Advances in Science and Technology
Research Journal
Volume 10, No. 32, Dec. 2016, pages 73–81
DOI: 10.12913/22998624/65138

Research Article

SPECTRAL METHODS IN POLISH EMOTIONAL SPEECH RECOGNITION

Paweł Powroźnik[1], Dariusz Czerwiński[1]

[1] Institute of Computer Sciences, Lublin University of Technology, Nadbystrzycka 36, 20-618 Lublin, Poland, e-mail: pawel.powroznik@pollub.edu.pl, d.czerwinski@pollub.pl

Received: 2016.06.17
Accepted: 2016.09.08
Published: 2016.12.01

ABSTRACT
In this article the issue of emotion recognition based on Polish emotional speech signal analysis was presented. The Polish database of emotional speech, prepared and shared by the Medical Electronics Division of the Lodz University of Technology, has been used for research. Speech signal has been processed by Artificial Neural Networks (ANN). The inputs for ANN were information obtained from signal spectrogram. Researches were conducted for three different spectrogram divisions. The ANN consists of four layers but the number of neurons in each layer depends of spectrogram division. Conducted researches focused on six emotional states: a neutral state, sadness, joy, anger, fear and boredom. The averange effectiveness of emotions recognition was about 80%.

Keywords: artificial neural network, spectrogram, emotional speech recognition.

INTRODUCTION

For humans, speech is the main tool for communication. Factors like age, language, emotions, gender of speaker and many others can influence the features of speech [14]. The above mentioned factors give additional information for listener, but through emotion some specific value can be added and communicated. Obviously, information which is conveyed by voice intonation have more then only textual meaning. The same sentences pronounced with different emotions can have completely different meaning [15].

It is well-known that a sentence spoken without any emotion cannot transfer extra information to the listener, but for systems constructed for automatic speech recognition it is a dream situation. In other words, emotional states caused essential changes in speech parameters, which deteriorate the accuracy of speech recognitions systems [7].

The largest problem for emotional speech recognition applications is the number of different emotional states. It is not trivial to construct model which will focused on all of emotions so usually researches consider states such as: joy, sadness, boredom, fear, anger [4, 7].

In this article the usage of spectrograms in Polish emotional speech recognition will be shown. Authors focused on above mentioned emotional states adding a neutral state – state emotionally unmarked [3, 7, 14]. In the researches Polish Emotional Speech Database was used. This base was prepared by Lodz University of Technology. All database contains 240 records prepared by professional actors (4 men and 4 woman). Each speaker pronounced 5 different sentences in 6 mentioned above emotional states [7].

The subject of presented research was to find if voice spectral analysis connected with artificial neural networks is enough to effectively recognize the speaker emotional state. The second objective was to determine the optimal input parameters and the whole structure of used artificial neural networks.

DESCRIPTION OF DATABASE

In researches connected with emotional speech analysis the Berlin Database of Emo-

tional speech is commonly used. Abovementioned database contains recordings in seven emotional states: fear, anger, boredom, joy, sadness, disgust and neutral state. Recordings were prepared by ten professional actors of both sexes [1]. If Polish emotional speech is considered, researchers rather used database prepared by Medical Electronics Division of the Lodz University of Technology. This base, the same as Berlin database, was prepared by eight professional actors: four women and four men. Collected files were recorded in six emotional states, that is: joy, anger, boredom, fear, sadness and neutral state [8]. The whole database contains 240 records saved in the 'wav' format sampled with 44.1 kHz frequency and the bit rate of 16 bps. This database includes the following statements: 'I stop shaving from today on', 'Johnny was at the hairdresser's' today, 'They have bought a new car today', 'This lamp is on the desk today' and 'His girlfriend is coming here by plane'.

VOICE SPECTRAL ANALYSIS

The most commonly used tools in speech signal processing are methods connected with time and frequency. The sets of time-frequency methods are large but can be divided into two main groups: time – frequency representations and time – scale representations [10]. Those methods can be interpreted as short – time frequency analysis, because they allow to estimate speech signal in a finite time intervals. This estimation is carried out based on signal's fragments cut by window function [18].

The voice spectral analyzes of Polish emotional speech is the main issue of the following research. In general spectrogram there is a visual representation of signal amplitude spectrum for each time, when the signal is determined. It is constructed by dividing the signal into specific parts. For each part the amplitude of harmonic components are counted. The frequency and time are arguments for spectrogram [6].

