

ARTIFICIAL NEURAL NETWORK APPLICATION FOR THE TEMPORAL PROPERTIES OF ACOUSTIC PERCEPTION

**Miloš Simonović¹, Marko Kovandžić¹, Vlastimir Nikolić¹,
Mihajlo Stojčić², Darko Knežević²**

¹Faculty of Mechanical Engineering, University of Niš, Niš, Serbia

²Faculty of Mechanical Engineering, University of Banja Luka, Banja Luka, Republika
Srpska, Bosnia and Herzegovina

Abstract. *Though acoustic perception is well established in literature, it seems to be insufficiently implemented in practice. Experimental results are excellent but a lot of issues arise when it comes to the application in real conditions. Using artificial neural networks makes acoustic signal processing very comfortable from the mathematical point of view. However, a great job has to be done in order to make it possible. The procedure includes data acquisition, filtering, feature extraction and selection. These techniques require much more resources than mere artificial neural networks and this represents a limiting factor for the implementation. The paper investigates the complete procedure of acoustic perception, in terms of time, in order to identify limitations.*

Key Words: *Perception, Temporal Properties, Localization, Filtering, Neural Networks*

1. INTRODUCTION

Acoustic perception is a good alternative to the visual perception in engineering applications, with respect to simplicity, reliability and price. There are a lot of techniques of acoustic observation where each of them assumes specific preconditions to be implemented. In accordance with the preconditions the methods provide limited results. It is necessary to combine several methods of acoustic observation in order to overcome these limitations. Filtering is, for example, an inevitable method of signal processing in real conditions (presence of disturbances) [1]. The function of filter is to suppress the disturbances while the signal is left unchanged. In such a manner, the filter enhances discriminative capacity (amount of useful information) in the signal.

Received April 15, 2019 / Accepted July 28, 2019

Corresponding author: Miloš Simonović

Faculty of Mechanical Engineering, University of Niš, A. Medvedeva 14, 18000 Niš, Serbia

E-mail: milos.simonovic@masfak.ni.ac.rs

In order to act against acoustic object, it is necessary to identify it. The procedure is called acoustic recognition because it assumes that what is currently being received corresponds in some way to something that has already been processed in the past [2]. The problem belongs to the general category of artificial intelligence problems, namely pattern recognition. In the experiment, artificial neural networks were chosen, as a proven pattern recognition tool, for processing acoustic signals [3]. Though neural networks can be trained to perform filtering [4], preprocessing, before the signal is presented to a neural network, is inevitable. Except filtering, the procedure includes normalization (scaling) [5], feature extraction and feature selection [6]. The last two of them decrease computational complexity by combining several features in a new one, with a higher discriminative capacity.

Determination of object position, relative to some reference frame, based on acoustic signals is called acoustic localization. The procedure is implemented, in various forms, in science and practice [7-12]. From the mathematical point of view, acoustic localization belongs to two categories. Near field acoustic localization is performed if the sound source is in the surroundings of the microphones and far field localization when the source is far away from them. The second is much simpler (computationally) but it gives less information, only the direction of sound source. The first method provides spatial coordinates of the sound source using time difference of arrival (TDOA) between microphones [13] as a reference. Analytical solution requires solving system of hyperbolic equations, which is not a trivial problem even for a minimal configuration [14]. In the experiment, the problem is solved by processing TDOAs using feed-forward neural network. Again, the signals were previously processed using different techniques including obtaining TDOAs [15, 16].

2. ACOUSTIC PERCEPTION

The experiment investigates two separate elementary processes of acoustic perception, namely acoustic recognition and acoustic localization, separately. But, in both, artificial neural network is employed as a signal processing tool because of its simplicity, universality and excellent performance. Neural networks are built of simple processing elements, called artificial neurons, which are inspired by biological nervous cells. Neurons are spatially organized in layers and different layers may perform different transformation on input signals. The neurons in a layer are not interconnected and, depending on the way they are connected between layers, there are several network topologies. Neural networks have ability to establish complex nonlinear relationship between input and output variables, by adjusting weights between neurons. The adjustment is done not by an explicit programming but through an adaptive method called learning algorithm, using training data set as a reference. Thanks to the massive parallelism in data processing, neural networks have excellent speed of performance, in the phase of exploitation. The use of neural networks spreads in modern engineering permitting us to investigate a wide range of problems [17], but it also demonstrates a superior accuracy with respect to alternative methods as evident in [18].

