

# The Modern State of Human-Robot Interaction: From Collaboration to Cooperation

Arkady S. Yuschenko

Bauman Moscow State Technical University (BMSTU), Robotic Systems and Mechatronics Department, Moscow 105005, Russia

## Abstract

The main trend of modern robotics is an active introduction of robotics into different areas of human activity. The first attempts of the man working together with robots in the same working space were realized on industrial conveyor. Here appeared the new term Collaborative Robots. But lately the term was widely applied for robotic devices with elements of AI capable to autonomous work under human control. Examples are the robots for medical operations and rehabilitation procedures, for operations in dangerous environment in space, underwater, etc. Such operations are much more complicated than usual labor in industry. The new term was Interactive Collaborative Robotics. We suppose that real human-robot interaction in such systems is more close to cooperation between "intelligent" robotic system and human-operator than collaboration in usual labor operations.

The new type of human-robot interaction has raised the new problems. The first is the problem of communication which now resemble the bilateral dialogue. Special knowledge base and problem-oriented language most convenient for user are necessary for speech dialogue mode of control. In usual dialogue between peoples we use the natural relations of space and time such as close, distantly, slow, rapidly and so on. It is examples of linguistic variables formalized as fuzzy variables. Modern fuzzy logic theory allows to form the production rules of behavior and to construct the logical inference. Here is necessary to mention the founder of the theory L. Zadeh and among numerous investigators in our country D.A. Pospelov and others. As for medium of cooperative systems realization the artificial neural networks seem the most suitable. We pay attention that some psychological concepts are also presented in such approach. Among them are psychophysiological scales which determinate the human perception. Fuzzy logic and neural networks seems a base to formalize the human – robot interaction in cooperative systems. It is also the way to the well ground theory of such systems elaboration taking into attention the restrictions both robot and human-operator.

The paper contains some examples of human-robot cooperation systems. Among them is the "robot-assistant" of surgeon controlled with a speech dialogue system and equipped with a convolution neuro-network to recognize the necessary medical instrument. Another example is the system of "intuitive programming" of manipulation robot based on the augmented reality.

## Introduction

### Co-operative robotics as a new stage of human-robot relations

The main trend of the modern stage of scientific and technological development is the socialization of robotics. The first step in this area was "collaborative robots", or Co-Robots. In accordance with the international standard ISO 8373: 2012-03 "Robots and robotic devices" (2012), collaborative technological (robotic) operations include such operations which a person performs together with a robot in the same workspace. Robotic systems that allow such human-robot interaction were called collaborative. Initially, this term was used in relation to industrial manipulators, which were used to help a person when performing work on the conveyor. The basic principles of this relatively new direction are robot safety for user and possibility of using a robotic system without a complicated training in robotics and programming. The leading enterprises in robotics now produced the industrial robots satisfied these requirements and market of Co-robots extends rapidly.

Success of collaborative robotics have strengthened the robotization trend in more complicated area of human activity such as medical operations and rehabilitation procedures, operations in dangerous environment in space, underwater, etc. The collaborative robots acquired the intellectual component to a considerable degree and possibilities for autonomous work together with human user.

## Publication History:

Received: November 07, 2022

Accepted: December 06, 2022

Published: December 08, 2022

## Keywords:

Collaborative robots, Human operator, Robotic system, Verbal dialogue, Emotional support, Augmented reality, Neural networks.

In connection a new term "Interactive Collaborative Robot" (ICR) have appeared. For some years ago, the annual international scientific and technical conference "Interactive Collaborative Robotics" has been held to discuss the new problems, connected with theory and application of ICR. Now we may determine some branches in this new direction in robotics.

