Automated system of audio components analysis and synthesis

Viktoriia V. Voitko, Svitlana V. Bevz, Sergii M. Burbelo, Pavlo V. Stavytskyi, Bogdan Pinaiev, et al.

Viktoriia V. Voitko, Svitlana V. Bevz, Sergii M. Burbelo, Pavlo V. Stavytskyi, Bogdan Pinaiev, Zbigniew Omiotek, Doszhon Baitussupov, Aigul Bazarbayeva, "Automated system of audio components analysis and synthesis," Proc. SPIE 11045, Optical Fibers and Their Applications 2018, 110450V (15 March 2019); doi: 10.1117/12.2522313

SPIE.

Event: 18th Conference on Optical Fibers and Their Applications, 2018, Naleczow, Poland

Automated system of audio components analysis and synthesis

Viktoriia V. Voitko*[1], Svitlana V. Bevz[1], Sergii M. Burbelo[1], Pavlo V. Stavytskyi[1], Bogdan Pinaiev[1], Zbigniew Omiotek[2], Doszhon Baitussupov[3], Aigul Bazarbayeva[4]
[1]Vinnytsia National Technical University, Vinnytsia, Ukraine; [2]Lublin University of Technology, Lublin, Poland; [3]Kazakh Academy of Transport and Communication, Almaty, Kazakhstan; [4]Scientific Center of Pediatrics and Child Surgery, Almaty, Kazakhstan

ABSTRACT

The work is aimed to research the process of synthesis and analysis of musical compositions. The main goal is to increase speed and quality of this process. This is done by combining sound analysis algorithms with music synthesis technologies. By this method music taken by the device microphone is processed by the mobile application and then transformed to the note sequence which could be edited by the user.

Keywords: mobile application, Android, musical sounds analysis, spectrogram, music recognition

1. INTRODUCTION

There are a lot of applications for music creation on the market, but they require deep knowledge music theory and professional software. It will be difficult for users to "digitize" melodies from their mind using existing software as it requires certain level of skills to use it. It is hard to find an application which allows to record a voice melody by the microphone and then transform it into the form eligible for editing and play with a certain musical instrument[1,2]. The developed software system allows to solve described problem and provides simple to use tool for creating musical compositions without deep knowledge of the music theory.

The research goal is to increase automation level of music creation process by combining sound analysis algorithms with music synthesis technologies. It allows to transform voice sound into music sequence played by specific musical instrument. The research object is technology stack for music synthesis and analysis. The research subject are software and algorithms for automation of the music creation process. The main research goal is to develop an automated mobile system with using of modern technologies for music synthesis and analysis.

2. CURRENT PROBLEM STATE ANALYSIS

Popularity of mobile devices is increasing every day. Mobile operating systems such as Android and IOS became the main players on the market overtaking desktop ones. This happened because of wide popularity of developed mobile applications and large number their users. There are dozens of applications for music creation on the market. Some of them allows to recognize existing musical compositions others provide functionality for their creation. Such software requires different level of competency in the music industry.

There are several competitors of the developed mobile software system. First of them is Shazam – application which can identify music, movies, advertising, and television shows. It is based on a short sample played and using the microphone on the device. The software is available for Android, macOS, iOS, watchOS and tvOS.

Second competitor is Music Maker Jam. This application allows user to create his own musical compositions by using prerecorded samples. It provides wide range of music styles which could be applied to the sequence. Some of them are paid and others are free. One of the primary feature of Music Maker Jam is adding effects to the composition in real time.

*dekanfki@i.ua

Optical Fibers and Their Applications 2018, edited by Ryszard S. Romaniuk, Waldemar Wójcik, Andrzej Smolarz, Proc. of SPIE Vol. 11045, 110450V

Final competitor described in the article is Yandex.Music. This service allows to search and legally for free listening to musical compositions, albums and collections of music tracks. One of the newest features of Yandex.Music if music recognition feature. It allows user to detect metadata of the recorded music fragment. The comparative description of existing solutions with the developed application is given in Table 1.

