Natural Language Human-Machine Interface Using Artificial Neural Networks

Maciej Majewski and Wojciech Kacalak

Technical University of Koszalin,
Department of Mechanical Engineering,
Raclawicka 15-17, 75-620 Koszalin, Poland
{maciej.majewski, wojciech.kacalak}@tu.koszalin.pl

Abstract. In this paper there is a natural language interface presented, which consists of the intelligent mechanisms of human identification, speech recognition, word and command recognition, command syntax and result analysis, command safety assessment, technological process supervision as well as human reaction assessment. In this paper there is also a review of the selected issues on recognition of speech commands in natural language given by the operator of the technological device. A view is offered of the complexity of the recognition process of the operator's words and commands using neural networks made of a few layers of neurons. The paper presents research results of speech recognition and automatic recognition of commands in natural language using artificial neural networks.

1 Intelligent Two-Way Speech Communication

If the operator is identified and authorized by the natural language interface in Fig. 1, a produced command in continuous speech is recognized by the speech recognition module and processed to the text format. Then the recognised text is analysed with the syntax analysis subsystem. The processed command is sent to the word and command recognition modules using artificial neural networks to recognise the command, which next is sent to the effect analysis subsystem for analysing the status corresponding to the hypothetical command execution, consecutively assessing the command correctness, estimating the process state and the technical safety, and also possibly signalling the error caused by the operator. The command is also sent to the safety assessment subsystem for assessing the grade of affiliation of the command to the correct command category and making corrections. The command execution subsystem signalises commands accepted for executing, assessing reactions of the operator, defining new parameters of the process and run directives [4]. The subsystem for voice communication produces voice commands to the operator [5].