The most important temporal characteristic of the artificial neural network is temporal resolution. It is a measure of precision with respect to the operating time. Temporal resolution is limited by computational power available and number of calculations. The transfer of signals from the layer with i to the layer with j neurons is described by [19]

$$Y_j = W_{ij} X_i, \quad X_j = f(Y_j) \quad (1)$$

According to the matrix algebra, the computational complexity of the first equation is $O(i \times j)$ and computational complexity of the second is $O(j)$, as element wise function. The total computational complexity of the transition between two layers is $O(i \times j)$.

Both experiments of acoustic perception deal with corrupted signals. In order to cancel the negative effect of disturbances, a digital IIR filter of second order was employed [13]. The transfer function of the filter in general is

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 - a_1 z^{-1} - a_2 z^{-2}} \quad (2)$$

Computational complexity of the digital filter is equal to its order. On the whole length of the signal it is $O(N \times r)$ or simply $O(N)$, because $r \ll N$, where N is the signal length and r is filter order. The experiment searches for an efficient procedure of filter design with the goal of improving the accuracy of acoustic perception.

Duration of the training phase has no influence on the neural network performance, in the phase of exploitation, but it still requires some resources. Most of the literature about the neural networks deals with computing problems but when it comes to the practical implementation the crucial phase of training is acquiring training samples. Without accurate and properly sampled data, it is not possible to perform a correct training procedure. There is no general rule for data acquisition and feature selection because any implementation requires specific solutions and innovative approach. The experiment searches for efficient methods of acquiring data for the purpose of acoustic perception. The solutions were limited to the regular, simple and cheap equipment.

2.1. Acoustic recognition

Categorization (taxonomy) is essential for acoustic recognition, as for all cognitive processes. It provides generalization rules that are used for making decisions [20]. Categories (classes) are groups of objects that have similar characteristics in some frame of reference. A categorization, or division of objects into classes, enables the observer to make predictions of unobserved characteristic of the object based on previous experience. The process, where general rules are derived from specific examples, is called abstraction. Taxonomy of the phenomena is never unambiguous. Except from the object properties, it depends on the observer properties and his experience.

Several clues of a sound can be used for acoustic categorization. In the living world, most of the acoustic sensations are attributed by pitch, timbre, loudness and duration. Engineering practice suggests different perceptual qualities of sound for different applications. For instance, temporal characteristics, like variance, zero-crossing rate and silence ratio, in combination with spectral properties, like harmonic ratio and sub-band energy, are used for discrimination between speech and music. The most important perceptual property of sound is surely frequency spectrum (envelope) and its derivatives (power spectral density, spectral centroid, spectral irregularity, odd and even harmonics). It is perceived as timbre or tone color by living beings. The easiest way to obtain frequency spectrum of the sample, $s[k]$ is to process it using fast Fourier transform [21]

$$S_k = \sum_{n=0}^{N-1} s_n e^{(-2ikn\pi/N)}, \quad 0 < k < N-1 \quad (3)$$

The first half of the fast Fourier transform, for the case of real valued signals, is conjugate complex of the second half. That is why one of the halves can be neglected without loss of information. Eq. (3) is standard tool for digital signal processing because it provides frequency response with a lower computational complexity, $O(N \log N)$, in comparison with the standard Fourier transform, $O(N^2)$. At the same time, the procedure provides satisfactory result. The obtained frequency spectrum was presented to a feed-forward neural network for pattern recognition.

2.2. Near field acoustic localization

People and animals are able to point to the horizontal direction that the sound is coming from using slightly different signals that arrive at each ear [13]. For the vertical direction, spectrum features, produced by a sound reflector (pinna), is used as the auditory cue. Artificial devices perform localization based on different acoustic clues. The working principle is strongly dependent on the number of microphones. Monoaural (by one microphone) localization, is performed based on energy drop or spectrum deformation, as the sound propagates through the medium. The localization ability of binaural systems can be established by the learning procedure through the repetition of movement. An alternative is employing head related transfer function (HRT), which captures transformations of a sound wave propagating from the source to the microphone. But the most frequent and most valuable clue for acoustic localization is a lag between the signals collected at different spatial positions (TDOA).