Table 1. Comparative description of existing solutions with the developed application.

Product feature	Shazam	Music Maker Jam	Yandex.Music	Developed application
Music recognition	1.0	0.0	1.0	1.0
Ability to create custom music	0.0	1.0	0.0	1.0
Range of available musical instruments	0.0	0.4	0.0	0.8
User experience	0.8	0.6	0.7	0.9
Total:	1.8	2.0	1.7	3.7

Music recognition and custom music creation features are represented in binary format where 1 means feature presence and 0 – its absence. Other features evaluated in a range from 0 to 1. Thus, after feature comparison of the given software developed application scored 3.7 points which means that such software is relevant for development.

3. DEVELOPMENT OF THE SYSTEM MODELS AND ALGORITHMS

One of the main advantages of the developed system of analysis and synthesis of musical sounds is the use of the combined method of creating musical compositions. It is based on the technologies musical sounds synthesis and algorithms for their analysis.

The developed system provides automation of the audio components handling[3] and consists of two modules:

- module for musical sounds synthesis;
- module for musical compositions comparison and analysis.

The module for synthesizing musical sounds is responsible for generating musical instrument sound. This module provides music creation mode. This mode has a sequencer functionality and allows to create music using criterion modeling principles[4] and logic-mathematical apparatus for formalization and description of information processes[5,6]. It is represented as a two-dimensional space in which the OX scale is a timeline and OY is responsible for the height of the notes. Thus, when placing a mark in a certain area of the working space at a specific moment of time, a certain note will be played that corresponds to the position of the given mark.

Synthesis of musical sounds is based on OpenSL ES technology – a multi-platform application programming interface for working with sound for embedded systems – smartphones, tablet computers, gaming consoles. This technology provides the ability to play and record sound, playback surround sound, customize musical effects, support interactive music and ringtones, and more. The object model of the OpenSL ES is based on the OpenMax software interface, which provides an abstraction level to unify the process of working with audio, video and photo formats. The Native Development Kit (NDK) for the Android operating system includes an adapted implementation of this interface and is designed to develop digital audio workstations (DAWs), synthesizers, drum machines, and audio effects[7-9].

Sound analysis module is responsible for input signal sampling. It returns arrays with sound signal representation in the time domain. It is necessary to use a discrete Fourier transform (DFT) to obtain the frequency characteristics of signals deployed in time. One of the most popular numerical algorithms for DFT calculation is the fast Fourier transform (FFT). In fact, the FFT is represented by a set of algorithms. Among them, the most commonly used variants of the Kulu-Tukey algorithm are used. The basis of this algorithm is the principle of "Divide and rule". In the course of calculations, recursive decomposition of the original DFT into small parts is used[10-12]. Music creation process is displayed in the Fig. 1.

