Robert Hanus
Damian Mazur
Christian Kreischer   *Editors*

Methods and Techniques of Signal Processing in Physical Measurements

Springer

Lecture Notes in Electrical Engineering

Volume 548

Lecture Notes in Electrical Engineering (LNEE) is a book series which reports the latest research and developments in Electrical Engineering, namely:

- Communication, Networks, and Information Theory
- Computer Engineering
- Signal, Image, Speech and Information Processing
- Circuits and Systems
- Bioengineering
- Engineering

The audience for the books in LNEE consists of advanced level students, researchers, and industry professionals working at the forefront of their fields. Much like Springer's other Lecture Notes series, LNEE will be distributed through Springer's print and electronic publishing channels.

More information about this series at http://www.springer.com/series/7818

Robert Hanus · Damian Mazur ·
Christian Kreischer
Editors

Methods and Techniques of Signal Processing in Physical Measurements

Editors
Robert Hanus
Faculty of Electrical and Computer
Engineering
Rzeszów University of Technology
Rzeszów, Poland

Damian Mazur
Faculty of Electrical and Computer
Engineering
Rzeszów University of Technology
Rzeszów, Poland

Christian Kreischer
Faculty of Electrical Engineering
Electrical Machines and Drive Systems
Helmut Schmidt University/University
of the Federal Armed Forces Hamburg
Hamburg, Hamburg, Germany

About This Book

This book presents selected issues of modern electrical metrology in the field of sensor technology, signal processing, and measurement systems. The presented articles discuss theoretical problems and applications regarding measurements in electrical engineering, mechanics, telecommunications, medicine, geology, as well as in the aviation and transport industries. This book provides researchers and practitioners with information on the current state of knowledge in these areas. It can also be a source of new ideas for further development and cooperation.

This book presents selected papers of the XXII International Seminar of Metrology "Methods and Techniques of Signal Processing in Physical Measurements" (MSM2018) held on September 17–20, 2018, in Rzeszów-Arłamów, Poland. The conference was organized by the Rzeszow University of Technology, Department of Metrology and Diagnostic Systems (Poland), and Lviv Polytechnic National University, Department of Information Measuring Technology (Ukraine).

Contents

Measurements of Gas Phase Velocity in Liquid Metal by Means of Ultrasonic Pulse-Echo Method

Artur Andruszkiewicz[1(✉)] and Kerstin Eckert[2]

[1] Wrocław University of Science and Technology, Wybrzeże Stanisława Wyspiańskiego 27, 50-370 Wrocław, Poland
`artur.andruszkiewicz@pwr.edu.pl`
[2] Faculty of Mechanical Science and Engineering, Technische Universitätv Dresden, George-Bähr- Strasse 3c, 01069 Dresden, Germany

Abstract. The subject of the article concerns the measurement of the gas phase velocity in the liquid metal by the means of ultrasonic pulse-echo method. Velocity measurements in the case of ultrasonic sensors in the bottom of the container and on its side wall are discussed. In the case of the bubbles swarm movement to the velocity measurement, the method of cross-correlation of the signals was presented. The article also described the main sources of errors in the pulse-echo method in velocity measurements.

Keywords: Liquid metal · Gas phase velocity · Ultrasonic pulse-echo method

1 Introduction

Two phase liquid metal-gas flows are present in many technological processes. Examples may be metallurgical processes of secondary metallurgy of steel. Degassing, purification or refining processes for liquid metal are directly related to its blowing with argon. The effectiveness of refining depend among others on the dimensions of rising bubbles, velocity and the area of their rising in liquid metal and the gas phase flow stream. Knowledge of these quantities can allow for optimization and control of metallurgical processes. In order to obtain a full data bank about rising gas bubbles, model studies of two phase flows in a liquid metal-gas systems are necessary. Among methods of studying two phase flows, developing very dynamically and taking on a high importance, are ultrasonic techniques. These include the Doppler method (UDV) and the pulse-echo method. The first one developed at the Forschungszentrum Dresden-Rossendorf allows the determination of both the velocity of rising gas bubbles and the velocity of liquid. Among the published papers using the UDV method should mention the work carried out by Eckert, Gerbeth and Zhang, the results of which are presented in the articles [1–4]. The second method – ultrasonic pulse-echo method – was developed at TU Dresden as a part of the SFB 609 project "Elektromagnetische Strömungsbeeinflussung in Metallurgie, Kristallzüchtung und Elektrochemie". Published papers on the use of this method can be cited as follows [5–7].

The aim of the article is to present methods of gas phase velocity measurements using the pulse-echo method and ultrasonic flaw detector. They were developed for a model reactor in which the liquid metal was a GaInSn eutectic and the gas-argon phase was pressed through a nozzle. Knowing the velocity of the rising gas phase in the liquid metal, it is possible to calculate its residence time and thus achieve the proper refining effect by the argonation.

2 The Principle of Using the Pulse-Echo Method in the Two Phase Flows

Ultrasonic pulse-echo method is widely used in non-destructive studying of the materials – ultrasonic flaw detection. It is the basic method of detecting discontinuities of materials, providing information about their location, dimensions or spatial orientation. Assuming that the gas phase is a discontinuity in the flow, the pulse-echo method can also be used for studying two-phase flows, including liquid metal-gas. The principle of measurements presented on the basis of the real echogram shown in Fig. 1 consists in determining the transit time of the ultrasonic wave pulse from the transducer to the rising bubble and back to the transducer. The broadcasting impulse from the transducer installed in the container wall, after passing through coupling layer and the front wall, is reflected from the gas bubbles and the rear walls of the container and as an echo it returns back to the transducer. In Fig. 1 the main echoes in the path of wave transit are marked. Knowing the time of tB echo transit from rising bubbles and ultrasonic wave velocity in the liquid metal and the container wall, it is possible to determine the position of the gas bubble in the liquid metal as well as to calculate the bubble velocity if several ultrasonic sensors are installed on the side wall of the container or in case of a sensor installed in the bottom of the container.
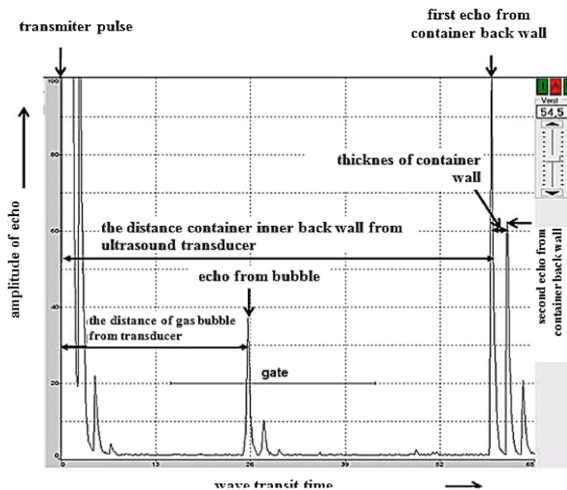
Fig. 1. The principle of measurement using the pulse-echo method.

3   Measuring Stand

The studies were carried out on the measuring stand shown in Fig. 2.
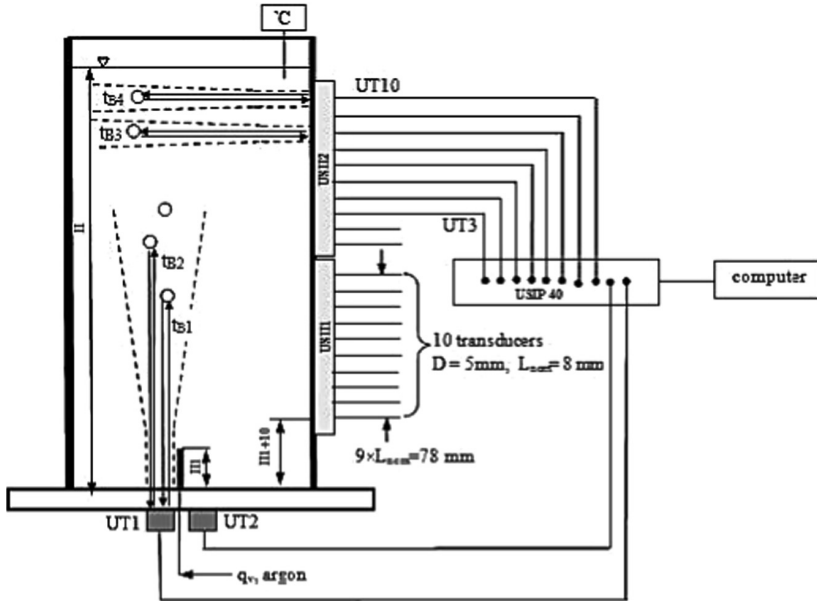
Fig. 2.   Scheme of the measuring stand.

A container with an inner diameter $Dw = 80$ mm was filled with liquid metal GaInSn to a height $H = 210$ mm. Argon in the form of gas bubbles was forced into the container through a nozzle, which height was $H1 = 20$ mm and the diameters of the nozzles depending on the measurement being carried out were: 0.5/0.7/0.9 and 1.2 mm. Argon volume flow qv was measured with a flow meter with a thermal sensor (Mass-Flo, MKS Instruments) class 1 with ranges: for the rising single bubbles 10 sccm, for a bubble chain 500 sccm. In order to determine the physical quantities of liquid metal, its temperature was measured with a Pt100 resistance thermometer.

Measurements of the transit time of the echo from the bubbles were made using two methods:

– Ultrasonic transducers installed on the side wall of the container or
– Ultrasonic transducers installed in the bottom of the container.

In the first method, the transducers were mounted in an ultrasonic head. The measurements were made using two heads (USH1, USH2) installed one above the other, consisting of 10 transducers each, with a frequency $f = 15$ MHz, diameter $D = 5$ mm, placed at nominal distances of 8 mm from each other. The first ultrasonic head (USH1)

was mounted 10 mm above the gas inlet nozzle. Ultrasonic heads were connected to a 10 channel USIP 40 Box ultrasonic flaw detector.

In the second method, two transducers (UT1 and UT2) were installed in the bottom of the container, while for the remaining 8 ultrasonic flaw detector inputs, transducers from the USH1 and USH2 were connected. Such a method of connections also allowed simultaneous measurement of the rising velocity with both methods and thus the control of measurement results at selected heights.

Table 1. Ultrasonic field parameters for GaInSn at 20 °C, transmitter diameter D = 5 mm and wave frequency 15 MHz

	GaInSn
Density ρ, kg/m^3	6361
Surface tension σ, N/m	0.533
Dynamic viscosity μ, kg/ms	2.2×10^{-3}
Ultrasonic wave velocity c, m/s	2748
Ultrasonic wave length λ, mm	0.420
Length of the near field lo, mm	34.3
Ultrasonic beam divergence angle φ	2.6

Table 1 presents basic physical quantities of GaInSn liquid metal with the parameters of the ultrasonic field produced in it.

4 Measuring Principle

The USIP 40 flaw detector, to which the ultrasonic transducers were connected, was controlled from a computer using the UltraPROOF program (GE Inpection Technologies GmbH). This program allowed for the digital recording of measurements signals with the resolution of the echoes transit time δt_e= 2.5 ns guaranteed by the flow detector manufacturer. At the same time it controlled the measurement time and the archiving of the results. The total measurement time was taken equal to T_m = 200 s. The cooperation of the flaw detector with the computer provided, in the real-time, on the computer monitor, a continuous observation of the amplitude of the echoes from all the reflectors in the path of the wave transit. This allowed for an optimal selection of the frequency of the transmitting ultrasonic pulses, the level of signal amplification and the position of measurement gate. It should be noted that the ultrasonic flaw detector records only those signals whose amplitude exceeds the gate's height from Fig. 1. After exceeding this height, the computing system records the wave transit time from the transmitter to the reflector and back to the transmitter. In the USIP 40 Box flaw detector it was possible to set the gate to any height, which with appropriate signal amplification increased the measuring range to bubbles with very small dimensions, whose echoes are small. In the studies, the height of the gate was set by controlling the level of

amplitudes from rising bubbles on the echogram and amplifying them so that their amplitudes exceeded the gate's height and at the same time were higher than the amplitudes of the echoes from any measurement disturbances. By adjusting the length of the gate and its position at a given height, the signal recording location from rising bubbles was determined. In the measurements, the length of the gate was adjusted so that for transducers mounted on the container wall, signals from the flowing bubbles in the area of ±30 mm from the argon inlet could be recorded. In the case of transducers mounted in the bottom of the container, the set length of the gate allows to collect signals from the bubbles on the distance about 190 mm from the bottom of the container.

The frequency of repeating ultrasonic pulses has been adopted:

– For configuration of 10 IFF transducers = 14286 Hz (for one transducer there is respectively IFF/10 = 1428.6 Hz)
– For a system with two transducers in the bottom of the IFF container = 6667 Hz (for one transducer there is respectively IFF/10 = 676.7 Hz).

These were the maximum frequencies that could be achieved on the wave path: transducer – back wall of the container – transducer or transducer – container filling height – transducer. The sampling period was $\delta t = 0{,}7$ ms or $\delta t = 1{,}5$ ms. The reproduction of the measurements signals and the calculation of the flow parameters were made using the DasyLab v10.0 program.

4.1 Rising Rate Measurements with Transducers in the Bottom of the Container

Figure 3 shows the picture of mounted ultrasonic transducers in the bottom of the container and the interpretation of signals from the rising bubble. Figure 4 shows an example of real recorded signals from rising argon bubbles in the liquid metal. At x_{B1} height, the wave transit time on the path transducer – bubble – transducer is t_{B1}. After time $\Delta\tau = t_2 - t_1$, the bubble is at x_{B2} height and the wave transit time is t_{B2}. By measuring these values, with the known velocity of the ultrasonic wave in the fluid, the rising velocity is determined from Eq. 1:

$$w_z = c \cdot (t_{B1} - t_{B2})/2\Delta t \tag{1}$$
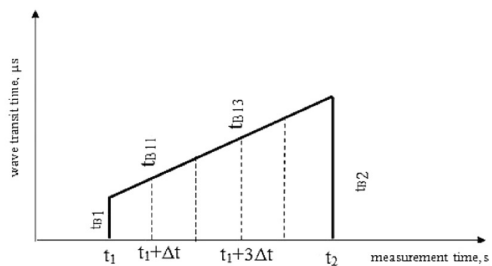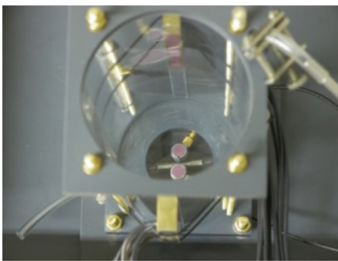
Fig. 3. Photo of transducers in the bottom of the container.
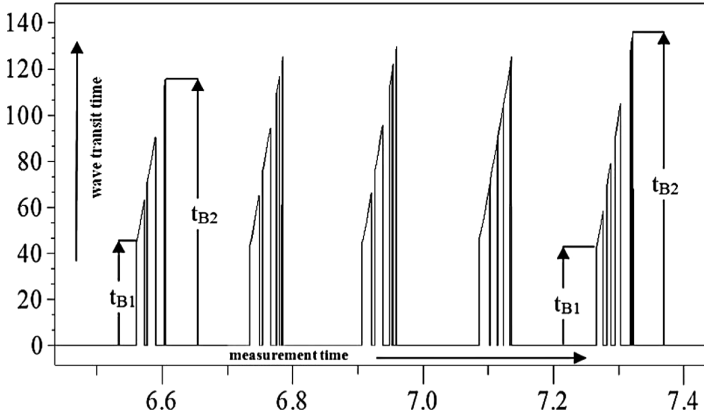
Fig. 4. Interpretation of signals from rising bubbles.

This method can determinate the velocity in the range of flow from individual bubbles to the flow of the bubble swarm. It is important to record the wave transit time from the same bubble at two different heights. The advantage of this method is that by dividing the measurement time $\Delta\tau = t_2 - t_1$ into intervals, for example, with an equal time step $\Delta\tau_i$ it is possible to determine velocity changes in the path of rising bubble. Figure 5 shows the changes of the bubble velocity on the rising path from $z_1 = 36,6$ mm to $z_2 = 85,5$ mm with the time step $\Delta\tau = 45$ ms on the background of the average rising rate calculated on this path.
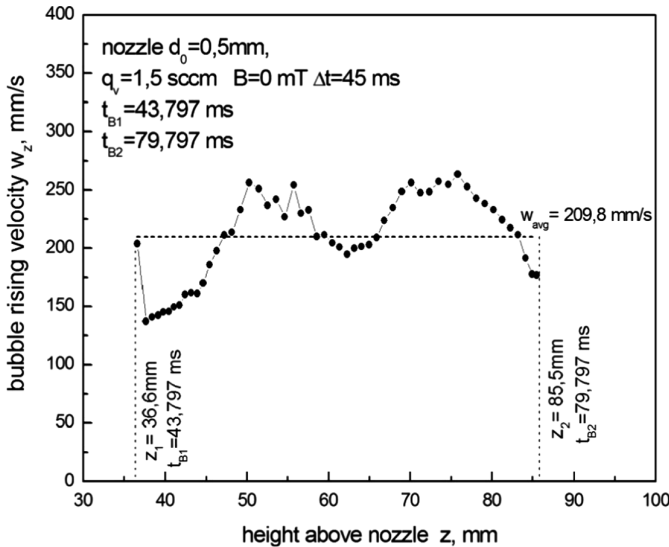
Fig. 5. The dependence of the velocity of rising bubbles on the flow path for the nozzle.

The range of changes in the speed of the bubbles ranged from $w_{zmin} = 137.1$ mm/s to $w_{zmax} = 263.5$ mm/s with the average rising velocity on this path $w_{avg} = 209.8$ mm/s.

4.2 Rising Rate Measurements with the Transducers on the Side Wall of the Container

A photo of the measuring stand with heads mounted on the side wall of the container with examples of actual signals from rising bubbles together with their interpretation is shown in Fig. 6.
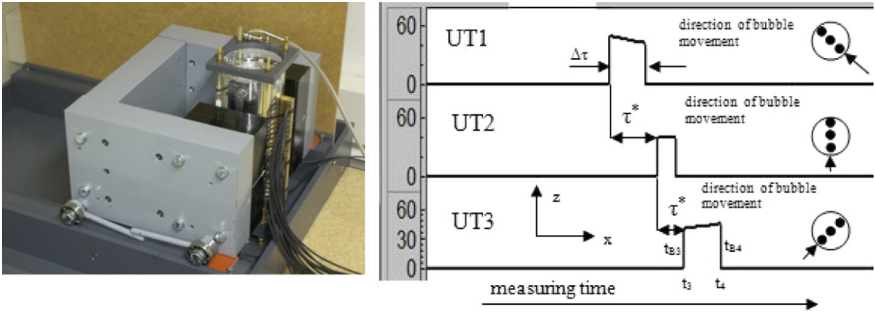
Fig. 6. Measuring stand with examples of real measuring signals: t_{B3}, t_{B4} – wave transit time from transducer to bubble and back to transducer, τ^*- bubble transit time from first to second transducer (time of signals delay), $\Delta\tau$–bubble transit time through ultrasonic field

The rising time is determinated from Eq. 2,

$$w_z = L/\tau^* \tag{2}$$

in which: τ^* - bubble transit time between transducers (time of signals delay),

   L- real distance between ultrasonic transducers.

In the case when bubble rising across the ultrasonic field (UT3 in Fig. 6) from the recorded signals it is also possible to determine the second component of velocity, in the x direction, according to Eq. 3:

$$w_x = c \cdot (t_{B4} - t_{B3})/2(t_4 - t_3) \tag{3}$$

Figure 7 shows an example of the dependence of bubbles velocity on the path of their rising path for a nozzle with dimensions up to $d_o = 1.2$ mm. The average speed and the standard deviation $\sigma(w)$, indicated in this figure, were calculated for 30 rising bubbles. The graphs clearly shows changes in the velocity of rising bubbles along the rising path which indicates that the path of their movement is not a straight line and the bubbles rising spirally or zigzag. An exemplary movement path for three selected bubbles is shown in Fig. 8.
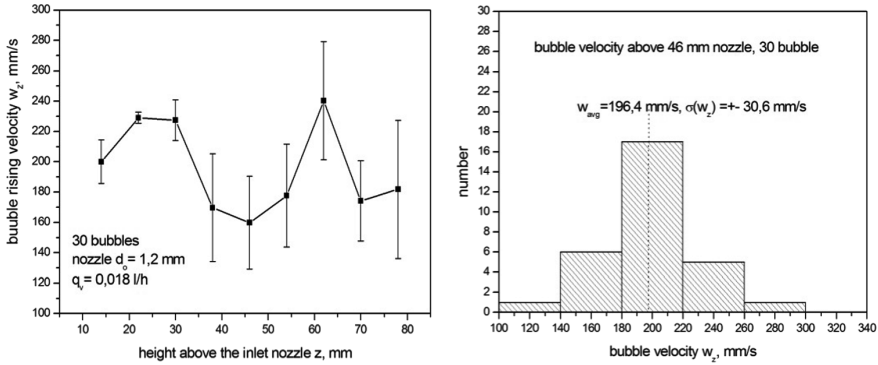
Fig. 7. The dependence of the velocity of argon bubbles on the rising path in GaInSn liquid metal
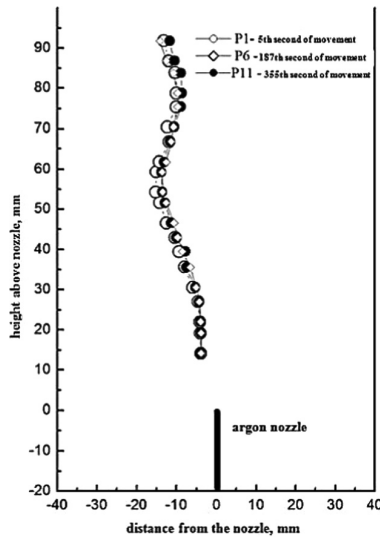
Fig. 8. An example of a bubble movement path in liquid metal for the flow of argon equal to 0.018 l/h

In the case when it is not possible to distinguish between individual signals recorded by ultrasonic transducers, that is in the flow of bubble swarm or in the chain flow, the bubble velocity is determined from the normalized cross-correlation function of signals ρz by calculating the delay time of τ* as the maximum of this function according to equation:

$$\rho_z = \frac{\hat{R}_{z_{i+1}z_i}}{\sqrt{\hat{R}_{z_iz_i}(0)\hat{R}_{z_{i+1}z_{i+1}}(0)}} \qquad (4)$$

where:

- $\hat{R}_{z_{i+1}z_i}$ cross-correlation function of signals from two consecutive sensors z_i, z_{i+1}
- $\hat{R}_{z_iz_i}$ autocorrelation function of signals.

The use of signal delay time for calculations has the advantage that the ρ_z coefficient values give information about the similarity of the registered signals to bubbles. At values of ρ_z close unity, transducers record signals from the same bubble. When $\rho_z= 0$ there is no correlation between the signals. In the literature [8] it is assumed that there is a similarity between signals when $\rho_z = 0.4$. Figure 9 presents examples of signals from 4 ultrasonic sensors together with signal correlation functions, while Fig. 10 illustrates the dependence of bubbles velocities on their rising path calculated from signals cross-correlation function.
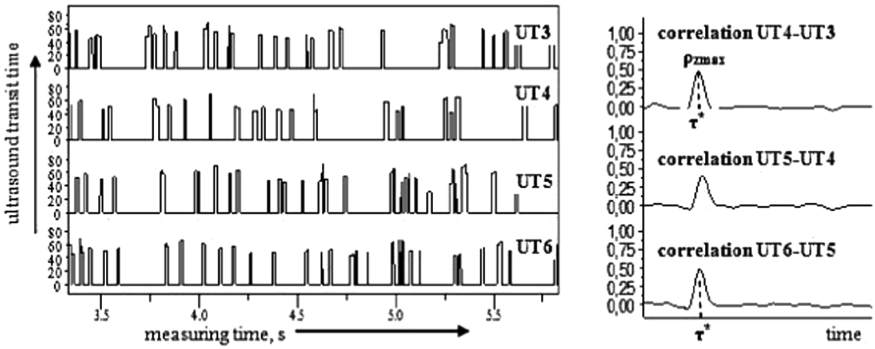
Fig. 9. Examples of signals along with correlation functions.

5 Error Sources

The highest influence on the accuracy of the velocity determination from Eq. 2 has the correct determination of the distance L between the ultrasonic transducers. This is due to the accuracy of their assembly in the ultrasonic head. The nominal distance between transducers is $L = 8$ mm. In the case of imprecise mounting, ultrasonic rays emitted from them are not parallel and deviations of parallelism can reach up to 10% (GE Inspektion Technologies GmbH). Lack of characteristics (calibration) of ultrasonic transducers leads to significant errors in determining the velocity of rising bubbles. For example, for ultrasonic head USH1, with the obtained transducers characteristics{UT1; UT2;UT3;…UT10}, correct distances between them have been deteriorated to mm {8,2; 8,3; 7,5; 8,4; 7,8; 7,6; 8,6; 7,9; 8}. Systematic errors in the calculations of bubbles velocities, when accepting distances between $L = 8$ mm, not taking correct distances, may reach up to 7%.

The second component of the error is related to the accuracy of determining the signals delay time. In the case of reading the delay time of signals $\tau^* = \tau_1 - \tau_2$ as in the Fig. 6, the standard uncertainty of type B of this time is:
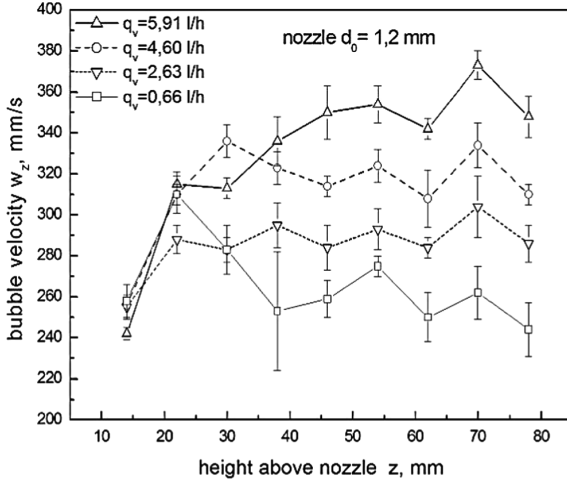
Fig. 10. Exemplary dependencies of bubble velocities on their rising path calculated from the signals cross-correlation function.

$$u_{B_{\tau^*}} = \sqrt{\left(u_{B_{\tau_1}}\right)^2 + \left(u_{B_{\tau_2}}\right)^2} \tag{5}$$

The time limiting errors of τ_1 and τ_2 are equal and their values correspond to the sampling period of $\delta t = 0.7$ ms. Standard uncertainties of type B of the times τ_1 and τ_2 are amount to $u_{B_{\tau_1}} = u_{B_{\tau_2}} = \frac{\delta t}{\sqrt{3}} = 0.405$ ms. The standard uncertainty of the signal
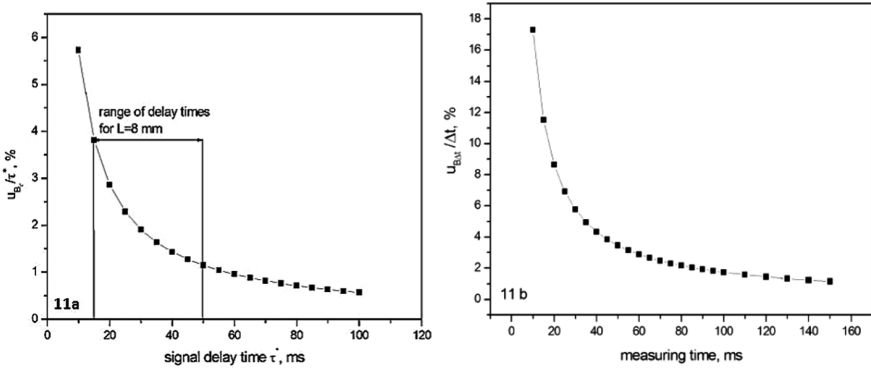
Fig. 11. Relative standard uncertainties of type B when sensors are mounted on the side wall of the container and its bottom.

delay time is $u_{B_{\tau^*}} = 0,573$ ms. The relative dependence of this uncertainty on the delay of signals is shown in the Fig. 11a.

With average signals delay times of 15–50 μs and the distances between the sensors $L = 8$ mm, the relative standard uncertainties of the delay time vary within the limits of 1.5%–3.8%. Adoption of the distance between individual transducers $L = 8$ mm for assembly in the ultrasonic head was mainly due to the fact that despite quite high uncertainty resulting from the sampling period, especially for high blistering velocities, rising bubbles are found in the ultrasonic field of each transducer. Thus the recorded signals come from practically every moving bubble which is very important when determining the flow parameters of bubbles rising spirally or zigzag.

When measuring with transducers in the bottom of the container, the relatively standard uncertainty of the velocity, determined from Eq. 1, can be presented as:

$$\frac{u_{B_{w_z}}}{w_z} = \sqrt{\left(\frac{u_{Bc}}{c}\right)^2 + \left(\frac{u_{Bdt}}{dt}\right)^2 + \left(\frac{u_{B\Delta t}}{\Delta t}\right)^2} \tag{6}$$

At a constant GaInSn temperature, the uncertainty of the ultrasonic wave velocity is to be neglected, also the uncertainty of the dt – difference of the echo transit time from the bubble can be neglected because the resolution is $\delta t_e = 2.5$ ns. The main component of uncertainty is the error of determining the measurement time t related to the period of signal sampling. For transducers mounted in the bottom of the container, this period was $\delta t = 1.5$ ms. Assuming that the limiting error of the time reading Δt is $2\delta t = 3$ ms, the standard uncertainty of type B is: $\frac{u_{B\Delta t}}{\Delta t} = \frac{2\delta t}{\sqrt{3}} = 1,73$ ms.

Figure 11b shows the dependence of the relative uncertainty of velocity on the measurement of Δt. The errors of the method are also related to the observation time of the Tm signals and the values of the correlation function ρ_z. As Shu [8] showed in his paper, the standard uncertainty of type A of signal time delay is a function of two components:

$$u_{A_{\tau^*}} = \sqrt{var(\tau^*)} \sim \sqrt{\frac{1}{T_M}} \tag{7}$$

$$u_{A_{\tau^*}} \sim \sqrt{\frac{1}{\rho_z} - 1} \tag{8}$$

The maximum number of samples possible to calculate the correlation function in the DasyLab program is 215 which at the signal sampling interval $\delta t = 0.7$ ms gives the observation time $T_m = 2^{15} * 0,7$ ms $= 22,9$ s. As can be seen on the Eq. 8, the adoption of such a long observation time effectively reduces the uncertainty of the standard signal delay time. Values of the correlation coefficient ρ_z close to unity also reduce this uncertainty. The ρ_z coefficient values show the similarity of the recorded signals and they are related to the distance between the transducers – the smaller is the distance the higher are the ρ_z values. The choice of the distance between the

transducers in the ultrasonic head L = 8 mm was also associated with the minimization of this uncertainty component.

6   Summary

The article presents the methods of measuring and calculating the velocity of the rising gas phase in the GaInSn liquid metal by means of an ultrasonic pulse-echo method. The method of determining the velocity in the case when ultrasonic sensors are mounted in the bottom of the container and in case when they are installed on the side wall is described. The sources of method errors are also presented. The knowledge of the velocity of the rising gas phase can be for examples useful in steel refining processes, to calculate the residence time of gas bubbles in the liquid metal and thus to increase the efficiency of this process. The pulse-echo method can be used in measurements, especially those in which the liquid phase is non-transparent, to determine other flow parameters such as: the frequency of gas bubbles, their dimensions or the flow area of the second phase in the liquid metal.

References

1. Zhang, Ch.: Liquid metal flows driven by gas bubbles in a static magnetic field. Dissertation TU Dresden, Fakultät Maschinenwesen (2009)
2. Zhang, Ch., Eckert, S., Gerbeth, G.: Gas and Liquid velocity measurements in bubble chain driven two-phase flow by means of UDV and LDA. In: 5th International Conference on Multiphase Flow, ICMF 2004, Yokohama, Japan (2004)
3. Zhang, Ch., Eckert, S., Gerbeth, G.: Experimental study of single bubble motion in a liquid metal column exposed to a DC magnetic field. Int. J. Multiph. Flow **31**(7), 824–842 (2005)
4. Eckert, S., Gerbeth, G.: Messung von Geschwindigkeitsfeldern in Flüssigmetallen mit der Ultraschall-Doppler-Methode. Tech. Mess. **79**(9), 410–416 (2012)
5. Sommerlatt, H.-D., Andruszkiewicz, A.: Dynamic measurement of particle diameter and drag coefficient with ultrasonic method. Archives Acoust. **3**, 293–304 (2008)
6. Andruszkiewicz, A., Eckert, K., Eckert, S., Odenbach, S.: Gas bubble detection in liquid metals by means of the ultrasound transit-time-technique. Eur. Phys. J. Spec. Topisc **220**, 53–62 (2013)
7. Richter, T., et al.: Measurements of the diameter of rising gas bubbles by means of ultrasound transit time technique. Magnetohydrodynamics **53**(2), 383–392 (2017)
8. Shu, W.: Durchflussmessung in Rohren mit Hilee von künstlichen und natürlichen markierungen, Dissertation T.U. Karlsruhe (1987)

Heat Resistant Monitoring System
for Medical Sterile Containers

Lukas Böhler[1,2]([✉]), Mateusz Daniol[1,2], Anton Keller[1], and Ryszard Sroka[2]

[1] Aesculap AG, Am Aesculap-Platz, 78532 Tuttlingen, Germany
`lukas.boehler@aesculap.de`
[2] AGH University of Science and Technology, Adama Mickiewicza Ave. 30,
30-059 Kraków, Poland

Abstract. In this paper we focus on the issue of the increasing amount of treatments in modern hospitals and the need to increase on the one hand the efficiency and on the other hand to decrease costs without decreasing the safety. Therefore, this research is about the development of a tracking and sensor system, applicable in medical sterile containers including a thermal insulation protecting the electronics against the sterilisation process. Using this system can increase the efficiency by tracking the containers within a hospital and signalizing its status. For this reason different tracking technologies were discussed and sterilisation tests with several RFID transponders were made. Furthermore, the shown sensor system can increase the safety by providing data about the last sterilisation process and the overall state of the sterile container.

Keywords: Tracking · Sensors · RFID · Sterilisation · Safety · Digitalisation

1 Introduction

Based on the increasing living conditions and improvements in healthcare, people in industrial countries are getting older and older. The disadvantage of this trend and our modern way of living is the fact, that diseases, physical defects and injuries increase. Due to this, hospitals need to increase their efficiency and the amount of treatments continuously. Therefore, modern hospitals are planned as medical centers, providing a multitude of medical disciplines. However, the bigger a hospital gets, the more complex logistic processes and the planning structure needs to be. Like in industrial companies, items need to be tracked, tasks need to be optimized and time-consuming tasks need to be reduced to maintain a high efficiency of the medical staff [1,2].

Beside the medical devices, sterile containers, containing medical instruments, small devices and tools, need to be managed and organized in the daily workflow of a hospital [3]. Because of the high number of containers, storing the needed equipment for surgeries, a system needs to be found, to track the containers inside a hospital and provide additional data to increase the safety.

1.1   Digitalisation in Hospitals

The digitalisation is a trend caused by the increasing possibilities of modern technologies and the expanding internet. The combination of cost efficient electronics and the possibility of connecting devices and sharing data was the next step in the digital age, starting more than half a century ago. Since the early 2000s, mobile applications started influencing the daily life more and more. This resulted in the so called digitalisation of many sectors in our life [4]. Connected sensor networks [5], decentralised availability of data and steady developments of electronic modules, which became less expensive and more efficient, can be used to increase efficiency in industry.

But in healthcare, this trend is still in its infancy. In the past years small approaches in this sector were made, but there is still a big potential to support the medical staff in their daily work [6]. Wearable medical devices and connected syringe pumps are a few examples of the growing digitalisation in hospitals. However, logistic problems regarding the asset management still need to be solved. A tracking of medical equipment, like its state of the art in industry, can increase the efficiency of the hospital logistics by decreasing the wasted time, spend with searching for needed goods, and allow a better planning of surgeries or regulate the supply chain. To achieve a tracking of needed goods, many approaches were made. In general two main fields can be named: tracking of non-sterile devices e.g. monitors or defibrillators and tracking of sterile goods like medical sterile containers.

Beside the tracking, also providing additional data about items can increase the safety and the efficiency. Hereby, sensors can transmit the state of a medical device, its health state and also planned uses in the close future. Such information can help the staff to sort out damaged or to-be-maintained items and get and overview of available items.

1.2   State of the Art

The current state of the art is the separation of electronics from sterilisable housings. An example for this are surgical drills, where the control electronics and the battery can be removed before the sterilisation and put back to the device with a sterile adapter like shown in Fig. 1.
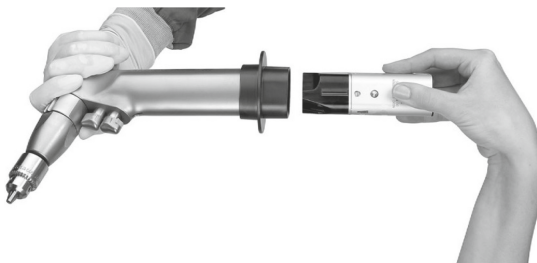
Fig. 1. Sterile adapter for inserting non-sterile battery into sterile device

The problem with this kind of treating surgical devices is the danger of contaminating the sterile environment. Further is the device not ready to use right after a sterilisation. There are already approaches to insulate electronics in such way, that they can stay inside the device, however, these are still quite bulky at the moment. For medical sterile containers, also some approaches exist, to implement sensors indicating the sterile state [7–9], but on the market just systems with an analogue mechanism like bi-metal switches can be found. Hereby, the reliability of the indication is controversial [10] due to a switching in a vague defined temperature area.

For tracking medical goods inside a hospital, most approaches use bulky transponders for Wi-Fi or Bluetooth [11–13], however, none of these modules can be sterilised, therefore, their use is limited to expensive devices, which don't need to be sterile. Identifying containers and providing a rough location, barcodes are used and read out after key steps in the work-cycle of the containers. New approaches also use labels or transponders, transmitting data over radio frequency [14,15]. However, in this case just very few manufacturers give information about the use in hospital environment regarding durability for sterilisations.

1.3   Issues in Medical Environment for Electronics to Overcome

The use of electronic devices in healthcare is since many decades state of the art. However, a steam sterilisation is needed for surgical instruments, implants and devices, which can not use a sterile cover. Because of that, most electronic systems for surgeries are designed, to be covered with a sterilisable or pre-sterilised cover or case. The reason of this is the fact, that during a steam sterilisation three critical factors occur:

1. Saturated steam, which can damage the electronics
2. Temperatures up to 135 °C, which damages electronics and many insulations
3. Fast pressure change from vacuum to 3 bar, which can damage insulations

In Fig. 2 a standard sterilisation cycle is shown. The whole sterilisation needs approx. 30 min.
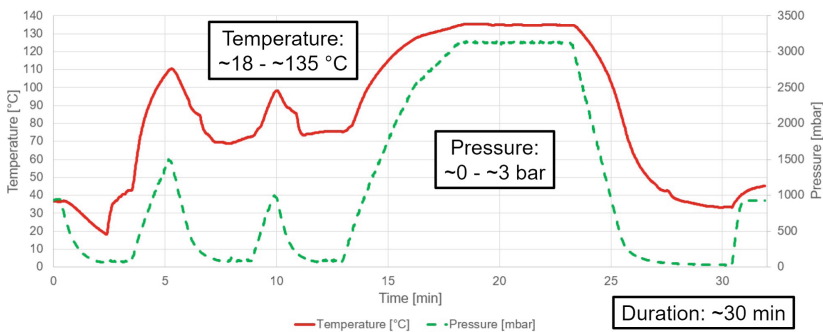
Fig. 2. Temperature and pressure curve of standard steam sterilisation cycle

As we can see in Table 1, a temperature higher than 85 °C is held longer than 15 min during a steam sterilisation. Furthermore, the holding time of 135 °C is approx. 5 min, which leads to serious damages in most electronic systems.

Table 1. Durations of critical temperatures during a sterilisation

Temperature	Time (min)
>85 °C	15:32
>100 °C	12:09
>130 °C	06:54
135 °C	05:14

Beside the issue of very high temperatures, additional electronic systems used in already existing products like medical sterile containers need to be design in such a way, that they don't affect the function or hygienic conditions. Therefore, new designs need to be adapted to available space and prevent the danger of contamination or incomplete sterilisations.

1.4   Review of Tracking Technologies

For tracking and localising items inside a building, several technologies can be used. In the past years mainly five different technologies were evaluated:

1. Wi-Fi
2. ZigBee
3. Bluetooth
4. Ultra Wide Band (UWB)
5. Radio Frequency Identification (RFID)

The indoor localisation is based on measuring the signal strength or in the case of UWB on the time, the signal needs from the sender to the receiver. The more receivers are placed around a sender, the more accurate the position can be calculated. Beside the RFID tracking, all systems need an active power source to send their data. In RFID active transponders, powered by a battery, can act like a regular sender. However, passive transponders don't need a power source because the signal from the reader module is reflected from the transponder. Therefore, RFID is the only system which is not depending on a battery.

Wi-Fi tracking modules can use the already existing wireless network. However, the accuracy of these modules is quite low and the costs are very high [16,17]. ZigBee modules have a very low power consumption and low costs. A drawback is the fact, that the used frequency of ZigBee causes interferences with other devices, using the same frequency. Like ZigBee, Bluetooth systems have very low power consumptions and also low costs. However, to increase the accuracy of the tracking, a multitude of transceivers is needed and because of the

widely use of Bluetooth modules, interferences can occur [18]. The technology of UWB is using a very high frequency between 3,1 GHz and 10,6 GHz. Therefore, almost no interferences can occur and a localisation is also possible, if the module is in a closed container with a very good accuracy. The disadvantage are the very high costs [19–22]. In passive RFID, a multitude of different transponders for different ranges and applications can be used. Hereby, the transponders have very low costs. The drawbacks are the high costs of the reader hardware and the limited accuracy [14, 23, 24].

The described tracking technologies have advantages and disadvantages. For the use of tracking medical sterile containers inside a hospital, the accuracy is less important than the caused costs of the modules. Therefore, Bluetooth and RFID modules fit the requirements the most [25–27]. Because of the need to equip a very high number of containers and the fact of needed sterilisations, RFID transponders seemed to be the best choice for a first evaluation of the usability.

1.5 RFID Transponders

The use of RFID transponders allows a readout of the stored information in the range of several meters, depending of the used frequency range. For a close readout, also possible with a smartphone, Near Field Communication (NFC) tags with a frequency of 13,56 MHz can be used. The benefit hereby is, that no expensive reader hardware needs to be used. For a tracking or readout over a higher distance, Ultra High Frequency (UHF) transponders with a frequency of 868 MHz (EU) or 915 MHz (USA) can be used. However, for the use of these transponders, reader systems with high costs need to be installed.

The benefit of the use of RFID transponders is the possibility to track assets inside a hospital by installing gates in defined spots inside the building. As soon a tagged item passes, the information can be displayed in a central system, containing also the stored information of the transponder. Further, the stored data can be updated e.g. to change the status of a medical container, indicating a surgery is planned. Another benefit is, that RFID transponders are produced in a multitude of different shapes, like overmolded transponders or thin labels, also providing different frequencies like UHF and NFC. Some transponders also can handle high temperatures and are protected against water.

2 Materials and Methods

For making a first prototype of the sensor module, the hardware components had to be chosen regarding the identified requirements. According to experts from the field of medical sterile containers crucial parameters of the sterilisation process are the temperature and the corresponding time. Furthermore, a high risk of container damages is the penetration of the sterile filters or a deformation of the container itself. An exemplary container and a tray filled with medical instruments is displayed in the following Fig. 3. For tracking the container inside a hospital, different technologies had to be evaluated.

Fig. 3. Sterile container with tray including medical instruments

2.1   Sensors

For monitoring the sterilisation process, a *KTY81* thermistor with a close range was chosen. To increase the accuracy a temperature range up to $150\,°C$ was used because of the maximum sterilisation temperature of approx. $135\,°C$. For reading out and logging the sensor values in specified time steps, a regular *ARM Cortex-M0+* was used. To detect a successful sterilisation, the measured values are compared with defined minimum temperature times similar to Table 1. This enables a fast and reliable check and decrease the possibility of false indications. Additional a sensor for detecting damages of the container had to be found. To detect possible damages of the case, a accelerometer for ±200 g was used, to capture serious drops of the container. These drops can harm the metal tank, the upper shell or the devices inside the container. Another risk is the penetration of the filters, placed in the upper shell like shown in Fig. 4, preventing the inside of the container to get contaminated.
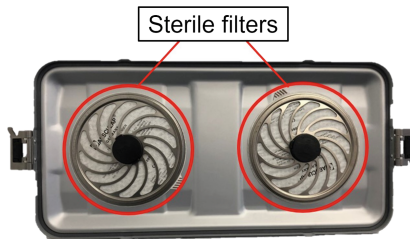
Fig. 4. Upper shell with two sterile filters

A damage of this filter can occur, if an instrument inside the container pierces the filter because of a drop or roll over. To detect such an incident, the microcontroller is programmed to recognize a roll over with data of the chosen accelerometer.

2.2 Communication

The collected data of the sensor module need to be provided to a smart device e.g. a smartphone or a tablet. The ability of the module to communicate with a multitude of different devices was a key requirement, because of the aim to reduce the costs for the hospital. Therefore, a Bluetooth Low Energy (BLE) module was chosen. An advantage of this is the very low power consumption, increasing the overall lifetime of the sensor module. Furthermore, Bluetooth modules are able to send signals out of a closed medical sterile container because of the high frequency.

2.3 Thermal Insulation

One of the main problems was the high temperature, exceeding the limits of most electronic components. Regular electronic modules withstand temperatures up to 85 °C. Electronics for military or automotive can reach temperatures up to 125 °C. However, to lower the costs of these modules, the aim was to use regular components.

To achieve this, a thermal insulation had to be found. In most cases, silicone or epoxy resins are used to protect electronics from heat or vapour. However, for this sensor module a thin insulation is needed to fit into the sterile containers without influencing the arrangement of the instrument tray. To minimize the insulation thickness, a material was searched with a very low thermal conductivity.

In the field of building insulation and fire protection, a material called *Aerogel* is used as a protection against high temperatures [28, 29]. The thermal conductivity of this material is very low [30, 31], compared to the regular insulations for electronics, like shown in Table 2:

Table 2. Comparison of thermal conductivities

Material	Thermal conductivity (W/m*K)
Silicone	0,2
Epoxy resin	0,2
Aerogel	0,02

Another benefit of aerogel is the fact, that its density is very low which reduces the weight of the whole module significantly. Because of the fast pressure change from vacuum to 3 bar inside the autoclave, the aerogel needs to be protected. For this purpose, a thin outer layer of silicone or epoxy resin can be used.

2.4 Design of Sensor Module

To combine the needed sensors and communication module in a compact design, a first approach is displayed in Fig. 5. Beside the temperature and acceleration

sensor, also a pressure sensor can be used to evaluate the sterilisation procedure. As power supply either a high temperature or a rechargeable battery can be used. The use of regular batteries, especially with Li-ion technology, must be prevented because of the high dangers caused by high temperatures [32,33]. To increase the lifetime of this battery, the insulation of the module needs to limit the temperature increase inside the battery to a minimum [34]. To lower the power consumption also the use of a passive shocksensor is possible to detect shocks with a defined minimum force. Hereby, an internal switch is closed as soon as a minimum force in one direction is reached. This can be read out from a low-power microcontroller and displayed as detected shock. In contrast to an accelerometer, the applied force cannot be measured. However, the benefit of such a sensor is, that no power is consumed as long the defined force is not reached.
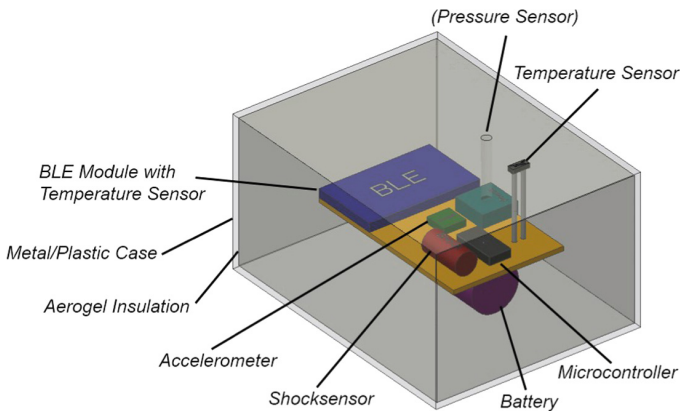
Fig. 5. Insulated sensor module for sterile containers

2.5   Tests of Thermal Insulations

For testing the aerogel insulation, different test setups were used to evaluate the needed thickness of the insulation. Furthermore, tests with one epoxy resin module and mixtures of aerogel and silicone were made. For the measurements of the thicker insulations like the first aerogel and the epoxy resin module, a circuit board with a Bluetooth module, a temperature sensor and a high temperature battery were used. In the following tests with thinner insulations, temperature strips with defined temperature indicators were used.

To compare the new insulation material with the state of the art, an epoxy resin insulation with a thickness of approx. 25 mm was tested first. Hereby, also the internal temperature was logged with a temperature sensor, connected to a Bluetooth module, powered by a high temperature battery, shown in Fig. 6, during multiple sterilisations.
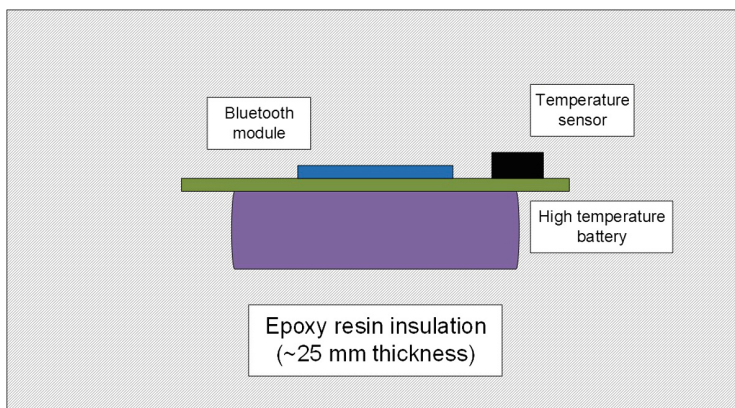
Fig. 6. Cross section of test module with epoxy resin insulation

Comparable to the test setup with the epoxy resin insulation, the first aerogel insulation was tested with the Bluetooth logging module, shown in Fig. 7. Hereby, the insulation thickness was lowered to 20 mm. Because of the structure and the unknown efficiency of the insulation, the module was tested in a heat chamber at continuously 135 °C until an internal temperature of 85 °C was reached.
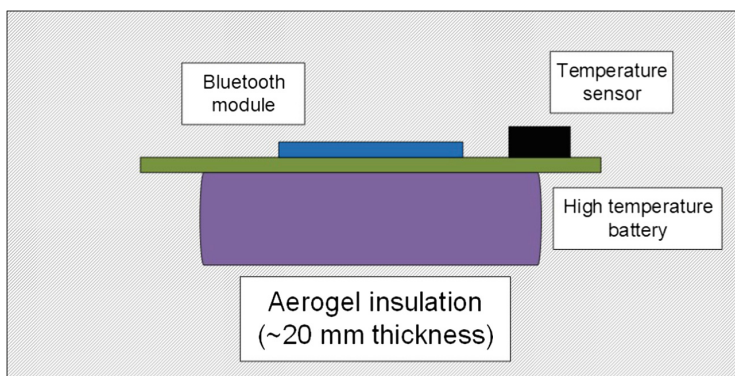
Fig. 7. Cross section of test module with aerogel insulation

Because of the previous results, further tests with aerogel were made. To evaluate the needed thickness of the insulation, two more test modules with an insulation thickness of 7 mm aerogel and 6 mm aerogel in combination with 1 mm silicone were made. Due to the thin insulation, no Bluetooth module was used. In this case temperature strips with indicators between 71 °C and 110 °C were insulated and checked after the test. The modules are shown in Figs. 8 and 9. To prevent the insulations from damages by the pressure change and

steam, the tests were performed in a heat chamber with a temperature curve, close to a sterilisation. However, because of the slow change in temperature, the cycle took approx. 1,5 h.
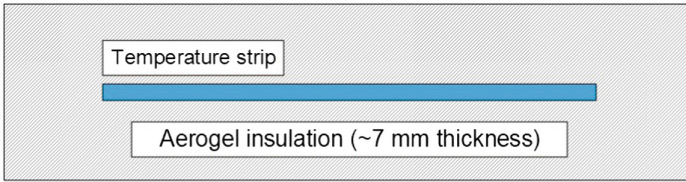
Fig. 8. Cross section of test module with aerogel insulation and temperature strip
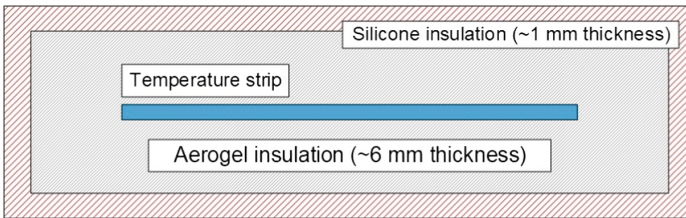
Fig. 9. Cross section of test module with aerogel and silicone insulation and temperature strip

2.6   Tests of RFID Transponders

To evaluate the reliability of different RFID transponders and labels, three different types were chosen, shown in Fig. 10. All three types can handle high temperatures and are protected against the steam. The first transponder is an overmolded UHF tag from the company *SAG*. Its maximum storage temperature is 125 °C. Four equal transponders were tested. The second transponder is a UHF RFID label from *Fujitsu*. The maximum storage temperature is 150 °C. From this type one label was tested. The last transponder is a very thin hybrid label from *SMARTRAC*, operating in UHF and NFC frequency ranges. Its maximum operation temperature is 85 °C. The maximum storage temperature is not given in the datasheet, however, the label is able to withstand much higher temperatures (approx. 125°C–150 °C). Four of these hybrid labels were tested.

Multiple transponders of each type were placed inside a sterile container and sterilised over a multitude of cycles. The aim was, to determine the maximum cycles, the transponders can handle.
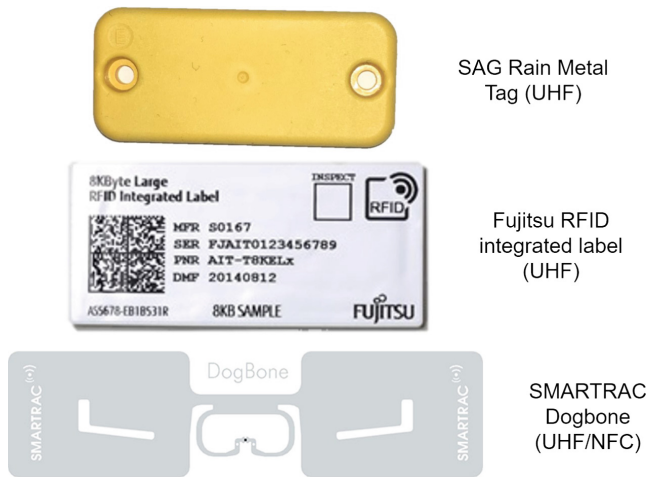
Fig. 10. Used RFID transponders

To read out the RFID transponders a *ID ISC.MRU102* mid range UHF reader from the company *FEIG ELECTRONIC*, shown in Fig. 11 was used.

Fig. 11. Used RFID mid range reader

3   Results

The results of the performed tests show, if it is possible to insulate electronic modules for the medical sector in such way, that a sterilisation of the modules can be performed without damaging the electronics and without the need of a bulky insulation.

3.1   Effectiveness of Thermal Insulations

The results of the test with an epoxy resin thermal insulation show, that the internal temperature does not exceed 81 °C during and after the sterilisation,

shown in Fig. 12. Furthermore, the temperature rises very slow, therefore, the maximum temperature inside the module is reached after the cycle.
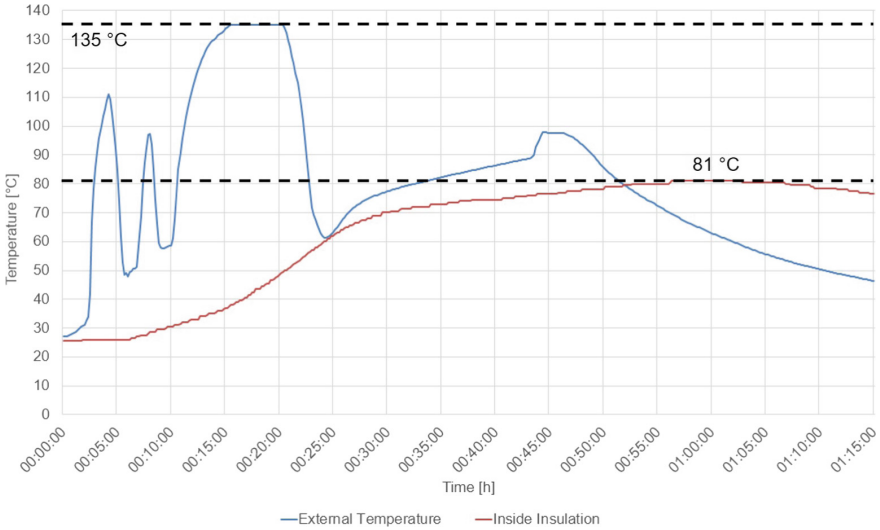
Fig. 12. Temperature curve of 25 mm epoxy resin insulation

To test the durability of the module, several sterilisation cycles were performed. In Fig. 13 a comparison of the module before the first sterilisation and after 21 cycles is shown. Because of possible air encapsulations and the inflexible epoxy resin, the insulation cracked during the 21. sterilisation.
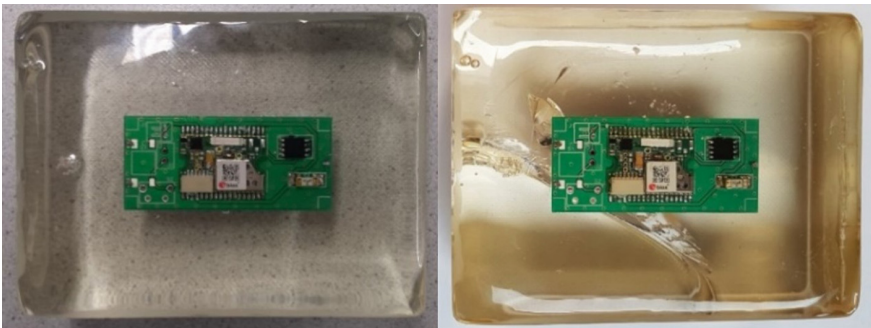
Fig. 13. Epoxy resin insulation before (left) and after (right) 21 sterilisation cycles

Therefore, the good thermal insulation of the epoxy resin can't compensate the disadvantages of high costs, a difficult processing to avoid air encapsulations and, in comparison to aerogel, high thermal conductivity.

The first test with an aerogel insulation showed, that this material is able to delay the rise of the internal temperature in an environmental temperature of 135 °C significantly. Under this conditions with an insulation thickness of 20 mm, the internal temperature reached the critical temperature of 85 °C after 50 min, shown in Fig. 14.
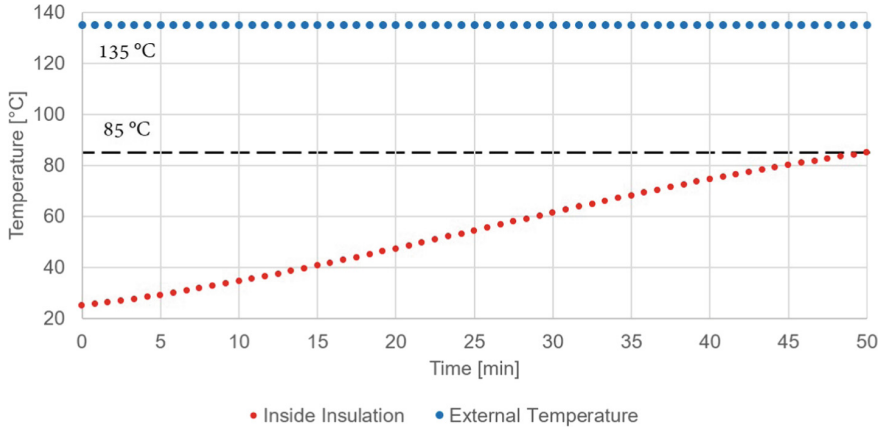
Fig. 14. Temperature curve of 20 mm aerogel insulation

Because of the very promising results of the first test, the insulation thickness was lowered significantly and tested with a changing temperature, similar to a sterilisation as displayed in Fig. 15.
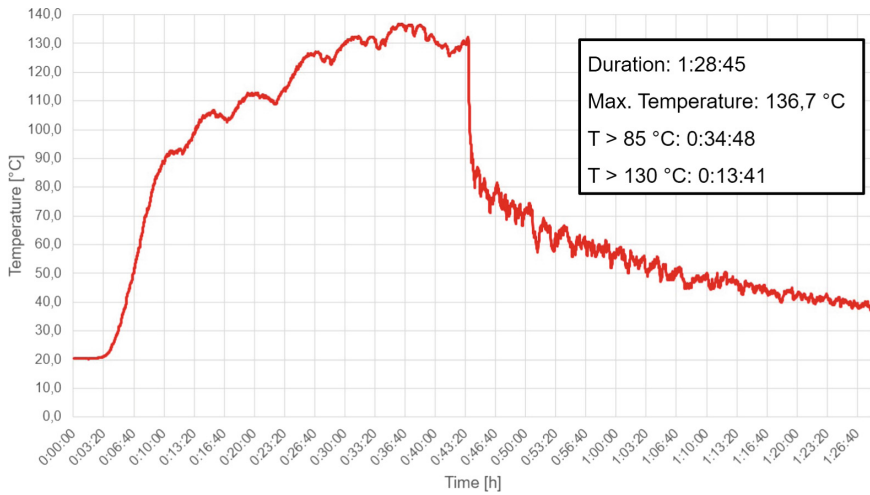
Fig. 15. Temperature curve of heat chamber for temperature strip tests

The result of this tests showed, that both, the insulation with 7 mm aerogel and the combination of 6 mm aerogel and 1 mm silicone, kept the internal temperature between 71 °C and 77 °C. The comparison of the temperature strip before and after the test is displayed in the following Fig. 16.

Fig. 16. Temperature indication before (up) and after (down) temperature test cycle

3.2   Sterilisability of RFID Transponders

During the sterilisation test, the functionality of the RFID transponders was tested in defined steps. Because of parallel tests in the autoclave, it wasn't possible to perform a test after each sterilisation. However, the number of sterilisation cycles without damages could be determined with a satisfying accuracy. In the following Table 3 the results of the performed test are shown.

Table 3. Result of sterilisation tests of RFID transponders

Transponder	Number of cycles
SAG rain metal tag	approx. 100
Fujitsu RFID integrated label	approx. 200
SMARTRAC dogbone	approx. 200

The transponders from the company SAG were damaged after approx. 100 cycles. The labels from Fujitsu and SMARTRAC were able to perform approx. 200 cycles before no communication could be performed anymore.

4   Discussion

The results show, that an insulation with a thin layer of aerogel already can protect electronic modules from the high temperatures, applied during a steam sterilisation. Therefore, the use of electronics inside a medical sterile container

can be achieved, providing sensor data from the last sterilisation, indicating a successful procedure, and detecting possible damages during the transport of the container. Furthermore, tracking modules can be implemented inside the containers, allowing a localisation inside a hospital to support the medical staff and increase the efficiency.

However, further tests with the insulation needs to be done, proving the usability in a real sterilisation. For a lifetime as long as possible also long time tests with a multitude of sterilisations need to be performed. With these future tests, it should be possible to estimate the lifetime of the insulation for the use in medical sterile containers.

Beside the tests of the insulation itself, also the sensor module needs to perform multiple sterilisation cycles to show, that the used sensors can determine a successful sterilisation and also detect failed sterilisations or possible damages of the container.

The use of RFID transponders can be a useful addition to the current used barcodes because of the ability to store data which can be updated. This provides the possibility to support the asset management inside a hospital and track containers with a satisfying accuracy. To increase the number of possible sterilisations, insulations of the transponders should be tested. Also the implementation of transponders in exchangeable parts of the container can be investigated. Hereby, the influence of the insulation on the read range of the transponders needs to be investigated.

To find alternative tracking possibilities, the accuracy of UHF RFID, UWB and Bluetooth needs to be compared. A key factor needs to be the cost effectiveness regarding the achieved accuracy in a metallic environment, like in a hospital container storage [35–41].

5   Conlusion

The results of this work show, that the use of alternative insulation materials enables possibilities to bring forward the digitalisation in hospitals and increase the amount of additional data for the medical and technical staff. This can increase the efficiency, lower costs and increase the time of treating patients instead of using precious time in searching goods and checking the inventory and state in documentations. As the tests show also the use of RFID labels for medical equipment is possible even in the case of a high number of steam sterilisations and therefore fitting the requirements for the use on medical sterile containers.

References

1. Choi, J.H., Fortsch, S.M., Park, I., Jung, I.: Efficiency of U.S. hospitals between 2001 and 2011. Managerial Decis. Econ. **38**(8), 1071–1081 (2017). https://doi.org/10.1002/mde.2846
2. Moatari-Kazerouni, A., Bendavid, Y.: Improving logistics processes of surgical instruments: case of rfid technology. Bus. Process Manag. J. **23**(2), 448–466 (2017). https://doi.org/10.1108/BPMJ-06-2016-0127

3. Diamant, A., Milner, J., Quereshy, F., Xu, B.: Inventory management of reusable surgical supplies. Health care Manag. Sci. (2017). https://doi.org/10.1007/s10729-017-9397-3

4. Chang, Y., Dong, X., Sun, W.: Influence of characteristics of the internet of things on consumer purchase intention. Soc. Beh. Personal. Int. J. **42**(2), 321–330, (2014). https://doi.org/10.2224/sbp.2014.42.2.321

5. Sisinni, E., Depari, A., Flammini, A.: Design and implementation of a wireless sensor network for temperature sensing in hostile environments. Sens. Actuators A: Phys. **237**, 47–55 (2016). https://doi.org/10.1016/j.sna.2015.11.012

6. Yin, Y., Zeng, Y., Chen, X., Fan, Y.: The internet of things in healthcare: an overview. J. Ind. Inf. Integr. **1**, 3–13 (2016). https://doi.org/10.1016/j.jii.2016.03.004

7. Childers, R.W., Henniges, B., Hassler, W., Blandino, T., Jeng, D., Morris, R.F.: Sterilization container with battery powered sensor module for monitoring the environment in the container (2015). WO002015138461A1

8. Childers, R.W., Chmelar, E.V., Dudycha, A., Henniges, B., Miller, M., Moaiery, A., Purrenhage, B.J.: Sterilization container capable of providing an indication regarding whether or not surgical instruments sterilized in the container were properly sterilized. WO002014159696A1 (2014)

9. Wehrle, C., Nonnenmann, M.: Surgical container contents detection system (2016). WO002016188959A1

10. Schuster, S.: Medizinischer sterilbehälter und verfahren zum bestimmen des sterilisationsstatus eines medizinischen sterilbehälters (2013)

11. Dobircau, A., Folea, S., Valean, H., Bordencea, D.: Indoor localization system based on low power wi-fi technology. In: 2011 19th Telecommunications Forum (TELFOR) Proceedings of Papers, pp. 317–320. IEEE (2011). https://doi.org/10.1109/TELFOR.2011.6143553

12. Torteeka, P., Chundi, X., Dongkai, Y.: Hybrid technique for indoor positioning system based on wi-fi received signal strength indication. In: 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN), pp. 48–57. IEEE (2014). https://doi.org/10.1109/IPIN.2014.7275467

13. Oosterlinck, D., Benoit, D.F., Baecke, P., van de Weghe, N.: Bluetooth tracking of humans in an indoor environment: an application to shopping mall visits. Appl. Geograph. **78**, 55–65 (2017). https://doi.org/10.1016/j.apgeog.2016.11.005

14. Hanada, E., Hayashi, M., Ohira, A.: Introduction of an rfid tag system to a large hospital and the practical usage of the data obtained. In: 2015 9th International Symposium on Medical Information and Communication Technology (ISMICT), pp. 108–111. IEEE, Piscataway, NJ (2015). https://doi.org/10.1109/ISMICT.2015.7107508

15. Lee, C., Palaniappan, S.: Effective asset management for hospitals with rfid. In: 2014 IEEE International Technology Management Conference (ITMC), pp. 1–4. IEEE, Piscataway, NJ (2014). https://doi.org/10.1109/ITMC.2014.6918596

16. Turgut, Z., Aydin, G.Z.G., Sertbas, A.: Indoor localization techniques for smart building environment. Proced. Comput. Sci. **83**, 1176–1181 (2016). https://doi.org/10.1016/j.procs.2016.04.242

17. Chen, C., Han, Y., Chen, Y., Liu, K.J.R.: Indoor gps with centimeter accuracy using wifi. In: 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), pp. 1–4. IEEE, Piscataway, NJ (2016). https://doi.org/10.1109/APSIPA.2016.7820842

18. Alarifi, A., Al-Salman, A., Alsaleh, M., Alnafessah, A., Al-Hadhrami, S., Al-Ammar, M.A., Al-Khalifa, : H.S.: Ultra wideband indoor positioning technologies: Analysis and recent advances. Sensors (Basel, Switzerland) **16**(5) (2016). https://doi.org/10.3390/s16050707

19. Gunia, M., Protze, F., Joram, N., Ellinger, F.: Setting up an ultra-wideband positioning system using off-the-shelf components. In: 2016 13th Workshop on Positioning, Navigation and Communications (WPNC), pp. 1–6. IEEE (2016). https://doi.org/10.1109/WPNC.2016.7822860

20. Jiang, L., Hoe, L.N., Loon, L.L.: Integrated uwb and gps location sensing system in hospital environment. In: I. Staff (ed.) 2010 5th IEEE Conference on Industrial Electronics and Applications, pp. 286–289. I E E E, [Place of publication not identified] (2010). https://doi.org/10.1109/ICIEA.2010.5516828

21. Mahfouz, M.R., Kuhn, M.J., Wang, Y., Turnmire, J., Fathy, A.E.: Towards sub-millimeter accuracy in uwb positioning for indoor medical environments. In: 2011 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems, pp. 83–86. IEEE (2011). https://doi.org/10.1109/BIOWIRELESS.2011.5724360

22. Monica, S., Ferrari, G.: Accurate indoor localization with uwb wireless sensor networks. In: Reddy, S.M. (ed.) 2014 IEEE 23rd International WETICE Conference (WETICE), pp. 287–289. IEEE, Piscataway, NJ (2014). https://doi.org/10.1109/WETICE.2014.41

23. Ben Kilani, M., Raymond, A.J., Gagnon, F., Gagnon, G., Lavoie, P.: Rssi-based indoor tracking using the extended kalman filter and circularly polarized antennas. In: 2014 11th Workshop on Positioning, Navigation and Communication (WPNC), pp. 1–6. IEEE (2014). https://doi.org/10.1109/WPNC.2014.6843305

24. Wang, Z., Ye, N., Malekian, R., Xiao, F., Wang, R.: Trackt: accurate tracking of rfid tags with mm-level accuracy using first-order taylor series approximation. Ad Hoc Netw. **53**, 132–144 (2016). https://doi.org/10.1016/j.adhoc.2016.09.026

25. Bisio, I., Sciarrone, A., Zappatore, S.: Asset tracking solution with ble and smartphones: an energy/position accuracy trade-off. In: 2015 IEEE Global Communications Conference (GLOBECOM), pp. 1–6. IEEE, Piscataway, NJ and Piscataway, NJ (2015). https://doi.org/10.1109/GLOCOM.2015.7417782

26. Bisio, I., Sciarrone, A., Zappatore, S.: A new asset tracking architecture integrating rfid, bluetooth low energy tags and ad hoc smartphone applications. Pervasive Mobile Comput. **31**, 79–93 (2016). https://doi.org/10.1016/j.pmcj.2016.01.002

27. Shirehjini, A.A.N., Yassine, A., Shirmohammadi, S.: Equipment location in hospitals using rfid-based positioning system. IEEE Trans. Inf. Technol. Biomed. **16**(6), 1058–1069 (2012). https://doi.org/10.1109/TITB.2012.2204896

28. Liang, Y., Wu, H., Huang, G., Yang, J., Wang, H.: Thermal performance and service life of vacuum insulation panels with aerogel composite cores. Energy Build. **154**, 606–617 (2017). https://doi.org/10.1016/j.enbuild.2017.08.085

29. Nocentini, K., Achard, P., Biwole, P.: Hygro-thermal properties of silica aerogel blankets dried using microwave heating for building thermal insulation. Energy Build. **158**, 14–22 (2018). https://doi.org/10.1016/j.enbuild.2017.10.024

30. Karami, S., Motahari, S., Pishvaei, M., Eskandari, N.: Improvement of thermal properties of pigmented acrylic resin using silica aerogel. J. Appl. Polym. Sci. **135**(1), 45,640 (2018) https://doi.org/10.1002/app.45640

31. Kiil, S.: Quantitative analysis of silica aerogel-based thermal insulation coatings. Progress Organic Coat. **89**(26–34), 26–34 (2015). https://doi.org/10.1016/j.porgcoat.2015.07.019. http://www.sciencedirect.com/science/article/pii/S030094401530062X

32. Feng, X., et al.: Characterization of large format lithium ion battery exposed to extremely high temperature. J. Power Sour. **272**, 457–467 (2014). https://doi.org/10.1016/j.jpowsour.2014.08.094

33. Wang, Q., Ping, P., Zhao, X., Chu, G., Sun, J., Chen, C.: Thermal runaway caused fire and explosion of lithium ion battery. J. Power Sour. **208**, 210–224 (2012). https://doi.org/10.1016/j.jpowsour.2012.02.038

34. Meekhun, D., Boitier, V., Dilhac, J.M.: Study of the ambient temperature effect on the characteristics and the lifetime of nickel-metal hydride secondary battery. In: IEEE Electrical Power & Energy Conference (EPEC), 2009, pp. 1–4. IEEE, Piscataway, NJ (2009). https://doi.org/10.1109/EPEC.2009.5420772

35. Calderoni, L., Ferrara, M., Franco, A., Maio, D.: Indoor localization in a hospital environment using random forest classifiers. Expert Syst. Appl. **42**(1), 125–134 (2015). https://doi.org/10.1016/j.eswa.2014.07.042

36. Fisher, J.A., Monahan, T.: Evaluation of real-time location systems in their hospital contexts. Int. J. Med. Inf. **81**(10), 705–712 (2012). https://doi.org/10.1016/j.ijmedinf.2012.07.001

37. Han, G., Klinker, G.J., Ostler, D., Schneider, A.: Testing a proximity-based location tracking system with bluetooth low energy tags for future use in the or. In: 2015 17th International Conference on E-health Networking, Application & Services (HealthCom), pp. 17–21. IEEE (2015). https://doi.org/10.1109/HealthCom.2015.7454466

38. Hosaka, R., Murohashi, T.: Experimental trial to detect medical engineering equipments in hospital by passive uhf rfid tag. ISMICT, 81–84 (2013). https://doi.org/10.1109/ISMICT.2013.6521704

39. Ionescu, G., Martinez de la Osa, C., Deriaz, M.: Improving distance estimation in object localisation with bluetooth low energy. SENSORCOMM, 45–50 (2014)

40. Malajner, M., Planinsic, P., Gleich, D.: Uwb ranging accuracy. In: Liatsis, P., Uus, A., Miah, S. (eds.) 2015 International Conference on Systems, Signals and Image Processing (IWSSIP 2015), pp. 61–64. IEEE, Piscataway, NJ and Piscataway, NJ (2015). https://doi.org/10.1109/IWSSIP.2015.7314177

41. Seong, J.H., Choi, E.C., Lee, J.S., Seo, D.H.: High-speed positioning and automatic updating technique using wi-fi and uwb in a ship. Wirel. Pers. Commun. **94**(3), 1105–1121 (2017). https://doi.org/10.1007/s11277-016-3673-2

Autoclave Sterilization Powered Medical IoT Sensor Systems

Mateusz Daniol[1,2(✉)], Lukas Böhler[1,2], Anton Keller[1], and Ryszard Sroka[2]

[1] Aesculap AG, Am Aesculap-Platz, 78532 Tuttlingen, Germany
`mateusz.daniol@aesculap.de`
[2] Department of Metrology and Electronics, AGH University of Science
and Technology, A.Mickiewicza Ave. 30, 30-059 Cracow, Poland

Abstract. The purpose of this study is to explore the possibilities of harvesting the thermal energy from steam sterilization process to power the IoT sensor node. Thermoelectrical generators based heat recovering have been used for powering IoT sensor nodes. The design process of the TEG based energy harvesting application is described in details. All vital parts of the system like choosing the suitable TEG module, heat storage material, power storage device, a power management system as well as insulation material to create the temperature gradient across the TEG were precisely described. The temperature-voltage characteristics of the module are analyzed within the test setup of standard steam sterilization. Power consumption of a CC2650 Bluetooth module is analyzed and optimized to maximize the power efficiency and the lifetime. During this study self powered Bluetooth IoT sensor node was developed. Power consumption software optimization have been applied resulting in the lifetime of over 10 days after single sterilization cycle.

Keywords: Energy harvesting · Steam sterilization · Sensors · IoT

1 Digitalization and Common Issues with Powering Surgical Tools

Digital healthcare is a very fast growing area of medical industry with a huge potential of delivering secure and high-quality patient care as well as driving greater business efficiency. Many reports suggest that digitalization is becoming a new business opportunity for healthcare industry involving Internet of (Medical) Things, Big Data and automatization. This could allow the connection between patients, healthcare professionals, manufacturers and providers [1,2]. There are strong indications that bringing the idea of traceable and connected devices to the operating room can improve significantly the efficiency and safety of surgical procedures [3,4]. Moreover, some papers highlight the direction towards digitalization in hospitals management systems to create so-called "Smart Connected
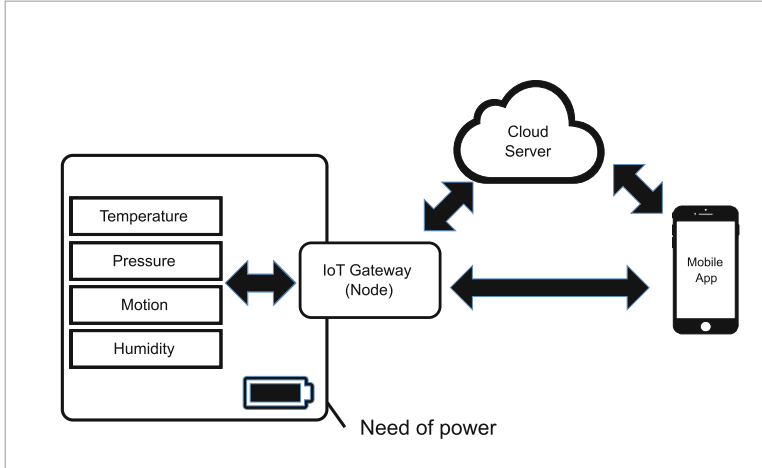
Fig. 1. Idea diagram of typical IoT architecture

Hospitals" with sensor nodes and tracking systems. To make this possible many different technologies are suggested for medical IoT, like RFID, ZigBee, Narrow Band Bluetooth, Bluetooth-Low-Energy [5]. Medical and surgical tools incomparably differ from other IoT enabled devices for consumer markets due to the need of sterilization. However, each of these devices and technologies require a power source. In the case of surgical tools and containers the power source needs to stand multiple steam sterilizations.

1.1   IoT Node

We can describe the Internet of Things as a network of physical devices embedded with electronics, sensors, software and connectivity which allows data exchange between these objects. Each object has its own unique identifier and is able to inter-operate within the existing internet infrastructure. The scheme of IoT system is shown in the Fig. 1. IoT node (gateway) equipped with a given set of sensors can communicate with cloud server with Wi-Fi connectivity or with the smartphone App with Bluetooth-Low-Energy. The data are then stored on the cloud server and synchronized among the all connected apps and devices. In this work we have used a SensorTag2 from Texas Instruments as a IoT node platform with Bluetooth-Low-Energy connectivity.

1.2   Steam Sterilization

In hospital environment there is an emphasis on clean, sterile environment. Special case in here is an operating room, where the sterile zone is needed for the patient safety. Every single medical instrument used during the surgical procedure needs to be maintained and sterilized according to the standards. Steam

sterilization, described in detail in ISO:17665 standards, involves instrument exposure on the temperatures ranging from 30 to 140 °C and the pressure from −1 to 3 Bars. The temperature inside the sterilization chamber is shown in the Fig. 2. During the sterilization process, the temperature reaches the maximum of 135 °C for 7 min. At the same time the pressure reaches its maximum at 3,1 Bar. These two factors causes the chamber environment to be very harsh for any electronic devices and power sources. Thus it was crucial to make a review on possible high temperature energy storage technologies, and find a new method to power the electronics and charge the energy storage during the steam sterilization.
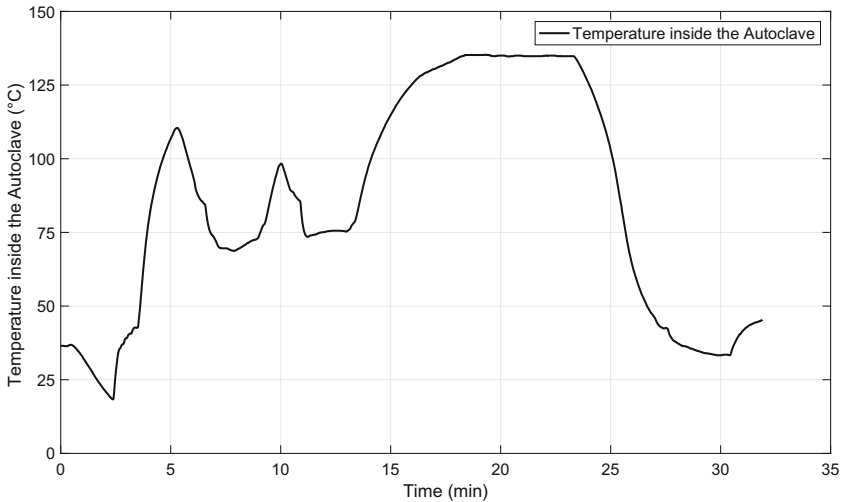
Fig. 2. Typical changes of the temperature inside the autoclave during steam sterilization

1.3   High Temperature Power Sources

High temperature inside the autoclave narrows significantly the choice of technologies for energy storage. In general, we can distinguish following energy storage systems (ESS) available on the market:

– Lithium primary cells
– Lithium secondary cells
– Supercapacitors
– Lithium-ion supercapacitors
– Energy harvesting systems

Lithium based batteries are known as a good and reliable power storage for consumer electronics with operating temperature range from −20 to 60 °C. Some of the lithium-ion chemistries can handle the temperatures of up to 150 °C, like

$Li-SOCl_2$ or $Li-CuO$ [6], however they are non-rechargeable and their energy density of 500 and 300 $Whkg^{-1}$ respectively determines very low maximum constant current and fast drain after working in high temperature environment. The secondary lithium-based cells in general cannot withstand the temperatures of a steam sterilization. When a cell is heated above 130 °C, and the heat cannot be dissipated by the cell, the exothermic processes will proceed causing rapid temperature rise which usually leads to gas release and explosion of the cell [7,8]. This process is called thermal runaway. Hammami et al. suggests, that thermal runaway can occur spontaneously after the cell reaches the temperature of 80 °C [9]. Recent advances in battery manufacturing processes led to development of small factor cylindrical cells with stable chemistry which prevents them from thermal runaway. However, the capacitance is limited to 3 mAh.

The second common type of ESS are electrolytic double layer capacitors (EDLC) which can also form supercapacitors. Supercapacitors usually employ acetonitrile or an organic carbonate solvent as the electrolytes. However, the boiling point of acetonitrile is only 82 °C and reaching this temperature causes the evaporation of the solvent and increase of the internal cell pressure which can lead to explosion. Recently, FastCAP Systems revealed a new technology of supercapacitors with an operational temperature up to 150 °C [10], however the cost effectiveness of this solution for the IoT node remains disputable. The other issue which needs to be taken into consideration is a self-discharge rate which can be high in the case of thermally stressed supercapacitors.

The special case of energy storage device is a lithium-ion supercapacitor, which is a hybrid of an EDLC and a lithium-ion battery. It is composed of a negative electrode doped with lithium ions (typical battery material), an activated-carbon positive electrode (typical capacitor material), and an organic electrolyte containing a lithium salt. Recent research shows that the exotermic reactions inside the Li-Ion supercapacitor begin at 90 °C leading to thermal runaway at 175 °C [11].

In conclusion, high temperature ESS are still in an early stage of development and suffer from limited thermal, chemical, and electrochemical stability at increased temperatures, resulting in their short lifetimes [6]. To prevent this, the energy storage in the sterilizable IoT node needs to be insulated from the high temperature and the self-discharge needs to be compensated in sterile, maintenance-free way. One way to recharge ESS is to use the energy harvesting techniques and recover the energy from the sterilization process itself.

1.4   Energy Harvesting

Waste heat recovery became recently an important topic causing rapid development in energy harvesting solutions.

The heat can come also from the steam sterilization process inside the autoclave. One of the most robust and reliable method of recovering the energy from the heat is using thermoelectrical generators (TEGs). Each TEG module is then composed of pairs of TE couples connected together electrically in series

and thermally in parallel, which directly convert a thermal energy that passes through them into electricity based on Seebeck's effect. This is shown in Fig. 3.
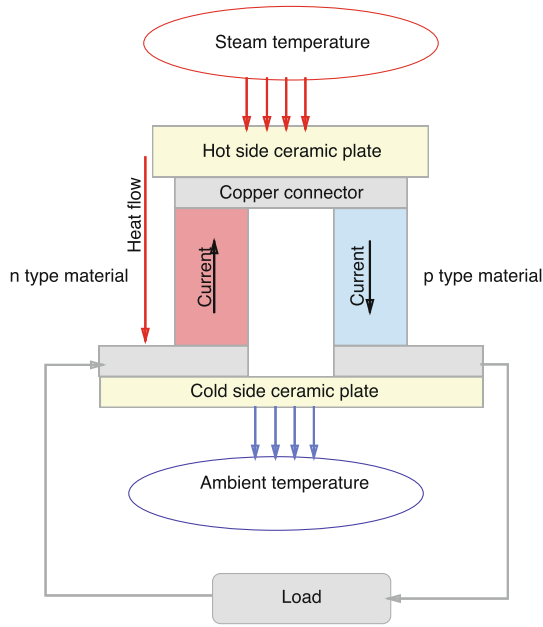
Fig. 3. Scheme of a single thermoelectric pair. It comprises of n-type and p-type semi-conductors. Heat flows from hot side to cold side and electrical current is flowing from n-type to p-type material due to temperature gradient

The efficiency of TEGs is usually relatively low. Studies shows an average efficiency of 8.45% of the TEG modules available on the market [12]. However, newest technologies can improve the efficiency up to 20% [13]. One way to do this, is to apply the oscilating temperature on the hot side of the TEG as some authors suggest [14]. This can be the case where the oscilating temperature of steam sterilization can improve the energy harvesting efficiency. Despite their robustness, TEGs are limited to the applications in extreme environments like space exploration, automotive industry, aircraft, military, and heavy industry. This was caused mainly by their relatively low efficiency and high cost of the modules [13]. During the last years an increased need of self-powered sensors is observed. Not only for industrial applications but as well for IoT nodes, wearables, surveillance systems, and everywhere where there is a need for stable, compact and maintenance-free energy source. Despite this fact applications of thermal energy harvesting in medical industry are still in their early infancy and are mainly focused on implantable or wearable devices [15,16]. One of the most important factors affecting the energy harvesting efficiency is matching the TEG to the working environment. In the case of the autoclave the temperatures varies

from 20 to 140 °C, with the maximum temperature gradient of 70 °C. Usually TEG modules are optimized for the high ambient temperatures as well as high gradients. For instance, low cost Bi_2Te_3 based modules are usually working with temperature gradients up to 300 °C. Thus, the steam sterilization environment needs to be considered as a low gradient conditions. Literature shows that in such a conditions it is possible to use thermoelectrical coolers as a TEG [17]. To choose the optimal TEG for given application thermal resistance of the heatsink and thermal resistance of the TEG needs to be known. The thermal resistance of the heatsink is taken usually from the producer specifications or material datasheet. However, the thermal resistance of the TEG needs to be calculated, and the most fitting TEG module needs to be chosen. The thermal resistance of the TEG can be estimated from the parameters given by the TEG manufacturer and in general is described as:

$$\theta_m = \theta_k \sqrt{1 + ZT_1}$$

where: θ_k - Thermal resistance of the heatsink

Z - figure of merit of the thermoelectric materials of the generator

T_1 - Temperature of the heat source

Matching these parameters with appropriate TEG ensures the thermal resistance match between TEG and heatsink and maximum power generation in given environmental conditions [18].

The work principle of TEG implies a need to generate a heat flow across the TEG in the environment with set ambient temperature. For this the heatsink made of aluminium can be thermally coupled with TEG on the hot side and a heat storage material on the cold side. Many different materials were reviewed according their parameters, heat accumulation, heat transfer and the cost effectiveness. Recently there are numerous heat storage materials available on the market, including phase changing materials with great thermal accumulation properties [19]. However, according to the dynamical changes of the temperature inside the autoclave within the cycle, it was essential to choose the material with not only high specific heat but also with high heat conductivity.

1.5   Power Conditioning Circuits

The energy from heat recovery of the TEG is linked to the energy storage usually via a DC-DC converter. Currently available state-of-the-art DC-DC converters can be divided in two categories: actively controlled with single inductor and passively switched coupled inductors without active control circuit. The single inductor circuits, due to their active control circuits, can match the input impedance and boost-up ratio dynamically to allow so called maximum power point tracking (MPPT). MPPT checks the output of the TEG module, compares it to energy storage voltage then fixes what is the best power that TEG module can produce to charge the storage and converts it to the best voltage to get maximum current into energy storage [20]. The disadvantage of single inductor architecture is relatively high minimum input voltage of 100 mV and a cold start

input voltage of 330 mV due to the poweing additional passive cold-start voltage converter. A market example of such an architecture is bq25504 from Texas Instruments.

Table 1. Comparison of bq25504 and LTC3108

Parameter	bq25504	LTC3108
Switching control	Active feedback control	Passively controlled switch
Startup voltage	330 mV	Down to 20 mV
Min. input voltage	100 mV	Down to 20 mV
Max. input voltage	3 V	Down to 400 mV
Efficiency	Up to 85%	Up to 50%
Voltage regulation	Controlled boost ratio	Linear drop-out

On the other hand, the coupled inductors based, fully passive circuit, DC-DC converter has a very small input startup voltage of 20 mV. The disadvantage of this topology is lower efficiency. This is caused by two main factors. Firstly the inefficiencies of passive switching circuits, inductor couplings and rectification required after the boosting stage. Secondly, using the fixed boost ratio defined by coupled inductors results in a boosted voltage which is much higher than the energy storage voltage for higher input voltages. Thus, a passively controlled low-dropout regulator is used, which causes linear efficiency decrease for higher input voltages. A market example of such an architecture is LTC3108 by *Linear Technology*. The comparison of these two chips are presented in Table 1.

As literature shows, the advantage of the bq25504 single inductor based solution is its ability to dynamic adaptation of MPPT algorithms resulting in higher efficiency, the tradeoff however is that it needs higher startup voltage to power up active circuit elements [21]. For this project the efficiency and wide input voltage was critical therefore the bq25504 has been chosen for further development.

2   Materials and Methods

The aim of this research was to design a maintenance-free, sterilizable IoT sensor node using energy harvesting from sterilization cycles. The whole system is composed from TEG module, DC-DC boost-converter, power control circuit, energy storage device, communication module and sensors serving as a load. The architecture diagram of the system is shown in the Fig. 4. The output of the TEG is connected to the boost converter and power management system. This allows to feed the load with sufficient power or to store it in a power buffer. As a sensor module Texas Instruments SensorTag2 with CC2650 Bluetooth chip has been used. Small Murata UMAC040130A003TA01 hybrid li-ion battery served as power buffer. In the experimental part of this work, the power generation from the steam sterilization and software optimizations of IoT module were examined.
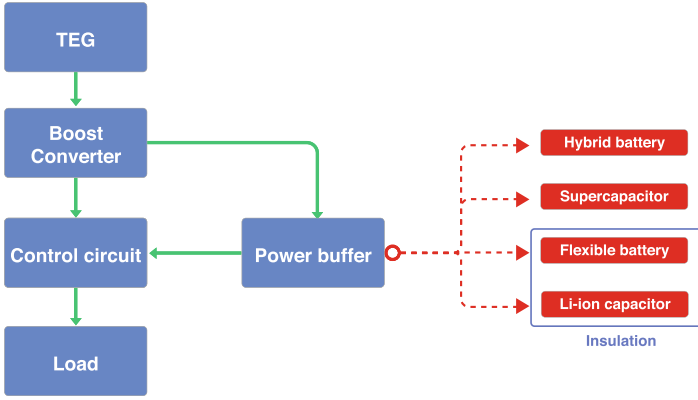
Fig. 4. Block diagram of the prototype

Two main challenges of the prototype design were to overcome. First of all to protect the electronics and energy storage from high temperatures, and to maximize the efficiency of thermal energy conversion among the TEG. The other issue was to optimize the energy consumption of Bluetooth-Low-Energy according to the average use-case of surgical containers. The module should be charged after the single steam sterilization cycle and work for at least 5 days after the sterilization. The cross-section of the first prototype is shown in Fig. 5. In this study, two *Ferrotec Nord TMG-127-0.4-1.6* connected in series have been used. The whole module was insulated using aerogel-based materials which provides extremely low heat conductance as 0.02 [W/m*K]. The output of the energy harvesting module is connected to the bq25504 Boost-up converter with MPPT algorithm and power storage maintenance control circuits. The bluetooth module based on CC2560 chip and temperature sensors are connected to the power buffer through the control circuit.
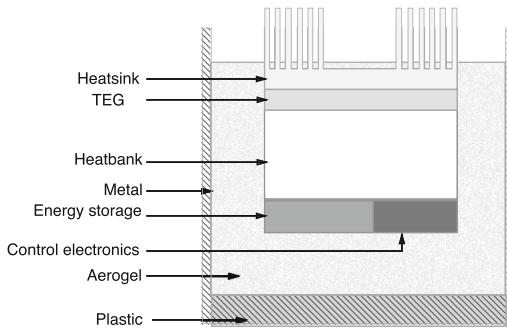
Fig. 5. Cross-section of the prototype of the self powered IoT node

2.1   Software Optimization

The lifetime of a Bluetooth-Low-Energy device is determine how much power the device's components consume. Thats why it was essential to evaluate and to choose the optimum device hardware setup regarding the microprocessor, radio and the sensors.

In the case of Bluetooth-Low-Energy devices the counter-intuitive rule of low-power communication is that listening for data is much more expensive in the case of energy than sending data. This is because of longer intervals when the radio needs to be turned on. In this case, the optimization was to find out the appropriate interval on advertising wake-up packets which tells the receiving device that it is a time to receive the data. This technique, described in [22], is called sampled listening. The other optimization is made by periodically wake-ups of the device from the stand-by mode to check if there is any pending connection requests. If so, it starts to do sampled listening, but only for a short while. The last two improvements were regarding the data rate and the responsiveness. The data rate is a parameter indicating how often the data will be sent by the Bluetooth module and responsiveness tells about how often the module needs to respond to the commands.

2.2   Test Setup

We have examined the prototype regarding efficiency of power generation, power consumption and estimated lifetime of the IoT node after single sterilization cycle. During the efficiency tests the module was placed inside the climate chamber which have simulated the temperature changes during the steam sterilization. The temperature on the module surface as well as inside the module and the voltage on the TEG was measured. The power consumption tests involved the measurement of a power consumption of CC2650 chip with different sensors attached to it and in different operating modes. The module was powered with the single coin cell battery and the measurements were performed by taking the current characteristic in several operating modes. Finally, the optimizations of the chip software was made to maximize the power efficiency and to extend the lifetime of the whole IoT node after single sterilization. The power consumption before and after the optimizations was compared.

3   Results

3.1   Power Generation

Power generation tests results are shown in the Fig. 6. The temperature in the climate chamber simulated the steam sterilization process (red line) and reaches 135 °C after 20 min. At that time the voltage of the TEG reaches its maximum of 2.3V (green line). The temperature inside the module reaches 80 °C (blue line). Next, the chamber is being opened, which is visible on the diagram as a sudden drop of the temperature inside the chamber. The heatflow changes its
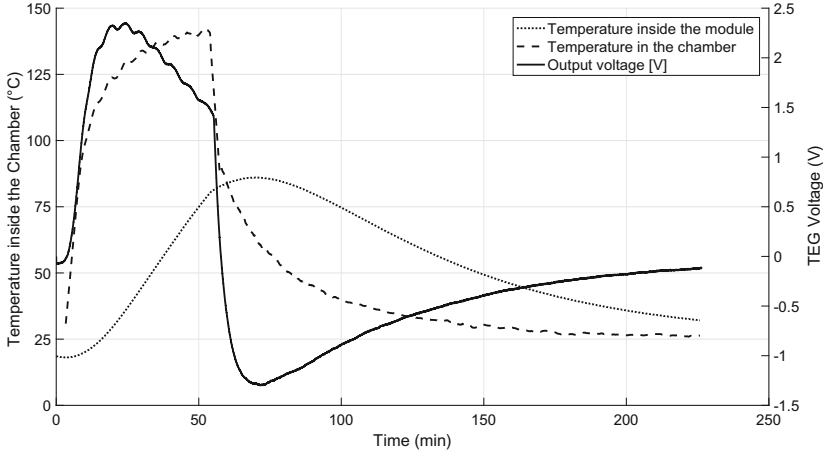
Fig. 6. Temperature and pressure data of steam steriliation process

direction and it goes from heatbank to the outside of the module causing the TEG to change its polarity and reach the −1.4 V. After that the module is cooling down for the next 140 min and the voltage generated on the output of TEG is slowly decreasing until it reaches the value of 0.08V which is the terminal test condition as the boost converter stops working. Using the low power rectification and conditioning circuit allows to charge the ESS even when the voltage of TEG is negative.

3.2   Power Consumption

The power consumption (in mA) of the CC2650 varies depending on the current operating mode and sensors. The several different measurements have been taken, as it is shown in Table 2. All of the measurements were taken using standard demo-software available from *Texas Instruments*. It is worth to mention, that this software was not optimized on power consumption, and its main aim is to demonstrate all of the features of the *CC2650 SensorTag*. Thus the average power consumption during the tests was 6.12 mA, caused mainly by the motion sensor. There was significant power consumption during the startup of the module, thus it is not recommended to switch it to the off state.

As it is shown in Table 2, the power consumption without any optimizations is unacceptable in the case of maintenance-free energy harvesting based application. The standard average is a state when the SensorTag2 with CC2650 is operating with a standard software and with standard sensor setup.

3.3   Power Optimization

After the optimizations described in Sect. 2.1, the power consumption has been lowered more than over 95% to 0.3mA. Reducing the advertisinig time interval

Table 2. Power consumption of evaluated CC2650 SensorTag

Mode	Power consumption (mA)
Powering On	12
Standby	0.24
Discovery	2.0
All sensors off	0.24
Temperature sensor	0.80
Barometric sensor	0.50
Motion sensor	4.16
Standard average	6.12

improved the power efficiency the most, lowering the energy consumption to 0.4 mA. Awaking the device periodically to check the connection request instead of constant monitoring allowed to fall below 0.05 mA of average current consumption. The last two optimizations lowered the power power consumption to the level of 0.03 mA. This prolonged the lifetime of the module after a single sterilization to over 10 days, which exceeded the first assumptions of this project and is enough to power the IoT node in during the sterilization cycles interval.

4 Discussion

Numerous reports, market researches and papers show that digitalization and IoT will spread through medical industry. What is most important some studies shows that hospital management information specialists are ready for adopting the IoT technology in their workplace [23]. In this study we have proved, that it is possible to power the sensor IoT nodes even in such a harsh environment as a steam sterilization. More over, it was possible to store the harvested energy in a save way. This allows us to think about the digitalization of the surgical equipment. Every single medical sterile container, medical instrument or power device needs to be sterilized on a regular basis. This opens a lot of possibilities for the applications like surveillance the sterile status of a container. Integrating ultra low-power shock sensors brings us the possibility to log the critical events during the transport. However, further design optimizations and tests of energy-harvesting modules need to be done, as well as optimization of their power efficiency. With these issues resolved, it should be possible to reduce the cost and size of the single module. These two factors plays the key role in medical and hospital logistics. The results of this work show, that implementing energy harvesting powered IoT nodes opens possibilities to increase the digitalization level in the medical facilities resulting in better management, faster logistics operations, increase in sterile safety. As an end result it could lower the cost and increase the safety of medical treatment.

5   Summary

The aim of this study was to examine the possibility of energy harvesting from steam sterilization process for medical IoT sensor nodes. The results show, that powering IoT sensor node from autoclave sterilization is possible. Choosing the right thermoelectrical generator plays a crucial role in the whole design of energy harvesting device. Together with a heat storage and insulation materials the physical dimensions and the design can be determined. As it was reviewed, power storage devices like lithium based batteries and supercapacitors needs to be insulated and connected with a TEG via DC-DC converter including power management system. Several different power management topologies have been reviewed. It was examined that it is possible to store the energy in safe way in a lithium-based safe batteries or supercapacitors protected with an aerogel-based insulation. To optimize time of life of the sensor and communication electronics several optimizations have been made according to the software of the module. It reduced the energy consumption to the average of $0.04\,\text{mA}$. Thus, the power generated by a single procedure is sufficient to power the Bluetooth-Low-Energy module and sensors for over 10 days.

References

1. David, A.: The Digital Future of Healthcare (2017)
2. Kanan, R., Elhassan, O.: Batteryless radio system for hospital application. In: Proceedings of 2016 SAI Computing Conference, SAI 2016, pp. 939–945 (2016). https://doi.org/10.1109/SAI.2016.7556093
3. Huang, A., et al.: The SmartOR: a distributed sensor network to improve operating room efficiency. Surgical Endoscopy Other Interv. Techn. (2017). https://doi.org/10.1007/s00464-016-5390-z
4. Mahfouz, M., To, G., Kuhn, M.: Smart instruments: wireless technology invades the operating room. In: RWW 2012—Proceedings: 2012 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems, BioWireleSS 2012 (2012). https://doi.org/10.1109/BioWireless.2012.6172743
5. Zhang, H., Li, J., Wen, B., Xun, Y., Liu, J.: Connecting Intelligent Things in Smart Hospitals using NB-IoT. IEEE Int. Things J. **4662**(c), 1–11 (2018). https://doi.org/10.1109/JIOT.2018.2792423
6. Lin, X., Salari, M., Arava, L.M.R., Ajayan, P.M., Grinstaff, M.W.: High temperature electrical energy storage: advances, challenges, and frontiers. Chem. Soc. Rev. **45**(21), 5848–5887 (2016). https://doi.org/10.1039/C6CS00012F. http://xlink.rsc.org/?DOI=C6CS00012F
7. Bandhauer, T.M., Garimella, S., Fuller, T.F.: A critical review of thermal issues in lithium-ion batteries. J. Electrochem. Soc. **158**(3), R1 (2011). https://doi.org/10.1149/1.3515880. https://doi.org/10.1149/1.3515880
8. Spotnitz, R., Franklin, J.: Abuse behavior of high-power, lithium-ion cells. J. Power Sour. **113**(1), 81–100 (2003). https://doi.org/10.1016/S0378-7753(02)00488-3
9. Hammami, A., Raymond, N., Armand, M.: Runaway risk of forming toxic compounds. Nature **424**, 635 (2003). https://doi.org/10.1038/424635b. https://doi.org/10.0414/424635b

10. Cooley, J., Signorelli, R., Green, M., Sasthan, P., Deane, C., Wilhelmus, L.A.: Power system for high temperature applications with rechargeable energy storage (2012)
11. Smith, P.H., Tran, T.N., Jiang, T.L., Chung, J.: Lithium-ion capacitors: electrochemical performance and thermal behavior. J. Power Sour. **243**, 982–992 (2013). https://doi.org/10.1016/j.jpowsour.2013.06.012. https://doi.org/10.1016/j.jpowsour.2013.06.012
12. Chen, W.H., Wu, P.H., Wang, X.D., Lin, Y.L.: Power output and efficiency of a thermoelectric generator under temperature control. Energy Convers. Manag. **127**, 404–415 (2016). https://doi.org/10.1016/j.enconman.2016.09.039. https://doi.org/10.1016/j.enconman.2016.09.039
13. Champier, D.: Thermoelectric generators: a review of applications (2017). https://doi.org/10.1016/j.enconman.2017.02.070
14. Chen, W.H., Huang, S.R., Wang, X.D., Wu, P.H., Lin, Y.L.: Performance of a thermoelectric generator intensified by temperature oscillation. Energy **133**, 257–269 (2017). https://doi.org/10.1016/j.energy.2017.05.091. http://linkinghub.elsevier.com/retrieve/pii/S0360544217308460
15. Kanan, R., Bensalem, R.: Energy harvesting for wearable wireless health care systems. In: 2016 IEEE Wireless Communications and Networking Conference, vol. 2016-Septe, pp. 1–6 (2016). https://doi.org/10.1109/WCNC.2016.7565034. http://ieeexplore.ieee.org/document/7565034/
16. Lundager, K., Zeinali, B., Tohidi, M., Madsen, J., Moradi, F.: Low power design for future wearable and implantable devices. J. Low Power Electron. Appl. **6**(4) (2016). https://doi.org/10.3390/jlpea6040020
17. Buist, R.J., Lau, P.G.: Thermoelectric power generator design and selection from TE cooling module specifications. In: XVI International Conference on Thermoelectrics, 1997 Proceedings ICT 1997. (616), pp. 551–554 (1997). https://doi.org/10.1109/ICT.1997.667589
18. Lineykin, S., Ruchaevsky, I., Kuperman, A.: Analysis and optimization of TEG-heatsink waste energy harvesting system for low temperature gradients. In: 2014 16th European Conference on Power Electronics and Applications, EPE-ECCE Europe 2014 pp. 1–10 (2014). https://doi.org/10.1109/EPE.2014.6910778. http://www.scopus.com/inward/record.url?eid=2-s2.0-84923902652&partnerID=40&md5=6115ae5a48219981242ca2d3f23a1686
19. Alva, G., Liu, L., Huang, X., Fang, G.: Thermal energy storage materials and systems for solar energy applications. Renewable and Sustainable. Energy Rev. **68**, 693–706 (2017). https://doi.org/10.1016/j.rser.2016.10.021. https://doi.org/10.1016/j.rser.2016.10.021
20. Laird, I., Lu, D.C.: High step-up DC/DC topology and MPPT algorithm for use with a thermoelectric generator. IEEE Trans. Power Electron. **28**, 7 (2013). https://doi.org/10.1109/TPEL.2012.2219393
21. Thielen, M., Sigrist, L., Magno, M., Hierold, C., Benini, L.: Human body heat for powering wearable devices: from thermal energy to application. Energy Convers. Manag. **131** (2017). https://doi.org/10.1016/j.enconman.2016.11.005
22. Hui, J.W., Culler, D.E.: IPv6 in low-power wireless networks. Proc. IEEE **98**(11), 1865–1878 (2010). https://doi.org/10.1109/JPROC.2010.2065791
23. Muhammad, A.P., Akram, M.U., Khan, M.A.: Survey based analysis of internet of things based architectural framework for hospital management system. In: 2015 13th International Conference on Frontiers of Information Technology (FIT), pp. 271–276 (2015). https://doi.org/10.1109/FIT.2015.54. http://ieeexplore.ieee.org/document/7421012/

Using an Input Data Segregation Algorithm to Minimise the Error of the Fuzzy Controller in the Metrological Correction System of Electric Energy Meters

Bartosz Dominikowski[(✉)], Krzysztof Pacholski, and Piotr Woźniak

Lodz University of Technology, Żeromskiego Street 116, 90-924 Lodz, Poland
{bartosz.dominikowski,krzysztof.pacholski,
piotr.wozniak.l}@p.lodz.pl

Abstract. The authors of this paper presented the possibility of using a fuzzy controller in the conversion factor correction system associated with the energy meter's current channel. The accuracy of non-adaptive fuzzy controllers is significantly affected by the relevant expert knowledge in the form of rules stored in the database. In order to increase fuzzy controller accuracy, the k-means clustering method was used to group the input data of the controller (peak value of the output signal of the energy meter's current transducer and its derivative). This analysis can be conducted to extract central points that represent particular input data groups. Based on computer testing of fuzzy controller output signals performed by the authors, the assignment of membership functions to the central points of the input data groups should be done by the expert at the beginning while designing the rules. Additionally, this paper presents the possibilities of tuning the fuzzy controller by changing its parameters.

Keywords: K-means clustering · Fuzzy controller · Gain corrector

1 Introduction

Based on the Polish Standard PN-EN 50463-2 "Railway applications – Energy measurement on board a train", DC energy meters were designed in 2013 to record the energy consumption of electric multiple units [1]. The costs incurred for the energy consumed are significant in the budget of each railway carrier, so the accuracy of determination is very important. An indication of the energy consumption of the on-board meter installed in the electric locomotive cab enables the driver to use the correct technique for train driving (the so-called eco-driving). Railway carriers frequently compare electricity consumption of similar railway vehicles operating on a route in order to avoid uneconomical transport operations.

Modern electric rolling stock is equipped with three-phase inductive drives [2]. The rotational speed of asynchronous traction motors is controlled by DC/AC semiconductor-type voltage bridge inverters. The method that is used to control the operation of the electric locomotive drive is based on a sine wave controller with a Space Vector Pulse Width Modulation (SVPWM), allowing the generation of a sinusoidally variable current

flow at the output terminals of the inverter (on the AC side) [2]. The input circuit of this converter (on the DC side) is connected to the contact line via a pantograph. At the movable contact point of the electric locomotive and the contact line, a dynamically changing current peak occurs due to the non-linear switching load (high crest factor – CF_i) [3]. This current is measured by a current transducer whose output terminals are connected to the input of the energy meter's current channel via an operating amplifier.

In the electric energy transducer, the energy measurement is subject to an error that can be determined with the following relationship [3]:

$$\delta_{wm} = \frac{(CF)_u (CF)_i}{U_m I_m PF} \delta_{ADC} \tag{1}$$

where: $(CF)_u$, $(CF)_i$ – crest factor of the voltage and current signals of low-voltage input signals of the electric energy meter; U_m and I_m – voltage and current amplitudes of the above-mentioned input related to rated values; PF – power factor given as the quotient of active power and apparent power; δ_{ADC} – resolution error of the ADC converter in the energy meter's input channels.

In the measuring circuits of the electric energy meters, it is only possible to influence the peak of the output signals of the current and voltage transducers by changing their gain. Operation of the meter with input signals with values close to the reference voltage level of the ADC converter reduces the error δ_{wm} (relationship (1)) during the electric energy measurement. The deviation of the contact line supply voltage from its rated value ($U_N = 3$ kV) is minor. Therefore, the voltage transducer is connected to the input channel of the electric energy meter via an operating amplifier with a selected fixed gain so that the value of the input signal is close to the upper limit of the measuring range of the energy meter's voltage channel. Yet, the change of the peak value of the output signal of the current transducer covers the entire conversion range of the current input channel of the electric energy meter. To ensure that electric energy is measured with high accuracy under railway traction operating conditions, the output signal of the current transducer should be amplified continuously to reach the upper limit of the input range of the ADC converter in the energy meter's input channel.

Manufacturers of electric energy meters use operating amplifiers at the input of the current channel, with programmable gain factor or those operating in parallel structure with selected fixed conversion factors. The authors of this paper proposed the use of smart correction that allows continuous determination of the conversion factor of the energy meter's current channel. The operating principle of the fuzzy corrector is explained later in this paper.

2 Fuzzy Controller in the Automatic Gain Control System for the Input Values of the Electric Energy Meter

For the conversion factor corrector system associated with the current channel of the electric energy meter, the authors used three fuzzy controllers operating in parallel structure (marked in Fig. 1 as F_1, F_2 and F_3). These controllers differ in the input voltage ranges and the corresponding gain values. The division into voltage input sub-

ranges results in faster attainment of the desired value level (close to the reference voltage of the ADC converter of the energy meter's input channel) by the rising and falling edges of the signal being measured.

The input signal of the fuzzy controllers is the vector x, which contains the following values: V_{in} – instantaneous peak output voltage of the current transducer (marked in Fig. 1 as CT) and its derivative determined numerically in the microprocessor system (MCU) and converted into geometrical degrees. The output signal of the above-mentioned smart controller is a scalar value representing the desired gain of the current signal of the energy meter's input channel (marked in Fig. 1 as Ch_1).
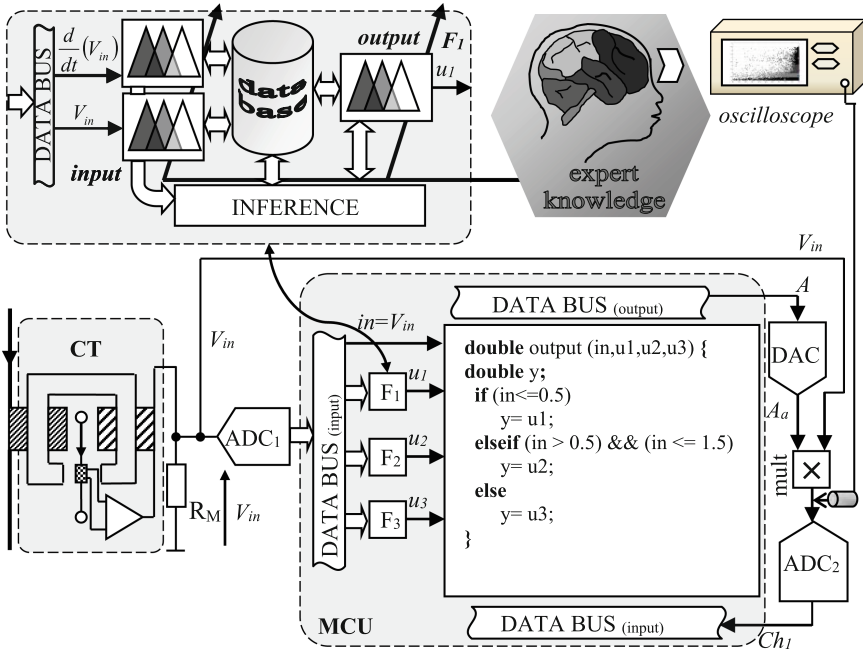
Fig. 1. Block diagram of information propagation in the fuzzy controller in a smart gain correction system

The operating principle of the controller is presented in a few steps below. The input vector $- x = \begin{bmatrix} v_{in} & \frac{d}{dt}(v_{in}) \end{bmatrix}$ of the fuzzy controller is sent to the fuzzifier to be transformed into a fuzzy set with a specific membership function. The role of the used membership functions is played by the Gaussian functions defined by the relationship [4] that, according to the expert, are suitable to achieve a smooth and continuous hypersurface of the fuzzy controller input/output [4].

$$\mu_{Gauss}(in,\, a,\, b) = \exp\left[-\left(\frac{in-b}{a}\right)^2\right] \qquad (2)$$

where: b – the centre of the membership function (core), a – width of the membership function, in – input variable of the fuzzy controller.

In the fuzzy inference system, the expert enters complex, intuitive heuristic rules into the database (Fig. 1 – Database), given by the relationship [4, 5]:

$$IF \underbrace{(V_{in} \text{ is } big) \text{ } AND \left(\frac{d}{dt}(V_{in}) \text{ is } big\right)}_{p},$$
$$THEN \underbrace{(A \text{ is } small)}_{k}$$

(3)

where: V_{in} – crisp value of the voltage output signal of the current transducer (marked in Fig. 1 as CT), $\frac{d}{dt}(V_{in})$- derivative of the above-mentioned signal (expressed in geometrical degrees), big and small – linguistic values defined as fuzzy sets obtained from the universe (space within which the set is defined): V_{in}, $\frac{d}{dt}(V_{in})$ and A.

The rule used to process the fuzzy information given by the relationship (3) includes: input values of the controller specified in the complex conjunctive antecedent (p) connected using the AND conjunction, whereas the consequent (k) indicates the corresponding gain of the measuring channel – A.

In the fuzzy controller of the corrector system, Mamdani's inference method was used, consisting in connecting antecedents and consequents (defined by the symbols p and k – relationship (3)) with the use of the t-norm operator. On the basis of numerous experimental computer tests in the computational software – Matlab/Simulink – using various implication operators, the best results of the input quantity were obtained while using the Larsene-type t-norm operator, i.e. PROD (product of the antecedents and consequents [10]). The final stage of the process that determines the smart conversion factor of the current input channel in the electric energy meter is to have it converted in the defuzzifier into a crisp value. In this block, due to a short response time, the so-called Centre of Maximum method was selected, as it performs the operation of determining the weighted average of the outputs for which the resulting membership function reaches a maximum of [8, 9].

The operating principle of smart gain corrector is as follows:

1. initial measurement of the peak value of output voltage V_{in} of the resistor R_M of the energy meter's current transducer (CT) by the ADC_1 system (Fig. 1) and saving it to the processor memory (MCU) and numerical determination of its derivative;
2. conversion of the input data into the output value A, representing the desired gain, by the fuzzy controller blocks;
3. conversion of the digital output quantity of the fuzzy controller A into an analogue value A_a by the DAC converter (Fig. 1);
4. product operation of the analogue signals $V_{in} \times A_a$ and a monolithic multiplier (MULT – Fig. 1);
5. measurement of the output quantity of the MULT multiplier by the ADC_2 converter (Fig. 1) in the input channel Ch_1 of the electric energy meter.

2.1 Error Minimisation for the Fuzzy Controller

The main factors determining the accuracy of input signal conversion into output signals in the rule-based fuzzy controllers are, but not limited to, an appropriate number of fuzzy rules stored in the controller's base and appropriate spacing of the cores of the membership functions of fuzzy sets (in the rules) along the input and output axes [4]. For generation of a fuzzy controller input/output hyperspace in the gain corrector, the applied method often consists in numerical computation of a series of gain values based on the relationship $A(j) = level/V_{in(j)}$ [V/V] (where 'level' is the desired input voltage of the ADC_2 converter in the current channel Ch_1 of the electric energy meter) for the input signal Vin changing at a fixed step – j, from 0 V to 5 V. At the determined gain points $A(j)$, the cores $Core(\mu_{V(j)}(V_{in}(j))) = 1$ of the membership function $\mu_{Vin(j)}$ of the peak voltage value of the energy meter's output current transducer are established in the fuzzy sets $V(j)$. The same number of cores is spaced at $90\%j$ steps and $V_{in}(j)$ for the axis of the fuzzy controller input space. The lines drawn from the vertices of the membership function cores create an evenly-intersecting orthogonal grid of the controller plane. In the case of curved-out input/output areas of the fuzzy system, shorter distances between the nodes are used in such sectors (the nodes created by the rule stored in the database, given by the relationship (3) [4]. If the inferencing fuzzy controller of the gain corrector is so designed, it becomes a universal system able to operate in each measuring channel equipped with an ADC converter. Low accuracy of the solution is its disadvantage. According to the authors, the parameters of the fuzzy controller conversion function should be appropriately selected for a specific application, ensuring more accurate operation.

Due to the high dynamics of changes in the actual current of the contact line under load applied by modern asynchronous drives, the value of V_{in} should be continuously amplified to the value level:

- $level_1$ = 4 V (i.e. 80% of the reference voltage equal to 5 V of the ADC_2 converter in the current channel Ch_1 of the electric energy meter) – for the slope in relation to the time axis of this signal converted into geometrical degrees: from 0° (constant component) to 75°;
- $level_2$ = 3 V (i.e. 60% of the reference voltage equal to 5 V of the ADC_2 converter in the current channel Ch_1 of the electric energy meter) – for the signal with a slope of >75°.

The fuzzy controller presented in this paper was designed to operate in the automatic input-gain control system of the traction current channel of the electric energy meter. For this purpose, a suitable expert with knowledge and experience in the field of traction drives was selected. The expert analyses the input data of the fuzzy controller by isolating the most common clusters of information and selecting their representative and then transferring the set of rules to the base. With this analysis, it is possible to minimise the fuzzy controller error by locating the nodes of the input space division grid $X \left(X \in V_{in} \times \frac{d}{dt}(V_{in}) \right)$ of the fuzzy controller cluster centres. For this purpose, the fuzzy system designer uses the analysis method of k-means clustering. The number of data groups (parameter k) to be obtained is entered by the expert at the beginning of the k-means algorithm. In the next step, the k representatives of the emerging groups are randomly selected so that they are as far apart from each other as possible. In the next step

of the k-means algorithm, all elements of the input universe of discourse of the fuzzy controller $-X$ are assigned to the closest initial group (cluster). For each group, based on the arithmetic mean of the coordinates of the elements included, the centre c_i ($i = 1, …, k$) (centroid) is determined. The next step is to recalculate the allocation of elements to the groups based on the computed distances from the determined centroids. The new centres are calculated as long as there is data migration between the adjacent groups. In the k-means method, an optimum division of data into clusters is provided by the determination of such groups that minimise the criterion function given by the relationship [6]:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \underbrace{\left\| x_i^{(j)} - c_j \right\|^2}_{dm} \tag{4}$$

where: dm – distance measure, k – number of groups (clusters), c_j -centroid for the group j, n – size of the set (groups)

The operating mode of the k-means method is shown in Fig. 2 for the input space of the fuzzy controller, created on the basis of the actual current signal of the contact line under load (this waveform was published in [7]) and its derivative. The effect of the applied data clustering is data grouping into clusters from G_1 to G_5.
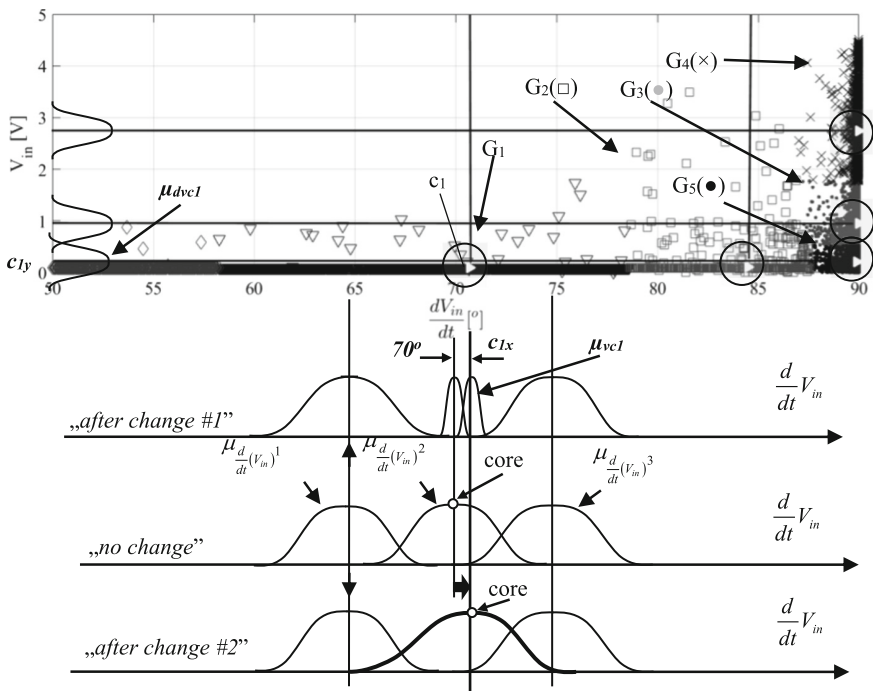
Fig. 2. Data grouping of the input space of the fuzzy controller using the k-means method, the symbols marked on the graph indicate the data included in the group: G_1. -▲, G_2. -◊, G_3.- □, G_4. $_5$ - ●, G_6. -×, the centroids of each group (C_1, …, C_5) are marked as ► (filled with white color)

Between the cores with the values 60, 65, 70 and 75° of the membership function: $\mu_{\frac{d}{dt}(V_{in})_1}$, …,$\mu_{\frac{d}{dt}(V_{in})4}$ (Fig. 2: axis $\frac{d}{dt}(V_{in})$ titled "before change") designated as "before change"), determined classically for the designed fuzzy controller (without preliminary input data analysis), there is G_1 as an example data group (elements marked with ▲), as shown in Fig. 2. The group is represented by the centroid c_1 with its coordinates (c_{1x}, c_{1y}). According to the authors, the expert should introduce here new membership functions in the input space of the fuzzy controller μ_{vc1} and μ_{dvc1} with the core for the V_{in} axis in the place of the X-coordinate c_{1x} = 72 V and for the $\frac{d}{dt}(V_{in})$ axis for the Y-coordinate c_{1y} = 0.2° (Fig. 2 axis $\frac{d}{dt}(V_{in})$ designated as "after change 1"). The output space of the smart controller remains unchanged. In the controller rule, the function μ_{vc1} is related to the function μ_A (fuzzy gain output function) so that the product of their cores depending on the input voltage slope V_{in} is close to the $level_1$ or $level_2$ limit. This type of procedure should be performed for the remaining determined data groups (G_2, …, G_k). It is essential not to exceed the admissible number of rules stored in the fuzzy controller database, as given by the relationship [4]:

$$N_R = (N_S)^{N_{in}} \tag{5}$$

where: N_R - number of fuzzy rules, N_{in} - number of model inputs, N_S - number of input fuzzy sets (identical for each input of the fuzzy controller).

If it is not possible to provide an additional membership function due to the limitation resulting from the relationship (5), the existing cores should be moved accordingly so that the closest one is at the coordinates specifying the centroid of the given group, as shown graphically in Fig. 2 (axis $\frac{d}{dt}(V_{in})$ designated as "after change 2").

The authors completed a computer simulation of the classically designed fuzzy controller with appropriately spaced membership functions of the universe of discourse. Computer analysis is aimed at verifying the usefulness of the k-means method to obtain segregation of controller input data before entering the rules into the database of the smart system.

Figure 3a shows two input signals V_{in1}, V_{in2} of the ADC_2 converter, obtained by computer simulation, which are the product of the input value V_{in} (experimental linear signal with a slope of 50°) with the gain generated by the smart corrector. The quantity V_{in1} is a response to the gain of the fuzzy controller whose parameters, such as spacing of function cores, the associated input plane X, selection of the membership function type, implication and t-norm operators, were not properly selected and tuned. The signal V_{in2} was generated based on the expert's additional knowledge of the most common data clusters in the input space of the fuzzy controller.
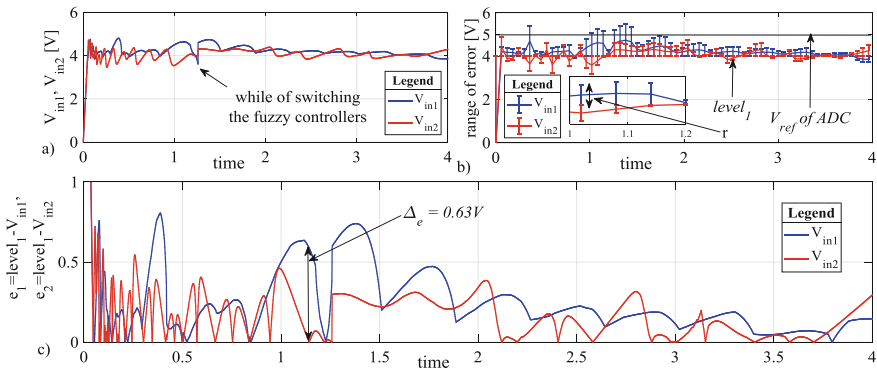
Fig. 3. Signal waveforms: (a) input V_{in1}, V_{in2} of the ADC_2 converter, (b) error range, (c) the difference between the signal and the full matching – output $level_1$ and the multiplier

The waveform V_{in2} (Fig. 2), generated by the classically designed fuzzy controller with tuned parameters and excellent knowledge, is qualitatively better than the waveform V_{in1}. Properly selected membership functions provided the elimination of the peak while switching between the adjacent controllers operating in parallel structure – see Fig. 3a. Figure 3b shows the same waveforms as in Fig. 3a with marked vertical ranges r whose length is equal to the difference between the desired input value of the ADC_2 converter and the value generated by the smart gain corrector (Fig. 3c).

In the computer simulation program, a set of linear signals of different slopes (from 0° to 89° at a step of 5°) was applied to the input terminals of the classically designed smart conversion corrector with tuned parameters. The tests indicate that the corrector with the tuned parameters made it possible to reduce the difference between the $level_1$ = 4 V and the input signal (V_{in2}) of the ADC_2 converter by approaching the level1 by the maximum value of Δ_e = 0.63 V in relation to the output signal (V_{in1}) of the classically designed corrector (increasing the use of the dynamic range of ADC_2 by 15.75%).

The actual output signal from the traction current transducer CT [7] with a maximum raising rate of 420 V/μs was applied to input terminals of the gain corrector in Fig. 1. Reducing the required voltage input level to $level_2$ for signals featuring such high dynamics protects against the loss of the information measured by the measuring system. Figure 4 shows two waveforms (V_{in3} and V_{in4}) being the product of the response of the fuzzy controller with classically selected and tuned parameters (based on the results obtained from the k-means algorithm) and the voltage signal across the resistor R_M of the CT transducer as generated by the actual current of the traction contact line under load.

In Fig. 4, the horizontal line marked with the variable $level_2$ means the limit value
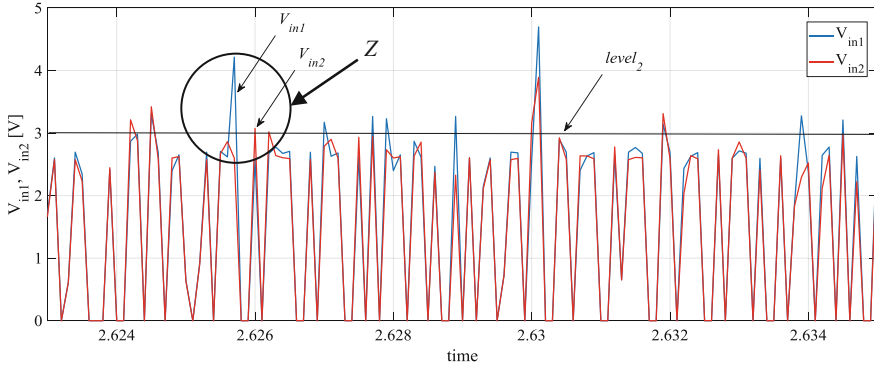
Fig. 4. Voltage waveforms: V_{in3}, V_{in4} of input ADC_2 transducer of the current ch1 channel (Fig. 1) of the electric energy meter, resulting from the product with the conversion factor of the fuzzy controller featuring the input/output plane designed classically and tuned based on the results of the k-means algorithm.

for dynamic waveforms, enabling the measurement without any loss of the measured information (signal value higher than the reference voltage of the ADC_2 converter). The correctness of the appropriate spacing of support points (nodes) of the input/output plane of the fuzzy controller operating in the gain corrector is proved by the waveform V_{in2}, located below the value $level_2$ in Fig. 4.

