=1, how='all')

similar_games = rated_by_similar_users.columns
for game in similar_games:
    movie_rating = rated_by_similar_users[game]
    total_scores = 0
    count_num_of_scores = 0
    for user in similar_users.index:
        if not pd.isna(movie_rating[user]):
            score = similar_users[user] * movie_rating[user]
            total_scores = total_scores + score
            count_num_of_scores = count_num_of_scores + 1
    recommended_item_score[game] = \
        total_scores / count_num_of_scores
```

Fig. 9. Code snippet of the user-based collaborative filtering recommendation algorithm. Used similarity measure is Pearson correlation.

In the case of item-based collaborative filtering the steps of the algorithm are changed in some point. In this algorithm after calculating the similarity between the items, selecting the set of games which were rated by the target user. In the same way, the set of games which the target user did not rated are also defined. The next step is predicting the similarity between the rated games and the not rated games. This correlation between items is calculating by the average of ratings and the similarity scores (how the items are similar) using weights. After that we sort the values from highest to lowest and pick the first top N items, which are the most like the target user's rated games.

Equation (4) used for calculating the average of ranks using similarity scores as weights:

$$\text{game rank} = \frac{\sum \text{rating of game} \times \text{similarity score}}{\text{all ratings}} \quad (4)$$

where *rating of game* is the ratings for boardgame given by target user, *similarity score* is the similarity measure between the rated and not rated games and *all ratings* is the number of rated games by the target user. The relevant code snippet can be found in Fig. 10.

```
ratings_ICF = ratings_ICF \
    .pivot_table(index='Title', columns='user', values='rating')
ratings_ICF = ratings_ICF \
    .subtract(ratings_ICF.mean(axis=1), axis = 0)
item_similarity_c = \
    pd.DataFrame(data=cosine_similarity(ratings_ICF.fillna(0)), \
    index=ratings_ICF.index, columns=ratings_ICF.index)
non_rated_games_by_test_user = \
    pd.DataFrame(ratings_ICF[test_user_id].isna()).reset_index()
non_rated_games_by_test_user = non_rated_games_by_test_user
non_rated_games_by_test_user = \
    non_rated_games_by_test_user[non_rated_games_by_test_user \
    [test_user_id]==True]['Title'].values.tolist()
rated_games_by_test_user = pd.DataFrame(ratings_ICF[test_user_id] \
    .dropna(axis=0, how='all').sort_values(ascending=False)) \
    .reset_index().rename(columns={1:'rating'})

for non_rated_game in non_rated_games_by_test_user:
    similarity_score_with_non_rated_game = \
        item_similarity_c[[non_rated_game]].reset_index() \
        .rename(columns={non_rated_game:'similarity_score'})
    similarity_with_test_user_rated_games = \
        pd.merge(left=rated_games_by_test_user, \
        right=similarity_score_with_non_rated_game, \
        on='Title', how='inner') \
        .sort_values('similarity_score', ascending=False)[:topN]
    predicted_rating = round(np.average(
        similarity_with_test_user_rated_games[test_user_id],
        weights=similarity_with_test_user_rated_games \
        ['similarity_score'], 6)
    rating_prediction[non_rated_game] = predicted_rating
```

Fig.10. Code snippet of the item-based collaborative filtering recommendation algorithm. Used similarity measure is cosine similarity.

IV. TESTING AND RESULTS

During development occurred challenges, needed to solve and test to find out the most appropriate version of the investigated terms. Furthermore, it was necessary to test how the prepared recommendation systems perform and which one has the most accurate suggestions.

A. Optimize ratings dataset

The size of the dataset containing the scraped games, the users, and the given ratings to games by the users occurred memory problems during processing it with the recommendation algorithm. The error was occurred during the calculations of the similarity matrix.

For optimizing these problems some limitations were introduced in data cleansing process to reduce the size of the

file. In addition, after data was loaded for the recommendation phase the following further steps were taken:

- data types were converted to type takes up less memory
- introduced garbage collector to delete non-used variables, thereby freeing up memory

For determine the used memory by the file a Pandas function was used. The main aim was reaching the limit memory usage be around 35-40 megabytes which can still be handled by the computer performing the operation. At the beginning the size of the file which will be using for the similarity matrices was around 123 megabytes, but after the transformations and preparations theses measure only counts 39 megabytes which can be handled by the computer already, without throwing any error.

Another solution would be for this problem is using Spark or Map Reduce technology or acquiring a more powerful computer for higher performance calculations.

B. Test and evaluate recommendation systems

In order to see how the recommendation systems perform testing, the evaluating are important steps. In present work the created systems are working with offline data, which means not real-time user interactions and ratings are the basis of the recommendations, instead pre-saved historical data was scraped from the website.

Big disadvantages of offline dataset are that hard to evaluate them, because of the lack of user interactions. Thus, the recommender's influence on user behavior cannot be directly measured.

As the goal of the offline evaluation is to filter algorithms, the data used for the offline evaluation should match as closely as possible the data the designer expects the recommender system to face when running it online [2].

In the present research work picked test users will be determined than the games which was rated by them will be compared with the recommended items. For measuring accuracy, the Mean Absolute Error and Root Mean Squared Error will be used [2]. Accuracy is one of the most discussed properties measuring while testing recommendation systems, but there are simple equations to work with.

For testing the systems according to what results it gets one user is selected and recommend him or her items with the systems. For this test case user *buffobass* was chosen. This user's 29 ratings can be found in the dataset, the system is working with. It can be said about this user that likes boardgames which genres are strategy, family, party or thematic. It is also known that he or she rather likes the more complex games with longer playing time. The preferred mechanisms in the games the card/drafting games, hand management games, set collection, dice rolling and cooperative games (there are many preferable, but these are the most frequents). The first 20 games rated by the test user can be seen in Fig. 11. with some comments to highlight the common properties.

Title	Comment
Kemet	Strategy, Hand management, Drafting
Horrified	Thematic, Family, Cooperative
Cosmic Encounter	Strategy, Thematic, Hand management
Castles of Mad King Ludwig	Strategy, time 90 minutes
Ora et Labora	Strategy, time 60-180 minutes
Ticket to Ride: Nordic Countries	Family, Hand management, Drafting
Captain Sonar	Thematic, Party
Legendary: A Marvel Deck Building Game	Thematic, Card game
T.J.M.E Stories	Thematic, Dice rolling, Cooperative
Secret Hitler	Party, Card game
Deception: Murder in Hong Kong	Party
My City	Strategy, Family
Fantasy Realms	Family, Card game, Drafting, Set collection
KLASK	Family, Party
Onitama	Abstract, Hand management
Exit: The Game – The Abandoned Cabin	Thematic, Cooperative
Century: Spice Road	Family, Card game, Hand management
Bärenpark	Family, Drafting, Set collection
Awkward Guests	Thematic, Family, Card game, Hand management
Meadow	Strategy, Family, Hand management, Drafting

Fig. 11. The first 20 games rated by the test user.

Investigating the recommended items, it can be said, that the suggested games for the most part have the same genres as the rated games by the test user. The recommended items have also higher playing time, with complex rules and with similar mechanics what the test user prefers. It is also an interesting connection that because of the test user liked one game with content of trains therefore 3 out of 4 recommendation systems recommended games with the same train/railways content, with similar mechanisms and genres.

Fig. 12. shows the recommendations of the item-based recommendation system made with cosine similarity, where some comments were added, to describe the main common properties of the games.

Item-based with cosine	Comment
Commands & Colors: Ancients	Dice Rolling, Hand Management
Agricola: All Creatures Big and Small	Strategy
For Sale	Family, Hand management, Card game
Samurai	Strategy, Hand management, Set collection
Flamme Rouge	Family, Hand management
Dune	Thematic, Strategy, time 120-180 minutes
Railways of the World	Strategy, time 120 minutes
Ticket to Ride	Family, Hand management
Memoir '44	Dice rolling
1960: The Making of the President	Strategy, Hand management

Fig. 12. Top 10 recommendation with item-based recommendation system with cosine similarity.

By using the different type of recommendation systems, it can be observed that even changing only the similarity parametrization the results will be differ from each other, but the recommended items fit to expectations.

V. CONCLUSION

This work presents different types of recommender systems that have been developed for recommending board games. As it can be seen in the presented test case the created recommendation systems based on item-based collaborative filtering are able to recommend items which suits the user's preferences.

The other types investigated recommendation systems have coverage with these suggestions, but mostly the recommended games vary from system to system. But it is clear, that the resulted items own very similar properties in both the rated games by the test user and the recommended ones.

Introducing more testing and evaluation techniques and by drawing conclusions from the results more accurate and user-friendly recommendation systems may be built.

REFERENCES

- [1] „Board Game Geeks,” Available: <https://boardgamegeek.com/>. accessed 23 September 2022.
- [2] F. Ricci, L. Rokach, B. Shapira és P. B. Kantor, Recommender Systems Handbook, New York: Springer, 2011.
- [3] L. Sharma és A. Gera, „A Survey of Recommendation System: Research Challenges,” International Journal of Engineering Trends and Technology, volume 4, issue 5, pp. 1989-1992, 2013.
- [4] M. S. P. Babu és B. R. S. Kumar, „An Implementation of the User-based Collaborative Filtering Algorithm,” International Journal of Computer Science and Information Technologies, volume 2, issue 3, pp. 1283-1286, 2011.
- [5] R. Mitchell, Web Scraping with Python, Sebastopol: O'Really Media, 2018.
- [6] J. Han, M. Kamber és J. Pei, Data Mining: Concepts and Techniques, Cambridge: Elsevier Science & Technology, 2011.
- [7] Amy GrabNGoInfo, „Medium,” Available: <https://medium.com/grabngoinfo/recommendation-system-user-based-collaborative-filtering-a2e76e3e15c4>. accessed 25 September 2022.
- [8] Amy GrabNGoInfo, „Medium,” Available: <https://medium.com/grabngoinfo/recommendation-system-item-based-collaborative-filtering-f5078504996a>. accessed 25 September 2022.
- [9] M. Qutbuddin, „Towards Data Science; An Exhaustive List of Methods to Evaluate Recommender Systems,” 20 April 2020. Available: <https://towardsdatascience.com/an-exhaustive-list-of-methods-to-evaluate-recommender-systems-a70c05e121de>. accessed 28 September 2022.

Virtual Reality Game Development with External Hand Motion Sensor

Erik Dobra
Alba Regia Technical Faculty
Óbuda University
Székesfehérvár, Hungary
erik.dobra@stud.uni-obuda.hu

Éva Hajnal
Alba Regia Technical Faculty
Óbuda University
Székesfehérvár, Hungary
hajnal.eva@amk.uni-obuda.hu

Abstract— In this paper the addition of a hand recognition device to VR glasses was demonstrated. In the investigation a leap motion sensor was added to a pair of Windows Mixed Reality VR glasses, which allows hand tracking without a hand controller. The implemented solution was tested in a sample application. The theme of the game is horror, which encourages the player to react quickly. Three alternative implementation of moving in virtual space in the application was tested. The development of the control unit redesigned for virtual space was done in a Unity based project. The system allows players to perform specific gestures using the thumb and index finger, specifically to press buttons. The software using the built-in profiling system was evaluated, which shows what future optimizations are possible within the project. In summary, we have succeeded in creating a hand recognition system that tracks accurate hand movements in real time, which is a significant improvement in hand recognition implementation compared to the driver produced for the system.

Keywords— VR, Virtual Reality, Hand-Tracking, Leap Motion, Windows Mixed Reality, Unity.

I. INTRODUCTION

Virtual reality is an environment simulated by a computer that you can interact with.

VR replaces reality with an imaginary world that can be seen through a special head-mounted device (HMD). In this environment, hand tracking is an important task [1-10]. Many devices do not have this functionality, or it is not sufficiently accurate for an industrial application. It is common that hand motion tracking is designed to follow a special hand controller, whereas many applications would require the use of someone's naked hand. Our goal was to create a VR game that is both fun to use and opened up a new way for users to use handheld interactions and in-game movement within the game. It is a very important question, because VR sector is currently at a turning point. Developers are now having to adapt to the hardware, and in the same way, the hardware is still evolving. Sometimes it is very straightforward to look at the problem from a developer's point of view: how can we create a software that meets the needs using specific hardware, how difficult is it, how can it be done. The basic idea with VR is to translate as many sensory experiences as possible into virtual space and to make them immersive for the user. Bringing our limbs into the virtual space is a milestone to this long way.

Our target is not only the entertainment industry, but can also benefit many other industries in the field of simulation [9]. We know of applications in the hazardous or heavy industry from our present studies. There is a wish that our system could be used for other industries daily practice later on.

Also, there is an increased tendency for people to stay more and more indoors (specifically because of the last years' health situation) and consequently people are looking for ways to break away from everyday reality, so there is an increased willingness to connect to virtual reality. The research questions are the ways of hand control implementation methods, their usefulness from the aspect of the user and from the aspect of system's performance. The structure of this paper is the next: section II. and III. contain applied hardware and software solutions, section IV. deals with the investigation of the usefulness of the introduced solutions and chapter V. contains the Conclusion.

II. HARDWARE IMPLEMENTATION AND TEST ENVIRONMENT

By our plan a Leap Motion control unit was required physically connecting to the HMD head unit. Based on the manufacturer's recommendation, the Leap Motion was mounted in the middle of the front of the HMD. For this purpose, a special adapter was designed. According to the manufacturer's website [1] it was modelled in Fusion 360 [2]. A frame around the model was drawn and then cut out the port location (Figure 1). Along the frame, two bumps on each side were created, and a support element to fit the curve of the VR Headset, which fits at the points of the bumps to ensure it is held in place.

Design was created in Ultimaker Cura [3], which implements the printing mesh for the model. The elements were then 3D printed (Fig. 2) [4-8].

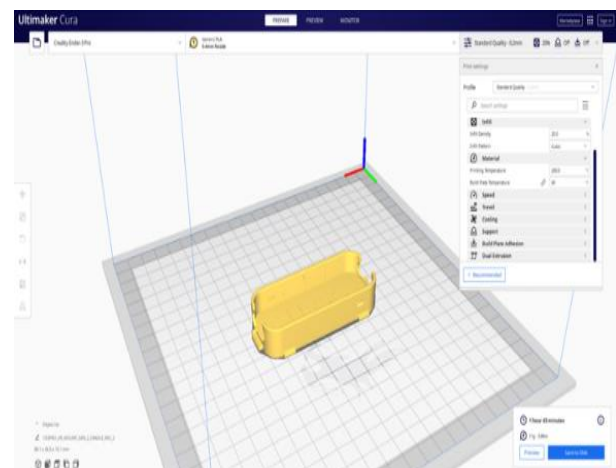


Figure 1: Modelling in Ultimaker Cura software.



Figure 3: 3D Printed Leap Motion holder from different angles.

After that we had to plan of use cases, which VR gestures could be implemented if You use your own hands in the virtual space. First is to direct the movements in the labyrinth. There were three implementations of player movement that were examined.

The second is the fine handling of objects in the VR space.

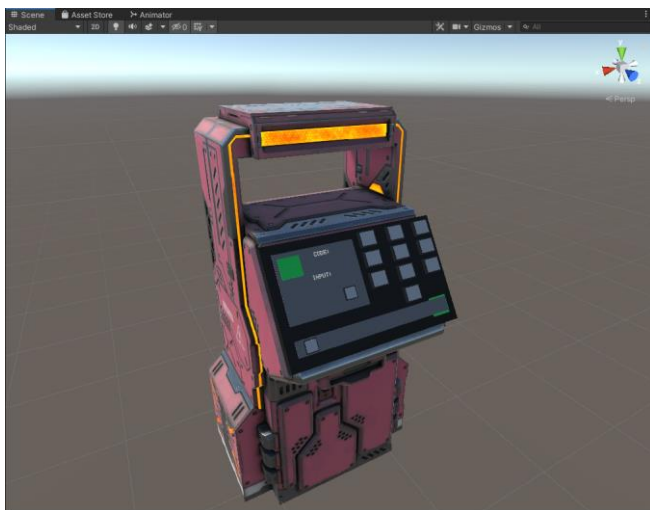


Figure 4: Interactive buttons within the game.

An ATM-like model was created with different operating controls on them (Fig. 3). The question is whether fingers can be recognized and interacted with. First implementations were buttons, next to the buttons, sliding cubes were also placed that need to be dragged to continue the game.

There is also a life level indicator system (Fig 4.) that monitors the rotation of the carpal bones on the left arm and appears when you turn your palm towards you. The player has three lives, and this is presented in VR.



Figure 2: Life indicator fitted to wrist movement.

III. HAND MOVEMENT RECOGNITION

A. Hand-Bone technique

It is based on the recognition of the rotation of the hand.

The most important of these is the calculation of rotation vectors - Euler angles or quaternions.

The Euler angle theorem is useful and important because it describes the rotation of a body with help of three angles. Since three perpendicular coordinate axes can be distinguished in space (X, Y, Z), three separate rotations can be performed about each axis. An acceptable alternative is to use quaternions, which allow us to perform any rotation, regardless of orientation, with a given angle and rotation value given by the vector. The first solution was preferred, as its calculation is more complex, but Unity Engine has a service for this calculation.

This principle could be used to control our character, and knowing that the hand recognition implementation was successful, we dared to think boldly and implement the movement of the planes to control the movement manually.

As we know, it can rotate on three axes, which helps it to move in space. Yaw, Pitch, Roll, i.e. lateral, vertical and longitudinal rotation. If we put our hands out in front of us, we can rotate our wrists in these three planes, because wrist movement is a movement with 3 degrees of freedom. If we map this logic into the above-mentioned hand recognition sensor, we can not only in principle, but also in practice, make the movement work with "gestures" in the absolute space

B. Magicsee R1 controller

The first, in which we used a tiny Magicsee R1 joystick ring [11-13]. This ring has an analogue lever that can be hidden on the inside or outside of the hand, which doesn't interfere with hand recognition, so it retains the advantage of both worlds, which is its own specialty. Fine hand movement can still be used to solve in-game tasks, but movement is handled by the joystick and headset.

This is a safe approach to the problem that will ensure proper, guaranteed operation.

C. Virtual VR keystrokes with hand recognition

Another, more daring approach, which is not provided by the vendors, is to push the boundaries of the framework with our own implementation. As some elements are not



Figure 5: Activating hand gesture illustration.

implemented in the Unity development environment, creative solutions were introduced in one or two cases for direct the movements in the labyrinth.

Here's how it works:

If the player extends the thumb and index finger (forming a pistol) with both hands, a script is activated, which results in a virtual keystroke (in my case "W") (Figure 5).

A keystroke will force the virtual keyboard to move forward, and then you can use the north point of the headset to move forward. If one of your hands is no longer showing the gesture, the forward progress will stop.

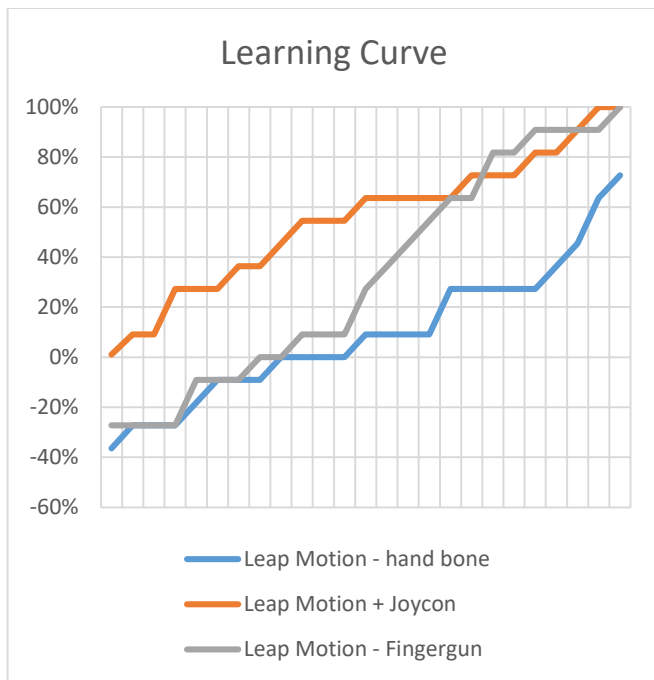
It is important to note, once again, that this tricky solution is not native, it is not feasible by default in the current development environment, but it is representative of what this build would be capable of. For me, this was important to point out because it was conceptually suitable, but a solution that could affect the stability of the computer would not be considered as final (in my case, virtual keystrokes meant that even if I minimized the program, it still monitored the hand movements and guaranteed keystrokes outside the program) [14-16].

IV. COMPARATIVE ANALYSIS OF MOTION CONTROL METHODS

Learning curves of the above-mentioned three solutions were analyzed in a non-representative investigation. The implemented three different locomotion technologies were compared by handling speed using the timer implemented in the game. Each successful run was recorded, measured and saved in a table, rounded to minutes for the simplicity. The 25 runs of each type were normalized with eq. 1 to show the performance that can be achieved with these three types of movement methodology.

TABLE I. RESULTS OF THE INVESTIGATION OF THREE IMPLEMENTED METHODOLOGY

Leap Motion Fingergun (C method)		Leap Motion hand bone (A method)		Leap Motion Joycon (B method)	
z_i	F3 /min	z_i	F2 /min	z_i	F1 / min
-27%	19	-36%	20	1%	16
-27%	19	-27%	19	9%	15
-27%	19	-27%	19	9%	15
-27%	19	-27%	19	27%	13
-9%	17	-18%	18	27%	13
-9%	17	-9%	17	27%	13
-9%	17	-9%	17	36%	12
0%	16	-9%	17	36%	12
0%	16	0%	16	45%	11
9%	15	0%	16	55%	10
9%	15	0%	16	55%	10
9%	15	0%	16	55%	10
27%	13	9%	15	64%	9
36%	12	9%	15	64%	9
45%	11	9%	15	64%	9
55%	10	9%	15	64%	9
64%	9	27%	13	64%	9
64%	9	27%	13	73%	8
82%	7	27%	13	73%	8
82%	7	27%	13	73%	8
91%	6	27%	13	82%	7
91%	6	36%	12	82%	7
91%	6	45%	11	91%	6
91%	6	64%	9	100%	5
100%	5	73%	8	100%	5



6. Figure: Learning rates calculated by eq.1

The following is the formula for the calculation:

$$z_i = -\frac{x_i - x(\max_n)}{x(\max_n) - x(\min_n)} \quad (1)$$

where the x_i is the actual measured value, n is the set with the smallest interval and the smallest variance (this is F1 in this case)

In each case, the test was executed 25 times to see the potential for improvement. The best results were 5 minute with fingergun and joycon solutions. The learning curve is better at joycon solution (Fig. 6.). As the figure above shows, the most difficult movement technique to learn was the wrist movement based one, which made it even more difficult to learn due to the absolute orientation (lack of dynamic north point selection mentioned above). The joycon solution is the most effective, but it still does not give up the idea of having empty hands, so the golden mean is the "fingergun" method, which is more difficult to learn than joycon, but guarantees similar results (Fig 6.).

V. CONCLUSION

Our goal was to solve the problem of using hands in the VR environment, replacing handheld controllers. My choice of topic was motivated by the fact that there are many applications where this might be necessary. This was supported by the literature research, but it is not solved in many HMD devices. Where it is solved, the accuracy of hand tracking and gesture recognition typically leaves something to be desired. As part of this exercise, several devices were tested and selected a hand recognition device, the Leap Motion sensor, to pair with my existing VR glasses. To fit it together, a 3D printed mounting element was made to secure the device. We reviewed the literature to see what options were available if we could incorporate the hand recognition device for in-game movement and interaction. Then a virtual console was implemented that could be controlled by buttons and a special display that worked based on gesture evaluation. We implemented 2+1 separate types of solutions for controlling spatial movement and compared their usability. I created a game that randomly generated a maze and included all the features. Based on my tests, I was able to implement a low-latency hand recognition algorithm. As an extension, I proposed additional new program features that could have specific behaviors. We hope that our investigations will be useful in VR practice, especially in the industry.

ACKNOWLEDGMENT

The paper was supported by HHTDK-21 grant.

REFERENCES

- [1] "Tracking | Leap Motion Controller | Ultraleap." <https://www.ultraleap.com/product/leap-motion-controller/> (accessed Nov. 10, 2021).
- [2] "Fusion 360 | 3D CAD, CAM, CAE, & PCB Cloud-Based Software | Autodesk." <https://www.autodesk.com/products/fusion-360/overview?term=1-YEAR&tab=subscription&plc=F360> (accessed Oct. 02, 2022).
- [3] "Ultimaker Cura: Powerful, easy-to-use 3D printing software." <https://ultimaker.com/software/ultimaker-cura> (accessed Oct. 02, 2022).
- [4] D. Holmes et al., Usability and performance of Leap Motion and Oculus Rift for upper arm virtual reality stroke rehabilitation Communication Style and Exercise Compliance in Physiotherapy (CONNECT) View project PhD Motion Tracking in Digital Images View project Usability and performance of Leap Motion and Oculus Rift for upper arm virtual reality stroke rehabilitation. 2016. [Online]. Available: <https://www.researchgate.net/publication/312868469>

- [5] A. G. D. Corrêa, N. R. Kintschner, V. Z. Campos, and S. M. Blascovi-Assis, "Gear VR and leap motion sensor applied in virtual rehabilitation for manual function training* an opportunity for home rehabilitation," in *PervasiveHealth: Pervasive Computing Technologies for Healthcare*, Sep. 2019, pp. 148–151. doi: 10.1145/3364138.3364169.
- [6] D. E. Holmes, D. K. Charles, P. J. Morrow, S. McClean, and S. M. McDonough, "Using fit's law to model arm motion tracked in 3D by a leap motion controller for virtual reality upper arm stroke rehabilitation," in *Proceedings - IEEE Symposium on Computer-Based Medical Systems*, Aug. 2016, vol. 2016-August, pp. 335–336. doi: 10.1109/CBMS.2016.41.
- [8] "Oculus Quest 2 Hand Tracking Has So Much Untapped Potential." <https://www.thegamer.com/oculus-quest-2-hand-tracking-has-so-much-untapped-potential/> (accessed Nov. 10, 2021).
- [9] F. Nainggolan, B. Siregar, and F. Fahmi, "User Experience in Excavator Simulator using Leap Motion Controller in Virtual Reality Environment," in *Journal of Physics: Conference Series*, Jul. 2020, vol. 1566, no. 1. doi: 10.1088/1742-6596/1566/1/012093.
- [10] F. Fahmi, K. Tanjung, F. Nainggolan, B. Siregar, N. Mubarakah, and M. Zarlis, "Comparison study of user experience between virtual reality controllers, leap motion controllers, and senso glove for anatomy learning systems in a virtual reality environment," in *IOP Conference Series: Materials Science and Engineering*, May 2020, vol. 851, no. 1. doi: 10.1088/1757-899X/851/1/012024.
- [11] Valve Corporation, "Valve Index Controllers: Your hands, in VR," <https://www.valvesoftware.com/en/index/controllers>, Jun. 28, 2019.
- [12] G. 'Pierre-Loup A., "Valve Index controller Extender," https://github.com/ValveSoftware/IndexHardware/blob/master/Controller/index_controller_example.stl, Jun. 25, 2019.
- [13] W. 'Mason, "Palmer Luckey Speaks Candidly About Xbox One Controllers Being Bundled with the Rift," <https://uploadvr.com/palmer-luckey-speaks-candidly-about-xbox-one-controllers-being-bundled-with-the-rift/>, Jun. 12, 2015.
- [14] "Quest Gets Hand Interaction SDK & Tracked Keyboard SDK." <https://uploadvr.com/quest-interaction-sdk-tracked-keyboard-sdk/> (accessed Apr. 17, 2022).
- [15] C. Khundam, V. Vorachart, P. Preeyawongsakul, W. Hosap, and F. Noël, "A comparative study of interaction time and usability of using controllers and hand tracking in virtual reality training," *Informatics*, vol. 8, no. 3, Sep. 2021. doi: 10.3390/informatics8030060.
- [16] "Leap Motion Hand Tracking vs Oculus Quest One vs Vive One | PaleBlue." <https://pale.blue/2020/07/09/leap-motion-hand-tracking-vs-oculus-quest-one-vs-vive-one/> (accessed Nov. 10, 2021).

POSTER SESSION III.
Geodesy and Geoinformatics

Lever arm calculating algorithm for efficient and precise vector determinations

Gergely LÁSZLÓ

Institute of Geoinformatics
Óbuda University, Alba Regia
Technical Faculty
Székesfehérvár, HUNGARY

Department of Orthophotogrammetry,
Envirosense Hungary Ltd.
Debrecen, HUNGARY
laszlo.gergely@amk.uni-obuda.hu

Gábor Péter MOLNÁR

Institute of Geoinformatics
Óbuda University, Alba Regia
Technical Faculty
Székesfehérvár, HUNGARY
molnar.gabor@amk.uni-obuda.hu

Gabriella BOR

Department of Orthophotogrammetry,
Envirosense Hungary Ltd.
Debrecen, HUNGARY
gabriella.bor@envirosense.hu

Abstract— In the orthophoto processing one of the key steps is the aerial triangulation during which the external orientation coefficients of the photographs are determined by adjustment. In order to be able to do this with sufficient accuracy, in addition to knowing the internal orientation elements of the cameras, we need the eccentricity vectors of the GNSS antennas, IMU and sensors determined in relation to each other, i.e. the lever arms. Characteristic points of the airplane and the installed instruments are measured using engineering geodesic method, while vectors are traditionally determined using a geometric technique.

In this paper we will present an algorithm which uses an analytical geometric approach for much quicker and more precise results.

Keywords— *lever arm, programming, analytical geometry, photogrammetry*

I. INTRODUCTION

Remote sensing is used for data collection when we cannot or do not want to work directly in the area to be surveyed. In order for a large part of the field measurements to be substituted, it is important that the accuracy of the remotely sensed data does not fall short of that available with field data collection procedures. To fulfill this aspect, our remotely sensed data must meet several criteria. For example, they are suitable if the followings are known:

- ground sample distance (GSD)
- depending on the final product, image overlap relative to each other
- sensor parameters known with sufficient accuracy
- coordinates of data collection locations determined with high precision and the spatial angle of the collected data in some known projection system, and designated direction
- and, of course, the appropriate data collecting conditions

If we are talking about photo-based aerial remote sensing, then we can ensure the desired field resolution by knowing the parameters of the measuring camera (internal orientation elements: pixel size, sensor size, focal length, principal point

coordinates, distortions parameters), based on a topography model, and calculating the appropriate flight height.

Some of the required camera parameters (pixel size in millimeters, sensor size, nominal focal length) are determined by the camera and lens type, while the rest (principal point coordinates, actual focal length, distortions parameters) require camera calibration.

To plan data collection with the required overlap, in addition to the method of use of the final product (e.g.: orthophoto, true ortho), which determines the amount of overlap to be used, we also need the calibrated parameters of the camera sensor and the previously calculated flight height.

Adequate data acquisition conditions can be achieved if a sensor suitable for the purpose is used under ideal environmental conditions (weather, time of day).

Knowing the data collection locations is important because in photogrammetry - or in any remote sensing procedure - the basic task is to be able to produce field coordinates from the image coordinates of recordings, thus making it possible to carry out measurements on them. To do this, we need to determine exactly where the projection center of the camera was located in space (interpreted in a given projection) at the time of recording. If we also know the orientation of the image plane, then we know all the external orientation elements ($x, y, z; \omega, \phi, \kappa$) in order to perform field coordinate calculations from pixel coordinates.

The problem with the above point is that the external orientation elements can only be determined indirectly, with separate sensors (GPS, IMU), not directly in the projection center or in the image plane, so the measured values are always burdened by eccentricity. For smaller, more compact systems, such as UAV survey systems, these eccentricity vectors are defined and published by the manufacturer, but for large-scale aerial photography, especially if the sensor configuration changes frequently, they need to be determined and checked regularly.

In this study, the authors present the traditional, geometric way of determining the eccentricity vectors, as well as the developed algorithm, which automatically enables the calculation of these vectors using an analytical method.

II. LEVER ARM MEASUREMENT

We can perform the measurement when the sensors are installed in the aircraft. The measurement is carried out using an engineering geodesic method, typically with a total station. In order for the eccentricity vectors and the rotation angles to be properly interpreted, in addition to the sensors, a reference plane and a reference direction must also be measured. In our case, the reference plane was the base plate holding the sensors, while the direction was the longitudinal axis of the fuselage. Overall, the following points are measured:

- base plate flat, with three points (bore holes A, B, C) (Figure 1)



1. Figure Points on plate, boreholes: point A, B and C

- plane longitudinal axis, with two points (t1, t2) (Figure 2)



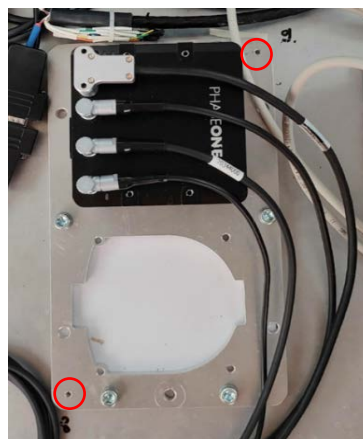
2. Figure Points on plane, endpoints of axis line

- inertial measurement unit, depending on the type, with one or two points (IMU) (Figure 3)



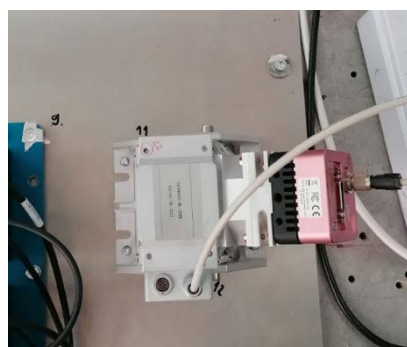
3. Figure Point on IMU

- camera (fixing frame), with two points (cph1, cph2) (Figure 4)



4. Figure Points of camera

- hyperspectral camera, with two points (hs1, hs2) (Figure 5)



5. Figure Points on Hyper

- two GPS antennas, with one point each (ant1, ant2) (Figure 6)



6. Figure Points of the antennas

During the measurement, we stand next to the plane with the instrument in such a way that both the outer and inner points are visible at the same time, in five rounds it takes place in one face, using a propped-up miniprism, with a 300 mm extension rod at some points if necessary.

III. PROCESSING WITH TRADITIONAL, GEOMETRIC METHOD

We start the processing by averaging the five-round measurement. Thus, if at any point we experience an error in the measurement results, we have the option to remove such values.

The coordinate list obtained in this way is imported point by point into AutoCAD, then for easier work we also place labels next to the points with point numbers and then group them by using layer allocation.

By connecting the points of the plane, we create a plate, then we also connect the axis points so that we can use them as a reference line.

We create a user coordinate system by shifting so that the XY plane contains the plane of the base plate, and the origin coincides with one of the measured bore holes (e.g., point A).

After that, the axis-line is moved with its midpoint to the origin, then its start and end points are rotated to the XY plane by zeroing the Z coordinates.

In the next step, we rotate the +X axis in the direction of the aircraft axis so that it faces the propeller.

Using the measured IMU points and the available technical drawings, we draw the reference point of the IMU, to which we move the origin using a repeated coordinate system shift.

In this final, directional reference system, based on the sensor points, using the dimensions of blueprints and 3D models, we determine the position of the camera focal points and the center of mass of the hyperspectral camera with various editing steps.

As a final step, the X, Y, Z coordinates of the transformed points are manually read point by point and recorded in a coordinate list, which can be used to perform the trajectory processing.

IV. PROCESSING USING ANALYTICAL METHOD

It can be seen from the previous chapter that the geometric method can be quite time-consuming and, because of its nature, is not necessarily the most accurate due to rounding and certain simplifications (e.g. we assume that the axes of the cameras are parallel to the axes of the transformed coordinate system).

It can also be observed that the whole process consists of displacements and rotations that can be combined into a 3D Helmert transformation (without scaling), and some edits that can be converted into polar point calculations and rotations.

Taking these into account, the algorithm is built from the following analytical geometric calculations:

The plane normal of the base plate are calculated from the coordinates of point A, B and C, (situated on the base plate) using the difference vectors of the position vectors of these points

$$\mathbf{n} = \mathbf{v}_{AB} \times \mathbf{v}_{AC}$$

This \mathbf{n} plane normal should be normalized, simply dividing by its length:

$$\hat{\mathbf{n}} = \mathbf{n}/|\mathbf{n}|$$

Using the t_1 and t_2 axis endpoints of the airplane, the axis vector can be expressed:

$$\mathbf{t} = \mathbf{t}_2 - \mathbf{t}_1$$

and normalizing it, we get the normalized airplane axis vector:

$$\hat{\mathbf{t}} = \mathbf{t}/|\mathbf{t}|$$

We could project this normalized airplane axis vector to the plane defined by the base plate, and we get the \mathbf{f} vector, what is on the plane and defines the forward direction.

$$\mathbf{f} = \hat{\mathbf{t}} - \hat{\mathbf{n}} \cdot (\hat{\mathbf{n}} \cdot \hat{\mathbf{t}})$$

where we use the dot product of vectors.

The cross product of the normalized $\hat{\mathbf{f}}$ vector and the $\hat{\mathbf{n}}$ plane normal vector is perpendicular to both of them:

$$\hat{\mathbf{p}} = \hat{\mathbf{f}} \times \hat{\mathbf{n}}$$

We need a rotational matrix, that transforms any vector to the newly defined coordinate system. This matrix would transform the $\hat{\mathbf{p}}, \hat{\mathbf{f}}$ and $\hat{\mathbf{n}}$ vectors to the base vectors of the newly defined coordinate system:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{R} \cdot \begin{bmatrix} \hat{\mathbf{p}}_x & \hat{\mathbf{f}}_x & \hat{\mathbf{n}}_x \\ \hat{\mathbf{p}}_y & \hat{\mathbf{f}}_y & \hat{\mathbf{n}}_y \\ \hat{\mathbf{p}}_z & \hat{\mathbf{f}}_z & \hat{\mathbf{n}}_z \end{bmatrix}$$

where \mathbf{R} matrix is simply the inverse of the right hand side matrix composed of the $\hat{\mathbf{p}}, \hat{\mathbf{f}}$ and $\hat{\mathbf{n}}$ vectors.

V. OPERATION OF PROGRAM CODE

The program itself was created using Python language, as it is easy to program, extremely flexible and widespread nowadays.

Among the optional function libraries, we used numpy module for mathematical and linear algebra functions, while matplotlib module was used to plot the result.

The first step after calling the modules is to import the properly prepared coordinate file. Data lines should be Tab-separated with *ID, description, X, Y, Z* order, where the description corresponds to the previously described abbreviations. The program automatically recognizes and handles the case when the IMU was measured with two points, and then fills a list with the data, which will be used in the coming steps.

In the next step, based on vectors *AB* and *AC*, the plane *ABC* is calculated, its normal vector, length and, based on these, its normal unit vector and the distance of the calculated plane from the origin.

This is followed by the calculation of a direction vector based on the axis points *t1, t2*, and then the determination of a vector in a plane, perpendicular to the axis.

After determining the length of the normal vector, we form a normal unit vector from it.

Based on the normal unit vector of the plane and the axis, we get the projection direction vector of the axis, from which, as a check, if we calculate its length, we should get units.

Using these three unit-vectors, we create the rotation matrix.

If we measured with a two-point IMU determination, the program calculates from two points, if we measured the type where only one is enough, we determine the IMU reference with a Z offset, and then perform the transformation to this point, to all measured elements, with the rotation matrix.

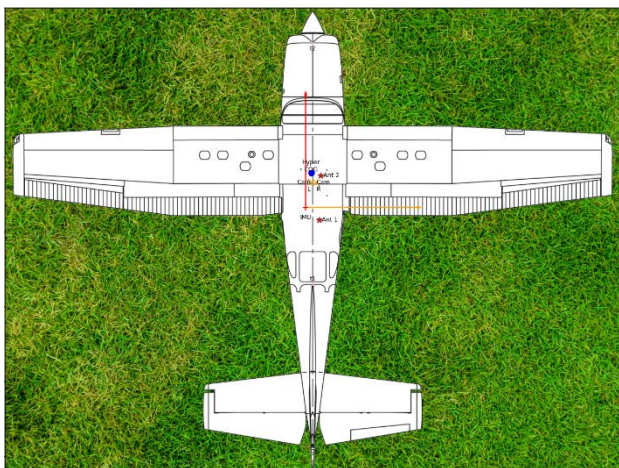
In the transformed system, we determine the camera focal points and the location of the center of mass of the hyperspectral sensor.