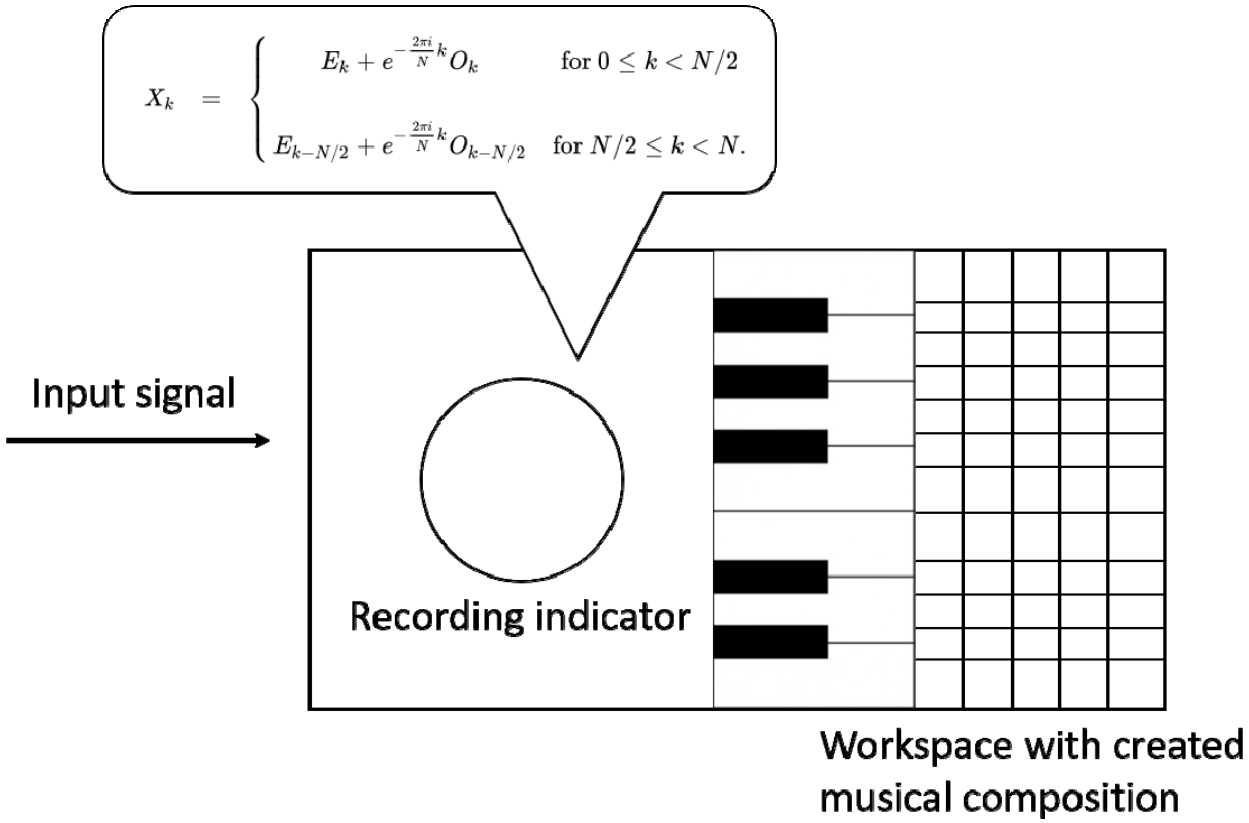

Figure 1. Music creation process.

To simplify the search for musical compositions, their signatures are used as keys in the hash table. The keys correspond to the time values when the set of frequencies for which the signature was found, appeared in the composition, and the identifier of the composition itself (for example, the title of the song and the name of the artist).

As a result of the identification process, the procedure for recognizing music is performed by the method of analysis and synthesis of musical sounds, the model of which is shown in Fig. 2:

1. Record the original sound.
2. Transformation of the signal into the frequency representation.
3. Creating an imprint for comparison.
4. Saving to the database in the required format