The quality of the smart amplification corrector with the tuned parameters, as shown in Fig. 1, is much better than that of the classically designed corrector system, with respect to the dynamically changing signal at the input terminals. For this signal, the computer program was used for numerical determination of the maximum difference between the required $level_2 = 3$ V, which was equal to approx. 1.2 V for the classically designed controller (Fig. 4, marked with the z symbol), while the perfect knowledge corrector reduced that value to the level of 0.2 V, within the same area of operation.

3   Summary

This paper presents a block diagram and the operating principle of the conversion factor corrector system [11] associated with the electric energy meter's current channel. The proposed system is suitable for continuous generation of a value representing the desired gain of the input signal of the energy meter's current channel. To design the fuzzy controller, intuitive rules of conduct formed by the expert were used. The experienced operator's excellent knowledge acquired by means of an analysis of input data clusters provides significant support in improving the quality of output signal generation by the smart corrector. The method selected to group the input data of the k-means fuzzy controller allowed the appropriate setting of the membership function cores along the axis of the input plane. The completed computer testing of the rule-

based gain control algorithm with appropriately shifted cores of the membership functions have shown effectiveness with respect to every input signal.

The application of the Gaussian rule-based gain control function allows significant limitation of the peak of the transducer input voltage (Fig. 3a) of the electric energy meters' current channel at the instant of switching between the controllers operating in parallel structure.

References

1. Norma PN-EN 50463-2: Zastosowania kolejowe – Pomiar energii na pokładzie pociągu. Część 2 – Pomiar energii. [In Polish: Railway applications - Energy measurement on the train board. Part 2 - Energy measurement] (2013)
2. Biliński, J., Buta, S., Gmurczyk, E., Kaska, J.: Nowoczesny asynchroniczny napęd z hamowaniem odzyskowym produkcji MEDCOM do zmodernizowanych elektrycznych zespołów trakcyjnych serii EN57AKŁ. Technika Transportu Szynowego 12/2004, pp. 20–25 (2004). [In Polish: Modern asynchronous electric drive with regenerative braking production by MEDCOM to modernized electric multiple units series EN57AKŁ]
3. Bolikowski, J.: Podstawy projektowania inteligentnych przetworników pomiarowych wielkości, Monografie nr 68, Zielona Góra (1993)
4. Piegat, A.: Fuzzy Modeling and Control. Springer, Heidelberg (2001)
5. Driankov, D., Hellendoorn, H., Reinfrank, M.: An Introduction to Fuzzy Control. Springer, Heidelberg (1993)
6. Wu, J.: Advances in K-means Clustering A Data Mining Thinking. Springer, Heidelberg (2012)
7. Dominikowski, B., Pacholski, K.: Korekcja właściwości metrologicznych przetworników szybkozmiennych sygnałów elektronicznych. Zeszyty Naukowe Wydziału Elektrotechniki i Automatyki Politechniki Gdańskiej 49, 21–24 (2016). [In Polish: Correction of metrological properties for fast changing electronic signal of transducer]
8. Babuška, R.: Fuzzy Modeling for Control. Springer, Heidelberg (1998)
9. Peña-Reyes, C.A.: Coevolutionary Fuzzy Modeling. Springer, Heidelberg (2004)
10. Zhang, H., Liu, D.: Fuzzy Modeling and Fuzzy Control. Birkhäuser, USA (2006)
11. Alegre Pérez, J.P., Celma, S., López, B.C.: Automatic Gain Control Techniques and Architectures for RF Receivers. Springer, Heidelberg (2011)

Investigation of the Instrumental Components in Uncertainty of Extreme Random Observations

Mykhaylo Dorozhovets[1(✉)], Ivanna Bubela[2,3], and Anna Szlachta[1]

[1] Rzeszow University of Technology, Rzeszow, Poland
{michdor,annasz}@prz.edu.pl
[2] National University - Lviv Politechnic, Lviv, Ukraine
popovych@ukr.net
[3] State Enterprise "Scientific-Research Institute of Metrology of Measurement and Control Systems" (DP NDI "Systema"), Lviv, Ukraine

Abstract. In this paper the instrumental components in the uncertainty of extreme observations are analyzed and quantitatively evaluated. In practice this method can be used to evaluate the uncertainty results of testing products, when during testing the most informative parameter is not the arithmetic mean but the extreme (minimal or maximal) observation. In the paper two main components of the uncertainty for such testing are studied: the statistical component - the variation of the measured parameter of a few tested specimens and the instrumental component - uncertainty of the measurement result of the appropriate parameter for each tested specimen. It is shown that the uncertainty of extreme observations depends in different ways on systematic and random effects in the measurements. If the standard uncertainty evaluated using the type B method (instrumental components) does not exceed (approximately) 1/3 of the standard uncertainty determined using the type A method (deviation values of observations), then the value of a coefficient which is used to calculate one-side expanded uncertainty of extreme observation can be determined approximately using a simplified method based on the ratio of both components of the standard uncertainty. The results of the research can be used to evaluate the uncertainty results in the quality testing of a wide variety of products in industry, agriculture and medicine when the result of the test depends on the minimum or maximum value of the parameter in the tested specimens.

Keywords: Measurement · Random · Systematic effects · Uncertainty
Extreme (minimal, maximal) observation · Monte carlo method

1 Introduction

In all industries it is very important to perform the continuous testing of the quality of the process of making of the wares or products. Such testing is carried in accordance with the requirements of regulatory documentations, which are established by the legislation and which take into account the specifics of product. Usually only limited number of the specimens (typically $n = 5 .. 10$) of the product which are selected randomly are tested

[1, 2]. In particular this applies to testing in which the specimens are destroyed. The testing procedures are based on the statistical processing results of the measurement of of the appropriate parameter of these specimens [1, 2]. Therefore the quality of the measurements may directly affect the result of testing the product quality.

In practice of quality testing there are cases when the extreme observation: minimal value $x_{min} = x_1$ - first after sorting or maximal value $x_{max} = x_n$ - last after sorting (Fig. 1) is most important from the point of view of product quality. For example in testing the mechanical properties of some products such as the elongation and tensile strength of wire rope or another metal and plastic product etc. only the minimal observation (most unreliable) from a few tested specimens is taken into account [3]. When testing medications, foods, etc. for which the maximal content of harmful contaminations is most important, only the observation with a maximal value of a testing parameter should be compared with the critical limit value. The result of such testing is obtained by comparing the appropriate extreme observation with an allowable (critical) value (Fig. 1).
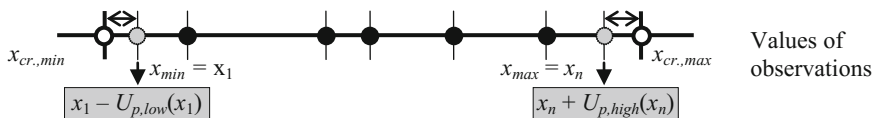
Fig. 1. Comparison of extreme observations with allowable (critical) values

The problem of testing quality based on the extreme observations in its essence is very close to the problem of testing outliers [4–7]. In both tasks more often the best estimate of location m (for example, arithmetic mean) and scale s (for example, experimental standard deviation) parameters of a sample of size n are determined and the value of the ratio $|x_* - m|/s$ (where x_* is suspected observation, minimal or maximal) is compared with a known critical value x_{cr}. However, between these issues there are significant differences, among others in the following.

(i) When we test the outliers, it doesn't matter what observation (minimal or maximal) will be detected as an outlier and will be rejected. Instead, depending on the content of the quality testing of a particular product only one side defines extreme observation: minimal (if the value of the tested parameter satisfies condition $x_{min} \geq x_{cr,min}$) or maximum (if the value of the tested parameter satisfies condition $x_{max} \leq x_{cr.max}$), but not both. For example, in the tests of mentioned above the mechanical properties of some products such as elongation and tensile strength, metal or plastic products only the minimal observation is taken into account and the maximal observation in this case is not analyzed. When testing for the content of harmful contaminations in medications, foods, etc. the maximal observation is compared with the maximal critical value, the minimal observation in this case is not analyzed.

(ii) The difference in these problems lies also in their ultimate goal: in outliers testing the detected observation with a "gross error" is removed from the sample and further processing is carried out without this observation. Here critical

values which are used for comparison with suspected observations depend only on the distribution, number of registered observations and significant level. In other words the critical value has a statistical meaning.

When we test the quality of a particular product, it is clear that critical value depends on the qualitative requirements, but also often depends on the requirements of health and life security. When during testing the defective specimen is found often not only this specimen but the whole party of products may also be rejected.

(iii) Classical testing methods of outliers are based on statistical analysis of random observations and do not take into account the systematic effects in observations. Because testing the quality of the products is based on the results of measurement of appropriate parameter for the few specimens, therefore the instrumental components of uncertainty in the measurement results, which are caused by random and systematic effects, must be taken into account. Sometimes, the uncertainty in measurements is comparable to the effect of the variation of observations.

In this article the issues concerning the impact of random and systematic effects in measurement results onto correctness of the test result, which is based on an extreme observation, are analyzed. An important aspect of this issue is that when we compare the extreme value with a critical value we must take into account the uncertainty of measurement result of extreme observation:

$$x_{\min} - U_p(x_{\min}) \geq x_{cr,\min} \text{ or } x_{\max} + U_p(x_{\min}) \leq x_{cr,\max} \tag{1}$$

where $U_p(x_{\min})$ and $U_p(x_{\max})$ are the expanded uncertainties of $x_{\min}$ and $x_{\max}$.

Because in such testing the key result is the extreme value of a few observations, therefore it is impossible to use directly the classic methodology [8] of Type A evaluation of uncertainty, which is applied to the arithmetic mean value but not to the extreme value.

We assume in the following that observations are obtained from the population with a symmetric distribution. Therefore, the standard uncertainty of the maximal observation will be the same as that for the minimal: $u(x_1) = u(x_n) = u(x_*)$, and also the left one-side expanded uncertainty $U_p(x_{\min})$ of minimal observation and right one-side of maximal observation $U_p(x_{\max})$ are equal: $U_p(x_{\min}) = U_p(x_{\max}) = U_p(x_*)$.

For the given distribution $p_x(x)$ it is easy to determine the distribution $p_{x_*}(x_*)$ of extreme observation x_* by well known dependence based on order statistics [9], here first $x_* = x_1$ or last $x_* = x_n$. In [10, 11] it was shown, that taking into account only the statistical dispersion of the observations (without instrumental component) expectation and standard deviation of extreme observation are equal to:

$$\mu_{x_*} = \mu_x + m_{0,1}\sigma_x; \quad \sigma_{x_*} = \sigma_{0,1} \cdot \sigma_x, \tag{2}$$

where μ_x and σ_x are the expectation and standard deviation of x; $m_{0,1}$ and $\sigma_{0,1}$ are the expectation and standard deviation of x_* when distribution of population $p_x(x)$ is normalized: expectation $\mu_x = 0$, standard deviation $\sigma_x = 1$.

Using estimated values of arithmetic mean $\bar{x}$ of μ_x and experimental standard deviation s_x we can evaluate the estimate m_{x_*} of x_* and its standard uncertainty [10]:

$$m_{x_*} \approx \bar{x} + m_{0,1} s_x; \quad u_A(x_*) \approx \sigma_{0,1} \cdot s_x, \tag{3}$$

However, because the parameters of population (μ_x, σ_x) are unknown, calculation of the expanded uncertainty $U_p(x_*)$ using distribution $p_{x_*}(x_*)$ and estimated values $\bar{x}$ and s_x for a small number of observations ($n \leq 5 .. 10$) may be very inaccurate. Also it is not possible to use asymptotic dependencies $p_{x_*}(x_*)$ for a large $n \to \infty$ [12], because the number of observations is usually limited.

For indirect evaluation of expanded uncertainty $U_p(x_*)$ the following ratio:

$$z_{x_*} = \frac{x_* - \bar{x}}{s_x} \tag{4}$$

can be used. Distribution $p_{x_*}(z_{x_*})$ does not depend on $\bar{x}$ and s_x but depends only on the distribution of population $p_x(x)$ and number of observations n.

The critical value $k_{z x_*}(p)$ of z_{x_*} for the one side confidence level p can be calculated from the solution of the following equation:

$$F_*(k_{z x_*}(p)) = 1 - p. \tag{5}$$

where $F_*(k) = \int_{-\infty}^{k} p_{x_*}(z) dz$ is a cumulative function.

From (4) and (5) using (3) we can determinate the approximate value of one-side expanded uncertainty $U_p(x_*)$ of x_*:

$$U_p(x_*) \approx (k_{z x_*}(p) + m_{0,1}) \cdot s_x = k_{x_*}(p) \cdot s_x, \tag{6}$$

where $k_{x_*}(p) = k_{z x_*}(p) + m_{0,1}$.

The range of random value z_* in (3) is independent of the type of $p_x(x)$ and equals $-(n-1)/\sqrt{n} \leq z_1 \leq -1/\sqrt{n}$ ((for $z_{x_*} = z_{x_n}$ this range is symmetrical) and consists of $n - 2$ parts [10]. For example, when $p_x(x)$ is normal and $n = 5$ then the distribution of $p_{x_1}(z_1)$ consists of the three parts:

$$p_{x_1}(z_1) = \frac{5\sqrt{5}}{2\pi} \begin{cases} \sqrt{1 - \left(\frac{\sqrt{5}}{4} z_1\right)^2}, & -\frac{2}{\sqrt{5}} \cdot 2 \leq z_1 - \leq -\frac{2}{\sqrt{5}} \cdot \sqrt{\frac{3}{2}}, \\ -\sqrt{1 - \left(\frac{\sqrt{5}}{4} z_1\right)^2} - \sqrt{\frac{5}{3}} \cdot \frac{\sqrt{5}}{4} z_1, & -\frac{2}{\sqrt{5}} \cdot \sqrt{\frac{3}{2}} \leq z_1 \leq -\frac{2}{\sqrt{5}} \cdot \sqrt{\frac{2}{3}}, \\ \sqrt{\frac{5}{3}} \cdot \frac{\sqrt{5}}{4} z_1 + \sqrt{1 - \left(\frac{\sqrt{5}}{4} z_1\right)^2} \cdot \left[\frac{1}{3} - \frac{6}{\pi} \phi_3(z_1)\right], & -\frac{2}{\sqrt{5}} \cdot \sqrt{\frac{2}{3}} \leq z_1 \leq -\frac{2}{\sqrt{5}} \cdot \frac{1}{2}, \\ 0, & \text{otherwise}. \end{cases} \tag{7}$$

where $\phi_3(z_1) = \int_{\arcsin\left(\sqrt{\frac{5}{3} \cdot \frac{\sqrt{5}}{4} z_1} \Big/ \sqrt{1 - \left(\frac{\sqrt{5}}{4} z_1\right)^2}\right)}^{\arcsin\left(\frac{1}{3}\right)} \arcsin\left(\sqrt{2} \, arctg(x)\right) \cos(x) dx.$

For a limited number of observations (for example $n = 5 .. 10$) the most important is the first part of this range within bounds [10]:

$$z_{b,1} = -(n-1)/\sqrt{n}, \quad z_{b,2} = -\sqrt{(n-1)(n-2)/2n}. \tag{8}$$

For normally distributed $n = 5$ observations the cumulative function is [10]:

$$F_{x_*}(z_{x_*}) = \frac{5}{\pi}\left[\frac{\sqrt{5}}{4}z_{x_*}\sqrt{1 - \left(\frac{\sqrt{5}}{4}z_{x_*}\right)^2} + \arcsin\left(\frac{\sqrt{5}}{4}z_{x_*}\right) + \frac{\pi}{2}\right]. \tag{9}$$

Using (9) in (4) gives the values of $k_{z_*}(p)$, which for $n = 5$ and $n = 10$ are shown in the first row of Table 1. To verify the correctness of the obtained results we compared the value of $k_{z_*}(p)$ (4) with the critical value $Gr(n, p)$ which is used to detect outliers for the normally distributed observations by the one sided Grubbs' test [4]:

$$Gr(n,p) = \frac{n-1}{\sqrt{n}} \cdot \sqrt{\frac{t^2_{\frac{1-p}{n},n-2}}{n - 1 + t^2_{\frac{1-p}{n},n-2}}}, \tag{10}$$

where $t_{\frac{\alpha}{2},n-2}$ is a square of the value of the Student's t - distribution with n - 2 degrees of freedom that corresponds to the $\alpha/2 = (1 - p)/2$ quantile.

In general, for an arbitrary distribution $p_x(x)$ determining distribution $p_{x_*}(z_{x_*})$ and also $k_{zx_*}(p)$ is quite a complex task. The values of $k_{zx_*}(p)$ and $k_{x_*}(p)$ for population distributions other then the normal, with different number of observations ($n = 5..10$) have been determined by the Monte-Carlo method [13].

Obviously, random Δr and systematic Δs effects in measurement results

$$y_i = x_i + \Delta r_i + \Delta s \tag{11}$$

cause different effects on the uncertainty of extreme observations. These effects deform the distribution $p_x(x)$ of the observations (in general $p(y) \neq p(x)$), as well as the distribution $p_*(z_*)$ of ratio (4) and, as a result, the values of $k_{zx_\bullet}(p)$ and $k_{x_\bullet}(p)$. Therefore, the impact of these effects will be studied separately. We assume that the component of uncertainty in measurements is evaluated by method A – standard uncertainty $u_A(x_*) = s_x$ and by the method B - combined standard uncertainty $u_{cB}(x_*)$.

1.1   Analysis of the Random Effects

Here it is assumed that measurement results y_i of all n investigated specimens with random values x_i have the independent random instrumental component Δr_i, i.e.:

$$y_i = x_i + \Delta r_i. \tag{12}$$

In (12) the minimal $y_* = y_1 = \min(x_i + \Delta r_i)$ and maximal $y_* = y_n = \max(x_i + \Delta r_i)$ values, the arithmetic mean $\bar{y} = \bar{x} + \bar{\Delta}r$ (where $\bar{\Delta}r$ is an arithmetic mean of Δr) and the

standard deviation $s_y = \sqrt{s_x^2 + s_r^2}$ (where s_r is a standard deviation of Δr) of observations depend on both random components. Therefore the ratio z_{y_*} is not equal to the ratio z_{x_*} (4):

$$z_{y_*} = \frac{y_* - \bar{y}}{s_y} \neq z_{x_*}. \tag{13}$$

Due to the mutual independency of x and Δr with distribution $p_r(\Delta r)$ the density of the distribution $p_y(y)$ of sum (12) is a convolution of both distributions: $p_y(y) = p_x(x) \otimes p_r(\Delta r)$. Distribution $p_y(y)$ at the first step is used for the determination of the distribution $p_{y_*}(z_{y_*})$ of ratio (13), which at the second step is used to determine the value of the values of $k_{zy_*}(p)$ from (5). The expanded uncertainty $U_p(y_*)$ of y_* is calculated from (6).

This problem is simplified when both distributions $p_x(x)$ and $p_r(\Delta r)$ are normal, then distribution $p_y(y)$ is also normal, like distribution $p_{y_*}(z_{y_*})$. Due to this the value of $k_{zy_*}(p)$ is the same as $k_{zx_*}(p)$. When $p_x(x)$ is normal and $p_r(\Delta r)$ is uniform, or vice versa, then the distribution $p_y(y)$ of sum (11) becomes a Flatten-Gaussian distribution, the properties of which are analyzed in detail in [13, 14]. But for the calculation of $k_{zy_*}(p)$ and $k_{y_*}(p)$ we need to determinate the distribution $p_{y_*}(z_{y_*})$ of the ratio (13). In practice the calculation of the distribution $p_{y_*}(z_{y_*})$ in such cases is very complicated.

To solve these problems and for test the quality of the results obtained we used the Monte Carlo (MC) method [15]. The number of MC trials is 10^5, expectation $m_x = 0$ standard deviation $\sigma_x = 1$; $p_x(x)$ and $p_r(\Delta r)$, normal and uniform in four combinations, ratios $\alpha_r = \sigma_r/\sigma_x : \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1$, one-sided confidence level $p = 0.90; 0.925; 0.95; 0.975; 0.99$.

The values of $k_{y_*}(p)$ for $n = 5$ and $n = 10$ are presented in Table 1. In this table we can see that when $n = 5$ for each value of confidence level p, for all values of ratio $\alpha_r = \sigma_r/\sigma_x$ and all combinations of distributions the differences between the $k_{y_*}(p)$ and $k_{y_*}(p)|_{\alpha_r=0}$ (first rows in Table 1) do not exceed approximately 0.7%.

When $n = 10$ the maximal difference does not exceed about 9%. When the ratio $\alpha_r = \sigma_r/\sigma_x$ does not exceed 1/3, which in practice is usually met, then the difference between the $k_{y_*}(p)$ and $k_{y_*}(p)|_{\alpha_r=0}$ (rows in Table 1 for $\alpha_r = 0$) does not exceed approximately 3%. Besides, when the convolution of $p_x(x)$ and $p_r(\Delta r)$ is not normal, the difference between the $k_{y_*,norm}(p)$ and $k_{y_*,norm}(p)$ (row in Table 1 for $\alpha_r = 0$ and normal $p_x(x)$) does not exceed about 1.2% when $n = 5$ and 9.5% when $n = 10$. It is well known [8] that when the number of the observation is limited ($n \leq 5 .. 10$) the value of an experimental standard deviation s_y and also of a Tape A standard uncertainty $u_A(y)$ has essential uncertainty: for $n = 5$ observations it is 36% and for $n = 10$ observations it is 24%. Therefore in the first approximation it is possible to use $k_{y_*}(p) \approx k_{y_*,norm}(p)$ from a normal distribution (10), because using this assumption caused uncertainty a few times smaller than the Type A uncertainty $u_A(y)$.

Table 1. The values of $k_{y_*}(p)$ as functions of $\alpha_r = \sigma_r/\sigma_x$, $n = 5$, $n = 10$.

p	0.90	0.925	0.95	0.975	0.99	0.90	0.925	0.95	0.975	0.99
α_r	$p_x(x)$- normal, $p_r(\Delta r)$- uniform, $n = 5$					$p_x(x)$- normal, $p_r(\Delta r)$- uniform, $n = 10$				
0	1.602	1.635	1.671	1.715	1.749	2.036	2.098	2.176	2.290	2.410
1/5	1.600	1.632	1.670	1.714	1.749	2.036	2.099	2.176	2.287	2.407
1/4	1.602	1.634	1.671	1.714	1.747	2.035	2.097	2.177	2.290	2.409
1/3	1.602	1.636	1.673	1.716	1.750	2.033	2.094	2.170	2.287	2.406
1/2	1.601	1.635	1.672	1.716	1.749	2.028	2.087	2.166	2.279	2.398
1/1	1.597	1.631	1.668	1.711	1.746	1.998	2.057	2.134	2.245	2.369
	$p_x(x)$-uniform, $p_r(\Delta r)$- uniform, $n = 5$					$p_x(x)$-uniform, $p_r(\Delta r)$- uniform, $n = 10$				
0	1.583	1.623	1.664	1.715	1.583	1.824	1.883	1.962	2.090	2.234
1/5	1.583	1.621	1.665	1.713	1.749	1.842	1.900	1.977	2.101	2.247
1/4	1.585	1.623	1.665	1.711	1.747	1.856	1.912	1.989	2.108	2.251
1/3	1.585	1.623	1.664	1.711	1.748	1.873	1.930	2.006	2.127	2.261
1/2	1.586	1.622	1.664	1.712	1.746	1.911	1.966	2.04	2.153	2.279
1/1	1.591	1.625	1.663	1.709	1.745	1.955	2.012	2.085	2.199	2.325
	$p_x(x)$ uniform, $p_r(\Delta r)$ - normal, $n = 5$					$p_x(x)$ uniform, $p_r(\Delta r)$ - normal, $n = 10$				
0	1.583	1.623	1.664	1.715	1.750	1.824	1.883	1.962	2.090	2.234
1/5	1.583	1.623	1.667	1.715	1.749	1.845	1.902	1.979	2.101	2.245
1/4	1.583	1.624	1.667	1.715	1.749	1.845	1.902	1.979	2.101	2.245
1/3	1.583	1.623	1.665	1.713	1.747	1.879	1.936	2.010	2.131	2.270
1/2	1.586	1.623	1.663	1.711	1.746	1.916	1.972	2.047	2.158	2.291
1/1	1.595	1.629	1.668	1.713	1.747	1.992	2.050	2.129	2.245	2.371

1.2 Analysis of the Systematic Effects

The systematic effect Δs with distribution $p_s(\Delta s)$ causes the identical (but unknown) offset in all observations:

$$y_i = x_i + \Delta s. \tag{14}$$

Therefore, the mean value also has the same offset $\bar{y} = \bar{x} + \Delta s$ but the standard deviation s_y remains without change: $s_y = s_x$. Therefore the value of the ratio z_{y_*} (4) does not change, i. e., is equal to the ratio z_{x_*}:

$$z_{y_*} = [y_* - \bar{y}]/s_y = [x_* + \Delta s - (\bar{x} + \Delta s)]/s_x = z_{x_*}. \tag{15}$$

As a result, the density of distribution of z_{y_*} is the same as distribution of z_{x_*}, i.e., $p_{y_*}(z_{y_*}) = p_{x_*}(z_{x_*})$. In any experiment after calculating the mean value $\bar{y} = \bar{x} + \Delta s$ and

standard deviation of observations $s_y = s_x$ the value of the extreme observation y_* may be presented as:

$$y_* = \bar{x} + z_{x_*} \cdot s_x + \Delta s = \bar{x} + \left(z_{x_*} + \Delta s / s_y\right) \cdot s_y = \bar{x} + z_{ys_*} \cdot s_y, \tag{16}$$

where $z_{ys_*} = z_{x_*} + \Delta s / s_y$.

Assuming mutual independence of z_{x_*} and Δs the distribution $p_{ys_*}\left(z_{ys_*}\right)$ is a convolution of $p_{x_*}\left(z_{x_*}\right)$ and $s_y \cdot p_s\left(\Delta s / s_y\right)$. For example, when distribution $p_s(\Delta s)$ is uniform (range $\pm\Delta_{\lim}$, standard deviation $\sigma_{\Delta s} = \Delta_{\lim} / \sqrt{3}$) then for distribution $p_{x_*}\left(z_{x_*}\right)$ (7) and cumulative function (8) the convolution $p_{xs_*}\left(z_{xs_*}\right) = p_{x_*}\left(z_{x_*}\right) \otimes \left(s_x \cdot p_s\left(\Delta s / s_y\right)\right)$ is given as:

$$p_{ys_*}\left(z_{ys_*}\right) = \frac{s_y}{2\sqrt{3}\sigma_{\Delta s}} F_{x_*}\left(z_{y_*} + \frac{\sqrt{3}\sigma_{\Delta s}}{s_y}\right), \ p_{ys_*}\left(z_{ys_*}\right) = \frac{s_y}{2\sqrt{3}\sigma_{\Delta s}} F_{x_*}\left(z_{y_*} + \frac{\sqrt{3}\sigma_{\Delta s}}{s_y}\right)$$
$$= z_{y_{\min}}. \tag{17}$$

The critical value $k_{zys_*}(p)$, which reflects the systematic effects, may be determined after substitution (17) in (5) as:

$$\frac{s_y}{2\sqrt{3}\sigma_{\Delta s}} \cdot \int\limits_{z_{y_{\min}}}^{k_{zys_*}(p)} F_{x_*}\left(z_{y_*} + \frac{\sqrt{3}\sigma_{\Delta s}}{s_y}\right) dz_{xs_*} = 1 - p. \tag{18}$$

Table 2 presents the values of $k_{zys_*}(p)$ when $p_s(\Delta s)$ is uniform and $p_x(x)$ is normal ($p_{x_*}\left(z_{x_*}\right)$ and α_s : $1/5$, $1/4$, $1/3$, $1/2$, 1. By calculation of discrete convolution the critical values $k_{zys_*}(p)$ were determined when $p_s(\Delta s)$ is normal (Table 2). Because the value of $k_{zys_*}(p)$ depends on the ratio $\alpha_s = \sigma_{\Delta s} / s_y = u_{cB}(y) / u_A(y)$ for arbitrary combinations of distributions $p_x(x)$ and $p_s(\Delta s)$ analytical calculation of the critical value $k_{zys_*}(p)$ is very difficult or often even impossible. If the systematic effect is not significant: $\sigma_{\Delta s}/s_y << 1$, then $k_{zys_*}(p) \rightarrow k_{zx_*}(p)$, where values of $k_{zx_*}(p)$ are determined when $\alpha_s = 0$. Values of $k_{zx_*}(p)$ are shown in the first row of Table 2.

The value of $k_{zys_*}(p)$ may be expressed by the known value of $k_{zx_*}(p)$ as:

$$k_{zys_*}(p) = k_{zx_*}(p) \cdot \sqrt{1 + \left(\beta_p \frac{\sigma_s}{\sigma_y}\right)^2} = k_{zx_*}(p) \cdot \sqrt{1 + \left(\beta_p \frac{u_{cB}(y)}{u_A(y)}\right)^2}, \tag{19}$$

where

$$\beta_p = \frac{u_A(y)}{u_{cB}(y)} \sqrt{\left(\frac{k_{zys_*}(p)}{k_{zx_*}(p)}\right)^2 - 1} = \frac{1}{\alpha_s} \sqrt{\left(\frac{k_{zys_*}(p)}{k_{zx_*}(p)}\right)^2 - 1} \tag{20}$$

Table 2. The values of $k_{zys_*}(p)$, $n = 5$.

p	0.90	0.925	0.95	0.975	0.99
$\alpha_s = 0$	-1.602	-1.635	-1.674	-1.717	-1.748
α_s	$p_s(\Delta s)$- uniform				
1/5	-1.667	-1.727	-1.789	-1.874	-1.954
1/4	-1.721	-1.776	-1.842	-1.936	-2.024
1/3	-1.803	-1.865	-1.941	-2.045	-2.144
1/2	-1.990	-2.064	-2.154	-2.276	-2.393
1/1	-2.634	-2.737	-2.856	-3.020	-3.175
α_s	$p_s(\Delta s)$- normal				
1/5	-1.669	-1.718	-1.779	-1.870	-1.969
1/4	-1.708	-1.763	-1.833	-1.938	-2.055
1/3	-1.784	-1.850	-1.935	-2.063	-2.209
1/2	-1.961	-2.050	-2.165	-2.341	-2.544
1/1	-2.561	-2.724	-2.936	-3.262	-3.639

is the factor by which the impact of the systematic effect on the combined standard uncertainty of extreme observation y_* is taken into account:

$$u_c(y_*) = \sqrt{u_A^2(y) + (\beta_p u_{cB}(y))^2} = u_A(y)\sqrt{1 + \left(\beta_p \frac{u_{cB}(y)}{u_A(y)}\right)^2}. \tag{21}$$

In many practical measurement situations the knowledge about the distribution $p_s(\Delta s)$ of the systematic effect is limited, therefore as uncertainty $u_A(y)$ Type B evaluation of standard uncertainty $u_{cB}(y)$ is also not precise. Thus the value of β_p in (20) depends on the ratio $k_{zys_*}(p)/k_{zx_*}(p)$ but $k_{zys_*}(p)$ in turn depends on the ratio of the standard uncertainties $u_{cB}(y)/u_A(y)$ and also distributions $p_x(x)$ and $p_s(\Delta s)$. For these reasons in general the value of β_p in (20) may be determined only approximately. As a result of this the one-side limit $y_{*,\lim}(p)$ of extreme y_* in (16) may also be determined only approximately:

$$y_{*,\lim}(p) \approx \bar{y} + k_{zx_*}(p) \cdot u_c(y_*)\cdot, \tag{22}$$

The approximated values of β_p for the four combinations of the uniform and normal distributions $p_x(x)$ and $p_s(\Delta s)$ when $\alpha_s : 1/5, 1/4, 1/3, 1/2, 1$, $n = 5$, and $n = 10$ are determined by the Monte Carlo method (number of trials 10^5) and are presented in Table 3. In practice the value of the component of standard uncertainty $u_{cB}(y)$ evaluated by method type B usually satisfies the condition: $u_{cB}(y)/u_A(y) \approx \le 1/3$. From Table 3 we can see that under such conditions, the changes of the values β_p relative to the central values β_p when $\alpha_s = 1/4$ are: from 4.2% to 8.6% when $n = 5$ and from 2.8% to 5% when $n = 10$. These differences are a few times less than uncertainty of

Type A uncertainty $u_A(y)$. Therefore it is possible to use the interpolated value of β_p, when the value of ratio $u_{cB}(y)/u_A(y) = a_s$ differs from the values presented in Table 3.

The algorithm of determination of the one-side limit $y_{*,\lim}(p)$ of extreme y_* is described below. At first after estimation of the standard uncertainties $u_A(y)$ and $u_{cB}(y)$ is determined the ratio $u_{cB}(y)/u_A(y) = a_s$. During the second step this value is used to choose (in Table 3) the appropriate value β_p. On the third step the one-side limit $y_{*,\lim}(p)$ of extreme y_* is determined by (22) using combined uncertainty $u_c(y_*) = \sqrt{u_A^2(y) + (\beta_p u_{cB}(y))^2}$. One-side expanded uncertainty $U_p(y_*)$ is determined from (6).

Table 3. The values of β_p, $n = 5$, $n = 10$.

p	0.90	0.925	0.95	0.975	0.99	0.90	0.925	0.95	0.975	0.99
α_s	$p_x(x)$ - normal, $p_s(\Delta s)$- uniform, $n = 5$					$p_x(x)$ - uniform, $p_s(\Delta s)$- normal, $n = 5$				
1/5	1.548	1.701	1.903	2.202	2.491	1.390	1.517	1.723	2.062	2.502
1/4	1.570	1.697	1.849	2.095	2.330	1.429	1.540	1.718	2.014	2.391
1/3	1.549	1.646	1.769	1.949	2.124	1.440	1.539	1.687	1.939	2.254
1/2	1.474	1.541	1.624	1.745	1.868	1.399	1.489	1.614	1.816	2.068
1	1.305	1.342	1.385	1.449	1.515	1.248	1.327	1.432	1.599	1.804
	$p_x(x)$ - normal, $p_s(\Delta s)$- normal, $n = 5$					$p_x(x)$ - uniform, $p_s(\Delta s)$- uniform, $n = 5$				
1/5	1.461	1.613	1.817	2.173	2.586	1.447	1.590	1.804	2.115	2.444
1/4	1.479	1.617	1.797	2.105	2.467	1.488	1.600	1.772	2.024	2.287
1/3	1.470	1.588	1.748	2.006	2.314	1.496	1.579	1.705	1.891	2.09
1/2	1.412	1.513	1.645	1.858	2.113	1.452	1.505	1.583	1.706	1.841
1/1	1.247	1.333	1.443	1.618	1.825	1.305	1.333	1.370	1.428	1.498
p	0.90	0.925	0.95	0.975	0.99	0.90	0.925	0.95	0.975	0.99
α_s	$p_x(x)$ - normal, $p_s(\Delta s)$- uniform, $n = 10$					$p_x(x)$ - uniform, $p_s(\Delta s)$- normal, $n = 10$				
1/5	1.135	1.182	1.239	1.407	1.527	1.458	1.481	1.496	1.525	1.534
1/4	1.173	1.216	1.268	1.399	1.505	1.435	1.462	1.491	1.527	1.544
1/3	1.187	1.226	1.272	1.365	1.444	1.400	1.433	1.470	1.449	1.546
1/2	1.172	1.201	1.229	1.285	1.334	1.322	1.367	1.414	1.472	1.524
1/1	1.067	1.081	1.094	1.113	1.127	1.150	1.202	1.262	1.340	1.412
	$p_x(x)$ - normal, $p_s(\Delta s)$- normal, $n = 10$					$p_x(x)$ - uniform, $p_s(\Delta s)$- uniform, $n = 10$				
1/5	1.124	1.182	1.248	1.423	1.569	1.458	1.472	1.478	1.483	1.479
1/4	1.159	1.209	1.274	1.420	1.467	1.448	1.456	1.467	1.472	1.473
1/3	1.171	1.219	1.282	1.406	1.531	1.411	1.426	1.431	1.436	1.432
1/2	1.145	1.195	1.256	1.362	1.475	1.348	1.357	1.362	1.358	1.344
1/1	1.029	1.081	1.143	1.240	1.342	1.195	1.202	1.200	1.187	1.163

Because, as was mentioned above from (19) and (20) the values of the β_p and $k_{zys_*}(p)$ may be determined only approximately, the quality of the proposed approximation was tested by the Monte-Carlo method (10^5 trials). The number of observations: $n = 5$ and $n = 10$. Four combinations of normal and uniform distributions $p_x(x)$ and $p_s(\Delta s)$ of both components were investigated. In each investigation the values of β_p were taken from Table 3 for five values of confidence level p: 0.90, 0.925, 0.95, 0.975, 0.99 and five values of ratio α_s : $^1/_5$, $^1/_4$, $^1/_3$, $^1/_2$, 1. The random observations were generated with standard deviations s_y and $\sigma_{\Delta s}$ of both components which corresponded to the value of β_p. The critical values $k_{zys_*}(p)$ were determined from (19) using corresponding values of β_p and α_s.

Table 4. The values of p_{MC}

p	0.90	0.925	0.95	0.975	0.99	0.90	0.925	0.95	0.975	0.99
α_s	$p_x(x)$ - normal, $p_s(\Delta s)$- uniform, $n = 5$					$p_x(x)$ - uniform, $p_s(\Delta s)$- normal, $n = 5$				
1/5	0.902	0.929	0.956	0.980	0.992	0.901	0.926	0.951	0.977	0.991
1/4	0.903	0.930	0.955	0.979	0.992	0.903	0.928	0.953	0.977	0.992
1/3	0.908	0.934	0.959	0.981	0.992	0.905	0.929	0.953	0.977	0.991
1/2	0.911	0.935	0.959	0.980	0.991	0.907	0.931	0.954	0.977	0.991
1	0.914	0.938	0.959	0.979	0.990	0.905	0.929	0.953	0.977	0.990
	$p_x(x)$ - normal, $p_s(\Delta s)$- normal, $n = 5$					$p_x(x)$ - uniform, $p_s(\Delta s)$- uniform, $n = 5$				
1/5	0.90	0.926	0.952	0.977	0.992	0.900	0.927	0.952	0.978	0.992
1/4	0.903	0.928	0.954	0.978	0.991	0.902	0.929	0.955	0.980	0.992
1/3	0.905	0.93	0.954	0.978	0.991	0.904	0.930	0.956	0.980	0.992
1/2	0.907	0.932	0.954	0.977	0.991	0.909	0.933	0.957	0.979	0.991
1	0.908	0.931	0.954	0.977	0.990	0.909	0.934	0.957	0.980	0.991
p	0.90	0.925	0.95	0.975	0.99	0.90	0.925	0.95	0.975	0.99
α_s	$p_x(x)$ - normal, $p_s(\Delta s)$- uniform, $n = 10$					$p_x(x)$ - uniform, $p_s(\Delta s)$- normal, $n = 10$				
1/5	0.900	0.925	0.951	0.976	0.991	0.902	0.926	0.951	0.977	0.991
1/4	0.901	0.925	0.950	0.976	0.990	0.904	0.929	0.953	0.977	0.990
1/3	0.902	0.927	0.951	0.976	0.991	0.906	0.930	0.954	0.974	0.991
1/2	0.905	0.930	0.954	0.978	0.991	0.908	0.932	0.955	0.978	0.991
1	0.906	0.931	0.955	0.978	0.991	0.904	0.929	0.953	0.977	0.991
	$p_x(x)$ - normal, $p_s(\Delta s)$- normal, $n = 10$					$p_x(x)$ - uniform, $p_s(\Delta s)$- uniform, $n = 10$				
1/5	0.900	0.924	0.950	0.976	0.990	0.902	0.926	0.951	0.976	0.990
1/4	0.901	0.926	0.951	0.976	0.990	0.903	0.927	0.952	0.976	0.990
1/3	0.900	0.925	0.951	0.976	0.991	0.903	0.927	0.952	0.976	0.991
1/2	0.901	0.926	0.950	0.975	0.990	0.905	0.930	0.954	0.979	0.992
1	0.904	0.928	0.952	0.976	0.991	0.907	0.931	0.955	0.979	0.993

During investigations the estimates p_{MC} of confidence levels were determined when y_* achieved the critical limits $y_{*,\lim}(p)$ the values of which are calculated by (22) according to the known values of $k_{zys_*}(p)$. The estimated values p_{MC} are given in Table 4.

From Table 4 we can see that when the number of the observation is $n = 5$ estimated confidence levels p_{MC} differ from assumed values p less than 1.6% when $p_x(x)$ is normal and $p_s(\Delta s)$ is uniform. In others cases the differences between p_{MC} and p do not exceed 1%. Thus simulation test results confirm the effectiveness of the proposed method.

1.3   Conclusion

In cases when the results of testing product quality depend on the extreme (minimal or maximal) value of the parameter in the tested specimens, for uncertainty evaluation of the test result we must use another methodology than the classic methodology [8] of Type A evaluation of uncertainty, which is applied to the uncertainty of the arithmetic mean.

Because the uncertainty of testing result depend not only on the dispersion of the measured parameter of a few specimens but also depends on the random and systematic effects in measured results, therefore these effects should be taken into account. Random and systematic effects in measured results affect the uncertainty of the extreme values differently. Random effects alter the distribution of the measured observations, and systematic effects alter the distribution of the ratio $z_{x_*} = (x_* - \bar{x}) / s_x$. Therefore different analyses of random and systematic effects are used in this article.

The key parameter determining the critical values for the extreme values is $k_{z_*}(p)$, whose value depends on the combination of distributions $p_x(x)$, $p_r(r)$ and $p_s(\Delta s)$ and the values of the standard deviation of the tested parameter (σ_x) and measurement results ($\sigma_r, \sigma_{\Delta s}$).

The exact analytical solution for the evaluation of the uncertainty of extreme values by the determination of $k_{z_*}(p)$ is usually possible only in limited cases, such as when measured observations have a normal distribution. In other cases, the Monte-Carlo method should be used. The influence of different combinations of distributions (normal - uniform) of tested parameters and random and systematic components and also the influence of the values of standard deviation of both components on the uncertainty of extreme value were investigated.

The research results have shown that if the random and systematic components in measured results match to conditions: $\alpha_r = \sigma_r/\sigma_x \approx \leq 1/3$, $\alpha_s = \sigma_{\Delta s}/\sigma_x \approx \leq 1/3$ then the approximate values of $k_{z_*}(p)$ may be determined from a normal distribution (10) ($k_{y_*}(p) \approx k_{y_*,norm}(p)$). As a result of this approximation, an additional component of uncertainty occurs, but its value is at lest a few times less than the uncertainty of experimental standard deviation s_y and also Type A uncertainty $u_A(y)$ when the number of observations is limited ($n = 5..10$).

During practical tests it is usually possible to choose measurement instruments which provide the conditions $s_r/s_x \approx \leq 1/3$ and $u_{cB}(y) / u_A(y) \approx \leq 1/3$, therefore the approximation $k_{y_*}(p) \approx k_{y_*,norm}(p)$ does not cause a significant increase in the uncertainty of extreme observation.

References

1. D 638 Test Method for Tensile Properties of Plastics Annual Book of ASTM Standards, Vol 08.01
2. ASM International. Tensile Testing, Second Edition (2004)
3. GOST 11262-80, GOST 26277-84, GOST 12423-66. Ukraine standards of testing methods and conditions of plastic materials and products
4. Grubbs, F.S.: Sample criteria for testing outlying observation. Ann. Math. Statist., 27–58 (1950)
5. Dixon, W.J.: Analysis of extreme values. Ann. Math. Statist. **21**, 488–506 (1950)
6. Hawkins, D.: Identification of Outliers. Chapman and Hall, London (1980)
7. Burke, S.: Missing Values, Outliers, Robust Statistics & Non-parametric Methods. Statistics and data analysis. LC•GC Europe Online Supplement (2001)
8. Guide to the Expression of Uncertainty in Measurement. GUM. First ed. 1993 ISO Switzerland, last corrected ed. JCGM BIPM 100 (2008)
9. Fisz, M.: Probability Theory and Mathematical Statistics. Willey, London (1963)
10. Dorozhovets, M,. Popovych I., Warsza Z.L.: Method of evaluation the measurement uncertainty of the minimal value of observations and its application in testing of plastic products. Advanced Mechatronics Solutions. Advances in Intelligent Systems and Computing. Springer International Publishing Switzerland Vol. 393, 421–430 (2016)
11. Dorozhovets, M., Bubela, I.: Computing uncertainty of the extreme values in random samples. Int. J. Comput. **15**(2), 127–135 (2016)
12. Gumbel, E.: Statistics of Extremes, New York (1962)
13. Dietrich, C.F.: Uncertainty, Calibration and Probability. The Statistics of Scientific and Industrial Measurement. Second Edition. The Adam Hilger Series on Measurement Science and Technology, p. 535 (1991)
14. Dorozhovets, M., Popovych, I.: Processing of the random observations with Flatten-Gaussian distribution by approximate order statistics method. In: Proceedings of the 2015 IEEE 8th International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, IDAACS 2015. Vol. 1, pp. 149–152
15. JCGM 101: Evaluation of measurement data—Supplement 1 to the 'Guide to the Expression of Uncertainty in Measurement'—Propagation of distributions using a Monte-Carlo method. BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP and OIML (2008)

Eliminating the Inertial Forces Effects on the Measurement of Robot Interaction Force

Piotr Gierlak[✉], Andrzej Burghardt, Dariusz Szybicki, and Krzysztof Kurc

Faculty of Mechanical Engineering and Aeronautics,
Department of Applied Mechanics and Robotics,
Rzeszow University of Technology, 35-959 Rzeszow, Poland
{pgierlak,andrzejb,dszybicki,kkurc}@prz.edu.pl

Abstract. The paper presents the problem of eliminating the effects of inertial forces on the measurement circuit of the forces of interaction between a robotic manipulator and its environment. Inertial forces are among the most significant interferences in force measurement circuits. The causes and effects of this interference type are discussed. A concept is discussed, based on the measurement of acceleration of a manipulator end-effector. A test rig is discussed, as used for verification testing of a procedure for eliminating the effect of inertial forces on the measurement of interaction forces. Verification test results are shown which proved that the procedure is effective. The procedure developed in this work is highly significant for practical applications in robotized machining.

Keywords: Force measurement · Inertial force · Robotics

1 Introduction

Modern industrial robotic manipulators have an increasing number of applications in processes which require control over how robots interact with their environment. A robot's tool mounted on the end-effector interacts with the environment, and the type of tool depends on the process made with the robot. Industrial manufacturing has been increasingly reliant on robotized grinding and polishing, casting edge deburring and light milling [1–3]. The quality of a machined component defined by the component's dimensional accuracy and surface roughness largely depends on the tool feed rate, its rotational speed and pressure force. The tool feed rate and rotational speed are relatively easy to achieve and maintain. The feed rate depends on the motion of the robotic arm on which a tool is mounted and the rotational speed is controlled with a frequency converter. In practical applications, the main problem is to keep the pressure force value constant. This force is applied by controlling the robotic arm, and the interaction force feedback is provided by a force sensor installed between the robotic arm and the tool [4, 5].

This work was motivated by research for aviation industry into the robotization of machining by the application of hybrid position/force control strategies [6–8]. Figure 1 shows a diagram of a cast of aircraft engine casing, the so-called diffuser, machined on a robotic workstation as part of the research this paper concerns [9]. The issues encountered during the process performance prompted a search for methods capable of improving the feedback signals of the interaction forces between a cutting tool and the machined component.
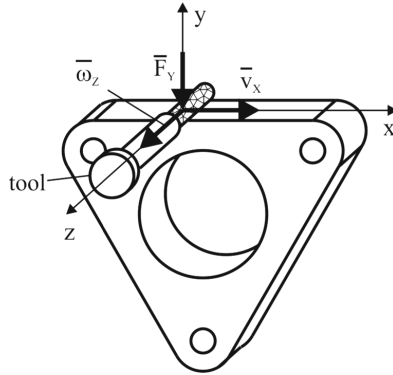
Fig. 1. Schematic diagram of a diffuser with a cutting tool and process parameters

Main issues tackled during the application of force feedback include [10,11]:

– the effect of tool weight on the measured interaction forces;
– the effect of gyroscopic phenomena occurring during the rotational motion of cutting tools;
– the effect of inertial forces on the interaction forces measured by the robot's sensor.

Elimination of the effects of tool weight on measured interaction forces is relatively straightforward. The tool weight can be measured by the force sensor and its effect can be eliminated. Commercial solutions of robotic workstations are provided with procedures for automatic elimination of tool weight effects on measured interaction forces. Furthermore, the effects of gyroscopic phenomena is negligible as rotating cutting tools are relatively light and the forces resulting from the phenomena are approximately three orders of magnitude lower than the forces required in manufacturing processes. However, the effect of inertial forces on measured interaction forces can result in serious issues with force feedback loops [12–15]. These issues must be clarified with mechanical engineering (Fig. 2).

Assume that tool tip (point A) moves along a path at a constant velocity $v(t)$ (Fig. 1). The tool tip moves along straight sections of the path with no acceleration, whereas when it moves along a curvilinear section of the path with a curve radius $\rho = |AK|$ it features normal (centripetal) acceleration $a_n(t_1)$ as a
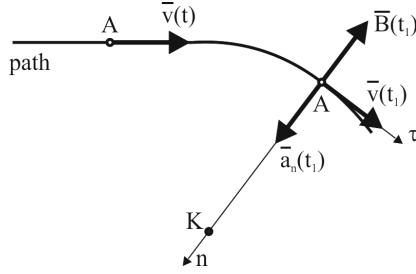
Fig. 2. Motion of the tool tip along a path

result of a change of velocity direction. Assume that the tool tip has a mass m. As such, the tip is subject to an inertial force $B(t_1) = ma_n(t_1)$ with a sense opposite to the sense of acceleration. Even though it is a fictitious force, it is detected by the force sensor. If the tool tip moves along said path without any interaction with the environment, the force sensor picks up the inertial force. This is similar to the condition under which the tool tip travels along the aforementioned path in interaction with the contact surface. Under this condition, the force picked up by the force sensor can be higher or lower than the actual interaction force, which depends on the sense of the fictitious force interfering with the force measurement. The force sensor can only record the resultant of both forces, and the control system 'understands' that the signal sent by the sensor represents the interaction force. In extreme cases, the inertial force exceeds the interaction force during the motion of the tool tip along a curvilinear section of the path at a high centripetal acceleration, and the control system responds to this feedback by moving the tool away from the surface in an attempt to reduce the force. This attempt is ineffective because the force sensor picks up the inertial force, not the actual interaction force. This destabilizes the system and results in a loss of contact between the tool and the machined surface.

2   Solution Concept

The aforementioned problem of the effect of inertial forces on the measurement of interaction forces can be eliminated if the inertial force value is known. To achieve such condition, the tool mass and the tool's acceleration must be known. While the tool mass can be determined easily, the determination of the actual acceleration of the robotic manipulator's end-effector is encumbered with numerous problems. Firstly, it is standard in robotics to measure the velocities of the links [16] and determine the positions of the links by integration. This solution is sufficient to achieve the feedback necessary to generate motion control. Hence, the necessity of knowing the acceleration values is an above-standard requirement for commercial solutions.

Several methods exist to facilitate the determination of acceleration of a robotic manipulator end-effector. One method assumes that the motion of a robot is in perfect conformity with the preset motion path and the velocity setting; hence, it is feasible to assume that the actual acceleration matches the acceleration generated by the preset trajectory of motion. Inherent control errors, however, make this assumption highly unrealistic and the approach can only be utilized for rough estimation of the effect of inertial forces. The second method is based on an approach by which velocity feedbacks are differentiated to derive acceleration values. This entails a problem of differentiating the feedbacks burdened with measurement noise, which results in very low quality of acceleration output. The third method is to equip the robotic manipulator end-effector with an acceleration sensor in order to determine acceleration by direct measurement. The third method was chosen for and presented in this paper because it offers the best acceleration feedback quality.

3   Analysis of Inertial Forces During Diffuser Machining

Figure 3a shows the geometric features of the diffuser [17], whereas Fig. 3b shows the points of measured interaction forces, the actual tool pressure forces and the inertial forces which occur during motion along the diffuser's curves. Figure 4 shows time trends related to the acceleration caused by the preset motion path (Fig. 4a), the force sensor output force value (Fig. 4b), the inertial force present during motion along the preset motion path (Fig. 4c) and the actual interaction force (Fig. 4d).
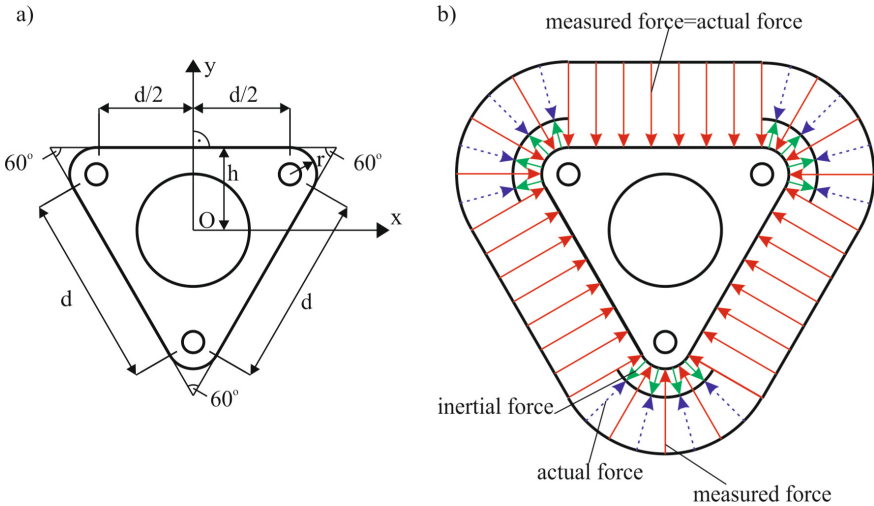
Fig. 3. Force analysis: (a) geometric features of the diffuser; (b) distribution of force along the diffuser's circumference
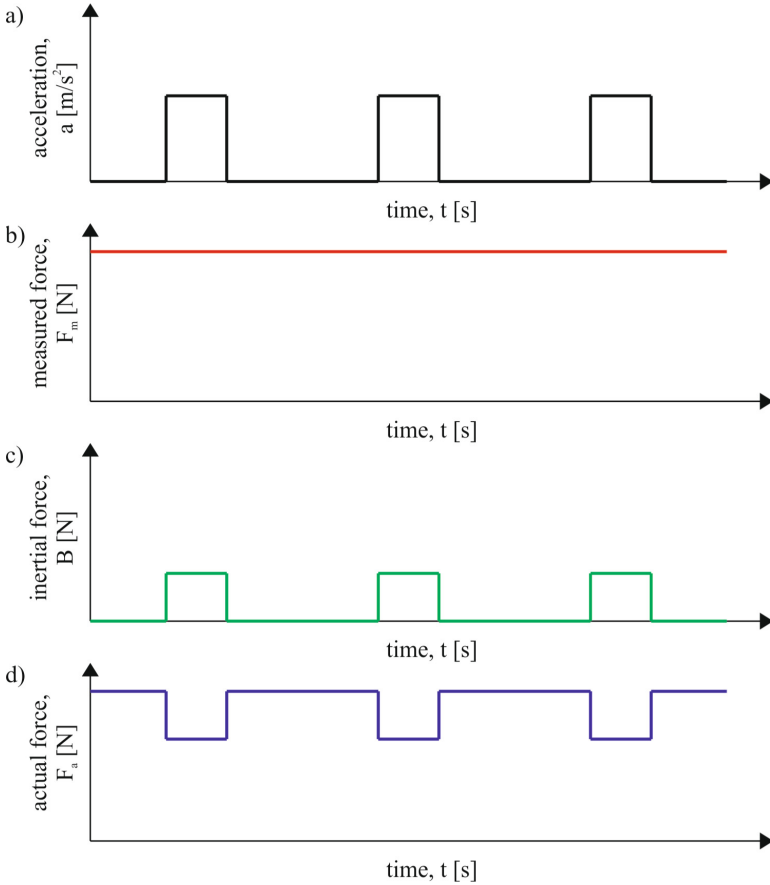
Fig. 4. Signals: (a) normal acceleration; (b) force sensor output force value; (c) inertial force; (d) interaction force

This analysis applies to an ideal case, that is the motion of the tool at a known acceleration value equal to the acceleration caused by strict conformity of the motion to the preset motion path. It is known that actual values of acceleration, interaction force and inertial force will feature slightly different trends.

4   Test Rig

The tests as part of the procedure for elimination of the effect of inertial forces on the measurement of the interaction forces were completed with a proprietary Cartesian robotic manipulator with three degrees of freedom [18]. The Cartesian robotic manipulator was designed for research e.g. into robotization of machining processes. It facilitates two independent horizontal translations and one vertical translation of the end-effector. The manipulator is provided with linear motion

modules, comprising of guides with trucks driven by geared DC motors with ball screws. To apply a position control loop, the Cartesian robotic manipulator was equipped with encoders installed on the DC motor drive shafts. To apply a force control loop required to control the forces of the robot's interaction with the environment, the end-effector was provided with FTD-Gamma SI-130-10 force sensor manufactured by ATI. The force sensor is capable of picking up three components of force and three components of torque. The force sensor was equipped with a tool chuck with integrated cutting tool. To measure acceleration, the end-effector was provided with a MEMS 633 6-DOF sensor manufactured by TE Connectivity. The test rig is shown in Figs. 5 and 6.

5   Procedure Verification

Four cases were tested to verify the procedure:

– Case 1: motion along a path conforming to the diffuser's shape with no contact of the cutting tool with the machined surface, with no inertial force effect adjustment;
– Case 2: motion along a path conforming to the diffuser's shape with no contact of the cutting tool with the machined surface, with inertial force effect adjustment;
– Case 3: motion along a path conforming to the diffuser's shape with the cutting tool pressed onto the machined surface, with no inertial force effect adjustment;
– Case 4: motion along a path conforming to the diffuser's shape with the cutting tool pressed onto the machined surface, with inertial force effect adjustment.

The test parameters are shown in Table 1 and Fig. 7 shows the results provided by testing the cases. The motion along the path curvatures occurred at the 2nd, the 5th and the 8th second, respectively. During the motion along the preset curvatures at the preset velocity, the normal acceleration a_n was approximately 0.1 m/s^2, which, given the tool mass is equal to 10 kg, gave an inertial force of approximately $B = ma_n = 1$ N. The inertial force was near zero during the motion along the preset straight sections of the path.

For Case 1, Fig. 7a shows that during the cutting tool's movement along a curve the force picked up by the force sensor increased despite the lack of contact between the cutting tool and the diffuser. This was caused by the effect of the inertial force. No effect of the inertial force is shown on the chart in Fig. 7b which shows the results for Case 2. The actual interaction force value was near zero, while its slight fluctuations were caused by measurement errors. The comparison between Case 1 and Case 2 demonstrated that the procedure's performance was correct. The results obtained in Case 3 and 4 were significant in practical terms since they exemplified the conditions of the robotic cutting tool in contact with the diffuser. Figure 7c shows that the measured force was nearly constant during the motion of the robot. However, it contained a given inertial force and the
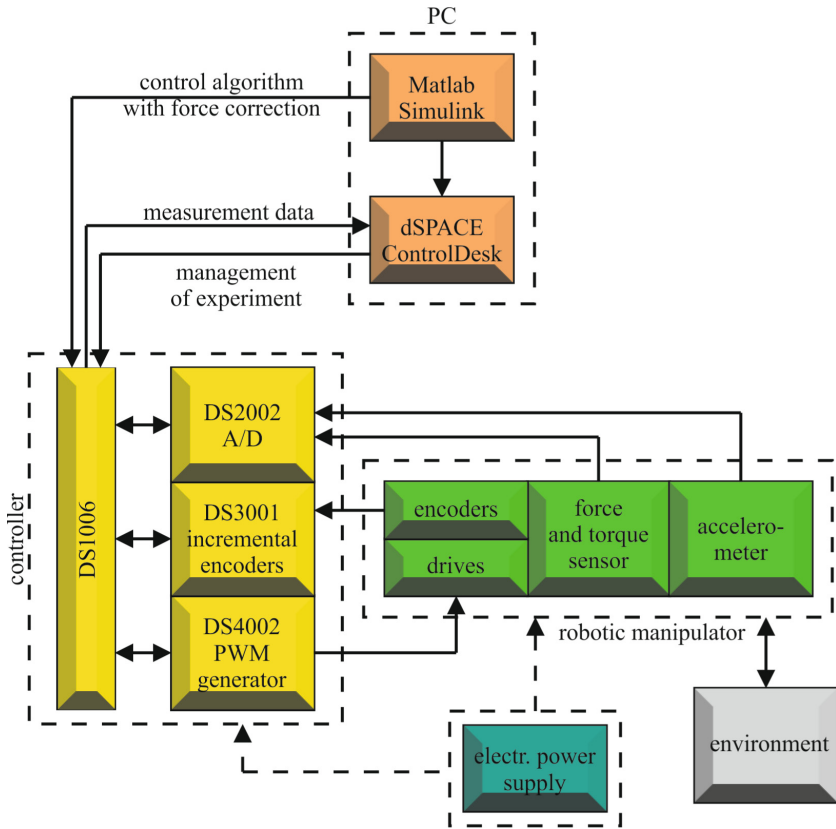
Fig. 5. Test rig diagram
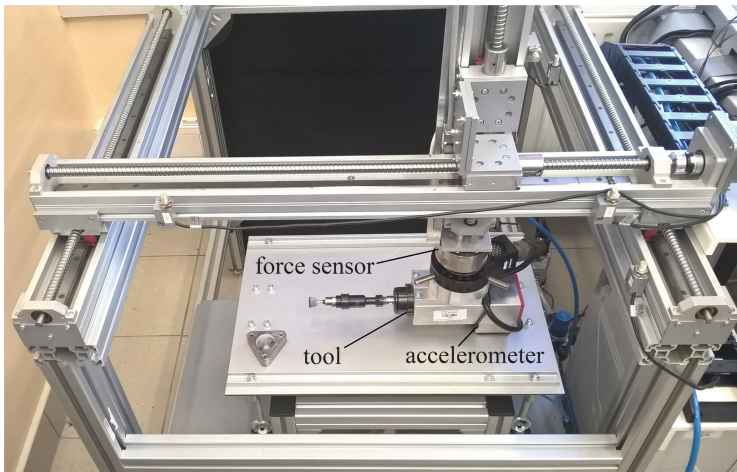
Fig. 6. Test rig overview

Table 1. Verification test parameters

Feed rate m/s	Tool mass kg	Diffuser curvature radius m	Desired pressure force N
0.0283	10	0.008	3
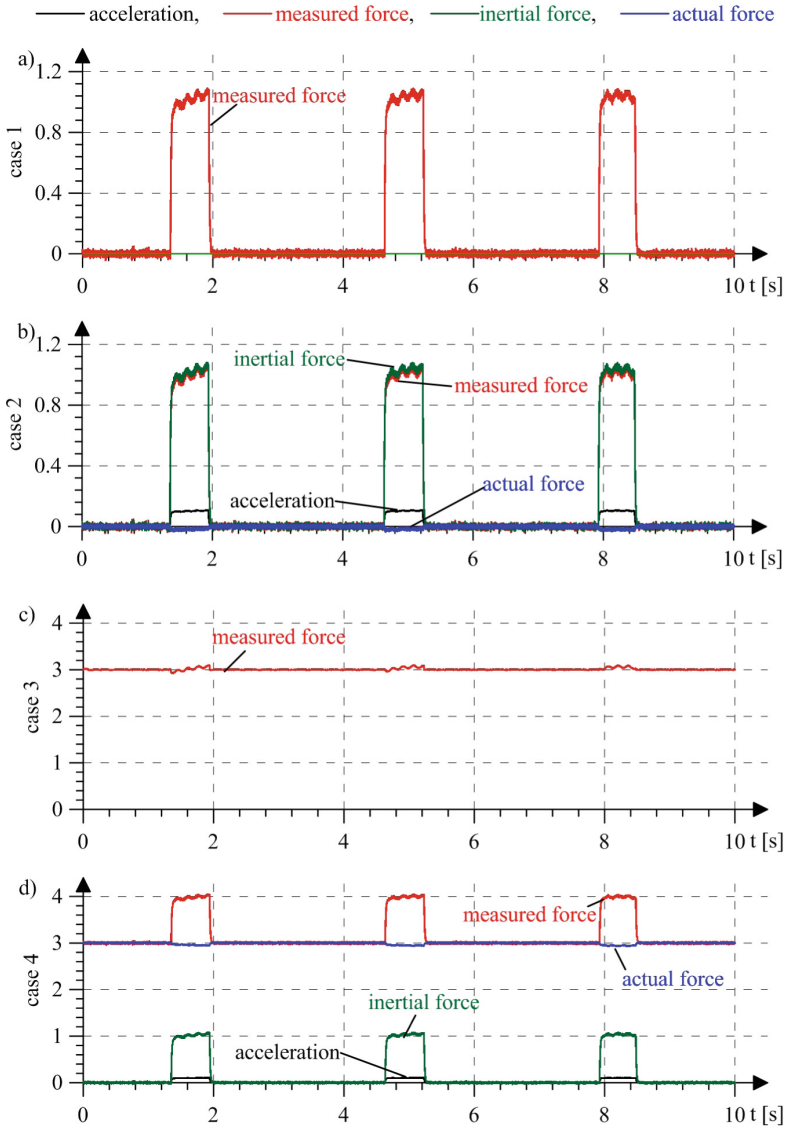
Fig. 7. Procedure verification results: (a) Case 1; (b) Case 2; (c) Case 3; (d) Case 4

actual interaction force was unknown. The robot's control system provided that the measured force was steady, but the interaction force was unstable. Figure 7d shows the results for Case 4. It is evident that the force measured by the force sensor periodically increased since the robot's control system was not stabilizing the measured force - it was stabilizing the actual interaction force as offset by the inertial force. Note that the actual interaction force in Cases 1 and 3 was unknown. The force picked up by the force sensor, the measurement output of which was interfered by the inertial force, was the only known value.

6   Conclusions

An analysis of the test results for the four cases allows a conclusion that the developed procedure for elimination of the effect of inertial force on the measured interaction forces performed well and improved the output quality in the force feedback loop.

Test parameters used during the verification tests represented actual conditions of industrial processing. Pressure force values amounting to several newtons and feed rate values amounting to several dozen mm/s are applied e.g. in grinding, polishing and deburring. The tool's mass is a total of the masses of the cutting tool, the electrospindle and the electrospindle's coupler on the robotic arm. It may vary from a few kilograms in small robots to dozens of kilograms in large, heavy-duty robotized machine tools. On the basis of the authors' research and the solutions applied it can be concluded that commercial solutions of robotic manipulator control are affected with certain defects, usually occurring at low interaction forces. This work is an attempt to eliminate one of the most significant interferences related to force feedback loops. The procedure developed herein is highly significant to practical applications of robotized machining.

References

1. Robin, V., Sabourin, L., Gogu, G.: Optimization of a robotized cell with redundant architecture. Robot. CIM-Int. Manuf. **27**(1), 13–21 (2011). https://doi.org/10.1016/j.rcim.2010.06.010
2. Barnfather, J.D., Goodfellow, M.J., Abram, T.: A performance evaluation methodology for robotic machine tools used in large volume manufacturing. Robot. CIM-Int. Manuf. **37**, 49–56 (2016). https://doi.org/10.1016/j.rcim.2015.06.002
3. Johansson, R., Nilsson, K., Robertsson, A.: Force control. In: Nee, A. (eds) Handbook of Manufacturing Engineering and Technology. Springer, London (2015). https://doi.org/10.1007/978-1-4471-4670-4_108
4. Stefanescu, D. M., Anghel, M. A.: Electrical methods for force measurement- a brief survey. Measurement **46**(2), 949–959 (2013). https://doi.org/10.1016/j.measurement.2012.10.020
5. Cutkosky, M.R., Provancher, W.: Force and tactile sensing. In: Siciliano B., Khatib O. (eds) Springer Handbook of Robotics. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-32552-1_28

6. Gierlak, P.: Hybrid position/force control in robotised machining. Sol. St. Phen. **210**, 192–199 (2014). https://doi.org/10.4028/www.scientific.net/SSP.210.192

7. Gierlak, P., Burghardt, A., Szybicki, D., Szuster, M., Muszyska, M.: On-line manipulator tool condition monitoring based on vibration analysis. Mech. Syst. Signal Proc. **89**, 14–26 (2017). https://doi.org/10.1016/j.ymssp.2016.08.002

8. Hendzel, Z., Burghardt, A., Gierlak, P., Szuster, M.: Conventional and fuzzy force control in robotised machining. Sol. St. Phen. **210**, 178–185 (2014). https://doi.org/10.4028/www.scientific.net/SSP.210.178

9. Burghardt, A., Szybicki, D., Kurc, K., Muszyska, M., Mucha, J.: Experimental study of Inconel 718 surface treatment by edge robotic Deburring with force control. Strength Mater **49**(4), 594–604 (2017). https://doi.org/10.1007/s11223-017-9903-3

10. Kumar, H., Sharma, C., Kumar, A., Arora, P. K.: Retrospective investigations of force measurement. MAPAN **30**(4), 291–302 (2015). https://doi.org/10.1007/s12647-015-0148-y

11. Kumar, R., Pant, B. D., Maji, S.: Development and characterization of a diaphragm-shaped force transducer for static force measurement. MAPAN, **32**(3), 167–174 (2017). https://doi.org/10.1007/s12647-017-0207-7

12. Shimachi, S., Hakozaki, Y., Tada, T., Fujiwara, Y.: Measurement of force acting on surgical instrument for force-feedback to master robot console. In: International Congress Series, vol. 1256, pp. 538–546. Elsevier (2003). https://doi.org/10.1016/S0531-5131(03)00356-X

13. Shimachi, S., Fujiwara, Y., Hakozaki, Y.: New sensing method of force acting on instrument for laparoscopic robot surgery. In: International Congress Series, vol. 1268, pp. 775–780. Elsevier (2004). https://doi.org/10.1016/j.ics.2004.03.310

14. Shimachi, S., Kameyama, F., Hakozaki, Y., Fujiwara, Y.: Contact force measurement of instruments for force-feedback on a surgical robot: acceleration force cancellations based on acceleration sensor readings. In: Duncan, J.S., Gerig, G. (eds.) Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005. MICCAI 2005. LNCS, vol. 3750, pp. 97–104. Springer, Berlin, Heidelberg (2005). https://doi.org/10.1007/11566489_13

15. Tutak, J. S.: Virtual reality and exercises for paretic upper limb of stroke survivors. Tehniki vjesnik-Technical Gazette **24**(2), 451–458 (2017). https://doi.org/10.17559/TV-20161011143721

16. Giergiel, J., Kurc, K.: Identification of the mathematical model of an inspection mobile robot with fuzzy logic systems and neural networks. J. Theor. App. Mech. **49**, 209–225 (2011)

17. Burghardt, A., Kurc, K., Szybicki, D., Muszyska, M., Nawrocki, J.: Software for the robot-operated inspection station for engine guide vanes taking into consideration the geometric variability of parts. Tehnicki Vjesnik-Technical Gazette **24**(2), 349–353 (2017). https://doi.org/10.17559/TV-20160820142224

18. Gierlak, P.: The manipulator tool state classification based on inertia forces analysis. Mech. Syst. Signal Proc. **107**, 122–136 (2018). https://doi.org/10.1016/j.ymssp.2018.01.002

System for Adhesion Control of Car Steering Column Couplers

Józef Grzybowski[1(✉)] and Dawid Janeczko[2]

[1] Rzeszów University of Technology, al. Powstańców Warszawy 12, 35-959
Rzeszów, Poland
`jozef.grzybowski@prz.edu.pl`
[2] PILC, al. gen. Okulickiego 20, 35-222 Rzeszów, Poland
`dawid.janeczko@pilc.pl`

Abstract. This paper presents designed and implemented system to measurement of parameters of car steering column couplers. Described system is operating as a quality control machine and it is a part of production line in severe environmental conditions. This paper consist of diagnostic and metrological requirements, design solutions, registered measurement parameters and system software.

Keywords: Distortion measurement · Welding seam detection
Industrial measurement system software

1   Introduction

Couplers made of steel and rubber are critical parts of car steering column. Mentioned coupler is a component consisting of two steel sleeves connected by rubber sleeve characterized by appropriate adhesion obtained during vulcanization process. The main measurement task is to perform test of each coupler. Test shall consist of compressing the coupler with strength 4 kN and measurement of distortion of rubber connection. Measurement system consists of device to adhesion control and controlling software which is connected with SQL database on the workplace server. The main task of the system is the 100% control of performed measurements as well as necessity of full identification of each coupler which has to be marked with individual permanent sign. Measurements are more difficult because tests are performed in the increased temperature environment caused by the temperature of couplers after vulcanization process and because of vulcanization fumes. This article describes designed and implemented solution of mentioned system.

2   Method of Controlling and Marking the Car Steering Column Couplers

The system to full control (100%) of adhesion of steel-rubber couplers, which is operating on the production line was designed for Sanok Rubber S.A. company. System consists of units shown in the Fig. 1.
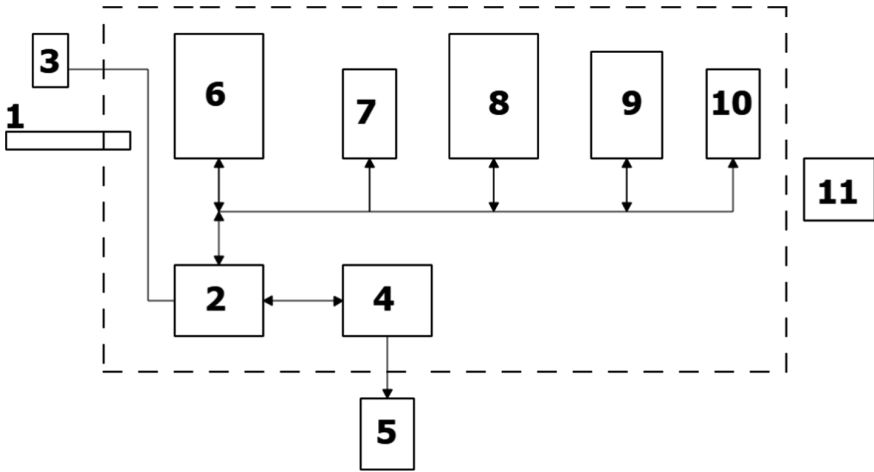
Fig. 1. Block diagram of the system to control of adhesion of car steering column couplers.

1. Operator's table
2. PLC controller
3. PLC controller's terminal
4. PC computer
5. Barcode scanner
6. Coupler's distortion measurement unit
7. Incorrect coupler's ejector
8. Weld detecting and coupler positioning unit
9. Laser marking machine
10. Correct coupler's ejector
11. Box for correct couplers.

Operating of the system shown in the block diagram (Fig. 1) is controlled on two levels. The first level is operated by PLC controller (2), which is presenting on the terminal (3) process parameters, the second level using industrial computer (4) connected with rubber type scanner (5). Couplers are placed on the operator's table (1), using transmission belt are moved to the measurement device. Distortion measurement unit is applying squeezing force (4 kN) which is determined by strain gauge transducer. Distortion of the coupler is determined as a displacement of the fulcrum related to LD-100 laser distance transducer produced by OMRON company. Measured values are transmitted to industrial computer (4) by the measurement unit. Distortion characteristic of each coupler is determined based on 50 measurements and then transmitted to workplace server via internet connection. Simultaneously level of distortion is compared with values set in controller's software, in case the item is incorrect, it is rejected (7). The next task of the system is to recognize the weld on the surface of the coupler and proper positioning of the item in order to do permanent marking using laser marking machine. After making the marking sign, the coupler is ejected to the box (11) by the pneumatic ejector (10).

3   Production, Metrological and Diagnostic Requirements

Vulcanizing machine releases 32 couplers in one series, subsequently two operators manually cut all overflows of the rubber and put couplers into holes in transporting belt. Duration of one operation of putting coupler into belt do not exceed 3 s. Machine is collecting 64 pieces of correct couplers in box at the end of line. Device to control of adhesion of car steering column couplers has to provide respectively.

1. measurement and evaluation of distortion of each coupler
2. rejection of incorrect coupler
3. detection of weld on the surface of the coupler
4. positioning of the coupler in order to correct marking
5. laser burning of pictogram on the surface of coupler
6. ejection of marked coupler to box with correct items.

Points 1 and 4 listed above are the main metrological tasks during operation of the machine. In the first case metrological task is realized by distortion measurement unit placed behind inspection doors E1 (Fig. 3). Described unit consists of pneumatic actuator generating pressure force of 4 kN on the tested coupler, measured by strain gauge force transducer with signal conditioner. Measurement of displacement of distorted coupler is realized using laser transducer LD100 made by OMRON company. Test cycle consisting of 50 measurements allows to prepare report with graph shown in
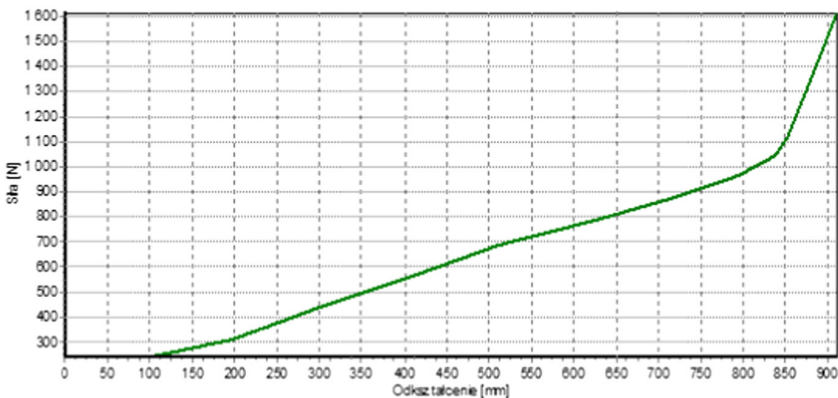
Fig. 2. Chart showing characteristic of distortion for correct coupler (force N/2, distortion mm × 100).

Fig. 2. The limit value for the compliant coupler is set depending on the type of rubber. The result of the measurement is presented by the terminal placed next to the operator table [3]. For the incorrect coupler, PLC controller starts procedure for rejecting wrong item by the activating actuator placed behind inspection door E3 (Fig. 3).

The next metrological task is detection of weld placed on the surface of the coupler. It is realized by the device placed behind inspection door E3 (Fig. 3). Coupler is rotated among three laser transducer type LD100, two of them are placed in front of coupler and generates differential signal of fluctuations of coupler's longitudinal axis and the third is sending auxiliary signal to identification of welding seam.

ARM microcomputer system is filtering signals from transducers, signals are disrupted by vibrations and by changes of position of coupler. Based on comparison of this filtered signals system is detecting welding seam on the coupler [1]. Charts with graphic visualization of described signals are shown in Fig. 4.
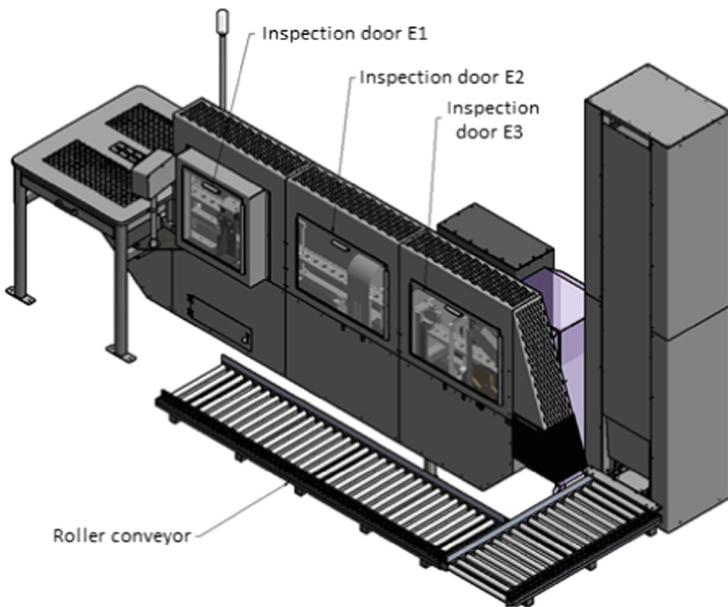
Fig. 3. Visualization of machine to 100% control of adhesion of car steering column couplers.

Disorder visible on the upper chart (Fig. 4) is used to comparison and stopping rotation of the coupler, moved to marking device placed behind inspection door E3. Burning of individual pictogram by the laser on the surface of each coupler is the end of the identification and marking process (Fig. 5). After that coupler is moved to box with correct items.
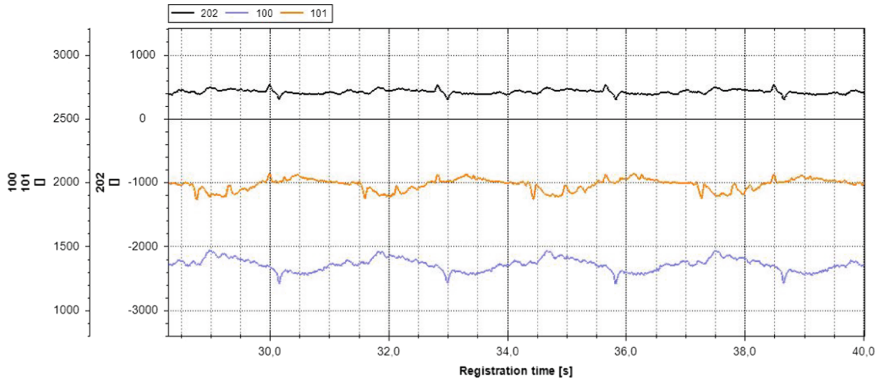
Fig. 4. Chart of filtered signals generated by the system detecting welding seam. 100, 101 – opposite transducers, 202 – received differential signal used to detection of welding seam.

Fig. 5. Car steering column couplers, left – item after test, dissection – negative result of test. Right – correct item with burned individual pictogram (QR code).

Figure 6 shown below presents machine with automatic device used to provide boxes for couplers.

Described machine is counting correct couplers. When amount of 64 pieces is reached, box with couplers is moved by the roller conveyor to prepared place into the box. At the same time, empty box is prepared for next batch of correct couplers.

4   Controlling Software

Controlling software consists of PLC driver software as well as industrial PC computer software. Display of operator's terminal is shown in Fig. 7.

Do avoid setting incorrect composition of rubber mixture, it is loaded direct based on scanned barcode. Controller software is supporting setting screen, which consists of tolerance and calibration factors. Connection with industrial computer provides

Fig. 6. Machine to 100% control of adhesion of car steering column couplers.
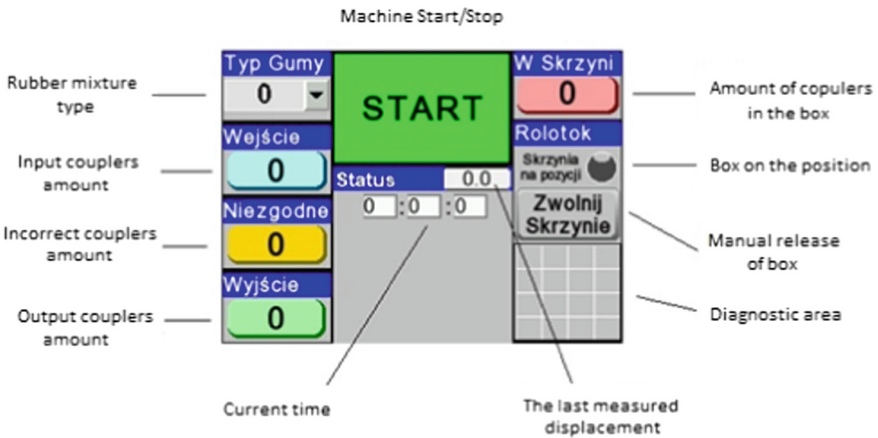
Fig. 7. Terminal display interface.

transmission of data from each measurement to further processing. Industrial computer software is converting data to text format, archiving on the local disc as well as sending data records to local workplace server (Fig. 8).

Subsequent component of the software is Quality Management System personnel PC computer software. It allows to access to workplace computer and real time preview of production as well as access to archive records.
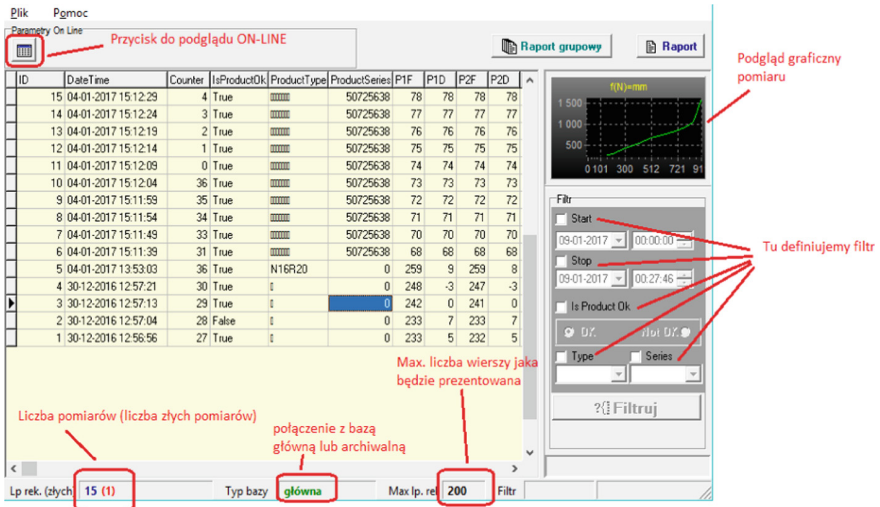
Fig. 8. Screen of PC computer with software for quality control.

5   Summary

On the primary stage of project, system based on visual method of detecting welding seam has identification errors level at 0,4%, however system based on the laser transducer doesn't show and identification errors. System works on the production line in the Sanok Rubber S.A. During one shift, machine is testing average more than 5 thousand couplers [2]. Currently device is operating with supervision to avoid contaminant of optical component by the vulcanization fumes. Laser transducers characterizes with high resistance to severe conditions of operating, optical parts of laser marking machine is susceptible to vulcanization fumes.

References

1. Baranowski, J.: Odtwarzanie stanu systemów dynamicznych z dyskretnych danych pomiarowych. In: Kacprzyk, J. (eds): Innowacyjne rozwiązania w obszarze automatyki, robotyki i pomiarów, praca zbiorowa red. nauk. Janusz. Warszawa (2011)
2. Babiasz, E.: Przewodnik zapewnienia bezpieczeństwa w konstrukcjach pokładowych urządzeń elektronicznych - dokument RTCA DO-254/ED-80EUROCAE. Zeszyty Naukowe Politechniki Rzeszowskiej, Mechanika nr 83. Rzeszów (2011)
3. Homik, W., Grzybowski, J.: System pomiaru parametrów kinematycznych wiskotycznego tłumika drgań skrętnych. Metody obliczeniowe i badawcze w rozwoju pojazdów samochodowych i maszyn roboczych samojezdnych. In: Materiały XVII Konferencji SAKON 06. Rzeszów (2006)

The Algorithm for Automatic Determination of Human Step Phases

Svitlana Herasymenko, Piotr Bieńkowski, and Piotr Kopniak[✉]

Institute of Computer Science, Lublin University of Technology,
ul. Nadbystrzycka 38d, 20-618 Lublin, Poland
`p.kopniak@pollub.pl`

Abstract. The article presents an algorithm for automatic determination of steps and individual phases of human gait based on data from an inertial sensor. To obtain the necessary data, measurements were made using Xsens inertial sensors from the MTw Awinda set. Changes in the feet rotation angles during gait and acceleration data were investigated. After a thorough analysis of the data, an algorithm for determining the step phases was implemented, based on which an application was created for automatically dividing gait into steps and steps into individual phases, which allows later analysis of gait parameters, e.g. for medical diagnosis. The main function of the application is a graphical presentation of the algorithm's operation, by generating a graph of data along with the selected step phases.

Keywords: Motion capture · Human gait · Algorithm · Step phases · Inertial sensor

1 Introduction

Nowadays, the subject of human gait research is quite often discussed. Gait evaluation is part of such research. Modern technologies, such as vision systems based on recordings from cameras, optoelectronic systems, portable systems based on various types of sensors, allow accurate recording and analysis of the moving human body, which is widely used in biomechanics, medicine, sports, computer graphics [1–5].

The division of the step into phases is an inseparable element of the gait analysis, whereby even the smallest changes taking place in the moving human body can be noticed, which gives completely new possibilities in medical and scientific research.

Very important effects of gait tests are the possibility to analyse pathological gait, comparing it to the normal gait, which may increase the accuracy of the diagnosis made by doctors and help to track the progress of rehabilitation [6].

1.1 Review of the State of Gait Research

There is a diverse scientific approach to pathological and physiological gait. According to Sutherland and his colleagues, mature gait can be determined by the regularity, the speed of movement, the length of the step, the duration of the support phase and the relationship between the width of the step and the intercostal distance of the pelvis [7].

Mature gait is characterised by functional mechanisms that Burnett and Johnson are seeking to address, including: achieving foot and knee movement sequences, reducing stride width, coordinating upper limb and torso movements, and starting pressure on the heel while putting feet on the ground with dorsally folded toes [8].

The conditions that Gage set [9] in order to achieve the correct gait consist of: adequate energy consumption and step length, correct foot position at the beginning of the step, proper foot position at the end of the step and proper body position during the support phase.

The subject of work [10] is the gait research and analysis system. During the tests, acceleration measurements were made at various anatomical points, and three-axis acceleration sensors were used for this. Data transmission took place in the Bluetooth standard. During the work carried out, software was created to record, process and analyse data. Gait tests were carried out on 17 people. As a result, parameters related to events in the gait cycle were obtained, such as: the cycle period of the walk, the time of the support phase, the time of the transfer phase and the range of accelerations.

The next study [11] describes an inertial measurement system. At the beginning the terminology of the gait cycle and the method of performing the tests are described. The system itself is used to analyse human gait and determine its main parameters, such as: step duration, duration of the support phase and transfer for the left and right legs. Several Inertial Measurement Units (IMUs) were attached to the feet of the subject. Using a program, the signal from the IMU is analysed, and to be exact the acceleration value by means of which the main gait parameters were calculated.

In [12], the APAS motion analysis system and Kistler dynamometric platforms were used to determine kinematic quantities and ground reaction forces during normal walking. The research was carried out on 15 volunteers, who were healthy people. Different gait speeds were tested (slow, normal and fast gait), and the obtained results were analysed in the original GAIT program written in the MatLab environment.

1.2 Objective of the Work

Scientific research shows that the problem of gait analysis is very important for medical diagnostics and science, for example the development of robotics. Therefore, the aim of the study is to examine healthy human gait based on data obtained from Xsens inertial sensors from the Awinda MTw set and the development of an algorithm for computer gait analysis. The acceleration data

measured by inertial sensors in three axes and Euler angles, i.e. rotation angles around each of the three coordinate axes in three-dimensional space, will be used for analysis. This article puts forward the thesis that the acceleration and the legs angular position data registered by inertial sensors are sufficient to determine individual phases of the human step. A scientific hypothesis was also assumed that checks whether it is possible to create an algorithm based on foot inclination data, which will enable automatic detection of individual phases of the human step.

2   Measuring System and Recorded Data

MTw inertial sensors from Xind's Awinda system were used for the tests. Xsens offers various motion registration systems based on inertial sensors. The company mainly deals in the production of two main families of motion sensors, MEMS, MTx and MTw. In the study the MTw sensors were used, shown in Fig. 1 [2].

Fig. 1. Xsens MTw inertial sensor [13]

The MTx and MTw sensors are slightly smaller and lighter, therefore they are used to measure the movement of the human body. The MTx and MTw sensors use a Velcro strap set so that they can be attached to the limbs and torso or can be placed in the pockets of a special elastic uniform from the same supplier. MTw sensors have their own power supply and can communicate wirelessly with the Awinda radio protocol as opposed to MTi and MTx sensors, which are connected by cable to a portable central unit that records and transmits data to a computer via a Bluetooth interface or a wired serial interface via USB.

MTw sensors are small, completely wireless and accurate recorders, which, similarly to MTx sensors, provide data about the orientation in 3D space, as well as kinematic data, that is: three-dimensional rotation, three-dimensional acceleration, three-dimensional data on the Earth's magnetic field. Devices of

this type are used to record the movement of parts of the human body, which can be used during training, pedestrian navigation, in robotics, etc.

Sample parameters of MTw sensors are presented in Table 1.

Table 1. Example parameters of MTw sensors [13]

Name of feature	Value
Static accuracy of measurement of roll and pitch angles	$< 0.5°$
Dynamic accuracy of measuring angles	$1.5°$
Accuracy of deflection measurement	$<1.0°$
Speed of gyroscope measurement	to 2000 $°/s$
The maximum acceleration measured: $50\,\text{m/s}^2$	to 160 m/s^2
Working temperature range	-25 to $80\,°\text{C}$
Output frequency (measurements)	to 1000 Hz
Communication interfaces	RS232/485/USB
Mass	27 g

All Xsens inertial sensors work in the same way. The sensors that have been mentioned before have an Inertial Measurement Unit (IMU), i.e. a measuring device for recording spatial location based on magnetometers, gyroscopes and accelerometers.

The tests were carried out on 5 healthy people (3 males and 2 females aged from 25 to 43 years), each of them made 10 passes on the route of about 10 m, i.e. 10 steps to one side and the other; they were all wearing shoes.

The above-mentioned sensors were used for the tests, which were mounted on the right and left leg on the foot, calf and thigh, in the places shown in Fig. 2 (on the left side). Each sensor takes Euler's acceleration and angular data for three axes X, Y and Z at 100 Hz using the MT Software Suite program provided by the manufacturer. For each sensor, the arrangement of each axis is identical, the sensor axes can be seen in Fig. reffig:attachsensors (right).

3 Step Division into Phases and Data Analysis Algorithm

In order to develop an algorithm for dividing the set of data recorded during a persons walk into steps and determining their phases, it is necessary to analyse the gait cycle itself. The walking cycle consists of movements made during a single step (Fig. 3). One cycle is the time between the two closest moments when the foot touches the ground. Each cycle begins with the stance phase and ends with the swing phase. During the support phase, which occupies about 60%, the limb is in contact with the ground. From the time the foot is lifted, the transfer phase starts and lasts until the foot is put back again; this phase is 40% of the walking cycle [11].
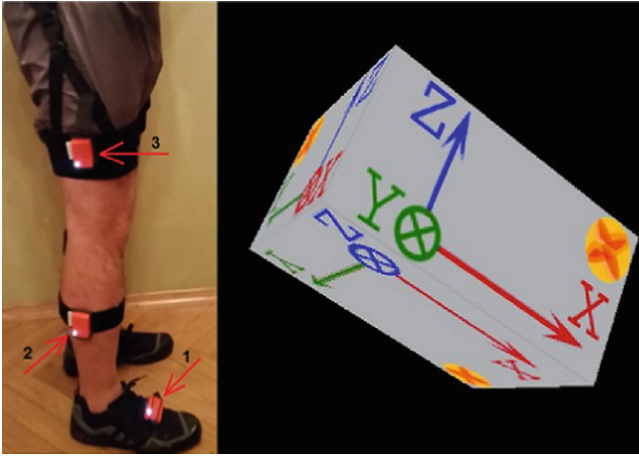
Fig. 2. Points of attachment of the sensors (on the left): sensor No. 1 on the foot, sensor No. 2 on the calf, sensor No. 3 on the thigh; sensor axes (on the right) [15]

The support phase can be divided into 5 stages [11,14]:

1. Initial contact – heel contact with the ground, the centre of gravity of the body at this point is in the lowest position.
2. Loading response – the depreciation phase starts from the moment when the foot is flat against the ground, taking over the entire weight of the body.
3. Midstance – the middle phase of support, followed by a moment when the opposite limb begins to move and ends when the centre of gravity of the whole body is above the foot.
4. Terminal stance – the final support phase, ending when the opposite limb touches the ground.
5. Preswing period – the phase before transfer, it starts when the opposite limb touches the ground.

The transfer phase is divided into three stages [17]:

1. Initial swing – lifting the toes, starts when the foot leaves the ground and ends when the knee bends as much as possible.
2. Midswing – the central phase of transmission, lasts until the tibial bone is placed vertically.
3. Terminal swing – the final phase of the transfer, the limb moves ahead of the torso and then slows down, preparing to take over the whole body weight, that is the end of the walking cycle. This phase overlaps with the first cycle of the support phase.

3.1   Dividing a Step Into Phases

In order to implement the step division algorithm, the analysis of data received from the sensor, which was located on the foot, more precisely during one step,
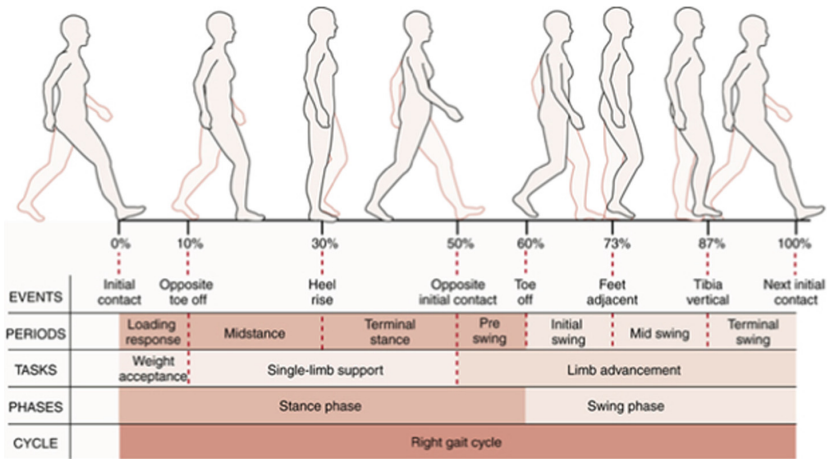
Fig. 3. Gait cycle [16]
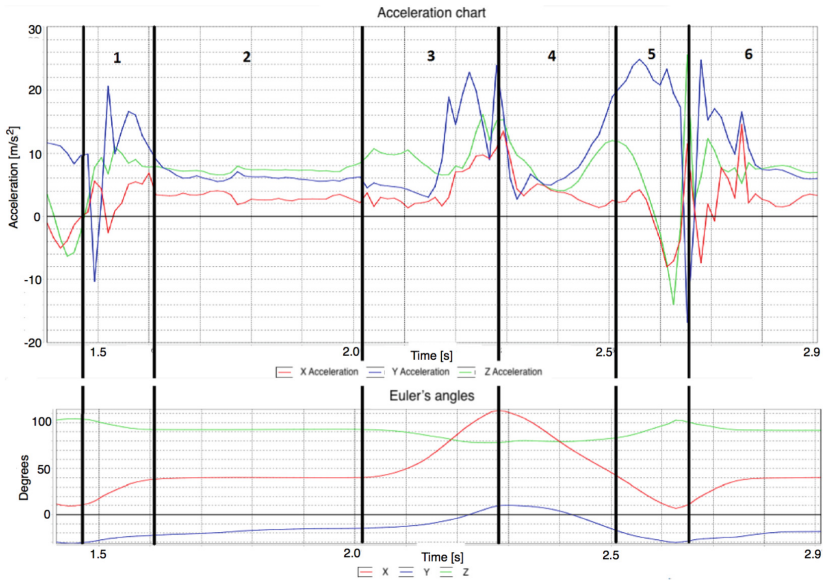
Fig. 4. One step divided into parts, acceleration graph (top), Euler angle graph (bottom)

was carried out. The data presented in the chart (Fig. 4) have been divided into parts:

1. heel contact with the ground (Initial contact), at which point also the support phase begins; the foot is being lowered, the acceleration in each axis is evened out, for the opposite foot it is the moment of lifting the fingers off the floor,

from the Euler angle graph it can be deduced that in the X axis (marked in red) the tilt angle changes slightly, the foot, starting from the heel, begins to tilt until it touches the floor;

2. the entire foot is adjacent to the ground, acceleration is constant, but not zero, because the sensor records terrestrial acceleration, with respect to the ground the acceleration is close to zero with slight vibration, which means that the foot has a constant speed, and the cycle is Midstance;

3. the moment when the heel begins to rise, the angle with respect to the X axis increases (Terminal stance), the heel is at the highest point at the highest angle of inclination, with high acceleration, the fingers start to detach from the ground (Pre swing);

4. lifting the fingers (Initial swing), this is the beginning of the transfer phase, the foot begins to move forward, changing the tilt angle, at some point the foot is again horizontal, the inclination of the sensor in the X axis is about 40 degrees (Mid swing);

5. lowering the leg, the toes are raised and the angle is below the initial, the heel is lower than the fingers, the vertical acceleration decreases and at one point is below zero, which means that the speed decreases before putting the heel back on the ground (Terminal swing);

6. the beginning of the next step.

3.2 Description of the Algorithm

The algorithm operates on the basis of the Euler angle data obtained with a sensor located on the foot. The data of rotation angles around the X axis are analysed, i.e. changes in angles showing the bending of the foot during walking.

The algorithm works in such a way that it analyses all measurements one by one, starting from the first, comparing each measurement with the previous one, in addition to the first measurement.

The algorithm used a deflection a value denoting a deviation from the starting point up or down (Fig. 5). The value of 5 degrees is the optimal deflection value.

Taking into account the deflection, three cases should be considered:

– case 1 the current point is below the starting point from which the deflection is subtracted, in Fig. 5 this is the range below the blue horizontal line;
– case 2 the current point is between the deflection points, in Fig. 5 it is the range between the green and blue horizontal lines.
– case 3 the current point is above the starting point to which the deflection is added; in Fig. 5 this is the range above the green horizontal line.

Figure 6 shows a UML activity diagram showing how to choose the right case. It consists of three conditional blocks, each one corresponding to one case described above. Depending on the fulfillment of the appropriate conditions in each block, the identifiers of the measurements are added to the list, which indicate the points at which the corresponding step phases start or end.
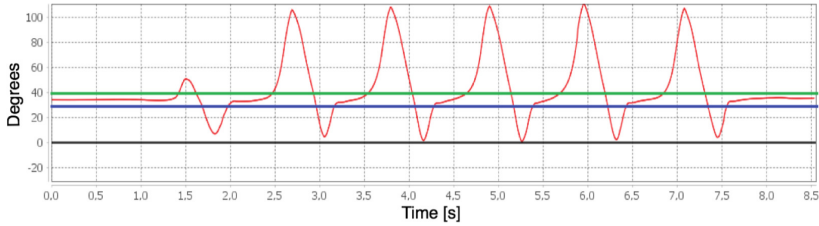
Fig. 5. Lower (blue line) and upper (green line) deflection

First Conditional Block

The first block corresponds to case 1, the second block corresponds to case 2, and the third block is case 3. The first conditional block will be executed first because the starting points in the graph are between the deflection values, and the first and second condition will not be met. The next one will execute the third block, just as in the time point of 1.4 s (Fig. 5) the chart exits above the upper deflection. Then the values are again found between the deflections, i.e. the second block will be made. After that, the values will exceed the lower deflection and at this moment the first block will be executed, etc.

The first conditional block is shown in Fig. 7. Five operations are performed in parallel in this block:

– a condition is checked if the actual value is smaller than the previous value. If this is the case, the measurement identifier at the current point is saved in the stanceInitId variable. This block will be executed several times, until the graph is below the value of the lower deflection, and as a result of this the stanceInitId variable will be assigned a value in the minimum point, i.e. the point which marks the beginning of the support phase and the beginning of the step (Fig. 13),
– a measurement identifier is added to the list, denoting the beginning of the transfer phase; this is the point shown in Fig. 8. The value of this variable was assigned during the execution of unit No. 3,
– the value true is assigned to the isNearSwingInit variable; this variable will be used in block No. 3.
– the value true is assigned to the isLikeInit variable; this variable will be used later in block No. 2,
– the current angle value is assigned to the previous variable, which will be the previous value at the next execution of any block.

Second Conditional Block

In Fig. 9, a second conditional block is shown, which is executed when the current angle value lies between deviations. Five operations are performed in parallel in this block:
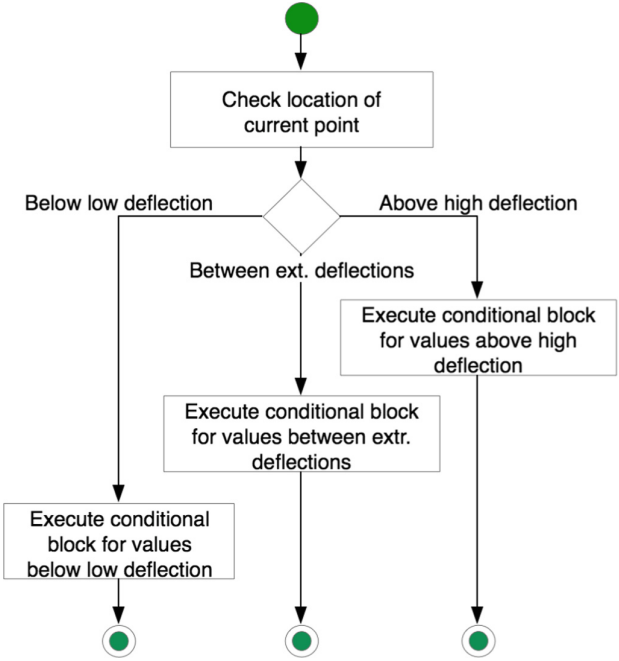
Fig. 6. UML activity diagram showing the method of choosing the appropriate algorithm
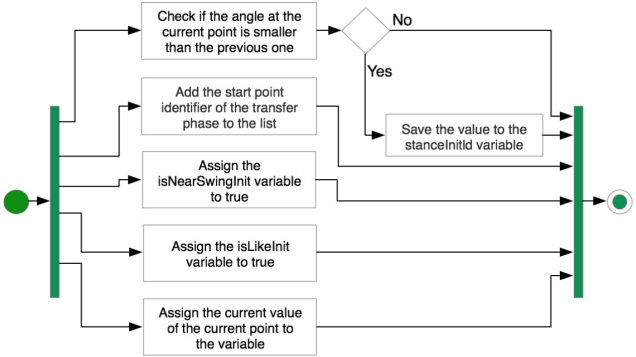
Fig. 7. The first conditional block

– the isLikeInit condition is checked. The value of this variable is set to true when executing blocks 1 and 3. This condition marks the beginning of the second block execution, at which point the variable value is set to false to prevent further unnecessary identifiers from being written to the list, as a result the identifiers will be added to the list at the appropriate points shown in Fig. 10.
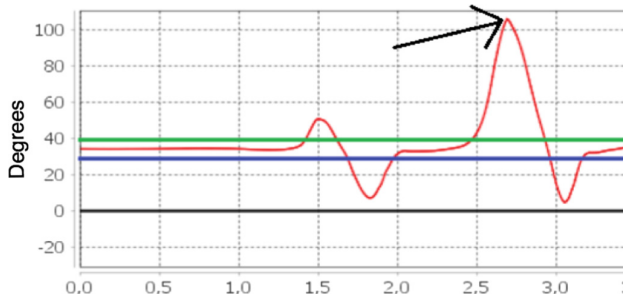
Fig. 8. A point on the graph indicating the beginning of the transfer phase

– the identifier of the beginning of the support phase is added to the list of identifiers, saved during the execution of block No. 1,
– the identifier of the beginning of the transfer phase is added to the list of identifiers, saved during the execution of block No. 3,
– the value true is assigned to the isNearSwingInit variable; this variable will be used in block No. 3, the current angle value is assigned to the previous variable, which will be the previous value at the next execution of any block.

Third Conditional Block

The third conditional block is shown in Fig. 11. Five operations are performed in parallel in this block:

– the isNearSwingInit condition is checked, the value of this variable is set to true in other condition blocks, this condition for one block execution series will
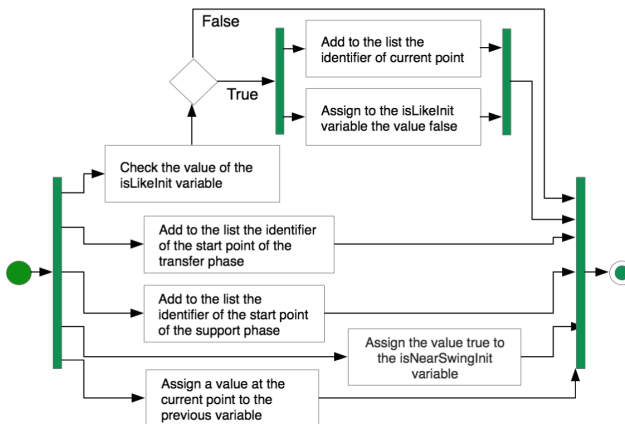
Fig. 9. The second conditional block

be met only once at the first execution, because this variable is immediately assigned the false value to prevent adding to the list unnecessary identifiers. At this moment, the measurement identifier will be added to the list at the point shown in Fig. 12,
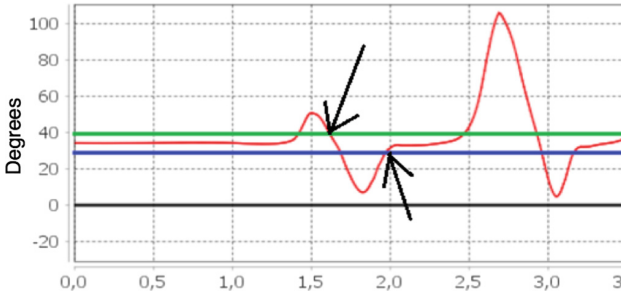
Fig. 10. Points on the chart that indicate different stages of the step
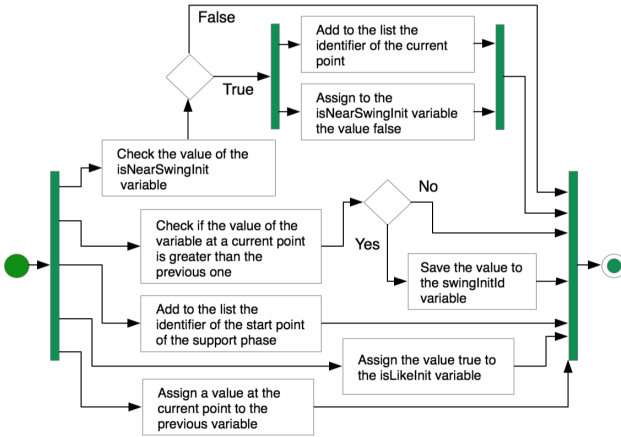
Fig. 11. The third conditional block

- the condition is checked if the actual value is greater than the previous value. If the condition is met, the measurement identifier at this point is stored in the swingInitId variable. This block will be executed several times, until the graph is above the deflection value, and as a result the variable swingInitId will be assigned the value of identifier at the maximum point, i.e. the point which marks the beginning of the transfer phase (Fig. 8),
- a measurement identifier is added to the list, denoting the beginning of the support phase (Fig. 13). The value of this variable was assigned for the time of execution of block No. 1,

– the value true is assigned to the isLikeInit variable; this variable will be later used in block No.2,
– the current angle value is assigned to the previous variable, which will be the previous value at the next execution of any block.
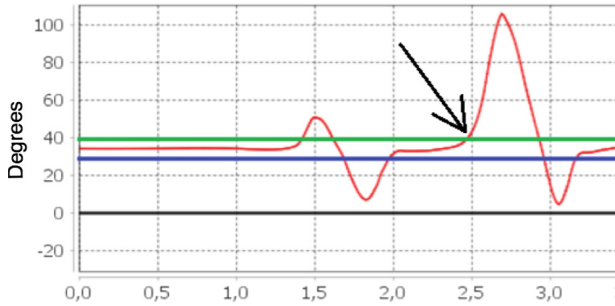
Fig. 12. A point on the graph marking approach to the beginning of the transfer phase
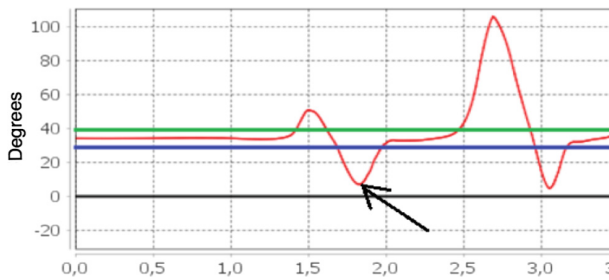
Fig. 13. A point on the graph indicating the beginning of the support phase

After completing all the series of conditional blocks, the measurement identifiers were added to the three lists at the points corresponding to the beginnings of the support phase (stanceInitIdList list), the beginning of the transfer phase (swingInitId list) and the beginning of the individual steps (listPeriodsIdList list). Depending on which list the identifiers are downloaded from when drawing the chart, vertical lines are added to the chart:

– the thicker black line indicates the beginning of the support phase (beginning of the step),
– the grey line, a little thinner than the line of the beginning of the step, marks the beginning of the transfer phase,
– the barely visible grey line means individual stages of the step.

4   Algorithm Tests

In order to test the analytical application that implements the described algorithm, the data from the measurements made for the first group of 5 healthy people (3 males and 2 females aged from 25 to 43 years) were used. For each person, measurements were made for the right foot, or, in total, 5 measurement cycles. Using the MT Manager program (an application provided by the Xsens sensor manufacturer), the Euler angle data were exported to text files. The content of the files has been formatted in accordance with the requirements of the created application. Figure 14 shows graphs of the Euler angle data, obtained by using an analytical application. Each graph shows the division into individual stages of the step.

Afterwards, the verification were conducted with group of 24 healthy people (males and females aged from 20 to 21 years). All test confirmed that the developed algorithm worked well.

At the end the two abnormal gaits were examined, i.e.: parkinsonian gait and neuropatic gait known as steppage gait, too. Characteristics of these gait is described among others by Tarrulli [18]. Patients with Parkinson Disease are hunched over, initiate slowly, and take short strides. Neuropathic gait is caused by weakness of foot dorsiflexion. Such gait type is characterised by leg lift high enough so that the foot does not drag on the floor.

The diagrams (Fig. 14) also show the points of division of steps into individual phases in the places described in Sect. 3.1. After testing the application, it can be concluded that the algorithm for automatic division of the step into phases works well.

Figure 15 shows graphs for segmentation comparison of the Euler angle data measured during normal and abnormal gaits. As we can see in the figure the characteristics of the inclination angles in cases of people with neurological disorders differ from the healthy person's gait. Both in the case of neuropathic walking and Parkinson's walk, there is basically no rotation of the foot at the beginning of the support phase, ie there is no standing on the heel.

Changes in the angle of rotation of the foot during neurotic walking have a similar sequence as in the case of a healthy person walking. The angle variation is smaller. During the support phase, the foot does not float on the fingers and is carried violently. Despite this, the algorithm worked correctly and the segmentation was done appropriately. For each step the algorithm correctly determined all phases of the step

Parkinson's walk is characterized by short steps. Low foot lift and shuffling on the ground causes irregularities in the recorded signal. Also the angles of rotation of the foot have small values - only about 20°. In spite of this, it is possible to determine some phases of gait using the developed algorithm. For each step, the beginning of the transfer phase has been correctly defined, i.e. the moment when the heel begins to rise (Terminal stance), the moment when the heel is at the highest point at the highest angle of inclination (Pre swing) and the point before putting the heel back on the ground (Terminal swing).
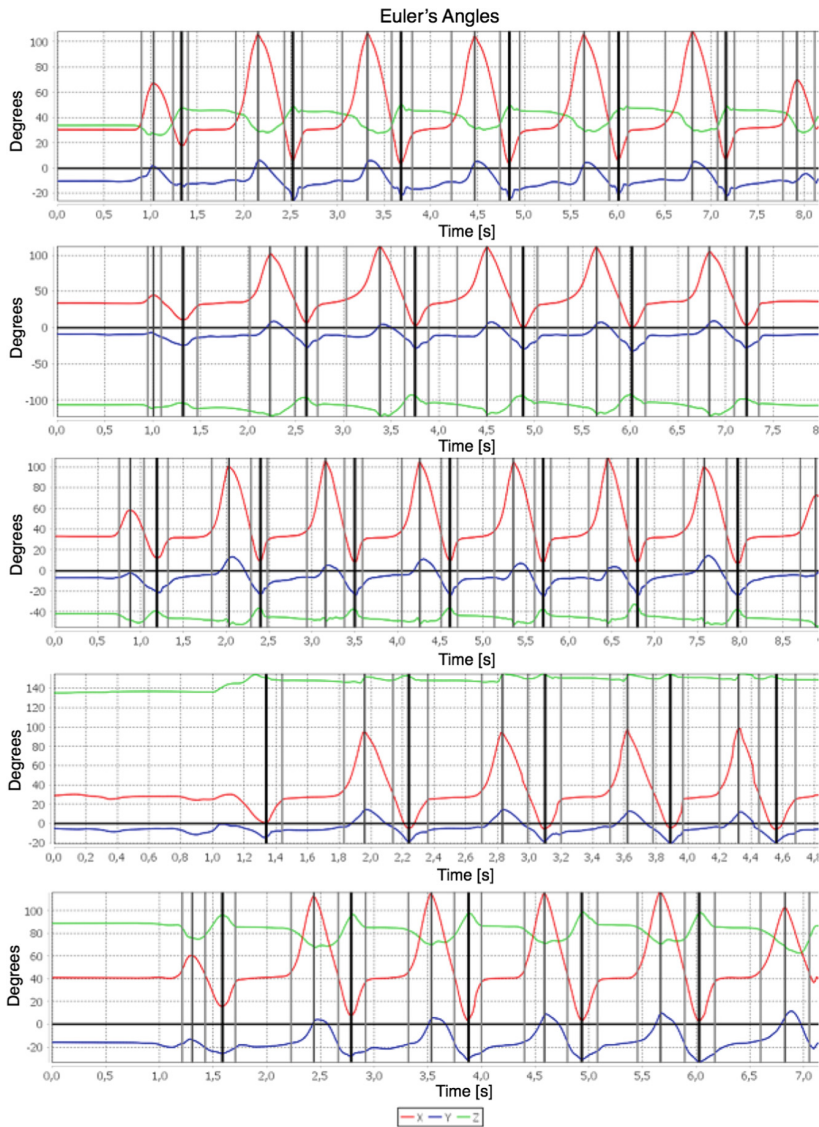
Fig. 14. Graphs of Euler angles from the sensors on the right leg of each subject (the first the first graph from above, the second the second graph from above, etc.)

It follows that the algorithm should be additionally modified in order to examine some cases of walking people with neurological disorders very different from healthy people walking. Such modifications are foreseen in the future.
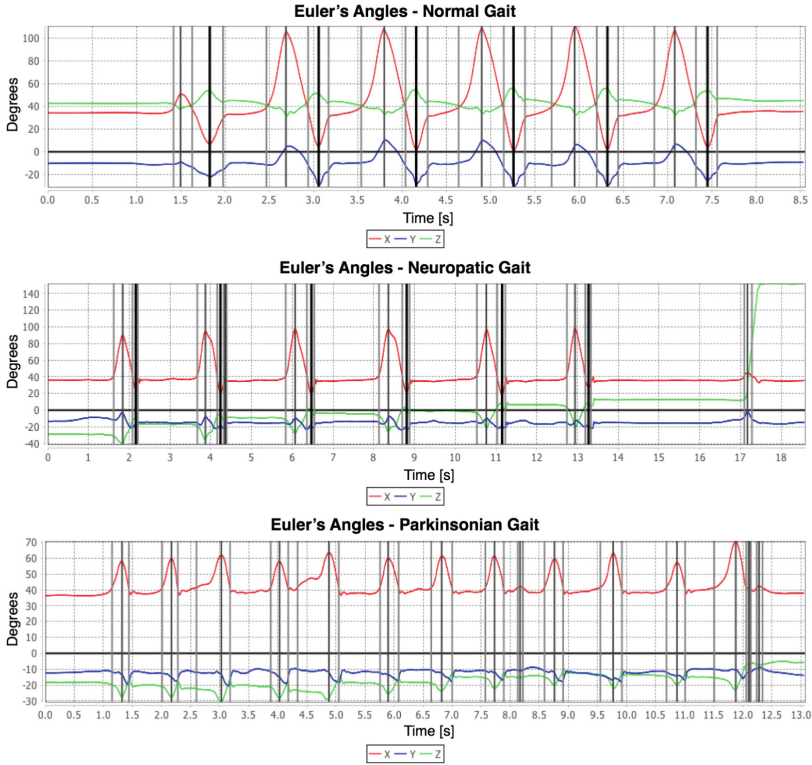
Fig. 15. Comparison of Euler's angles of three gait types: normal gait, neuropatic gait and parkinsonian gait

5   Summary

The above studies were aimed at analysing phase sequence of human gait and developing an algorithm for automatic division of gait data into steps and their phases. The algorithm has been implemented in the proprietary analytical application that will be able to visualise the individual phases through graphical presentation of data plots. At the beginning, we describe the way in which a person should move in a way that is considered to be correct. Next, the algorithm for dividing the steps into individual phases is presented. The application was tested on the data obtained during the research.

All motion data were obtained through a sensor attached to the right foot of a subject. The stage of downloading the relevant data required for testing was not without problems. The most important of the errors encountered was the problem that appeared after downloading the data from the sensor. It turned out that the data set was incomplete. This was caused by the fact that during the test, there were interruptions in the signal (loss of connection with the sensors and the receiver). The research was conducted during a natural human walk

over a distance of 10 m. Possible measurements on the treadmill could enable synchronisation of data with the reference camera and a more detailed analysis of the collected data.

To determine the individual phases of human steps, the data of accelerations and angular position, registered by inertial sensors, were enough to confirm the earlier thesis.

The assumed hypothesis that the creation of an algorithm that will allow automatic detection of individual phases of the human step is possible using data on the angle of inclination of the foot has been confirmed. The proof of this is an application with an implemented algorithm, which, as was assumed at the beginning, divides the human gait into individual phases. The performed tests of the application show its correct operation.

Based on the data on the steps and their phases obtained as a result of the developed algorithm, further research may be conducted on:

- distinguishing the left leg from the right,
- distinguishing a woman's walk from a mans,
- distinguishing the walk of a healthy person from a person with a walking disorder,
- recognition of gait phases of people with neurological disorders,
- determining if there is progress in a patients rehabilitation by comparing previous data with current ones.

References

1. Chen, S., Lach, J., Lo, B., Yang, G.: Toward pervasive gait analysis with wearable sensors: a systematic review. IEEE J. Biomed. Health Inform., 1521–1537 (2016)
2. Kopniak, P.: Construction, principle of operation and application of the Xsens motion capture system. Logistyka (3), 3049–3058 (2014)
3. Kopniak, P.: Java wrapper for Xsens motion capture system SDK. In: 7th International Conference on Human System Interactions (HSI), pp. 106–111. IEEE, Lisbon, 16–18 June 2014
4. Kopniak, P., Kaminski, M.: Natural interface for robotic arm controlling based on inertial motion capture. In: 9th International Conference on Human System Interactions (HSI), Portsmouth, pp. 110–116, 06–08 July 2016
5. Kopniak, P., Bocian, K.: Differentiating horse walk and trot on the basis of inertial motion capture. In: 9th International Conference on Human System Interactions (HSI), Portsmouth, pp. 270–276, 06–08 July 2016
6. Drubicki, M., Szymczyk, D., Snela, S., Dudek, J., Chuchla, M.: Objective, quantitative gait analysis methods in clinical practice. Przeglad Medyczny Uniwersytetu Rzeszowskiego (4), 356–362 (2009). Rzeszow
7. Sutheland, D.H., Olshen, R., Cooper, L., Woo, S.L.-Y.: The development of mature gait. J. Bone Joint Surg. (1980)
8. Burnett, C.N., Johnson, E.W.: Development of gait in childhood. Dev. Med. Child Neurol. (1971)
9. Gage, J.: The Treatment of Gait Problems in Cerebral Palsy. Mac Keith Press, London (2004)

10. Maj, P., Barczewska, K., Drozd, A., Kowalski, J.: Accelerometer gait examination and analysis. In: PAK, vol. 58, nr 4/2012 (2012)
11. Korotkin, D., Kuznetcov, A.: Inertial measurement system for human gait analysis. In: Proceeding BodyNets 2013 Proceedings of the 8th International Conference on Body Area Networks, pp. 414–419 (2013)
12. Śniezek, A., Mezyk, A., Michniuk, R.: Analysis of the dynamics and kinematics of normal walking. Aktualne Problemy Biomechaniki, 1/2007 (2007)
13. MTw Development Kit Lite (legacy product) (2018). https://www.xsens.com/products/mtw-development-kit-lite. Accessed 15 Mar 2018
14. Chwała, W., Serafin, W., Ruchlewicz, T.: Biomechanical analysis of gait disturbances in a person subjected to hip arthroplasty. Acta Sci. Acad. Ostroviensis **13**, 85–102 (2003)
15. Quick start for MTi Development Kit (2018). https://base.xsens.com/hc/en-us/articles/209580825-Quick-start-for-MTi-Development-Kit. Accessed 15 Mar 2018
16. Kernozek, T.W., Willson, J.D.: Gait. http://clinicalgate.com/gait-2/. Accessed 15 Mar 2018
17. Zak, E.: Gait phases and muscular activity. http://www.fizjoterapeutom.pl/files/5/fazy_chodu_aktywnosc_miesniowa.pdf. Accessed 15 Mar 2018
18. Tarulli, A.: Gait disorders. In: Neurology. Springer, Cham (2016)

A Procedure for Validating Impedance Parameters of HF/UHF RFID Transponder Antennas

Piotr Jankowski-Mihułowicz[1][(✉)] , Mariusz Węglarski[1],
and Wojciech Lichoń[2]

[1] Department of Electronic and Telecommunications Systems,
Rzeszow University of Technology, Pola 2, 35-959 Rzeszow, Poland
{pjanko,wmar}@prz.edu.pl
[2] Talkin Things, Pulawska 182, 02-670 Warsaw, Poland
wojciech.lichon@talkinthings.com

Abstract. The performance of automatic identification in every RFID system is strongly dependent on proper operation of the transponders that are used to mark different kind of objects. The impedance matching between chip and connected antenna is the most significant component determining the design quality of transponder internal circuitry, and hence influencing overall system parameters such as shape and dimensions of interrogation zone, level of identification efficiency, etc. Taking into consideration the various types of RFID systems, the problem has to be considered differently with respect to the operating frequency. Moreover it has to be treated in a different way than it is known from the classical theory of typical radio communication systems. The authors have proposed and developed their own method for validating impedance parameters of RFID transponder antennas operating in the regular HF and UHF bands. It is based on a generalized model of the RFID transponders dedicated to different standards. The developed test procedure consists of four steps involving antenna designing, manufacturing, measuring and validating processes. The practical usefulness of the proposed method is confirmed by experiments conducted with using representative examples designed in research and development projects realized with partners from the industry.

Keywords: RFID · HF · UHF · Transponder · Antenna · Impedance Validation

1 Introduction

Progress in the electronics, telecommunications, computer science, robotics and control engineering as well as in many other areas of the technology has a significant impact on changes in the human environment [1]. As a consequence of these alternations, the growing number of intelligent technology implementations with ability to support an advanced automatic identification of objects is observed [2, 3]. The radio frequency identification (RFID) technology is commonly used in such systems [4–6].

The RFID systems are placed in a group of radio equipment devices [7]. They use bands (LF, HF, UHF, etc.) and operating frequencies that are commonly available for industrial-scientific-medical (ISM) and many other purposes [8]. Therefore, the band and operating frequency constitute basic factors influencing the differentiation between types of RFID systems. They subsequently determine a different approach in terms of considering the performance essence of the systems.

Irrespective of the mentioned frequency bands, a software and hardware parts may be distinguished in any RFID system. The software serves for both direct control of individual digital devices as well as for managing the whole system. The hardware part is composed of a read/write device (RWD) with at least one antenna and a single or many electronic transponders which are used to mark objects.

The RFID transponder consists of a chip with a connected antenna [7]. The impedance matching of these two components is the basic requirement for the proper operation of this electronic device. This problem has to be considered differently with respect to constructions that are commonly used in RFID systems operating in HF [9–11], UHF [12–18] or multi bands [19]. It should be emphasized that the efficiency of automatic identification of electronically marked objects is determined by the proper operation of the transponders – the design quality of transponder internal circuitry has significant impact on parameters of RFID systems, such as shape and dimensions of interrogation zone (IZ), satisfactory level of identification efficiency, etc. [20].

It should be also noted that phenomenon of impedance matching between the transponder antenna and chip has to be considered in a different way than it is in the classical theory of typical radio communication systems. First of all, the impedance of the chip is represented by two parts: imagine and real, and it differs from the pure real value of 50 Ω or 75 Ω that is commonly considered in classical radio devices. Moreover, this impedance varies with the amount of power that is transferred from the antenna to the chip. Secondly, the complex impedance of the antenna is vulnerable to the environmental conditions such as kind of material which the marked object is made of [11, 16–18], fluctuations of ambient temperature [21] and many others. Thirdly, connectors (N, SMA, UFL, etc.) are not applicable to the components of the transponder and, on the other hand, the type of interconnection structure in the chip has significant impact on the impedance of the antenna [22] and methods that are used to measure this parameter [23]. Finally, the antenna design determines the selection of measuring equipment.

The method for validating impedance parameters of RFID transponder antennas operating in the regular HF and UHF bands has been developed taking into consideration the above mentioned conditions. It has been worked out on the basis of the authors' experience that had been growing and expanding for several years during numerous projects related to industrial implementations of RFID devices. The proposed method can be adapted to both scientific as well as industrial research and development laboratories. Since it arises from the generalized model of the RFID transponder (Sect. 2), the impedance parameters of the device can be validated on the basis of algorithm yielded from the model synthesis (Sect. 3). The practical usefulness of the developed method is confirmed on the basis of a few representative examples (Sect. 4), which are the result of the research and development cooperation between the Department of Electronic and Telecommunications Systems (DETS) in Rzeszow University of Technology (RUT) and industry partners.