Evolution of remotely controlled handling systems has led to the emergence of a new type of system, which could be called *interactive robotic remote-control system*. A well-known example of such a system is the Da Vinci robotic surgical system, which facilitates the work of a surgeon (Figure 1). This system is not a robot in generally, but rather a smart robotic tool in hands of a professional surgeon. But it is an ICR with human-robot interaction (HRI) at the level of direct control

**\*Corresponding Author:** Prof. Arkady S. Yuschenko, Bauman Moscow State Technical University (BMSTU), Robotic Systems and Mechatronics Department, Moscow 105005, Russia, E-mail: [arkadyus@mail.ru](mailto:arkadyus@mail.ru)

**Citation:** Yuschenko AS (2022) The Modern State of Human-Robot Interaction: From Collaboration to Cooperation. Int J Comput Softw Eng 7: 182. doi: <https://doi.org/10.15344/2456-4451/2022/182>

**Copyright:** © 2022 Yuschenko. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

and partial automation. Suppose that it is not a collaboration in usual sense as the technical partner is rather passive one.



Figure 1: RTS “Da Vinci”.

Another area in the field of collaborative robotics, based on the direct inclusion of a person in a robotic system, is exoskeletons. Such devices are the medium to enhance the physical capabilities of a person, as well as for the rehabilitation of patients with impaired motor functions. The usage of these devices in medicine to restore such functions requires the usage of new information technologies. Among them is the complex processing of bio-potentials of muscles of human limbs for robotic system control [3]. An example of such system is the HAL project (Figure 2). In essence, this is a human-robot robotic system, controlled by a person with direct movements of his limbs, or bio-potentials, corresponding to such movements. The term “collaborative” also applied to such systems, although it might be more correct to talk about a *robotic system with biotechnological control*.



Figure 2: Exoskeleton HAL.

The question now is what type of HRI is close to real interaction. We suppose that it is a robot-human system with both active participants – as human as robot and with forms similar of interaction between peoples. Firstly, it is a bilateral speech dialogue between human-user and robot-assistant. So we may speak of the *dialogue mode of robotic system control* [4]. The effectiveness of such dialogue is possible to raise using an emotional support and gesture [5,6]. Such modes of robot control need the elements of artificial intelligence of robot as a human assistant. Now we include the robot in the execution in the

implementation of complex operations together with the person. The HRI now resemble the relation between two partners where one is master and the second is assistant of the first one. We suppose that more correct name of such human-robot system is a *co-operative robotic system*. That is the sense of the modern trend in robotics - the evolution from collaborative Robotics to co-operative robotics.

### Co-operative robotic systems as a new class of RTS

A new class of *co-operative robots*, which maintain active interaction with human-operator during the operation may be defined as a “*partner robots*”. As an example, one could name rescue mobile robots that independently explore a certain territory after an accident or disaster. Such robots could not be only observers, but they are able to carry out some manipulation operations in the outside world under the control of their master-operator. In this case, the robot has a system for recognizing and analysing environmental objects, perceives the operator’s speech instructions, specifying them if necessary. Than the robot is capable to plan and perform complex operations independently in the workspace under human supervision.

Functional diagram of a co-operative robotic system is shown on Figure 3. In fact, the diagram consists of three main blocks – a human-operator’s model, an autonomous robot with elements of artificial intelligence, and a block of human-robot interface with enhanced intellectual capabilities. The robot control system is equipped with the necessary external information sensors (vision system, laser scanners, etc.) and internal sensors of robot control system. That is necessary to create an adequate image of the world, i.e. the external environment in which the activity occurs, including the operator himself. That is the base for operations planning in human-robot interface.

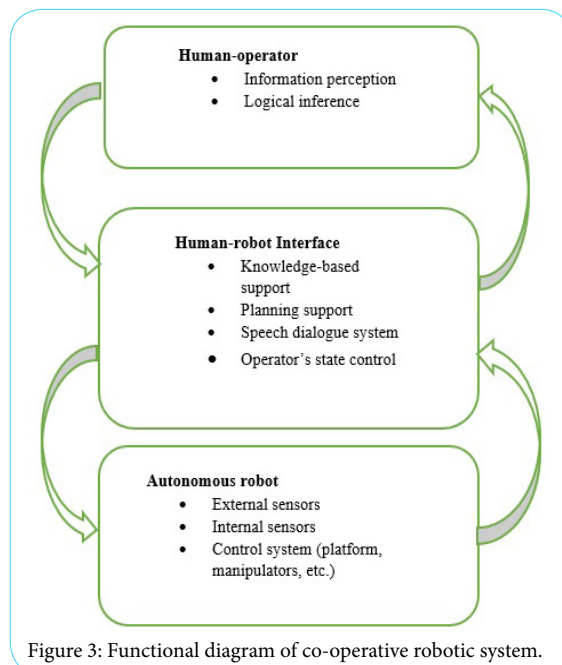


Figure 3: Functional diagram of co-operative robotic system.