SHORT – TIME FOURIER TRANSFORM (STFT)

STTF fulfil the main role in speech signal analysis as well as spectrograms. These two methods can be included to time – frequency representation group [9]. STFT can be treated as special case of Gabor transformation [18]. Its definition for continuous signal *x(t)* has a form in frequency domain as follows [10]:

$$STFT_x^F(t,f) = e^{-j2\Pi 2\Pi}\int_{-\infty}^{+\infty} X(\theta)W*(\theta - f)e^{j2\Pi 2\Pi}d\theta \tag{1}$$

and in time domain:

$$STFT_x^T(t,f) = \int_{-\infty}^{+\infty} x(\tau)w*(\tau - t)e^{-j2\Pi 2\Pi}d\tau \tag{2}$$

where: *w(t)* is window function of the Fourier spectrum *W(f)* and *X(f)* is spectrum of the analyzed signal.

For digital speech signal analysis particular significance has discrete form of the first equation [10]:

$$STFT(n,k) = \sum_{m=-\infty}^{+\infty} x(m)w*(n-m)e^{-j\left(\frac{2\Pi}{N}k\right)m} \tag{3}$$

For window function of N real, non-zero samples above equation has the form as follows [10]:

$$STFT(n,k) = \sum_{m=0}^{N} w(m)x(n-m)e^{-j\left(\frac{2\Pi}{N}k\right)m} \tag{4}$$

where: $n = 0, N, 2N, \ldots, M\text{-}N$; $k = 0, 1, 2, \ldots N\text{-}1$; M – the number of analyzed samples.

All calculation are performed based on fourth equation. Parts of speech signal x(n) are consecutively cuts by window function w(n) and for each part Discrete Fourier Transform (DFT) is calculated. Using STFT for discrete signal a spectrogram can be defined as follows[10]:

$$S(n,k) = |STFT(n,k)|^2 \tag{5}$$

The selection of resolution in the time and frequency domain has a main influence on spectrogram quality. Wide window in STFT guarantee high resolution in frequency domain but narrow window increase resolution in time domain. This effect has justification in time – spectrum correlation for window function. To obtain a high time resolution window function w(n) with a small number of elements should be applied. If the number of elements for window function – *N,* is small, the DFT calculation, which is performed for successive frequencies mutually distant for $\Delta f{\cdot}f_p/N$ (f_p *– sampling frequency*), will be carried out with large increase in frequency. An additional disadvantage will be the occurrence of the blur effect in spectrum, which will be caused by a large width of main leaf in amplitude frequency characteristics for time window. For high *N value,* the resolution in frequency axis will increase and the width for main leaf in amplitude frequency characteristics will decrease. The disadvantage is that the calculation for DFT will be performed with big time step $\Delta t=N/f_p$, which will have a negative influence on spectrogram precession [2]. The overlapping method in STFT counting process is used to improve the spectrogram quality [11].

In the Figure 1 the overlapping method is shown. Window function, consists of six samples, cuts parts of the analysed signal. In this example the overlapping value is 50% ($N_o = 3$ samples). It can be easily noticed that overlapping samples number can be $N_o = 1$ to $N_o = N - 1$.

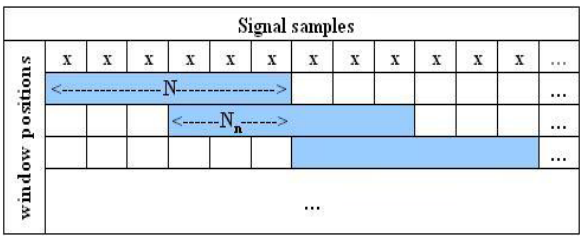

Fig. 1. Example of overlapping process

SELECTION OF SPECTROGRAM'S PARAMETERS

Getting the spectrogram which enables an efficient inference is associated with selection of spectrogram's parameters such as: resolution in time domain, window function or window width. The best resolution in time domain can be achieved by usage the maximum overlapping $N_o = N - 1$. The frequency resolution is directly proportional to number of elements of window function *N* [5, 10]. Applying the maximum overlapping seems to be the most desirable situation but this method is connected with significant increase in computational effort. Selecting the length of window defines the frequency resolution according to the above-mentioned relation: $\Delta f{\cdot}f_p/N$. The appropriate selection of window length (*N*) is more complex [5, 10]. If the signal is modulated the length of time interval should be defined as follows: the quotient of the mean width of the frequency *B* to time *A* should be equal to the quotient of the frequency rate increase to the time at which it occurred [10]:

$$\frac{B}{A} = \frac{\Delta f}{\Delta t} \tag{6}$$

where:

$$B = \sqrt{\frac{1}{E}\int_{-\infty}^{+\infty} f^2 \left|W(f)^2\right| df} \tag{7}$$

is mean square frequency width for window function *w(t)* in Fourier spectrum *W(f)* and:

$$A = \sqrt{\frac{1}{E}\int_{-\infty}^{+\infty} t^2 \left|w(t)^2\right| dt} \tag{8}$$

is mean square time width but [10]:

$$E = \sqrt{\frac{1}{E}\int_{-\infty}^{+\infty} \left|w(t)^2\right| dt} = \sqrt{\frac{1}{E}\int_{-\infty}^{+\infty} \left|W(f)^2\right| df} \tag{9}$$