```

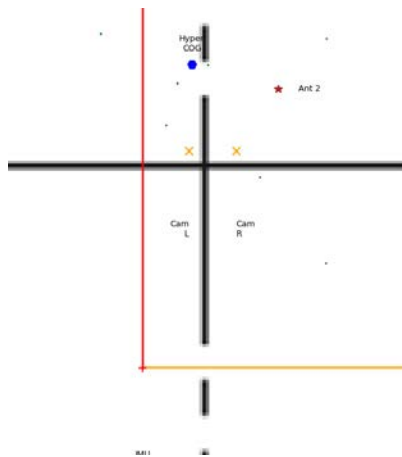
[-3.277 -2.644 -0.825]
A [ 0.42038378  0.24041505 -0.03908122]
B [ 0.4215701  0.7561452  -0.03908122]
C [-0.09698058  0.76687788 -0.03908122]
t1 [ 0.14326238 -1.55210035  0.73907638]
t2 [0.14326238  3.29543082  0.71210277]
IMU [0.  0.  0.023]
cph1 [ 0.26852985  0.43664198 -0.0303994 ]
cph2 [ 0.05231898  0.55699464 -0.02935162]
hs1 [0.0784955  0.65341512  0.05481418]
hs2 [0.14932891  0.69373265  0.0544624 ]
ant1 [ 0.27975696 -0.25416175  0.99509714]
ant2 [0.31202478  0.63951048  1.13808918]
-222.05255335325126
Cam_b [ 0.10638737  0.49705626 -0.02935162]
Cam_j [ 0.21446146  0.49658036 -0.02935162]
Hyp COG [0.11260394  0.69651104  0.00363829]
Pont   X       Y       Z
IMU    0.0     0.0     0.0
Hyper COG      0.113   0.697   0.004
Cam b   0.106   0.497  -0.029
Cam j   0.214   0.497  -0.029
Ant e   0.312   0.64    1.138
Ant h   0.28    -0.254  0.995
    
```

7. Figure Results in Python

In the last step, using the top view diagram of the given airplane type (Figure 8), we draw the locations of the points proportionally, label them with a symbol key (Figure 9). In parallel with this print out the transformed coordinates of the points as a formatted character string on the screen (Figure 7).



8. Figure Location of measured points I.



9. Figure Location of measured points II.

VI. SUMMARY AND POSSIBLE DEVELOPMENTS

It can be said about the completed program that, since it performs the previously manual, geometrically solved task with automatic calculation, it gives results much faster, provides an immediate visual control option, and can also consider aspects in the calculation that were not possible until now due to simplifications, therefore more accurate results are obtained.

In the future, we plan for the program to be able to handle, preferably automatically recognize, the different configurations in the two planes used by the company, and to be able to compare, based on previous measurements, which points' positions have changed to an extent that exceeds the expected measurement accuracy, thus helping the possible error detection.

VII. REFERENCES

- [1] János Katona: Parcel analysis for the general application of remote sensing monitoring, In: AIS 2019 : 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University, Székesfehérvár: Óbudai Egyetem, pp 69-70 (2019)
- [2] Lucas Grégory; Halász László; Solymosi József. Exploring the capacities of airborne technology for the disaster assessment, HADMÉRNÖK 8 : 3 pp. 74-91. , 18 p. (2013)
- [3] Lucas Grégory, Considering time in orthophotography production: from a general workflow to a shortened workflow for a faster disaster response, International Archives of Photogrammetry and Remote Sensing (2002-) XL3 : W3 pp. 249-255. , 7 p. (2015)
- [4] Schickler, W., and A. Thorpe, 1998. Operational procedure for automatic true orthophoto generation, International Archives of Photogrammetry and Remote Sensing, 32(Part 4):527-532.
- [5] Balázsik V., Tóth Z., Abdurahmanov I. Analysis of Data Acquisition Accuracy with UAV INTERNATIONAL JOURNAL OF GEOINFORMATICS (1686-6576): 17 1 pp 1-10 (2021)
- [6] Kugler, Zsófia ; Tóth, Zoltán ; Szalay, Zsuzsa ; Szagri, Dóra ; Barsi, Árpád Supporting microclimate modelling with 3D UAS data acquisition IDŐJÁRÁS / QUARTERLY JOURNAL OF THE HUNGARIAN METEOROLOGICAL SERVICE 123 : 3 pp. 279-294. , 16 p. (2019)
- [7] A. Ládai, C. Toth, Z. Tóth INDOOR MAPPING WITH AN OMNIDIRECTIONAL CAMERA SYSTEM: PERFORMANCE ANALYSIS INTERNATIONAL ARCHIVES OF PHOTOGRAMMETRY AND REMOTE SENSING (2002-) XLIII-B1-2022 pp. 347-352. , 6 p. (2022)

Comparison of lever arm determination methods in terms of their effect of the accuracy of aerial triangulation

Ágnes KRAUSZ
 Institute of Geoinformatics
 Óbuda University, Alba Regia Technical Faculty
 Székesfehérvár, HUNGARY
krauszagi@gmail.com

Gergely LÁSZLÓ
 Institute of Geoinformatics
 Óbuda University, Alba Regia Technical Faculty
 Székesfehérvár, HUNGARY

Department of Orthophotogrammetry,
 EnviroSense Hungary Ltd.
 Debrecen, HUNGARY
laszlo.gergely@amk.uni-obuda.hu

Abstract— With the help of aerial remote sensing, especially when we are talking about large aircraft surveys, we can collect a large amount of information from a large area in a short time. In order for the resulting, in our case, photo-based (photogrammetric) data to be used for engineering purposes, it is important to ensure their accuracy. One of the components of this, in addition to an appropriately accurate camera model, is knowing the 3D position of the photographs and their spatial position with great accuracy. Knowing these, we can perform aerial triangulation. Since these data cannot be measured directly, we need to know the eccentricity vectors of the measuring devices (GNSS/IMU) with respect to the main points of the cameras.

In this study, we will examine the effect of the accuracy of different measurement methods (engineering geodesy, laser scanning, software estimation) on the results of aerial triangulation.

Keywords— lever arm, aerial triangulation, photogrammetry, laser scanning, engineering geodesy

I. INTRODUCTION

Remote sensing is used for data collection when we cannot or do not want to work directly in the area to be surveyed. In order for a large part of the field measurements to be substituted, it is important that the accuracy of the remotely sensed data does not fall short of that available with field data collection procedures. To fulfill this aspect, our remotely sensed data must meet several criteria. For example, they are suitable if the followings are known:

- ground sample distance (GSD)
- depending on the final product, image overlap relative to each other
- sensor parameters known with sufficient accuracy
- coordinates of data collection locations determined with high precision and the spatial angle of the collected data in some known projection system, and designated direction
- and, of course, the appropriate data collecting conditions

If we are talking about photo-based aerial remote sensing, then we can ensure the desired field resolution by knowing the parameters of the measuring camera (internal orientation elements: pixel size, sensor size, focal length, principal point coordinates, distortions parameters), based on a topography model, and calculating the appropriate flight height.

Some of the required camera parameters (pixel size in millimeters, sensor size, nominal focal length) are determined by the camera and lens type, while the rest (principal point coordinates, actual focal length, distortions parameters) require camera calibration.

To plan data collection with the required overlap, in addition to the method of use of the final product (e.g.: orthophoto, true ortho), which determines the amount of overlap to be used, we also need the calibrated parameters of the camera sensor and the previously calculated flight height.

Adequate data acquisition conditions can be achieved if a sensor suitable for the purpose is used under ideal environmental conditions (weather, time of day).

Knowing the data collection locations is important because in photogrammetry - or in any remote sensing procedure - the basic task is to be able to produce field coordinates from the image coordinates of recordings, thus making it possible to carry out measurements on them. To do this, we need to determine exactly where the projection center of the camera was located in space (interpreted in a given projection) at the time of recording. If we also know the orientation of the image plane, then we know all the external orientation elements $(x, y, z; \omega, \phi, \kappa)$ in order to perform field coordinate calculations from pixel coordinates.

The problem with the above point is that the external orientation elements can only be determined indirectly, with separate sensors (GPS, IMU), not directly in the projection center or in the image plane, so the measured values are always burdened by eccentricity. For smaller, more compact systems, such as UAV survey systems, these eccentricity vectors are defined and published by the manufacturer, but for large-scale aerial photography, especially if the sensor configuration changes frequently, they need to be determined and checked regularly.

In this study, the authors present the traditional, geometric way of determining the eccentricity vectors, as well as the developed algorithm, which automatically enables the calculation of these vectors using an analytical method.

II. LEVER ARM MEASUREMENT WITH TOTAL STATION

We can perform the measurement when the sensors are installed in the aircraft. The measurement is carried out using an engineering geodesic method, typically with a total station. In order for the eccentricity vectors and the rotation angles to be properly interpreted, in addition to the sensors, a reference plane and a reference direction must also be measured. In our case, the reference plane was the base plate holding the sensors, while the direction was the longitudinal axis of the fuselage. Overall, the following points are measured:

- base plate flat, with three points (bore holes A, B, C) (Fig. 1)



2. Figure Points on plate, boreholes: point A, B and C

- plane longitudinal axis, with two points (t1, t2) (Fig. 2)



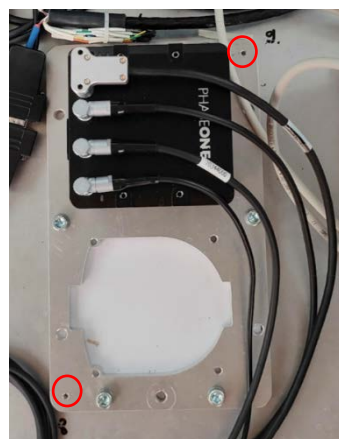
3. Figure Points on plane, endpoints of axis line

- inertial measurement unit, depending on the type, with one or two points (IMU) (Fig. 3)



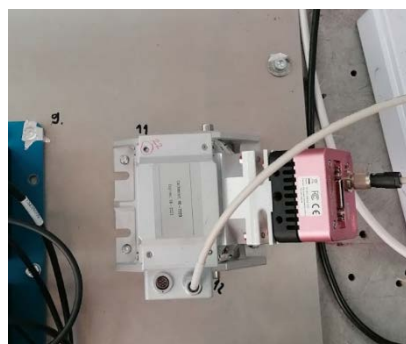
1. Figure Point on IMU

- camera (fixing frame), with two points (cph1, cph2) (Fig. 4)



4. Figure Points of camera

- hyperspectral camera, with two points (hs1, hs2) (Fig. 5)



5. Figure Points on Hyper

-In Figure 6. two GPS antennas, with one point each (ant1, ant2)



6. Figure Points of the antennas

During the measurement, we stand next to the plane with the instrument in such a way that both the outer and inner points are visible at the same time, in five rounds it takes place in one face, using a propped-up miniprism, with a 300 mm extension rod at some points if necessary.

III. PROCESSING WITH TRADITIONAL, GEOMETRIC METHOD

We start the processing by averaging the five-round measurement. Thus, if at any point we experience an error in the measurement results, we have the option to remove such values.

The coordinate list obtained in this way is imported point by point into AutoCAD, then for easier work we also place labels next to the points with point numbers and then group them by using layer allocation.

By connecting the points of the plane, we create a plate, then we also connect the axis points so that we can use them as a reference line.

We create a user coordinate system by shifting so that the XY plane contains the plane of the base plate, and the origin coincides with one of the measured bore holes (e.g., point A).

After that, the axis-line is moved with its midpoint to the origin, then its start and end points are rotated to the XY plane by zeroing the Z coordinates.

In the next step, we rotate the +X axis in the direction of the aircraft axis so that it faces the propeller.

Using the measured IMU points and the available technical drawings, we draw the reference point of the IMU, to which we move the origin using a repeated coordinate system shift.

In this final, directional reference system, based on the sensor points, using the dimensions of blueprints and 3D models, we determine the position of the camera focal points and the center of mass of the hyperspectral camera with various editing steps.

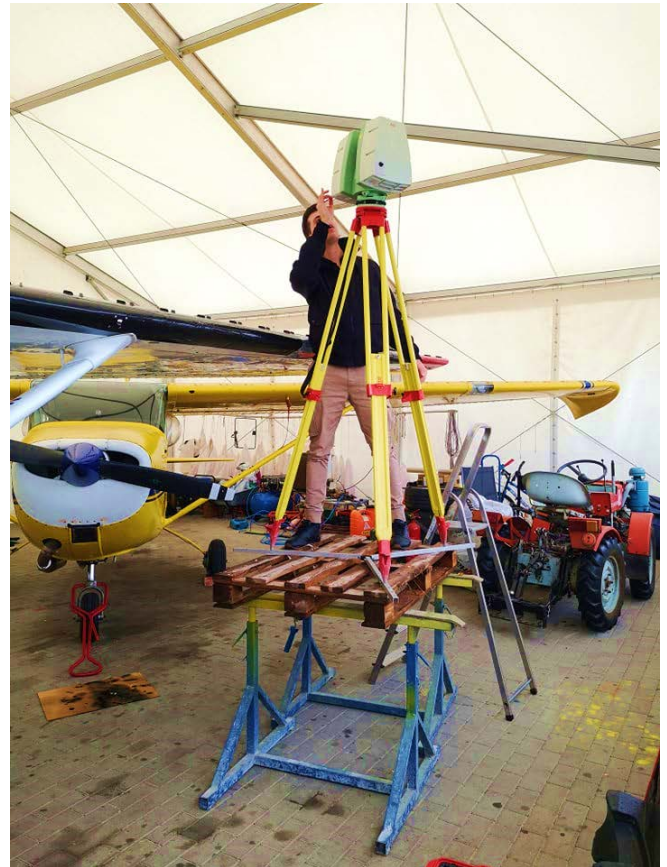
As a final step, the X, Y, Z coordinates of the transformed points are manually read point by point and recorded in a coordinate list, which can be used to perform the trajectory processing.

IV. LEVER ARM MEASUREMENT WITH LASER SCANNER

Laser scanner measurement differs fundamentally from traditional geodetic surveys in that, while the latter measures selected points, with scanning we only have the option of choosing from which point and which part of the space we

want to measure, and the selection of the points we need can only be done during processing.

The planning of the measurement is also different, because due to its structure, it is almost impossible to measure an aircraft from one position in such a way that the GNSS antennas placed on top and the sensors that are of utmost importance for the task are visible, so the measurement must be performed from at least two positions, one of which is also it should be of a high height, relatively close to the plane, as shown in Figure 7.



7. Figure Measurement from elevated station

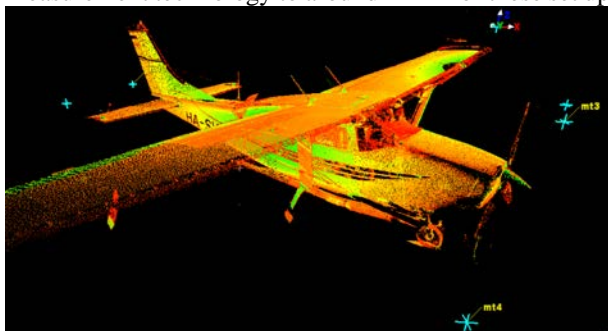
When measuring the interior of the aircraft, the scanner must be set up with a different approach than a total station, because here, in contrast, it is also important that the sensors are visible as much as possible on the point cloud, so that their geometry can be determined as precisely as possible, for the calculation of the characteristic points. This last clause is particularly important because, while a total station can measure a point in a cover (e.g., a drill hole) by raising a prism pole, these points can only be determined indirectly during scanning, based on the surrounding point cloud points. The resolution of the scanning was set to 1x1 mm interpreted at the farthest point of the plane. After the measurement, photographs were also taken for subsequent coloring of the point cloud.

The first step after the measurement was to import the data into the program used for processing (Leica Cyclone). Since the plane was not surveyed from one point of view, it was necessary to transform the stations into each other. Since we used targets during the measurement, which the program recognizes and can automatically use as reference points for the operation, we also have control option over the step. The

point cloud (as in Fig. 8) was matched to targets with a mean error of between 1-3 mm, while the automatic point matching procedure between the point clouds gave the following results:

- o Overlap Point Count: 244133
- o Overlap Error Statistics
- o RMS: 0.00849298 m
- o AVG: 0.00400456 m
- o MIN: 5.20021e-07 m
- o MAX: 0.0938572 m

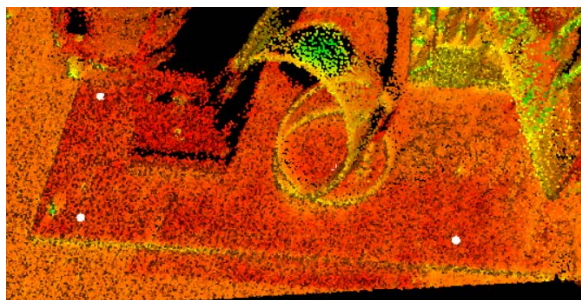
The results tally with the expected values given by the specifications of the instrument and indicate the reliability of the measurement technology to around 4 mm for these set up.



8. Figure Transformed point cloud image of the plane

The transformation of the point cloud was followed by coordinate measurement. This was done in such a way that after identifying the location of the points to be measured, we sorted them together with their immediate surroundings, and then selected with the greatest possible accuracy the point that was most likely the closest to the reference point. After that, we queried the coordinates with object info.

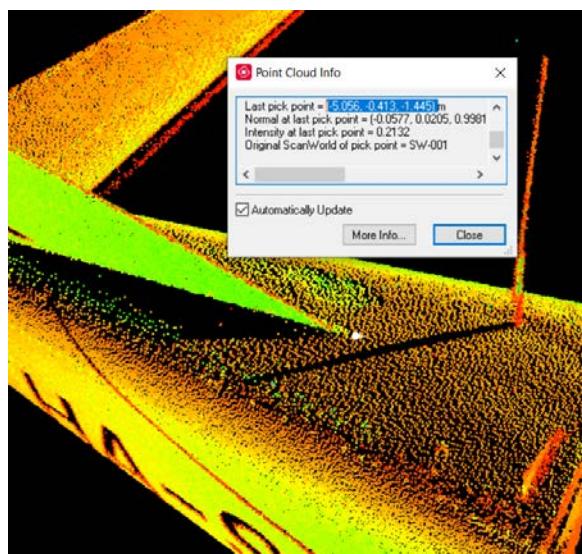
The following figures 9-15. show the measured point on the point cloud:



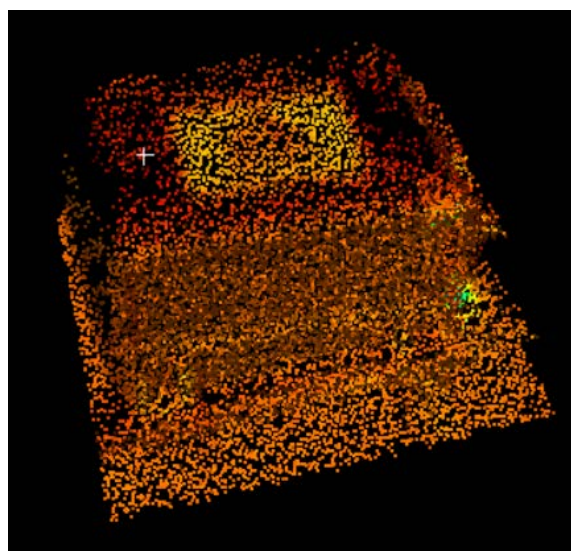
9. Figure Point cloud points on plate



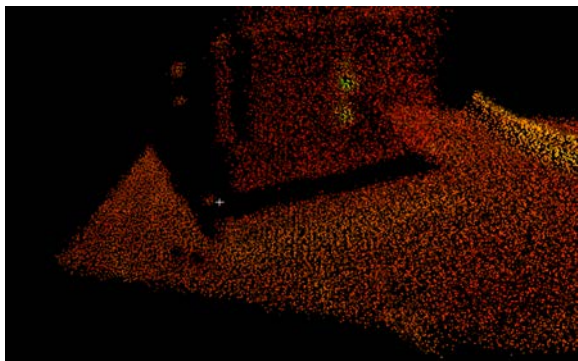
10. Figure Front axis point on the point cloud



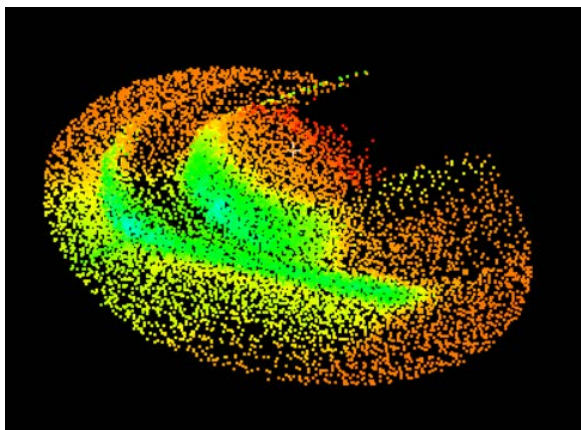
11. Figure Rear axis point on the point cloud



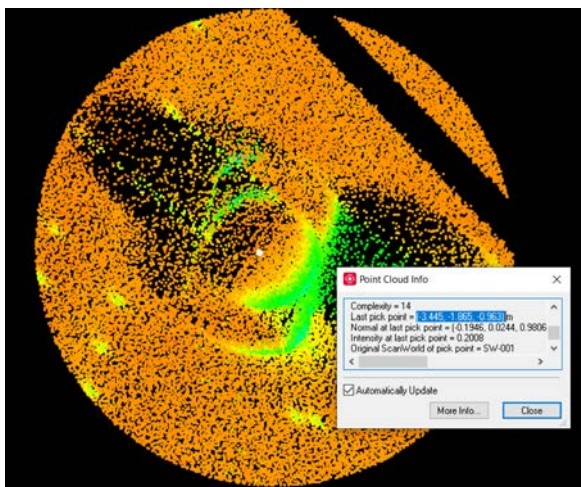
12. Figure IMU and its reference point



13. Figure Camera with measurement point



14. Figure Points of the rear GNSS antenna



15. Figure Point cloud image of the front GNSS antenna

The coordinates were processed and transformed into the IMU origin coordinate system as described in the total station method.

Point	x	y	z
IMU	0	0	0
Ant b	-0.2773	-0.2831	0.9723
Ant f	-0.3091	0.6114	1.1143
Cam l	-0.1063	0.4680	-0.0152
Cam r	-0.2132	0.4681	-0.0152

1. Table 3D vectors of the Total Station measured sensors (Ant-GNSS Antenna, Cam - Camera, b - Rear, f - Front, l - Left, r - Right)

Pont	x	y	z
IMU	0	0	0
Ant b	-0.2794	-0.2839	0.9740
Ant f	-0.3106	0.6103	1.1161
Cam l	-0.1090	0.4712	-0.0118
Cam r	-0.2108	0.4669	-0.0143

2. Table 3D vectors of the Laser Scanned sensors

Point	Δx	Δy	Δz	$\Delta 3D$
IMU	0.0000	0.0000	0.0000	0.000
Ant b	0.0021	0.0008	-0.0017	0.003
Ant f	0.0015	0.0011	-0.0018	0.003
Cam l	0.0027	-0.0032	-0.0034	0.005
Cam r	-0.0024	0.0012	-0.0009	0.003

3. Table Differences of the two measurements

The differences in Tables 1-3. show that the results of the two determinations are very close to each other, the difference is 3 mm on average.

V. EFFECTS OF MEASURED LEVER ARM VECTORS ON AERIAL TRIANGULATION RESULTS

To carry out the test, we had to choose a flight project that was suitable for testing in terms of its scope and other parameters with the given configuration. Our choice fell on the Nagykáta settlement project, as it had the right configuration and was flown in the afternoon on 04/14/2022, a week earlier than the measurement 04/20/2022, under ideal weather conditions. Since it is a small settlement, the aerial triangulation of the 187-frame project can be done in about 10 minutes, which is useful for such research if it needs to be repeated several times. After selecting the project, we calculated the external orientation elements of the photographs based on the flight trajectory with the sensor coordinates created in IMU system for both cases.

Afterwards, to make the results as comparable as possible, we carried out the identification of Ground Control Points (GCP) on a preliminary aerial triangulation and determined the AT parameters during the test. To not distort the results in any way, the camera calibration was not allowed to be refined during the process.

The test parameters were set as follows:

Number of photos : 187
 Number of strips : 7
 Photo scale : 1:13643
 Mean terrain height [m] : 156

Strategy parameter:

Overview # to start : 5
 Overview # to stop : 0
 # of pyramid levels to be processed : 6
 # of tie points to be measured in last level : 3
 Use manual measurements as TPC : ON
 Create numeric point id : ON
 TiePointCenterPattern : 4 x 4

