

Intelligent Speech-Based Interactive Communication Between Mobile Cranes and Their Human Operators

Maciej Majewski^(✉) and Wojciech Kacalak

Faculty of Mechanical Engineering, Koszalin University of Technology,
Raclawicka 15-17, 75-620 Koszalin, Poland

{maciej.majewski,wojciech.kacalak}@tu.koszalin.pl

Abstract. In this paper, an overview of human-machine interactive communication for controlling lifting devices is presented, covering also the integration with vision and sensorial systems. Following a general concept, and motivation towards intelligent human-machine communication through artificial neural networks, selected methods are proposed, which provide further directions both of recent as well as of future research on human-machine interaction. The aim of the experimental research is to design a prototype of an innovative interaction system, equipped with a speech interface in a natural language, augmented reality and interactive manipulators with force feedback. The presented research offers the possibility of motivating and inspiring further development of the intelligent speech interaction system and methods that have been elaborated in this paper.

Keywords: Intelligent interface · Neural networks · Interactive system · Speech communication · Intelligent control · Natural language processing

1 The Design of an Innovative Human-Machine Interface

The most up-to-date artificial intelligence-based technologies find their application in the process of designing modern systems for controlling and supervising machines. An example are vision systems - machine vision, augmented reality, voice communication as well as interactive controllers providing force feedback. The design and implementation of intelligent human-machine interactive communication systems is an important field of applied research. Recent advances in development of prototypes of human-machine speech-based interfaces are described in articles in [1–3].

The presented research involves the development of a system for controlling a mobile crane, equipped with a vision and sensorial system, interactive manipulators with force feedback, as well as a system for bi-directional voice communication through speech and natural language between an operator and the controlled lifting device [4]. The system is considered intelligent, because it is capable of learning from previous commands to reduce human errors.