2   Impedance Model of HF/UHF RFID Transponder

Internal constructions as well as operation principles of transponders are strongly dependent on the considered frequency band and the carrier frequency (in the HF band f_0 is equal 13.56 MHz, in the UHF band: f_0 = 860–960 MHz depending on world regions). Nevertheless, the generalized model is proposed for the desired measuring process that could be used for validating required parameters in any RFID system (Fig. 1).
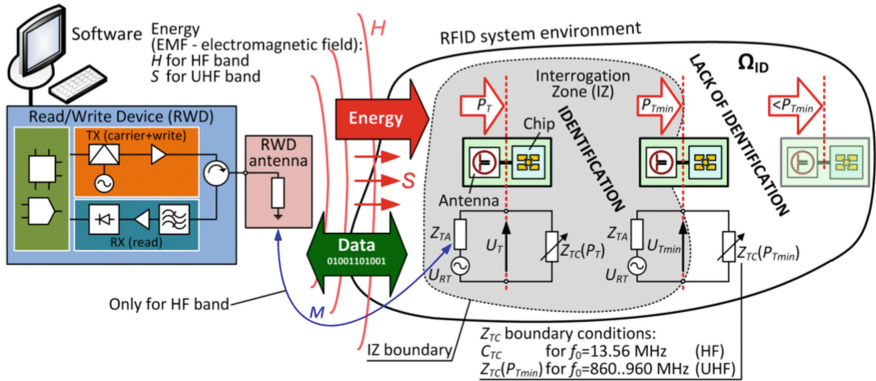
Fig. 1. Impedance model of HF/UHF RFID transponder in RFID system

In the model, transponder blocks are respectively represented by impedance of the antenna (Z_{TA}) and the chip (Z_{TC}). The impedance Z_{TA} has an inductive nature and is constant for a given operation frequency of a transponder. In the serial equivalent, it can be written as:

$$Z_{TA} = R_{TA} + \mathrm{j}X_{TA} = R_{TA} + \mathrm{j}\omega L_{TA} \tag{1}$$

where: R_{TA} denotes the resistance of a transponder antenna, X_{TA} – reactance which can be expressed by the inductance L_{TA}, $\omega = 2\pi f_0$ describes pulsation.

On the other hand, the chip impedance has a capacitive character and it varies with parameters of the electromagnetic field (EMF). In the serial equivalent, it can be described by the following dependence:

$$Z_{TC} = R_{TC} + \mathrm{j}X_{TC} = R_{TC} + \frac{1}{\mathrm{j}\omega C_{TC}} \tag{2}$$

where: R_{TC} denotes the resistance of an active chip, X_{TC} – reactance which can be expressed by the chip capacitance C_{TC}.

The HF RFID systems operate in the space $\mathbf{\Omega_{ID}}$ that is characterized by an inhomogeneous magnetic field (expressed by the magnetic field strength H) and strong coupling (expressed by the mutual inductance M) between RWD and transponder antennas. The inhomogeneous magnetic field generated in the RWD antenna vicinity is a medium for both energy transfer and wireless communications. The minimum

magnetic field strength H_{min} at which the correct data transmission between the RWD and transponder takes place is an EMF boundary condition that characterizes the interrogation zone and read/write range in this inductively coupled RFID system [24].

These conditions are significantly different in the UHF band. A far-field region is used in UHF RFID systems and the radiated wave can be considered locally as a plane wave. In this region vectors of electric and magnetic field strength are perpendicular both to each other and to the direction in which the wave disperses. Since, the radiated electromagnetic wave of power density S is energy medium supplying transponders, the minimum power density S_{min} describes an EMF boundary condition that characterizes the IZ in UHF RFID system.

Despite the differences in the operation principles of HF and UHF RFID systems, the voltage induced in the transponder antenna when it is in the EMF of RWD antenna can be represented by the source U_{RT} in the generalized model. The electromagnetic induction phenomenon determines energy transmission without any galvanic connections and is represented by the voltage U_T that is induced at terminals of the connected antenna as well as by the power P_T received in the transponder. The electronic chip is designed to be supplied by the minimal voltage U_{Tmin} (HF band) or the minimal power P_{Tmin} (chip sensitivity in the UHF band) which are enough for activating internal circuits of the transponder that is located at the IZ boundary. The U_{Tmin} and P_{Tmin} parameters are used to determine characteristic values of the chip impedance (Z_{TC} boundary conditions: C_{TC} for the $f_0 = 13.56$ MHz [25], $Z_{TC}(P_{Tmin})$ for the $f_0 = 860$–960 MHz [26]) that are the first criterion in the antenna designing process.

In the HF RFID systems, the transponder antenna is made in a form of a rectangle [9], a square [10], a circle [11] or another polygon loop [19], which is small in relation to the wavelength λ (because λ is about 22 m). The impedance matching is obtained for the parallel resonance between the inductance L_{TA} and the capacitance C_{TC} of an active chip. Besides L_{TA} inductance and R_{TA} resistance (for the carrier frequency $f_0 = 13.56$ MHz), the database of parameters designated to antenna of RFID transponder operating in the HF band should be completed with quality factor Q_{TA}:

$$Q_{TA} = \frac{\omega L_{TA}}{R_{TA}} \tag{3}$$

In RFID systems of the UHF band, the transponder antennas are made in different shapes and technologies [12–19] and their sizes are always matched to the wavelength. Furthermore, the full impedance matching of antenna and chip is obtained when $Z_{TA} = Z_{TC}^*$ at the chip sensitivity P_{Tmin} (where * indicates the complex conjugate). The quality of this matching is defined on the basis of the power transfer coefficient τ which is defined as follows:

$$\tau = \frac{4\mathrm{Re}(Z_{TA})\mathrm{Re}(Z_{TC})}{\mathrm{Re}(Z_{TA} + Z_{TC})^2 + \mathrm{Im}(Z_{TA} + Z_{TC})^2} \tag{4}$$

The database of impedance parameters for the RFID transponder antennas operating in the UHF band consists of real and imaginary part of the Z_{TA} impedance that is determined for the frequency $f_0 = 860$–960 MHz (depending on world regions).

In both frequency bands, application conditions of RFID systems are the second key determinant for designing antennas – there is no universal RFID transponder, which could be used to mark any objects. Such a transponder should be designed for an object, in view of many conditions of its performance in the RFID system.

3 Validation of Impedance Parameters

3.1 Algorithm

The algorithm (Fig. 2) that is dedicated to validate the impedance parameters of transponder antenna operating in the HF or UHF band is elaborated on the basis of the model presented in Sect. 2. It consists of four steps carried out successively: (1) antenna designing, (2) its manufacturing, (3) measuring of the specified impedance parameters and (4) the parameter experimental verification. The determinants described in the model (Z_{TC} boundary and application conditions of RFID system) are the input data for the first and fourth stage of the algorithm.
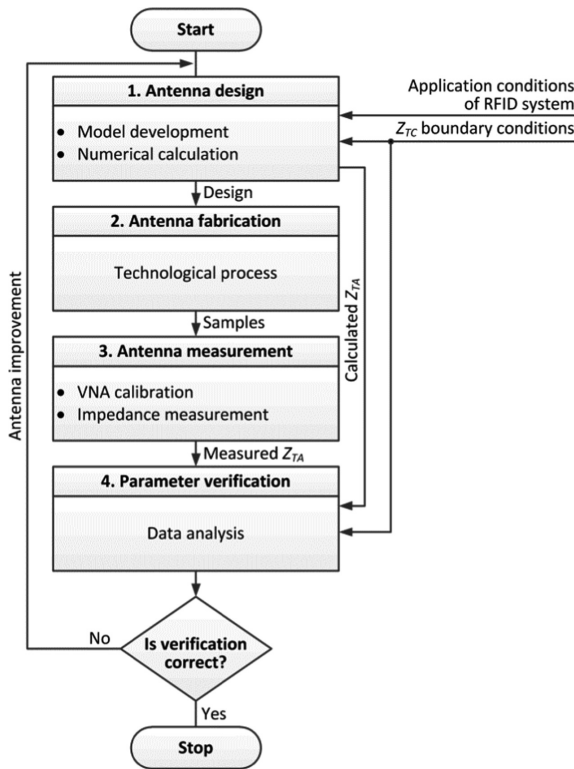
Fig. 2. Alghoritm of proposed validation process

It should be noted that the stages of the antenna design/fabrication and the parameter verification have to be carried out by using advanced know-how, software and technological processes that are available in a R&D laboratory. It involves designers' key abilities essential for method implementation such as the selection of suitable apparatus, the calibration of the used equipment as well as the art of measuring impedance parameters of RFID transponder antennas by using sophisticated devices.

3.2   Antenna Measurement

Model. Methods and procedures that are commonly used for 50/75 Ω antennas operating in typical radio communications systems (such as: TV, GSM, UMTS, LTE, WiFi and many others) cannot be used for measuring parameters of HF or UHF RFID transponders. It is primarily due to the discussed phenomenon of untypical matching of impedance Z_{TA} and Z_{TC} (at the Z_{TC} boundary conditions), and also balanced HF [9–11] and UHF [12–18] antenna designs. Therefore the suitable test procedure for the antenna under test (AUT) consists in realization of indirect differential measurements of impedance parameters by using two 50 Ω coaxial ports (P1, P2) of a vector network analyzer (VNA) [27, 28] and a dedicated passive differential probe (PDP) with the S-S (signal to signal) contact tips [23].
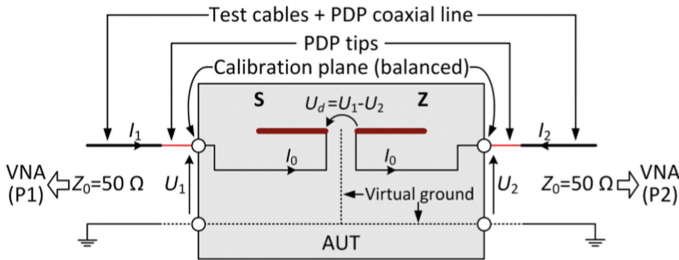
Fig. 3. Network model for AUT measurement

The measuring process can be described by the model (Fig. 3) in which a typical linear 2-port network is characterized by the impedance or scattering matrix (denoted as: **Z**, **S**). The S-parameters measurements does not provide the impedance (1) reading directly from the testing device. Hence it is necessary to determine the differential impedance Z_d and since the AUT current $I_0 = I_1 = -I_2$, it can be calculated from the following dependence:

$$Z_d = \frac{U_d}{I_0} = \frac{U_1 - U_2}{I_0}$$

(5)

where:

$$\begin{cases} U_1 = I_0 Z_{11} - I_0 Z_{12} \\ U_2 = I_0 Z_{21} - I_0 Z_{22} \end{cases} \tag{6}$$

On the basis of (5) and (6), the Z_d depends only on impedance parameters of the modeled network:

$$Z_d = Z_{11} - Z_{12} - Z_{21} + Z_{22} \tag{7}$$

The direct relation between $\mathbf{Z}$ and $\mathbf{S}$ matrix is given by [23].

$$\mathbf{Z} = Z_0(\mathbf{1} - \mathbf{S})^{-1}(\mathbf{1} + \mathbf{S}) \tag{8}$$

where: the system reference impedance Z_0 is 50 Ω and $\mathbf{1}$ denotes the unit matrix.

On the basis of (7) and (8), the differential impedance Z_d is finally given as follows:

$$Z_d = \frac{2Z_0(S_{12}S_{21} - S_{11}S_{22} - S_{12} - S_{21} + 1)}{(1 - S_{11})(1 - S_{22}) - S_{12}S_{21}} \tag{9}$$

Measurement Setup. In relation to the proposed network model, the two coaxial 50 Ω line (VNA/probe test cables, PDP) have to be used in the measurement procedure (Fig. 4).
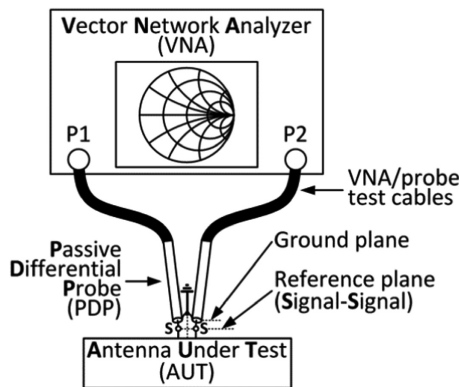
Fig. 4. Block diagram of measurement setup

The AUT should be matched to the chip impedance Z_{TC}. The typical value of chip resistance Re(Z_{TC}) equals from a few to tens of Ω for the chip sensitivity P_{Tmin}. The value of chip reactance Im(Z_{TC}) that is equal typically a few hundred ohms depends mostly on an internal capacitance that accumulates energy necessary for supplying the circuitry of transponder [26]. Therefore thin semi-rigid coaxial cables that extend tips to at least several mm should be used in the impedance measurements of the balanced

antenna. Such a PDP construction constitutes an acceptable compromise which concerns the probe casing impact on impedance measurements and also possibility of reaching the VNA calibration in the reference plane [23, 29, 30].

Calibration Procedure. In most similar cases, the calibration procedure is performed by using open/short port extension methods [15, 17, 29, 30]. It is used to shift the calibration plane to the tips of the PDP. In the first step, standard calibration is conducted at the end of the VNA/probe test cables. Next, the port extension technique (typically open method) is used to shift the calibration plane to probe tips (reference plane). It should be emphasized, that the discrepancies between measuring paths of the PDP significantly influences the accuracy of S parameter measurements and hence the value of the differential impedance Z_d (9). Moreover, the port extension is a less accurate correction of the VNA [31], because it does not remove mismatch effects induced by adding the PDP.

Higher accuracy of conducted measurements can be achieved by attempting to calibrate the both paths of the VNA by using specialized calibration standards that are implemented in the PDP [17, 23]. In this case, the calibration process involves also the PDP characterization as well as using the embed/de-embed features [32]. The problem is, however, that the PDP has no output terminal of ground and also both coaxial lines of this probe are permanently connected to each other. The calibration with the use of prepared short-open-load (SOL) elements that are soldered at the ends of the P1 and P2 lines [17], is not very accurate and leads to destruction of PDPs. On the other hand, it is possible to design a model of the probe in the specialized software [33] but it is time-consuming and often unattainable (e.g. the probe parameters cannot be found in the producers' specifications, the experimental verification of the model is impossible). Bearing in mind the mentioned problems, calibration substrates with short, open, load and thru (SOLT) standards [34, 35] that are usually dedicated for S-parameter calibrations and time-domain reflectometry (TDR) impedance validations are used in the procedure proposed by the authors of the paper. This method does not influence the destruction of PDPs.
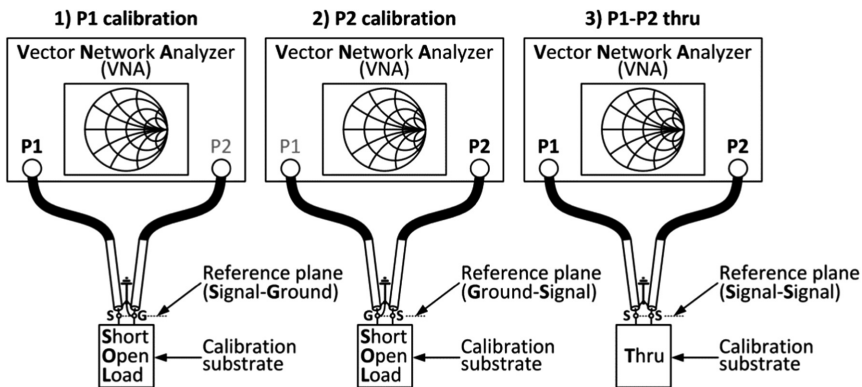
Fig. 5. Block diagram of proposed calibration procedure

In the proposed method, the configuration SG (Signal-Ground) for the P1 calibration and GS (Ground-Signal) for the P2 calibration is alternately created for the PDP (Fig. 5). These configurations are realized by shorting one of the probe tips to the ground. This converts the PDP into a 50 Ω probe. In the first step, the tip of the P2 probe is shorted and then the SOL calibration for P1 is conducted. In the second step, the similar calibration is performed for P2 at the shorted tip of P1 probe. In the third step, the thru calibration is realized for SS (Signal-Signal) connection established between tips of the probes. In effect, this procedure allows to perform calibration and to move the measurement reference point to the PDP tips. The proposed method of removing the PDP errors can be supported by various calibration substrates and VNAs. The details of the method and an example of calibration are presented in Sect. 4.2. After calibration, the **S** scattering matrix is measured and results are used in the impedance calculations on the basis of (9) and (1) equations. These parameters are necessary for designing efficient antennas.

4   Results

4.1   Measurement Setup

The practical usefulness of the developed method has been presented on the basis of two representative examples, successively for the considered HF and UHF frequency bands. The measurement process has been done by using the test stand prepared in the RFID laboratory at the DETS RUT (Fig. 6).
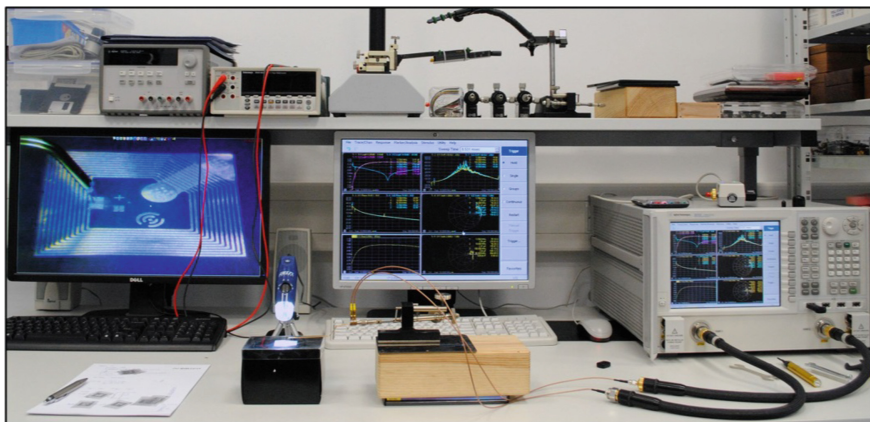
Fig. 6. Test stand – example for HF band

The *S*-parameters of AUT were measured by using the VNA Keysight PNA-X N5242A and commercially available probe (Micromanipulator 44-8000-D-NA) [36]. In the test stand, the probe is mounted on an economical manipulator (Micromanipulator model 110) which is feature with 3-axis direct leadscrew-leadnut drives (resolution 2.2 μm, 10 mm max travel each axis) [36]. The analyzer and probe are connected by: the Keysight 85131F flexible cable set 3.5 mm (VNA test ports) to

3.5 mm (PDP cables), and the manually prepared RG178 UMC (probe connectors) to SMA (VNA flexible cable set).

4.2 Calibration

The test stand was calibrated at the probe tips by using the PacketMicro TCS50 substrate (Fig. 7). This calibration substrate is produced on the polished alumina (size: $17.3 \times 9.4 \times 0.6$ mm). It contains open, short, thru, 50 Ω, and 100 Ω gold contacts (accuracy: <0.5% for the 50 Ω) with GS/SG configuration and probe pitches from 0.2 mm to 1.5 mm [35]. The size of contacts and pitch of the selected probe and calibration substrate is compatible with the typical lead spacing of RFID chips dedicated to the HF as well as UHF bands and provided in various types of packages.
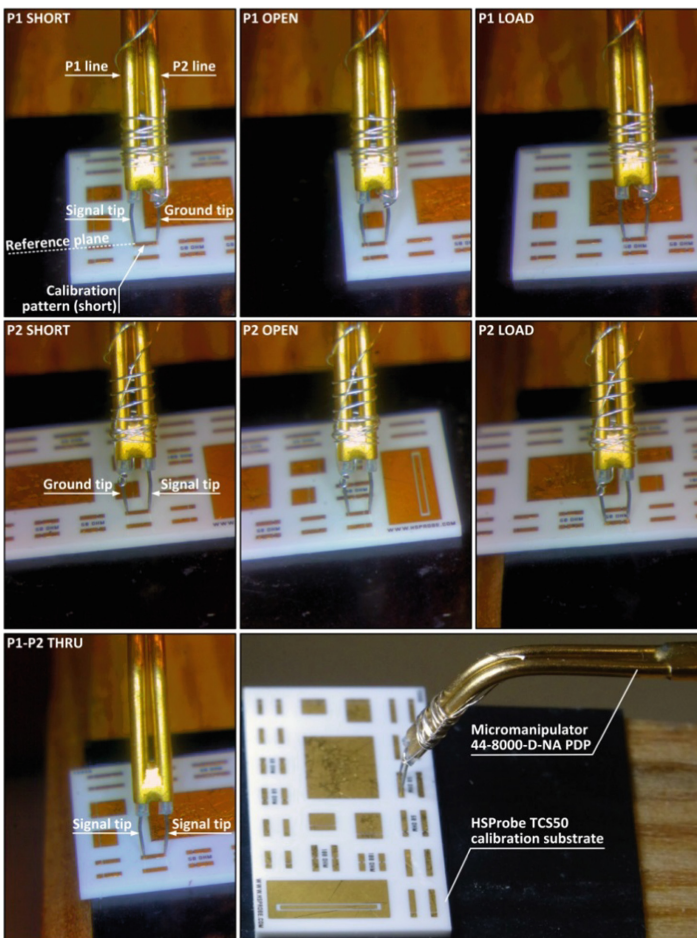
Fig. 7. PDP calibration

The calibration procedure on the laboratory stand is conducted in three steps (Fig. 7) that are discussed in details in Sect. 3.2. In order to meet the requirements assumed for the test stand, the calibration is made in the frequency band from 10 MHz

to 4 GHz (Fig. 8). The configurations SG for P1 and GS for P2 are prepared with using a thin silver plated wire. The shorting between selected signal and ground is realized by wrapping the wire around one of the probe tips and then around the PDP body.
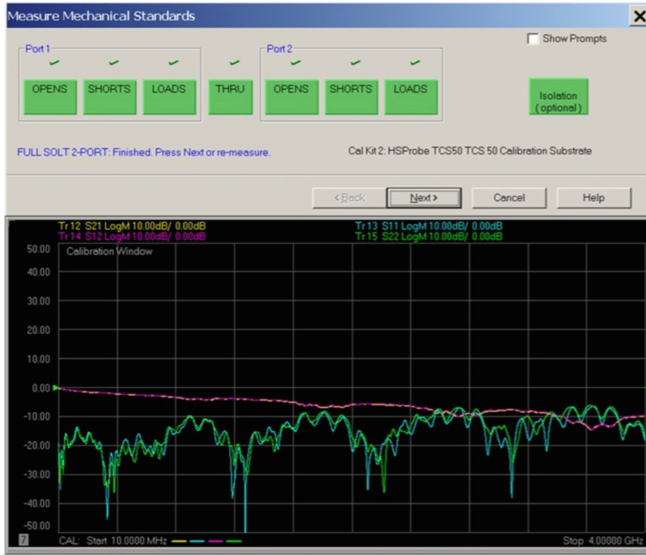
Fig. 8. Full SOLT 2-port calibration results

After this procedure, the P1–P2 thru configuration is represented only as a dot on the Smith chart (Fig. 9). It means that the reference plane is moved to the probe tips. Therefore the VNA calibration allows accurate and repetitive measurements of the differential impedance (9) in the test stand.
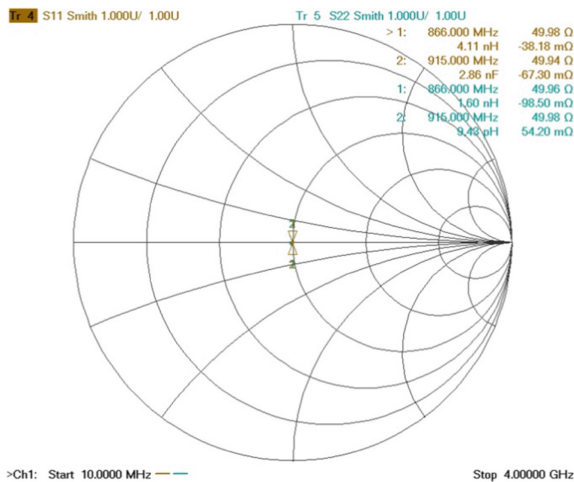
Fig. 9. Smith chart with P1–P2 thru

4.3   Example of HF Transponder Antenna

The stages: "3. Antenna measurement" and "4. Parameter verification" (Fig. 2) are discussed on the basis of two practical examples. The numerical results are compared with data obtained during measurements. The antennas have been designed and simulated by using Mentor Graphics HyperLynx 3D EM (HL3DEM) software.

The antenna of HF transponder designed for Talkin Things Company is analyzed as the first example. It is in the shape of a square of size 19 × 19 mm and is presented in Fig. 10. The impedance matching is provided for the chip NXP NT2H1001, NTAG 210 μ in FFC bump package. This chip complies with the requirements of the following communication protocol: ISO/IEC 14443 Type A, NFC Forum Type 2 (C_{TC}= 50 pF) [37]. The numerical model of the antenna has been elaborated in HL3DEM software tool, in the RFID laboratory at the DETS RUT. The test samples of inlays have been manufactured by selected supplier in China.
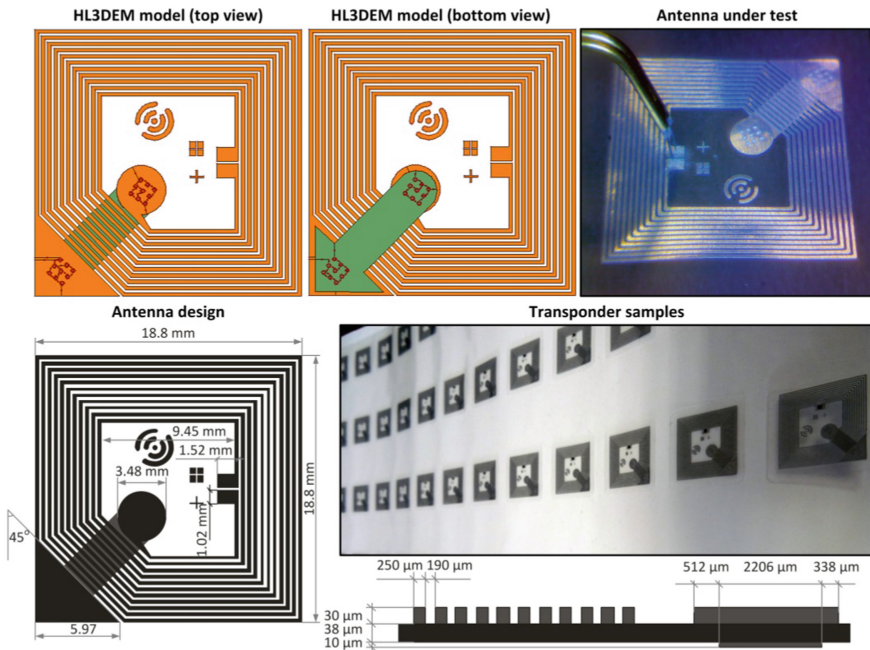
Fig. 10. Talkin Things 19 × 19 mm HF transponder

The model consists of two layers: the upper layer with RFID chip and windings of antenna, and the bottom one with bypass trace of the inductive loop. Moreover, real localization of vias, metal logotype of the company owner and contact pads to which the chip is attached are implemented in the model in order to gain the accuracy of the calculations. The PET film (thickness 38 μm, dielectric constant ε_r= 3.4 and loss tgδ = 0.0026) is considered as a substrate for the modeled antennas.

Constant width of traces (0.25 mm) and trace to trace spacing (0.19 mm) are given into the antenna loop design. The thicknesses of the upper and bottom conductive

layers are assumed to be equal to 30 μm and 10 μm respectively. Due to the lack of detailed information about the technological process defined as aluminum etching, the typical conductivity value ($\sigma = 3.5 \cdot 10^7$ S/m) of such material is assumed in the calculations.

The basic antenna impedance parameters for the given type of the RFID transponders have been calculated with using the prepared numerical model. They have been determined for the carrier frequency $f_0 = 13.56$ MHz. Obtained results are compared with the measurement data that are averaged for 11 gathered samples (Table 1).

Table 1. Calculated and measured parameters of HF transponder antenna

Parameter	HL3DEM calculation results	Measurement results[a]
L_{TA}, μH	2.56	2.55 ± 0.01
R_{TA}, Ω	4.32	5.11 ± 0.61
Q_{TA}, -	50.49	42.52 ± 5.09

[a]Average of 11 samples and expanded uncertainty for a coverage factor $k = 2.28$ and a level of confidence $p \cong 95\%$

The measured average value of the inductance L_{TA} is convergent with the calculations obtained on the basis of the numerical model. The reported inaccuracy of the quality factor is caused by the different values established for the resistance R_{TA}. It is due to the lack of information about the real conductivity of the metal that is used in the technological process of manufacturing the tested antennas. It is suggested that the purity of aluminum alloy that is used to form metal layers in the measured samples is slightly smaller than it is assumed in the model (than $\sigma < 3.5 \cdot 10^7$ S/m).

4.4   Example of UHF Transponder Antenna

The second example concerns an antenna design for a semi-passive RFID sensor transponder operating in the UHF band (Fig. 11). The demonstrator of the circuit has been elaborated in the research/development project under the title "The development of zero-energy quantum system with active packets complex" (POIR.01.01.01-00-0407/16) realized for an industry partner (Aluron) and financed by the government (NCBR). The impedance matching in the design is provided for the chip AMS SL900A in QFN16 package [38] for which the dependence $Z_{TC}(P_{Tmin})$ at $f_0 = 860$–960 MHz is known [26]. The SL900A is an EPC global Class 3 transponder chip which is fully EPC Class 1 compliant (standardized by the ISO/IEC 18000-63) with additional custom commands for extended functions (RFID sensor transponder).

As previously, the numerical model of the antenna has been elaborated in HL3DEM software tool, in the RFID laboratory but the measuring samples of the demonstrator have been manufactured in the laboratory of integrated electronic micro - and nanotechnology HYBRID at the DETS RUT. The process of cutting out metal layers has been realized by using PCB plotter LPKF ProtoMat S100. The modeled elements of the transponder and its antenna are designed on selected epoxy laminate

which is suitable for advanced RF applications (ISOLA FR408: substrate thickness 0.51 mm, copper thickness 35 μm, dielectric constant $\varepsilon_r = 4.19$, loss tangent $tg\delta = 0.0102$ for $f = 1$ GHz) [39].
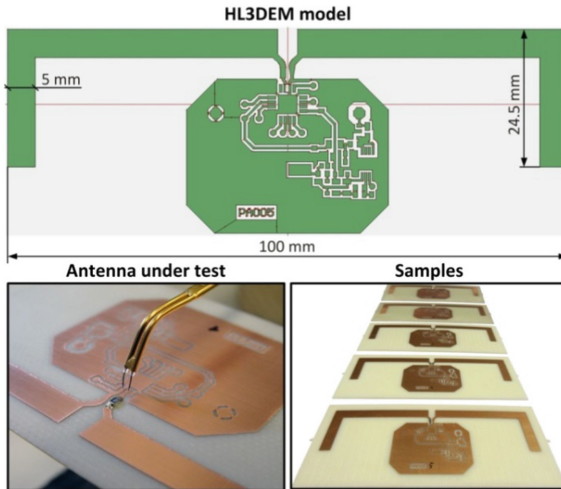
Fig. 11. UHF transponder

The basic antenna impedance parameters for the given type of the RFID transponders have been calculated with using the prepared numerical model. They have been determined in the frequency band 800–1000 MHz. Obtained results are compared with the measurement data that are averaged for 5 gathered samples (Fig. 12).
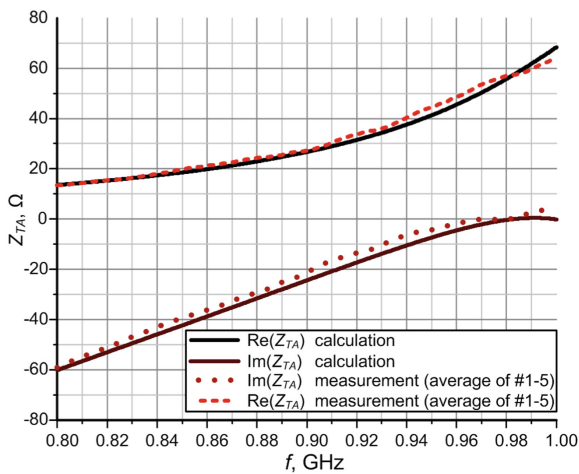
Fig. 12. Calculated and measured parameters of UHF transponder antenna

In this case, the obtained results of the impedance Z_{TA} show a satisfactory convergence (Pearson correlation coefficient is equal 0.938 for real part and 0.995 for imaginary part of impedance Z_{TA}) suitable for analyzing the impedance matching between antenna and selected chip in the whole RFID UHF band (860–960 MHz).

5   Conclusion

At the beginning it should be emphasized that there is no universal RFID transponder which could be used to mark any objects and to work in any standard of the radio frequency identification. Every transponder should be specially designed for an object, in view of many conditions of its performance in the RFID system. Moreover all steps of creating new efficient RFID devices involve using advanced know-how, software tools and technological processes that are available only in R&D laboratories. But designers' abilities (such as selection of suitable apparatus, the calibration of the used equipment, the art of measuring parameters in the time and frequency domain) are the most important for implementing successively research and development procedures in the considered scope.

Since the test methods that are commonly used for 50/75 Ω antennas operating in typical radio communications systems cannot be used for determining parameters of HF or UHF RFID transponders, the authors have proposed algorithm consists in realization of indirect differential measurements of impedance parameters by using two 50 Ω coaxial ports P1 and P2 of a VNA and dedicated PDP with the S-S contact tips. In particular, the question of equipment calibration was taken into consideration in the paper. Bearing in mind the known problems of common methods, the solution that does not interfere in the PDP structure was revealed and the calibration substrates with short, open, load and thru standards that are usually dedicated for *S*-parameter calibrations and TDR impedance validations were used. After calibration, the **S** scattering matrix was measured and results were used in the calculations of the impedance that is necessary for designing efficient antennas.

The details of the method were explained on the basis of two representative examples, successively for the considered HF and UHF frequency bands. The measurement process was done by using the test stand prepared in the RFID laboratory at the DETS RUT and samples were prepared in cooperation with the authors' partners form the industry. The measurement results were compared with data obtained for numerical models in the Mentor Graphics HyperLynx 3D EM. Generally, the measured values were convergent with the calculations despite some discrepancies resulting from restricted information about parameters that are not revealed by producers.

References

1. Plunkett, J.W.: Plunkett's Telecommunications Industry Almanac 2018, 2018th edn. Plunkett Research, Houston, USA (2017)
2. Ustundag, A., Cevikcan, E.: Industry 4.0: managing the digital transformation, 1st edn. Cham, Switzerland (2018)
3. Greengard, S.: The Internet of Things. The MIT Press, London, GB (2015)
4. IDTechEx: continued growth as market for RFID exceeds $10bn milestone. ID World Magazine, 38–39, Dec 2015
5. Das, R., Harrop, P.: RFID forecasts, players and opportunities 2014–2024. Report, IDTechEx (2014)
6. Ustundag, A.: The Value of RFID. Benefits vs. Costs. Springer-Verlag, London (2013). https://doi.org/10.1007/978-1-4471-4345-1
7. Finkenzeller, K.: RFID Handbook – Fundamentals and Applications in Contactless Smart Cards, Radio Frequency Identification and Near-Field Communication, 3rd edn. Wiley (2010)
8. CEPT ERC: ERC recommendation 70-03, Relating to the use of Short Range Devices (SRD). Electronic Communications Committee (2017)
9. Fernández-Salmerón, J., Rivadeneyra, A., Rodríguez, M.A.C., Capitan-Vallvey, L.F., Palma, A.J.: HF RFID tag as humidity sensor: two different approaches. IEEE Sens. J. **15**(10), 5726–5733 (2015). https://doi.org/10.1109/JSEN.2015.2447031
10. Jankowski-Mihułowicz, P., Kalita, W., Skoczylas, M., Węglarski, M.: Modelling and design of HF RFID passive transponders with additional energy harvester. Int. J. Antennas Propag., 1–10 (2013). Article ID 242840. https://doi.org/10.1155/2013/242840
11. Saghlatoon, H., Mirzavand, R., Honari, M.M., Mousavi, P.: Investigation on passive booster for improving magnetic coupling of metal mounted proximity range HF RFIDs. IEEE Trans. Microw. Theory Tech. **65**(9), 3401–3408 (2017). https://doi.org/10.1109/TMTT.2017.2676095
12. Zamora, G., Zuffanelli, S., Aguila, P., Paredes, F., Martin, F., Bonache, J.: Broadband UHF-RFID passive tag based on split-ring resonator (SRR) and T-match network. IEEE Antennas Wirel. Propag. Lett. Accepted for publication (2018). https://doi.org/10.1109/lawp.2018.2800166
13. Jankowski-Mihułowicz, P., Kawalec, D., Węglarski, M.: Antenna design for semi-passive UHF RFID transponder with energy harvester. Radioengineering **24**(3), 722–728 (2015). https://doi.org/10.13164/re.2015.0722
14. Zamora, G., Zuffanelli, S., Paredes, F., Martı, F., Bonache, J.: Design and synthesis methodology for UHF-RFID tags based on the T-match network. IEEE Trans. Microw. Theory Tech. **61**(12), 4090–4098 (2013). https://doi.org/10.1109/tmtt.2013.2287856
15. Lu, Y., Basset, P., Laheurte, J.M.: Performance evaluation of a long-range RFID tag powered by a vibration energy harvester. IEEE Antennas Wirel. Propag. Lett. **16**, 1832–1835 (2017). https://doi.org/10.1109/LAWP.2017.2682419
16. Ramirez, R.A., Rojas-Nastrucci, E.A., Weller, T.M.: UHF RFID tags for On-/Off-metal applications fabricated using additive manufacturing. IEEE Antennas Wirel. Propag. Lett. **16**, 1635–1638 (2017). https://doi.org/10.1109/LAWP.2017.2658599
17. Zhang, Y.J., Wang, D., Tong, M.S.: An adjustable quarter-wavelength meandered dipole antenna with slotted ground for metallically and airily mounted RFID tag. IEEE Trans. Antennas Propag. **65**(6), 2890–2898 (2017). https://doi.org/10.1109/TAP.2017.2690535

18. Sohrab, A.P., Huang, Y., Hussein, M.N., Carter, P.: A hybrid UHF RFID tag robust to host material. IEEE J. Radio Freq. Identif. **1**(2), 163–169 (2017). https://doi.org/10.1109/JRFID.2017.2765623

19. Alibakhshi-Kenari, M., Naser-Moghadasi, M., Sadeghzadeh, R.A., Virdee, B.S., Limiti, E.: Dual-band RFID tag antenna based on the Hilbert-curve fractal for HF and UHF applications. IET Circuits Devices Syst. **10**(2), 140–146 (2016). https://doi.org/10.1049/iet-cds.2015.0221

20. Jankowski-Mihułowicz, P., Węglarski, M: Definition, characteristics and determining parameters of antennas in terms of synthesizing the interrogation zone in RFID systems. In: Crepaldi, P.C., Pimenta T.C. (eds.) Radio Frequency Identification, Chapter 5, pp. 65–119. INTECH, Rijeka, Croatia (2017). https://doi.org/10.5772/intechopen.71378

21. Taoufik, S., Dherbécourt, P., El Oualkadi, A., Temcamani, F.: Reliability and failure analysis of UHF RFID passive tags under thermal storage. IEEE Trans. Device Mater. Reliab. **17**(3), 531–538 (2017). https://doi.org/10.1109/TDMR.2017.2733519

22. Bauernfeind, T., Renhart, W., Alotto, P., Bíró, O.: UHF RFID antenna impedance characterization: numerical simulation of interconnection effects on the antenna impedance. IEEE Trans. Magn. **53**(6), 1–4 (2017). https://doi.org/10.1109/TMAG.2017.2655883

23. Jankowski-Mihułowicz, P., Pitera, G., Węglarski, M.: The impedance measurement problem in antennas for RFID technique. Metrol. Meas. Syst. **21**(3), 509–520 (2014). https://doi.org/10.2478/mms-2014-0043

24. Jankowski-Mihułowicz, P.: Field conditions of interrogation zone in anticollision radio frequency identification systems with inductive coupling. In: Turcu, C. (eds.) Radio Frequency Identification Fundamentals and Applications Bringing Research to Practice, Chapter 1, pp. 1–26. INTECH, Rijeka, Croatia (2010)

25. Rizkalla, S., Prestros, R., Mecklenbräuker, C.F.: De-embedding transformer-based method for characterizing the chip of HF RFID cards. In: IEEE Wireless Power Transfer Conference (WPTC), pp. 1–4, Taipei (2017). https://doi.org/10.1109/wpt.2017.7953814

26. Jankowski-Mihułowicz, P., Węglarski, M.: Determination of passive and semi-passive chip parameters required for synthesis of interrogation zone in UHF RFID systems. Elektronika ir Elektrotechnika **20**(9), 65–73 (2014). https://doi.org/10.5755/j01.eee.20.9.5007

27. Meys, R., Janssens, F.: Measuring the impedance of balanced antennas by an S-parameter method. IEEE Antennas Propag. Mag. **40**(6), 62–65 (1998). https://doi.org/10.1109/74.739191

28. Peruzzi, M., Masson, F., Mandolesi, P., Perotoni, M.: Technique for measurement of UHF RFID balanced antennas. Electron. Lett. **54**(2), 59–60 (2018). https://doi.org/10.1049/el.2017.3989

29. Qing, X., Goh, C.K., Chen, Z.N.: Impedance characterization of RFID tag antennas and application in tag co-design. IEEE Trans. Microw. Theory Tech. **57**(5), 1268–1274 (2009). https://doi.org/10.1109/TMTT.2009.2017288

30. Cai, C., Hong, W., Deng, L., Li, S.: Impedance measurement of RFID tag antenna based on different methods. In: IEEE 5th International Symposium on Electromagnetic Compatibility, pp. 1–4, Beijing (2017). https://doi.org/10.1109/emc-b.2017.8260455

31. Agilent Technologies: Advanced Calibration Techniques for Vector Network Analyzers. Modern Measurement Techniques for Testing Advanced Military Communications and Radars. 2nd edn. Agilent Technologies (2006)

32. Keysight Technologies: De-Embedding and Embedding S-Parameter Networks Using a Vector Network Analyzer. Application Note, 5980-2784EN, Keysight Technologies, USA (2017)

33. Wang, Q., Gao, Y., Fan, J., Drewniak, J., Zai, R.: Differential probe characterization. In: IEEE International Symposium on Electromagnetic Compatibility (EMC), pp. 780–785, Ottawa (2016). https://doi.org/10.1109/isemc.2016.7571748
34. Cascade Microtech: Probe Selection Guide – Impedance Standard Substrate (ISS). ProbeGuide-1017, Cascade Microtech (2017)
35. PacketMicro: TCS50 Calibration Substrate, S-Parameter Calibration and TDR Impedance Validation, Santa Clara, USA (2017)
36. Micromanipulator: Probe Tips and Probe Holders. Reference Manual, Carson City, Nevada, USA (2013)
37. NXP: NT2L1001_NT2H1001, NTAG 210μ, NFC Forum Type 2 Tag compliant IC with 48 bytes user memory. Product data sheet, Rev. 3.0, 343930 (2016)
38. AMS: SL900A EPC Class 3 Sensory Tag Chip – For Automatic Data Logging. Datasheet, v1-07 (2016)
39. Jankowski-Mihułowicz, P., Lichoń, W., Pitera, G., Węglarski, M.: Determination of the material relative permittivity in the UHF band by using T and modified ring resonators. Int. J. Electron. Telecommun. **62**(2), 129–134 (2016). https://doi.org/10.1515/eletel-2016-0017

Efficiency Problem of FMCG Identification in HF RFID System with Multiplexed Antennas for Commercial Refrigerator

Piotr Jankowski-Mihułowicz[1(✉)] , Mariusz Węglarski[1],
Wojciech Lichoń[2], and Marcin Pilarz[2]

[1] Department of Electronic and Telecommunications Systems,
Rzeszow University of Technology, Pola 2, 35-959 Rzeszow, Poland
{pjanko,wmar}@prz.edu.pl
[2] Talkin Things, Pulawska 182, 02-670 Warsaw, Poland
{wojciech.lichon,marcin.pilarz}@talkinthings.com

Abstract. The overriding goal of the research discussed in the article is to propose an RFID system with the set of multiplexed RWD antennas that could be implemented in a commercial refrigerator. The need to solve this type of a problem has been reported by the one of the authors' partner from the industry. The main assumptions for designing the demonstration system consist in the possibility of obtaining 100% efficiency for the identification of electronically marked bottles with beverages as well as lack of significant interferences in the construction of the refrigerator that could disturb the operating and maintaining conditions. According to these requirements and meeting the partner's needs and expectations, several configurations of the set of multiplexed RWD antennas have been designed, simulated and examined on the experimental stand in the authors' laboratory. On the basis of obtained results the conclusions have been drawn regarding the construction of the RFID system, its implementation in the commercial refrigerators and suggestions for structural changes in the devices at manufacturing process. In addition, the usefulness of the proposed system should be considered in a broader context since the transponders applied to the bottles could be used in the whole life cycle of a FMCG product (production, logistics, marketing services, recycling and others).

Keywords: RFID · Read/write devices · Multiplexed antennas
Refrigerator

1 Introduction

Automation of object identification processes is an important factor that leads to improvements of economic activity in the logistics, industry, trade, transport and other various socio-economic areas. These changes are especially visible in the scope of intelligent systems, which operation is covered by the current trends in Internet of Things (IoT) or automation and data exchange in manufacturing technologies (Industry 4.0) [1, 2]. Nowadays, the radio frequency identification (RFID) technology is more and more often used in this kind of advanced developments [3] and its usability is

confirmed by growing number of practical implementations [4, 5]. It results from availability of system devices on the market and also forecast in terms of their applicability within the next years [6]. On the other hand, better recognition of the essence pertinent to the operation of these devices as well as methods of determining their parameters confirm the statement in technological terms [7].

The problem presented in this paper is inscribed in the conception of IoT that is dedicated for the fast moving consumer goods (FMCG). According to this idea, commonly used bar codes will be replaced by RFID transponders in the nearest future, and the electronic product code (EPC) defined by the EPCglobal Tag Data Standard is dedicated for these aims [8]. The development works are conducted to ensure that automatic identification will be effectively and smoothly applied to the FMGC in the supply chain management [9].

The need to conduct the presented research and development works has been announced by one of the authors' partners active in the FMCG industry. The problem has been connected with selecting appropriate RFID transponders operating in various frequency bands (HF or UHF) as well as practical methods of dealing with such kind of tasks. It is especially important in the scope of automation effectiveness when a whole life cycle of a FMCG product is taken into consideration (production, logistics, marketing services, recycling and others). The works related to solving this issue have been divided into stages and have been realized under the cooperation between FMCG industry partners as well as the Department of Electronic and Telecommunications Systems (DETS) in Rzeszow University of Technology (RUT) and Talkin Things Company. Although both HF and UHF systems can be implemented in the solution under development, the first one is used in the elaborated demonstrator due to the requirements set by the industrial partner.

The overriding goal of the stage discussed in this article is to propose a set of multiplexed RWD antennas dedicated to a commercial refrigerator. The possibility of obtaining 100% efficiency for identification of electronically marked bottles with beverages is assumed for the demonstration RFID system. A remote automatic inventory of the objects inside the refrigerator can be performed by means of such a system in the nearest future. Further, sales reports and orders can be generated and directed to a distribution center of desired FMCG products.

The problem of RFID system implementation in refrigerators can be encountered in the subject literature. It is considered for example in aspects of specialized medical implementations (location based blood bag management [10], medication monitoring [11, 12], and other intelligent systems with refrigerator [13]) but also some publications refer to industrial applications (e.g. traceability system of milk samples [14]) as well as to household equipment (e.g. IoT refrigerator at smart home [15–20]). Most of these cases, however, concern the manual registration of electronically marked products in an RFID system (the single RFID system). The solution presented by the authors is unheard of in the subject technology. It is based on an automatic recognition of a large number of electronically marked objects (the anti-collision RFID system) that additionally can be located in a large space. The authors' concept of determining the three-dimensional interrogation zone (IZ) by using the Monte Carlo method in inductively coupled anti-collision RFID systems is applied in the solution [21].

2   HF RFID System with Multiplexed Antennas

A typical single or anti-collision RFID system consists of transponder(s) and read/write device (RWD) with its antenna(s) [3]. The RWD connected to a computer and their software form a management centre whereas the electronic transponders are used for marking objects.

The HF RFID systems (the carrier frequency of $f_0 = 13.56$ MHz) operate in the space Ω_{ID} that is characterized by an inhomogeneous magnetic field (expressed by the magnetic field strength H) and strong coupling (expressed by the mutual inductance M) between antennas of the arrangement component (Fig. 1). The inhomogeneous magnetic field generated in the RWD antenna vicinity is a medium for both energy transfer and wireless communication. The communication mechanisms in the HF band are implemented in adequate protocols (typically ISO/IEC 14443, 15693, 18000-3).
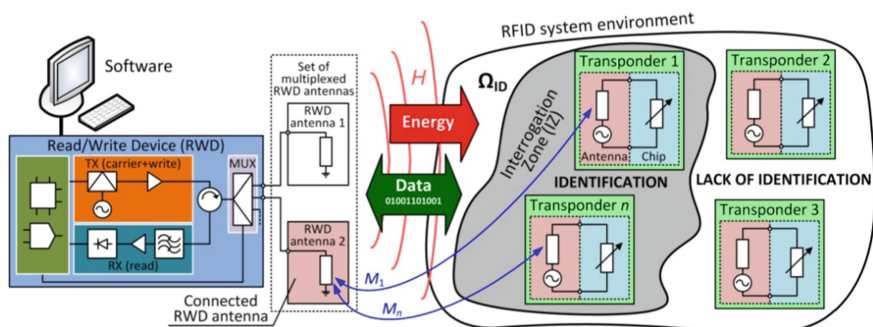
Fig. 1.   Anti-collision HF RFID system with multiplexed antennas

The communication process can be carried out only when transponders are in the interrogation zone. It means that the operational capability of RFID systems is characterized by the IZ, which covers problems of energy and communication activity with regards to a specified hardware and software configuration. It also determines and comprehensively describes possibilities of an RFID system application in desirable automated processes.

The effectiveness of the anti-collision RFID system is defined on the basis of the IZ synthesis for a group of n-transponders which are located in the space Ω_{ID}, and for a given automatic identification process (for example objects that are localized in a commercial refrigerator). The interrogation zone is determined for an assumed efficiency of identification η_{ID}:

$$\eta_{ID} = \frac{l_{IDOK}}{n} \cdot 100\% \tag{1}$$

where l_{IDOK} is the number of transponders for which desired read/write operations are properly executed.

From a practical point of view, the IZ should be as large as possible regardless of the variable location/orientation of activated transponders. However in the RFID system with only one RWD antenna, it is not possible to identify all transponders at any point in the space Ω_{ID}. The problem stems from the fact that energy can be conveyed from the read/write device to transponders only on a limited distance [22]. The energy is not transferred (like in a far-field region) but stored in an inhomogeneous magnetic field (near-field region). At the present stage of knowledge, it is the main reason why practical implementations of anti-collision identification are restrained.

In the HF RFID systems, the wavelength λ is about 22 m for the carrier frequency of $f_0 = 13.56$ MHz. Therefore, the RWD antennas are made in a form of loop (typically as a square [23], rectangle [24], circle [25] or other polygons [26]), which is small in relation to λ. The primary way to enlarge the size of interrogation zone consists in conveying more energy to transponders that are freely deployed and oriented in the space Ω_{ID}. The greatest flexibility in developing RFID applications and shaping the IZ space can be achieved by using the system in which a multiplexer (MUX) and more than one RWD antenna is applied [21]. So, the different configurations of multiplexed RWD antennas are considered in this paper in order to present the described idea of FMCG identification in a commercial refrigerator.

3 Research Problem

The demonstration RFID system under test is build on the base of glass door merchandiser (single swing glass door refrigerator with hydrocarbon refrigerant and LED lighting from the True Food Service Equipment, model GDM-26-LD [27]). This kind of device is widely used in bars, restaurants, retail shops and markets (Fig. 2).

Fig. 2. GDM-26-LD
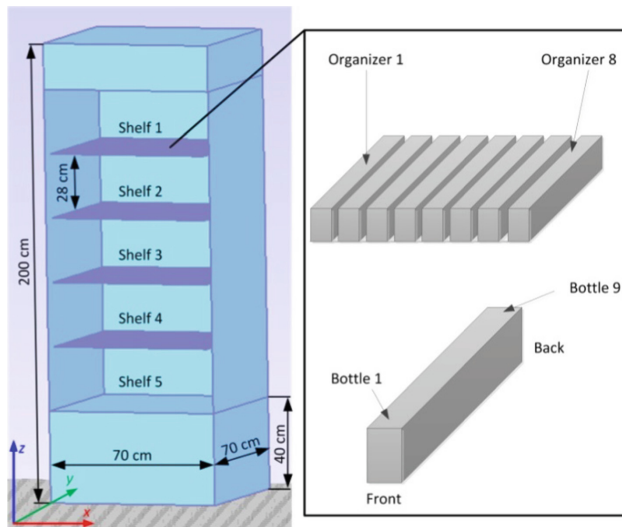
Fig. 3. Refrigerator model with element numbering

The inside of this commercial refrigerator is made of metal: the side walls and the upper part – white-painted aluminum sheet, the bottom – stainless steel sheet. The casing is made of plastic (laminated vinyl), the appearance of which can be adapted to the customer's requirements. The space between the internal and external walls is filled with a polyurethane foam seal (Ecomate). The tested model of the refrigerator has 5 shelves manufactured from metal wire covered with a layer of white plastic.

A selected batch of 360 bottles is the exemplary research object in the complete RFID identification process on the demonstration stand. Each of the 500 ml bottles is factory-filled with liquid. They are placed in the same orientation in plastic gravity feed organizers that are dedicated equipment of GDM-26-LD. It is possible to place maximum of eight 9-bottle cassettes on one shelf, hence, for the whole refrigerator, it gives the number of 360 objects intended for the process of anti-collision RFID identification (Fig. 3).

All of the bottles in the demonstration system under consideration are marked with HF RFID transponders. The NXP I-CODE SLIX chip that is used in the selected class of electronic tags (model SL2 S2002/SL2 S2102) operates in accordance with the ISO/IEC 15693 protocol (Fig. 4). The transponders are placed in the middle of the bottle labels. Extra paper stickers with information about location in the refrigerator as well as with the last 3 bytes of the unique identifier (UID of RFID transponders) are attached to all of the bottles. The coordinates of an object in the space are described by the shelf number (P1…5), the cassette number (K1…8) and the position of a bottle in it (B1…9). The use of these stickers provides the opportunity to determine the location of unidentified bottles during the experimental research. It should be emphasized, that other standards of RFID systems could be used in the glass door merchandiser with regards to requirements of a target application or a desired period of life cycle of FMCG products. Semi-passive or passive transponders operating with ISO/IEC 14443 protocol of the HF band or ISO/IEC 18000-63 for the UHF band as well as dual-band HF + UHF models (Fig. 5) and other contemporary RFID devices are considered in other research projects conducted in cooperation with authors' partners from the industry.

The realized preparatory activities constitute the basis for conducting research and development works. Furthermore, the method developed in [21] is implemented in application experiments what confirms its usefulness in this scope. Various configurations of multiplexed RWD antennas dedicated to RFID systems of the HF band are subjects of the conducted analysis. The designs that are the most perspective from the practical point of view are presented in Sect 4. Even though, they are considered for a representative case of one shelf, the system for entire refrigerator can be elaborated by

Fig. 4. Selected organizer with electronically marked bottles
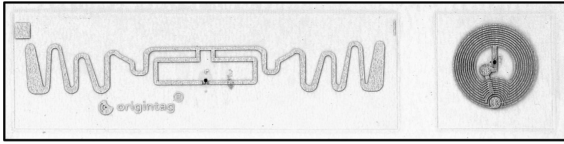
Fig. 5. POIR.01.01.01-00-0226/15 hybrid RFID transponder

increasing number of RWD multiplexer inputs and duplicating the solution for sub-sequent shelves. The numerical model EMCoS Antenna Virtual Lab is prepared for each configuration. Calculations of the current distribution and the inhomogeneous magnetic field are made using the TriD solver based on the Method of Moments (MoM). The necessity to carry out such extensive preparatory and calculation steps was forced by the diversity of materials from which the refrigerator is made. Moreover, the potential modification of the refrigerator structure is analyzed in the research works, because the components of RFID systems are considered to be possible for future integration with construction of the device. The positive results of numerical calculations yield the design brief for preparing and assembling the set of multiplexed RWD antennas, and then for conducting tests in a demonstration RFID system. Although demonstration installation was tested for 100% identification efficiency, the proper configuration set of multiplexed RWD antennas that would enable periodic removal of the shelves was also searched for. This is caused by additional hygienic-sanitary requirements that have to be met during the operational maintenance of the refrigerator inside. Equipment and software of Feig company were used in this stage of the works.

4   Results

4.1   Gate Type of Multiplexed RWD Antennas

Model. According to the specified assumptions, the space between shelves of the glass door merchandiser is the most prospective location for the set of multiplexed RWD antennas. The basic conception consists of four opposite antennas that form two orthogonal gates (Fig. 6). In such a configuration, each pair of antennas is dedicated to identify bottles with selected orientation in which the central plane of label is set in parallel to the loop arrangement. Since there is no need to interfere in the construction of the commercial refrigerator when installing the antennas, it should be considered as a crucial advantage for the proposed solution. It should be also noted that the possibility of removing the shelves during the devices operation is retained.

Two numerical models (for Gate 1: antenna 1 and 2, for Gate 2: antenna 3 and 4) are considered for the proposed concept. In the first case (Gate 1), the loops are moved 10 cm from the door and from the rear wall of the refrigerator whereas in the second case (Gate 2 v.1) there is the 1 cm separation of antennas from metal walls. The components of the arrangement are modeled as loops made of a coaxial cable with a core diameter of 0.3 mm to which a 50 Ω current source with amplitude of 1 A is connected.
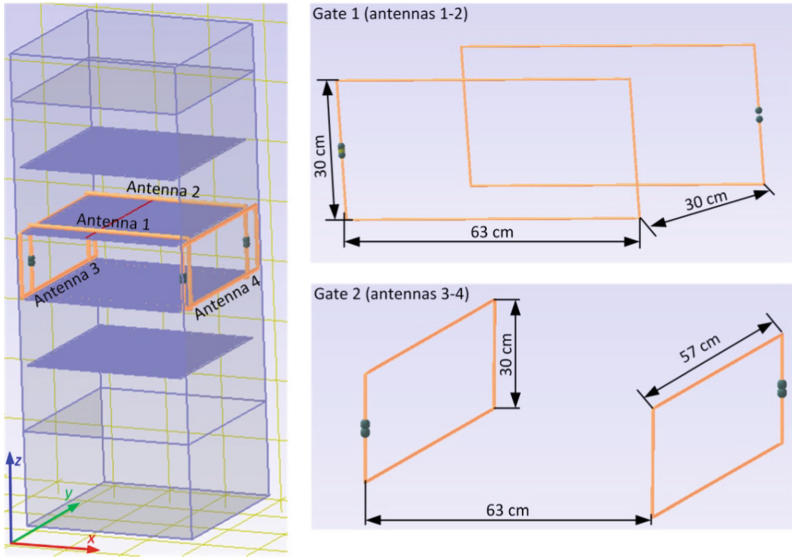
Fig. 6. Model of gate type antennas

Numerical Calculation. In the next step, the magnetic field strength distribution inside the refrigerator is determined with respect to the x-z plane for the Gate 1 (Fig. 7) and y-z for the Gate 2 (Fig. 8). The locations where magnetic field strength may be sufficient to supply passive RFID transponders properly ($H > H_{min}$) are estimated on the basis of the obtained results. It should be noted that the data are analyzed qualitatively. Such an assessment is caused by a significant simplification of the refrigerator numerical model. But, it allows for efficient calculation and justifies decision to move on to the experiment phase in the demonstration RFID system.
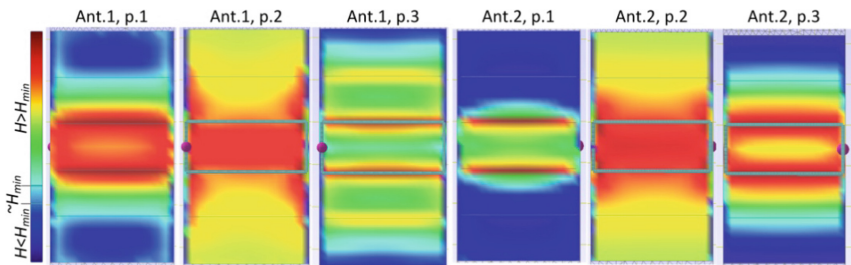
Fig. 7. Total magnetic field strength (x-z plane, Gate 1, p. 1–5 cm from the door, p. 2 – middle, p. 3–5 cm from the rear wall)
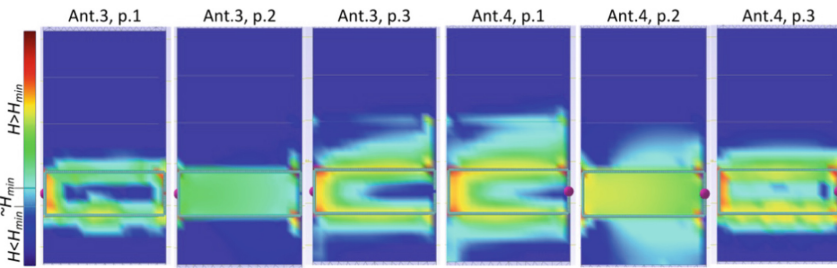
Fig. 8. Total magnetic field strength (*y-z* plane, Gate 2 v.1, p. 1–5 cm from the left side, p. 2 – middle, p. 3–5 cm from the right side)
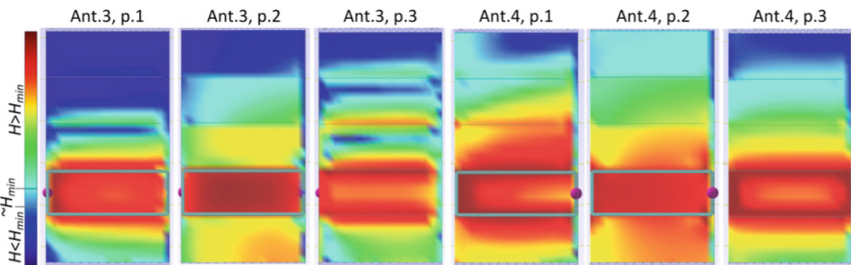
Fig. 9. Total magnetic field strength (*y-z* plane, Gate 2 v.2, p. 1–5 cm from the left side, p. 2 – middle, p. 3–5 cm from the right side)

The magnetic field intensity distribution is determined for the carrier frequency of $f_0 = 13.56$ MHz, and the calculations are carried out for selected planes related to the door and side walls of the refrigerator. For the antennas located parallel to the door (Fig. 7), the magnetic field strength should be sufficient to supply passive RFID transponders. In the case of the second pair (Fig. 8), eddy currents are generated in the side walls which make it impossible to implement an object identification process correctly – the value of magnetic field intensity is lower than H_{min} for most locations in the considered space.

The identification of objects oriented perpendicular to the door can be implemented in two other ways that should lead to improvement the second case. The first one consists in increasing the separation of antennas 3 and 4 from the metal walls of the refrigerator. As a result, the induction of eddy currents is significantly lower. Hence, it reduces the impact of this phenomenon on impedance parameters of the antennas. The limitation of the exhibition space and the possibility of damaging antennas during maintaining the device (e.g. when pushing/pulling out a bottle with a drink) are significant disadvantages of the proposed solution. The second way involves the necessity to make changes in the construction of the glass door merchandiser. For example, if the metal side walls were replaced with plastic sheets, it would resolve all the problems with the antenna disruption issue.

New numerical models for improved constructions of the Gate 2 are proposed to fulfill above mentioned remarks. The separation of 10 cm from metal walls of the refrigerator to the 3rd and 4th antenna is taken into account in the Gate 2 v.2 version
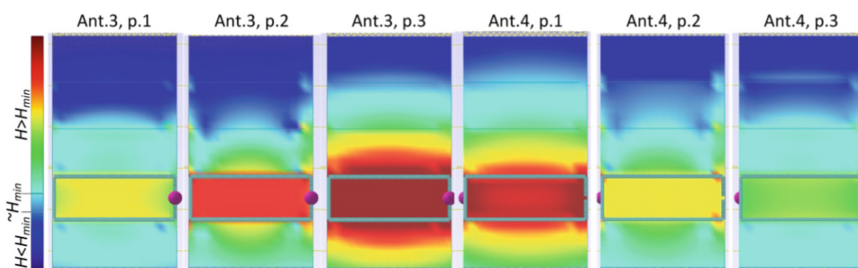
Fig. 10. Total magnetic field strength (*y-z* plane, Gate 2 v.3, p. 1–5 cm from the left side, p. 2 – middle, p. 3–5 cm from the right side)

(Fig. 9). It can be noticed when comparing obtained results with the Gate 2 v.1 (previously described as Gate 2) case (1 cm separation – Fig. 8) that the space – in which transponders are supplying with sufficient power – is enlarged significantly. The next modification consists in changing the glass door merchandiser construction and is described as the third version of the Gate 2 model (Gate 2 v.3). A new type of materials that are commonly used to build side walls of some commercial refrigerators is introduced in the model (Fig. 10). Plastic planes with a thickness of 1 cm and relative permittivity $\varepsilon_r = 2.3$ are analyzed in this case. Although the antennas are placed as in the Gate 2 v.1 model (Fig. 8), it can be noticed that the dimensions of the space where the passive RFID transponders are powered properly is also increased significantly. Hence, the modification of the glass door merchandiser construction is the best choice in the discussed implementation of the RWD antenna set. It guarantees the possibility of identifying bottles in any location and orientation.

Experiment. The experimental verification of the multiple identification process was the goal of this research stage. It was carried out in the prepared demonstration system based on the commercial glass door merchandiser. A set of multiplexed RWD antennas were installed in the refrigerator according to the model Gate 1/Gate 2 v.1. The experimental works were divided into two steps. In the first one, the arrangement of the 1st and 2nd antenna – Gate 1 that is dedicated to identify objects with labels oriented towards the door was subjected to the verification. In the second step, 3rd and 4th antenna – Gate 2 that is responsible for identifying orthogonally oriented bottles was tested.

The antenna loops are made of coaxial cable RG-58 in plastic cable raceways constituting supporting construction (Fig. 11). The vertical wires are run parallel to metal shelf mounting rails while the horizontal wires – above and below surfaces of the shelves. In this way, unused parts of the internal space are occupied while the exhibition area in the refrigerator is not reduced at all. The terminals of the each loop are attached to a dynamic antenna tuner (FEIG ID ISC.DAT). This device with automated matching circuits is designed for tuning RFID antennas with an operating frequency of 13.56 MHz. It enables to match antennas to impedance of 50 Ω with the use of dedicated control software tool. All matching circuits are connected by coaxial cables to the selected ports of the 8-channel multiplexer (FEIG ID ISC.ANT.MUX). This device also provides power and control commands to the dynamic antenna tuner.
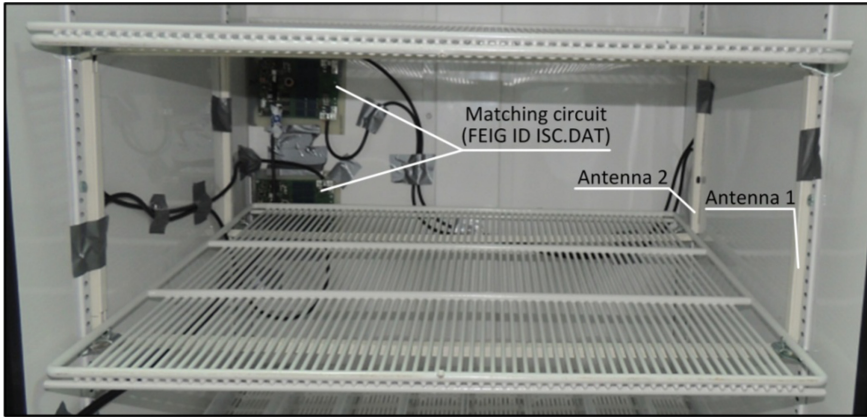
Fig. 11. Gate 1 of multiplexed RWD antennas

The multiplexer is connected to the antenna port in the ISO/IEC 15693 long range RWD module (FEIG ID ISC.LRM2500-B). The RWD module provides output power equal 4 W and is controlled by ID ISOStart V09.09.10 software. The read/write device is set in Host Mode during the measurement procedure, what enables to send orders manually.

RFID transponders attached to the bottles are directed towards the door (Fig. 12a) or a side wall of the refrigerator (Fig. 12b). Conducted experimental tests allow to verify the RFID system's ability to identify electronically marked objects. The procedure consisted in implementation of the inventory round as well as the localization of the recognized objects on the basis of UID. The obtained results are summarized in Tables 1, 2, 3, 4, 5 and 6. The OK means UID recognition and NOK – lack of object identification in the inventory round.

The case of RFID transponders directed towards the door is verified in the first step (Tables 1, 2 and 3). The inventory round is carried out during the tests separately for the each antenna of the Gate 1 and results are presented in Table 1 (for the 1st antenna) and 2 (for the 2nd antenna). The sum of the cases from the switched antennas is collected in Table 3. It can be seen that all objects on the test shelf are identified. In addition, partial areas of correct identification obtained from the individual antennas overlap themselves. It proves the reliability of the proposed solution.

Fig. 12. Electronically marked bottles with transponders directed towards: (a) door, (b) side walls of the refrigerator

Table 1. Results for labels oriented towards the refrigerator door (antenna 1)

	K1	K2	K3	K4	K5	K6	K7	K8
B1	OK	OK	OK	OK	OK	OK	OK	OK
B2	OK	OK	OK	OK	OK	OK	OK	OK
B3	OK	OK	OK	OK	OK	OK	OK	OK
B4	OK	OK	OK	OK	OK	OK	OK	OK
B5	OK	OK	OK	OK	OK	OK	OK	OK
B6	OK	OK	OK	OK	OK	OK	OK	OK
B7	OK	OK	OK	OK	OK	OK	OK	OK
B8	OK	OK	OK	OK	OK	OK	OK	OK
B9	OK	OK	OK	NOK	NOK	OK	OK	OK

Table 2. Results for labels oriented towards the refrigerator door (antenna 2)

	K1	K2	K3	K4	K5	K6	K7	K8
B1	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B2	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B3	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B4	NOK	OK	NOK	OK	NOK	OK	OK	OK
B5	OK	OK	NOK	OK	NOK	OK	OK	OK
B6	OK	OK	OK	OK	OK	OK	OK	OK
B7	OK	OK	OK	OK	OK	OK	OK	OK
B8	OK	OK	OK	OK	OK	OK	OK	OK
B9	OK	OK	OK	OK	OK	OK	OK	OK

Table 3. Summary results for labels oriented towards the refrigerator door (Gate 1)

	K1	K2	K3	K4	K5	K6	K7	K8
B1	OK	OK	OK	OK	OK	OK	OK	OK
B2	OK	OK	OK	OK	OK	OK	OK	OK
B3	OK	OK	OK	OK	OK	OK	OK	OK
B4	OK	OK	OK	OK	OK	OK	OK	OK
B5	OK	OK	OK	OK	OK	OK	OK	OK
B6	OK	OK	OK	OK	OK	OK	OK	OK
B7	OK	OK	OK	OK	OK	OK	OK	OK
B8	OK	OK	OK	OK	OK	OK	OK	OK
B9	OK	OK	OK	OK	OK	OK	OK	OK

The operation of the set of multiplexed RWD antenna (in the range of the Gate 1) is verified in a similar way for the second spatial orientation of bottles. In this case RFID transponders are located towards the side wall of the refrigerator. As before, the inventory round is carried out separately for the each antenna (Table 4 and 5). The results obtained for this orientation are much worse than in the previous study. This is due to the fact that energy of the magnetic field produced by the RWD antennas is less

Table 4. Results for labels oriented towards the side wall (antenna 1)

	K1	K2	K3	K4	K5	K6	K7	K8
B1	NOK	NOK	OK	OK	OK	OK	OK	OK
B2	OK	NOK	OK	NOK	NOK	NOK	NOK	NOK
B3	NOK	OK	OK	NOK	OK	NOK	NOK	NOK
B4	NOK	NOK	NOK	NOK	NOK	OK	OK	OK
B5	NOK	NOK	OK	NOK	OK	NOK	NOK	NOK
B6	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B7	OK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B8	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B9	OK	NOK	NOK	NOK	NOK	NOK	NOK	NOK

Table 5. Results for labels oriented towards the side wall (antenna 2)

	K1	K2	K3	K4	K5	K6	K7	K8
B1	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B2	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B3	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B4	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B5	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B6	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B7	NOK	NOK	OK	NOK	NOK	NOK	NOK	NOK
B8	NOK	NOK	NOK	OK	NOK	OK	OK	OK
B9	OK	OK	NOK	NOK	NOK	NOK	OK	NOK

Table 6. Summary results for labels oriented towards the side wall (Gate 1)

	K1	K2	K3	K4	K5	K6	K7	K8
B1	NOK	NOK	OK	OK	OK	OK	OK	OK
B2	OK	NOK	OK	NOK	NOK	NOK	NOK	NOK
B3	NOK	OK	OK	NOK	OK	NOK	NOK	NOK
B4	NOK	NOK	NOK	NOK	NOK	OK	OK	OK
B5	NOK	NOK	OK	NOK	OK	NOK	NOK	NOK
B6	NOK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B7	OK	NOK	OK	NOK	NOK	NOK	NOK	NOK
B8	NOK	NOK	NOK	OK	NOK	OK	OK	OK
B9	OK	OK	NOK	NOK	NOK	NOK	OK	NOK

absorbed by transponder antennas. Hence, the value of voltage that is induced at the antenna terminals of transponder is not sufficient to power an RFID chip. Only 35% of the electronically marked bottles are identified in the presented case (Table 6).

Studies on the antenna arrangement that is dedicated to recognizing objects with RFID transponders oriented towards the side walls of the refrigerator are planned in the second stage of the experimental works. The 3rd and 4th antenna is made according to assumptions considered in the first model and placed in the close proximity of the side walls. Unfortunately, it is not possible to tune these antennas to the impedance of 50 Ω by using the selected automated matching circuits. It is caused by the proximity of the

metal components of the refrigerator. It yields the effect of reducing the loop induc-
tance to a level that could no longer be compensated by the elements built into the
dynamic antenna tuner (tuning range inductance 0.7–2.5 µH).

It can be stated on the basis of the conducted research that the modification of the
refrigerator structure proposed in the Gate 2 v.3 model is compulsory. It is a factor that
determines the high efficiency of automated identification of electronically marked
objects regardless of the bottle position and orientation in the refrigerator equipped
with the discussed set of multiplexed RWD antennas.

4.2   L Type of Multiplexed RWD Antennas

Numerical Model. The new conception is proposed on the basis of previously elab-
orated solutions. The multiplexed set with two spatial L-shaped antennas is considered
in order to simplify the structure of RFID system (Fig. 13). It is assumed that despite
using a reduced number of necessary RFID devices the improved construction is
enough to enable recognizing variously oriented objects.

The numerical model of multiplexed set with the L type RWD antennas is devel-
oped on the basis of the proposed conception. Specific conditions for integrating the
new construction with the glass door merchandiser are also taken into account. The
model includes the previously discussed parameters of structural materials and antenna
power supply.
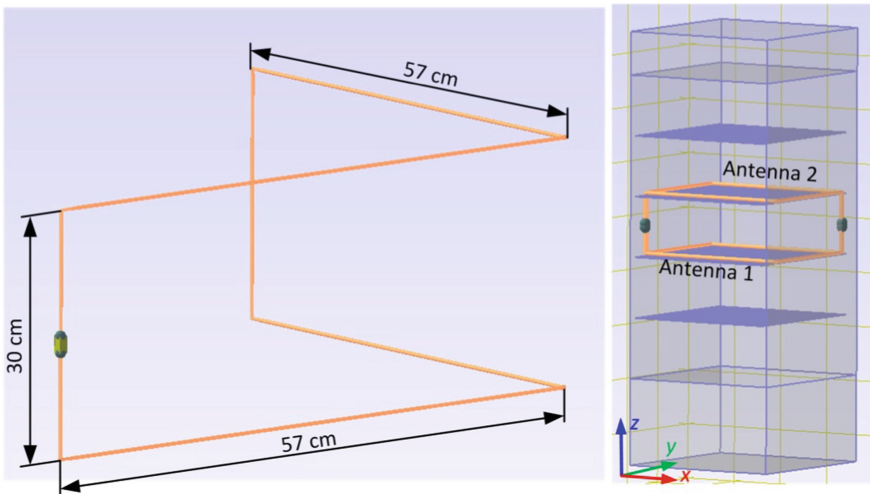
Fig. 13.   Model of set with L type antennas

Numerical Calculation. As it is described in the previous models, the magnetic field
strength distribution is determined for the carrier frequency of $f_0 = 13.56$ MHz. Due to
the shape of the antennas, calculations are carried out for the x-z (Fig. 14) and y-z plane
(Fig. 15).
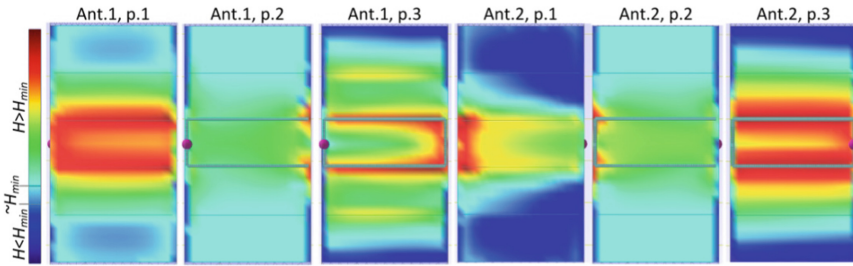
Fig. 14. Total magnetic field strength (*x-z* plane, L v.1, p. 1–5 cm from the door, p. 2 – middle, p. 3–5 cm from the rear wall)
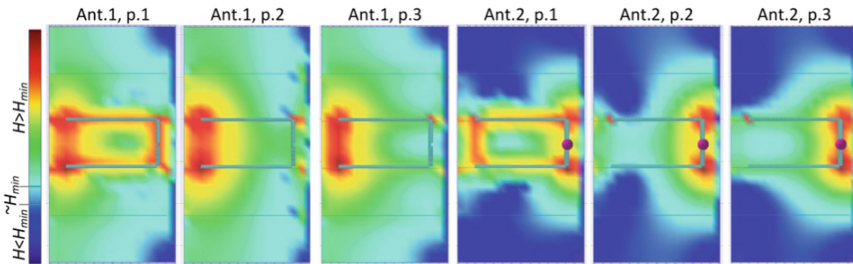
Fig. 15. Total magnetic field strength (*y-z* plane, L v.1, p. 1–5 cm from the left side, p. 2 – middle, p. 3–5 cm from the right side)
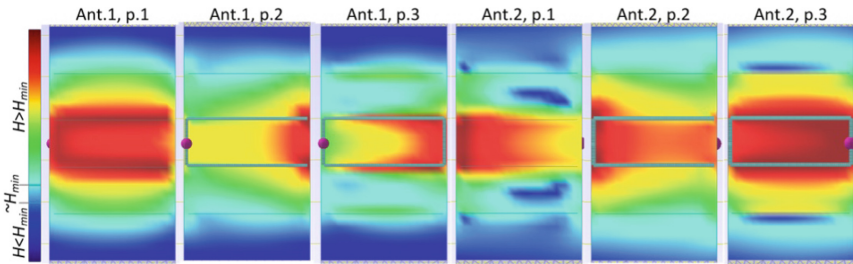
Fig. 16. Total magnetic field strength (*x-z* plane, L v.2, p. 1–5 cm from the door, p. 2– middle, p. 3–5 cm from the rear wall)

The analysis of calculations shows that it is possible to recognize the objects with the transponders that are oriented in *x-z* plane (parallel to the door of the refrigerator). However, it should be noted that there are areas on the shelf where transponders may remain without power. In the case of *y-z* plane, object identification is only possible in the close proximity to the RWD antenna. It should be noticed that changes in refrigerator construction could improve the RFID system also in this conception. If metal elements were replaced with plastics the identification effectiveness could be increased for all orientations of market objects. Considerations of the spatial separation between antennas and metal walls are not practical here due to the operational aspects of the glass door merchandiser.

Fig. 17. Total magnetic field strength (*y-z* plane, L v.2, p. 1–5 cm from the left side, p. 2 – middle, p. 3–5 cm from the right side)
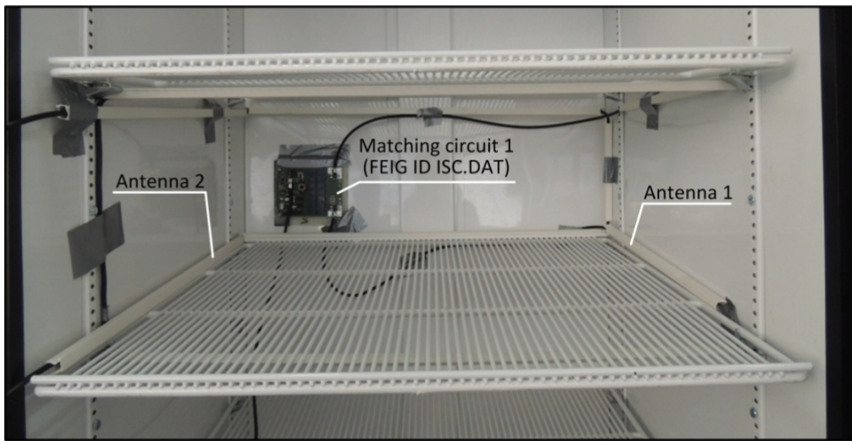
Fig. 18. L type of multiplexed RWD antennas

The above mentioned observations are the basis for developing the next numerical model (L v.2) with some modifications of the refrigerator construction. Instead of metal side walls, a casing made of plastic is considered. Hence the walls are modeled with a plane of 1 cm thickness and relative permittivity of $\varepsilon_r = 2.3$. The calculation results are compiled for the *x-z* (Fig. 16) and *y-z* plane (Fig. 17).

A significant increase in the geometry of the area in which it is possible to supply passive RFID transponders with energy is noticeable in the calculation results. It should be noted, however, that despite the modification of the refrigerator construction, the distribution of the magnetic field around the antenna loop is not as homogeneous as it is in the concept of Gate 2 v.2.

Experiment. The effectiveness of identification process with regards to the new model was also verified by experimental works. Multiple objects were subjected to identification processes in the modified demonstration system. In the first approach, the antenna arrangement was made in accordance with the model shown in Fig. 13. Unfortunately, an automatic impedance matching in such a system was incorrect

despite many experimental attempts. The inner side of single swing glass door is covered with a transparent antibacterial layer of metal and it was identified as the reason of the troubles. It strongly affected the antenna impedance. Hence, one of the 1st antenna's arms was shortened for solving the problem. As a result, the loop was separated from the fridge door (Fig. 18). Since the material parameters of the antibacterial layer are not designated and the numerical calculations have only qualitative nature it was decided not to prepare the next version of the model for these changes.

Table 7. Summary results for labels oriented towards the refrigerator door

	K1	K2	K3	K4	K5	K6	K7	K8
B1	OK	OK	OK	OK	OK	OK	OK	OK
B2	OK	OK	OK	OK	OK	OK	OK	OK
B3	OK	OK	OK	OK	OK	OK	OK	OK
B4	OK	OK	OK	OK	OK	OK	OK	OK
B5	OK	OK	OK	OK	OK	OK	OK	OK
B6	OK	OK	OK	OK	OK	OK	OK	OK
B7	OK	OK	OK	OK	OK	OK	OK	OK
B8	NOK	NOK	NOK	OK	OK	OK	OK	NOK
B9	NOK	OK	NOK	OK	OK	OK	OK	OK

Table 8. Summary results for labels oriented towards the side wall

	K1	K2	K3	K4	K5	K6	K7	K8
B1	OK	OK	OK	NOK	NOK	NOK	NOK	NOK
B2	OK	OK	OK	NOK	OK	NOK	NOK	NOK
B3	OK	OK	OK	NOK	OK	NOK	OK	NOK
B4	NOK	NOK	NOK	NOK	OK	OK	OK	OK
B5	NOK	NOK	OK	NOK	NOK	NOK	OK	NOK
B6	NOK	NOK	OK	NOK	NOK	NOK	NOK	NOK
B7	OK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B8	OK	OK	OK	NOK	NOK	NOK	NOK	NOK
B9	OK	OK	OK	NOK	NOK	NOK	OK	OK

As previously, the inventory procedure was carried out separately for each of the antennas and summary results of experiments for different orientation of transponders are compiled in Table 7 and 8. As it was expected, the best identification efficiency (about 92%) is obtained for the case in which RFID transponders are directed towards the door of the refrigerator. The effectiveness at orthogonal spatial orientation of objects is significantly worse and it is equal only 39%. These results coincide with their numerical representation.

4.3   X Type of Multiplexed RWD Antennas

Numerical Model. An arrangement consisting of two mutually intersecting antenna loops is the last of the considered configuration (Fig. 19). It is assumed once again that despite using a reduced number of necessary RFID devices the improved construction is enough to enable recognizing variously oriented objects.
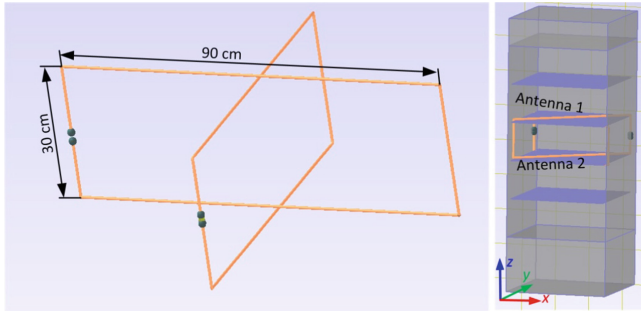
Fig. 19. Model of X type antennas

The numerical model of multiplexed set with the X type RWD antennas is developed on the basis of the proposed conception. Specific conditions for integrating the new construction with the glass door merchandiser are also taken into account. The model includes the previously discussed parameters of structural materials and antenna power supply.

Numerical Calculation. As it is described in the previous models, the magnetic field strength distribution is determined for the carrier frequency of $f_0 = 13.56$ MHz. Due to the shape of the antennas, calculations are carried out for the *x-z* (Fig. 20) and *y-z* plane (Fig. 21).

It can be concluded on the basis of obtained data that it should be possible in the case under consideration to identify any objects inside the modeled refrigerator regardless of orientations of transponders. According to predictions, the magnetic field
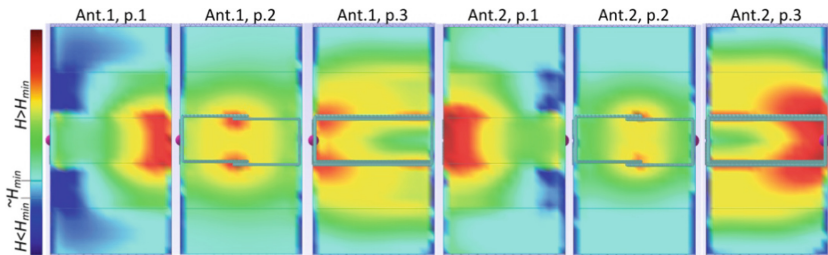
Fig. 20. Total magnetic field strength (*x-z* plane, p. 1–5 cm from the door, p. 2 – middle, p. 3– 5 cm from the rear wall)

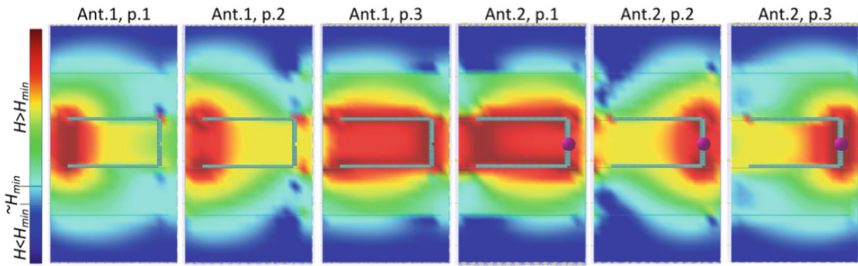**Fig. 21.** Total magnetic field strength (*y-z* plane, p. 1–5 cm from the door, p. 2 – middle, p. 3–5 cm from the rear wall)

strength has the highest values in the immediate vicinity of the antenna loop. Moreover, it does not significantly decrease inside the entire shelf.

Experiment. The effectiveness of identification process with regards to the last model was also verified by experimental works. Multiple objects were subjected to the identification processes in the modified demonstration system (Fig. 22).

Fig. 22. Experimental stand of multiplexed set with X-type antennas and group of marked bottles in considered case of RFID process

Table 9. Summary results for labels oriented towards the refrigerator door

	K1	K2	K3	K4	K5	K6	K7	K8
B1	OK	OK	OK	OK	OK	OK	OK	OK
B2	OK	OK	OK	OK	OK	OK	OK	OK
B3	OK	OK	OK	OK	OK	OK	OK	OK
B4	OK	NOK	OK	OK	OK	OK	OK	OK
B5	OK	OK	OK	NOK	OK	OK	OK	OK
B6	OK	OK	OK	OK	OK	OK	OK	NOK
B7	OK	OK	OK	OK	OK	OK	OK	OK
B8	OK	OK	OK	OK	OK	OK	OK	OK
B9	OK	OK	OK	OK	OK	OK	OK	OK

Table 10. Summary results for labels oriented towards the side wall

	K1	K2	K3	K4	K5	K6	K7	K8
B1	OK	NOK	NOK	NOK	NOK	NOK	NOK	NOK
B2	OK	OK	NOK	OK	NOK	NOK	NOK	NOK
B3	OK	OK	OK	OK	OK	OK	OK	NOK
B4	OK	OK	OK	OK	OK	OK	NOK	NOK
B5	OK	OK	OK	OK	OK	OK	OK	NOK
B6	OK	OK	OK	OK	OK	OK	OK	NOK
B7	OK	OK	OK	OK	OK	OK	OK	NOK
B8	OK	OK	OK	OK	OK	OK	OK	NOK
B9	OK	OK	OK	OK	OK	OK	OK	OK

As before, the inventory procedure was carried out separately for each of the antennas and summary results of experiments for different orientations of transponders are compiled in Table 9 and 10. The identification efficiency is equal to 96% for the *x-z* plane and 74% for the *y-z* plane.

It can be concluded on the basis of the conducted research that the X type of multiplexed RWD antennas is characterized by high application potential. However, the construction of the glass door merchandiser would need to be changed significantly. So, the practical implementation of this solution in contemporary commercial refrigerators is not achievable, because one of the basic requirements is not met – it has to be possible to remove the shelves from the device.

5   Conclusion

The implementation of RFID system in the glass door merchandiser was the aim of the conducted research. The possibility to identify any oriented electronically marked bottles with only one transponder attached to the central point of product label was considered in this paper. Taking into account economic and marketing benefits it does seem to be rational to equip commercial refrigerators with the complete set of multiplexed antennas regardless of the overall construction of the devices that are commonly used in the FMCG distributors. Nevertheless the necessity to complicate the hardware part of the system (increased number of multiplexer inputs and antennas with tuning circuits, the need to use a long range read/write device with higher output power in the transmitter) is one of the identified problems that has to be overcome. Moreover, operating and maintaining requirements (removed shelves, free access to bottles, etc.) of the refrigeration equipment have to be included in a designing solution. On the other hand, the implementation of RFID system for conducting the automated identification process only for one selected orientation of electronically marked bottles is the reasonable application (it meets the industry partner's needs and expectations). However, it is necessary to use an appropriate structural construction of the RWD antenna system. The possibility to use the gate that consists of antennas placed in the space between the shelves confirms the usefulness of such an application. It gives 100% efficiency of identification process (labels oriented towards the refrigerator door) and

the shelves can still be removed from the refrigerator. This configuration confirms the requirements of the assumed FMCG identification in a commercial glass door merchandiser specific for the considered application. It should be emphasized that the RFID transponders applied to the bottles can be used in the whole life cycle of a FMCG product (production, logistics, marketing services, recycling and others) for different tasks.

The obtained test results should constitute the initial conditions for the next stage of research and development works. The target construction of the set of multiplexed RWD antennas can be designed taking into account conducted analysis with regards to manufacturing potential, for implementation in a commercial refrigerator. In addition, it can be stated on the basis of the presented results that modification of the device construction is a crucial requirement and important factor for developing RFID applications in this scope. If materials used to build its interior were changed, it could be possible to synthesize an effective antenna system for the RFID identification of electronically marked objects regardless of their location and orientation. And then, although both L and X type of multiplexed RWD antennas is characterized by high application potential and it is possible to reduce the complexity of RFID system by using such an improved construction, the first shape is better choice according to simulation results and manufacturing requirements of integration.

References

1. Ustundag, A., Cevikcan, E.: Industry 4.0: Managing The Digital Transformation, 1st edn. Cham, Switzerland (2018)
2. Greengard, S.: The Internet of Things. The MIT Press, London, GB (2015)
3. Finkenzeller, K.: RFID Handbook – Fundamentals and Applications in Contactless Smart Cards, Radio Frequency Identification and Near-Field Communication, 3rd edn. Wiley (2010)
4. IDTechEx: Continued growth as market for RFID exceeds $10bn milestone. ID World Mag., 38–39 (2015)
5. Ustundag, A.: The Value of RFID. Benefits vs. Costs. Springer, London (2013). https://doi.org/10.1007/978-1-4471-4345-1
6. Das, R., Harrop, P.: RFID Forecasts, Players and Opportunities 2014-2024. Report, IDTechEx (2014)
7. Jankowski-Mihułowicz, P., Węglarski, M.: Definition, characteristics and determining parameters of antennas in terms of synthesizing the interrogation zone in RFID systems. In: Crepaldi, P.C., Pimenta, T.C. (eds.) Radio Frequency Identification, chapter 5, pp. 65–119. INTECH, Rijeka, Croatia (2017). https://doi.org/10.5772/intechopen.71378
8. GS1: GS1 System Architecture Document – How GS1 Standards fit together. Release 6.0. GS1 (2017)
9. Traub, K.: The GS1 EPCglobal Architecture Framework. Ver. 1.6. GS1 (2014)

10. Kim, D.S., Yoo, S.K., Kim, H., Chang, B., Bae, H., Kim, S.: Location based blood bag management using active RFID and ubiquitous sensor network. In: 6th International Special Topic Conference on Information Technology Applications in Biomedicine, pp. 320–322. Tokyo (2007). https://doi.org/10.1109/itab.2007.4407413

11. Suzuki, T., Oyama, Y., Nakauchi, Y.: Intelligent medicine case system with distributed RFID readers. In: IEEE Annual International Conference of the Engineering in Medicine and Biology (EMBC), pp. 344–347. Buenos Aires (2010). https://doi.org/10.1109/IEMBS.2010.5627662

12. Chen, Y.C., Hsieh, M.F., Wang, C.C., Lee, H.R.: RFID-based intelligent systems for home-healthcare. In: IEEE International Conference on Consumer Electronics (ICCE), pp. 1–2. Las Vegas (2007). https://doi.org/10.1109/icce.2007.341502

13. Shadangi, V., Jain, N.: Medical internet refrigerator. In: IEEE International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), pp. 363–366. Kumaracoil (2015). https://doi.org/10.1109/iccicct.2015.7475305

14. Carreno Laguna, J., Garcia Higuera, A., Zangroniz Cantabrana R., de las Morenas, J.: Comprehensive traceability system of milk samples using RFID. In: European Conference on Smart Objects, Systems and Technologies (SmartSysTech), pp. 1–8. Osnabrück, Deutschland (2012)

15. Qiao, S., Zhu, H., Zheng, L., Ding, J.: Intelligent refrigerator based on internet of things. In: IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), pp. 406–s 409. Guangzhou (2017). https://doi.org/10.1109/cse-euc.2017.262

16. Shweta, A.S.: Intelligent refrigerator using Artificial Intelligence. In: IEEE 11th International Conference on Intelligent Systems and Control (ISCO), pp. 464–468. Coimbatore (2017). https://doi.org/10.1109/isco.2017.7856036

17. Hachani, A., Barouni, I., Ben Said, Z., Amamou, L.: RFID Based Smart Fridge. In: IEEE 8th IFIP International Conference on New Technologies, Mobility and Security (NTMS), pp. 1–4. Larnaca (2016). https://doi.org/10.1109/ntms.2016.7792472

18. Desai, S.: Understanding IoT management for smart refrigerator. Int. J. Sci. Res. Dev. **1**(2), 1–4 (2016)

19. Floarea, A.D., Sgârciu, V.: Smart refrigerator: A next generation refrigerator connected to the IoT. In: IEEE 8th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), pp. 1–6. Ploiesti (2016). https://doi.org/10.1109/ecai.2016.7861170

20. Khan, A.M., Khaparde, A., Savanur, V.P.: Self-aware inventory system based on RFID, sensors and IBM security directory integrator. In: IEEE International Conference on Inventive Computation Technologies (ICICT), pp. 1–4. Coimbatore (2016). https://doi.org/10.1109/inventive.2016.7976074

21. Jankowski-Mihułowicz, P., Węglarski, M.: Interrogation zone determination in HF RFID systems with multiplexed antennas. Arch. Electric. Eng. **64**(3), 459–470 (2015). https://doi.org/10.2478/aee-2015-0035

22. Jankowski-Mihułowicz P.: Field conditions of interrogation zone in anticollision radio frequency identification systems with inductive coupling. In: Turcu, C. (ed.) Radio Frequency Identification Fundamentals and Applications Bringing Research to Practice, Chapter 1: 1–26, Intech (2010)

23. Reinhold, C., Scholz, P., John, W., Hilleringmann, U.: Efficient antenna design of inductive coupled RFID-systems with high power demand. J. Commun. **2**(6), 14–23 (2007)

24. Xu, C., Yan, Y., Liu, X.: Design of a long-range rectangular coil antenna for RFID access control system. In: International Conference Antennas and Propagation, pp. 420–423. Loughborough, UK (2013). https://doi.org/10.1109/lapc.2013.6711933

25. Aerts, W., Mulder, E.D., Preneel, B., Vandenbosch, G.A.E., Verbauwhede, I.: Dependence of RFID reader antenna design on read out distance. IEEE Trans. Antennas Propag. **56**(12), 3829–3837 (2008). https://doi.org/10.1109/TAP.2008.2007378

26. D'hoe, K., Goemaere, J.P., Stevens, N., Nauwelaers, B., De Strycker, L.: Automated design of an HF RFID loop antenna based on parametric geometry modification. In: IEEE International Conference on RFID (IEEE RFID), pp. 1–7. Orlando, FL, USA (2014). https://doi.org/10.1109/rfid.2014.6810705

27. True: GDM-26-LD – Glass Door Merchandiser, Swing Door Refrigerator with LED Lighting. True Food Service Equipment, O'Fallon, MO, USA (2015)

Measurement Methods for End Winding Vibrations of Large Electrical Machines

Christian Kreischer[(✉)]

Faculty of Electrical Engineering, Electrical Machines and Drive Systems,
Helmut Schmidt University, Holstenhofweg 85, 22043 Hamburg, Germany
christian.kreischer@hsu-hh.de

Abstract. In this paper measuring methods for the detection of end winding vibrations on large machines are presented and special requirements with regard to monitoring and diagnostic systems are discussed.

Large end winding vibrations due to switching operations, grid disturbances or mechanical loosening may cause partial conductor breaks and damage of the end winding insulation. The damaged insulation can lead to a short circuit with damage or even complete destruction of the electrical machine. The resulting costs for operating failure and repair exceed the costs for continuous monitoring of the end winding condition by a factor of 1000 for large electrical machines like turbo-generators in power plants.

In the first part of this paper, the causes and effects of end winding vibrations are discussed. Afterwards, modern methods for measuring end winding vibrations are presented.

Subsequently, a combined method for modal operating mode analysis is presented, taking into account a learned reference state. This modal view allows to draw conclusions about vibration excitation and phenomena such as operation close to resonance. However, the plausibility and spatial arrangement of the individual bar vibrations used for a modal transformation must be critically examined, since errors in the measurement chain and an unsuitable sensor configuration lead to misinterpretations. The modal consideration of the end winding vibration thus always represents a supplement to other observation forms, such as the Fourier spectrum of a single bar vibration. By learning a reference state with the help of neural networks, even small changes in vibration behavior can be detected.

Finally, the paper gives an outlook on the future developments and requirements of corresponding measurement systems

Keywords: End winding · Vibration · Monitoring · Fiber optics
Accelerometer · Modal analysis

1   Introduction

Through comprehensive measurement, monitoring and diagnosis of large electrical machines, costs can be saved through longer operating times and plannable maintenance work. The end windings represent two of the most sensitive machine components. In recent years, the continuous vibration monitoring of end windings has been

used more and more frequently. The development opportunities and potential savings in this area are correspondingly high.

The forced, damped operating vibrations of an end winding are stimulated both by electromagnetic forces between the bars and by the laminated core vibrations caused by the air gap field.

The electromagnetic forces in the end winding area are determined by the current through the stator winding and the magnetic field, which in turn is generated by the currents of all windings within the end area of the machine. The forces are therefore approximately proportional to the square of the currents occurring in the electrical machine.

In recent years an intensive development in the field of vibration measurement, vibration monitoring and vibration diagnosis can be observed. In the past, only very large machines were usually monitored for vibration. Continuous monitoring of shaft and bearing vibrations, for example, has become established in all large turbo-sets. The development of new and more compact measuring and monitoring systems has made vibration diagnosis beneficial for other machine components as well.

For the evaluation of a vibration, the recorded measuring signal is usually divided into its frequency components according to amplitude and phase by means of a Fast Fourier Transform (FFT).

Due to the increasing number of old turbo-generators and new operating requirements resulting from the liberalization of the electricity market, age-related damage is increasingly occurring in the end winding areas, which necessitates greater monitoring effort and the development of more precise diagnostic tools. Currently only relatively few end windings are continuously monitored, whereby the need for condition monitoring in this area continues to grow. Sensors for vibration monitoring in these areas must be high-voltage resistant and insensitive to strong magnetic fields. Conventional inductive or capacitive vibration transducers do not meet these criteria and therefore cannot be used for vibration monitoring. Fiber-optic accelerometers, on the other hand, are ideal. They are used primarily to monitor the radial vibrations of individual end winding bars. For cost reasons, not every bar can be equipped with accelerometers, which makes an accurate condition assessment difficult.

2   Causes and Effects of End Winding Vibrations

The vibration of an end winding depends on its structure and vibration excitation. The structure has stiffness and damping properties, which depend on design, production process, thermal condition and the aging condition of the end winding. The end winding's temperature depends on the operating condition of the machine with its characteristic operating variables such as speed, active power, reactive power and terminal voltage. With changing operating conditions, not only direct changes in excitation but also delayed temperature changes occur, which have a reversible effect on the stiffness of the end winding. In addition, irreversible structural changes can occur as a result of aging and grid disturbances.

A knowledge of the excitation mechanisms is essential for the evaluation of end winding vibrations, therefore the most important types of excitation are briefly described here.

2.1   Stator Core Excitation

In stationary operation, the usually dominant vibration excitation is caused by the magnetic tensile forces of the air gap field between rotor and stator of the machine. This force deforms the stator into an ellipse, which rotates synchronously with the air gap field. The stator bars are wedged into the slots of the laminated stator core and therefore the end winding is excited to vibrate. Figure 1 shows the stator core deformation for a two-pole synchronous machine in generator operation, whereby the rotor field is built up by an exciter winding fed with direct current. Due to the load, a phase shift occurs between rotor and stator fields, resulting in an angular offset between the rotor axis and the maximum stator core deformation. Due to the tensile forces acting on both sides, the dominant frequency of the stator core vibration corresponds to the product of the number of poles and the grid frequency.
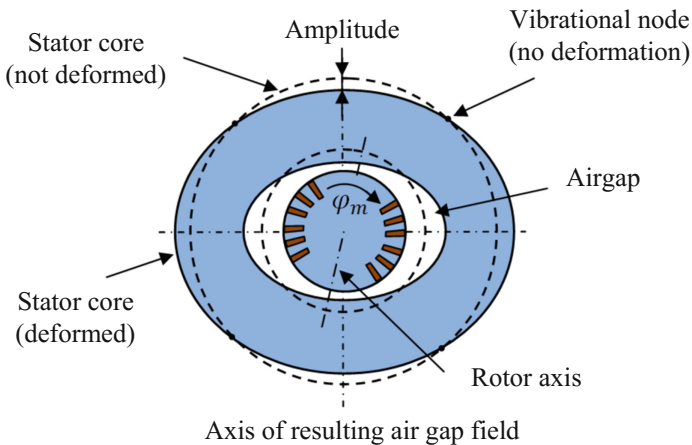
Fig. 1. Laminated stator core deformation as a result of magnetic tensile forces

2.2   Electromagnetic Forced Excitation

Another excitation is caused by the electrical current and the magnetic field in the end winding area. Since the end winding field as well as the resulting force are both proportional to the phase currents, the resulting electromagnetic force depends quadratic on the phase currents as a first approximation. In rated operation, the influence of this force is comparatively low in the end winding area, whereby high currents with considerable forces can occur during switching operations or operating faults in the grid, which contribute significantly to aging or loosening of the end winding structure.

2.3   Coupled Vibrations

Mechanical vibrations of the shaft can be transmitted to the end winding via the bearings, housing and stator core. Under certain circumstances, neighboring machine vibrations can also be coupled in via the foundation. Vibration excitation of the shaft can be caused, for example, by a mechanical unbalance or misalignment of the rotor, by a direction-dependent stiffness of the rotor or by a magnetic unbalance within the machine. The rotor imbalance S causes a rotating deformation during operation as shown in Fig. 2.
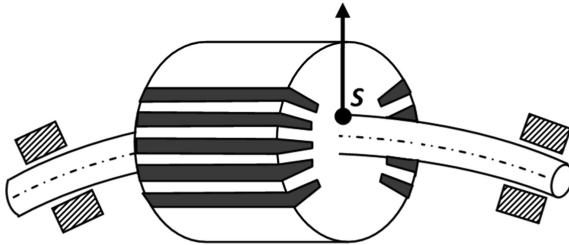
Fig. 2. Rotor imbalance and direction-dependent rotor stiffness

2.4   Effects of End Winding Vibrations

In [1] it is shown that any loosening of individual components can be detected by periodically repeated determination of the natural frequencies. Furthermore, it is explained that the reduced natural frequency of an end winding provides information about an initial damage and allows conclusions to be drawn about the degree of damage. If a two-pole machine in the 50 Hz grid has a two-node mode shape close to 50 Hz or a four-node mode shape close to 100 Hz, increased vibrations are to be expected during operation. To determine the natural frequencies, a so-called bump test can be used, which can only be carried out at standstill and with the machine opened. Since the natural frequencies shift due to the temperature dependence during operation, there must be a sufficient distance to the operating frequencies or their multiples. To determine the eigenmodes and eigenfrequencies of generator end windings, the bump test is nowadays used as a common method. In most cases, 12 accelerometers are installed at selected bar end connections, whereby the sensors are evenly distributed over the circumference. Fiber optic sensors have become established for measurement, as they combine several significant advantages for both bump tests and permanent monitoring of the stator end windings.

According to [2], the vibration of end windings may cause damage, which can be observed in cracks in the bonding and loosening of the bandages. As a result of progressive loosening, friction dust forms, which can reduce the high voltage resistance of the bar insulation and may lead to a short circuit with severe consequences. Each vibration monitoring system essentially consists of a measured value acquisition unit, a computer system and analysis software. In addition to the actual vibrations, additional parameters such as active power, reactive power and excitation current must be

recorded. It is also recommended to measure coolant temperatures, stator current, stator voltage and gas pressure, because "a well-founded machine diagnosis without knowledge of the associated operating parameters and states is generally not possible" [2].

Naghashan [3] investigates the ageing of winding bar insulations by means of partial discharge measurements on artificially aged bars and thus makes a significant contribution to the evaluation of permissible bar stresses. According to [3], thermal and mechanical loads are the most common causes of insulation damage, with bending stresses resulting from electromagnetic forces occurring mainly at the ends of the bars and in the slot exit region of the machine and increased temperatures causing insulation fatigue. An important proof provided by Naghashan is that the causes of damage of various stresses can be distinguished by partial discharge measurements. Furthermore, it is determined that the partial discharge activity contains a lot of information about the ageing state of the insulation and can therefore be used for diagnostic investigations.

The influence of mechanical loads on the bar insulation in the end winding area underlines the importance of suitable measuring and diagnostic tools in order to detect system changes as early as possible before the fatigue of the bar insulation becomes noticeable through increased partial discharge activity.

3 Measuring Systems for End Winding Vibrations

The measurement of end winding vibrations in the area of large machines is very demanding. In order to not reduce the high voltage resistance between the insulated bars due to measurement setup, no metallic components must be present in the sensor or in the supply lines. In addition, a low sensor weight is important in order to not influence the vibration behavior. When selecting sensors, attention should be paid to durability, good retrofitting and reliability. To obtain a reliable phase reference in the measurement data, the measurement must be triggered accordingly. With regard to a common evaluation of several measuring signals, a time-synchronous measurement is also necessary. Fiber-optic sensors of suitable design and installation meet all requirements and are particularly well suited for use in large machines.

In the following, the most important measuring systems and concepts for recording end winding vibrations on a fiber-optic basis are presented. While the initially presented measuring system still has an internal mechanics, efforts exist to use fiber optic sensors with Bragg gratings in the end winding area.

For a vibration assessment of the machine condition, individual vibrations must be reliably detected in a first step. While transducers for shaft and bearing vibrations are usually based on a conventional inductive or capacitive measuring principle, special requirements are placed on the measuring setup in the end winding area.

3.1 Fiber-Optical Accelerometers with Mechanics

Compared to other vibration measurement systems, fiber-optic accelerometers have high voltage resistance and magnetic insensitivity required in the end winding area.

Structure and Functional Principle. The fiber-optic system essentially consists of a fiber-optic acceleration sensor, a fiber-optic conductor and an electronic coupler. The overall length of the system is generally pre-assembled by the manufacturer.

The fiber-optic accelerometer consists of a small optical sensor head made of a non-conductive material. The sensor head contains an oscillating system with a mirror.

The fiber optic conductor is designed as a glass fiber, which establishes an optical connection between the sensor head and the electronic coupler. The glass fiber is protected from mechanical loads by a Teflon jacket surrounding it.

The electronic coupler typically has a housing made of weak magnetic metal. The electronics consists of an emitter and a receiver of the optical signals. The housing also contains an amplifier and a filter for direct signal processing. The processed analog measuring signal can be picked up via a plug connection at the end of the coupler. The solid flange design of the electronic coupler also serves as a gas-tight feed through.

The coupler's electronics generate a light signal which is transmitted to the sensor head via the broadband optical fiber. The optical signal is deflected at the mirror of the oscillating system and reflected to the electronic coupler. The angle of deflection is proportional to the excitation force or acceleration. The reflected optical response signal is converted into an analog measuring signal by the electronic coupler and then amplified and filtered. The output signal is a voltage proportional to the acceleration, which can be further processed using conventional analog signal technology.

Due to the electrically non-conductive fiber optics, the sensor can be mounted directly on high-voltage components of the generator. Fiber optics are not only used for the protection of measurement technology, but also for personal protection. In addition, the optical signal routing is insensitive to interference from electromagnetic fields.

Installation and Sensor Positioning. Easy retrofitting is particularly important for the installation of measuring systems, as there are usually no design solutions for the installation and placement of the sensors inside the machines. Therefore, the installation of fiber optic systems is often very individual.

When installing the glass fibers, care must be taken that they cannot move freely. This prevents the glass fibers in the mostly turbulent interior of the generator from rubbing against each other, which in turn can lead to sensor failure. The acceleration sensors for monitoring a stator end winding are usually installed in areas with high vibration amplitudes. These include switching connections and winding connections.

Depending on the vibration problem, different positioning strategies for the sensors in axial, radial or tangential direction are available. Already known damages can be specifically monitored by sensors near the loosening.

3.2   Accelerometers with Fiber Bragg Grating

A new development in the field of vibration detection are optical fibers with integrated fiber Bragg gratings (FBG). An FBG consists of several periodically consecutive layers of different refractive index n, which are fired into the glass fiber by means of a strong UV laser and a phase mask. Incident light within a narrow wavelength range is reflected by the grating.

Figure 3 shows the cross-section of a monofiber with an inscribed Bragg grating with a grid spacing of Λ. The different refractive properties of the materials in the optical waveguide are determined by refractive indices $n_1, \ldots, n_3$. The light output P_E is coupled into the fiber optic cable. The FBG reflects or filters the power P_R, whereby the non-reflected power P_A exits at the end of the optical waveguide.
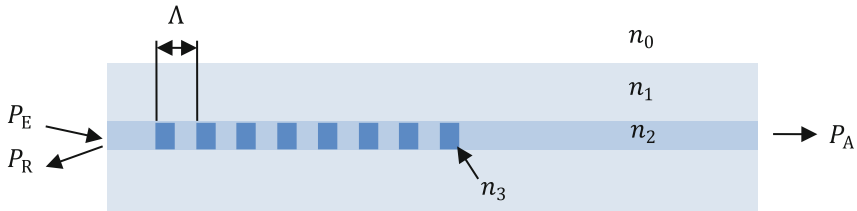
Fig. 3. Schematic representation of an optical waveguide with fiber Bragg grating

The average wavelength λ_R of the bandwidth reflected in a monofiber can be determined according to the Bragg condition as follows:

$$\lambda_R = \frac{n_2 + n_3}{2} \cdot 2\Lambda \tag{1}$$

The bandwidth depends on the length of the FBG and the refractive index gradient between the individual layers.

When coupling in a wide input spectrum, a spectrometer can be used to determine the grating distance and thus the mechanical strain at the measuring position. A sensor based on FBG can therefore be used in the same way as conventional strain gauges to detect deformations and also offers the essential advantage of immunity to interference from high alternating electrical and magnetic fields. Another advantage of this sensor type is the possibility of integrating up to 24 sensors within one optical fiber [6].

In addition to the direct measurement of deformations, other quantities can also be measured with the aid of optical fibers using FBG. Alternating electric fields and high voltages can be determined by measuring the field strength-dependent material expansion of a piezoelectric crystal. Magnetic alternating fields can be measured by the elastic deformation of magnetostrictive materials, and conclusions can be drawn about the temperature in the area of the FBG via the thermal expansion of the optical waveguide. For precise measurement, it is important to clearly separate the different effects from each other, which can be achieved during temperature measurement, for example, by means of a decoupling glass tube around the sensitive area.

4   Analysis of Vibrational Data

The decomposition of a spatially distributed vibration into its modal forms enables a more precise analysis of the different types of excitation. The vibration of a measuring point can be represented both by a superposition of the individual vibration harmonics

and by a superposition of the modal forms. By means of a known sensor configuration within an end winding plane with at least five sensors aligned in the same direction (radial, tangential or axial), it is possible to convert the real measured data of operating vibrations into modal forms by means of a transformation presented in [4]. The sensors should be distributed as evenly as possible over the circumference and attached to bars with similar vibrational behavior, e.g. verified by bump tests. For the machine's switching, this means that the sensors should be mounted in the middle of a phase group if possible. For the modal analysis it is crucial that the measurement of all sensor positions is done at the same time and with a defined phase reference.

In addition to consider the end winding as a closed system, individual bar monitoring with the currently used tools for vibration analysis, such as trend diagrams, plots over a specific operating parameter, polar diagrams and the Fourier spectrum, should also be part of the entire diagnostic system.

For an evaluation of individual bar vibrations, the mode of operation must be taken into account, since the vibration amplitudes significantly depend on the operating parameters speed, active power and excitation current.

In addition to vibration the amplitudes, the vibration phases must also be considered. Larger changes in vibration behavior can also be determined by a trend analysis. Due to the dependencies on the operating parameters, direct monitoring on the basis of limit values is difficult.

However, changes can be easily identified by modelling a reference vibration behavior. For a modal vibration evaluation, the trend representations of the individual vibration signals are particularly important if atypical modal forms occur. In this case, an expert is necessary to determine whether these modal shapes are caused by a defect in the sensor chains or by local damage to the observed end winding.

For a separation of operation-dependent and structural changes, the healthy vibrational behavior of the end winding has to be known. For example, the use of neural networks which are trained with a reference behavior (vibration and operational data) allows the detection of structural changes of the end winding [5].

5   Future Developments and Requirements

The right definition of sensor position and orientation are crucial for the subsequent evaluation of the measurement data. This information must be reliably recorded during installation and taken into account during subsequent analysis.

In order to be able to make statements about the ageing of the end winding from the measurement, the sensor conditions must be known. The required measuring range is typically between 0 and 200 Hz. Explosion protection must be ensured if the measuring system is to operate with critical cooling media such as hydrogen. In this case, the measurement setup must also have either a suitable radio connection through the machine housing or a gas-tight lead-through of the measuring lines.

A minimum requirement for future end winding monitoring systems is a time- and phase-related vibration measurement including the most important operating parameters of the machine (rotor speed, active power, reactive power, bar temperature, excitation current, terminal voltages and currents) every 100 μs. The data is first stored

into a ring buffer. To reduce the measured data, a frequency analysis (FFT) over a short time range of e.g. 500 ms has to be done. The determined frequency components can be checked for values exceeding limits. In addition, the time signal can also be assessed in terms of short-term amplitude changes.

The raw data are further analyzed and stored in case of detected anomalies. This serves the detection of network disturbances with transient processes and enables the estimation of short-term stresses. In addition, the time data for the operating parameters with the same phase reference have to be analyzed and stored.

For normal measurement data, it is sufficient to store the most important frequency components of the vibration data according to amplitude and phase as well as the sum of the residual amounts averaged over a period between 1 s and 1 min. In addition, the operating data averaged over time must be stored.

6   Conclusion

In summary, it can be stated that the condition monitoring of the end winding vibrations of large machines can detect damage at an early stage. For a separation of operation-dependent and structural changes, the use of neural networks is recommended, which can detect deviations from a reference behavior. Limit values can be defined for both absolute deviations and deviation gradients. This enables continuous monitoring.

For the diagnosis of the vibration, it can be transferred to the modal range and analyzed, whereby the different excitation influences can be separated from each other and the causes of a vibration change can be identified. In particular, electromagnetic operating excitations can be separated from rotor influences and resonance phenomena can be detected. The combination of neural monitoring of modal analysis therefore complements each other very well.

By detecting damage at an early stage, action measures can be planned in time, the availability of the machine can be increased and costs due to unexpected down times can be saved.

For a comprehensive machine diagnosis, however, it is not only the end winding that should be considered. Further electrical, pressure, temperature and vibration measurements can provide valuable information for unusual operating conditions and help to narrow down the causes.

References

1. Intichar, L.: Eigenfrequenz- und Eigenformbestimmung an Generator-Wickelköpfen. In: Symposium Schwingungsdiagnose, Potsdam (2006)
2. Frerichs, D.: Überwachung der Wickelkopfschwingungen von Generatoren mit faseroptischen Beschleunigungsaufnehmern. In: Symposium Schwingungs-diagnose, Potsdam (2006)
3. Naghashan, M.R.: Untersuchungen zur Teilentladungsaktivität von maschinentypischen Hochspannungsisolierungen. Dissertation am Lehrstuhl für Hochspannungstechnik, Universität Dortmund (1996)

4. Kreischer, C., Kulig, S., Thien, D.: Modal analysis of operational end winding vibrations. In: IEMDC 2011, Niagara Falls, ON, pp. 1207–1212 (2011). https://doi.org/10.1109/iemdc.2011.5994775
5. Kreischer, C., Golebiowski, M.: Monitoring of end winding vibrations using neural networks. Przegląd Elektrotechniczny **89**(11) (2013)
6. Strack, S., Weidner, J. R., Bosselmann, T., Villnow, M., Willsch, M.: Faseroptische Messtechnik zur Online-Überwachung und Diagnostik bei großen Generatoren in Kraftwerken. In: ITG-Fachtagung: Sensoren und Messsysteme, Nürnberg (2010)

Non-contact Robotic Measurement of Jet Engine Components with 3D Optical Scanner and UTT Method

Krzysztof Kurc(✉) , Andrzej Burghardt , Piotr Gierlak ,
and Dariusz Szybicki

Rzeszow University of Technology, al. Powstańców Warszawy 12,
35-959 Rzeszów, Poland
{kkurc,andrzejb,pgierlak,dszybicki}@prz.edu.pl

Abstract. This paper presents a method for the robot-assisted geometric inspection of an aircraft engine turbine stator segment, involving two robots. The first robot was an ABB IRB 1600 with an optical 3D scanner. The second robot was an ABB IRB 140, to automatically inspect the stator vanes at 168 points by the application of a UTT method. If the casting geometry tolerances are met, characteristic coordinates of points across the casting are determined for their further use during an alternative robot-assisted vane wall thickness measurement process. The operating principle of the test stand measurement system is presented, with a specific focus on the measurement strategy. The results of the wall thickness measurements performed on stator vanes are presented in the report. The correctness of the solution has been proved with scans and measurements of two turbine rotor guide vane segments of an aircraft engine provided by courtesy of Consolidated Precision Products Poland sp. z o.o.

Keywords: 3D scanning measurement · UTT measurement · Robot
Aircraft engine · Geometry inspection · Vane thickness measurement

1 Introduction

The quality control processes for workmanship related to aircraft engines are undertaken at the manufacturing and assembly stages of component production. The significant time required for these processes, sometimes involving repetitive measurements, has prompted their robotization. Measurements, diagnostics, identification and quality control has been considered in many research papers, [1–10], which feature completed or simulated solutions.

Stators are manufactured by precision casting. Low thrust aircraft engines feature monolithic stator castings with a number of vanes, which range from ten to several dozen. The processing conditions for these castings make it difficult to manufacture large stators for large aircraft engines. In this case, a stator is built as an assembly of ten or more segments, each with 3 to 6 vanes. The casting weight is a decisive factor when selecting the stator casting method. To reduce the casting weight requires the use of thin-wall cored castings with a wall thickness of less than 1 mm. Given the

manufacturing difficulties, complexity and technical requirements, these castings are among the most technologically advanced products, and are usually protected (along with their manufacturing technology) as the intellectual property of the manufacturer; hence their manufacturing methods are not widely published. Precision-made polycrystalline castings of the large thin-wall segments for a turbine stator of an aircraft engine are prone to deformation at the casting solidification stage. Hence, 20% of the castings have defects, caused by many factors. The presented robot-assisted geometric inspection of a stator segment was intended as preliminary geometric verification of the casting by comparison with a CAD model (master). The geometric verification process also identified characteristic coordinates for points on the casting for their further use during a robot-assisted vane wall thickness measurement process involving ultrasonic sensors.

The turbine rotor guide vanes (stator) of a turbofan engine are components which shape the engine gas path. The vanes are also highly exposed to dynamic loads from the hot gas stream (with temperatures ranging from 950 °C to 1650 °C) [11]. Hence, it is essential to improve the high temperature resistance of rotor blades by using ceramic thermal barriers and by cooling the stator during operation. The cooling of the component requires a complex cored casting, where the core forms the internal passageways that transfer the cooling medium and also reduce the component weight. The cooling process results in a temperature gradient across the casting wall section. This temperature gradient induces heavy thermal stresses, which may contribute to failure of the casting (by fracture) and its protective thermal barrier (by chipping, cracking or flaking). The component weight can be reduced, and the stator segment and vane cooling system performance improved, by making the vane casting walls thinner. It is thus critical to assure the dimensional repeatability of wall thickness during the manufacturing process. Among the available NDT (non-destructive test) methods suitable for wall section inspection, e.g. active thermography [12], CT (computer tomography) [13], and eddy current defectoscopy [14], ultrasonic measurement methods [15] are the most popular.

In general, three measurement methods are used during ultrasound measurements: echo [16], shade [17] and resonance [18]. The method selected for the measurement depends on the workpiece size and material. The choice of test medium is also important, and depends on the measurement method used and the environmental conditions. Manual defectoscopy uses gels [19], while robotic or automated systems tend to use water and oils [20, 21]. However, if the measurement temperatures are low, silicone gels are used [22]. In this paper, water was used as the medium.

The need for continuous quality control and the high number of defined measurement points increase the process time, which in turn results in higher costs, and hence leads to the search for automated [20] or robotic solutions [21]. Solutions are also used which include dedicated 2, 3 and 4-axis robots [23] and mobile dedicated robots for special non-destructive testing [24]. In addition to the dedicated robotic structures, there are also a small number of systems based on standard 6-axis robots [25].

This paper presents a developed and implemented robotic inspection station to assess the workmanship quality of rotor castings for jet engine turbines, based on a robot-assisted geometry inspection system.

2   Robot-Assisted Geometry Inspection

The robot-assisted geometry inspection was performed with an ABB IRB 1600 robot manipulator equipped with an ATOS Core 135 optical 3D scanner (Fig. 1) and ATOS Professional control and measurement software suite. With a specific control program developed in ATOS Professional to run the control cabinet commands, the ABB IRB 1600 robot manipulator facilitated all the required positioning and orientation of the optical 3D scanner and automatic measurements [26].
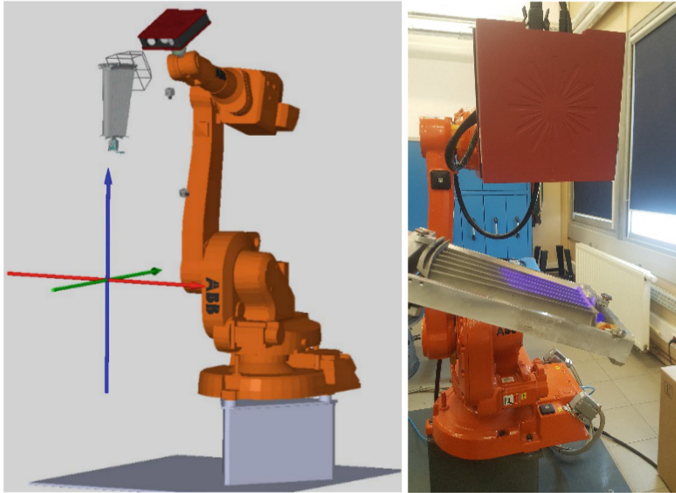
Fig. 1. Virtual and actual geometry inspection stand.

The ATOS Professional software suite can store the spatial measurement positions and orientations of the 3D scanner in order to automate and provide repeatability of successive component inspection runs. When each measurement is complete, the scanner must be moved and rotated to areas not recorded in the preceding scanning run. The individual measurements are automatically processed within a common coordinate system, which produces a complete cloud of 3D points.

The ATOS Core 3D scanner is a triangulated system of stereoscopic cameras (i.e. to calculate the intersection of a specific plane with a beam in three-dimensional space). The scanner projects a system of bands on the inspected surface of the workpiece. The projected bands are recorded by the two stereoscopic cameras (Fig. 2), providing a phase shift image from the sine distribution of intensity on the camera detectors. The ATOS Core uses multiple heterodyne phase shifts to provide the maximum sub-pixel accuracy. Separate 3D coordinates are calculated for each camera detector pixel from optical transformation equations.

This gives a calculated polygonal grid which circumscribes the free surfaces and geometric elements. The polygonal grid is verified by comparing the reproduced
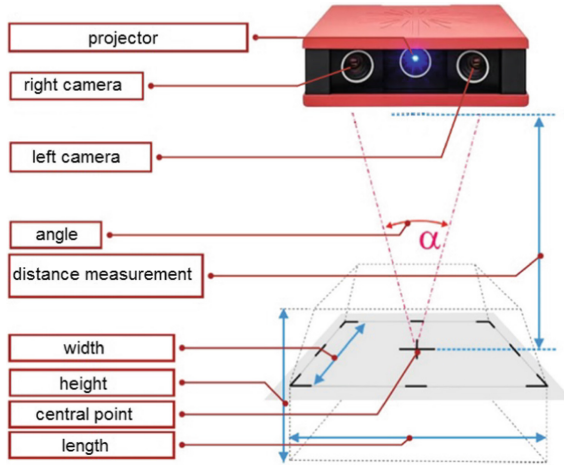
Fig. 2. ATOS Core measurement concept.

surface with the engineering drawing or directly with a set of CAD data [27, 28]. The software can also implement 3D surface analysis and 2D analysis of sections or points.

Each single measurement generates up to 5 million data points. The scanner records only those points visible to both cameras in a single scan. In order to digitize a complete object, several individual measurements are required from different angles. Based on reference points (circular markers), which are attached directly to the object or to the measuring plate or a fixture, Atos transforms these individual measurements automatically into a common global coordinate system.

The Atos Core 135 three-dimensional scanner is equipped with projector and camera lens setups (Table 1).

Table 1. ATOS Core 135 system configurations.

Measuring area	135 × 100 mm
Working distance	170 mm
Point spacing	0.05 mm
Resolution	5 Mpix
Sensor dimensions	206 × 205 × 64 mm
Weight	2.1 kg
Power supply	90–230 V AC
Operating temperature	+ 5 °C up to + 40 °C, non-condensing

There are already several fields where three-dimensional scanning is an established method of data acquisition. In mechanical engineering, 3D scanners are often used for workpiece inspection, deformation analysis, reverse engineering and reengineering of moulds and dies, and general quality control procedures [29, 30]. Civil engineering also

uses three-dimensional scanning during building inspections, custom fit furniture design, and cultural heritage protection and renovations. The textile industry uses scanning for digitalization of the human figure in custom fit product design. The movie industry also widely uses three-dimensional scanning for various CGI effect creation processes. Its use has also spread to the marketing and advertisement industries, [31–34]. This has already caused some degree of specialization by scanner manufacturers regarding the field of use.

The detailed inspection results can be presented in the form of reports, which may include screenshots, measurement charts, diagrams, text and graphic elements. The results can be represented in a graphical format, edited in the user interface or exported as PDF files.

3   Determination of Actual Measurement Point Coordinates

A number of unfavourable phenomena occur during the solidification of the large thin-wall polycrystalline castings of stator segments. These include shrinkage, stress, and ultimately, deformation. Once cleaned, the casting is subject to a preliminary visual inspection. If the casting passes, it is scanned to produce a complete cloud of 3D points with defined parameters (Fig. 3).

PROPERTIES
»
Information ◄
General ▼

Name   Vane

Comment

BASIC POLYGONISATION:
creation tolerance = 0.002727 mm
merge tolerance = 0.005453 mm
maximum gap = 0.237785 mm
maximum edge length = 0.713356 mm
unit edge length = 0.059446 mm
'Polygonize large data volumes' is ON

POSTPROCESSING:
smooth (size = 4, mode = 1, tolerance = 0.005453 mm)
thin (maximum edge length = 0.475571, tolerance = 0.000909 mm)
smooth (size = 4, mode = 1, tolerance = 0.002727 mm)
thin (maximum edge length = 0.475571, tolerance = 0.000454 mm)

Fig. 3. Parameters of a completed 3D scan.