```

TiePoint density           : default
Thinout size in pixel     : 50
Try to track points       : ON
  from upper levels:     : FBM FBM LSM
Matching sequence         : SKIP FBM LSM
Number additional threads  : 1
Cache size per thread     : 10 MByte

Working directory         : .\
DEM for Initialization    :Nagykata_DTM_UTM34.dtm

Matching parameter :
-----

FBM-Method:

Matching window           : 100 x 100
Operator window size     : 5 x 5
Non-max. suppr. window size : 5
Threshold for correlation coefficient : 0.92
Window for correlation coefficient : 5 x 5
Wmin                      : 0.10
Qmin (roundness)         : 0.50
Parallax bound           : 30
Epipolar line to be used : ON
Distance to epipolar lines : 0.50
LSM-Method:

Correlation coefficient   : 0.93
Template size [pixel]    : 21
Edge size [pixel]        : 6
# of iterations           : 20
    
```

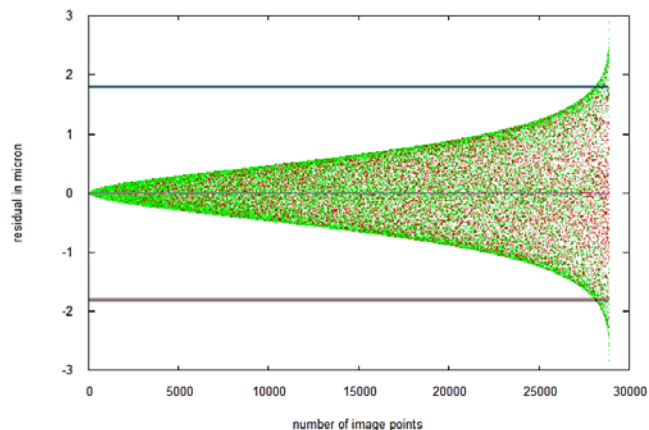
Control parameter for block adjustment:

```

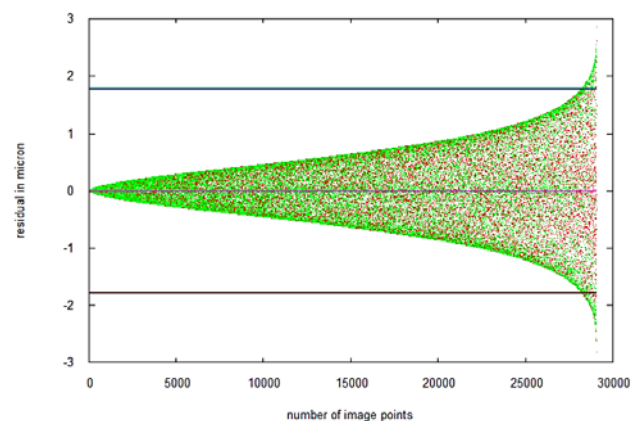
-----
Self-calibration         : OFF
GNSS-Mode                : ON
Drift-Mode               : ON
drift per strip          : ON
drift for X, Y, Z        : ON,ON,ON
enable shifts only       : OFF
IMU-Mode                 : ON
IMU-Boresight            : ON
Earth's curvature correction : ON
Atmospheric correction   : ON
Do not eliminate manual points : ON
Do not eliminate GNSS    : ON
    