For discrete signal analysis (4) the parameter *N* is responsible for resolution in frequency domain [10]. Times resolution can be increased by taking maximum overlapping but as it was mentioned above it is connected with significant calculation incensement. In order to obtain high time resolution without overlapping time windows with a small number of elements should be used but it will cause small frequency resolution [16, 17]. In that case the resolution in frequency

domain can be increased by complementing window by sequence of zeros. It means that if window function *w(m)* takes non-zero values for $m=0,1,2, \ldots, N_p - 1$ the window should be complemented by zeros sequence to *N* length. Based on (4) equation the STFT will be as follows [4, 10]:

$$STFT(n,k) = \sum_{m=0}^{N-1} w_u(m)x(n=m)e^{-j\left(\frac{2\Pi}{N}k\right)m} \tag{10}$$

where:

$$n = \sum_{i=1}^{N} x_i w_i + b \tag{11}$$

CONDUCTED RESEARCHES

The main aim of the conducted researches was to identify emotion based on Polish speech signal processing. These researches are focused on the use of spectrogram and artificial neural networks (ANN) to achieve abovementioned goal. As it was mentioned above, speech signals were obtained from database prepared by Medical Electronics Division of the Lodz University of Technology.

Researches were conducted in two ways. The first was regardless of speaker's sex and the second taking into account the above division. The first step in the conducted researches was pre-processing. The values of amplitude of particular samples have undergone the process of normalization and reached values between -1 to 1. The second step was framing. The whole signal was divided into small frames with time length 20 ms. To reduce the discontinuities at the edges of frames the Hamming window was used.

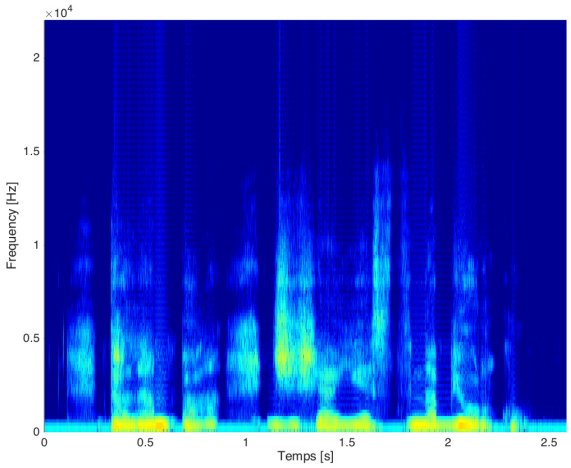

Fig. 2. Frequency and time spectral representation

The next step was to use the Fast Fourier Transform (FFT) for transforming each segment of speech signal to its frequency domain from discrete time domain.

After mentioned above transformation the spectrograms were created in MatLab application. The example of time-frequency signal representation in shown in Figure 2. The spectrograms were created based on Hamming window of 128 size with 50 % overlapping.

FEATURE EXTRACTION

The main goal of feature extraction process was preparation the inputs for ANN. This process was as follows:

- The spectrogram was converted into grey scale.
- Achieved spectrogram was converted into binary. The values below threshold was changed into 0 and values greater than threshold into 1.
- The whole spectrogram was divided into matrix as follows: 3x3, 4x4, 5x5 separately. In Figure 3 the division by matrix 4x4 is shown. The researches were conducted for each of the above-mentioned divisions.
- All the values in each sub areas were summed and became input parameters for ANN.