The first task of HRI is presentation of a visual image of the situation as to operator as well to robotic control system. The image is possible to clarify in the speech human-robot dialogue if it necessary.

The human-robot interface also supports operator in planning operations in accordance with set tasks using the problem-oriented knowledge-based system. Another function of the system is control

of the operator's state himself. That aspect is crucial important for tasks of moving objects control. The human-robot interface function is also to support the human-robot speech dialogue including the blocks of speech recognition, speech synthesis and speech control. The system makes it possible to apply the speech command to control the robot and to form the feedback from robot to operator also in speech form.

The main functions of user-operator are perception of presented information to form an image of situation in the whole including as environment as the condition of the controlled robot. Another function of operator is the logical inference to form the corresponding instructions to robot. Result is a control action interpreted in speech or another possible form based on a logical inference.

To design the robotic system as before is necessary to choose the adequate mathematical apparatus to formalize all the elements of the functional diagram including human-operator. The problem arose is well known. The usual mathematical mediums such as differential equations etc. are not efficient to formalize the human-operator model. Suppose now we have an opportunity to solve the task via fuzzy logic approach. It is necessary to underline the importance of psychology aspects of HRI in cooperative robotic system/

### Fuzzy logic and HRI

To organize a speech dialogue between user and robot, natural the spatial-temporal relations can be used [2]. Remind that the introduced natural relations include fuzzy relations of the position and orientation of objects, which are defined by membership functions. These functions correspond to the perception of above relations by a person. The appropriate membership functions are the formalization of the corresponding psychophysiological scales that determine the capabilities of human receptors. To complete the description, the so-called intentional relations was added to such relations, such as to be inside, to be on the same plane, in touch one another, etc. Using the necessary means of observation (CVS, laser rangefinders, ultrasonic sensors of near location), as well as membership functions contained in the knowledge base, cooperative robot can describe the situation independently in terms of natural spatial relations and report it to the operator in a verbal form. The same relations used by operator prove to determine the tasks for the robot in the process of dialogue.

The current situation, including  $M$  objects and the robot itself, was described by a system of binary frames containing the fuzzy relations between all the objects of the scene: ( $\langle \text{object } m \rangle, \langle \text{relation} \rangle, \langle \text{object } n \rangle$ ),  $m, n = 1, 2, \dots, M$ . If the fuzzy binary relationships between all the objects that the robot can observe during movement was determined in advance, we got a fuzzy spatial-semantic network, or a fuzzy map. Now it is possible to navigate the robot along the observed benchmarks.

The combination of designations (names) of given objects in the space of the working scene and fuzzy relations between them consist a dictionary of the formal language to describe the situation. Using the terms introduced by D.A. Pospelov, one could name such language as situational one. It is a dialogue language between the cooperative robot and the user during their interaction.

The functional diagram of the dialogue control system presented on Figure 4.

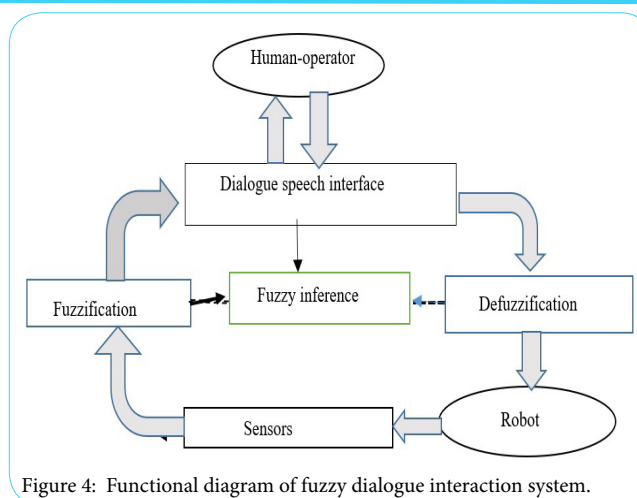


Figure 4: Functional diagram of fuzzy dialogue interaction system.