Atos Professional has an efficient feature called Fast Inspection, which facilitates the display of a colour map showing the deviations from the preset geometric tolerances (Fig. 4) and inspecting the measurement features in real time, directly after each scanning run. The measured results are updated and displayed during the measurement cycle in relation to an imported or manually input measurement plan and a CAD model.
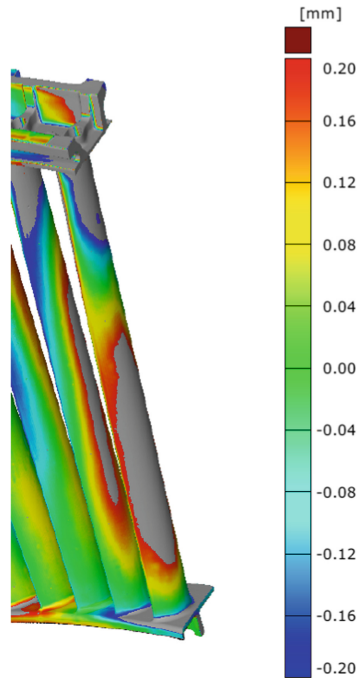
Fig. 4. Colour map of preset geometric tolerance deviations.

The quality control functions for the analysis of airfoils and turbine vanes or blades include checks of: camber line, geometric solid mass centre, turbine blade and vane profile thickness, chords, angle of twist, attack radius, trail radius, etc. For the inspection of castings, a program was developed that generates 42 sections of stator segment vanes, with a comparison of the CAD model with the generated 3D point cloud, the camber line being reproduced at the selected points of the vane section components with geometric tolerance deviations (Fig. 5).

The 3D point cloud fitted to the CAD model as above is used downstream to determine the actual coordinates of the measurement points required to verify the vane wall thickness.

Each of the stator vanes is a precision thin-wall casting and requires the verification of wall thickness at 28 measurement points. The points are determined with utmost precision, and currently the inspection process requires manual marking with a special template. This method has been acceptable so far, since the vane wall thickness has been measured manually, with a hand-held flaw detector. If the section thickness inspection process is conducted with a robot and the UTT method, the points must be input into a controller to manage their locations and measurement. Hence the point coordinates are derived from the 3D scan (the 3D point cloud), generated as described above.
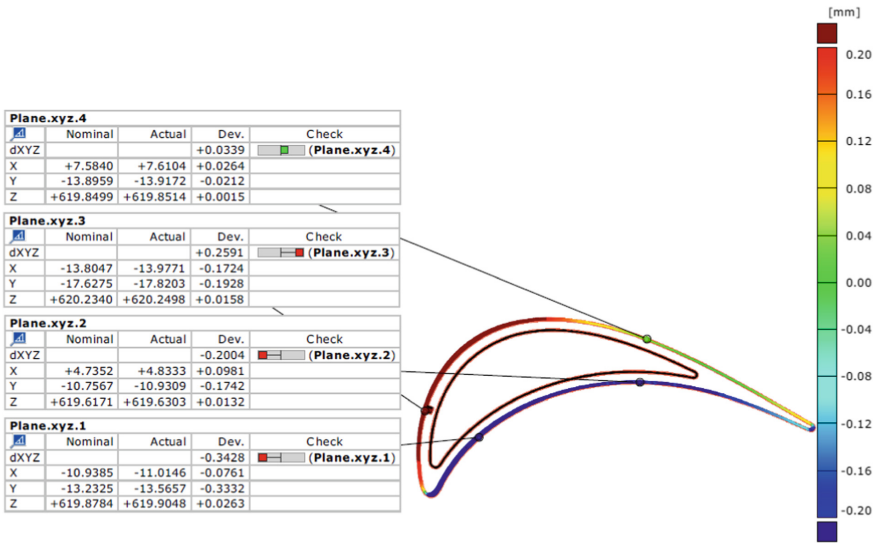
Fig. 5. One of the 42 sections of a stator segment vane.

The nominal workpiece (the CAD model) has the nominal coordinates determined (Fig. 6) for the first 7 points on the vane of the inspected stator, with the actual coordinates for the casting being generated automatically (Fig. 6).
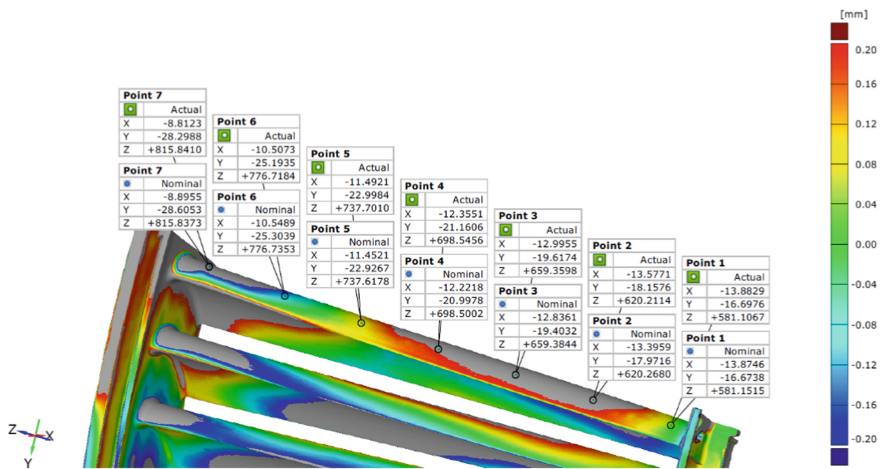
Fig. 6. Nominal and actual measurement points for the generation of motion trajectory applied in the stator vane wall thickness measurement process.

The origin of the coordinate system is selected (offset) during the determination of the actual coordinates of the geometric measurement system (Fig. 1) to match the coordinate system applied to measure vane wall thickness (Fig. 7).

Fig. 7. Robot-assisted vane thickness measurement stand.

4 Robotic Vane Thickness Measurement Stand

The 168 specific coordinates of the actual measurement points, referenced to the nominal points (Fig. 6), are produced for a single complete part of the stator and are input to the controller of the robotic measurement stand (Fig. 7) [35]. The specific coordinates permit the generation of the motion trajectory for the ABB IRB 140 to perform the UTT measurement [36].

The software was designed to assume a robot operating on a fixed object, i.e. the inspected stator, with the robot tools being UT scanning probes (Fig. 8). This assumption was critical to the programming of the solution and greatly facilitated the design process.

The workpiece and its coordinate systems are moved, while both tools are fixed in place. This helps to design the motion paths faster and to modify them with greater ease [37].

The measurement heads (both left and right-hand ones) were permanently fastened and had constant 3D coordinates. The actual measurement point coordinates (Fig. 6) were variable and automatically determined by the robot manipulator with its optical 3D scanner. This rapid determination of the measurement coordinates at 168 points allowed further automatic definition, with a measurement point on the trajectory being defined in each coordinate system, with several waypoints for the approach to the UT scanning probe (Fig. 9b).
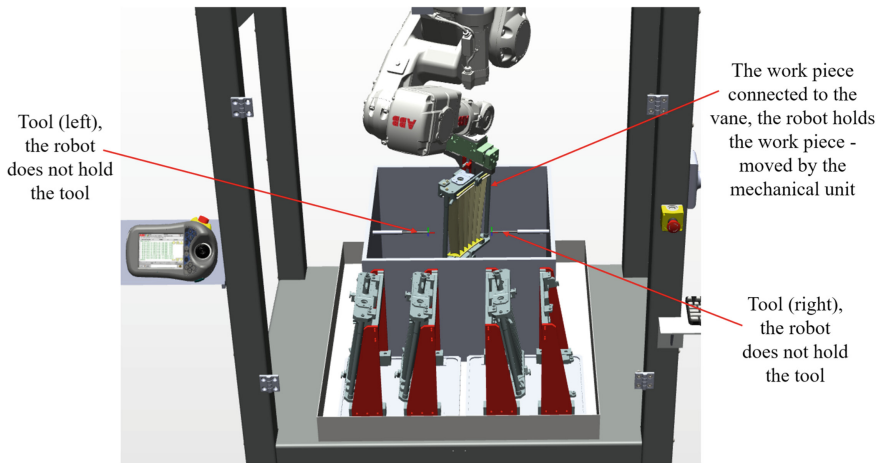
Fig. 8. Layout of the coordinate systems and tools on the test stand.
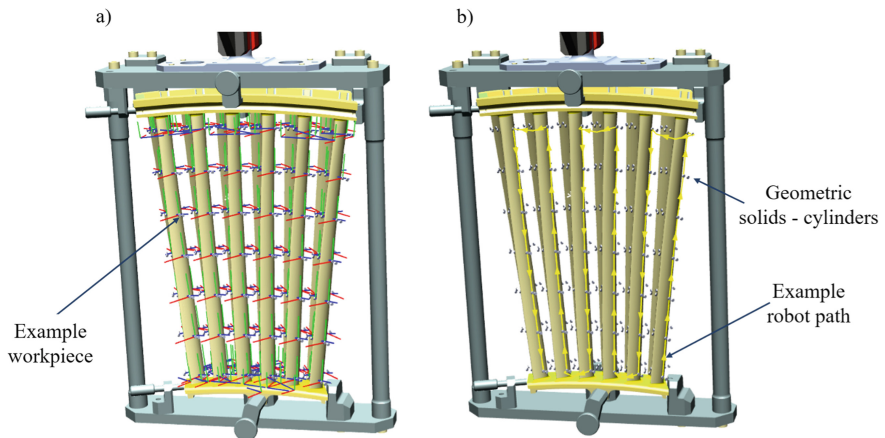
Fig. 9. Measured element: (a) Fixture stator with a stator and defined coordinate systems; (b) Fixture cassette with a stator and an example of the scanning path.

5 UT Measurements of Stator Casting Quality

The application of UT measurements for the determination of material section thickness was based on the physical phenomena of wave propagation, i.e. the reflection (echo) of the wave incident to a medium with physical and chemical characteristics that differ from the tested material. The reflection (echo) point was the interface between the test material and the medium. The phenomenon was caused by the change in the acoustic resistance of the wave. The reflected wave amplitude increased with increases in wave resistance difference between the two media. Therefore, if an echo of the sound wave was present in the investigated medium, there must have been a discontinuity

within the medium. If the time from transmitting the sound wave in the tested medium to the return of the wave reflected by that discontinuity was measurable, and the wave propagation velocity in the given material was known, then the distance covered by the wave was calculable and hence the material thickness could be determined.

The produced robotic test station application used a UT wave transceiver formed by two piezoelectric elements. The location of the piezoelectric elements was determined by the form of the inspected casting. The location of the piezoelectric transceivers was determined by simulation in RobotStudio.

The robotic measurement of vane wall thickness involved the verification of the casting section thickness along a sequence of measurement points. The couplant used was water. An example of the measurement signal is shown in Fig. 10. The visualised wave trend featured three characteristic areas. Area 1 was the wave echo caused by the design of the UT probe. This trend fragment had no relevant measurement data. Area 2 was the echo at the interface between the couplant and the outer casting surface. Area 3 was the echo at the interface between the casting and the couplant.
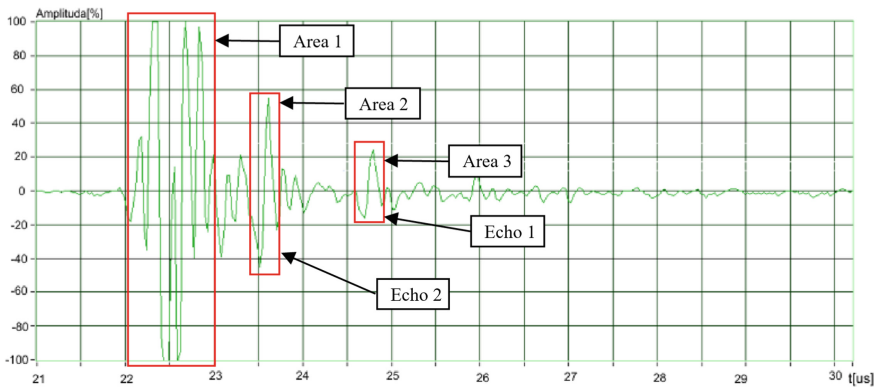
Fig. 10. Example measurement signal showing characteristic areas.

The thickness was measured with upper signal amplitudes, which eliminated the measurement of the distance between the upper and lower halves of the amplitude. The measurement was performed by determining the time of signal transition between echo 1 and echo 2. Three level indicators, or gates (Fig. 11), were used in this measurement. Gate 1 was the detection of the echo maximum and it automatically determined the signal analysis area on the x-axis. Based on the gate 1 position value, the coordinate x value of gate 2 was dynamically determined. This facilitated locating the echo 1 amplitude value (Fig. 1) within gate 2. In the next step, the gate 2 value was used to automatically determine the coordinate x value of gate 3. This facilitated locating the echo 2 amplitude value (Fig. 1) within gate 3. Thanks to the dynamic gate positioning, the measurement system was immune to the effects of changes in the distance of the UT transceiver probe from the test item.

The coordinate y values of gates 1 and 2 were selected experimentally to have echo values higher than the values assumed for the gate level. The lower limit must be higher
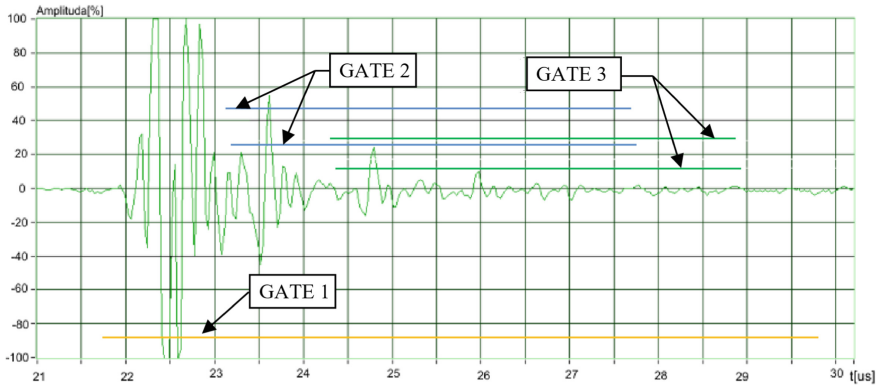
Fig. 11. Example of the measurement signal with characteristic areas.

than the signal variation amplitude. During the experiments, the y value was 20% of the maximum signal amplitude.

The applied measurement method required calibrating the UT wave propagation velocity for the test item material. The calibration process for the test stand discussed here was automatic: when a standard item thickness was input, the calibration function automatically determined the coefficient of wave propagation velocity.

The software suite of the test stand featured an advanced UT defectoscope functionality and digital connectivity with the robot controller. The application level allowed management of the UT signal parameters. The software was able to store the recorded measurement results as a HTML report and the settings as an INI file.

The software allowed adjustment of the UT transceiver probe output pulse frequency, the pulse generation repeat interval, and the pulse generation voltage. The
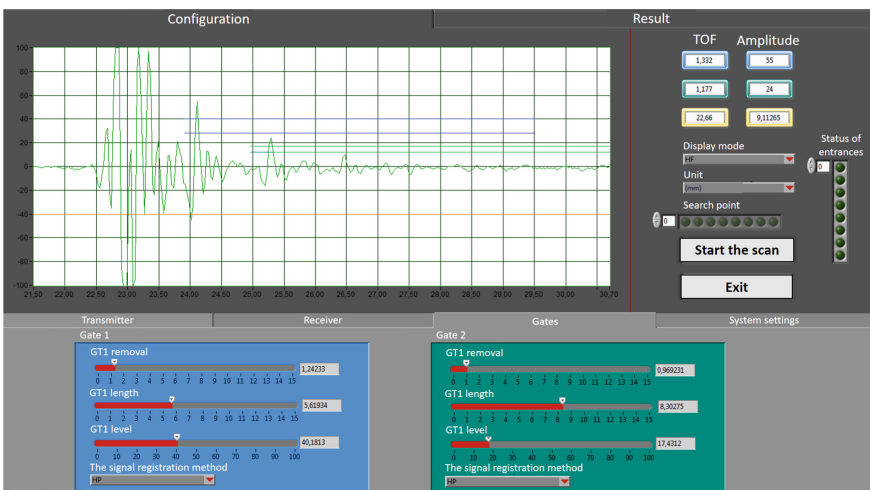
Fig. 12. Measurement software: user panel view.

software also permitted selection of: analysed signal range, signal recording start delay, signal gain (dB) value, frequency band, signal attenuation, and signal averaging method. Figure 12 shows an example of the ImageSonic software user panel.

An example of the measurement report (Fig. 13) was generated as a HTML file at the end of the measurement. The output measurement report featured test item iden-

	Blade 1	Blade 2	Blade 3	Blade 4	Blade 5	Blade 6	Mean	Deviation	Minimum	Maksimum
1	0,68145	1,6107	0,92925	0,8673	1,42485	1,06864	1,09703	0,353769	1,6107	0,68145
2	0,712425	0,665963	0,68145	0,68145	0,665963	0,68145	0,68145	0,0169657	0,712425	0,665963
3	1,03766	0,727912	1,54875	0,712425	0,80535	0,68145	0,918925	0,334425	1,54875	0,68145
4	1,40936	0,92925	0,665963	0,7434	0,92925	1,70363	1,06347	0,40643	1,70363	0,665963
5	1,37839	1,51777	0,665963	1,4868	1,54875	0,68145	1,21319	0,421837	1,54875	0,665963
6	0,712425	1,17705	0,68145	1,31644	0,898275	0,882787	0,944738	0,253541	1,31644	0,68145
7	0,9912	0,898275	0,7434	0,712425	1,13059	0,712425	0,864719	0,172577	1,13059	0,712425

Fig. 13. Example measurement report.

tification, thickness values at individual measurement points on the vanes, and statistical data (mean, deviation, minimum and maximum values per data line).

6   Conclusion

This paper presents a robotic geometry inspection stand equipped with a scanner device, a GOM software suite and an ABB robot. The paper discusses a method of automatic geometry scanning, and the results produced by that process are compared with a measurement model (nominal) to perform an initial geometry verification of the casting and determine 168 specific coordinates for the actual measurement points that are necessary to generate the motion trajectory for the other robotic test stand, used for the UTT measurement of the precision thin-wall casting of the vanes. The correctness of the solution has been proved with scans and measurements of two turbine rotor guide vane segments of an aircraft engine provided by courtesy of Consolidated Precision Products Poland sp. z o.o. The maximum correction of the measurement point coordinates for the two stator segments was approximately 0.4 mm. The paper presents a robotic test stand for UTT quality control. The design and fabrication process of the test stand was not discussed here. The operating principle of its measurement system is presented, with a specific focus on the measurement strategy. Following a verification of the presented solutions, the measurement time per single stator was reduced from 6 h to 15 min. With the future development of optical scanning systems and the increasing accuracy of scanned data, the use of this system will undoubtedly widen to include other areas of inspection, as an alternative to conventional measurement methods.

References

1. Gierlak, P., Burghardt, A., Szybicki, D., Szuster, M., Muszyńska, M.: On-line manipulator tool condition monitoring based on vibration analysis. Mech. Syst. Signal Process. **89**, 14–26 (2017)
2. Giergiel, J., Kurc, K.: Identification of the mathematical model of an inspection mobile robot with fuzzy logic systems and neural networks. J. Theor. Appl. Mech. **49**, 209–225 (2011)
3. Tutak, J.S.: Virtual reality and exercises for paretic upper limb of stroke survivors. Tehnički vjesnik **24** (Supplement 2), 451–458 (2017)
4. Kurc, K., Szybicki, D., Burghardt, A., Muszyńska, M.: The application of virtual prototyping methods to determine the dynamic parameters of mobile robot. Open Eng. **6** (1), 55–63 (2016)
5. Tutak, J.S.: Design of ELISE robot for the paretic upper limb of stroke survivors. J. Vibroengineering **18**(6), 4069–4085 (2016)
6. Kohut, P., Kurc, K., Szybicki, D., Cioch, W., Burdzik, R.: 1823. Vision-based motion analysis and deflection measurement of a robot's crawler unit. J. Vibroengineering **17**(8), 4112–4121 (2015)
7. Kohut, P., Holak, K., Dworakowski, Z., Mendrok, K.: Vision-based measurement systems for static and dynamic characteristics of overhead lines. J. Vibroengineering **18**(4), 2113–2122 (2016)
8. Trojnacki, M., Dąbek, P.: Determination of motion parameters with inertial measurement units–Part 2: algorithm verification with a four-wheeled mobile robot and low-cost MEMS sensors. In: Mechatronics-Ideas for Industrial Application, pp. 253–267. Springer, Cham (2015)
9. Kohut, P., Holak, K., Martowicz, A.: An uncertainty propagation in developed vision based measurement system aided by numerical and experimental tests. J. Theor. Appl. Mech. **50** (4), 1049–1061 (2012)
10. Burghardt, A., Szybicki, D., Kurc, K., Muszyńska, M., Mucha, J.: Experimental Study of Inconel 718 surface treatment by edge robotic Deburring with force control. Strength Mater. **49**(4), 594–604 (2017)
11. Sieniawski, J.: Nickel and titanium alloys in aircraft turbine engines. Adv. Manuf. Sci. Technol. **27**(3), 23–33 (2003)
12. Carl, V., Becker, E., Sperling, A.: Thermography inspection system for gas turbine blades. In: 7th European Conference on non-destructive testing, pp. 2658–2665. Copenhagen (1998)
13. Kilian, D.: 3D tomography of turbine blades. In: Proceedings of the International Symposium of Computerized Tomography for Industrial Applications, vol. 31, pp. 1–17. Berlin (1999)
14. Le Bihan, Y., Joubert, P.Y., Placko, D.: Wall thickness evaluation of single-crystal hollow blades by eddy current sensor. NDT E Int. **34**(5), 363–368 (2001)
15. Lane, C.: The development of a 2D ultrasonic array inspection for single crystal turbine blades. Springer Theses. pp. 63–79. Springer International Publishing, Switzerland (2014)
16. Krause, M., et al.: Comparison of pulse-echo methods for testing concrete. NDT E Int. **30**(4), 195–204 (1997)
17. Chakroun, N., Fink, M.A., Wu, F.: Time reversal processing in ultrasonic nondestructive testing. IEEE Trans. Ultrasonics Ferroelectrics Freq. Control **42**(6), 1087–1098 (1995)
18. Schindel, D.W., Hutchins, D.A., Grandia, W.A.: Capacitive and piezoelectric air-coupled transducers for resonant ultrasonic inspection. Ultrasonics **34**(6), 621–627 (1996)

19. Garnier, C., Pastor, M.L., Eyma, F., Lorrain, B.: The detection of aeronautical defects in situ on composite structures using non destructive testing. Compos. Struct. **93**(5), 1328–1336 (2011)

20. D'orazio, T., Leo, M., Distante, A., Guaragnella, C., Pianese, V., Cavaccini, G.: Automatic ultrasonic inspection for internal defect detection in composite materials. NDT e Int. **41**(2), 145–154 (2008)

21. Tanase, L., Margaritescu, M.: Quality control of product/process using nondestructive control and possibilities of robotic investigation. In: SIOEL'99: Sixth Symposium on Optoelectronics, vol. 4068, pp. 364–375. International Society for Optics and Photonics (2000, February)

22. Yao, Y., Liu, S., Zhang, W.: Regeneration of silica gel using high-intensity ultrasonic under low temperatures. Energy Fuels **23**(1), 457–463 (2008)

23. Steiner, K.V.: Defect classifications in composites using ultrasonic nondestructive evaluation techniques. In: Damage Detection in Composite Materials. ASTM International (1992)

24. Kersting, T., Schönartz, N., Oesterlein, L., Liessem, A.: High end inspection by filmless radiography on LSAW large diameter pipes. NDT E Int. **43**(3), 206–209 (2010)

25. Mineo, C., Pierce, S.G., Nicholson, P.I., Cooper, I.: Robotic path planning for non-destructive testing–A custom MATLAB toolbox approach. Robot. Comput. Integr. Manuf. **37**, 1–12 (2016)

26. Burghardt, A., Kurc, K., Szybicki, D., Muszyńska, M., Szczęch, T.: Robot-operated inspection of aircraft engine turbine rotor guide vane segment geometry. Tehnicki Vjesnik-Technical Gazette **24**(Supplement 2), 345–348 (2017)

27. Liu, B., et al.: Virtual plate pre-bending for the long bone fracture based on axis pre-alignment. Comput. Med. Imaging Gr. **38**(4), 233–244 (2014)

28. Lavoué, G., Tola, M., Dupont, F., Lavou, G.: MEPP-3D mesh processing platform. In: GRAPP/IVAPP, 206–210 (2012)

29. Brajlih, T., et al.: Possibilities of using three-dimensional optical scanning in complex geometrical inspection. Strojniški vestnik-Journal of Mechanical Engineering **57**(11), 826–833 (2011)

30. GOM. Quality control of injection moulded parts, from http://www.gom.com/EN/B0C.html. Accessed 31 May 2009

31. Dúbravčík, M., Kender, Š.: Application of reverse engineering techniques in mechanics system services. Proced. Eng. **48**, 96–104 (2012)

32. Paulic, M., Irgolic, T., Balic, J., Cus, F., Cupar, A., Brajlih, T., Drstvensek, I.: Reverse engineering of parts with optical scanning and additive manufacturing. Proced. Eng. **69**, 795–803 (2014)

33. Pogacar, V.: Integrated renaissance of design. In: Proceedings of the 1st DAAAM International specialized conference on additive technologies, pp. 9–12. Celje (2007)

34. Min, Z.H.O.U.: A new approach of composite surface reconstruction based on reverse engineering. Proced. Eng. **23**, 594–599 (2011)

35. Burghardt, A., Kurc, K., Szybicki, D., Muszyńska, M., Szczęch, T.: Monitoring the parameters of the robot-operated quality control process. Adv. Sci. Technol. Res. J. **11**(1), 232–236 (2017)

36. Burghardt, A., Kurc, K., Szybicki, D., Muszyńska, M., Nawrocki, J.: Robot-operated quality control station based on the UTT method. Open Eng. **7**(1), 37–42 (2017)

37. Burghardt, A., Kurc, K., Szybicki, D., Muszyńska, M., Nawrocki, J.: Software for the robot-operated inspection station for engine guide vanes taking into consideration the geometric variability of parts. Tehnicki Vjesnik-Technical Gazette **24**(Supplement 2), 349–353 (2017)

New Measures of Operational Readiness for Multi-states Avionics Integrated Systems with Reduced Efficiency

Jerzy Lewitowicz, Mariusz Zieja, Andrzej Szelmanowski,
and Andrzej Pazur[(✉)]

Air Force Institute of Technology, Warsaw, Poland
poczta@itwl.pl

Abstract. The paper presents a method of analysis and determination of partial operational readiness coefficients and a way of its assessment, as well as guidelines for making decisions in the field of rational management and efficient use of the aircraft with equipment that is partially non-airworthy under conditions of the occurrence of, e.g. armed conflict, state threat, and crisis state. It was indicated that for efficient management of the operation of military aircraft and helicopters under conditions of ensuring flight safety, it is necessary to support a decision-making process with the use of IT systems that allow for current determination of the partial and complete operational readiness level of individual aircraft on-board systems with many states with reduced usability.

Keywords: Integrated communications systems ·
Possibilities and conditions of forming operational readiness aircraft
Reliability centered maintenance

1 Introduction

The operational readiness of aircraft is a level of readiness of on-board systems, as well as aviation personnel, aviation and engineering service, logistic equipment, and any equipment that is necessary to learn, prepare, and perform the initial operating capability and to develop the ways of checking the ability to perform them [4, 5, 13]. In the Polish military aviation, in the conditions of peace, the applied priority is to absolutely ensure the flight safety, while resigning from the performance of air tasks with the use of the un air worthy aircraft. It is possible to undertake the implementation of these tasks, when the aircraft is improved and restored to the state of airworthiness. In the situations of a direct risk, it may be necessary to use the aircraft with equipment usable for the selected mission implementation, including damage to the on-board devices, which do not directly affect the flight safety and will not be used during the performance of a given task [17, 18].

Such cases occur in the conditions, when the mission performance requires the usability of the selected on-board systems and devices, which are used for the implementation of, among others, a search and rescue mission (SAR) and a combat search and rescue mission (CSAR). The situation may relate to the on-board systems with many partial usability states using component devices participating in the selected modes of the system operation, depending on the type of the performed task [10, 12].

The aircraft readiness is a property characterising the usability to undertake an air task immediately or at a given time t with a forecast of its successful completion within the time interval τ. The aircraft can perform various aerial operations and be in one of the selected reliability states (airworthiness, non-airworthiness, partial airworthiness) and of the operational states (combat duty performance, aircraft subject to operation, renovation and diagnosis, etc.). The appropriate technical condition of a given aircraft is a condition insufficient to perform the above-mentioned tasks. As a condition for the air task implementation, it is necessary to perform logistic undertakings providing the condition of readiness. Therefore, readiness is considered in three aspects as: initial readiness, technical readiness and operational readiness [2, 5].

Readiness as a measure of quality of the aircraft and its related operation system, determines the ability to provide the military aircraft or helicopter operation within the determined time intervals and the ability to the maximum operation time within the considered operation period, timely task performance, and also to the maximum time of the combat duty while minimising the preparation time to start random tasks occurring in the intervention systems, among others, defensive, protection and armed conflict ones.

The aircraft readiness measures in a given operation system are the probabilistic characteristics of the time of the aircraft and operation system staying in the states ensuring the ability to operate in the desirable states.

The Air Force Institute of Technology (AFIT), as the first one in Poland, has developed and built an Integrated Communication System (ZSŁ) and it has taken the measures aimed at development of new operational readiness measures, systems increasing the reliability and functionality of the communication system with many states with reduced usability [8, 10, 19].

2 Integrated Communication System as a Complex Avionic System

In order to improve the situational awareness of the crews operating *Mi8/Mi17/Mi24 and W-3PL* helicopters, the integrated communication system provides a minimum set of air and tactical radios, which is necessary to implement the task (Fig. 1). It protects overt and covert communication with the use of frequency coding (the so-called TRANSEC) and speech and data encryption (the so-called COMSEC). Owing to their help, the crew during the whole performed flight is provided with communication and control of their parameters during the flight. The status of use of individual on-board radios is illustrated on the communication control panels and/or multifunction monitors [2, 8].
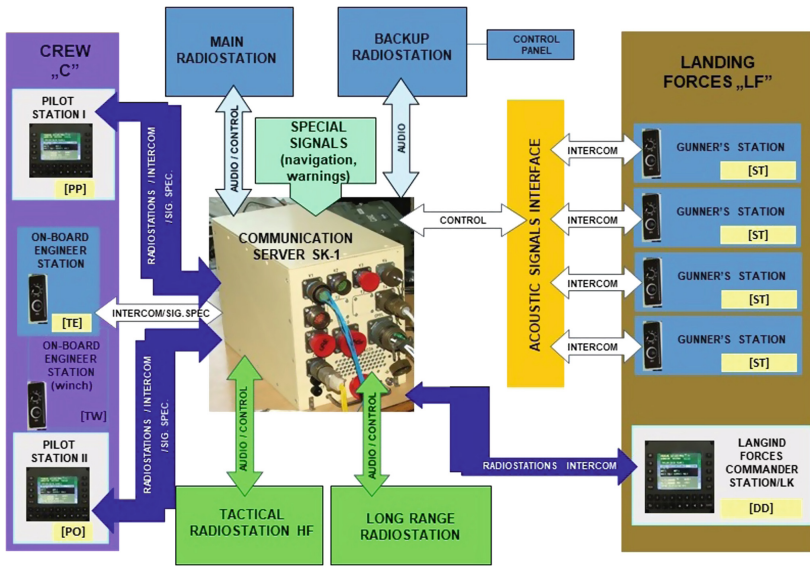
Fig. 1. View of the integrated communication system architecture [8]

The integrated communication system, like modern Western solutions [6, 13], is based on IT technologies and it constitutes an analogy of a computer system based on a digital data exchange bus in accordance with the adopted standard, among others, MIL-1553B (is a military standard the mechanical, electrical, and functional characteristics of a serial data bus.).

The main element of the integrated communication system is a communication server (SK-1), which is the "heart of the system", and it provides the control and management of the on-board radio communication network. It controls the internal and external communication system in the crew circuit, as well as within the crew circuit via communication control panels or multifunction monitors. It provides the helicopter crew with special signals, including disconnectable navigation signals (e.g. marker) and non-disconnectable –warning signals (e.g. dangerous flight altitude) [7, 8, 15].

The integrated communication system is designed to provide the external communication with subscribers outside the helicopter (air traffic control, command posts, army subunits, other aircraft) and internal communication on the helicopter board between the crew members and the troop compartment. The connections are implemented from the designated communication positions without the necessity of replacing the headset–, regardless of the type of the carried-out radio communication. The internal communication system consists of communication positions equipped with the communication control panels and user plates.

The internal communication system provides the crew with the playback of special signals (coming from the transponder systems, radio altimeter), voice signals

Fig. 2. Architectural view of the integrated communication system on *W-3PL* aircraft [8]

(RI65 voice warning system) and signals from navigation devices (VOR/LOC, TACAN, MARKER, ARK/UD/2). The external communication system consists of four air and tactical radios providing the frequency band coverage in the range of 1.6÷400 MHz, the use of which is available from the one selected among three communication control panels and/or multifunction monitors mounted on the boards of *Mi8/Mi17/Mi24 and W-3PL* helicopters (Fig. 2).

The application software, which was developed by AFIT specialists, is implemented on the communication server [8–10].

The communication server elements during operation are subject to ageing, degradation and wear processes. The provision of radio communication and flight safety requires the fulfilment of many design requirements by the communication server, and above all –the appropriate level of its reliability and adequate life and durability [10].

In order to test the application software of individual air and tactical radios, the Integration Station [9, 14], the task of which is to optimise the developed communication systems in terms of their architecture, organisation and detailed management of individual modes of operation (Fig. 3).

Fig. 3. View of the Integration Station of the integrated communication system [8]

3   Determination of Readiness Measures of Aircraft

The determination of readiness measures includes a phase of the identification process and creation of calculation algorithms as well as a phase of measurements and data collection necessary to determine the values of the parameters used in the algorithms [1, 2, 6, 11]. The above-discussed sets of operational states of aircraft and the distinguished states of airworthiness create the states of readiness. The sets of states that allow for the air task performance or operation in the determined time interval are called task readiness states, and those which allow for correct operation at a given moment t are called technical readiness states. The sets of states, which give the opportunity to start the task performance after the set time are the initial readiness states, the sets providing the opportunity to start the task after the set time are known as potential readiness states [5].

The aircraft readiness can be treated as the probability of an event that the aircraft being at the moment t in technical readiness started the implementation of the air task reported at the t after the time $\Delta\Theta$ shorter than Θ, and it will perform the task within the time interval τ, which means:

$$G(t, \Theta, \tau) \,=\, K_{t_1} \,\cdot\, \mathrm{P}[\Delta\Theta \leq \Theta, \tau] \tag{1}$$

where:

K_{t_1} – technical readiness; $\mathrm{P}[\Delta\Theta \leq \Theta]$ – probability of readiness to perform the air task.

In case of high values $t(t \to \infty)$ the above-described relationship can be presented in the following form:

$$G(\infty, \Theta) \,=\, K_{t_1} \left\{ 1 - \exp\left[- \int_0^\Theta \lambda(\vartheta)\mathrm{d}\vartheta \right] \right\} \tag{2}$$

where: $\lambda(\vartheta)$ – intensity of changing the readiness states within the time interval $[t, \Theta]$; ϑ – current time for the time interval $\Delta\Theta$.

In the operation process, the aircraft can be in the states of airworthiness or in the states of non-airworthiness. From the states of airworthiness, it is possible to distinguish a subset of the states of airworthiness of the aircraft to perform a given task and a subset of the states of airworthiness to operation in case of the task performance, but without the possibility of its completion. Among the non-airworthiness states of the aircraft to be used, one can distinguish a subset of such states, in which –within the framework of the time reserve, an appropriate maintenance task can be performed, which introduces the aircraft into a subset of usability –the task performance or into the subset of usability states to perform this task [5, 19].

An example can be the state of potential readiness of the aircraft to perform a given air task for the set time reserve, which is included in the set of states of airworthiness and non-airworthiness or in the set of states with the insufficient usability potential for proper operation in order to perform the task.

It is forecasted that within the time reserve, the service position will be started and the service introducing the aircraft into the state of readiness for the air task

performance will be implemented [2, 9]. The shorter the time reserve, the higher the degree of readiness. The state of readiness at this reserve is a state of task readiness of the aircraft for a given task.

The operation strategy according to the status of the reliability control level (Eng. Reliability-Centered Maintenance –RCM) is used for systems that are required to have high operational reliability due to flight safety. The research and assessment of reliability within the strategy is conducted with the use of statistical methods for events and computer simulation technology methods, the so-called reliability testing programming. The purpose of using this method is not to ensure the maximum reliability of devices on the aircraft board, but to provide it on such a level that is required by the function implementation by the object under given conditions of the task performance [18].

4   New Operational Readiness Measures for the Complex System

The integrated communication system has various types of the operation modes, with the use of component devices for their implementation. It results in the fact that it can be treated as a complex system having many states of partial usability, which enable the implementation of selected air tasks and missions (Fig. 4).
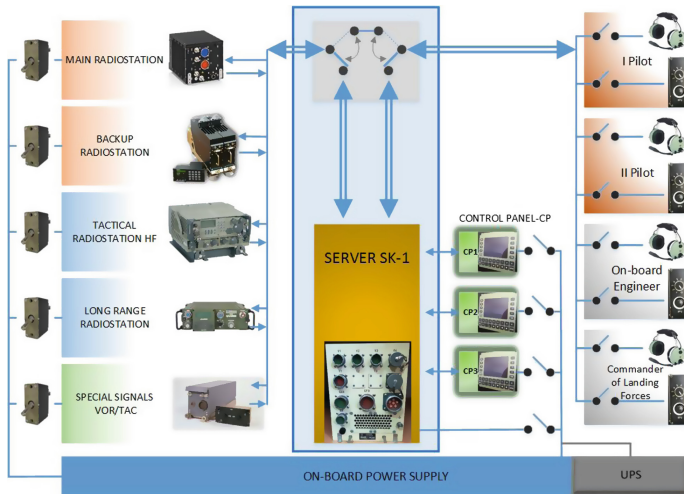
Fig. 4.   The operation organisation of the integrated communication system [8]

For example, in order to implement the search and rescue tasks (SAR), the system must have two efficient air radios, however, in combat operations (CSAR), a necessary condition is to provide the covert (coded) air communication. In addition, with the use of troop groups, the aircraft crew must be in contact with the troop compartment in order to transfer commands and supplementary information to the landing group commander and on-board shooters.

4.1 Tasks Performed in Peacetime and the States of Higher Combat Readiness by the Integrated Communication System

In the time of peace (P state), the integrated communication system with an open architecture that allows for mounting both on the military and civil aircraft is provided for the following tasks [8]:

- The SAR rescue mission (complete usability of the system is required, with no use of covert/coded communication);
- Transport of living force (required communication with the troop compartment);
- Freight transport (communication with the troop compartment is not required);
- Air operations in P time (covert/coded communication is not required).

In the event of threat, crisis, catastrophes, natural disasters and war (W state), the integrated communication system is intended for the following tasks [8]:

- CSAR mission (required complete usability of the system, covert/coded communication is used);
- Freight transport (communication with the troop compartment is not required);
- Transport of living force (communication with the troop compartment is required);
- Other air tasks in W time (covert/coded communication is not required).

On the basis of the analysis of the above tasks, it is possible to distinguish the states of usability of the integrated communication system, and hence, several configurations covering a specific set of component devices with the set states of usability, for which the operational readiness, which is required to perform particular tasks included in a set of P time and W time tasks, will be maintained.

The configurations were presented in Table 1.

On the basis of the obtained analyses, it is possible to distinguish the following categories of the tasks implemented by the integrated communication system with many states with reduced usability [8]:

A—category of air tasks possible to be performed by the system, i.e. complete state of usability (all elements of the integrated communication system are efficient);
B—category of air tasks possible to be performed by the integrated communication system, in case, when it is in the state of reduced usability, i.e. one inefficient device managing the operation of the entire system (unfit SK-1);

Table 1. Configurations of a set of component devices with usability states

Configuration of the system	Required composition of the System	Performed task	Probability/period of occurrence
System with SK-1	Integrated communication system	Flight SAR, CSAR	System with SK-1
System without covert/coded communication	Integrated communication system	Freight transport	System without tactical communication
System without tactical communication	Air radio without the integrated communication system	Living force transport	System without SK-1
System without SK-1	Air radio without the integrated communication system	Other air tasks in W time	System without SK-1

C—category of air tasks possible to be performed by the integrated communication system, in case, when it is in the state of reduced usability with one inefficient device of the system (unfit tactical communication radio);

D—category of air tasks possible to be performed by the integrated communication system, in case, when it is in the state of reduced usability with many inefficient component devices of the entire system, i.e. one inefficient device managing the operation of the entire system and one inefficient device of the system (unfit SK-1 and tactical communication radio).

The integrated communication system staying in the category A means that it can perform all the tasks provided for the aircraft, e.g. (SAR, CSAR) with the use of air and tactical communication, a change in the overt and covert communication parameters with the use of the communication control panels, the radio communication with the troop compartment is carried out during the task performance.

The integrated communication system staying in the category B means that it can perform only the tasks provided for the aircraft, which do not require the use of SK-1 server, e.g. task performance with the use of air communication, but without the need to use PSŁ-1 communication control panels (communication parameters are constant or introduced during flight with the use of the backup radio panel).

The integrated communication system staying in the category C means that it can perform only the tasks provided for the aircraft, which do not require the use of the tactical radio, e.g. living force transport, freight transport. The use of the air radio with the necessity of using SK-1 server (management of the system operation and recording of messages) and the communication control panels (change in radio parameters and radio frequency ranges during the task performance).

The integrated communication system staying in the category D means that it can perform only the tasks provided for the aircraft, which do not require the use of SK-1 server and the tactical radio, e.g. task performance with the use of a backup air radio, without the need to use the communication control panels (communication parameters are constant or introduced during flight with the use of the radio control panel).

4.2   Applied Operational Readiness Measures in the States of Higher Combat Readiness

In the available specialist and standardisation literature [20–23], the mathematical relationship describing the determination of the operational readiness coefficient of the system with the use of the probabilities of staying in only one of two possible states: in the state of usability or in the state of unfitness, are available [16].

For a system (or device composed of many elements) with one state of usability, the operational readiness is specified in the following form [19]:

$$G(t, \tau) = G_F(t) \cdot G_Z(\tau) = K_g(t) \cdot R(\tau) \qquad (3)$$

where: $G(t, \tau)$ – operational readiness of the system determined for the selected moment of time t and the selected time interval of the task duration τ; $G_F(t)$ – functional readiness of the system determined for the selected moment of time t; $G_Z(\tau)$ –

task readiness of the system determined for the selected time interval of the task duration τ; $K_g(t)$ functional readiness coefficient; $R(\tau)$ – system reliability function determined for the selected time interval of the task duration τ.

For a system (or device composed of many elements) with one state of usability, the functional readiness coefficient can be determined in the form of the probability or time of staying in the state of usability in relation to the total probability or time of the system staying in the states of usability and unfitness [19]:

$$G_F(t) = K_g(t) = \frac{p_Z(t)}{p_Z(t) + p_N(t)} \text{ lub } G_F(t) = K_g(t) = \frac{T_Z(t)}{T_Z(t) + T_N(t)} \quad (4)$$

where: $p_Z(t)$ – probability of the system staying in the state of usability in the selected moment t;

$p_N(t)$ – probability of the system staying in the state of unfitness in the selected moment t; $T_Z(t)$ – average time of the system staying in the state of usability to the selected moment t; $T_N(t)$ – average time of the system staying in the state of unfitness to the selected moment t.

The data needed to determine the time of staying in individual operational states are obtained from IT systems (data base of the Integration Station of the Integrated Communication System of the Avionics Division of AFIT) or paper records kept by aviation units, which use the integrated communication system.

For a system (or device composed of many elements) with one state of usability, the task readiness coefficient can be determined on the basis of knowledge of the reliability function determined for the time interval of the task performance [19]:

$$G_Z(\tau) = R(\tau) = e^{-\lambda \cdot \tau} \quad (5)$$

where: λ – intensity of the system damage determined on the basis of the operational data or determined at the design stage (within the framework of the implementation of the adopted reliability management strategy).

On the basis of the above, the operational readiness of the system can be determined in the following form [19]:

$$G(t, \tau) = K_g(t) \cdot R(\tau) = \frac{p_Z(t)}{p_Z(t) + p_N(t)} \cdot e^{-\lambda \cdot \tau} \quad (6)$$

Example:
For the data, determined on the basis of the analysis, collected in the IT system of the Integration Station of the Integrated Communication System, for the integrated communication system operated in *Mi8/Mi17/Mi24* helicopters, analysed as a set of three components connected in series in the reliability chain (SK-1 server, tactical radio and the so-called core of the system including other devices, among others, communication control panels, headphones, microphones, user plates), the following was adopted:

- The functional readiness coefficient of the system in the state of complete usability is: K_g= *0.900;*

- The reliability function coefficient of the system in the state of complete usability is:
  $R(\tau) = 0.729$

  hence, the calculated operational readiness is:

$$G(t, \tau) = K_g(t) \cdot R(\tau) = 0,900 \cdot 0,729 = 0,656 \qquad (7)$$

The operational readiness of the system (or device), determined in the above manner, with one state of usability specifies the probability of the event that the integrated communication system will be usable at a given time t and that it will maintain this usability during the task performance with the selected time interval τ.

4.3 New Operational Readiness Measures in the States of Higher Combat Readiness

A new approach to determining the operational readiness measures is based on an attempt of using the current standard-sanctioned description of the class of devices of the integrated communication system, operating in many states with reduced usability. It applies both to one device and the entire system, which consists of many devices, i.e. with one state of usability or many states of usability [3, 19].

The proposed new measures relate to the system description (or device composed of many elements), which in addition to the state of complete usability, where all its elements are efficient, can stay in one of the states of reduced usability, which allows the performance of the selected tasks, for which all elements of the system are not required to be in the state of its usability.

For the system (or device composed of many elements) with many states of usability (complete and reduced usability), the functional readiness coefficient can be determined in the form of probability or time of staying in the state of usability in relation to the joint probability of time of the system staying in the states of usability and unfitness:

$$K_g(t) = \frac{\sum\limits_{i=1}^{i=M} p_{Zi}(t)}{\sum\limits_{i=1}^{i=M} p_{Zi}(t) + \sum\limits_{i=1}^{i=M} p_{Ni}(t)} \quad \text{lub} \quad K_g(t) = \frac{\sum\limits_{i=1}^{i=M} T_{Zi}(t)}{\sum\limits_{i=1}^{i=M} T_{Zi}(t) + \sum\limits_{i=1}^{i=M} T_{Ni}(t)} \qquad (8)$$

where:

$\sum\limits_{i=1}^{i=M} p_{Zi}(t)$ – the sum of probabilities of the system staying in the states of complete and reduced usability at a given time t; $\sum\limits_{i=1}^{i=M} p_{Ni}(t)$ – the sum of probabilities of the system staying in the states of unfitness at a given time t; $\sum\limits_{i=1}^{i=M} T_{Zi}(t)$ – the sum of average times of the system staying in the states of complete and reduced usability at a

given time t; $\sum\limits_{i=1}^{i=M} T_{Ni}(t)$ – the sum of probabilities of the system staying in the states of unfitness to a given time t.

The above relationship can be presented in a detailed form showing individual probabilities of the system (or device composed of many elements) staying in the state of complete usability and in the states of reduced usability:

$$K_g(t) = \frac{p_{Z1}(t) + p_{Z2}(t) + \ldots + p_{ZM}(t)}{\sum\limits_{i=1}^{i=M} p_Z(t) + \sum\limits_{i=1}^{i=M} p_N(t)} \tag{9}$$

where: $p_{Z1}(t)$ – probability of the system staying in the *1st* state of usability (complete usability); $p_{Z2}(t)$ – probability of the system staying in the *2nd* state of usability (reduced usability); $p_{ZM}(t)$ – probability of the system staying in the *M* state of usability (reduced usability), which can be then written as separate components of the sum in the following form:

$$K_g(t) = \frac{p_{Z1}(t)}{\sum\limits_{i=1}^{i=M} p_Z(t) + \sum\limits_{i=1}^{i=M} p_N(t)} + \frac{p_{Z2}(t)}{\sum\limits_{i=1}^{i=M} p_Z(t) + \sum\limits_{i=1}^{i=M} p_N(t)} + \ldots + \frac{p_{ZM}(t)}{\sum\limits_{i=1}^{i=M} p_Z(t) + \sum\limits_{i=1}^{i=M} p_N(t)} \tag{10}$$

On the basis of the analysis of the above relationship, it can be concluded that the readiness coefficient for the system (or device composed of many elements) with many states of usability (complete and reduced usability) can be presented as the sum of component coefficients:

$$K_g(t) = K_{g1}(t) + K_{g2}(t) + \ldots + K_{gM}(t) \tag{11}$$

where:

$K_{g1}(t)$ – readiness coefficient of the system staying in the *1st* state of usability (complete usability); $K_{g2}(t)$ – readiness coefficient of the system staying in the *2nd* state of usability (reduced usability); $K_{gM}(t)$ – readiness coefficient of the system staying in the *M* state of usability (reduced usability).

$$K_g(t) = \sum\limits_{i=1}^{i=M} K_{gi}(t) \tag{12}$$

In order to emphasise the significance of the states of usability, the operational readiness coefficient can be presented in the form dependent on the weight of the partial operational readiness coefficient:

$$K_g^*(t) = W_1 \cdot K_{g1}(t) + W_2 \cdot K_{g2}(t) + \ldots + W_M \cdot K_{gM}(t) \tag{13}$$

where: W_1 – weight of the readiness coefficient of the system staying in the *1st* state of usability (complete usability); W_2 – weight of the readiness coefficient of the system staying in the *2nd* state of usability (reduced usability); W_M – weight of the readiness coefficient of the system staying in the M state of usability (reduced usability), which can be written in the following form:

$$K_g^*(t) = \sum_{i=1}^{i=M} W_i \cdot K_{gi}(t) \tag{14}$$

The functional readiness coefficient measures in the modified version can be written in the following form: for the weight $W_1 = 1$ and other weights $W_2 \div W_M = 0$.

Then, the new measure adopts the form of the functional readiness coefficient, previously used for the system staying only in the state of complete usability.

The measures for the functional readiness coefficient of the system with many states of reduced usability can be written in the form of a matrix:

$$\begin{bmatrix} K_{g1}^*(t) \\ K_{g2}^*(t) \\ \cdots \\ K_{gM}^*(t) \end{bmatrix} = \begin{bmatrix} K_{g1}(t) & 0 & \cdots & 0 \\ 0 & K_{g2}(t) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & K_{gM}(t) \end{bmatrix} x \begin{bmatrix} W_1 \\ W_2 \\ \cdots \\ W_M \end{bmatrix} \tag{15}$$

then, the operational readiness can be presented in the matrix form:

$$\begin{bmatrix} G_{11}(t,\tau) \\ G_{12}(t,\tau) \\ \cdots \\ G_{1M}(t,\tau) \end{bmatrix} = \begin{bmatrix} K_{g1}(t) & 0 & \cdots & 0 \\ K_{g1}(t) & K_{g2}(t) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ K_{g1}(t) & K_{g2}(t) & \cdots & K_{gM}(t) \end{bmatrix} x \begin{bmatrix} R_{11}(\tau) \\ R_{12}(\tau) \\ \cdots \\ R_{1M}(\tau) \end{bmatrix} \tag{16}$$

The operational readiness determined for the system with many states and reduced usability presented in the matrix form allows to assess the use of the system in aerial operations, for which the usability of all component elements of the system is not required.

For these elements, the matrix record allows to determine the required reliability value for the complex operational readiness value. The reliability levels can be determined with the use of a matrix of the operational readiness coefficients after its reversal (matrix reversibility condition must be met), which can be presented in the following form:

$$\begin{bmatrix} R_{11}(\tau) \\ R_{12}(\tau) \\ \cdots \\ R_{1M}(\tau) \end{bmatrix} = \begin{bmatrix} K_{g1}(t) & 0 & \cdots & 0 \\ K_{g1}(t) & K_{g2}(t) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ K_{g1}(t) & K_{g2}(t) & \cdots & K_{gM}(t) \end{bmatrix}^{-1} x \begin{bmatrix} G_{11}(t,\tau) \\ G_{12}(t,\tau) \\ \cdots \\ G_{1M}(t,\tau) \end{bmatrix} \tag{17}$$

The determination of the required reliability values allows to specify its level at the design stage or at the operational stage (by introducing correction changes, i.e.

replacement of elements with reduced reliability into elements with reliability determined in the above manner. The new measures may constitute a tool for supporting the operation process according to the strategy with controlled reliability [2, 5, 6, 18].

Example:
For the data, determined on the basis of the analysis, collected in the IT system of the Integration Station of the Integrated Communication System, for the integrated communication system operated in *Mi8/Mi17/Mi24* helicopters, analysed as a set of three components connected in series in the reliability chain (SK-1 server, tactical radio and the so-called core of the system including other devices, among others, communication control panels, headphones, microphones, user plates), the following was adopted:

- The functional readiness coefficient of the system in the state of complete usability is: $K_g = K_{g1} = 0.900,$ and in case of reduced usability, it is: $K_{g2} = 0.050;$ $K_{g3} = 0.030; K_{g4} = 0.010;$
- The reliability function coefficient of the system in the state of complete usability is: $R(\tau) = R_{11}(\tau) = 0.729,$ and in case of reduced usability, it is: $R_{12}(\tau) = 0.810;$ $R_{13}(\tau) = 0.810; R_{14}(\tau) = 0.900$

thus, the calculated operational readiness of the system in the state of complete usability is:

$$G(t, \tau) = K_g(t) \cdot R(\tau) = 0,900 \cdot 0,729 = 0,656 \tag{18}$$

The use of the new measure allows to present the operational readiness of the integrated communication system staying in many states of reduced usability in the following form:

$$\begin{bmatrix} G_{11}(t,\tau) \\ G_{12}(t,\tau) \\ G_{13}(t,\tau) \\ G_{14}(t,\tau) \end{bmatrix} = \begin{bmatrix} K_{g1}(t) & 0 & 0 & 0 \\ K_{g1}(t) & K_{g2}(t) & 0 & 0 \\ K_{g1}(t) & K_{g2}(t) & K_{g3}(t) & 0 \\ K_{g1}(t) & K_{g2}(t) & K_{g3}(t) & K_{g4}(t) \end{bmatrix} x \begin{bmatrix} R_{11}(\tau) \\ R_{12}(\tau) \\ R_{13}(\tau) \\ R_{14}(\tau) \end{bmatrix} \tag{19}$$

where: $G_{11}(t, \tau)$ – operational readiness of the system staying in the state of complete usability (all elements are usable); $G_{12}(t, \tau)$ – operational readiness of the system staying in the state of reduced usability (unfit SK-1 server); $G_{13}(t, \tau)$ – operational readiness of the system staying in the state of reduced usability (unfit tactical radio); $G_{14}(t, \tau)$ – operational readiness of the system staying in the state of reduced usability (unfit SK-1 server and tactical radio).

For the adopted values of the functional readiness coefficient and reliability determined on the basis of the data obtained from the IT system of the Integration Station of the Integrated Communication System, the operational readiness values of the system were obtained with the use of the following relationship (19):

$$\begin{bmatrix} G_{11}(t,\tau) \\ G_{12}(t,\tau) \\ G_{13}(t,\tau) \\ G_{14}(t,\tau) \end{bmatrix} = \begin{bmatrix} 0,656 \\ 0,697 \\ 0,721 \\ 0,730 \end{bmatrix} = \begin{bmatrix} 0,900 & 0 & 0 & 0 \\ 0,900 & 0,050 & 0 & 0 \\ 0,900 & 0,050 & 0,030 & 0 \\ 0,900 & 0,050 & 0,030 & 0,010 \end{bmatrix} x \begin{bmatrix} 0,729 \\ 0,810 \\ 0,810 \\ 0,900 \end{bmatrix}$$

(20)

The new measures also allow to implement a reverse process consisting in determining the necessary reliability level of components at the set operational readiness of the system.

The values of the required reliability of individual configurations for the adopted operational readiness level of the system (70% in the state of complete usability, 80% in the state of reduced usability without SK-1 server and the tactical radio), can be determined with the use of the relationship (17) and presented in the matrix form:

$$\begin{bmatrix} R_{11}(\tau) \\ R_{12}(\tau) \\ R_{13}(\tau) \\ R_{14}(\tau) \end{bmatrix} = \begin{bmatrix} 0,778 \\ 0,900 \\ 0,900 \\ 0,900 \end{bmatrix} = \begin{bmatrix} 0,900 & 0 & 0 & 0 \\ 0,900 & 0,050 & 0 & 0 \\ 0,900 & 0,050 & 0,030 & 0 \\ 0,900 & 0,050 & 0,030 & 0,010 \end{bmatrix}^{-1} x \begin{bmatrix} 0,700 \\ 0,745 \\ 0,772 \\ 0,800 \end{bmatrix}$$

(21)

For the increased requirements in the field of the reliability values for the adopted operational readiness level of the system (70% in the state of complete usability, 85% in the state of reduced usability without SK-1 server and the tactical radio), the reliability level values of individual elements of the system can be determined with the use of the relationship (17) and presented in the matrix form:

$$\begin{bmatrix} R_{11}(\tau) \\ R_{12}(\tau) \\ R_{13}(\tau) \\ R_{14}(\tau) \end{bmatrix} = \begin{bmatrix} 0,833 \\ 0,900 \\ 0,900 \\ 1,000 \end{bmatrix} = \begin{bmatrix} 0,900 & 0 & 0 & 0 \\ 0,900 & 0,050 & 0 & 0 \\ 0,900 & 0,050 & 0,030 & 0 \\ 0,900 & 0,050 & 0,030 & 0,010 \end{bmatrix}^{-1} x \begin{bmatrix} 0,750 \\ 0,795 \\ 0,822 \\ 0,850 \end{bmatrix}$$

(22)

On the basis of the analysis of the obtained results, it can be stated that in order to obtain the operational readiness at the level of 75% (0.750) for the system in the state of complete usability and 85% (0.850) for the system in the state of reduced usability (without SK-1 server and the tactical radio) should be completely usable (not damaged).

The obtained results indicate the achievement of limit reliability values and they can be used for assessing the operation process with the adopted strategy for managing their reliability.

5   Conclusion

In the military aviation, the proposed new measures can be used in the reliability management strategy for operation of aviation systems in the crisis states. However, it requires additional analytical and verification work that is currently conducted in AFIT. Every aircraft, even older generation ones, after installing of a modern communications system shall expand its performance capabilities with previously unavailable applications, which is why the design of such a system in terms of hardware, i.e., purchasing equipment, does not currently present much of a problem, whereas, developing adequate, efficient and reliable software, which meets the requirements in the scope of ensuring internal and external communication on-board an aircraft, is a big challenge. Such a task was attempted by the AFIT, which was the first in Poland to perform integration on modernized helicopters of the Land Forces Command. The possession of such a station allowed AFIT to integrate new radio-electronic equipment while modernizing a W3PL helicopter. An integrated communications system is operated on Mi8/Mi17/Mi24 and W3PL helicopters, which performed combat tasks, i.a., for PKW (Polish Military Contingent) in Iraq and Afghanistan. The target task for an integrated communications system is obtaining a fully functional version on-board an aircraft. Optimizing the elements of an integrated communication system at an integration station is always limited relative to the organization of a system on-board an aircraft. Nonetheless, it allows testing a series of solutions, which are too expensive to be implemented on-board an aircraft for research purposes. One of the main tasks of such optimization is to achieve high operating capability of the constructed equipment.

Improvement and optimization of the operating process of integrated communications systems requires new evaluation methods and IT systems supporting the decision-making process in the scope of defining and shaping the operating readiness.

The data accumulated in the system allows the determination of the intervals between failures, damage intensities and the damage probability distributions, and on that basis, factors characterizing the operating capability of individual elements of a ZSŁ system. The constructed model enables current evaluation and forming of the operating capability of an SK-1 communication server, which is the main component of a ZSŁ system. The proposed new measures for determining the operational readiness for the integrated communication system with many states of reduced usability, and the developed method of analysis can be used in the reliability management strategy for rational (scientific) operation of the on-board aviation systems.

References

1. Barlow, R.E., Proschan, F.: Statistical Theory of Reliability and Testing Probability Models, Eds. Holt, Rienhart and Wilson, New York (1975)
2. Kececioglu, P.: Maintainability, Availability and Operational Readiness Engineering Handbook, Eds. Prentice Hall, New Jersey, (1995)
3. Knopik, L., Migawa, K.: Multi-state model of maintenance policy. Maint. Reliab. **20** (1), 125–130 (2018). Poland

4. Knopik, L., Migawa, K., Wdzięczny, A.: Profit optimization in maintenance system. Polish Martime Res. **1**(89), 193–98 (2016). Poland
5. Lewitowicz, J.: The basics of aircraft operation, Volume 3, Aircraft operation systems Eds. AFIT, Warsaw, Poland (2006)
6. Stalberg, Frederick R.: Handbook of Reliability, Availability, Maintainability and Safety in Engineering Design, Springer, London (2009)
7. Michalak, S., Pazur, A., Szelmanowski, A.: AFIT's laboratory test equipment to optimise the integrated avionics systems for polish military aircrafts. In: IEEE International Workshop on "Metrology for Aerospace", Benevento, Italy (2014)
8. Pazur, A.: Testing the reliability of communication systems based on a specialized communication server, Eds. AFIT, Warsaw, Poland (2010)
9. Pazur, A.: Technology No. 50/43/15 Service and repair of the integrated helicopter communication system Mi8, Mi17, (Mi171 V), Mi24 (every 2 years of operation), Eds. AFIT, Warsaw, Poland (2015)
10. Pazur, A., Szelmanowski, A.: AFIT's laboratory test equipment to optimize the integrated communication systems for polish military helicopters. In: IEEE International Workshop on Metrology for Aerospace, Benevento, Italy (2014)
11. Sherwin, D.J., Bossche, A.: The Reliability, Availability and Productiveness of Systems, Springer (1993)
12. Pazur, A., Szelmanowski, A., Kowalczyk, H., Janik, P.: The polish electronically integrated avionics systems for military aircraft. In: 3rd IEEE International Workshop on "Metrology for Aerospace", Florence, Italy (2016)
13. Collinson, R.P.G.: Introduction to Avionics. Springer, Netherlands (2011)
14. Pazur, A.: Technology no. 106/43/2012 of the Integrated Avionic System of the W3PL helicopter level „D" communication server SK-1 ver.2. AFIT, Warsaw, Poland (2012)
15. Szelmanowski, A.: Integration Standpoint of Avionics Systems Based on Digital Data Buses. AFIT, Warsaw, Poland (2004)
16. Restel, F.: The Markov reliability and safety model of the railway transportation system. Safety and reliability: methodology and applications. In: Proceedings of the European Safety and Reliability Conference, ESREL 2014, CRC Press/Balkema, pp. 303–311, Wroclaw, Poland (2015)
17. Zieja, M., Ważny, M., Stępień, S.: Distribution determination of time of exceeding permissible condition as used to determine lifetimes of selected aeronautical devices/systems. Maint. Reliab. **18**(1), 57–64 (2016). Poland
18. Wróblewski, M.: Operation Strategies of Aircraft – MSG3, Eds. Military University of Technology, Warsaw (2017)
19. Woropay, M., Żurek, J., Migawa, K.: Model of assessment and development of operational readiness of the traffic jam subsystem in the transport system, Eds. ITE, Radom (2003)
20. Standard PN-IEC 60300 3–9 Reliability Management. Guide of Applications. Risk Analysis in Technical Systems
21. Standard PN-EN ISO 31000:2012 Risk Management, Principles and Guidelines
22. Methodology of Risk Management Aviation of The Polish Armed Forces, Ministry of National Defence, Warsaw (2010)
23. Basics of Risk Management in Aviation, Command of the Air Force, DWLOP 55/2010, Warsaw (2010)

Synthesis of Three-Phase Waveforms Using the Selected Microprocessor – Analysis of Constraints

Romuald Masnicki(✉) and Adam Minczyk

Gdynia Maritime University, Morska 81-87, 81-225 Gdynia, Poland
romas@am.gdynia.pl

Abstract. Waveform generators employed in student laboratories are often purchased instruments, while some their functionalities are not required. An important part of many test sets are multi-phase generators with an analog or digital structure. Designed based on the known concept, the construction of a three-phase sinusoidal signal generator and selected test results are presented. The requirements for properties of the output signals of the generator have been specified. The phase shift between sine waves are kept precisely at 120° and the amplitudes of the signals are controlled within the adopted frequency range. Low distortion of signals was also obtained. The model of the developed three-phase digital generator was described and the results of selected research were presented and then discussed. A short analysis of constraints related to the generators' properties, mainly resulting from the features of the elements used in the project, was also carried out.

Keywords: DDS technology · Three-Phase wave generator · Launchpad board

1 Introduction

Waveform generators are widely used in scientific research, student laboratories and in tests on production lines [1–4]. Sine waves are extensively utilized signals as the standard signals in electronic circuit performance test or parameter measurement [5]. Generators of sinusoidal signals should generate a harmonic waveform with the lowest possible distortion and constant frequency and amplitude.

A big advantage of analog generators is low distortion of the output signal. Changing the frequency of the generated waveform without changing its other parameters, which requires changing the values of the frequency setting elements, is already a problem. This problem is significantly increased in case of an analog three-phase signal generator. Modern generators also allow to program the shape of the output waveform. It is very difficult to design an analog generator system that would form waveforms with pre-set parameters. It is much easier to do this with digital generators. Parameters of the generated signal are much easier to control using the digital method of analog signal generation. The most popular digital generation waveform technology is the direct digital synthesis (DDS). The DDS has been long recognized as a superior technology for generating highly accurate, low-distortion output waveforms.

Precise, high-cost DDS generators are purchased to be used in laboratories however often not all of their functions are fully utilized. To meet the requirements of student laboratories, a simplified sine wave generator has been developed in order to achieve required precision, simple structure and easy operation. It can also generate square wave, triangle wave and waves of other shapes, while other functionalities not commonly used were removed.

The DDS is a method of generating an analog wave, usually a sine wave, by generating a signal that changes over time in a digital form, followed by a digital-to-analog conversion. DDS architecture employs digital signal processing (DSP) techniques to generate a digital sine wave representation which is referenced to a highly-stable reference clock. With DDS technology, one cycle of the desired waveform is represented digitally in waveform memory. The contents of the waveform memory are read sequentially, which produces a stream of digital data representing the desired waveform. The stream of data drives the input of a digital-to-analog converter (DAC), whose analog output is a series of voltage steps approximating the desired waveform. Finally, the DAC output is applied to a low-pass filter (LPF) to smooth the voltage steps and produce the final output waveform. As operations in a DDS device are mainly digital, they can offer fast switching between output frequencies, frequency resolution and wide spectrum operation. The ability to generate and accurately control waveforms with different frequencies and profiles has become a key feature of DDS generators.

The block diagram of the DDS generator is shown in Fig. 1. It consists of a microcontroller unit (MCU) with the program controlling the output of data from LUT (Look–up table) memory, LUT memory containing values of waveform samples, DAC converter and low-pass filter (LPF). The LUT is a commonly used technique that replaces the runtime computation of data with a simpler and much more efficient data memory indexing operation.

The MCU controls the addresses of subsequent samples and the time interval between samples. The shape change of signal can be made by changing the memory content of the samples. By changing the time interval between the samples, the frequency of the output signal is adjusted. To change the output signal amplitude, also a LUT technique can be used. The advantage of this method is the ability to generate waveforms of various shapes that are stored in the memory - the LUT table.
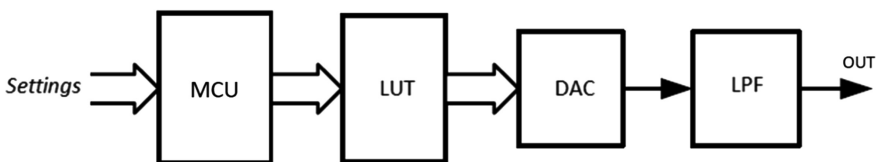
Fig. 1. Simplified block diagram of DDS generator.

In the generator, the diagram of which is presented in Fig. 1, in contrast to many DDS solutions presented in the literature, it is not possible to adjust the phase of the output signal. This simplification of the system is intentional and is related to the target application of the developed generator in the 3-phase version.

The concept of DDS generator presented above can be used to build a three-phase sinusoidal generator. The frequency of its output signals is adjustable in the range of 45–55 Hz, while the amplitudes of these signals vary within the range of 0.8–1.0 of their nominal values. This generator will be implemented as a very simple model of an alternating current generator, a part of the power plant model with several generating sets operating with safe signal levels. The generator has been designed to be used in student laboratories. After introducing simple program changes, the generator can also be used to generate three-phase waves with other shapes, such as rectangular or triangular.

2   Construction and Principle of Operation

The main requirement of the project is to obtain the basic functions of the three-phase sine wave generator, i.e., stable generation of three waveforms mutually offset by 120° and digital control of amplitude and frequency of output signals, in the adopted ranges.

The block diagram of the designed generator is shown in Fig. 2. The waveforms appearing on OUT1, OUT2 and OUT3 outputs are controlled by the MCU equipped with the microprocessor unit (MPU). The parameters of the output signals, i.e. their frequency and amplitude, can be adjusted by the user in the set range using the pushbuttons. The digital representations of the instantaneous values of the output signals are stored in the microcontroller memory. One data set contains data corresponding to the instantaneous values of three signals shifted by 120°. The sequence of rewriting the set of data from the memory to individual BUFF buffers is repeated at the time specified by the MPU timer. The destination buffer for the relevant data is specified by the active one of the lines LE1–LE3 (Latch Enable). The DAC converter in each channel converts digital data into instantaneous values of the generated signal. The quantization noise in the generated signals is reduced in the low-pass filters LPF1–LPF3.
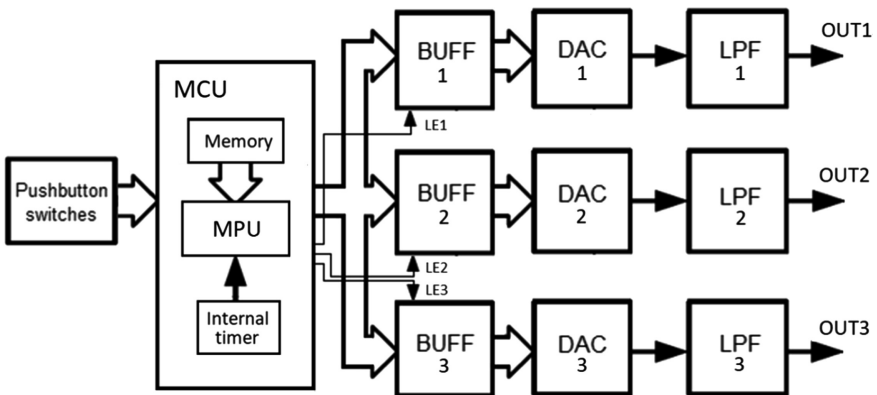

Fig. 2. Simplified block diagram of three-phase DDS generator.

2.1   Construction

The electrical diagram containing the basic generator circuits is shown in Fig. 3. MSP-EXP430G2 LaunchPad Development Kit serves as the MCU. It is an easy-to-use Texas Instruments microcontroller development board for low-power and low-cost MSP430G2x family microprocessor units (MPUs). It has, among others, on-board emulation for programming and debugging and features a 14/20-pin DIP socket.

The MPU onboard MSP-EXP430G2 LaunchPad, used in the DDS generator design - MSP430G2553 from Texas Instruments, is based on 16-bit RISC architecture (Reduced Instruction Set Computing). It has followings basic parameters:

- two 16-bit timers,
- wide clock frequency range up to 16 MHz,
- two 8-bit I/O ports,
- 16 kB Flash memory,
- 512 B RAM memory.

Rich resources of the processor, at a low cost of the starter kit, give the possibility to create a variety of layouts. The selected microprocessor is part of a large family of MSP430 microcontrollers with low-power, high integration and high performance. Free programming environments (e.g. CCS or IAR) enable easy programming in both high-level language and assembler using the USB interface.

From the MPU's internal memory, organized as a LUT memory, using a P2 digital port, 8-bit data is sent to three SN74A373 buffers, activated by the MPU's P1.4–P1.7 lines for data distribution into 3 channels. Digital data is converted into analog signals in DAC0808 converters. Their output signals have negative values, hence the circuits on LM324 operational amplifiers realize the shifting of these signals so that their mean values are equal to zero. At the same time, these amplifiers function as first order LPF filters with a cut-off frequency of around 200 Hz. In the expected use of the generator, for a narrow range of frequency changes in the output signals, as well as due to identical filter parameters in all 3 channels, the effect of phase characteristics of the FDP filter on the relationship between the signals was considered negligible.

Three buttons S1–S3 are used to adjust the frequency and amplitudes of the signals generated. It was assumed that the frequency of the output signals may vary in the range of 45–55 Hz, while the amplitudes in the range of 80–100%, with their nominal values of 5 V. The frequency control resolution was set to 0.1 Hz and the amplitude control was pre-set with the step 0.05 V (5% of adjustment range). Frequency change (S1, S2) consists in changing the number entered in the TA0 timer register, and the amplitude adjustment (manipulation of the S3 pressing time) is done by changing the current LUT table to another LUT containing samples with different values.

Visible in the diagram in Fig. 3, identical markings of lines in individual blocks (MCU, BUF and DAC) correspond to the interconnections of these lines between the blocks.

Figure 4 shows a view of the generator prototype, developed and tested as part of the student project.

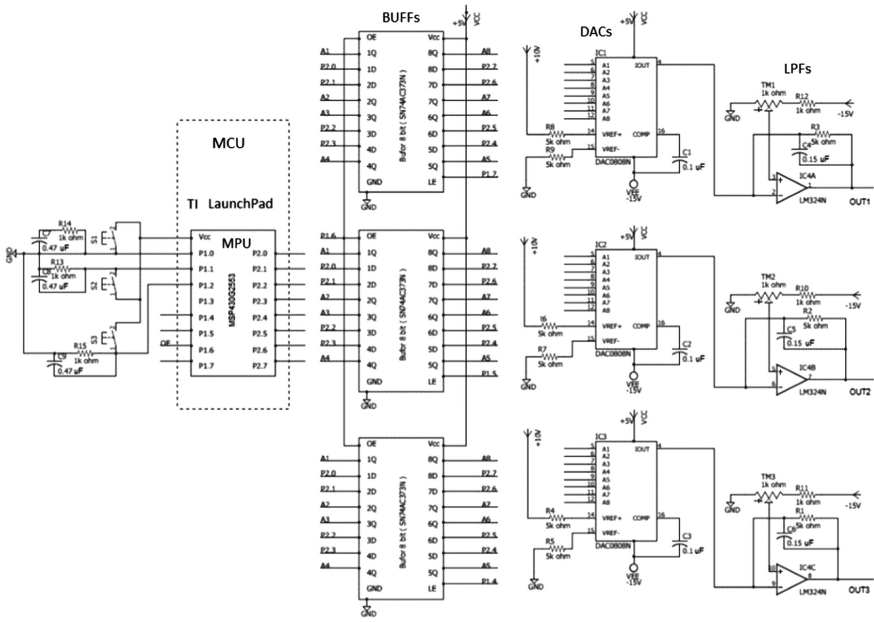**Fig. 3.** The electrical schematic of the three-phase DDS sine wave generator.
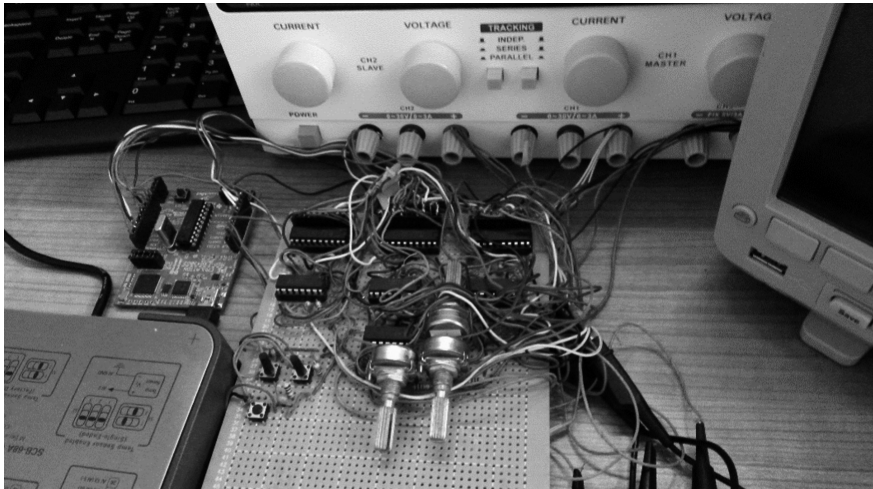

Fig. 4. The view of the generator prototype.

2.2 Algorithms

The algorithms implemented in the program executed by the MPU are shown in Fig. 5. Main loop of the program (Fig. 5a) is an empty loop. After initializing the necessary processor resources, it waits for any interrupt. In Fig. 5b, the subroutine of outputting

data from the LUT memory to BUFF1–BUFF3 buffers (Fig. 2). The interrupt is initiated cyclically by the TA0 timer in the MPU. The TA0 timer, operating in the up-mode, determines the moments in which the digital samples are sent via the output port P2 to BUFF buffers. In this way, the frequency of the DDS generator signals is determined. The interrupt service subroutine, initiated by pushbuttons connected to P1.0 or P1.1 lines (Fig. 5c), allows to change the frequency of signals generated. The activated pushbutton on the P1.2 line causes an another interrupt service subroutine to be executed (Fig. 5d) and change of amplitude of output signals in accordance with established algorithm by switching the LUT tables of the samples being sent.
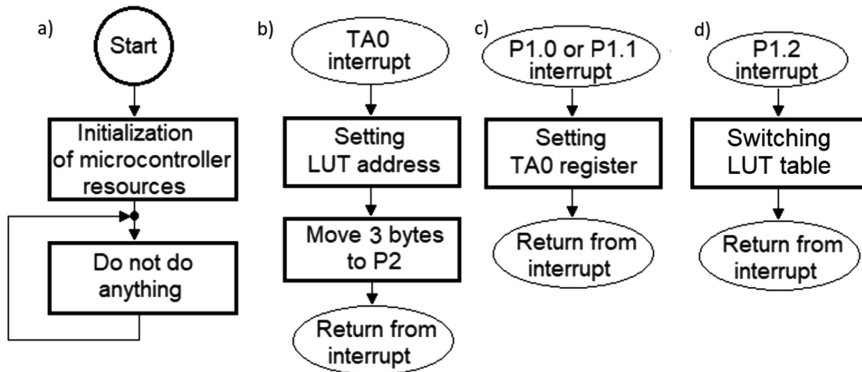

Fig. 5. Main algorithms of the MPU program.

The frequency of the generator is controlled by the TA0 timer. If the generator does not operate at a frequency which high enough, the integer number set in TA0 register has to be changed. The smaller the number, the higher the frequency.

The MPU is equipped with a memory, the capacity of which is sufficient to store the values of the output signal samples. The program executed by the MPU rewrites the values stored in its memory to the BUFF buffers, and further to the DAC converters, in the timing of the sampling signal. This signal comes from an internal MPU timer. By changing the frequency of the sampling signal, a constant number of samples can be kept for the period of the generated signal, and thus also the THD distortion factor.

In this solution, the resolution of frequency settings is also closely related to the number of samples per period of generated waveform. In addition, the high resolution of the frequency puts the requirements on the MPU, which must have a short enough cycle to be able to rewrite the values of samples to the BUFF buffers during one period of the sampling signal.

3   Analysis of Constraints

The TA0 register is reset at the beginning of the interrupt service, so the time interval ΔT_s between successive samples can be determined using the formula (1).

$$\Delta T_s = \frac{N}{f_{clk}} \quad (1)$$

where: N - data entered into the TA0 timer register, f_{clk} – the MPU clock frequency.

Then, the frequency f_s of the signal on the DDS generator output can be expressed by the formula (2).

$$f_s = \frac{1}{k \cdot \Delta T_s} = \frac{f_{clk}}{k \cdot N} \quad (2)$$

where: k - number of samples during the signal period generated.

A single cycle of sending data from the LUT memory to the P2 port for 3 channels, with $f_{clk} = 16$ MHz, requires a time T_c of about 11.1 μs. This value was obtained as measurement result for implemented TA0 interrupt service subroutine.

It is obvious that the condition (3) must meet:

$$\Delta T_s > T_c \quad (3)$$

From Eqs. (1) and (3) it can be concluded that $N > 178$.

For $f_s = 55$ Hz (the most unfavorable case) the maximum value of k cannot exceed 1635.

For the above conditions referring to $f_{clk} = 16$ MHz and for boundary frequencies f_s (45 and 55 Hz), the relations between N and k are shown in Fig. 6. It shows that for a better resolution of frequency control f_s, the number k of samples per period of generated signal should be chosen as less than 200. Then the change N causing the unitary change f_s is the largest.
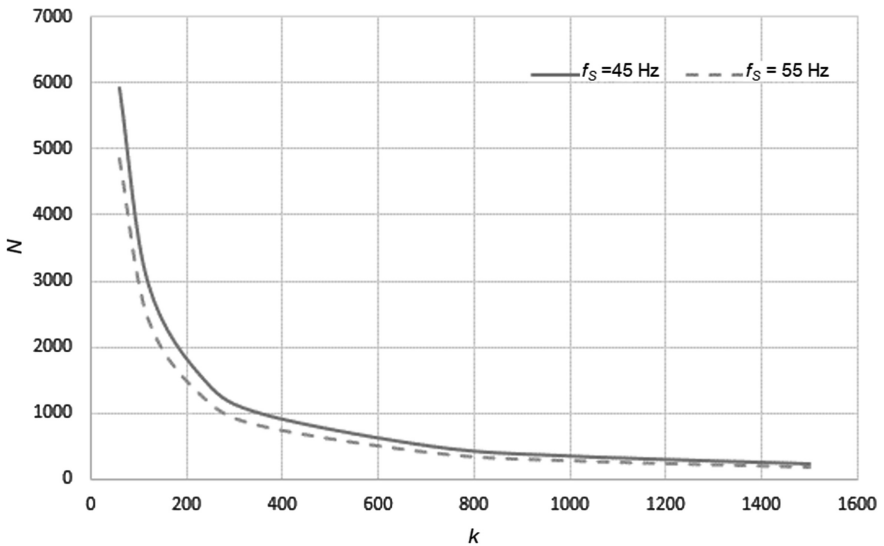

Fig. 6. A graph of dependencies between N and k for boundary frequencies f_s.

The interrupt service time T_c for sending samples to 3 channels depends on f_{clk}. For f_{clk} = 16 MHz, i.e. the maximum frequency of the MPU's clock, used in the generator, T_c is measured as about 11.1 µs.

4   Tests

The functional tests of three-phase generator were carried out using a digital oscilloscope (OSC) MSO2024 Tektronix and a computer set equipped with data acquisition board (DAQ) NI PCIe6321 working in LabVIEW environment. The properties of tested generator – DUT (Device under Test) – were examined in full range of adjustable parameters – frequency and amplitude of three-phase signals.

Generator tests included, among others, determining the duration of a single data transmission cycle T_c for different MPU clock frequencies f_{clk}, as well as the maximum number of samples k_{max} during the signal. The results are summarized in Table 1.

Table 1. List of selected generator parameters for f_s = 55 Hz, at different MPU clock rates f_{clk}.

f_{clk} [MHz]	T_c [µs]	k_{max}
1	209	87
8	22	827
12	15	1216
16	11.1	1635
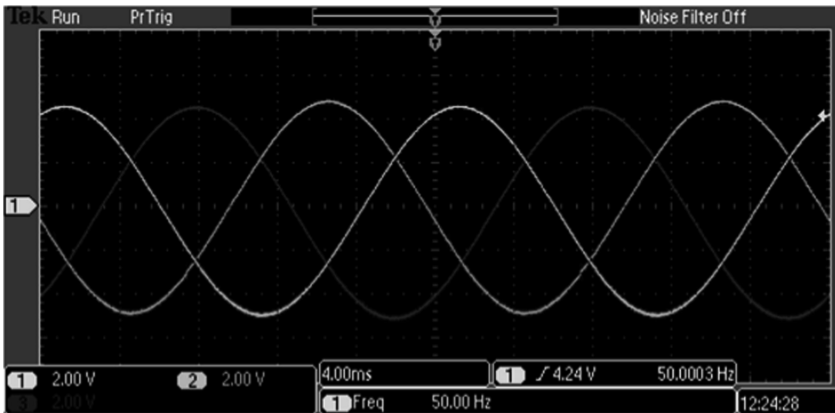
Fig. 7. Exemplary waveforms obtained from the generator.

Exemplary three-phase waveforms from the generator are shown in Fig. 7. They were taken for f_s = 50 Hz, k = 360 and f_{clk} = 16 MHz.

Figure 8 presents the empirically determined dependence, binding N - the number to set to TA0 – with the value of the set frequency f_s of the generated signal. The relations
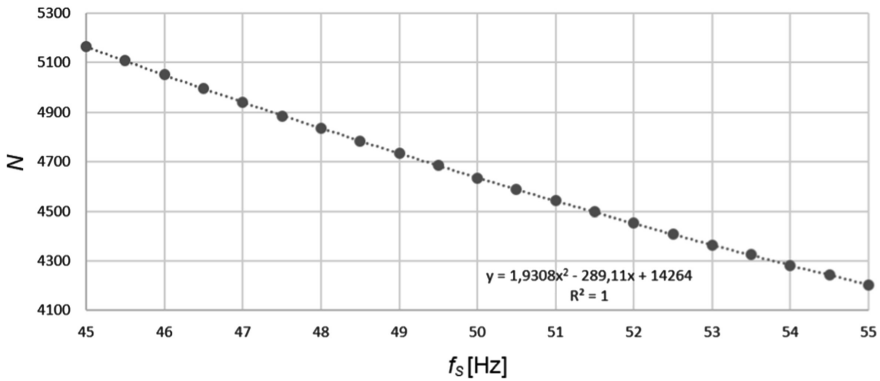
Fig. 8. Exemplary relations between N and f_s; $k = 60$, $f_{clk} = 16$ MHz.

were measured with a step of 0.5 Hz and, applying interpolation, they can be used to denote N to achieve the required resolution of control f_s.

The harmonic analysis of the exemplary output signal from the generator is shown in Fig. 9. Quite low level of the THD factor and the presence of harmonics related to the number of signal samples pay attention. In the example shown, THD is 0.44% and the harmonics level around 60th is visibly increased compared to other harmonics. The level of the harmonics can be reduced by applying the DAC with higher resolution.
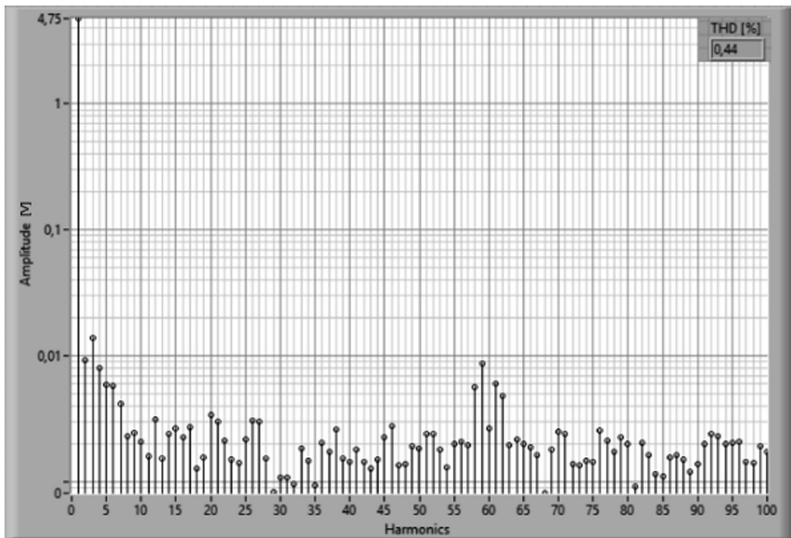
Fig. 9. Example of harmonic content analysis for a selected phase signal at $f_s = 55$ Hz of a three-phase DDS generator; $k = 60$, $f_{clk} = 16$ MHz.

5   Conclusion

A key requirement in many applications is the ability to generate and control waveforms at various frequencies. DDS based sine generators offer substantial performance improvements, at reduced costs, over conventional analog function generators.

The Direct Digital Synthesis is a technique of generating an analog waveform, including the shape of a sine wave, using a time-varying signal in a digital form and a DAC converter. DDS technique is gaining wide popularity and acceptance from the industrial community to achieve programmable analog outputs with accuracy and high resolution.

The article presents a three-phase generator designed to meet specific requirements. The low cost waveform generator has been developed with several advantages such as simple structure and easy operation. It can produce three-phase sine wave signals to satisfy the requirements of student laboratories. The resources of the selected microprocessor are fully sufficient for the designed generator. In order to improve the generator parameters, mainly to increase the resolution of amplitude control, higher resolution DACs (for example 10-bit) should be used. It will also reduce the share of higher harmonics in the output signals. This will be introduced in the final version of the generator.

The developed generator will be used to model phenomena in a ship power plant using low-voltage signals.

References

1. Platiše, U.: Analog three-phase sine wave generator, Document No. 001-35339, Rev. C (2015). www.cypress.com
2. Callegaro, L., D'Elia, V.: A synchronized two-phase sinewave generator for AC metrology system compensations. IEEE Trans. Instrum. Meas. **49**(2), 320–324 (2000)
3. Bilgin, S., Üser, Y., Oktay, M.: Low cost laboratory type signal generator using DDS method. Int. J. Eng. Appl. Sci. (IJEAS) **8**(4), 59–65 (2016)
4. Maśnicki, R., Jurek, K.: Constructions of power generators developed for laboratory applications. Prz. Elektrotech. **91**, 30–33 (2015). (in Polish)
5. Bell, B.A.: Standards for waveform metrology based on digital techniques. J. Res. Nat. Inst. Stand. Technol. **95**(4), 377–405 (1990)

Operational Estimating of Arcs Voltage of Arc Steel Furnace

Yaroslav Marushchak[1,2]([✉]), Mykhaylo Dorozhovets[1,2]([✉]),
and Damian Mazur[1]

[1] Rzeszow University of Technology, 35-959 Rzeszow, Poland
{jamaru,michdor}@prz.edu.pl
[2] National University - Lviv Politechnic, Lviv, Ukraine

Abstract. In this paper the possibility of improving accuracy of maintaining the regime of melting metal by appropriate control of the arc voltages in electric arc furnaces is presented. For this purpose the methodology of the determination of the own and mutual inductances of the three flexible cables, are proposed and analyzed. This methodology based on realization experiments of the three-phase short network, as well as three biphasic short network when in the third phase arc is absent. The proposed methodology was implemented to determine the parameters of current propulsion of electric arc furnace type DSP-100. The determined cabel parameters (resistances and inductances) are used to form the control effect of the automatic control system by moving the electrodes and hence the power of the arcs of each phase.

Keywords: Arc furnace · Determination · Inductances · Voltages
Experiment · Control

1 Introduction

Almost all electric arc furnaces, in which the melting of the charge is carried out for the purpose of obtaining electro technical steels (Fig. 1), are functioning due to the automatic control system (ACS) which provide appropriate displacements of the electrodes.

These control systems provide an automatic ignition of the arc in each of the three phases of supply the arcs, as well as maintenance of arc capacities in accordance with the directive schedule of the melting process. This is realizing due to the using of the impedance principle of regulation [1–4], which consists in the fact that the control voltage of the electrodes drive is formed as the voltage of the discrepancy:

$$\Delta_j = a_j \cdot I_{dj} - b_j \cdot U_{dj} \tag{1}$$

where I_{dj}, U_{dj} are current values of currents and voltages of arcs in the j phase; a_j, b_j are coefficients (set point) that define the condition $\Delta_j = 0$, when a certain arc length is set and the appropriate power in the furnace's space is entered.

The measuring of currents I_{dj} is not problematic, but voltages U_{dj} cannot be measured directly. This is due to the fact, that the arc is burning in a closed furnace
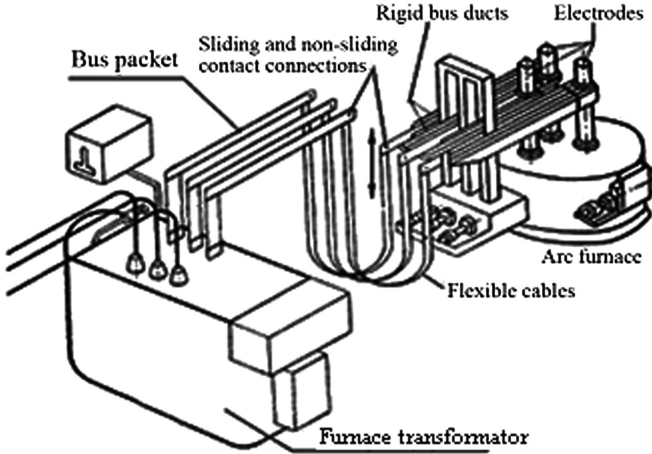
Fig. 1. Components of the arc steel furnace.

space and connecting the measuring apparatus to the end of the electrode with the arc is not possible. Therefore, the phase voltages are measured between the phase output of the furnace transformer and the zero point formed by the three arcs on the charge.

Voltage sequences are determined according to known parameters of flexible cables (short network) and electrodes of each phase, and then the U_{dj} voltage is formed which ensures the functioning of the ACS by the displacement of the electrodes. Thus, the accuracy of determining the parameters of flexible cables with the electrodes of each phase influences on the power supply into the furnace space, and on the efficiency of an arc steel furnace (ASF) exploitation. Typically, the values of these parameters are given in the reference literature [5, 6] and they are reduced to the resistances and reactances of the short network r_{kj}, x_{kj}, and the electrode r_e. It is assumed that these parameters are the same in each phase.

The electrodes in different phases worn out in different ways during the technological process of melting steel in the ASF, and therefore the electrode holders occupy different levels. Consequently, the configuration of the flexible cables of each phase will be different. In addition, the effect of mutual induction and the phenomenon of the transfer of active power between the phases, that accompanying it, are strongly manifested between the phases of the short network. Thus, a short network should be taken into account by using of its inductances L_{k11}, L_{k22}, L_{k33}, mutual inductances $L_{k12} = L_{k21}$, $L_{k23} = L_{k32}$, $L_{k31} = L_{k13}$ and the appropriate resistances r_{k1}, r_{k2}, r_{k3}. Then the inductance of each phase of the short network could be represented by an equivalent value:

$$\left.\begin{aligned} L_{eq1} &= L_{k11} - L_{k12} - L_{k13} + L_{k23} \\ L_{eq2} &= -L_{k12} + L_{k13} + L_{k22} - L_{k23} \\ L_{eq3} &= -L_{k23} - L_{k13} + L_{k12} + L_{k33} \end{aligned}\right\} \qquad (2)$$

Equation (2) are obtained by solving magnetic connections between the three phases of the cables. Determination of the cable parameters is possible according to the results of measurements during a series of experimental researches, which should be performed after each electrode replacement at least in one phase, or more often. Own and mutual inductances are most easily determined through reactancesre. They make sense only with sinusoidal currents and voltages. Unfortunately, conducting of special experiments not always could be realized due to material losses caused by lowering the efficiency of the ASF. These losses are increasing with capacity of the furnace. Therefore, the problem of measuring the voltage of arcs by data of passive experiment without interference with the technological process is actual for heavy electric arc furnaces.

2 Proposed Methodology to Determination of Arc Voltages

2.1 Determining of Arc Voltages Using Measurement Results Obtained from Active Experiment

The condition of sinusoidal currents and voltages in a short network is not fulfilled during arcs burning. Realization of different combinations of short networks and idle stroke (absent arc) in phases leads to possibility of elimination the distortion of currents and voltages from the lower side of the furnace transformer. It is possible to realize the following experiments on the current three-phase arc furnace, in which the currents and voltages will be sinusoidal, in particular, the three-phase short network, as well as three biphasic short network when there is no arc in the third phase. Under these conditions, the voltage between the phase of the furnace transformer at the zero point accessed by the furnace design can be written in a complex form. So for a three-phase short network might be written:

$$
\left.
\begin{aligned}
\dot{U}_1 &= \dot{I}_1 \cdot (r_{k1} + j\omega L_{k11}) + \dot{I}_2 \cdot j\omega L_{k21} + \dot{I}_3 \cdot j\omega L_{k31} \ , \\
\dot{U}_2 &= \dot{I}_2 \cdot (r_{k2} + j\omega L_{k22}) + \dot{I}_1 \cdot j\omega L_{k12} + \dot{I}_3 \cdot j\omega L_{k32} \ , \\
\dot{U}_3 &= \dot{I}_3 \cdot (r_{k3} + j\omega L_{k33}) + \dot{I}_1 \cdot j\omega L_{k13} + \dot{I}_2 \cdot j\omega L_{k23} \ .
\end{aligned}
\right\}
\tag{3}
$$

For short network in phases 2 and 3 during the idle stroke in phase 1:

$$
\left.
\begin{aligned}
\dot{U}_2' &= \dot{I}_2' \cdot (r_{k2} + j\omega L_{k22} + j\omega L_{k23}) \ , \\
\dot{U}_3' &= \dot{I}_2' \cdot (r_{k3} + j\omega L_{k33} + j\omega L_{k32}) \ .
\end{aligned}
\right\}
\tag{4}
$$

For short network in phases 3 and 1 during the idle stroke in phase 2:

$$
\left.
\begin{aligned}
\dot{U}_1'' &= \dot{I}_1'' \cdot (r_{k1} + j\omega L_{k11} + j\omega L_{k13}) \ , \\
\dot{U}_3'' &= \dot{I}_3'' \cdot (r_{k3} + j\omega L_{k33} + j\omega L_{k31}) \ .
\end{aligned}
\right\}
\tag{5}
$$

For short network in phases 1 and 2 during the idle stroke in phase 3:

$$\left.\begin{array}{l} \dot{U}_1''' = \dot{I}_1''' \cdot (r_{k1} + j\omega L_{k11} + j\omega L_{k12}) , \\ \dot{U}_2''' = \dot{I}_2''' \cdot (r_{k2} + j\omega L_{k22} + j\omega L_{k21}). \end{array}\right\} \tag{6}$$

During such experiments, it is necessary to ensure the same position of the electrode holders of the corresponding phase, as in the case of a three-phase short network experiment, and of two-phase short network experiments. Under these conditions, the corresponding own and mutual inductances will have the same value for different experiments in Eqs. (3)–(6). Corresponding resistances will be the same. It should be noted that r_{kj} also includes resistances of electrodes. Resistances are determined by the wattmeter and ammeter measurements in each phase.

The values of six unknown inductions L_{k11}, L_{k22}, L_{k33}, L_{k12}, L_{k23}, L_{k13} could be calculated as the solutions of the Eqs. (3)–(5). We can write these equations in projections on the imaginary axis of the complex plane:

$$\left.\begin{array}{l} U_1\sin\alpha_1 = I_1 r_{k1}\sin\beta_1 - I_1\omega L_{k11}\cos\beta_1 - I_2\omega L_{k12}\cos\beta_2 - I_3\omega L_{k31}\cos\beta_3, \\ U_2\sin\alpha_2 = I_2 r_{k2}\sin\beta_2 - I_1\omega L_{k12}\cos\beta_1 - I_2\omega L_{k22}\cos\beta_2 - I_3\omega L_{k23}\cos\beta_3, \\ U_3\sin\alpha_3 = I_3 r_{k3}\sin\beta_3 - I_1\omega L_{k13}\cos\beta_1 - I_2\omega L_{k2}\cos\beta_2 - I_3\omega L_{k33}\cos\beta_3, \\ U_2'\sin\alpha_2' = I_2'(\omega L_{k22} + \omega L_{k23}), \\ U_3'\sin\alpha_3' = I_2'(\omega L_{k33} + \omega L_{k23}), \\ U_1''\sin\alpha_1'' = I_1''(\omega L_{k11} + \omega L_{k13}) . \end{array}\right\} \tag{7}$$

where $\alpha_1 \div \alpha_3$, $\alpha_2' \div \alpha_3'$; α_1'' – arguments (angles) of voltage vectors $\dot{U}_1 \div \dot{U}_3$, $\dot{U}_2' \div \dot{U}_3'$, $\dot{U}_1''$ in accordance; $\beta_1 \div \beta_3$ – arguments (angles) of current vectors $\dot{I}_1 \div \dot{I}_3$.

Having solved the Eq. (7), we obtain:

$$L_{k13} = \frac{I_2\cos\alpha_2(I_2 r_{k2}\sin\alpha_2 - U_2\sin\beta_2 - U_2'I_2\cos\alpha_2\sin\beta_2/I_2' +}{\omega I_1\cos\alpha_1(I_2\cos\alpha_2 - I_3\cos\alpha_3 + I_1\cos\alpha_1)} \rightarrow$$
$$\rightarrow \frac{+ I_3 r_{k3}\sin\alpha_3 - U_3\sin\beta_3 - U_3'I_3\cos\alpha_3\sin\beta_3/I_2' -}{} \rightarrow \tag{8}$$
$$\rightarrow \frac{-I_1\cos\alpha_1\left(I_1 r_{k1}\sin\alpha_1 - U_1\sin\beta_1 - U_1''I_1\cos\alpha_1\sin\beta_3/I_3\right)}{} ,$$

$$L_{k12} = \frac{I_1 r_{k1}\sin\alpha_1 - U_1\sin\beta_1 - U_1''I_1\cos\alpha_1\sin\beta_1/I_1'' - (I_3\cos\alpha_3 - I_1\cos\alpha_1)\omega L_{k13}}{\omega I_2\cos\alpha_2} \tag{9}$$

$$L_{k23} = \frac{I_3 r_{k3}\sin\alpha_3 - U_3\sin\beta_3 - U_3'I_3\cos\alpha_3\sin\beta_3'/I_2' - I_1\cos\alpha_1\omega L_{k13}}{\omega(I_2\cos\alpha_2 - I_3\cos\alpha_3)} \tag{10}$$

$$L_{k11} = U_1''\sin\alpha_1''/(\omega I_1'') - L_{k13}, \tag{11}$$

$$L_{k22} = U_2' \sin \alpha_2' / (\omega I_2') - L_{k23} \tag{12}$$

$$L_{k33} = U_3' \sin \alpha_3' / (\omega I_2') - L_{k23} \tag{13}$$

Using condition $\beta_1 = 0$ (vector of current $\dot{I}_1$ is placed on the axis of real numbers) form a triangle of currents, which are flowing at a three-phase short network (Fig. 2), the values of angles β_2 and β_3 are equal:
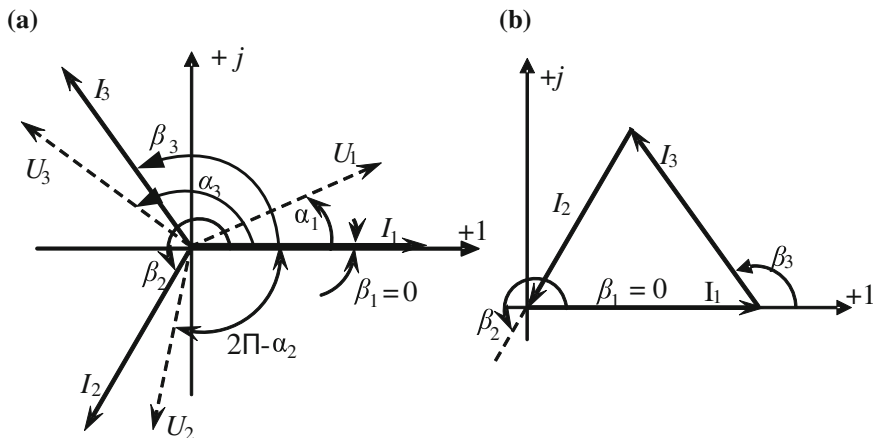
Fig. 2. Vector diagram of the currents of three-phase short network.

$$\beta_2 = \pi + arccos\left[(I_1^2 + I_2^2 - I_3^2)/(2I_1I_2)\right], \beta_3 = \pi - arccos\left[(I_1^2 + I_3^2 - I_2^2)/(2I_1I_3)\right] \tag{14}$$

The values of angles α_i for vector diagram presented on Fig. 2 could be defined as:

$$\left.\begin{aligned} \alpha_1 &= arccos[P_1/(U_1I_1)] + \beta_1, \\ \alpha_2 &= arccos[P_2/(U_2I_2)] + \beta_2, \\ \alpha_3 &= arccos[P_3/(U_3I_3)] + \beta_3 , \end{aligned}\right\} \tag{15}$$

$$\left.\begin{aligned} \alpha_2' &= arccos\left[P_2'/(U_2'I_2')\right], \\ \alpha_3' &= arccos\left[P_3'/(U_3'I_3')\right], \\ \alpha_1'' &= arccos\left[P_1''/(U_1''I_1'')\right]. \end{aligned}\right\} \tag{16}$$

where P_j are measurements of the wattmeter in phases during corresponding experiments, U_j, I_j are measurements of corresponding voltmeters and ammeters.

2.2   Implementation of Proposed Method Obtained from Active Experiment

The proposed methodology was implemented to determine the parameters of short network ASF −1.5. According to the reference data [6] for this furnace:

$$r_k^d = 1.12 \cdot 10^{-3} \text{ Ohm}, \ x_k^d = 4.15 \cdot 10^{-3} \text{ Ohm}.$$

The values of the parameters were averaged (the average arithmetic value for the three phases) for a comparative analysis of these data with the results obtained on the basis of experimental researches.

The value of $r_k^e = 1.08 \cdot 10^{-3}$ Ohm for resistance to the results of experimental research. In this case, the relative deviation from the reference data is small and equals <4%.

The following values for the inductances of the cables ASF −1.5 are obtained:

$$L_{k11} = 14.2 \cdot 10^{-6} \text{ H}, \ L_{k22} = 15.2 \cdot 10^{-6} \text{ H}, \ L_{k33} = 11.92 \cdot 10^{-6} \text{ H},$$
$$L_{k12} = 4.90 \cdot 10^{-6} \text{ H}, \ L_{k23} = 4.80 \cdot 10^{6} \text{ H}, \ L_{k13} = 3.45 \cdot 10^{-6} \text{ H}.$$

According to these values the equivalent inductances (2) are:

$$L_{eq1} = 10.65 \cdot 10^{-6} \text{ H}, \ L_{eq2} = 8.95 \cdot 10^{-6} \text{ H}, \ L_{eq3} = 8.57 \cdot 10^{-6} \text{ H}.$$

Then average arithmetic value is $L_{av}^e = 9.39 \cdot 10^{-6}$ H, that corresponds to reactance $x_{av}^e = 2.95 \cdot 10^{-3}$ Ohm. In this case, the deviation is almost 30%.

If not take into account mutual inductances, then

$$L_{av}^e = (L_{k11} + L_{k22} + L_{k33})/3 = 13.77 \cdot 10^{-6} \text{ H}.$$

In this case, the deviation is almost 4%. The references give the values of their own inductances of the cables, which entails an increasing of the error in determination of the arc voltage.

2.3   Determining of Arc Voltages Using Measurement Results Obtained from Passive Experiment

Measurement of the cable parameters according to the considered algorithm requires specially planned experimental research (active experiment). Due to the peculiarities of the technological process of steel smelting in the ASF, it always has breaks due to the loading of the furnace, the introduction of special components for obtaining a certain mark of steel, removal of slag from the surface of the molten charge, etc. During such breaks it is possible to organize a series of experiments to clarify the parameters of the cable. It's clear that an active experiment takes some time, and therefore it can affect the performance of the ASF. If the measure of operation efficiency of an electric arc furnace has a higher priority, it is proposed to determine the parameters of the current

transmission to be carried out according to the measurement during a passive experiment without interference in the technological process.

During the passive experiment there is a continuous control of the instantaneous value of the phase voltage from the low side of the furnace transformer (u_1, u_2, u_3) and the derivatives of the instantaneous values of arc current ($di_{1\partial}/dt$, $di_{2\partial}/dt$, $di_{3\partial}/dt$) [7] (Fig. 3).
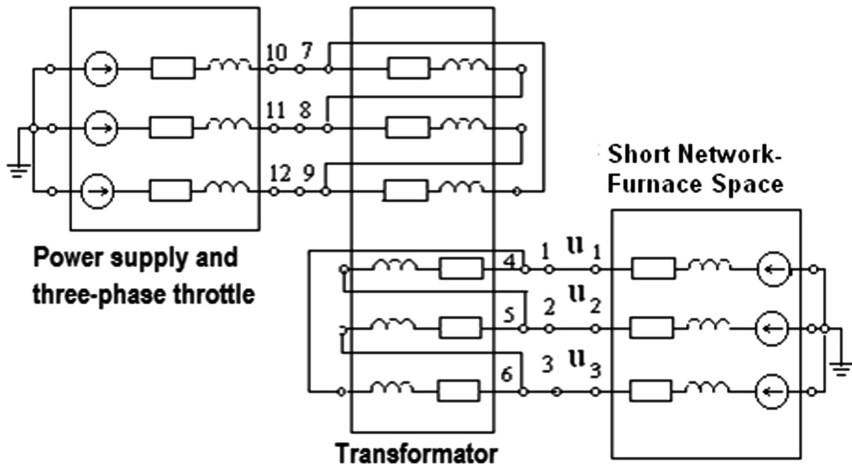

Fig. 3. The scheme of passive experiment in arc furnace.

The instantaneous values of these voltages can be written in the form:

$$
\left.
\begin{aligned}
u_1 &= i_{1\partial} \cdot r_1 + L_{k11} \cdot \frac{di_{1\partial}}{dt} + L_{k21} \cdot \frac{di_{2\partial}}{dt} + L_{k31} \cdot \frac{di_{3\partial}}{dt} + u_{\partial 1}, \\
u_2 &= i_{2\partial} \cdot r_2 + L_{k22} \cdot \frac{di_{2\partial}}{dt} + L_{k32} \cdot \frac{di_{3\partial}}{dt} + L_{k12} \cdot \frac{di_{1\partial}}{dt} + u_{\partial 2}, \\
u_3 &= i_{3\partial} \cdot r_3 + L_{k33} \cdot \frac{di_{3\partial}}{dt} + L_{k23} \cdot \frac{di_{2\partial}}{dt} + L_{k13} \cdot \frac{di_{1\partial}}{dt} + u_{\partial 3},
\end{aligned}
\right\}
\tag{17}
$$

where $u_{\partial i}$ are the instantaneous arc voltages in appropriate phases.

It should be noted that the electric arc of the alternating current has an resistance character, as indicated by its dynamic volt-ampere characteristics [4, 8, 9] for different lengths in relation to the electric arc furnace type DSP-100 (see Fig. 4).

At time $t_j^{(1)}$ and $t_j^{(2)}$, when the current of j - th phase passes through zero for a positive (at $t_j^{(2)}$) and negative (at $t_j^{(1)}$) the values of its derivative, on the basis of (17) taking into account that the voltage drops on the resistances and the voltage of the arcs in this case are equal to zero, the equations for instantaneous arc voltages in appropriate phases at these times have the following form
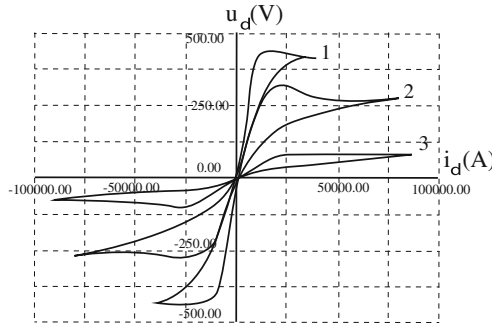
Fig. 4. Dynamic volt-ampere characteristics for different lengths in relation to the electric arc furnace type DSP-100.

$$
\left.
\begin{aligned}
u_1\left(t_1^{(1)}\right) &= L_{k11} \cdot \frac{di_{1\partial}}{dt}\bigg|_{t=t_1^{(1)}} + L_{k21} \cdot \frac{di_{2\partial}}{dt}\bigg|_{t=t_1^{(1)}} + L_{k31} \cdot \frac{di_{3\partial}}{dt}\bigg|_{t=t_1^{(1)}}, \\
u_1\left(t_1^{(2)}\right) &= L_{k11} \cdot \frac{di_{1\partial}}{dt}\bigg|_{t=t_1^{(2)}} + L_{k21} \cdot \frac{di_{2\partial}}{dt}\bigg|_{t=t_1^{(2)}} + L_{k31} \cdot \frac{di_{3\partial}}{dt}\bigg|_{t=t_1^{(2)}}, \\
u_2\left(t_2^{(1)}\right) &= L_{k22} \cdot \frac{di_{2\partial}}{dt}\bigg|_{t=t_2^{(1)}} + L_{k12} \cdot \frac{di_{1\partial}}{dt}\bigg|_{t=t_2^{(1)}} + L_{k32} \cdot \frac{di_{3\partial}}{dt}\bigg|_{t=t_2^{(1)}}, \\
u_2\left(t_2^{(2)}\right) &= L_{k22} \cdot \frac{di_{2\partial}}{dt}\bigg|_{t=t_2^{(2)}} + L_{k12} \cdot \frac{di_{1\partial}}{dt}\bigg|_{t=t_2^{(2)}} + L_{k32} \cdot \frac{di_{3\partial}}{dt}\bigg|_{t=t_2^{(2)}}, \\
u_3\left(t_3^{(1)}\right) &= L_{k33} \cdot \frac{di_{3\partial}}{dt}\bigg|_{t=t_3^{(1)}} + L_{k23} \cdot \frac{di_{2\partial}}{dt}\bigg|_{t=t_3^{(1)}} + L_{k13} \cdot \frac{di_{1\partial}}{dt}\bigg|_{t=t_3^{(1)}}, \\
u_3\left(t_3^{(2)}\right) &= L_{k33} \cdot \frac{di_{3\partial}}{dt}\bigg|_{t=t_3^{(2)}} + L_{k23} \cdot \frac{di_{2\partial}}{dt}\bigg|_{t=t_3^{(2)}} + L_{k13} \cdot \frac{di_{1\partial}}{dt}\bigg|_{t=t_3^{(2)}}.
\end{aligned}
\right\} \tag{18}
$$

It is well known [10–12] that the duration of displacement of the elements of the short network due to the functioning of the system of automatic control of the motion of the electrodes, which causes changes in the values of inductances, is much greater than the period of the supply voltage. This allows for several voltage periods to assume parameters as constant. Having solved the system of Eq. (18), we obtain the values of own and mutual inductances for the cables, which are used to determine the voltages of the arcs u_d of the Eq. (18). Parameters r_1, r_2, r_3 are practically constant, given in the passport of the cables, or can easily be determined experimentally by the data of the wattmeter and the corresponding ammeters during the technological short networks required to ignite the arcs at the beginning of the technological process.

Thus, the time required to obtain the information necessary for correction of the values of inductances is commensurate with the period of instantaneous supply voltages, and therefore, until the specified parameters of the cables have not changed, the

instantaneous values of the voltage of the arcs of the three phases of the $u_{\vartheta i}$ are calculated using these parameters on the basis of the Eq. (18). The instantaneous values of the $u_{\vartheta i}$ over one or several periods make it possible to determine their effective values, which are used by the ACS to move the electrodes to realize more accurate maintenance of a given length, and hence the power, arcs of each phase.

3 Results

For the electric arc furnace type DSP 100 N3A the values of their own and mutual inductances of the short network were calculated using the proposed methodology. Numerical values of phase voltages and derivative currents were determined as a result of processing of oscillograms, which were obtained experimentally on an existing ASF
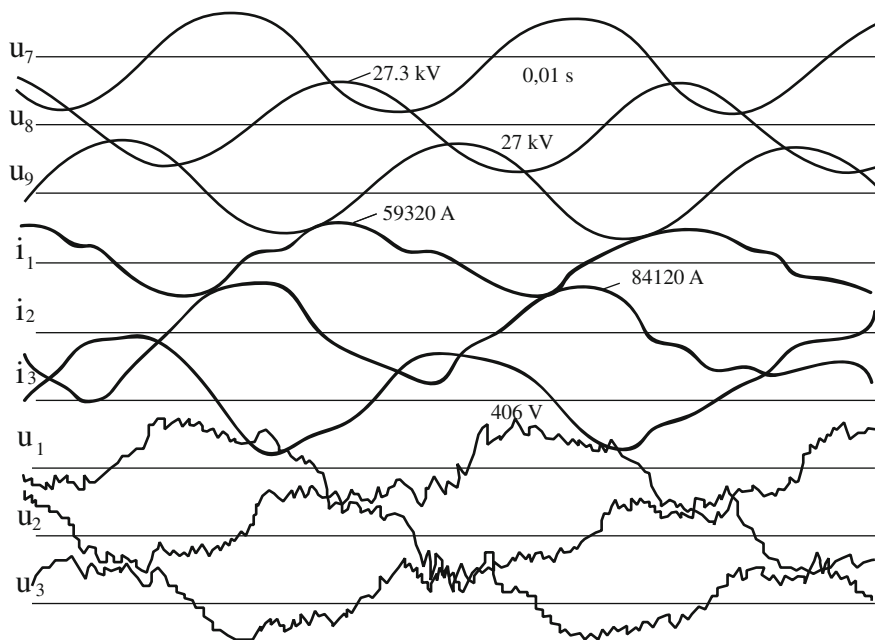
Fig. 5. Experimental oscilloscope diagrams for ASF.

(see Fig. 5).

As a result, the following values of the inductances were obtained for the position of the electrodes, in which the oscillogram was recorded:

$L_{k11} = 19.6 \cdot 10^{-6}$ H, $L_{k22} = 2.6 \cdot 10^{-6}$ H, $L_{k33} = 21.5 \cdot 10^{-6}$ H,

$L_{k12} = L_{k21} = -1.5 \cdot 10^{-6}$ H, $L_{k23} = L_{k32} = -3.5 \cdot 10^{6}$ H, $L_{k13} = L_{k31} = 6.7 \cdot 10^{-6}$ H.

These values correspond to the following values of equivalent inductances calculated according to the formulas (2):

$$L_{eq1} = 10.9 \cdot 10^{-6} \text{ H}, L_{eq2} = 7.3 \cdot 10^{-6} \text{ H}, L_{eq3} = 16.8 \cdot 10^{-6} \text{ H},$$

They have such reactances for sinusoidal current:

$$x_{eq1} = 3.42 \cdot 10^{-3} \text{ Ohm}, x_{eq2} = 2.29 \cdot 10^{-3} \text{ Ohm}, x_{eq3} = 5.28 \cdot 10^{-3} \text{ Ohm}.$$

Then the arithmetic mean is $x_{av} = 3.66 \cdot 10^{-3}$ Ohm. If you calculate x^*_{aver} without taking into account the mutual inductances then $x^*_{av} = 4.58 \cdot 10^{-3}$ Ohm.

Compare the results with passport values ($x = 3.66 \cdot 10^{-3}$ Ohm) it makes no sense because, as a rule, these values are calculated from data from experimental studies for short-circuit experiments when introduced ASF into operation. Obviously, this corresponds to some initial position of a short network and does not take into account the change in the configuration of the ASF coupling phase.

Regarding the current values of arc currents, which are also used by such ACSs it should be noted that for chipboard capacity up to 100 t the currents of arcs are measured from the low side using current transformers, and for furnaces with a capacity of more than 100 t the current measurement is carried out, or from the high side of the furnace transformer for using a current transformer with further recalculation of the measurement results, or from the low side using the Rogovsky belt with filters.

4   Conclusion

Taking into account the parameters of cables in full volume obtained on the basis of a series of active research experiments for ASF small capacity allows you to get complete information about the value of the parameters of a short network. Due to this, the determination of the voltage of the arc, which takes part in the formation of control influence ACS the displacement of electrodes is carried out more precisely than on the basis of the passport data of the parameters of the short network of the electric arc furnace. This increases the accuracy of maintaining the arc length and provides more effective introduction of melting.

For ASF of large capacity it is expedient to use the data of the passive experiment, as a result, the parameters of the phase of the current drive are determined operatively, taking into account both their own and the mutual inductances. Due to this, it is possible to more accurately measure the instantaneous arc voltage, which calculates their updated values. These operating values serve to form the control effect of the ACS by moving the electrodes, and hence the power of the arcs of each phase.

Thus, control is adequate to the real state of arcs in the furnace space and the implementation of the technological process of melting steel in accordance with the policy values of the procedure for the melting of this steel grade is provided.

References

1. Bowman, B., Kruger, K.: Ark furnace physics. Stahleisen Verlag, Dusseldorf (2009)
2. Lozynski, O., Marushchak, Y.: Sytuacyjne keruvannia v dugovyh staleplavylnyh pechah. NU „Lvivska Politehnika", Lviv (2004). (in Ukrainian)
3. Sokolov, M.M., Grasevich, V.N.: Elektrooborudovaniye mehanizmov elektrotermicheskih ustanovok.- M. Eniergoatomizdat (1983). (in Russian)
4. Balan, R., Maties, V., Hancu, O., Stan, S., Ciprian, L.: Modeling and control of an electric arc furnace. In: Control & Automation, 2007, MED 2007, Mediterranean Conference, pp. 1–6 (2007)
5. Mineyev, A.R., Korobov, A.I., Pogrebinskiy, M.Y.: Modelirovaniye lektrotehnologicheskih processov i ustanovok [Modelling of electrotechnical processes and installations]. Sputnik, Moscow (2004). (in Russian)
6. Dancis, Y.B., Kacievich, L.S., Shilov, G.M.: Korotkie sieti i eliektrichieskiye paramietry dugovyh eliektropiechiej. Spravochnik. M. Mirtallutgia (1984). (in Russian)
7. Denys, B.D., Marushchak, Y.Y.: Equipment for determination of the electrical characteristics of the current-conducting wire and arcing voltage of the three-phase electric arc furnace. A.C. No. 1086557 SSSR, МКИ Н 05В7/148 (1984)
8. Zhao, H.: NN-Based approximate model control for the EAF electrode regulator system. Math. Prob. Eng. **2013**, 874890 (2013)
9. Gomes, A.A., Durango, J.J.M., Mejia, A.E.: Electric arc furnace modeling for power quality analysis. In: ANDESCON 2010, pp. 1–6. IEEE (2010)
10. Lozynskyy, O.Y., Marushchak, Y.Y.: Some aspects of control actions formation in electromechanical systems. Electrotekhnika J. **1999**(5), 52–56 (1999)
11. Zhao, H.: Approximate model control for the EAF electrode regulator system. Math. Prob. Eng. **2013**, 874890 (2013)
12. Liand, L., Mao, Z.: A direct adaptive controller for EAF electrode regulator system using neural networks. Neurocomputing **82**, 91–98 (2012)

Modeling and Analysis of the AFPM Generator in a Small Wind Farm System

Damian Mazur[(✉)], Lesław Gołębiowski, Andrzej Smoleń,
Marek Gołębiowski, and Zygmunt Szczerba

Rzeszow University of Technology, ul. Wincentego Pola 2,
35-959 Rzeszow, Poland
`mazur@prz.edu.pl`

Abstract. This paper reports measurements and simulations performed on the generator - matrix converter system designed for small vertical axis wind turbine. The presented coreless axial flux permanent magnet generator has 3 kW of rated power at 300 rev/min. The measurement system constructed for examination of the generator under investigation as well as the measured characteristics were presented. The system of grid connected matrix converter designed for the presented generator was simulated.

Keywords: AFPM generator · Wind farm system · Wind turbine

1 Introduction

The paper presents the design and test results of a generator dedicated especially for wind power plants. Unfortunately, wind turbines generate relatively slow speed range on the shaft. For small units, it can be several hundred revolutions per minute, for large dozen revolutions per minute. In order to match the turbine speed to the rated parameters of a classic generator, multiplicators must be used. Mechanical gears, however, worsen the overall efficiency of the power plant, increase noise and cost of structures reducing its reliability. Therefore, the search goes towards the development of a generator with a significantly low rotational speed. The magnitude of the voltage generated in fact does not depend on the angular velocity, but on the peripheral speed at the location of the winding and the magnet. In order to reduce the speed, the diameter of the machine needs to be increased to ensure sufficient linear speed. The development of neodymium magnet technology has caused a revolution in the field of electrical machines. Currently created machines based on the classic magnet-winding -movement solution are more simple and smaller designs with similar power [1]. The design of the generator shown in the paper belongs to a dedicated system for applications in wind energy, it is a flat system called the "pancake", multi-pole, three-phase with a capacity of 3 KW and rated speed of 300 1/ min.

2   Construction of the Generator

Figure 1 shows the most important elements in the generator arrangement. The stator disc in which the flat stator windings are flooded in the perimeter part is marked with the letter "s". The external rotor discs must at the same time ensure load transfer from the turbine profiles.
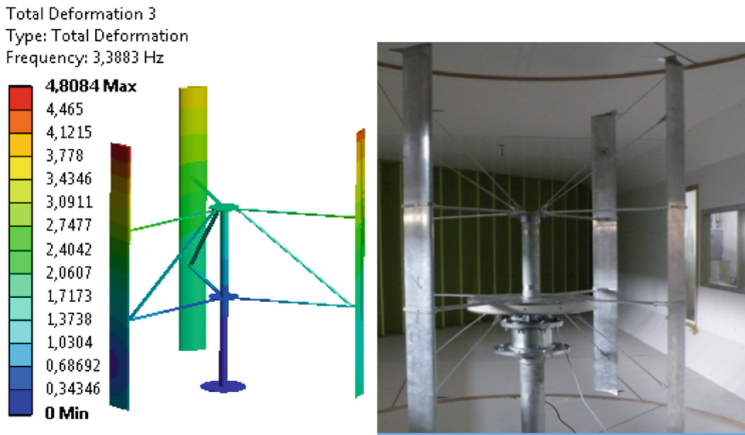
Fig. 1. Distribution of the magnetic induction vector on the cylindrical surface passing through the center of the active part of the winding and construction of the generator with wind turbine.
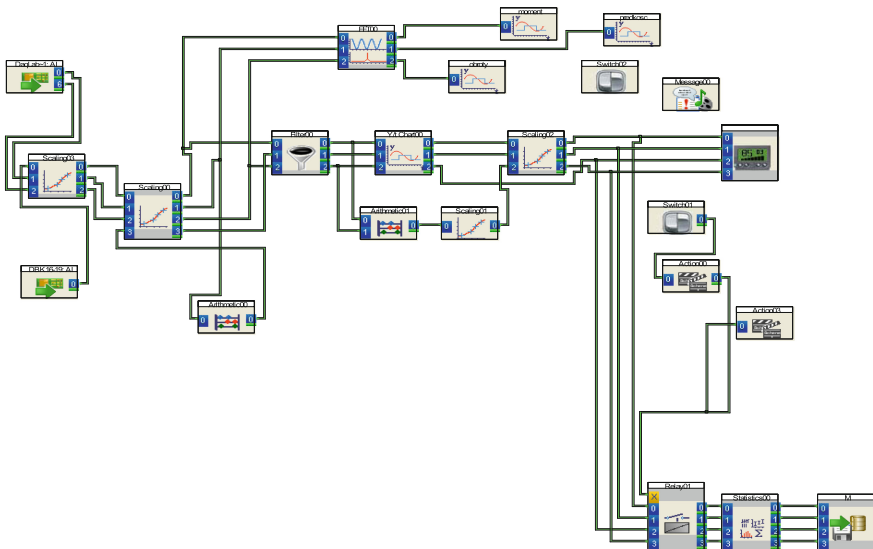
Fig. 2. Measurement system diagram created in DasyLab.

The presented generator is adapted to cooperate with small vertical turbines, in principle it creates an inseparable whole structure of the head. It is also possible to use it for a system with a horizontal axis. The disc form allows it to be embedded in the flow area [2]. It should be noted here that the generator shown has a very small starting torque, no latching torque and better smoothness of operation at very low rpm. Low power units work in a varying rotating system, so in order to cooperate with the network, an inverter is required. The quality of generated energy returned to the network depends only on the quality of the inverter (Fig. 2).

3 Results

The tests were carried out on a research site, where the generator was driven by an asynchronous motor controlled by an inverter in order to remove the some features of the generator. Then, a rotation reduction adapted to the generator was applied. Between the gearbox and the generator, a dynamometer was set up to measure the mechanical torque. In addition, the rotational speed of the generator was optically measured. The output of the generator was connected to a three-phase rectifier with filtration, obtaining a constant component, in order to more conveniently measure the electrical parameters of the generator [3]. The power was determined based on the measurement of the load current and the output voltage on the regulated active load. All results were recorded using a 16-bit USB-231 measurement system, operated from the DasyLab measurement software. The developed program carried out all the required measurements, calculated the necessary quantities like the power generated and the mechanical power supplied. On this basis, it determined the efficiency of the generator in real time. In addition, it carried out the spectral analysis of signals measured using Fourier fast transformation module. All parameters were measured, calculated, recorded in instantaneous values, as well as averaged samples in files separately.

The tests were carried out for two combinations of stator pole winding connections. One in serial mode giving a nominal voltage of 680 V (RMS) and the other being a parallel combination generating a nominal voltage of 340 V (RMS). Figure 3 shows the graph of changes in the output voltage in an unloaded system for both combinations of stator winding combinations (U and Ur lines), and the power of mechanical and aerodynamic losses of the generator itself. A characteristic feature of this solution are very small mechanical and aerodynamic losses, especially for low rotational speed.

Figure 4 shows the family of load characteristics for constant rotational speeds. The load curves are linear in nature, the characteristic feature is that the generator can be used in very wide limits of rotational speed variation. Useful speed is already 60 revolutions per minute. In the case of using an inverter, the generator can be used effectively within 60–300 or more of the rotational speed. It translates into a large range of wind speed variation.
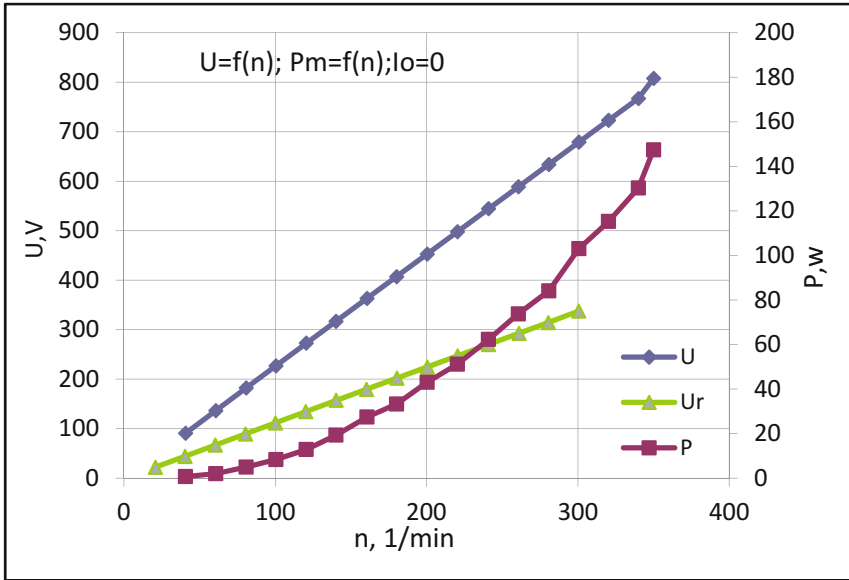
Fig. 3. Idle voltage characteristics and the power of mechanical losses.
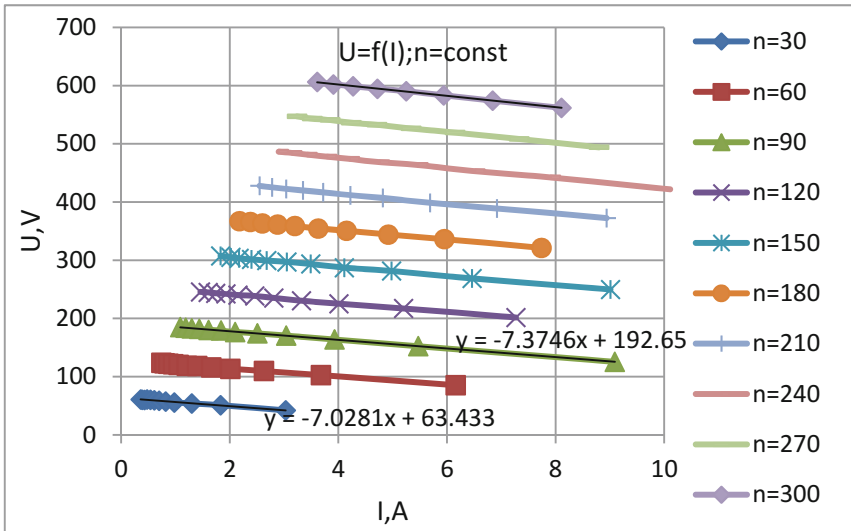
Fig. 4. External characteristics of the generator for the series winding combination.