ARTIFICIAL NEURAL NETWORKS

In this article artificial neural networks have been created in MatLab application. The mathematical formulation on ANN is as follows [13]:

$$n = \sum_{i=1}^{N} x_i w_i + b \tag{12}$$

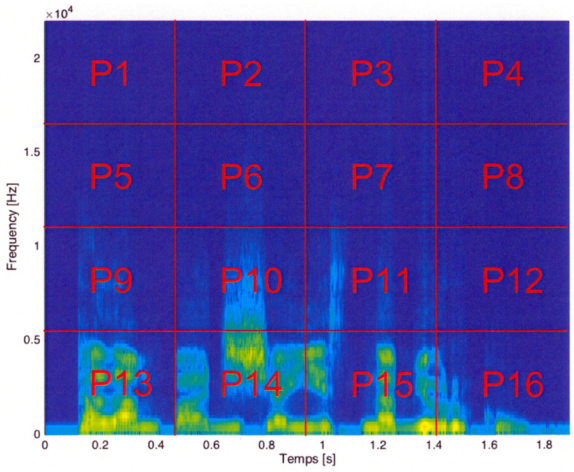

Fig. 3. The example of spectrogram division

Depends on spectrogram division three structures of ANN were used. Considering 3x3 division the neural network had 4 layers. The input layers were constructed with 10 neurons – 9 elements from spectrogram division and the 10th was speaker's sex, if researches were conducted with sex division or bias in other case. The ANN had two hidden layers, 20 neurons each and output layer constructed with 6 neurons corresponding 6 emotional states. The neurons were activated by sigmoidal function and the whole network was taught by backpropagation algorithm.

When the signal spectral representation was divided by 4x4 matrix, the changes, comparing to previously described ANN, was in input layer and hidden layer. All other parameters remain unchanged. In this case the input layer consisted of 17 neurons. The additional neuron was as in the previous case, either speaker gender or bias. The hidden layers had 34 neurons each. The exampled ANN architecture was shown in Figure 4.

The last researches were conducted for 5x5 spectrogram divisions. The input layer was constructed with 26 neurons. The first hidden layer consists of 20 neurons the second of 10 neurons. The output layer and other ANN parameters was the same as in previous cases.

The artificial neural network was trained by setting the network parameters and stopping criteria. As it was mentioned, the back propagation algorithm was used. The sets of desired output and inputs were introduced to the network to learn the data's relationships. Error correction of this data set was generated by using local training approach. The main difference between traditional back propagation algorithm and the used one is usage of semi-supervised teaching technique. The main role of hidden layers was to adjust weights connected with each input nodes. For error calculation the root mean square error was used. If this value was not satisfied, the algorithm propagated error from output layer to input layer. The algorithm was working until the mean square error was not satisfied.

ACHIEVED RESULTS AND DISCUSSION

The database prepared by Medical Electronics Division of the Lodz University of Technology was used in conducted researches. This database contains 240 files recorded in six emotional states: joy, anger, fear, boredom, sadness and neutral state. The experimental results for each spectrogram division is shown in Figures 5-7.

Tables 1÷3 show the confusions matrix of the proposed algorithm. The above-mentioned tables were prepared with gender division and without. Each column represents the instance of an original class, each row represents emotion predicted by ANN.

It can be easily seen that the worse efficient was achieved if women were test group in researches. The average ANN effectiveness