Notice that the diagram is differ radically from a usual control system because it does not include as an object under control as well as subject controller. The system includes two partners where the human operator is the master and the second partner the robot is an assistant. Both partners possess his (it's) own information system, base of knowledge and fuzzy logic inference system. As shown on Figure 4 the main way of dialogue control is the direct speech control by user on the base of a problem-oriented language. But it is possible to realize the autonomous fuzzy control on the base of the production rules partially known beforehand and partially formed during the user's work by artificial neuro network (dotted line). Such rules are compiled in the knowledge base of HRI in form "if the situation is  $S_i$  than behavior may be  $B_j$  with reward  $R_j$ ,  $j=1,2,\dots,k$ ". The choice depends of result of the elementary operation or their consequence for complicated operations planning. Such approach is usual for reinforcement learning, developed by R.S. Sutton and A.G. Barto [7].

The main part of HRI is a bilateral speech dialogue system consisting of speech recognition part, speech reproducing part and dialogue management subsystem. For the means of speech recognition, today it seems the most effective way to use the deep neural networks Deep Neural Network (DNN), which have greater capabilities than conventional neural networks. However, neural networks are not in position to simulate voice signals directly. Therefore, in order to use the ability of neural networks to classify, a combined DNN-HMM model, including DNN and the Hidden Markov Model (HMM) have been proposed [8]. The hidden Markov model have been used to describe the dynamics of voice signals. The probability of the observed vectors had been calculated by neural networks. This model has advantages in recognition accuracy over other methods for recognizing continuous speech. Neural network training is realized using the backward error propagation (BP) algorithm.

The dialogue process consists of the requests between the human user-master, making the request, and the assistant -robot, responding to the request dependently on the context of the discourse. The dialogue management system could create the following dialogue states according to its current state, or by the request of the subject. To solve the problems of monitoring the state of the dialogue and managing the dialogue, one could apply the technique of state automata [8]. The system recognizes user requests using the DNN-HMM algorithm and generates a response to the request. In practice, in an object-oriented dialogue system for controlling a robot, the

number of client input requests is limited (usually the number of requests is from 20 to 30). Therefore, it is advisable to develop in advance the main dialogue scenarios depending on the operations performed. For example, analysis of the current situation, the feasibility of the task, checking the status of the robot, etc. This simplifies the task of dialogue managing. It should be noted that the operator's capabilities are significantly expanded if it is possible to connect the dialogue system to the Internet, which contains the necessary service for speech synthesis.

To represent the image of the current situation in an interactive control system, it is possible to use the spatial-semantic representation of space. It includes a description of the objects of the working scene and the space-temporal relations between them as mentioned before [2], which could be the basis of a more general approach to the organization of dialogue management.

HRI for co-operative robot may also apply gesture [6,9]. It made it possible to simplify the human-robot interaction in situations when speech dialogue is impossible. Another possibility to make the interaction more natural and clear both for human-user and for robot is the supplement of the speech dialogue with the analysis of the emotional state of user as well demonstration of the conventional "face" of robot representing the state of environment, or the robotic system itself, or external danger etc. [5]

For facial expressions of human face describing a facial expression coding system was developed (FACS) [10]. The next investigations made it possible the automation of the process [11]. This technology declared a facial state in terms of typical action units (AU) which describe movement of distinctive facial points of a separate facial muscle. Now the mimic expression may be depicted at the human-like robot "face" or at the screen of the user's controller. Figure 5 a and Figure 5b) shows the results of human operator face analysis and Figure 5c) depicted the artificial face of human-like robot Alex with expression corresponding the emotional state of "surprise" [5].

It seems possible to interpret the AU technology also in natural linguistic variables. Such formalization will make it possible to present the "emotion state" of robot-assistant as a result of fuzzy logic inference on the base of the real situation evaluation. The reverse task determined the emotional state of user himself is also of interest mainly in extreme and dangerous situations.