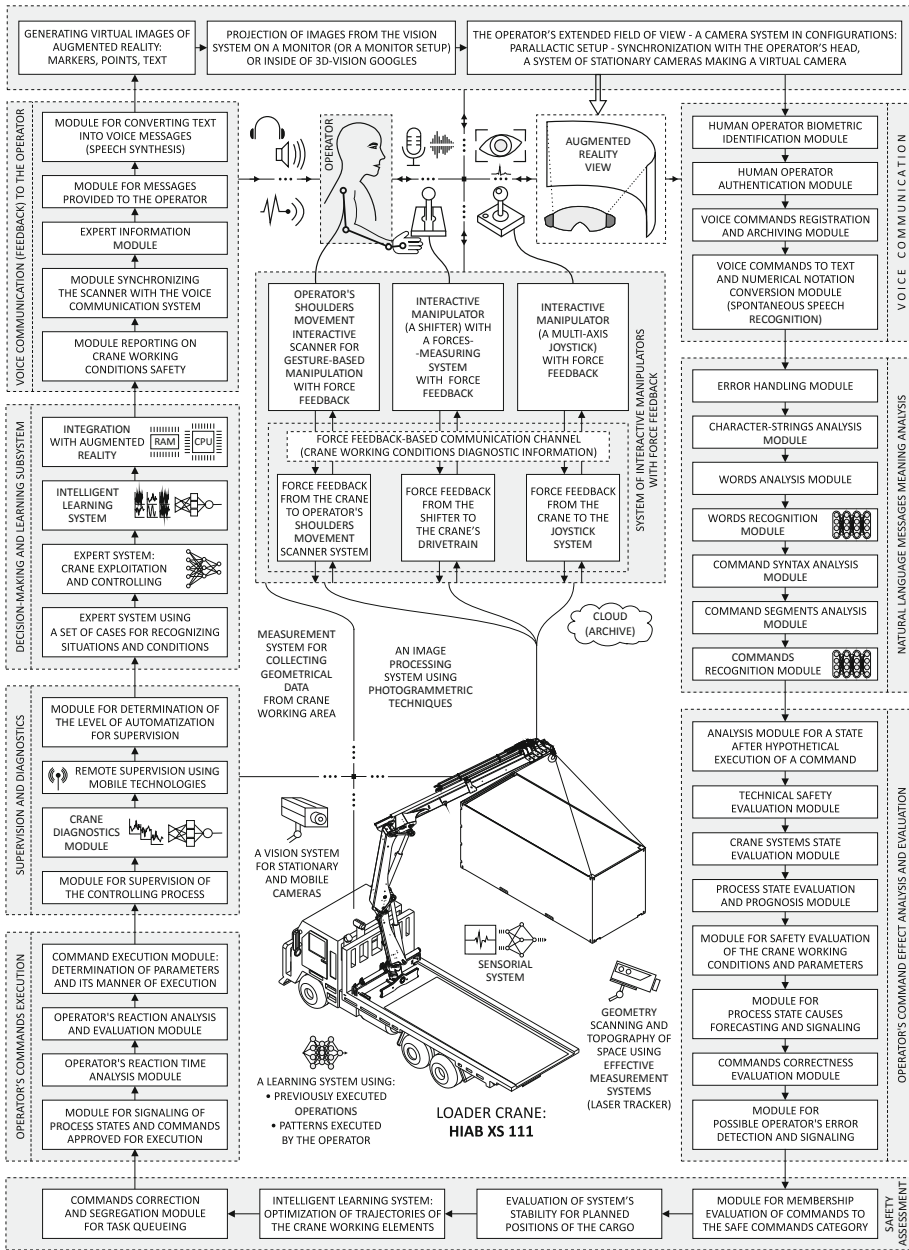


Fig. 1. Designed structure of an innovative system for interaction of the loader crane (Hiab XS 111) with its operator equipped with a speech interface, vision and sensorial systems, and interactive manipulators with force feedback.

The ARSC (Augmented Reality & Smart Control) prototype control system uses: intelligent visual-aid systems based on augmented reality, interactive manipulation systems providing force feedback, as well as natural-language voice communication techniques. We propose a new concept which consists of a novel approach to these systems, with particular emphasis on their ability to be truly flexible, adaptive, human error-tolerant, and supportive both of human-operators and data processing systems. The concept specifies integration of a system for natural-language communication with a visual and sensorial system.

The proposed interactive system (Fig. 1) contains many specialized modules and it is divided into the following subsystems: a subsystem for voice communication between a human-operator and the mobile crane, a subsystem for natural language meaning analysis, a subsystem for operator's command effect analysis and evaluation, a subsystem for command safety assessment, a subsystem for command execution, a subsystem of supervision and diagnostics, a subsystem of decision-making and learning, a subsystem of interactive manipulators with force feedback, and a visual and sensorial subsystem. The novelty of the system also consists of inclusion of several adaptive layers in the spoken natural language command interface for human biometric identification, speech recognition, word recognition, sentence syntax and segment analysis, command analysis and recognition, command effect analysis and safety assessment, process supervision and human reaction assessment.

2 Meaning Analysis of Commands and Messages

The concept of the ARSC system includes a subsystem of recognition of speech commands in a natural language using patterns and antipatterns of commands, which is presented in Fig. 2.

In the subsystem, the speech signal is converted to text and numerical values by the continuous speech recognition module. After a successful utterance recognition, a text command in a natural language is further processed. Individual words treated as isolated components of the text are subsequently processed with the modules for lexical analysis, tokenization and parsing. After the text analysis, the letters grouped in segments are processed by the word analysis module. In the next stage, the analyzed word segments are inputs of the neural network for recognizing words. The network uses a training file containing also words and is trained to recognize words as command components, with words represented by output neurons.

In the meaning analysis process of text commands (Fig. 3A) in a natural language, the meaning analysis of words as command or message components is performed. The recognized words are transferred to the command syntax analysis module which uses command segment patterns. It analyses commands and identifies them as segments with regards to meaning, and also codes commands as vectors. They are sent to the command segment analysis module using encoded command segment patterns. The commands become inputs of the command recognition module. The module uses a 3-layer Hamming network to classify the

