

**ITC 3/46**

Journal of Information Technology  
and Control  
Vol. 46 / No. 3 / 2017  
pp. 372-381  
DOI 10.5755/j01.itc.46.3.17582  
© Kaunas University of Technology

**The Use of Wavelet Transformation in Conjunction  
with a Heuristic Algorithm as a Tool for Feature Extraction  
from Signals**

Received 2017/03/23

Accepted after revision 2017/07/03


<http://dx.doi.org/10.5755/j01.itc.46.3.17582>

# The Use of Wavelet Transformation in Conjunction with a Heuristic Algorithm as a Tool for Feature Extraction from Signals

**Dawid Połap and Marcin Woźniak**

Institute of Mathematics, Silesian University of Technology, Kaszubska 23, 44-100 Gliwice, Poland  
e-mail: Dawid.Polap@polsl.pl, Marcin.Wozniak@polsl.pl

Corresponding author: Dawid.Polap@polsl.pl

The use of algorithms is helpful in analysis of various data samples. Examples of these applications are sound and graphic processing that are used in authentication and analysis of images. In this work, we propose technique of extracting data from image file created based on voice sample. Proposed method makes use of mathematical model of wavelet transformation and its graphical visualization like scaleogram combined with computational intelligence methods like neural network and heuristic algorithm. In order to verify operation of proposed technique we present experimental results.

**KEYWORDS:** Neural networks, speaker recognition, wavelet transformation, heuristic algorithm, feature extraction, image processing.

## 1. Introduction

Nowadays, audio signals are very common for their applications in many areas of today life. First of all, search for new forms of security and identity verification is necessary. Classical solutions do not support fast and reliable verification, but cause queues of workers in large corporations. It is one of motivations to improve and seek alternative solutions. Of course, it is

not only identity verification but also other fields where any other types of signals are used. One of applications of similar analysis was proposed in [3], where authors presented use of spectrograms for music classification. Again in [21], the idea to use audio signal analysis as tool in heavy industry was presented. The main application of these method was intended

for metal stamping. Signals find their use also in soft *computing* like user identity verification in smartphones [5, 12] and speaker recognition for micro manipulation of robots [20]. It is highly recommended that all systems work automatically and that was presented in [1]. The idea of automatic speech recognition was presented for under-resourced languages [14], which can data structures for knowledge storing [6, 7].

Application is among topics, which scientists are focused on, but on the other hand, theoretical aspects of analysis and its functionality are very important. In [16], authors analyzed high resolution property of group delay function in terms of signals processing and for contrast of all theoretical aspects, in [8] wavelet decomposition for vibration simulation signal was shown.

The use of signals requires many other tools like processing and classification methods. The most popular classifiers are grouped into two type - fuzzy and neural. In the first type, fuzzy approach is based on rough set and in [11], authors presented method for fast image classification based on that idea. It may be hard to define a specific value in fuzzy controller, that is why some researches work on definition of *less, more* etc. [15]. The second type is neural network which is simplified model of human brain. Some application of such solution is presented in [13] where authors showed the idea of classification using incomplete or missing data. All possible solutions are created based on processing methods [10] and modeling as well as some rules which are called programming languages [18, 4]. Mentioned rules are related to data record, which is required for certain embodiments of technology. And in [14] number of advantages and disadvantages of software-solution architectures were presented. Similarly, heuristics are currently used in various models to simulate and optimize [2, 17].

In this paper, we would like to connect few of these ideas linked in effective tool for feature extraction from *voice* signals. We presented selected application (like voice verification) as a tool for testing our approach and an indication of its advantages.

## 2. Wavelet Transform

The history of wavelet theory dates back to early twentieth century, when Hungarian mathematician, Alfred Haar presented the first known wavelet (today known

as Haar wavelet) [9]. The author applied orthonormal wavelet to show an example of system in interval  $<0,1>$ . That wavelet is function defined as follows

$$\psi(x) = \begin{cases} 1 & \text{for } 0 \leq x < \frac{1}{2} \\ -1 & \text{for } \frac{1}{2} \leq x < 1 \\ 0 & \text{for others } x \end{cases} \quad (1)$$

70 years later, French geophysicist, Jean Morlet has introduced formal definition of wavelet.

**Definition 1.** Wavelet is a function  $\psi_{a,b}:R \rightarrow R$  coming from parent function  $\psi$  with scaling  $a$  and shift  $b$  (in effect giving entire family of functions). Wavelet of scale  $a$  and shift  $b$  is defined as follows

$$\psi_{a,b}(x) = \psi(2^a x + b). \quad (2)$$

The creation of formal definition of wavelet allowed for rapid growth of mathematical aspects associated with it. Within few years, researchers have worked on models applicable for transformation and its fundamental operations.