The basic approach of estimating TDOA is using cross-correlation function [13]

$$R_{ij}(l) = \frac{1}{N} \sum_{k=0}^{N-1} s_i[k] s_j[k-l] \quad (4)$$

as argument l that maximizes its value within the range of possible lags

$$\Delta t_{ij} = \frac{1}{f_s} \operatorname{argmax}(R_{ij}[l]), \quad -\frac{l_{\max}}{2} \leq l \leq \frac{l_{\max}}{2} \quad (5)$$

where N is the signal length and is the range of expected lags. It has to be chosen in accordance with the measuring range of the experimental setup. Since the expression in the bracket has N^2 multiplications, $N-1$ additions and one division, its computational complexity is $O(N^2)$. The sum has to be evaluated $T+1$ times plus searching for the maximum in the range of possible lags. The final estimation of computational complexity of cross-correlation procedure with maximum allowed lag is $O(N^2 \times T)$. A good approximation of cross-correlation function can be obtained using inverse discrete Fourier transformation

$$R_{ij}(l) \approx \frac{1}{K} \sum_{k=0}^{K-1} \psi(f) R_{ij}(f) e^{\frac{j2\pi fl}{K}} \quad (6)$$

where $R_{ij}(f)$ is cross-power spectral density (XPSD)

$$R_{ij}(f) = S_i(f) \bar{S}_j(f) \quad (7)$$

Since frequency spectrum, $S(f)$ is twice shorter than original signal, $s[k]$, and the computational complexity of cross correlation procedure, $O(N^2 \times T)$ Eq. (6) can save a lot of computational effort. To be implemented it has to provide acceptable approximation error.

Function $\psi(f)$ is called windowing function; its role is to highlight some of the spectrum features in order to improve its discriminativity. Different windowing functions (Table 1) are intended for different purposes but the choice among them is still ambiguous.

Table 1 Windowing functions

Window	Window Function
Cross correlation	$\psi_{CC} = 1$
Roth window	$\psi_{Roth} = 1/S_{ii}(f)$
PHAT window	$\psi_{PHAT} = 1/ S_{ij}(f) $
SCOT window	$\psi_{SCOT} = 1/\sqrt{S_{ii}(f)S_{jj}(f)}$

3. EXPERIMENTAL SETUP

The experiment investigated two separate experiments. In both cases, the acoustic signals were processed using regular processor Intel (R) Celeron (R) CPU N3350 @ 1.10 GHz.

3.1. Acoustic recognition

For the training of the feed-forward neural network in acoustic recognition 500 sound samples were recorded and collected from the Internet. The samples were processed by a human listener, using specially designed software with the possibilities of playing, visualizing and separating parts of interest from the rest of the content. At the end, each of the samples, 1 s long at the frequency of 44.1 kHz, contained only consistent content that can be uniquely labeled with one label. To simulate different levels of abstraction, sound samples were chosen from three categories. The first category was made of 32 recorded sound samples, all produced by the same cricket. This category was most specific among three in the implemented category-abstraction space. The second category was consisted of 261 sound samples produced by different fly individuals that belong to several subgroups and families. This category represented the middle of category-abstraction space. And the last category was made of 392 acoustic samples of different backgrounds, starting from human voices, animals, natural phenomena up to the different engines, vehicles, machines, musical instruments and other technical devices. The samples were grouped in the category, simply named "Sound", which represented the most abstract category in the category-abstraction space.

Amplitude-frequency spectrum, as a recognition clue, was calculated by Eq. (3) and the recognition was performed using a feed-forward neural network with the sigmoid activation function and the back-propagation algorithm with momentum. The number of neurons in the input layer, for the constant sampling frequency, was determined by the signal length. The number of outputs was equal to 3, which is the number of signal categories. The rest of the network configuration and training parameters were obtained as a result of examination. The network was tested by 200 samples, not used in the training procedure.