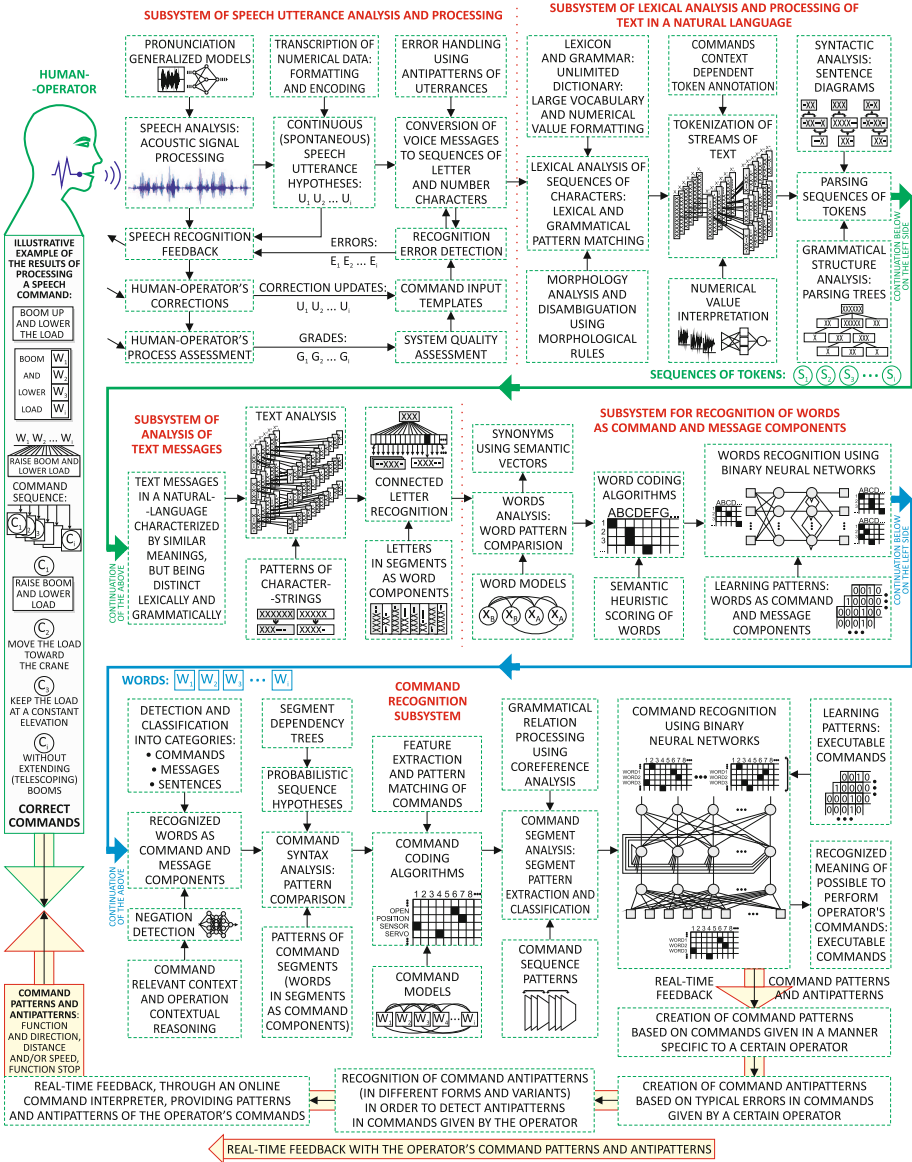


Fig. 2. A concept of a system of recognition of speech commands in a natural language using patterns and antipatterns of commands.

command and find its meaning (Fig. 3B). The neural network of this module uses a training file with meaningful executable commands.

The proposed method for meaning analysis of words, commands and messages uses binary neural networks (Fig. 3A and B) for natural language