Figure 5 shows the family of load characteristics in the case of parallel connection of stator windings. As we can see the voltage has fallen by a half, but the current efficiency is greater. The coefficients at the approximation function are three times smaller than the configuration from Fig. 4. Such a system is less sensitive to the voltage

drop on the windings at higher loads. These two operating modes allow for more flexible adaptation to the inverter and turbine.

Figure 6 shows the variability curve of the generator efficiency as a function of the rotational speed at a constant resistive load. The efficiency of the presented machines is very high and reaches 95%. It is flat in most of the working range.
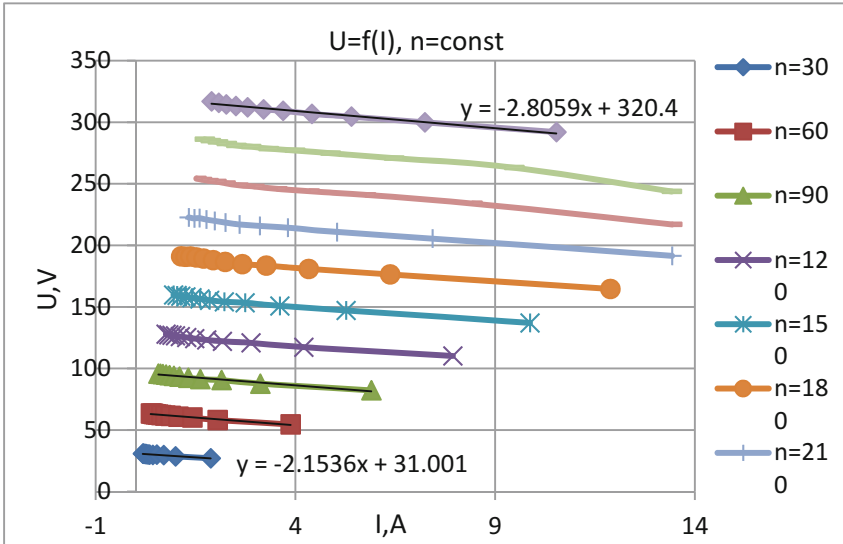
Fig. 5. Load characteristics for parallel combination of stator windings.
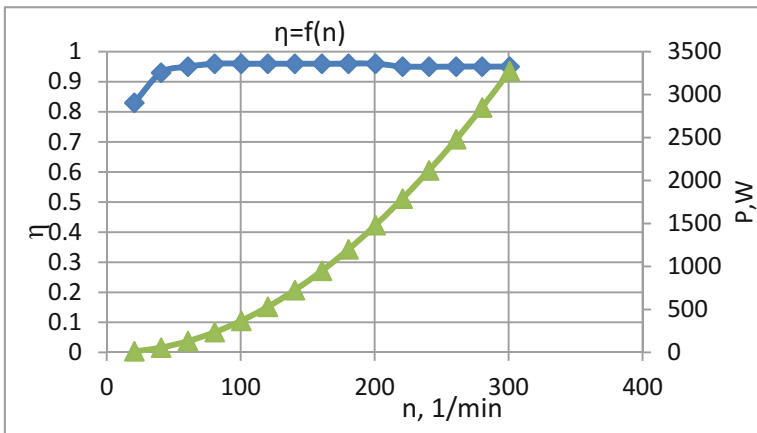
Fig. 6. The graph of generator efficiency as a function of the variability of the speed at constant active load.

Figure 7 presents the dependencies of variability of generator output voltage for two different cases of active load. Changing of the load for nominal conditions by 100% results in a small change in the output voltage of the generator, which proves its high stiffness, i.e. low relative dynamic resistance.

Figure 8 shows the distribution of the magnetic induction vector on the cylindrical surface passing through the center of the active part of the winding. This distribution confirms the correct the magnets magnetizing and the correct definition of boundary conditions of the Master/ Slave type.
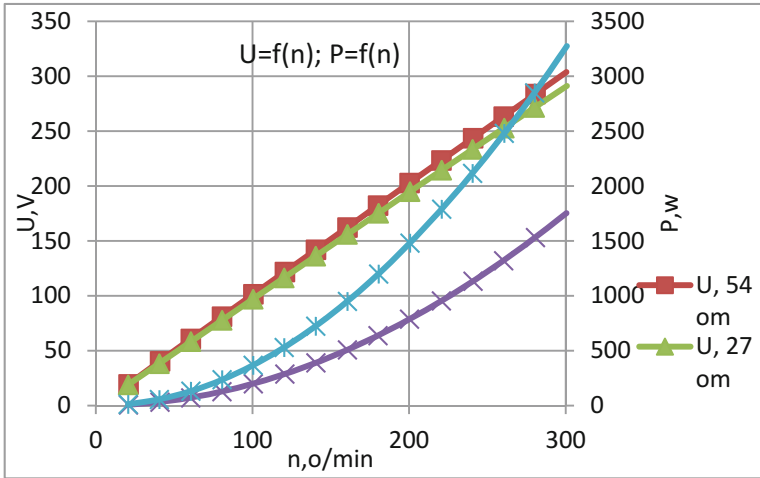
Fig. 7. The graph of voltage and power variation for two different constant active loads.
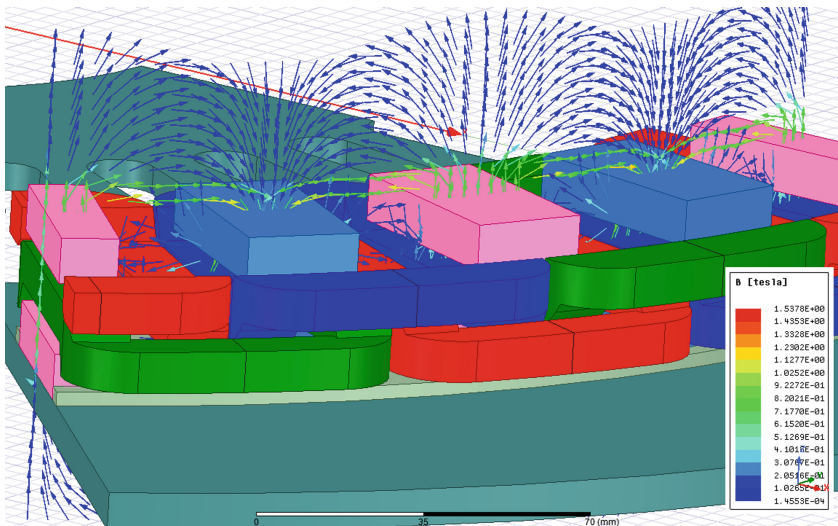
Fig. 8. Distribution of the magnetic induction vector on the cylindrical surface passing through the center of the active part of the winding.

4 Turbine-Generator System Powered by Power Grid via a Matrix Converter

The matrix converter is designed to connect generator phases directly to grid voltages. It is assumed that the grid is sufficiently rigid, and proper switching on is to ensure proper generator voltage for maintaining the set speed *w0_zad*. This speed results from the conditions of cooperation with a turbine and therefore the moment loading the generator from the side of the turbine *mobc* should be taken into account [4].

Figures 9, 10, 11 presents the results of simulation of dynamic states of the axial generator when powered from a matrix converter. The necessary voltages at the
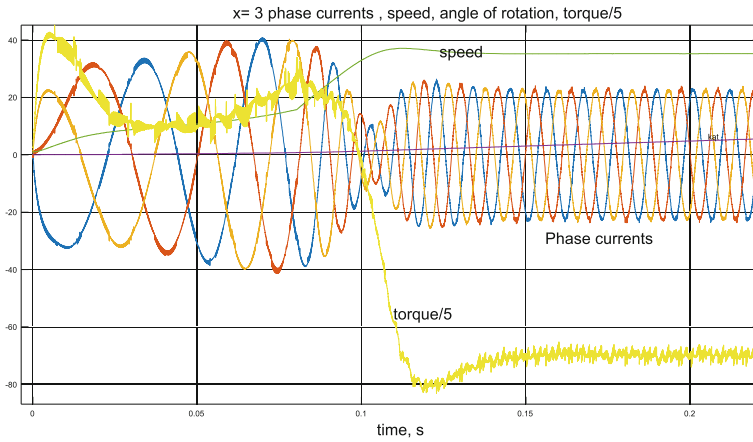
Fig. 9. The waveforms of phase currents, rotational speed, mechanical angle of rotation and electric torque (/5) at the set speed *w0_ad* = 31.4 [rad/s] and load moment *mobc* = −350 Nm.
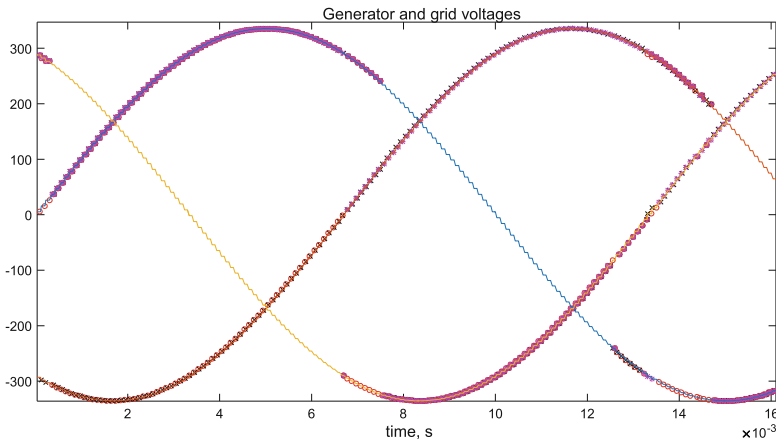
Fig. 10. The sequence of grid voltages (solid line) and the way of connecting the generator phases to these voltages: generator phase A - red o, phase B - black x, phase C - magenta *.

generator phases are calculated using the prediction method. These voltages are obtained by directly connecting the generator phases to the mains voltage via a matrix converter. Figure 9 shows start-up of the generator to the set speed w0_zad = 31.4 rad/s. Driving torque from the wind turbine was assumed as mobc = −350 Nm. Figure 10 shows how the phases of the generator are connected to the mains phases during the start-up. Figure 11 shows the waveforms of voltages obtained in this way on the phases of the generator.
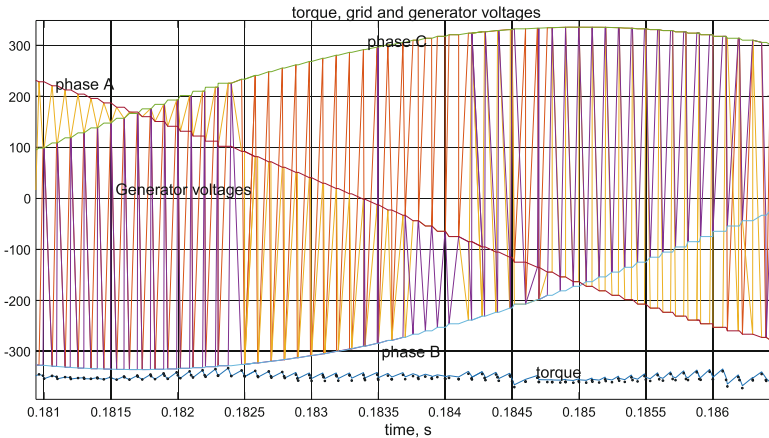
Fig. 11. The sequence of grid voltages of the phase A, B, C generator voltages based on grid voltages through a matrix converter and electromagnetic torque.

5   Conclusion

The coreless Axial flux permanent magnet generator of 3 kW rated power was presented. This generator was designed to work in the small wind turbine - matrix converter system. The properties of the discussed generator such as big ratio of machine diameter to its axial length, high efficiency, low rotational speed, lack of cogging torque makes it especially suitable for proposed application. The measurement system created for the investigated generator was presented [5]. The idle and load characteristic measured for two windings arrangements proved the generator design and its suitability for the intended application. Further investigation of the generator working in the grid connected matrix converter system was carried out by means of simulation. Those investigations showed very good performance of the designed system.

References

1. Kamper M.J., Wang R-J., Rossus F.G.: Analysis and performance of axial flux permanent-magnet with air-cored nonoverlapping concentrated stator windings. IEEE Trans. Ind. Appl. **44**(5) (2008)

2. Wang R-J., Kamper M.J.: Calculation of eddy current loss in axial field permanent-magnet machine with coreless stator. IEEE Trans. Energy Convers. **19**(3) (2004)
3. Wang R-J., Kamper M.J., Van der Westhuizen K., Gieras J.F.: Optimal design of coreless stator axial flux permanent-magnet generator. IEEE Trans. Magn. **41**(1) (2005)
4. Chirca M., Breban S., Oprea C.A., Radulescu M.M.: Analysis of innovative design variations for double-sided coreless-stator axial-flux permanent-magnet generator. IEEE (2014). 978-1-4799-4775-1/14/$31.00
5. Chan, T.F., Lai, L.L.: Axial flux permanent magnet synchronous generator for a direct-coupled wind turbine system. IEEE Trans. Energy Convers. **22**(1), 86–94 (2007)

Assessment of Image Processing Methods for the Determination of Propagation of Squat-Type Defects in Rails

Eligiusz Mieloszyk, Anita Milewska, and Sławomir Grulkowski[✉]

Faculty of Civil and Environmental Engineering, Gdańsk University
of Technology, Narutowicza Str. 11/12, 80-233 Gdańsk, Poland
{eligiusz.mieloszyk,anita.milewska,slawi}@pg.edu.pl

Abstract. We demonstrate the idea of squat-type defect measurement in the rail and the concept of tracking of the defect development using the techniques of image acquisition and image processing as well as the methods of metric spaces. We introduce the concepts of a set diameter $\delta(A)$ and the metric ρ_1, which come from the properties of plane figures, to compare and to observe the development of the defects.

We characterize the feasibility of the method to determine the dynamics of the defect development. The tests have shown that it is possible to apply the method with a camera during current diagnostic procedures provided that the distance to the rail is similar. Normalized metric enables easy comparison of the results and allows for the assessment of the reliability of the rails. The advantages of the method include simplicity and ability to observe the defects during the entire cycle of their development, which makes it possible to take the diagnostic decisions at the appropriate time.

Keywords: Squat-type defects · Image processing methods · Rolling contact fatigue (RCF) defects · Measurement of rail defects

1 Introduction

The defects in the rails can be classified as the defects at or outside the rail ends, and the defects caused by rail damaging, the defects developing during rail joining or rail repair. The rolling contact fatigue (RCF) defects are the most frequent defects located outside the rail ends. This group includes squat-type defects, which are indicated as 227-defects (Fig. 1).

Squat-type defect develop at the running surface of the rail in the region of the contact between the rail and the wheel. Squat-type defect has the form of a darker
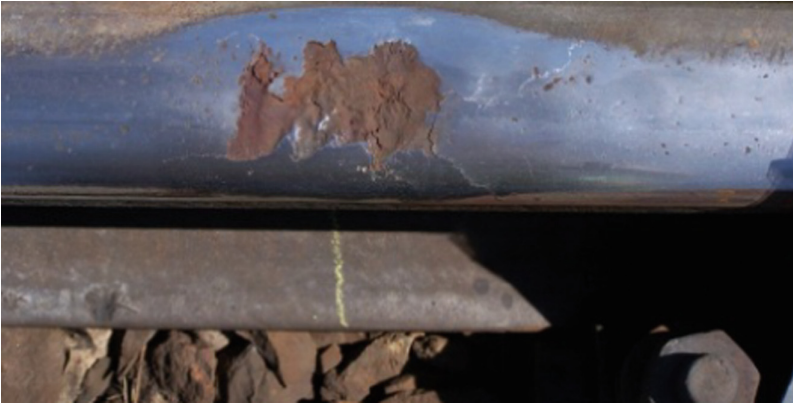
Fig. 1. Squat-type rail defect.

localized depression containing cracks with semi-circular arc or V-shape (Fig. 2). Squat can be located randomly as a single or multiple defects, and it is usually detected in the rail joints (in that case squat is classified as the defect due to the process of rail joining) [1].

Fig. 2. Typical forms of squat-type defects: V-shaped crack (left) and arc-shaped crack (right).

The number of squat-type defects reaches ca. 48% of all rail defects in some railway lines. The diagnostics of squat-type defects is challenging due to the physical effects and technical limitations such as the lack of devices enabling the measurement of the defects during rail exploitation.

The details of mechanisms of generation and development of squat-type defects are actually not known. In this project, we tried to analyze that challenges.

The investigations of defect development during the exploitation are difficult due to the limitations of observation automation of those processes and due to the inability to assess the early-stage defects. In fact, it was reported that high-speed photography cannot achieve sufficient level of accuracy. Therefore, the dynamics of the development of defects is not a criterion of their repair due to the problems with the measurements and the analysis [1–3].

2 Assessment of the Squat - Type Defects

We performed the measurements of the squat-type defects in rails of the running train line. We measured the parameters such as the length of depression, length of defect, the depth of depression and the depth of defect (Fig. 3).

Fig. 3. Measurement of defects.

The experimental section of the length of 31,5 km is located in the track #1 and #2 of the train line #131 Chorzów Batory – Tczew. The analysis was performed for the track #2. The rails were made of the R260 steel manufactured by the Katowice Steelworks (K). The rails were fastened to the sleepers using K-type rail clips. General repair of railroad surface in 2012 caused changes in the track #2 design at the analyzed section.

The studies showed that the rate of the defects' development is different at various locations in both rails [2, 3]. It means that the rails are more vulnerable to defect development (Table 1). The technical and exploitation conditions impacting the rate of defect development include: the type of rail, rail steel grade, type of track, type of sleepers, condition of other elements of rail surface, the type of rail traffic, the load of the line, train line geometry, brake and start zones.

From practical point of view it is important to automatize the measurements and to assess the dynamics of defect development in rail (Fig. 4) and to use the results for forecasting the maintenance works. This is crucial since the vehicles use the tracks with defected rails, which impacts the safety of passengers and freight.

3 Idea of the Measurement and Acquisition of Data on Squat - Type Defects

The cameras can be used to acquire the images of the squat-type defects. Formally, both the cameras and the photographs are non-metric [4].

We used KODAK camera to check the usefullness of those devices to measure plane linear deformations such as scratches, cracks, contours etc.

Table 1. Development of 227-squat defects

Defect number	Rail (R-right, L-left) (number of items)	Initial measurement		Measurement after 6 months		Measurement after 15 months		
		Length of depression [mm]	Length of defect [mm]	Length of depression [mm]	Length of defect [mm]	Length of depression [mm]	Length of defect [mm]	Depth of degression/defect [mm]
1	R	45	20	50	20	65	25	1,0
2	R	135	65	140	65	140	65	3,0
3	R	50	0	50	0	50	0	<1,0
4	L	50	0	65	10	65	45	6,0
5	R	120	0	120	0	130	10	1,0
6	L (2)	120	50	120	105	130	115	4,2
7	R	65	5	75	5	75	5	1,0
8	L	140	35	220	40	220	160	4,6
9	R	50	0	50	0	50	10	<1,0
10	L (3)	140	15	150	40	150	40	4,2
11	L (2)	100	15	100	15	115	30	2,8
12	L	55	0	90	0	100	0	<1,0
13	R	90	10	90	15	100	40	5,2
14	L (2)	115	30	130	60	145	85	3,2
15	R (4)	70	5	85	5	85	5	1,6
16	R	140	25	140	25	150	25	1,7

Fig. 4. Development of defect in the rail over a period of 15 months of rail exploitation.

We used line segments of different lengths and circles of different radii as models. To perform the measurement, we took a photo of a model line segment of the length of 100 mm from the distance 1,3 m using the focal distance of the camera of 200 mm. the photographs were taken at different inclination angles of the camera axis with respect to the CCD array. The most precise results were obtained for photographs taken in the plane parallel to the model plane since the principles of perspective projection that decide about reliable mapping of the model on the photograph [5]. The obtained errors were of the order of 1%. The largest error (10%) was observed for the inclination angle of 20°. The error of 4% was detected for inclination angle of 10°.

From the geometrical point of view, the squat-type defects can be treated as the plane set (plane geometric shape) A if the depth (usually 3 mm) is not taken into

account. This set can be easily identified (imaged) with the camera during the entire period of the experiment. How can the obtained images be used to assess the propagation of the defect? This can be achieved with the aid of the concept of metric spaces, the quantities and their properties. This allows to propose the following methodology.

In a metric space (X, ρ) [6] one define the diameter of the set A in X [7]. In the case of the set $A \neq \emptyset$ its diameter is the finite or infinite number $\delta(A)$ given by:

$$\delta(A) = \sup_{x,y \in A} \rho(x, y) \tag{1}$$

(sup indicates supremum).

Additionally, we have $\delta(\emptyset) = 0$. As a consequence, $0 \leq \delta(A) \leq \infty$ and for a bounded set A we obtain $\delta(A) < \infty$.

One can directly conclude from the definitione of the diameter of the set that

$$if\ A \subset B, then\ \delta(A) \leq \delta(B). \tag{2}$$

For two arbitrary non-empty sets A, B so that $A \cap B \neq \emptyset$ the following relations holds

$$\delta(A \cup B) \leq \delta(A) + \delta(B). \tag{3}$$

In a metric space (X, ρ) one can introduce a metric ρ_1 given by

$$\rho_1(x, y) = \frac{\rho(x, y)}{1 + \rho(x, y)}\ for\ x, y \in X \tag{4}$$

with the following interesting property [8]

$$0 \leq \rho_1(x, y) < 1 \tag{5}$$

The property of the metric ρ_1 is important from practical point of view since $\rho_1 \in\ <0, 1)$(Fig. 5).

The quantities changing in time for a given defect are the metric parameters of the defect (Table 1), as shown in the photograph (see Fig. 4).

We take the photo of the defect u in time t_0, so that $u(t_0)$ is imaged according to the rules, and the image (photograph) $z(t_0) \leftrightarrow A(t_0)$ is assigned to the defect u. The defect $u(t_0)$ corresponds now to the two-dimensional figure (plane set) $A(t_0) \subset R^2$. In this way, we introduces the correspondence between (t_0), $z(t_0), A(t_0)$, therefore $u(t_0) \leftrightarrow z(t_0) \leftrightarrow A(t_0)$. In this geometric shape $A(t_0)$ one can further assign the quantities: diameter $\delta(A(t_0))$ of the set $A(t_0)$, its area $\mu(A(t_0))$, center of gravity $S(t_0)$, and the distance between selected (characteristic) points $P(t_0), Q(t_0) \in A(t_0)$ in different metrics.

This procedure can be repeated in time instances $t_i = t_0 + \Delta t_i, i = 1, 2, 3, \ldots, n$ during defect analysis $u(t_i), i = 1, 2, 3, \ldots, n$. We will obtain finite sequences of the quantities $\delta(A(t_i)), \mu(A(t_i)), S(t_i), \rho(P(t_i), Q(t_i))$dla i $= 1, 2, 3, \ldots$, n, which enable acquisition of the analysis of the defect development in time.

One can introduce different metrics in a set $X = R^2$. In particular, rail metrics given by the equation

$$\rho(x,y) = \begin{cases} |x_1 - y_1| + |x_2| + |y_2| & for \; x_1 \neq y_1 \\ |x_2 - y_2| & for \; x_1 = y_1 \end{cases} \tag{6}$$

is an interesting metric to determine the development of the squat-type defect. The rail metric is especially useful to determine $\rho(P(t_i), Q(t_i))$ for i = 1, 2, 3, …, n or $\rho_1(P(t_i), Q(t_i))$ for i = 1, 2, 3, …, n.

Table 2 includes metric $\rho_1(x,y)$ calculated for the data from Table 1. This allows an easy comparison of the results. Additionally, selected results in the form of normalized trends of defect development are demonstrated in Fig. 5. Varying dynamics at different stages of defect development is shown here.

The parameters defined earlier can characterize the development of the defect and reveal different forms of the defect allowing for a more comprehensive quantitative analysis of the data. The meaning of the parameters of $\delta(A)$, $\mu(A)$ is demonstrated in Figs. 6, 7 and 8.

Figure 6 shows that in some situations the lengths and corresponding metric ρ_1 cannot be the objective parameters for shape assessment and defect development. In that case, $\delta(A)$ should be included in the measurement. The length of the defect is shorter than $\delta(A)$ in Figs. 6 and 8. The significance of that parameter is presented here. Since the defects are of different and very irregular character, a more precise quantity is

Table 2. Metric ρ_1 for the defects mentioned in Table 1

Defect number	Rail (R-right, L-left) (number of items)	Initial measurement ρ_1	Measurement after 6 months ρ_1	Measurement after 15 months ρ_1
1	R	0,9524	0,9524	0,9615
2	R	0,9848	0,9848	0,9848
3	R	0,0000	0,0000	0,0000
4	L	0,0000	0,9091	0,9783
5	R	0,0000	0,0000	0,9091
6	L (2)	0,9804	0,9906	0,9914
7	R	0,8333	0,8333	0,8333
8	L	0,9722	0,9756	0,9938
9	R	0,0000	0,0000	0,9091
10	L (3)	0,9375	0,9756	0,9756
11	L (2)	0,0000	0,0000	0,9375
12	L	0,0000	0,0000	0,0000
13	R	0,9375	0,9375	0,9677
14	L (2)	0,0000	0,0000	0,0000
15	R (4)	0,9091	0,9375	0,9756
16	R	0,9677	0,9836	0,9884
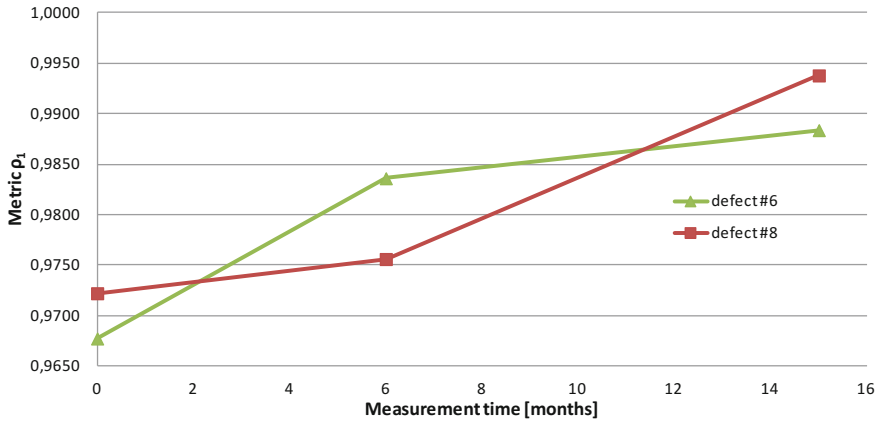
Fig. 5. Metric ρ_1 for two selected defects (# 6 and 8 in Table 2) measured three times.

Fig. 6. Defect with a short length and a significant width (the difference between defect length and its diameter $\delta(A)$)

the defect area $\mu(A)$. The significance of the area $\mu(A)$ is given in Figs. 6, 7 and 8. On the other hand, the significance of the train metric in the measurement procedure of the squat-type defects in rails is revealed in Figs. 7 and 8.

Fig. 7. Defect with a significant length and a significant width.
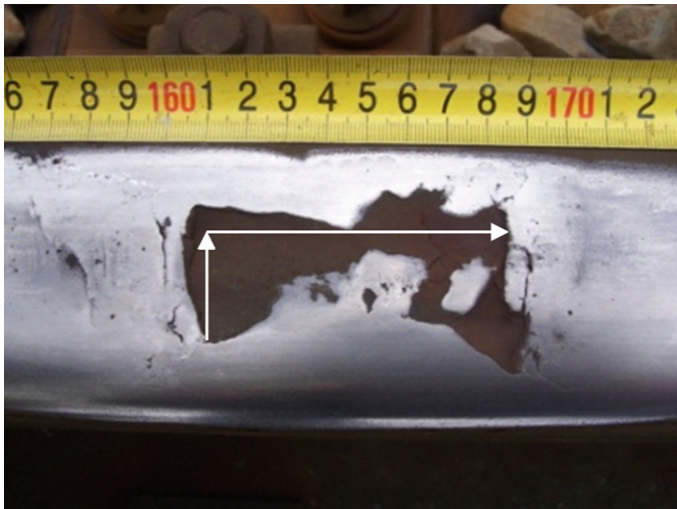
Fig. 8. Defect with a significant length, a very irregular shape and a small area.

4   Conclusion

The images of the defects $u(t_i), i = 1, 2, 3, \ldots, n$ are the starting points of the proposed method.

The method enables registration, identification, characterization and archiving the defect during current maintenance and diagnostic works.

The method does not require any sophisticated equipment or highly experienced specialists.

The photograph of the defect should be taken in the plane parallel to the surface of the object. This provides homogenous scaling of the entire photo and minimizes the errors [9].

The studies showed that the photo can be taken in a standing position to improve the acquisition of the images.

The proposed methodology is not subjective and does not involve labor consumption of the applied technologies.

The analytical apparatus is relatively simple allowing for the assessment of the trends in dynamics of defect development.

The proposed method using metric spaces is effective in determination of dynamics of defect development.

The results can be compared using the metric ρ_1 with the property (5).

The analyses of the metric quantities for $A(t_i), i = 1, 2, 3, \ldots, n$ offer opportunities to use the results for the determination of the reliability of rails.

The proposed methodology opens new gates towards development of the devices for quantitative analysis of the squat-type defects.

The analyses shown in this paper represent contribution for the explanation of the mechanism of generation and development of squat-type defects of rails.

The advantage of the method is the possibility to perform the measurements during regular exploitation of the rails in the track.

However, the method is not able to determine the depth of defect.

References

1. Zariczny, J., Grulkowski, S.: Observation as the fundamental tool in the diagnostics of the rails. Trans. Infrastruct. **6**, 10–15 (2014). (in Polish)
2. Zariczny, J., Grulkowski, S.: Assessment criteria of the exploitation duration of the rails. Rail Trans. Tech. **9**, 4245–4255 (2012). (in Polish)
3. Zariczny, J., Grulkowski, S.: Characterization of the rail defects detected in the train line # 131 Chorzów Batory–Tczew, with particular attention to the 227-squat defetcs. Scientific and Technical Papers of SITK RP Branch in Krakow, no. 3, 349–363 (2012). (in Polish)
4. Bernasik, J.: Lectures on Photogrametry. Kraków (2006). (in Polish)
5. Bieliński, A.: Descriptive geometry. Warsaw University of Technology Publishers, Warsaw (2015). (in Polish)
6. Jänich, K.: Topology. Springer, New York (1998)
7. Engelking, R.: General Topology. Heldermann Verlag, Berlin (1989)
8. Mieloszyk, E.: Non-classical operational calculus in application to generalized dynamical systems. Polish Academy of Sciences Scientific Publishers, Gdańsk (2008)
9. Koc, W., Chrostowski, P.: Computer-aided design of railroad horizontal arc areas in adapting to satellite measurements. J. Transp. Eng. ASCE **140**(3), 1–8 (2014)

The Multiplatform Environment
for Simulation and Features Estimation
of Mixed-Signal Devices

Krystyna Maria Noga[2] , Beata Palczynska[1,2(✉)] ,
and Romuald Masnicki[2]

[1] Gdansk University of Technology, 11/12 Gabriela Narutowicza Street, 80-233
Gdansk, Poland
`beata.palczynska@pg.edu.pl`
[2] Gdynia Maritime University, Morska 81-87, 80-225 Gdynia, Poland
`k.noga@we.am.gdynia.pl`, `romas@am.gdynia.pl`

Abstract. The use of simulation laboratories is gaining popularity in the domains of engineering programs. However, the experience in teaching shows that the simulation itself is not very effective in didactic processes. Teaching processes in the field of specialist subjects, designed for students of technical universities, should be based on direct operations performed by the student on real devices. At the same time, at the later stages of didactic processes, modern computer tools and techniques that enable modeling, simulations and measurements of projected systems or devices cannot be omitted. The article presents an example of applications of computer technology in the analysis of system properties at the stage of their designing, commissioning and testing of prototype properties. Based on the chosen hybrid system, in mixed analog-digital technology, the novel techniques of testing its functional properties were presented. The multiplatform combined with graphical programming and simulation software and hardware allows comparing a schematically captured circuit with a prototype of the same design. The platform based on NI myDAQ instruments and the Multisim Circuit Design software is presented. To illustrate its capabilities, the generator of the sawtooth wave with a mixed analogue-digital structure was tested. The integration between NI Multisim and NI myDAQ makes possible to correlate real and simulated measurements in a single interface. The possibility of presented techniques combined use allows optimizing the processes of designing, commissioning, testing and teaching the properties of electronic circuits.

Keywords: Digital signal processing · Multisim · LabVIEW software

1 Introduction

The development of computer techniques in the field of designing, simulating and testing the properties of electronic circuits contributes to the simplification and optimization of the processes of starting new systems. Figure 1 shows typical procedures implemented during the design and evaluation of the properties of the electronic system

using computer techniques. Designing a device with new original functional features begins long before appearing on the market. Its properties are expressed in the assumptions for the new construction. Recognition of problems usually leads to many different concepts of how to organize the device configuration. Choosing the optimal concept is a difficult task. Lack of simulation (Fig. 1 - part of the algorithm marked in red) caused the necessity to construct a prototype device for each concept. It was expensive and long work. The introduction of simulation to the design algorithm makes it possible to achieve a positive project result faster, while saving time and money in a much more effective way.

In the classic approach, only actual conventional instruments were used in the testing procedures of the device being built, there was no simulation stage, it appeared along with the development of computer techniques. These techniques are successfully introduced at the stage of: - testing the physical model, - testing the system being built, - to emulate the devices used in the study of the designed system.

Nowadays, simulation studies are becoming more important in many areas of engineering activity at the design stage. Computer techniques, like simulations, measurements and control, become an important element of didactics. Also many student projects focus on the construction of a device with predefined properties and functions. Simulations are also applicable here. Many bad solutions and concepts can be excluded and omitted in this way. The use of simulation laboratories is gaining popularity in the domains of engineering programs. However, the experience in teaching supported by numerous studies [1–8] shows that the simulation itself is not very effective in didactic processes. The simulation becomes effective when it is accompanied by the student's activity in the study of actual systems. Currently, there are many programs that enable simulation of electronics, electrical engineering, digital technology and digital processing systems, e.g. SPICE, Multisim, LabVIEW, Vissim, Mathcad, Matlab [9, 10]. The SPICE and Multisim environments, with the SPICE program being the core program and the next version of Multisim, are one of the more frequently used simulation software in research, design and teaching [10].

The cross platform combined with integrated software and hardware, allows to compare real and simulated measurements in a single interface. The platform connects the theoretical model of the tested device to real components in simulation and integrates physical analysis at once in experimentation.

In the paper the cross platform for simulation and measurement of mixed-signal devices based on National Instruments (NI) myDAQ instrumentation and the Multisim Circuit Design software is presented. To illustrate its capabilities, the generator of the sawtooth wave with a mixed analogue-digital structure was tested. The integration of these environments into one workspace enables simulation and emulation of the operation of models of various systems, as well as testing of relevant real systems. This creates the conditions for the implementation of the didactic program, in which the student combine traditional hands-on activity with the capabilities of modern computer tools.

This paper is organized as follows. In Sect. 2, the NI myDAQ design template in Multisim is described. Section 3 presents the results of simulation using NI myDAQ instruments in the Multisim environment. The comparison of simulation and results of real data generating are shown in Sect. 4. Concluding remarks are drawn in last Sect. 5.
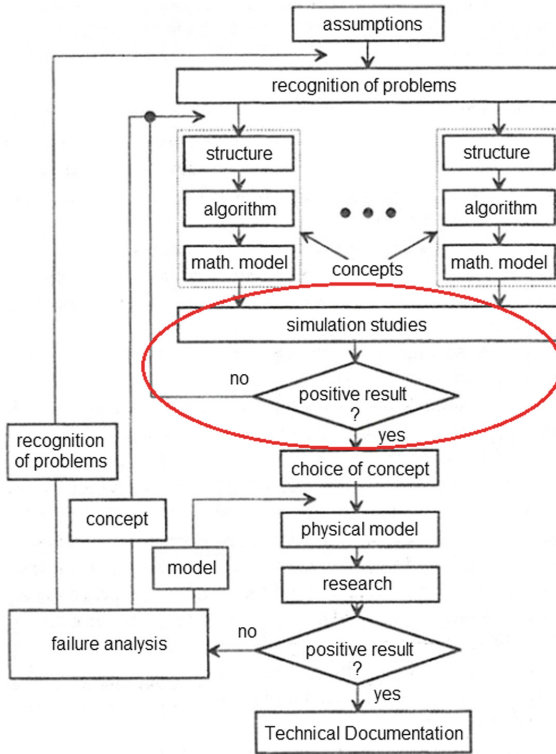
Fig. 1. An algorithm for designing and implementing an electronic system [11].

2   The NI myDAQ Design Template

The Multisim Circuit Design software is an attractive tool supporting the design of circuits. It allows to enter a schematic diagram of a simulated system that can be connected to various measuring instruments, e.g. oscilloscope, logic analyzer, voltmeter, ammeter, digital multimeter, etc. An important feature of the Multisim environment is the possibility of extending the package with the VHDL and Verilog hardware description interpreter module, as well as Utilboard and Utilroute programs for creating printed circuit boards based on the scheme introduced in the Multisim program. It is also the ability to modify the vast majority of component models by the user in accordance with specific own needs. The user, using the elements available in the library, can create his own element in the form of a sub-circuit or a model introduced using the SPICE system syntax. Multisim has the user database in which files containing a data of components created by a given user are stored; it can be used to store new components.

Multisim, depending on the type of tested circuits, offers different types of simulators, which are based on different programming languages of the equipment, e.g. SPICE, VHDL, Verilog [9]. Coordination of communication between models in

SPICE, VHDL and Verilog takes place automatically, full cooperation is ensured. Multisim is based on industry standard SPICE 3F5. It supports models created using standard SPICE syntax. A model can be created using the Model Makers, by assigning values to the parameters of a primitive model, or by creating a sub-circuit model. A primitive model is a model that is defined by a collection of parameters. Many electronics devices are not represented by primitives but are still well suited as SPICE models. Sub-circuit models are used to capture the characteristics of these models. In Multisim, it is possible to import/export a SPICE netlist. The ability to open *.cir* netlists allows for the generation of simple schematics upon import (see Fig. 2).
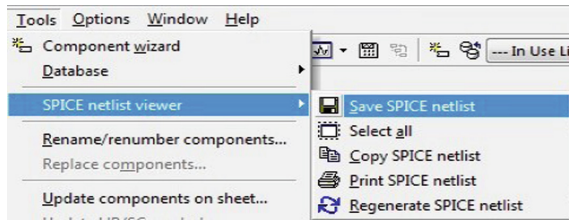
Fig. 2. The toolbar of a Multisim front panel.

2.1 Schematic Capture of a Mixed-Signal Device

The diagram of a system with a mixed analogue-digital structure is shown in Fig. 3. The designed system (*simulated/real device* in Fig. 3) generating the sawtooth wave is built using the following components. In order to generate the assumed waveform, two reversing, synchronous binary counters based on integrated circuit 74193 N, are connected in series, so that they work according to graph, as shown in (1).

$$255_{10} \rightarrow 254_{10} \rightarrow \ldots \rightarrow 2_{10} \rightarrow 1_{10} \rightarrow 0_{10} \rightarrow 255_{10}\ldots \tag{1}$$

The f_{clk} clock signal (*out_1*) controls the state of the counters. The counters work according to the required graph, i.e. they count down, operating in the subtraction mode. Counter outputs (8 lines) are connected to the inputs of an 8-bit digital-to-analog converter DAC 0808. In the device under consideration, the frequency of the sawtooth waveform (Fig. 3) is equal to $f_{clk}/512$.

The classic approach to testing the prototype of the designed device needs the use of an independent digital signal generator (generating f_{clk} signal in Fig. 3) and oscilloscope (connected to *out_1* and *out_2* test points) as well as appropriate test procedures performed by test staff. Of course, a complete prototype of the device must be made before. Only then construction errors can be detected, committed during the design and construction of the device's prototype. Any changes to the prototype made, including changes in the settings or processing characteristics of the individual functional blocks, require tedious operations usually associated with the exchange (soldering) of the relevant components of the system.
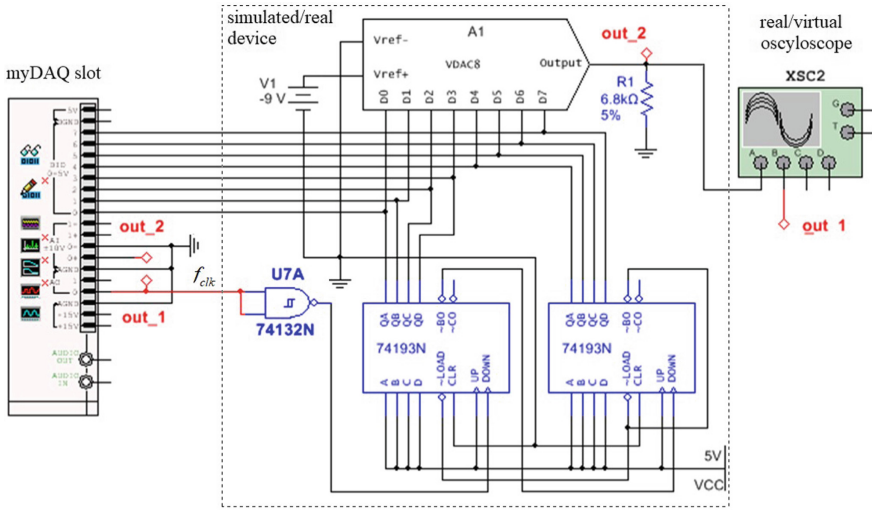
Fig. 3. The diagram of a generator of a sawtooth wave.

Simulations of the designed device in Multisim, before making a prototype, eliminate the majority of design errors and perform functional tests of the device. The design template found in Multisim main toolbar has the myDAQ connecting slots (see Fig. 3). The template allows to draw wires connecting a captured schematic with the lines on the virtual myDAQ device. The ideal elements available in the Multisim libraries do not always correctly represent the real systems [7, 8]. The models of actual elements available in the Multisim environment were used to build the sawtooth generator. Actual elements include the occurrence of transient states that are not significant or are suppressed in real systems but may appear in simulations using ideal elements. The methods of simulation of basic static and dynamic parameters are also important. Therefore, the image from the oscilloscope (see Fig. 4) shows so-called pins. They can be eliminated by using an appropriate filter on the D/A converter output. The application of Schmidt's gate (see Fig. 3) was forced to obtain the right steepness of rectangular waves from the signal generated by NI myDAQ.

In the same way it would be connected a real breadboard circuit of prototype device to the actual myDAQ device. It can fully simulate the hardware instrumentation (a signal generator, an oscilloscope) within the virtual environment. To generate the clock signal the test system uses the National Instruments myDAQ virtual instrument: a function generator (see Fig. 5). The output signal (*out_2*) can be presented on a virtual oscilloscope *XSC2* (see Fig. 4). At the same time, the signal *out_2* is given to the inputs of the myDAQ virtual oscilloscope.

2.2   Instrumentation Simulated Using NI myDAQ

The Multisim software integrates myDAQ instruments within the design template. There are eight different measurement and signal generation tools: a digital multimeter

(DMM), an oscilloscope (SCOPE), a function generator (FGEN), a Bode analyzer (Bode), a dynamic signal analyzer (DSA), an arbitrary waveform generator (ARB), a digital reader (DIGIN) and writer (DIGOUT). A specific instrument can be enabled or disabled in simulation.
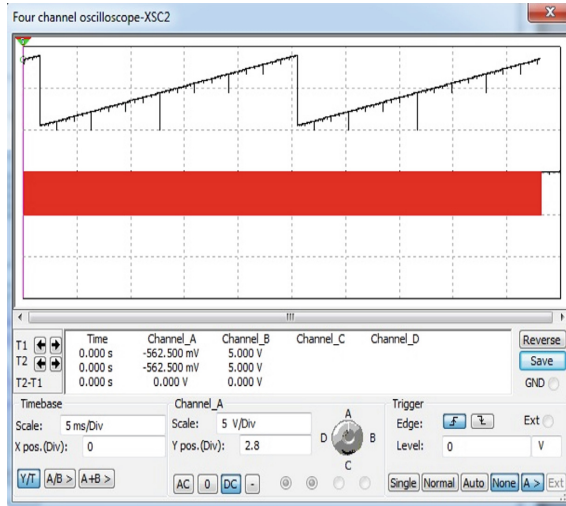
Fig. 4. The Multisim plots of an output sawtooth wave (black line) and a clock signal (red line).
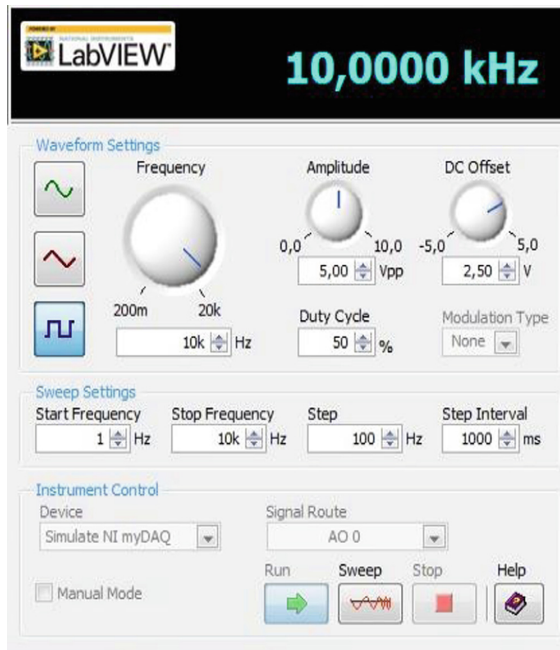
Fig. 5. The myDAQ virtual generator of the rectangular signal.

3   The Schematic Capture Simulation

The simulation is performed using NI myDAQ instruments in the Multisim environment. The function generator FGEN and the oscilloscope SCOPE should be enabled during a simulation. The simulated data appear in the front panel of a virtual SCOPE (see Fig. 6).

After successfully simulating the schematic capture, it could acquire the real data from myDAQ and compare the results.
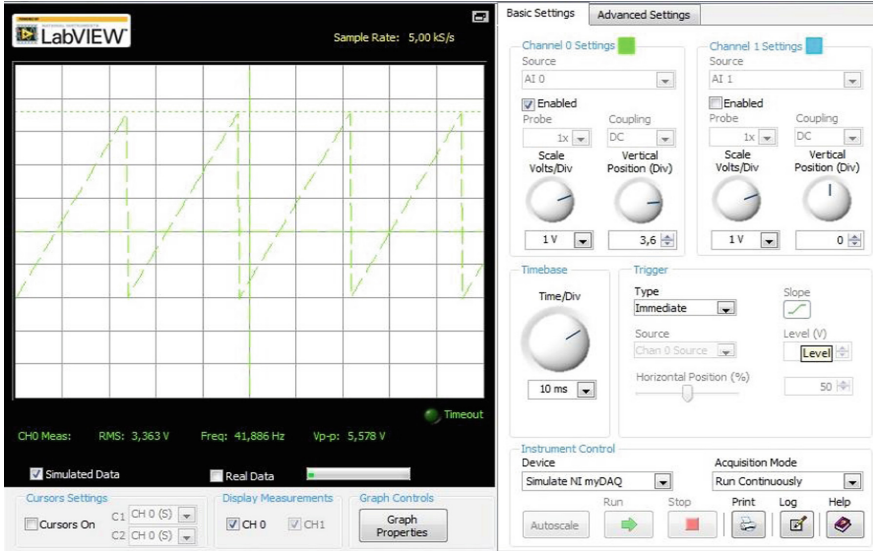
Fig. 6. The front panel of the SCOPE with the sawtooth waveform of the schematic capture simulation

4   Testing a Prototype Device

After construction the prototype of the device with the structure corresponding the schematic capture in Multisim, it should be connected to the myDAQ right slots. In the Multisim environment, under Instrument Control in the front panels the device corresponding to "NI myDAQ" instead "Simulate NI myDAQ" has to be chosen (see Fig. 7).

Fig. 7. Selecting myDAQ to generate real data.

The NI myDAQ is a data acquisition device that provided analog input (AI), analog output (AO), digital input and output (DIO), audio, power supplies, and digital multimeter (DMM) functions in a compact USB device. The input channels can be configured either as general-purpose high-impedance differential voltage input or as audio input. The analog inputs are multiplexed, meaning a single analog-to-digital converter (ADC) is used to sample both channels. They can measure up to ± 10 V signals, at the rate of up to 200 kS/s per channel, so they are useful for waveform acquisition. Analog inputs are used in the DMM, the SCOPE, the DSA and the Bode instruments. There are two analog output channels on NI myDAQ. These channels can be configured as either general-purpose voltage output or audio output. Both channels have a dedicated digital-to-analog converter (DAC), so they can update the output signals simultaneously. In general-purpose mode analog outputs can be updated at the rate of up to 200 kS/s per channel, making them useful for waveform generation up to ± 10 V. They are used in the FGEN, the ARG and the Bode instruments.

During testing prototype of the device and a real data generation, the FGEN instrument is generating an analogue clock signal f_{clk} and a sawtooth signal from a digital-to-analog converter DAC 0808 (*out_2*) is given on the analog inputs of the myDAQ. The SCOPE instrument is running.

Figure 8 illustrates the simulated data in Multisim against the real data obtained during testing of the device's prototype. The visible shift between the two waveforms results from the difference in the frequency of the f_{clk} signals.
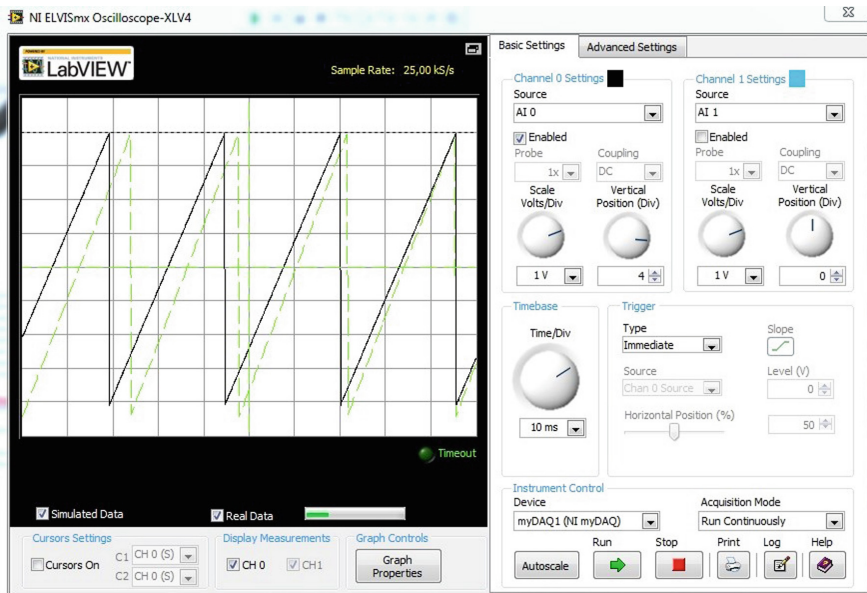
Fig. 8. The sawtooth waveforms obtained from simulation (green line) and real data (black line) generated during testing of the prototype.

The results of operation of real system could be also compared with the waveform plot from a stand-alone oscilloscope (see Fig. 9). The waveform obtained in the myDAQ SCOPE overlaps the waveform observed on the oscilloscope.
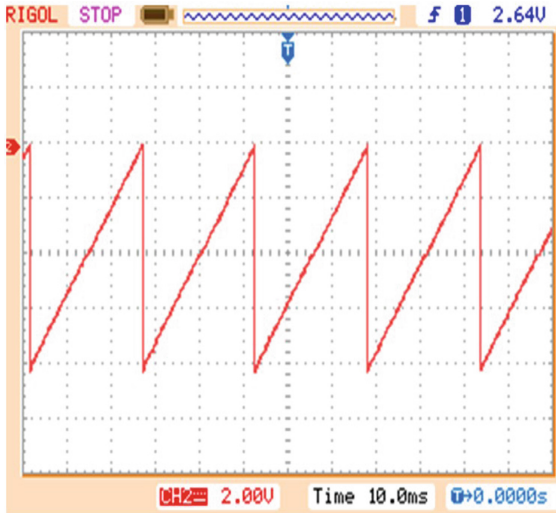
Fig. 9. The sawtooth waveforms obtained from a stand-alone oscilloscope.

5   Conclusion

The nowadays technology provides tools that enable effective testing and correcting the structure of designed devices. The classic approach to the design and testing of new devices is replaced by computer-assisted operations, greatly facilitating the processes of designing and implementing new solutions. In this way, a large part of the process of testing the properties of the actual systems is transferred from the hardware testing area to the virtual test, i.e. software area.

The Multisim programming environment in cooperation with virtual instruments operating in LabVIEW is a cross platform that creates new possibilities in effective designing and testing of both analog and digital devices as well as mixed-signal devices.

Didactic values of the cross platform should also be appreciated. In addition to familiarizing with new technologies, including simulation and device emulation techniques, students can conduct research on the properties of real elements and devices, with a small additional workload.

References

1. Taher, M., Khan, A.: Comparison of simulation-based and hands-on teaching methodologies on students' learning in an engineering technology program. In: QScience Proceedings of Engineering Leaders Conference. http://dx.doi.org/10.5339/qproc.2015.elc2014.58 (2014). Accessed 05 Feb 2018

2. Mahata, S., Maiti, A., Maiti, C. K.: Cost effective web based electronics laboratory using NI Multisim, LabVIEW and Elvis II. In: International Conference on Technology for Education (T4E), pp. 242–243 (2010)

3. Dai, K., Zeng, S., Huang, L., Wang, N.: The application of mixed software simulation platform based on Multisim and MATLAB for electronic specialty experiment teaching. In: International Conference on Computational Intelligence and Software Engineering (CISE), pp. 1–4 (2009)

4. Yang-Mei, L.,Bo, C.: Electronic circuit virtual laboratory based on LabVIEW and Multisim. In: 7th International Conference on Intelligent Computation Technology and Automation (ICICTA), pp. 222–225 (2014)

5. Azaklar, S., Korkmaz, H.: A remotely accessible and configurable electronics laboratory implementation by using LabVIEW. Comput. Appl. Eng. Educ. **18**(4), 709–720 (2010)

6. Palczynska, B., Noga, K.M.: Teaching digital filters design in electrical engineering. In: Fifth International Symposium Communication Systems Networks and Digital Signal Processing, pp. 866–869. University of Patras (2006)

7. Noga, K.M., Radwanski, M.: Using the virtual model in teaching digital signal processing. In: Technological Developments in Education and Automation, pp. 195–200. Springer (2010)

8. Noga, K.M., Radwanski, M.: Our experiences in teaching of digital logic. In: Innovations in E-learning, Instruction Technology, Assessment and Engineering Education, pp. 237–242. Springer (2007)

9. Hulewic, A., Krawiecki, Z.: Simulation programs for electronic analog circuits, (in Polish). Pozn. Univ. Technol. Acad. J. Electr. Eng. **88**, 57–66 (2016) (Poznan)

10. National Instruments Multisim http://www.ni.com/multisim/. Accessed 05 Feb 2018

11. Bolikowski, J.: Basics of designing intelligent measuring transducers of electrical quantities, (in Polish). Wydawnictwo Wyższej Szkoły Inżynierskiej w Zielonej Górze, Zielona Góra (1993)

Monitoring the Parameters of Industrial Robots

Paweł Obal, Andrzej Burghardt[(✉)], Krzysztof Kurc,
Dariusz Szybicki, and Piotr Gierlak

Rzeszow University of Technology, al. Powstańców Warszawy 12, 35-959
Rzeszów, Poland
{p.obal,andrzejb}@prz.edu.pl

Abstract. The article presents two commercial tools used to monitor the parameters of robot operation. They are briefly characterised, and their advantages, disadvantages and limitations described. The indicated shortcomings led to the development of a proprietary program, which allows for monitoring the parameters of ABB robot operation. The application's task is to collect data about the parameters of the robot's movements and store them as variables intended for further use. The study method was to capture the data transmitted between the controller and the program. The *Wireshark* software was used to identify the manner of communication, syntax and content of the data packages transmitted. The obtained data enabled the writing of a program for communication and registration of the defined robot parameters using a Matlab script file. In order to check the correctness of application performance, a numerical experiment was carried out, which involved connecting to a virtual controller, and an experiment involving connecting to a real controller. The performed tests confirmed the correctness of the developed application.

Keywords: ABB robots · Wireshark · Robot operation monitoring
RobotStudio · ABB test signal viewer · Matlab

1 Introduction

Monitoring the parameters of robot operation is carried out in order to manage the production process, to control the movement of components or details, as well as to monitor the alarms. To achieve this, the robotic stations are integrated with SCADA systems [1]. Apart from SCADA systems, Manufacturing Execution Systems (MES) and FIX software can also be used for the monitoring purposes [7]. There is also the option of developing your own applications using industrial network standards - the TCP/IP protocol and Ethernet standard [2, 3]. Additionally, metering of the robotic stations allows for the analysis or identification of vibrations in order to minimise tool damage [4]. Other reasons for controlling the status of robot performance are the economic aspects, and the selection of parameters in order to reduce the demand for energy [6]. This work considers the issue of monitoring the parameters of robot performance in real time using dedicated external software.

The ever-higher costs of human labour and technological advances are the main factors leading to the increased implementation of robots in technological processes. The robotisation of an ever-wider range of tasks requires constant development of both software and hardware. Thus, robot manufacturers equip their controller software with additional options facilitating the programming of specific technological processes.

The development of software, and the diversity of robotised processes, are often associated with the necessity to select a series of parameters and factors. They can affect the decision whether to manage additional devices cooperating with the manipulator or extend the possibilities of robot movement programming. Some examples of robotised applications requiring appropriate parameter selection are:

- painting,
- welding,
- sealing,
- cutting,
- force control applications.

The first four of the above-mentioned applications have been used in industry for years. Their programming requires the introduction of values of a series of parameters according to the applied technology. The force control applications have been developed and implemented recently [5]. ABB robots provide a good example as they can be equipped in the *Force Control Pressure* application. It allows for the control of the force exerted by the robot tool on the environment. The application performance scheme is presented in Fig. 1.
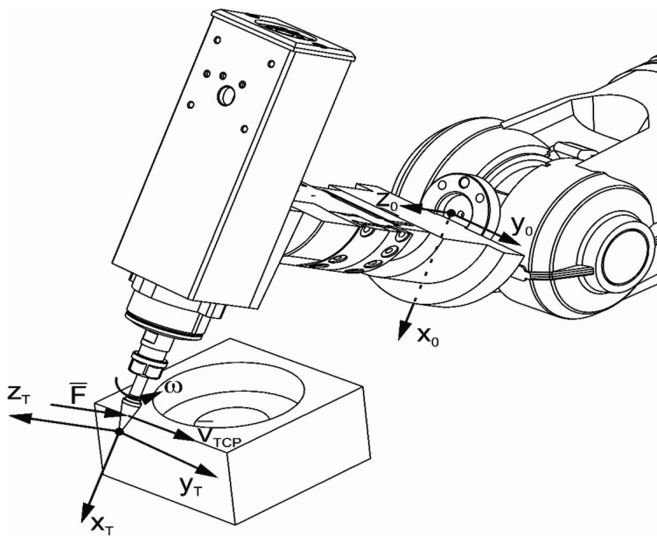
Fig. 1. The scheme of *FC Pressure* strategy

The application is useful in compensating for errors in the positioning of the processed detail and when it is necessary to maintain constant contact power between the tool and the detail. Using the *Force Control* application requires the determination of a series of parameters, such as: speed changes (threshold [%]), force change rate [N/s], dumping [%], maximum time of waiting for contact (Timeout [s]), percentage of the set force with which the robot initiates the process (Zero Contact Force [%]), the level of filtration for the low-pass filter (Noise level [Hz]).

There is no information for how to select the coefficients for a specific process. Each process requires the execution of time-consuming and costly tests. During the tests, the robot performance parameters must be monitored and recorded to build a knowledge base.

The article presents two commercial tools used to monitor the parameters of robot operation. They have been briefly characterised and their advantages, disadvantages and limitations have been enumerated. The indicated shortcomings let to the development of a proprietary program, which allows for the monitoring of the parameters of ABB robot operation.

2   Review of Tools for Parameter Monitoring

Monitoring the performance of ABB robots is possible by means of two programs. The first one is *Signal Analyzer*, being part of the *RobotStudio* software. *RobotStudio* is software intended to program and simulate the performance of ABB industrial robots. The communication between *RobotStudio* and the robot's controller takes place over an Ethernet connection. The *Signal Analyzer* program screen is presented in Fig. 2.
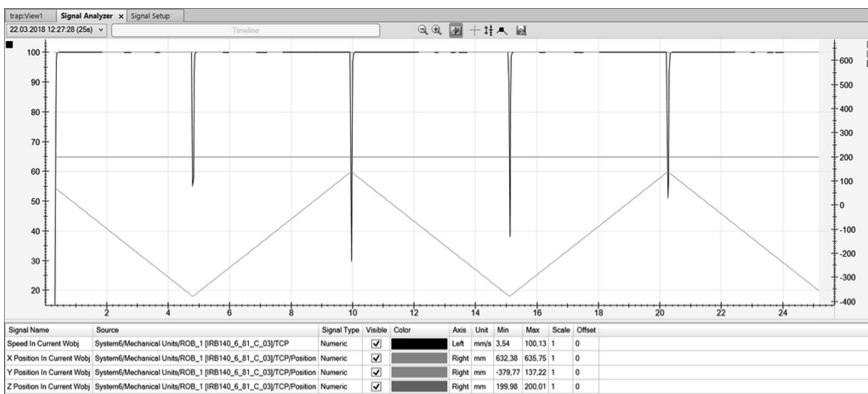
Fig. 2. *Signal Analyzer* program screen.

This program allows for real-time registering of the performance parameters of robots connected to controllers. These parameters include: the tool centre point (TCP) position, its current speed, orientation, total power generated by the motors, controller input and output status etc. The data are visualised in the form of a graph generated in real time. The program also allows for recording the registered data in the form of a text file.

The second program is an independent application *ABB Test Signal Viewer*, which allows for reading the values of so-called Test Signals. The application communicates with the controller by means of the TCP/IP protocol. The program screen is presented in Fig. 3.
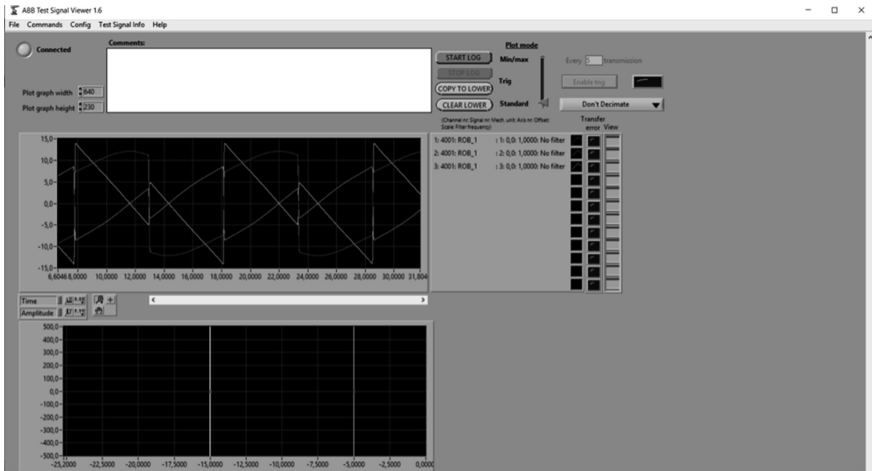
Fig. 3. Test Signal Viewer application screen

The program supports twelve channels to which any test signal can be assigned. The following test signals can be identified: angle parameters of movement of selected axes, drive and brakes control system signals etc. The signal description is included in the manuals provided by the manufacturer. The data are visualised in the form of a graph generated in real time. They can be saved in the form of a text file.

The two mentioned commercial programs allow for preview and registering of only selected parameters of ABB industrial robots. They do not offer the possibility of data processing in real-time or streaming them to another program. What is more, Test Signal Viewer has the measurement time limited to twenty-five seconds. The lack of the ability to process the data in real-time prevents these tools from being used in research on coefficient adaptation algorithms in robotic applications or custom control systems for ABB robots. Thus, an attempt has been made to write our own application to monitor the parameters of robot performance.

3   A Proprietary Application for Monitoring the Parameters of Robot Performance

The indicated limitations of the available programs or add-ons were the source of inspiration for developing our own software. It was assumed that the application will allow for the monitoring, recording and sending of the performance parameters of a robotic station.

The first step was to decipher the way in which the robot controller communicates with the commercially available software. To acquire this information, the data package

sniffing method was used regarding the data sent between *Test Signal Viewer* installed on a PC and the IRC 5 robot controller. The *Wireshark* program was used for this purpose. The program screen is presented in Fig. 4.

Fig. 4. Wireshark program screen

The program captures all of the data packets sent over a selected connection. It records the following information at the moment of package being sent: the sender and recipient's addresses, type of communication protocol, number of the communication port and the transmitted data. The program is equipped with a set of tools for analysing the captured data. Thanks to the fact that the communication between the robot controller and *Test Signal Viewer* is not encrypted, it was possible to identify the principles of the communication between them. It was enough to change the values of the data in the program and observe in what form this change is reflected in the information sent to the controller, and the response. It let us establish that each measuring channel is configured by sending the data frame presented in Table 1.

Table 1. Configuration data frame

Information	Starting word: number 1	Channel number	Test signal number	Name of task controlling the mechanical unit	Number of mechanical unit axis	Sampling time
Data type	int 32bits	int 32bits	int 32bits	40 characters in ASCII code	int 32bits	Single

Such a frame must be sent separately to every measuring channel used. If the configuration data are incorrect, the controller sends an error message back in the form of a character string. When all of the channels are set, data collection from the controller can commence. Data collection is initiated with sending of the data frame presented in Table 2.

Table 2. Data frame initiating the measurement

Information	Starting word: number 2	Number of channels	12 integers containing the active channel numbers
Data type	int 32bits	int 32bits	int 32bits

The last 12 integers tell the controller which channels are to be sent and in what order. After receiving this information, the controller starts sending the data to the computer in the form presented in Table 3.

Table 3. Measurement data frame

Information	Starting word: number 7	Number of samples	Channel number	Data package number	Data
Data type	int 32bits	int 32bits	int 32bits	int 32bits	Single

For each channel, from time to time, the controller sends buffered measurement data. At the beginning of each frame, there is information about the number of transmitted samples. This number depends on the sampling time and the time elapsed from sending the previous frame. Apart from this, each sent data frame is numbered sequentially in order to monitor whether there are no communication errors. The controller can simultaneously send up to several hundred samples. In order to stop the measurement, the controller must receive the frame presented in Table 4.

Table 4. Data frame stopping the measurement

Information	Starting word: number 3	Number of channels	12 integers containing the active channel numbers
Data type	int 32bits	int 32bits	int 32bits

The only difference from the initiating data frame is the starting word value. The obtained information facilitated the development of our own application that reads the values of test signals.

Having the knowledge about the information from the packages that can be obtained from the robot controller, the next step was to write the application. The *Matlab* environment was chosen because it provides the opportunity not only to monitor the mechanical unit's status but also to analyse the signals. In the future it will be used to create our own algorithms of force control. It also has a comprehensive library of calculation tools. The application performance scheme is presented in Fig. 5.
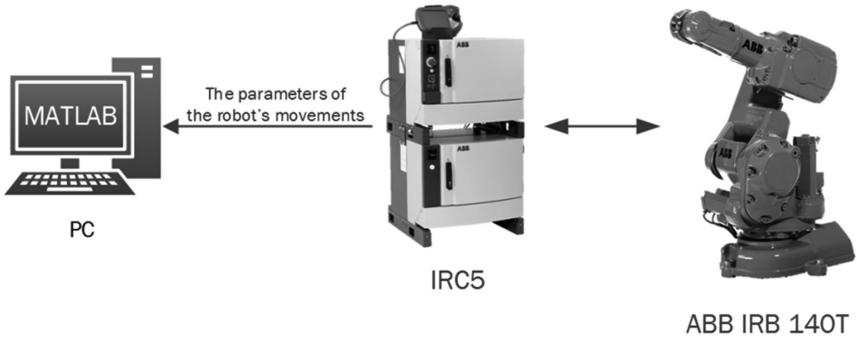
Fig. 5. The application performance scheme

The application's task is to collect data about the parameters of the robot's movements and store them as variables intended for further use. A script has been written in *Matlab* that allows for the reading of the robot's performance parameters using a network connection.

4   Verification of the Solution

The correctness of the obtained solution was first checked using a virtual controller, and then using an actual mechanical unit. First, it is required to define the network parameters of the connection, that is, the robot's network address and port. The test signals are available on port number 4011. Next, the measurement time is provided, expressed in seconds. In order to trigger the selected measuring channels, a value of 1 must be assigned to variable *channel_(channel_number)_Activ*. A value of 0 means that the channel should remain inactive. Each active channel requires the provision of
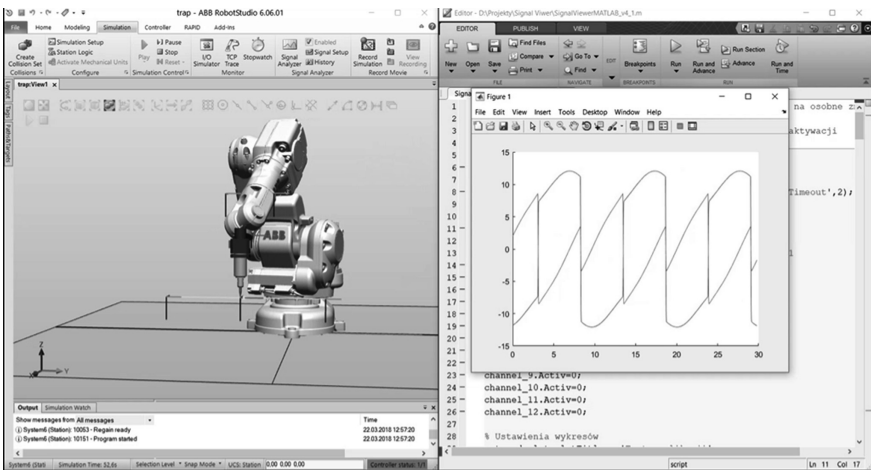

Fig. 6. Application performance test with virtual controller in RobotStudio

information about the test signal number and the name of the mechanical unit controlling task. Moreover, if the measurement is to concern, for example, angle parameters of a robot axis movement, the axis number and the sampling time must be defined. After running the script, the configuration data are transmitted to the controller. If there are no errors, the measurement is initiated. As an example of application performance, we carried out tests which involved making connection with a virtual controller running in the *RobotStudio* program (Fig. 6).

A station with an IRB 140 robot was constructed in *RobotStudio* software and programmed to make a linear movement between two points. The application read test signals carrying information about the angular velocity of the manipulator's first and second axes. The received data were registered and presented in the diagram (Fig. 7).
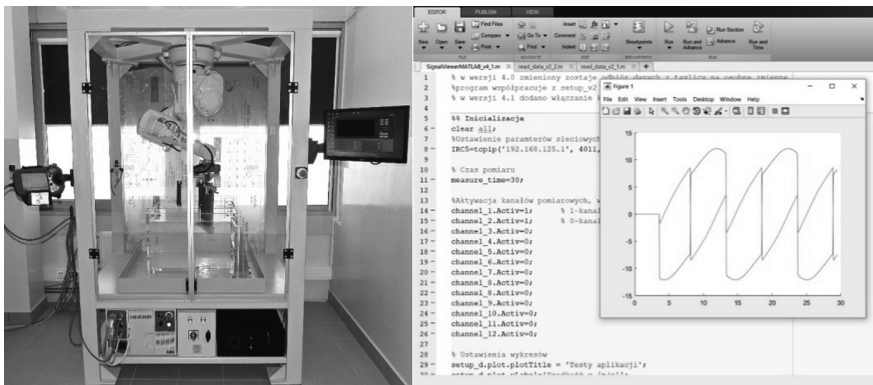

Fig. 7. Application performance test with actual controller

In the next step, acting similarly to the connection with the virtual controller, that is, setting the network parameters, signal numbers and time, a connection was established with an actual controller. An IRB 140 robot was programmed to make a linear movement between two points. The application read the test signals carrying information about the angular velocity of the manipulator's first and second axes. The received data were registered and presented in the diagram.

To sum up, the test conducted in both the virtual environment and using an actual robot confirmed the correctness of the applied solution. The developed software allows for monitoring the defined parameters of robotic station performance using the *Matlab* software.

5   Summary

Monitoring the parameters of robot performance is a very important issue, especially at the time of the fourth industrial revolution. The idea of Industry 4.0 considers the possibility of supervising the work of all devices taking part in the production process. This work presents commercial software solutions used in monitoring the parameters of

ABB robot performance. The shortcomings and limitations of the discussed programs led to the development of proprietary software. The study used the method of capturing the data transmitted between the controller and the *Test Signal Analyzer* program. The identified manner of communication was used in further works involving the development of dedicated software. In order to check the correctness of the application written in the *Matlab* environment, a numerical experiment was carried out, which involved connecting to a virtual controller and an experiment involving connecting to a real controller.

The performed tests confirmed the correctness of the developed application, which connects with the robot controller and registers the defined data. Test signals in older versions of the controller operational system (*RobotWare*) allowed for reading the parameters of the system's external axes movement and signals from the force sensor. In the most current version of *RobotWare 6.06*, the quantity of test signals increased. The added signals facilitate reading the parameters of the robot's movement. It can be presumed that ABB is going to develop the test signal capabilities and, thus, the methods of their reading will gain popularity.

References

1. Kost, G., Reclik D.: Aplikacja systemu InTouch do monitorowania stanu zrobotyzowanego gniazda produkcyjnego. Instytut Automatyzacji Procesów Technologicznych i Zintegrowanych Systemów Wytwarzania Politechniki Śląskiej w Gliwicach, 59–68, Gliwice (2009)
2. Hui-Ping, Liu, Dai-Min, C., Miao, Y.: Communication of multi-robot system on the TCP/IP In: mechatronic science, electric engineering and computer (MEC). In: International Conference on. IEEE, 1432–1435 (2011)
3. Jiang, Y.F.: The application of the TCP/IP on the robot system. Appl. Mech. Mater. **159**, 351–354 (2012)
4. Gierlak, P., Burghardt, A., Szybicki, D., Szuster, M., Muszyńska, M.: On-line manipulator tool condition monitoring based on vibration analysis. Mech. Syst. Signal Process. **89**, 14–26 (2017)
5. Burghardt, A., Szybicki, D., Kurc, K., Muszyńska, M., Mucha, J.: Experimental study of Inconel 718 surface treatment by edge robotic Deburring with force control. Strength Mater. **49**(4), 594–604 (2017)
6. Burghardt, A., Kurc, K., Szybicki, D., Muszyńska, M., Szczęch, T.: Monitoring the parameters of the robot-operated quality control process. Adv. Sci. Technol. Res. J. **11**(1), 232–236 (2017)
7. Tao, Z., Ueno, H.: Knowledge model-based heterogeneous multi-robot system implemented by a software platform. Knowl. Based Syst. **20**(3), 310–319 (2007)

Signal Reconstruction from Sparse Measurements Using Compressive Sensing Technique

Beata Palczynska[1,2](✉) ![ORCID]

[1] Gdansk University of Technology, 11/12 Gabriela Narutowicza Street,
80-233 Gdansk, Poland
`beata.palczynska@pg.edu.pl`
[2] Gdynia Maritime University, Morska 81-87, 80-225 Gdynia, Poland

Abstract. The paper presents the possibility of applying a new class of mathematical methods, known as Compressive Sensing (CS) for recovering the signal from a small set of measured samples. CS allows the faithful reconstruction of the original signal back from fewer random measurements by making use of some non-linear reconstruction techniques. Since of all these features, CS finds its applications especially in the areas where, sensing is time consuming or power constrained. An electromagnetic interference measurement is a field where the CS technique can be used. In this case, a sparse signal decomposition based on matching pursuit (MP) algorithm, which decomposes a signal into a linear expansion of element chirplet functions selected from a complete and redundant time-frequency dictionary is applied. The presented paper describes both the fundamentals of CS and how to implement MP for CS reconstruction in relation to non-stationary signals.

Keywords: Compressive sensing · Matching pursuit

1 Introduction

Compressive sensing (CS) is a novel research area, which has been widely developed in recent years [1–10]. Solid mathematical foundations, determining the conditions for reliable reconstruction of signals based on significantly fewer measurements (in relation to Shannon's requirements) and proposals for a wide range of solutions concerning mainly the measurement matrix and conditions of a reconstruction have opened up a wide range of new opportunities and challenges for many applications. Theory of CS is based on two main principles: signal sparsity and incoherent sensing. The term sparsity means that an examined signal may be modeled with a small number of components taken from a large dictionary. A signal can either have sparse or compressible representation in original domain or in some transform domains. Given a sparse signal, its information content can be recovered even from what could appear to be an incomplete set of measurements, at the expense of a greater computational effort at reconstruction stage.

The CS measurements are non-adaptive, i.e., not learning from previous measurements. The resulted fewer compressive measurements can be easily stored or

transmitted. CS allows the faithful reconstruction of the original signal back from fewer random measurements by making use of some non-linear reconstruction techniques. Since of all these features, CS finds its applications especially in the areas where, sensing is time consuming or power constrained [11].

An electromagnetic interference measurement is a field where the CS technique can be used. Many signals, e.g., encountered in an electromagnetic environment of a ship, are modeled as the chirplet-type functions [12]. In this case, a sparse signal decomposition based on matching pursuit (MP) algorithm, which decomposes a signal into a linear expansion of element chirplet functions selected from a complete and redundant time-frequency dictionary can be applied. The presented paper describes both the fundamentals of CS and how to implement MP for CS reconstruction in relation to non-stationary signals. The work is preliminary in nature; it explores the possibilities of using a novel technique in the measurement of electromagnetic interference, supported by exemplary results.

The paper is organized as follows. Section 2 contributes the compressive sensing background. In Sect. 3, a CS reconstruction approach based on MP procedure is presented. Section 4 shows the exemplary simulation results. Concluding remarks with directions for future work are drawn in Sect. 5.

2 Compressive Sensing Concept

CS is a process of measuring sparse signals and then reconstructing these signals with incomplete data (in comparison to classical measurement methods). CS - data acquisition process requires a relatively small amount of work. It is performed a small number of measurements. A reconstruction is a complex process, although its implementation is based on processing a small amount of data on the analyzed signal. The CS signal processing scheme contains both acquisition and reconstruction models (see Fig. 1).
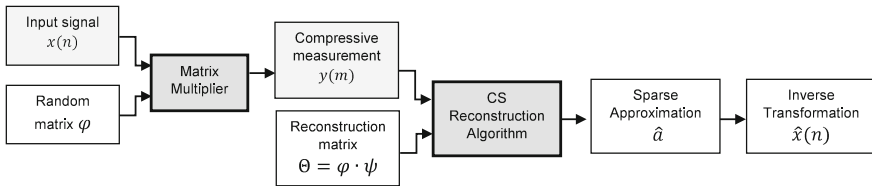
Fig. 1. A signal processing scheme based on CS [13].

Essentially, the CS theory involves three stages:

- determining the sparse decomposition of a signal,
- designing applicable compression representing matrix, which well approximates the original signal of length N for the least M coefficients,
- applying corresponding reconstruction algorithm, which can reconstruct original signal from observed M coefficients that are gained from step 2.

A sparsity of the signal and an incoherence during sensing are basis conditions for using CS.

2.1 Sparse Decomposition

Sparsity is an intrinsic property of the signal for which, all information contained in the signal can be represented only by means of several significant components. If signal $x \in R^N$ of length N can be transformed in sparse transform matrix ψ to K orthogonal vectors, where $K \ll N$, that x is K – sparse. The signal x can be represented as a linear combination of columns of ψ as [14]

$$x = \sum_{i=1}^{K} a_i \cdot \psi_i = \psi \cdot a \tag{1}$$

where, $a \in R^N$ is the sparse coefficients of length N, ψ_i are the basis vectors in sparse matrix ψ.

2.2 Measurement

CS works by taking fewer random measurements that are non-adaptive. The CS acquisition model can be defined mathematically by [1]

$$y = \varphi \cdot x \tag{2}$$

where, $\varphi \in R^{MxN}$ is a random measurement matrix and $y \in R^M$ is the measurement vector of length M.

The number of measurements taken are much lesser than the length of input signal, i.e. $M \ll N$. Taking into account Eq. (1) the measurement process is described by following expression [1]

$$y = \varphi \cdot \psi \cdot a = \Theta \cdot a \tag{3}$$

where, $\Theta \in R^{MxN}$ is a reconstruction matrix.

The process of signal acquisition depends on two bases, the transform matrix ψ (related to sparsity of signal) and the random measurement matrix φ (used in measurement process). The size of measurement matrix and hence the number of measurements is proportional to the sparsity of input signal. To further reduce the number of measurements which are necessary for proper reconstruction, the measurement matrix must be incoherent with basis in which signal has sparse representation. The relation for finding the coherence between two matrices is given in [2]

$$\mu(\varphi, \psi) = \sqrt{N} \cdot \max_{0 \le i,j \le n} \left| \varphi_i, \psi_j \right| \tag{4}$$

The coherence should be as small as possible. The CS theory states that there exists an overmeasuring factor $c > 1$ such that only $M = cK$ incoherent measurements y are required to reconstruct x with high probability [1, 2].

2.3 Reconstruction Model

The inputs to the reconstruction algorithm are the measurement vector y and reconstruction matrix Θ (see Fig. 1). The original signal can be reconstructed from compressive measurements by solving the inverse problem of formula (2), which gives infinite number of possible solutions. In such cases, signal recovery can be regarded as ℓ_1- norm optimizing solution, as shown in (5) [4, 5].

$$\hat{a} = \frac{\arg\min}{a} \| a \|_1 \ subject\ to\ y = \Theta \cdot a \tag{5}$$

where, $\hat{a}$ is the estimate of a and $\| a \|_1$ denotes the ℓ_1 – norm of a and

$$\ell_1 : \| a \|_1 = \sum_i |x_i| \tag{6}$$

CS requires less samples for reconstruction than the conventional sampling, so the required dimension of the measurement vector y is estimated and is found to be [5]

$$M \geq c \cdot \mu(\varphi, \psi) \cdot K \cdot \log N \tag{7}$$

The integer M in (7) indicates the number of samples such a way that it is very likely that a K – sparse signal can be successfully reconstructed from y. Besides the number of the incoherent measurements, another factor that affects the reconstruction is the measurement matrix φ, which must satisfy the restricted isometry property (RIP) [11]. The RIP is used to ensure that the pair of bases with the number less than or equal to K approximately behaves like an orthonormal system [15].

3 Matching Pursuit Reconstruction Approach

The optimization problem expressed in (5) can be solved using nonlinear greedy algorithm – matching pursuit (MP). The MP iterative algorithm decomposes a signal y into a linear expansion of element functions $g_{\gamma n}$ selected from a complete and redundant time-frequency dictionary Θ. It is approximately given by the following equation [16]

$$y = \sum_{n-0}^{M-1} a_n \cdot g_{\gamma n} \tag{8}$$

where, a_n is a sparse coefficient of the signal in time frequency domain.

The dictionary Θ is the reconstruction matrix in (3), expressed as follows [14]

$$\Theta = \begin{bmatrix} G_{11} & \cdots & G_{1N} \\ \vdots & \ddots & \vdots \\ G_{M1} & \cdots & G_{MN} \end{bmatrix} \tag{9}$$

and

$$g_{\gamma i} = [G_{1i}, \ldots, G_{Mi}]^T \qquad (10)$$

A dictionary as a set of the time-frequency atoms $g_{\gamma n}$ is generated by translating (u), scaling ($s > 0$) and modulating (ξ) a single elementary function, as shown in (11) [17].

$$g_{\gamma}(t) = \frac{1}{\sqrt{s}} \cdot g\left(\frac{t-u}{s}\right) \cdot e^{j\xi t} \qquad (11)$$

The aim of the reconstruction algorithm is to solve the K non-zero coefficients of a which subjects to (3), on condition that a is a K – sparse representation of the signal. In each iteration, the solution is updated by selecting only those columns of reconstruction matrix Θ, which are highly correlated with the measurement signal y.

The first step ($i = 0$) of the MP procedure consists in selecting an elementary function $g_{\gamma 0}$ from the dictionary so that the amplitude of the inner product $\left|\langle R^0 y, g_{\gamma 0} \rangle\right|$ between this function $g_{\gamma 0}$ and the signal y is the largest. Then y can be decompose into

$$y = \langle y, g_{\gamma 0} \rangle g_{\gamma 0} + R^1 y \qquad (12)$$

Subsequently, the residual signal $R^1 y$ obtained after extracting the approximation of y in the direction of $g_{\gamma 0}$ from y, is decomposed in a similar way. Iterative procedures are executed to the next residua, as follows [17]

$$\begin{cases} R^0 y = y \\ R^{i+1} y = R^i y - \langle R^i y, g_{\gamma i} \rangle g_{\gamma i} \end{cases} \qquad (13)$$

In each iterative step the atom is chosen such that [17]

$$g_{\gamma i} = argmax_{g_{\gamma i}} \left|\langle R^i y, g_{\gamma i} \rangle\right| \qquad (14)$$

After M - iterations the signal y can be expressed by a sum of residues.

$$y = \sum_{n-0}^{M-1} (R^n y - R^{n+1} y) + R^M y \qquad (15)$$

Considering the Dependence (13) Is Obtained

$$y = \sum_{n-0}^{M-1} \langle R^n y, g_{\gamma n} \rangle \cdot g_{\gamma n} + R^M y \qquad (16)$$

In this manner, the signal y is expressed as a sum of atoms that best fit its residues.

$$y = \sum_{n-0}^{M-1} a_n \cdot g_{\gamma n} + R^M y \qquad (17)$$

The procedure is implemented iteratively until $\|R^M y\|$ reaches a predefined threshold or $M > K$. Then the decomposition coefficients can be described by

$$\widehat{a}_n = \langle R^n y, g_{\gamma n} \rangle, \; n = 0, 1, \ldots, M - 1 \tag{18}$$

4   Simulation Results

This section presents the results of simulating experiments carried out using LabVIEW programming environment. The LabVIEW Project, in which the CS algorithm is implemented, consists of three main parts, responsible for a signal generation, a compressive sampling acquisition and a signal reconstruction.

Figure 2 shows the time-waveform of the tested signal x, which is composed of four chirps and Gaussian noise components, for which the signal to noise ratio is equal to 10 dB. In the CS acquisition block, a set of single random (scalar) measurements was performed, where each measurement represents a random projection of the signal onto a single scalar value. The random measurement matrix φ is created based on Bernoulli distribution. The linear Gaussian chirplets were chosen as time-frequency of atoms in the dictionary Θ that used in the MP reconstruction algorithm. Figure 3 shows the results of reconstructions of the tested signal for different number of Bernoulli measurements for the same of the sparsity level K equals 9.

An accuracy of a reconstruction can be defined as percentage of tested signal's variance (energy) explained by the CS reconstruction. In the case of a small number of measurements, the obtained representation of the signal represents about 35% of the total signal's energy (see Fig. 3a). Increasing the number of measurements allows for detection of broadband components, related to the presence of noise in the signal (see Fig. 3b). However, up to 50 atoms explain only 60% of the total energy. The reconstruction algorithm using of 100 measurements in this case represents over 75% of the
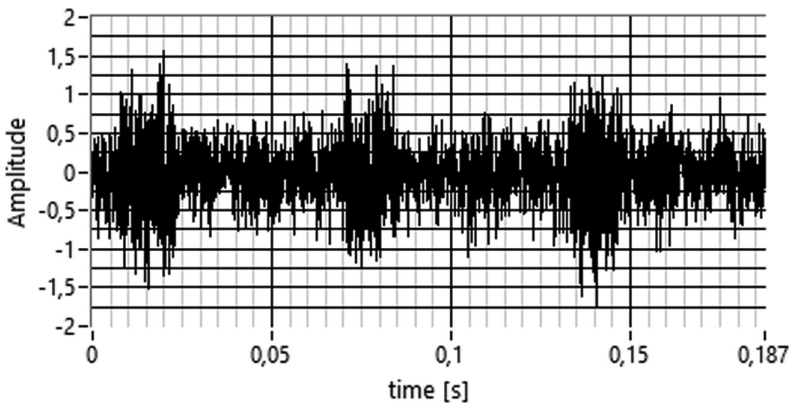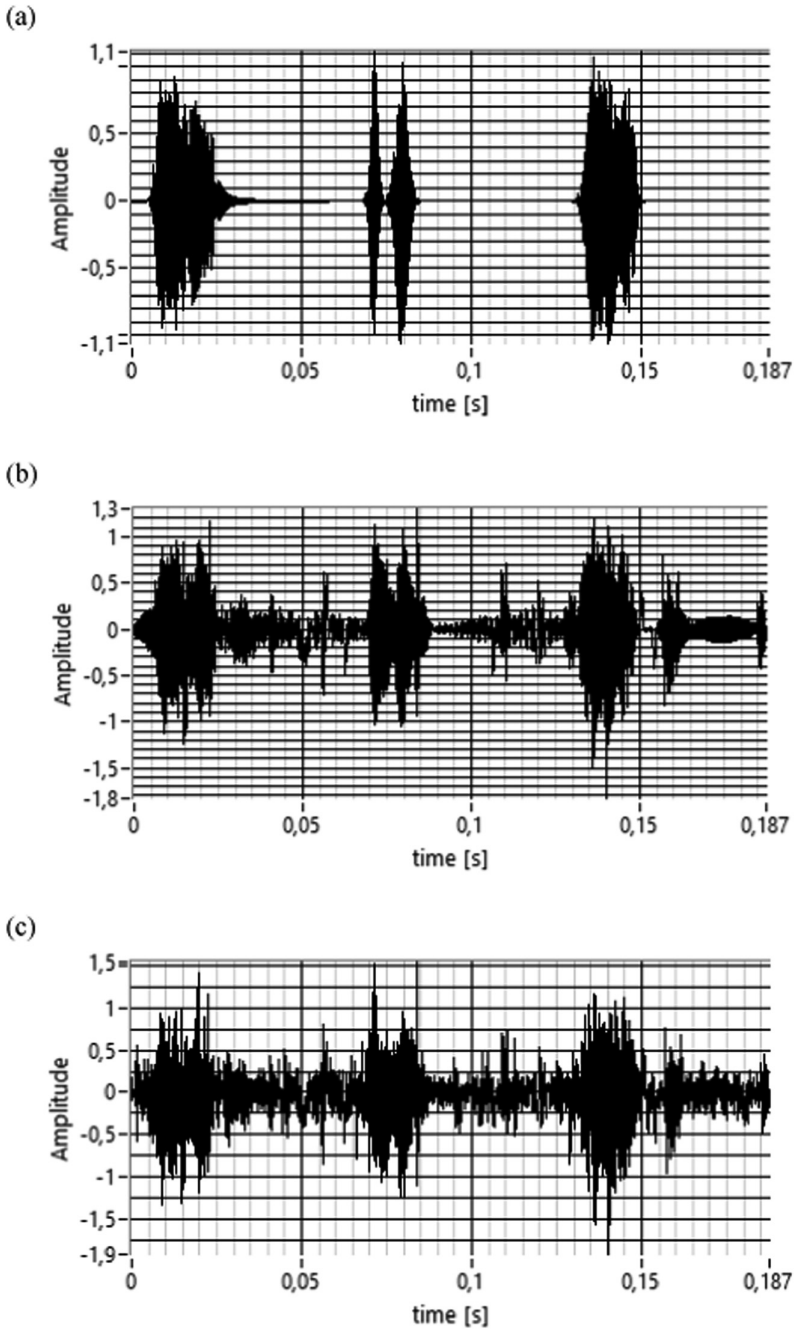

Fig. 2. The waveform of original signal.

Fig. 3. The waveforms of reconstructed signal for number of measurements equal 10 (a), 50 (b) and 100 (c) respectively.

energy of the original signal. The next increasing the number of measurements does not cause a noticeable increase in energy of the reconstructed signal (see Fig. 3c).

The simulation results show that it is possible to provide a proper reconstruction of the signal with a few incoherent measurements when the signal of interest is sparse in some dictionary. However, in the future, other cases should be considered, e.g. related to different types of test signals.

5   Conclusion

This paper reviews the CS concept and describes the implementation of CS signal reconstruction by the MP algorithm. The MP provides comparable and accurate results in reconstructing the noiseless input signal. The noisy signal reconstruction case is under consideration.

CS allows the accurate reconstruction of the original signal back from fewer random measurements by making use of some non-linear reconstruction techniques. Because of all these features, CS finds its applications especially in the areas where, taking measurements is too expensive or time consuming, e.g., medical imaging, electromagnetic measurements and where, sensing is power constrained, etc. There are many opportunities for future research.

References

1. Donoho, D.L.: Compressed sensing. IEEE Trans. Inf. Theory **52**, 1289–1306 (2006)
2. Baraniuk, R.G.: Compressive sensing - lecture notes. IEEE Signal Process. Mag. **24**(4), 118–121 (2007)
3. Candes, E.J., Wakin, M.B.: An introduction to compressive sampling: a sensing/sampling paradigm that goes against the common knowledge in data acquisition. IEEE Signal Process. Mag. **25**, 21–30 (2008)
4. Candès, E.J., Romberg, J., Tao, T.: Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. IEEE Trans. Inf. Theory **52**, 489–509 (2006)
5. Candès, E., Tao, T.: Near optimal signal recovery from random projections: universal encoding strategies. IEEE Trans. Inf. Theory **52**, 5406–5425 (2006)
6. Candes, E., Wakin, M.: An introduction to compressive sampling. IEEE Signal Process. Mag. **25**, 21–30 (2008)
7. Sankaranarayanan, A.C., Turaga, P.K., Chellappa, R., Baraniuk, R.G.: Compressive acquisition of linear dynamical systems. SIAM J. Imaging Sci. **6**, 2109–2133 (2013)
8. Dai, W., Milenkovic, O.: Subspace pursuit for compressive sensing signal reconstruction. IEEE Trans. Inf. Theory **55**(5), 2230–2249 (2009)
9. Sejdića, E., Orovićb, I., Stanković, S.: Compressive sensing meets time–frequency: an overview of recent advances in time–frequency processing of sparse signals. Digit. Signal Proc. **77**, 22–35 (2018)
10. Baraniuk, R.: An Introduction to compressive sensing. http://legacy.cnx.org/content/col11133/1.5/. Accessed 21 May 2018
11. Candès, E.J.: The restricted isometry property and its implications for compressed sensing. Comptes Rendus Math. **346**(9), 589–592 (2008)

12. Yoo, J., Becker, S., Monge, M., Loh, M., Candès, E.J., Emami-Neyestanak, A.: Design and implementation of a fully integrated compressed-sensing signal acquisition system. In: Proceedings of 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5325–5328. IEEE, Kyoto (2012)
13. Rani, M., Dhok, S.B., Deshmukh, R.B.: A systematic review of compressive sensing: concepts, implementations and applications. IEEE Access **6**, 4875–4894 (2018)
14. Gribonval, R., Nielsen, M.: Sparse representations in unions of bases. IEEE Trans. Inf. Theory **49**(12), 3320–3325 (2003)
15. Wang, Z., Lee, I.: Sorted random matrix for orthogonal matching pursuit. In: International Conference on Digital Image Computing: Techniques and Applications, pp. 116–120. NSW, Sydney (2010)
16. Mallat, S.: A Wavelet Tour of Signal Processing: The Sparse Way, 3rd edn. Elsevier, Burlington (2009)
17. Palczynska B.: Identification of the time-vary magnetic field sources based on matching pursuit method. Energies **10**(5), article no 655 (2017)

Power Grid Frequency Estimation Based on Zero Crossing Technique Using Least Squares Method to Approximate Sampled Voltage Signal Around Zero Level

Eligiusz Pawłowski[(✉)]

Faculty of Electrical Engineering and Computer Science, Lublin University of Technology, Nadbystrzycka Str. 38A, 20-618 Lublin, Poland
e.pawlowski@pollub.pl

Abstract. The article presents the application of the least squares method to the estimation of voltage transition points through the zero level in order to determine the frequency of the power grid system. A linear approximation of the quantized voltage signal samples near zero has been applied, which ensures effective noise suppression and reduces measurement errors. An even sampling of the sinusoidal voltage disturbed by noise was assumed, and on this basis the corresponding mathematical relations were derived. The presented algorithm for processing signal samples is computationally simple and can be easily implemented in a microprocessor system. The dependence of the accuracy of the signal frequency measurement on: SNR signal ratio, ADC converter resolution, signal sampling rate and the number of samples used for approximation were investigated. The simulation tests of the presented method were carried out and the results were presented, which enable proper design of the measurement system (ADC converter resolution, sampling rate, number of points to approximate) depending on the expected signal SNR ratio and the expected accuracy of measurements. In order to verify the presented method in practice, a measuring system was implemented using the National Instruments NI USB 6009 Data Acquisition Card and a personal computer. The measurement algorithm was implemented in the LabVIEW environment. The structure of the program was presented and the method of implementation of the most important parts of the algorithm was discussed, as well as examples of measurement results. The developed method can be used to build an independent measuring instrument or it can be an element of a larger measuring system.

Keywords: Frequency measurement · Power grid frequency · Zero crossing technique · Least squares method · Linear approximation · Sampled signal

1 Introduction

Frequency is one of the most important voltage parameters in the power system and is measured for a number of reasons, such as: supervision and regulation of the generator unit work, power flow control in the system, energy quality assessment, overload

protection, audio recording authentication and many other. Each of these applications requires measurement of frequencies in accordance with the requirements in terms of measurement accuracy, measuring range, averaging time, measurement speed and possible recording of results. To assess the quality of electricity in accordance with IEC Standard 61000-40-30 [1], the frequency of the power grid should be measured at intervals of 10 s. For this purpose, it is necessary to measure the duration of the total number of mains voltage periods that occurred in each subsequent time interval equal to 10 s. The uncertainty of frequency measurements should not exceed ± 0.01 Hz for the A-measurement class and ± 0.05 Hz for the S-measurement class, in the 42.5 Hz to 57.5 Hz measuring range. Completely different requirements for example are found in the application for authentication of audio recordings, where the frequency measurement of the power network with the accuracy of ± 0.0005 Hz is applied, measurements are taken and recorded at a rate of 10 measurements per second [2]. Due to such a variety of applications and requirements, many methods of frequency measurement have been developed. It should be noted that due to large errors at small frequencies, the classic direct counting method [3] is not used to measure the power grid frequency. At present, the methods that dominate in these applications use digital processing of instantaneous voltage of the power network.

A wide variety of data processing algorithms is used. For this purpose we use, among others: convolution of the signal with the Walsh function [4], correlation of the signal with sine/cosine functions [5], decomposition of the signal into two orthogonal components [6], demodulation of orthogonal signals [7], calculation of the second derivative of the sine function [8, 9], optimization of the solution of the second order differential equation [8], effect of signal spectrum leak [10], FFT transformation [10, 11], derivative of the phase of the rotating voltage phasor [11], development of a sinusoidal function into the Taylor series [12] and many others. A large group of current used methods performs frequency measurement based on voltage transition points through zero levels [13–21].

All these methods can be divided in terms of the number of samples used. Some methods use all signal samples for one or several periods of voltage [4–12], other methods use only a few samples covering part of the period in any of its fragments. The simplest methods in this group use only three samples [13], another 4 samples [14, 15] and more [16]. Another group of methods uses only samples in the zero crossing area, with the zero crossing being determined by the least of two [17] or more samples [16]. The zero crossing can also be determined by a hardware method [18] with digital measurement of the signal period [19, 20]. Many of the above methods also use the method of least squares to approximate samples of instantaneous values of the signal [16], to approximate the instantaneous phase of the signal [5–9, 19] or to approximate the spectrum of the signal [10–12].

The work presents a frequency estimation algorithm using the least squares method for linear approximation of a voltage signal to determine zero crossing points. Attention was focused on the possibility of implementing the method in a microprocessor system, which requires the development of a computationally simple algorithm, with the lowest possible sampling rate and a small number of signal samples included in the approximation.

2   Hardware Realization of the Zero Crossing Method

2.1   Basics of the Zero Crossing Method

Measuring the signal frequency by the zero crossing method is indirect measurement. This measurement consists in calculating the frequency as the inverse of the period of the signal, wherein the period of the signal is measured by the direct method as the time between two consecutive instances of the signal's passing through the zero level with the same derivative sign. The principle of measurement is presented in Fig. 1. We assume the model of the measured signal $v_s(t)$ in the form of a pure sine wave:

$$v_s(t) = V_{max} \sin(2\pi f_s t + \varphi_0), \tag{1}$$

where V_{max} is the amplitude of the signal, f_s is the measured frequency and φ_0 is the initial signal phase.
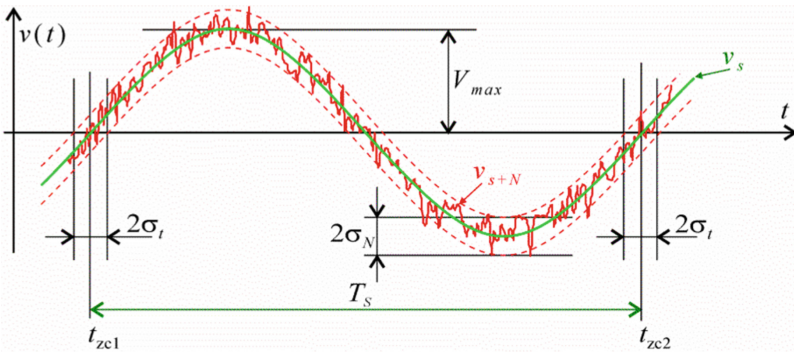

Fig. 1.   Measurement of the signal period in the presence of disturbances.

The signal (1) has a period T_s determined by two consecutive instances t_{zc1}, t_{zc2}, in which this signal exceeds the zero level with a positive derivative sign (Fig. 1). The period $T_s = t_{zc2} - t_{zc1}$ can be measured with the classical hardware digital method in the arrangement shown in Fig. 2.
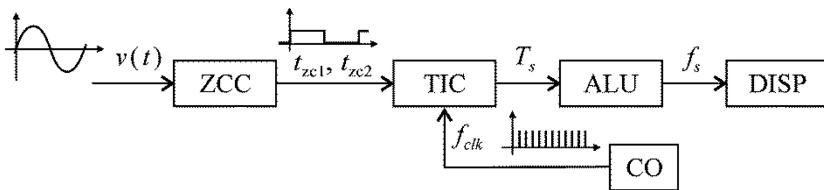

Fig. 2.   Block diagram of the system for frequency measurement by hardware method.

The zero-crossing circuit ZCC includes the Schmitt trigger, which sets the t_{zc1}, t_{zc2} instances. The period T_s is measured in the time interval counter TIC by the digital method by filling it with N_s impulses of the reference frequency f_{clk} from the clock oscillator CO. The f_s frequency of the signal $v_s(t)$ is calculated in the arithmetic logic unit ALU according to the relation (2) and presented on the display DISP.

$$f_s = 1/T_s = f_{clk}/N_s \tag{2}$$

2.2   Accuracy Analysis

Accuracy of the measurement by hardware method is limited by three errors: quantization error arising during the counting operation of the pulses in the TIC counter, error of the clock frequency f_{clk} and the error of determining the instances t_{zc1}, t_{zc2} of the signal passing through zero. The first two of these errors can be effectively limited to small values by known methods [3, 21]. The biggest problem in the classical hardware method is the error of determining the t_{zc1}, t_{zc2} instances arising in the ZCC input circuit, resulting from the occurrence of interference adding up to the measured signal [3]. Partially, this error can be limited by using special constructions of the ZCC input system, including the Schmitt trigger with the appropriate selected hysteresis, amplifiers with automatic gain control, galvanic isolation systems, low-pass filters, etc. [3, 18].

Accuracy of the Zero Crossing Points Determination. However, for a signal disturbed by noise, accurate determination of t_{zc1}, t_{zc2} is impossible, and the measurement result is always subject to an additional error, called a trigger error [3]. For the purpose of further analysis, it was assumed that the frequency of the signal $v_{s+N}(t)$ is measured, which is the sum of the sine wave signal and noise:

$$v_{s+N}(t) = v_s(t) + N(0, \sigma_N) = V_{max} \sin(2\pi f_s + \varphi_0) + N(0, \sigma_N), \tag{3}$$

where $N(0, \sigma_N)$ is a Gaussian noise with a normal distribution with a zero expected value and a standard deviation σ_N. Based on Fig. 1 it can be stated that instances t_{zc1}, t_{zc2} are random variables with a standard deviation σ_t [21]:

$$\sigma_t = \sigma_n/S, \tag{4}$$

where S is the rate of the increment of the signal $v_s(t)$ at the point of its passage through the zero level:

$$S = \left.\frac{dv_s}{dt}\right|_{t=t_{zc1},t_{zc1}} = 2\pi f_s V_{max} \cos(2\pi f_s t + \varphi_0)|_{t=t_{zc1},t_{zc1}} = 2\pi f_s V_{max}. \tag{5}$$

The level of interference in the measured signal $v_s(t)$ is conveniently expressed by the signal to noise ratio $SNR = V_{RMS}/\sigma_N = V_{max}/\sqrt{2}\sigma_N$, where V_{RMS} is the root-mean-square value of the signal $v_s(t)$. In practice, the signal-to-noise ratio is usually given on

a logarithmic scale $SNR_{dB} = 20\log(SNR)[dB]$, however, for the convenience of transformations, the linear scale will be used later. By substituting relation (5) to (4), after simple transformations we get:

$$\sigma_t = 1/\left(2\sqrt{2}\pi f_s SNR\right). \tag{6}$$

Accuracy of the Period Measurement. Because the period T_s is the difference between two random variables $t_{zc2} - t_{zc1}$, the standard deviation σ_T of the period T_s results from the uncertainty of determining its beginning t_{zc1} and the end t_{zc2}. Taking into account the dependence on the variance of the sum of two random variables, we get:

$$\sigma_{Ts}^2 = 2\sigma_t^2(1 - r), \tag{7}$$

where r is the correlation coefficient between t_{zc1} and t_{zc2}. For a disturbance in the form of white noise with unlimited bandwidth and for noise with a band limited to $f \geq 10\, f_s$, we can in practice assume $r \approx 0$ [21] and the relative standard deviation of the period measurement $\sigma_{rel\, Ts}$ will be equal to:

$$\sigma_{relTs} = \sigma_{Ts}/T_s = 1/(2\pi SNR). \tag{8}$$

Possibility of the Error Reduction by the Averaging. Based on the obtained dependence (8) it can be concluded that for sinusoidal signals disturbed by noise (3) the trigger error does not depend on the measured frequency f_s, but only on the SNR. In literature, it is often stated, for example, that for $SNR = 100$ ($SNR_{dB} = 40$ dB) this error will be 0.16% [3, 21]. This error can be reduced by measuring the average period. By extending the measurement, on the basis of n consecutive instances $t_{zc1}, t_{zc2}, \ldots, t_{zcn}$, the passing of the signal $v_s(t)$ through the zero level, one can determine the average value of the period $T_{AVG} = (t_{zcn} - t_{zc1})/(n - 1)$. The obtained measurement result will be burdened with relative standard deviation $\sigma_{rel\, TAVG}$ with the value $(n - 1)$ times smaller than $\sigma_{rel\, Ts}$ [3]:

$$\sigma_{relT_{AVG}} = \sigma_T/(n - 1)T_s = 1/(2(n - 1)\pi SNR). \tag{9}$$

Based on n consecutive instances $t_{zc1}, t_{zc2}, \ldots, t_{zcn}$ the phase of the least squares signal [19] can also be approximated, which gives better results than normal averaging. The error of frequency measurement by the zero crossing method can also be reduced by determining the instances t_{zc1}, t_{zc2} by digital processing of the quantized signal samples $v_s(t)$. In this work the method of least squares has been used for this purpose.

3 Software Realization of the Zero Crossing Method

3.1 Measuring System for Software Realization of the Zero Crossing Method

The application of the least squares method to determine the instances of the signal passing through the zero level requires sampling and processing of the signal into digital form at strictly defined moments spaced evenly over time. The simplified block diagram of the respective system is shown in Fig. 3. The signal conditioning systems that are necessary in the actual measurement system are omitted, but they are not important during the analysis of the metrological properties of the presented method. A sinusoidal signal $v(t)$ disturbed by the noise described by the dependence (3) is given to the ADC converter input. Signal frequency pulses f_{clk} from the CO generator determine the instances $t_1, t_2, \ldots, t_k$ in which the ADC converter takes k samples $v_1, v_2, \ldots, v_k$ of the input signal. Subsequent samples $v_1, v_2, \ldots, v_k$ are equidistant over time by one period of the clock signal $T_{clk} = 1/f_{clk}$. The number of samples k must be chosen such that the samples $v_1, v_2, \ldots, v_k$ cover at least one complete period of the signal T_s.
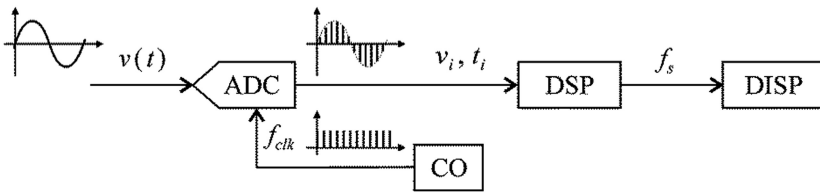
$$v_i = v_{s+N}(t_i), t_i = iT_{clk} \tag{10}$$

Fig. 3. Block diagram of the system for measurement by least squares method.

Based on the sequence of values of samples $v_1, v_2, \ldots, v_k$ and temporal moments $t_1, t_2, \ldots, t_k$ in the digital signal processing system DSP, zero crossing points t_{zc1}, t_{zc2}, are determined, and finally the signal frequency f_s (11) is calculated.

$$f_s = \frac{1}{T_s} = \frac{1}{t_{zc2} - t_{zc1}} \tag{11}$$

3.2 Algorithm for Determining the Zero Crossing Point by the Least Squares Method

The method of determining the zero crossing point by the least squares method is explained in Fig. 4. It presents an example of a signal fragment $v_s(t)$ in the ± 1 ms time interval around the zero crossing point. The signal frequency $f_s = 50$ Hz and the amplitude $V_{max} = 1$ V were assumed. The samples v_i show the signal v_{s+N} disturbed by noise with the $SNR_{dB} = 20$ dB, the clock frequency $f_{clk} = 5$ kHz, the instances t_i thus are equidistant by the period $T_{clk} = 0.2$ ms (10). The method of least squares determines

the straight line $y = ax + b$, approximating six samples of signal $v_1, v_2, \ldots, v_6$ in instances $t_1, t_2, \ldots, t_6$. Six samples $v_1, v_2, \ldots, v_6$ were selected from among all k samples in such a way that the two middle samples v_3 and v_4 differed in the sign, making them near the zero crossing point. The moment of passing through zero t_{zc} is determined based on parameters a, b of the approximating line $y = ax + b$:

$$t_{zc} = -\frac{b}{a}. \tag{12}$$

The directional coefficient a and the free expression b of the straight line $y = ax + b$ approximating the set of measurement data containing n pairs of numbers x_i, y_i can be determined by the method of least squares according to known dependencies [21]:

$$a = \frac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}, \tag{13}$$

$$b = \bar{y} - a\bar{x} = \frac{1}{n}\sum_{i=1}^{n} y_i - a\frac{1}{n}\sum_{i=1}^{n} x_i. \tag{14}$$
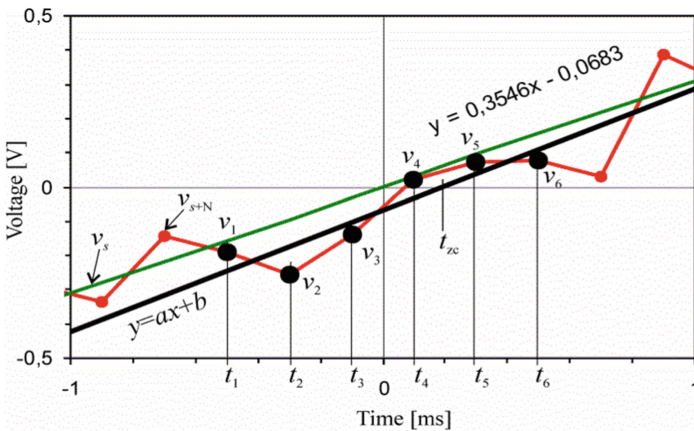
Fig. 4. The principle of determining the zero crossing point by the least squares method.

In the case under consideration, the values of y_i are equal to the following values of the signal samples $v_1, v_2, \ldots, v_n$, and the values x_i are equal to the next instances $t_1, t_2, \ldots, t_n$, in which these samples were taken. The signal is sampled uniformly over time, i.e. that the sampling times are equidistant from one another by one period T_{clk} of the clock signal (Fig. 3):

$$x_i = t_i = iT_{clk}, y_i = v_i. \tag{15}$$

The calculation of the directional coefficient a can be greatly simplified taking into account the known patterns for sums of finite numeric series:

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}, \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}. \tag{16}$$

By substituting the dependences (15) and (16) in the formula (13), after appropriate transformations we get a much simpler form of the expression for the desired directional coefficient a of the approximation line:

$$a = \frac{1}{T_{clk}} \frac{6}{n(n+1)(n-1)} \sum_{i=1}^{n} (2i - n - 1)v_i = \frac{1}{c_n T_{clk}} \sum_{i=1}^{n} d_i v_i, \tag{17}$$

where:

$$c_n = \frac{n(n+1)(n-1)}{6}, d_i = 2i - n - 1 \tag{18}$$

The c_n and d_i (18) coefficients are integers and uniquely determined only by the number of measuring points n included in the calculations. Dependence (17) shows that in the presented algorithm the weight averaging of successive samples of the signal v_i is carried out in the instances t_i with the coefficients of d_i. An exemplary shape of the weighing function for $n = 6$ and $n = 7$ is shown in Fig. 5. The weighing functions used are sawtooth with zero average value. Table 1 presents the values of coefficients c_n and d_i for several examples of n values. It should be noted that for an odd number of samples n, the value of the central coefficient d_i for $i = (n + 1)/2$ is equal to zero, i.e. the middle sample is not used in calculations. For this reason, it is preferable to use even numbers of n samples. The free expression b of the straight line equation is calculated from the relationship (14), which has a relatively simple form and does not need to be transformed.
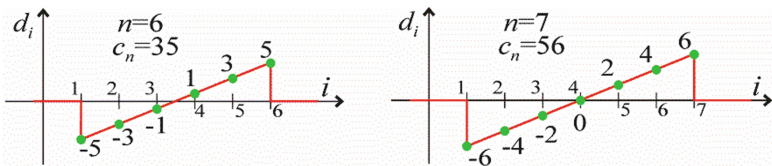

Fig. 5. Coefficients d_i of the sawtooth weight function for $n = 6$ and $n = 7$.

Table 1. Values of coefficients c_n, d_i for the least squares method.

n	c_n	d_i							
4	10	−3	−1	−1	−3				
5	20	−4	−2	0	2	4			
6	35	−5	−3	−1	1	3	5		
7	56	−6	−4	−2	0	2	4	6	
8	84	−7	−5	−3	−1	1	3	5	7
9	120	−8	−6	−4	−2	0	2	4	6	8

3.3   Assessment of Zero Crossing Method Accuracy

The uncertainty of determining the straight line coefficients a, b can be determined from known formulas [21] for estimated standard deviation

$$s_a = s_y \sqrt{\frac{n}{D}}, s_b = s_y \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{D}}, \tag{19}$$

where:

$$s_y = \sqrt{\frac{1}{n-2} \sum_{i=1}^{n} (y_i - b - ax_i)^2}, D = n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2 \tag{20}$$

Substituting (15), (16), (24) we get:

$$s_a = s_y \frac{1}{T_{clk}} \sqrt{\frac{12}{n(n^2-1)}}, s_b = s_y \sqrt{\frac{2(2n+1)}{n(n-1)}}, \tag{21}$$

Analyzing the form of dependence (21) it can be concluded that the uncertainty of determining the coefficients a, b decreases with increasing the number of n points used for approximation. In addition, the uncertainty of determining the coefficient a, decreases with increasing the period of time T_{clk} between samples. In conclusion, extending the approximation window width $W_a = nT_{clk}$ increases the accuracy of zero crossing points.

However, it is important to remember that you can approximate a sine wave with a straight line only near its zero crossing point. Figure 6 shows the linearity error of the sine wave δ_{lin} (22) with the frequency $f_s = 50$ Hz depending on the time relative to the zero crossing point. It should be noted that for an approximation window lying symmetrically around the zero crossing, with a width $W_a = 2$ ms (e.g.: $n = 5$, $f_{clk} = 2$ kHz) the sine wave linearity error already exceeds 1%. Therefore, extending the approximation window width W_a, a reasonable compromise should be maintained.

$$\delta_{lin} = \frac{x - \sin x}{\sin x} \cdot 100\% \tag{22}$$
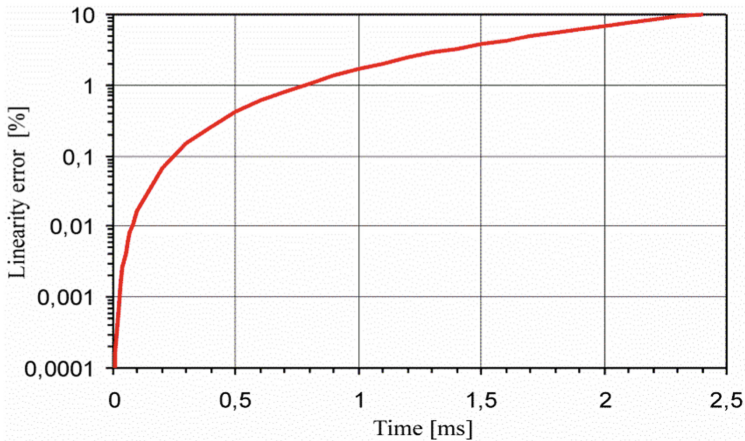
Fig. 6. Linearity error of 50 Hz sine wave as function of time after zero-crossing point.

The uncertainty of the zero crossing point t_{zc} can be determined using known dependencies on the propagation of uncertainty in indirect measurements. Having (12) we can write [21]:

$$s_{tzc}^2 = \left(\frac{\partial}{\partial a}\left(-\frac{b}{a}\right)s_a\right)^2 + \left(\frac{\partial}{\partial b}\left(-\frac{b}{a}\right)s_b\right)^2 + 2\frac{\partial}{\partial a}\left(-\frac{b}{a}\right)\frac{\partial}{\partial b}\left(-\frac{b}{a}\right)s_{ab}, \tag{23}$$

where s_{ab} is the estimated covariance [21]. After appropriate transformations we get:

$$s_{tzc}^2 = \left(\frac{b}{a^2}s_a\right)^2 + \left(\frac{1}{a}s_b\right)^2 - 2\frac{b}{a^3}s_{ab}. \tag{24}$$

3.4 Calculating the Frequency of the Signal

The principle of determining the signal frequency f_s is shown in Fig. 7. It shows the same signal as in Fig. 4, but in a longer time interval equal to 30 ms. The least squares zero crossing algorithm was applied twice and the coefficients for two straight lines $y = a_1 x + b_1$ and $y = a_2 x + b_2$ were calculated, thanks to which the position of two adjacent transitions through zero t_{zc1}, t_{zc2} determining the period of signal T_s which ultimately makes it possible to calculate signal frequency f_s:
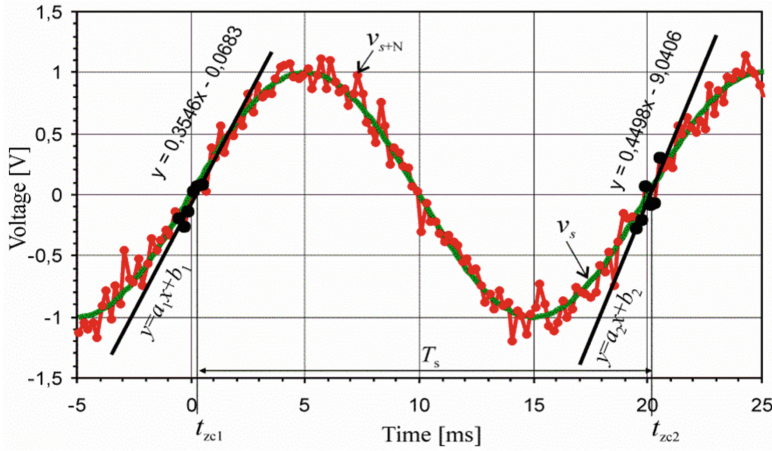
Fig. 7. The principle of frequency measurement using the zero crossing least squares method.

$$f_s = \frac{1}{T_s} = \frac{1}{t_{zc2} - t_{zc1}} = \frac{1}{-\frac{b_2}{a_2} + \frac{b_1}{a_1}}. \tag{25}$$

4   Simulation Investigation of the Method Properties

In order to verify the derived analytical relationships and confirm the properties of the presented method, appropriate simulation tests were carried out. Simulated a sinusoidal signal with amplitude $V_{max} = 1$ V and frequency $f_s = 50$ Hz with added noise level $SNR_{dB} = 20 \ldots 80$ dB. Signal samples were generated with the clock frequency $f_{clk} = 500$ Hz $\ldots$ 20 kHz and the frequency of the signal f_s was calculated based on the different number of samples $n = 4, 6, 8$. For each combination of SNR, f_{clk} and n parameters, a series of 256 sets of samples was performed, each set of samples had other random noise values. For each data series, 256 signal frequency values f_s were determined and ultimately the mean squared error was calculated as the estimated standard deviation s_f and relative mean squared error δ_f expressed in % (26).

$$s_f = \sqrt{\frac{1}{256} \sum_{i=1}^{256} (f_i - 50\,\text{Hz})^2}, \delta_f = \frac{s_f}{50\,\text{Hz}} \cdot 100\%, \tag{26}$$

The obtained simulation results are presented in the form of diagrams in Figs. 8, 9, 10, 11 and 12.

4.1 Influence of the Interference Level on the Accuracy

Figure 8 shows the relationship between the mean square error δ_f (26) and the level of interference for the SNR_{dB} ratio in the range of 2 … 80 dB, at different sampling frequencies f_{clk} = 1 kHz, 2 kHz, 5 kHz and the number of samples n = 4, 6, 8. All the graphs confirm the expected conclusion that for larger SNR ratios the δ_f errors decrease. However, for a different number of samples n than the f_{clk} sampling rate, this relationship changes. For a small SNR_{dB} = 20 … 30 dB errors are large and depend a little on f_{clk}. For signal to noise ratios above SNR_{dB} > 30 dB the errors strongly decrease for higher sampling rates (f_{clk} = 2 kHz, 5 kHz), and to a small extent they change for small f_{clk} = 1 kHz.

The graphs presented in Fig. 9 allow to better assess the effect of the number of samples n for a given sample rate f_{clk}. A larger number of samples (n = 8) provides smaller errors only for the high sampling rate f_{clk} = 5 kHz (Fig. 9c). For slower sampling f_{clk} = 1 kHz (Fig. 9a) smaller errors are obtained for a smaller number of samples n = 4, but for a small SNR_{dB} ratio in the range of 20 … 30 dB differences are small.

The graphs presented in Figs. 8, 9 confirm the conclusion that the method should be applied in a linear sinusoidal range near the zero crossing (Fig. 6), where the linearity errors (22) are small. Slower sampling f_{clk} = 1 kHz with a large number of samples n = 8 extends the approximation window width W_a = 7 ms beyond the linear part of the sinusoid (Fig. 6) and causes an increase in errors. However, for a strongly disturbed signal (SNR_{dB} = 20 … 30 dB) this problem becomes less important because the noise level σ_N in the signal reaches the level of sinusoidal linearity (22) and the width of the approximation window W_a is not as critical anymore.

### 4.2 Influence of the Sampling Rate and Number of Samples

Figure 10 shows the relationship between the mean square error δ_f (26) and the sampling rate for f_{clk} in the range of 500 Hz … 20 kHz, SNR_{dB} ratio in the range of 20 … 80 dB and the number of samples n = 4, 6, 8. In Fig. 10a it can be seen that for strong interference (SNR_{dB} = 20 dB) the errors do not depend to a small extent on the sampling rate, but generally a large number of samples n = 8 gives smaller errors. For smaller disturbances (Fig. 10b, c) and for low sampling rates f_{clk} = 500 Hz … 2 kHz lower errors are obtained for a small number of samples n = 4. However, for faster sampling f_{clk} = 2 kHz … 20 kHz it is more favorable to have a greater number of samples n = 8 (Fig. 10b).

The diagrams presented in Fig. 11 allow to better assess the effect of the SNR interference level for a given number of samples n. For strong interference SNR_{dB} = 20 dB increasing the sampling rate increases errors, confirming the effect of the T_{clk} sampling period on the uncertainty of the coefficient a of the approximating line (21).

However, for smaller disturbances (SNR_{dB} = 40, 60 dB), too slow sampling f_{clk} = 500 Hz… 2 kHz increases errors. Increasing the number of samples n reduces errors, which is particularly evident in severe disturbances (SNR_{dB} = 20 … 30 dB). This confirms the conclusions drawn from the dependence on the uncertainty of determining the coefficients a, b of the approximating straight line (21). However, an

(a)

(b)

(c)
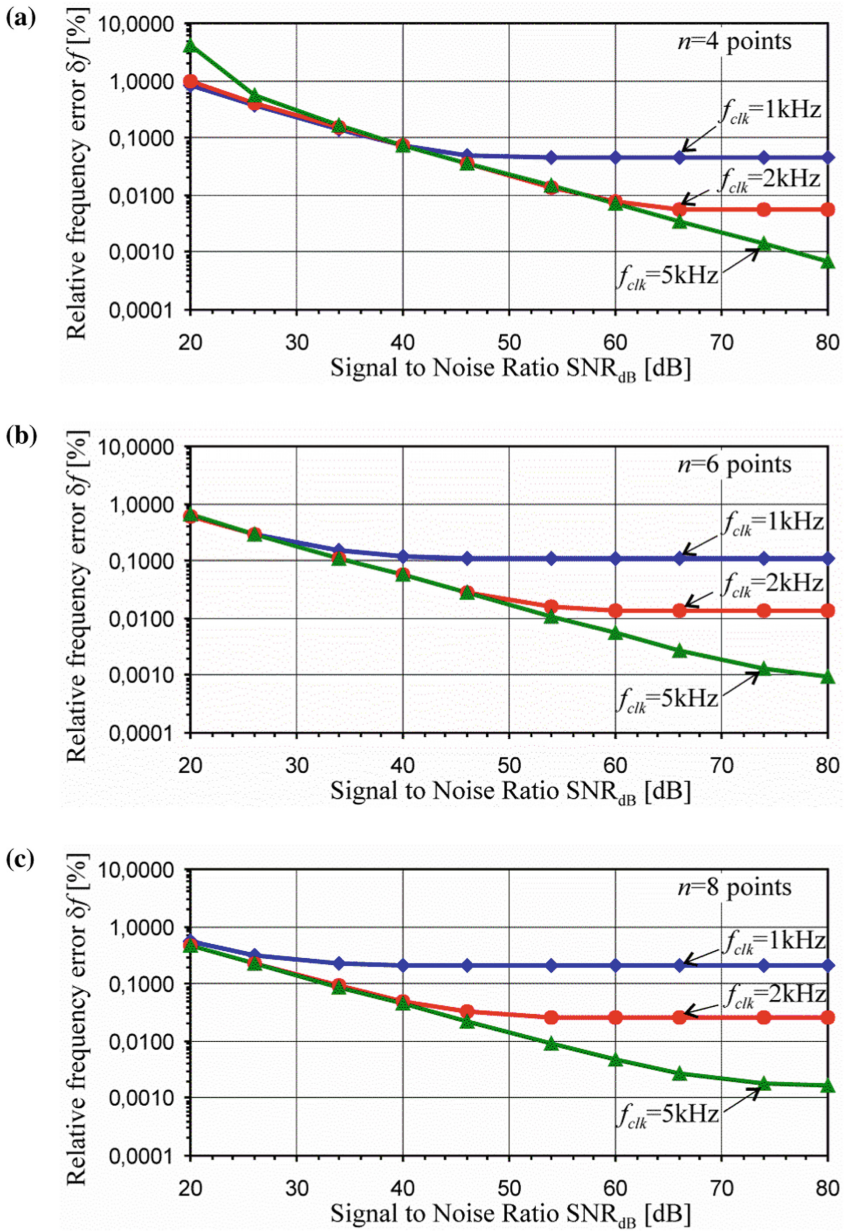
Fig. 8. Relative frequency error δf versus signal to noise ratio $SNR_{dB} = 20\ldots 80$ dB for different numbers of samples $n = 4, 6, 8$ and sampling ratio $f_{clk} = 1$ kHz, 2 Hz, 5 kHz.
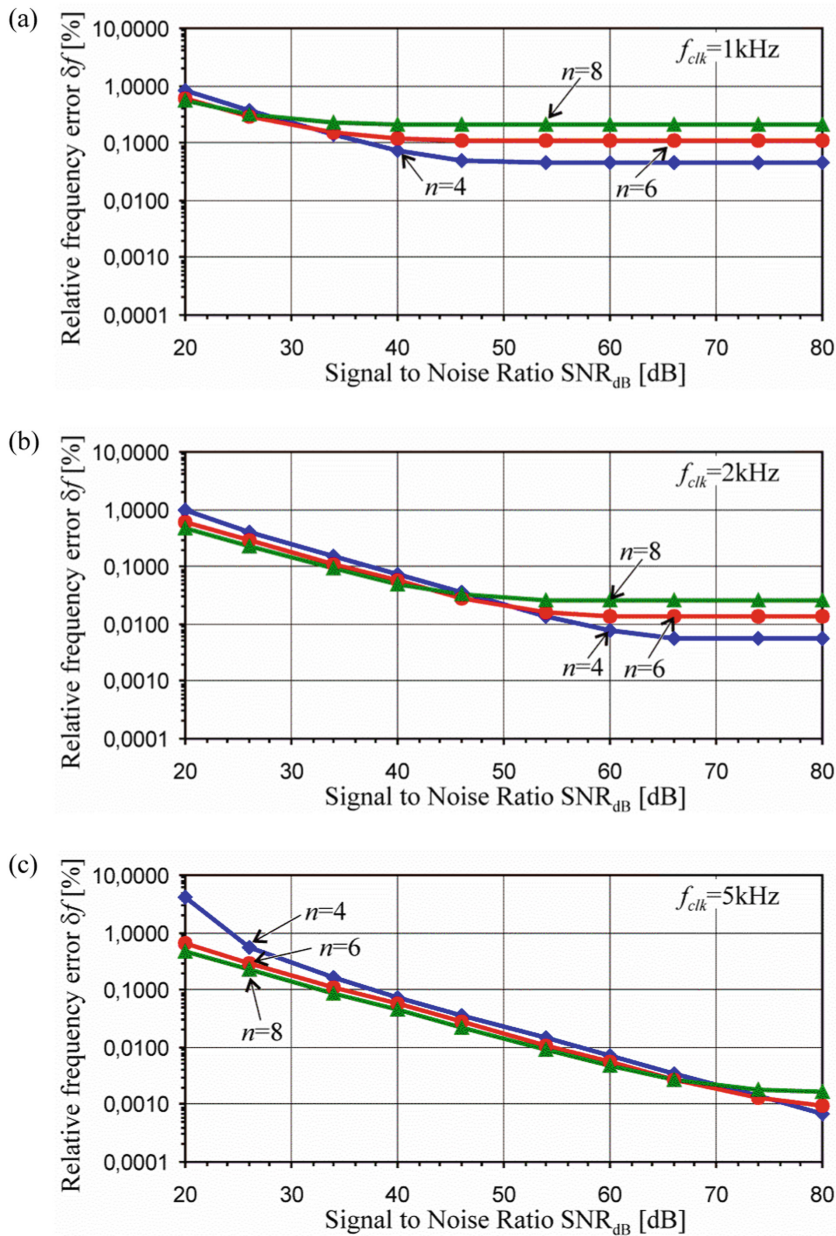
Fig. 9. Relative frequency error δf versus signal to noise ratio SNR_{dB} = 20 … 80 dB for different sampling ratio f_{clk} = 1 kHz, 2 Hz, 5 kHz and numbers of samples n = 4, 6, 8.