### 2.1 Continuous Wavelet Transform

In wavelet representation, signal is projected onto continuous family of frequencies. In most cases, bandwidth is scaled in scale equal to 1 and shifted by parent function. An example of such function is  $\text{sinc}(x)$ , what can be presented as

$$\psi(x) = 2\text{sinc}(2x) - \text{sinc}(x) = \frac{\sin(2\pi x) - \sin(\pi x)}{\pi x}, \quad (3)$$

where

$$\text{sinc}(x) = \begin{cases} \frac{\sin(\pi x)}{\pi x} & \text{for } x \neq 0 \\ 1 & \text{for } x = 0 \end{cases}. \quad (4)$$

Generalizing, subspaces in scale  $a$  are generated by positively defined function (called child function) and presented as

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right), \quad (5)$$

where  $\left(\frac{x-b}{a}\right)$  is called core of transformation, and  $a \neq 0$ .

**Definition 2.** Core of transformation is called function  $\psi(x)$  that satisfies condition of admissibility

$$\int_0^\infty \frac{|\psi(\omega)|}{\omega} d\omega < \infty, \tag{6}$$

where  $\psi(\omega)$  is Fourier transform for wavelet  $\psi(x)$ . Eq. (6) shows that function must satisfy the following equality

$$\int_{-\infty}^\infty \psi(x) dx = 0. \tag{7}$$

**Definition 3.** Continuous wavelet transform is called integral transformation defined as

$$\hat{s}_\psi(a, b) = \frac{1}{\sqrt{a}} \int_{-\infty}^\infty s(x) \psi\left(\frac{x-b}{a}\right) dt. \tag{8}$$

**Theory 1.** Inverse wavelet transform is defined as

$$s(x) = \frac{1}{\int_0^\infty \frac{|\psi(\omega)|}{\omega} d\omega} \int_0^\infty \frac{1}{a^2} \int_{-\infty}^\infty \hat{s}_\psi(a, b) \psi\left(\frac{x-b}{a}\right) db da \tag{9}$$

**Proof.** Let us describe kernel by Eq. (6). Eq. (8) can be defined as a real product of signal  $s(x)$  with function  $\psi_{a,b}(x)$  as

$$\hat{s}_\psi(a, b) = \int_{-\infty}^\infty \psi'_{a,b}(x) s(x) dx. \tag{10}$$

When function satisfies condition of admissibility (see Eq. (6)), signal  $s(x)$  can be obtained  $\hat{s}_\psi(a, b)$  from wavelet transform to yield Eq. (9).

Ultimately, wavelet transform for signal  $s(x)$  returns number of coefficients that depend on parameters  $a$  and  $b$ . These coefficients are measures of similarity between signal and wavelet.

### 2.2 Discrete Wavelet Transform

For discretization of wavelet transform, its parameters  $a, b$  and wavelet itself must be given. Parameter of scaling takes form  $a = a_0^m$  and is shifting  $b = nb_0$ , just as we define wavelet

$$\psi_{m,n} = a_0^{-m/2} \psi\left(\frac{t-nb_0}{a_0^m}\right), \quad n, m \in Z. \tag{11}$$

**Definition 4.** Discrete wavelet transform is defined as follows

$$\hat{s}_\psi(m, n) = \int_{-\infty}^\infty \psi'_{m,n}(x) s(x) dx. \tag{12}$$

**Definition 5.** Inverse wavelet transform for discrete signal is defined as

$$s(x) = k_\psi \sum_m \sum_n \hat{s}_\psi(m, n) \psi_{m,n}(x), \tag{13}$$

where  $k_\psi$  is constant for normalization purpose.

#### 2.2.1 Mallat's Algorithm

In 1988, Stephen Mallat proposed an iterative algorithm to determine coefficients of wavelet transform for signal  $s(x)$ . Proposed algorithm is composed of two stages – decomposition and reconstruction of the signal, and may be used in order for distribution of various filters.

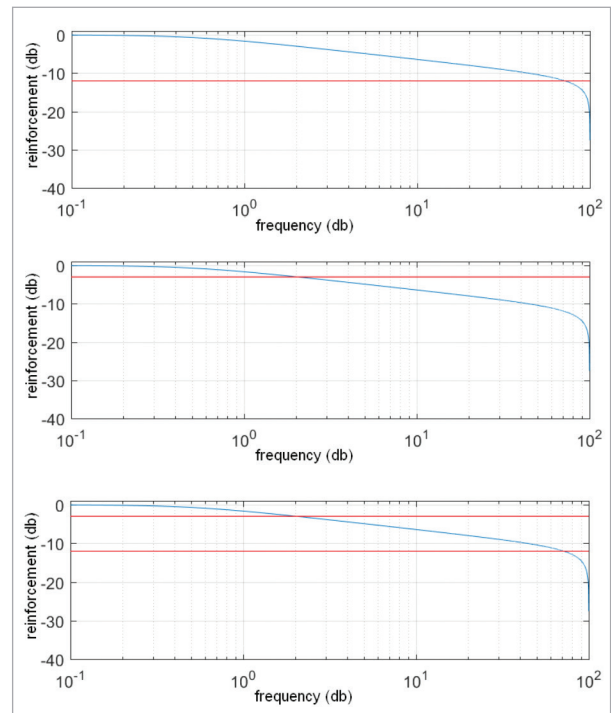
**Definition 6.** Filter is function which attenuates signal according to frequency value. There are three types of these filters