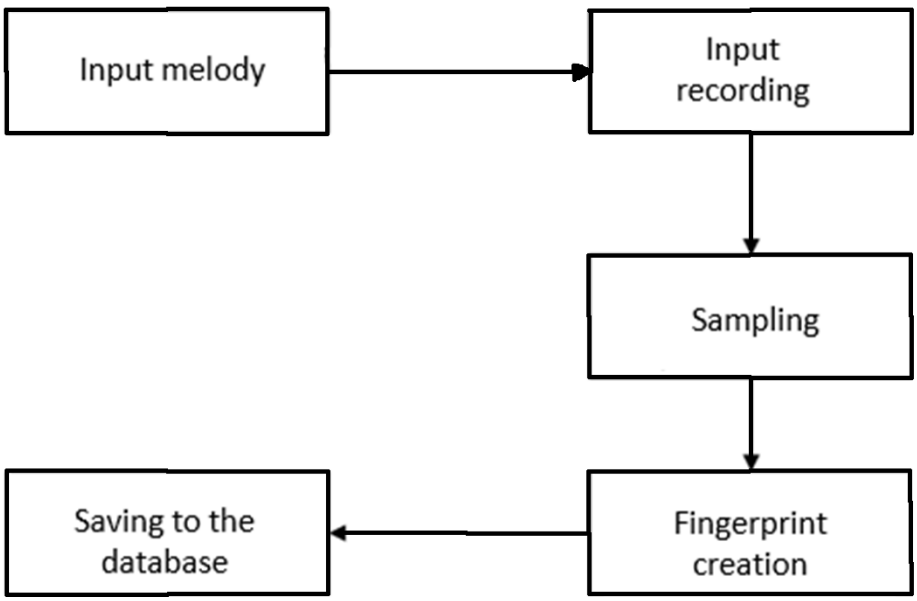

Figure 2. The model of music recognition process.

The next step is to identify composition wave amplitudes and create fingerprints. This allows to recognize music even with a high noise level. The model of the music synthesis and analysis is displayed in the Fig. 3.

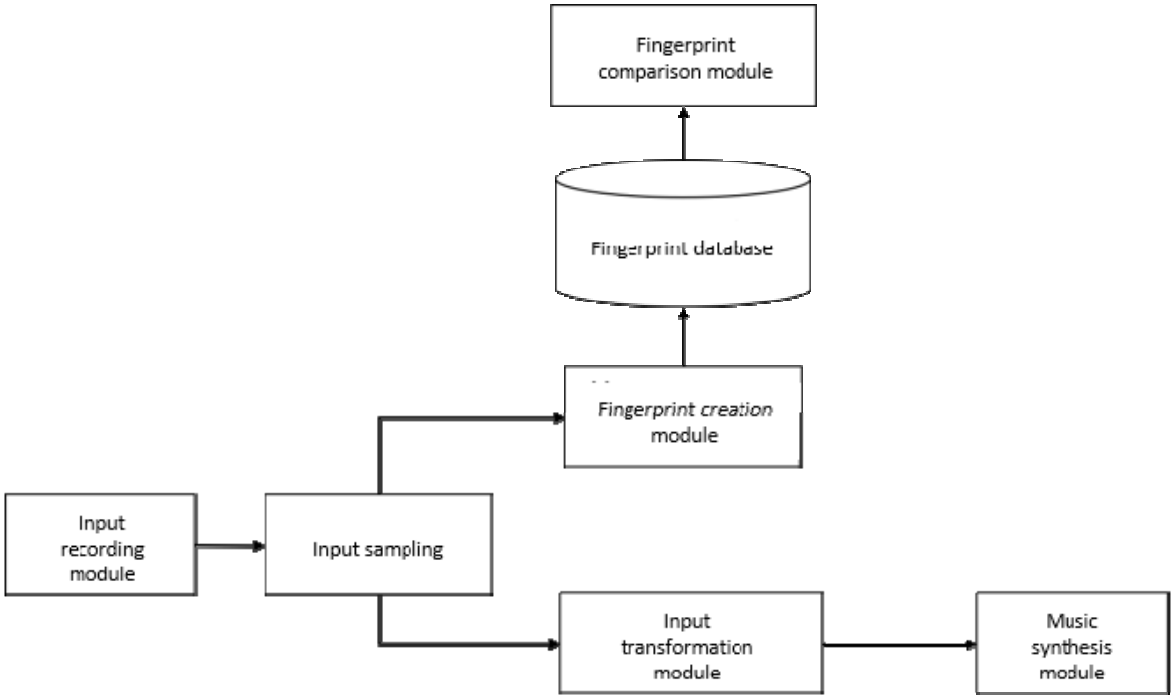

Figure 3. The model of the music synthesis and analysis.

Comparison of complete spectrographs causes performance issues, since it processes a few megabytes of unmanaged data. The presence of third-party noise is also negatively affects the algorithm effectiveness. However, when playing music tracks with different kinds of noise, the peaks of spectrograms are best kept. The peaks are represented as local maximums of the amplitude. As a result, spectrograms are obtained by a "constellation" with such peaks. Therefore, when searching for a result for the requested segment of a melody it is necessary find the corresponding track in the

database, in which the coincidence of spectrogram peaks is the largest. In addition, there is a possibility to save hash values based on the peaks instead of raw peak data. It allows to optimize the matching process[13-17].