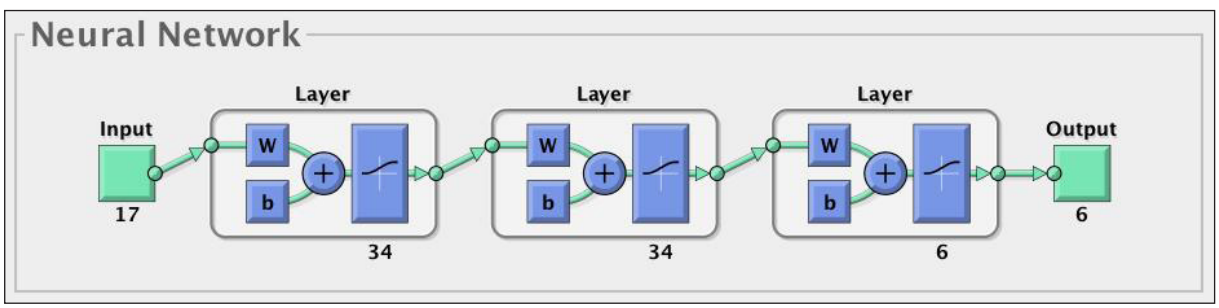

Fig. 4. The example of structure of used ANN

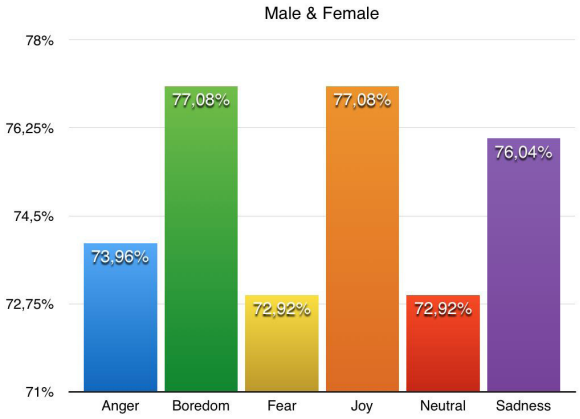

Fig. 5. ANN effectiveness for 3x3 spectrogram division without gender division

Table 1. Confusion matrix for 3x3 spectrogram division without gender division

Recognized emotion	Input emotion					
	Anger	Boredom	Fear	Joy	Neutral	Sadness
Anger	**73,96**	1,04	7,29	9,38	3,13	2,08
Boredom	3,13	**77,08**	3,13	1,04	10,42	8,33
Fear	8,33	2,08	**72,92**	8,33	2,08	2,08
Joy	6,25	2,08	9,38	**77,08**	1,04	2,08
Neutral	6,25	10,42	4,17	2,08	**72,92**	9,38
Sadness	2,08	7,29	3,13	2,08	10,42	**76,04**

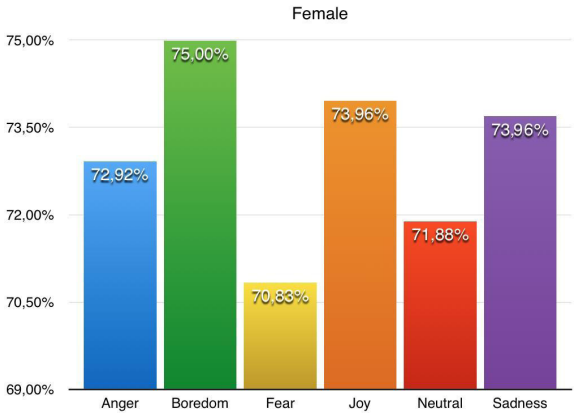

Fig. 6. ANN effectiveness for 3x3 spectrogram division for female

Table 2. Confusion matrix for 3x3 spectrogram division for female

Recognized emotion	Input emotion					
	Anger	Boredom	Fear	Joy	Neutral	Sadness
Anger	**72,92**	2,08	6,25	8,33	4,17	3,13
Boredom	2,08	**75**	4,17	2,08	8,33	8,33
Fear	9,38	2,08	**70,83**	8,33	3,13	2,08
Joy	7,29	2,08	10,42	**73,96**	2,08	2,08
Neutral	5,21	10,42	3,13	4,17	**71,88**	10,42
Sadness	3,13	8,33	5,21	3,13	10,42	**73,96**

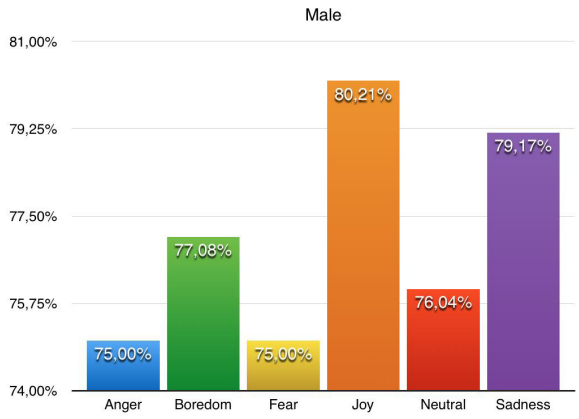

Fig. 7. ANN effectiveness for 3x3 spectrogram division for male

Table 3. Confusion matrix for 3x3 spectrogram division for male

Recognized emotion	Input emotion					
	Anger	Boredom	Fear	Joy	Neutral	Sadness
Anger	**75**	3,13	7,29	6,25	3,13	2,08
Boredom	1,04	**77,08**	2,08	2,08	7,29	6,25
Fear	8,33	5,21	**75**	5,21	2,08	2,08
Joy	8,33	2,08	8,33	**80,21**	1,04	2,08
Neutral	4,17	7,29	4,17	3,13	**76,04**	8,33
Sadness	3,13	5,21	3,13	3,13	10,42	**79,17**

was in this case about 73%. In male test group this effectiveness was 77% and if both groups were considered ANN correctly predict average 75%.

A little bit more effective was the ANN if 4x4 spectrogram division was considered. In this case the best results were achieved. All ANN prediction was shown in Figures 8 to 10. In this case researches were conducted with mentioned above division. In tables 4÷6 the achieved confusions matrixes were shown.