```

Looking at the results in Figures 16-18, in both cases the Sigma 0 value of aerial triangulation was 0.8 microns, which corresponds to 0.2 pixels. This means an average accuracy of 1 cm for a GSD of 5 cm, which is very good.

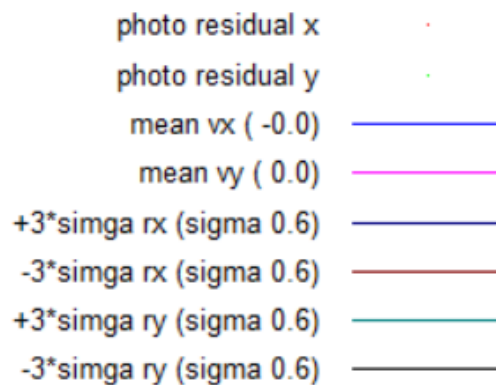
The residuals of the automatic point matching are also very close to each other in the two cases, based on the histogram no visual difference can be established.



16. Figure Photo residual histogram plot for automatic measured points for Total Station

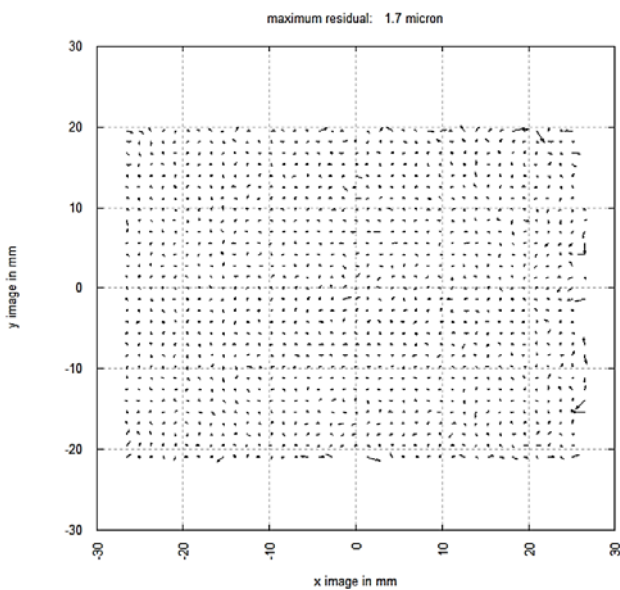


17. Figure Photo residual histogram plot for automatic measured points for Laser Scanner

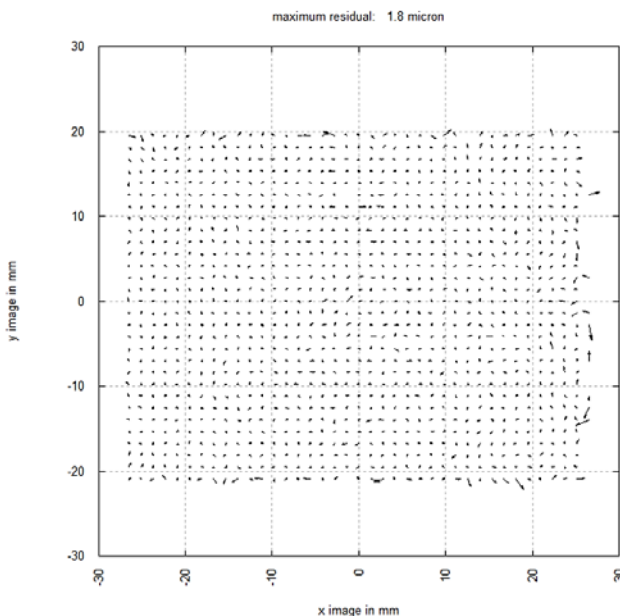


18. Figure Legend for the two histogram plots above

There is a difference of 0.1 micron in the residuals of the camera model in favor of the total station determination method (Fig. 19-20), which is also negligible.



19. Figure Image residual plot for automatic measured points for Total Station



20. Figure Image residual plot for automatic measured points for Laser Scanner

By analyzing the numerical results on Table 4, there are thousandths of micron differences in favor of the laser-scanned method, which can be explained by the fact that the GCP designation was refined on that data set.

		Total Station	Laser Scanner
sigma θ in micron		0.8	0.8
number block points		8112	8161
number of images		187	187
number of iterations		12	12
redundancy		33859	34036
mean std dev ground	x	0.009 [m]	0.009 [m]
	y	0.009 [m]	0.009 [m]
	z	0.026 [m]	0.026 [m]
mean std dev ori	omega	1.6 [mdeg]	1.6 [mdeg]
	phi	1.6 [mdeg]	1.6 [mdeg]
	kapp	0.7 [mdeg]	0.7 [mdeg]
mean std dev ori	x	0.019 [m]	0.019 [m]
	y	0.019 [m]	0.019 [m]
	z	0.014 [m]	0.014 [m]
rms image points	x	0.594 [micron]	0.590 [micron]
	y	0.603 [micron]	0.597 [micron]
rms control in image	x	1.683 [micron]	1.675 [micron]
	y	2.176 [micron]	2.137 [micron]
rms control in terr	x	0.008 [m]	0.008 [m]
	y	0.007 [m]	0.008 [m]
	z	0.013 [m]	0.013 [m]
max res. control	x	0.015 [m]	0.015 [m]
	y	-0.015 [m]	-0.015 [m]
	z	-0.036 [m]	-0.036 [m]

4. Table Comparison of the AT results of the two measuring methods

VI. SUMMARY, POSSIBLE DEVELOPMENTS

The lever arm measurement is traditionally done with a total station, but the question arose as to whether it could be an advantage if it were done with a laser scanner procedure. The advantage of the latter is that we don't have to bother with identifying the points in the field, this step can be solved in office conditions on the ready-made point cloud. The list of advantages ends here, but there are several important aspects among its disadvantages:

1. the planning of the measurement requires even more attention and therefore time,
2. the measurement with maximum accuracy settings takes longer than with a measuring station, even with 5 rounds
3. it is not enough to perform the measurement from one station, at least one more position with a view of the top of the plane from a high elevation is necessary
4. if there are several stations, in order to connect them, then the placement of the targets, requires special planning and time
5. target measurement is required at each station, which consumes additional time
6. since even with careful planning, there can be important points that cannot be seen, so additional manual measurements with a ruler or a measuring station may be necessary (which calls into question the viability of the whole method in terms of simplicity)
7. for using several stations, alignment inaccuracy affects the final accuracy
8. during processing, the identification of points (especially those of devices with low reflectance) and

the construction of separately measured parts are time-consuming

9. if there is no reference measurement, it is easy to make (blunder) mistakes

Based on one pair of measurements, it cannot be stated that laser scanning has the right to exist in the topic of lever arm measurement, but neither can the opposite. A new pair of measurements is definitely necessary, preferably with a new reference area for aerial triangulation. With the new measurement, it would be worthwhile to try to perform scanning from only one position, and by placing reflectors on the antennas and internal hidden points, we would "see" it, thus eliminating a significant part of the listed disadvantages (3-8). With this method, we can perhaps get a more complete picture of the applicability of laser scanner measurement for this purpose.

VII. REFERENCES

- [1] János Katona: Parcel analysis for the general application of remote sensing monitoring, In: AIS 2019 : 14th International Symposium on Applied Informatics and Related Areas organized in the frame of Hungarian Science Festival 2019 by Óbuda University, Székesfehérvár: Óbudai Egyetem, pp 69-70 (2019)
- [2] Lucas Grégory; Halász László; Solymosi József. Exploring the capacities of airborne technology for the disaster assessment, HADMÉRNÖK 8 : 3 pp. 74-91. , 18 p. (2013)
- [3] Lucas Grégory, Considering time in orthophotography production: from a general workflow to a shortened workflow for a faster disaster response, International Archives of Photogrammetry and Remote Sensing (2002-) XL3 : W3 pp. 249-255. , 7 p. (2015)
- [4] Schickler, W., and A. Thorpe, 1998. Operational procedure for automatic true orthophoto generation, International Archives of Photogrammetry and Remote Sensing, 32(Part 4):527-532.
- [5] Balázsik V., Tóth Z., Abdurahmanov I. Analysis of Data Acquisition Accuracy with UAV INTERNATIONAL JOURNAL OF GEOINFORMATICS (1686-6576): 17 1 pp 1-10 (2021)
- [6] Kugler, Zsófia ; Tóth, Zoltán ; Szalay, Zsuzsa ; Szagri, Dóra ; Barsi, Árpád Supporting microclimate modelling with 3D UAS data acquisition IDŐJÁRÁS / QUARTERLY JOURNAL OF THE HUNGARIAN METEOROLOGICAL SERVICE 123 : 3 pp. 279-294. , 16 p. (2019)
- [7] A. Ládai, C. Toth, Z. Tóth INDOOR MAPPING WITH AN OMNIDIRECTIONAL CAMERA SYSTEM: PERFORMANCE ANALYSIS INTERNATIONAL ARCHIVES OF PHOTOGRAMMETRY AND REMOTE SENSING (2002-) XLIII-B1-2022 pp. 347-352. , 6 p. (2022)

Design of Cemetery Cadastre of Saint Vince Home of the Sisters of Charity in Csákvár

Regina Hanna BALATON

Institute of Geoinformatics
Óbuda University, Alba Regia
Technical Faculty
Székesfehérvár, HUNGARY
reginahannabalaton@gmail.com

Andrea PÓDÖR

Institute of Geoinformatics
Óbuda University, Alba Regia
Technical Faculty
Székesfehérvár, HUNGARY
andrea.podor@amk.uni-obuda

Abstract— In this article we describe the survey of a cemetery cadastre and the construction of grave site database. The problems encountered in the creation of the database are discussed and suggestions are made for the database structure. In creating the database, we have tried to implement a structure that best fits the existing register.

The plan was to measure, map and create of a geospatial database of the cemetery plots of the Beloved Sisters of the Saint Vince Home in Csákvár. Attaching the data belonging to the corresponding individuals in the geospatial database and assigning the corresponding images. During the implementation of the task, I considered the requests of the home and relatives.

The basis of this article was my thesis, the topic of which was to create a digital cemetery cadaster, with the preparation of which I measure and map the cemetery plot of the Dear Sisters of the Saint Vince Home in Csákvár, with which I intend to pay tribute to the memory of the deceased sisters, and to provide a complex register for their living relatives, with the help of which the reconstruction of the plot may later become a simpler process.

During the survey of the cemetery, UAV drone was used for survey, and, in some places, a measuring tape was used due to the vicissitudes of the vegetation. During the implementation, my goal was to be able to provide a more accurate registration picture from the existing data thanks to GIS.

Keywords—cadaster, GIS, database

I. INTRODUCTION

As an introduction to the task, we would like to say a few words about cemeteries and registration forms in Hungary.

Humanity has been preoccupied with the relationship between life and death since the beginning, which is why burials played a significant role already in prehistoric times, which later underwent further significant developments until today. Unfortunately, burials in mass graves occurred countless times during various battles, wars, and epidemics, during which the identification of the deceased later caused significant problems.

The cemetery as a concept is defined as the place where the deceased can find eternal rest. In Hungary, two types are

distinguished: public cemeteries and those operated by the church.

Nowadays, thanks to developing technologies, a digital version of the register of cemeteries is introduced in more and more places in the form of cadaster management. If the existing data is transferred in this case, a lot of new opportunities will open for the drivers and their relatives. One such example is the fact that the identification of graves can be tracked locally and spatially, which is a huge advantage in the case of a larger cemetery. The details of the relatives will also become easier to follow, if, for example, a grave becomes obsolete, or any other problems arise in connection with it.

The relationship between geospatial informatics and records can already be linked since the appearance of the concept, since GIS software has been dealing with the management of large amounts of information into quality data from the very beginning (Dobesova, 2016, Balla et al, 2020).

This connection between the existing cemetery registration system to be implemented and the GIS software can significantly simplify the handling and management of the data, and additional functions can be introduced easily as a result (Márkus, 2009).

The paper-based register data, on which the register system can be built, in many cases complicates the processing process, because in many cases the church registers and the cemetery versions contain significant differences, and there are also many typos in them.

II. THE SURVEY

A. UAV drone measurement

The beginning, the survey was conducted on October 22, 2021. We have determined the location of the ground control points (GCP) in the field with the Leica GS18 T RTK GPS. For GCPs we used neon orange paint “X” marks so they could be clearly visible on the orthophoto. During the measurement, we used RTK (Real Time Kinematic) continuous, real-time positioning method from among the GNSS (Global Navigation Satellite System) measurement types to measure the exact location of the connection points with the aforementioned instrument (Duleba et al, 2021).

The survey was carried out in leafless vegetation conditions. We have used DJI Phantom 4 drone to take the aerial photographs and perform the flight. Before the start of the flight, we went around the affected area, and then we determined the position of the control points on the ground as described above, because they are essential for the production of the orthophoto from the aerial photos during processing.

To take the aerial photos, we set the height of the flight to 30 meters due to the tree cover of the location, and to execute the flight, we used a predetermined flight plan.

The first Figure shows the image of the flight plan during the survey process, on which the flight data can also be extracted.



1. Figure - An image of the flight plan during the process

B. Additional measurement

The second option to fix our measure is use another measure type. During the preliminary processing, it became obvious that there are points that are not visible on the orthophoto and cannot be saved from the point cloud either, because they are overgrown by evergreen vegetation. Due to the problem, we have decided that an additional measurement would be necessary, but since only a few points were missing to create the register in full, we have decided to perform an orthogonal measurement in relation to an existing point taken from an orthophoto to determine the position of the points. The supplementary measurement was carried out on March 9, 2022.

The crosses were located entirely in one row, so the ordinate values were zero for all points. During the measurement, we have kept measurement notes on the affected areas.

We took photographs on the day of the survey, which appear as hyperlinks in the database. The order of the images and thus the numbering of the crosses were also numbered according to the direction of the flight plan. Photographs were attached to the database attribute table as displayed hyperlinks and additional descriptive data for the crosses.

For the photographing a Huawei P30 Lite smartphone was used. we did the serial numbering of the photographs on the spot in order to speed up the subsequent work process.

III. PROCESSING

The process of office processing consisted of several steps. The processing began with the selection of the mass of analog data collected by the staff of the Loving Home.

After the successful systematization, we have selected the most suitable software to complete the task.

After all this, the orthophoto was created from the aerial photos, which served as the graphical basis of the database. Once this was completed, the construction of the database began, followed by the replacement of the missing, difficult-to-extract parts that emerged later from the additional measurements.

A. Data selection

The data was sent to us broken down into employee orders of the Szent Vince Otthon of the Catholic Charity Service in Csákvár, which included a total of 19 orders, which were as follows:

- Civilians
- Congregation of the Daughter of the Divine Redeemer
- Daughters of Charity of St. Vincent de Paul Sisters of Mercy
- Kalocsai School Sisters
- Society of Sisters of Christ the King
- Anunciata Sisters
- Sisters of the Poor named after Our Lady, Szeged
- Orsolyaite Sisters of the Order of St. Orsolya
- Franciscan Sisters of the Daughter of St. Francis of Assisi
- Hungarian Congregation of Servite Sisters
- Sisters of the Cross
- Congregation of the Sisters of St. Dominic, Dominican Sisters
- Premontre Women's Order of Canons
- Congregation of Our Lady of Angers, Named for the Love of the Good Shepherd Good Shepherd Sisters
- Sisters of Mercy of Sztalmári
- Order of Saint Elizabeth
- English Misses
- Eternal worshipper
- Society of Daughters of the Heart of Jesus

Over a long period of time, after several reviews, we have paired the data with the data set we have recorded in the cemetery, thus creating a table of attachable attribute data. During the selection and organization of the data, our work was significantly influenced by the fact that there were matching names and corresponding dates of death in the set. Another problem was that the missing plates could be clearly found in several places in the cemetery, which also caused data loss. One of the main problems was caused by name variations, which are also part of the attribute data table. This problem caused the greatest difficulties during the sorting process. The reason is that the names of the given persons

were recorded differently in the various registers at different periods of time.

B. Database building

We have designed and created the geodatabase in ArcMap program of the ArcGIS software package distributed by ESRI. As a first step, we have created a File Geodatabase, then the different layers within it. During the selection of the layers, we created a polygonal surface in which the parcels are demarcated per row, and then a point layer, which marks the center of the crosses. we assigned the various attribute data to the point interface, which were as follows:

- age
- parcel id
- line number
- serial number
- name
- passed away
- lived (x years)
- image

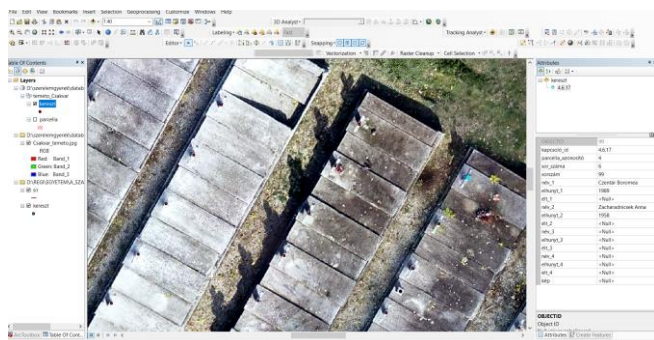
If several elements belong to a point, in those cases these columns are multiplied, usually one or two deceased people belong to a point, but there are cases when four.

Among the data types of the descriptive data, we used text type for the switch element, name, and image columns, using different lengths. We set the width of the four name and image columns to 50, and the switch element to 15. It was necessary to use the text data format in the id column because the numbers of the switch elements are separated by periods (for example, in the case of the first individual, the switch element is: 4.1.1), so text was the most appropriate choice for the program to handle this as well.

We used short integer data format in the parcel identifier, row number, row number, deceased and alive columns, because in these the descriptive data is just a single number.

In the names of the records of the attribute table, we made sure that they did not contain accents anywhere so that the program could handle them, so as a solution to the explanation, we filled in the alias column in every case for easier understanding.

The second figure illustrates the editing of the database, where you can view the filling of the attribute data while displaying the position of the points on the map.



2. Figure - Editing the location of crosses and filling in descriptive data

After that, I attached the Excel of the pre-selected, organized attribute data to the previously given table using the Relate command using the linking id.

Over a long period of time, after several reviews, I paired the data with the data set I recorded in the cemetery, thus creating a table of attachable attribute data. During the selection and organization of the data, my work was significantly influenced by the fact that there were matching names and corresponding dates of death in the set. Another problem was that the missing plates could be clearly found in several places in the cemetery, which caused data loss. One of the main problems was caused by name variations, which are also part of the descriptive data table. This problem caused the greatest difficulties during the sorting process. The reason is that the names of the given persons were recorded differently in the various registers at the given moments in time.

The elements of the table of attachable descriptive data are as follows:

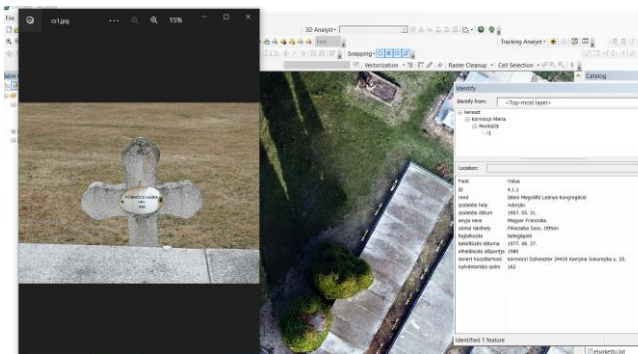
- ID
- name
- order
- Place of birth
- Date of birth
- His mother's name
- last place of residence
- occupation
- date of moving in
- date of death
- known relative
- registration number

IV. RESULT

The result of the project can be considered a fresh, up-to-date, correct database, easily transparent and editable at any time, with an expandable map base, which can provide an excellent basis for the relatives and the staff of the charity home for the reconstruction of the existing cemetery.

Thanks to the displayed images, the current state of the graves can be viewed, as well as information about the individuals found in the cemetery.

The third figure shows the result of the project. Explaining how the database stores the attribute data of the individuals and how the associated images are displayed.



3. Figure - The result of the project

V. CONCLUSION

Nowadays, digital cemetery cadastres designed, created, and managed in GIS software contain a lot of advantages and opportunities for users, and they can also be implemented in several designs to meet the desired needs.

Databases manage data easily and more transparently than their paper-based predecessors, in addition they offer many options to simplify the register, they can provide an up-to-date picture of the different statutes of limitations, however, if they become available to the living relatives, then finding the graves in the cemetery it also becomes easier for them.

Nowadays, a web page can easily be built from the created database after saving the appropriate file format, which makes it available to anyone to track the status of the final resting place of their deceased relative.

ACKNOWLEDGMENT

We would like to say a few words of thanks to all the people who provided any kind of help so that this project could be realized, thank you for all the mental and physical support, all the kind words and deeds!

First of all, we owe special thanks to my external consultant, Máté Lehoczky, who provided me with his own "instrument park" during the measurement, and also provided a lot of support with my questions regarding the production of the point cloud and orthophoto.

REFERENCES

- [1] Balla, D., Zichar, M., Tóth, R., Kiss, E., Karancsi, G., & Mester, T. (2020). Geovisualization techniques of spatial environmental data using different visualization tools. *Applied Sciences*, 10(19), 6701. J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [2] Dobesova, Z. (2016). Teaching database systems using a practical example. *Earth Science Informatics*, 9(2), 215-224. K. Elissa, "Title of paper if known," unpublished.
- [3] Duleba E ; László, G. Tóth, Z.: Establishment of GNSS calibration network. In: Csuka, Ildikó; Simon, Gyula (szerk.) AIS 2021-16th International Symposium on Applied Informatics and Related Areas – Proceedings Székesfehérvár, Magyarország : Óbudai Egyetem (2021) pp. 93-96. , 4 p. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, "Electron spectroscopy studies on magneto-optical media and plastic substrate interface," *IEEE Transl. J. Magn. Japan*, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].
- [4] Márkus, B. (2009). *Térinformatika*, NyME GEO jegyzet.

Conversion of Laplace equation from the spherical coordinate description to Cartesian coordinate system

Gibárt Gilányi

Óbuda University Doctoral School of Applied Informatics and
Applied Mathematics
Budapest, Hungary
gilanyi.gibart@gmail.com

Gábor Molnár

Óbuda University, Alba Regia Technical Faculty
Institute of Geoinformatics
Székesfehérvár, Hungary
molnar.gabor@amk.uni-obuda.hu

Abstract—

In this article, we show how the solution of the Laplace's equation in spherical coordinate and its coefficients can be rewritten in Cartesian coordinates, and we show the necessary normalization of these coefficients to ensure the identity of these solutions with the pure Cartesian solution.

The solution of the Laplace equation in spherical coordinates is obtained by separating the longitude, colatitude and radius variables. An elementary solution – a basis function - is the product of the negative integer exponent of the radius coordinate, the associated Legendre function of the colatitude coordinate, and the harmonic function of the longitude coordinate.

Using symbolic computations in MATLAB®, we generated these functions. The associated Legendre functions were created with the Rodrigues formula. Within the symbolic module we were able to simplify the functions with trigonometric transformations and convert the spherical coordinates into Cartesian coordinates.

As a result, the obtained formulas that could be directly compared to the pure Cartesian solution based on the 3D Taylor series, and this enabled us to calculate the normalization coefficients. The coefficients we obtained were the Gauss normalization coefficients, instead of Schmidt normalization coefficients, regularly used in geoid calculations.

Keywords — Laplace-equation, Cartesian coordinates, Legendre equation, Rodrigues formula, Matlab® symbolic tool

I. INTRODUCTION

The Laplace equation is a partial differential equation that is used to describe many phenomena in physics.

Laplace considered Newton's relation to gravity, the power law of the form $\sim 1/r^2$, which states that the gravitational acceleration produced by a mass placed at the origin of a coordinate system at a given point in space can be described as a vector whose direction from the given point is toward the origin of the coordinate system, and whose magnitude is proportional to the reciprocal of the square of the distance of the point from the origin.

Laplace, in considering this vector, found that the vector at any point in space can be produced as a vector formed from the partial derivatives of a scalar function with respect to the Cartesian coordinates. He also made another observation: the sum of the second derivatives of this scalar function (with respect to Cartesian coordinates) is zero at any points where

there are no masses. The expression of this last relation is the Laplace-Poisson differential equation.

$$\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2} + \frac{\partial^2 U}{\partial z^2} = 0$$

The Laplace equation helps to describe many phenomena in physics. The gravity example mentioned above is valid not only for a single point of mass, but the basis for describing the gravitational field of bodies with complex density distributions, such as the Earth or the planets of the Solar System. Laplace's equation can also be used to describe the Earth's magnetic field, since a magnetic scalar potential can be defined, the gradient of which gives rise to the magnetic induction field. In classical physics, the Laplace equation is used to describe static electric fields [1].

II. LAPLACE'S EQUATION IN SPHERICAL COORDINATES

In the description of Laplace's equation in spherical coordinates, the coordinates are the radius (r - the distance from the origin), the polar angle (ϑ - colatitude) and the longitude (λ - azimuth) angle. The equation takes the following form:

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial U}{\partial r} \right) + \frac{1}{r^2 \sin \vartheta} \left(\sin \vartheta \frac{\partial U}{\partial \vartheta} \right) + \frac{1}{r^2 \sin^2 \vartheta} \left(\frac{\partial^2 U}{\partial \lambda^2} \right) = 0$$

This is solved by the method of separation of variables, (Fourier method) i.e., a solution is sought which is the product of a r -dependent function $R(r)$, a ϑ -dependent function $P(\vartheta)$ and a λ -dependent function $\Lambda(\lambda)$. Substituting back and performing the necessary transformations, we obtain the following conditions for each function:

The form of the function $R(r)$ must be either r^2 or r^{-n-1} , where n is an integer.

The function $\Lambda(\lambda)$ must be either $\cos(m \cdot \lambda)$ or $\sin(m \cdot \lambda)$, and since the function must be periodic in 2π , m is also an integer. From these two conditions it follows that the function $P(\vartheta)$ must be given by the following, so called Legendre's differential equation:

$$\frac{1}{\sin \vartheta} \frac{\partial}{\partial \vartheta} \left(\sin \vartheta \frac{\partial P(\vartheta)}{\partial \vartheta} \right) + \left[n \cdot (n + 1) - \frac{m^2}{\sin^2 \vartheta} \right] P = 0$$

Its solution involves two integers. This occurred to Schrödinger, who was able to relate it to the concept of electron shells, where n is the principal quantum number and m is the azimuthal quantum number. In this way, the solutions

of the Laplace equation with spherical coordinates are also solutions of the time-independent Schrödinger equation [2].

The solution of the Legendre differential equation is called a Legendre polynomial for given n and $m = 0$, and an associated Legendre function, for $m \neq 0$, which can be obtained by the Rodriguez formula:

$$P_{n,m} = \frac{(1 - \cos^2 \vartheta)^{m/2}}{2^n \cdot n!} \cdot \frac{\partial^{n+m} (\cos^2 \vartheta - 1)^n}{\partial (\cos \vartheta)^{n+m}}$$

The consequence of the derivation in the formula shows that non-zero elements are obtained only for $m \leq n$ [3].

III. SCHMIDT NORMALIZATION AND QUASI-NORMALIZATION

Schmidt proved in an article [4] about geomagnetism, that the Legendre polynomials besides being orthogonal should also be normalized. The Schmidt quasi-normalisation factors can be calculated as follows:

$$\begin{array}{ll} \text{if } m \neq 0 & \text{if } m = 0 \\ k_n^m = \sqrt{2 \frac{(n-m)!}{(n+m)!}} & k_n^0 = \sqrt{\frac{(n-m)!}{(n+m)!}} \end{array}$$

The Schmidt quasi-normalization factors are not completely normalized harmonics. This condition is satisfied if we multiply the normalizing factor by the square of the Legendre function and then take its integral over the range [-1 1] to obtain 1. This can be obtained in spherical coordinates as follows.

$$k_n^m = 4\pi / \int_0^{2\pi} \int_0^\pi P_n^m \sin \vartheta \, d\vartheta \, d\lambda$$

That means, the completed Schmidt normalization is:

$$\begin{array}{ll} \text{if } m \neq 0 & \text{if } m = 0 \\ k_n^m = \sqrt{(2n+1) * 2 \frac{(n-m)!}{(n+m)!}} & k_n^0 = \sqrt{(2n+1) * \frac{(n-m)!}{(n+m)!}} \end{array}$$

IV. LAPLACE'S EQUATION IN CARTESIAN COORDINATES

To get the potential of a mass distributed around the origin of the coordinate system, a 3D Taylor series approximation can be applied. The potential is the infinite sum of n^{th} order multipole sources multiplied by the appropriate basis functions, which are the n^{th} partial derivatives of the $1/r$ function. If only one unit mass point is considered, whose coordinates are x, y and z , the n^{th} order member of the series is the following:

$$U^{(n)} = (-1)^n \sum_{i+j+k=n} x^i y^j z^k \frac{1}{n!} \cdot \frac{\partial^n}{\partial X^i \partial Y^j \partial Z^k} \left(\frac{1}{\sqrt{X^2 + Y^2 + Z^2}} \right)$$

Where X, Y, Z are the locations where we want to find the potential of the mass [5].

V. METHODS

To be able to compare the solution of Laplace equation in spherical and Cartesian coordinates, we transform the spherical coordinates to Cartesian coordinates. By using the appropriate trigonometric functions and ordering the equations, both the coefficients and the basis functions can be transformed. The transformation between the Cartesian and spherical coordinates is described by the following equations:

$$x = r * \sin(\vartheta) * \cos(\lambda)$$

$$y = r * \sin(\vartheta) * \sin(\lambda)$$

$$z = r * \cos(\vartheta)$$

in addition:

$$r = (x^2 + y^2 + z^2)^{0.5}$$

A. Theoretical solution

The solution of Laplace's equation with spherical coordinates is described by the following equation:

$$U(r, \vartheta, \lambda) =$$

$$\sum_{n=0}^{\infty} \sum_{m=0}^n r^{-(n+1)} [C_{n,m} \cos(m\lambda) + S_{n,m} \sin(m\lambda)] P_{n,m}(\vartheta)$$

The coefficients $C_{n,m}$ and $S_{n,m}$ contain the coordinates of the source of the potential. The coefficients can be calculated by the following method:

$$C_{n,m} = r^{(n)} * \cos(m\lambda') * P_{n,m}(\vartheta')$$

$$S_{n,m} = r^{(n)} * \sin(m\lambda') * P_{n,m}(\vartheta')$$

where r, ϑ' and λ' are the spherical coordinates of the unit mass. The following operations are recommended to convert the solutions of coefficients and basis functions to the Cartesian coordinate system:

- In the case where m is greater than 1, it is important to convert the trigonometric functions of multiple angles into single angles.
- Solve the following equation $1 - \cos^2(\vartheta')^{(k/2)} = \sin^k(\vartheta')$, and then convert every $r * \sin(\vartheta') * \cos(\lambda')$ and $r * \cos(\vartheta') * \cos(\lambda')$ expressions into x and y coordinates.
- After calculating all possible x and y coordinates, the $\sin^k(\vartheta')$ should be converted into $1 - \cos^2(\vartheta')^{(k/2)}$. Then it is possible to convert every $r * \cos(\vartheta')$ into z coordinate.
- Finally, the remaining r 's in the equation must be converted to $(x^2 + y^2 + z^2)^{0.5}$

B. Solutions using Matlab©

The calculations were done within Matlab© R2022a software. Within the symbolic module, the expressions mentioned in the previous subsection can be obtained.

Within the symbolic module we used the following commands:

- `syms`: create symbolic variables
- `simplifyFraction`: Simplify symbolic rational expressions. The `simplifyFraction` command is more useful than the `simplify` command in most cases. The `simplify` command created unfavorable trigonometric transformations, like multiple angles.
- `subs`: Symbolic substitution. We use this command to perform the converting operations.
- `expand`: Expand expressions and simplify inputs of functions by using identities [6].

First, we had to express the Rodrigues formula:

$$\begin{aligned}
 f &= (\text{costheta}^2 - 1)^n; \\
 f2 &= \text{diff}(f, \text{costheta}, n + m) \\
 f2 &= \text{subs}(f2, \text{costheta}, \cos(\vartheta)); \\
 P(n, m) &= (-1)^m * \left(\frac{(1 \cos^2(\vartheta))^{\frac{m}{2}}}{2^n * \text{factorial}(n)} \right) * f2;
 \end{aligned}$$

As mentioned before, besides the basis functions, the $C_{n,m}$ and $S_{n,m}$ coefficients also must be obtained. To convert to cartesian coordinates, according to the previous chapters, we performed the following operations:

$$\begin{aligned}
 C_{n,m} &= \text{subs}(C_{n,m}, (\sin^2(\vartheta))^{\frac{m}{2}}, (\sin(\vartheta))^m); \\
 C_{n,m} &= \left(\text{subs}(C_{n,m}, [(r * \sin(\vartheta) * \cos(\lambda))^{\frac{m}{2}}, (r * \sin(\vartheta) * \sin(\lambda))^{\frac{m}{2}}], [x^m, y^m]) \right) \\
 C_{n,m} &= \text{subs}(C_{n,m}, \sin^{2*k}(\vartheta), (1 - \cos^{2*k}(\vartheta))); \\
 C_{n,m} &= \left(\text{subs}(C_{n,m}, [(r * \cos(\vartheta))^{\frac{m}{2}}], [z^m]) \right) \\
 C_{n,m} &= \text{subs}(C_{n,m}, r, (x^2 + y^2 + z^2)^{0.5})
 \end{aligned}$$

VI. RESULTS

Table 1. shows some of the $C_{n,m}$ and $S_{n,m}$ coefficients expressed with spherical and Cartesian coordinates.

Coeffs	Spherical coordinates	Cartesian coordinates
$C_{1,0}$	$r * \cos(\vartheta)$	z
$C_{1,1}$	$-r * \cos(\lambda) * \sin(\vartheta)$	$-x$
$S_{1,1}$	$-r * \sin(\lambda) * \sin(\vartheta)$	$-y$
$C_{2,0}$	$\frac{3 * r^2 * \cos^2(\vartheta)}{2} - \frac{r^2}{2}$	$-\frac{x^2}{2} - \frac{y^2}{2} + z^2$
$C_{2,1}$	$-3 * r^2 * \cos(\lambda) * \cos(\vartheta) * \sin(\vartheta)$	$-3 * x * z$
$S_{2,1}$	$-3 * r^2 * \cos(\vartheta) * \sin(\lambda) * \sin(\vartheta)$	$-3 * y * z$
$C_{2,2}$	$3 * r^2 * \cos^2(\lambda) * \sin^2(\vartheta) - 3 * r^2 * \sin^2(\lambda) * \sin^2(\vartheta)$	$3 * (x^2 - y^2)$
$S_{2,2}$	$6 * r^2 * \cos(\lambda) * \sin(\lambda) * \sin^2(\vartheta)$	$6 * x * y$

Table 1. $C_{n,m}$ and $S_{n,m}$ coefficients expressed in spherical and Cartesian coordinates

Substituting back the coefficient to the potential formula we obtained the following equation:

$$\begin{aligned}
 U(R, \theta, \Lambda) &= \\
 \sum_{n=0}^{\infty} \sum_{m=0}^n & R^{-(n+1)} [C_{n,m} \cos(m\Lambda)] P_{n,m}(\Lambda) + \\
 & R^{-(n+1)} [S_{n,m} \sin(m\Lambda)] P_{n,m}(\Lambda)
 \end{aligned}$$

The coefficient contains the coordinates of the mass point, i.e. the source, and R, θ, Λ are the locations where we want to get the potential of the mass.

Now we were able to transform any expression written in spherical coordinates into Cartesian coordinates.

For example, let's compare the second order term of the potential of a unit mass located in (x, y, z) point. From one side this can be expressed using the coefficient and basis functions according to Table 1, as this solution is based on the spherical coordinates formulation.

From other side this second order term can be expressed using the Taylor series approximation:

$$\begin{aligned}
 U^2 &= \sum_{i+j+k=2} x^i y^j z^k \cdot \frac{1}{2} \cdot \frac{\partial^2}{\partial X^i \partial Y^j \partial Z^k} \left(\frac{1}{R} \right) = \\
 &= \frac{\left(\begin{aligned} &2 * X^2 * X^2 - X^2 * Y^2 - X^2 * Z^2 + 6 * X * Y * X * Y \\ &+ 6 * X * Z * X * Z - Y^2 * X^2 + 2 * Y^2 * Y^2 - Y^2 * Z^2 \\ &+ 6 * Y * Z * Y * Z - Z^2 * X^2 - Z^2 * Y^2 + 2 * Z^2 * Z^2 \end{aligned} \right)}{2 * (X^2 + Y^2 + Z^2)^{\frac{5}{2}}}
 \end{aligned}$$

However, these two solutions should be identical, and so the solutions of the spherical coordinate solution should be normalized.

To ensure the identity of the two approaches, the spherical coordinate solutions should have been multiplied with the subsequent coefficients:

$$\begin{aligned}
 &\frac{k_2^{0^2} * (x^2 + y^2 - 2 * z^2) * (X^2 + Y^2 - 2 * Z^2)}{4} \\
 &+ 9 * k_2^{2^2} * (X^2 - Y^2) * (x^2 - y^2) + 36 * X * Y * k_2^2 * x * y \\
 &+ 9 * X * Z * k_2^{1^2} * x * z + 9 * Y * Z * k_2^1 * y * z
 \end{aligned}$$

where k_0, k_1 and k_2 are the normalizing factors, in this case:

$$k_2^0 = 1, \quad k_2^1 = \sqrt{\frac{1}{3}}, \quad k_2^2 = \sqrt{\frac{1}{12}}$$

These normalizing factors proved to be the Gaussian normalizing factors.

The advantage of the solution starting with the spherical coordinate description is that it requires less terms to be summed, $2n+1$ elements for a given order, while, if $n > 2$ the Cartesian coordinate solution requires more elements ($0.5n^2 + 1.5n + 1$) to be summed. Table 2. shows some examples, how many terms has to be summed for the n^{th} order solution.

n	number of added elements using the Taylor series based Cartesian coordinates	number of added elements using the spherical coordinate system
1	3	3
2	6	5
3	10	7
4	15	9
5	21	11
6	28	13
n	$0,5n^2 + 1,5n + 1$	$2(n + 1)$

Table 2. Number of the independent terms for a given order using the Taylor series based Cartesian coordinates and using the spherical coordinate system

As the chosen example for $n=2$ shows, the spherical method consists of $2n+1 = 5$ elements, while the Cartesian method consists of $0,5n^2 + 1,5n + 1 = 6$ independent elements.

We could also check the correctness of the quasi-normalization factor in Matlab© symbolic tools. The quasi-normalization solution was the same as the Taylors series based Cartesian solution when the equations were properly solved.

VII. CONCLUSIONS

We have shown in the Results (Table 2), that we could successfully express the coefficients using Cartesian coordinates. The expressions written in Cartesian coordinates are simpler and are easier to handle.

We got the result, that instead of the Schmidt normalization – widely used in gravity problems – the Gauss

normalization should be used to ensure the identity of the two approaches. To understand the consequences, further investigation is needed.

REFERENCES

- [1] Laplace, Pierre Simon: *Traité de mécanique céleste*. L'Imprimerie de Crapelet: Chez J.B.M. Duprat, Paris, 1827. <https://library.si.edu/digital-library/book/traite-de-mecanique-celeste>
- [2] LeGendre, Adrien-Marie: *Sur l'attraction des sphéroïdes homogènes*. Mémoires de Mathématiques et de Physique, présenté à l'Académie royale des sciences, **10**:411-435, Paris, 1785.
- [3] Rodrigues, Olinde: *De l'attraction des sphéroïdes*. Correspondence sur l'École Impériale Polytechnique, (Thesis for the Faculty of Science of the University of Paris), **3** (3): 361–385, Paris, 1816.
- [4] Schmidt, A., 1917. Erdmagnetismus, *Enzykl. Math. Wiss.*, **6**, 265–396.
- [5] Simonyi K.: *Elméleti villamosságtan*. Tankönyvkiadó, Budapest 1973, **6**, p. 118.
- [6] MATLAB© and Statistics Toolbox Release 2022a, The MathWorks, Inc., Natick, Massachusetts, United States.

Comparative Analysis of several contour line generation methods and softwares

G. Nagy*, Zs. Ungvári†

*Óbuda University, Alba Regia Technical Faculty, Institute of Geoinformatics,

† Eötvös Lóránd University, Faculty of Informatics, Institute of Cartography and Geoinformatics

nagy.gabor@amk.uni-obuda.hu

ORCID ID: 0000-0001-9453-2291

ungvarizs@map.elte.hu

ORCID ID: 0000-0001-6084-2195

Abstract—Contour lines are a traditional method of the visualization of terrain. The geographic information systems use digital elevation models, but the visualization of the terrain sometimes needs contour lines which are generated from the digital elevation model. This paper compare and analyze some methods and software which can generate contour line maps from a digital elevation model. The results are compared with other contour line generation methods and the classical contour line maps.

Index Terms—contour line, digital elevation model, geographic visualization

I. INTRODUCTION

The contour lines are a frequent method to show the terrain surface. In the past, the classical paper-based maps stored the knowledge about the terrain elevation in this way. Nowadays, we create contour line maps from the digital elevation models for visualizing the elevation.

In this paper we study some program and method, and analyze the created contour lines.

II. SOURCE DATA AND THE STUDY SITES

The study based on an 25 meter resolution digital elevation model, which covered the (400000, 0) and (1000000, 400000) positions in the hungarian national grid (EOV). The elevation model was made from SRTM[1], [2] (Shuttle Radar Topography Mission) one arcsec (about 30.9 and 21 meters in this latitude) resolution filled[3] dataset. This dataset contains total area of Hungary and some territory of the surrounding countries.

We chose twenty study sites in this dataset, in different topography regions. The Figure 1. shows the location of these study sites and the Table I. gives the corner coordinates and the countour line interval, which will be applied in the following studies. The study sites are not only from Hungary, some mountain area (1, 12, 18) are from neighbouring countries. Every study sites is an 5000x5000 meters square.

III. DATA PROCESSING

Contour lines are generated by different software and options. The created contour lines are stored in SHP (shape) files.

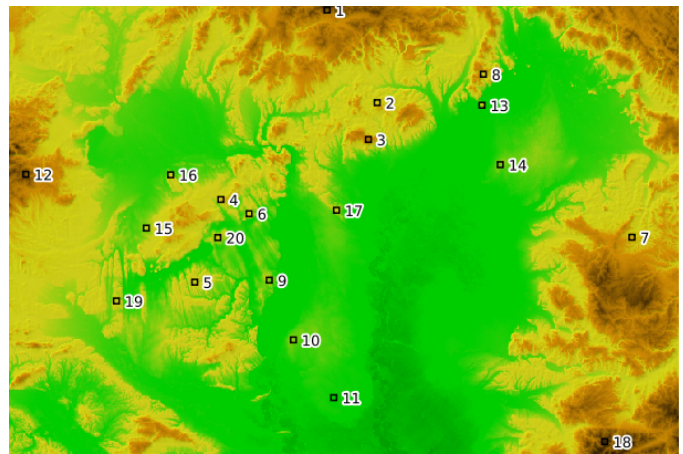


Figure 1. The study sites with the site number and the 5x5 km squares of these areas. The southwest corner is the (400000,0) and the northeast corner is the (1000000,400000) point in the Hungarian National Grid (EOV)

Table I
THE LOCATION OF THE STUDY SITES AND INTERVALS OF CONTOUR LINES

id	minY	minX	maxY	maxX	interval
1	682300.0	394500.0	687300.0	399500.0	50.0
2	727300.0	311500.0	732300.0	316500.0	20.0
3	719400.0	278600.0	724400.0	283600.0	25.0
4	587200.0	224800.0	592200.0	229800.0	10.0
5	563800.0	150800.0	568800.0	155800.0	5.0
6	612500.0	212100.0	617500.0	217100.0	10.0
7	955600.0	191000.0	960600.0	196000.0	10.0
8	822400.0	337000.0	827400.0	342000.0	20.0
9	630600.0	152600.0	635600.0	157600.0	5.0
10	652200.0	99100.0	657200.0	104100.0	5.0
11	688400.0	47400.0	693400.0	52400.0	5.0
12	412400.0	247400.0	417400.0	252400.0	50.0
13	821200.0	309200.0	826200.0	314200.0	20.0
14	837600.0	256000.0	842600.0	261000.0	5.0
15	520400.0	199200.0	525400.0	204200.0	20.0
16	542200.0	246900.0	547200.0	251900.0	10.0
17	690800.0	215200.0	695800.0	220200.0	10.0
18	931200.0	8000.0	936200.0	13000.0	50.0
19	493600.0	134000.0	498600.0	139000.0	10.0
20	584500.0	190700.0	589500.0	195700.0	5.0

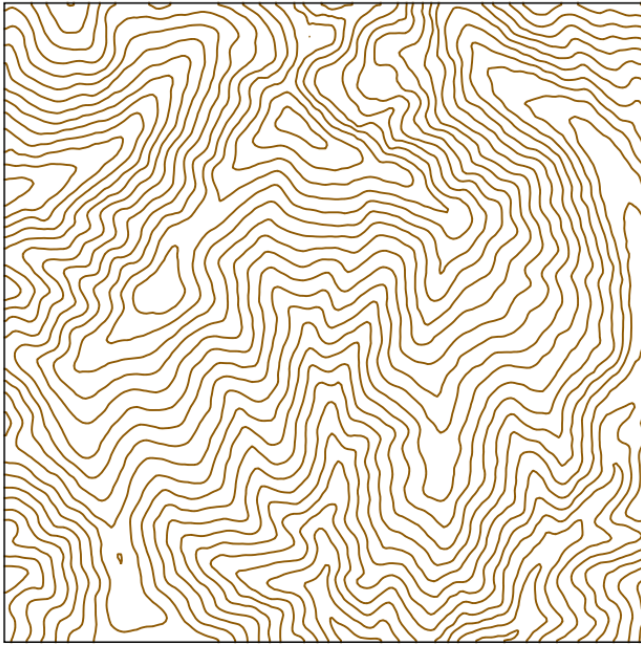


Figure 2. Contour lines in site 1, generated by GDAL `gdal_contour` program

The first set of contour lines is the `gdal_contour` program from the GDAL[4] (Geospatial Data Abstraction Library). This is an element of an open source toolkit, the program can be used in command line. Some other program (for example QGIS) use this tool for creating contour lines from an elevation model. (An example in the site 1 is in the Figure 2.)

We also have created countour lines with Global Mapper[5]. The second set of contour lines was created with default setting. (An example in the site 2 is in the Figure 3.)

The third set of countour lines was created with the the smooth option. This option provides curved countour lines. (An example in the site 7 is in the Figure 4.)

IV. COMPARATIVE ANALYSIS

We can see some difference between the contour line sets. The difference is smaller in the mountain areas, for example in Figure 6., and bigger in plains, for example in Figure 8.

The hills (see in Figure5.) are between in the mountains and plain areas. The Figure 7. presents different regions in one view (mountains is east part, and areas with smaller slope angles in west).

We also compare the generated contour lines with classical topographic maps, the 1:10000 scale “civilian” maps, in the tile number 54-242. Figure 9. and Figure 10. show these comparisons.

The left sides of these figures are the classical topography maps (which contains more content than the contour lines), and the right sides are the `gdal_contour` generated contour lines.

The topography map made from an different elevation model, this may cause mostly the differences.

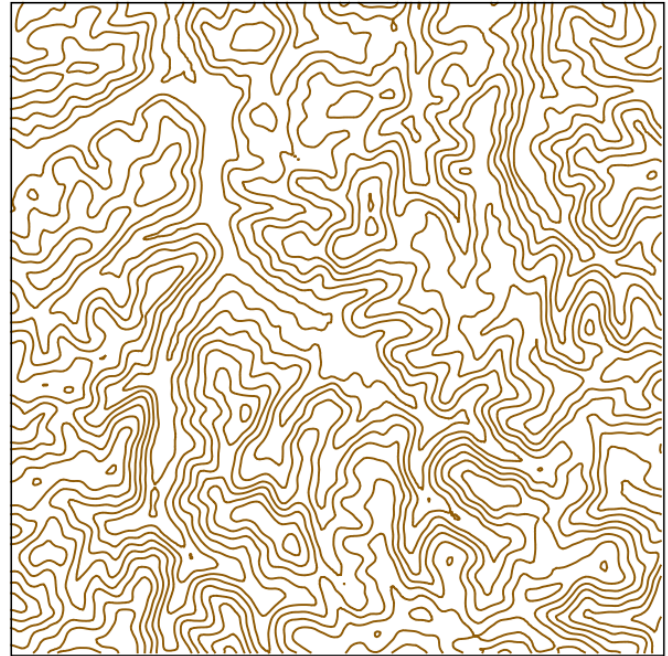


Figure 3. Contour lines in site 2, generated by Global Mapper program

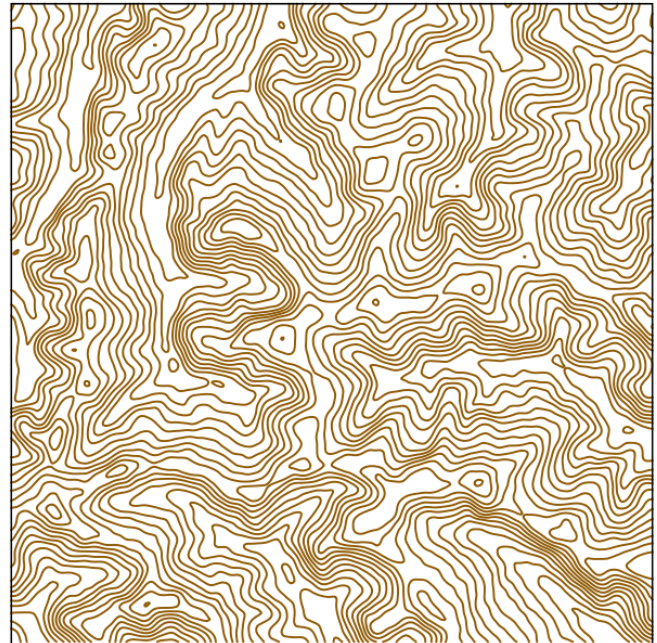


Figure 4. Contour lines in site 7, generated by Global mapper program with smooth option

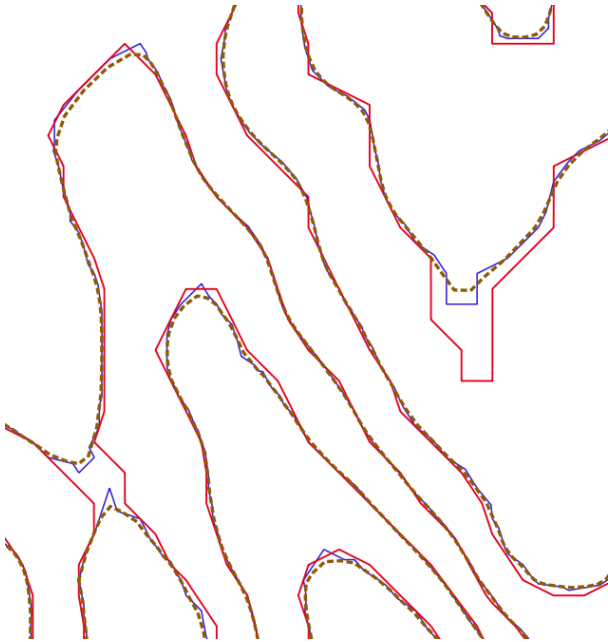


Figure 5. Different contour lines with 5 meter interval in a part of site 5 (approx 566500, 154100 EOJ coordinates). The red line is from the gdal_contour, the blue is from the Global Mapper. The brown dotted line is the smoothed version of the Global Mapper contours.

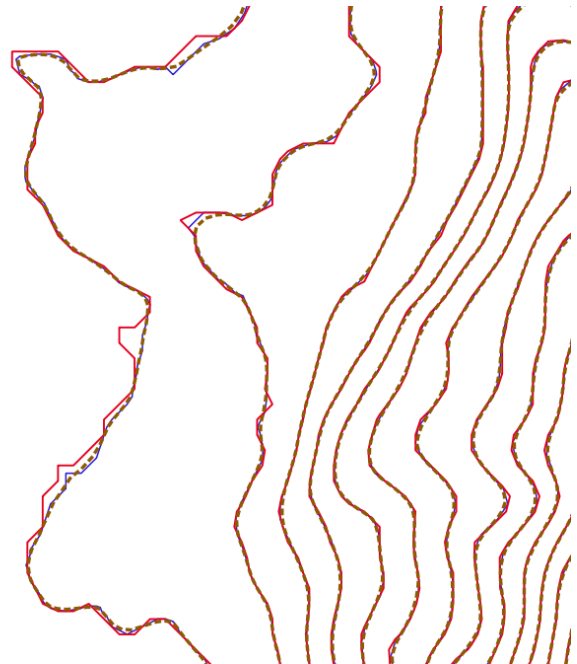


Figure 7. Different contour lines with 10 meter interval in a part of site 4 (approx 588200, 228600 EOJ coordinates). The red line is from the gdal_contour, the blue is from the Global Mapper. The brown dotted line is the smoothed version of the Global Mapper contours.

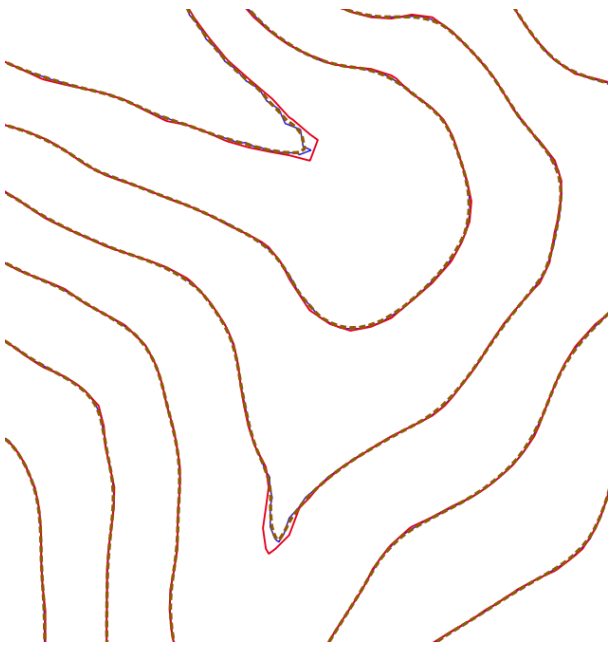


Figure 6. Different contour lines with 50 meter interval in a part of site 1 (approx 685800, 397600 EOJ coordinates). The red line is from the gdal_contour, the blue is from the Global Mapper. The brown dotted line is the smoothed version of the Global Mapper contours.

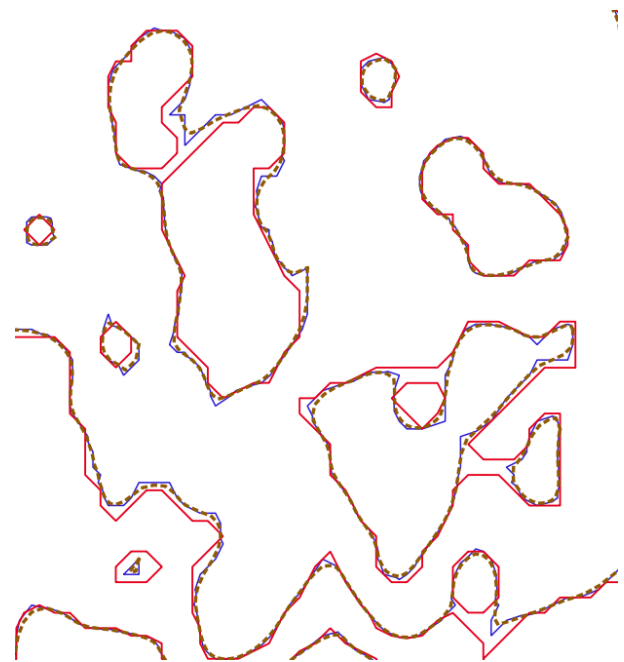


Figure 8. Different contour lines with 5 meter interval in a part of site 14 (approx 840500, 258300 EOJ coordinates). The red line is from the gdal_contour, the blue is from the Global Mapper. The brown dotted line is the smoothed version of the Global Mapper contours.

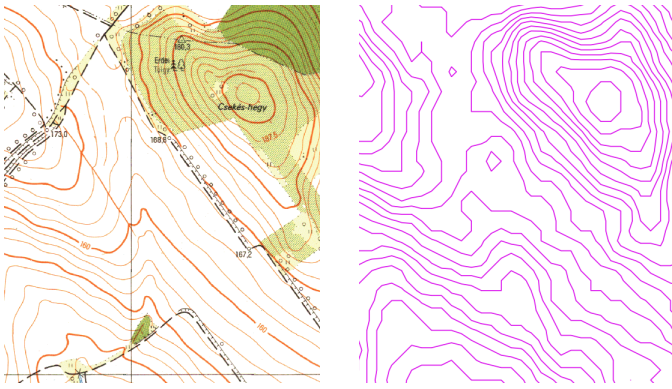


Figure 9. Contour lines in a topographic map (1:10000 scale), and generated by `gdal_contour` from STRM model. The interval of contour lines is 2.5 meters on both maps

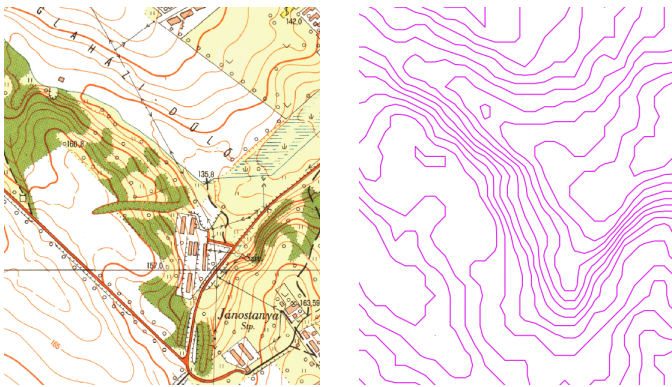


Figure 10. Contour lines in a topographic map (1:10000 scale), and generated by `gdal_contour` from STRM model. The interval of contour lines is 2.5 meters on both maps

V. CONCLUSION

The contour lines, which have been created by different programs or parameters, have some differences. The differences are larger in plain areas, the contour lines in mountain regions (larger slope angle) are very similar from different methods.

Smoothing options can create better look, because the classical contour lines (and the real terrain) is also smoothed.

The optimal settings of the contour line generation is depend from the terrain of the visualized area.

REFERENCES

- [1] J. J. Van Zyl, "The Shuttle Radar Topography Mission (SRTM): a breakthrough in remote sensing of topography," *Acta Astronautica*, vol. 48, no. 5, pp. 559–565, 2001.
- [2] E. Rodriguez, C. S. Morris, and J. E. Belz, "A global assessment of the SRTM performance," *Photogrammetric Engineering & Remote Sensing*, vol. 72, no. 3, pp. 249–260, 2006.
- [3] A. Jarvis, H. I. Reuter, A. Nelson, E. Guevara *et al.*, "Hole-filled SRTM for the globe Version 4," available from the *CGIAR-CSI SRTM 90m Database* (<http://srtm.csi.cgiar.org>), 2008.
- [4] F. Warmerdam, "The geospatial data abstraction library," in *Open source approaches in spatial data handling*. Springer, 2008, pp. 87–104.
- [5] D. B. Rahe, "Product review: Global mapper gis software," *Soil Survey Horizons*, vol. 49, no. 2, pp. 40–40, 2008.

POSTER SESSION IV.

Education, Enterprise and Economics

Entrepreneurial activity and ecosystem in Hungary

Ildikó Petőné Csuka
Alba Regia Technical Faculty Óbuda
University
Óbuda University
Székesfehérvár, Hungary
email: csuka.ildiko@amk.uni-obuda.hu

Abstract- The article gives an introduction into the world of entrepreneurships. The author analyses the changes in entrepreneurial activity in Hungary in the last two decades, mainly of late years and provides an assessment of the impacts of COVID-19 on level of entrepreneurship. Hungary's entrepreneurial activity is investigated comparing it to European countries. The article construes the local entrepreneurial ecosystem with international outlook. A further research will analyse the entrepreneurial attitudes and ambitions of students in different aspects.

Keywords—entrepreneurship, entrepreneurial ecosystem, entrepreneurial activity, start-up

I. INTRODUCTION

The author introduces us into the world of entrepreneurships in Hungary with analysing the Hungarian entrepreneurial ecosystem and attitudes in international comparison. The changes of the late years are construed in the Hungarian entrepreneurial activity and attitudes and the consequences and conclusion are drawn from the analysis. The article shows further research direction: into analysis of the entrepreneurial attitudes and inclination of students and into the education of entrepreneurial skills.

The systematic literature review methodology is applied during this research (Fig.1.). The methodology is:

- searching the available scientific bulletin and publication,
- monitoring the quality of the paper,
- choosing the acceptable ones,
- comparing the research results objectively,
- and drawing and formulating the results in publication unbiasedly.

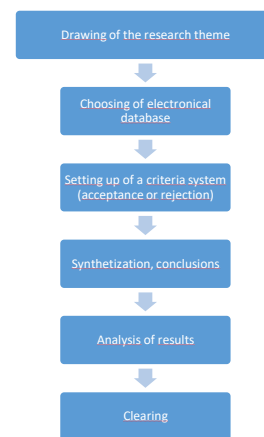


Fig. 1. Systematic literature review methodology

II. ANALYSIS OF ENTREPRENEURIAL ATTITUDES AND ACTIVITY

A. Measures of entrepreneurship

A lot of entrepreneurship-indexes have been created over the last two decades measuring the performance of the economy. The most important index is the connection between economic growth and entrepreneurship. Carey and Turik [8] defined four measure-indexes for entrepreneurships:

1. turbulence (the sum of entries and exit)
2. change of the size distribution of firms
3. number of market participants
4. self-employment.

The self-employment attitude researches were introduced by the OECD in the late 1990s. In 2006 OECD brought a steering group into being to initiate an entrepreneurship indicator. The generated entrepreneurship indicator includes [1]:

- *determinants of entrepreneurship*: regulation, R&D, entrepreneurial capabilities, culture, access to finance and market conditions,
- *entrepreneurial performance*: firms, employment, wealth,
- *impact of entrepreneurship*: employment, economic growth, poverty reduction.

Defining of the indicators were relatively easy but to compare cross-country data is more difficult.

Eurostat regularly gives reports on entry and exit statistics, business start-up and other entrepreneurship related data. These sources give information for policy makers on the quantity of entrepreneurship rather than its quality. These

data and measure indexes do not take into account the differences in environmental factors, the different efficiency and the level of institutional setup. In developed market economies the small and medium-sized enterprise sector is given more attention because it has been recognised that it plays a significant role in domestic production and in the mitigation of the unemployment problem due to its ability to soak up the workforce, and so due to its economic power it can potentially play a stabilising role in the economy [6]. In 1999 the Global Entrepreneurship Monitor (GEM) defined an entrepreneurship measure to compare entrepreneurial activity across countries.

B. GEM report

The analysis of entrepreneurial attitudes and activity is based on the GEM (Global Entrepreneurship Monitor) reports. This monitor is an international entrepreneurial research which has been taking for two decades and more than one hundred countries has been taking part in it. The research has got 4 primary purposes:

- analysis and comparison of the levels of attitudes, activities, capabilities and aspiration in different countries
- analysis of the entrepreneurial activity’s impact on economical growth
- factor analysis of promotional and impedimental elements
- identification of incentive and stimulating economic policy tools.

The analysis of entrepreneurial activity is not only an indirect purpose having an important role in generating value-added and new workplaces but also conduces to the growth of economy and society. The extent of positive impacts depends on the entrepreneurial ecosystem (Fig. 2).

Entrepreneurship			
Attitudes	Activity	Capabilities	Aspiration

Fig. 2. Factors of an entrepreneurship impacting on entrepreneurial ecosystem

C. Economic growth according to GEM

GEM measures the levels of entrepreneurial activity since 1999 and 120 economies have been taking part in the research. GEM judges the economic growth of different countries based on classification of WEF (World Economic Forum) and determines different improvement goals according to the maturity levels [10].

- **Factor-driven economies** (for example: Vietnam, Burkina) are least developed. The most impulsive force of these economies are the natural resources and low-qualified workforce. Workforce is mainly applied in agriculture and raw material extracting industrial sector.
- **Efficiency-driven economies** (for example: China, Egypt) are increasingly competitive, increased product quality and more-efficient production

processes, technological adaption and the increased market-size are typical characteristics.

- **Innovation-driven economies** (for example: Singapore, Sweden) are most developed. More knowledge-intensive sectors and the service sector are dominant. Business development and innovation support are necessary in these economies to growth.

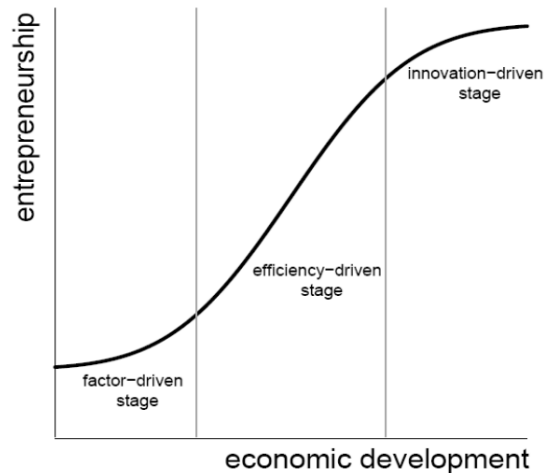


Fig. 3. Relationship of entrepreneurship and economic development [1]

Between entrepreneurship and economic development there is an S-shaped relationship according to Porter [7] (Fig.3.). For example if we move from the efficiency-driven stage to the innovation-driven stage entrepreneurship plays a more important role but it increases at a decreasing rate. Innovation is essential for economic development and innovation is generated by entrepreneurs.

TEA (Total Early-stage Entrepreneurial Activity) index (see explanation of TEA in chapter III) was measured in 2005 in terms of economic development. An U-shape phenomenon was figured out showing the relationship between the entrepreneurship and the development of a country [8]. Economies at a lower level of development have higher rate necessity entrepreneurs. The necessity-driven entrepreneurial activity decreases as economies move from the factor-driven stage to the efficiency-driven stage. In the innovation driven stage entrepreneurial activity increases and there is also a high rate of startups.

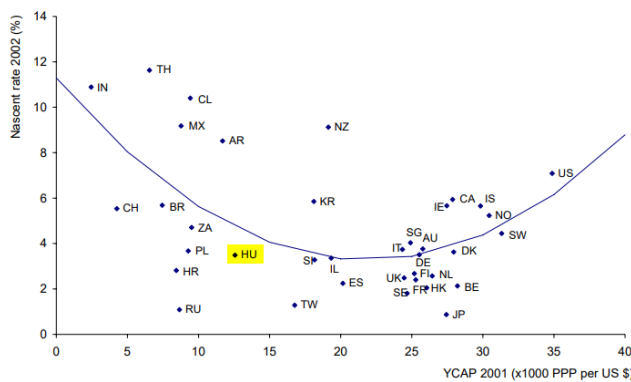


Fig. 4. Nascent entrepreneurship versus per capita income, the U-curve [8]

A low rate of entrepreneurship may also be a consequence of low economic growth levels, which also reflect higher levels of unemployment. Entrepreneurial opportunities are not just the result of the push effect of (the threat of) unemployment but also of the pull effect produced by a thriving economy as well as by entrepreneurial activities in the past. There is a strong positive correlation between entrepreneurship and economic growth [8] (Fig.4.).

The framework of GEM research is based on three data sources:

- on APS (Adult Population Survey) queried 2000 adults in each country taking part in the research. It is a standardized, representative primary research according to age and gender. APS surveys the attitudes, behaviours and expectations of individuals. APS informs on hidden economy, give information about unrecorded economic activities and jobs, about the individual decisions whether to start or continue a business.
- on NES (National Expert Survey): half structured interviews are made with at least 36 entrepreneurship-experts which provides an opportunity to assess the entrepreneurial context. This context influences the individual decision whether to start a new business and decisions to sustain or grow the business. The entrepreneurial context means access to finance, the education and skills of the population, and social support to entrepreneurship, the extent of bureaucracy and taxes and the development of infrastructure. The environment of entrepreneurship has a substantial impact on the likelihood to start and to grow a business.
- on national statistical data, official government statistics, such as business registrations, OECD reports and the other international data resources. National reports provide information on the annual research results.

III. GEM METHODOLOGY

The national environment of entrepreneurship is assessed by nine “Entrepreneurial Framework Conditions” (Fig. 5.). These conditions incorporate the ease of access to finance to social support for entrepreneurship, the access to

entrepreneurial education and the availability and cost of essential business services. The conditions are assessed in the NES by national experts. Figure 5. shows the GEM Conceptual Framework: the relationship between entrepreneurship and its local environment. The social, cultural, political and economic environment (entrepreneurial ecosystem) directly and indirectly influences entrepreneurship. The jobs, value-added and incomes generate entrepreneurial activity. The reverse is also true: if entrepreneurship of a country operates successfully that stimulates economic development, adds new value and creates jobs. The economy, politics and society have a profound impact on education, which basically determines entrepreneurship and the willingness of becoming an entrepreneur. At the same time, however, skills and competences also count.

Several EU-surveys show that 15-20% of secondary school students formerly involved in mini company programmes will establish an enterprise later. Students involved in entrepreneurial education will have developed business knowledge and sense, which means improving their basic skills and attitude including creativity, initiative, endurance, teamwork, understanding risks and developing a sense of responsibility. This is the entrepreneurial mindset that will make action from ideas and significantly improve employability [4].

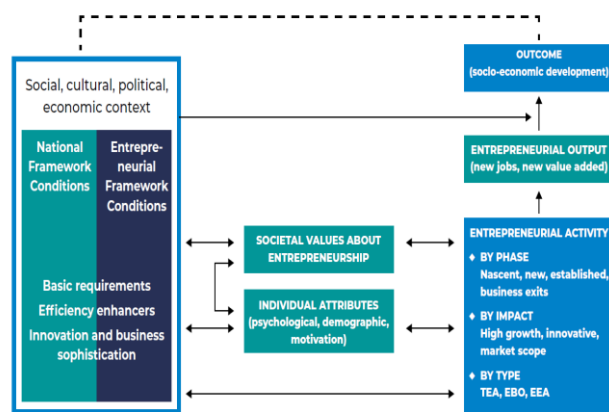


Fig. 5. GEM Conceptual Framework (Source: Global Entrepreneurship Monitor 2020/2021 Global Report)

There are a lot of different notions for entrepreneurship in the literature. Entrepreneurship according to GEM methodology is only the active behaviour, so the activity of someone who is committed to start or run a business. This methodology does not count in the notion of entrepreneurship recognizing a new opportunity for a new start-up, or thinking about a new business.

Figure 6. demonstrates the indicators of GEM entrepreneurship according to the periods of an enterprise, from the conception to an established business. One of the GEM indicator is the *Total early-stage Entrepreneurial Activity (TEA)*. This is the percentage of adults who start or manage a new business (nascent entrepreneurs) and the new business owner (baby businesses – aged less than 3,5 years). Potential entrepreneur who recognize opportunities for a new business and intend to start a new business is not involved in

TEA. There is another important GEM indicator: the level of *Established Business Ownership (EBO)*. This is the percentage of adults who own and manage an established business (where wages/salaries are paid for 42 months or more). Figure 3.2 shows that new or established business owner can exit the business at different stages.

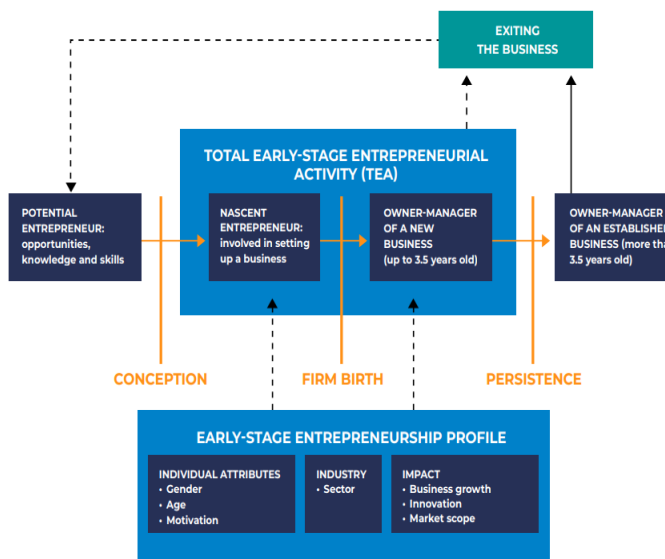


Fig.6. Entrepreneurial stages and GEM entrepreneurship indicators (Source: Global Entrepreneurship Monitor 2020/2021 Global Report)

IV. ENTREPRENEURIAL ECOSYSTEM AND ATTITUDES IN HUNGARY

Hungarian adults aged 18-64 were questioned in the GEM research and the results shows that to have an entrepreneurship is attractive however Hungarian adults are less willing to start a new business. The reason of that is the lack of experience and knowledge. Only 36% of Hungarian adults feel to have the abilities and knowledge to start and run a business (GEM Report 2020/2021). It means that business knowledge should be educated.

The entrepreneurial activity is higher among the men, among the age of 25-44 and among adults with higher qualification. The proportion of adults planning to start a business is lower than the average of countries taking part in the GEM research.

The strongest motivation to start a business is the own livelihood in Hungary. The innovation activity of Hungarian entrepreneurs is low according to GEM results. The major part of nascent entrepreneurs and new business owners (Total early-stage Entrepreneurial Activity) distributes, trades with generally known products and services and use established technologies (more than three quarters). The customers of Hungarian entrepreneurs are from local region and not abroad.

The Hungarian entrepreneurial ecosystem can be listed into the lower middle segment of European economies according to the NECI Index (4,5), however in relation to Middle-East Europe Hungary is in a good position considering entrepreneurial ecosystem. According to the results the entrepreneurial education is unsatisfactory. The market openness, the dynamics of the market, the government support

policy, and the R&D transfer are judged in disadvantageous situation. The key factor in the development of the entrepreneurial activity is strengthening of the entrepreneurial training, the increase of entrepreneurial education on different educational levels. The third of the Hungarian adult population think that there are good opportunities to start a new business in Hungary (GEM Report 2021/2022)

V. IMPACT OF THE COVID-19 ON HOUSHOLD INCOME

More than 130.000 individuals have participated in APS interviews inquired about the entrepreneurial impacts of Covid.-19 in 2020.

Level A >\$40,000	Level B >\$20,000-<\$40,000	Level C <\$20,000
Canada	Belarus	Brazil
Finland	Chile	Colombia
France	Croatia	Dominican Republic
Germany	Cyprus	Egypt
Ireland	Greece	Guatemala
Israel	Hungary	India
Italy	Kazakhstan	Iran
Japan	Latvia	Jamaica
Luxembourg	Lithuania	Mexico
Netherlands	Oman	Morocco
Norway	Panama	South Africa
Qatar	Poland	Sudan
Republic of Korea	Romania	
Saudi Arabia	Russian Federation	
Sweden	Slovak Republic	
Switzerland	Slovenia	
United Arab Emirates	Spain	
United Kingdom	Turkey	
United States	Uruguay	

Fig. 7. Economies in GEM (Source: Global Entrepreneurship Monitor 2020/2021 Global Report)

Economies were categorized into levels A, B and C (Fig. 7.).

- Level A is GDP per capita of more than \$40,000 (high income economies),
- Level B is a GDP per capita of between \$20,000 and \$40,000 (middle income economies), and
- Level C is GDP per capita of less than \$20,000 (low income economies).

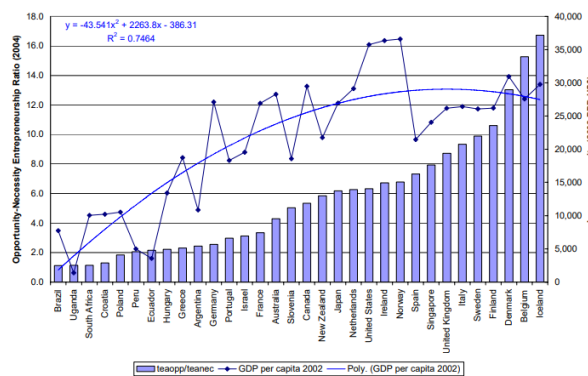


Fig. 8. Opportunity-Necessity Entrepreneurship Ratio and Income per capita (Source: Global Entrepreneurship Monitor 2004 Global Report)

Figure 8. shows the relationship between the opportunity-necessity entrepreneurship ratio and country income: there is a positive relationship between income level and the entrepreneurship ratio. The higher a country's per capita income, the more entrepreneurship is motivated by

recognizing an economic opportunity instead of necessity. As visible on Figure 8 Hungary is in the lower-middle segment of the analysed countries by GEM that means that Hungarian entrepreneurs are motivated from necessity instead of recognition of a business opportunity.

According to the results of the research in 2020/2021 Covid had a great negative impact on the income of households and mainly on households in the poorest financial situation. There were a deep fall in the income of poorer households and in the high-income economies There were the lowest proportions of interviewed individuals who reported that their household income decreased strongly. Hungary is among the middle level income economies so there was a great fall in household incomes as the impact of Covid-19 (Fig. 9.).

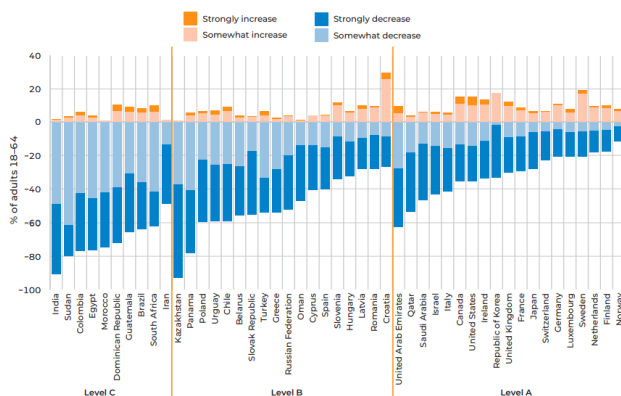


Fig. 9. The impact of COVID-19 on household income (Source: Global Entrepreneurship Monitor 2020/2021 Global Report)

The impact of Covid-19 on household incomes is much more lower in high income economies because of the higher economic stimulus packages.

Attitudes and perceptions influence on the level of the entrepreneurship. Entrepreneurial attitudes and perceptions can vary across the different countries and regions. Successful entrepreneurs can play an important role in a region because their behaviour strongly influence the decisions of individuals whether to start a new business. Covid-19 made starting a new business more difficult.

VI. CONCLUSION

The socio-economic development of a nation is dependent on knowledge, innovation, and entrepreneurship to a great extent. The extent of positive impact on economic growth depends on the factors of entrepreneurial ecosystem: attitudes, activities, capabilities, aspiration. The Hungarian entrepreneurial ecosystem can be listed into the lower middle segment of European economies according to the NECI Index and GEM Reports but in relation to Middle-East Europe Hungary is in a good position considering entrepreneurial ecosystem.

The entrepreneurial ecosystem influences entrepreneurs: the jobs, value-added and incomes generate entrepreneurial activity. The reverse is also true: if entrepreneurs of a country thrive that encourages economic development, adds new value and creates jobs. The society and economy have a great impact on education, which

basically determines entrepreneurship and the willingness of becoming an entrepreneur.

Covid had a negative impact on the income of households and mainly on households in the poorest financial situation. Since the COVID-19 innovation has become more important. One of its initiatives is Global Innovators, a community of start-ups and scale-ups focused on technological and business model innovations. Education of entrepreneurial knowledge play a very important role, moreover there is a positive relationship between post-secondary education and high start-up skills. The educated potential entrepreneurs can cope with start-up difficulties better than non-educated people.

REFERENCES

- [1] Ács, Zoltán J. and Szerb, László (2009): The global entrepreneurship index (GEMINDEX) . Jena Economic Research Papers in Foundations and Trends in Entrepreneurship
- [2] Ahmad, N. and A. Hoffman (2007), A Framework for Addressing and Measuring Entrepreneurship. Paris: OECD Entrepreneurship Indicators Steering Group.
- [3] Carree, M.A. and A.R. Thurik. (2002). "The impact of entrepreneurship on economic growth," in D.B. Audretsch and Z.J. Acs (eds), Handbook of Entrepreneurship Research, Boston/Dordrecht: Kluwer Academic Publishers.
- [4] Czeglédi, Csilla és Hajós, László és Kollár, Péter és Kuna, Zsuzsanna és Mikáczó, Andrea és Varga, Erika (2018): Motivations and attitudes to entrepreneurship in Hungary based on a 2015 examination. Studia Mundi – Economica, 5 (3). pp. 136-147. ISSN 2415-9395
- [5] Global Entrepreneurship Monitor Global Reports from 1999 until 2021/2022 (last unloaded: 9/10/2022)
- [6] Parragh, Bianka (2017) Competitiveness and Economic Stimulus. New Dimensions and Instruments of Monetary Policy. Polgári Szemle: Gazdasági és Társadalmi Folyóirat, 13 (Spec.). pp. 151-166.
- [7] Porter, M. E., J. D. Sachs, and J. W. McArthur (2002), 'Competitiveness and stages of economic development'. In: M. E. Porter, J. D., Sachs, P. K. Cornelius, J. W. McArthur, and K. Schwab (eds.): The Global Competitiveness Report 2001–2002. New York: Oxford University Press, pp. 16–25.
- [8] R. Thurik, S. Wennekers, L. Uhlaner (2002): Entrepreneurship and economic performance: a macro perspective, Published Economics, Business; Scales research reports
- [9] Wennekers, S, Van Stel, A, Thurik, R. and Reynolds, P. 2005, Nascent Entrepreneurship and the Level of Economic Development, Small Business Economics, 24(3), p. 293-309.
- [10] Horváth, Krisztina and Szerb, László (2015): GEM 2015 Hungary: The situation of entrepreneurs and the entrepreneurial ecosystem in 2015, GEM 2015, Pécsi Tudományegyetem, Közgazdaságtudományi Kar

MOTIVATION POTENTIAL IN HIGHER EDUCATION AFTER COVID-19 BASED ON PERSONALITY TRAITS

Timea Bekk
Alba Regia Technical Faculty
Obuda University
Szekesfehervar, Hungary
bekk.timea@amk.uni-obuda.hu

Monika Pogátsnik
Alba Regia Technical Faculty
Obuda University
Szekesfehervar, Hungary
pogatsnik.monika@amk.uni-obuda.hu

Abstract— Covid19, which started in 2019 and continues to have an impact today, has brought many changes to the world of education, including higher education. With the return of face-to-face education, it was no longer possible to continue teaching where we left off. Students found it difficult to engage in community life and apply the necessary learning methods. With the return of face-to-face teaching, it is a common experience that students are less motivated and difficult to motivate. In our research, we looked at which methods are suitable to help students learn, arouse their interest, and motivate them. We investigated the differences between the types defined by the MBTI personality traits of the students, and which motivational methods are more successful. Our research consisted of two elements, a preliminary qualitative study, and a quantitative study. The interviews provided input from the students' side about motivation, which was processed to produce the questionnaire used in our quantitative research for the second phase of the study. From the results of the research, we found that different types of people can be motivated by different methods. In addition to teaching and learning methods, the university environment, extra-curricular activities, and the personal relationship between teacher and student play a major role in motivation.

Keywords— motivation, Covid effect, personality traits, higher education

I. CHANGES IN LEARNING AND TEACHING AFTER COVID-19