### ICR as a surgeon assistant

As an example of interactive cooperative robotic system we present the robot-surgeon assistant [12]. The simplified diagram of the system presented on Figure 6. The aim of the system is to help the surgeon during the medical operation. It has to recognize the voice request to find and to give the surgeon the necessary instrument. The system contains manipulator KUKA LBR IIWA 14, CV system with a static camera and another miniature camera located on the last link of manipulator (Figure 7), image processing system and voice dialogue system. To control the position of operator we used stationary KINECT.

The surgeon asks to give him the necessary medical instrument. The command was proceeding by dialogue system including the speech recognition and speech reproduction subsystems realized as mention above as a combined DNN-HMM model, including Deep Neural Network and Hidden Markov Model. The system made it possible the speech surgeon-robot dialogue for the task clarification if necessary. The result of the voice recognition system is definition of the subject, being specified by surgeon. The next task of the interactive cooperation system was to find the necessary object on the surgeon's table using computer vision and image processing system. The last stage was to capture the desired object by manipulator and to transfer it to surgeon.

The problem of object recognition conditionally divided into several stages: the stage of searching of the object, the stage of object classification and finally the stage of the object position and orientation on the table. To save time we proposed to use a convolutional neural network consisting of two parallel channels of information processing [13]. One of the channels finds the allocation area of the desired object using a convolution neural network scanning the workspace with a sliding window (Figure 8). Another convolutional neural network solves the problem of distinguishing objects in each of the subdomains and classifying the objects. Each of the neural networks were trained separately by the back propagation method. The approach turned out to be quite effective when the system observes several different objects among which the desired object is located. When the position of the desired instrument on the table is determine the task of capture it by manipulator is not complicated.

The last task of ICR system is to find the surgeon himself, to detect the position of his hand using KINECT and to plan the necessary trajectory to give him the instrument without any damage.