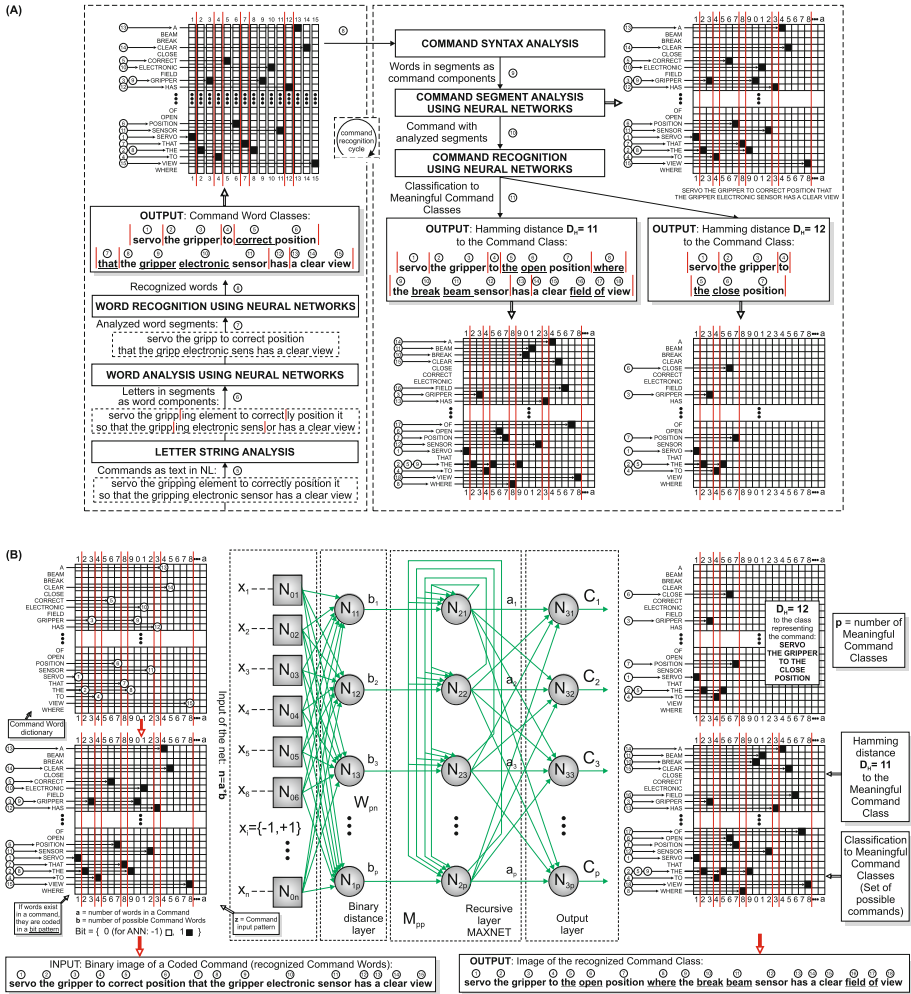


Fig. 3. (A) Block diagram of a meaning analysis cycle of an exemplary command, (B) Illustrative example of recognition of commands using binary neural networks.

understanding. The motivation behind using this type of neural networks for meaning analysis [5] is that they offer an advantage of simple binarization of words, commands and sentences, as well as very fast training and run-time response. The cycle of meaning analysis for an exemplary command is presented in Fig. 3A. The proposed concept of processing of words and messages enables a variety of analyses of the spoken commands in a natural language.