- lowpass – signal is passed only below particular value,
- highpass – signal is passed only above particular value,
- bandpass – signal is passed only between particular values.

Illustration of these filters is shown in Fig. 1.

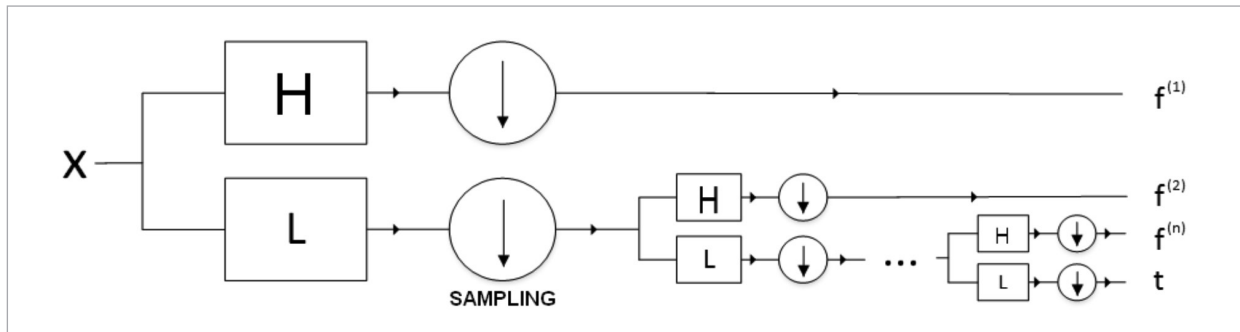
**Figure 1**

Visualization of filters actions. Starting from top we can see highpass, lowpass and bandpass filters



**Figure 2**

Graphical representation of Mallat's algorithm, which helps to speed up wavelet decomposition and reconstruction



The algorithm assumes distribution of signal into two components – lower and high half of the signal bandwidth using suitable filters. The author of the algorithm proposed to use spline function to decompose (high-pass and low-pass filter). Then, convolution with the signal is applied with Fourier transform. Filter functions are described as

$$H(\omega) = \sqrt{\frac{S_{2n}(\omega)}{2^{2n}S_{2n}(2\omega)}} \tag{14}$$

$$L(\omega) = \exp(-i\omega) \overline{H(\omega + \pi)}, \tag{15}$$

where  $S_{2n}$  is spline function of third degree.

This method uses pyramidal decomposition scheme, that is, in each iteration an additional element or row is added until the last point of analyzed signal. Graphical representation of Mallat's algorithm is shown in Fig. 2.

### 3. Heuristic Algorithm

Complex problems require efficient ways for solving, that will not enlarge execution time, consumption of computing power, etc. One of these, inspired by nature, are called heuristic algorithms.

**Definition 7.** Heuristic algorithm is a method of finding optimal solution for given problem, however with condition that returned solution might be improved.

In each heuristic algorithm, we assume that individuals are interpreted as points in  $n$ -dimensional space and described as  $\mathbf{x} = (x_1, x_2, \dots, x_n)$ . Number of individuals in population is defined as  $k$  and number of iterations  $T$  is constant. Algorithm is based on modeling movement of

individuals for the best adaptation, which is determined by function called fitness function.

#### 3.1 Wolf Search Algorithm

Wolves are individuals that live in herds that try to steer clear of enemies while searching for food. Mathematical model of behavior of wolves has been described as a heuristic algorithm by Tang et al. [19] Because of the fact that wolves live in herd, model assumes that the greater the distance, the less attractive is place. Each individual is moving in search for food, movement of which is modeled by

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \beta_0 \exp(-r^2) (\mathbf{x}_j^t - \mathbf{x}_i^t) + \gamma, \tag{16}$$

where  $t$  is iteration of the algorithm,  $\beta_0$  is parameter defining incentive of place,  $\mathbf{x}_j^t$  is the closest neighbor of  $\mathbf{x}_i^t$  with better value of fitness function and  $\gamma$  is random value in range  $\langle 0,1 \rangle$  and  $r$  is distance between these two wolves defined as

$$r = d(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{\sum_{p=1}^n (\mathbf{x}_{i,p} - \mathbf{x}_{j,p})^2}. \tag{17}$$