3.2. Near field acoustic localization

The acoustic source was driven along the training path hanged on three strings. The opposite end of each string was wrapped around a step motor driven pulley. Three of these pulleys were geometrically placed in vertices of horizontal, approximately 4.35 m edge long equilateral triangle, above the acoustic source, building a simple routing mechanism (Fig. 1). A meaningful winding and unwinding of the pulleys were used for achieving any specified location of the sound source within a selected range. The step motors were driven by Arduino CNC driver and the driver is governed using PC through USB connection and ATmega328P microcontroller.

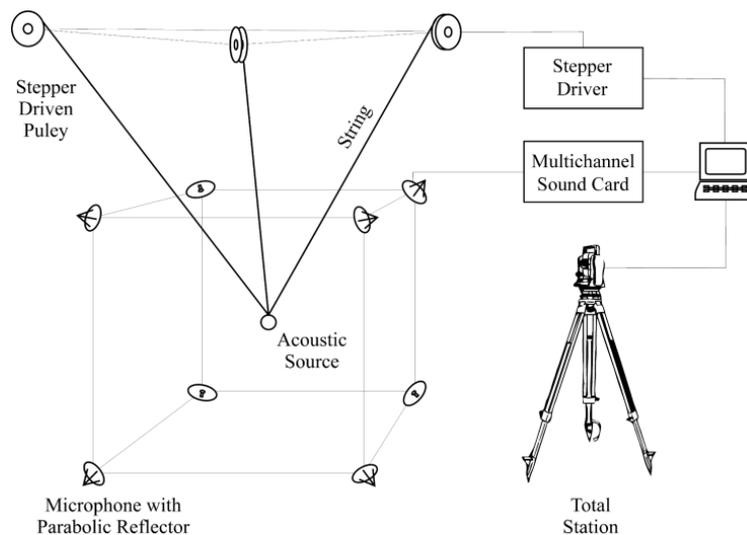


Fig. 1 Sketch of experimental setup for near field acoustic localization

Random number generator was employed for selecting 500 spatial locations within a 1.6 m edge long cubic space, below the three pulleys. The locations were intended for training feedforward neural network in near field acoustic localization. Before training was performed, the locations were ordered using genetic algorithm with the objective of minimizing training route.

The source was stopped, at each of 500 spatial locations and the sound sample was recorded for a period of 3s. For this purpose, the array of 8, low cost, mini spy microphones, was designed and connected to a PC through 8 channel TerraTec EWS88 MT sound card. These microphones were spatially displaced at vertices of 2m edge length cube, symmetrically around the sound source moving zone, with purposely chosen tolerance of ± 20 mm. Parabolic reflector dishes were applied, on each microphone, as an audio signal mechanical amplifier.

The first second, of each acoustic sample, was recorded before the sound source has been turned on as a representative of the noise that exists in the room independently of the sound source activity. The rest of each signal was recorded for a period of 2 s, after the emitter has been activated, as a representative of corrupted signal. The samples were

recorded by cheap equipment in a highly reverberant room full of interfering sound sources (fans and step motors) so they were expected to be corrupted with a high level of noise. To cancel the negative effect of disturbances on TDOA estimation accuracy, the signals were filtered by a suitable second order IIR filter. The filter was designed through the iterative steps of evolutionary strategy in order to minimize the mean absolute TDOA estimation error over the recorded collection of samples. The error was calculated as a difference between TDOA estimated using the preprocessed signal and the theoretical TDOA calculated based on the sound speed and geometrical relations between acoustic components. Except for filter coefficients, the chance was taken for the rest of preprocessing procedure to be configured. Finally, the genotype of the complete preprocessing consisted of 8 variables. The first five of them were digital filter coefficients while one real variable more was employed for determining optimal range of lags to be tested in cross-correlation procedure. Two integer variables were used to make choice among windowing functions and nonlinear operators. The algorithm was started with an initial population of 50 individuals each of them represented one preprocessing configuration. The genotype of the initial individuals was chosen randomly within the logical range of values. The termination condition was formulated as a maximum number of successive evolutions with no improvement in the mean absolute TDOA error. The best configuration was employed for preprocessing in the experiment of acoustic localization.