J. Wang et al. (Eds.): ISNN 2006, LNCS 3973, pp. 1161–1166, 2006.

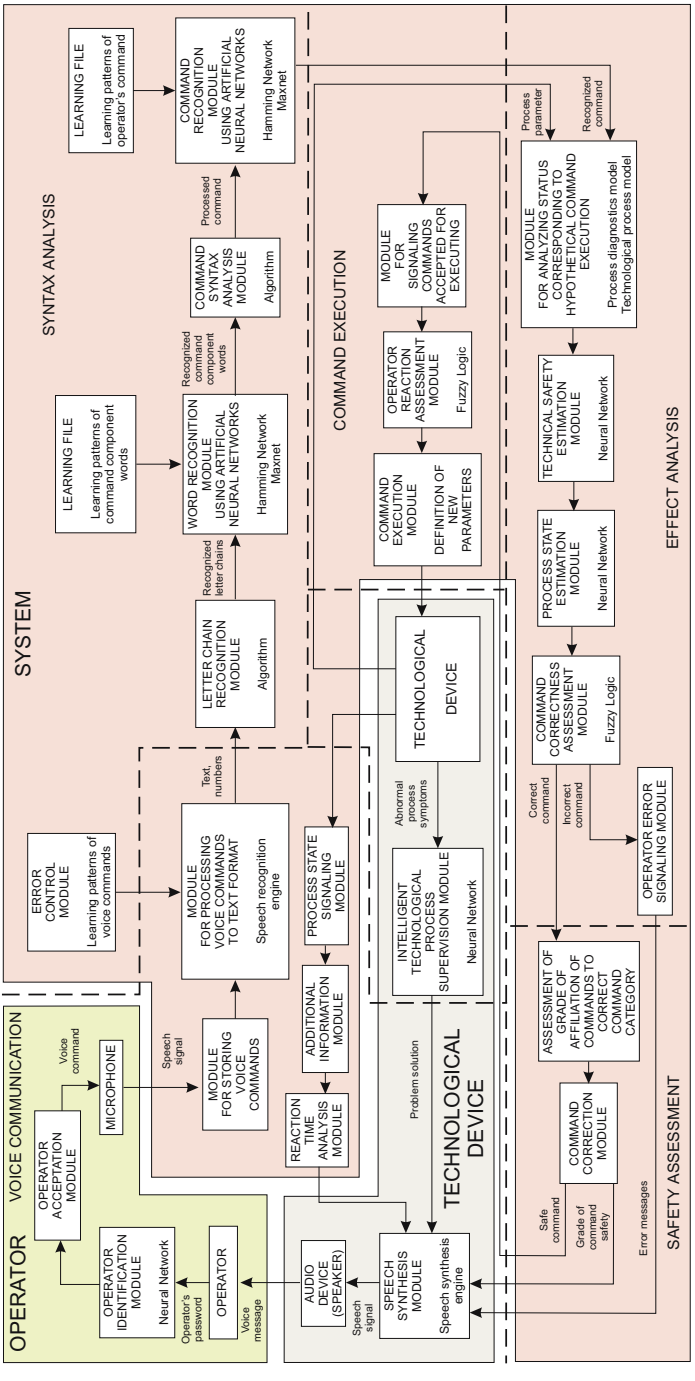

Fig. 1. Architecture of the natural language human-machine interface

2 Recognition of Commands in Natural Language

In the automatic command recognition system as shown in Fig. 2, the speech signal is processed to text and numeric values with the module for processing voice commands to text format. The speech recognition engine is a continuous density mixture Gaussian Hidden Markov Model system which uses vector quantization for speeding up the Euclidean distance calculation for probability estimation [1,2]. The system uses context dependent triphonic cross word acoustic models with speaker normalization based on vocal tract length normalization, channel adaptation using mean Cepstral subtraction and speaker adaptation using Maximum Likelihood Linear Regression. The separated words of the text are the input

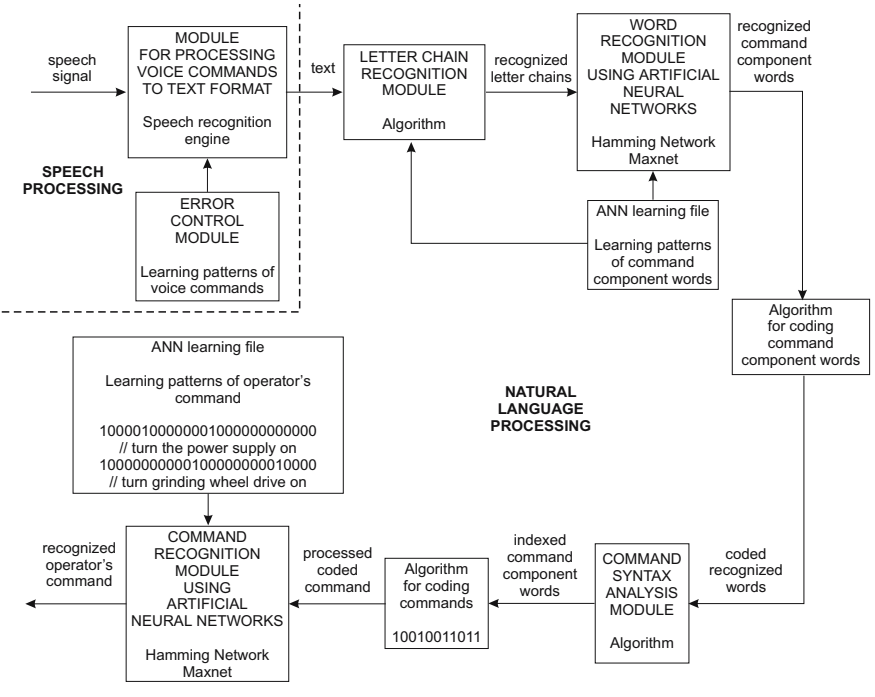

Fig. 2. Scheme of the automatic command recognition system

signals of the neural network for recognizing words. The network has a training file containing word patterns. The network recognizes words as the operator's command components, which are represented by its neurons. The recognized words are sent to the algorithm for coding words. Next the coded words are transferred to the command syntax analysis module. It is equipped with the algorithm for analysing and indexing words. The module indexes words properly and then they are sent to the algorithm for coding commands. The commands

are coded as vectors and they are input signals of the command recognition module using neural network. The module uses the 3-layer Hamming neural network in Fig. 3, either to recognize the operator's command or to produce the information that the command is not recognized. The neural network is equipped with a training file containing patterns of possible operator's commands [3].