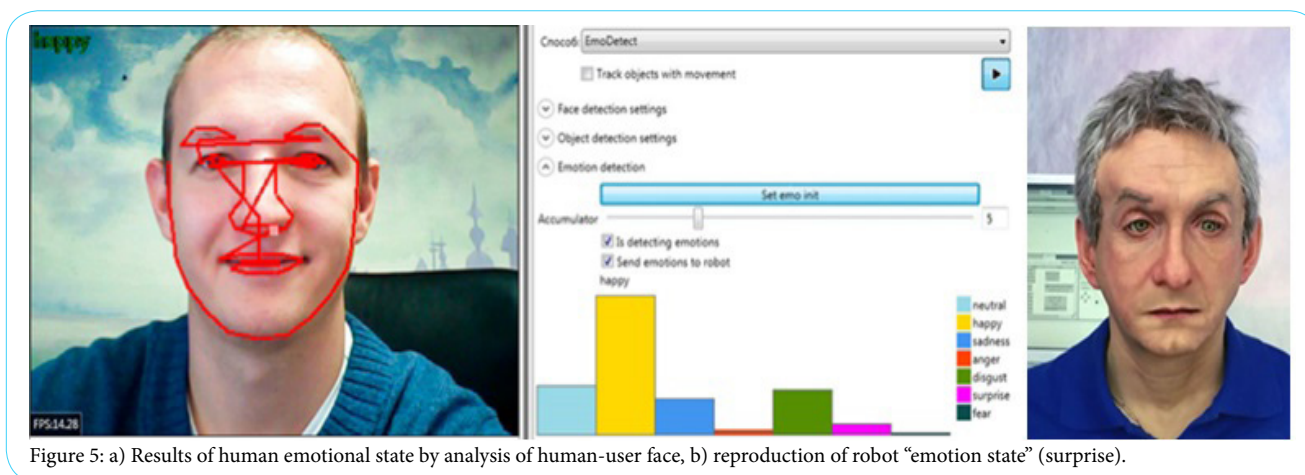
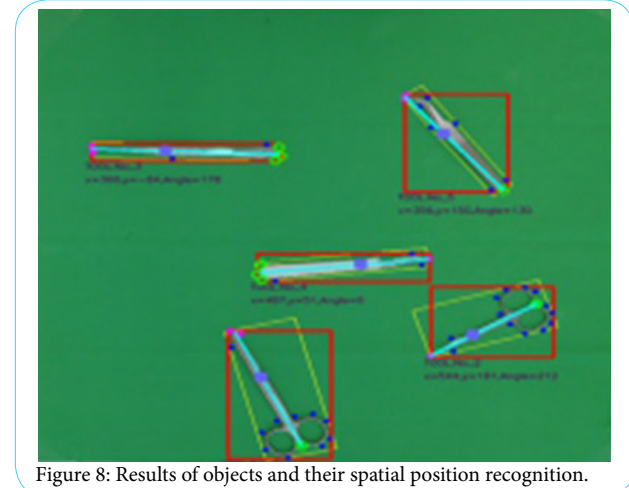
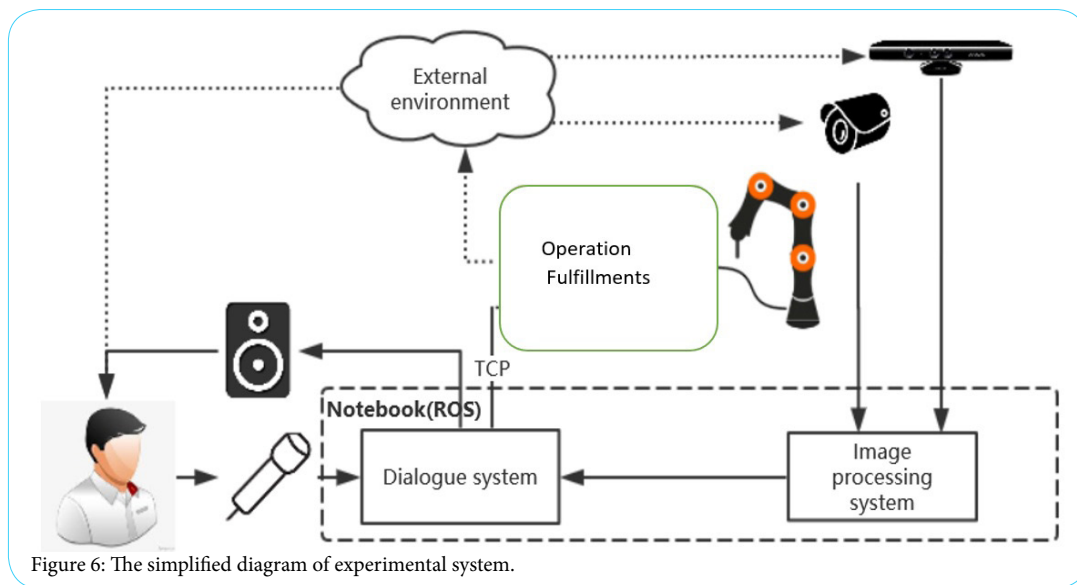


Figure 5: a) Results of human emotional state by analysis of human-user face, b) reproduction of robot "emotion state" (surprise).



Experiments showed the effectiveness of the proposed ICR [12]. We suppose that the structure of the system is general for wide class of cooperative manipulation systems.

### The use of augmented reality and intuitive programming

Another way of human interaction with a collaborative robot is the use of augmented reality at the stage of robot motions programming [14]. Currently, the most common way to program a Co-Robot is the direct capture of a working tool, held by the manipulator along the desired path, which is remembered in robot control system (Figure 9). It is much safer for operator to use for programming the augmented reality when operator sets desired movement without real manipulator activation. In that case he observes the desired motion of 3D manipulator model on a transparent screen together with real environment. This is achieved through the usage of transparent augmented reality glasses, being a miniature screen on which the movement of the manipulator model is displayed. At the same time, operator sees the real space in which the motion is planned. This method may be called as *programming a virtual Co-robot by demonstration*.



Such approach may be effective for operations performing in a limited space, or associated with the risk of damage the surrounding objects and manipulator itself. Operator can use his own experience and intuition to choose the safe and effective trajectory of robot movement. In this regard, this method is sometimes called an intuitive programming [15].