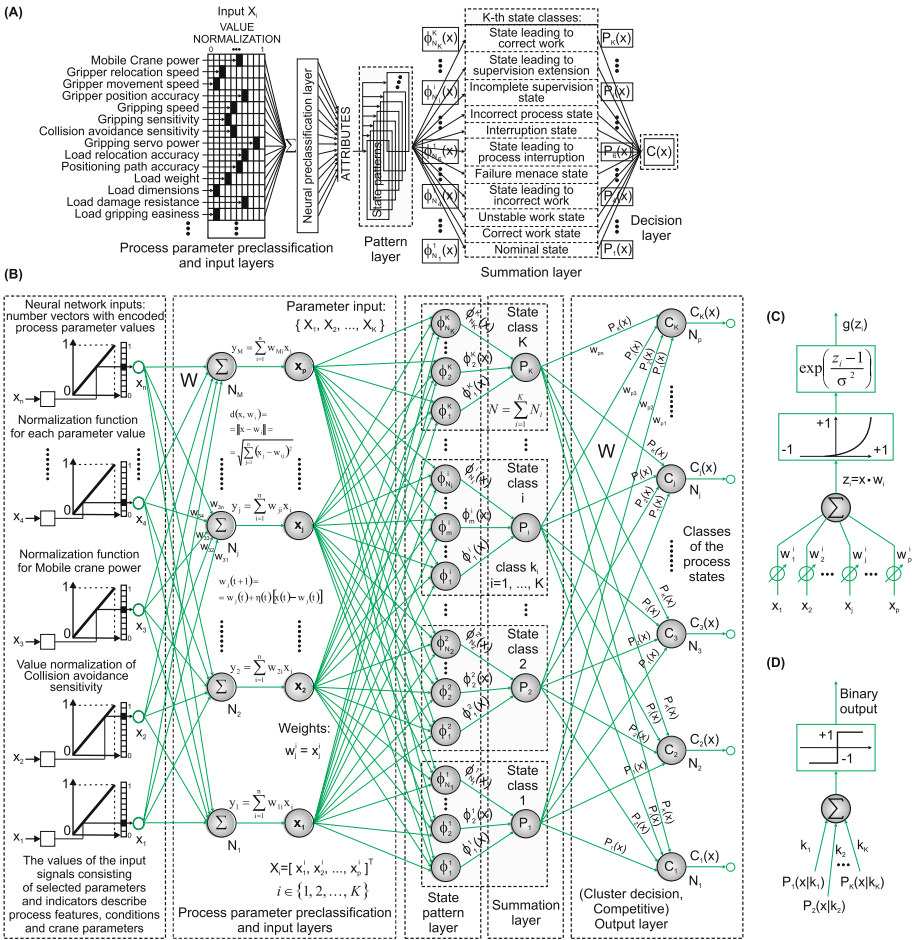


Fig. 4. (A) Hybrid neural model of effect analysis and safety assessment of commands in a cargo manipulation process, (B) The architecture of the hybrid neural network used, (C) Neuron of the pattern layer, (D) Neuron of the P of the output layer.

3 Effect Analysis and Safety Assessment of Commands

The problem of effect analysis and safety assessment of commands can be solved with hybrid neural networks. The proposed method (Fig. 4A) uses developed hybrid multilayer neural networks consisting of a modified probabilistic network combined with a single layer classifier. The probabilistic network is interesting, because it is possible to implement and develop numerous enhancements, extensions, and generalizations of the original model [6]. The effect analysis and safety assessment of commands is based on information on features, conditions and parameters of the cargo positioning process. The developed hybrid network