(a)

(b)

(c)
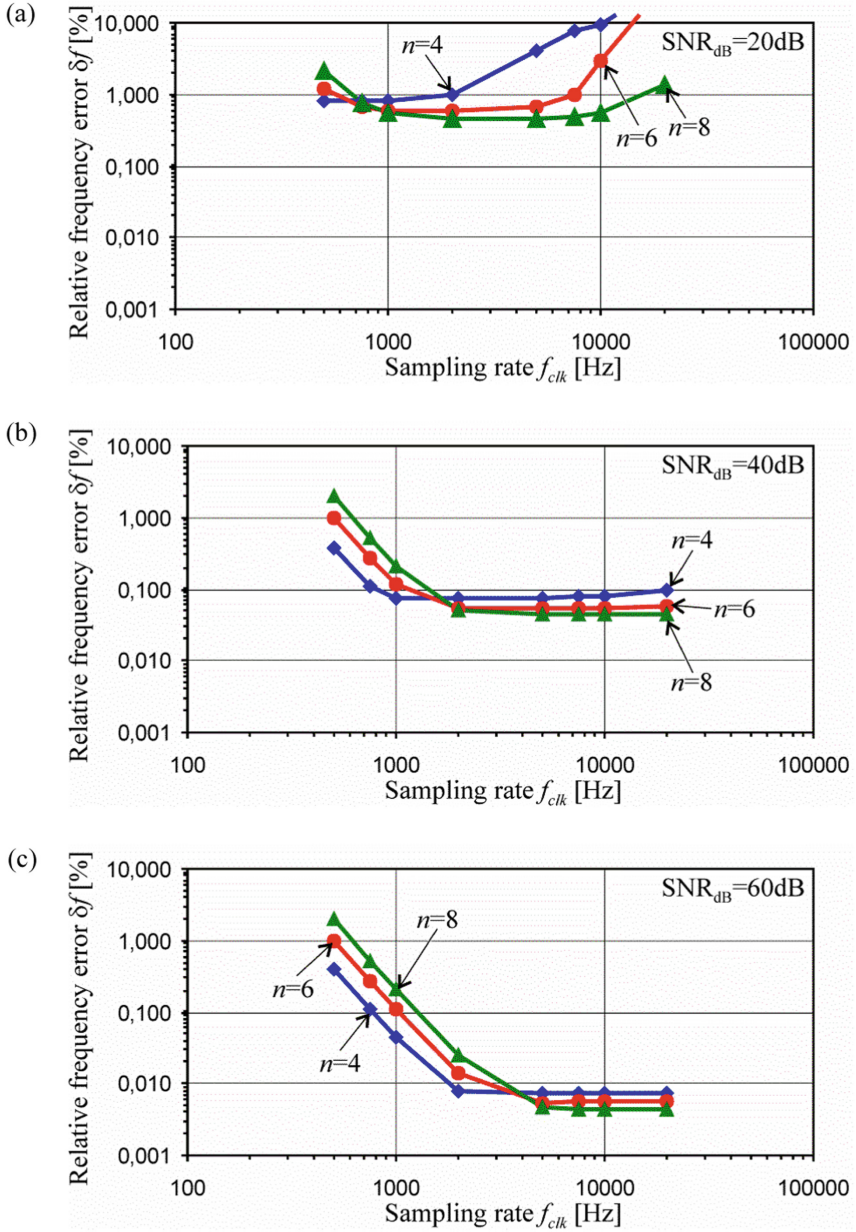
Fig. 10. Relative frequency error δf versus sampling ratio f_{clk} = 500 Hz … 20 kHz, for different values of signal to noise ratio SNR_{dB} = 20, 40, 60 dB and numbers of samples n = 4, 6, 8.

(a)

(b)

(c)
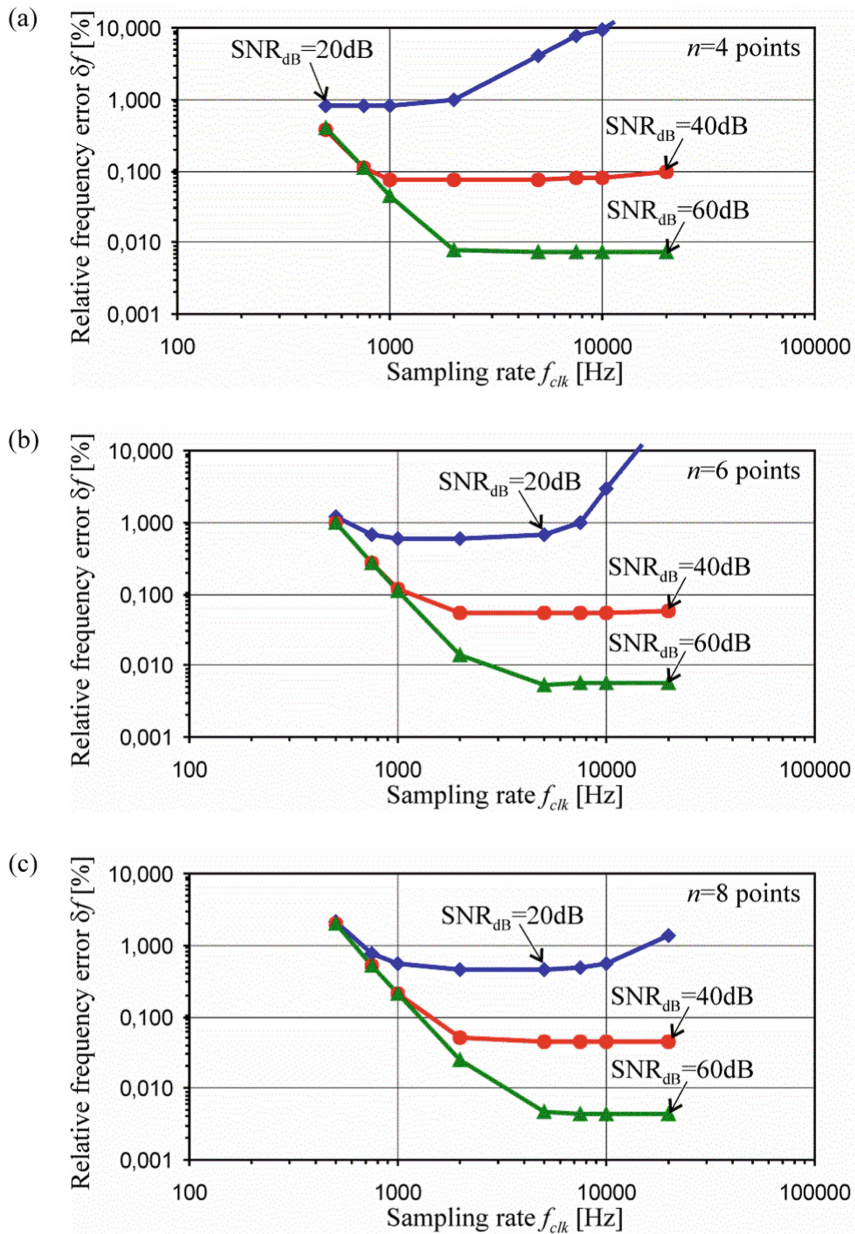
Fig. 11. Relative frequency error δf versus sampling ratio f_{clk} = 500 Hz … 20 kHz, for different numbers of samples n = 4, 6, 8 and signal to noise ratio SNR$_{dB}$ = 20, 40, 60 dB.

excessive number of samples with slow sampling excessively extends the approximation window and sinusoidal linearity errors (22) increase the measurement uncertainty. For strong interference SNR_{dB} = 20 dB, measurement errors below 0.5% for n = 8 and sampling rate f_{clk} = 2 … 5 kHz can be obtained (Fig. 10c). For smaller interference SNR_{dB} = 40 dB, errors reach below 0.05% for n = 8 at f_{clk} = 2 kHz and do not decrease at increasing sampling rate. Further reduction of interference to SNR_{dB} > 60 dB reduces errors below 0.005%, but requires an increase of the sampling rate to f_{clk} = 5 kHz.

(a)

(b)
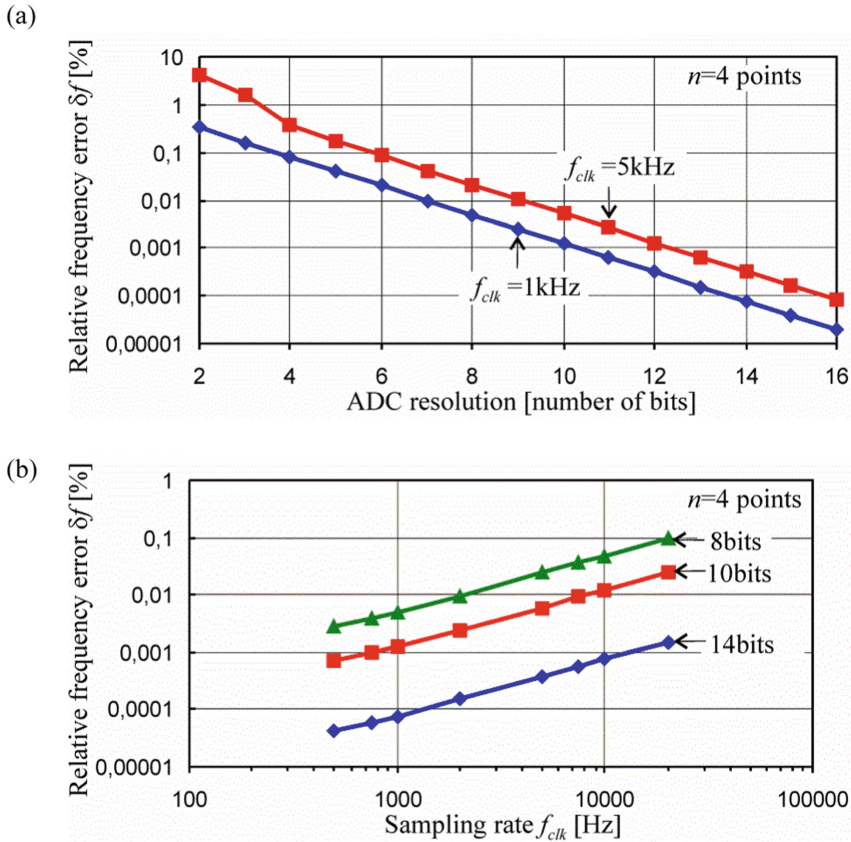
Fig. 12. Relative frequency error δf versus ADC resolution B = 2 … 16 bits (a) and versus signal sampling ratio f_{clk} = 500 Hz … 20 kHz (b) for number of samples n = 4.

4.3   The Effect of the ADC Resolution on the Accuracy

In Fig. 12, the relationship between the mean square error δ_f (26) and the number of bits B = 2… 16 is shown, at sampling rate f_{clk} = 500 Hz … 20 kHz, for the number n = 4.

The influence of the transducer's resolution on the accuracy of the measurements was simulated by adding a square-wave noise to the sinusoidal signal, simulating the

quantization noise. The width of the rectangular distribution A for the simulated quantization noise was calculated according to the following: $A = \frac{\text{FSR}}{2^B}$, where B is the number of bits, FSR is the full scale measuring range of the converter. FSR $= 2V_{max}$ is assumed, i.e. the signal is fully matched to the converter scale range.

From the graphs in Fig. 12a, the error decreases 10 times when the resolution increases by 3 bits. For example, for a 14-bit converter, the error is less than 0.0001% for $f_{clk} = 1$ kHz. Figure 12b shows the dependence of errors on the sampling rate. Errors increase when increasing the sampling rate f_{clk}, which is consistent with (21), where T_{clk} is in the denominator. It should be noted that the noise level of the ADC converter is significantly lower than the level of interference occurring in the power network voltage. The signal to noise ratio SNR_{dB} for a B-bit converter can be calculated from a known formula: $\text{SNR}_{\text{dB}} = 6.02B + 1.76 \, \text{dB}$. For example, a 10-bit converter introduces quantization noise of $\text{SNR}_{\text{dB}} = 62$ dB, and for a 14-bit converter $\text{SNR}_{\text{dB}} = 86$ dB.

5   Measurements in the Real-World System

The presented method was used practically in the arrangement shown in Fig. 13. The measurements were made in a single-phase 230 V/50 Hz network. Transformer TR provides galvanic isolation and adjusts the voltage of the power network ($V_{\max} \approx 325$ V) to the measuring range of the ADC (FSR $= \pm 20$ V). The NI USB 6009 Data
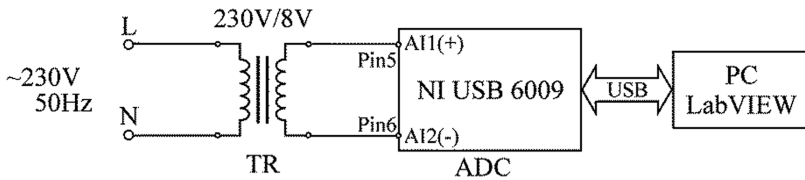
Fig. 13. Simplified flowchart of the measurement system for power grid frequency estimation.

Acquisition Card with 14-bit ADC converter was used, measurements were made in the differential mode on the ± 20 V range. Data collected by the ADC converter is transmitted via USB interface to a PC computer with LabVIEW software installed for its further processing according to the algorithm presented above.

The measurement algorithm was implemented in the LabVIEW environment. Figure 14 shows the fragment of the algorithm designating the first passage through zero t_{zc1}. In the WHILE loop, further signal samples are analyzed to determine the position of the samples lying around the zero crossing. The fixed number of n samples (e.g. 8 samples in Fig. 14) is then multiplied by the coefficients d_i and divided by c_n. Finally, the values of coefficients a (17) and b (14) and the instance of passing through zero t_{zc1} (12) are calculated. Figure 15 shows the program panel, on which the current result of the signal frequency measurement f_s and the graph in time function and histogram are presented. Exemplary measurement results shown in Fig. 15 was

Fig. 14. Selected part of the LabVIEW diagram for zero-crossing point t_{zc1} estimation.

Fig. 15. Panel of the LabVIEW diagram for power grid frequency estimation with exemplary measurement results for 3 night hours before 2 a.m. 2018-03-10.

obtained by the 3 night hours before 2 a.m. 2018-03-10. The first experiments carried

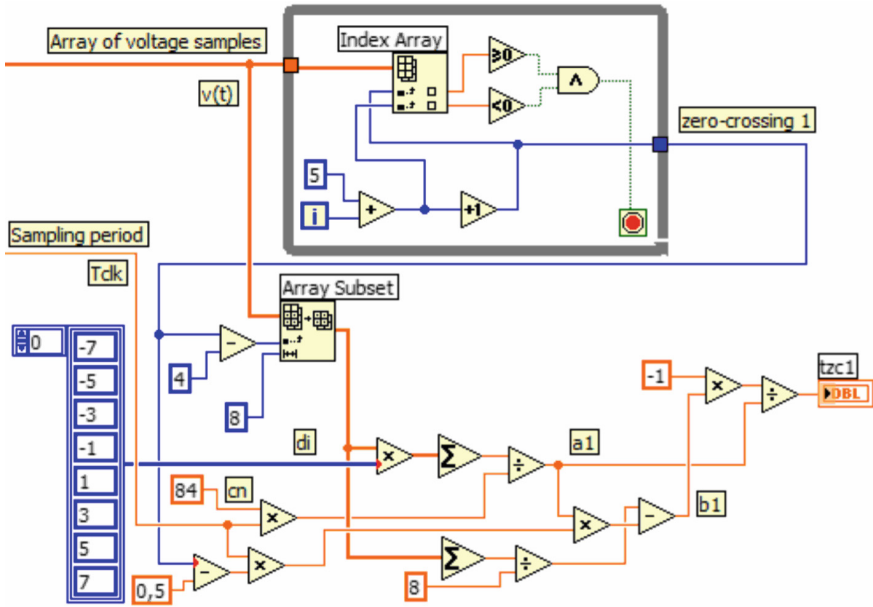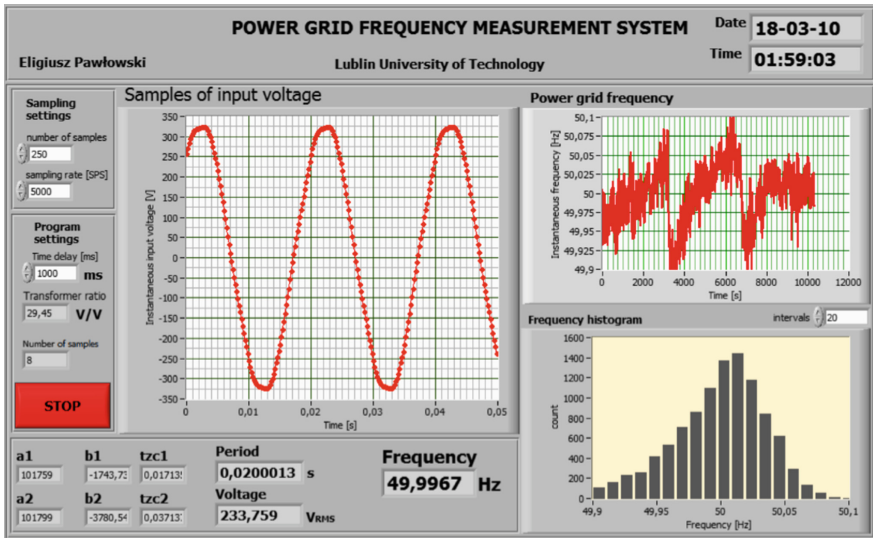out confirm the correctness of the method. Accurate tests to determine errors in the actual measurement system will be the subject of further work.

6 Summary

The proposed power grid frequency measurement algorithm uses the least squares method to linearly approximate the disturbed signal in order to determine zero crossing points. The appropriate mathematical relations were derived and numerical simulations were carried out to investigate the dependence of measurement error on transducer resolution, sampling rate, level of interference and the number of samples included in the approximation algorithm. A small measurement error can be obtained by selecting the algorithm parameters depending on the level of interference. For low noise level (SNR_{dB} = 60 … 80 dB) should use fast sampling (f_{clk} > 5 kHz) and at least 8 signal samples, which allows to achieve errors below 0.005%. By accepting a larger error rate, you can simplify the calculation of up to 6 or 4 samples, but the sampling rate should be reduced so that the approximation window falls on the linear part of the sine wave. For high noise levels (SNR_{dB} ≈ 40 dB), measurement errors below 0.05% can be obtained with approximation of 8 samples and f_{clk} > 2 kHz. The approximation window should not be wider than approx. 3 ms around the zero crossing. For very high level of interference (SNR_{dB} ≈ 20 dB) you can get errors of 0.5% for 8 samples and f_{clk} = 2 … 5 kHz. For very strong interferences, increasing the sampling rate above f_{clk} > 5 kHz is not beneficial because the approximation window is too narrowed and the errors increase.

The ADC converter introduces additional quantization noise and increases measurement errors, but the 14-bit converter already provides SNR_{dB} ≈ 86 dB, resulting in errors of only 0.00005%, which can be considered negligible in relation to method errors.

The built-in measurement system allowed for the practical verification of the proposed method. The initial measurements carried out confirmed the correctness of the developed algorithms. Further research is planned for the experimental evaluation of measurement errors.

References

1. IEC 61000-4-30:2015, Electromagnetic compatibility (EMC)-Part 4-30: Testing and measurement techniques - Power quality measurement methods
2. Liu, Y., Yuan, Z., Markham, P.N., Conners, R.W., Liu, Y.: Application of power system frequency for digital audio authentication. IEEE Trans. Power Deliv. **27**(4), 1820–1828 (2012)
3. Fundamentals of the Electronic Counters, Application Note 200, Electronic Counter Series, Hewlett Packard Co. (1997)
4. Lobos, T.: Nonrecursive methods for real-time determination of basic waveforms of voltages and currents. IEE Proc. C Gener. Transm. Distrib. **136**(6), 347–352 (1989)

5. Phadke, G., Thorp, J.S., Adamiak, M.G.: A new measurement technique for tracking voltage phasors, local system frequency, and rate of change of frequency. IEEE Trans. Power Appar. Syst. **102**(5), 1025–1038 (1983)

6. Moore, P.J., Allmeling, J.H., Johns, A.T.: Frequency relaying based on instantaneous frequency measurement. IEEE Trans. Power Deliv. **11**(4), 1737–1742 (1996)

7. Akke, M.: Frequency estimation by demodulation of two complex signals. IEEE Trans. Power Deliv. **12**(1), 157–163 (1997)

8. Lobos, T., Rezmer, J.: Real-time determination of power system frequency. IEEE Trans. Instrum. Meas. **46**(4), 877–881 (1997)

9. Zhang, C., Tan, J., Kirby, B., Bo, Z.: A derivative based instantaneous frequency tracking algorithm. In: 2008 43rd International Universities Power Engineering Conference, Padova, pp. 1–3 (2008)

10. Girgis, A.A., Ham F.M.: A new FFT-based digital frequency relay for load shedding. IEEE Trans. Power Appar. Syst. **101**(2), 433–439 (1982)

11. Sachdev, M.S., Giray, M.M.: A least error squares technique for determining power system frequency. IEEE Trans. Power Appar. Syst. **104**(2), 437–444 (1985)

12. Hwang, J.K., Markham, P.N.: Power system frequency estimation by reduction of noise using three digital filters. IEEE Trans. Instrum. Meas. **63**(2), 402–409 (2014)

13. Cai, H.: Fast frequency measurement algorithm based on zero crossing method. In: 2010 2nd International Conference on Computer Engineering and Technology, Chengdu, pp. V4-606-V4-608 (2010)

14. Vizireanu, D.N.: A simple and precise real-time four point single sinusoid signals instantaneous frequency estimation method for portable DSP based instrumentation. Measurement **44**(2), 500–502 (2011)

15. Alegria, F.C., Molino-Minero-Re, E., Shariat-Panahi, S.: Evaluation of a four-point sine wave frequency estimator for portable DSP based instrumentation. Measurement **45**(7), 1866–1871 (2012)

16. Mendonça, T.R.F., Pinto, M.F., Duque, C.A.: Least squares optimization of zero crossing technique for frequency estimation of power system grid distorted sinusoidal signals. In: 2014 11th IEEE/IAS International Conference on Industry Applications, Juiz de Fora, pp. 1–6 (2014)

17. Maru, K., Fujii, Y., Hessling, J.P., Shimada, K.: Frequency estimation method for measuring time-varying single frequency from digitized waveform. In: 2009 4th IEEE Conference on Industrial Electronics and Applications, Xi'an, pp. 3687–3689 (2009)

18. Wall, R.W.: Simple methods for detecting zero crossing. Industrial Electronics Society, 2003. In: IECON'03, The 29th Annual Conference of the IEEE, Vol.3, pp. 2477–2481 (2003)

19. Pawłowski, E., Warda, P.: Power grid frequency measurement in LabVIEW environment using the least mean squares method to signal phase approximation in the presence of noise. In: 2017 International Conference on Electromagnetic Devices and Processes in Environment Protection (ELMECO), Lublin, pp. 1–4 (2017)

20. Świsulski, D., Hanus, R., Zych, M., Petryka, L.: Methods of measurement signal acquisition from the rotational flow meter for frequency analysis. Eur. Phys. J. Web Conf. **143**, 02124 (2017)

21. Taylor, J.R.: An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements. 2nd edn. University Science Books, Sausalito, California (1997)

Measurements of Gas Streams with Disturbed Velocity Distribution by the Means of an Elbow Flowmeter

Piotr Piechota, Piotr Synowiec, Artur Andruszkiewicz[(✉)],
and Wiesław Wędrychowicz

Wrocław University of Science and Technology,
Wybrzeże Stanisława Wyspiańskiego 27, 50-370 Wrocław, Poland
`artur.andruszkiewicz@pwr.wroc.pl`

Abstract. The subject of the article concerns the measurement of flow streams with disturbed velocity distribution by the means of elbow flowmeter. The velocity disturbances were forced by a different setting of the aperture located at the inlet to the installation. In the article, flow coefficients for the elbow flowmeter for exemplary disturbances of velocity distribution were determined and deviations in relation to the flow without disturbances were calculated. For the exemplary disturbance, numerical calculations were performed, the results of which were compared with the actual results of the experiment.

Keywords: Elbow flowmeter · Flow stream · Non-invasive measurement

1 Introduction

Among the many types of flowmeters for continuous flow streams measurement, the elbow flowmeters deserve attention [1–6]. Their advantage is that they are non-invasive flowmeters – they do not interfere with the flow and do not cause any additional pressure loss. It is possible to measure the flow rates of both liquids and gases in a wide range of Reynolds numbers and also to measure solid phase stream, for example in pneumatic transport systems. In the each transport installation there is an elbow and by measuring the differential pressure between the outer and inner side of the elbow and determining the flow coefficient, it is possible to continuously measure the medium stream. Very often it happens that measurements are made in the absence of sufficiently long straight sections before the elbow or immediately after the aperture or the narrowing in the pipeline. In such cases, the velocity distribution of the medium in the pipeline cross section is disturbed. It would then be necessary to introduce a correction to the equation of characteristic of elbow flowmeter due to the change in the flow coefficient factor and related with the non-symmetrical velocity distribution. The aim of the article is to determine flow coefficients for an exemplary velocity disturbance and to define corrections related to this disruption in the air installation flow. The article also carried out simulations of the work of elbow flowmeter at the disturbed velocity distribution and the obtained results were compared with the results from experiments.

2 Measuring Principle

The principle of measurement is the proportionality of the stream flow from the square root of the fluid pressure difference between the outer and the inner side of the elbow. The difference is caused by the centrifugal force acting on the fluid flowing in the elbow.

The characteristic equation can be presented as follows:

$$q_{mk} = C_k \cdot A \cdot \sqrt{\Delta p_k \cdot \rho_k} \tag{1}$$

In which:

q_{mk} – mass flow,
C_k –flow factor depending on the Reynolds number, cross-section of the elbow,
Δp_k –static pressure difference on the elbow,
ρ_k- air density before the elbow

As it results from Eq. 1, in order to be able to measure the stream flow, the value of the coefficient C_k from the Reynolds number and the minimal value of this number should be determined starting from which the value of the flow coefficient is constant.

Determination of the C_k coefficient requires determination of the mass flow by another method, for example orifice method. Knowing the actual mass flow in the installation from the orifice q_z, the value of coefficient C_k is expressed by the equation:

$$C_k = \frac{q_z}{\frac{\pi D_k^2}{4} \sqrt{\rho_k \Delta p_k}} \tag{2}$$
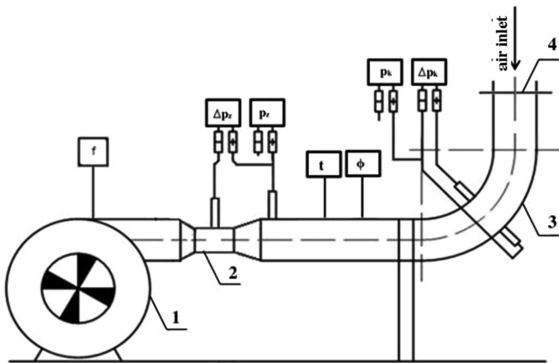
Fig. 1. Scheme of the measurement stand: 1 – centrifugal fan, 2- Venturi tube, 3 – elbow flowmeter, 4 – aperture simulating flow disturbances.
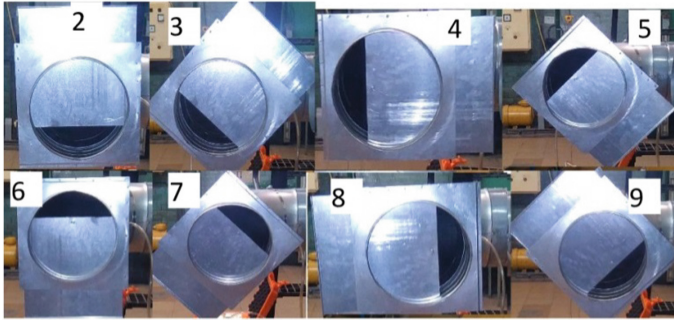
Fig. 2. Positions of apertures disturbing the velocity distributions.

The measurements were taken at the stand shown if Fig. 1. In the picture, in Fig. 2, the position of the aperture disturbing the velocity distributions are shown.

The differential pressure Δp_k on the elbow, with diameter $D_k = 0.388$ m, and the static pressure p_k were measured and recorded every 5 s for each aperture position using a digital manometer with a resolution of 1 Pa. The recording time was 300 s. At the same time and with the same measurement step, the real mass flow with a Venturi tube with a flow coefficient $C_z = 0.925$ and an internal diameter $d_z = 0.271$ m was measured. The air temperature in the installation was measured with a Pt100 metal resistance thermometer and a relative humidity ϕ with a thermo-hygro-barometer. The air stream flow was set by changing frequency from 45 Hz to 5 Hz with a 5 Hz step using an i55A power inverter. The first series of measurements was made in the installation without an aperture, for 9 different air mass flow and the flow coefficient for C_k elbow was calculated. The mass flow was changing in the range (460÷5300) kg/h and the Reynold number (72000÷460000). Then measurements for each aperture setting, as in Fig. 2, were made. Unfortunately, due to the large resistances, the maximum mass flow was not obtained and measurements were made for 8 or 7 flow streams. For these measurements and each position of the aperture, the C_k coefficients for the elbow, from the Eq. 2, were also calculated.

3   Results from the Measurements and Simulations

Figure 3a shows the obtained values of the flow coefficient depending on the Reynolds number for measurements without an aperture and for selected 5 position of the aperture. It shows that the values of flow coefficients stabilize from certain Reynolds numbers. For measurements without the aperture, these values are of the order of $Re = 148000$ and for some positions of the aperture these values are approximately $Re = 200000$. However, there are differences in the flow coefficient in the relations to the measurement without the aperture. In the case of the 9th position of the aperture, the C_k coefficients are greater than for the measurement without it and its average value is $C_k = 1.16$, while in the remaining cases it is smaller. The measurements also showed that in the case of measurements 2÷7, the characteristics of the C_k coefficient versus the

Reynolds number are close to each other and are within the limits of $C_k = (0,79 \div 0,87)$. This is shown in Fig. 3b, which shows the calculated mean values of the C_k coefficient
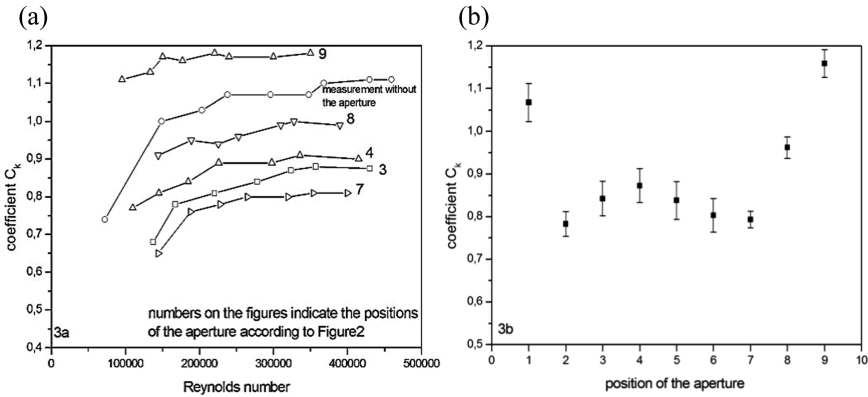
Fig. 3. The dependence of the C_k flow coefficient for the elbow on the Reynold number (a) and the position of the aperture disturbing the velocity distribution (b).

along with the standard deviations of this coefficient for all aperture positions.

Figure 4 shows the dependence of the relative percentage deviation of the flow coefficient for the elbow on the measurement without an aperture for its various positions. These deviations range between $(-27 \div -18)\%$ for positions from 2 to 7, for positions 8 it is -9.9% and for 9 $+8.5\%$. This figure also shows exemplary characteristics of the elbow
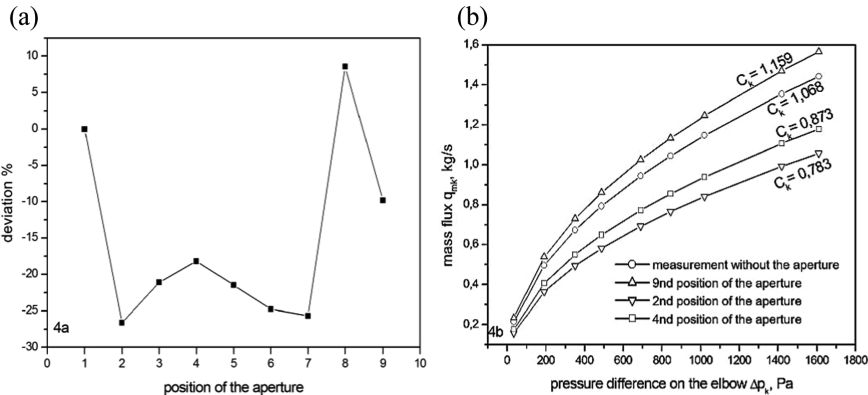
Fig. 4. The dependence of the flow coefficient deviation for the elbow on the position of the aperture (4a) and selected characteristics of the elbow flowmeter (4b).

flowmeter calculated from Eq. 1 for selected aperture positions.

Figure 5 shows exemplary results of the pressure distribution after the aperture, in position 2, obtained from simulations and comparison of the results of numerical calculations and experimental measurements for this position of the aperture, as well as measurements without apertures. Numerical calculations were made in the Fluent program [7], using the Navier-Stokes (N-S) equations of conservation of fluid momentum and flow continuity equation. The difference in obtained results does not exceed 10% in both cases.
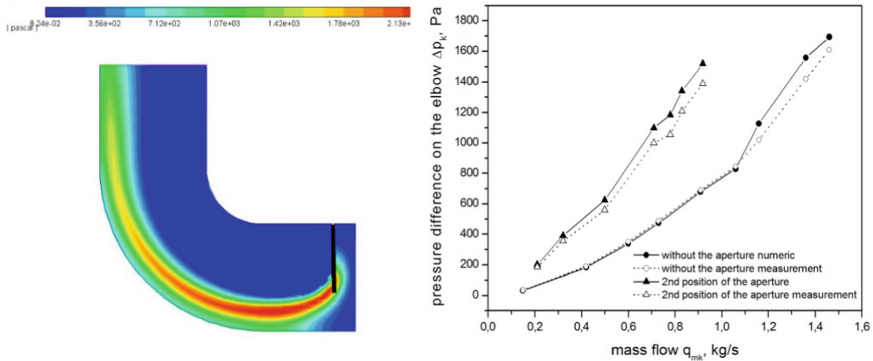
Fig. 5. Example results of pressure distribution for the 2nd aperture position and mass flow $q_m = 0.68$ kg/s along with the comparison of numerical calculations with the results obtained from measurements.

4   Summary

The measurements showed out how the value of the flow coefficient for the elbow flowmeter changes in a case of a disturbed velocity distribution. Variable velocity distributions were forced by a different setting of aperture shown in Fig. 2. Calculations of C_k flow coefficient values showed that its values may vary up to about 27% in relation to the measurement without disturbance. This requires introducing a large correction to the stream flows values. The performed numerical calculations showed that the results were consistent with the experimental studies – the differences were not greater than 10% and for some of the pressure values were practically coincided. However, values from numerical calculations have always been greater than in the case of measurements. Therefore, it seems, that numerical simulations can be used to determine correction factors in the case of a disturbed velocity distribution. The study's results presented in the article are just a small part of the work on the measurement of flow streams with various types of flowmeters in a conditions that differ from the manufacturer's recommendations.

References

1. Synowiec, P., Wędrychowicz, W., Andruszkiewicz, A.: Analysis of the possibilities of using eblow flowmeter to measure the mass flow in the flow of gas – solid. Elektronika **10**, 89–90 (2016)
2. Grzeszczak, J., Kamiński, M., Andruszkiewicz, A.: Elbow flowmeter for measuring gas flow streams. Elektronika **6**, 45–46 (2014)
3. Sudo, K., Sumida, M., Hibara, K.: Experimental investigation on turbulent flow in a circular-sectioned 90- degree bend. Exp. Fluid. **25**(1), 42–49 (1998)
4. Rup, K., Malinowski, Ł.: Sarna, P: measurement of flow rate in square- sectioned duct bend. J. Theor. Appl. Mech. **49**(2), 301–311 (2011)
5. Feng, B., Wang, S., Shengqiang, L., Yang, X., Shengyao, J: Experimental and numerical study on pressure distribution of 90° elbow for flow measurement. Sci. Technol. Nucl. Install. **2014**, ID 964585, 7 p
6. Mrowiec, A.: Optimization of the elbow flowmeter. Ph.D. thesis. Opole University of Technology (2013)
7. Fluent 12.0 User's Guide, Fluent Inc. (2009)

Analysis of the Accuracy of Liquid Flow Measurements by the Means of Ultrasonic Method in Non-standard Measurements Conditions

Piotr Piechota[(✉)], Piotr Synowiec, Artur Andruszkiewicz, and Wiesław Wędrychowicz

Wrocław University of Science and Technology,
wybrzeże Stanisława Wyspiańskiego 27, 50-370 Wrocław, Poland
`piotr.piechota@pwr.edu.pl`

Abstract. The article presents the results of ultrasonic flow measurements performed after the hydraulic elbow. Ultrasonic flowmeter with applied head set in accordance with the Z-type was used to carry out the measurements. The results of flow measurements after the hydraulic elbow were referenced to measurements made on a straight section of the pipeline before the elbow, where the flow was stabilized, and the velocity profile was symmetrical with respect to the pipe-line axis. Measurements, with the maintaining a constant volumetric stream flow, were made for 12 different angles of the flowmeter head settings in 16 distances from the hydraulic elbow. The results of the measurements were compared with the velocity values obtained from the flow simulation performed in the ANSYS CFX program. On the basis of the comparison of the measurement results with the simulation results, and also based on the analysis of the velocity profiles, it was found that at the appropriate angle of the head setting, measurements can be made using the ultrasound method at a distance smaller than the one described in the standards. The optimal location of the measurement can be selected on the basis of a computer flow simulation, which is a representation of geometry and measurement conditions. This action scheme can be used in the flow measurements, which are carried out after the obstacle which is disturbing the flow, in the pipelines with large diameters (for example power plants, electrical power and heating plant, chemical industry) where finding a straight section with a length of 15–20 pipeline diameters is problematic.

Keywords: Ultrasonic flowmeter · Turbulent flow · Hydraulic elbow

Nomenclatures

D   diameter of the pipeline
K   turbulent kinetic energy
Re   Reynolds number
U   flow velocity
U_j mean flow velocity component in the x_j coordinate direction
S_{ij} mean strain rate tensor
t   transit time of the ultrasonic wave

x_j space coordinate component j = 1, 2, 3
Q   flow rate
α angle of the flowmeter probe setting
ε turbulent dissipation rate
μ_t turbulent eddy viscosity
ρ density
τ_{ij} total stress tensor
τ_{tij} turbulent Reynolds stress tensor
ω specific turbulent dissipation rate.

1   Introduction

Systems for a stream and flow velocity measurements of liquids and gases have a big importance in industrial processes. Their most desirable features include high accuracy, reliability and non-invasiveness. The group of measurement methods that meet these standards includes the ultrasonic stream flow measurement method. To perform measurements, the results of which are presented in this article, used a transit-time type flowmeter.

1.1   Measurement of Liquid Flow by the Means of Ultrasonic Method

The main advantages of the ultrasonic method are the lack of flow disturbances and pressure losses caused by the use of a flowmeter (ultrasonic). An unquestionable advantage of this method is also the simple assembly of ultrasonic heads on the outer surface of the pipeline, without need to interfere in the flow installation. Flow measurement results obtained using an ultrasonic flowmeter have a relatively small error. When measured with an ultrasonic flowmeter, the permissible error are within the range of 0.5%–1.5% and depends on the type of flowmeter, its metrological properties, as well as the condition of the pipeline in which the flow takes place. Obtaining such measurement accuracy is possible when maintaining the appropriate straight pipeline sections, suggested by the manufacturer, directly before and after the measuring location. The article, however, presents the results of measurements made directly after the obstacle that is disturbing the flow. Such a situation is often encountered in measurement practice, when it is impossible to preserve the required straight pipeline sections given by the manufacturer or standards [1].

1.2   Operating Principle of the Transit-Time Type Ultrasonic Flowmeter

The aforementioned transit-time method consist in determining the value of the flow stream based on the difference in transit times of ultrasonic waves between two sensors. [2, 3] Transit-time flowmeters are used to measure the flow of clean liquids and liquids that contain small amounts of solids or air bubbles [4]. Various configurations of the mutual positions of the sensors are possible, which gives the device users the freedom

to take measurements. In the case of made measurements, the flowmeter heads were mounted on the pipeline walls and arranged along its radius, in accordance with the Z-type system, which is shown in Fig. 1.
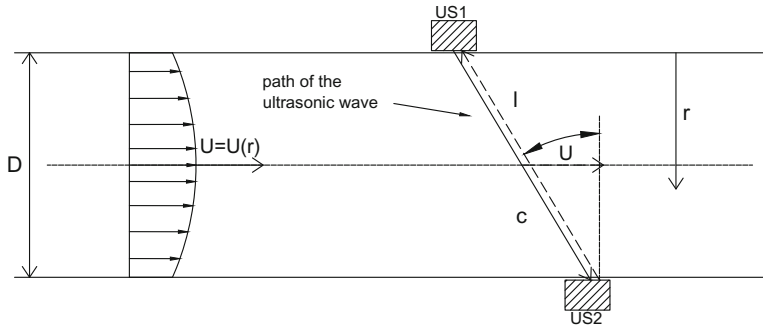
Fig. 1. Scheme of setting the measuring heads of the transit-time type ultrasonic flowmeter in accordance with the Z-type system

The transit time of the ultrasonic wave according to the flow direction describes the Eq. (1), while the transit time of the ultrasonic wave in the direction opposite to the flow is described by Eq. (2) [5].

$$t_1 = \frac{l}{c + U \cdot \cos \alpha} + \sum_{i=1}^{n} \frac{l_i}{ci} \tag{1}$$

$$t_2 = \frac{l}{c - U \cdot \cos \alpha} + \sum_{i=1}^{n} \frac{l_i}{ci} \tag{2}$$

The flow velocity of the liquid in the ultrasonic path determined from Eqs. (1) and (2) is given by the Eq. (3)

$$U = \frac{l}{2 \cdot \cos \alpha} \cdot \left(\frac{1}{t_1} - \frac{1}{t_2}\right). \tag{3}$$

2 Carrying Out Measurements

2.1 Methodology of Measurements

Stream flow measurements were made at a constant flow rate $Q = 16$ m^3/h. Two ultrasonic flowmeters were used for the measurements. One of them was installed before the hydraulic elbow and its location was not changed. Such arrangement of this flowmeter on the straight section of the pipeline, in the place where the flow was not disturbed, allowed for reference to the second flowmeter, installed after the hydraulic

elbow. The position of the flowmeter installed after the elbow was changed to determine the influence of the distance from the hydraulic elbow and the angle of heads positioning on the liquid flow obtained as a result of the measurement. For this purpose, measurements were taken in 16 measurement diameter (0D-15D), as shown in Fig. 2. The distance between successive measurement diameter was equal to the outer diameter of the pipeline and was D = 60.2 mm.
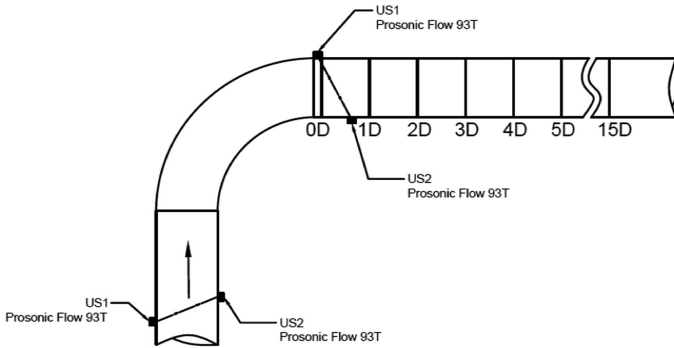
Fig. 2. Scheme of the ultrasonic flowmeters installation, showing the measurement diameter after the hydraulic elbow

For each measurement diameter 12 measurement series (angle of the ultrasonic head in the range $\alpha = 0°-360°$) lasting 10 min were made. The indications of both flowmeters were recorded every 1 s. After completing a single measurement at a given measuring location, the flowmeter heads were rotated for a 30° around the horizontal axis of the pipeline (Fig. 3).
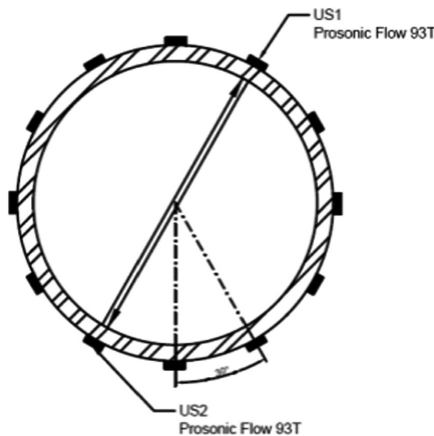
Fig. 3. Scheme showing the changes in the angle of the ultrasonic flowmeter heads

2.2 Results of Measurements

The following charts (Fig. 4) presents the results of the flow velocity measurements, before and after the hydraulic elbow. The following measurement diameterwere con-
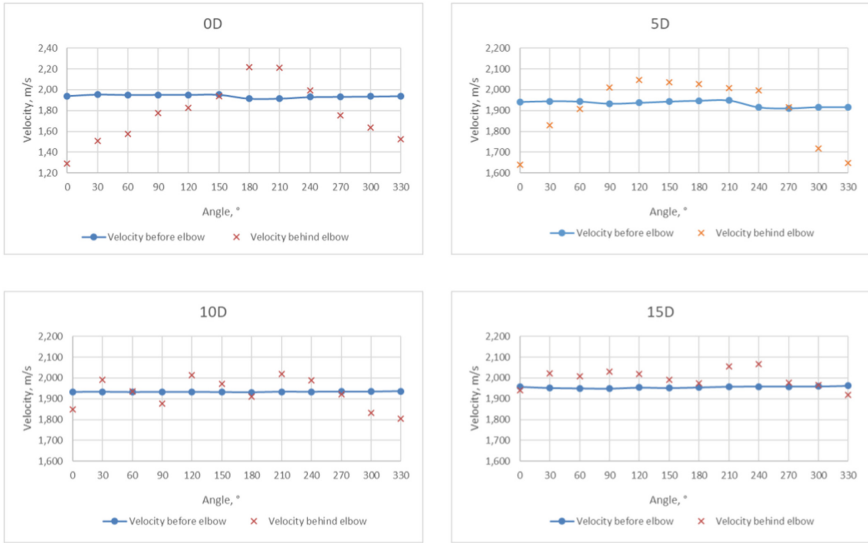
Fig. 4. Charts of flow velocity dependence U (in the next measurement diameters behind the hydraulic elbow 0D-15D) of the angle of ultrasonic head positioning α.

sidered as representative: 0D, 5D, 10D and 15D.

Analyzing the presented charts (Fig. 4) it should be noted that the velocity values measured before the hydraulic elbow do not significantly change. This is a consequence of the constant stream flow of liquid at which the measurements were made. Differences between the velocity values measured after the elbow and before elbow decreasing with increasing distance from the hydraulic elbow. Velocities differ the most in the 0D measurement diameter. In the section 15D the differences are very small – the calculated relative error reaches a maximum of 5.5%, while the mean value of the relative error is 2.6%.

In cross-sections 0D and 5D, the velocity fluctuations measured after the elbow are very large – for the extreme angles of 0° and 330° the velocity values are the smallest, while for the angles of 120°, 150° and 240° are the closest to the velocity values measured before the elbow. In cross-sections 10D and 15D, the velocity values measured after the elbow oscillate around the velocity values before the elbow and there are not significant changes depending on the angle.

3   Computer Simulation of Flow

Reproducing the geometry of the pipeline on which the measurements were made, a numerical grid was created on which the simulation in the ANSYS CFX program was performed. The measuring installation was divided into 3 elements (before the elbow, hydraulic elbow, after the elbow) and in each of them a hexagonal grid was generated (Fig. 5).
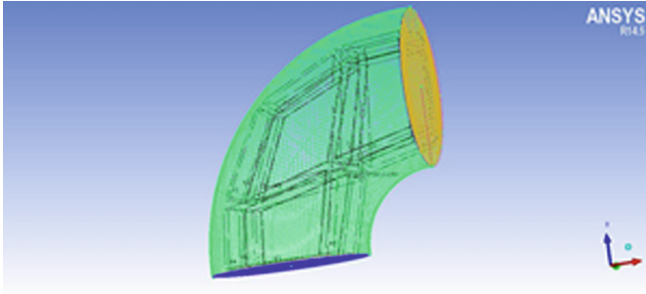
Fig. 5. Hexagonal grid generated for the hydraulic elbow

Simulation was performer by mapping the flow conditions from measurements. The selected material that was set for the created geometry was water at 18°C. The simulation does not take into account the thickness of the pipeline walls through which heat transport took place. This phenomenon was omitted due to the negligible effect of heat exchange on the velocity and the nature of the flow. That is why the Isothermal heat exchange model was chosen.

The flow, for which the measurements were made, was a highly turbulence flow. The Reynolds number, calculated according to the Eq. (4) for the results of the measurement of velocity after the elbow, was an average of 98000. While reproducing the flow during the simulation preparation, this value was used when declaring the simulation conditions and choosing the appropriate turbulence model. Using the results from the measurements, the average flow velocity before the hydraulic elbow was also calculated. The calculated velocity served as the boundary condition at the inlet. At the outlet, the boundary condition was set using the Average Static Pressure option and Relative Pressure = 0. On the remaining walls the Wall condition was imposed.

$$Re = \frac{\rho U D}{\mu} \tag{4}$$

As the turbulence model, the k-equation model, which is suitable for describing turbulent flow, was used. The k-equation model is a two-equation model. The k-Eq. (5) is a model of the transport equation for the turbulent kinetic energy, and the ε-Eq. (6) is a model for the dissipation rate of turbulent kinetic energy [6].

$$\frac{\partial \rho k}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho U_j \frac{\partial k}{\partial x_j} - \left(\mu + \frac{\mu_\tau}{\sigma_k}\right)\frac{\partial k}{\partial x_j}\right) = \tau_{tij}S_{ij} - \rho\varepsilon + \phi_k \qquad (5)$$

$$\frac{\partial \rho \varepsilon}{\partial t} + \frac{\partial}{\partial x_j}\left(\rho U_j \varepsilon \frac{\partial k}{\partial x_j} - \left(\mu + \frac{\mu_\tau}{\sigma_\varepsilon}\right)\frac{\partial \varepsilon}{\partial x_j}\right) = C_{\varepsilon 1}\frac{\varepsilon}{k}\tau_{\tau ij}S_{ij} - C_{\varepsilon 1}f_2\rho\frac{\varepsilon^2}{k} + \phi_\varepsilon \qquad (6)$$

3.1   Simulation Results

The choice of the appropriate turbulence model proved to be crucial for correctness of the simulation performance. Its results in the form of velocity profiles are shown in Fig. 6.
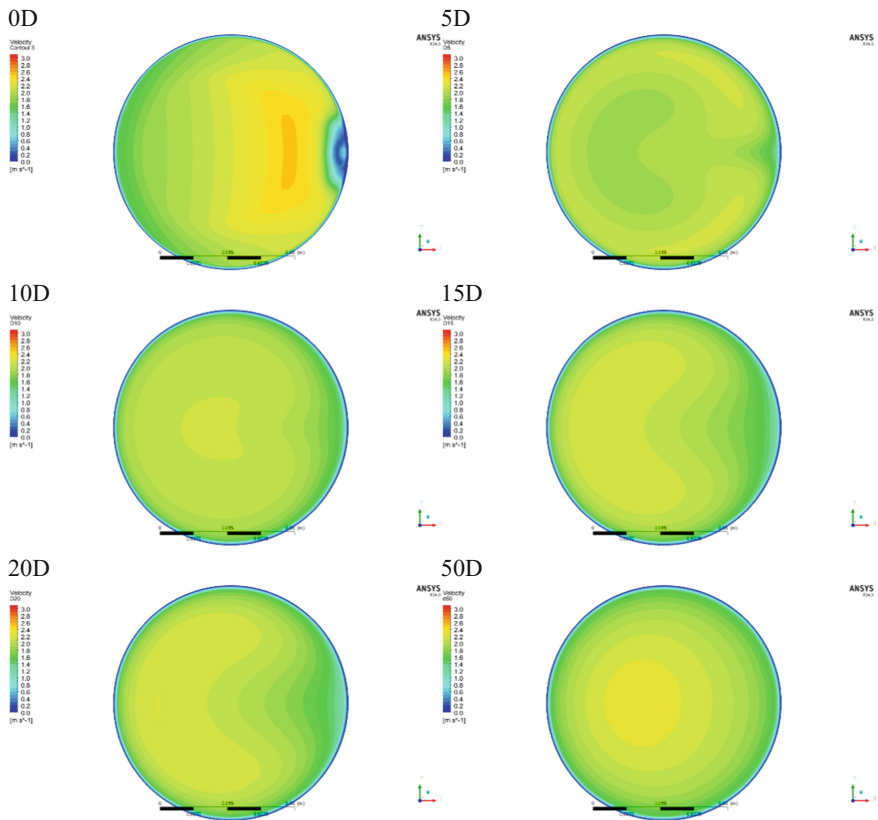
Fig. 6. Juxtaposition of velocity profiles for selected measurement cross-sections

Based on the analysis of the presented velocity profiles, it can be concluded that the flow can be considered for stabilized or symmetrical only for the 50D cross-section. For cross-sections 5D, 10D, 15D and 20D, individual velocity layers are very similar. Only for the 0D cross-section the velocity values deviate significantly from the others. In this measurement diameter the flow is strongly disturbed after it emerges from the hydraulic elbow. The resulting vortex structures have a significant influence on the velocity value. The flow disturbances in the 0D measurement diameter are visible on the graph (Fig. 9) showing the flow velocity distribution in the pipeline. In the 5D, 10D, 15D and 20D cross-sections, the propagation of vortices is visible. The vortex structure is only blurred in the 50D cross-section.

However, based on the analysis of velocity profiles, we cannot extract information on the flow velocity measured using the ultrasonic method. For this purpose, lines that illustrate the path of ultrasonic wave transitions have been created. They were shown in Fig. 7. Such lines were created for all angles of flowmeter head setting (Table 1).
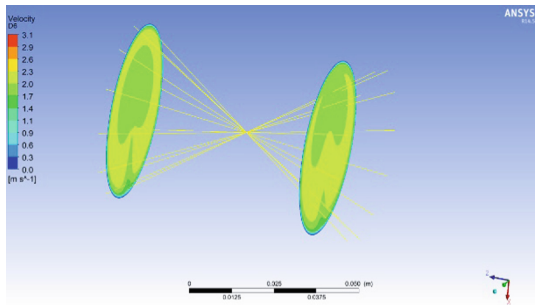
Fig. 7. Created an imitation of the path of ultrasonic wave transition

Table 1. Comparison of measured values with simulation results

D	U [m/s]/Angle	0°	30°	60°	90°	120°	150°	180°	210°	240°	270°	300°	330°
0	behind the elbow	1,29	1,51	1,58	1,77	1,83	1,94	2,22	2,21	1,99	1,76	1,64	1,52
	k-ε Model	1,75	1,79	2,01	2,19	2,21	2,15	1,64	2,15	2,21	2,19	2,01	1,79
5	behind the elbow	1,64	1,83	1,91	2,01	2,05	2,04	2,03	2,01	2,00	1,92	1,72	1,65
	k-ε Model	1,88	1,94	1,94	1,94	1,94	1,94	1,87	1,94	1,94	1,94	1,94	1,94
10	behind the elbow	1,85	1,99	1,94	1,88	2,01	1,97	1,91	2,02	1,99	1,92	1,83	1,80
	k-ε Model	1,93	1,96	1,99	2,00	1,99	1,97	1,94	1,97	1,99	2,00	1,99	1,96
15	behind the elbow	1,94	2,02	2,01	2,03	2,02	1,99	1,97	2,05	2,07	1,98	1,97	1,92
	k-ε Model	1,89	1,94	2,00	2,01	2,00	1,94	1,90	1,94	2,00	2,01	1,99	1,94

The chart presented below as Fig. 8 shows the relative error that would be made assuming the values obtained from simulation as a measured velocity.
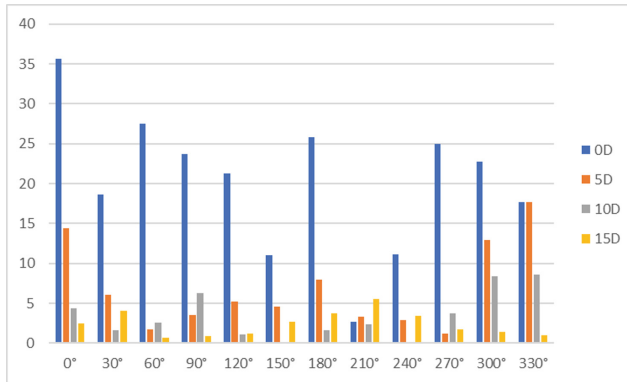
Fig. 8. The relative error for the velocity equal to the velocity from the simulation
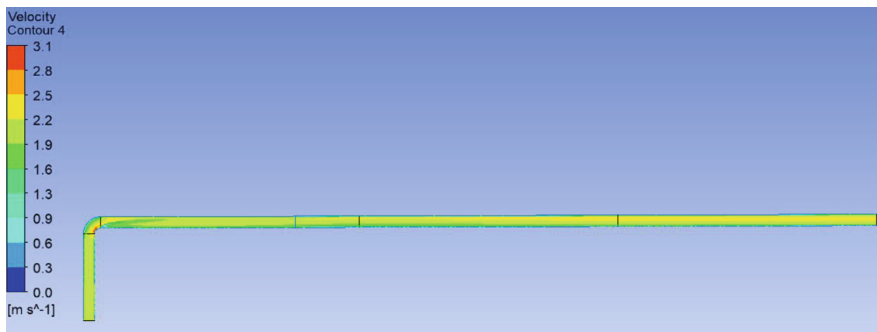
Fig. 9. Graph showing the velocity distribution in the pipeline in the whole analyzed flow length

It is easy to see that the velocity values from the simulation best match to the measured values for an angle of 210°. For the 10D and 15D measuring diameters, the data from simulation coincides with the measurement values.

4   Conclusion

The purpose of the article was to show that the making velocity measurements behind an obstacle disturbing the flow, at a distance less than the one recommended in the instruction or standard, can be a method used in measurement practice. This could greatly facilitate measurements in large diameter pipelines when it is difficult to find the correspondingly long straight pipeline sections. Analyzing the measurement data and the results of the computer flow simulation it was found:

- For measurement diameters at a distance from the hydraulic elbow further than 10D, the velocity values are very close to the velocity values measured before the elbow. For these diameters, the selection of the angles of flowmeter head setting

does not have a large impact on the measured velocity value. It coincides with the knowledge acquired thanks to the implementation of velocity profiles in the ANSYS program. In the aforementioned measurement diameters, the flow trends to stabilize and the velocity profile begins to take a symmetrical shape with respect to the pipeline axis.

- For measurement diameters below 10D, it is necessary (for example based on computer simulation or analyzing the mathematical equations describing the velocity distribution) determine the optimal angle of the flowmeter heads, because the differences in the velocity between the individual angles are very large.

- The simulation performed in ANSYS program reflects the flow that is being analyzed. Simulation data coincides with the measurement results for diameters 5D, 10D and 15D. Due to the high impact of vortex structures created in the hydraulic elbow on the flow in the 0D cross-section located immediately after the elbow, the results of the simulation significantly differ from the results of measurements in this cross-section. This indicates that, given only this cross-section, the used turbulence model does not well represent the actual flow.

The simulation can be made before the measurements if there is a disturbed velocity distribution (for example through a hydraulic elbow, throttle or other elements of the installation). Considering velocity profiles, we can choose the area in which the flow will stabilize, which will allow to make measurements with a relatively small error for the optimal angle of ultrasonic heads setting. However, for a simulation that should accurately describes the actual flow, a suitable dense numerical grid, a suitably selected turbulence model and boundary conditions in line with real flow are needed.

The advantage of the transit-time ultrasonic flow measurement method is non-invasive and high accuracy. This method finds application in the measurement of pure liquid flow, without the presence of solid particles. It should be remembered, however, that the obtained velocity value is the average flow velocity in the ultrasonic path. The obtained velocity value depends on the angle of ultrasonic heads positioning and velocity distribution in a given measurement diameter. The average flow velocity values in a given measurement diameter can be obtained with greater precision using multipath ultrasonic flowmeter [7, 8].

References

1. Hydrometry - Measurement of the flow rate using the transition time (transit time) of the ultrasound, PN-EN ISO 6416:2018-01 (2018)
2. Moore, P., Brown, G., Simpson, B.: Ultrasonic transit-time flow meters modelled with theoretical velocity profiles methodology. Measurement Sci. Technol. (2000)
3. Iooss, B., Lhuillier, C., Jeanneau, H.: Numerical simulation of transit-time ultrasonic flowmeters: uncertainties due to flow profile and fluid turbulence. Ultrasonics **40**(9), 1009–1015 (2002)
4. Cichoń, M.: Non-invasive Flow measurements, i.e. a Dissertation with Myths - part II. Automatyka i Pomiary No. 1, 11–13 (2015)
5. Qiang, Chen: Weihua, Li, Jiangtao, Wu: Realization of a multipath ultrasonic gas flowmeter based on transit-time technique. Ultrasonics **54**, 285–290 (2014)

6. Bardina, J.E., Huang, P.G., Coakley, T.J.: Turbulence modeling validation, testing, and development, NASA Technical Memorandum (1997)
7. Kurniadi, D., Trisnobudi, A.: A multipath ultrasonic transit time flow meter using a tomography method for gas flow velocity profile measurement. Part. Part. Syst. Charact. **23** (3–4), 330–338 (2006)
8. Rychagov, M., Tereshchenko, S.: Multipath flowrate measurements of symmetric and asymmetric flows. Inverse Probl. **16**(2) (2000)

Marker Method of Determining the Speed of Fluid Movement During the Flow Visualization

Mariusz R. Rzasa[✉] and Beata Czapla-Nielacna

Opole University of Technology, 45-271 Opole, Poland
m.rzasa@po.opole.pl, b.czapla@doktorant.po.edu.pl

Abstract. This paper contains a description of a method applied for determination of local velocities of liquid flow around a vortex generator. The postulated method involves the visualization of von Karman Vortex Street and determination of the local velocities of the liquid. The investigation is based on capturing images of tracing particles introduced into a liquid flow. The study is performed for various exposition times and the traces of the particle motion are registered in the image captures. The analysis of the images leads to the determination of the length of the traces and on this basis the velocity of the liquid flow can be calculated. The presented method yields an accuracy in the range of a few per cent in relation to the measured value. This paper also contain the description of the measurement and design of the experimental setup. The results of the study are presented along with the assessment of the measurement uncertainty.

Keywords: Liquid flow visualization · Liquid flow measurement
Vortex flow meter

1 Introduction

The mathematical description of liquid flow around various types of obstacles forms one of the challenging issues in research. Liquid flow around an obstacle is often accompanied by disturbance, and this process leads to the formation of vortex patterns. The frequency of vortices and their intensity vary depending on the liquid flow rate and the type of the obstacle. This phenomenon is applied in vortex flowmeters [1].

The operating principle of a vortex flowmeter involves the generation of regular vortices. They are formed solely downstream of an obstacle. The shape of vortices also directly affects the sensitivity and measuring range of a flowmeter. The relation between the vortices and flow velocity is expressed by the Strouhal number [1, 2]:

$$St = \frac{f \cdot d}{v} \tag{1}$$

where: St –Strouhal number, f – frequency of generated vortices, d – dimension characteristic to an obstacle, v – mean flow velocity.

The basic advantage of vortex flowmeters is associated with very small pressure drops resulting from their installation. The narrow measurement range forms the basic drawback and it profile of the vortex generator plays the principal role on it [3, 4]. The search for new profiles is associated with the need of assessing the regularity of vortices and the phenomena accompanying them. One of the manners in which this type of flow can be researched involves visualization of the liquid flow.

A variety of visualization methods are applied nowadays in combination with computer image analysis. The first vortex flowmeters were built on the basis of studies involving liquid flow visualizations. In the 15th century Leonardo da Vinci used flow visualization during observations of water flow in a river and drew profiles of the spontaneously forming vortices [5]. In 1878 Vincent Strouhal discovered that that the tone height generated by a wire exposed to wind is proportional to the diameter of the wire and the wind speed. In 1954 Anatol Roshko suggested the application of the phenomenon of von Karman vortex street generation in vortex flowmeters [6]. In 1911 on the basis of observations made by Hiemenz regarding the formation of vortices forming behind a cylinder streamlined by a flow. Theodore von Karman defined the theory of vortex street formation [7]. Despite the fact that this visualization did not provide an adequate level of precision. This approach involved the observation of a large flow region and phenomena accompanying the formation of vortices. For this reason it offered an approach to study that is still common followed [10].

The aim of this work is to present a measurement method allowing to determine the velocity of liquid movement on the basis of the visualization image of the velocity distribution around the vortex generator used in the vortex flow meter. A marker method was used to visualize the motion of the liquid. The work contains a description of the method enabling the determination of local velocities for the liquid flow in the visualization process. The paper presents the results of tests and the selection of appropriate parameters was determined taking into account the measurement errors.

2   Flow Visualization

At present various flow visualization methods are applied in the research. Most experiments are based on the use of a variety of tracer particles seeded in a flow, followed by the observations concerned with the way in which they are carried along with gas or liquid. The use of tracer particles provides a manner which observations can be performed with the purpose of measuring local velocities of gas or liquid flow. Particle Image Velocimetry (PIV) employs a registration method involving the capture of multiple frames and an analysis of optical autocorrelation. Series of images are taken using high-speed recording equipment and a small set of regions of interest is determined by examining correlations [8, 11]. A modification of the PIV is offered by the computational method using Digital Particle Image Velocimetry (DPIV). This approach involves a search for velocity vectors in the liquid flow using digital image correlation. Tracer particles are introduced into liquids, including plant pollen and plastics whose density is similar to liquid so as to enable flow observations and an analysis of the behavior of films formed in the liquid. In this method a very important aspect is related to the selection of the size and density of the tracers. It is necessary to

apply tracer particles so as not to cause disturbance in the liquid flow. Their density needs to be similar to the density of the liquid or their mass needs to be small enough to ensure that they are freely carried along with the liquid.

The visualization of gas flows commonly utilizes tracer particles in the form of clouds of smoke or small liquid droplets such as aerosols. Flow visualization of liquid flow often employs a variety of dyes paints and small gas bubbles, e.g. polymer solid particles, whose density is similar to that of the liquid.

This paper present a method of measurement that is similar to PIV yet it applies another technique of calculating the value of the local velocity. The method developed by the present authors is based on the observation of flow of tracer particles that are introduced into the liquid. However the tracers used in the study have considerably greater dimensions compared to the ones utilized in the PIV method. The calculation of the velocity of the tracer particles carried with the flow involves an analysis of flow images on the basis of just a single image capture. In this case, image correlation involving at least two successive photographs does not need to applied, as it is in the case of the PIV method. The method developed by the authors applies the analysis of the flow of tracer particles that is registered using a long exposure time. For a long exposure time, it is understood that the exposure time is such that the markings in the registered image have marked the trace of the path it has traveled, and not just its shape has been registered.

2.1   Method Utilized for Liquid Flow Measurement

The velocity of the liquid was determined on the basis of images captured with an adequately long exposure time (Fig. 1). The tracer particles carried along a plane that is illuminated by a laser light lead to light reflection in the direction of the installed camera. The marker particles used to measure were black polyamide with dimensions from μm to 1 mm. The long exposure time results of the exposure to light of the subsequent image pixels along with the displacement of the tracers. As a consequence the characteristic trace of the motion can be followed and recorded for a further analysis.

Fig. 1.   Image of tracer particle flow.

The path travelled by a tracer particle forms the basis on which the velocity and liquid flow direction can be calculated. This velocity is derived from the formula:

$$v = \frac{k \cdot p}{t} \tag{2}$$

where: v – velocity of liquid flow, k – scale coefficient, p – number of pixels, t – exposition time.

The scale coefficient was derived on the basis of a capture of a reference length that was located directly in the experimental section (Fig. 2). The scale coefficient is formed by the quotient of the distance obtained from the reference marker and the number of pixels corresponding to this distance.
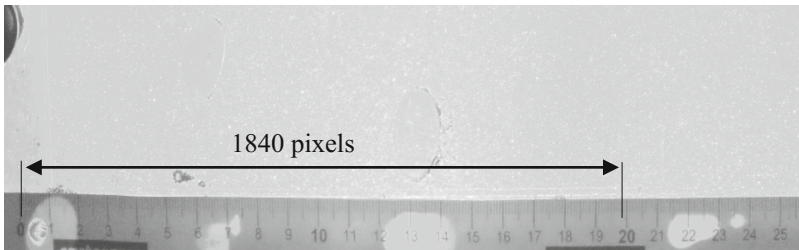
Fig. 2. Paths travelled by tracer particles.

2.2 Experimental Setup

A setup comprising Particle Image Velocimetry [9] system was designed and built with the purpose of testing flow around various vortex generators (Fig. 3).
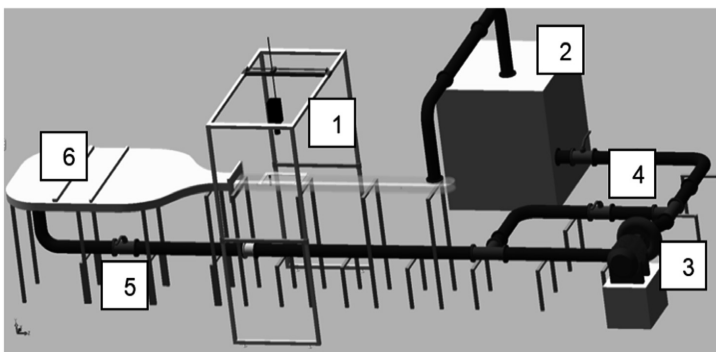
Fig. 3. Experimental setup.

A transparent channel (1) with a rectangular cross-section and dimensions of 50.0×250 mm forms the main element of this setup. The tunnel was fabricated from

acrylic glass and the top cover can be removed so as to install various types of vortex generators. The water feed into the tunnel comes from a liquid vessel (2) with a capacity of 1m3. The liquid flow is generated by a pump (3) with the capacity of 17 m^3 and a maximum lifting height of 18 m. The pumping arrangement is driven by a 3 kW electric motor. The water in the system circulates in a closed system and its flow rate is controlled by a throttling valve (5). The pressure in the system is regulated by a valve (4) that is installed on the pump's bypass. To ensure the uniform liquid velocity at the inlet to the channel a flow straightener is used (6). This role is played by a channel with an adequate profile whose effect results in flow conditioning by increasing the flow velocity along the walls of the channel and eliminating turbulence. A camera was installed above the experimental section to provide the registration of the flow of the indicators. The images were recorded using ILCE-7SM2 digital camera from Sony. It is designed to operate functions such as 120fps Full HD recording. In addition still photos can be taken with the exposition times from 1/8000 s to 30 s.

Fig. 4. Experimental section.

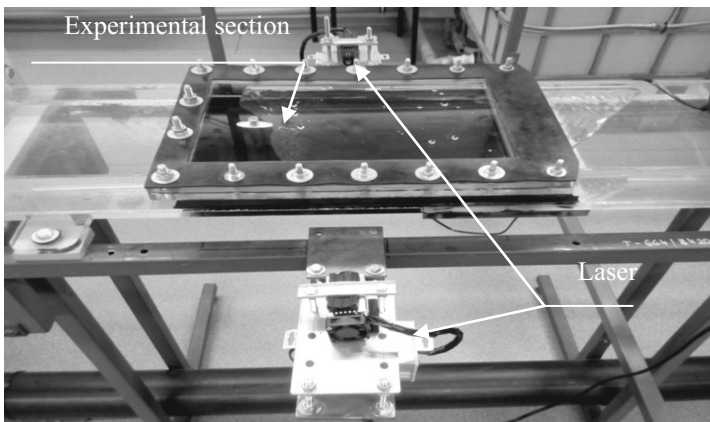The experimental section is formed by a pipe section with a length of 450 mm (Fig. 4). This section is illuminated in the horizontal plane from two directions. The light sources include two linear lasers of 2 W peak power, each one operating at a wavelength of 520 nm. The lasers were installed along the sides of the channel in such a way that the liquid film with a thickness of 1 mm was illuminated in the middle of the channel height.

3   Results and Discussion

The length of the traces left by the particles is relative to the velocity of the flow as well as image exposition times (Fig. 5). For the case of long expositions time and considerable velocity of the tracing particles the length of the trace is often considerable. The use of longer traces affects the accuracy of the measurements yet, it can lead to an

increase of the measurement uncertainty. Liquid flow in a channel forms a three-dimensional process. Beside the flow in the horizontal plane, particles also travel in the vertical direction. For the case when a long path is traveled by a tracing particle throughout a single exposure, there is a risk that it will leave the cross-section that is illuminated by the laser light before the exposure time expires. As a result the trace of the particle in the image will be reduced.

Since the postulated method does not offer the possibility of analyzing all traces of the particles registered during the exposure time, it is also necessary to perform an analysis of the effect of the exposure time on the uncertainty of the measurements. If the exposure time is too long, uncertainty is likely to increase. In contrast too short exposure times result in very short tracks traveled by the particles, in particular for low flow rates. Because the image resolution has a big influence on the measurement error, it is important to choose the right time of exposure. In extreme cases the difference amounting to several pixels can lead to significant measurement error.

(a)   (b)
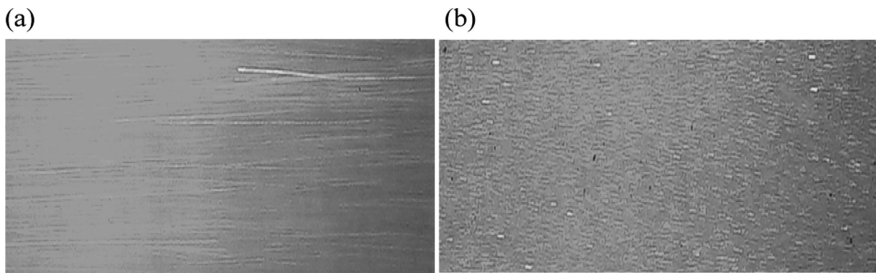
Fig. 5. Images of tracing particle paths for: (a) long exposure time 1/10 s and high velocity of the particles, i.e. 799 cm/s; (b) short exposure time 1/50 s and small velocity of the particles, i.e. 47 cm/s.

Due to the effect of the exposure time on the measurement uncertainty, we decided to determine the recommended exposure times depending on the velocity of the tracing particles. A series of measurements into liquid velocity was undertaken for various settings of the exposure time. The velocity of the tracing particles was adjusted in the range from 10 mm/s to 650 mm/s. The exposure times were set stepwise in the range from 1/5 s to 1/50 s. Standard uncertainty was determined for the particular image series.

$$U = \sqrt{\frac{\sum\limits_{n=1}^{N} (v_i - v_s)^2}{N(N-1)}} \tag{3}$$

where: v_i – flow velocity for i-th trace of the particle, v_s – mean velocity, N – number of measurements.

Figure 6 contains the results of the assessment of the uncertainty relative to the exposure times depending on the velocity of the tracing particle flow. The values of

this uncertainty are expressed in per cent with the purpose of providing an easier analysis of the effect of this velocity on the measurement uncertainty. The values in per cent are related to the mean velocity of the tracing particles.
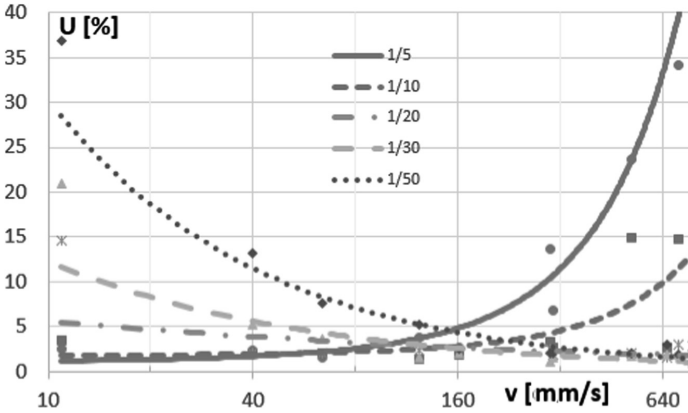
Fig. 6. Dependence of measurement uncertainty on the velocity of tracing particles for various exposition times.

$$U[\%] = \frac{U}{v_s} \cdot 100\% \tag{4}$$

As we can see in the presented characteristics, the exposition times below 1/10 s offer the maintenance of the small values of the measurement uncertainty, i.e. below 100 mm/s. The use of exposition times that are shorter than 1/30 s is recommended for the liquid velocities below 500 mm/s.

Table 1. Recommended exposition times depending on the liquid velocities.

Recommended exposition time (s)	Mean velocity of tracing particles (mm/s)
1/5	<100
1/10	<150
1/20	>100
1/30	>150
1/50	>400c

On the basis of this analysis, it was possible to determine the recommended exposition times for various velocities of tracing particle flow. The values are summarized in Table 1. When we compare the values found in this table, we realize that there are at least two recommended exposition times for the majority of the measurement ranges. The results demonstrate that there is a considerable degree of freedom in the process of selecting the exposition times yet, certain threshold values need to be maintained. The threshold values may include the value of 130 mm/s, for which the

exposition time should not be less than 1/20. The second limit value is 300 mm/s, above which the exposition time should be less than 1/30. In the 130–300 mm/s range, the recommended exposition time is 1/30 to 1/10 mm/s.

4 Conclusion

The measurement method presented in the work can be used to determine the velocity of liquid flow in the visualization processes. The only limitation is the use of transparent liquids. Due to the fact that many transparent liquids exist in technical conditions, its range of use is very large. The presented approach offers the possibility to determine liquid flow velocity with a precision in the range of up to a few per cent in relation to the velocity of the liquid flow. The uncertainty of such measurements is considerably relative to the liquid velocity. An adequate selection of the exposition time forms a vital aspect in ensuring the required accuracy of the measurements; yet it is relative to the velocity of the analyzed phenomenon. The analysis described in this paper provides valuable insights into the selection of exposition times in studies involving visualization of the liquid flow.

References

1. Ford, C.L., Winroth, P.M., Alfredsson, P.H.: Vortex-meter design: the influence of shedding-body geometry on shedding characteristics. Flow Meas. Instrum. **59**, 88–102 (2018)
2. Venugopal, A., Agrawala, A., Prabhu, S.V.: Review on vortex flowmeter—designer perspective. Sens. Actuators A Phys. Indie (2011)
3. Gandhi, B.K., Singh, S.N., Seshadri, V., Singh, J.: Effect of bluff body shape on vortex flow meter performance. Indian J. Eng. Mater. Sci. **11**, 378–384 (2004)
4. Singh, S.N., Seshadri, V., Swaroop, A.: Effect of size and shapeof the bluff body on Strouhal number in pipe flow. In: 20th National Conference on FMFP (1993)
5. Bartoli, G., Borsani, A., Borri, C., Martelli, A., Procino, L., Vezzosi, A.: Leonardo, the wind and the flying sphere. Italy, EACWE 5 (2009)
6. Roshko, A.: On the development of turbulent wakes from vortex streets. NACA Report, No 1191 (1954)
7. Honda, S., Yamasaki, H.: Vortex shedding in a three-dimensional flow through a circular pipe. In: Proceedings of X IMEKO Word Congress, Prague (1985)
8. Hanus, R., Petryka, L., Zych, M.: Velocity measurement of the liquid–solid flow in a vertical pipeline using gamma-ray absorption and weighted cross-correlation. Flow Meas. Instrum. **40**, 58–63 (2014)
9. Zych, M., Hanus, R., Vlasák, P., Jaszczur, M., Petryka, L.: Radiometric methods in the measurement of particle-laden flows. Powder Technol. **318**, 491–500 (2017)
10. Peng, J., Fang, M.: Response of a dual triangulate bluff body vortex flowmeter to oscillatory flow. Flow Meas. Instrum. **35**, 16–27 (2014)
11. Jaszczur, M., Nowak, R., Szmyd, J., Branny, M., Karch, M., Wodziak, W.: An application of SPIV technique to experimental validation of the turbulence model for the air flow in the intersection of the mining face with the ventilation gallery. J. Phys: Conf. Ser. **318**, 052043 (2011)

Measuring Methodology of Conducted Disturbances Generated by Avionic Systems in Aircrafts

Wiesław Sabat[(✉)], Dariusz Klepacki, Kazimierz Kuryło, and Kazimierz Kamuda

Rzeszów Univesity of Technology, W. Pola 2, 35-959 Rzeszów, Poland
{Wsabat,Dklepa,Kkurylo,kazik}@prz.edu.pl

Abstract. The problems connected with measurements of conducted disturbances emission for on-board devices installed in aircraft have been presented in this paper. The electromagnetic environment of aircraft has been characterized. Referring to the requirements determined in Section 21 of the RTCA/DO-160G standard, the categories of avionics systems, acceptable disturbance values for individual categories, construction and requirements for test setup have been described. For the dedicated drive for the unmanned aerial vehicle, the results of the emission measurement were presented using the elaborated scripts in the Rohde&Schwarz EMC32 program.

Keywords: Electromagnetic compatibility (EMC) · Avionics
Disturbances measurements

1 Introduction

Aircrafts, as well as increasingly used unmanned aerial vehicles belong to the class of means of transport, which require 100% reliability of operation in all phases of flight. Electrical and electronic avionic systems currently play a very important role in the safety ensuring of aircraft. The pilot instruments on board, radio navigation devices, air communication systems, assistance systems and automatic flight controls are the basic equipment of every modern air vehicle. For proper operation of the above-mentioned systems (very often placed in close proximity and limited space) it is required to fulfill the EMC standards. The problem of electromagnetic compatibility of on-board systems is solved in the design and starting phase of a new device. The analysis process of conducted and radiated electromagnetic disturbances generated by electrical and electronic devices is subject of strict regulations. The civil aviation standards in the field of electromagnetic compatibility determine the rules and procedures for conducting tests that would allow to obtain unambiguous and repeatable results [1].

Every electrical or electronic system designed for an aircraft must now meet the requirements of the RTCA/DO-160 standard. The essential requirements for equipment and systems used in a civil aircrafts were collected in 26 sections of the standard. Its European version (in direct translation) is included in the EUROCAE - ED14 standard [2]. The fast progress in the field of wireless technologies, the need to consider their

interaction with associated as well as non-associated devices and systems were the reasons of changes in the current version G of the standard, introduced in 2010 [3, 4]. The Sections 1–14 relate to the environmental exposures, Sections 15–25 relate to the electromagnetic compatibility of on-board avionics devices.

2 Permitted Levels of Conducted Emission for Aircraft According to the RTCA/DO160G Standard

An important aspect in the design of on-board avionics systems is the limitation of the level of emission of electromagnetic disturbances generated by the devices to the values specified in the standard and providing them with the required level of immunity. The records included in Section 21 of the RTCA DO-160G standard formulate requirements concerning the method of measuring conducted and radiated emissions. This section defines the test setups, measurement procedure, category of equipment and control - measurement apparatus used in tests.

In the case of equipment used in civil aviation, it is currently divided into 6 categories (B, L, M, H, P, Q), with a distinction between power ports and communication interfaces. The category B is the least restrictive category. It concerns the broad sense of a wide class of on-board electrical and electronic equipment. It is to create an acceptable electromagnetic environment and maintain emissions at a strictly defined level. Referring to power ports and interfaces, it should be note that devices in this category have the highest emissions. Category L has been defined for zones where electronic systems, components and systems are most often installed in the aircraft. Category M refers to equipment that is installed near zones where electromagnetic fields can penetrate through various types of airframe structure (windows, doors), while category H refers to equipment that is installed in the receiving zones of sensitive wireless systems. This category applies in particular to accessories that are installed outside the aircraft. Category Q is a new category that was introduced in the last amendment of the standard and defines the requirements for the equipment used in the VHF bands and GPS navigation systems.

In the on-board aircraft systems there are multiple, complex wiring systems in the form of harnesses – from this reason the measurement with artificial measuring networks would introduce some technical complications. According to the procedural requirements for measuring of conducted emissions at power and communication ports, the current clamps are used in the frequency range from 150 kHz to 152 MHz.

According to the provisions in the standard, it is recommended that such measurement should be carried out in a shielded room or an anechoic chamber. The curves presented in Fig. 1 refer to the emission level of conductive disturbances measured with current clamps for power circuits as well as in Fig. 2 they refer to communication and control interfaces, respectively.
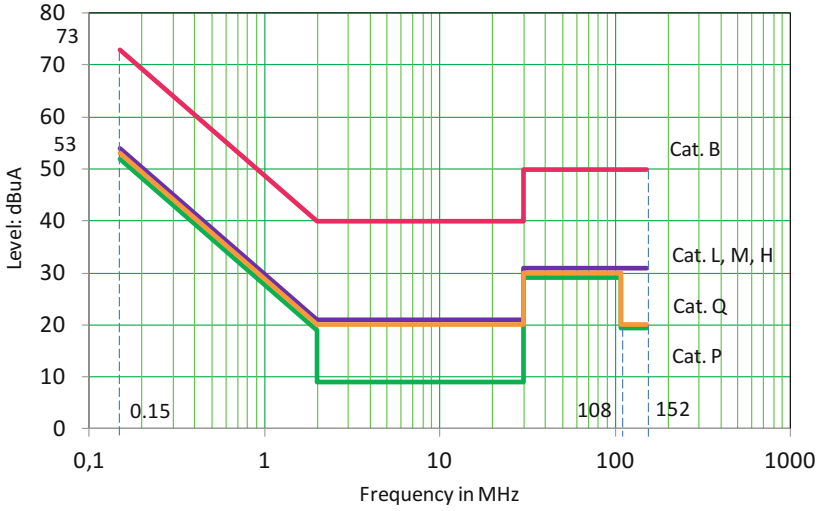
Fig. 1. Permissible conducted emission levels for various categories of equipment in accordance with the requirements of the RTCA DO-160 standard for power ports.
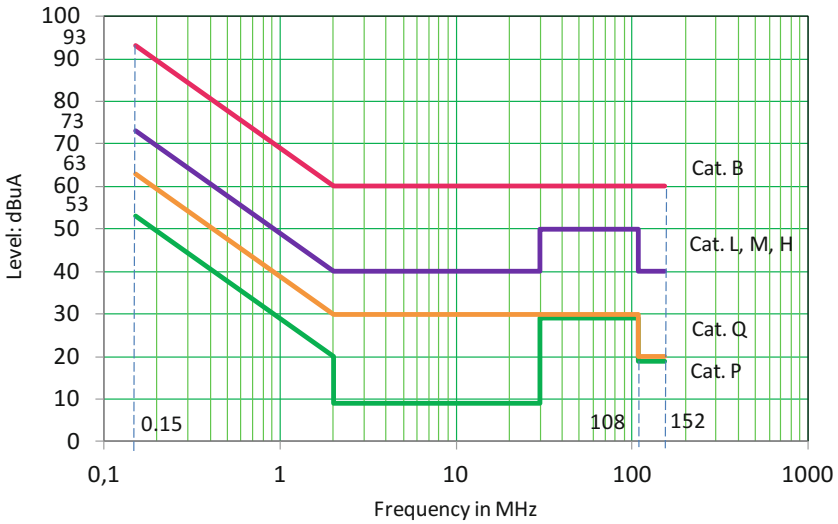
Fig. 2. Permissible conducted emission levels for various categories of equipment in accordance with the requirements of the RTCA DO-160 standard for communication and control ports

3   Test Setup Requirements

The principles for emission measuring of electromagnetic disturbances are strictly determined by procedures specified in the RTCA DO-160G standard. Such requirements are a consequence of the need to obtain representative, repeatable results of measurements at equivalent stands. The measurements for determination of the emission level of disturbances require the use of a certified set of instruments. In terms of measuring of the emission of conductive and radiated disturbances, the necessary minimum equipment of the laboratory is an anechoic or reverberation chamber equipped with appropriate technical infrastructure, a measuring receiver, a set of antennas, current clamps, artificial networks and structured cabling. In order to obtain reliable measurement results, all basic equipment must have appropriate calibration certificates issued by independent notified bodies.

In accordance with the provisions in the standard the tested elements and components are set on a table with normalized geometric size and (depending on the requirements) are placed directly on a table covered with a metal plate. The metal plate is the reference ground and - if not otherwise stated - its surface shall not be less than 2.25 m^2, the width shall not be less than 0.76 m. The brass plate should have a thickness not less than 0.63 mm, copper plate - not less than 0.25 mm and non-ferritic steel plate - not less than 1 mm. The through resistance of the grounding plate should not exceed 0.1 m/cm^2. The resistance between the device casing and the earth plate should not exceed 2.5 mΩ for DC current. The plate should be connected to the metal wall of the screened chamber or to the floor that the transition resistance to this cabin for the DC current should be not more than 2.5 mΩ. Tapes connecting the earth plate with the cabin wall should be located at spaces of no more than 0.9 m. This connection should be made with a continuous metal strip with a length to width ratio of no more than 5:1.

Current clamps in the measurement of conducted emission for power circuits as well as communication interfaces must be placed at a distance of 5 cm from the tested object (Fig. 3). The cabling of both power and communication interfaces must be positioned not less than 0.1 m from the reference ground edge. The power ports of the tested device must be connected indirectly to the power supply via an artificial network. Such a network, connected into the circuit, allows for stabilization of the impedance of the power circuit, and thus for obtaining unanimous results. During measurement, the network measurement ports must be closed with 50 Ω terminators.

The measurement of conducted disturbances is carried out using a specialized instrument, which is a measuring receiver. The receiver must be equipped with detector of peak value. It detects the peak value of the envelope of the modulated intermediate frequency signal and must have the adequate to the standard band selectivity for the RBW filters in the considered frequency ranges (Table 1). The use of the measuring application allows in this case for automatic change of receiver settings and acquisition of data without interrupting the measurement.
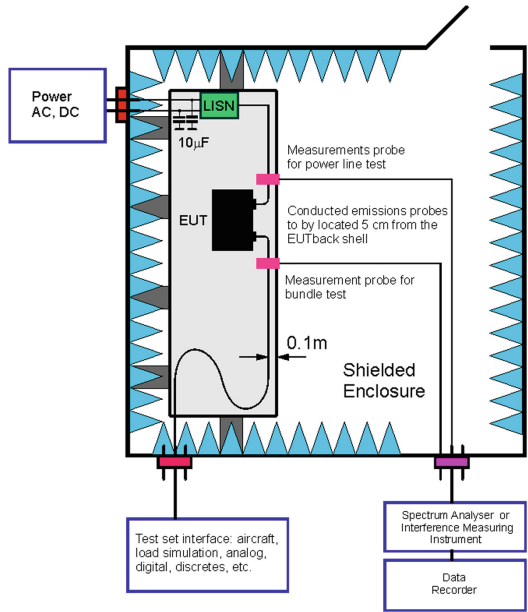
Fig. 3. Stand for measuring the emission of conducted disturbances according to RTCA DO-160G standard.

Table 1. Bandwidth selectivity of the RBW filter for the EMI receiver in the measurement of conducted emission, in accordance with the requirements of the RTCA DO-160 standard [1]

Frequency band (MHz)	RBW (kHz)	Minimum Dwell time (ms)	Minimum measurements time (s/kHz)
0,15 ÷ 30	1	15	0.015
30 ÷ 152	10	15	1.5

4 Idea of the Measurement Script in the R&S EMC32 Software

Rohde&Schwarz's EMC32 software package is a universal software application that allows to build automated measuring scripts to measure electromagnetic disturbances with various transducers and supervise the apparatus for generation of standardized conducted and radiated electromagnetic disturbances. To carry out the measurement with this application, a proper measurement script should be built before the test [5].

In the case of measurement of conducted disturbances, first group of necessary devices is defined in the Device List tab and Hardware Setup is built for a given hardware configuration. After selecting the Hardware Setup tab, a hardware configuration is proposed for the used apparatus set. The person who configures the system has to select elements that are physically defined in the system for each graphic element that are presented as a system component. In the case of measuring the emission of

conducted disturbances, the type of transducer should be chosen which the disturbances will be analyzed. It can be a current clamp or an artificial measuring network (impedance stabilizer). For each of these elements, after declaring it in Device list, in the Correction Tables/Transducer menu, a file with converter conversion factors should be created [6].

The values of such coefficients are provided by the manufacturer together with the calibration certificate. In many cases, the data are delivered on electronic media in the form of files that can be directly copied to the system (Fig. 4).
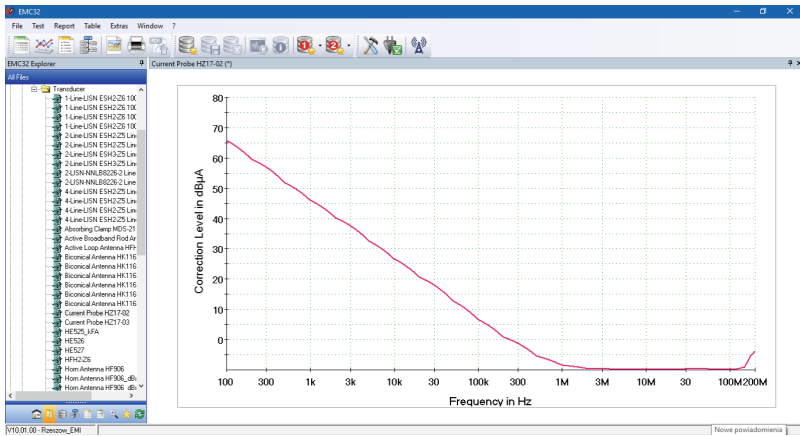
Fig. 4. Calibration curve of the Rohde&Schwarz EZ-17 current probe used during measurements for a model object.

In the next step, the connection paths are defined. Physically, in files for specific signal paths, insertion loss values for individual connections are stored. Files with measured values of attenuation are stored in the *Corrections Tables ‖ Attenuation* folder. Files with attenuation of connecting elements are obtained by measurements using a spectrum analyzer or a vector analyzer. The attenuation of cables and connectors can also be measured using a signal generator, a power probe and an EMC32 program with the *Signal Path ‖ Calibration function*. All connections with HF cables, which are carried out in the laboratory, have strictly defined files with appropriate values of their attenuation. All files can be found in the *Correction Tables ‖ Attenuation* menu [6].

In the last step, the element of the disturbances analyzing system is selected. In this case, it is the ESU26 receiver from Rohde&Schwarz. It allows to measure the electromagnetic disturbances in the frequency range from 20 Hz to 26.5 GHz.

According to the logic of the software, in order to start the configuration process of the measurement template, it is necessary to specify in the General Settings menu the Hardware Setup on which the measurement template will be built (Fig. 5).

The frequency range in which the measurements will be carried out and the minimum and maximum frequency values to be displayed in the graphical window of the
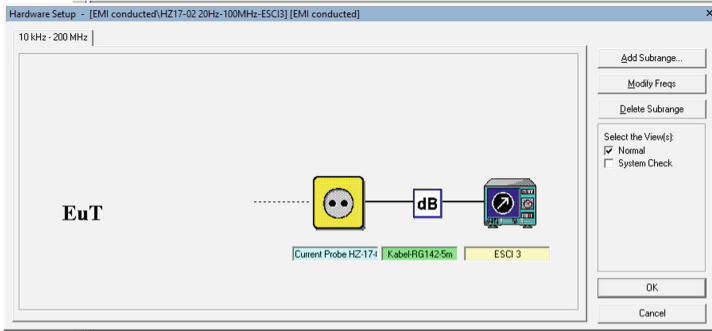
Fig. 5. Hardware setup in the R&S®EMC32 program to measure the emission of conducted disturbances.

analysis are defined in this menu. During defining the range of the analyzed frequencies, the application automatically checks the technical capabilities of the measurement system components in the scope of their measurement band (Fig. 6). In the event of non-compliance, the operator receives an appropriate message.
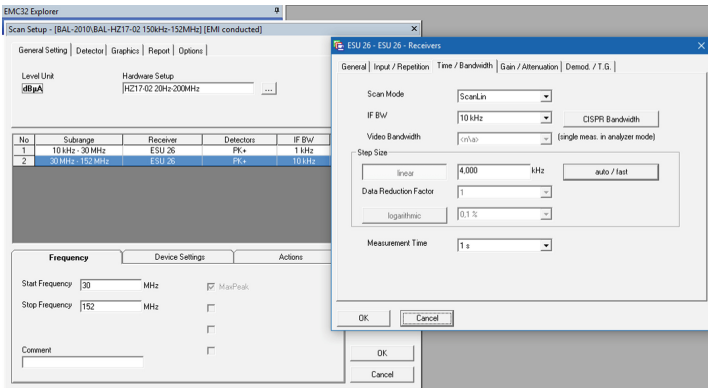
Fig. 6. Rohde&Schwarz's EMC32 configuration windows. Fragment of the script for measuring the emission of conducted disturbances with the use of current clamps.

Each created script is saved in a file with the name given by the operator or automatically by the system. When running a measurement script, it is possible to define the range of analyzed frequencies and limit curves, allowed for particular categories of system components. Exceeding the limit values by the resultant curve presenting the level of emission even at one point (for one frequency) is equal to the negative result for the tested object.

5 Analysis of Disturbances in Accordance with the Guidelines of the RTCA DO-160 Standard for Drive of Unmanned Aerial Vehicles

Practical aspects of the measurement of conducted disturbances using the elaborated script were presented on the example of an electric drive for unmanned aerial vehicle. Such object, moving in space, must meet the requirements of the RTCA DO-160 standard. As part of the project R10 0026 06 "Micro-hybrid drive unit for Unmanned Aerial Vehicle", a brushless electronic-controlled motor with a nominal power of 1.6 kW was developed and dedicated to unmanned aerial vehicles [6–8].

The specificity of the BLDC brushless DC motor is the necessity to use the electronic key set to supply its windings, properly commutating the energy from the DC power source. Depending on the nature of the drive and technical requirements, the set of keys creating the electronic commutator converts energy into voltage and current waveforms of a trapezoidal shape (PMDCBMTC - Permanent Magnet Direct Current Brush-less Motor with Trapez Control) (Fig. 7). The process of energy commutation in the electronic commutator is connected with the process of fast switching of current with significant values.
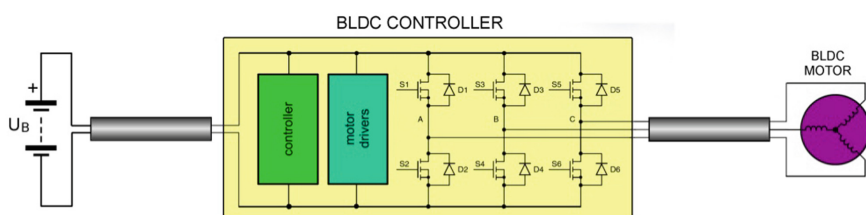
Fig. 7. Block diagram of the drive with a BLDC motor.

In the tested solution, when the system is powered from a LiPo battery (Lithium iron Phosphate battery) with a nominal voltage of 29.6 V (2×14.8 V) and a total capacity of 6.6 Ah, the current in each group of valves reaches its peak value 50 A at nominal engine load (Figs. 8 and 9).

As is shown in the presented current and voltage waveforms at the output of the BLDC motor controller, the process of switching of transistors is the cause of pulse changes in current and voltage waveforms. These short-term, periodic changes - especially in the voltage waveform - are the cause of the generation of additional harmonics in the course of voltage and current, and in the total form of conducted and radiated disturbances.

The tests of conducted disturbance for a model object were made in accordance with the requirements of Section 21 of the DO-160G standard. In the laboratory stand for the analysis of conducted disturbances (in the frequency range from 150 kHz to 152 MHz) the artificial measuring network NNBL8226-2 from Schwarzbeck (Fig. 10),
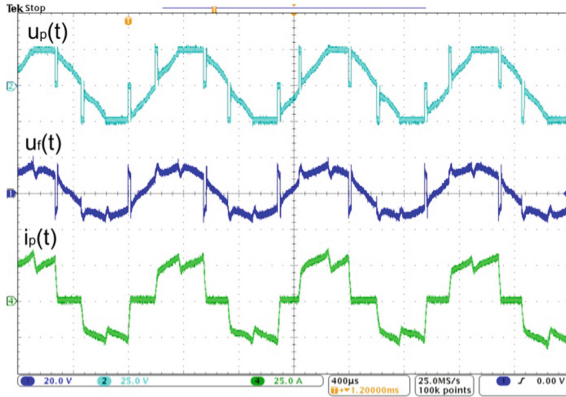
Fig. 8. Measured waveforms of the phase $u_f(t)$ and line $u_p(t)$ voltages and line current for the nominal motor load: $n_n = 8000$ rpm, $P_n = 800$ W, $M_o = 0.95$ Nm on the output of the motor controller (full control state of the controller).
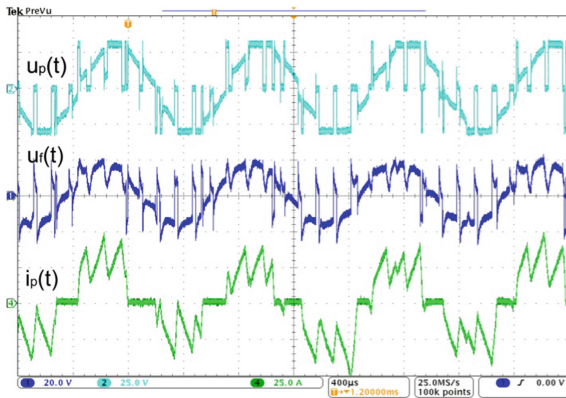
Fig. 9. Measured waveforms of the phase u_f (t) and line u_p (t) voltages and line current for the nominal motor load: $n_n = 8000$ rpm, $P_n = 800$ W, $M_o = 0.95$ Nm on the output of the motor controller, (PWM regulation state).

current clamps EZ17, ESU26 measuring receiver and the R&S®EMC32 program from Rohde&Schwarz were used.

The measurement was carried out in the operating state of the drive, when the highest level of disturbances is generated. This is the case when for a fully charged accumulator batteries the controller adjusted the rotational speed of the unit, switches to the PWM operating state.

In accordance with the guidelines of the DO-160 standard, measurement of the level of emission of conductive disturbances must be carried out in the supply circuits of the tested object (battery - controller) and in load circuits (controller - BLDC motor). Referring to the acceptable levels specified in the standard, in the case of the tested
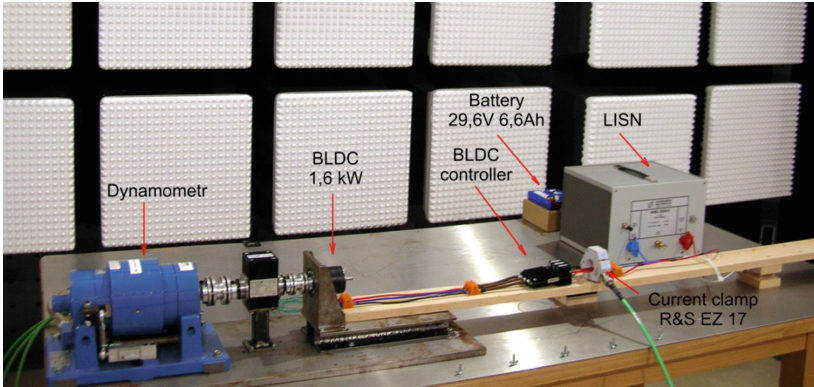
Fig. 10. Laboratory stand for the measurement of conducted disturbances in the frequency range from 150 kHz to 152 MHz. Emission measurement in the driver of power circuit.
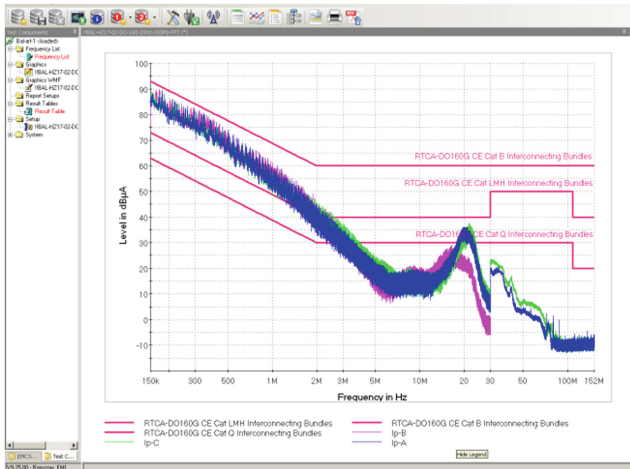
Fig. 11. Measurement results of emission level of conducted disturbances in PWM regulation state for nominal motor load: $P_n = 800$ W, $n_n = 8000$ rpm, $M_o = 0.95$ Nm - BLDC motor for individual phases.

drive, the emission levels will be exceeded if the object is included in the L, M, and H categories (Figs. 11 and 12). In the case of an unmanned aerial vehicle made from composite materials, passing the tested drive to this class of equipment will be legitimate, because sensitive navigation and communication devices may interact due to the proximity of the airframe location, conditioned by geometric dimensions of such object. Category B refers to the equipment that is placed in an adequate distance from sensitive equipment in shielded spaces without apertures.
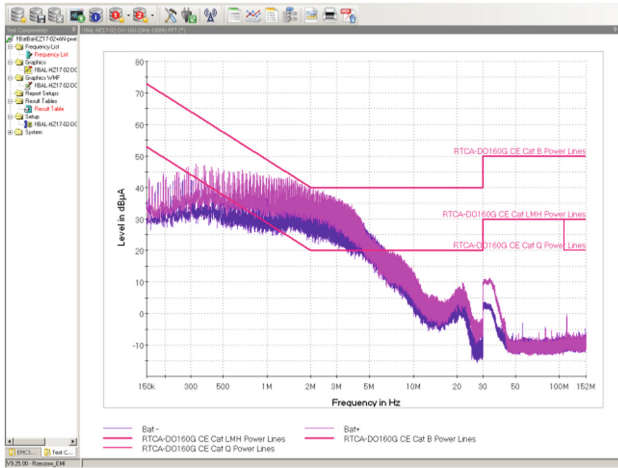
Fig. 12. Measurement results of conducted disturbances level in PWM regulation state for nominal motor load: P_n = 800 W, n_n = 8000 rpm, M_o = 0.95 Nm - baterry - driver circuit for plus and minus poles.

6   Conclusion

A high degree of participation of avionics and electrical components in flight safety in all its phases, makes it necessary to provide these systems with the required level of immunity to the standard types of electromagnetic disturbances characteristic of this environment and to limit their emission to the value specified in the standard. In the case of on-board electrical and electronic devices, these requirements are specified in the applicable RTCA/DO-160G standard from 2010. Records included in Section 21 of the RTCA DO-160G standard formulate requirements for the measurement of conducted and radiated emissions. This section defines the requirements relating to the measurement stand, measurement procedure, equipment category and control and measurement equipment used for the testing. As this paper presents, for obtaining the unambiguous and reproducible results the appropriate laboratory stand is required. Control of emissions of electrical and electronic on-board devices allows creating a friendly electromagnetic environment in the limited space of the aircraft. As it is shown at the beginning of the paper, in the modern aircraft are devices and systems where the difference between transmitted and received signals is several dozen decibels. As was shown on example of the drive for unmanned aerial vehicle, every device where an intentional or unintentional state of energy commutation the generation process of electromagnetic disturbances have to take into account.

References

1. RTCA DO-160, Environmental Conditions and Test Procedures for Airborne Equipment, Published by: Radio Technical Commission for Aeronautics, December (2010)
2. EUROCAE ED 14, Environmental Conditions and Test Procedures for Airborne Equipment, Published By: European Organization for Civil Aviation Equipment (EUROCAE), January 2015
3. Advisory Circular 21-16G, RTCA Document DO-160 versions D, E, F, and G, Environmental Conditions and Test Procedures for Airborne Equipment, June 2011
4. Borgstrom, E.J.: New EMC Requirements for Commercial Avionics: RTCA/DO-160G; Interference Technology, May 2011
5. R&S® EMC32 V10.30.10, Online help for EMC, AMS, WMS, Rohde&Schwarz GmbH&Co.KG, Munich, Germany, https://www.rohde-schwarz.com/us/product/emc32-productstartpage_63493-8674.html, last accessed 2018/05/05
6. Report "Analysis of the conditions of emission of electromagnetic disturbances in the control system of an unmanned aerial vehicle" Project: Hybrid drive unit for unmanned aerial vehicle (in Polish), Narodowe Centrum Badań i Rozwoju, PBR VI, N R10 0026, pp. 1–42 June 2009
7. Sabat, W., Klepacki, D.: Mechanism Analysis of Generation And Propagation of Electromagnetic Interferences (Emi) in Unmanned Aerial Vehicle; Microelectronic Materials And Technologies, Monograph No. 232, Koszalin Technical University Monograph Series, pp. 353–360, Koszalin (2012)
8. Sabat, W., Kuryło, K., Klepacki, D., Kamuda, K.: The Measurement Methods of Electromagnetic Disturbance Generated by Aircraft Avionics Systems (in Polish); Poznan University of Technology, Academic Journals Electrical Engineering z.89, 135–145 (2017)

The Concept of Measuring Luminous Flux Distribution Emitted from Sports Facilities Using Unmanned Aerial Vehicles

Magdalena Sielachowska(✉), Damian Tyniecki,
and Maciej Zajkowski

Bialystok University of Technology, Wiejska 45D Street,
15-001 Bialystok, Poland
{m.sielachowska,d.tyniecki}@doktoranci.pb.edu.pl,
m.zajkowski@pb.edu.pl

Abstract. The dynamic development of cities is associated with a significant increase in the luminance of architectural buildings, advertisements or usable spaces. Objects that emit a significant amount of luminous flux to the environment are also various types of sports buildings, such as city stadiums, football pitches or tennis courts. The need to protect the environment determinates the need to reduce and prevent the effect of light pollution. The article analyzes the possibilities of using unmanned aerial vehicles to assess light pollution by measuring the distribution of luminous flux emitted by sports facilities. The method of light measurement based on a goniometric system using unmanned aerial vehicles was presented and their positioning in three-dimensional space was taken into account. A method for controlling the photometric probe and calibrating the measurement system has been proposed.

Keywords: Light pollution · Luminous flux · Unmanned aerial vehicles

1 Introduction

The topics related to light pollution have been very popular lately. Excessive light beam distribution into the upper half-space affects not only the way of seeing and receiving various types of objects, but above all disturbs the functioning of living organisms, such as plants, animals or humans. The unnaturally lit night sky causes changes in the behavior of fauna, as well as significantly affects the state of human health.

The problem of light pollution appears mainly in the areas of large urban agglomerations, where there is external utility lighting, backlit and translucent backlights, large format LED screens and the illumination of buildings.

One of the groups of objects that emits a significant amount of luminous flux into the atmosphere are sports facilities, including football stadiums, tennis courts, ski slopes, etc. Sports complexes from year to year are becoming more and more spectacular also in terms of lighting, which is usually caused by the need to conduct high-resolution television broadcasts of various types of games and competitions.

A significant level of light emanating from, among others, football stadiums, illuminates adjacent areas, even several hundred meters away. It is worth paying attention to the possibility of limiting unnecessary and harmful light distribution in the immediate surroundings, especially in the upper half-space. In order to be able to develop the correct method of reducing the light escaping towards the sky, the directions of propagation of the luminous flux from the tested object should be examined and characterized. Creating an accurate image of light distribution from a sports facility or from other large-size illuminated areas will allow to assess the degree of light pollution. In addition, it is possible to use such knowledge to assess the impact of luminous flux on objects and the light polluted sky on illuminating adjacent areas and roads, and consequently also on reducing electricity consumption.

2 Standards and Legal Acts

Lighting in sport is characterized by specific requirements. The proposed installation must ensure the safety of practicing a given discipline, meet the size of the facility and ensure the comfort of watching the game and its coverage.

Detailed recommendations regarding the design of sports lighting systems can be found in the documents published by Illuminating Engineering Society, among others in the IES RP-6-15 Sports and Recreational Area Lighting standard [1]. The RP-6-15 report focuses primarily on lighting requirements, defined for two categories: players and spectators, divided according to the class of the place of practicing a given discipline (Table 1).

Table 1. Class of play [1].

Facility	Class			
	I[a]	II[b]	III[c]	IV[d]
Professional	X			
College	X	X		
Semi-professional	X	X		
Sports clubs	X	X	X	
Amateaur leagues		X	X	X
High school		X	X	X
Training facilities			X	X
Elementary school				X
Recreational event				X
Social event				X

[a]Class I: Facilities with spectator capacity over 5000
[b]Class II: Facilities with spectator capacity under 5000
[c]Class III: Facilities with some provision for spectators
[d]Class IV: Facilities with limited or no provision for spectators

Apart from the regulations regarding the selection of lighting, the required levels of luminance, and illuminance, the document also brings up the topic related to light pollution. The specification indicates the need to minimize the light distribution beyond the area of sports facilities, which may obstruct the observation of the stars and interfere with the functioning of living organisms in the vicinity.

Currently, the binding standard for lighting sports facilities is PN-EN 12193: 2008 Light and lighting - Lighting in sport [2], which defines the requirements in terms of illuminance, uniformity, luminance and methods of measuring these quantities, inside and outside the premises, during the most popular European sports.

The requirements indicated in the standard [2] specify the maximum permissible value of interfering light, coming from external lighting installations divided into four environmental zones (Table 2):

- E1 – intrinsically dark areas (national parks, protected sites),
- E2 – low district brightness areas (industrial and rural areas),
- E3 – medium district brightness areas (industrial and residential suburbs),
- E4 – high district brightness areas (town centers and commercial areas).

The percentage of the luminous flux that is radiated above the horizon is referred as the ULR. It is defined as the percentage share of the luminous flux emitted by the luminaire or lighting system in the upper half-space in relation to the installed luminous flux.

Table 2. Maximum obtrusive light permitted for exterior lighting installations [2].

Environmental Zone	Light on properties		Luminaire intensity		Upward light
	E_V [lx]		I [cd]		ULR [%]
	Pre-curfew[a]	Post-curfew	Pre-curfew	Post-curfew	
E1	2	0	2500	0	0
E2	5	1	7500	500	5
E3	10	2	10000	1000	15
E4	25	5	25 000	2500	25

[a]In case no curfew regulations are available, the higher values shall not be exceeded and the lower values should be taken as preferable limits.

Professional sports organizations such as FIFA, FIH and NCAA create their own specifications and norms. The International Football Federation FIFA, in 2007 updated the set of technical recommendations and requirements for football stadiums, in which it stressed the need to limit artificial light distribution outside the sports facility.

Regulations, presented in the Football Stadiums Technical Recommendations and Requirements [3], are the basis for the design of a lighting system at football stadiums. Table 3 presents the required values of vertical and horizontal illuminance on the football pitch for events broadcast in the media. For the highest class V, which refers to

events on the international arena, the vertical illuminance values reach over 2,400 lux, and horizontal - minimum 3,500 lux. These are much higher numbers than in the case of PN-EN 12193: 2008 [2].

Table 3. Lighting specifications for televised events [3].

		Vertical illuminance			Horizontal illuminance			Properties of lamps	
		E$_V$ cam ave	Uniformity		Eh ave	Uniformity		Colour temperature	Colour rendering
Class	Calculation towards	Lux	U1	U2	Lux	U1	U2	Tk	Ra
Class V International	Fixed camera	2400	0,5	0,7	3 500	0,6	0,8	>4000	>65
	Field camera (at pitch level)	1800	0,4	0,65					
Class IV National	Fixed camera	2000	0,5	0,65	2 500	0,6	0,8	>4000	>65
	Field camera (at pitch level)	1400	0,35	0,6					

[a]All illuminance values indicated are maintained values. A maintenance factor of 0.7 is recommended; therefore initial values will be approximately 1.4 times those indicated above.

FIFA regulations also pay attention to the issue of light pollution. The report [3] divides the undesirable emission of light into the atmosphere into two categories: lighting sent outside the stadium, but located in its vicinity and reflections appearing on adjacent streets, dazzling drivers and pedestrians.

Table 4. Environmental impact [3].

Angle of Illumination	Distance from stadium perimeter	
Horizontal spill	50 m from stadium perimeter	25 lx
Horizontal spill	200 m further	10 lx
Maximum vertical	50 m from stadium perimeter	40 lx
Maximum vertical	200 m from stadium perimeter	20 lx

According to current norms and regulations, new design specifications should take into account the possibility of limiting the penetration of light outside sports facilities and glare inside, as this is crucial for the safety of road users and the comfort of life of people in the vicinity of this type of complexes. The lighting requirements of the sports objects presented in Table 3 are very high, especially in the context of obtaining the intensity distribution in the vertical plane, which leads to the necessity of positioning the lighting fittings so that a part of the outgoing luminous flux will always be emitted into the upper half space. There will therefore be light pollution, which should be minimized to the levels of illumination described in Table 4.

3   Assessment of Light Pollution

The emergence of new legal regulations concerning light pollution and all kinds of updates of existing norms with additional light pollution requirements force designers and investors to verify the criteria set out in the regulations. However, the correct assessment of the degree of pollution is still not the simplest, among others due to the fact that the regulations and standards in force do not explicitly determine the measurement methodology.

In the literature we can find information related to the methods of assessing the pollution of the night sky with artificial light. Most of them relate to visual assessment or measurement of the luminance level of the sky. In addition, various types of simulation programs are increasingly used to theoretically determine the degree of light pollution.

3.1   Observation Methods

Estimating the brightness of the night sky, consisting primarily of observing and assessing the visibility of individual astronomical objects, is one of the simplest methods of assessing light pollution. Depending on the degree of brightness and visibility of the constellations, the skies can be divided according to the six-level Berry scale [4] or the nine-degree Bortle scale [5]. Both scales are based on the magnitude of the star and represent the sky from the perfectly dark, when the constellations are clearly visible, and the pollution with artificial light does not occur, to the heavily polluted night sky.

The division of the night sky according to the Bortle scale looks as follows [5]:

- Class 1. Excellent dark-sky site – the zodiacal light is all visible, but no details are visible on the surface of the earth. A clearly visible atmospheric glow. The visual range of the weakest stars is 7.6–8.0 mag.
- Class 2. Typical truly dark site – the atmospheric glow is hardly visible, the zodiacal light casts clear shadows at dawn and at dusk. The visual range of the weakest stars is 7.1–7.5 mag. The environment is poorly visible.
- Class 3. Rural sky – zodiacal light clearly visible in spring and autumn. The visual range of the weakest stars is 6.6–7.0 mag. Poorly visible close surroundings.
- Class 4. Rural/suburban transition – over the horizon you can see the glows from the surrounding villages. The Milky Way is visible, but without a clear structure. The visual range of the weakest stars is 6.1–6.5 mag. The surroundings are clearly visible.
- Class 5. Suburban sky – Zodiacal light visible only partly on the darkest nights in spring and autumn. Light sources are visible in all directions. The visual range of the weakest stars is 5.6–6.0 mag.

- Class 6. Bright suburban sky – the zodiacal light is invisible, the Milky Way is only partially visible near the zenith. The visual range of the weakest stars is 5.1–5.5 mag.
- Class 7. Suburban/urban transition – the whole sky has a gray-white shade. Strong light sources are visible in all directions. The visual range of the weakest stars is 4.6–5.0 mag.
- Class 8. City sky – the sky shines with white or orange light. Most constellations are unrecognizable. The visual range of the weakest stars is 4.1–4.5 mag.
- Class 9. Inner-city sky – the whole sky is brightly lit, even at the zenith. Only the Moon, the planets and the brightest stars are actually visible from heavenly objects. The visual range of the weakest stars is at best 4.0 mag.

The above observation methods are useful for simple, amateur studies of the night sky, but they can't be the basis for a professional description of the degree of sky pollution with artificial light. They are a subjective assessment of the observer, which is influenced by factors such as age and visual quality.

3.2 Measurement Methods

The correct objective assessment of light pollution should be based on the measurement method by means of which the brightness of the sky can be described in an unambiguous and reproducible way. For this purpose, generally available photometric meters such as luxmeters, spectrophotometers or luminance meters are used.

The device available for measuring the brightness of the sky is a photometric meter Sky Quality Meter. With the SQM device you can get in just a few seconds the information about the amount of light that reaches the meter from a specific surface of the sky. The results, expressed in magnitude per second square arc, are obtained after aiming the entrance window on the selected sky area and triggering the start button [6]. In the article "Night sky photometry with Sky Quality Meter", Pierantonio Cinzano [7] has made a thorough photometric analysis based on synthetic photometry and laboratory tests and compared the SQM device with other systems used in light pollution measurements, such as Johnson's B and V bands, CIE photopic and CIE scotopic responses for typical spectra and the spectral mismatch. Figure 1 compares the SQM normalized response (dotted line) and the standard normalized responses of Johnson's B band, CIE scotopic, Johnson's V band and CIE photopic (dashed lines from left to right).

The results obtained by the authors turned out to be satisfactory, and the inaccuracy of the measurements was about 0.1–0.4 mag/arcsek^2. According to the manufacturer's data, measurement inaccuracies should oscillate within $\pm10\%$, or 0.1 mag/arcsek^2. An undoubted advantage of the Sky Quality Meter photometer is its size, price and simple, intuitive operation. All these properties make the device widely used by scientists, researchers, astronomers, but also amateurs.
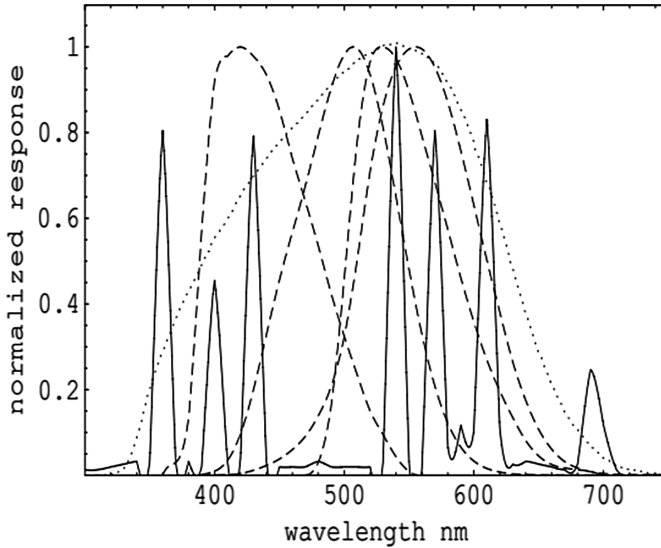
Fig. 1. SQM normalized response (dotted line), standard normalized responses of Johnson's B band, CIE scotopic, Johnson's V band and CIE photopic (dashed lines from left to right) and emission spectra of an HPL mercury vapour lamp (solid line) [7].

Similar opportunities are offered to Apple smartphone users, the Dark Sky Meter application. The cameras of the latest iPhone models (iPhone 4S and newer) are sensitive enough to read the values of the brightness of the sky already at 21 mag/arcsek^2, with a practical limit of around 28 mag/arcsek^2 [8]. Declared accuracy of measurements is ±0.2 m/arc sec, and typical absolute calibration differences from 20% to 30%, depending on the phone model. Additionally, using the application, the angle of inclination of the device, moon phase and weather conditions (cloudiness) and GPS location are recorded. The creators, also tempted to set up a public database - Globe at Night, which provides measurement data, sent by users from around the world.

Another device used to measure the luminance of the night sky is IYA lightmeter (International Year of Astronomy 2009). The IYA light-meter, shows on Fig. 2, is designed for long-term monitoring of light pollution. The device uses photoelectric current measurements of a solar cell, and then converts them to illuminance and irradiance values. Due to its construction, the photometer can be used in various weather conditions, it is fully waterproof, requires no maintenance, operates independently of the season and temperature, and has a high sampling rate of 1 Hz [9].

The IYA lightmeter is characterized by a sensitivity of around 10–5 lux to over 200,000 lux with a resolution of 1%, thus it is able to cover the full range of human perception of light with a single sensor [8].

The degree of light pollution can also be estimated using commonly used luminance and luxmeters. Matrix luminance meters, such as LMK Mobile Air, record several million luminance measurement points in one exposure. After the lens shutter is triggered, a photo is taken, which is then converted to the appropriate format. In a

Fig. 2. The IYA Lightmeter device [10].

dedicated computer program, the result is displayed in the form of a map of the luminance distribution of the recorded image. Spot meters require a much larger number of measurements, as they are able to indicate the luminance value only for one selected point. Instead, they enable an accurate result to be obtained at a given measurement point. The same applies to classic luxmeters. After a series of measurements, it is possible to obtain an accurate distribution of light intensity on a given surface. What's more, luxmeters belong to a group of very popular and widely used photometers. They are therefore widely available at various prices, sizes and different parameters. This means that each user can adjust the meter ideally to the equipment they have and the research they intend to carry out.

4   The Concept of the Measurement Method

Measurement methods, described in the third chapter, are based on the assessment of the level of light pollution with an artificial night sky in a given area. However, none of the presented procedures can be used to determine the quantitative pollution from a selected object or light source.

The author's measurement method presented in this article will allow for precise determination of the degree of light pollution that is generated by individual architectural buildings, sports facilities or selected lighting fittings. This will allow for unambiguous verification of parameters included in standards and technical specifications. It will enable, among others, assessment of the ULR coefficient, which determines the percentage of luminous flux escaping into the upper half-space (Table 2). First of all, however, it will allow to assess light pollution from sports facilities, in this case football stadiums. The proposed measurement method will help to check parameters such as unwanted light emission to the atmosphere (Table 4).

4.1   Spatial Distribution of Luminous Flux

In order to assess the light pollution coming from the analyzed sports object, it is necessary to measure the spatial distribution of the luminous flux. For this purpose, it was assumed that the analyzed football stadium is a lighting fitting, for which the basic way of presenting photometric properties is to present its luminous intensity distribution.

The photometric shape was treated as a closed surface, formed by the ends of sections with a common beginning in the luminous center of the luminaire, whose length and spatial direction corresponds to the value of candlepower [cd] in this direction [11].

In order to present the light distribution of the light-optical system, photometric measurements of the lighting fitting are usually made or simulation calculations are carried out. In the case of measurements, it is necessary to define the correct geometry of the measurement system. The key parameter is the distance of the photometer from the tested luminaire, which affects the measurement uncertainty.

The error resulting from the inverse square law is less than 1% if the distance r from the light source is greater than the fivefold largest dimension d of the light source. Assuming that the roof opening of the stadium is a lighting fitting with dimensions of 117×81 m, the largest dimension d is a roof diagonal with a length of 142 m. An error value below 1% can be obtained for distances r = 710 m (Fig. 3).
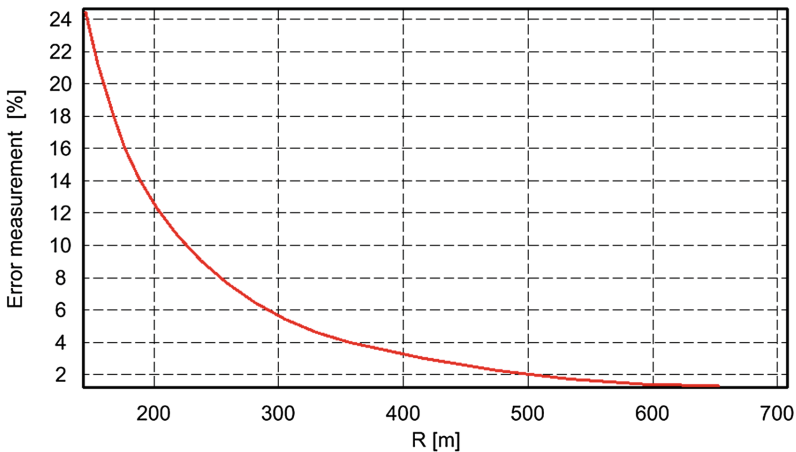
Fig. 3. Light error measurement plot as a function of distance for the analyzed case.