TDOAs were processed using a feed-forward neural network. The number of neurons in the input layer was determined by the number of employed microphones. In order to achieve the best accuracy, all redundant pairs that correspond to the certain number of microphones were employed as inputs.

$$n = M(M - 1) / 2 \quad (8)$$

where M is number of microphones. The number of outputs was always 3 because the spatial position was determined by 3 independent coordinates. The rest of the network configuration (number of hidden layers and artificial neurons in them) and training parameters were obtained as the result of examination. Network performance was tested along 126 spatial locations, from the same space, which were not used in the training procedure.

4. EXPERIMENTAL RESULTS

In both experiments, the processing time of neural network and of filtering duration was hard to notice in comparison to the time it takes for the preprocessing techniques. This is in accordance with the theoretical predictions about the computational complexity of these procedures.

4.1. Acoustic recognition

The best recognition accuracy was achieved using a feed-forward neural network with 50 neurons in a single hidden layer. The network was trained using learning coefficient 0.025, momentum factor 0.999 and 400 training epochs. The overall accuracy was around 92% (Fig. 2). The result was estimated as satisfactory since it was achieved in the presence of disturbances.

		Target			
		C	F	S	
Recognized	C	53	0	1	98.1%
	F	0	55	9	85.9%
	S	0	6	76	92.7%
		100%	90.2%	88.4%	92.0%

Fig. 2 Confusion matrix as result of the acoustic recognition

From the point of duration, the most important phase of preprocessing, for the acoustic recognition, was fast Fourier transform. Fig. 3. represents the duration of the fast Fourier transform procedure with respect to the signal length while Fig. 4. represents the recognition error with respect to the same parameter.

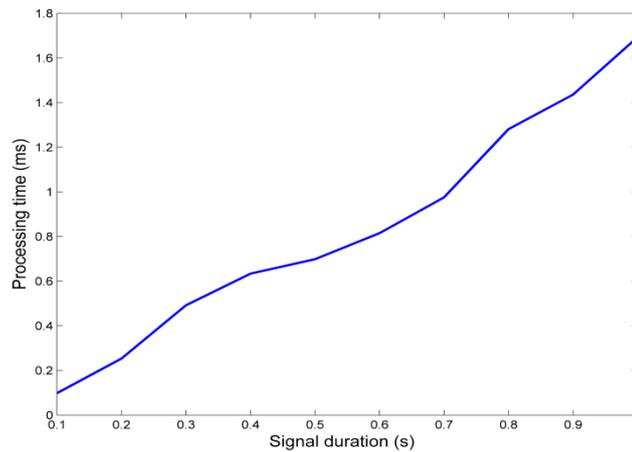


Fig. 3 Fast Fourier transform duration with respect to the signal length

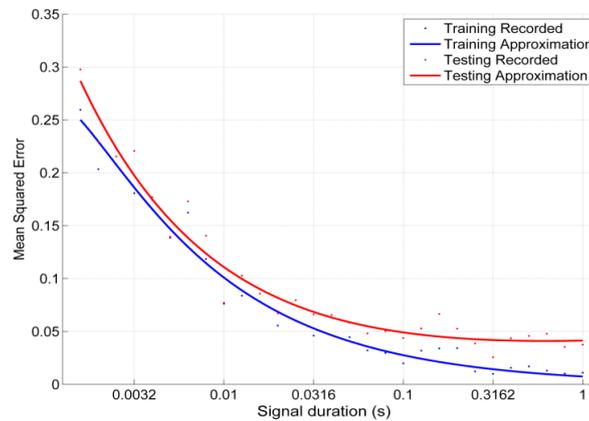


Fig. 4 Mean squared error with respect to the signal length

4.2. Near field acoustic localization

Geometrical relations between the experimental components and the realized spatial positions of the acoustic source were precisely measured using Total Station Sokkia SET630R. The instrument provides a laser measurement of distances with the accuracy of ± 3 mm, at the used range of lengths, memorization and automatic data transfer, through the RS-232 port, to PC. After the realized positions were compared to the given coordinates, the resulting mean absolute error, achieved by the routing mechanism over the whole collection of audio samples, was approximately 10 mm. The accuracy of the mechanism was evaluated as satisfactory regarding the near field acoustic localization because the sound source diameter was approximately 40 mm. Since the training positions were randomly chosen, the realized positions, precisely determined by total station, were adopted as reference for calculating theoretical (reference) TDOA-s.