The Covid-19 pandemic, which began in 2019 and continues to affect us today, has brought many changes to the world of education, including higher education. Students have learned to study online, and educators have learned to teach online, moving to a completely different way of life, teaching, and learning environment. A good example of this is the application of gamification [1], which was tried out at the Óbuda University Alba Regia Technical Faculty.

With the return of face-to-face teaching, it was not possible to continue teaching where we left off. The students also had difficulties getting involved in community life and applying the necessary learning methods. They use a different learning method live than the one they used for two years in the online space.

A prevailing attitude in higher education is that [2] are expected to think and behave as adults, to participate in the educational process as responsible persons with the ability to self-regulate their learning. Some instructors only expect students to 'give back the learning material, perpetuating learning habits that are only in line with external motivations and assessment situations [3].

We can look at the concept of motivation from several angles. Motivation influences every single action, preceded by an action or decision, or choice [2]. As Bigg mentions it, "Motivation is a result of good learning, not a condition of it" [4]. It is important that students can be motivated by their teachers and if students learn well and achieve good results, teachers will be motivated as well [2]. Therefore, it is necessary to develop a new teaching-learning method to better trigger and maintain their interest and motivation to learn [5].

Students enrolled full-time in university undergo big changes, especially in the first year of university life. The majority of students have lived at home during their previous studies at secondary school, always having a secure background with "nothing to do but study". Their parents helped them with everything else. When they left home, they had to learn not only a new learning regime but also self-sufficiency. It was a lot of change for the young people at once, and after Covid, it was even harder for them to pick up the pace.

Several studies have been carried out on the learning characteristics of students. The result was that learning orientations about learning and knowledge are the learning characteristics that influence the methods that regulate learning [6]. Thus, in higher education, it is necessary to change teaching methods to adapt teaching to students' views. In this way, we can better motivate them and make learning more interesting.

During the COVID-19 epidemic, students spent about one and a half years of their secondary education studying at home. They were completely isolated from each other, there was less opportunity for face-to-face contact, and they kept in touch online. Students who were just starting their university studies also had to relearn the rules of social networking. Students who were starting their second year of studies also had a difficult time, because they had their first year online and did not have the opportunity to learn the rules of higher education and to develop a trusting relationship with their peers.

With the return of face-to-face teaching, the general experience is that students are less motivated and difficult to motivate. Many students struggle with a sense of everyday drabness and monotony. They feel that even when they have completed a task, they have no sense of achievement and that there is always another task to be done.

In our research, we looked at which methods are suitable to help students learn, arouse their interest, and motivate

them. We examined the differences between the types defined by the MBTI personality test [7] and the personality traits of the students and examined which motivational methods are more successful.

In our research, we were looking for answers to the question of how we could facilitate students' learning and what motivational methods we could use to appeal to them.

We divided our research into two parts.

- Qualitative research in focus groups
- Quantitative research with questionnaires

II. QUALITATIVE RESEARCH METHOD AND RESULTS

In our qualitative research, students of Óbuda University Alba Regia Technical Faculty were interviewed through focus group discussions. The first element of the research aimed to use an informal focus group discussion as an opportunity to get to know the target group of the research better and to prepare the discussion for the compilation of the quantitative questionnaire. The focus groups were attended by 5-10 participants from several years of several courses and they answered the following question:

"What would motivate you more in your learning?"

We were able to interview full-time students from the following degree courses:

- Technical Management BSc I, II, III
- Electrical Engineering BSc I, III
- IT Engineering BSc I, III
- Land Surveying Engineering BSc I, II, III

This represented a total of 9 focus groups and 104 students.

III. RESULTS OF THE QUALITATIVE RESEARCH

The importance of practice-oriented education was mentioned by all groups of students interviewed. Students find it difficult to learn theory, and they need to see how they can apply what they have learned. They could even imagine doing this with several factory visits.

In addition to practice-oriented teaching, a good theoretical curriculum with lots of explanatory texts is very important for many groups. As an idea, the students suggested that if they could receive the theoretical material before the lecture, it would help them a lot. They could bring the printed outline to the lectures and write their notes on it. In today's fast-paced world, it is difficult to keep the attention of students, and different interactive learning materials or lecture methods could be used to add color to a lesson. For each topic, they would like the lecturer to be more focused, drawing their attention to where the topic will be relevant in practice.

The strict lecturers mentioned earlier were also mentioned during the focus group discussions, that it is not enough to 'give back the material, but that they need consistent rigor. "Rigour does not equal consistency."

There was divided opinion among the groups on how to explain the theory in detail. It depends on the personality type who needs more detailed explanatory lectures and which personality type learns theory easily through

exercises. What is important, according to the students, is that the trainers are good at motivating students (and need to) with different techniques to get them to attend lectures. During the semesters, two theoretical final papers are usually written. As an option, it has been suggested to replace at least one of the final papers with some practical exercises. Students also suggested the specific idea that one of the papers could be written at home, but that this should be much more difficult than writing it at the university. Where appropriate, they should research a specific topic. During the semester, if they could get extra points for doing an extra assignment and have it count towards the result of the essay, it would also give the students an external motivation. In one group, students also mentioned the grade offered as a motivating goal for them.

In group discussions, students mentioned project tasks as something they liked and could do more of, getting to know each other better and learning to work together or alongside each other in teams, in addition to learning. It also develops their personality; they learn empathy and tolerance. The possibility of team building was raised with one group.

During Covid, online test-taking, practice and debriefing were also used. This was very much appreciated and several groups specifically mentioned that they had learned a lot from the practice tests. One group also mentioned independent work in this area. This is very useful and they can learn a lot from it. The reward for completing the tasks was highlighted by the students.

The next big topic was scholarships. Here, opinions were very divided in the groups, as some were very concerned about this issue and others were completely unmotivated. What came up in several discussions was the calculation of the scholarship. In its current form, this was not considered appropriate by several groups, and ideas were put forward on how it could be well adapted so that the scholarship could also be a good external motivator. Besides the amount, another problem for students is that the money does not arrive in their accounts on time. As prices change over the years, it would be good if the amount of the scholarship could also increase.

For three groups, the development and extension of infrastructure have been raised. Respondents think that more and more modern practical equipment is needed for the training.

The third major theme is the relationship between teachers and students, both in and out of the classroom. Most of the groups mentioned the importance of good relations between trainers and students. Personal contact outside the classroom is very important so that the student can ask for help with the current course material and assignments. The interviews also revealed that students are more courageous to ask for help when they are on a first-name basis with the instructor than when they are in a private relationship.

A predictable rhythm was also highlighted by many. They like to see what will happen when because then they can plan.

IV. QUANTITATIVE RESEARCH METHOD AND RESULTS

Our quantitative research was carried out using Google Forms, an online interface that was easily accessible to the target group via the link provided. The questionnaire was filled in by 104 full-time students of the Óbuda University Alba Regia Technical Faculty (20 students in technical management BSc, 50 students in land surveying BSc, 18 students in electrical engineering BSc, 46 students in IT engineering BSc).

A personality test preceded the answers to our questions on student motivation. Nowadays, there are several excellent internationally accepted personality tests, also in Hungarian, to choose from. After much deliberation, we chose the validated Myers-Briggs Type Indicator (MBTI) personality test in online format [8], [9]. It was considered a good choice because it examines many aspects of an individual's personality, and the official Hungarian version is freely available (<https://www.16personalities.com/hu>). The theory underlying the test is that the seemingly random variation in behavior is quite orderly and consistent. People differ systematically in what they perceive and how they reach conclusions, and they differ in their interests, reactions, values, motivations, and abilities.

The test consists of 60 statements, the answers are given on a scale of 1 to 7, with which answers the respondent agrees. Circles indicate the values, and their size varies. In the end, the program shows you which personality type you belong to, based on your answers.

The personality test distinguishes between four categories:

- The first pair of psychological preferences, shows where the respondent focuses his or her attention and where he or she draws his or her energy from. Does the respondent prefer to spend time in the external world of people and things (Extraversion / Extraversion / E) or the world of inner thoughts and images (Introversion / Introversion / I)?
- The second pair of psychological preferences is the realistic perception of the environment and intuition. Which information does the individual pay more attention to, that which is received through the five senses (Realistic / Sensing / S) or does he pay more attention to that which is intuitively perceived based on the information received (Intuitive / Intuition / N)?
- The third preference pair shows how the respondent makes his/her decisions. Do you place more weight on objective principles and impersonal facts (Logical / Thinking / T) or do you prefer to pay more attention to personal concerns and the people involved (Intuitive / Feeling / F)?
- The fourth pair of preferences describe how the individual prefers to present him/herself to the outside world. Does he or she prefer a more structured and decisive behavior (Planning / Judging / J) or a more flexible and adaptable behavior (Researching / Perceiving / P)?

In the above categories, the test gives the individual measurement results as a percentage at the end of the test followed by a detailed description of the personality type identified by the test.

The MBTI distinguishes 4 types per category in 4 broad categories (16 personality types in total). These are discussed below in the focus of our topic, according to the motivation that characterizes each type [9]:

1. Analysts:

- Architect (INTJ) - They seek the cause of events, and they brainstorm. They are visual types. They look for connections.
- Scientist (INTP) - More ideas are important to them. They can be motivated by specific readings and tasks.
- Commander (ENTJ) - They filter the material and learn quickly. They like to work in teams, brainstorm, and initiate discussions.
- Debater (ENTP) - They are motivated by complex tasks, they use a logical approach to solve the problem. Tasks should be conceptually based, pay attention to multiple solutions, and link the material to the application.

2. Diplomats:

- Advocate (INFJ) - Motivated by the prospect of solving humanity's problems. They like to push boundaries, listening to their intuition. They tend to work independently and prefer theoretical tasks. They like to work in teams and are visual types.
- Mediator (INFP) – They are concerned about human problems; they value information about human behavior. They have an eye for understanding the goal. They need to use visual materials in lessons and see multiple perspectives on a task.
- Protagonist (ENFJ) – The context is important for them and to help human development, the curriculum should be taught. Active learning is useful for them.
- Campaigner (ENFP) - Issues related to people are important to them. They are motivated by caring about people. Logical connections are not important in learning, they can learn large amounts of material. They like tasks that require creativity.

3. Guards:

- Logistician (ISTJ) - Goal setting, practical training, and teamwork are important to them.
- Defender (ISFJ) - They put a lot of emphasis on theoretical knowledge to understand how they can use it in practice. Examples are important to them; the curriculum should include details and facts for them.
- Executive (ESTJ) - Arousing interest, good theoretical material, video, active involvement in the lesson is motivating for them.
- Consul (ESFJ) - These types of people are people-oriented, and theoretical knowledge is important to them. They do not need logical information; they can easily memorize details. They learn better when they are emotionally involved.

4. Explorers:

- Virtuoso (ISTP) - They have a technical mind, they are constantly looking for tools to improve.

- Adventurer (ISFP) - If they are aesthetically pleasing to the subject, they are enthusiastic learners. They do not need logical thinking; they can do mechanical memorization. A lot of lecture material, and emotion is important for them.
- Entrepreneur (ESTP) - Practical solution seekers. They can memorize small details. They learn from examples and participation in discussions. They like group work.
- Entertainer (ESFP) - They grasp new material quickly when they see it applied, especially if it entertains them. They like to communicate actively in the process and can remember a lot of information without logical thinking. They are visual types, and examples are important.

Table 1 shows the distribution of the 104 participants in the survey by MBTI type. The largest proportion of the sample was 'Diplomats', N=43 (41%). Within the main category, the most typical personality types were 'Protagonists' (N=16) and 'Advocates' (N=15). 'Analysts' N=24 (23%), 'Guardians' N=20 (19%) and 'Explorers' N=17 (16%).

TABLE 1 DISTRIBUTION OF THE SAMPLE BY MBTI TYPE (PERSONS)

Analysts	24	Diplomats	43	Guardians	20	Explorers	17
Architect INTJ	8	Defense lawyer INFJ	15	Logistic ISTJ	3	Virtuoso ISTP	3
Scientist INTP	8	Mediator INFP	9	Defender ISFJ	7	Adventurous ISFP	4
Commander ENTJ	4	Protagonist ENFJ	16	Leader ESTJ	4	Entrepreneurial ESTP	2
Debater ENTP	4	Campaigner ENFP	3	Consultant ESFJ	6	Entertaining ESFP	8

The second element of the quantitative research was a questionnaire, in which the first item was the personality type code obtained because of the MBTI test. Six questions were formulated in the questionnaire using the results of the qualitative research. On a Likert scale of 1-7, respondents were asked to rate the extent to which each question influenced their motivation to learn. 1 being the least and 7 being the most important, most influential on their learning. We chose the 1-7 scale because we felt that they could not necessarily express their feelings accurately on the 1-5 scale, and the 1-10 scale was very broad.

We asked the students the following questions:

"How much does it affect your learning:

1. your relationship with your tutors
2. the extra-curricular activities you do together
3. the scholarship
4. the practical orientation of the teaching
5. attending a good theoretical lecture
6. your relationship with your peers"

V. RESULTS OF THE QUANTITATIVE RESEARCH

The results of our quantitative research were analyzed in Excel. Six charts were created for the six questions, shown in Figures 1-6, where the horizontal axis shows the personality types, and the vertical column shows the average scores from 1-7.

The first question examined the importance of good rapport with instructors, with an average score of 4.95 on a

scale of 1-7. The first chart (Figure 1) clearly shows that for the personality types Adventurous, Scholar and Leader, it is most important to have a good relationship with the instructor (mean 5.75). It is also important to be available to them outside of class if they have questions about the assignment.

The least important for the Commanding and Virtuoso personality types is the need to maintain regular, direct contact with their respective instructors (mean: 3.50).

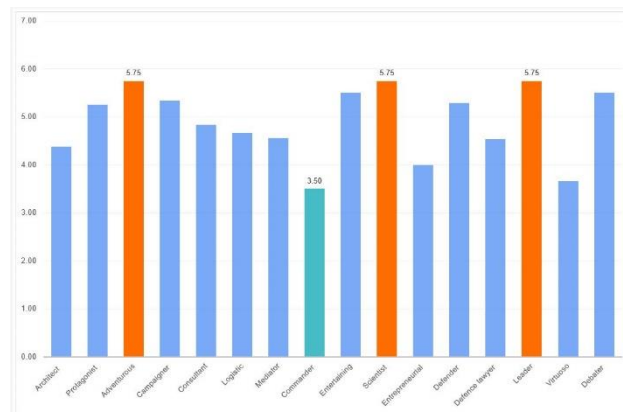


Figure 1: How much does your relationship with your teachers influence your learning

Figure 2 illustrates personality types, for whom the extra-curricular activities are important. The average value for the whole sample was 4.13. Here, the value for the Leader personality type is the highest (mean: 6.00). This shows that it is important for him to have contact with everyone. The Campaign Leader and Fun personalities also have high scores, and they may be influenced in their learning by the company they can fit in with, and the extent to which they can have a student life outside of the compulsory study programs at university. Those with the lowest mean score here are the Virtuoso personality (mean: 1.67) and are less likely to be motivated by this.

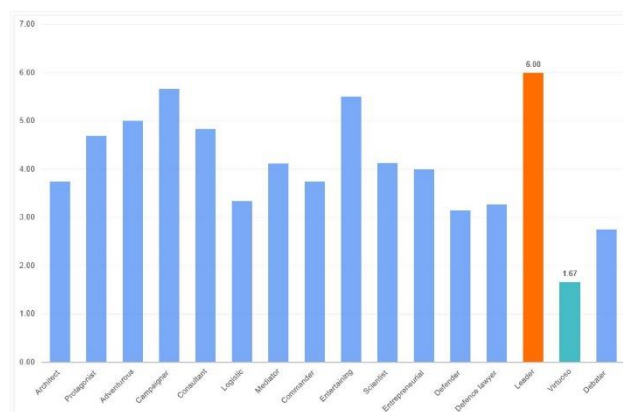


Figure 2: How much does the extra-curricular program influence your learning?

The question of the importance of scholarships was also shared across the qualitative focus group discussions (Figure 3). The average score across the whole sample was 4.60. Scholarships had the least impact on learning for the Virtuoso-type students (mean: 2.67). They were not motivated by any of the questions asked. Here, the highest score was for the Commander's type (7.00).

During the focus group discussions, the students told us that their parents' main motivation was money, they worked a lot when they were children, and they did not see a positive outcome. They had a variety of mental and physical problems and illnesses. They could not enjoy the use of money to the extent that they worked for it. During the interviews, students said that other types of recognition were important in addition to financial recognition. For example, feedback from managers, and flexible working hours [10].

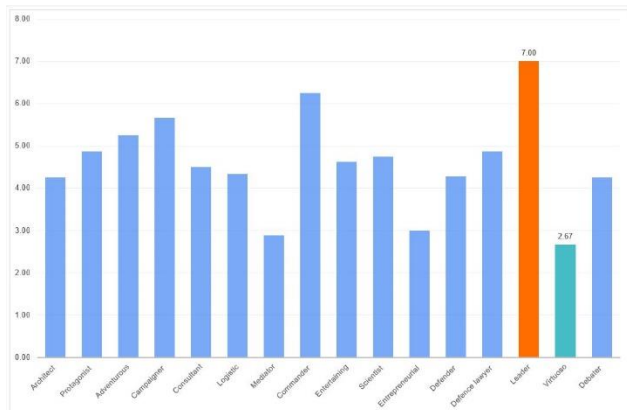


Figure 3: How much does the scholarship affect your learning?

The highest scores, and hence the most important motivational element for this generation, is practice-oriented training. They memorize, learn and comprehend better when they can see the connections and the meaning of what they have learned in practice. The average score for the whole sample was 5.81. The lowest value here was also 4.50, which is the highest minimum value when considering the other factors examined (Figure 4). The importance of this was also mentioned by all teams during the focus group discussions.

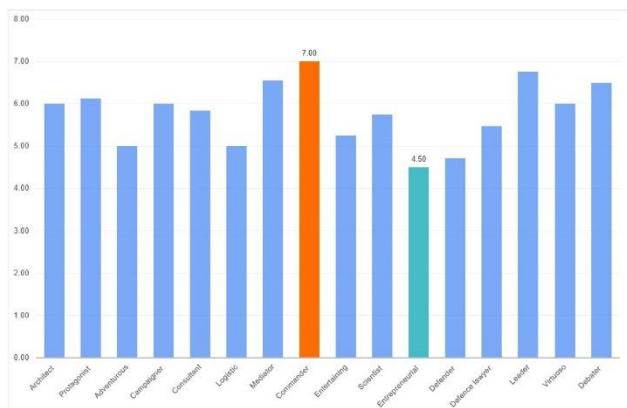


Figure 4: How much does the practical orientation of education influence your learning?

When asked how much a good theoretical curriculum influences learning, there is no extreme difference. The average score across the whole sample was 4.79. Those who are highly motivated by this are Logistic personalities (mean: 6.67). This is shown in Figure 5. The other personality types do not show a large variance.

The importance of good relationships with peers in learning was assessed in the next question (Figure 6). The average score across the whole sample was 5.21.

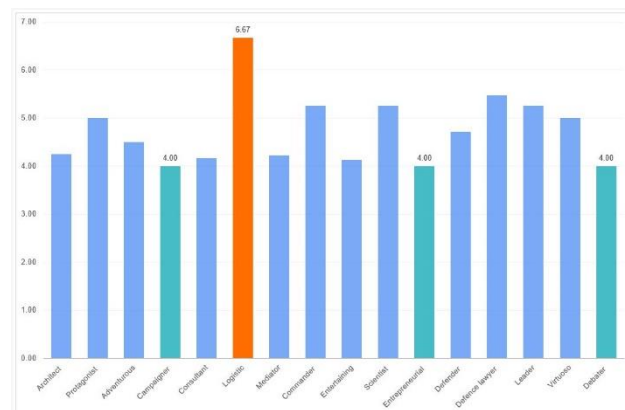


Figure 5: How much does attending a good theory lecture influence your learning?

Relationships with peers were the most important for the Campaign Leader (mean: 6.67) and the least important for the Virtuoso personality (mean: 2.00). The results of our study suggest that this personality type is best motivated by practice-oriented training.

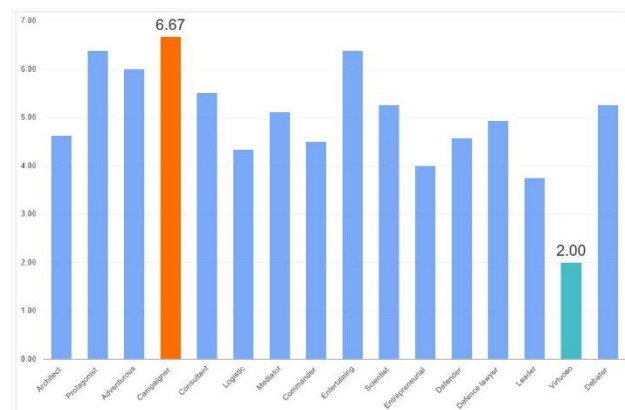


Figure 6: How much does your relationship with your peers influence your learning?

The last question was open-ended "What other motivational opportunities help you to learn?" This question was optional. Compared to our previous experience, when we asked an optional open question in the questionnaires, not many people answered, but this time almost everyone did.

The students consider it important to have a good timetable so that they can fit in their leisure activities with their studies. The sport was highlighted in particular.

Intrinsic motivations included having a vision of what they want to achieve in life and what drives them to move forward. Positive feedback after a paper or report makes them feel good that it was worth learning. What several students highlighted was literacy. If they study a lot, they become more literate in each subject. Goal setting and the desire to improve are important. External motivations included the importance of theoretical and practical education.

The following were highlighted by several students: the instructor's preparation for the lessons, competition in the subject matter, and completion of tests. Offering a grade at the end of the semester. Mutual respect between instructors and students. Showing in practical classes what and where you can use what you have learned, gaining extra points for

the exam. Providing accurate information on what the student can find where. Transparency and comprehensibility of the course material.

In our fast world, technology is also rapidly evolving, it is important that when students graduate from university, they can work immediately in the workforce with the tools they use in the workplace, not learning outdated methods. Several students wrote this.

Students also set goals for themselves, for example, one student wrote that he was motivated by the fact that if he studied for half an hour, he would allow himself to play a video game for half an hour. This is how he rewards himself.

Another important motivating factor is human relations. Here several people wrote about the parent-child relationship. Even at this age, it can be observed that the pupil is not only learning for himself but also wants to please his mother and father. The family background is very important for them.

The relationship with their classmates and classmates, both in learning and in leisure activities. Students prefer to study in small groups and work together to solve problems. In general, group members who study together also do leisure activities together. Conversations were also highlighted, informal casual discussions with each other on specific topics. The third human relationship highlighted was the relationship with the teachers. It is important to dare to ask for help if they get stuck.

The last motivational incentive highlighted was money. In the academic world, it is the scholarship, in the working world, it is the salary. Today's students are much more likely than in previous years to take jobs alongside their studies to finance their studies and support their livelihoods, and possibly their families. There are also frequent examples of students moving from full-time to part-time study to find a full-time job. In the world of work, a secure financial future and advancement are important goals and motivations for them in their studies. A clear tangible vision.

Two students wrote the following in response to the open question:

"Ancient and medieval Chinese culture and mythology have been an integral part of my life for nearly a decade and a half. From their stories, I have come to understand that knowledge is one of the most important things in a person's life and that knowledge is a respectable, authoritative quality. Besides, I have often been assigned to work as a teacher, mainly tutoring primary and secondary school students in all kinds of subjects, by organizing my school. It makes me feel good to have knowledge that I can use to help others in their studies. Also, my language skills have given me quite good job opportunities (3 foreign languages)."

"I set myself goals that are within reach. I try to do my best, but I often run into difficulties. For example, I had an exam here at my university that I studied a lot for, and I was one of the best in that subject, I helped a lot my peers' week after week with the papers and I almost failed the exam. If something unfair happens to me, it affects my motivation. I'm motivated by being a role model for other people, making my mum and dad proud of me, and much more. I have ambitious plans to make an impact in my local community."

VI. SUMMARY

In our research, we were looking for answers to the question of how we can motivate students studying at university and make learning interesting for them. We used the MBTI personality test to find the most effective motivational tools for each personality type. The sample of our study is a small segment of university students, general conclusions can only be drawn based on larger, multi-institutional further research.

However, it is already clear from this small study sample of 105 students that we can no longer simply apply the previous, mainly frontal learning methodologies to Generation Z students and their successors. University education also needs to be radically renewed, and alternative methodological techniques need to be used to arouse and maintain young people's interest and to ensure that they leave their chosen university with a positive feeling and knowledge that is useful for the labor market.

There is also a common perception in the world of education that young people are not hard-working and persistent enough, that they are incapable of hard work, and are generally characterized as less capable than the older generation. This may be because the standards set by the older generation of teachers are no longer valid. The young generation of the 21st century is no less interested than their predecessors, but their attitudes to learning need to be recognized, and teaching thinking and methods need to be transformed so that higher education can continue to innovate in the fast-changing world of the 21st century.

REFERENCES

- [1] Módné Takács, J., Pogátsnik, M., Kersánszki, T. (2022): Improving Soft Skills and Motivation with Gamification in Engineering Education. In: Auer, M.E., Hortsch, H., Michler, O., Köhler, T. (eds) *Mobility for Smart Cities and Regional Development - Challenges for Higher Education*. Vol 389. Springer, Cham. https://doi.org/10.1007/978-3-030-93904-5_81
- [2] Jármai, E. M., Végh, Á. (2017): Motivációról a felsőoktatásban – az oktatói és tanulási motiváció kapcsolata (About motivation in higher education – Relationship between learning and teaching motivation). SJE International Scientific Conference "Value, quality and competitiveness - challenges of the 21st century", Komárom. URL: <https://bit.ly/3NYPDtD> (downloaded 12.04.2022)
- [3] Győrfyné Kukoda, A. (2012): Should higher education students be educators or pedagogues? *School Culture* 11. pp. 48-56. URL: <https://bit.ly/3uyz4Nu> (downloaded: 12.04.2022.)
- [4] Rodgers, T. (2016): Measuring Value Added in Higher Education. In PUSZTAI Gabriella, Bócsi Veronika, Ceglédi Tímea (eds.) *Higher Education & Society. Approaches to empirically capturing the institutional contribution*. Partium Könyvkiadó - Personal Problems Solution - Új Mandátum Könyvkiadó, Oradea - Budapest, 2016. 373 p. pp. 90-111. (93) URL: <http://mek.oszk.hu/16000/16053/16053.pdf> (downloaded 27.06.2022)
- [5] Juhász, Cs. (2017): *Central European Publications*, 10(2), pp. 131-141.
- [6] Kálmán, O. (2009): URL: <https://bit.ly/3SQq6Vd> (downloaded 27.06.2022).
- [7] Erős, I., Jobbágy, M. (2001): The Myers-Briggs Type Indicator (MBTI) in Hungary. URL: <http://www.mentalhub.hu/mbti.pdf> (downloaded 12.04.2022).
- [8] Villányi, G. (2015): MBTI URL: <https://bit.ly/3rat36U> (downloaded 27.06.2022)
- [9] Unicheck (2017): MBTI Types: Educating Various Types of Students URL: <https://bit.ly/3RdjMFV> (Retrieved 11.07.2022)
- [10] Magasvári, A., Szilágyi, T. (2019): In. *Customs and Financial Investigation Section, Bp., pp. 142-156*. <https://doi.org/10.37372/mrtvpt.2019.1.9>

The changes in the concept and content of the mechanical engineer training from the 3rd industrial revolution to industry 4.0

Bráda Csaba
Alba Regia Technical Faculty
Obuda University
Szekesfehervar, Hungary
brada.csaba@amk.uni-obuda.hu

Monika Pogatsnik
Alba Regia Technical Faculty
Obuda University
Szekesfehervar, Hungary
pogatsnik.monika@amk.uni-obuda.hu

Abstract—In a three-step study, the training system of mechanical engineering was examined from several points of view to better meet the expectations of the labor market. In the first step of the research, we define the framework of the "playing field" in which we examine the possibilities of adapting higher education to the needs of customers. We give an overview of the changes in the concept of engineer and mechanical engineer, based on the articles on the defining lexicons of the given era. In addition to the official wording, we were also interested in how the professional community involved in the training defines the identity of the engineer and the mechanical engineer. Among the mechanical engineering students, we asked for an explanation of the concepts and related activities in an explanatory association questionnaire. In the third step, from the period of the strong development of the Hungarian industry, from the beginning of the 1970s, the training of mechanical engineers is examined, its changes in time, and its content. The change of the officially defined concepts over time indicates the expansion of the content elements, which can be traced back to technical and technological development.

Keywords—mechanical engineering, technical higher education, labor market needs.

I. CHANGES IN THE INDUSTRIAL ENVIRONMENT IN THE 21ST CENTURY

Industrial revolutions fundamentally change the technologies of the past, radically transforming working conditions and people's way of life. In our study, we compare the period of the 3rd industrial revolution with the industrial environment of our time, the period of the 4th industrial revolution, focusing on the profession of mechanical engineering [1], examining the evolution of the profession's requirements and changes in training.

The great achievement of the first industrial revolution was the use of steam power and the mechanization of production in the 18th century. The second industrial revolution 19-20. the period of the turn of the century, when electricity was discovered, and mass production accelerated with production lines. The third industrial revolution began in the 1970s with the appearance of partial automation implemented with programmable memory controllers and computers, thanks to the introduction of these technologies it was possible to automate entire production processes. Currently, the fourth industrial revolution is taking place, which is also referred to as "Industry 4.0". The networking of manufacturing systems results in the creation of smart

factories, in which manufacturing systems and people communicate with each other through a network, and production is almost automatic [2][3].

For the successful implementation of Industry 4.0, there is a great need for specialists to build and maintain new smart factories [4]. Therefore, mechanical engineering students must master a combination of classical mechanical engineering and informatics [1]. The industrial sector is looking for specialists who are prepared for the new challenges that arise [5]. Therefore, higher education must provide a training environment where future professionals in the industrial sector can acquire the necessary competencies. Our goal is to identify the specific competencies that mechanical engineers of the future will need from the perspective of Industry 4.0.

Many studies deal with the mapping of the so-called 21st-century competencies, in the case of engineering careers, for example, the most frequently mentioned competencies include motivation, dedication, the ability to change, organizational awareness: analytical thinking, creativity, logical and mathematical skills, cooperation skills, problem-solving [6]. In addition to the importance of soft skills, it is also worth reviewing the professional skills (hard skills) that appear as new elements in the Industry 4.0 period and that must also appear in higher education.

II. THE PURPOSE AND METHOD OF THE RESEARCH

Among the objectives of higher education institutions, the commitment to quality education appears in different formulations. In general terms, this objective sounds like graduates can meet labour market expectations with high-quality professional knowledge.