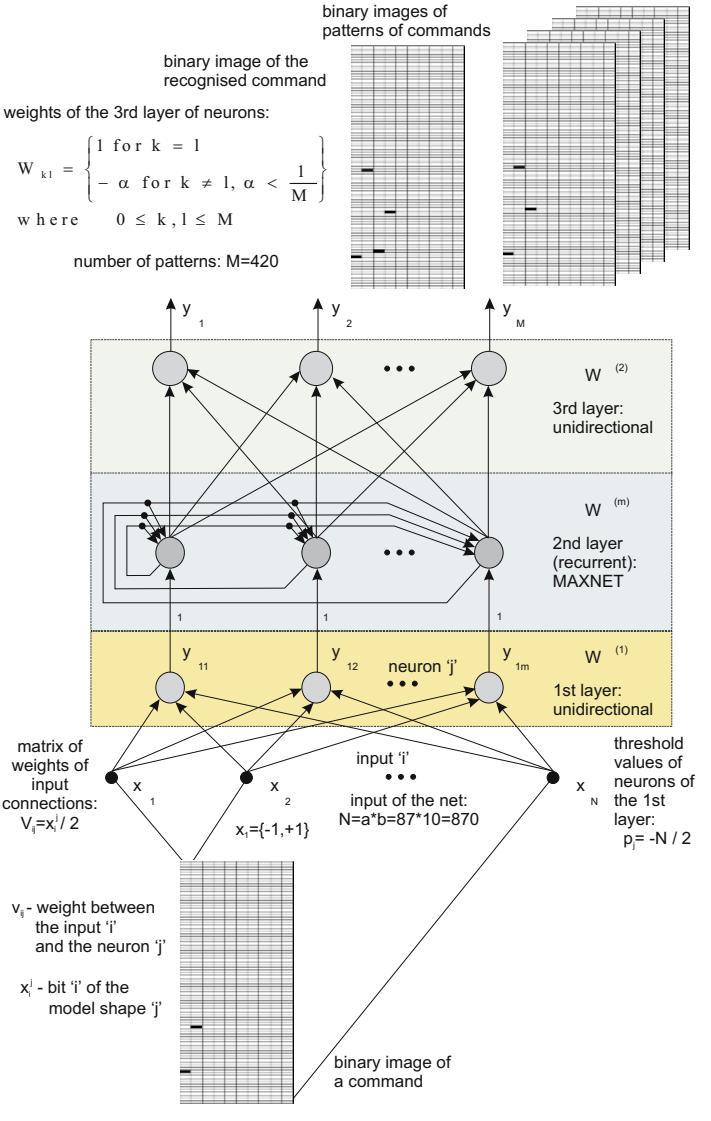

Fig. 3. Scheme of the 3-layer neural network for automatic command recognition

3 Research Results of Automatic Command Recognition

As shown in Fig. 4a, the speech recognition module recognizes 85-90% of the operator's words correctly. As more training of the neural networks is done, accuracy rises to around 95%. For the research on command recognition at different noise power, the microphone used by the operator is the headset. As shown in Fig. 4b, the recognition performance is sensitive to background noise. The recognition rate is about 86% at 70 dB and 71% at 80 dB. Therefore, background noise must be limited while giving the commands. For the research on command