The hunt of wolves is defined as stalking process which consists of three phases. The first of them is called initiative stage which is understood as movement of individual within his sight in search of a better place. This step is modeled as

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \alpha v \gamma, \tag{18}$$

wherein  $v$  is velocity of wolf. The second stage is called passive and it means to stay in current location. The last stage is escape which is known as movement when

enemy will be in vicinity, modeled by

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^t + \alpha s y, \quad (19)$$

where  $s$  is step size. The complete algorithm is presented in Algorithm 1.

#### Algorithm 1 Wolf Search Algorithm

```

1: Start
2: Define all the parameters – the number of iterations  $T$ ,
   number of wolves  $n$ , radius of the visual range  $r$ , step size
    $s$ , velocity coefficient  $\alpha$ , coefficient of enemy appearance
    $p_\alpha$  and fitness function  $f(\cdot)$ 
3: Generate an initial population of wolves at random
4:  $t := 0$ 
5: while  $t \leq T$  do
6:   for each wolf  $\mathbf{x}_i^t$  in population do
7:     Prey initiatively using Eq. (17)
8:     Generate new position  $\mathbf{x}_i^{t+1}$  according to Eq. (15)
9:     if  $d(\mathbf{x}_i^t, \mathbf{x}_i^{t+1}) < r \wedge f(\mathbf{x}_i^{t+1}) > f(\mathbf{x}_i^t)$  then
10:      Stay in current position
11:     end if
12:     Find a new position randomly
13:     Generate random number  $\beta \in [0, 1]$ 
14:     if  $\beta > p_\alpha$  then
15:       Escape from the current position using Eq. (18)
16:     end if
17:   end for
18:   if  $t \neq T$  then
19:     Take 20% of the best adapted wolves to the next
     iteration
20:   end if
21:    $t ++$ ,
22: end while
23: Return the best wolf as a solution
24: Stop

```

## 4. Feature Extraction

Our method for feature extraction from signal sample is based on idea of extraction information from images. In order to present the signal in graphical form, we use scaleograms and visualization of sound sample in frequency-time domain. Created images become search space in which individuals of a given population in heuristic algorithm move and search for important information.

### 4.1 Signal Visualization

One way to visualize wavelet transform for a given signal is scaleogram, where wavelet transform is dis-

played using energy density. Scaleogram is built on two axes: time representing (OX axis) and scale (axis OY). Each value on such spread-axis corresponds to wavelet coefficients, which are illustrated by different shades of brightness. Another form of visual representation can be done using signal frequency and time domain. Example of scaleogram and visualization of wavelet are illustrated in Fig. 5–7.

Figure 5

Representation of the sentence “Han Solo” by using the scaleogram

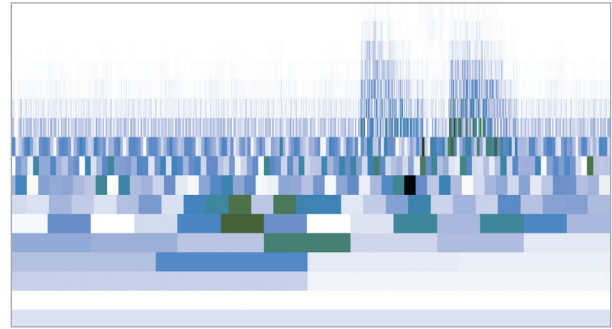


Figure 6

Linear representation of the sentence “Han Solo” in frequency-time domain

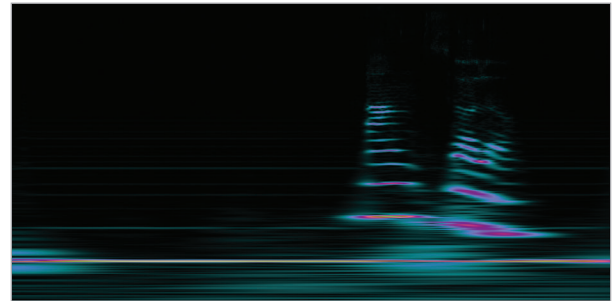
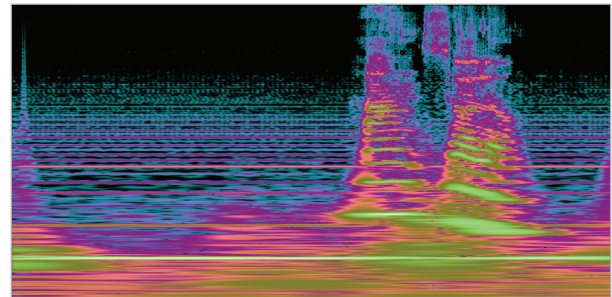


Figure 7

Logarithmic representation of the sentence “Han Solo” in frequency-time domain



### 4.2 Extraction of Important Information

Visualization of signal allows to use it as search space for applied heuristic method. Each individual will correspond to coordinates of points represented by pixel. In algorithm we will look for most valuable points (pixels) of scalogram image. In this way, algorithm returns specific set of points, which correspond to contained information (or features) in given sound sample.