Provided algorithm shows that user is able either to create new musical sequence or to recognize existing one. It allows to provide a flexibility of the application. It is possible to use described features separately from each other if needed. However, music creation process significantly accelerates when using these features together (Fig.4).

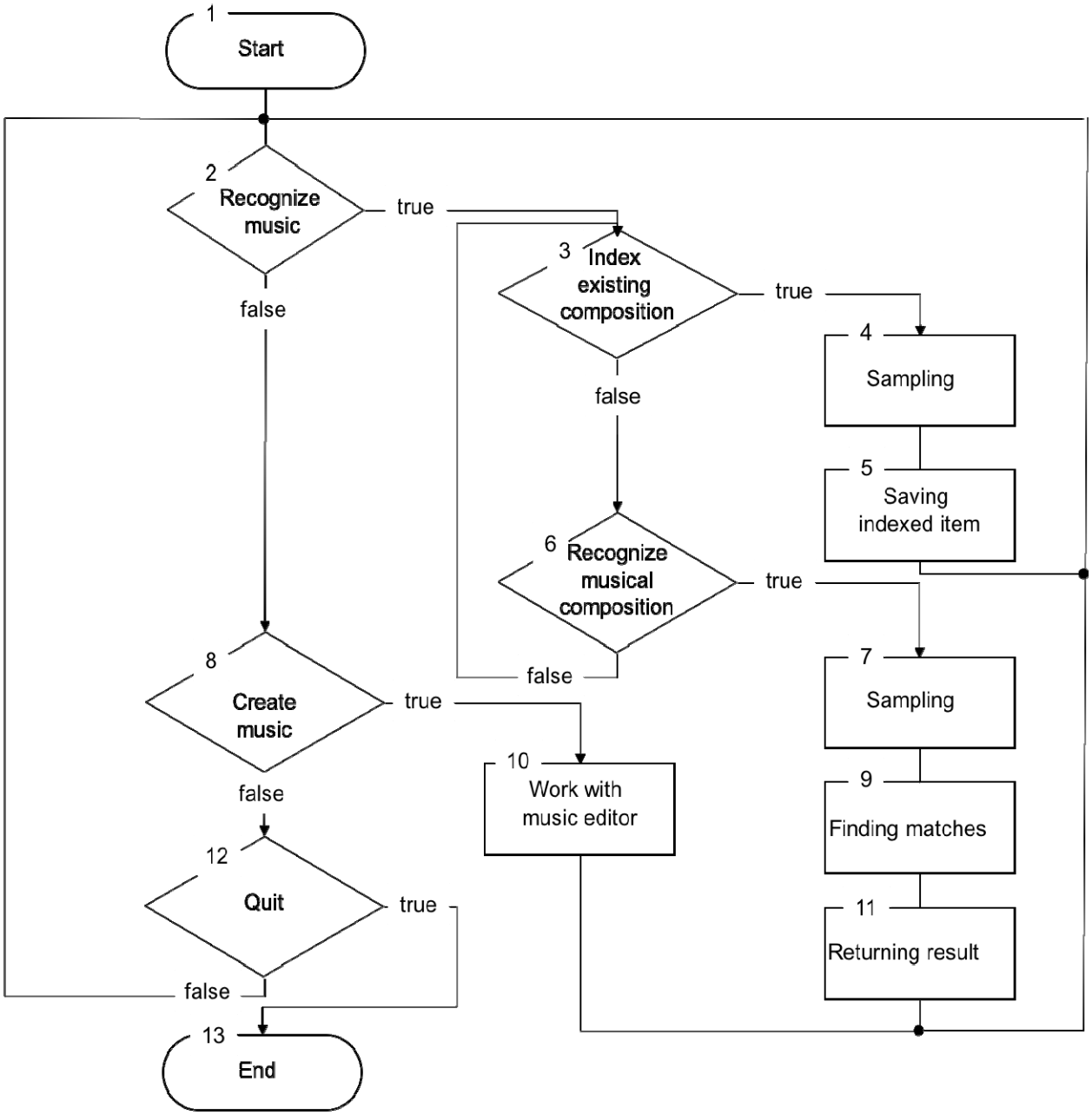

Figure 4. Sound synthesis and analysis system algorithm.