During our research, we are looking for an answer to the question of where and what modifications will help in the training system for mechanical engineering students to meet the labour market expectations of the 4th industrial revolution better than they currently do. The diversity of partner needs and changes over time make it necessary to periodically review, evaluate and modify training courses as necessary.

As a first step, we determined the framework of the "playing field" in which we examine the possibilities of adapting higher education to customer needs. We give a brief overview of the changes in the concept of engineer and mechanical engineer profession, based on the entries of the

defining lexicons of the given era. In addition to the official formulations, we were also interested in how the professional community involved in the training defines the identity of the engineer and the mechanical engineer profession. The change over time of the officially defined concepts, - from now on, concerning mechanical engineering - clearly refers to the expansion of content elements, which can be traced back to technical and technological development [7].

During the definition and interpretation of the concept and content of the expression engineer and, within that, the mechanical engineer, the question arose as to how well the students who will obtain a mechanical engineering degree are aware of the basic concepts of their future profession. As a second step, we, therefore, surveyed 32 mechanical engineering students. In a paper-based, four-question, association questionnaire, we asked for the definitions of the concepts and related activities. The questionnaire asked about the concept of engineer, mechanical engineer, and knowledge of the categories of engineer and mechanical engineer. When defining the term engineer, mechanical engineer, keywords were used for the evaluation, while a simple quantitative evaluation took place in the case of the categories.

In the third step, we examined mechanical engineering education and its changes in time and content from the period of strong development of the Hungarian industry, from the beginning of the 1970s. For the investigation, we used the training documents of the Technical University of Heavy Industry in Miskolc. During the examination of this chapter, among the qualitative methods, we performed the content analysis of the available documents, as well as the examination of the data appearing in the documents as databases.

III. CONCEPT OF ENGINEER, MECHANICAL ENGINEER

The word engineer began to spread in England in the 13th century, where the root engine denoted a military invention, a military structure. The engineer designated the person who invented, created, and professionally operated these devices. With the development of technology, the content of the word gradually expanded, in addition to military equipment, it extended to the creation and operation of all kinds of machines and mechanical structures [8]. The spread of the term and the interpretation of its content to this day can be linked to the first industrial revolution.

The Hungarian word “mérnök” (engineer) can originally be derived from the verb “mérni” (measure) and refers to the activity of surveying. The concepts of mechanical engineer and chemical engineering were created during the first industrial revolution. In the 19th century, engineering activities were expanded to include the categories of architect, smelter engineer, mining engineer, and surveying engineer, and in the 20th century, among others, the categories of electrical engineer, transportation engineer, and mechanical engineer. In Hungary, by 2022, the number of engineer categories reached 18. Each engineering category contains additional subcategories which typically appear as specializations in university education.

The official description of the word engineer, based on the 1916 edition of the Hungarian Révai encyclopaedia, is the following "an engineer is a person who deals with the application of technical tools and sciences and has obtained a

qualification for this". The content of the concept did not change much in the 1962 New Hungarian Lexicon or the entries of the latest editions of other Hungarian lexicons. Each definition links engineering to university or college education.

When we try to define the concept of a mechanical engineer also based on the 1916 edition of Révai's encyclopaedia "a mechanical engineer has obtained a higher theoretical qualification and received his training at a technical college".

It is interesting to observe the appearance of machine builders in the use of words, which very well shows the level of technical development of the time. Since mechanical devices and machines were already being manufactured and operated at this time, machine building was an accepted and established profession, which included several professions. These can generally be divided into two main groups, according as they serve the machine-making industry with their dexterity, craft, or art, and those who serve the machine-making industry with their theoretical and practical training acquired in scientific educational institutions. The first is machine workers, so-called machine locksmiths, turners, mechanics, etc. and the latter are machine technicians and mechanical engineers [9].

Engineering education in Hungary started in the 19th century at two locations in Selmečbánya, where smelter and mining engineers were trained, and in Budapest, at the King Joseph University, where there were five departments and each of them inaugurated the students who passed the exam as engineers with a separate diploma [10]:

- engineering
- mechanical engineering
- architectural engineering
- chemical engineering and
- provided economic engineering qualifications.

In 1913, training took place in 2 locations in Hungary in a total of 7 engineer categories. This number grew 11 in 1962, nowadays the FEOR-08 (Hungarian professions list) enumerate 18 engineering occupations, felvi.hu (Hungarian University admission web site) announced 26 different basic technical courses in 2022. The various engineer categories are summarized in Table I.

We can observe not only the expansion of the various engineering categories but also the change in their content. The content elements are practically related to the job, which appears in the training system in the form of specializations. The mechanical engineer category is currently assigned 45 jobs based on FEOR data (Table II).

For each job, a general technical knowledge material appears, which defines engineering, and a knowledge material belonging to the specialization, which provides a great deal of help for the filling of a given job.

The jobs of students who graduated with a mechanical engineering major based on the 2019 data of the Graduate Career Tracking System (hereinafter referred to as DPR), the largest number of graduates were employed in the fields of design engineering, construction engineering, process engineering, project management, and building mechanical engineering.

TABLE I. ENGINEER CATEGORIES IN 1962 AND TODAY

(Berei, 1962) Engineer categories	FEOR-08 (2022)	BSc technical courses (felvi.hu, 2022)
General or cultural engineering	Mining engineering	Materials engineering
Architectural engineer	Smelter and materials engineer	Bioengineering major
Transportation plant engineer	Food industry engineer	Safety engineering major
Mechanical engineer	Wood and light industry engineer	Energy engineering major
Electrical engineer	Architectural engineer	Architectural engineering major
Chemical engineer	Civil engineer	Civil engineering major
Smelter engineer	Chemical engineer	Wood engineering major
Mining engineer	Mechanical engineer	Mechanical engineering major
Surveying engineer	Electrical engineer (energy engineer)	Industrial product and design engineering major
Agricultural Mechanical Engineer	Electrical Engineer (Electronic Engineer)	Vehicle Operating Engineering
Forestry and wood industry engineer	Telecommunications engineer	Vehicle engineering major
	Agricultural engineer	Light industrial engineering major
	Forestry and nature conservation engineer	Environmental engineering major
	Landscape and garden architect	Transportation engineering major
	Urban and transport planning engineer	Logistics engineering major
	Surveying and GIS	Mechatronics engineering major
	Graphic and multimedia designer	Molecular bionics engineering major
	Technical manager major	Aeronautical engineering major
	Quality assurance engineer technical geoscience major	Test engineering major
		Fire protection engineering major
		Chemical engineering major
		Electrical engineering major
		Electrical engineering major
		Water management engineering major

TABLE II. MECHANICAL ENGINEER JOBS BASED ON FEOR (2010)

Automotive component design engineer	Lubrication engineer
Car design engineer	Construction engineer
Building mechanical engineer	Research engineer, mechanical engineering
Building mechanical design engineer	Air technology system design engineer
Power engineering plant engineer	Mechatronics engineer
Development engineer mechanical engineering	Deep sea mechanical engineer
Process engineer, mechanical engineering.	Agricultural machinery consulting engineer
Heating equipment, ventilation equipment engineer	Agricultural mechanical engineer
Mechanical plant engineer	Assembly engineer, machine production
Mechanical engineering plant engineer	Aircraft engineer
Mechanical researcher	Multiplication process engineer
Automotive engineer	Ventilation equipment engineer
Automotive engineer	Tool developer
Car constructor	Machine tool manufacturer design engineer
Machine design engineer	Tool engineer
Ship constructor	Tool technologist engineer
Hangar engineer	Tool design engineer
Efficiency development engineer, mechanical engineering	Product developer, mechanical engineering
Welding engineer	Cylinder head production engineer
Production engineer, machine manufacturing	Production planning engineer, mechanical engineering Test engineer, machine manufacturing
Hydraulic engineer	Turbine engineer
Vehicle mechanical engineer	Maintenance engineer, mechanical engineering

Mechanical engineering training has expanded not only in terms of its content but also in terms of training locations, currently training in this major is taking place in 14 higher education institutions in Hungary. We reviewed what specializations are available in each institution, which is summarized in Table III. The specialization allows for more in-depth knowledge of a narrower field of study within the bachelor's degree in mechanical engineering. The most frequently advertised specializations are partially consistent with the professional map that also represents industrial needs, i.e. specializations are most typical of the basic mechanical engineering course, for which the most jobs appear according to the DPR results. The most common specialization: vehicle industry with various names (7 locations), machine designer (5 locations), material technology (4 locations), building engineer (4 locations), machine manufacturing technology (4 locations), mechatronics (3 locations), process technology (3 locations).

TABLE III. MECHANICAL ENGINEERING TRAINING PLACES AND SPECIALIZATIONS IN HUNGARY (SOURCE: FELVI.HU, 2022)

Training place	Optional specializations
University of Technology and Economics of Budapest, Faculty of Mechanical Engineering	material technology, building engineering, process technology, mechanical engineering development machine manufacturing technology, machine designer, a mathematical engineer
University of Debrecen, Faculty of Technology	materials technology, building engineering, machine designer, vehicle industry process designer, operator-maintainer
University of Dunaujváros	maintenance; mechatronics
ELTE Faculty of Informatics (Szombathely)	not defined
Hungarian University of Agriculture and Life Sciences (MATE), (Kaposvár, Gödöllő)	building engineer, motor vehicle technology; machine manufacturer; IT Engineer
University of Miskolc, Faculty of Mechanical Engineering, and Informatics (Miskolc, Sátoraljaújhely)	materials technology; machine manufacturing technology; machine designer; engineering modelling; quality assurance; machine tool and target machine designer; chemical engineering
Neumann János University GAMF Faculty of Technology and Informatics (Kecskemét)	materials technology and quality; weapons designer and manufacturer; production IT; mechatronic; plastic processor
University of Nyíregyháza	mechanical engineering technology; vehicle mechanic
Óbuda University Alba Regia Technical Faculty, Bánki Donát Faculty of Mechanical and Safety Engineering	process technology, CAD-CAM-CNC; weapon and ammunition technical; machine design; combat vehicle technical; vehicle technology; aeronautical
Pannon University, Faculty of Engineering (Veszprém, Zalaegerszeg)	production design; mechatronics; silicate industry machinist; technical logistics
The University of Pécs, Faculty of Engineering and Informatics,	building engineer; machine structure and process designer
István Széchenyi University, Faculty of Mechanical Engineering, Information Technology and Electrical Engineering (Győr)	mechanical engineering technology; automotive technology; vehicle manufacturing
Faculty of Engineering, University of Szeged,	process engineering; mechanical maintenance

IV. THE CONCEPT OF ENGINEER AND MECHANICAL ENGINEER IN THE REFLECTION OF STUDENT VIEWS

We examined the extent to which mechanical engineering students are aware of the concept of their studied profession and their labour market opportunities. We surveyed the mechanical engineering students of the Alba Regia Faculty of Technology at Óbuda University. We asked the students to fill out an associative questionnaire consisting of 4 questions. 32 undergraduate mechanical engineering students of Óbuda University filled out the short questionnaire, which was evaluated by matching questions 1 and 2 to key words, while questions 3 and 4 was evaluated by simple quantification. The sample consisted of 19 third-year, 5 second year, and 8 first-year students.

Before filling in, the students were informed that filling in the questionnaire had neither positive nor negative consequences. It is not a problem if they cannot give a complete or correct answer to a given question. The purpose

of the questionnaire is to establish a general level of knowledge.

The concept of an engineer: in the first question, we asked the students to define the concept of an engineer in their own words. During the evaluation, the keyword was higher education. The results are presented in Table IV.

Most respondents do not associate the general concept of engineering with high-level professional knowledge, but consider calculation, measurement, and planning to be the defining characteristic of an engineer. 53% of the students made a clear reference only and exclusively to calculation, planning, and design. 25% of the students, in addition to referring to the degree, also consider calculation, planning, and dimensioning as important defining elements of engineering in the other supplement.

 TABLE IV. THE CONCEPT OF THE ENGINEER (1ST QUESTION)

Answer category	Number of answers
reference to education without comments	4
reference to education with other additions	8
reference to calculation, planning	17
cannot be interpreted	3
total answer	32

The concept of a mechanical engineer: in the second question, we asked the respondents to define the concept of a mechanical engineer. Keywords for this question were mechanical structures, design, operation, and maintenance. The results are presented in Table V.

When defining the term mechanical engineer, the picture is much more favourable than in the case of the general engineering term. The vast majority of the surveyed university students referred to working with mechanical structures. Based on the previous question, planning almost naturally appears among the answers in many cases. However, virtually no one covered the entire vertical of the mechanical engineer's activity. The large area of maintenance and operation was practically omitted from the answers.

 TABLE V. THE CONCEPT OF THE MECHANICAL ENGINEER (2ND QUESTION)

Answer category	Number of answers	Answer category	Number of answers
does not include keyword	14	does not include keyword	12
reference to mechanical structures	18	reference to design, production	20
total answers	32	total answers	32

Engineer categories: we asked the students to list the well-known engineer categories in the third associative question. The results are presented in Figure 1.

Two things are evident from the answers. One is that there are well-known categories of engineers, and there are lesser-known ones, as well as those referring to a similar field of activity, such as the architect and civil engineer. They do not deal with the differences between them since they are not involved.

The other thing that can also be seen from the answers is that the students think in non-engineering categories, i.e., qualifications and the jobs that can be assigned to them. This is indicated by the large number of responses that can be classified into other non-main categories, as the engineer titles appeared here, which mean a qualification related to a job.

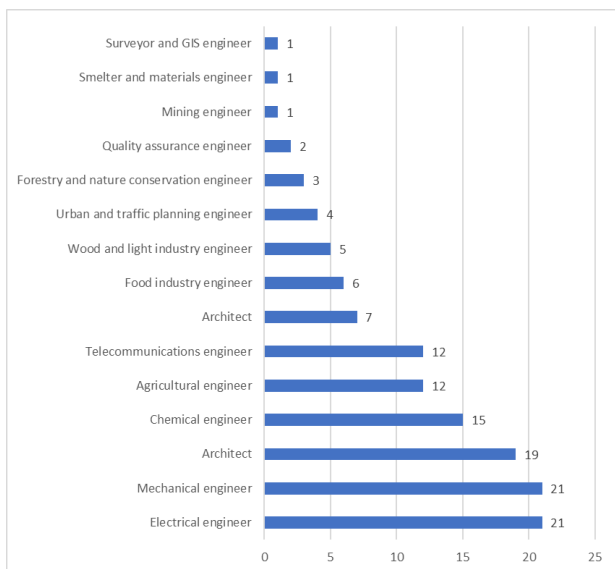


Fig. 1. Listed engineer categories (number of mentions)

Mechanical engineering jobs: Our fourth question was to list the mechanical engineering jobs known to the students. Results are presented in Figure 2.

Of the 45 jobs listed in the FEOR that can be filled with a mechanical engineering degree, only 6 was mentioned in the responses. Possible reasons for this could be:

- professional interests,
- limited experience gained in a practical place,
- knowledge of a narrower professional environment.

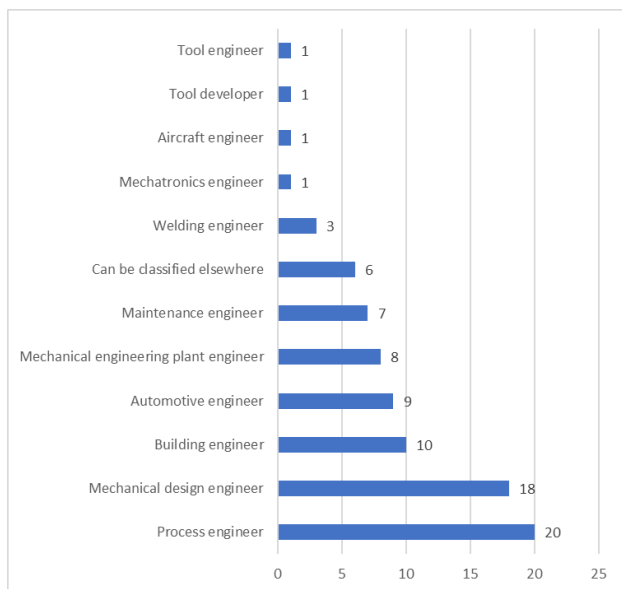


Fig. 2. Listed mechanical engineer jobs (number of mentions)

V. TIME AND CONTENT CHANGES IN MECHANICAL ENGINEERING TRAINING

There is always a reason behind the change that inspires development. In the area we are examining, the root cause is technical development and scientific progress. Newer and newer scientific discoveries and technical solutions have made it necessary for the concept of engineering and mechanical engineer to be constantly changed and expanded. Although there is always a part of the curriculum that loses its importance, the knowledge expected of a mechanical engineer is increasing exponentially. There is also a broadening of knowledge, in addition to the knowledge of traditional mechanical structures, knowledge of the operation of electrical machines, or the knowledge and application of various information technology hardware and software are now expected.

How can the system of mechanical engineering training follow this expansion of knowledge, and how can the system incorporate it into itself? We conducted a comparative analysis of the subject matter of mechanical engineering training in 4 periods based on the available data from the Miskolc Heavy Industry Technical University:

- 1970/71. The period following the end of strong domestic industrialization [11].
- 1988/89 The period of completion of socialist planned management [12].
- 2004/2006 The period following the accession to the European Union. The period of European legal harmonization [13].
- 2014/15. The period of the last major educational reform [14].

In each of the 4 periods, we examined the evolution of the number of theoretical and practical hours in a semester. We also examined the distribution of the basic and other supplementary subjects, as well as the change in the total number of hours of different subjects. Since the total number of hours of the traditional mechanical engineering training and today's BSc training is different, for the sake of comparability, we compared the part of the total training time per semester and its development to establish trends. The typical average number of hours per semester is between 29 and 28.5 hours (Fig.3).

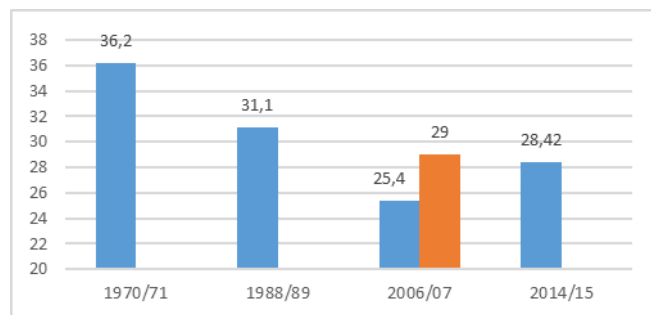


Fig. 3. The average total training time per semester in a week

Examining the four topics, a radical decrease in the number of lessons is typical for basic professional subjects. Computer technology appeared as a new basic subject in the

1988/89 academic year, but the number of hours of which, in contrast to the other basic subjects, increased in each new curriculum. The ratio of theory and practice, on the other hand, has increased in favour of the theory, which may be due to the increase in the number of students in the courses, the limitations of practical capacities, and the high cost.

VI. CONCLUSION, FURTHER RESEARCH DIRECTION

Based on the student questionnaires, it can be established that the students have general and inaccurate ideas regarding their education and the labour market opportunities related to their education. It would be important to eliminate these misunderstandings already in the first semester, for which the support classes provide a good opportunity. We consider it important that our students have accurate knowledge of their chosen profession and the possibilities inherent in it.

We can be sure that technical progress will not slow down. The knowledge material of the technical sciences will expand more and more quickly and to an ever-greater extent, which should be able to be passed on to the students during the training to meet the expectations of the labour market even as fresh graduates. It is also certain that due to the inertia of the training system, many contents that are already considered "outdated" or that have lost their practical significance due to changes in technology remain in education. Taking all of this into account, there are three alternatives for solving the problem.

1. Increasing the training time, i.e., longer time to transfer the increased knowledge material. This option is not possible due to the current legal framework. The duration of the BSc course is 7 semesters. All content must be condensed into this.
2. Increasing the training time through specialization. In this case, during the available 7 semesters, the currently basic subjects would be taught, at the end of which the student would obtain a general engineering degree. After that, the specialization would take place and the content of the specialization would be taught, like the current MSc training.
3. The third option is content selection. The selection factor can be, among other things, the use value of the knowledge, or the overlap, which is the knowledge

that can be acquired even while studying modern techniques and technologies and does not necessarily require training in a separate class time frame. Using the results and findings of this study, we would like to investigate the third alternative in the further parts of the research.

The next planned step of the research is the mapping of workplaces and jobs requiring a degree in mechanical engineering in a geographically limited area (Székesfehérvár and its surroundings), which aims to determine the professional expectations, knowledge, and skills related to the job. Evaluation of the results, a compilation of proposals for inclusion in the training system of the Alba Regia Technical Faculty of the Óbuda University.

REFERENCES

- [1] Fernandez-Miranda, S., Marcos, M., Peralta, M., Aguayo, F. (2017): The challenge of integrating industry 4.0 in the degree of mechanical engineering. *Procedia Manufacturing*, 13:1229–1236.
- [2] The Boston Consulting Group (2015): *Industry 4.0: The future of productivity and growth in manufacturing industries*.
- [3] Beszédes, B., Széll, K., Györök, Gy. (2020): A Highly Reliable, Modular, Redundant and Self-Monitoring PSU Architecture. *Acta Polytechnica Hungarica* 17 : 7 pp. 233-249. , 14 p. (2020)
- [4] Fomunyam, K. G. (2019): Education and the fourth industrial revolution: Challenges and possibilities for engineering education. *International Journal of Mechanical Engineering and Technology (IJMET)*, 10(8), 271-284.
- [5] Kulacki, F.A. (1996): The Education of Mechanical Engineers for the 21st Century. *JSME international journal. Ser. A, Mechanics and material engineering*, Volume 39, Issue 4, Pages 467-478.
- [6] Dankó, C.T. (2019): Important competencies in the engineering and IT world. URL: <https://gramontinternational.com/hu/fontos-kompetenciak-a-mernok-es-it-vilagban/> (Last download: 06/19/2022)
- [7] Szász, G. (1994): The technical level of socialist-oriented countries 1945–1990. The F4632. s. A study was carried out in the framework of OTKA.
- [8] Zhang, C., & Yang, J. (2020): *A history of mechanical engineering*. Springer.
- [9] Révai (1911–1935): *Révai Great Lexicon*, Budapest, Révai Brothers Literary Institute Joint Stock Company.
- [10] Tarsoly, I. (ed.) (1996-2000): *Hungary in the XX. century*, volume IV. Szekszárd, Babits Publishing House.
- [11] Technical University of Heavy Industry: *Yearbook 1970/71*, Miskolc.
- [12] Technical University of Heavy Industry: *Yearbook 1988/89*, Miskolc.
- [13] Technical University of Heavy Industry: *Yearbook 2004/06*, Miskolc.
- [14] Technical University of Heavy Industry: *Yearbook 2014/15*, Miskolc.

Online examination from the point of view of teachers

Nikoletta Tolner
Alba Regia Technical Faculty
Obuda University
Szekesfehervar, Hungary
tolner.nikoletta@amk.uni-obuda.hu

András Dávid
Alba Regia Technical Faculty
Obuda University
Szekesfehervar, Hungary
david.andras@amk.uni-obuda.hu

Judit Módné Takács
Alba Regia Technical Faculty
Obuda University
Szekesfehervar, Hungary
modne.t.judit@amk.uni-obuda.hu

Monika Pogátsnik
Alba Regia Technical Faculty
Obuda University
Szekesfehervar, Hungary
pogatsnik.monika@amk.uni-obuda.hu

Abstract— *As a result of the COVID19 pandemic, education has been forced into the online space. The delivery of the course material and the assessment of the students' acquired knowledge also had to be solved via online interfaces. At Óbuda University, we used the MOODLE system and the MS Teams application to implement online education and examinations. During the research, we got to know the opinions and experiences of the teachers regarding online examinations with the help of an online questionnaire. We were looking for answers from teachers of higher education about their experiences in the online exams:*

- *how suitable the instructors found it to be*
- *would instructors use it even when it is no longer mandatory*
- *what were its positives and negatives*
- *how the teachers made sure of the identity of the students taking the exam*
- *what methods the teachers used to prevent the use of unauthorized devices*
- *from the teachers' point of view, what technical problems arose during the online examinations*

Exams in the online space presented the institutions with many problems to be solved. During the research, we wanted to explore these by interviewing the teachers. In our quantitative research, we asked the teachers of Óbuda University. The online questionnaire received 61 responses in the 2nd semester of the 2021/22 academic year.

Keywords— *online exam, higher education, views of instructors.*

I. INTRODUCTION

Online tools were used in higher education even before the turn of the millennium, but they were not yet widely accepted [1]. Online education played an important role due to the impact of the COVID-19 pandemic in the 2020s. Since then, several studies have dealt with the methodological issues of online education, with the impact this form of education had on the lives of universities and students [2] [3] [4] [5] [6].

This completely new situation necessitated the application of new methods and techniques in the field of examinations as well. Many articles have been published on the subject, but we still do not have a sufficient amount of data on their correctness and applicability. It was important during this period to share our experiences. [7] The epidemic situation demanded changes in all areas of life,

and education was no exception. In a very short time, it was necessary to switch from face-to-face education to online form [8]. Conducting the exams online required a lot of preparation from the instructors, and several problems had to be solved [9] [10].

In the course of our research, we asked the teachers' opinions about online exams, and in the essay, we systematized and analyzed the answers we received. The questionnaire contained 19 questions, which were open-ended and closed-ended. There were questions where the answers had to be marked on a 5-item Likert scale. Except for two questions, where the email address can be entered, and where we asked about specific technical problems, it was mandatory to answer the other questions.

II. ANSWERS TO THE QUESTIONNAIRE

A total of 533 teachers work at the 7 faculties of Óbuda University, of which 61 filled out the questionnaire. The staff of 61 also includes external teachers. We also took these into account when analyzing the data, since they used the university systems and filled out a questionnaire related to their university subject. If external teachers are not counted, then university teachers have approx. 10% filled out the questionnaire. This is a relatively low turnout. Among the faculties, the Alba Regia Faculty of Technology received 26 answers, which represents 42.6% of the answers.

A. DEMOGRAPHIC DATA

Among the respondents, the proportion of women and men is 42.6%-57.4%. Regarding the age distribution, 60.7% of the teachers belong to the 43-57 age group. The second largest group is the 28-42-year-olds with 21.3%. Those aged over 58 make up 18% of the respondents, and not a single answer was received from the youngest age group. Two teachers from the group over 65 filled out the questionnaire. For the age groups, we followed the division according to generation. Based on the number of examinees, the picture is very mixed. Groups of over 90 people are quite rare, which was also since most of the respondents came from the Technical Faculty of Alba Regia, and this is where the smaller groups are typically found. Up to 90 people (for groups growing by 10 members), the distribution of groups is very similar.

B. DISTRIBUTION BY EXAMINATION TYPE

As for the type of exam, typically for exams with more than 50 people, the written exam was used, or perhaps a verbal exam was added. For smaller groups, the verbal exam was typical. The majority of respondents, 32.52%, chose the verbal exam, 20.33% preferred the written exam, and 9.15% used a mixture of verbal and written exams.

C. INTERFACE OF THE EXAMINATION

The university recommended Moodle and MS Teams for the exam interface, and this appears in the majority of answers (Figure 1). Moodle was chosen by 55.7% of respondents and MS Teams by 45.9%.

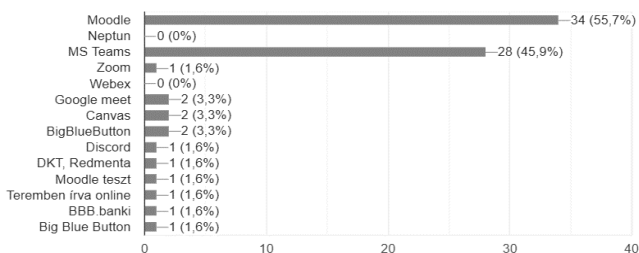


Fig. 1. Distribution according to the surface of the exam

D. VERIFICATION OF STUDENTS' IDENTITY

We expected individual answers to the question about checking the identity of the students. Most people relied on the Moodle login ID and password, which meant the Neptun ID and password since most people used Moodle as the interface. In the verbal exam, the teachers knew the students in the small groups, so no other identification was necessary. For larger groups, students were asked for some form of identification. There was only one teacher who did not deal with this.

E. EVALUATION OF THE ONLINE EXAM AND POSSIBLE FUTURE USE

4.9% of the respondents found the online exam not good at all and 6.6% rather bad (Figure 2). They would not use the online exam later either.

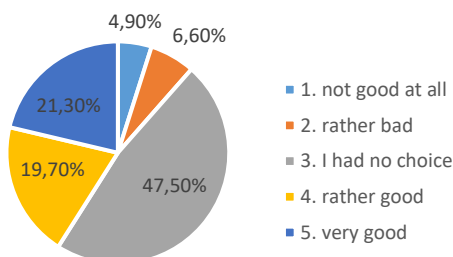


Fig. 2. Distribution according to the favorable opinion of the online exam

Among those who answered that they had no other choice, 73% would not use this form of examination in the future. Of those who answered 4 to this question, only one person would not use this form of examination later. Anyone who liked the online exam will definitely use it in the future, with one uncertain exception. 47.5% of respondents would not use the online exam in the future, 14.8% were unsure, and only 37.7% would use it later.

F. TECHNICAL PROBLEMS

50.8% of the respondents had no technical problems during the online exam, and 44.3% had only minor problems that they could solve themselves. Perhaps this is not surprising from the point of view of technical education. The majority of those with problems drew attention to the shortcomings of the Internet connection. Some instructors mentioned that they had problems with online exams because of Moodle or MS Teams. There were problems with the sound and picture quality, but not in a particularly large number. They would not use the online exam not because of technical problems, but rather because of the lack of personal contact.

G. ONLINE EXAM PREPARATION

Those who needed much less preparation (Figure 3) for the online exam than for the in-person exam would stay with this form of exam later on, and they had no technical problems. Those who needed the same amount of time to prepare for the two types of exams had very few technical problems, yet they wrote that they no longer wanted to use the online exam. A greater proportion of those who require more preparation would choose the personal exam. More than half of the respondents (54.1%) needed more time to prepare for the online exam than for the in-person exam.

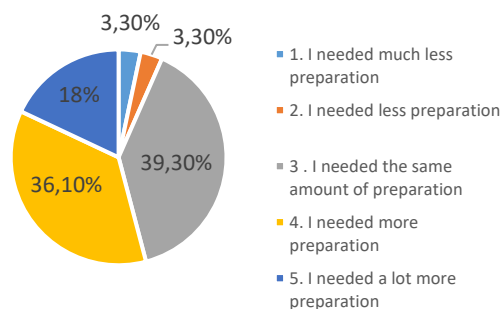


Fig. 3 Distribution of answers based on preparation for the online exam

H. STUDENTS' PERFORMANCE

The teachers who answered that the students performed worse - this is 6,5% of the respondents - in the online exam compared to the personal exam (Figure 4), all needed more preparation. Not sure if they would use this exam in the future. The teachers who thought that the students performed much better (13,1%) all took the exam in the form of a test.

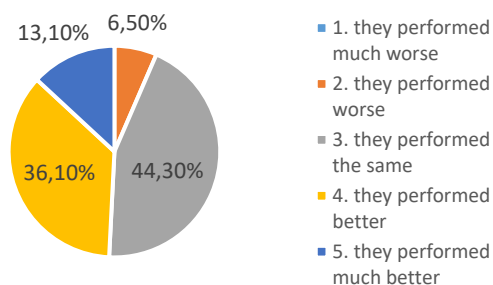


Fig. 4. Distribution based on students' performance

I. APPLIED EXAMINATION METHODS

Among the examination methods used, the most popular are verbal responses via the online interface and the test (Figure 5). This result is also due to the fact that most of the instructors conducted the online exams using Moodle and MS Teams. In general, it can be said that the instructors used several methods together, the verbal response via the online interface was the only method where no other method was already used.

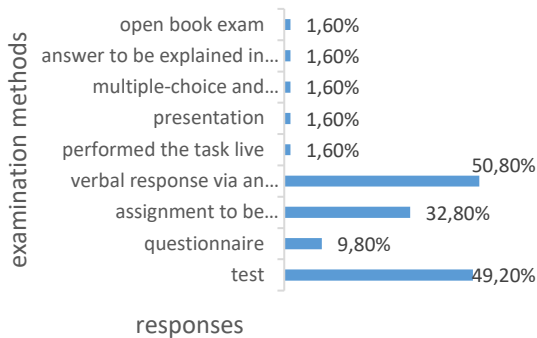


Fig. 5. Distribution according to examination methods

J. METHODS USED TO PREVENT CHEATING

In order to avoid cheating, in the case of tasks and tests to be solved during the exam, most people decided to leave a short time for the solution (Figure 6). Almost no one singled out just one solution, rather a combination of them was used. The fact that the exam was mainly about practical questions was also very popular. After that, the instructors used continuously cameras, and name-based tasks and randomly asked questions in almost the same proportion.