In case of scaleogram, the most important information are dark areas, so we use heuristic to search for pixel i.e. using brightness feature. This allows us to define fitness function

$$f_1(\mathbf{x}) = \frac{\delta(\mathbf{x}) + \eta(\mathbf{x})}{2}, \tag{20}$$

where  $\delta(\mathbf{x}) = \max(R(\mathbf{x}), G(\mathbf{x}), B(\mathbf{x}))$  and  $\eta(\mathbf{x}) = \min(R(\mathbf{x}), G(\mathbf{x}), B(\mathbf{x}))$  wherein  $R(\cdot)$ ,  $G(\cdot)$  and  $B(\cdot)$  are related to color components (red, green and blue) of a given pixel. Defined in this way function refers to HSL model from which other properties may be used, like hue defined as

$$f_2(\mathbf{x}) = \begin{cases} 60^\circ \left[ \frac{G(\mathbf{x}) - B(\mathbf{x})}{\delta(\mathbf{x}) - \eta(\mathbf{x})} \pmod{6} \right] & \text{if } \delta(\mathbf{x}) = R(\mathbf{x}) \\ 60^\circ \left[ \frac{B(\mathbf{x}) - R(\mathbf{x})}{\delta(\mathbf{x}) - \eta(\mathbf{x})} + 2 \right] & \text{if } \delta(\mathbf{x}) = G(\mathbf{x}). \\ 60^\circ \left[ \frac{R(\mathbf{x}) - G(\mathbf{x})}{\delta(\mathbf{x}) - \eta(\mathbf{x})} + 4 \right] & \text{if } \delta(\mathbf{x}) = B(\mathbf{x}) \end{cases} \tag{21}$$

The above function characterize property which is perceived as one of basic color components in RGB model. Another function for saturation can be defined as

$$f_3(\mathbf{x}) = \frac{\delta(\mathbf{x}) - \eta(\mathbf{x})}{1 - |\delta(\mathbf{x}) + \eta(\mathbf{x}) - 1|}. \tag{22}$$

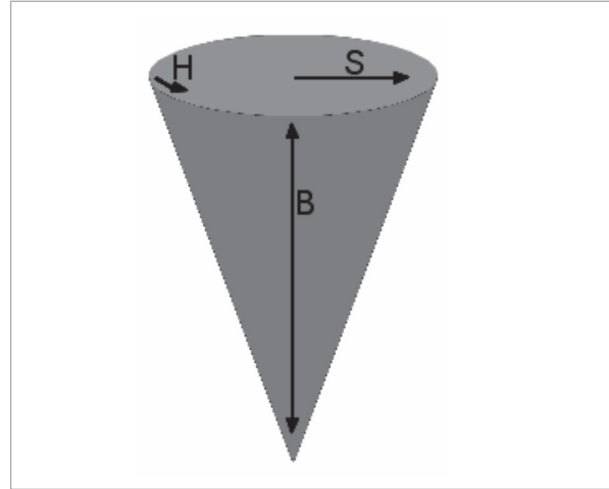
All of these three functions are used in formal description of HSL model, which is in a form of cone representing colors as compositions of three components: hue, saturation and brightness (also called lightness). The cone representation is shown in Fig. 3.

Defined fitness function can be used to extract information from image. Fitness function is a flexible element of proposed approach, so we can modify it to various initial conditions. Found points  $x_i$  in image can form a feature vector as follows

$$[f_i(x_1), f_i(x_2), \dots, f_i(x_n)]. \tag{23}$$

Figure 3

Cone, which represents color space to build the model for application in decision support system



This type of vector will represent specific values in terms to selected fitness function  $f_i$ . Number of elements depends on selection of individuals in heuristic algorithm. Proposed heuristic returns only  $n$  best individuals. The greater number of features, the higher should be number of individuals in predetermined population.

## 5. Speaker Recognition as Testing Tool

Proposed solution for feature extraction from graphical representation of audio signal is efficient for voice verification or recognition. In this section, we present two examples of verification methods that can benefit from proposed approach to feature extraction.