Each audio file is represented in a fingerprint format. The application compares fingerprints while new input is recorded and returns the most relevant result. Music recognition algorithm is described in the Fig. 5.

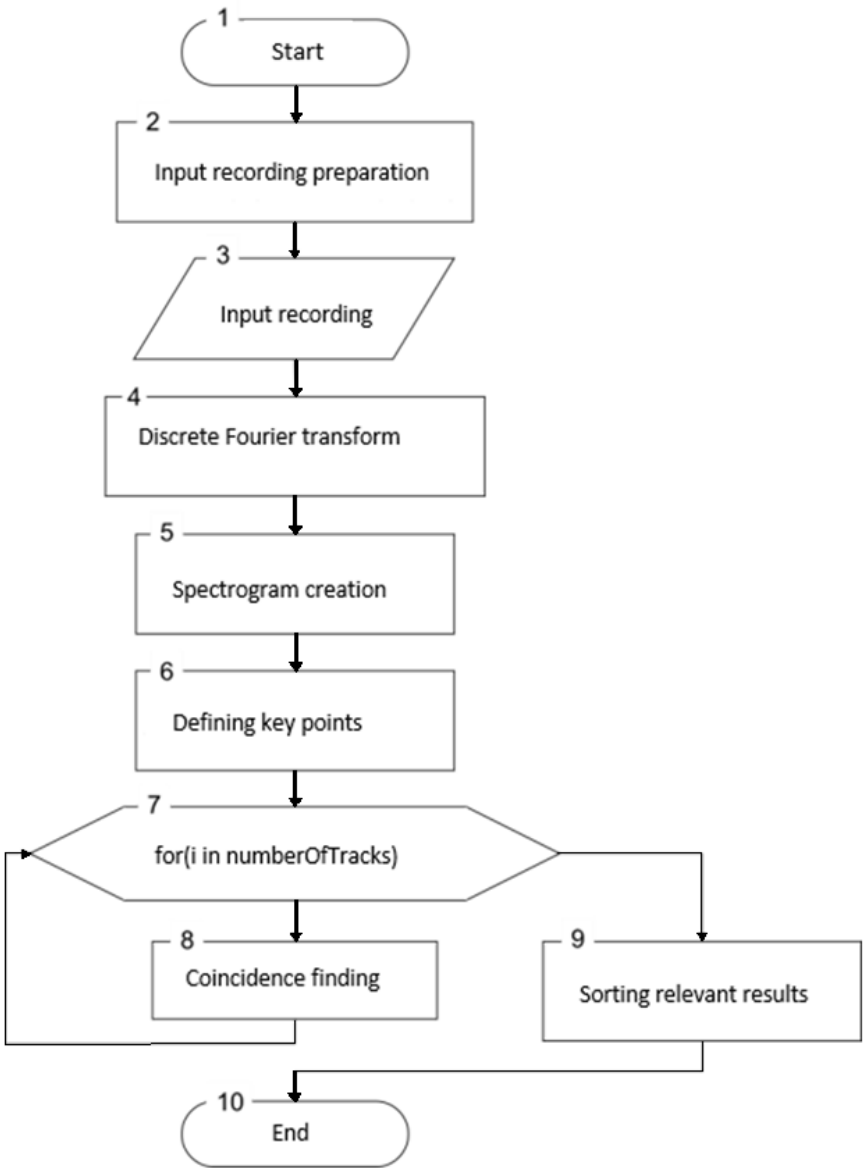

Figure 5. Music recognition algorithm.