The path of the acoustic source, between training positions, was optimized using evolutionary algorithm. The procedure reduced the total length of the training path for 4-5 times. The shorter training path did not only have an influence on a shorter time required for the routing mechanism to complete it but it also affected a better positioning accuracy. The reason for this is that the routing mechanism used in the experiment, governed by the winding strings, made a higher error with a longer movement.

All the audio samples collected contained a part recorded before the sound source was activated and the part after it was started. The amplitude frequency spectrums of these are evaluated directly, using fast Fourier transform, and averaged over the whole collection of samples. Subtracting frequency spectrum of noise from the frequency spectrum of corrupted signal resulted in the frequency spectrum of a clear signal, without noise. The ratio estimated between the dominant frequencies of the clear signal and noise was around 0.1 which is equal to the SNR ratio of -20 dB, in the logarithmic scale. According to the literature [22], the minimum SNR ratio, which provides meaningful TDOA estimation, is in the range between -13 dB and -13.5 dB. The SNR evaluated suggested that the collected audio samples, in the experiment of acoustic localization, contained too much noise to be useful for TDOA estimation. The same conclusion was obtained based on the dependency of the mean absolute TDOA estimation error with respect to the sample length (Fig. 5). The diamonds represent the error obtained using raw signals, without any preprocessing, for 8 different lengths of acoustic sample. The TDOAs were calculated by cross-correlation in time domain. The approximation line, between them, was obtained using two terms exponential function. The curve shows increasing of the mean absolute TDOA error with the signal length, which is against logical assumption that more data should provide a better result.

The influence of preprocessing, on the level of mean absolute TDOA estimating error, is demonstrated by the cyan curve marked with triangles. The approximation line was obtained in the same way as previous. The line constantly decreases with the length of acoustic samples. The results presented, in Fig. 5, proved the necessity of preprocessing in the acoustic localization procedure over employed collection of samples.

According to the graph of processed signal, in Fig. 5, the length of 80ms was adopted as optimal for obtaining TDOAs in the experiment of acoustic localization. Further increasing of the signal length, despite higher computational complexity, gave no significant improvement of accuracy. The minimal TDOA estimation error achieved in

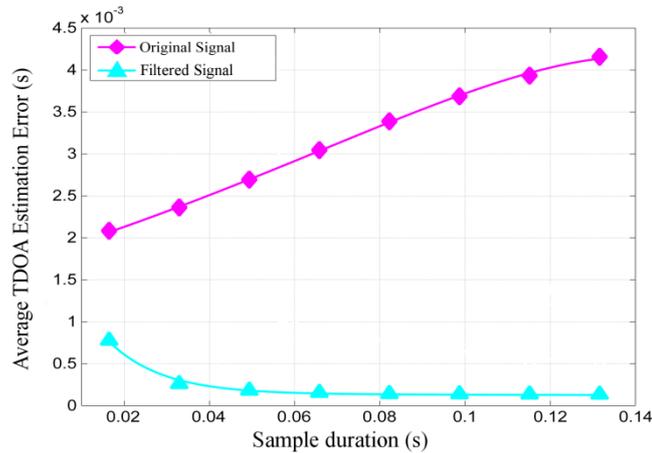


Fig. 5 Average TDOA estimation error with respect to the sample length

the experiment was approximately 0.13ms. For the sound velocity of 334.33 m/s, at the temperature of 5°C that ruled during the experiment. the mean absolute error corresponded to the length of 45 mm. This was estimated as satisfactory, regarding the acoustic source diameter too. The average duration of TDOA estimation, which assumed preprocessing (filtering) and calculating of lags between microphones, was just under 0.1s per location. Fig. 8 represents TDOA processing time, for 8 microphones, with respect to the signal length while Fig. 6 presents TDOA processing time with respect to the number of microphones. Both graphs were obtained with the signal length of 0.8s.