### 5.1 Template Matching

One of the easiest and most popular methods of comparison of two vectors is their pattern matching. Pattern will be called vector formed by  $k$  samples, where each value is averaged in suitable way and some errors of measurement are specified. Pattern vector can be determined in various way, for example, arithmetic mean of  $k$  samples

$$\left[ \sum_{j=1}^k \frac{1}{k} f_i^j(x_1), \sum_{j=1}^k \frac{1}{k} f_i^j(x_2), \dots, \sum_{j=1}^k \frac{1}{k} f_i^j(x_n) \right], \tag{24}$$

where  $f_i^j$  is value  $f(x_i)$  for  $j$ -th sample. For this vector, comparison must be evaluated relative to error, because it is unlikely to achieve identical results for two different samples. For this purpose, we propose to construct error function  $\delta(\cdot)$  dependent on all variables in vectors as follows

$$\delta(\vec{v}) = \sum_{j=1}^n \|v_j - v_j^*\|, \quad (25)$$

where  $\vec{v} = [f_i(x_1), f_i(x_2), \dots, f_i(x_n)]$  and  $v_j$  indicate  $j$ -th element in this vector. As  $v_j^*$ , we denote pattern vector described in Eq. (22) and  $\|\cdot\|$  is function which can be described in any way satisfying condition  $\|\cdot\|: R \rightarrow [0, \infty)$ . In this discussion, we assume that this function is described in the following way

$$\|x_j\| = \frac{|x_j|}{n^2}. \quad (26)$$

Comparison of two vectors will consist of two stages. In the first one, error value is calculated by Eq. (25) and in next one, these value is verified within range of acceptance, which will be adjusted based on value of this vector and error function for a particular problem.

Another example is creation of two vectors, which contain the most diverging from each other

$$\left\{ \begin{array}{l} \left[ \min_{j=1, \dots, k} [f_i^j(x_1)], \min_{j=1, \dots, k} [f_i^j(x_2)], \dots, \min_{j=1, \dots, k} [f_i^j(x_n)] \right] \\ \left[ \max_{j=1, \dots, k} [f_i^j(x_1)], \max_{j=1, \dots, k} [f_i^j(x_2)], \dots, \max_{j=1, \dots, k} [f_i^j(x_n)] \right] \end{array} \right\} \quad (27)$$

This approach is indicated that if the value is not between minimum and maximum of variables, the vector

is discarded. In the case of too many restrictions on acceptance, both vectors can be extended to a certain percentage of error adjusted in relation to given problem.

## 6. Experiments

Measurement accuracy and effectiveness of proposed method was analyzed using 200 audio samples with the sentence "Han Solo". The tests were performed for both presented methods of speaker recognition. Experiments were performed for population of 200 and 300 individuals, where only the top 15% of them have been returned. Number of selected individuals forced length of vectors as 30 and 45. As fitness function three variants were selected described by Eq. (20)–(22).

For these parameters, all samples were processed by each method. Thereafter, each sample was tested for these patterns and specific fitness function. Results of measurements are shown in Tab. 1 and Tab. 2. For first examined method with error measures, accuracy for almost all fitness functions becomes worse with increasing amount of individuals. Accuracy for second method remains at almost constant level, about 83% for each of fitness functions.

Randomness of heuristic algorithm and sound samples influence operation of methods so we can not predict constant accuracy for other functions. Of course tests were carried out on a small number of samples with short recording time so accuracy of these results is very high.

Average accuracy of methods in term of applied fitness function is shown in Fig. 4. For used functions the best

**Table 1**

The obtained accuracy of the described feature extraction method in the speaker recognition application for population consisting of 200 individuals

Method	Fitness function	Population	Number of the best ones	Classified samples		Accuracy
				Correctly	Incorrectly	
Measurement error with respect to Eq. (26)	Eq. (20)	200	30	163	37	81,5%
	Eq. (21)	200	30	179	21	89,5%
	Eq. (22)	200	30	145	55	72,5%
Interval method with respect to Eq. (27)	Eq. (20)	200	30	176	24	88%
	Eq. (21)	200	30	165	35	82,5%
	Eq. (22)	200	30	147	53	73,5%

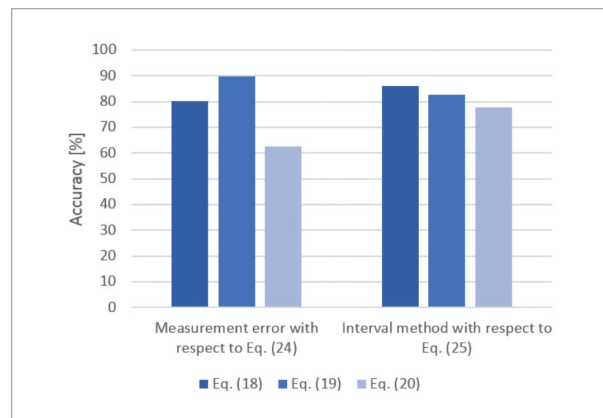
**Table 2**

The obtained accuracy of the described feature extraction method in the speaker recognition application for population consisting of 300 individuals