In this case also the worse results were achieved in female tested group. The average ANN effectiveness in this case was about 76%. In male group it was over 80% and if both groups were considered the average result was almost 79%. The results and confusions matrixes for 5x5 spectrogram divisions were shown in Figures 11 to 13 and Tables 7 to 9 respectively.

Also in this case the experiments conducted on female group gave the worse results. ANN effectiveness was about 74%. In male group this value was about 77% and in researches without gender division it was 76%.

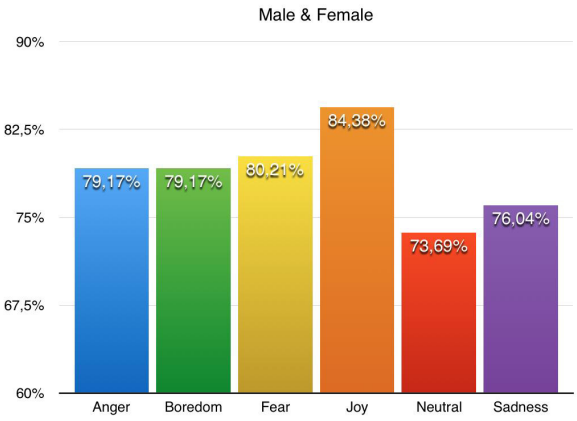

Fig. 8. ANN effectiveness for 4x4 spectrogram division without gender division

Table 4. Confusion matrix for 4x4 spectrogram division without gender division

Recognized emotion	Input emotion					
	Anger	Boredom	Fear	Joy	Neutral	Sadness
Anger	**79,17**	2,08	5,21	7,29	4,17	2,08
Boredom	2,08	**79,17**	3,13	1,04	8,33	8,33
Fear	6,25	1,04	**80,21**	4,17	2,08	2,08
Joy	4,17	2,08	7,29	**84,38**	3,13	2,08
Neutral	5,21	7,29	2,08	2,08	**73,96**	9,38
Sadness	3,13	8,33	2,08	1,04	8,33	**76,04**

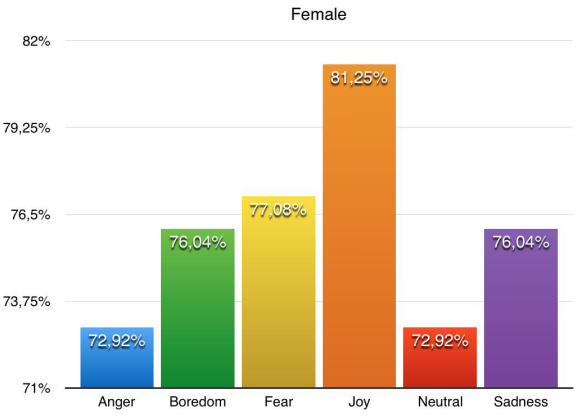

Fig. 9. ANN effectiveness for 4x4 spectrogram division for female

Table 5. Confusion matrix for 4x4 spectrogram division for female

Recognized emotion	Input emotion					
	Anger	Boredom	Fear	Joy	Neutral	Sadness
Anger	**72,92**	2,08	4,17	7,29	2,08	3,13
Boredom	2,08	**76,04**	2,08	1,04	9,38	6,25
Fear	9,38	3,13	**77,08**	5,21	1,04	3,13
Joy	7,29	1,04	8,33	**81,25**	1,04	2,08
Neutral	5,21	8,33	2,08	3,13	**72,92**	9,38
Sadness	3,13	9,38	6,25	2,08	7,29	**76,04**

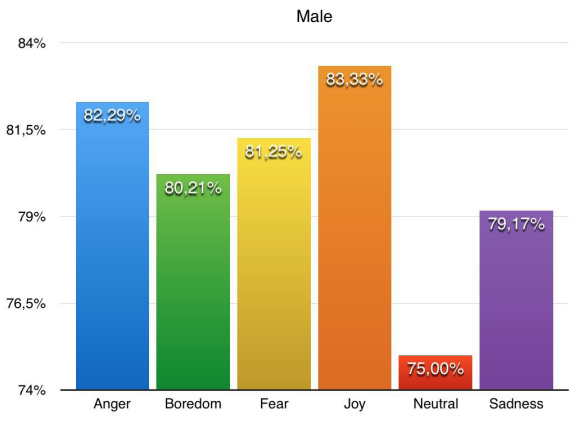

Fig. 10. ANN effectiveness for 4x4 spectrogram division for male

Table 6. Confusion matrix for 4x4 spectrogram division for male

Recognized emotion	Input emotion					
	Anger	Boredom	Fear	Joy	Neutral	Sadness
Anger	**82,29**	1,04	6,25	4,17	4,17	2,08
Boredom	2,08	**80,21**	2,08	4,17	7,29	7,29
Fear	4,17	6,25	**81,25**	4,17	1,04	1,04
Joy	5,21	1,04	6,25	**83,33**	1,04	3,13
Neutral	2,08	8,33	2,08	2,08	**75**	7,29
Sadness	4,17	3,13	2,08	2,08	11,46	**79,17**