4. CONCLUSIONS

Automated software system for music synthesis and analysis was developed. It is based on the combined method which includes music recognition algorithms and technologies for its creation. The given approach increases speed and quality of music creation process which allows users to be more productive while making music. The application was developed for mobile operation system Android using Java, Kotlin and C++ programming languages.

REFERENCES

[1] Martin, R., [Agile Software Development, Principles, Patterns, and Practices 1st Edition], Pearson, London, 529 (2002).

[2] Lloyd, W., [Music in Sequence: Complete Guide to MIDI Sequencing. Paul Terry], Musonix Publishing, London, 112 (1996).

[3] Romanyuk, A. N., Bevz, S. V. and Burbelo, S. M., "Building automation of the computer systems of management reporting," Proceedings of the 2011 International Siberian Conference on Control and Communications, SIBCON 2011, (2011).

[4] Bevz, S., Voytko, V., Burbelo, S., et al., "Criterion modeling in the control problems," Proceedings of the 2011 International Siberian Conference on Control and Communications, SIBCON 2011, (2011).

[5] Bevz, S., "Logic-mathematical apparatus of data processing used in information technology of web-portal development," CEUR Workshop Proceedings, (2014).

[6] Lezhniuk, P., Bevz, S. and Piskliarova, A., "Evaluation and forecast of electric energy losses in distribution networks applying fuzzy-logic," Proceedings of the IEEE Power and Energy Society 2008 General Meeting: Conversion and Delivery of Electrical Energy in the 21st Century, PES, (2008).

[7] Lezhnyuk, P. D., Kulik, V. V. and Burikin, O. B., [Mutual influence of electric networks in the process of optimal control of their modes], UNIVERSUM, Vinnytsya, 123 (2008).

[8] Melnikov, N. A., "Matrix analysis of electrical circuits," Energy 232, (1972).

[9] Vedmitskyi, Y. G., Kukharchuk, V. V. and Hraniak, V. F., "New non-system physical quantities for vibration monitoring of transient processes at hydropower facilities, integral vibratory accelerations," Przeglad Elektrotechniczny 93(3), 69-72 (2017).

[10] Kukharchuk, V. V., Kazyv, S. S. and Bykovsky, S. A., "Discrete wavelet transformation in spectral analysis of vibration processes at hydropower units," Przeglad Elektrotechniczny 93(5), 65-68 (2017).

[11] Kukharchuk, V. V., Hraniak, V. F., Vedmitskyi, Y. G., et al., "Noncontact method of temperature measurement based on the phenomenon of the luminophor temperature decreasing," Proc. SPIE 10031, (2016).

[12] Kukharchuk, V. V, Bogachuk, V. V., Hraniak, V. F., et al., "Method of magneto-elastic control of mechanic rigidity in assemblies of hydropower units," Proc. SPIE 10445, (2017).

[13] Azarov, O. D., Krupelnitskyi, L. V., Komada, P., "AD systems for processing of low frequency signals based on self calibrate ADC and DAC with weight redundancy," Przeglad Elektrotechniczny 93(5), 125-128 (2017).

[14] Azarov, O. D., Dudnyk, O. V., Kaduk, O. V., et al., "Method of correcting of the tracking ADC with weight redundancy conversion characteristic," Proc. SPIE 9816, (2015).

[15] Azarov, O. D., Murashchenko, O. G., Chernyak, O. I., et al., "Method of glitch reduction in DAC with weight redundancy," Proc. SPIE 9816, (2015).

[16] Ali Bagher, B., Wójcik, W., Mamyrbayev, O., et al., "Speech recognizer-based non-uniform spectral compression for robust MFCC feature extraction," Przeglad Elektrotechniczny 94(6), 90-93 (2018).

[17] Mashkov, V., Smolarz, A., Lytvynenko, V., et al., "The problem of system fault-tolerance," Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Srodowiska 4, 41-44 (2014).