Method	Fitness function	Population	Number of the best ones	Classified samples		Accuracy
				Correctly	Incorrectly	
Measurement error with respect to Eq. (26)	Eq. (20)	300	45	158	42	79%
	Eq. (21)	300	45	180	20	90%
	Eq. (22)	300	45	105	95	52,5%
Interval method with respect to Eq. (27)	Eq. (20)	300	45	168	32	84%
	Eq. (21)	300	45	166	34	83%
	Eq. (22)	300	45	164	36	82%

**Figure 4**

Average accuracy of detected features during verification by application of various fitness functions



results were obtained for Eq. (21) and the worst for Eq. (22). The reason for differences is primarily a function, and more specifically adaptation to search for specific features. For graphical representation of audio signal, specific areas of images were searched and selection of most precise fitness function improves results.

## 7. Conclusions

Presented solution for feature extraction from audio signal based on image processing and heuristic method proved to be a good solution for that purpose. We have tested our proposition in terms of speaker recognition where decision is made on basis of voice sam-

ples. The use of heuristic algorithm allows to detect specific, unique (because of randomness of population) sequence of positions in images what can be used in real-time recognition, which could find application in many practical purposes. Especially in large corporations where access is granted under certain identity verification conditions.

In corporations we can use a sample mechanism, to collect samples from workers, which will be stored in a database system from which trained architectures can extract samples for verification. A user coming to work can be verified by ad-hoc recording of the voice, which will be forwarded to the verification system. Result is returned to entrance, where user gain access or is denied from entering.

Proposed solution depends on fitness function what allows for feature extraction not only of audio signals but also all kinds of signals, but also many other objects like 2D graphics. Interesting approach may be the usage of other transforms as well as fitness functions. In future research we would like to examine the impacts of application of various heuristic algorithms. Additionally, use of similar solutions to other classifiers could be attractive in terms of accuracy.

## Acknowledgments

Authors acknowledge contribution to this project from the "Diamond Grant 2016" No. 0080/DIA/2016/45 funded by the Polish Ministry of Science and Higher Education.