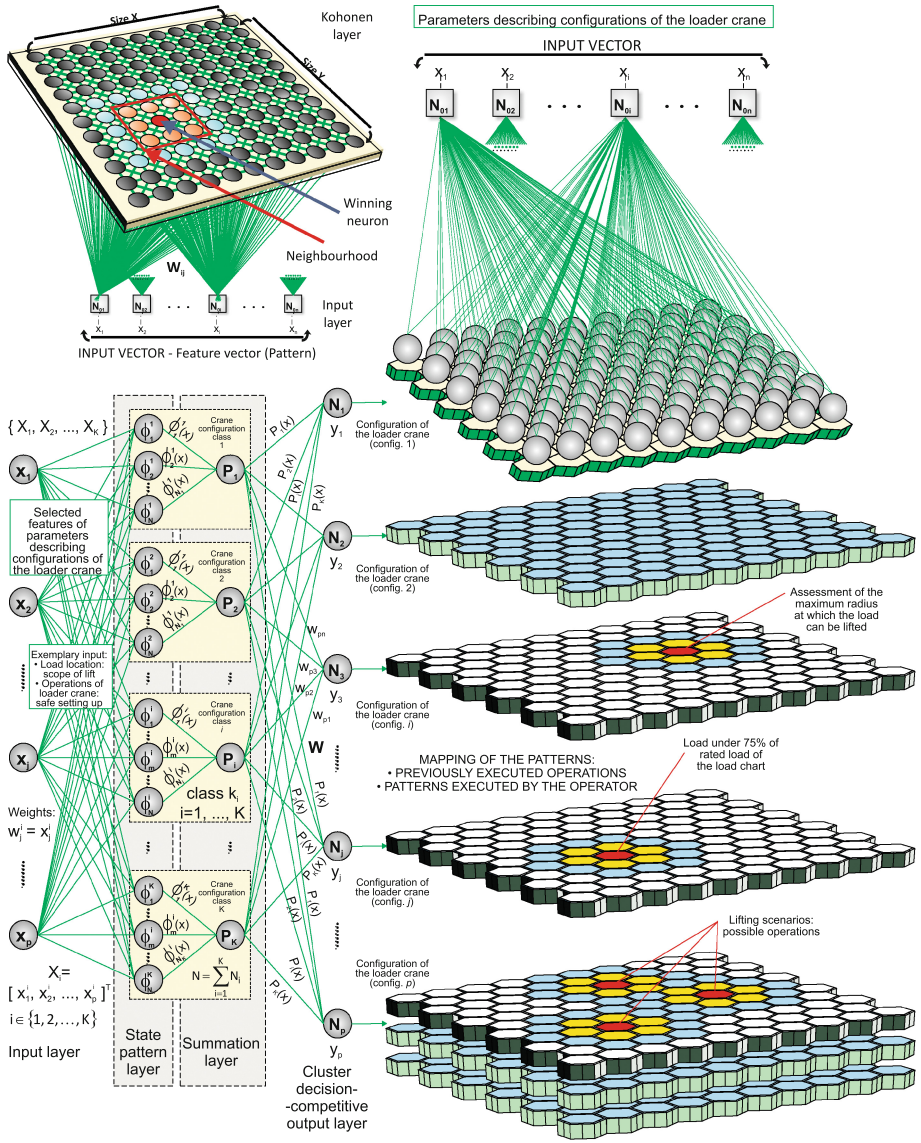


Fig. 5. Proposed learning systems using previously executed operations and patterns executed by the operator.

(Fig. 4B, C and D) is applied for classification of the cargo manipulation process state.

The proposed innovative speech interface is equipped with learning systems using previously executed operations and patterns executed by the operator. The developed learning systems are based on proposed hybrid neural networks

(Fig. 5) consisting of self-organizing feature maps (Kohonen networks [7]) combined with a probabilistic classifier. The inputs of the hybrid networks contain selected features of the parameters describing configurations of the loader crane. The outputs represent individual configurations of the crane which provide self-organizing feature maps of the previously executed operations and patterns executed by the operator.

4 Conclusions and Perspectives

The designed interaction system is equipped with the most modern artificial intelligence-based technologies: voice communication, vision systems, augmented reality and interactive manipulators with force feedback. Modern control and supervision systems allow to efficiently and securely transfer, and precisely place materials, products and fragile cargo. The proposed design of the innovative AR speech interface for controlling lifting devices has been based on hybrid neural network architectures. The design can be considered as an attempt to create a new standard of the intelligent system for execution, control, supervision and optimization of effective and flexible cargo manipulation processes using communication by speech and natural language.

Acknowledgements. This project is financed by the National Centre for Research and Development, Poland (NCBiR), under the Applied Research Programme - Grant agreement No. PBS3/A6/28/2015.

References

1. Kacalak, W., Majewski, M., Budniak, Z.: Interactive systems for designing machine elements and assemblies. *Manag. Prod. Eng. Rev.* **6**(3), 21–34 (2015). De Gruyter Open
2. Kumar, A., Metze, F., Kam, M.: Enabling the rapid development and adoption of speech-user interfaces. *Computer* **47**(1), 40–47 (2014). IEEE
3. Ortiz, C.L.: The road to natural conversational speech interfaces. *IEEE Internet Comput.* **18**(2), 74–78 (2014). IEEE
4. Majewski, M., Kacalak, W.: Intelligent speech interaction of devices and human operators. In: Silhavy, R., et al. (eds.) *CSOC 2016. AISC*, vol. 465, pp. 471–482. Springer, Switzerland (2016)
5. Majewski, M., Zurada, J.M.: Sentence recognition using artificial neural networks. *Knowl. Based Syst.* **21**(7), 629–635 (2008). Elsevier
6. Specht, D.F.: Probabilistic neural networks. *Neural Netw.* **3**(1), 109–118 (1990). Elsevier
7. Kohonen, T.: *Self-Organization and Associative Memory*. Springer, Heidelberg (1984)