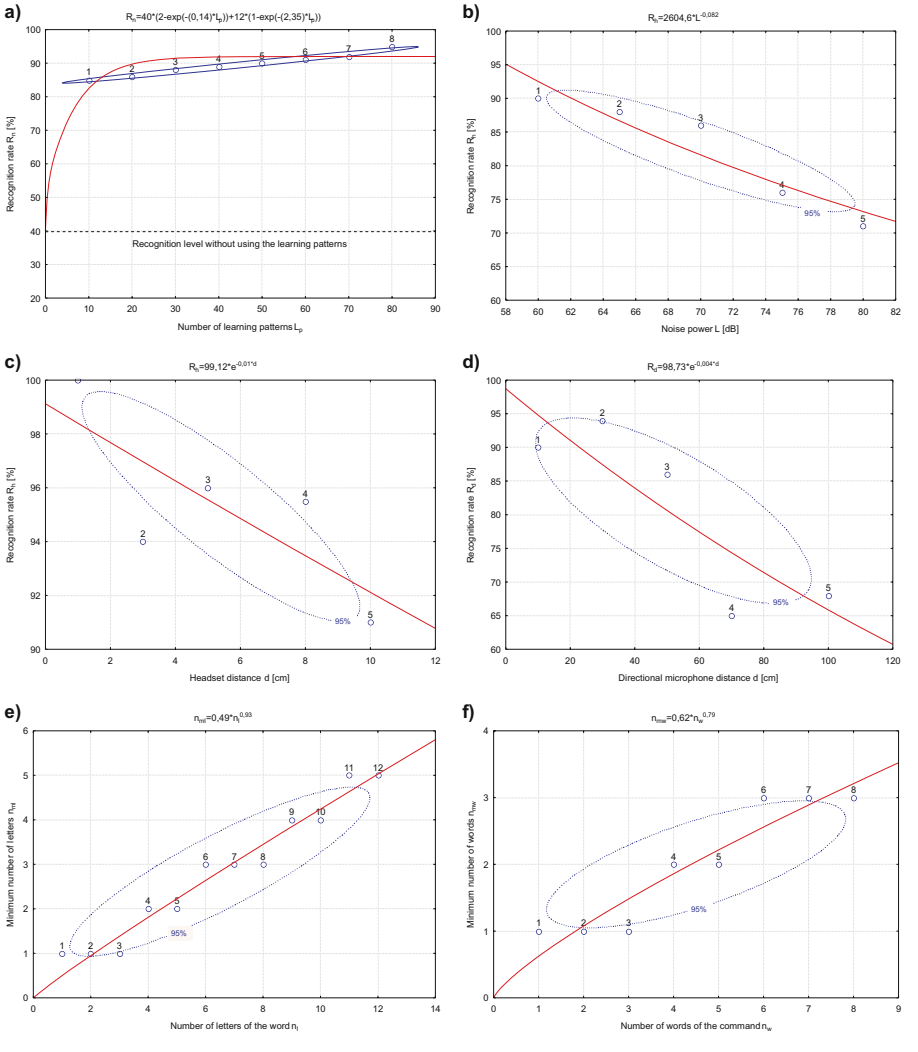

Fig. 4. Speech and command recognition rate

recognition at different microphone distances, the microphone used by the operator is the headset. As shown in Fig. 4c, the recognition rate decreases when the headset distance increases. The recognition rate has been dropped for 9% after the headset distance is changed from 1 to 10 cm. Also for the research on command recognition at different microphone distances, the microphone used by the operator is the directional microphone. As shown in Fig. 4d, the recognition rate after 50 cm decreases reaching rate about 65%. As shown in Fig. 4e, the ability of the neural network to recognise the word depends on the number of letters. The neural network requires the minimal number of letters of the word being recognized as its input signals. As shown in Fig. 4f, the ability of the neural network to recognise the command depends on the number of command component words. Depending on the number of component words of the command, the neural network requires the minimal number of words of the given command as its input signals.

4 Conclusions and Perspectives

The condition of the effectiveness of the presented system is to equip it with mechanisms of command verification and correctness. In operations of the automated technological processes, many process states and various commands from the operator to the technological device can be distinguished. A large number of combined technological systems characterize the realization of that process. In complex technological processes, if many parameters are controlled, the operator is not able to analyze a sufficient number of signals and react by manual operations on control buttons. The research aiming at developing an intelligent layer of two-way voice communication is very difficult, but the prognosis of the technology development and its first use shows a great significance in efficiency of supervision and production humanization.

References

1. Kacalak, W., Majewski, M.: Automatic Recognition of Voice Commands in Natural Language Using Artificial Neural Networks. Artificial Neural Networks in Engineering ANNIE 2004 Conference, St. Louis. ASME Press, New York (2004) 831-836
2. Majewski, M., Kacalak, W.: Intelligent Human-Machine Voice Communication System. Engineering Mechanics International Journal 12(3) (2005) 193-200
3. Majewski, M., Kacalak, W.: Intelligent Layer of Human-Machine Voice Communication. 9th International Conference KES2005, Melbourne, Australia (2005), Lecture Notes in Computer Science 3683. Springer-Verlag (2005) 930-936
4. Majewski, M., Kacalak, W.: Intelligent Human-Machine Speech Communication System. International Conference on Intelligent Computing ICIC2005, Hefei Anhui, China (2005) 3441-3450
5. O'Shaughnessy, D.: Speech Communications: Human and Machine. IEEE Press, New York (2000)