## References

1. Besacier, L., Bernard, E., Karpov, A., Schultz, T. Automatic Speech Recognition for Under-Resourced Languages: A Survey. *Speech Communication*, 2014, 85-100. <https://doi.org/10.1016/j.specom.2013.07.008>
2. Brociek, R., Słota, D. Application and Comparison of Intelligent Algorithms to Solve the Fractional Heat Conduction Inverse Problem. *Information Technology and Control*, 2016, 45(2), 184-194. <https://doi.org/10.5755/j01.itc.45.2.13716>
3. Costa, Y. M., Oliveira, L. S., Carlos, C. N. An Evaluation of Convolutional Neural Networks for Music Classification Using Spectrograms. *Applied Soft Computing*, 2017, 52, 28-38. <https://doi.org/10.1016/j.asoc.2016.12.024>
4. Chodarev, S., Kollar, J. Extensible Host Language for Domain-Specific Languages. *Computing and Informatics*, 2016, 35(1), 84-110.
5. Damaševičius, R., Maskeliūnas, R., Venčkauskas, A., Woźniak, M. Smartphone User Identity Verification Using Gait Characteristics. *Symmetry*, 2016, 8(10), 100. <https://doi.org/10.3390/sym8100100>
6. Gabryel, M. The Bag-of-Features Algorithm for Practical Applications Using the MySQL Database. *Lecture Notes in Computer Science*, 2016, 9693, 635-646. [https://doi.org/10.1007/978-3-319-39384-1\\_56](https://doi.org/10.1007/978-3-319-39384-1_56)
7. Gabryel, M. A Bag-of-Features Algorithm for Applications Using a NoSQL Database. *Communications in Computer and Information Science*, 2016, 639, 332-343. [https://doi.org/10.1007/978-3-319-46254-7\\_26](https://doi.org/10.1007/978-3-319-46254-7_26)
8. Griffiths, K. R., Hicks, B. J., Keogh, P. S., Shires, D. Wavelet Analysis to Decompose a Vibration Simulation Signal to Improve Pre-distribution Testing of Packaging. *Mechanical Systems and Signal processing*, 2016, 76, 780-795. <https://doi.org/10.1016/j.ymsp.2015.12.035>
9. Haar, A. Zur Theorie der Orthogonalen Funktionensysteme. *Mathematische Annalen*, 69(3), 1910, 331-371. <https://doi.org/10.1007/BF01456326>
10. Kameoka, H. Non-negative Matrix Factorization and Its Variants for Audio Signal Processing. *Applied Matrix and Tensor Variate Data Analysis*. Springer Japan, 2016, 23-50. [https://doi.org/10.1007/978-4-431-55387-8\\_2](https://doi.org/10.1007/978-4-431-55387-8_2)
11. Korytkowski, M., Rutkowski, L. Fast Image Classification by Boosting Fuzzy Classifiers. *Information Sciences*, 2016, 327, 175-182. <https://doi.org/10.1016/j.ins.2015.08.030>
12. Martisius, I., Damasevicius, R. A Prototype {SSVEP} Based Real Time {BCI} Gaming System. *Computational Intelligence and Neuroscience*, 2016, 639, 3861425:1-3861425:15. <https://doi.org/10.1155/2016/3861425>
13. Nowicki, R. K., Scherer, R., Rutkowski, L. Novel Rough Neural Network for Classification with Missing Data. *21st International Conference on Methods and Models in Automation and Robotics (MMAR)*, 2016, 820-825. <https://doi.org/10.1109/MMAR.2016.7575243>
14. Porubaen, J., Bačíková, M., Chodarev, S., Nosal, M. Teaching Pragmatic Model-Driven Software Development. *Computer Science & Information Systems*, 2015, 12(2), 683-705. <https://doi.org/10.2298/CSI-S140107022P>
15. Scherer, R., Rutkowski, L. A Fuzzy Relational System with Linguistic Antecedent Certainty Factors. *International Conference on Neural Networks and Soft Computing*, 2003, 563-569. [https://doi.org/10.1007/978-3-7908-1902-1\\_86](https://doi.org/10.1007/978-3-7908-1902-1_86)
16. Sebastian, J., Kumar, M., Murthy, H. A. An Analysis of the High Resolution Property of Group Delay Function with Applications to Audio Signal Processing. *Speech Communication*, 2016, 81, 42-53. <https://doi.org/10.1016/j.specom.2015.12.008>
17. Słota, D., Brociek, R. Application of Real Ant Colony Optimization Algorithm to Solve Space Fractional Heat Conduction Inverse Problem. *Communications in Computer and Information Science*, 2016, 639, 369-379. [https://doi.org/10.1007/978-3-319-46254-7\\_29](https://doi.org/10.1007/978-3-319-46254-7_29)
18. Sulír, M., Nosál, M., Porubän, J. Recording Concerns in Source Code Using Annotations. *Computer Languages, Systems & Structures*, 2016, 46, 44-65. <https://doi.org/10.1016/j.cl.2016.07.003>
19. Tang, R., Fong, S., Yang, X. S., Deb, S. Wolf Search Algorithm with Ephemeral Memory. *7th International Conference on Digital Information Management*, 2012, 165-172. <https://doi.org/10.1109/ICDIM.2012.6360147>
20. Terada, H., Makino, K., Nishizaki, H., Yanase, E., Suzuki, T., Tanzawa, T. Positioning Control of a Micro Manipulation Robot Based on Voice Command Recognition for the Microscopic Cell Operation. *Advances in Mechanism Design II*, Springer International Publishing, 2017, 73-79.
21. Ubhayarantne, L., Pereira, M. P., Xiang, Y., Rolfe, B. F. Audio Signal Analysis for Tool Wear Monitoring in Sheer Metal Stamping. *Mechanical Systems and Signal Processing*, 2017, 85, 809-826. <https://doi.org/10.1016/j.ymsp.2016.09.014>

---

## Summary / Santrauka

The use of algorithms is helpful in analysis of various data samples. Examples of these applications are sound and graphic processing that are used in authentication and analysis of images. In this work, we propose technique of extracting data from image file created based on voice sample. Proposed method makes use of mathematical model of wavelet transformation and its graphical visualization like scaleogram combined with computational intelligence methods like neural network and heuristic algorithm. In order to verify operation of proposed technique we present experimental results.

---

Algoritmų panaudojimas yra naudingas įvairių duomenų pavyzdžių analizėje. Tokio panaudojimo pavyzdžiai yra garsinis ir grafinis apdorojimas, naudojamas nuotraukų autentifikacijoje ir analizėje. Straipsnyje autoriai siūlo duomenų išgavimo iš nuotraukos failo, sukurto pagal balso pavyzdį, techniką. Siūlomas metodas panaudoja matematinį bangelės transformacijos modelį ir tokią jo grafinę vizualizaciją, kaip skaleograma, apjungtą su kompiuterinio intelekto metodais, pavyzdžiui, neuroniniu tinklu ir euristiniu algoritmu. Straipsnyje pateikiami ir eksperimentiniai rezultatai, parodantys siūlomos technikos veikimą.