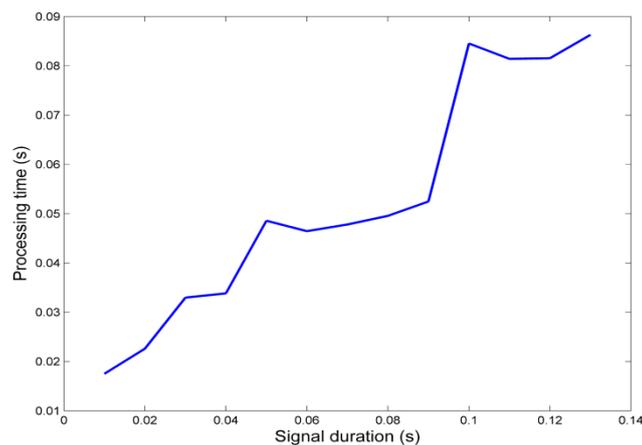


Fig. 6 TDOA processing time with respect to the signal length

After testing different configurations, the artificial neural network was adopted with 10 neurons in a single hidden layer. The network was trained using learning coefficient 0.7, momentum factor 0.9 and 4000 training epochs. The final result is presented in Fig.

7. Average deviation from the actual path was 35.7 mm which is even lower than the mean absolute error of input TDOAs. The accuracy was result of redundant microphones.

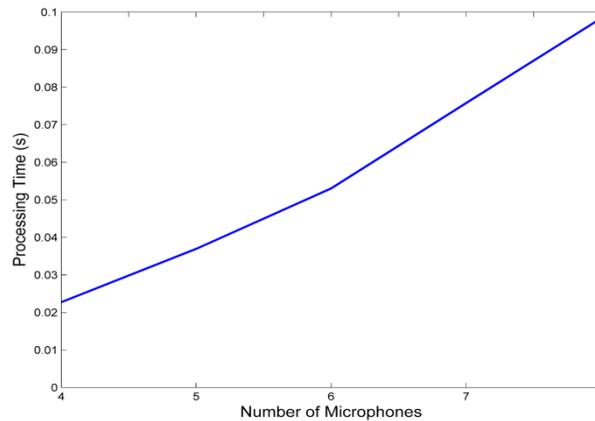


Fig. 7 TDOA processing time with respect to the number of microphones

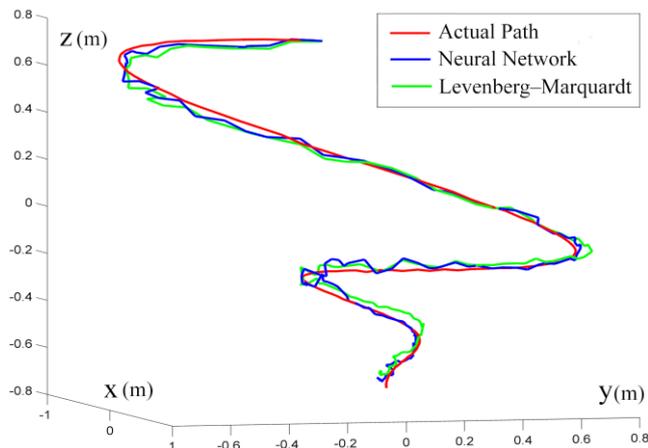


Fig. 8 Actual path and estimated paths of acoustic source

3. CONCLUSION

The most demanding procedure of the acoustic recognition was fast Fourier transform. On the described processor, it lasted about 500 times shorter than signal itself which leaves the possibility to employ the rest of computational power for improving recognition accuracy. One of the techniques is overlapping audio signals that result in raising temporal resolution. The processing time of near field localization is conditioned with the complexity of cross-correlation (3). On the employed processor, it was at the limit of real time application.

The experiment confirms theoretical assumption that the temporal resolution of acoustic perception, by artificial neural networks, strongly depends on the feature extraction procedure. The paper indicates crucial implementation problems of the acoustic perception, which are omitted in literature, and gives some solutions.