In all the researches neutral state was the least likely recognizable emotional state. It was frequently confused with sadness and boredom. If the best results are considered the ANN effectiveness is about 80 % which is a satisfactory result.

From the experimental results it can be confirmed that time – frequency domain spectral representation is efficient tool for visualization of different approach to Polish emotional state identification. The results which were achieved are satisfactory. Moreover, it was shown that ANN are good tool for spectrogram processing.

CONCLUSIONS

The conducted researches presented a novel approach for detecting human emotions based on Polish emotional speech by using voice spectral representation. The important features from spectrogram were extracted and a new architecture for ANN was presented. The structure of ANN allows to reduce the classification process difficulty

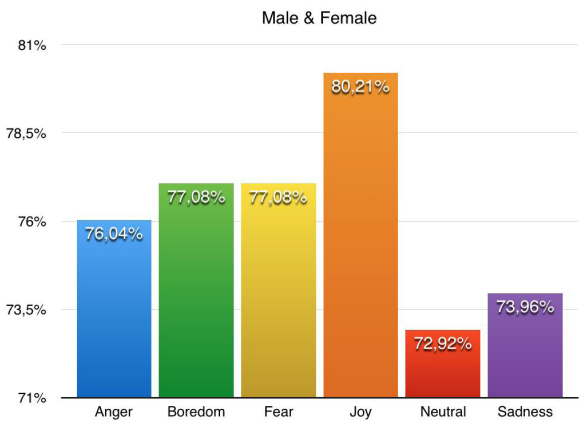

Fig. 11. ANN effectiveness for 5x5 spectrogram division without gender division

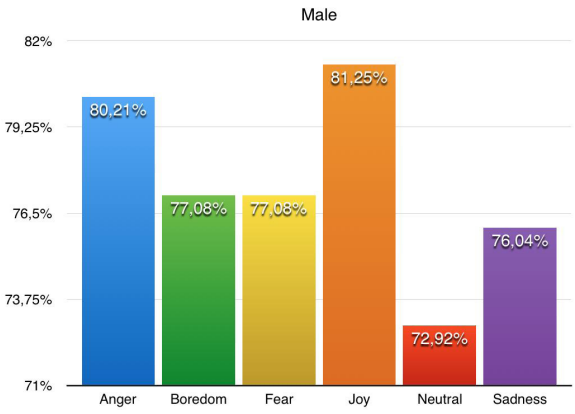

Fig. 13. ANN effectiveness for 5x5 spectrogram division for male

Table 7. Confusion matrix for 5x5 spectrogram division without gender division

Recognized emotion	Input emotion					
	Anger	Boredom	Fear	Joy	Neutral	Sadness
Anger	**76,04**	1,04	6,25	7,29	2,08	3,13
Boredom	2,08	**77,08**	3,13	2,08	9,38	9,38
Fear	10,42	1,04	**77,08**	8,33	1,04	1,04
Joy	5,21	1,04	7,29	**80,21**	2,08	1,04
Neutral	5,21	12,5	3,13	1,04	**72,92**	11,46
Sadness	1,04	7,29	3,13	1,04	12,5	**73,96**

Table 9. Confusion matrix for 5x5 spectrogram division for male

Recognized emotion	Input emotion					
	Anger	Boredom	Fear	Joy	Neutral	Sadness
Anger	**80,21**	2,08	6,35	5,21	3,13	3,13
Boredom	1,04	**77,08**	3,13	1,04	8,33	7,29
Fear	6,25	4,17	**77,08**	4,17	3,13	4,17
Joy	6,25	1,04	4,17	**81,25**	1,04	3,13
Neutral	3,13	9,38	3,13	4,17	**72,92**	6,25
Sadness	3,13	6,25	4,17	4,17	11,46	**76,04**