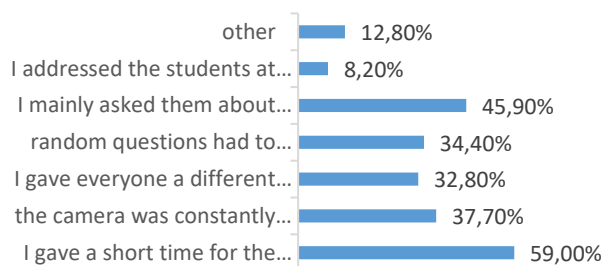


Fig.6. Distribution according to the methods used to prevent cheating

K. ADVANTAGES AND DISADVANTAGES OF THE ONLINE EXAM

Figure 7 shows the negatives that the teachers experienced with the online exam. The listed reasons were highlighted according to the frequency of the answers. Only 24% of the teachers think that there were no negative aspects of the online exam. Among the negative qualities, 43% of the respondents singled out cheating, the use of other aids, and the possibility of group work as the most critical problem. Surprisingly, 32% of teachers highlighted the impersonal nature of this format as a negative. 27% of the respondents felt that the preparation was much more time-consuming compared to the traditional form of education, but many also added that this would probably only mean extra work in the initial period and that it would improve in

the long term because the prepared tests and tasks in the future can also be used.

Details from the teachers' answers:

- "Lack of personal contact"
- "I am not sure that the result reflects the actual knowledge."
- "Although it was not necessary to prepare more for the verbal exam than before, the preparation of the tests took a lot of time."



Fig. 7. Disadvantages of online exams

20% of the respondents concluded that there is no positive effect of the online exam. Surprisingly, this position was taken by the younger generation, as 45% of these respondents belong to Generation Z, while the remaining 55% belong to Generation Y. Figure 8 shows the advantages of the online exam from the instructor's point of view, also with the help of a word cloud.



Fig. 8. Advantages of the online exams

According to 46% of the respondents, the most advantageous aspect of this form was the comfort, within which the time savings associated with travel, the stress-free environment for teachers and students, and the flexibility were highlighted. 65% of the respondents are from Generation Y. Of those who highlighted this advantage, 35% of them would not gladly use the form of the online exam despite this advantage. 30% of teachers enjoyed the quick, automatic assessment option. Unfortunately, only 11% of the respondents highlighted as a positive the possibility of getting to know and trying different methodologies, which would be very important from an educational point of view. Surprisingly, 60% of those who have drawn attention to this advantage are of the Baby Boomer generation. Details from the teachers' answers:

- "It could be solved from home. Most of the tasks were evaluated automatically, so I spent less time on corrections."
- "There was no need to waste paper."
- "My digital competencies have improved."
- "customizable time frame - in the case of correspondence students, the exam can be organized at the most suitable time for the student."

III. 3. EXAMINATION OF CONNECTIONS

Figure 9 shows that 60% of those who filled out the questionnaire belong to the Y generation. Among the advantages of the online examination, 72% highlighted convenience, flexibility, the possibility of simple and quick evaluation, and the possibility of automation. Based on these, it is not surprising that almost 70% of this age group chose the written exam.

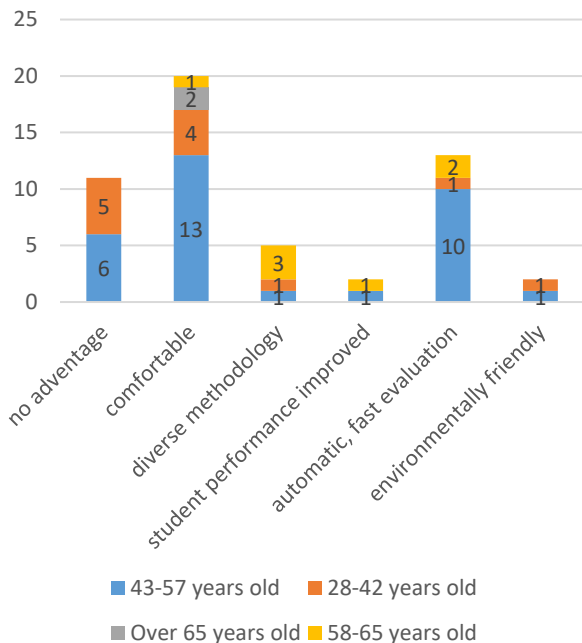


Fig. 9. Generational distribution of the positives of the online exam

Examining the methods used by instructors to check students' identity, we found $r=-0.731$, a strong relationship, which indicates an opposite orientation. So those who do not find the user authorization management options of the platform they use to be sufficient, used video-based, face image, and voice identification instead. Also, the correlation between these two measured variables may point to the exact opposite, that if the possibility of facial verification was not provided or adequate, then they used authorization-based verification of the applications they used.

If we examine the correlation between not checking the identity of the students and the better performance experienced when conducting the online assessment, we find that the correlation is $r=0.342$. This indicates a medium-strength relationship with unidirectional co-movement. So, if the instructors did not find a suitable way of checking, or if they could not use the opportunities provided by digital tools, they believed that it could be concluded that the students had an advantage and thus better results were achieved. On the other hand, looking at the correlation between the more time-consuming preparation and the possibility of joint work as negative factors, $r=-0.367$. From this, it can be concluded that there is a moderately strong correlation in the opposite direction between these two factors, meaning that it was not the efforts to conduct the exam safely that required more time for preparation.

	comfortable	automatic fast	would rather use it	would definitely use it
comfortable	1	-0,444	-0,328	0,270
automatic fast	-0,444	1	0,320	-0,024
would rather use it	-0,328	0,320	1	-0,245
would definitely use it	0,270	-0,024	-0,245	1

Fig 10. Correlation coefficients between the positives highlighted during the online exam and future application

Based on the analysis of the coefficients (Figure 10), $r=-0.444$, comfort, flexibility, home atmosphere, and a stress-free environment as positive characteristics and the positive aspect of using online applications capable of automatic evaluation are moderately strongly related, indicating an opposite relationship in terms of direction. If we add the non-forced use of these variables to the examination in the future, it can be established that the correlation between the automatic evaluation provided by the technique, faster repair, and the future application is $r=-0.024$, so there is no linear relationship, these two characteristics are uncorrelated. So, it is not necessarily because of this feature that instructors would definitely choose the online exam format in the future. When examined with the variable "It would rather apply", $r=0.320$ indicates a medium-strength, one-way relationship.

A. ANALYSIS BY GENDER

For women, the average answer to the question of how nice the online exam was was 3.69, for men it was 3.29. Technical problems produced an average score of 1.57 for both women and men. Examining the preparation time compared to the traditional exam, we found that the average for women was 3.42, while that of men was (Figure 11)

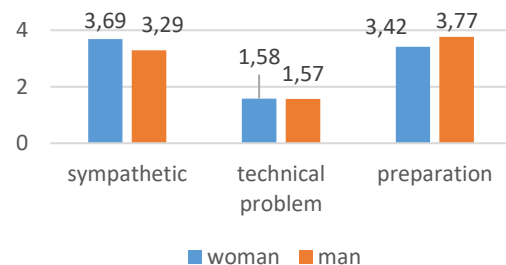


Fig. 11 Analysis by gender

Correlation coefficients:

- Gender-sympathetic: -0.192
- Gender-technical problem: -0.004
- Gender preparation: 0.186

Examining the relationship between a few more parameters:

- Sympathetic-technical problem: -0.138
- Sympathetic preparation: -0.446
- Technical problem preparation: 0.159

It can also be seen from the calculated correlation coefficients that there is no close relationship between the genders and the given aspects. The two parameters between which there is a moderately strong relationship are the preference of the online exam and the time required to prepare for the online exam, $r=-0.446$. This means that the less preparation someone needed to prepare for the online exam, the more enjoyable the online exam was for them.

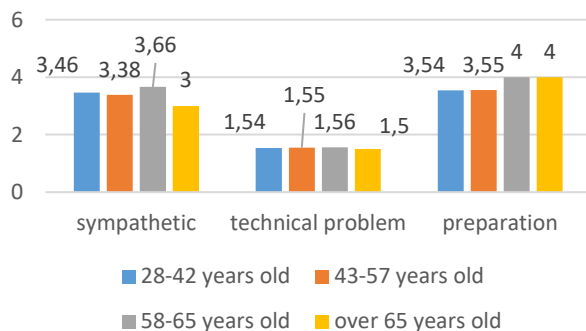
B. ANALYSIS BY AGE

Analyzing likeability, the average for the 28-42 age group is 3.46, for the 43-57-year-old group the average is 3.38, for the 58-64-year-old group the average is 3.66, while for the over 65-year-old group it is 3.0. The technical problems, an average of around 1.5 was found for all age groups. In the case of preparation, an average of 3.5 was obtained for the 28-42 and 43-57 groups, and 4.0 for the 58-64 and over 65 age groups.

Correlation coefficients:

- Between age and sympathy: 0.0
- Between age and technical problem: 0.0
- Between age and preparation: 0.1

Examining the correlation coefficients, we also conclude that there is no correlation between age and the given aspects (sympathetic, technical problem, preparation).



IV. SUMMARY AND FURTHER RESEARCH

During our research, we asked the teachers of Óbuda University about their experiences with online exams. During the analysis by gender, we did not find a strict correlation between the likeability of online exams, the technical problems that arise and the time required to prepare for online exams, and the gender of the respondent. Examining the previous aspects based on age, we again did not find a strict relationship between the respondent's age and the given aspects. Although these results may also be the result of low involvement. Among the negatives of the online exam, most people highlighted the possibility of cheating and the lack of personal contact. Among the positives of the online exam, a surprising number of people mentioned convenience, which in most cases meant getting rid of time-consuming traveling. Based on the responses we

received, it became clear to us that the majority of instructors did not like online testing, but discovered methods they would like to use in the future. Hopefully, in the future, we will have the opportunity to create a safe exam center that meets all needs, which will greatly contribute to increasing the acceptance of this form of the exam.

REFERENCES

- [1] Harasim, Linda (2000): Shift happens: Online education as a new paradigm in elearning. *The Internet and higher education*, 3(1-2), 2000. 41-61. p
- [2] Kopp Erika és Saád Judit (2021): A pandémia első hulláma a felsőoktatás-kutatások tükrében – Szakirodalmi áttekintés, *Neveléstudomány* 2021/3. szám 7-22. p.
- [3] Horváth László, Czirfusz Dóra, Misley Helga és N. Tóth Ágnes (2021): Alkalmazkodási stratégiák a távolléti oktatás során hallgatói, oktatói és intézményi szinten, *Neveléstudomány* 2021/3. szám 23-42. p.
- [4] Kálmán Orsolya, Eszes Fruzsina, Kardos Dorottya, Lénárd Sándor, Pálvölgyi Lajos és Szivák Judit (2021): Arctalanság a távolléti felsőoktatásban – Oktatók és hallgatók dilemmái és kritikus eseményei a távolléti oktatás első időszakában, *Neveléstudomány* 2021/3. szám 43-61. p.
- [5] Kálmán Orsolya, Eszes Fruzsina, Kardos Dorottya, Lénárd Sándor, Pálvölgyi Lajos és Szivák Judit (2021): A partnerség lehetőségei – Oktatók és hallgatók dilemmái a távolléti felsőoktatás második időszakában, *Neveléstudomány* 2021/3. szám 62-81.
- [6] Orosz Beáta – Molnár György (2021): A digitális átállás innovációs lehetőségei és kezdeti tapasztalatai. In: Kozma, Tamás; Juhász, Erika; Tóth, Péter (szerk.) *Társadalmi innováció és tanulás a digitális korban*, HERA Évkönyvek VIII, 2021, 266–277. p.
- [7] Prohászki Ágnes (2020): A tantermi és az on-line oktatás (tanítás és tanulás) összehasonlító elemzése, In: Benedek András (szerk), *Opus et Educatio* 2020, 7. évfolyam 3. szám)
- [8] Kálmán Botond – Tóth Arnold (2020): A covid-19 hatása a felsőoktatásra oktatói vélemények kérdőíves felmérése alapján, In: Korcsmáros E. (szerk.), *A J. Selye Egyetem 12. nemzetközi konferenciája*. Közgazdasági szekció, 209-225 p.
- [9] Vereckei Judit (2021): Személyes jelenlét nélküli felsőoktatás a vizsgaszervezés aspektusából a közoktatási vezetőképzésben, In: Benedek András (szerk), *Opus et Educatio* 2021, 8. évfolyam 2.
- [10] Beszedes, Bertalan; Széll, Károly; Györök, György (2021): Redundant Photo-Voltaic Power Cell in a Highly Reliable System *ELECTRONICS* 10 : 11 p. 1253 , 20 p.

Promotion of dual training form and its' tools at a large company

Zsuzsanna Felícia Pretz
University of Óbuda
Alba Regia Technical Faculty
Székesfehérvár, Hungary
pretz.zsuzska@gmail.com

Éva Takács
University of Óbuda
Alba Regia Technical Faculty
Székesfehérvár, Hungary
takacs.eva@amk.uni-obuda.hu

Abstract— The theoretical goal of our thesis was to popularize the dual training form and support the need of a dual brand within the company, based on the research we have prepared, and to make a proposal for its practical implementation. In connection with the dual brand, we have set up several hypotheses, which were verified in the thesis. For example, that the students who are working at the company would have a need to create such a brand and would be happy to show their belonging to the company by wearing a dual uniform.

To prove our hypotheses, we have prepared a market research questionnaire, the target group of which was dual students currently working at the company. After evaluating the answers to the questionnaire, we have verified our assumptions with the help of statistical calculations. After the need became clear, we have gathered information about the possibility of physical implementation.

Our research achieved its goal, because according to the students' opinion, it is necessary to create such an image at the company. At the end of our thesis, based on the results of the research, as the first step of practical implementation, we have prepared a couple of plans for the dual uniform.

Keywords— market research, dual training, uniform, cohesion, merchandising

I. RATIONAL FOR THE CHOICE OF TOPIC

The primary goal of our thesis is to popularize the dual training form [1], and we present the results of our research as a means of this. Harman Becker Automotive Manufacturing Ltd. (Figure 2) currently employs 66 students who study at six different universities (Figure 1), so the student body within the company is heterogeneous.



Figure 1 University partners in connection with the company

In order to make the days spent at work easier, the Education Group places great emphasis on strengthening student cohesion and creating a homogeneous Harman student body. As a means of symbolizing this togetherness, I would like to present the dual uniform plan and its implementation process, which I based on the results of the questionnaire We prepared.

II. INTRODUCTION OF THE COMPANY

The parent company Harman International is present on several continents, such as Asia, America, Africa, and Europe.

As the Hungarian subsidiary of Harman Becker Automotive Systems Ltd., has been on the domestic market for 27 years now, it started operating in 1995, and is now known as a stronghold of dual training.



Figure 2 Logo of the company
Source: www.harman.com

The dual partnership is profitable for both the company and the universities. Dual higher education in Hungary enjoys state tax support and is becoming more and more popular among students applying to university. From the company's point of view, acquiring professional staff is an essential part of the long-term strategy. The direct means of this is supporting the studies of students studying in dual education, making offers after graduation, and then employing the students as quickly as possible.

III. THE IMPORTANCE OF UNIFORMS

Our research was conducted among the dual students at Harman, the central theme of which is the promotion of the training by creating a dual brand.

An integral part of this is the creation of a dual "uniform", uniform clothing. Starting from the analogy of university promotional clothing, we conceived the idea that students have a need to express their affiliation [2]. They also like to wear clothes promoting their university or their faculty. The physical projection of the sense of togetherness would also be important within the company, and this can be achieved in the simplest and most spectacular way with the same clothing. In addition to university uniforms, many organizations still use uniforms today (Figure 3) [3].



Figure 3 Figure Different uniforms nowadays [4]

After assessing the needs of the students, they clearly consider it a good idea, and they would like to create and introduce such an image at the company. Of course, wearing the dual uniform would not be mandatory for students working at the company, we would like them to proudly wear the designed garments and use the accessories. The "uniform" would not only manifest itself in functional clothing, but we have also listed several accessories among the options, which are useful in everyday life, but can still be part of the brand. Everyone knows the classic festival promotional products, such as hats, sunglasses, mugs, glasses, water bottles, pens, or stress balls.

From now on, let's consider the dual students as consumers who form the target group of the brand's products. First of all, we have to convince them why this is beneficial for them, since the biggest advertisement for the company is satisfied students, who become influencers (Figure 4) for the future dual students [4].



Figure 4 An influencer

Source: <https://marketingsoul.hu/irasaink/elolvasom/te-tudod-hogy-mi-az-influencer-marketing-60>

IV. DUAL UNIFORM AS AN IMAGE CAMPAIGN

Image campaigns do not promote the specific product or its positive qualities, the focus here is on the brand itself. In this case, the brand, Harman itself, would be dual training. The best-known tool in our hands is the use of image transfer during such a campaign. We want to create non-descriptive advertising, on the contrary. Our goal is relationship-building communication. We want to embed in the public consciousness the clear connection between the attribute and the brand. This is image transfer (Figure 5). In this case, the attribute is the dual training itself, while the "brand" is Harman. The logic used by image transfer goes back to the basic logical relationship used during programming, the inference logic if-then [5].

IF this attribute (IF dual training)	→	THEN this brand (THEN Harman)
↑	↻	↓
THEN this attribute (THEN dual training)	←	IF this brand (IF Harman)

Figure 5 If-then logic for image transfer

V. MARKET RESEARCH WITH QUESTIONNAIRE

University major	Years																			
	2016		2017		2018		2019		2020		2021		2022		2023		2024		2025	
	I.	II.	I.	II.	I.	II.	I.	II.	I.	II.	I.	II.	I.	II.	I.	II.	I.	II.	I.	II.
Electric engineer	2 people		2 people		8 people		5 people		4 people		4 people									
			1 person		2 people		5 people		4 people		2 people									
					2 people		3 people		1 person											
					2 people		8 people		5 people		2 people									
Mechanical, Mechatronics engineer			5 people		2 people		1 person													
					4 people		2 people													
Technical manager					2 people		3 people		1 person											
IT engineer					5 people		2 people													
					2 people		1 person													
Finance and Accounting					1 person		2 people		5 people											
Economy					2 people		5 people													
					2 people		1 person													
Human Resources					2 people		1 person													

Figure 6 Distribution of dual students at the company

The above table shows the distribution of students according to their university majors (Figure 6), as well as the expiration date of their training period. We created a 10-question questionnaire [6] for dual students, which was completed by 64 people, thus the completion rate is 96.97%. Below we present the evaluation of the responses to the questionnaire and their results, followed by stochastic relationship analyses, which prove that it is necessary to create a dual image at the company. In the following, we will support the motivational effect associated with the creation of the dual brand with the results of the questionnaire.

A. Question #1

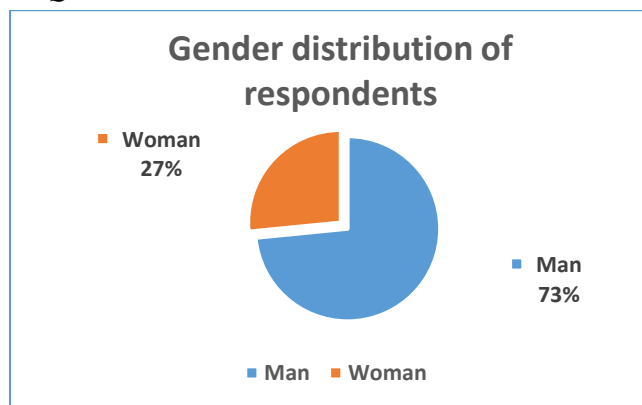


Figure 7 Gender distribution of respondents

The first and perhaps most important question referred to the gender distribution of the respondents (Figure 7). Considering the profile of the company, we can assume that fewer female students are employed by a company with a technical orientation. Based on the questionnaire, this was confirmed, as the gender distribution of the 64 respondents was 17 women and 47 men. So, slightly more than a quarter of the student population is female.

B. Question #2

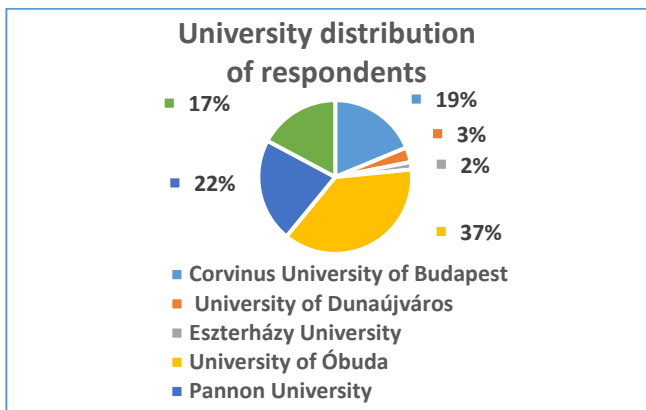


Figure 8 University distribution of respondents

Based on the answers received here, most students (24 people) joined the dual education program in connection with University of Óbuda. On the second place is Pannon University with 14 people, followed by Corvinus University of Budapest with 12 and then the University of Pécs with 11 people. The University of Dunaújváros and the Eszterházy University are also present with a small number of partners, which together represent 5% of the students (Figure 8).

C. Question #3

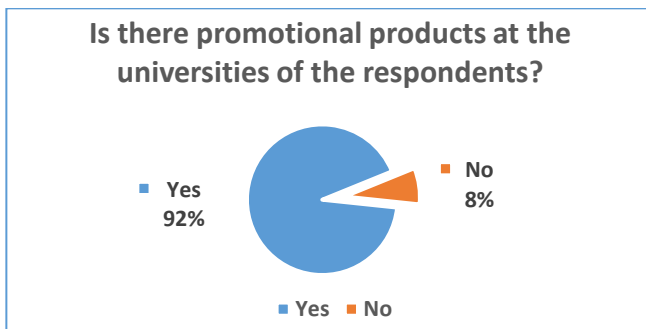


Figure 9 Does the applicants' university have promotional products?

According to 92% of those who filled in, their university has promotional products, the number of those who did not answer was 5 (Figure 9). The explanation for this is that they are probably not aware that such products are available at their university at all. So, all the universities partnering with Harman have promotional products.

D. Question #4

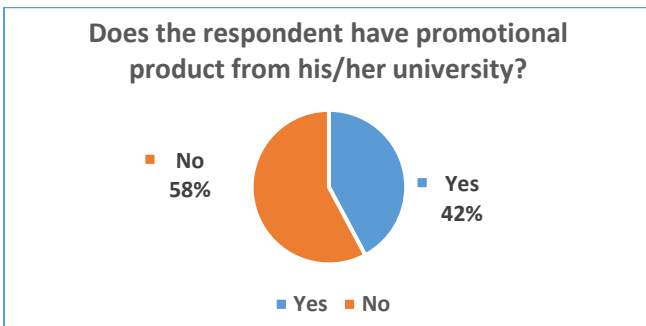


Figure 10 Does the applicant have a product promoting his/her university?

In terms of the number of responses, 27 people wear or use articles promoting their university/major. Unfortunately, non-respondents are in the majority (Figure 10), which is probably explained by the cost implications of these products since they can mostly be purchased at universities through online shops. It would certainly be motivating for most students to have access to such articles for free.

E. Question #5

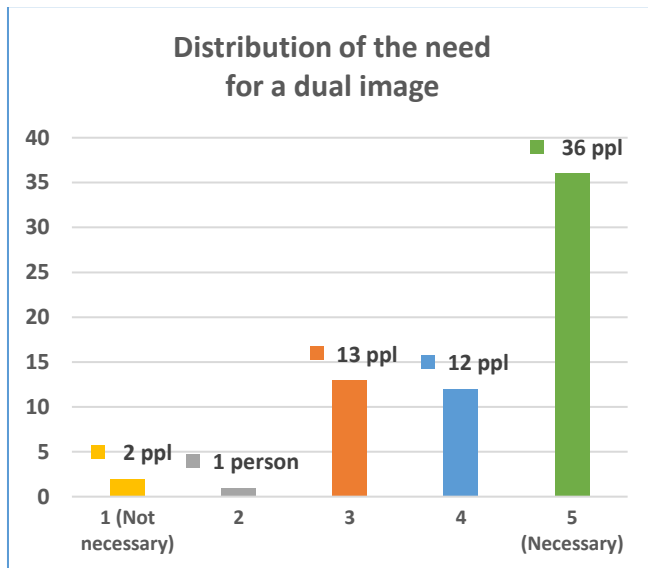


Figure 11 Do you think it would be necessary to create such a promotional image for students at Harman as well?

The answers showed (Figure 11) that the research made sense, as there was mostly positive feedback about the idea. Out of the 64, 48 gave a rating of four and five, and 13 gave a neutral answer.

F. Question #6

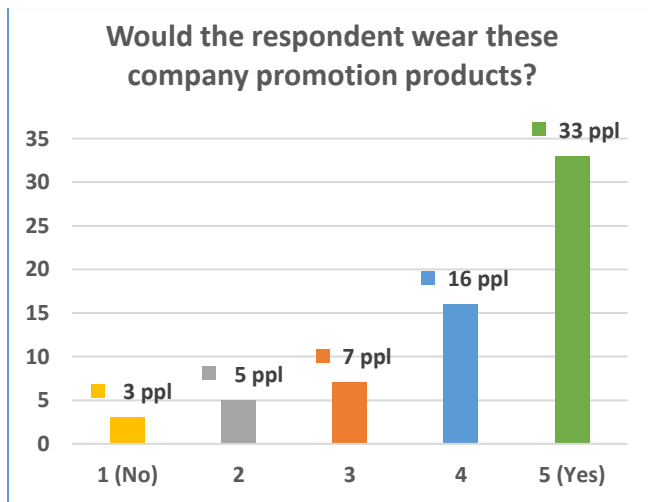


Figure 12 Would You wear these products?

From this diagram (Figure 12), we can read that most students would wear Harman's dual brand products, a total of 49 people out of 64, which is 76.56%. 7 students abstained, 3 refrained from wearing it.

G. Question #7

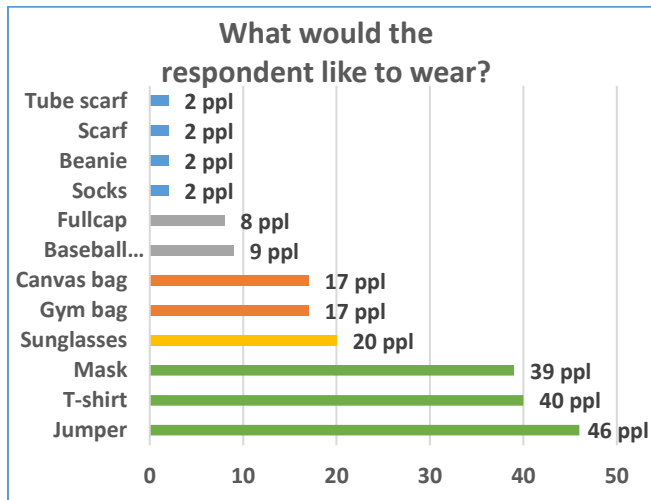


Figure 13 What would you like to wear?

Among the answers I gave in advance, the sweater, T-shirt and mask became the most popular (Figure 13). However, the students also expanded the range with their own creative ideas, there would be demand for socks, beanie hats and different scarves among the products of the dual brand.

H. Question #8

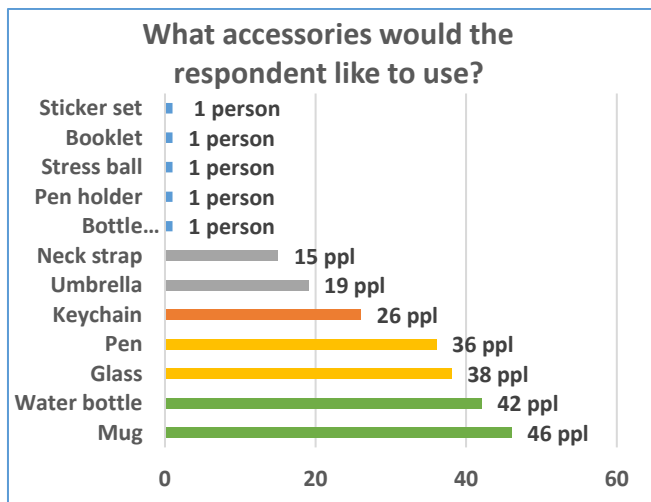


Figure 14 What accessories would you like to use?

In terms of accessories, the three most popular were mugs, water bottles and glasses. Among the answers added by the students were a beer opener, a pen holder, a stress ball, a notebook, and a set of stickers to put on the phone. These products can all be part of the dual product range (Figure 14).

I. Question #9

The company currently has promotional products, but it was important for us to find out what should be on the products so that they not only advertise the company, but also the dual education itself.

In this case, the top three answers were the abbreviation of the university, the #DUAL notation and the student's name options. In addition to these, the people filling in also added the name of the section (e.g.: HR, TEST), Harman word cloud with information about the company (Figure 15).

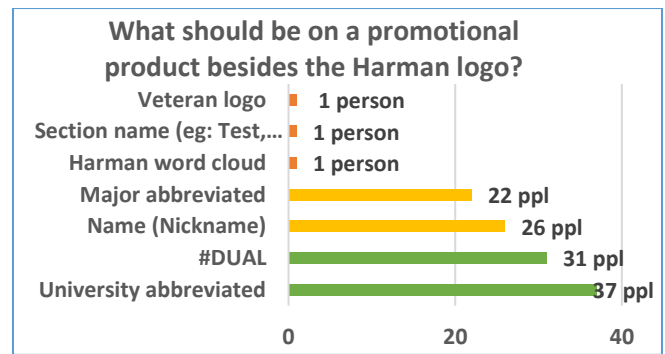


Figure 15 What should be on such a product besides the company logo?

I consider the most creative proposal to be the veteran logo, which would be deserved by the students who have already graduated. It is very positive that the students want to indicate that they once worked at the company in dual training even when they become "veterans". This can have an encouraging effect on the students who are still in training, as they could see how many former students work within Harman's ties even after graduation.

J. Question #10

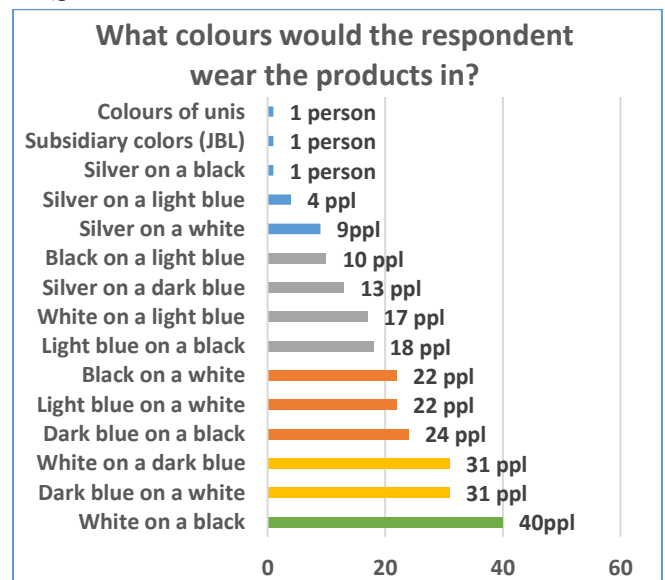


Figure 16 What colors would you like to wear the products in?

When choosing the colours, we combined the colours characteristic of the company, these are dark blue, light blue, white, black, and silver. It is important that the dual brand complements the corporate brand, so we didn't have to think much when choosing the defining colours (Figure 16). A white combination on a practical black base tops the list. There was a tie for second place between the options dark blue on a white background and its opposite, white on a dark blue background, with 31 responses to both. The third place was taken by dark blue on a black background, but we will have to think about this combination later, as it is not certain that the two colours would be able to prevail in relation to each other. Added answers include university and affiliate colours. However, in my opinion, these would not fully fit in to achieve the desired effect, since the university colours are used in the products distributed by the student's university.

VI. RESULTS OF STOCHASTIC RELATIONSHIP ANALYSIS

Based on the data extracted from the questionnaire related to dual image building, we examined 3 association relationships using the tools of stochastic relationship analysis. [7] Yule and Chuprov association coefficients were used to measure the closeness of the associations. In the case of stochastic relationship analyses, the y value must fall between 0 and 1, closeness between 0-0.3 is weak, between 0.3-0.7 is medium and between 0.7-1 is strong.

A. First hypothesis

In the first round, we examined to what extent the gender of the respondents is related to whether they have products that promote their university (Figure 17). We used the premise that these products are more popular among female fillers, however, the calculation below (Figure 18) shows that there is no strong relationship between gender and disposition.

Gender \ Ownership	Have	Does not have	Σ
Female	9	8	17
Male	18	29	47
Σ	27	37	64

Figure 17 The relationship between gender and disposition

Calculated from the values in the table above, we obtained the coefficient shown below:

$$y = \frac{f_{11} * f_{22} - f_{12} * f_{21}}{f_{11} * f_{22} + f_{12} * f_{21}}$$

$$y = \frac{9 * 29 - 8 * 18}{9 * 29 + 8 * 18}$$

$$y = 0,28$$

Figure 18 Application of Yule's association coefficient

As a result of the calculation, we can see that there is no close relationship between the investigated parameters.

Based on the values obtained here, it can be stated that there is a weak relationship between gender and provision among the respondents. This way we know that students prefer the products regardless of their gender.

B. Second hypothesis

University \ Affinity	1	2	3	4	5	Σ
BCE	0	1	5	2	4	12
DUE	0	0	0	0	2	2
EE	0	0	1	0	0	1
ÓE	0	0	5	4	15	24
PE	1	0	1	4	8	14
PTE	1	0	1	2	7	11
Σ	2	1	13	12	36	64

Figure 19 The relationship between university and affinity

Our second hypothesis was that university affiliation and affinity for promoting image are closely related. We examined this relationship with the help of Chuprov's association coefficient, which required the calculation of the marginal frequencies (Figure 20) and the values of the table above (Figure 19).

University	Marginal frequency $N_{ij} = \frac{n_i * n_j}{n}$	Chi value $X = \frac{N_{ij} - N_{ij}^*}{N_{ij}^*}$	University	Marginal frequency $N_{ij} = \frac{n_i * n_j}{n}$	Chi value $X = \frac{N_{ij} - N_{ij}^*}{N_{ij}^*}$
BCE1	$N_{ij} = \frac{2 * 12}{64} = 0,375$	$X = \frac{0 - 0,375}{0,375} = 0,375$	ÓE1	$N_{ij} = \frac{2 * 24}{64} = 0,750$	$X = \frac{0 - 0,750}{0,750} = 0,750$
BCE2	0,188	3,507	ÓE2	0,375	0,375
BCE3	2,438	2,692	ÓE3	4,875	0,003
BCE4	2,250	0,028	ÓE4	4,500	0,056
BCE5	6,750	1,120	ÓE5	13,500	0,167
DUE1	0,063	0,063	PE1	0,438	0,721
DUE2	0,031	0,031	PE2	0,219	0,219
DUE3	0,406	0,406	PE3	2,844	1,196
DUE4	0,375	0,375	PE4	2,625	0,720
DUE5	1,125	0,681	PE5	7,875	0,002
EE1	0,031	0,031	PTE1	0,344	1,251
EE2	0,016	0,016	PTE2	0,172	0,172
EE3	0,203	3,129	PTE3	2,234	0,682
EE4	0,188	0,188	PTE4	2,063	0,002
EE5	0,563	0,563	PTE5	6,188	0,107
Σ K _{hi} =	13,205		Σ K _{hi} =	6,423	
Chi ² =19,628					

Figure 20 Marginal frequencies and χ values belonging to the relationship between university and affinity

The X² (chi square) value shown in the formula was obtained from the summation of the chi values of the previous table, 'N' means the element number, which can be seen in the lower right corner of the table, which is, by definition, the same as the number of fillers. 's' means the number of rows, in this case we also must consider table number 5, so the value of s is 6, the value 't' indicates the number of columns, which means 5 columns in the table. This is how the following formula (Figure 21) is composed, from which the Cuprov association coefficient can be calculated:

$$T = \sqrt{\frac{X^2}{N * \sqrt{s - 1} * \sqrt{t - 1}}}$$

$$T = \sqrt{\frac{19,628}{64 * \sqrt{6 - 1} * \sqrt{5 - 1}}}$$

$$T = 0,262$$

Figure 21 Application of Chuprov's association coefficient

In this case too, the value falls between 0 and 0.3, thereby proving that there is a weak relationship between university affiliation and the affinity of creating an image that promotes the company. Therefore, we cannot say that, for example, only the students of Óbuda University would support the creation of a dual image, since we can conclude from the calculation that the students at the different universities all support the image plan.

C. Third hypothesis

And finally, but not least, we would like to show with another association study the strength of the relationship between the gender of the participants and the willingness to wear the products of the dual image (Figure 22).

Again, we used the premise that female fillers would prefer to wear products related to the dual image or would prefer to use the tools of the image. In this case, too, the calculation was performed using the Chuprov association coefficient.

Gender \ Affinity	1	2	3	4	5	Σ
Women	1	2	2	5	7	17
Men	2	3	5	11	26	47
Σ	3	5	7	16	33	64

Figure 22 Relationship between gender and affinity

To calculate the coefficient, it is necessary to calculate again the marginal frequencies and values of the above table, which can be seen in the table below (Figure 23):

	Marginal frequency	Chi value
Female1	0,797	0,052
Female2	1,328	0,340
Female3	1,859	0,011
Female4	4,250	0,132
Female5	8,766	0,356
Male1	2,203	0,019
Male2	3,672	0,123
Male3	5,141	0,004
Male4	11,75	0,048
Male5	24,234	0,129
Chi²= Σ Khí		1,214

Figure 23 Marginal frequencies and χ values for the relationship between gender and affinity

By calculating the coefficient (Figure 24), there is a weak relationship in this case as well. In this way, we can state that there is no clear connection between gender and the wearing of promotional products. So, we can't say that the products of the dual image are worn only by female fillers, and male avoid them.

$$T = \sqrt{\frac{X^2}{N * \sqrt{s - 1} * \sqrt{t - 1}}}$$

$$T = \sqrt{\frac{1,214}{64 * \sqrt{2 - 1} * \sqrt{5 - 1}}}$$

$$T = 0,097$$

Figure 24 Application of Chuprov's association coefficient II.

Based on the calculations, it can be concluded that the corporate students support the introduction of promotional products, they would wear and use them, regardless of gender or university.

VII. FROM IDEA TO IMPLEMENTATION

We talked about the implementation of the idea with a Sr. HR Manager of the company, in which we were informed about the design and production process of the current corporate uniform. About their ordering process and the possibility of introducing a dual brand. We then made a couple of sketches of the dual uniform (Figure 25 and Figure 26).



Figure 25 Plan of dual promotional products I.



Figure 26 Plan of dual promotional products II.

VIII. SUMMARY

At the end of our scientific student thesis, we managed to present our research in accordance with our objective, which is worthy of promoting the dual education and the Harman company. With our article, we want to show that it is not enough to win over students before admission, but that it is necessary to constantly maintain their interest and create a sense of community within the company. The results of our research proved to be useful for the company as well, as the Education Group works hard to recruit students. The uniform appearance also lends credibility at the open days, and the individual signs would facilitate the recruitment process. Our research is also useful for the university, as they see that their students spend their days at the company in an environment where they receive sufficient support.

SOURCES:

- [1] "Óbudai Egyetem AMK Duális Képzés," [Online]. Available: <http://amk.uni-obuda.hu/index.php/hu/felvetelizoknek/dualis-kepzes>.
- [2] N. Goldschmied, P. Furley and R. Bush, "Taylor&Francis Online - Critical review of uniform color effects in sports," [Online]. Available: <https://www.tandfonline.com/doi/full/10.1080/1750984X.2020.1855668>.
- [3] "Profession.hu - Egyenruhák világa: földön, vízben, levegőben," [Online]. Available: <https://www.profession.hu/cikk/egyenruhak-vilaga-foldon-vizen-levegoben>.
- [4] "Marketing Soul," [Online]. Available: <https://marketingsoul.hu/irasaink/elolvasom/te-tudod-hogy-mi-az-influencer-marketing-60>.
- [5] I. Sas, Reklám és pszichológia a webkorszakban - Upgrade 3.0, Kommunikációs Akadémia, 2012.
- [6] Z. F. Pretz, "_Duális_ arculatépítés," [Online]. Available: https://docs.google.com/forms/d/1ochRXgW_oMIOAUWjdSapJjp5yZ-x3mpS_alV2Q7tjGE/edit.
- [7] L. Kóczy Á., "Statisztika I. - 4. előadás – Kombinációs táblák elemzése," [Online]. Available: <https://uni-obuda.hu/users/koczy1/letolt/stat1-7.hu.pdf>.
- [8] "Google Képek - Egyenruha," [Online]. Available: https://www.google.com/search?q=egyenruha&tbm=isch&ved=2ahUKEwjR9sOz1cX2AhXRRcAKHcU9AJcQ2-cCegQIABAA&oq=egyenruha&gs_lcp=CgNpbWcQAzIHCCMQ7wMQJzIHCCMQ7wMQJzIFCAAQgAQyBQgAEIAEMgUIABCABDIGCAAQBRAeMgYIABAFEB4yBggAEAUQHjIGCAAQBRAeMgYIABAFEB46CAgAELEDEIMBOgsIAB.
- [9] "Harman Akadémia - Általános dokumentumtár," [Online]. Available: <http://moodle.harman.com/>.
- [10] "Pólómánia - Pólótervező," [Online]. Available: https://polomania.hu/polotervezo?gclid=CjwKCAjwxZqSBhAHEiwASr9n9FZd6o5Mn8VfWbPxY3y7H2MH0PVI8UHRxKaKIaC0J6liHoXHBnuszqBoCyaWQAvD_BwE.
- [11] "Textsite.info - Egyenruha definíciója," [Online]. Available: <http://hu.textsite.info/Egyenruha>.
- [12] G. Kaszás, Reklámcsinálás - A kis adrenalinjáték, S. Péter, Ed., Budapest: HVG Kiadó Zrt., 2014.
- [13] B. Krenn, "Semantic Scholar - The effect of uniform color on judging athletes' aggressiveness, fairness, and chance of winning.," [Online]. Available: <https://www.semanticscholar.org/paper/The-effect-of-uniform-color-on-judging-athletes%27-of-Krenn/48780a4d4ad46e2619ad56f3db14554f9af85ed3>.