The distance of more than half a kilometer from the examined object seems to be too far in this case, taking into account the presence of various types of obstacles, trees and buildings. What's more, none of the norms provides admissible values for such remote areas. The authors accepted 200 m as a sufficient measuring distance, mainly due to the requirements appearing in the report [3] (Table 4), in which the maximum illuminance values are given for 50 and 200 m.

4.2 The Concept of the Measurement System

The proposed concept of the luminous flux measurement system of a sports facility is based on the classical photometric method and is shown in Fig. 4. It uses a photometric head moving relative to a stationary radiation emitter that always faces the direction defined by the angles C-γ.
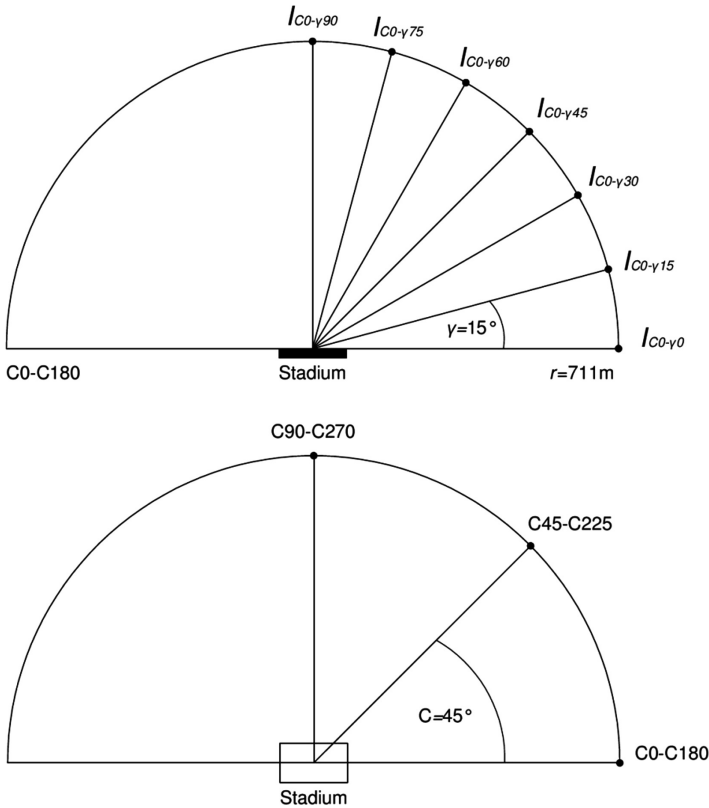
Fig. 4. The concept of measuring geometry of the light distribution from the stadium in the vertical and horizontal plane.

Due to the distance from which light measurements should be made, it is necessary to ensure that the photometer is moved after a radius of at least 200 m (depending on the size of the stadium). Therefore, it was proposed to carry out such measurements using an unmanned aerial vehicle (UAV). A photometer should be installed on the drone, which is usually a classic lux meter.

In addition, the photometer should be mounted on the so-called gimbal, i.e. a device with automatic positioning of the object mounted on it and allowing the measuring head to be directed towards the test object. Such solutions are characterized by the accuracy of positioning resulting from the resolution of positioning of stepper motors

that are used to drive the gimbal. Currently available solutions allow for setting the stepping motor shaft position with a resolution of more than 1°, and the repeatability of the setting is realized with incremental sensors. In photometric measurements such positioning accuracy is acceptable and sufficient in assessing the intensity distribution surface of the light-optical system.

The way to determine the position in the UAV space (drone with a photometer) is to use the GNSS receiver (Global Navigation Satellite Systems). In order to increase reliability and position accuracy, GNSS two (GPS + GALILEO) or three-system (GLONASS) receivers are used. In the literature [13] there are studies on the accuracy of positioning objects using GNSS systems. The errors resulting from accuracy coefficients (DOP dilution of precision [14]) should be analyzed (Fig. 5):

- GDOP geometrical dilution of precision,
- HDOP horizontal dilution of precision,
- VDOP vertical dilution of precision,
- TDOP time dilution of precision.
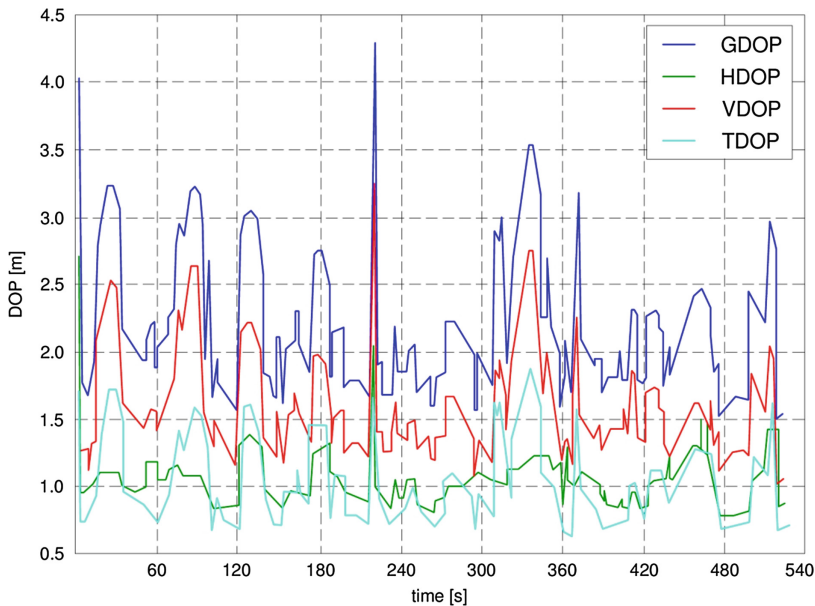
Fig. 5. Blur of the accuracy of geometric coefficients [13].

Measurement of luminous intensity should take place after determining a specific angular position of the photometer, which affects the location of the drone over the object and its stabilization. An important parameter when performing this type of measurements is also the time of measurement, and thus the time of determining the vertical and horizontal positions. On the basis of the geometric position deflection waveforms (Fig. 5) it appears that the weighting of the positioning accuracy of the

drone in the horizontal plane is maximally 1.75 m, while in the vertical plane it reaches 3.2 m. In the case of using a flying object with a photometer, it is expected that its angular positioning is inaccurate with respect to the sport object under test at the level of $\pm 0.141°$ in the C plane, and the positioning imprecision in the γ angle range will be ± 0.258.

5   Summary

The problem of light pollution is usually considered in the calculation of simulations and does not take into account the actual conditions that may arise, for example, from the imprecise placement and directing of luminaire installed, for example, at a football stadium. To realize the measurement, it is necessary to use a photometer that will measure the light. The authors propose the use of an unmanned aerial vehicle for measurements of light pollution, i.e. the assessment of the luminous flux emitted in the upper half-space (towards the sky). Errors resulting from UAV positioning in the indicated measurement geometry are typical for precision photogoniometric systems, where positioning accuracy and repeatability of $\pm 0.5°$ is acceptable. Due to the lack of reliable information from UAV manufacturers and equipment in the form of "gimbals" about their data related to positioning accuracy or repeatability, the assessment was made on the basis of known dependences and properties of the components used in them. It should be expected that the proposed method of photometric measurements can be developed not only to assess the degree of light pollution but also, for example, to audit functioning lighting systems in an open space or large-scale industrial facilities.

References

1. IES RP-6-15 Sports and Recreational Area Lighting, Illuminating Engineering Society, New York (2015)
2. PN-EN 12193:2008 Światło i oświetlenie - Oświetlenie w sporcie, PKN, Warszawa (2008)
3. FIFA: Football Stadiums Technical Recommendations and Requirements, Switzerland (2007)
4. Ściężor, T.: Almanach Astronomiczny na rok 2018. Polskie Towarzystwo Astronomiczne, Warszawa (2017)
5. Bortle, J.E.: Introducing the Bortle Dark-Sky Scale. Sky & Telescope, pp. 126–129, February 2001
6. http://www.unihedron.com/projects/darksky/Instruction_sheet.pdf. Accessed 30 June 2018
7. Cinzano, P.: Night Sky Photometry with Sky Quality Meter, ISTIL Internal Report n. 9, v.1.4 Thiene (2005)
8. Hanel, A., Posch, T., Ribas, S.J., Aube, M., Duriscoe, D., Jechow, A., Kollath, Z., Lolkema, D.E., Moore, C., Schmidt, N., Spoelstra, H., Wuchterl, G., Kyba, C.C.M.: Measuring night sky brightness: methods and challenges. J. Quantitative Spectrosc. Radiat. Transf. (2017)

9. Muller, A., Wuchterl, G., Sarazin, M.: Measuring the night sky brightness with the lightmeter. Revista Mexicana de Astronoma y Astrofisica **41**, 46–49 (2011)
10. http://hms.sternhell.at/lightwiki/images/0/0c/Lightmeter_1Euro_persp.png. Accessed 02 July 2018
11. Żagan, W.: Podstawy techniki świetlnej. Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa (2005)
12. Tabaka, P., Fryc, I.: Zależność poziomu zanieczyszczenia otoczenia światłem od kształtu krzywej światłości użytkowanej oprawy oświetleniowej. Kosmos problemy nauk biologicznych. T:64, **4**(309), 669–677 (2015)
13. Kowalik, R., Bieńczak, R., Komorek, A.: Wyznaczanie dokładności pozycji obiektów ruchomych w przestrzeni 3d. Autobusy **12**, 1038–1042 (2016)
14. Brandford, W., Parkinson. James. J. Spilker Jr.: Global Positioning System: Theory and Applications, vol. 1, American Institute of Aeronautics and Astronautics, Inc. Washington (1996)

Estimation of the Angular Positioning Inaccuracy in Avionics Helmet-Mounted Cueing Systems with Magnetic Method

Andrzej Szelmanowski, Mariusz Zieja, Andrzej Pazur[✉],
and Paweł Janik

Air Force Institute of Technology, Warsaw, Poland
poczta@itwl.pl

Abstract. The paper presents methods for estimating the inaccuracy of determining the angular position of the pilot's helmet dedicated to the NSC-1 Orion helmet-mounted cueing system with a magnetic method, intended for the W-3PL Głuszec multi-purpose helicopter with integrated avionics system ZSA. A flat coil with two separated rings, which generates a magnetic field with a curved profile from the axis of symmetry, developed in the Polish Air Force Institute of Technology (AFIT), was presented. The magnetic field curvature causing errors of determining the angular position of the pilot's helmet in relation to the aircraft cabin reference frame was described with the use of relationships formulated for the flat coil with two separated rings (inner and outer). The inaccuracies determined on the basis of computer simulations of the developed mathematical relationships were compared with experimental data obtained from the magnetic field measurements using the integrated ADIS 16405 triaxial sensor (with a measurement resolution of 0.5 mgauss/LSB). In order to improve the accuracy of determining the angular position, an original algorithm for specifying linear deviations of the pilot's helmet from its neutral position, determined during the adjustment process of the helmet (located on the axis of symmetry of the flat coil) was used. It allowed to determine the current curvature of the generated magnetic field and to make corrections.

Keywords: Avionics · Helmet-mounted cueing systems
Errors of the angular positioning · Magnetic field measurement
Flat coil profile modeling

1 Introduction

One of the most technologically modern and also the most important on-board systems of the contemporary multi-purpose military aircraft and helicopters is a helmet-mounted cueing system also called a helmet-mounted targeting system [2, 4, 19, 21]. The task of such a system includes the helmet-mounted indicating and tracking of a target, as well as helmet-mounted imaging of piloting-navigational and observation-targeting parameters. Such a system allows to guide the on-board weaponry systems with the use of the pilot's head movement (among others, including a moveable

shooting position with an on-board machine rifle or a gun, coordinators of infrared homing missiles as well as observation and targeting heads).

One of the methods used in the aviation helmet-mounted cueing systems is a magnetic method [17]. An example of the system with such a method is the JHMCS system built in F-16 multi-purpose aircraft, operated, among others, in the Polish Armed Forces Aviation. It uses an on-board magnetic field transmitter, built in the aircraft cabin, with a profile that allows to determine the angular position of the pilot's helmet in the system of coordinates associated with the aircraft cabin. The JHMCS system represents a new generation of the helmet-mounted cueing systems, which is currently expanded to the so-called virtual cockpit version [4, 21].

The element of the system mounted on the pilot's helmet is the so-called attachment that includes a projecting system of data imaging on a helmet-mounted display system, as well as a helmet position tracking system in relation to the aircraft cabin with a magnetic field sensor. The system is characterized by a field of view of 6°x 6°, with an image source in the form of a miniature CRT lamp and night vision with a field of 40° using NVG night vision goggles or an image from FLIR and LLTV heads [2, 4]. This system allows to present information to the pilot in the field of target data, in "air-air" and "air-ground" modes (in cooperation with LANTIRN suspended container designed for F-16 multi-purpose aircraft).

The JHMCS system (Fig. 1.) uses the magnetic system of tracking the pilot's helmet with active noise reduction, which has a wide operating range of 180° in azimuth and 90° in elevation. A zeroing line is determined in relation to the aircraft cabin with a 240 Hz update frequency [17]. The estimated accuracy of the target indication is of the order of ±0.5°. In the available advertising materials, there is no information on the algorithm for determining the helmet position on the basis of the analysis of components of the generated magnetic field.
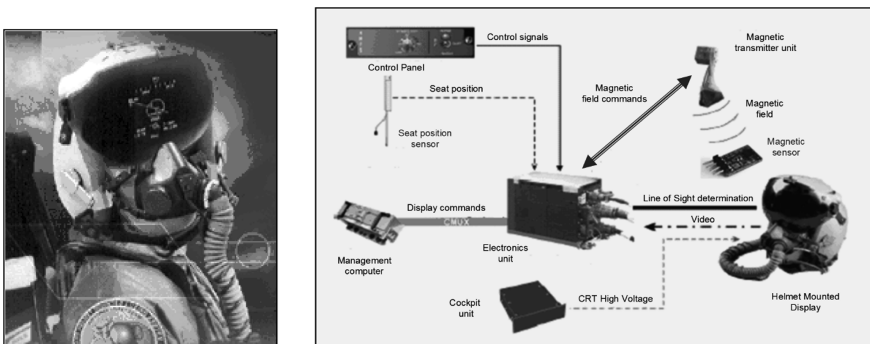
Fig. 1. View of the JHMCS helmet-mounted cueing system (left side) and the system of generating the magnetic field and managing on the board of the F-16 aircraft (right side) [17]

The literature analysis results in the fact that one of the main problems related to the combat use of the JHMCS class systems is the achievement of too little accuracy for precise determination of the pilot's zeroing line, currently required for new types of

on-board combat assets, among others, missiles aimed at targets indicated by laser and opto-electronic surveillance TV and FLIR high resolution devices [4, 15, 17].

In order to meet these needs, within the framework of the implementation of a research project carried out from the resources of the Ministry of Science and Higher Education of the Republic of Poland, for the research and development (R&D) purposes, a demonstrator of technologies in the scope of the NSC-1 helmet-mounted cueing system was built in AFIT (Fig. 2).

Fig. 2. View of the NSC-1 Orion helmet-mounted cueing system (left side) and the magnetic field generation system with the flat coil at the flight simulator station (right side) [15]

This system uses the magnetic method in order to determine the momentary angular position of the pilot's helmet in relation to the aircraft cabin [15]. As a magnetic field generator, the flat coil and specialized control system for supplying and measuring the magnetic field components, necessary to determine the system of coordinates associated with the aircraft cabin, were used. The developed algorithm for determining the helmet position in relation to the aircraft cabin is based on a quaternion calculus and the magnetic field components measured by a sensor built into the attachment mounted on the pilot's helmet [20].

The currently built NSC-1 helmet-mounted cueing system is used in AFIT to control the angular position of the observation-targeting head with TV/FLIR cameras, coupled with a flight simulator under laboratory conditions [21].

2 Magnetic Method and System Used in the NSC-1 System

The tested magnetic system used in the NSC-1 system consists of one flat coil (Fig. 3) generating an artificial magnetic field on the aircraft board [11, 13, 16, 18]. This coil (marked as C) contains 80 coils (wound in the form of a spiral with the radius $R_C = 18$ cm), has two separated rings: an inner ring (marked as A) and an outer ring (marked as B). The inner ring (A) contains 40 coils (wound in the form of a spiral with the radius $R_A = 9$ cm), however, the outer ring (B) contains 40 coils (wound in the form of a spiral with the radius $R_B = 18$ cm, from 9 cm to 18 cm).

Under laboratory conditions of AFIT, in order to supply the coil, a DC power supply with digital regulation and stabilization of the intensity of the current flowing in

the coil winding was used. In order to control the operation of the power supply system and the measuring system with the ADIS 16405 magnetic sensor [1, 7, 23], the specialised KG-1HC computer, developed in AFIT, which is an integral part of the NSC-1 helmet-mounted cueing system, was used [20].
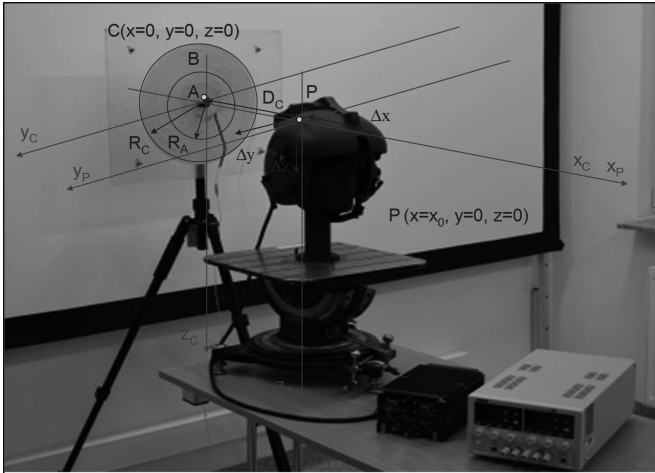
Fig. 3. View of the flat coil system with the inner and outer rings intended for magnetic field generation at the point of the ADIS 16405 helmet-mounted measuring system location [22]

The measurement cycle includes four successive stages of supplying the selected elements of the flat coil (inner and outer rings):

Stage 1.   Supplying of the inner ring (marked as A coil) and measuring of the field by the helment-mounted sensor in the state determined for the set angular position of the pilot's helmet;

Stage 2.   Supplying of the outer ring (marked as B coil) and measuring of the field by the helment-mounted sensor in the state determined for the set angular position of the pilot's helmet;

Stage 3.   Supplying of the inner and outer rings (marked as C coil) and measuring of the field by the helment-mounted sensor in the state determined for the set angular position of the pilot's helmet;

Stage 4.   Measurement of the environmental field (without coil supplying) by the helmet-mounted sensor in the state determined for the set angular position of the pilot's helmet.

The magnetic field generated by the flat coil is treated as a reference field, forming the coordinate system associated with the aircraft cabin [22].

By taking into account the environmental magnetic field measurement (implemented at stage 4), three measurement vectors (from stages 1, 2 and 3) containing differential signals free from any distortive impact of the external magnetic field (natural magnetic field of the Earth and the magnetic field generated by on-board

electric circuits and structural ferromagnetic elements of the aircraft, including the pilot's seat, engine, spars) are obtained [3, 8, 14].

3 Description of Magnetic Field Generated in the NSC-1 System

The magnetic field, generated at individual stages of the measurement cycle, includes the components oriented in the reference frame associated with the aircraft cabin, which are described with the use of the following relationships [22]:

$$[B_{SP}(A)] = [B_{SP}^X(A), B_{SP}^Y(A), B_{SP}^Z(A)] \tag{1}$$

where: $B_{SP}(A)$ – vector of the magnetic field generated by the inner ring (A coil); $B_{SP}^X(A)$, $B_{SP}^Y(A)$, $B_{SP}^Z(A)$ – components of the magnetic field generated by A coil.

$$[B_{SP}(B)] = [B_{SP}^X(B), B_{SP}^Y(B), B_{SP}^Z(B)] \tag{2}$$

where: $B_{SP}(B)$ – vector of the magnetic field generated by the outer ring (B coil); $B_{SP}^X(B)$, $B_{SP}^Y(B)$, $B_{SP}^Z(B)$ – components of the magnetic field generated by B coil.

$$[B_{SP}(C)] = [B_{SP}^X(C), B_{SP}^Y(C), B_{SP}^Z(C)] \tag{3}$$

where: $B_{SP}(C)$ – vector of the magnetic field generated by both rings (C coil); $B_{SP}^X(C)$, $B_{SP}^Y(C)$, $B_{SP}^Z(C)$ – components of the magnetic field generated by C coil.

The components of the generated magnetic field were determined at the measurement point $P(x_0, y_0, z_0)$, lying on the axis of symmetry of the flat coil and distant from the coil by D_C, characterising the so-called neutral position of the pilot's helmet in the adjustment process (Fig. 4).
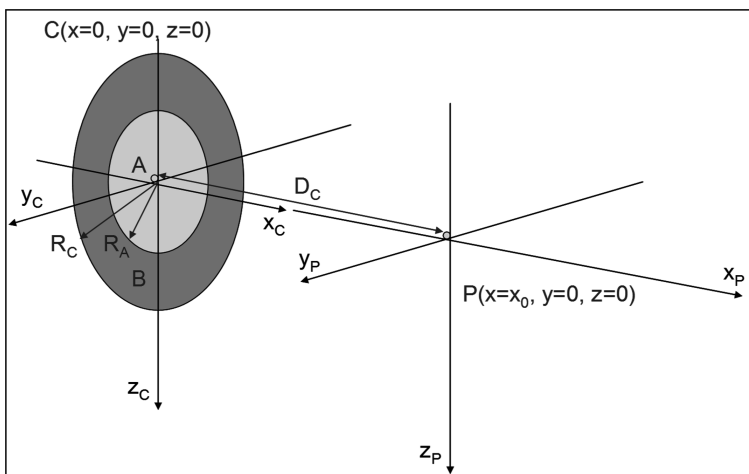
Fig. 4. Overview of the method for determining the components of the magnetic field generated by the flat coil at the point of the ADIS 16405 measuring system location [22]

For A coil (inner ring) supplied with the current of $I(A)$ intensity with the axis of symmetry oriented along the longitudinal axis of the aircraft (where $D_C = D(A)$, $R_A = R_i(A)$), the magnetic field components are defined in the following form [12]:

$$B_{SP}^X(A, x_0, y_0, z_0) = \mu \cdot I(A) \cdot A_3(A, x_0, y_0, z_0) \tag{4}$$

$$B_{SP}^Y(A, x_0, y_0, z_0) = 0 \tag{5}$$

$$B_{SP}^Z(A, x_0, y_0, z_0) = 0 \tag{6}$$

where:

$$A_3(A, x_0, y_0, z_0) = \frac{1}{2} \cdot \sum_{R_{\min}(A)=0\,\mathrm{cm}}^{R_{\max}(A)=9\,\mathrm{cm}} \frac{R_i^2(A)}{[(D(A))^2 + R_i^2(A)]^{3/2}} \tag{7}$$

For B coil (outer ring) supplied with the current of $I(B)$ intensity with the axis of symmetry oriented along the longitudinal axis of the aircraft (where $D_C = D(B)$, $R_B = R_i(B)$), the magnetic field components are defined in the following form [12]:

$$B_{SP}^X(B, x_0, y_0, z_0) = \mu \cdot I(B) \cdot A_3(B, x_0, y_0, z_0) \tag{8}$$

$$B_{SP}^Y(B, x_0, y_0, z_0) = 0 \tag{9}$$

$$B_{SP}^Z(B, x_0, y_0, z_0) = 0 \tag{10}$$

where:

$$A_3(B, x_0, y_0, z_0) = \frac{1}{2} \cdot \sum_{R_{\min}(B)=9\,\mathrm{cm}}^{R_{\max}(B)=18\,\mathrm{cm}} \frac{R_i^2(B)}{[(D(B))^2 + R_i^2(B)]^{3/2}} \tag{11}$$

For C coil (both rings) supplied with the current of $I(C)$ intensity with the axis of symmetry oriented along the longitudinal axis of the aircraft (where $D_C = D(C)$, $R_C = R_i(C)$), the magnetic field components are defined in the following form [12]:

$$B_{SP}^X(C, x_0, y_0, z_0) = \mu \cdot I(C) \cdot A_3(C, x_0, y_0, z_0) \tag{12}$$

$$B_{SP}^Y(C, x_0, y_0, z_0) = 0 \tag{13}$$

$$B_{SP}^Z(C, x_0, y_0, z_0) = 0 \tag{14}$$

where:

$$A_3(C, x_0, y_0, z_0) = \frac{1}{2} \cdot \sum_{R_{min}(C) = 0\,cm}^{R_{max}(C) = 18\,cm} \frac{R_i^2(C)}{[(D(C)^2 + R_i^2(C)]^{3/2}} \tag{15}$$

4 The NSC-1 System - Helmet Position Determining Algorithm

In order to determine the angular position of the pilot's helmet in relation to the aircraft cabin, the quaternion calculus describing the relations between the magnetic field components generated by the flat coil (or its rings) and components of the magnetic field measured by the sensor placed on the pilot's helmet can be used [9, 10, 20].

Between the components of the magnetic field generated by the flat coil (C) and the components of the magnetic field measured by the sensor placed on the pilot's helmet in the set angular position (represented by an elevation angle and an azimuth angle), there is [20]:

$$\{\,B_{PO}(C)\,\} \,\circ\, \{\,Q_{SP}^{PO}(C)\,\} = \{\,Q_{SP}^{PO}(C)\,\} \,\circ\, \{\,B_{SP}(C)\,\} \tag{16}$$

$$\{\,Q_{SP}^{PO}(C)\,\} = \{\,B_{PO}(C)\,\} \,\circ\, \{\,Q_{SP}^{PO}(C)\} \,\circ\, \{\,B_{SP}(C)\,\}^{-1} \tag{17}$$

where: $\{B_{SP}(C)\}$ – quaternion built from the components of the vector of the magnetic field generated by C coil; $\{B_{PO}(C)\}$ – quaternion built of magnetic field vector components measured by the ADIS 16405 sensor; $\{Q_{SP}^{PO}(C)\}$ – quaternion specifying the angular position of the pilot's helmet in relation to the aircraft cabin.

In order to solve the above equation, the original algorithm, developed in AFIT and based on the normalized quaternion properties, was used, and for it, the following occurs [20]:

$$[Q0_{SP}^{PO}(C)]^2 + [QX_{SP}^{PO}(C)]^2 + [QY_{SP}^{PO}(C)]^2 + [QZ_{SP}^{PO}(C)]^2 = 1 \tag{18}$$

where: $Q0_{SP}^{PO}(C)$, $QX_{SP}^{PO}(C)$, $QY_{SP}^{PO}(C)$, $QZ_{SP}^{PO}(C)$ – components of the quaternion defining the angular position of the pilot's helmet in relation to the aircraft cabin.

Based on the designated quaternion components, it is possible to determine the angles of spatial orientation of the pilot's helmet in relation to the aircraft in the form of:

– elevation angle of the pilot's helmet in relation to the cabin reference frame:

$$E_{SP}^{PO}(C) = ARC\,CTN\left(\frac{-\sqrt{1 - 4 \cdot \left[QX_{SP}^{PO}(C) \cdot QZ_{SP}^{PO}(C) + Q0_{SP}^{PO}(C) \cdot QY_{SP}^{PO}(C)\right]^2}}{2 \cdot \left[QX_{SP}^{PO}(C) \cdot QZ_{SP}^{PO}(C) + Q0_{SP}^{PO}(C) \cdot QY_{SP}^{PO}(C)\right]}\right)$$

$$\tag{19}$$

– azimuth angle of the pilot's helmet in relation to the cabin reference frame:

$$A_{SP}^{PO}(C) = 2ARC\ CTN \left(\begin{array}{c} \dfrac{\sqrt{1 - 4 \cdot \left[QX_{SP}^{PO}(C) \cdot QZ_{SP}^{PO}(C) + Q0_{SP}^{PO}(C) \cdot QY_{SP}^{PO}(C) \right]^2}}{2 \cdot \left[QY_{SP}^{PO}(C) \cdot QZ_{SP}^{PO}(C) - Q0_{SP}^{PO}(C) \cdot QX_{SP}^{PO}(C) \right]} + \\[4mm] + \dfrac{\left[(Q0_{SP}^{PO}(C))^2 + (QZ_{SP}^{PO}(C))^2 - (QX_{SP}^{PO}(C))^2 - (QY_{SP}^{PO}(C))^2 \right]}{2 \cdot \left[QY_{SP}^{PO}(C) \cdot QZ_{SP}^{PO}(C) - Q0_{SP}^{PO}(C) \cdot QX_{SP}^{PO}(C) \right]} \end{array} \right)$$

(20)

Preliminary testing accuracy of determining the spacial position's angles of the pilot's helmet for the NSC-1 helmet-mounted cueing system were performed on the test station, built in the AFIT for testing the flat coil used in the magnetic methods [5, 6].

The errors of determining the angles of spatial orientation of the pilot's helmet in relation to the aircraft cabin for the assumed inaccuracy of the magnetic field measurement for the ADIS 16405 sensor (Fig. 5) can be determined with the use of the above relationships, taking into account the sensor errors in the equations (4)÷(6), (8)÷(10), (12)÷(14) respectively for the components of the field generated by individual rings of the flat coil.

Fig. 5. View of the integrated ADIS 16405 triaxial sensor built on the pilot's helmet [20]

The error values for the selected angular positions of the pilot's helmet in relation to the aircraft cabin for the assumed inaccuracy of the magnetic field measurement by the ADIS 16405 sensor at the level of ±0.5 [mgauss] in the axes transverse to the axis of

Table 1. The errors of determining the elevation angle of the pilot's helmet in relation to the reference frame associated with the cabin resulting from the measurement error of the magnetic field generated by the coils

		Azimuth						
		−45°	−30°	−15°	0°	+15°	+30°	+45°
Elevation	+45°	0,058	0,058	0,058	0,058	0,058	0,058	0,058
	+30°	0,057	0,057	0,057	0,057	0,057	0,057	0,057
	+15°	0,057	0,057	0,057	0,057	0,057	0,057	0,057
	0°	0,057	0,057	0,057	0,057	0,057	0,057	0,057
	−15°	0,057	0,057	0,057	0,057	0,057	0,057	0,057
	−30°	0,057	0,057	0,057	0,057	0,057	0,057	0,057
	−45°	0,058	0,058	0,058	0,058	0,058	0,058	0,058

Table 2. The errors of determining the azimuth angle of the pilot's helmet in relation to the reference frame associated with the cabin resulting from the measurement error of the magnetic field generated by the coils

		Azimuth						
		−45°	−30°	−15°	0°	+15°	+30°	+45°
Elevation	+45°	0,081	0,081	0,081	0,081	0,081	0,081	0,081
	+30°	0,066	0,066	0,066	0,066	0,081	0,081	0,081
	+15°	0,059	0,059	0,059	0,059	0,081	0,081	0,081
	0°	0,057	0,057	0,057	0,057	0,081	0,081	0,081
	-15°	0,059	0,059	0,059	0,059	0,081	0,081	0,081
	-30°	0,066	0,066	0,066	0,066	0,081	0,081	0,081
	-45°	0,081	0,081	0,081	0,081	0,081	0,081	0,081

symmetry of the coil are shown in Table 1 (for the elevation channel) and in Table 2 (for the azimuth channel).

The obtained error values do not exceed the level of 0.1° and they can be reduced in the future by using a more accurate magnetic field sensor [21, 22].

5 Field Curvature Correction Algorithm in the NSC-1 System

The components of the generated magnetic field were determined at the measurement point P'$(x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z)$, lying outside the axis of symmetry of the coil and distant from the so-called neutral position of the pilot's helmet (at the point P(x_0, y_0, z_0)) by the deviations Δx, Δy and Δz (Fig. 6).
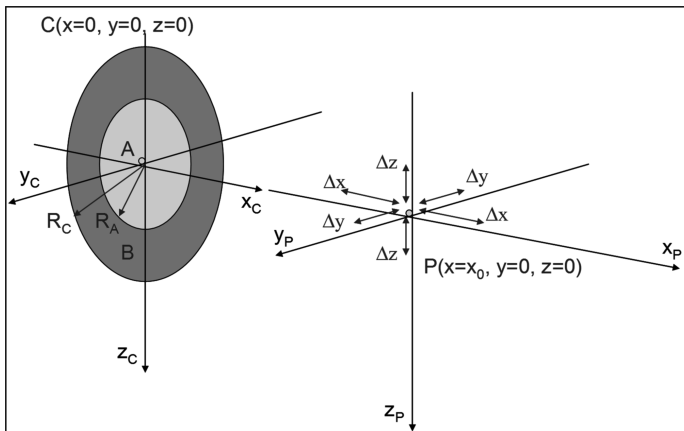
Fig. 6. Overview of the method for determining the components of the magnetic field generated by the flat coil at the point of the ADIS 16405 measuring system location [22]

The magnetic field generation and measurement take place analogically to the above-described case for the pilot's helmet placed in the neutral position. The measurement cycle consists of four successive stages of supplying the selected elements of the flat coil (inner and outer rings). In order to reduce interference (from the Earth's magnetic field and the magnetic field generated by the aircraft and its equipment), the environmental magnetic field is measured when the flat coil is not supplied.

For A coil (inner ring) supplied with the current of $I(A)$ intensity with the axis of symmetry oriented along the longitudinal axis of the aircraft, the magnetic field components are defined in the following form [22]:

$$
\begin{aligned}
B_{SP}^X(A, \Delta x, \Delta y, \Delta z) = {} & B_{SP}^X(A, x_0, y_0, z_0) + \mu \cdot I(A) \cdot [2A_1(A, x_0, y_0, z_0) \cdot \Delta x] + \\
& + \mu \cdot I(A) \cdot \left[A_2(A, x_0, y_0, z_0) \cdot \left(\Delta x^2 - \frac{1}{2}\Delta y^2 - \frac{1}{2}\Delta z^2 \right) \right] + \\
& + \mu \cdot I(A) \cdot \left[\frac{1}{6} A_4(A, x_0, y_0, z_0) \cdot \left(4\Delta x^3 - 6\Delta y^2 \Delta x^2 - 6\Delta z^2 \Delta x^2 \right) \right]
\end{aligned}
$$
(21)

$$
\begin{aligned}
B_{SP}^Y(A, \Delta x, \Delta y, \Delta z) = {} & B_{SP}^Y(A, x_0, y_0, z_0) + \mu \cdot I(A) \cdot [A_1(A, x_0, y_0, z_0) \cdot \Delta y] + \\
& + \mu \cdot I(A) \cdot [A_2(A, x_0, y_0, z_0) \cdot \Delta x \cdot \Delta y] + \\
& + \mu \cdot I(A) \cdot \left[\frac{1}{4} A_4(A, x_0, y_0, z_0) \cdot \left(\Delta y^2 + \Delta z^2 - 4\Delta x^2 \right) \cdot \Delta y \right]
\end{aligned}
$$
(22)

$$
\begin{aligned}
B_{SP}^Z(A, \Delta x, \Delta y, \Delta z) = {} & B_{SP}^Z(A, x_0, y_0, z_0) + \mu \cdot I(A) \cdot [A_1(A, x_0, y_0, z_0) \cdot \Delta z] + \\
& + \mu \cdot I(A) \cdot [A_2(A, x_0, y_0, z_0) \cdot \Delta x \cdot \Delta z] + \\
& + \mu \cdot I(A) \cdot \left[\frac{1}{4} A_4(A, x_0, y_0, z_0) \cdot \left(\Delta y^2 + \Delta z^2 - 4\Delta x^2 \right) \cdot \Delta z \right]
\end{aligned}
$$
(23)

where:

$$
A_1(A, x_0, y_0, z_0) = \frac{3}{4} \cdot \sum_{R_{\min}(A) = 0\,\text{cm}}^{R_{\max}(A) = 9\,\text{cm}} \frac{D(A) \cdot R_i^2(A)}{[(D(A))^2 + R_i^2(A)]^{5/2}}
$$
(24)

$$
A_2(A, x_0, y_0, z_0) = \frac{3}{4} \cdot \sum_{R_{\min}(A) = 0\,\text{cm}}^{R_{\max}(A) = 9\,\text{cm}} \frac{R_i^2(A) \cdot [4 \cdot D^2(A) - R_i^2(A)]}{[(D(A))^2 + R_i^2(A)]^{7/2}}
$$
(25)

$$
A_4(A, x_0, y_0, z_0) = \frac{15}{8} \cdot \sum_{R_{\min}(A) = 0\,\text{cm}}^{R_{\max}(A) = 9\,\text{cm}} \frac{R_i^2(A) \cdot [4 \cdot D^2(A) - 3 \cdot R_i^2(A)]}{[(D(A))^2 + R_i^2(A)]^{9/2}}
$$
(26)

For B coil (outer ring) supplied with the current of $I(B)$ intensity with the axis of symmetry oriented along the longitudinal axis of the aircraft, the magnetic field components are defined in the following form [22]:

$$
\begin{aligned}
B_{SP}^{X}(B, \Delta x, \Delta y, \Delta z) &= B_{SP}^{X}(B, x_0, y_0, z_0) + \mu \cdot I(B) \cdot [2A_1 (B, x_0, y_0, z_0) \cdot \Delta x] + \\
&+ \mu \cdot I(B) \cdot \left[A_2(B, x_0, y_0, z_0) \cdot (\Delta x^2 - \frac{1}{2}\Delta y^2 - \frac{1}{2}\Delta z^2) \right] + \\
&+ \mu \cdot I(B) \cdot \left[\frac{1}{6} A_4 (B, x_0, y_0, z_0) \cdot (4\Delta x^3 - 6\Delta y^2 \Delta x^2 - 6\Delta z^2 \Delta x^2) \right]
\end{aligned}
$$
(27)

$$
\begin{aligned}
B_{SP}^{Y}(B, \Delta x, \Delta y, \Delta z) &= B_{SP}^{Y}(B, x_0, y_0, z_0) + \mu \cdot I(B) \cdot [A_1 (B, x_0, y_0, z_0) \cdot \Delta y] + \\
&+ \mu \cdot I(B) \cdot [A_2(B, x_0, y_0, z_0) \cdot \Delta x \cdot \Delta y] + \\
&+ \mu \cdot I(B) \cdot \left[\frac{1}{4} A_4 (B, x_0, y_0, z_0) \cdot (\Delta y^2 + \Delta z^2 - 4\Delta x^2) \cdot \Delta y \right]
\end{aligned}
$$
(28)

$$
\begin{aligned}
B_{SP}^{Z}(B, \Delta x, \Delta y, \Delta z) &= B_{SP}^{Z}(B, x_0, y_0, z_0) + \mu \cdot I(B) \cdot [A_1 (B, x_0, y_0, z_0) \cdot \Delta z] + \\
&+ \mu \cdot I(B) \cdot [A_2(B, x_0, y_0, z_0) \cdot \Delta x \cdot \Delta z] + \\
&+ \mu \cdot I(B) \cdot \left[\frac{1}{4} A_4 (B, x_0, y_0, z_0) \cdot (\Delta y^2 + \Delta z^2 - 4\Delta x^2) \cdot \Delta z \right]
\end{aligned}
$$
(29)

where:

$$
A_1(B, x_0, y_0, z_0) = \frac{3}{4} \cdot \sum_{R_{\min}(B)=9\,\text{cm}}^{R_{\max}(B)=18\,\text{cm}} \frac{D(B) \cdot R_i^2(B)}{[(D(B)^2 + R_i^2(B)]^{5/2}}
$$
(30)

$$
A_2(B, x_0, y_0, z_0) = \frac{3}{4} \cdot \sum_{R_{\min}(B)=9\,\text{cm}}^{R_{\max}(B)=18\,\text{cm}} \frac{R_i^2(B) \cdot [4 \cdot D^2(B) - R_i^2(B)]}{[(D(B)^2 + R_i^2(B)]^{7/2}}
$$
(31)

$$
A_4(B, x_0, y_0, z_0) = \frac{15}{8} \cdot \sum_{R_{\min}(B)=9\,\text{cm}}^{R_{\max}(B)=18\,\text{cm}} \frac{R_i^2(B) \cdot [4 \cdot D^2(B) - 3 \cdot R_i^2(B)]}{[(D(B)^2 + R_i^2(B)]^{9/2}}
$$
(32)

For C coil (both rings) supplied with the current of $I(C)$ intensity with the axis of symmetry oriented along the longitudinal axis of the aircraft, the magnetic field components are defined in the following form [22]:

$$B_{SP}^{X}(C, \Delta x, \Delta y, \Delta z) = B_{SP}^{X}(C, x_0, y_0, z_0) + \mu \cdot I(C) \cdot [2A_1(C, x_0, y_0, z_0) \cdot \Delta x] +$$
$$+ \mu \cdot I(C) \cdot \left[A_2(C, x_0, y_0, z_0) \cdot \left(\Delta x^2 - \frac{1}{2}\Delta y^2 - \frac{1}{2}\Delta z^2 \right) \right] +$$
$$+ \mu \cdot I(C) \cdot \left[\frac{1}{6} A_4(C, x_0, y_0, z_0) \cdot (4\Delta x^3 - 6\Delta y^2 \Delta x^2 - 6\Delta z^2 \Delta x^2) \right]$$
(33)

$$B_{SP}^{Y}(C, \Delta x, \Delta y, \Delta z) = B_{SP}^{Y}(C, x_0, y_0, z_0) + \mu \cdot I(C) \cdot [A_1(C, x_0, y_0, z_0) \cdot \Delta y] +$$
$$+ \mu \cdot I(C) \cdot [A_2(C, x_0, y_0, z_0) \cdot \Delta x \cdot \Delta y] +$$
$$+ \mu \cdot I(C) \cdot \left[\frac{1}{4} A_4(C, x_0, y_0, z_0) \cdot (\Delta y^2 + \Delta z^2 - 4\Delta x^2) \cdot \Delta y \right]$$
(34)

$$B_{SP}^{Z}(C, \Delta x, \Delta y, \Delta z) = B_{SP}^{Z}(C, x_0, y_0, z_0) + \mu \cdot I(C) \cdot [A_1(C, x_0, y_0, z_0) \cdot \Delta z] +$$
$$+ \mu \cdot I(C) \cdot [A_2(C, x_0, y_0, z_0) \cdot \Delta x \cdot \Delta z] +$$
$$+ \mu \cdot I(C) \cdot \left[\frac{1}{4} A_4(C, x_0, y_0, z_0) \cdot (\Delta y^2 + \Delta z^2 - 4\Delta x^2) \cdot \Delta z \right]$$
(35)

where:

$$A_1(C, x_0, y_0, z_0) = \frac{3}{4} \cdot \sum_{R_{min}(C)=0\,cm}^{R_{max}(C)=18\,cm} \frac{D(C) \cdot R_i^2(C)}{[(D(C)^2 + R_i^2(C)]^{5/2}}$$
(36)

$$A_2(C, x_0, y_0, z_0) = \frac{3}{4} \cdot \sum_{R_{min}(C)=0\,cm}^{R_{max}(C)=18\,cm} \frac{R_i^2(C) \cdot [4 \cdot D^2(C) - R_i^2(C)]}{[(D(C)^2 + R_i^2(C)]^{7/2}}$$
(37)

$$A_4(C, x_0, y_0, z_0) = \frac{15}{8} \cdot \sum_{R_{min}(C)=0\,cm}^{R_{max}(C)=18\,cm} \frac{R_i^2(C) \cdot [4 \cdot D^2(C) - 3 \cdot R_i^2(C)]}{[(D(C)^2 + R_i^2(C)]^{9/2}}$$
(38)

The knowledge of mathematical relationships describing the individual components of the magnetic field generated by individual rings of the flat coil at the set measurement point P'($x_0 + \Delta x, y_0 + \Delta y, z_0 + \Delta z$), distant from the neutral position P(x_0, y_0, z_0), allows to determine the magnetic field curvature angles, and it is used in the developed algorithm for the correction of errors of determining the angular position of the pilot's helmet [22].

The error values before making corrections for the selected angular positions of the pilot's helmet in relation to the cabin and the adopted deviations: $\Delta x = 0$ cm, $\Delta y = 1$ cm, $\Delta z = 1$ cm, were presented in Table 3 (for the elevation channel) and in Table 4 (for the azimuth channel).

Table 3. The errors of determining the elevation angle of the pilot's helmet in relation to the reference frame associated with the cabin resulting from the curvature of the magnetic field generated by the coils

		Azimuth						
		−45°	−30°	−15°	0°	+15°	+30°	+45°
Elevation	+45°	2,765	2,765	2,765	2,765	2,765	2,765	2,765
	+30°	2,740	2,740	2,740	2,740	2,740	2,740	2,740
	+15°	2.721	2.721	2.721	2.721	2.721	2.721	2.721
	0°	2,704	2,704	2,704	2,704	2,704	2,704	2,704
	−15°	2,721	2,721	2,721	2,721	2,721	2,721	2,721
	−30°	2,740	2,740	2,740	2,740	2,740	2,740	2,740
	−45°	2,765	2,765	2,765	2,765	2,765	2,765	2,765

Table 4. The errors of determining the azimuth angle of the pilot's helmet in relation to the reference frame associated with the cabin resulting from the curvature of the magnetic field generated by the coils

		Azimuth						
		−45°	−30°	−15°	0°	+15°	+30°	+45°
Elevation	+45°	4,016	4,016	4,016	4,016	4,016	4,016	4,016
	+30°	3,213	3,213	3,213	3,213	3,213	3,213	3,213
	+15°	2,839	2,839	2,839	2,839	2,839	2,839	2,839
	0°	2,707	2,707	2,707	2,707	2,707	2,707	2,707
	−15°	2,839	2,839	2,839	2,839	2,839	2,839	2,839
	−30°	3,213	3,213	3,213	3,213	3,213	3,213	3,213
	−45°	4,016	4,016	4,016	4,016	4,016	4,016	4,016

The obtained errors of determining the position of the pilot's helmet reach the maximum values of 4° and significantly exceed the acceptable level adopted for the NSC-1 Orion helmet-mounted cueing system [20].

In order to reduce these errors, the correction method that uses the same measurement data without the additional method for determining the linear position of the pilot's helmet in relation to the aircraft cabin, was developed in the Avionics Division of AFIT.

The developed author's original correction algorithm is based on the determination of the deviations Δx, Δy and Δz in relation to the neutral position of the pilot's helmet with the use of the relationships defining the so-called standard of the vector of the magnetic field generated by individual coils A, B and C.

For A coil (supplied inner ring) with the axis of symmetry oriented along the longitudinal axis of the aircraft, the vector module of the generated magnetic field is described in the following form:

$$\left| [B_{SP}(A, \Delta x, \Delta y, \Delta z)] \right| = \sqrt{\left(B_{SP}^X(A, \Delta x, \Delta y, \Delta z) \right)^2 + \left(B_{SP}^Y(A, \Delta x, \Delta y, \Delta z) \right)^2 + \left(B_{SP}^Z(A, \Delta x, \Delta y, \Delta z) \right)^2}$$

$$(39)$$

which is, at the same time, equal to the vector module of the magnetic field measured by the sensor in the set linear position of the pilot's helmet in relation to the aircraft cabin:

$$\left| [B_{PO}(A, \Delta x, \Delta y, \Delta z)] \right| = \sqrt{\left(B_{PO}^X(A, \Delta x, \Delta y, \Delta z) \right)^2 + \left(B_{PO}^Y(A, \Delta x, \Delta y, \Delta z) \right)^2 + \left(B_{PO}^Z(A, \Delta x, \Delta y, \Delta z) \right)^2}$$

$$(40)$$

For B coil (supplied outer ring) with the axis of symmetry oriented along the longitudinal axis of the aircraft, the vector module of the generated magnetic field is described in the following form:

$$\left| [B_{SP}(B, \Delta x, \Delta y, \Delta z)] \right| = \sqrt{\left(B_{SP}^X(B, \Delta x, \Delta y, \Delta z) \right)^2 + \left(B_{SP}^Y(B, \Delta x, \Delta y, \Delta z) \right)^2 + \left(B_{SP}^Z(B, \Delta x, \Delta y, \Delta z) \right)^2}$$

$$(41)$$

which is, at the same time, equal to the vector module of the magnetic field measured by the sensor in the set linear position of the pilot's helmet in relation to the aircraft cabin:

$$\left| [B_{PO}(B, \Delta x, \Delta y, \Delta z)] \right| = \sqrt{\left(B_{PO}^X(B, \Delta x, \Delta y, \Delta z) \right)^2 + \left(B_{PO}^Y(B, \Delta x, \Delta y, \Delta z) \right)^2 + \left(B_{PO}^Z(B, \Delta x, \Delta y, \Delta z) \right)^2}$$

$$(42)$$

For C coil (both rings supplied) with the axis of symmetry oriented along the longitudinal axis of the aircraft, the vector module of the generated magnetic field is described in the following form:

$$\left| [B_{SP}(C, \Delta x, \Delta y, \Delta z)] \right| = \sqrt{\left(B_{SP}^X(C, \Delta x, \Delta y, \Delta z) \right)^2 + \left(B_{SP}^Y(C, \Delta x, \Delta y, \Delta z) \right)^2 + \left(B_{SP}^Z(C, \Delta x, \Delta y, \Delta z) \right)^2}$$

$$(43)$$

which is, at the same time, equal to the vector module of the magnetic field measured by the sensor in the set linear position of the pilot's helmet in relation to the aircraft cabin:

$$\left| [B_{PO}(C, \Delta x, \Delta y, \Delta z)] \right| = \sqrt{\left(B_{PO}^X(C, \Delta x, \Delta y, \Delta z) \right)^2 + \left(B_{PO}^Y(C, \Delta x, \Delta y, \Delta z) \right)^2 + \left(B_{PO}^Z(C, \Delta x, \Delta y, \Delta z) \right)^2}$$

$$(44)$$

The above relationships allow to determine the deviations Δx, Δy and Δz in relation to the neutral position of the pilot's helmet by solving the equation system (39) ÷ (44), after substitution of variables from the relationships (21) ÷ (23), (27) ÷ (29) and (33) ÷ (35).

For example, after taking into account only the linear parts from the development of individual components of the magnetic field into the Taylor series, the deviations Δx and Δy characterising the linear position of the pilot's helmet along the longitudinal and transverse axes of the aircraft in the curved magnetic field generated by the selected rings (inner and outer) of the flat coil can be determined with the use of the following relationships:

$$\Delta x^* = \frac{\left[(A_1(A,x_0,y_0,z_0))^2 \cdot (A_3(B,x_0,y_0,z_0))^2 - (A_1(B,x_0,y_0,z_0))^2 \cdot (A_3(A,x_0,y_0,z_0))^2\right]}{4 \cdot \left\{\left[(A_1(B,x_0,y_0,z_0))^2 \cdot A_1(A,x_0,y_0,z_0) \cdot A_3(A,x_0,y_0,z_0)\right] - \left[(A_1(A,x_0,y_0,z_0))^2 \cdot A_1(B,x_0,y_0,z_0) \cdot A_3(B,x_0,y_0,z_0)\right]\right\}} +$$
$$+ \frac{\left[(A_1(B,x_0,y_0,z_0))^2 \cdot (B_{PO}(A,x_0,y_0,z_0))^2 - (A_1(B,x_0,y_0,z_0))^2 \cdot (A_3(A,x_0,y_0,z_0))^2\right]}{4 \cdot \left\{\left[(A_1(B,x_0,y_0,z_0))^2 \cdot A_1(A,x_0,y_0,z_0) \cdot A_3(A,x_0,y_0,z_0)\right] - \left[(A_1(A,x_0,y_0,z_0))^2 \cdot A_1(B,x_0,y_0,z_0) \cdot A_3(B,x_0,y_0,z_0)\right]\right\}}$$

$$\tag{45}$$

$$\Delta y^* = \frac{\left[(B_{PO}(B,x_0,y_0,z_0))^2 - (A_3(B,x_0,y_0,z_0))^2\right]}{(A_1(B,x_0,y_0,z_0))^2} - \frac{[2 \cdot A_1(B,x_0,y_0,z_0)]}{(A_1(B,x_0,y_0,z_0))^2} \cdot \Delta x^* \tag{46}$$

The determined deviation values Δx^*, Δy^* and Δz^* allow to specify the curvature angles of the magnetic field at the set measurement point, in which the magnetic field sensor built on the pilot's helmet is located, and thus, to make the corrections to the determined elevation and azimuth angles of the pilot's helmet in relation to the cabin reference frame. The knowledge of the magnetic field curvature angles is used in the developed algorithm for correcting the errors of determining the angular position of the pilot's helmet.

The values of these errors after making corrections for the selected angular positions of the pilot's helmet in relation to the cabin and the adopted deviations: $\Delta x = 0$ cm, $\Delta y = 1$ cm, $\Delta z = 1$ cm, were presented in Table 5 (for the elevation channel) and in Table 6 (for the azimuth channel).

Table 5. The errors of determining the elevation angle of the pilot's helmet in relation to the cabin reference frame after the correction of the curvature of the magnetic field generated by the coils

		Azimuth						
		−45°	−30°	−15°	0°	+15°	+30°	+45°
Elevation	+45°	0,573	0,573	0,573	0,573	0,573	0,573	0,573
	+30°	0,563	0,563	0,563	0,563	0,563	0,563	0,563
	+15°	0,554	0,554	0,554	0,554	0,554	0,554	0,554
	0°	0,547	0,547	0,547	0,547	0,547	0,547	0,547
	−15°	0,554	0,554	0,554	0,554	0,554	0,554	0,554
	−30°	0,563	0,563	0,563	0,563	0,563	0,563	0,563
	−45°	0,573	0,573	0,573	0,573	0,573	0,573	0,573

Table 6. The errors of determining the azimuth angle of the pilot's helmet in relation to the cabin reference frame after the correction of the curvature of the magnetic field generated by the coils

		Azimuth						
		−45°	−30°	−15°	0°	+15°	+30°	+45°
Elevation	+45°	0,549	0,549	0,549	0,549	0,549	0,549	0,549
	+30°	0,481	0,481	0,481	0,481	0,481	0,481	0,481
	+15°	0,426	0,426	0,426	0,426	0,426	0,426	0,426
	0°	0,338	0,338	0,338	0,338	0,338	0,338	0,338
	−15°	0,426	0,426	0,426	0,426	0,426	0,426	0,426
	−30°	0,481	0,481	0,481	0,481	0,481	0,481	0,481
	−45°	0,549	0,549	0,549	0,549	0,549	0,549	0,549

The obtained errors of determining the angular position of the pilot's helmet reach the maximum values of 0.5° and do not exceed the acceptable level adopted for the NSC-1 Orion helmet-mounted cueing system [20]. The values of these errors can be reduced by using more developed models of the generated magnetic field (including further non-linear components from the development into the Taylor series).

6   Conclusion

The scientific aspect of the presented subject was to present a method for determining the angular position of the pilot's helmet in the NSC-Orion helmet-mounted cueing system, developed in AFIT for the W-3PL Głuszec multi-purpose helicopter. The carried-out simulation and verification tests were aimed at determination of the inaccuracy of specifying the angular position of the pilot's helmet in relation to the reference frame associated with the aircraft cabin, with the use of the magnetic method (with the magnetic system including the flat coil that generates the magnetic field with a curved profile from the axis of symmetry).

The carried-out error estimates in the range of the inaccuracy of determining the angular position of the pilot's helmet with the use of the magnetic method showed that for the flat coil with two separated rings (inner and outer), it is possible to correct the errors by determining the helmet linear deviations from the neutral position (determined during adjustment of the helmet-mounted system).

The advantage of the developed correction algorithm is the use of the same measurement data that are used in order to determine the angular position of the pilot's helmet for the curved field. It has been proven that the components of the magnetic field generated by individual rings of the flat coin allow to determine the linear deviations of the helmet from the neutral position, and thus, to specify the magnetic field curvature and to make its correction based on the magnetic field models developed in the Avionics Division of AFIT.

The developed computational algorithms, as implemented software applications, were used in the laboratory of the Avionics Division of AFIT for helmet-mounted control of the angular position of the observation-targeting and reflector-searchlight heads. It is anticipated that the positive results of these works will be applied on board of the W-3PL Głuszec helicopter with the NSC-1 Orion helmet-mounted cueing system, among others, for the implementation of SAR/CSAR search and rescue missions.

References

1. ADIS 16405, Analog Devices, Triaxial Inertial Sensor with Magnetometer, datasheet. www.analog.com
2. Boniol, F.: New challenges for future avionic architectures. In: Amine, A., Otmane, A., Bellatreche, L. (eds.) Modeling Approaches and Algorithms for Advanced Computer Applications. Studies in Computational Intelligence, vol. 488. Springer, Cham (2013)
3. Chu, Y.: Numerical calculation for the magnetic field in current-carrying circular arc filament. IEEE Trans. Magn. 34 (1998)
4. Collinson, R.P.G.: Introduction to Avionics. Springer, Dordrecht (2011)
5. COSMOL Multiphysics® Modeling Software. www.cosmol.com
6. Furlani, E.P.: The field analysis and simulation of a permanent-magnet bias-field device for magneto-optic recording. J. Phys. D Appl. Phys. **30** (1997)
7. IEEE Sens. J. Adv. Magn. Field Sens. **10**(6) (2010)
8. Jackson, J.D.: Classical Electrodynamics, 3rd Edn. Wiley (1998)
9. Simpson, J.C.: Simple analyctic expressions for the magnetic field od a circular current loop, NASA/TM-2013-217919. NASA, Kennedy Space Center, Florida (2001)
10. Selvaggi, J.P.: Multipole analysis of circular cylindrical magnetic systems. Rensselaer Polytechnic Institute, Troy, New York (2005)
11. Selvaggi, J.P., Salon, S., Kwon, O.-M Chari, M.V.K: Calculating the external magnetic field from permanent magnets in permanent-magnet motors-an alternative method. IEEE Trans. Magn. **40**(5) (2004)
12. Conway, J.T.: Exact solutions for the magnetic fields of axisymmetric solenoids and current distributions. IEEE Trans. Magn. **37**(4) (2001)
13. Kildishev, A.V.: Multipole analysis of an elongated magnetic source by a cylindrical sensor array. IEEE Trans. Magn. **38**(5) (2002)
14. Knoepfel, H.E.: Magnetic Fields: A Comprehensive Theoretical Treatise for Practical Use. A Wiley-Interscience Publication, Toronto (2000)
15. Lewitowicz, J., Szelmanowski, A., Pazur, A., Michalak, S., Janik, P.: Helmet-mounted display and cueing systems as pilot aids for aviation/maritime search and rescue missions. Sci. J. Marit. Univ. Szczec. **52**(124) (2017)
16. Schill, R.A.: General relation for the vector magnetic field of a circular current loop: a closer look. IEEE Trans. Magn. **39**(2) (2003)
17. Rash, C.: Helmet displays in aviation. Helmet mounted display. Design Issues for Rotary-Wing Aircraft, USA, Fort Rucker (2009)
18. Raab, R.E., De Lange, O.L.: Multipole Theory in Electromagnetism-Classical, Quantum, and Symmetry Aspects, with Applications. Oxford University Press (2005)
19. Sarma, K.R., Grothe, S., Gannon, A.: Avionics displays. In: Chen, J., Cranton, W., Fihn, M. (eds.) Handbook of Visual Display Technology. Springer, Cham (2016)
20. Szelmanowski, A.: Helmet-mounted cueing system for military helicopters with avionic integrated system, AFIT, monography, ISBN 978-83-61021-76-6, Warsaw (2013)

21. Szelmanowski, A., Janik, P., Sobielarski, M.: Possibilities of new technologies implementation in helmet-mounted cueing and display systems for Polish military helicopters. J. KONES Powertrain Transp. **23**(4) (2016)
22. Szelmanowski, A., Michalak, S., Witos, M., Pazur, A., Janik, P., Paterek, W., Gaj, A., Bulikowska, J.: Magnetic field sensors overview and example application in helmet mounted cueing system. In: 2nd IEEE Conference on Advances In Magnetics, La Thuile, Italy, Febuary, pp. 4–7 (2018)
23. Uchiyama, et al.: IEEE Trans. Magn. **48**(11), 3833–3839 (2012)

Robot-Assisted Quality Inspection of Turbojet Engine Blades

Dariusz Szybicki[✉], Andrzej Burghardt, Piotr Gierlak,
and Krzysztof Kurc

Rzeszow University of Technology,
Al. Powstancow Warszawy 12, 35-959 Rzeszow, Poland
{dszybicki, andrzejb, pgierlak, kkurc}@prz.edu.pl

Abstract. This paper presents the design and construction of a module for measuring the geometry and orientation of turbojet engine blades. The measurement module forms a component of a robotic processing station for grinding of turbojet engine blades. The robotic processing station comprises an industrial robot manipulator with a dedicated gripper, the measurement module with proximity sensors, and blade grinding tooling components. The measurement outputs were transferred by TCP from the measurement module to the controller of the robot manipulator. The measurement outputs could be applied in blade grinding, displayed on a control panel, or processed into a measurement test report.

Keywords: Robotic measurement · Blade inspection · Turbojet engine blades

1 Introduction

The development of industrial robotics focuses mainly on applications which require a high level of process-robot interaction. This includes turbojet engine blade measurement and grinding processes, where varying allowances must be removed. The processes include the robotic-assisted quality control of jet engine components presented in [1–3].

This paper proposes an alternative solution to the measurement module used for robot-assisted geometry measurement of turbojet engine blades. The measurement data output by the solution can be used for quality control or for the execution of or assistance in robotic machining. Aerospace robotic machining and the diagnostics of robotic processing station components are discussed in [4–8]. An application of a 3D scanner manufactured by GOM for the measurement of blade geometry during servicing and overhaul is presented in [9]. Blade geometry is measured in machining processes for cutting tool motion correction, for example by measurements made with machining head sensors, [10, 11] or the correction of a robot's TCP with a laser measurement system to improve grinding processes, [12, 13]. A grinding process can also be improved by laser scanning [14]. Laser triangulation systems for fast blade geometry measurements are presented in [15], with a mathematical description of the measurement method. The problems of approximation of the high volumes of

measurement output information obtained by scanning are considered in [16]. 3D laser systems which measure blade geometry and compare the measurement outputs to geometric masters (models) in robotic machining control applications are discussed in [17]. The problem of measurement uncertainty in vision systems is discussed in the paper [18].

To recapitulate, the problems of constructing a module dedicated to robot-assisted blade geometry verification remain valid and a part of the automation of machinery component fabrication and control processes.

2 Blade Characteristics

Blades are among the basic components of turbojet engines. The design of any modern turbojet engine is very complex, and the engines vary in the number and application types of blades [15]. A general classification of turbojet engine blades is as follows:

- compressor blades: these compress the air stream passing through the compressor; they are prone to foreign object damage (FOD), such as bird strikes,
- turbine blades: these transfer the active energy from the jet fuel combustion gas to the turbine rotor to propel it; they operate at high temperatures and in aggressive conditions.

Turbojet engine blades have to be manufactured using state of the art technologies to achieve the required high geometric accuracy and strength parameters. Turbojet engine blades are qualified as flight safety critical components, and are subject to many control procedures. The blade control procedures focus on the quality of castings and forgings, material structure, and fabricated geometric features. Turbojet engine blades are subject to torsion and bending by the aerodynamic forces and tension by the centrifugal forces. These mean that the blades the most stressed components in a turbojet engine, and require replacement at short intervals. The time and cost of blade production can be as much as 35% of the total time and cost of production of an entire turbojet engine [19]. A modern turbojet engine may have around 3000 blades. This sheer number of blades makes their durability a crucial criterion of the strength of the entire turbojet engine. The number of blades per turbojet engine means their manufacturing and quality control have become mass, automatic and repeatable processes.

The blade design (Fig. 1) depends on the blade type. A blade comprises an airfoil (the working part of the blade) and a blade locking piece. The crosssections of a blade airfoil are aerodynamically profiled and fabricated down to a geometric tolerance of 0.1 mm, at roughness levels (Ra) between 0.08 and 0.63 mm in the case of turbine blades. The blade locking piece fastens the blade to a component, and is fabricated to a geometric tolerance of 0.01 mm and Ra = 1.25 mm [20].

The concave portion of a blade is the high pressure surface while the convex portion is the low pressure surface. A blade features different edges: the leading edge (LE) is oriented towards the incoming gas flow, while the opposite edge is the trailing edge (TE). The quality control of these features involve geometry measurements of the blade by reference to the blade section, as shown in Fig. 2.
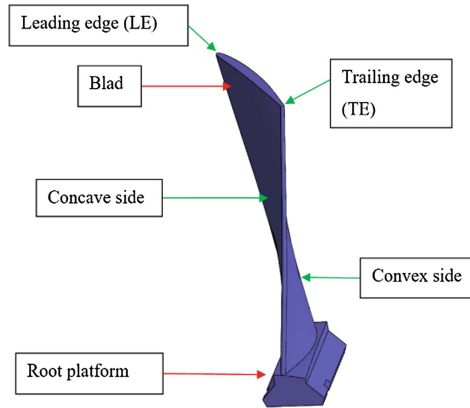
Fig. 1. Blade design.
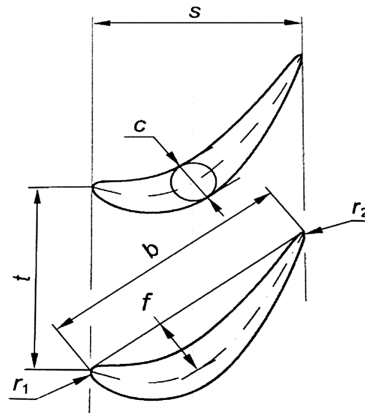
Fig. 2. Blade section [20].

The blade section has the following features:

- b – blade chord,
- c – blade thickness,
- f – section centre line,
- r1 – LE radius,
- r2 – TE radius.

The blade section is specifically twisted along its length and relative to the longitudinal axis of the blade; this twist is defined by the operating conditions of the blade and the stress it must withstand. Blades are dimensioned according to gas dynamic calculations. However, several general relationships can be defined for blade dimensioning, including the relationship of the blade thickness to the blade length (Fig. 3).
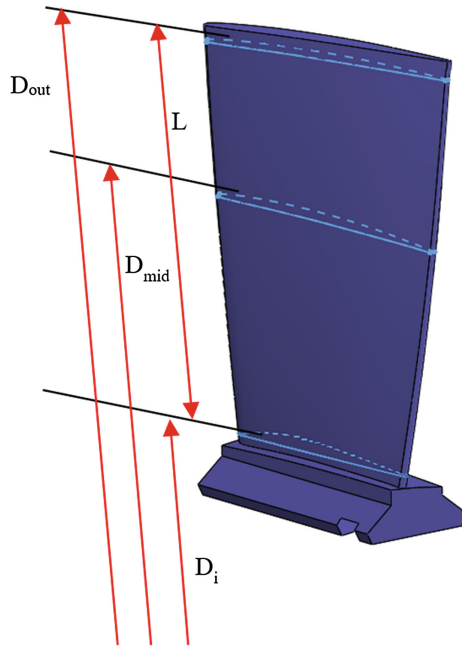
Fig. 3. Blade geometry vs. blade length.

The blade section thickness cmax for a turbine generator blade is within the following limits:

- root diameter (Din), $c1 = 0.2 - 0.3b$,
- middle diameter (Dmid), $c2 = 0.10 - 0.15b$,
- tip diameter (Dout), $c3 = 0.04 - 0.06b$.

3   Blade Geometry Measurement Methods

Several methods are available today by which the geometry and the twist angle of blades at individual blade sections can be measured. Geometry measurement methods generally include contact and non-contact methods.

Traditional geometry measurements rely on contact methods. The blade is measured by specifically designed instruments. The blade to be measured is aligned and secured in a fixture, following which contact sensors are used to measure the geometry. The measurement is made by bringing the measurement instrument's contact sensor head to the blade surface; if the measurement is done on both sides of the blade, the blade thickness can be determined. Legacy contact measurement solutions relied on a few specifically chosen measurement points, due to the immobility of the contact

sensors. Commercially available, dedicated measurement modules exist which enable automatic geometric measurements. One example is a module manufactured by Renishaw (Fig. 4).
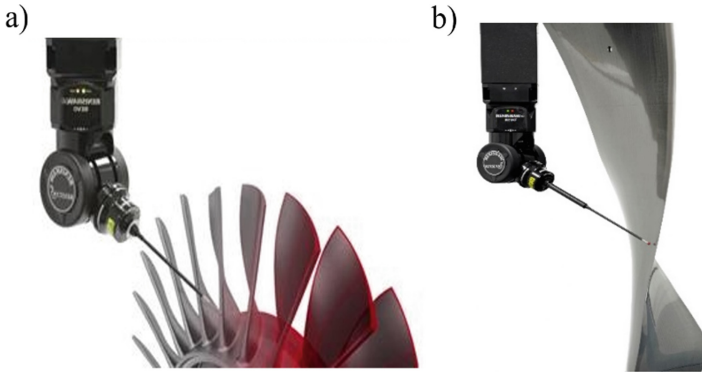
Fig. 4. REVO 5-axis contact sensor head applied in a blade geometry scan, (a) measurement of the rotor, (b) zoom.

In the solution proposed here, the REVO contact sensor head was installed in a coordinate measurement machine. The contact sensor head's paths of motion were generated by a dedicated software suite, APEXBlade. A drawback of this geometry measurement method was the need for manual fixing of the blades, the time required for the process, and the restricted data connectivity between the measurement system and other processing equipment. The overall solution was expensive due to the necessary cost of purchasing a CMM.

Non-contact geometry measurement methods can be based on laser proximity sensors, 2D and 3D laser scanners, and other types of 3D scanners.

Geometric measurement with laser triangulation sensors is a non-contact method. Laser triangulation sensors provide 2D scanning, by which the successive measurement points are offset by a fixed distance defined by the sensor resolution. Each measurement sensor has the distance to the laser sensor specified. The blade geometry is determined from the measurement outputs by a dedicated software suite.

The advantages of laser sensors include high accuracy, short measurement times and larger mounting distances from the test object than is possible with other types of non-contact sensors. Their weakness is that the LOS (line of sight) between the laser sensor and the test object must be clear and unobstructed at all times. Test results can be disturbed by suspended airborne dust or smoke, or a high gloss level of the blade being measured. However, the primary drawback of laser sensors is an accuracy level lower than available with contact geometry measurement methods. The application of laser sensors requires a precise definition of the blade basic surface location to enable the geometric measurement by a known distance.

Measurements of the blade geometry are often made by structural light 3D scanning. There are many commercial manufacturers of devices based on this method. One of these is the ATOS Core optical 3D scanner, manufactured by GOM (Fig. 5).
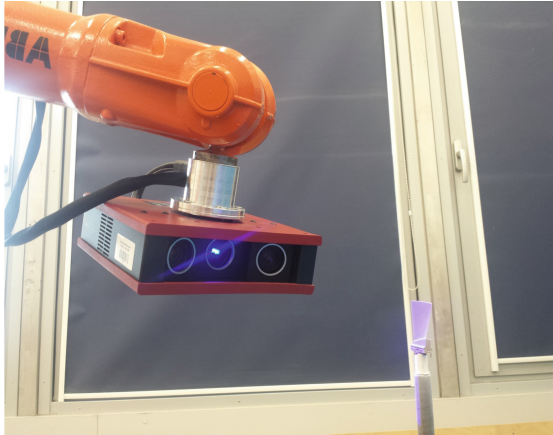
Fig. 5. Actual geometric measurement of a blade with the optical 3D scanner.

The ATOS Core optical 3D scanner features two stereo cameras, which work by the principle of triangulation, and one projector. The projector casts bands on the surface of the test object for the stereo cameras to capture. GOM uses blue LED projectors for scanning; this projected light emission spectrum enables precise geometric measurements under different lighting conditions.

The ATOS Core optical 3D scanner has the advantage of scanning small and moderately-sized objects at significantly reduced measurement and inspection times. This geometric measurement method produces a cloud of points, where the geometric data is compared to the CAD model (master) of the test object. Next, the dedicated software suite determines what differences, if any, exist between the actual test object and its CAD model. Optical 3D scanning has been gaining popularity in real-life applications; however, its use in the large-scale manufacturing of engine blades is difficult due to the high gloss levels of the product. To facilitate the geometric measurement of blades (Fig. 5) with this method, the high-gloss surfaces of each blade must be dimmed by applying a special powder. The powder reduces the measurement accuracy and makes the geometric measurement difficult in large-scale manufacturing. Another drawback of optical 3D scanners is the time required for measurement and subsequent polygonal digitization; it is longer than for laser sensor-based geometry measurements.

4   Measurement Module Design

The measurement module formed a component of a robotic processing station for grinding of turbojet engine blades (Fig. 6). The robotic processing station comprised an ABB industrial robot manipulator with a dedicated gripper, the measurement module, the ForceControl expansion for grinding, electrospindles, and a suite of other machining tools.
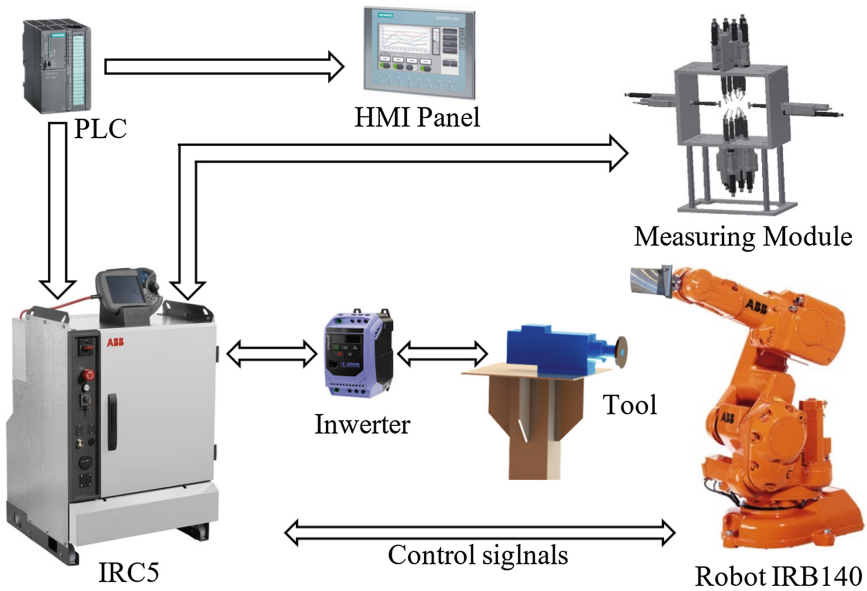
Fig. 6. Schematic diagram of the robotic processing station for grinding of turbojet engine blades.

A review of the available solutions for the quality control of turbojet engine blades and their deficiencies prompted the decision to develop a proprietary measurement module. Three measurement module concepts were prepared, and one of these was implemented. Each of the concepts had an accurate 3D-CAD model developed with a simulation of its functioning in RobotStudio, a software suite for offline robot programming. At the concept development stage it was assumed that the quality control of blades would be made by measuring the thickness at three blade sections (Figs. 2 and 3) and the chord at each blade section. The measurement sensors, the key elements of each of the three concepts, were functionally tested on a selection of blade types. The selection criteria of the final solution of the measurement module were: measurement accuracy, feasibility of communication with the robot controller over Ethernet TCP/IP, and price.

The first measurement model concept was based on 2D distance laser scanners and six laser spot sensors (Fig. 7). The measurement module with the 2D distance laser scanners was stationary. A blade was fastened in the gripper of an industrial robot manipulator with a repeatability rating of 0.06 mm. The position of the gripper with the blade was determined by measuring the distance with the laser spot sensors. The blade geometry, based on the measured distance, was determined with an algorithm implemented in the robot controller. This solution was tested with Keyence LJ-G080 series 2D scanners from Keyence.
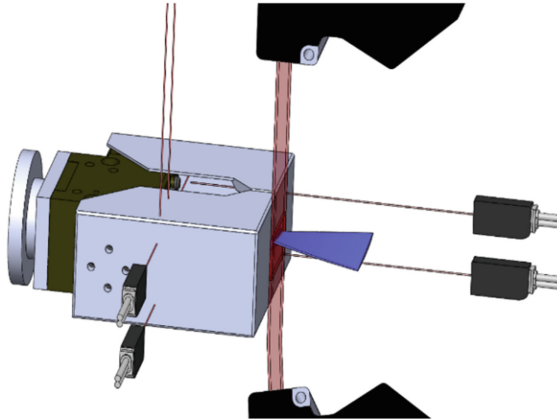
Fig. 7. Measurement module concept with 2D scanners.

The measurements with the sensors were based on a system of triangulation. The laser beam formed a series of points on the surface of the test object. The reflected laser light was picked up by a detector, which measured the reflected laser light's angle of incidence. The determined angle of incidence defined the distance from the surface of the test object. The main specifications of the Keyence LJ-G080 sensor were:

- Measurement distance: 80 ± 23 mm,
- Laser beam width: 32 ± 7 mm,
- Measurement accuracy: 0.05 mm.

The exact position of the gripper holding the blade was determined with 6 laser spot sensors installed in a linear arrangement on three walls of the measurement module. The laser spot sensors were a part of the determination of blade position within the measurement module. The test object positional data was sent to the robot controller, which enabled the correction of positioning errors imposed by the accuracy and repeatability of the robot manipulator. The laser spot sensors chosen were Keyence LV-S61.

The measurement model concept with the 2D scanners was tested. An example of the tests is shown in the photograph in Fig. 8a. The measured value master was the measurements made with a GOM 3D scanner and processed in ATOS Professional (Fig. 8(b)).

The tests demonstrated the feasibility of the measurement module concept with the 2D scanners. However, a number of problems were observed, which consisted of light reflections from the blades after grinding, which disturbed the measurement output data. The price of the solution was twice that of the finished solution.

The second measurement module concept (Fig. 9) was also based on 2D scanners, with the measurement module in motion relative to a stationary blade held by the robot manipulator gripper. The mobile measurement module solution was intended to eliminate the measurement errors imposed by the repeatability of the robot manipulator.
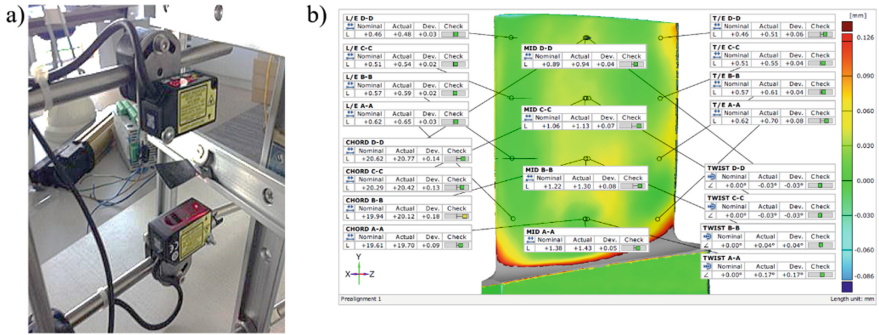
Fig. 8. (a) Testing of the solution with the 2D scanners, (b) Measurement results obtained with the GOM 3D scanner.
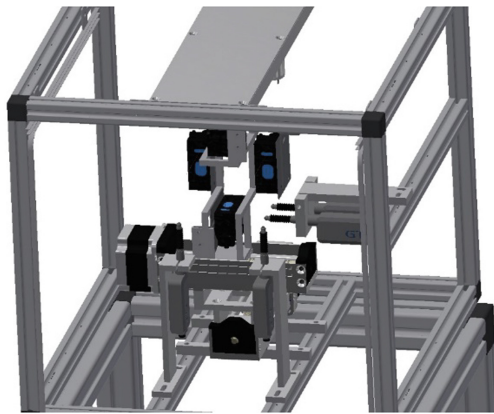
Fig. 9. The mobile measurement module concept.

The motion of the 2D scanners in this concept was made by a linear motion module, comprising aluminium sections, guides, and ball screws. The ball screws were driven by 57BYGH series stepper motors manufactured by WObit.

A CAD model, detailed engineering drawings and RobotStudio simulations were prepared for this measurement module concept. The concept was not implemented due to the issues with precision manufacturing and calibration of the linear motion modules. The price was higher than the price of the first measurement module concept.

The third measurement module concept was based on contact sensors. The concept featured a stationary measurement module with the blade (test object) moved by a robot manipulator. The repeatability error of 0.06 mm by the robot manipulator was deemed acceptable. The contact sensors chosen were GT2-A32 manufactured by Keyence, with the rated accuracy and travel of 0.003 mm and 32 mm, respectively. Figure 10 shows a photograph from the preliminary testing of this solution.
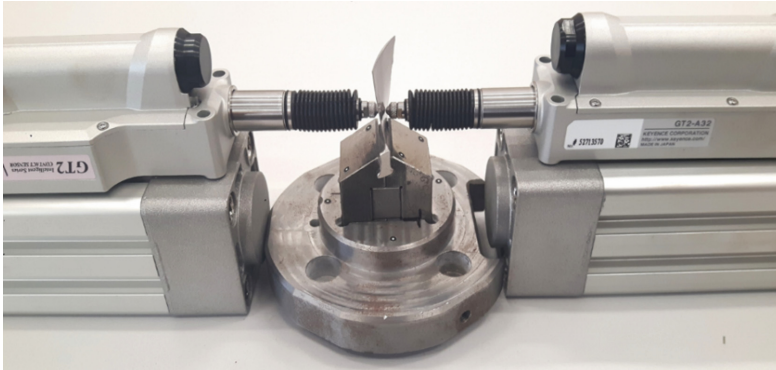
Fig. 10. The mobile measurement module concept. Geometry measurement with Keyence contact sensors.

The positive results of testing the solution with the contact sensors provided the foundation required for further development of the measurement module design (Fig. 11(b)). The contact sensor mounts were designed and fabricated as the casing of the measurement sensor, as shown in Fig. 11(a).
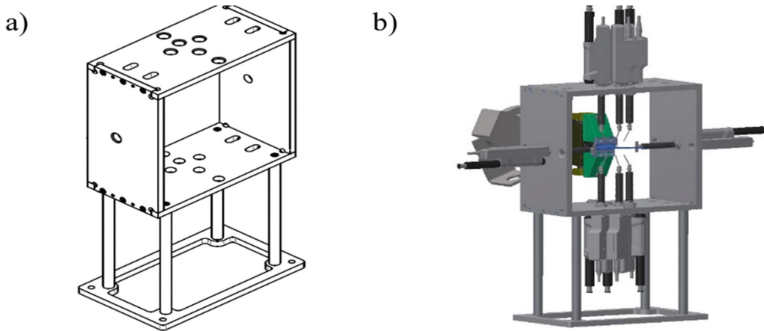
Fig. 11. (a) Part of the assembly drawing of the measurement casing, (b) CAD model of the measurement module.

The Keyence contact sensors were connected to a GT2-500 amplifier module, which could process signals from a maximum of five contact sensor heads. The amplifier model powered the contact sensor heads and processed the analogue electrical outputs from the contact sensor heads into the digital data required for the extension distance of each head. The digital data of the extension length of each head was then relayed to other devices connected to an Ethernet network via a DL-EN1 module. The data communication diagram is shown in Fig. 12.

GT2-A32   GT2-500   DL-EN1   Ethernet HUB   IRC5
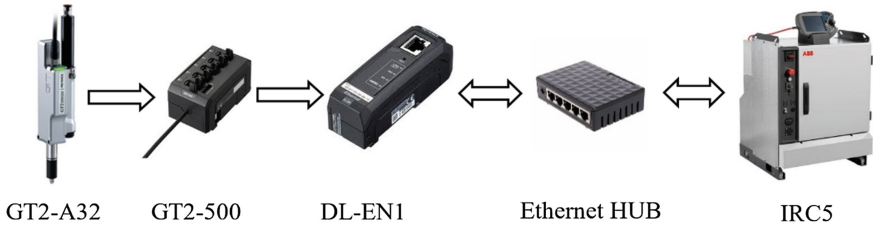
Fig. 12. Connection wiring diagram of the Keyence sensors and the robot controller.

The measured distance data was relayed over TCP/IP. The contact sensor head outputs were read by interrogating the data communication module with an ASCII command. Four command formats were available:

- command M0: read the measurement values from all connected sensors,
- command MS: read the measurement values and the output status of all connected sensors,
- command SR: read the operating parameters of a specific sensor,
- command SW: write the operating parameters of a specific sensor,
- command FR: read the decimal place count of the parameter of a specific sensor.

Each command format had a strictly defined content for the interrogation and reply data frames. The frame formats are shown in Fig. 13.
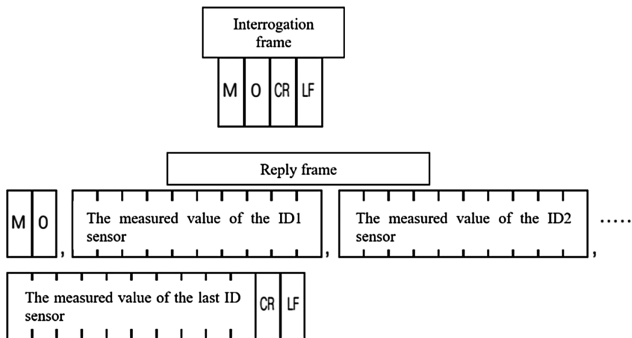
Fig. 13. Formats of the interrogation and reply data frames.

Each data frame ended with the control characters: CR (carriage return, ASCII 13) and LF (line feed, ASCII 10). Each applied GT2-A32 contact sensor featured a gas spring to control the motion of the contact sensing probe. The contact sensor heads required a compressed air supply, as shown in the diagram in Fig. 14.
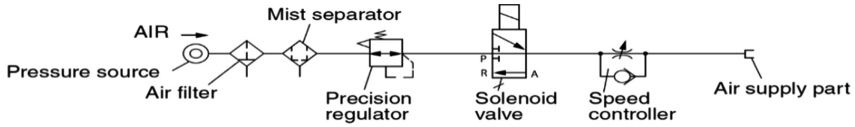
Fig. 14. GT2-A32 compressed air supply system.

The compressed air supply was controlled in this solution with a valve block manufactured by SMC (Fig. 15). The valve block was controlled by the robot controller over ProfiBus.

Fig. 15. Connection wiring diagram of the valve block and the robot controller.

The designed and fabricated measurement module is shown in Fig. 16(a). The measurement of a blade with the module is shown in Fig. 16(b).

a)   b)
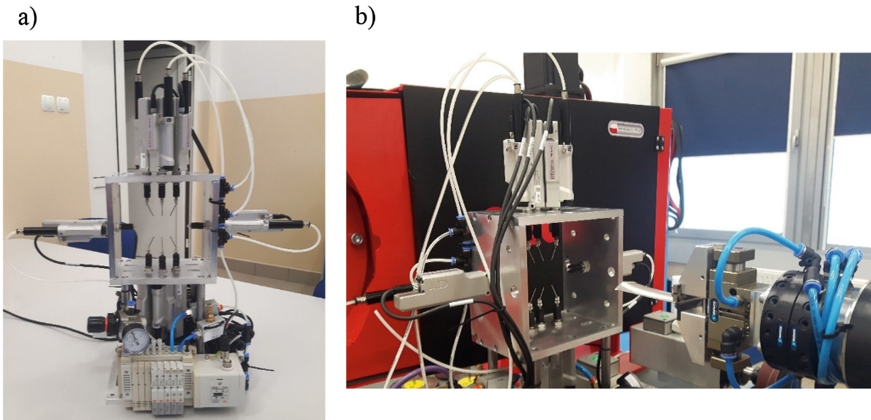
Fig. 16. Photographs of the fabricated measurement module, (a) Photographic overview of the module, (b) A blade introduced into the measurement module by the industrial robot manipulator.

The functional tests of the fabricated measurement module proved that the conceptual assumptions were valid. Once calibrated, the measurement module correctly measured the geometry of the blades by the distance detected by the contact sensor

heads. The distance data output was transmitted to the robot controller. A proprietary algorithm implemented in the robot controller could identify defective blades according to the measurements made by the contact sensor heads, display a measurement report and adjust the robot-assisted grinding process parameters to suit the measured geometric features.

5   Conclusion

This paper presents the process of designing and constructing a module for geometry and orientation measurement of turbojet engine blades. An overview of existing solutions was made and directed the authors to the deficiencies in commercially available measurement instruments. The authors proposed, modelled and preliminarily tested three measurement module concepts. Suboptimal solution selection criteria were adopted to qualify and fabricate the final concept for the measurement module. The designed and fabricated measurement module was intended for a robotic processing station for grinding of turbojet engine blades. The algorithm implemented in the robot controller to determine the blade geometry features based on contact sensor head measurements is a topic for future research papers. The functioning of the designed and fabricated robotic processing station for grinding of turbojet engine blades will be demonstrated to the public with the application of measurement feedback.

References

1. Burghardt, A., Kurc, K., Szybicki, D., Muszyńska, M., Szczęch, T.: Robot-operated inspection of aircraft engine turbine rotor guide vane segment geometry. Tehnicki Vjesn.-Tech. Gaz. **24**(Supplement 2), 345–348 (2017)
2. Burghardt, A., Kurc, K., Szybicki, D., Muszyńska, M., Nawrocki, J.: Software for the robot-operated inspection station for engine guide vanes taking into consideration the geometric variability of parts. Tehnicki Vjesn.-Tech. Gaz. **24**(Supplement 2), 349–353 (2017)
3. Burghardt, A., Kurc, K., Szybicki, D., Muszyńska, M., Nawrocki, J.: Robot-operated quality control station based on the UTT method. Open Eng. **7**(1), 37–42 (2017)
4. Gierlak, P., Burghardt, A., Szybicki, D., Szuster, M., Muszyńska, M.: On-line manipulator tool condition monitoring based on vibration analysis. Mech. Syst. Signal Process. **89**, 14–26 (2017)
5. Gierlak, P.: Hybrid position/force control in robotised machining. Solid State Phenom. **210**, 192–199 (2014)
6. Burghardt, A., Kurc, K., Szybicki, D., Muszyńska, M., Szczęch, T.: Monitoring the parameters of the robot-operated quality control process. Adv. Sci. Technol. Res. J. **11**(1), 232–236 (2017)
7. Burghardt, A., Szybicki, D., Kurc, K., Muszyńska, M., Mucha, J.: Experimental study of Inconel 718 surface treatment by edge robotic deburring with force control. Strength Mater. **49**(4), 594–604 (2017)
8. Hendzel, Z., Burghardt, A., Gierlak, P., Szuster, M.: Conventional and fuzzy force control in robotised machining. Solid State Phenom. **210**, 178–185 (2014)
9. Yilmaz, O., Gindy, N., Gao, J.: A repair and overhaul methodology for aeroengine components. Robot. Comput.-Integr. Manuf. **26**(2), 190–201 (2010)

10. Zhao, P., Shi, Y.: Posture adaptive control of the flexible grinding head for blisk manufacturing. Int. J. Adv. Manuf. Technol. **70**(9–12), 1989–2001 (2014)
11. Zhsao, P., Shi, Y.: Composite adaptive control of belt polishing force for aero-engine blade. Chin. J. Mech. Eng. **26**(5), 988–996 (2013)
12. Xu, X., Zhu, D., Zhang, H., Yan, S., Ding, H.: TCP-based calibration in robot-assisted belt grinding of aero-engine blades using scanner measurements. Int. J. Adv. Manuf. Technol. **90** (1–4), 635–647 (2017)
13. Li, W.L., Xie, H., Zhang, G., Yan, S.J., Yin, Z.: P: hand-eye calibration in visually-guided robot grinding. IEEE Trans. Cybern. **46**(11), 2634–2642 (2016)
14. Li, W.L., Xie, H., Zhang, G., Yan, S.J., Yin, Z.P.: 3-D shape matching of a blade surface in robotic grinding applications. IEEE/ASME Trans. Mechatron. **21**(5), 2294–2306 (2016)
15. Sun, B., Li, B.: Laser displacement sensor in the application of aero-engine blade measurement. IEEE Sens. J. **16**(5), 1377–1384 (2016)
16. Qi, L., Gan, Z., Yun, C., Tang, Q.: A novel method for aero engine blade removed-material measurement based on the robotic 3D scanning system. In: 2010 International Conference on Computer, Mechatronics, Control and Electronic Engineering (CMCE), vol. 4, pp. 72–75 (2010)
17. Zhang, Y., Chen, Z. T., Ning, T.: Efficient measurement of aero-engine blade considering uncertainties in adaptive machining. Int. J. Adv. Manuf. Technol. **86**(1–4), 387–396 (2016)
18. Kohut, P., Holak, K., Martowicz, A.: An uncertainty propagation in developed vision based measurement system aided by numerical and experimental tests. J. Theor. Appl. Mech. **50** (4), 1049–1061 (2012)
19. Godzimirski, J.: New technologies of aviation turbine engines. Works of the Institute of Aviation (in Polish) (4 (213)), 22–36 (2011)
20. Budzik, G.: Geometric accuracy of aircraft engine turbine blades. Rzeszow University of Technology Publisher (in Polish), 22–34 (2013)

Frequency Components of Signals Producing the Upper Bound of Absolute Error Generated by the Charge Output Accelerometers

Krzysztof Tomczyk$^{(\boxtimes)}$ and Marek Sieja

Faculty of Electrical and Computer Engineering, Cracow University of
Technology, Warszawska 24, 31-155 Krakow, Poland
{ktomczyk,msieja}@pk.edu.pl

Abstract. The paper presents an assessment of frequency components by the
time-frequency representation of signals with one constrain producing the upper
bound of the absolute error generated by charge output accelerometers. The
constraint concerns the amplitude resulting from the measuring range of an
accelerometer. This assessment was carried out by using a wavelet analysis
implemented in MATLAB. Mathematical basis regarding both modeling charge
output accelerometers and determining the absolute error were presented.
Shapes of signals producing the upper bound of error and results of analysis for
selected parameters of the accelerometer model are also presented and
discussed.

Keywords: Frequency component · Upper bound of error ·
Charge output · Accelerometer

1 Introduction

Signals with one constraint regarding the magnitude can be used for dynamic cali-
bration of measuring instruments through determination of the upper bound of error
[1–3]. In the case of an absolute error, its upper bound can be obtained for the final time
of instrument testing, and determining the signal shape that produces such an error is
related to the need to implement a dedicated computer program. Of course, it is also
necessary to develop the mathematical model of the reference system which should
meet the assumptions of the so-called non-distorting transformation [1].

In the case of such a task, assessing the content of the harmonic components of the
signal generating the upper bound of error seems to be interesting. In the work [4], only
relations between the amplitude and the frequency of the harmonic components of
signals with two constraints (regarding both the magnitude and velocity) were deter-
mined. Such signals generate the upper bound of dynamic error for the criterion of
maximum distance. This work, however, does not present the relationship between the
amplitudes of these components and the time of their occurrence.

In turn, the time-frequency analysis of signals with two constraints for the case of
integral-square error and on the example of the voltage output accelerometer is pre-
sented in the paper [5]. In this work, the special coefficients for assessing the

correctness of determining the signals generating the upper bound of error were proposed. These signals were determined by using the genetic algorithm [6–8], while the wavelet decomposition was used for their analysis [9–12].

The time-frequency analysis of signals with one constraint producing the upper bound of absolute error [1] on the example of charge output accelerometer [13, 14] is presented in this paper. This analysis was based on the wavelet transformation implemented in MATLAB. The Morlet wavelet was used to analyze the signal components.

2   Block Diagram of Proposed Analysis

Figure 1 shows the block diagram of proposed method for analysis of frequency components of signals producing the upper bound of absolute error generated by the charge output accelerometers.
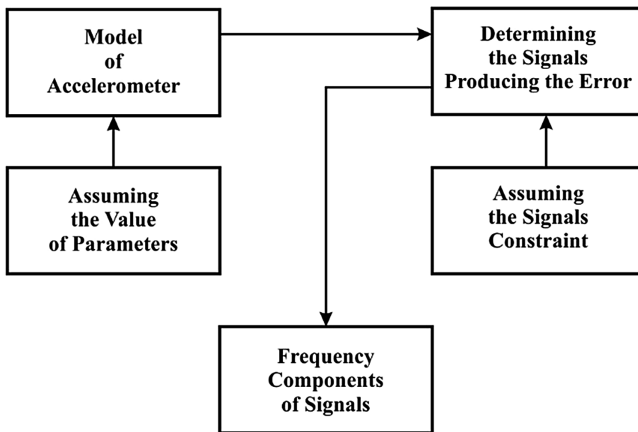
Fig. 1. Block diagram of proposed method. Own work.

The block described as *Model of Accelerometer* concerns the necessity of determining the model of the accelerometer selected for analysis. At this stage, the values of the accelerometer parameters which are the subject to analysis, are adopted. The block described as *Determining the Signals Producing the Error* refers to the algorithm that allows to determine the shape of the signal giving the upper bound of error for the assumed constraint of signal. The last block called *Frequency Components of Signals* refers to the analysis of the frequency components of signals producing the upper bound of error.

3   Model of Charge Output Accelerometer

The mathematical model of charge output accelerometer is described by the following transfer function [13, 14]

$$K(s) = \frac{(2\beta\omega_0 s^2 + \omega_0^2 s) S_V \tau}{\tau s^3 + (2\beta\tau\omega_0 + 1)s^2 + (\tau\omega_0^2 + 2\beta\omega_0)s + \omega_0^2} \tag{1}$$

where: β denotes damping ratio, $\omega_0 = 2\pi\omega f_0$ and f_0 is the non-damped natural frequency, $S_V = S_m S_e$ $[V/ms^2]$ is the voltage sensitivity, while S_m is the mechanical sensitivity, $S_e = k_p/C_t$ $[V/N]$ is the electrical sensitivity and $\tau = R_t C_t$. Parameters R_t and C_t represent the total resistance and capacity of the system that includes a charge output accelerometer, a cable and a voltage amplifier.

The state-space representation of (1) expressed by the observable canonical form is

$$K(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} \tag{2}$$

where $\mathbf{I}$ denotes the identity matrix while $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are represented by

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & -\omega_0^2/\tau \\ 1 & 0 & -(\tau\omega_0^2 + 2\beta\omega_0)/\tau \\ 0 & 1 & -(2\beta\tau\omega_0 + 1)/\tau \end{bmatrix},$$

$$\mathbf{B} = \begin{bmatrix} 0 & \omega_0^2 & 2\beta\omega_0 \end{bmatrix}^T, \tag{3}$$

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}.$$

4   Upper Bound of Absolute Error

The upper bound of absolute error is determined by the formula [1, 2]

$$E_A = A \int_0^T |k_a(t) - k_r(t)| dt \tag{4}$$

where A denotes the magnitude constraint, T is the time of accelerometer testing, while $k_a(t)$ and $k_r(t)$ are the impulse responses calculated on the basis of mathematical models of the accelerometer and its reference, respectively.

The mathematical model of reference is represented by the model of a higher-order filter. The maximum order of this filter is determined by the possibility of calculation the impulse response $k_r(t)$, as an inverse Laplace transformation calculated for the transfer function of the reference, using dedicated computer programs (MathCad, MATLAB e.t.c). The cut-off frequency f_c of this filter corresponds to the pass-band frequency of the accelerometer. An analog fifth-order Butterworth filter was used for the error calculations.

Denoting by $k_d(t)$ the difference between impulse responses included within the above integral, we can calculate the signal $x_u(t)$ that produce upper bound of absolute error as follows [1]

$$x_u(t) = A \cdot sign[k_d(T - t)] \tag{5}$$

5   Signals Decomposition

Figure 1 shows the shapes of signal $x_u(t)$ with one constraint that produce the upper bound of an absolute error for constant values: $S_V = 1 \, \text{V/ms}^2$ and $\tau = 0.05 \, \text{s}$ and for the change of two other accelerometer parameters as shown in Table 1.

Table 1. Parameters of the accelerometer for which the signals producing the upper bound of error were determined.

No.	Accelerometer parameters
1	$\beta = 0.05, f_0 = 1000 \, [\text{Hz}]$
2	$\beta = 0.05, f_0 = 2000 \, [\text{Hz}]$
3	$\beta = 0.1, f_0 = 1000 \, [\text{Hz}]$
4	$\beta = 0.1, f_0 = 2000 \, [\text{Hz}]$
5	$\beta = 0.3, f_0 = 1000 \, [\text{Hz}]$
6	$\beta = 0.3, f_0 = 2000 \, [\text{Hz}]$
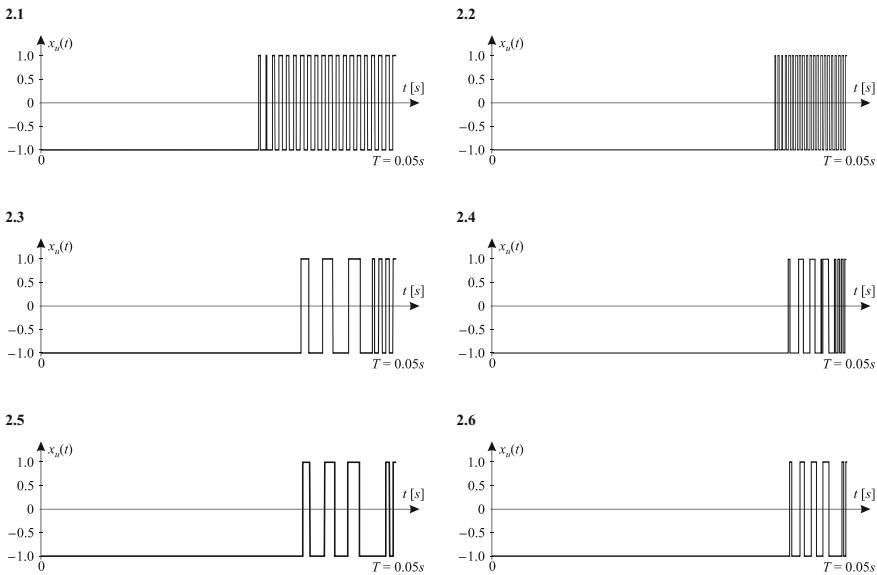
Fig. 2. Signals producing the upper bound of absolute error. Own work.

3.1
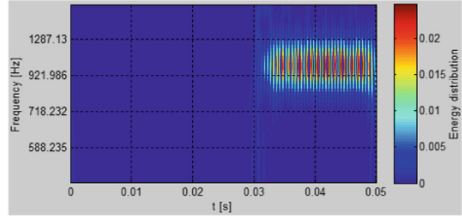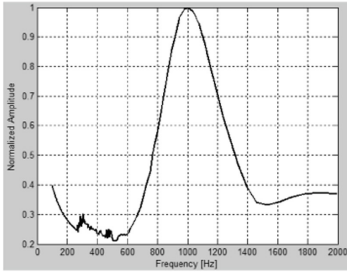
3.2
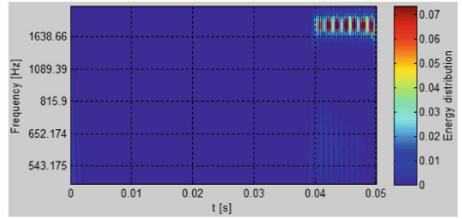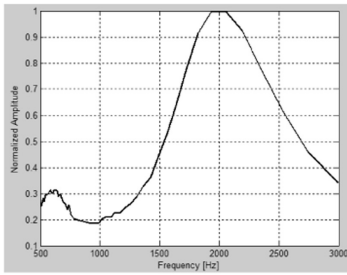
3.3
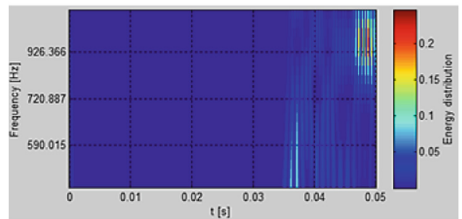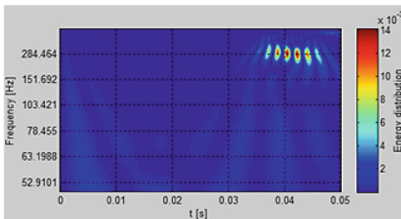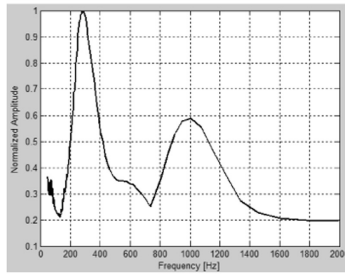
Fig. 3. Frequency components of signals $x_u(t)$ and corresponding energy distribution. Own work.

3.4
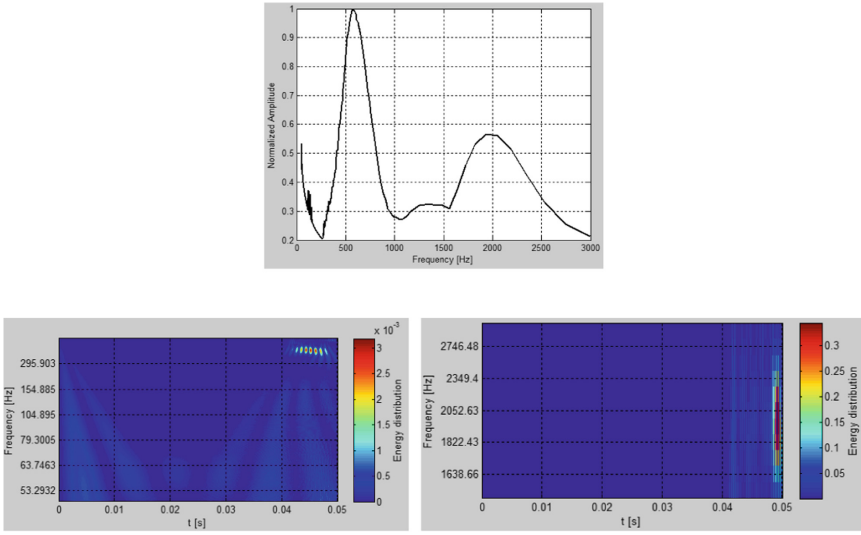
3.5
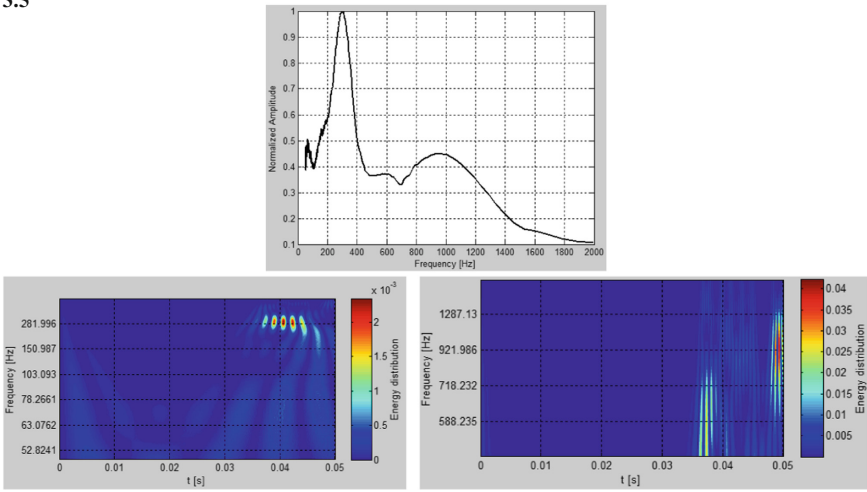
Fig. 3. (*continued*)

3.6
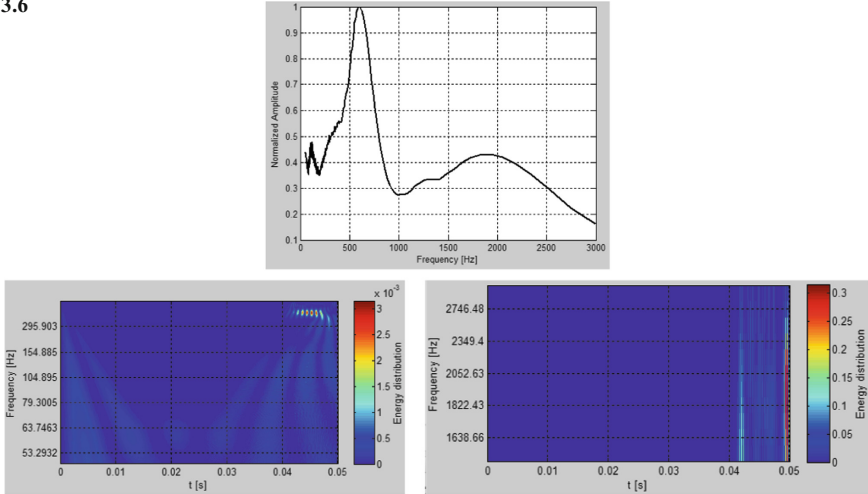
Fig. 3. (*continued*)

The numbering of particular signals shown in Fig. 2 corresponds to the parameter numbering from the first column in Table 1.

The change of the value of the parameter S_V does not affect the content of the harmonics of the signal $x_u(t)$, while the effect of the τ parameter on this content is insignificant and occurs in the range: (0.0001–0.001 s). Hence, this influence is not considered in this paper.

Figure 2 shows that with the increase of damping factor β, the number of time switchings of signal $x_u(t)$ decreases, while the increase of the non-damped natural frequency f_0 causes the time corresponding to the fixed value of this signal while increasing the number of its switching.

Figure 3 shows wavelet based decomposition of signals $x_u(t)$ from Fig. 2. The graphs which show the relationship between the *Normalized Amplitude* and *Frequency* allow to determine the dominant frequencies contained in the signal $x_u(t)$, while the graphs which represent the relationship between the *Frequency* and *Time* illustrate the time intervals in which dominant frequencies occur. Figure 3 uses numeration corresponding to the first column in Table 1, as in the case of Fig. 2.

In the case of Fig. 3(3.3–3.6) relationship between the *Frequency* and *Time* was determined separately for both dominant frequencies in order to more accurately presentation of *Energy distribution* for these components.

From Fig. 3(3.1 and 3.2) it results that for $\beta = 0.05$ one frequency is dominant and it is the non-damped natural frequency f_0. For $\beta = 0.1$ (Fig. 3(3.3 and 3.4)), the frequency corresponding to the pass-band (10%) of the accelerometer is dominant, which is also the cut-off frequency f_c of the filter being the reference to the error determination. Additionally, the influence of the frequency f_0 is also visible and it occurs near the end of the signal $x_u(t)$ duration. There is a frequency f_c immediately before this frequency.

From Fig. 3(3.5 and 3.6) it results, that the domination of the frequency f_c is apparent for $\beta = 0.3$.

In turn, the frequency amplitude f_0 is smaller by about 20% compared to the amplitudes in Fig. 3(3.3 and 3.4).

6   Conclusion

The paper examines the content of frequency components of signals producing the upper bound of an absolute error generated by the charge output accelerometer. Six signals obtained for different values of two parameters (damping ratio and non-damped natural frequency) of the accelerometer model were analyzed by means of the wavelet transformation.

As a result of the analysis, it was shown that the signals producing the upper bound of absolute error are dominated by two frequency components corresponding to the non-damped natural frequency and the accelerometer pass-band frequency. For small values of the damping ratio, the influence of non-damped natural frequency is dominant, while with the increase of this coefficient, the pass-band frequency gains the advantage.

The results of the analysis carried out in this paper may be useful from the point of view of checking the correctness of determining the signals which produce the upper bound of the absolute error.

References

1. Layer, E., Gawedzki, W.: Dynamics of Measurement Systems: Investigation and Estimation. Polish Scientific Publisher, Warsaw (1991)
2. Rutland, N.K.: The principle of matching: practical conditions for systems with inputs restricted in magnitude and rate of change. IEEE Trans. Autom. Control **39**, 550–553 (1994)
3. Hessling, J.P.: A novel method of estimating dynamic measurement error. Meas. Sci. Technol. **17**, 2740–2750 (2006)
4. Layer, E., Gawedzki, W.: Time-frequency properties of signals maximizing the dynamic errors. SAMS **11**, 73–77 (1993)
5. Tomczyk, K., Layer, E.: Energy density for signals maximizing the integral-square error. Measurement **90**, 224–232 (2016)
6. Goldberg, D.E.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, MA (1989)
7. Booker, L.B., Goldberg, D.E., Holland, J.H.: Classifier systems and genetic algorithms. Artif. Intell. **40**, 235–282 (1989)
8. Tomczyk, K.: Application of Genetic Algorithm to Measurement System Calibration Intended for Dynamic Measurement. Metrology and Measurement Systems, vol. XIII, pp. 93–103 (2006)
9. Mallat, S.A.: Tour of Signal Processing. Academic Press, New York (1998)
10. Huang, M.C.: Wave parameters and functions in wavelet analysis. Ocean Eng. **31**, 111–125 (2005)

11. Xiangcheng, M., Haibao, R., Zisheng, O., Wei, W., Keping, M.: The Use of the Mexican Hat and the Morlet Wavelets for Detection of Ecological Patterns. Plant Ecology, vol. 179, pp. 1–19. Springer (2005)
12. Bialasiewicz, J.T.: Wavelet-based approach to evaluation of signal integrity. IEEE Trans. Ind. Electron. **60**, 4590–4598 (2013)
13. Tomczyk, K.: Problems in modelling of charge output accelerometers. Metrol. Meas. Syst. **23**(4), 645–659 (2016)
14. Jyh-Cheng, Y., Ching-Bin, L.: System modeling and robust design of microaccelerometer using piezoelectric thin film. In: Proceedings of the IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems. Taiwan, pp. 99–104 (1999)

Differential Measurements of TGS Samples – Unbalanced Bridge Concept and Model

Mariusz Trybus[(✉)]

Institute of Mathematics and Applied Physics, Rzeszów Technical University,
Rzeszów, Poland
m_trybus@prz.edu.pl

Abstract. In this paper, we investigate a theoretical model of measurements of the phase transition in triglycine sulfate (TGS) samples using the idea of unbalanced bridge and four electrode sample setup. Four-electrode samples and differential measurements present some interesting results that can be explained by blurred phase transition in two areas of the sample. Our previous experiments proved some kind of order in the paraelectric phase, which resulted in a pyro-electric response to the fast temperature pulses above the critical temperature. We also observed a switching effect in the four-electrode samples. Many authors have investigated phase transition and structural changes of TGS and its doped crystals by various methods. We hope that the results of our model will encourage further researches to find the "evolution" of local order despite the order/disorder character of the phase transition in TGS single crystals.

Keywords: Phase transition · Pyroelectric sensor · Bridge measurements

1 Introduction

Triglycine sulfate (TGS) is one of the most comprehensively studied ferroelectric materials for infrared, non-cooled thermal detectors. TGS is a model uniaxial ferro-electric and is attractive because of its excellent pyroelectric properties, high figures of merit, and easy growth process that can be realized in a simple manner. There are various methods of growing TGS single crystals. The most popular (also used by our group) is growth from water solutions in a controlled evaporation process [1]. The elementary cell of the single crystal is monoclinic in both polar and non-polar phases. The ferroelectric properties of TGS were discovered in 1956 [2] and since that moment TGS has been assumed to undergo a second-order phase transition. The Curie temperature is about 49 °C.

In spite of intensive studies for over 60 years, TGS is still an interesting and "not fully discovered" material. Many authors have investigated the phase transition and structural changes of TGS and its doped crystals. Investigations have focused on the structural origin of its ferroelectricity and explanations for the mechanisms of the phase transition in the material. There are still questions even about the trigger of the phase transition in TGS [3, 4].

In our previous works, we investigated the pyroelectric phenomenon in TGS single crystals. We found some kind of order above the critical temperature [5]. A simple mathematical model of the switching effect in four-electrode samples [6] was suggested and we used compensation and differential methods [7] to investigate the pyroelectric phenomenon close to the Curie temperature in TGS samples. Some of the results regarding the pyroelectric properties might be evidence that in real TGS single crystals the nature of phase transition is more of the nucleation and growth. Researches on pyroelectric properties are very important from the application point of view. In our recent work [7], we have fabricated a set of samples with a four-electrode setup, as presented schematically in Fig. 1. We observed a switching pyroelectric current that might be proof of blurring of the critical temperature and existence of nucleation and growth processes in the crystal bulk. To confirm our observation, dielectric measurements were needed.

In order to measure the difference between the temperature characteristics of electric permittivity of two neighbouring areas of the four-electrode sample, we decided to implement the idea of an unbalanced bridge with two digital sine-wave generators.
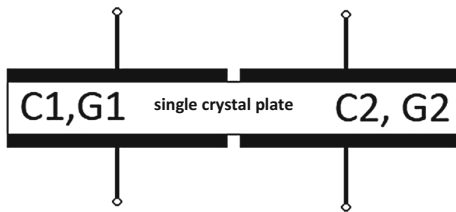
Fig. 1. Cross-section of four-electrode sample. C_1, G_1 and C_2, G_2 denote the conductivity and electric capacitance of the two sample areas respectively
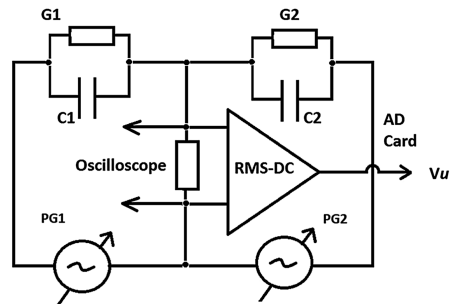
Fig. 2. Functional scheme of the measurement bridge.

The idea of the bridge is presented in Fig. 2, where G_1 & C_1 and G_2 & C_2, respectively, denote the conductance and capacitance of two neighbouring regions of the sample. The four-electrode sample is assumed to be placed in a thermostatic unit to ensure that the temperature of both crystal regions is the same. PG_1 and PG_2 are two digital power generators with regulated frequency, amplitude, and phase of the generated voltage waveform. RMS-DC is just a general idea of unbalance detection. The root mean square value (RMS) of the bridge unbalance sine voltage is converted to a direct current (DC) measured by an AD card.

The bridge can be balanced by changing the amplitude (V_1, V_2) and phase (Φ_1, Φ_2) of the output voltages of the power generators PG_1 and PG_2, respectively. In a state close to the balance, the ratio between the capacitances C_1 and C_2 should be determined based on measurements of voltages V_1 and V_2 and the unbalance voltage V_u. But in our case, we want to use the bridge for something different. We want to observe the difference in behaviour between two areas of one sample. If there is any delay or

different course of the phase transition of two neighbouring regions of the sample, we should detect the increasing unbalance voltage of the bridge V_u.

2   Mathematical Model

In order to check the concept of applying the unbalanced bridge in the measurements, we prepared a mathematical model in CAS MAXIMA. We assumed that neighbouring regions of the sample may vary in terms of their geometric parameters. We assumed

Table 1. List of parameters for model simulations.

Parameter	Unit	C_1 region	C_2 region
R – resistance of the sample	[Ω]	$(2 \div 6)10^6$	$(2 \div 6)10^6$
G – conductance of the sample 1/R	[S]	$(1.6 \div 5.0)$ 10^{-7}	$(1.6 \div 5.0)$ 10^{-7}
C_o – capacitance of the sample at room temperature	[F]	$(50 \div 60)$ 10^{-12}	$(50 \div 60)$ 10^{-12}
R_d – input resistance of unbalance voltage measurement channel	[Ω]	$(1 \div 10)10^6$	
E_1 – amplitude of PG$_1$ voltage	[V]	$1 \div 20$	
E_2 amplitude of PG$_2$ voltage	[V]	$1 \div 20$	
Φ_1 – phase of PG$_1$ voltage	[rad]	$0 \div 2\pi$	
Φ_2 – phase of PG$_1$ voltage	[rad]	$0 \div 2\pi$	
f – frequency of PG$_1$ and PG$_2$	[Hz]	1000	

that the temperature characteristics of electric permittivity may vary slightly (including blurring the temperature of phase transition). The parameters considered for model calculations (or its ranges) are listed in Table 1. The values of the parameters were established experimentally.

3   Model Equations

Assuming the parameters of elements given in Table 1 and configuration of the measurement system presented in Fig. 2, one can redraw the circuit for complex variables marked by underlining (Fig. 3) and define the circuit equation using the symbolic node potentials method given by Eq. 1.

$$\left(\underline{Y_1} + \underline{Y_2} + \frac{1}{R_d} \right) \underline{V_u} = \underline{E_1} \underline{Y_1} + \underline{E_2} \underline{Y_2} \tag{1}$$
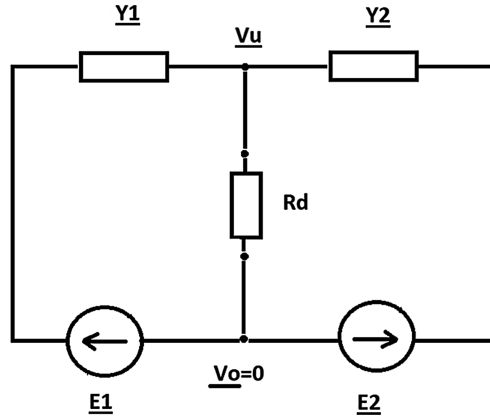
Fig. 3. Electric circuit of measurement bridge in complex domain. Complex values are underlined. Vu – complex value of reference node voltage, assumed to be 0; V_u – complex value of unbalance voltage; E_1 and E_2 – complex values of PG$_1$ and PG$_2$ generators' voltages; Y_1 and Y_2 – admittance of two neighbouring regions of the sample; R_d – input resistance of RMS-DC converter.

In order to simulate the behaviour of the temperature characteristics of each of the admittances Y_1 and Y_2, one has to consider the temperature dependence of each of the (region) capacitances C_1 and C_2. The mentioned relations are given by Eq. 2.

$$C = \varepsilon_o \varepsilon_r(T) \frac{S}{d} = \varepsilon_r(T) C_o, \underline{Y} = G + j2\pi f C \qquad (2)$$

where:

C – capacitance; C_o – capacitance at room temperature; ε_o – dielectric permittivity of the vacuum; ε_r – relative dielectric permittivity of the material; T – temperature; S – capacitors plates area; d – dielectric layer thickness; Y – admittance; G – conductance; f – frequency; j – imaginary unit.

According to thermodynamic theory, the electric relative permittivity ε_r depends on the temperature T. Below the critical temperature T_c, it is described by a function proportional to $1/(T_c–T)$ in the ferroelectric phase and by a function proportional to $1/(T–T_c)$ in the paraelectric phase.

$$C = \begin{cases} \frac{C_o}{2(T_c - T)}, & T < T_c \\[2mm] \frac{C_o}{(T - T_c)}, & T > T_c \end{cases} \qquad (3)$$

Keeping in mind all of the above and using Eq. 3 to calculate C_1 and C_2, it is possible to redefine the circuit equation Eq. 1, considering the critical behaviour of capacitive elements at the phase transition point.

$$\left(G_1 + G_2 + \frac{1}{R_d} + j2\pi f(C_1 + C_2)\right)\underline{V_u} = \underline{E_1 Y_1} + \underline{E_2 Y_2} \tag{4}$$

Equation 4 was used to determine and simulate the bridge unbalance voltage V_d given by the complex equation Eq. 5. The temperature characteristics of complex voltage V_u were calculated in the range of temperatures 45–55 °C.

$$\underline{V_u} = \frac{\underline{E_1 Y_1} + \underline{E_2 Y_2}}{G_1 + G_2 + \frac{1}{R_d} + j2\pi f(C_1 + C_2)} \tag{5}$$

Table 2. List of parameters used in model simulations.

Parameter	Unit	C_1 region
T_c – critical temperature of phase transition	[°C]	49.0
G – conductance of the sample	[S]	$1.67 \cdot 10^{-7}$
C_o – capacitance of the sample at room temperature	[F]	$50 \cdot 10^{-12}$
R_d – input resistance of unbalance voltage measurement channel	[Ω]	106
E_1 – amplitude of PG_1 voltage	[V]	5
E_2 amplitude of PG_2 voltage	[V]	5
Φ_1 – phase of PG_1 voltage	[rad]	0
Φ_2 – phase of PG_1 voltage	[rad]	2π
f – frequency of PG_1 and PG_2	[Hz]	1000
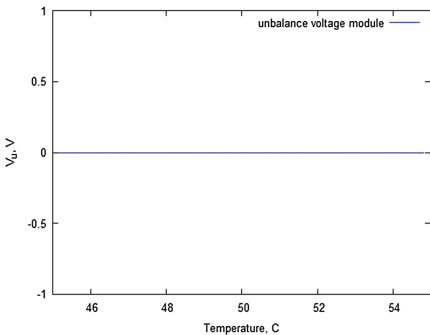
Fig. 4. Modulus of V_d (C_2 region parameters: Tc = 49.0 °C,   G = $1.67 \cdot 10^{-7}$S, C_o = 50 10^{-12} F)
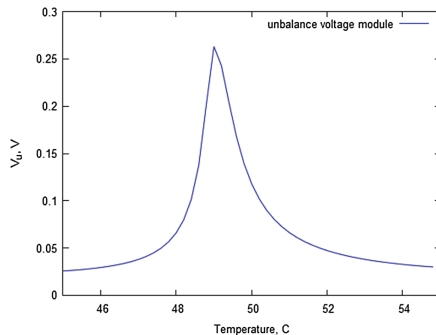
Fig. 5. Modulus of V_d (C_2 region parameters: Tc = 49.0 °C, G = $0.95 \cdot 10^{-7}$S, C_o = $45 \cdot 10^{-12}$ F)
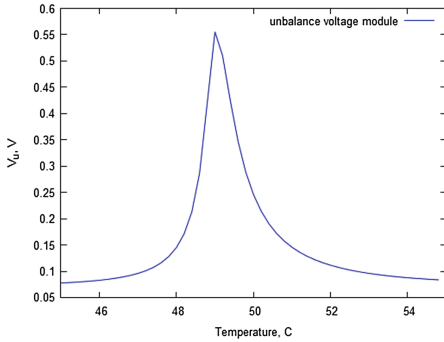
Fig. 6. Modulus of V_d (C_2 region parameters: $T_c = 49.0$ °C, $G = 0.83 \cdot 10^{-7}$S, $C_o = 40 \cdot 10^{-12}$ F)
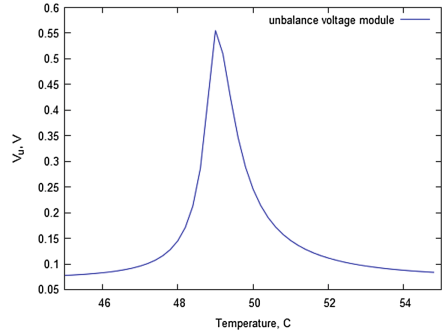
Fig. 7. Modulus of V_d (C_2 region parameters: $T_c = 49$ °C, $G = 1.11 \cdot 10^{-7}$S, $C_o = 45 \cdot 10^{-12}$ F)
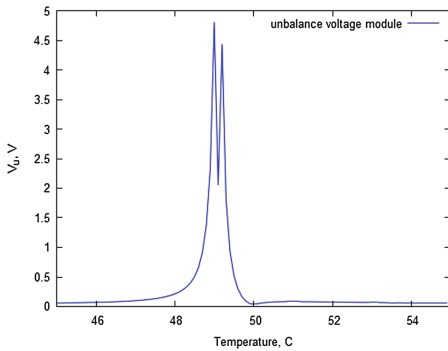
Fig. 8. Modulus of V_d (C_2 region parameters: $T_c = 49.2$ °C, $G = 1.11 \cdot 10^{-7}$S, $C_o = 45 \cdot 10^{-12}$ F)
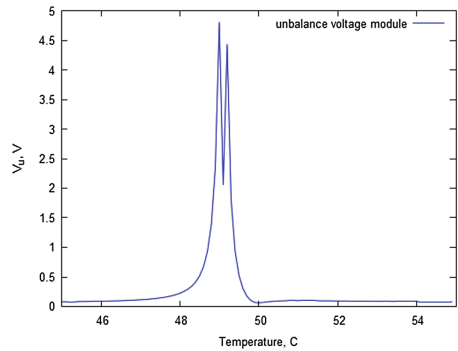
Fig. 9. Modulus of V_d (C_2 region parameters: $T_c = 49.2$ °C, $G = 0.83 \cdot 10^{-7}$S, $C_o = 40 \cdot 10^{-12}$ F)

4 Simulation

CAS MAXIMA was used to simulate the temperature characteristics of the unbalance voltage Vu. We assumed several possible setups for the bridge and parameters of the four-electrode sample. The results of the simulations of the modulus of V_u for the selected setups are presented in Figs. 4, 5, 6, 7, 8 and 9. Region C_1 was taken as the reference and its parameters were not changed during the simulation. The combined parameters of the circuit assumed in the model together with the parameters of the reference region C_1 are presented in Table 2.

5   Results and Discussion

The electric response of the TGS sample is a derivative of changes of numerous electric parameters of the single crystal (i.e. spontaneous polarization, electric permittivity and conductivity, etc.). Many authors study the pyroelectric response and dielectric properties of TGS in close proximity to the critical temperature T_c, because it may reveal the secret of the phase transition mechanism. In the case of complex phenomena, the various components may influence each other and sometimes weak components may be covered by stronger ones. In most such cases, differential or bridge methods are recommended for detection of weak elements of the signal.

The method presented in the paper using the idea of an unbalanced measurement bridge seems to be promising with regard to its high sensitivity and good temperature resolution observed in the model charts. In Fig. 4, one may observe the reaction of the C_1 and C_2 regions with the same electric parameters as are assumed in the model. There is no response in the differential unbalance voltage calculated from the simulation. The next figures, Figs. 5, 6 and 7, present the results of simulations for various electric parameters of the regions C_1 and C_2 but still with the same T_c in both regions. One can observe a single peak with limited amplitude. In the simulation results, the characteristic double peaks of unbalance voltage can be observed only in cases where the course of the phase transition is slightly different in each sample region. The amplitude observed in model calculations is one order of magnitude higher, which means the method should be sensitive to such issues in real measurements. Such behavior in measurement experiment should be the evidence of nucleation and growth type phase transition in the sample. The next step of our research is experimental verification of the model.

References

1. Renugadevi, R., Kanchana, G., Kesavasamy, R.: Growth and characterization of triglycine sulphate (TGS) single crystals. Elixir Cryst. Growth **55A**, 13033–13035 (2013)
2. Matthias, B.T., Miller, C.E., Remeika, J.P.: Ferroelectricity of glycine sulfate. Phys. Rev. **104**, 849 (1956)
3. Mnyukh, Y.: Second-order phase transitions, L. Landau and his successors. Am. J. Condens. Matter Phys. **3**(2), 25–30 (2013)
4. Drozhdin, S.N., Golitsyna, O.M.: Temperature and time behavior of the parameters of the domain structure of triglycine sulfate crystals near the phase transition. Phys. Solid State **54**(5), 905–910 (2012)
5. Trybus, M., Wos, B.: Dynamic response of TGS ferroelectric samples in paraelectric phase. Infrared Phys. Technol. **71**, 526–532 (2015)
6. Trybus, M., Paszkiewicz, T., Wos, B.: Observation of dynamics of hydrogen bonds in TGS crystals by means of measurements of pyroelectric currents induced by changes of temperature. Infrared Phys. Technol. **79**, 128–134 (2016)
7. Trybus, M.: Pyroelectric effect in tryglicyne sulphate single crystals – differential measurement method. Infrared Phys. Technol. **91**, 72–77 (2018)

The Mechatronic Device for the Hand and Forearm Rehabilitation

Jacek S. Tutak[✉] and Wojciech Kłos

Faculty of Mechanical Engineering and Aeronautics,
Department of Applied Mechanics and Robotics,
Rzeszow University of Technology, Al. Powstańców Warszawy 8,
35-959 Rzeszow, Poland
`tutak.j@prz.edu.pl, allegrowodec@wp.pl`

Abstract. The mechatronic device has been designed and implemented based on the comprehensive rehabilitation of the paretic upper limb. This system has been prepared for an individual approach to the recovery process including diagnostics, passive or active exercises with biofeedback and reports. The mechatronic device consists of a three basic cooperating parts: mechatronic frame with module for hand and forearm rehabilitation, glove for the hand rehabilitation and module for health hand. This mechatronic device was developed in order to realize a passive exercises and active exercises with paralysed limb using the healthy limb to conduct these exercises. A very important part of rehabilitation is to introduce biofeedback (e.g. visual and auditory) to motion exercises. This paper presents the main technical characteristics of the project, especially design, kinematics and dynamics of the device and the details of the hardware/software system. This paper suggests a new approach to the rehabilitation device for the spastic upper limb of stroke survivors. The functionality of the mechatronic device for hand and forearm rehabilitation have been presented during the first tests, and preliminary assessment of usability and acceptance is promising.