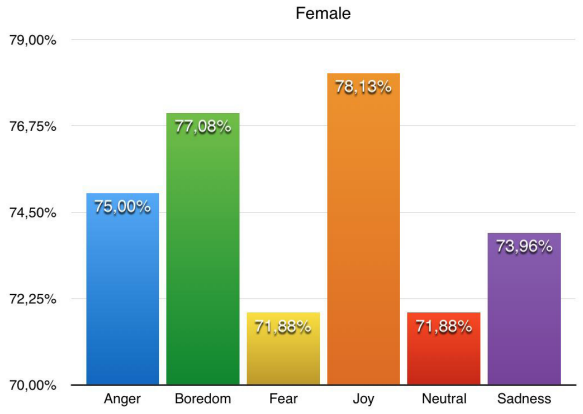

Fig. 12. ANN effectiveness for 5x5 spectrogram division for female

Table 8. Confusion matrix for 5x5 spectrogram division for female

Recognized emotion	Input emotion					
	Anger	Boredom	Fear	Joy	Neutral	Sadness
Anger	**75**	4,17	5,21	5,21	3,13	2,08
Boredom	1,04	**77,08**	5,21	6,25	9,38	10,42
Fear	7,29	1,04	**71,88**	5,21	2,08	1,04
Joy	6,25	1,04	9,38	**78,13**	4,17	1,04
Neutral	6,25	7,29	4,17	3,13	**71,88**	11,46
Sadness	4,17	9,38	4,17	2,08	9,38	**73,96**

and number of features which are inputs for ANN. The new way of features extraction was presented. This approach shows that based on analysis of speech signal time-frequency representation the emotional states can be properly identify.

REFERENCES

1. Berlin Database of Emotional Speech, www.expressive-speech.net.
2. Bracewell R. The fourier transform and its application. McGraw-Hill International Editions, Electric Engineering Series, Singapore, 2000.
3. Strona internetowa Instytutu Eletroniki Politechniki Łódzkiej (www.eletel.p.lodz.pl).
4. Dennis J., Tran H.D. and Chang E.S. Overlapping sound event recognition using local spectrogram features and the generalized hough transform. Pattern Recognition Letters, 34, 2013, 1085–1093.
5. Duan Z., Mysore G.J. and Smaragdis P. Speech enhancement by online non-negative spectrogram decomposition in non-stationary noise enviroments. IEEE Workshop on Application of Signal Processing to Audio and Acoustics, 2013.
6. Hang Q., Wang K. and Ren F. Speech emotion

recognition using combination of features. ICICIP 2013, 523–528.

7. Kamińska D. and Pelikant A. Zastosowanie multimedialnej klasyfikacji w rozpoznawaniu stanów emocjonalnych na podstawie mowy spontanicznej. IAPGOŚ, 03, 2012, 36–39.
8. Kamińska D., Sapiński T., Niewiadomy D. and Pelikant A. Porównanie wydajności współczynników perceptualnych na potrzeby automatycznego rozpoznawania emocji w sygnale mowy. Studia Informatica, 34, 2013, 59–66.
9. Kim E.H., Hyu K.H., Kim S.H. and Kwak Y.K. Speech emotion recognition using eigen-FFT in clean and noisy environments. 16th IEEE International Conference on Robots and Human Interactive Communication, Jeju, Korea, 2007.
10. Konratowski E. Czasowo – częstotliwościowa analiza drgań z wykorzystaniem metody overlapping. Logistyka, 3, 2014, 3104–3110.
11. Konratowski E. Monitoring of the multichannel audio signal, computional collective intelligence. Technologies and Applications, Lecture Notes in Artifical Intelligence, Springer Verlag, 6422, 2010, 298–306.
12. Kozieł G. Zastosowanie transformaty Fouriera w stenografii dźwięku. Studia Informatica, 32, 2011, 542–552.
13. McCulloch W. and Pitts W. A logical calculations of the ideas in nervous activity. Bulletin of Mathematical Biophisics, 5, 1943, 115–133.
14. Panda S.P. and Nayak A.K. Automatic speech segmentation in syllabe centric speech recognition systems. International Journal of Speech Technology, 9, 2016, 9–18.
15. Powroźnik P. Polish emotional speech recognition using artificial neural network. Advances is Science and Technology Research Journal, 8(24), 2014, 24–27.
16. Ramakrishnan S. Recognition of emotion from speech, A review. Speech Enhancement, Modeling and Recognition – Algorithms and Applications, 2012.
17. Szymczyk T. Rozpoznawanie tekstur z wykorzystaniem baz modeli. Prace Instytutu Elektroniki, 249, 2011, 95–115.
18. Zieliński T.P. Cyfrowe przetwarzanie sygnałów. Od teorii do zastosowań, WKiŁ, Warszawa 2009.