Acknowledgements: *This paper presents the results of the research conducted within the project "Research and development of new generation machine systems in the function of the technological development of Serbia" funded by the Faculty of Mechanical Engineering, University of Niš, Serbia.*

REFERENCES

1. McLoughlin, I., Zhang, H., Xie, Z., Song, Y., Xiao, W., Phan, H., 2017, *Continuous robust sound event classification using time-frequency features and deep learning*, Plos ONE, 12(9), pp 1-19.
2. Mcadams, S., 1993, *Recognition of sound sources and events*, Oxford University Press.
3. Bishop, C., 1995, *Neural networks for pattern recognition*, Oxford: Oxford University Press.
4. Michaelides, P.G., Tsionas, E. G., Vouldis, A. T., Konstantakis, K. N., Patrinos, P., 2018, *A Semi-Parametric Non-linear Neural Network Filter: Theory and Empirical Evidence*, Computational Economics, 51(3), pp 637-675.
5. Choi, J.Y., Hu, E.R., Perrachione, T.K., 2018, *Varying acoustic-phonemic ambiguity reveals that talker normalization is obligatory in speech processing*, Attention, Perception & Psychophysics, 80(3), pp. 784-797.
6. Flaszinski, M., 2016, *Introduction to Artificial Intelligence*, Springer, Cham.
7. Osamu, I., Masafumi, T., Tetsuya, N., 2003, *Sound Source Localization Using a Pinna-based Profile Fitting Method*, Ieice Transactions - IEICE, pp. 263-266.
8. Johnson, M. L., 2015, Systems and methods of processing information regarding weapon fire location using projectile shockwave and muzzle blast times of arrival data, Retrieved from <http://search.ebscohost.com/login.aspx?direct=true&db=edsppgr&AN=edsppgr.08995227&site=eds-live> (last access: 01.03.2019).
9. Rowell, C.R., 2014, *Three-Dimensional Volcano-Acoustic Source Localization at Karymsky Volcano, Kamchatka, Russia*, Journal of Volcanology and Geothermal Research, 283, pp. 101-115.
10. Martín, S. R., Genescà, M., Romeu, J., Clot, A., 2016, *Aircraft localization using a passive acoustic method. Experimental test*, Aerospace Science and Technology, 48, pp. 246-253.
11. Grabowski, K., 2016, *Time-distance domain transformation for Acoustic Emission source localization in thin metallic plates*, Ultrasonics, 68, pp. 142-149.
12. Tan, C., 2016, *A low-cost centimeter-level acoustic localization system without time synchronization*, Measurement, 78, pp. 73-82.
13. Kovandžić, M., Nikolić, V., Al-Noori, A., Ćirić, I., Simonović, M., 2017, *Near field acoustic localization under unfavorable conditions using feedforward neural network for processing time difference of arrival*, Expert Systems with Applications, 7(1), pp 138-146.
14. Park, C., Jeon, J., Kim, Y., 2014, *Localization of a sound source in a noisy environment by hyperbolic curves in quefreny domain*, Journal Of Sound And Vibration, 333, pp. 5630-5640.
15. Hing, C.S., 2005, *A comparative study of two discrete-time phase delay estimators*, IEEE Transactions on Instrumentation and Measurement, 54, pp. 2501-2504.
16. Khaddour, H., 2011, *A Comparison of Algorithms of Sound Source Localization Based on Time Delay Estimation*, Elektrovjeue, 2(1), pp. 31-37.
17. Babic, M., Cali, M., Nazarenko, I. et al., 2019, *Surface roughness evaluation in hardened materials by pattern recognition using network theory*, International Journal on Interactive Design and Manufacturing, 13(1), pp. 211-219.
18. Fragassa, C., Babic, M., Bergmann, C., Minak, G., 2019, *Predicting the tensile behaviour of cast alloys by a pattern recognition analysis on experimental data*, Metals, 9(5), 557.
19. Rojas, R., 1996, *Neural Networks*, Springer.
20. George, I., Cousillas, H., Richard, J., Hausberger, M., 2008, *A Potential Neural Substrate for Processing Functional Classes of Complex Acoustic Signals*, Plos ONE, 3(5), pp 1-10.
21. Smith, W. S., 1997, *Digital signal processing*, California Technical Publishing. San Diego.
22. Dhull, S., Arya, S., Sahu, O.P., 2010, *Comparison of time-delay estimation techniques in acoustic environment*, International Journal of Computer Applications, 8(9), pp 29-31.