Keywords: Rehabilitation device · Hardware and software system Biofeedback

1 Introduction

In recent years, a growing interest in mechatronic devices for rehabilitation can be seen. It is associated with the aging population and an increasing number of neurological diseases. This paper presents the mechatronic device for hand and forearm rehabilitation which is especially dedicated for stroke survivors.

Stroke is a global health-care problem that is common and serious. In most countries, stroke is the second or third most common cause of death. Stroke is caused by damaged neuron connections in brain. The consequence of a stroke is one-sided paralysis, known as hemiplegia (spasticity), or one-sided weakness, known as hemiparesis. Approximately 17 million people had a stroke in 2010 and 33 million people have previously had a stroke and are still alive. Very disturbing statistics are saying that

only about 12% of patients after stroke achieve satisfactory motor recovery of the upper limb [1]. The World Health Organization (WHO) estimates that stroke events in EU countries are likely to increase by 30% between 2000 and 2025 [2].

Most of them will require treatment and rehabilitation [3]. Rehabilitation is the process of re-learning of movement [4]. It rebuilds these motor patterns which were damaged during illness or an accident. The basis of this process is plasticity and regeneration of the nervous system [5, 6]. Complementary to the traditional rehabilitation process is the use of rehabilitation devices. These devices are develop in order to accelerate the regeneration process. Rehabilitation devices provide not only the physical exercises but also introduce a biofeedback. Treatment and rehabilitation is a long, complex and requires a high level of regularity.

Many universities and companies adapt existing robots for rehabilitation requirements [7, 8]. The others, want to create new rehabilitation mechatronic device [9].

There are already some devices for the rehabilitation of people with hand's disabilities, allowing to move fingers and hands, based on the splint system. An example of this type of device can be Amadeo produced by the Austrian company Trymotion GmbH. This device provides the passive exercise (the device guides the passive movement of the fingers), the active exercise (training is adjusted to the capabilities of the patient) and the interactive treatment (active exercises are performed in the virtual space) [10]. Another example of this type of device is Bi-Manu-Track, which is produced by the German company Reha-Stim. The device provides the ability to perform exercises of probation/supination of the forearm and flexion/extension of the wrist. The movements are carried out in a mirror system - forced movement of the injured limb is carried out using healthy limb through the device [11].

The aim of these mechatronic rehabilitation devices is not to eliminate the therapist but to help in the very hard and demanding work the therapist must perform [12–14]. Additionally, the purpose of this project was also to reduce the cost of the rehabilitation. All of these features will increase access to exercises with novel and time/cost-efficient approaches to the rehabilitation.

2   Materials and Methods

In order to meet the expectations of the therapists and the patients, this mechatronic rehabilitation device has been designed and implemented. As it turns out, there is a large demand for modern mechatronic solutions for patients with various neurological diseases. Interdisciplinary and mechatronic approach to design, gave us the opportunity to prepare a modular concept of a device for a comprehensive rehabilitation of the forearm and hand.

The aim of the project was to develop a device for the rehabilitation, which provides conducting the controlled passive and active exercises of the hand and forearm of the patient. Most of exercises happen on the transverse plane and on planes between the transverse plane and the sagittal plane. The passive rehabilitation is done by using determined by the therapist sequence of movements of servomechanisms, to which the patient's hand is connected. In the active rehabilitation or active in relieving, the patient is trying to accomplish the exercise considering the stimuli sent by the device.

This developed concept of hand's exercises includes the capability to mobilize the affected limb by the healthy limb. The mobilization involves using extensometers - force sensors that are in a healthy hand module, enabling a healthy hand to set the range of the rehabilitation of the defective hand. The developed mechatronic device is equipped with a diagnostic system and reporting system of the progress of the rehabilitation, assisted by feedback from the patient (biofeedback) [15–17]. The system has analysis functions of power consumption during loading servos, which allow physiotherapist to control the quality of the work of the patient. The rehabilitative device described in the article is dedicated, especially to patients with a history of neurological or orthopedic injury (Fig. 1).
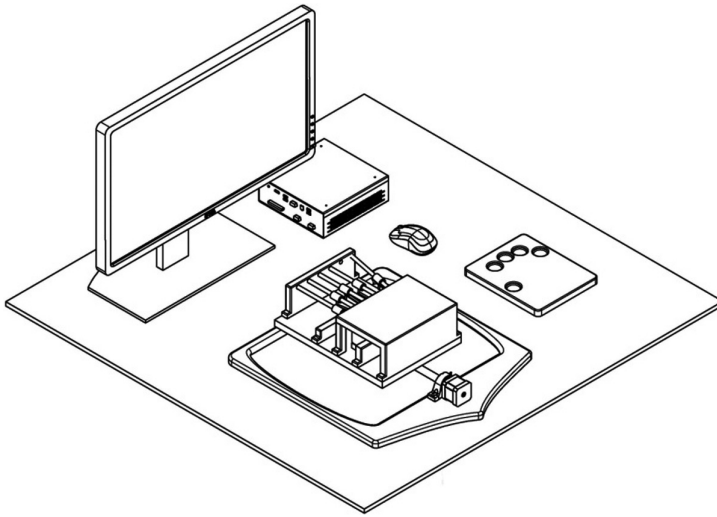
Fig. 1. The concept of the mechatronic device for rehabilitation.

The mechatronic rehabilitation device consists of:

- mechatronic frame with module for hand and forearm rehabilitation,
- glove with velcro for wrist and for each fingertip for the paralyzed hand,
- module for health hand,
- drive modules with sensors,
- hardware/software.

The device is equipped with a mechatronic frame, which consists of a base, to which a motor, connected with the ball-bearing shaft, is attached. The module for hand rehabilitation was placed on the shaft. The module for the rehabilitation of the hand includes a top and bottom base. The Velcro strap is placed on the upper base, which is simultaneously the clamping for the Velcro, sewn to the glove at the wrist of the rehabilitated hand placed in the glove. The mounted motor provides motion of the entire module to the rehabilitation of the hand, thus, providing the opportunity to conduct exercises of the forearm. Furthermore, the module for hand rehabilitation is responsible for the rehabilitation of the fingers. Shafts that cooperate with linear

bearings were placed on the inside of the module for the hand rehabilitation. The items of Velcro were attached by glue on the upper side of bearings. They are simultaneously the fastening for the Velcro placed on fingertips of the rehabilitated hand located in the glove. The metal screws were attached on the lower side of the bearings, in order to connect them with the cranks that together with servomechanisms form the crank - crosshead system. The servos are placed on the lower base of the described module. Considering the thumb, there is a possibility to change the settings of the shaft, thus introducing an additional range of exercises. The mechatronic frame is connected to the cables, i.a. with the right side module of the healthy arm, on which in the milled places for fingertips embedded five extensometers. With the module of healthy hand, it is possible to conduct and control the rehabilitation of the disabled hand by the other healthy hand. Work of the device is controlled through the CPU placed on the right side of the mechatronic frame and connected by a cable, and through the specialized pro-prietary computer program, through which it is possible to carry out the convalescence process. The described mechatronic device allows to carry out the rehabilitation, diagnostics and reporting. Moreover, information is displayed on the monitor on the left side of the CPU [18].

Figure 2 shows the CAD design of the device for the rehabilitation, developed in Inventor Professional 2017.
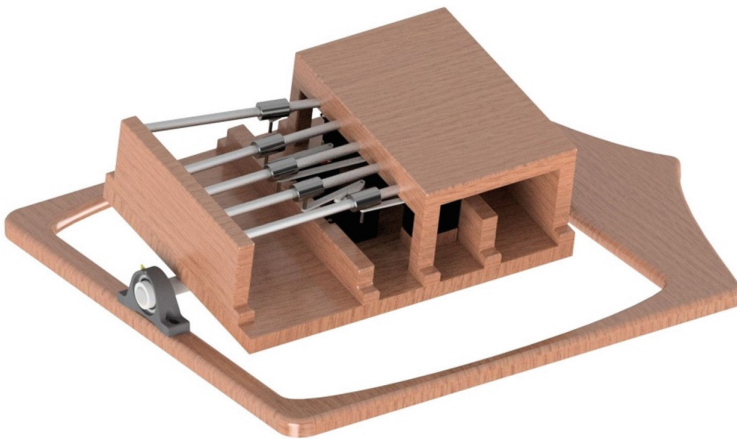
Fig. 2. The concept of the mechatronic frame with module for hand and forearm rehabilitation.

3   Kinematic and Dynamics Theory

In order to propose a design solution for this type of device as well as the selection of appropriate components for the client, it becomes necessary to conduct a detailed analysis of the kinematics and dynamics of the manipulator. The crank mechanism convert alternative translational movement to a continues rotating movement, or vice versa. The kinematic parameter of the mechanism: the linear displacements, the linear velocity and the linear acceleration of the fingers connection are obtained through relations (1), (2), (3) (Figs. 3 and 4).

$$x_M = r \cdot \left[1 - \cos(\varphi) + \frac{\lambda}{2} \sin^2(\varphi) \right], \tag{1}$$

$$v_M = r \cdot \omega \cdot \left[\sin(\varphi) + \frac{\lambda}{2} \sin^2(2\varphi) \right], \tag{2}$$

$$a_M = r \cdot \omega^2 \cdot [\cos(\varphi) + \lambda \cos(2\varphi)], \tag{3}$$

Taking into account the dependence determining the forces of inertia of the masses moving in a reciprocal motion is described with the formula:

$$F_{bp} = -m_{bp} \cdot r \cdot \omega^2 \cdot [\cos(\varphi) + \lambda \cos(2\varphi)], \tag{4}$$

and the masses moving in a rotational motion with:

$$F_{bo} = -m_{bo} \cdot r \cdot \omega^2, \tag{5}$$

where:

m_{bp} - a mass of elements in a slider-crank system making a to-and-fro motion
m_{bo} - a mass making a rotational motion

Equation describing the motion of the model [19]:

$$M_{bp} = \frac{F_{bp} r \sin\left(\varphi + \arccos\left(\sqrt{1 - \frac{\sin^2 \varphi}{n}} \right) \right)}{\sqrt{1 - \frac{\sin^2 \varphi}{n}}}, \tag{6}$$
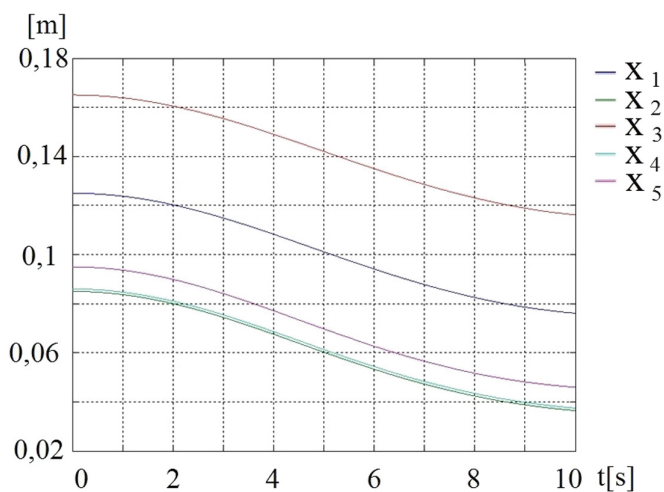
Fig. 3. Moving the sliders.
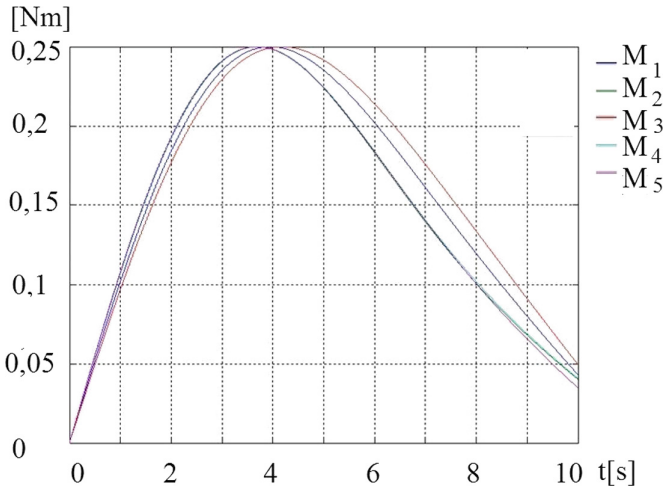
Fig. 4. Torque in each servo

4   The Most Important Components and Their Installation

The servomechanisms were used as drives of movable elements of the machine. The selection was based on the calculated moments of the analysis of the dynamics of the machine and the requirements of engineering [19–21]. Taking into consideration the future requirements of the device, including accuracy, safety and high reliability of future work, it was finally decided on the Hitec servos number HS-311. Drivers were chosen for servos. It was decided to use 6-channel servo controller Pololu Micro Maestro USB which operates digital and analog signals. The communication is carried out via the serial interface (UART). The interface allows to connect the Maestro to the CPU, which can be any microcontroller, in this case Rasbperry Pi. The operation of the servos is made by sending a special command character.

The Fig. 5 shows a diagram of an electronic board used to read-out the current of the individual servos. This information is very important from the point of view of the reporting the progress of the rehabilitation. The developed electronic circuit is designed to convert analog signals into digital. The voltage drop is measured on the resistor of the servomechanism that is proportional to the current flowing through the resistor. The servo signal is subjected to filtration by a simple low-pass filter consisting of a resistor and a capacitor. The signal after filtering, is amplified by the operational amplifier, and then goes to the correct input of the integrated circuit MCP3008 – of the converter ADC. ADC converts the voltage on the 10-bit variable that is sent to the Raspberry Pi. There are also processed the signals from the extensometer, which then goes to the
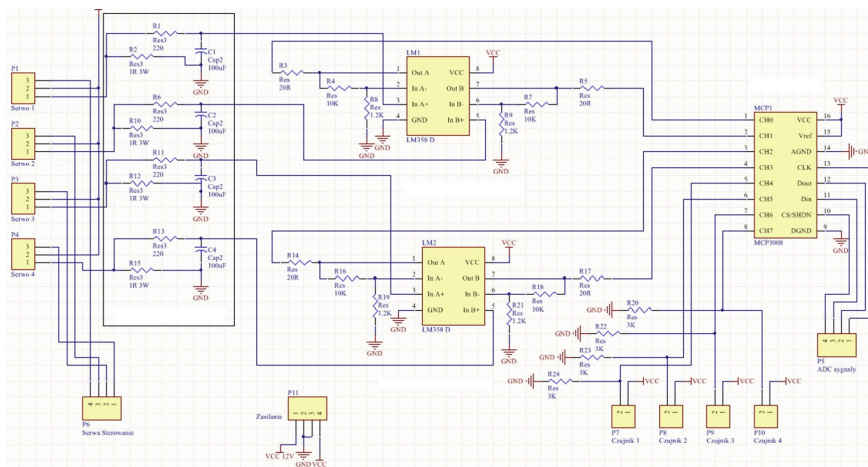
**Fig. 5.** The electronic diagram.

ADC [22]. The servo control program reads the current values of the voltages on the servos and analog values from the extensometer, and then sends the appropriate signals that control servos [23].

All electronic circuits are placed on a specially designed printed PCB board. It was decided to apply the THT and SMD technology. The board was prepared by the method of thermal transfer, and then etched in a solution of sodium persulfate. After cleaning the plate from the remaining toner, there were drilled holes and then soldered all the electronic components.

The most important element of the module of the healthy arm is the extensometer. The sensor reduces its resistance, while the force applied to the tip increases. Therefore, by using an analog and digital microcontroller it was possible to create a sensor that measures force.

The control of electronic components is performed by using a microcomputer Raspberry Pi model B. It is a microcomputer with the capabilities of a microcontroller. The GPIO connector (General Purpose Input/Output) mounted on the board can be used to connect sensors. The Raspberry is the ideal solution for this project because it gives the opportunity to work on LINUX and programme in C or Phyton.

A laboratory power feeder KA3005P was used to supply. It is a stabilized power feeder with programmable memory and smooth regulation of voltage in the range from 0 to 30 [V], and smooth regulation of current in the range of 0–5 [A]. It has anti-overtension and overload protection. Figure 6 shows a diagram of the device's developed connection.
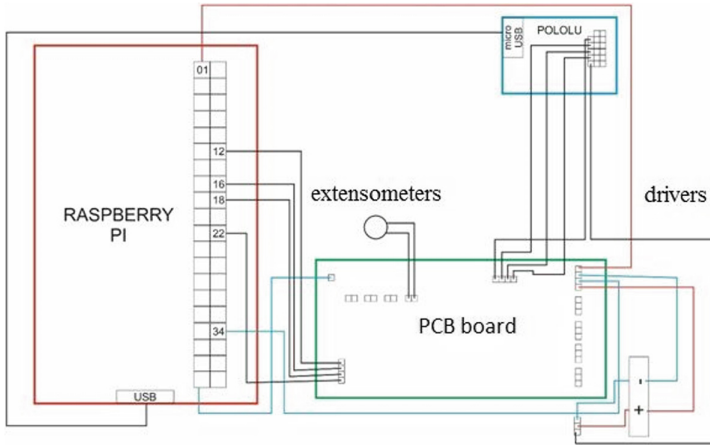
**Fig. 6.** The diagram of the device's connection.

5   The Software

An authorial application was written in Python programming language for the Linux platform for the Raspberry Pi microcomputer [24]. Python is a programming language of a high level of general purpose, includes a comprehensive suite of libraries, which causes that it is applicable both in education and commercial solutions. It is characterized by brevity and readability of a source code [25, 26]. Python is an interpreted language, i.e. a script or program can be written and used immediately without compiling to machine code.

To develop a graphical interface, a set of libraries pyQT was used. It allows the user to create friendly graphical interfaces for programs written in Python. Moreover, the program QT Creator was used to simplify the positioning of interface's graphical elements. It contains predefined functions, classes and objects, such as button, form, text box, etc. Figure 7 shows the main screen of the system developed to conduct passive exercises.

Number {1} indicates a bookmark with a choice of the type of the rehabilitation (active or passive exercises). The list of programs available in the passive rehabilitation is indicated by the number {4}. At the moment, the three programs for the rehabilitation of the fingers were developed, varying in the parameters of exercise (speed and range of motion) and implemented sequence - the order of movements of each finger. It is the movement of the fingers from the right to the left side and from the left to the right side, also flexing and straightening of all fingers simultaneously. It is also possible to compose individual exercises and add them to the list of programs. The clock visible in the middle of the screen shows the current exercise time {2}. The function button "Start" starts the previously selected program while the "Stop" button stop the current sequence. They are shown in the figure with the number {3}. "Exit" - this is the button corresponding to the exit from the system, it was determined by the number {5}.
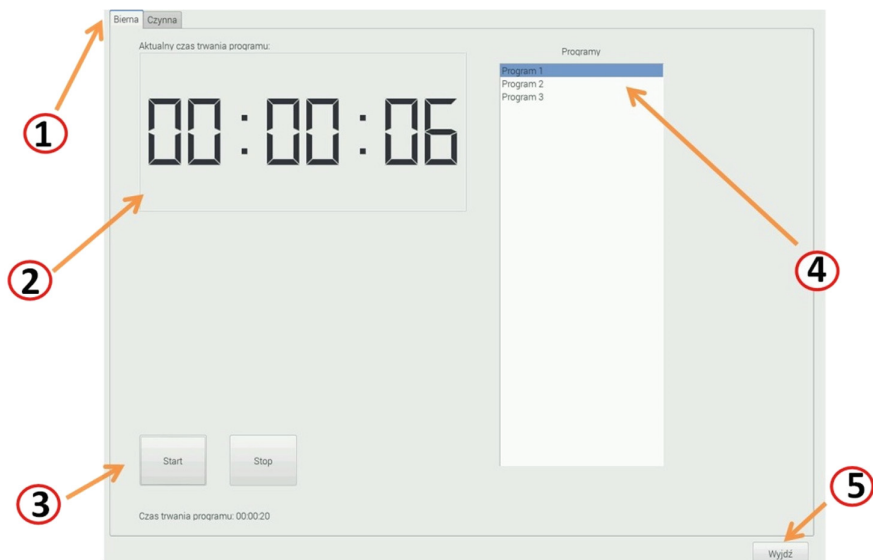
**Fig. 7.** The screen showing the passive rehabilitation.

Figure 8 shows another bookmark developed for the active rehabilitation. The patient realizes exercises of fingers. The diagrams that illustrate how much the patient affects the drive are placed in the central part of the active exercises' window.

The greater the force counteracting the movement of the servo, the higher the value can be seen in the diagram {3}. There is also the opportunity to present the reports of ranges of the movement.

Presented on the Fig. 5 graphs moments are very important from the point of view of the rehabilitation after neurological injuries where there is a spastic paralysis. This is the characteristic muscle contracture (clenched hand). A physiotherapist with a practice of gradually increasing the range of motion is trying to open the clenched hand of the patient. During the progress of rehabilitation, the spastic force must become smaller and therefore the range of motion of joints of each finger becomes increased.

It is possible to present all five fingers, or select a few of them. Figure 8 shows graphs for four fingers. The beam located on the charts, described by the number {2} indicates which graph is responsible for each finger. The bookmark used to change the level of rehabilitation is shown in the upper left corner.

A very important part of rehabilitation is to introduce biofeedback to motion exercises. This biofeedback is based on additional impulses for brain stimulation e.g. visual and auditory. With the aid of these stimulations, this mechatronic device is intended to speed up the process of recovery of damaged neuron connections in the brain. This biofeedback in virtual reality (VR) gives a possibility to compensate for some intrinsic limitations of traditional treatments. The patient has an opportunity to realize their own movements and discover reactions which are related to this movement. This mechatronic device contains a set of games in VR. These games were prepared to improve the mechanisms of perceptual and motor learning. Additional
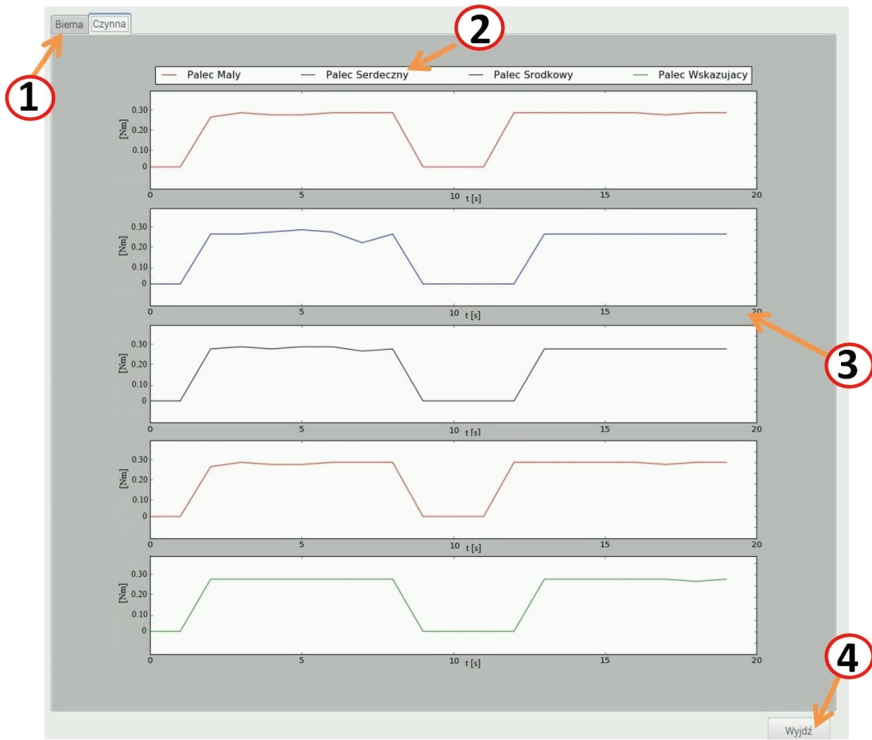
Fig. 8. The screen of the active rehabilitation with charts.

details about the other rehabilitation exercises in VR will be presented in a separate article.

This system has the possibility to generate reports. There is an option to generate a detailed report and a summary report of the rehabilitation exercises at any time during realization of the rehabilitation process. The report module was created based on the active exercises software program. A therapist can check the numerical and graphical results of rehabilitation collected over multiple weeks or months.

During designing of the mechatronic device, individualized approach toward each patient was not forgotten. Moreover, particular attention was devoted to the safety of the users (e.g. design for safety and software security). All of these features give a real chance to speed up the process of recovery of damaged neuron connections in the brain.

From the level of the active rehabilitation it is also possible to exit from the system, by clicking on the button "Exit" {4}.

6   Construction of the First Prototype and Its Testing

Wood was chosen as the basic building material of the prototype, due to its low cost and ease of modification. The wooden construction consists of six parts joined together by screws. The module drivers are servomechanisms, which are mounted by screws to the lower base of the rehabilitation hand module. Metal cranks were mounted to the servos. Cranks were cut from the metal sheet with a thickness of 3 [mm]. Then, there were drilled holes having a diameter of 2.5 [mm]. For the sake of safety, the crank in the next stage have been polished to avoid causing damage on the skin. Then, cranks were combined with the mounting screws that hold elements of the rehabilitated fingers. To fit the fingertips of each finger to the device, it was necessary to stick the screws with a diameter of 2.5 [mm] to linear bearings on shafts that are located between the vertical elements of wooden base. This developed system creates a slider-crank mechanism. The construction is light and simple. It weighs about 3 [kg]. The total cost of the prototype amounted to 400 (EUR). These are the costs of components of the built prototype. The estimated budget does not include the time spent on the construction of the prototype and on preparing the software. It is worth mentioning that currently the price of equipment for the rehabilitation of the upper limb on the market oscillate between 10 000 (EUR) up to 100 000 (EUR), it makes the purchase available for a larger audience. Figure 9 shows a module constructed for the hand rehabilitation.

Fig. 9. The prototype of the module for the hand rehabilitation.

Figure 10 shows the complete station consisting of a monitor, a wooden structure with a drive, power supply, box containing the electronics, and the module for the healthy hand.

During the first tests, the patient took the rehabilitation session, which consisted of the passive and active rehabilitation. The patient in the past suffered an injury of the spine. One of the consequences of the accident was the loss of muscle tension in the area of his left hand. The patient performed the series of exercises sessions in the range

Fig. 10. The prototype of the mechatronic device for rehabilitation.

of 2 weeks, the results of which are shown in Figs. 8 and 9. Each session began with the passive rehabilitation. The program for the passive rehabilitation contains three other programs. It includes the movement of the fingers from right to left and from left to right, then flexion and straightening of all fingers simultaneously. The passive rehabilitation consisted of the double repetition of the available programs. These exercises were regarded as a warm-up. Then they focus on the implementation of active exercises. The Fig. 10 presents the patient during one session of active exercises. In the upper left corner of the photo, can be seen graphs that are mapping of the torque applied by fingers. Details about these charts are also shown in Figs. 11 and 12. Figure 11 shows graphs during the active rehabilitation in the first days of exercises. The relationship between the moments of time can be observed on the graphs. The short spikes of moments are worth mentioning, because that means the patient is trying to straighten the fingers under load. Muscle tension of fingers are strong but short-lived. The patient has no possibility of the full range of movement for each of the fingers practiced.

Figure 12 shows graphs of the active rehabilitation in the last days of exercises. Based on the graph, it can be concluded that the patient after a test session has gained repeatability of the muscle tension. The movements of the fingers are stronger and longer, in relation to the first exercise using the device. Therefore, the range of motion was increased. Moreover, during the patient's cooperation with the device in both passive and active rehabilitation, there was no system errors. The device performs its determined movements. The station of the device is so simple that a patient who first uses it to rehabilitation, finds the system easy to operate and in an appropriate manner cooperates with the device.
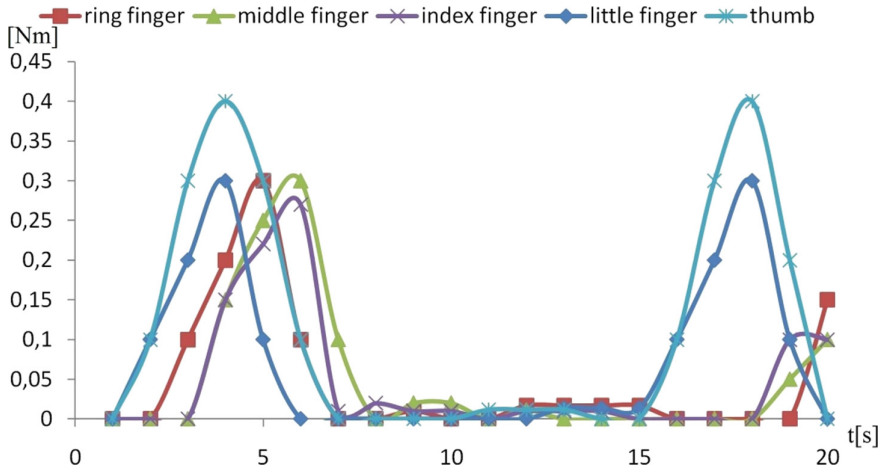
Fig. 11. The graph during the active rehabilitation on the first day of the rehabilitation.
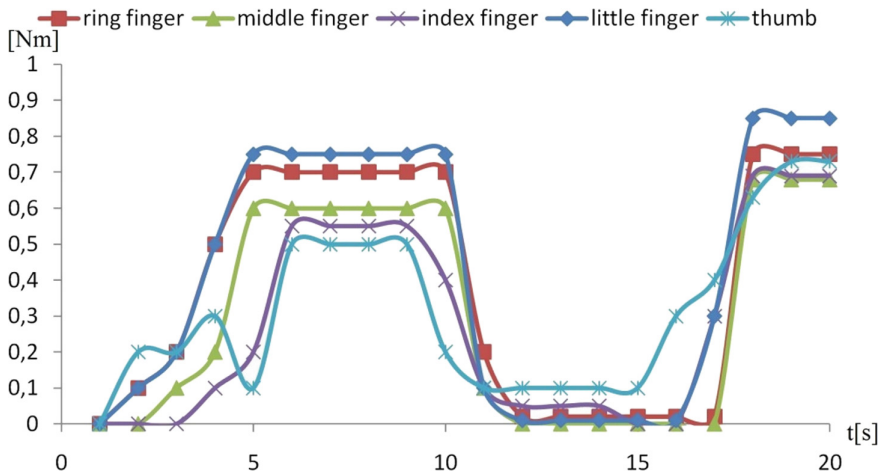
Fig. 12. The Graph during the active rehabilitation in the last days of the rehabilitation.

7   Conclusion

The article presents an innovative device, which allow to carry out the rehabilitation of the hand and forearm to help both doctors and patients. This developed device provides the ability to perform controlled passive and active exercises of the patient's hand and forearm, especially those affected by hemiplegia after diseases and neurological injuries (including a stroke). The passive rehabilitation is performed by determined by a therapist sequence of servo's movements, which are connected to the arm of the patient. In the active rehabilitation or in non-weight bearing exercises, the patient is

trying to make an exercise taking into account the therapist's suggestions and stimuli sent by the device.

It is worth noting, that the concept of the developed exercises of the hand includes the possibility of mobilizing the affected limb by the healthy limb. The mobilization involves using extensometers - force sensors located in the module of the healthy arm, allowing this hand to set rehabilitative exercises for the affected hand. This provides the possibility of faster recovery of the body.

An important element of the developed device is equipping it with diagnostic system and reporting the progress of rehabilitation, assisted by the feedback from the patient (biofeedback). The introduction of additional stimulation provides the ability to regain the lost physical fitness after injuries and neurological diseases, in a quicker and more interesting way. The system has features of analysis of the current consumption during loading servomechanisms, by which the physiotherapist can check the quality of the work of the patient.

It is also worth mentioning that there is the possibility of conducting rehabilitation exercises not only in rehabilitation centers, but also in private homes of patients. The device is small, lightweight, allowing to assemble, transport and place it in a desired position, in an easy way. The position requires only the constant access to the supply network. In addition, the proprietary software is very intuitive, and the information displayed is large and clear. While, the cost of the construction is more than 10 times low than other devices for the limbs' rehabilitation.

The conducted tests confirmed the previously established assumptions. The device successfully acts as a support tool for patients and physicians in the rehabilitation of the upper limb. The previously described patient, who in the past suffered an injury of the spine, while the consequence was the reduction of muscle tension around his left hand, developed significant progress in the course of the rehabilitation. Comparing the results obtained during the 2-week cycle of exercising, it significantly decreased muscle tension by changing the torque generated by the rehabilitated fingers in the area 0,3 [Nm], and has improved the range of motion of fingers in the field of 37 [% of the full open hand].

The developed prototype of the device is unilateral, which engages one side of the body, in this case the left hand. To encourage the potential consumer to purchase the device, there should be introduced the versatility, which guarantees the operation of the device in a bilateral way. This is a subject of ongoing work.

References

1. Kwakkel, G., Kollen, B.J., Grond, J., et al.: Probability of regaining dexterity in the flaccid upper limb: impact of severity of paresis and time since onset in acute stroke. Stroke **34**(9), 2181–2186 (2003)
2. Stroke Association, Research Spend in the UK: Comparing stroke, cancer, coronary heart disease and dementia 2014. http://www.stroke.org.uk/research-spend-uk. Accessed 9 Jan 2017
3. Dobkin, B.H.: Strategies for stroke rehabilitation. Lancet Neurol. **3**(9), 528–536 (2004)
4. Tutak, J.S.: Virtual reality and exercises for paretic upper limb of stroke survivors. TV-TG **24**(2), 451–458 (2017)
5. Hesse, S., Schmidt, H., Werner, C.: Machines to support motor rehabilitation after stroke: 10 years of experience in Berlin. J. Rehabil. Res. Dev. **43**(5), 671–678 (2006)
6. Budzik, G., Turek, P., Traciak, J.: The influence of change in slice thickness on the accuracy of reconstruction of cranium geometry. J Eng. Med. **231**(3), 197–202 (2017)
7. Fazekas, G., Horvath, M., Troznai, T., et al.: Robot-mediated upper limb physiotherapy for patients with spastic hemiparesis: a preliminary study. JJRM **39**(7), 580–582 (2007)
8. Lum, P., Burger, C.G., Van der Loos, M., et al.: MIME robotic device for upper-limb neurorehabilitation in subacute stroke subjects: a follow-up study. JRRD **43**(5), 631–642 (2006)
9. Gunasekara, M., Gopura, R., Jayawardena, S.: 6-REXOS: upper limb exoskeleton robot with improved pHRI. IJARS **4**, 1–13 (2015)
10. Kudasik, T., Libura, M., Markowska, O., Miechowicz, S.: Methods of reconstructing complex multi-structural anatomical objects with RP techniques. Bull. Polish Acad. Sci. Techn. Sci. **64**(2), 315–323 (2016)
11. Hesse, S.: Recovery of gait and other motor functions after stroke: novel physical and pharmacological treatment strategies. Restor. Neurol. Neurosci. **22**(3–5), 359–369 (2004)
12. Lo, H.S., Xie, S.Q.: Exoskeleton robots for upper-limb rehabilitation: state of the art and future prospects. Med. Eng. Phys. **34**(3), 261–268 (2012)
13. Pistohl, T., Joshi, D., Ganesh, G., et al.: Artificial proprioceptive feedback for myoelectric control. TNSRE **3**, 498–507 (2015)
14. Ozkul, F., Barkana, D.E.: Upper-extremity rehabilitation robot RehabRoby: methodology, design, usability and validation. IJARS **10**, 1–13 (2013)
15. Vieira, A., Gabriel, J., Melo, C., et al.: Kinect system in home-based cardiovascular rehabilitation. J Eng. Med. **231**(1), 40–47 (2016)
16. Kurc, K., Szybicki, D., Burghardt, A., et al.: The application of virtual prototyping methods to determine the dynamic parameters of mobile robot. Open Eng. **6**(1), 55–63 (2016)
17. Heinen, F., Lund, M.E., Rasmussen, J., et al.: Muscle–tendon unit scaling methods of Hill-type musculoskeletal models: an overview. J Eng. Med. **230**(10), 976–984 (2016)
18. Seung, S., Choi, H., Jang, J., et al.: Virtual wall–based haptic-guided teleoperated surgical robotic system for single-port brain tumor removal surgery. J. Eng. Med. **231**(1), 3–19 (2016)
19. Tutak, J.S., Wiech, J.: Horizontal automated storage and retrieval system. Adv. Sci. Technol. Res. J. **11**(1), 82–95 (2017)
20. Szuster, M., Gierlak, P.: Approximate dynamic programming in tracking control of a robotic manipulator. Int. J. Adv. Robot. Syst. **13**(16), 1–18 (2016)
21. Dong-Yuan, G., Xi-Fan, Y., Qing-He, Y., et al.: Robot sensor calibration via neural network and particle swarm optimization enhanced with crossover and mutation. TV-TG **21**(5), 1025–1033 (2014)

22. Kowalczyk, A., Szlachta, A., Hanus, R., Chorzępa, R.: Estimation of conditional expected value for expone ntially autocorrelated data. Metrol. Meas. Syst. **24**(1), 67–69 (2017)
23. Hendzel, Z., Burghardt, A., Gierlak, P., et al.: Conventional and fuzzy force control in robotised machining. Solid State Phenom. Trans. Tech. Publ. **210**, 178–185 (2014)
24. Tutak, J.S.: Design of ELISE robot for the paretic upper limb of stroke survivors. J. Vibroengineering **18**(6), 4069–4085 (2016)
25. Kowalczyk, A., Hanus, R., Szlachta, A.: Investigation of the statistical method of time delay estimation based on conditional averaging of delayed signal. Metrol. Meas. Syst. **18**(2), 335–342 (2011)
26. Hendzel, Z., Burghardt, A., Szuster, M.: Reinforcement learning in discrete neural control of the underactuated system. In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, Lotfi A., Zurada, Jacek M. (eds.) ICAISC 2013. LNCS (LNAI), vol. 7894, pp. 64–75. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38658-9_6

The Mechatronic Device Which Provides Comfort and Safety for the Elderly and Disabled People

Jacek S. Tutak[(✉)] and Wojciech Puzio

Faculty of Mechanical Engineering and Aeronautics,
Department of Applied Mechanics and Robotics,
Rzeszow University of Technology, Al. Powstańców Warszawy 8,
35-959 Rzeszow, Poland
`tutak.j@prz.edu.pl, wpuzio5@gmail.com`

Abstract. This article describes the device which provides comfort and safety for the elderly and disabled people. The system can be installed in any room, particularly in a flat or a house. The device is dedicated for the elderly and disabled people who are lonely. The project is based on the system that analyzes the pattern of the user's behavior. Signals are received from many sensors and transform to the form which is readable. Based on the collected data, it is possible to detect early symptoms of illness and to inform the family. The article describes in detail the developed system, including: its diagnostic capabilities (in case of decreased activity, longer stay in bed, consumption of less food, compared to previous period reported in the developed system, may indicate the occurrence of the first symptoms e.g. depression); reporting to the doctor's recommendations (in case the patient's compliance with the diet, movement or recommended amount of sleep); comfort (adjust lighting/temperature/humidity to the user's current behavior) or safety (information about dangerous situations, such as falling, or fading. The system observes a flat or a house while user is absent and can simulate presence through change light level or open/close windows). A great asset of the system is the low cost of implementation and no need to interfere with the existing infrastructure of the building.

Keywords: Mechatronic · Hardware and software system ·
Elderly and disabled people

1 First Section

One of the most frequently appearing problem are demographics and related to it changes. This results in an increase in the number of elderly people compared to other age groups. In 2012 there were many countries with similar statistics (e.g. in Europe, North America, Chile and Iran).

As a result, the average life span is extended. People have better life conditions and medical care. They more often achieve old age than before. As a result government must allocate more money and change health politics. Although people have longer life, they can't avoid most of adulthood diseases. Older people are more vulnerable to

environmental hazards. These factors often cause illness or death. However, all of the above can be offset by environmental adaptation or geriatric experience and knowledge. To slow down the speed of movement and cognitive abilities and to weaken the senses are natural things.

Among elderly people, the number of cases of depression and dementia is increasing. People need more care and control of them. In many cases this is not possible because the family does not have much money to hire professional medical care. In many countries, the government does not run programs that fund projects such as care for the elderly. Also relatives can not care for an older family member because they work and have many other responsibilities related to children [1, 2].

There was conducted market research to determine if there is a need for such a project. Smart home solutions are becoming more and more popular on the market. These types of products are dedicated especially for young and active people looking for technological innovations. Examples of such solutions are the following commercial projects (e.g. Control4 Smart Home, EasyLiving), patents (e.g. US 9230560 B2, US 6756998 B1) and universities projects (e.g. MavHome, House of Matilda) [3, 4].

Increasingly, even in less-developed systems, the user has the possibility to:

- check in on home from anywhere in the world
- with one touch, dim the lights, play music, turn up the heat, lock the doors and arm the security system;
- with one touch, put house to sleep, turn off all lights and lock all doors.

More complicated solutions use a combination of multidisciplinary technologies: artificial intelligence, multimedia technology, mobile computing, and robotics.

Moreover another solution like The Cerner Smart Room or IBM Smart Room are known solutions supporting patient care dedicated to applications in hospitals and focused on organizing work with the patient. As a result, it is known that the device described in this article can attract customers because there are no such solutions on the market.

This means that installing the above device in a room or apartment where an elderly person lives gives the access to health statistics Currently there are some works on home automation systems. However, this systems use sensors but does not save data. Our project receives signals from sensors and saves them to the database. In the next point we turn the information into a readable form. The results can be send to medical personnel and analyzed in term of health. There were used existing elements and the costs of the whole device was significantly decreased. All this is aimed at improving comfort, safety and health, especially for the elderly and the disabled.

2 Mathematical Model for Describing the State of Health of a Users

The purpose of this chapter is to suggest mathematical model for describing the state of health of a users of our system based on data from sensors over a given period of time (e.g. 1 year). This system is dedicated for the elderly and disabled people but another

people can also use this system. People of different ages and sexes have different susceptibilities to diseases, and diseases may act differently upon them. The potential users should be divided into subpopulations according to these variables. These subpopulations will be assumed homogeneous with respect to all demographic variables except age. Based on data from sensors, it is possible to index of health of population. This indicator includes also mathematical model. The state of health is best measured by the frequency and duration of dangerous symptoms (exceeding the defined limit values of sensor readings. It is symptoms of potential illness), by the severity of potential illness, and by the number of deaths. These components taken together give a comprehensive picture of health; separately, each describes an aspect of the state of health. This model is not specifically developed for a particular type of illness (defined limit values of sensors in the system.), the general line of approach applies equally well for any specific disease. The derived probability distribution characterizes the pattern of proneness and susceptibility of a subpopulation to disease; it also provides an easy means of calculating incidence and prevalence rates.

Consider the time interval of 1 year $(0, 1)$, and for each t, $0 \leq t \leq 1$, let the random variable $N(t)$ be the number of dangerous symptoms (overrun or low limit values of measurements from sensors) that an user has during the time interval $(0, t)$, with $N(0) = 0$. The purpose of this section is to derive the probability distribution of the random variable $N(t)$.

$$P_n(t) = \Pr\{N(t) = 0 | N(0) = 0\} \tag{1}$$

This presented probability function is an idealization of the proportion of people in the population having n dangerous symptoms, for $n = 0, 1, \ldots$, during the interval $(0, t)$.

Assumption from function (1) is that the probability of occurrence of an illness during the infinitesimal interval $(t, t+1)$ equals $\lambda_t h + o(h)$, where λ_t is a function of time t and $o(h)$ is a negligible quantity when h tends to zero. This means that the probability of an dangerous symptoms occurring within an infinitesimal time interval is a function of time and is independent of the number of previous dangerous symptoms. This assumption leads to a system of differential difference equations for $P_n(t)$. Consider two contiguous time intervals, $(0, t)$ and $(t, t+h)$. Exactly n dangerous symptoms can occur in the interval $(0, t+h)$ in three mutually exclusive ways: $(a)n$ dangerous symptoms will occur in $(0, t)$ and none in $(t, t+h)$ with a probability $P_n(t)[1 - \lambda_t h - o(h)]$. $(b)n - 1$ dangerous symptoms will occur in $(0, t)$ and one in $(t, t+h)$ with a probability $P_{n-1}(t)[\lambda_t h + o(h)]$; and $(c)n - 2$ dangerous symptoms or less in $(0, t)$ and two or more in $(t, t+h)$, with probability of $o(h)$. Consider these options together we get the formula:

$$P_n(t+h) = P_n(t)[1 - \lambda_t h - o(h)] + P_{n-1}(t)[\lambda_t h + o(h)] + o(h) \tag{2}$$

Transposing $P_n(t)$ dividing through by h, and taking the limit as h tends to zero, yield a system of differential difference equations.

$$\frac{d}{dt}P_0(t) = -\lambda P_0(t) \tag{3}$$

$$\frac{d}{dt}P_n(t) = -\lambda_t P_n(t) + \lambda_t P_{n-1}(t), n = 1, 2, \ldots \tag{4}$$

The first equation have solution:

$$P_0(t) = e^{-\int_0^t \lambda_\tau d\tau} \tag{5}$$

the remaining equations are solved successively to give the probabilities:

$$P_n(t) = \frac{e^{-\int_0^t \lambda_\tau d\tau}\left[\int_0^t \lambda_\tau d\tau\right]^n}{n!}, n = 1, 2\ldots \tag{6}$$

For a period of 1 year, for $t = 1$ the random variable N has distribution

$$P\{N = n\} = \frac{e^{-\int_0^t \lambda_\tau d\tau}\left[\int_0^t \lambda_\tau d\tau\right]^n}{n!}, n = 1, 2\ldots \tag{7}$$

Within the period of 1 year, the instantaneous probability $\lambda_t h + o(h)$ of occurring dangerous symptoms need to be dependent upon the time t and λ_t may be assumed to be constant. Under this assumption, we have ordinary Poisson distribution

$$P\{N = n\} = \frac{e^{-\lambda}\lambda^n}{n!}, n = 0, 1, 2, \ldots \tag{8}$$

Where λ is the expected number of dangerous symptoms during analysis period of time. To describe the health status of a users, we shall study the probability distribution of λ. The distribution of λ will be denoted by $g(\lambda)d\lambda$, the theretical proportion of people having the specified value λ. Since the sum of the proportions of individuals is unity, the function g satisfies the condition

$$\int g(\lambda)d\lambda = 1 \tag{9}$$

where the integral extends over all possible values of λ. The probability distribution of dangerous symptoms will be a weighted average of the probability function $g(\lambda)d\lambda$ employed as weights, that is:

$$P\{N = n\} = \int \frac{e^{(-\lambda)}\lambda^n}{n!}g(\lambda)d\lambda, n = 0, 1, 2, \ldots \tag{10}$$

The $g(x)d\lambda$ function is dependent upon the health condition of the particular group of users:

$$g(\lambda)d\lambda = \frac{\beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda}}{\Gamma(\alpha)} \tag{11}$$

where $\Gamma(\alpha)$ is defined by:

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy \tag{12}$$

The ranges of the constants for which (11) is defined are $\alpha \geq 0$ and $\beta > 0$. The function $g(\lambda)$ starts at $\lambda = 0$, increases as λ increases at a rate of:

$$\frac{d}{d\lambda} g(\lambda) = g(\lambda) \left[(\alpha - 1)\lambda^{-1} - \beta \right] \tag{13}$$

The expectation and variance of λ may be directly computed from (11):

$$E(\lambda) = \int_0^\infty \frac{\lambda(\beta^\alpha)}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} d\lambda = \frac{\alpha}{\beta} \tag{14}$$

1. and

$$\sigma_\lambda^2 = \int_0^\infty \left(\lambda - \frac{\alpha}{\beta} \right)^2 \left(\frac{\beta^\alpha}{\Gamma(\alpha)} \right) \lambda^{\alpha-1} e^{-\beta\lambda} = \frac{\alpha}{\beta^2} \tag{15}$$

The ratio $\frac{\alpha}{\beta}$ measures the average health of a population and the reciprocal of is the relative variance,

$$\frac{1}{\alpha} = \frac{\sigma_\lambda^2}{(E(\lambda))^2} \tag{16}$$

2. which is a measure of variation of health among individuals in the subpopulation relative to the mean health.
3. Assuming (11) as the function underlying the distribution of the population with respect to health condition, we have from (10) the probability function of the number of illnesses during the year:
4.

$$\begin{aligned} P\{N = n\} &= \int_0^\infty \frac{e^{(-\lambda)}\lambda^n}{n!} \left(\frac{\beta^\alpha}{\Gamma(\alpha)} \right) \lambda^{\alpha-1} e^{-\beta\lambda} d\lambda = \\ &= \frac{\Gamma(n+\lambda)}{n!\Gamma(\alpha)} \beta^\alpha (1+\beta)^{-(n+\lambda)}, n = 0, 1, 2, \dots \end{aligned} \tag{17}$$

This probability is the expected proportion of individuals in the population having n ilnesses during the year, taking inot account the variability among individuals in the population. The expected number of illnesses occurring to an individual in the sub-population is given by:

$$E(N) = \frac{\alpha}{\beta} \tag{18}$$

and the variance by:

$$\sigma_N^2 = \frac{\alpha(1+\beta)}{\beta^2} \tag{19}$$

5. Formula (17) represents a family of infinitely many probability distributions, depending upon the constants α and β. The health status of a subpopulation may best be described as a member of the probability distribution family for which α and β assume particular values, In order to estimate these values, it is necessary to know the observed frequency distribution of the number of illnesses occurring to the individuals of the subpopulation from which the mean $\bar{N}$ and variance S_N^2; of the number of illnesses are computed. Substituting $\bar{N}$ and S_N^2 for $E(N)$ and σ, respectively, in (18) and (19) and solving the resulting equations for $\sim$ and B give the estimates

$$\hat{\alpha} = \frac{\bar{N}^2}{(S_N^2 - \bar{N})} \tag{20}$$

$$\hat{\beta} = \frac{\bar{N}}{(S_N^2 - \bar{N})} \tag{21}$$

7. Using the estimated values $\hat{\alpha}$ and $\hat{\beta}$ (17), we have
8.

$$P\{N = n\} = \frac{\Gamma(n+\hat{\alpha})}{n!\Gamma(\hat{\alpha})} \hat{\beta}^{\hat{\alpha}} \left(1+\hat{\beta}\right)^{-(n+\hat{\alpha})}, n = 0, 1, 2, \ldots \tag{22}$$

The probability (22) multiplied by the total number of individuals in the subpopulation is the expected number of individuals having n illnesses during the year, for n = 0,1,2,... [5].

3   System Concept

The device supporting comfort and safety of older and disabled people consists of the control panel with a computer, a touch screen and the set of sensors (including motion sensors, force sensors, temperature and humidity sensors, open/closed door/windows sensors and barcode reader). The system has the ability to recognize any situation which can be dangerous for life and user's health (particularly while loss of consciousness or any illness symptoms) [6]. The device is dedicated for older and disabled people who live in solitude and don't take advantages of old people's home or family help. Simultaneously, they don't have money to hire the proper medical help. This solution supports not only users but also the family who has the access to observe the family member's behaviors and predict dangerous situations. Furthermore, the developed system can prepare reports on the recommendations of physicians, physiotherapists, diabetics. This is possible by tracking the motion activity of the user. A diet can be realized due to the sensors in fridge (the system recognizes when somebody takes something from a fridge) and barcode. It is possible to count calories and adjust them to the user. The software consists of two parts. The first part includes services which start while the system boot. They work all the time when device is on-line and there operate all connected devices. The second part provides the access to control the website and information from the database. For many customers the most essential is the comfortable interface. Fast development of the Internet gives the option to connect with the device via the remote access. The presented system is based on the client - server architecture. The server ensures the continuous service and gives permissions for the appropriate persons. The interface and the server is created by the framework called Flask. The framework allows building websites by Python programming language [7, 8]. Main advantages of this solutions are: making useful websites very quickly and many configuration options.

The device presented on the Fig. 1, shows the examplary room with whole installation and the most important elements. The control module includes computer and touch screen 1, which are the control module and the set of sensors in one part. The set of sensors consists of motion sensors 3, force sensors 4, cameras 5, sensors which respond on opening or closing doors 6, sensors related with recognizing the kind of food 7, radio sensors 8, temperature sensors 9 and humidity sensors 10. Motion sensors 3 are placed on the vertical runner 11 placed at 1.5 m height. Vertical runner 11 featured in toothed belt is driven by the electric motor 12 placed at the bottom. The motion sensor is simultaneously the passive detector of the infrared with LED light source 13 with electrical regulate pitch between floor and beam of light. Sensors that react on opening or closing are placed on the entrance door, windows, cupboards and the fridge door. Force sensors are located on a couch and other places where user sits down. Information about temperature and humidity come from sensors placed on wall, each of them in a different case. The data from all sources are streamed by the wire. The wire is installed under the roughcast. The touch screen is activated when sensors detect somebody nearby.
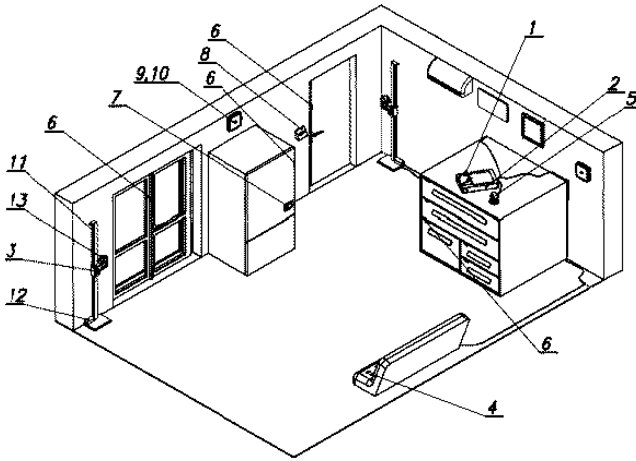
Fig. 1. Schema of the room where system was installed with the most important elements. (1 - computer with touch screen, 2 - case, 3 - motion sensors, 4 - force sensors, 5 - camera, 6 - open/closed sensors, 7 – sensor for the diet control, 8 - radio sensors, 9 - temperature sensor, 10 - humidity sensor, 11 - vertical runner, 12 - electrical runner motors, 13 - LEDs).

The most important feature of the system is to provide safety and comfort. Safety is understood as the ability to diagnose and report a patient's health as well as to protect a home from burglary. Comfort is understood from the point of view of the apartment itself, as well as from the point of view of ease of use [9–12].

Based on information gathered from presence detectors, the user's activity is monitored. In the event that the user drops his position relative to the presence sensor, the sensor moves along the guide down. By mounting the presence detector in the guide, it is possible to minimize the situation when the user is outside the presence sensor field. Such solutions allow the device to detect falls and loss of consciousness. Detecting a lack of movement by the device may indicate that it needs help and automatically informs the relevant person. Long-term limited activity of the user may indicate the presence of disease symptoms.

Thanks to the pressure sensors mounted in the seats, beds and chairs and cameras, the device not only allows to analyze the movement, but also to gather information about sleep patterns or correct positioning of the body in the correct position. This information can also be very important for the diagnosis of possible symptoms of illness.

With food sensors mounted on refrigerators or cabinets where food products are stored, there can be monitored eating habits, followed the recommended diet or regular meals. The device allows to generate reports that help to assess the current health.

With the dedicated software based on motion analysis, the device learns typical user behavior within a few days and can adjust the intensity of the light emitted by the LEDs mounted on the presence sensors not only to the current time of the day, but also to the situation. It is possible to increase the comfort of moving around the apartment at night in terms of the possibility of changing the gradual intensity of light as the user

wakes up and adjust the height and angle of incidence of light. In addition, the system seeks to recognize the user's behavior by adopting its current needs. The integrated lighting system allows to manage and adapt it to the user's mood or reduce the energy consumption.

Thanks to temperature sensors and humidity sensors, it is possible to increase the comfort of the rooms in which the device is installed by maintaining the appropriate values of these indicators.

Another feature of the solution is the ability to monitor and protect property. Ideal for this task are the presence sensors. With their help, it is possible to detect third-party traffic, launch cameras, and send relevant information to a pre-defined person, which may be at least the police. Open/closed sensors allow to check all windows and door whereas owner is absent. Moreover, the device has the ability to simulate the presence.

It is also worth emphasizing again that the system itself was designed not only for the user himself - an elderly or disabled person living alone, but also for his/her family. The system not only provides the ability to monitor and report on the daily behavior of the elderly, but it is also an ideal tool for alerting a family when a hazardous situation occurs (e.g. user fall) or alarms based on prediction of hazardous situations. Based on the analysis of the generated pattern of behavior and in case of sudden change, the first symptoms of the disease may be presumed (such as depression, long-term sleep, lack of recommended exercise or diet/medication limitations).

An important part of the system is to improve communication with medical personnel. This can be done through a video chat placed on the user interface or touch screen. Medical staff have access to data stored in the database through generated reports. They are especially helpful to doctors, rehabilitators, dieticians, to control compliance with their recommendations for users with small self-discipline. The information included in the reports may include physical activity in the home, regular medication and meals (particularly important for people with diabetes). These reports are based on information gathered from room sensors, door openers and other food lockers, and scanned food products using barcodes. The signs of the disease can also be seen in data from the pressure sensors in the bed and in the seats. A restless and erratic dream contributes to the feeling of being unwell or pain in the back of the spine. Demonstration of insomnia may also be an indication for medical personnel with various illnesses. And in conjunction with the data from other sensors, there can be found the cause. Insomnia affects more and more people and has many causes, including heart failure, hypertension, hyperthyroidism, stress or restless legs syndrome.

It is also worth noting that the described device is a modular system. It is possible to individually configure individual modules of the system to meet the needs of the user, with the possibility of retrofitting in the future with additional components.

A great asset of the presented solution is certainly the low cost of implementing the system and no need to interfere with the existing infrastructure of the building. The possibility of personalization of the described system exists not only from the point of view of its modularity but also from the point of view of the housing of the individual components depending on the arrangement of the rooms. Enclosures in which each module of the system can be modified depending on the decor of the room and the user's taste. This kind of capability was obtained even on the basis of the 3D printing used for the prototype [13–16].

For this type of project, it is also important to emphasize the possibility of fully protecting the user's privacy. The user decides what data and who can be accessed. This is a very important aspect of life. This also has a positive effect on the speed of convalescence of people with disabilities staying in their private homes, not in hospitals, with the ability to share basic living and behavioral data with authorized medical staff. The next very important thing is the modularity and compatibility of the system, as well as the fact that there is a possibility of cooperation of the described system with the solutions monitoring the basic parameters of life [17–21].

Control of the device as well as its communication with other devices is ensured thanks to the mounted central unit and dedicated software, which also ensures the realization of the described functions. The prototype uses the Raspberry Pi system, which has a great price advantage. The performance of the system and the ability to install Linux were the next criteria in favor of the selected electronics [22–25].

4   Software

The prototype software was written for a Linux distribution called Raspbian. This is a full fledged Debian based system. It provides the ability to run a server that supports client applications and to create the entire database required for operation. The task of the implemented algorithms is to read the corresponding values from all sensors and to save them in the database in a way that allows them to be analyzed later. The software consists of two parts. The first is a service running on the system, running all the time in the background, supporting the connected devices. The second part is server software that provides access to a web site, device control, and database information.

For the user the interface is a very important thing. Due to the rapid development of the Internet, access to the device not only via the touch screen, but also via the website is undoubtedly a great asset. In the presented device, the interface access was realized
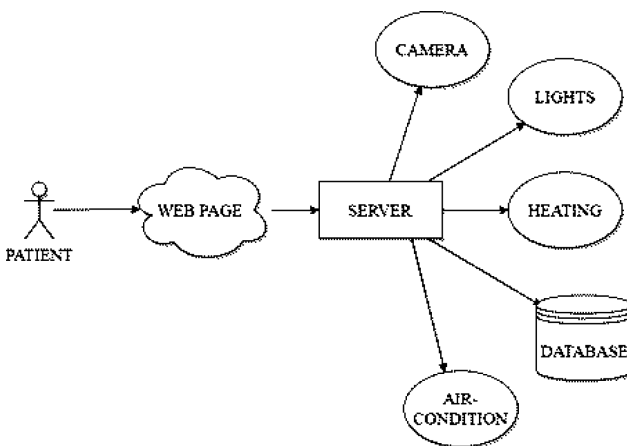
Fig. 2. Communication between patient and system.

in client-server architecture. The server ensures continuity of work and ensures proper verification of people who want to use the service (Fig. 2).

The server was based on a framework called Flask. It allows to build minimalist websites using the Python programming language. Its main advantages are the speed of creation and great configuration possibilities.

In the developed prototype, the user can access the interface through a web page and a display on the front panel. On the page there will be graphs drawn on the basis of data from the sensors. The user will also be able to interact via the control panel for household appliances, i.e. lighting, heating to match their preferences (Fig. 3).
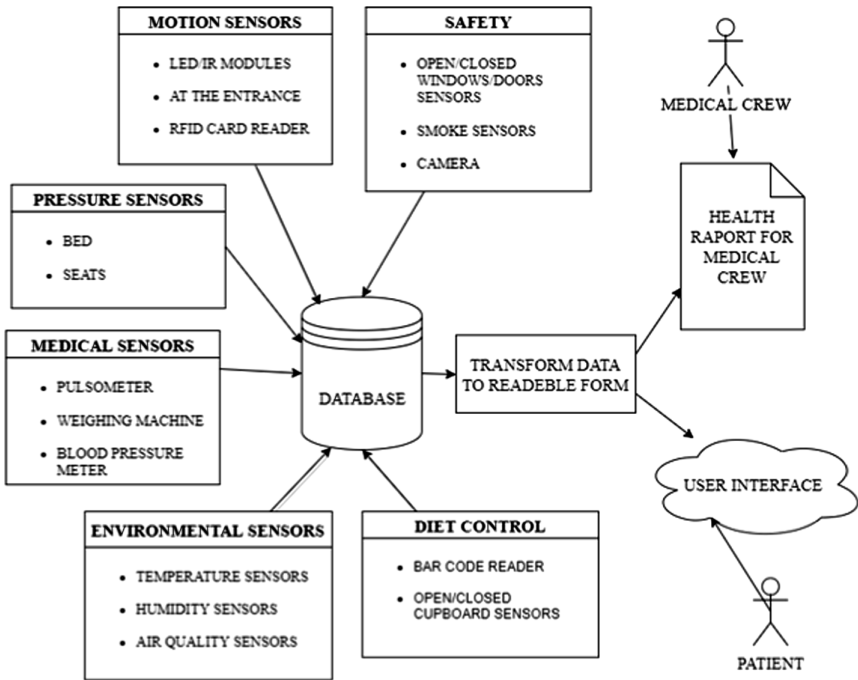
Fig. 3. Source of data which are saving in database.

5   User Interface

The user interface is implemented through the website (Fig. 4). Graphic design is based on the existing AdminLTE solution. It is distributed as open source software. In this project, the whole has been substantially simplified and modified to work in conjunction with the Flask framework. AdminLTE provides all necessary UI elements. There is a possibility to use icons, sliders, buttons, timelines, editors, tables and graphs. The JavaScript programming language and the jQuery library were used to develop the project.
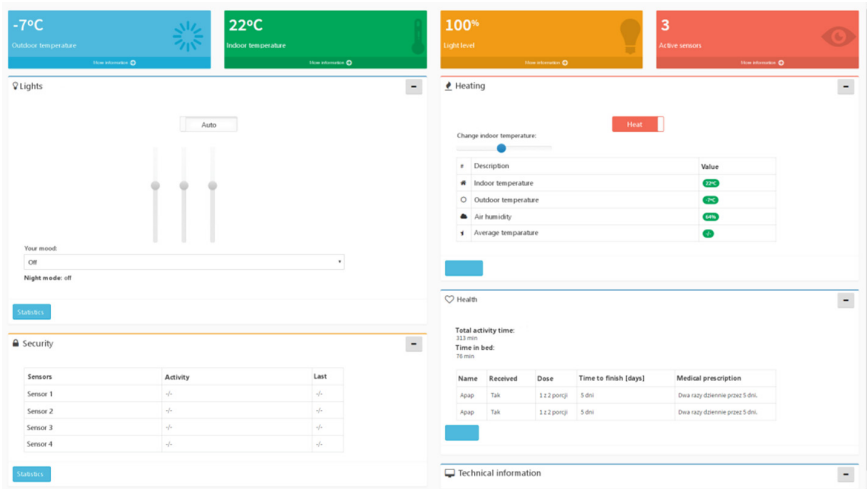
Fig. 4. The user interface is implemented through the website.

For ease of navigation, the user interface is divided into several sections, each of which is responsible for a separate component of the system. In addition to the configuration of the device in the section, statistics and the latest information gathered through the sensors are located. To illustrate how the section works, the article discusses only the selected ones: "Lighting," "Heating," "Safety," and "Health."

After entering the url, the user has the access to the "control panel" with which he or she can customize the room elements to his/her preferences. Each section is equipped with a button to the page with graphs showing the basic parameters and collected information.

Access to the home page is also displayed via mobile browsers. Scaling to display resolution is automatic (Fig. 5).

At the very top, the most basic information is presented in the headings (Fig. 5). In the first square the outside temperature is visible and the second contains the temperature inside the room. The third shows the percentage power of all light sources, when all are switched on with 100% brightness, the user also sees the value: 100%. The last square represents the number of sensors that detect the presence of a person at the same time

The first section deals with lighting, which provides an opportunity to interact with lighting elements. It is possible to adjust each light source separately or to select the automatic mode. In this mode, the intensity of each light source is determined by an algorithm that takes into account the motion detected by its individual sensor, the time of day, the time of year and the level of energy saving selected by the user. When the traffic is detected, the event information is also stored in the database. By switching to manual mode the sliders below are activated and the user gain the ability to control the power of individual LEDs according to his/her own preferences. By fitting the system in a variable color LED (RGB) bulb, users gain the ability to adjust the lighting to the mood.

The second section deals with heating. Due to this, it is possible to adjust the temperature of the air coming out of the blow. The main switch is responsible for switching the heating on or off. It has options such as heating and air conditioning. In heating mode the slider is activated by means of which the temperature can be adjusted. Below the slider is the table, the first line of the slider is updated on the current temperature in the room. The second line is the temperature outside. The third line contains the value of air humidity is a particularly important information for people with respiratory diseases. The last line shows the average daily temperature from inside the room. As in the previous section here also in the footer users will find a button with a link to the graphs showing the change of temperature outside in the given time period, the temperature from the inside, the air humidity and the average temperature.

The third section deals with patient safety. The information provided relates to the individual motion sensors. This information becomes particularly important in the absence of a patient. This section may also display information related to the safety of the patient and the elderly. In the event of a longer period of inactivity, the corresponding service may be notified at the time it was first registered. In addition, you have the ability to run simulations of your presence. This is additional protection against burglary when the user is away from home This mode is designed to simulate the presence of a household member based on the information gathered from the sensors when the automatic lighting mode is activated. Simulation of presence consists in triggering the light sequences from the time stored in the database.
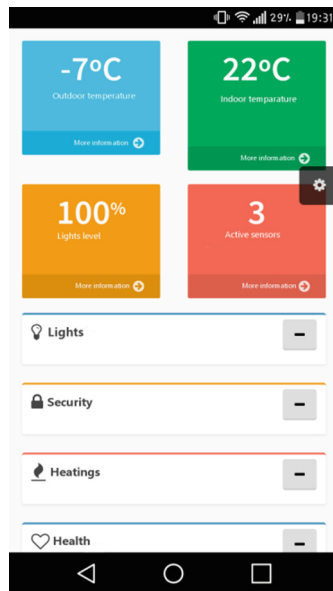
Fig. 5. The user interface is also displayed via mobile browsers. concept of the mechatronic device for rehabilitation.

The last section deals with health. It is responsible for providing information about the patient's state of health, through which it is possible to access information collected from individual sensors and to present in the context of diagnosis and improvement of health. Generated graphs refer to the patient's movement activity over the last 24 h or the selected time period. After a suitable time in the current graph (e.g. the last day), the chart will also show the average activity. This will allow users to observe specific behavioral changes. In this section it is also worth noting about the presented data from pressure sensors placed in the bed. Based on their analysis, the quality of sleep information can be extracted and, in case of problems with it, the diagnosis can be made, for example if a restless leg syndrome is suspected. Pressure sensors can also be a source of information that is especially important for patients who do not leave the bed. They suffer from severe problems, one of which is bedsores and painful wounds resulting from them. Pressure sensor data can help identify the places where the above-mentioned pressure is highest and try to lower it in individual parts of the body, for example by using appropriate pads or using modern technology in the form of variable pressure mattresses. The next element of this section is the presentation of results from a bar code reader that will allow the user to control the diet and regular intake of meals and medications. This section contains important information about the doctor's recommendations and medications. Data from medical sensors is also an important complement to the rest of the information. Parameters collected by medical sensors, such as pulse, blood pressure and weight, directly affect the patient's health and well-being. Particularly, systematic control over them avoids many civilization diseases such as hypertension or, in severe cases, keeping the disease within the limits.

6 Device Prototype

The first prototype of the device was built and installed in one of the apartments of the elderly. Commonly available and cheap components are available in most electronics stores. In spite of adopting the prototype of cheap modules, these components (motion sensors, pressure sensors and other previously mentioned components) fully meet the requirements and are able to extract the information needed to carry out health diagnostics and to improve the comfort of using the house appliances by an older person.

Important components used in the project:

- The Raspberry Pi model 2B, which is based on the SoC Broadcom BCM2836 chip and has 1 GB of ram. The Raspberry Pi model 2B is equipped with 40 GPIO pins, which can be expanded with the help of dedicated MCP23017, if needed. In addition, the computer has such interfaces as HDMI, 4x USB, Ethernet, microSD card, audio output
- DHT11 temperature sensors with integrated humidity sensors and DS18B20 temperature sensors. The DHT11 sensor has a temperature range of 0° to 50° C and a humidity of 20 to 90% RH. The system operates at a voltage of 3 V to 5.5 V. Its accuracy is ± 1° C and ± 4% RH for the humidity, respectively. The operating range of the DS18B20 is between -55° and 125° C. According to the manufacturer,

the supplied power supply should be in the range of 3 V to 5.5 V. Its accuracy is ± 0.5° C

- PIP-HRC 505 motion sensors are rated at 4.5 V to 20 V and range up to 7 m. The sensor's viewing angle is 100°
- heating and cooling of rooms will be carried out by means of air blowing at the set temperature
- lights module includes light source and motion detector.
- to recognize patient and other person presence there was used RFID reader, working with frequency equal 13,56 MHz. The device desires 3,3 V current voltage.
- FSR-402 are force sensors, which work in a range from 0.2 N to 20 N. To connect this sensor and microcomputer, converter A/C is necessary.
- case is designed in cad software and printed on 3D printers used Rapid Prototyping Methods
- modified power supply unit.

The cost of building a prototype and adaptation of the apartment was about 300 euros. The cost estimate does not include time spent on construction and software development. It is worth mentioning that the cost of energy and materials needed to make the device is low.

The great advantage of the device is the versatility and no interference with the existing infrastructure of the apartment. There are many possibilities to customize the device for different target groups. It will work well for older people requiring 24/7 care and younger people with disabilities.

The Fig. 6 presents the model of an exemplary room that has been installed in which selected components of the comfort and safety system of the elderly and the disabled are installed.

Fig. 6. Base module.

7   Results

The article also presents the first tests of the device. Verification was based on server and client checking. The collected data were from the 24-hour life of one of the monitored patients.

The beginning of the tests was to start the device and all the required services in the system. The device worked for one day and collected information from the attached sensors. At the same time, it performed his task by managing lighting, heating and collecting information about pressure. After the verification time has passed, the user interface has been validated.

Results of tests are presented for the each section.

7.1   Lights Part

The Fig. 7 presents details of the lights system (Fig. 8).
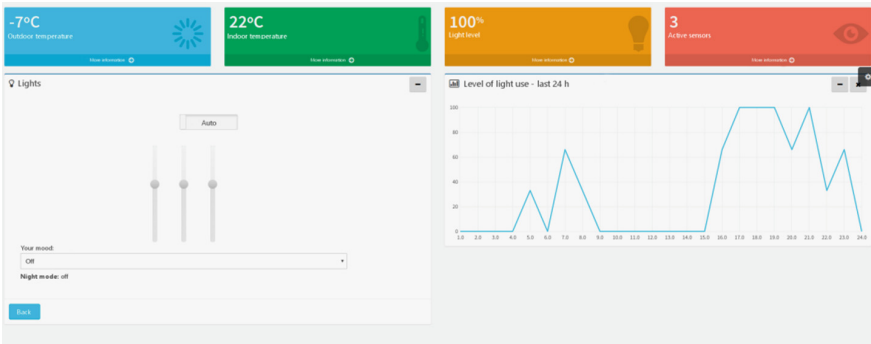
Fig. 7. Lights statistics.
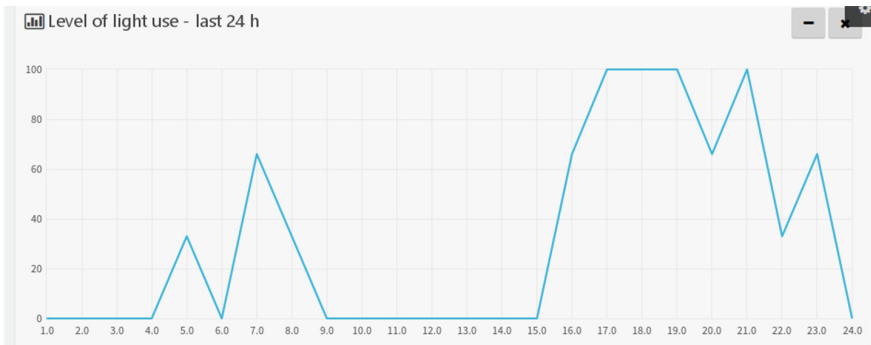
Fig. 8. Chart presented percent of count working light source for 3 connected LED's

7.2   Heating and Air-Condition Section

In the detail page the user can see information about temperature, humidity and other conditions related with the air state inside home (Fig. 9).
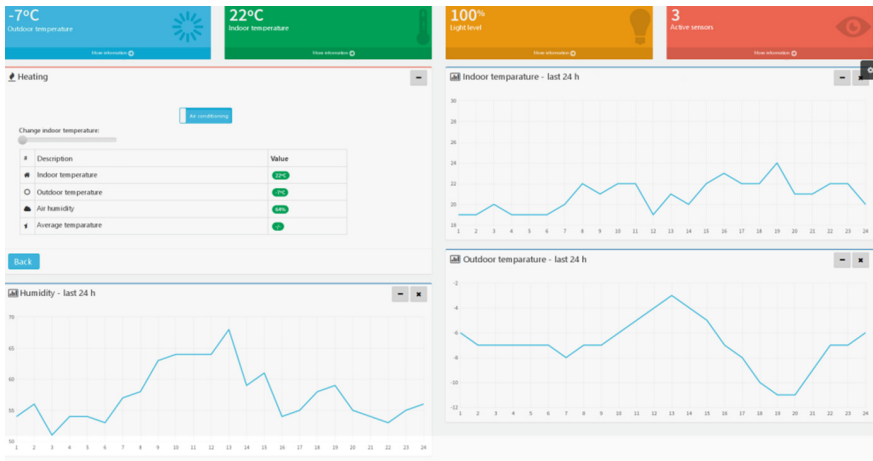
Fig. 9. Heating statistics.

The Fig. 10 includes charts based on data picked up in the day when the test was performed. The second chart contains temperature from a sensor placed outside home. The last chart presents changes of humidity value through the 24 h test. Based on this data, the current season is winter. From about 7 am to 11 am and from around 3 pm to 11 pm, the patient stayed in the flat. The patient prefers to have a higher temperature in the flat, around 22–23 ºC.

7.3   Safety Section

The Fig. 11 shows subpage with statistics based on information from motion sensors. On the charts user can see count of detected event. Tests supported only 3 sensors.

The test required to place sensors in a room, when any activity was detected, the system saved this information in the database. Charts shows count of detected activities in each day hours (Fig. 12).

7.4   Health Section

This section is the most important of the whole system. Here are presented information about patient health and predicted disease. Charts based on data from all sensors are presented in the context of health care (Fig. 13).
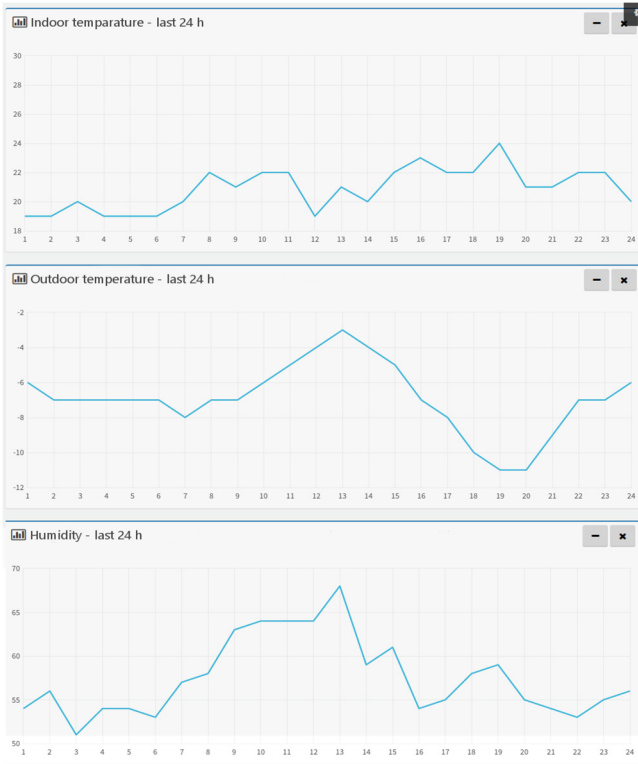
Fig. 10. Charts presented data picks up from temperature and humidity sensors.
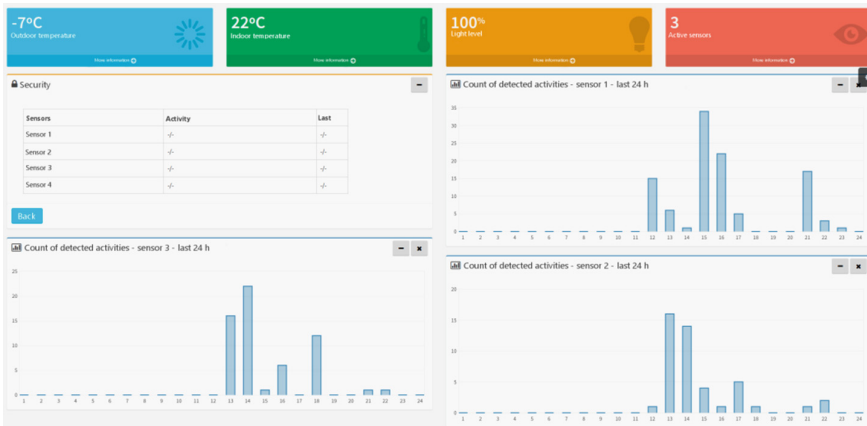
Fig. 11. Safety statistics.

At the above Fig. 13 the first chart concerns the proportion of the detected motion, pressure on bed and place where are force sensors. At the request of the medical staff it is possible to display data from the whole week. It is on the basis of this data that the potential threat can be determined as quickly and as accurately as possible. In addition, it should be clarified that all the presented charts are examples developed in any chosen period.
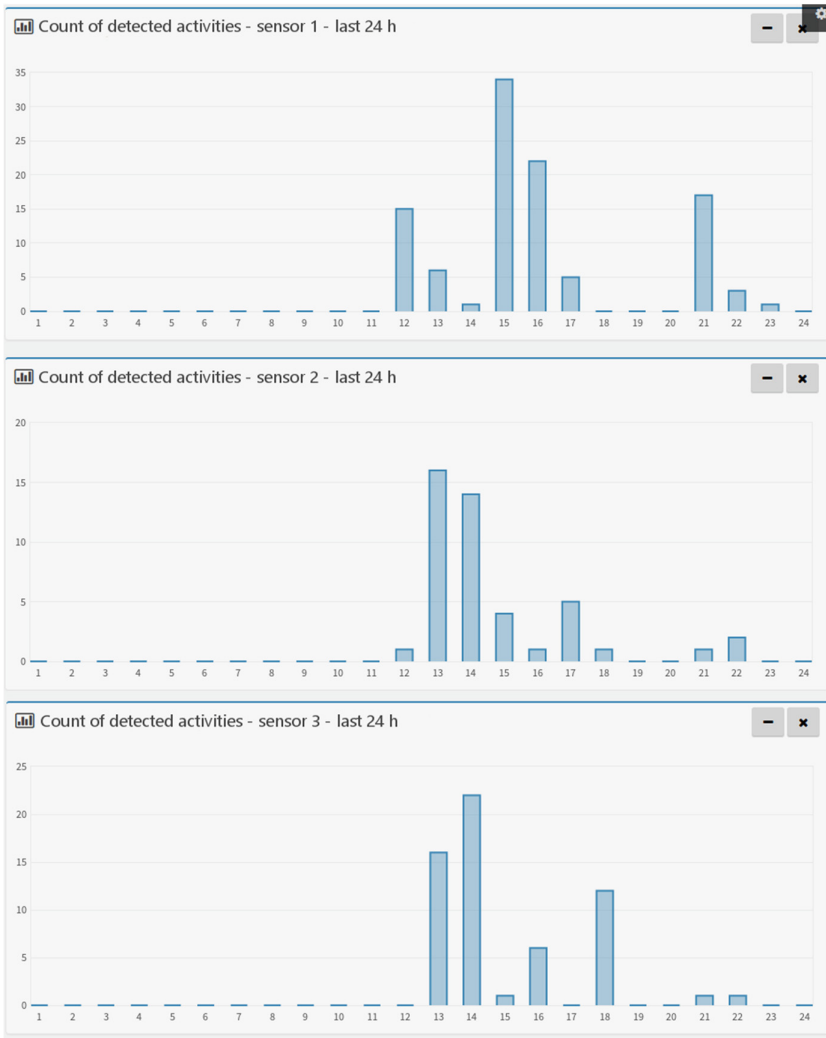
Fig. 12. Charts presented statistics related with motion sensors.

Fig. 13. Health statistics

8   Conclusion

By analyzing the current situation and needs of the elderly or disabled, and taking into account the existing home automation systems, the systems monitoring the vital parameters of individuals, the idea was developed to develop an innovative comfort, security and health care system for a social group living alone.

Compared to currently available solutions, the developed system not only enhances the comfort and safety of the user from the point of view of operating the equipment and staying in a given dwelling equipped with the developed system, but in particular is responsible for monitoring, diagnosing and reporting the emergence of hazardous situations involving the user (e.g. falls) or inform the family or health care provider about the appearance of a future illness (e.g. depression). The system analyzes the behavior patterns of the elderly or disabled and is based on data from many types of sensors located in the affected unit. On the basis of this, it is also possible to better control the treatment and convalescence of the user staying in his or her home.

The components include: control module with computer and display and sensor set, including presence sensors, cameras, door closers, food sensors, radio sensors, temperature sensors and humidity sensors.

Information obtained in this way originating, i.a. from motion sensors, pressure sensors and medical sensors allow to generate reports that are relevant for the patient, the family, or for physicians, rehabilitators, dieticians. The data on physical activity, sleep quality, meals and medication times, provide increased discipline by the patient himself to treat his illness. Very often, the lack of systematic and uncontrolled patient treatment may prejudice its success or failure. Providing selected data from reports as well as the ability to analyze real-time user activity is very important also for the user's family. An important part of the system is to improve communication with the medical personnel. The information included in the reports may include physical activity in the home, regular medication and meals (particularly important for people with diabetes).

These reports are based on information gathered from room sensors, door openers and other food lockers, and scanned food products using barcodes. The signs of the disease can also be seen in data from the pressure sensors in the bed and in the seats.

It is also worth noting that the described device is a modular system. It is possible to individually configure individual modules of the system to meet the needs of the user, with the possibility of retrofitting in the future with additional components.

The developed system features a user-friendly interface that is available through a touch screen, web page or mobile phone.

For patients and families, privacy is very important. All data collected is stored in a password protected database. If necessary, it is possible to encrypt data which, even in the case of unauthorized access, will preserve privacy.

Another feature of the solution is the ability to monitor and protect property. Ideal for this task are the presence sensors. With their help, it is possible to detect third-party traffic, launch cameras, and send relevant information to a pre-defined audience, which may be at least the police.

The closure sensors provide information that all windows and doors are closed when the user exits the room and records it using the radio sensor located at the exit door.

With dedicated software and LEDs mounted on the presence detectors, the device can automatically turn on and off the light by simulating the presence of in-room occupants, thereby reducing the risk of intrusion in the absence of the user.

A great asset of the presented solution is certainly the low cost of implementing the system and no need to interfere with the existing infrastructure of the building. Enclosures in which each module of the system can be modified depending on the decor of the room and the user's taste. This kind of capability was obtained even on the basis of the 3D printing used for the prototype. Its installation in the building does not require any interference with the existing building infrastructure.

The modular design allows almost infinite adding elements and collecting more data, depending on the requirements set by the physician and the patient's needs. The modularity and compatibility of the system, as well as the fact that there is a possibility of cooperation of the described system with the solutions monitoring the basic parameters of life, is also worth noting.

In addition to the previously mentioned benefits for the user and his family, the software takes into account the ecological aspect. Energy savings are achieved by adjusting the lighting to the user. Reducing brightness and starting at the right times gives measurable benefits. Also, the use of LED technology has a positive effect on reducing energy consumption.

Presented system is a prototype. The device has been tested and results are shown in this article. Also the whole project has been send to the Polish Patent Office for the protection of copyright.

References

1. Vieira, A., Gabriel, J., Melo, C., et al.: Kinect system in home-based cardiovascular rehabilitation. J. Eng. Med. **231**(1), 40–47 (2016)
2. Stroke Association, Research Spend in the UK: Comparing stroke, cancer, coronary heart disease and dementia 2014. http://www.stroke.org.uk/research-spend-uk. Accessed 9 Jan 2017
3. Raisul Alam, M., Bin Ibne Reaz, M., Alauddin Mohd Ali, M.: A Review of smart homes – past, present, and future. IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.) **42**(6), 1190–1203 (2012)
4. Lobaccaro, G., Carlucci, S., Löfström, E.: A review of systems and technologies for smart homes and smart grids. Energies **9**(348), 1–33 (2016)
5. Linder, F.E., et al.: An index of health: mathematical models. Washington, D.C.: U.S. Dept. of Health, Education, and Welfare. Public Health Service, Series 2, No. 5, May 1965
6. Tutak, J.S.: Design of ELISE robot for the paretic upper limb of stroke survivors. J. Vibroengineering **18**(6), 4069–4085 (2016)
7. Giergiel, J., Kurc, K.: Identification of the mathematical model of an inspection mobile robot with fuzzy logic system and neural networks. J. Theor. Appl. Mech. **49**(1), 209–225 (2011)
8. Hendzel, Z., Burghardt, A., Szuster, M.: Reinforcement learning in discrete neural control of the underactuated system. In: Rutkowski, L., Korytkowski, M., Scherer, R., Tadeusiewicz, R., Zadeh, L.A., Zurada, J.M. (eds.) ICAISC 2013. LNCS (LNAI), vol. 7894, pp. 64–75. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-38658-9_6
9. Dong-Yuan, G., Xi-Fan, Y., Qing-He, Y., et al.: Robot sensor calibration via neural network and particle swarm optimization enhanced with crossover and mutation. TV-TG **1**(5), 1025–1033 (2014)
10. Tutak, J.S.: Virtual reality and exercises for paretic upper limb of stroke survivors. TV-TG **24**(2), 451–458 (2017)
11. Hendzel, Z., Burghardt, A., Gierlak, P., et al.: Conventional and fuzzy force control in robotised machining. Solid State Phenomena. Trans Tech Publications, Vol. 210, pp. 178–185 (2014)
12. Fitzpatrick, P., Needham, A., Natale, L., Metta, G.: Shared challenges in object perception for robots and infants. Infant. Child Dev. **17**, 7–24 (2008)
13. Kudasik, T., Libura, M., Markowska, O., Miechowicz, S.: Methods of reconstructing complex multi-structural anatomical objects with RP techniques. Bull. Polish Acad. Sci. Techn. Sci. **64**(2), 315–323 (2016)
14. Kurc, K., Szybicki, D., Burghardt, A., et al.: The application of virtual prototyping methods to determine the dynamic parameters of mobile robot. Open Eng. **6**(1), 55–63 (2016)
15. Strumillo, P., Electronic interfaces aiding the visually impaired in environmental access, mobility and navigation. In: 3rd International Conference on Human System Interaction HSI 2010 − Conference Proceedings, pp. 17–24 (2010)
16. Loomis, J.M., Klatzky, R.L., Giudice, N.A. (2013) Representing 3D space in working memory: spatial images from vision, hearing, touch, and language. In: Lacey, S., Lawson, R. (eds) Multisensory Imagery. Springer, New York, NY (2013)
17. Hannaford, B., Ryu, J.H.: Time-domain passivity control of haptic interfaces. IEEE Tran. Robot. Autom. **18**(1), 1–10 (2002)
18. Tutak, J.S., Wiech, J.: Horizontal automated storage and retrieval system. Adv. Sci. Technol. Res. J. **11**(1), 82–95 (2017)
19. Adams, R.J., Hannaford, B.: Stable haptic interaction with virtual environments. IEEE Trans. Robot. Autom. **15**(3), 465–474 (1999)

20. Budzik, G., Turek, P., Traciak, J.: The influence of change in slice thickness on the accuracy of reconstruction of cranium geometry. J Eng. Med. **231**(3), 197–202 (2017)
21. Bu, N., Okamoto, M., Tsuji, T.: A hybrid motion classification approach for EMG-based human-robot interfaces using bayesian and neural networks. IEEE Trans. Robot. **25**(3), 502–511 (2009)
22. Kowalczyk, A., Hanus, R., Szlachta, A.: Investigation of the statistical method of time delay estimation based on conditional averaging of delayed signal. Metrol. Meas. Syst. **18**(2), 335–342 (2011)
23. Shin, S.Y., Kim, J.Y., Lee, S., Lee, J., Kim, S.-J., Kim, C.: Intentional movement performance ability (IMPA): a method for robot-aided quantitative assessment of motor function. In: Proceedings of IEEE International Conference on Rehabilitation Robotics (ICORR), pp. 1–6 (2013)
24. Kowalczyk, A., Szlachta, A., Hanus, R., Chorzępa, R.: Estimation of conditional expected value for expone ntially autocorrelated data. Metrol. Meas. Syst. **24**(1), 69–78 (2017)
25. Llewellyn, F.: Some fundamental properties of transmission systems. Proc. IRE **40**(3), 271–283 (1952)

Analysis of the Accuracy of Crime Scene Mapping Using 3D Laser Scanners

Tadeusz Wieczorek[✉], Roman Przyłucki, Joanna Lisok,
and Adrian Smagór

Department of Industrial Informatics, Silesian University of Technology,
Krasińskiego 8, 40-019 Katowice, Poland
tadeusz.wieczorek@polsl.pl

Abstract. The use of a 3D scanner at the crime scene i.e. the use of its measurements and the final spatial documentation of the crime scene, must meet the conditions of reliability and accuracy. Acquisition of spatial data can be important from the point of view of investigating and verifying hypotheses. Verification can be subjected to either the mutual location of the elements within the scene or their mutual distance. To test the accuracy of measurements on 3D scans a series of measurements was made at the actual scene, and then the same measurements were made in 3D scans. In the research, we considered that the accuracy of a crime scene mapping, depends on the surface characteristics. The measuring noise also increases with the distance from the scanner. The scanning density, as well can significantly limit the actual accuracy of the point measurement. Measuring distances based on indicating a point by the user is also characterized by an error resulting from the operator's competences, as well as the ambiguity of the position the characteristic point. It cannot be clearly stated what the real accuracy of mapping, with a 3D scanner, of a crime scene is.

Keywords: 3D scanner · Crime scene investigation · Scanner accuracy

1 Introduction

The aim of the research doing by authors under the R&D Project of the Polish National Research and Development Centre (DOB-BIO6/18/102/2014) was to analyze 3D laser measuring technologies that can be used during the crime scene investigation [1]. Currently, photography and video registration are the dominant techniques in this area. Despite the significant development of photographic techniques (including photogrammetry and digital photography), the limitation to acquisition and analysis of flat images remains important. Modern techniques analyzed in the project include:

- acquisition of spatial data by means of 3D scanning,
- linking the spatial model with photos,
- satellite measurements (GPS and analogous satellite location systems) for geographical localization of a crime scene or a road incident.

Visual documentation of such events is very important and 3D documentation should replace photographs, sketches and video records. Limitations resulting from

overrides indicate that it is advisable to combine the use of a terrestrial scanner with scanners with the greatest possible mobility and ease of use (portable scanners).

In the case of new technologies, the issue of acceptance of research methods in the investigative process and subsequent court trials. In the literature there is a standard of acceptance of scientific results in legal proceedings (Daubert's standard) - the name of which derives from a specific case, and is related to the precedent nature of law in the USA. The reliability and accuracy of the new test method used in the criminal process should be assessed by the judge in the following aspects [2]:

– the method is the subject of peer-reviewed scientific publications,
– its error range is known,
– have adopted standards describing its operation,
– meet with wide acceptance in the scientific community.

The use of a 3D scanner at the crime scene i.e. the use of its measurements and the final spatial documentation of the crime scene, must meet the conditions of reliability and accuracy. The basic purpose of the crime scene inspection is objective reconstruction of the event and securing material evidence for the needs of criminal investigations. The spatial scanner enables obtaining information about the surface in the 3-dimensional space. Only visible surfaces will be scanned. It requires to perform additional scans from other positions. Building a model of the scene based on many scans requires their subsequent submission (registration), which needs common areas to be scanned. The requirement of common areas demands appropriate planning of the scanning process. Registration also affects the accuracy of scene reconstruction.

2   The Forensics' Use of 3D Scanners in the World

Spatial scanners are an example of technology for documentation, which is gaining wider and wider recognition. Applications of scanners include technique, medicine, earth sciences, cultural heritage or forensics [3, 4]. However, the process of reaching their practical use in forensics was long and turbulent. Some doubts were raised, among others, by the accuracy of measurements obtained with the use of 3D scanners' computer tools, which was highlighted in a 350-page report prepared by the National Academy of Science (USA) in 2009 [5]. On this basis, the NIST Law Enforcement Standards Office and NIST Physical Measurement Laboratory in cooperation with Leica Geosystems have developed a calibration standard for scanner devices that ensures the reliability of measurements in forensic applications and more.

The research shows that 3D scanners in Europe have, among others, Police services of the following countries [1]: Switzerland (Kantonspolizei w St. Galler, Police Neu-châteloise – Police de la circulation), Great Britain (City of London Police, Metropolitan Police Traffic Unit), Germany (Hessisches Landeskriminalamt, Wiesbaden; LKA Bayern, Monachium; LKA Sachsen, Dresden), the Netherlands (Politie Rotterdam-Rijnmond, Krimpenaan den IJssel), and France (State Military Police - Gendarmerie). There are also countries in which Police services use some external entities with 3D scanners, such a country is e.g. Denmark. In addition, scanners are commonly used in forensic science in the USA.

Due to the fact that Germany is a federal country, local forensic laboratories have a large independence in the use of measurement techniques, including during the inspection of the crime scene. The Federal Criminal Police of Germany (Bundeskriminalamt, BKA) is involved only in international organized crime or at the request of the competent authorities of a Federal State or Federal Minister of Internal Affairs. In BKA, the most commonly used measuring devices used during visual inspections are 3D scanners and spherical cameras.

Police in the Netherlands are in possession of 3D scanners and use them during the inspection of the scene of the incident. The offense must, however, have to do with the deprivation of life (murder, serious traffic accident, plane crash, etc.). It is therefore about serious events. This equipment is used on large and hard to reach spaces or in areas particularly exposed to evidence obliteration. One cannot use the 3D scanner in the case of a so-called common crimes (e.g. burglary). The use of this technology is planned to be extended to include visual inspection of illegal synthetic drug laboratories.

In Switzerland, in recent years, new methods of 3D scanning in forensics have developed and gained significance. 3D laser scanning was considered an efficient research method. Scanners are used in areas related to forensic medicine, namely virtual autopsies of corpses, with the fact that typical 3D modelling of the body's interior is responsible a CT scanner, and 3D laser scanning refers only to the body surface and external injuries.

The greatest usage of 3D scanners may be observed by the American Police. Multidisciplinary Accident Investigation Team, working within the California Highway Patrol, is concerned mainly with crime detection of traffic accidents, their reconstruction and causes. Moreover, MAIT helps in other crime scenes, especially those in which firearms are used both by perpetrators or policemen. In the State of New Mexico (US Albuquerque Police), a 3D scanner is used by the Mobile Crime Lab at the sites of most serious crimes, usually accompanied by fire-arms usage. 3D scanning technology has recently started to be employed also in Europe by police detectives mostly in investigations carried out after explosions originated by organized criminals or terrorists [4].

3 Results of Research on the Accuracy of the Crime Scene Measuring

Acquisition of spatial data can be important from the point of view of investigating and verifying hypotheses. Verification can be subjected to either the mutual location of the elements within the scene or their mutual distance without having to leave the forensic laboratory, which can significantly speed up the investigation time. The other advantage of full crime scene registration is the possibility of critically analyzing the forensic hypotheses at the stage of their preparation without having to carry out numerous local visions. An example of this is the analysis of shots taken from a moving vehicle and its visualization. Similar 3D scanning techniques have also found application in the analysis of traffic accidents, facilitating the reproduction and simulation of the moment of occurrence of an event [6]. Another area of application for portable scanners is the registration of footprints, tire traces, both those collected from permanent and nondurable places such as the imprint in the snow. A separate area of application may also be issues of identification and adjustment of small elements such as dentition elements,

skull bone fit, or reconstruction of the facial appearance, but they require devices with very high accuracy, which was analyzed in [7].

Evaluating accuracy of 3D scanners some problems should be discussed. The scanning technique based on reflection beam analysis depends on such parameters of the scanned surface as its angle of inclination - relative to the laser beam (it is optimal to scan perpendicular surfaces with an angle limit value acceptable by the software) and reflectance in the infrared light range. Many manufacturers also point to an increase in measuring noise depending on the surface characteristics. Optical parameters of the object being scanned are very important. In order for the scanners to be able to correctly acquire data, the light beam displayed by them must be properly visible on the scanned surface. Therefore, surfaces that are highly reflective, transparent, absorbing light or causing its diffusion are difficult materials for scanning. During the tests, we checked how the examined scanners behave when trying to scan the surfaces generally considered as hard to scan materials. For example, materials that transmit a light beam are virtually invisible for scanners. Without applying, for example, matting powders, the scanner will not be able to capture their geometry. However, if used on the crime scene, usually such action will not be allowed. Resistance to weather conditions (low temperatures, humidity, and lighting) is also important in forensic applications. Analyzing the possibilities of accurate mapping of the crime scene, all these factors should be taken into account, as they affect the precision of measurements performed on scans.

To test the accuracy of measurements on 3D scans a series of measurements was made at the actual scene, and then the same measurements were made in 3D scans. In total, hundreds of measurements were made, in different configurations of both, crime or traffic incident scenes locations, and evidence of the crime discovered.

3.1 Possibilities of Measurements of Various 3D Scanners at the Crime Scene

The infrared camera connected to the scanner allows using the temperature reading as a texture superimposed on the spatial scan (Fig. 1).

Fig. 1. A combination of two armchairs - an employee sat on the right armchair before scanning. It is in parts adjacent to the body about 1 °C warmer.

Clouds of points that we obtain are assigned values, i.e. each point outside of its XYZ coordinates has an RGB value and additionally, information about the intensity of laser reflection from a given surface is also recorded. The intensity of the reflection allows to notice various types of cracks/changes not visible to the naked eye on a structure that would seem to be quite uniform. It is also easier to perceive bloody traces in places with insufficient light, where photo in the visible range may have difficulty with proper registration. Documentation of the crime scene performed with a laser scanner does not require lighting and can be done in absolute darkness. The view of the scan made in this way can be visualized as shades of gray as well as the intensity of reflection.

Attempts have also been made to use a 3D structural light scanner to scan shoe prints (Fig. 2). A Go!Scan 50 scanner (having accuracy 0,1 mm) was used for measurements. Faro Freestyle, DotProduct, Artec or Mantis Vision scanners are also often used in forensic applications. A few scans were performed by moving the scanner around the object being scanned. Partial scans were imported into the VXelements software, where they were combined into one unit. The combined cloud of points has been transformed into a grid of triangles, becoming a model of the test imprint. In order to confirm the high accuracy of the scan, a computer model of the shoe was made, and then both models were compared. The obtained accuracy for almost the entire shoe print is high.

Fig. 2. Visualization of scanned shoe prints.

3.2   Testing the Accuracy of Measurements on Scans

To test the accuracy of measurements on 3D scans a series of measurements was made on real and simulated crime scenes, and then the same measurements were made on 3D scans. The scenes were scanned using Leica both P40 and Z + F Imager 5010X. An example set of 10 measurements is presented in the paper. Real distances have been measured using laser distance meter with estimated error ±2 mm. The description of

the measurement points is summarized in Table 1. The measurements were repeated 5 times (average values in the table).

Table 1. Results of distance measurements.

Distance between	Real distance [mm]	Measured on scan [mm]
(a) Drawer handle - the leg of the chair	2270	2286
(b) Two cupboards corners distance	630	643
(d) Corners of the washing machine and the cabinet	195	198
(e) Corners of two upper cupboards	2875	2860
(f) Floor – the top of the cupboard	1999	1961
(i) Dimensions of the table	500 × 1500	498 × 1488
(j) Heel of the shoe - corner of the table	960	911
(o) Thickness of the table top	18	20
(p) The width of the door	940	949
(s) The width of the lamp	1265	1249

The accuracy of measurements performed on scans, obtained in practice, are different and depend on many factors. Some of them are discussed below.
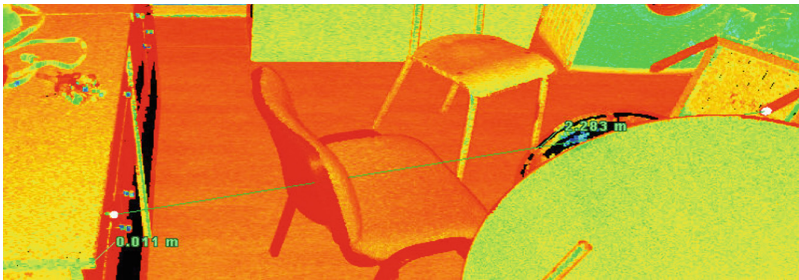
Fig. 3. For this case (a), the problem was to determine the points on the handle because it had a shiny surface.
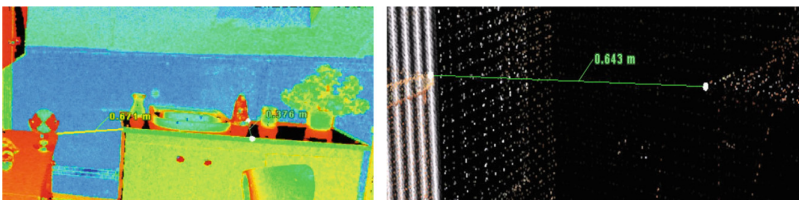
Fig. 4. In this case (b), the reason for the difficulty was the overlap of the edges on the scan, it was necessary to choose the appropriate perspective of measurement/observation.
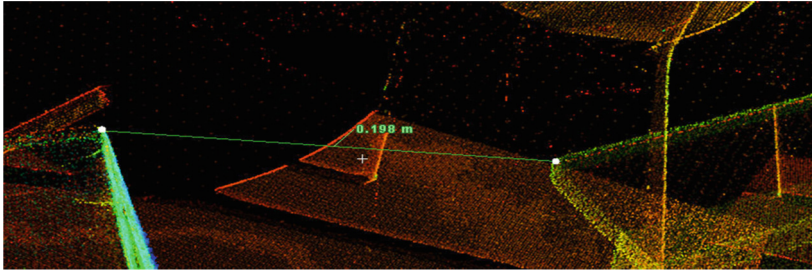
Fig. 5. For this measurement (d), the problem was to accurately determine the edge of the washing machine (mirror surface). It was necessary to change the perspective and use the representation in the intensity of reflection.
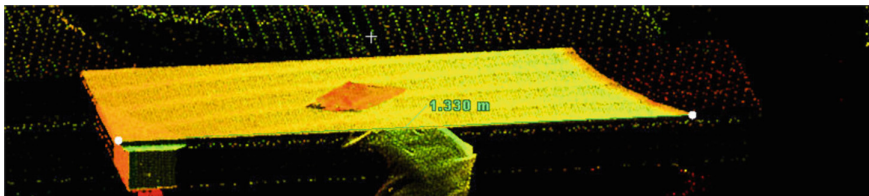
Fig. 6. The difficulty (i) consisted in overlapping of some edges and in the table area there were areas with different density of points. The effect of this is visible in the form of an untrue change in the table shape.
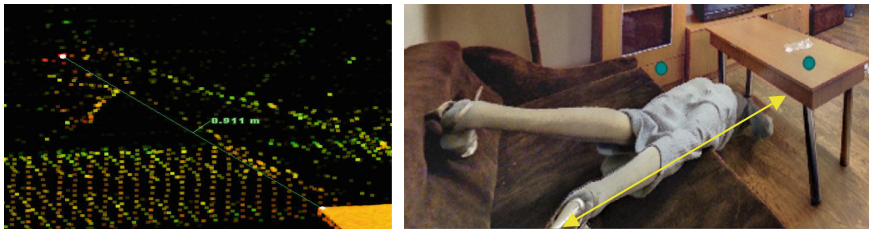
Fig. 7. For this measurement (j), the difficulty was that the heel, due to its small size, had few measuring points.

The lamp (s) was very difficult for measuring because it had all the features of hard-to-scan objects. The lamp housing was made of transparent plastic. They had rounded edges with two lines of refraction, they were scanned in a disadvantageous perspective and there were very few measuring points. Therefore the measuring errors are big.

Presented in the Table 1 and Figs. 3, 4, 5, 6 and 7 results are just exemplary. We have done wide research on accuracy of crime scenes mapping for scenes placed either in buildings or open areas (traffic incidents). Generally, it was found during the tests:

- Objects that are difficult to measure on the scan are, of course, those that are difficult to scan, i.e. objects that strongly reflect light (Fig. 8).

Fig. 8. An example of a blurring cloud of points on a shiny glossy kitchen countertop.

- Carefully choose the perspective to mark the measurement point because often the edges, determined by the points on the scan, overlap each other and further objects (e.g. the back edge of the cabinet) merge closer (the front edge of the cabinet).
- For small objects (e.g. thickness of the countertop) problems occur due to the scanner's resolution (measuring points are distant from each other by approx. 3 mm). Scanner's resolution can be enlarged but resulted files size are huge.Such large data files are difficult to process and store.
- When assessing the accuracy of measurements on the scan, the accuracy of measurements taken in the actual scene should also be taken into account.
- Real accuracy of crime scenes mapping is much lower than the precision and resolution declared by scanners' producers.

4   Summary

When considering the issue of accuracy one should distinguish two parameters presented by manufacturers - accuracy and precision. In the case of accuracy, an error is considered relative to the actual distance from the tested object. For precision, the repeatability of the test is considered. In the research, we considered accuracy of a crime scene mapping, in both meanings. The 3D scanning error depends on the surface characteristics. The ideal surface should reflect the laser beam well. With the reduction of the reflection coefficient, the measuring noise increases (other parameters may also change, such as the maximum scanning range). In some special cases, scanning of some objects may not be possible at all. The measuring noise also increases with the distance from the scanner. The Zoller + Fröhlich scanner documentation [8] also defines a linear error - this is the maximum error resulting from a non-linear distance mapping and should be considered as a component of the distance error.

The scanning density (resolution) determines how densely located scan points can be [9]. While the accuracy of the distance measurement using the scanner is related to

the position error of the measuring point (relative to any chosen reference point in the scene), the scanning density can significantly limit the actual accuracy of the point measurement (hence limiting the precision of measuring the distance between points). For example [3], for the FARO X 130 scanner, whose declared distance error is approx. 2 mm, for a distance of 40 meters, a single distance measurement falls on a surface of 6×6 mm (for the highest possible scanning density). In fact, a point can be located in a space section with dimensions of approx. $6 \times 6 \times 2$ mm. For smaller distances from the scanner, the diameter of the laser beam may be more important than the scanning density. For example, for a FARO X 130 scanner with a beam diameter of 2.25 mm and the densest sampling, the beam diameter is less than the spacing between points only for scanning surfaces at a distance greater than 7.2 m. When analyzing the granularity of the measurement (the minimum size of the voxel, including the measured point), the fact of the discovergency of the laser beam should also be taken into account. For example, in a FARO X 130 scanner it is typically 0.19 miliradians, which adds about 19 mm diameter of the beam at 100 m.

Another problem is the precision of the point indication itself. Manufacturers of scanners provide software for making measurements, and also provide sets of dedicated tools. Measuring distances based on indicating a point by the user is also characterized by an error resulting from the operator's competences, as well as the ambiguity of the position (and the possibility of indicating) the characteristic point. During the tests, efforts were made to select the measuring points in an unambiguous manner, but the unambiguous indication of the point was often only possible within a single scene.

Accuracy evaluation was also made for scans of open areas (traffic incidents scenes), but a problem of a smaller number of characteristic points that can be used for measurements, occurred stronger. The distance measurement errors obtained during the open area scenes research were significantly larger (up to 0,5 m at 100 m distance from the scanner). It can be assumed, however, that in this case the main source of errors was the difficulty in clearly indicating the end points, due to the low density of scanned points, at large distances from the scanner.

The assessment of the actual accuracy of the mapping of the crime scene is important, because of using 3D scans, as evidence, in the judicial process. The wide studies carried out by authors as part of a research grant, showed that the real accuracy of distance measurements on the scan is much lower than the accuracy declared by the scanners' manufacturers. The accuracy depends mainly on the type of material being scanned, proper determining the measuring points, choosing the appropriate perspective of measurement/observation, detection and removal of the overlap of the edges of various objects, the use of appropriate procedures for scanning transparent objects, areas in the scan with different density of points, small size objects measuring due to few scanning points, proper identification of an object edges due to scanning points blurring and from the investigator competences and experience. Therefore, it cannot be clearly stated what the real accuracy of mapping, with a 3D scanner, of a crime scene is.

References

1. Wieczorek, T., Zubańska, M., Wiciak, K., Szymczak, M.: Techniczne i prawne aspekty oględzin miejsca zdarzenia z wykorzystaniem skaningu 3D. In: J. Kosiński (eds.) Przestępczość teleinformatyczna, pp. 147–158, W-wo Wyższej Szkoły Policji w Szczytnie (2015)
2. Daubert Standard. https://www.law.cornell.edu/wex/daubert_standard. 26 Mai 2015
3. Kowalski, P., Skabek, K.: The accuracy of a scene geometry based on the terrestrial laser scanner measurements. In: Conference Proceedings, Man-Machine Interactions 5, vol. 1, pp. 408–417 (2017)
4. Wieczorek, T., Górawska, A.: Portable 3D scanners for crime scene investigations. Studia Informatica **8**(3), 132–141 (2017)
5. Committee on Identifying the Needs of the Forensic Sciences Community, National Research Council, Strengthening Forensic Science in the US: A Path Forward (2009)
6. Buck, U., Naether, S., Räss, B., Jackowski, C., Thali, M.J.: Accident or homicide—virtual crime scene reconstruction using 3D methods. Forensic Sci. Int. **225**(1), 75–84 (2012). (Online)
7. Sholts, S.B., Flores, L., Walker, P.L., Warmlander, S.K.T.S.: Comparison of Coordinate measurement precision of different landmark types on human crania using a 3D laser scanner and a 3D digitiser: implications for applications of digital morphometrics. Int. J. Osteoarchaeol. **21**, 535–543 (2011)
8. Zoller + Fröhlich: Understanding Imager 5010 accuracy specifications (2014)
9. Tahar, K.N.: Investigation on different scanning resolutions for slope mapping studies in Cameron highlands. Arab. J. Sci Eng. **40**, 245–255 (2015)

Application of Artificial Neural Networks in Identification of Geological Formations on the Basis of Well Logging Data – A Comparison of Computational Environments' Efficiency

Marcin Zych[1(✉)] , Gabriel Stachura[1], Robert Hanus[2] , and Norbert P. Szabó[3]

[1] AGH – University of Science and Technology, 30 Mickiewicz Av, 30-059 Kraków, Poland
`zych@geol.agh.edu.pl`, `gabrielstachura@gmail.com`
[2] Rzeszów University of Technology, 12 Powstańców Warszawy Av., 35-959 Rzeszów, Poland
`rohan@prz.edu.pl`
[3] University of Miskolc, Miskolc-Egyetemváros 3515, Hungary
`gfnmail@uni-miskolc.hu`

Abstract. The paper presents the application of artificial neural networks in lithology identification on the basis of well logging data. The problem is very important considering petroleum geophysics as it allows to find sweet spots - potential deposits of hydrocarbons (oil and gas). The use of advanced statistical methods such as artificial neural networks is expected to improve geological interpretation of geophysical data. Moreover, such methods are capable of dealing with big data sets since well logging provides more and more information about petrophysical (e.g. porosity, density, resistivity, natural gamma radiation, sonic wave propagation) and chemical rock properties (mineral content and element abundance). Therefore, the analyzed data comprises around 56000 records. Two different computational environments has been used in order to examine their efficiency in terms of accuracy of a lithological classification. Computation was done in R software, which is an open source environment, and STATISTICA v. 13 which is a commercial one. As an input, logging data from three boreholes drilled in the Baltic Basin, North Poland were used. The results show that R offers more possibilities of modification of a net. However, STATISTICA provides more user-friendly interface and better accuracy of lithology identification.

Keywords: Artificial neural network · Well logging · Lithological classification

1   Introduction

Exploration of hydrocarbons, which are nowadays an essential source of energy, is a complicated task. It involves considerable financial expenses and using advanced techniques during data acquisition, processing and interpretation [1–3]. Moreover, data interpretation is a complicated and subjective process [4–9]. Therefore, the authors' intention was to apply advanced statistical methods, which comprises artificial neural networks (ANN) to support decision-making processes. Commonly known type of ANN has been used – a multilayer perceptron (MLP) [10, 12].

Artificial neural networks are classified as so-called computational intelligence. These are methods that are currently developing rapidly, through the emergence of new algorithms and new applications. Apart from MLP, machine learning methods comprise i.e. Support Vector Machine (SVM), single decision tree, K-means method, deep learning and Kohonen networks [9–15]. The adaptability of ANNs can be improved by evolutionary computational techniques such as float-coded genetic algorithms [16]. These methods allows to do classification, nonparametric multiple regression and time series analysis [15]. Global optimization-assisted inversion procedures not only give more accurate and reliable solution than conventional linear approaches, but more information extracted from the well logging dataset in formation evaluation [17, 18].

Since machine learning methods are being developed rapidly, a significant progress has also been made with statistical analysis software. The methods are available in commercial software (MATLAB, STATISTICA, DTREG) as well as in open-source ones (R, Octave, SciLab, Python) [10, 12, 15, 19]. Software that works under GNU GPL license has an advantage of fast implementation of new methods and algorithms. However, commercial software is often more user-friendly and better optimized regarding computing efficiency.

Therefore the paper presents a comparison of two computational environments: R and STATISTICA (v. 13). An efficiency indicator is defined as a number of correctly identified samples (in %).

The paper tackles a problem of identification of geological formations on the basis of well logging data. Data that has been used was obtained from three boreholes (A, B and C) sited in the Baltic Basin, North Poland located in the vicinity of each other. In the geological profiles of each well six potential lithostratigraphic formations (some information on the type and age would be beneficial) has been identified. Data obtained from well A were used only during the process of learning. Data from well B and C were a subject of testing the net's classification accuracy.

2   Comparison of Computational Environments

2.1   R Software

R software is a part of a larger project R which bases on the GNU GPL license. It was originally established by Robert Gentleman and Ross Ihake from University in Auckland [15] and designed for statistical analysis. Since academic researchers from all over the world contribute to this project, it grows and develops very fast. CRAN server,

which was created especially for handling the project, holds currently over 12 500 packages. This is an advantage and a disadvantage of the software at the same time. The advantage is a possibility to download every package or create your own one which includes the newest statistical solutions. The main drawback is that descriptions of packages are often not plain enough to use them in a convenient way.

There are several packages that enables applying MLP networks, e.g. neuralnet, nnet, caret. Authors have chosen the most updated and well described one – package nnet developed by Brian Ripley and William Venables from University of Oxford in 2016 [19]. The package allows to create a model of a single-hidden-layer network and predict values for new data. Considering classification, sigmoid function is the only available activation function to apply on input. In output layer softmax activation function was used. It is possible to modify many parameters of network model, i.e. size of a network, maximum number of iterations during the process of training error reduction, maximum number of weights or apply weight reduction.

Nnet's efficiency was optimized in relation to the number of nodes in a hidden layer. It is one of the most important parameter of MLP network since it strongly influences computation rate. For this purpose, 500 models of a network have been generated. Relation was examined for number of hidden nodes from the range 1 to 19 with a step equal 2. Maximum number of iterations during the process of training error reduction was assumed to be 100. Results are shown at the chart in Fig. 1.
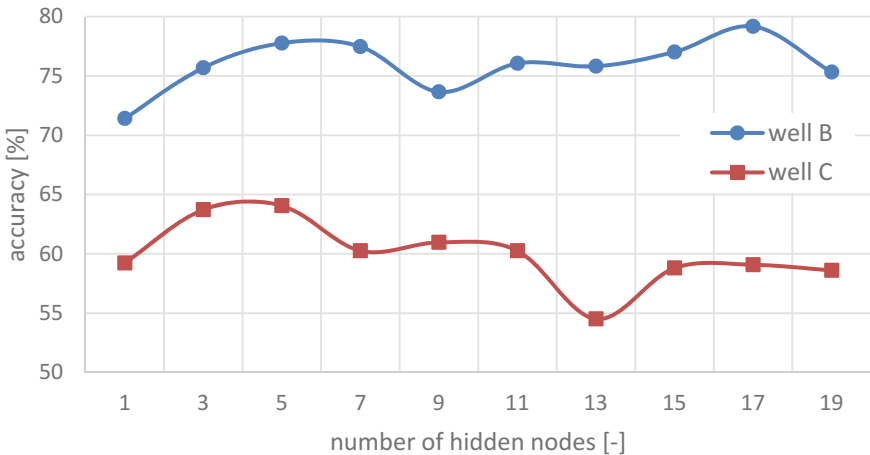

Fig. 1. Relation between classification accuracy and the number of neurons in a hidden layer for data from B and C wells in R software

On the basis of the above-presented chart, in case of well B one could notice two maxima – for 6 and 17 nodes in a hidden layer. The highest score of the correctly identified formations reached 79.1%.

The results for well C stands considerably worse. The best classification accuracy was achieved with 3 hidden nodes - accuracy ratio was equal 64%.

2.2   STATISTICA

STATISTICA offers for neural networks a separate module called SANN (STATIS-TICA Automatic Neural Networks) [20]. In a user-friendly way it allows to select data input, set the proportion in its separation into training, validation and testing sets and to select desirable net's parameters, e.g. number of hidden notes, type of a net topology (MLP or radial). The module offers automatic searching for a net based on user-defined parameters or more detailed searching, where a choice of preferable optimization algorithm is available. In this analysis an automatic searching was used. Class prediction for new data (well B and C) takes place in a different module called *Rapid Deployment of Predictive Models*. Previously saved net model is applied to selected dataset. As a result, a sheet is generated which contains in columns known lithology, a predicted one and logical value "correct" or "incorrect".

A relation between classification accuracy and a number of nodes in a hidden layer was analyzed. Default proportions in splitting the dataset into training (70%), validation (15%) and testing (15%) sets were used. A number of created nets was set to be 500. In order to compare algorithm's efficiency, activation function was the same as it was in R software (sigmoid function). The results are presented in chart Fig. 2.
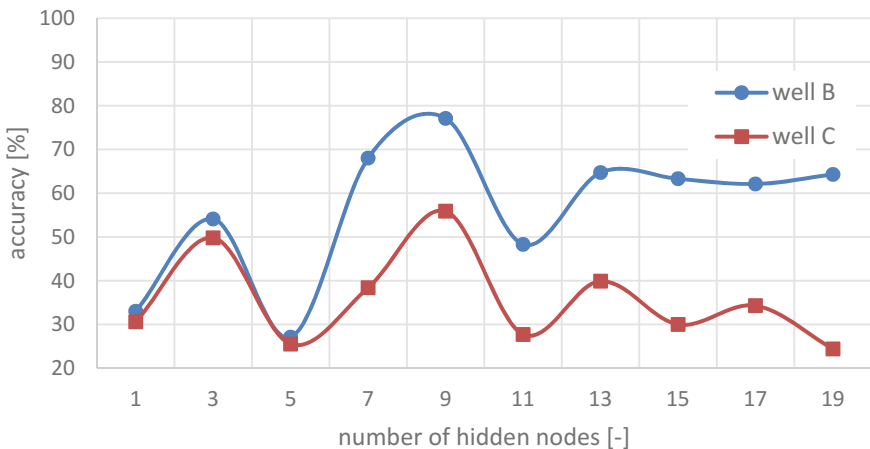
Fig. 2. Relation between classification accuracy and the number of neurons in a hidden layer for data from B and C wells in STATISTICA software. Sigmoid activation function was used

A big dispersion and instability of the results is noticeable for a number of neurons less than 12. It might be due to overlearning some net models. Overlearning occurs when during learning process a net structure matches actual dataset instead of generalized task. Then, despite small error while training, prediction is considerably incorrect. Accuracy ratio tends to be more stable for number of hidden nodes greater than 12 and is significantly lesser in case of well C. The highest accuracy (77.1% for well B and 55.9% for well C) comes out consistently for both wells at 9 nodes in a hidden layer.

Comparing the relation for both pieces of software, it is easy to notice that in case of STATISTICA a dispersion of the results is much bigger than in R. The dispersion is the highest for small values of a size of a hidden layer and for these values accuracy ratios from both pieces of software differ the most. Regarding well B, for a greater number of nodes in a hidden layer the accuracy ratio is similar in both computational environments – on average however, R achieves better accuracy by a few percent. A considerable difference occurs in well C – STATISTICA gives lesser classification accuracy by circa a half. The reason for this might be a different way of net's evaluation. In STATISTICA the best net model is chosen on the basis of the smallest values of errors (learning, testing and validation ones). Classification accuracy is calculated in a separate module. A script in R evaluates every net model immediately after training process. It is the accuracy ratio, not errors during training that are taken into consideration in this case.

The mentioned above SANN module provides bigger choice of activation functions than package nnet in R. These are: linear, sigmoid, hyperbolic tangent, exponential and sine function. It is worth mentioning, that MLP nets use only a part of a sine function that falls in an interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, where sine is a monotonic function [20]. Then, an influence of an activation function on classification accuracy has been examined. Constant value of a number of hidden nodes was used. Based on the chart in Fig. 2, it was assumed to be 9. A histogram in Fig. 3 presents how a type of activation function affects net's efficiency.
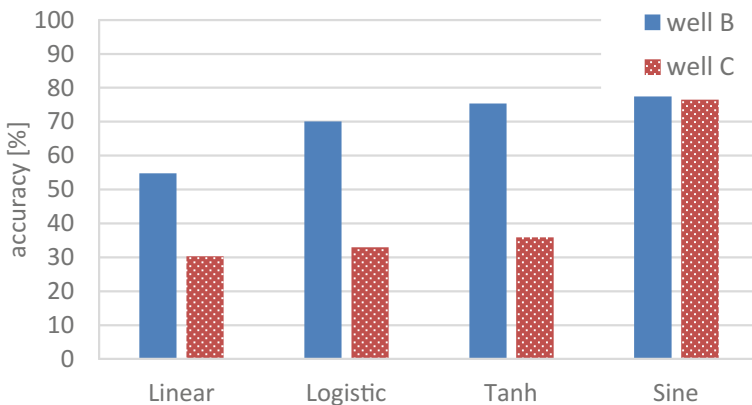
Fig. 3. Histogram presenting classification accuracy ratio for different activation functions for data from B and C wells

For both wells the highest accuracy (reaching almost 80%) was obtained by a net with a sine activation function. In case of well C the results are over twice as better as for other functions. Considering well B, the differences are considerably smaller – either a sigmoid function or hyperbolic tangent give accuracy ratio smaller by a few percent.

On the other hand, the lowest accuracy was achieved by a linear activation function. This is due to the fact that relations between log's values and rock mass are too complicated, i.e. normally approximated by multivariate non-linear probe response functions, to be described by such a simple formula.

3   Conclusions

Problem of ANN optimization in two computational environments - R (package nnet) and STATISTICA was presented in the paper. Optimization was carried out for a task of identification of geological formations in boreholes based on well logging data. It allowed to compare a quality and efficiency of statistical computations for both pieces of software. Based on the results of conducted investigations, one could conclude that for the same type of activation function (sigmoid function), better accuracy was obtained in STATISTICA software.

Additionally, for the sake of a bigger choice of activation functions in STATIS-TICA, an influence of a type of activation function on classification accuracy was examined. In this case the best results was achieved by a sine, which in an interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ is similar to a sigmoid function. Taking into consideration the fact, that STATISTICA, which is an example of commercial software, has got more user-friendly interface, one has to admit that its efficiency is better than in case of R. On the other hand, R is an open-space environment that is under the constant development by a scientific community. Therefore some solutions (especially new ones) are available there faster than in STATISTICA.

However, it is worth observing that both computational environments allow lithology classification on the basis of well logging data from the analyzed wells with satisfying accuracy (around 80%). Therefore ANN – MLP could be a remarkable tool supporting interpretation of well logging data.

References

1. Kaźmierczuk, M., Jarzyna, J.: Improvement of lithology and saturation determined from well logging using statistical methods. Acta Geophys. **54**, 378–398 (2006)
2. Puskarczyk, E., Jarzyna, J., Porębski, S.J.: Application of multivariate statistical methods for characterizing heterolithic reservoirs based on wire line logs – Example from the carpathian Foredeep basin (Middle Miocene, SE Poland). Geol. Q. **59**, 157–168 (2015)
3. Szabó, N.P.: Hydraulic conductivity explored by factor analysis of borehole geophysical data. Hydrogeol. J. **23**, 869–882 (2015)
4. Rogers, S.J., Fang, J.H., Karr, C.L., Stanley, D.A.: Determination of lithology from well logs using a neural network. Am. Assoc. Pet. Geol. Bull. **76**, 731–739 (1992)
5. Benaouda, D., Wadge, G., Whitmarsh, R.B., Rothwell, R.G., MacLeod, C.: Inferring the lithology of borehole rocks by applying neural network classifiers to downhole logs: an example from the ocean drilling program. Geophys. J. Int. **136**, 477–491 (1999)

6. Bhatt, A., Helle, H.B.: Determination of facies from well logs using modular neural networks. Pet. Geosci. **8**, 217–228 (2002)
7. Zhou, J., Yan, J., Pan, L.: Application on lithology recognition with BP artificial neural network. In: 3rd International Symposium on Intelligent Information Technology Application, IITA 2009. Pp. 56–59 (2009)
8. Parvizi, S., Kharrat, R., Asef, M.R., Jahangiry, B., Hashemi, A.: Prediction of the Shear Wave Velocity from Compressional Wave Velocity for Gachsaran Formation. Acta Geophys. **63**, 1231–1243 (2015)
9. Puskarczyk, E.: Applying of the Artificial Neural Networks (ANN) to identify and characterize sweet spots in shale gas formations. In: E3S Web of Conferences 35, 03008 (2018)
10. Jarzyna, J., Zych, M., Krakowska, P., Puskarczyk, E., Wawrzyniak-Guz, K.: Total organic carbon from well logging – statistical approach, Polish shale gas formation case study, Int J Oil, Gas Coal Technol (in printing – Forthcoming articles)
11. Roshani, G.H., Hanus, R., Khazaei, A., Zych, M., Nazemi, E., Mosorov, V.: Density and velocity determination for single-phase flow based on radiotracer technique and neural networks. Flow Meas. Instrum. **61**, 9–14 (2018)
12. Hanus, R., Zych, M., Kusy, M., Jaszczur, M., Petryka, L.: Identification of liquid-gas flow regime in a pipeline using gamma-ray absorption technique and computational intelligence methods. Flow Meas. Instrum. **60**, 17–23 (2018)
13. Roshani, G.H., Nazemi, E.: A novel dual-molality densitometer for gauging in annular two phase flows using radial basis function. Kerntechnik **83**, 145–151 (2018)
14. Al-Anazi, A., Gates, I.D.: On the capability of support vector machines to classify lithology from well logs. Nat. Resour. Res. **19**, 125–139 (2010)
15. Karatzoglou, A., Smola, A., Hornik, K., Zeileis, A.: kernlab – an S4 Package for Kernel methods in R. J. Stat. Softw. **11**, 1–20 (2004)
16. Tettamanzi, A.G.B., Tomassini M.: Soft Computing - Integrating Evolutionary, Neural, and Fuzzy Systems. Springer (2001)
17. Dobróka, M., Szabó, N.P.: Interval inversion of well-logging data for automatic determination of formation boundaries by using a float-encoded genetic algorithm. J. Pet. Sci. Eng. **86**, 144–152 (2012)
18. Szabó, N.P.: A genetic meta-algorithm-assisted inversion approach: hydrogeological study for the determination of volumetric rock properties and matrix and fluid parameters in unsaturated formations. Hydrogeol. J. **26**, 1935–1946 (2018)
19. Ripley, B., Venables, W., Package 'nnet', https://cran.r-project.org/web/packages/nnet/nnet.pdf. Accessed 15 Aug 2018
20. STATISTICA Help. http://documentation.statsoft.com. Accessed 15 Aug 2018

Author Index