The structure of the intuitive programming system under discussion [16] is shown on Figure 10. The system consists of three main parts: an operator interface, a portable single-board computer and a manipulator with its own control system. Operator is equipped with augmented reality glasses that allow one to see both the surrounding space and the additional information, which is generated by the augmented reality unit. The programming tool is a freely settable operator AR (Augmented Reality) pointer ensured as position as orientation of the controlled object [17]. The position and orientation of the AR-pointer in space is detected at each moment of time by a miniature pinhole camera placed directly on augmented reality glasses. The continuous video signal of the camera is projected by AR software unit onto the portable screen, on the plane of the glasses, without distorting the operator's field of view. The portable optically transparent screen provides visibility of the work stage through the display. After processing the received information by a wireless communication channel in the AR unit, it is entered into the manipulator motion planning unit. The latter is interconnected with a database including a 3D model of the manipulator itself, its capture, working tool, and manipulation objects. Reference points and trajectories obtained during the training of the robot and presented in the manipulator's own coordinate system are also entered into the database. The desired current position is transmitted to the manipulator by the wireless module from the side of the AR system, through the controller of the manipulator. The feature of the proposed solution is the ability for the operator to move freely in the workspace and inspect the workspace from different points of view.

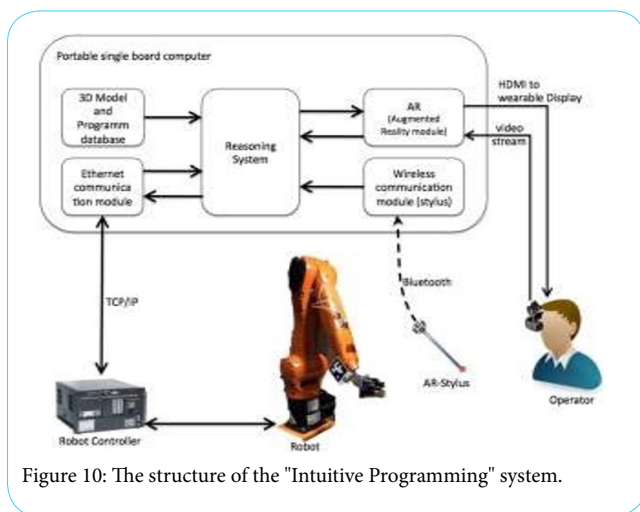


Figure 10: The structure of the "Intuitive Programming" system.

Unlike manipulator programming in off-line mode [14], there is no need to simulate the robotic cell itself in case we use the augmented reality. The visual 3D AR model of the manipulator's working tool or another manipulation object in the operator's field of view is superimposed on the image of the real working environment on the display screen. Due to this, operator is able to consider the possibility of a collision between the manipulator and environmental objects, and to ensure the desired orientation to capture the object if necessary. Experimental studies have confirmed the effectiveness of the proposed system for programming the movement of an industrial manipulation robot [16].

The approach under discussion prove operator to fulfil the stage of robot training in augmented reality. The same as in the example of robot-assistant above human operator stays the task for robot and

then controls the real movement with possibility to stop and reprogramming. But now the movement has to be determined directly in the robot working space. Such situation is convenient for industrial manipulation robots for assembling also for welding and cutting with manipulation robots. It seems of interest to supplement the AR with speech commands and bilateral speech dialogue as we discussed above. We suppose that such approach may be useful also for manipulators programming in previously undetermined situations after accidents and disasters.

It is necessary to underline that the human-robot interaction in cooperative systems is new and rather complicated mode of activity for human-operator. Perhaps the main obstacle for wide application of such systems would be not technical, but ergonomic and psychological problems of interaction in cooperative robotics [18].

## Conclusion

The emergence of collaborative robotics was a natural development of production systems that included robots in the direct working process. In essence, collaborative robotics today is a new stage in the development of technogeneous human environment, which is characteristic of the 4th industrial revolution, i.e. for a stage defined commonly as Industry 4.0. Given the diversity of human activities, which now includes robots, let us assume that we are at the stage in the development of our society when robots become its integral part. Some problems are arisen. First of all, the problems of human trust to such systems, especially when performing critical tasks. Starting from driving trains and ending with performing surgical operations associated with immediate risk to the patient. This problem could be solved only by achieving a greater anthropomorphism of interactive robots, in the ways of interacting with a person, ways of thinking, assessing the situation, making decisions. And on the other hand, a person should be ready to interact with an artificial partner, to understand his real opportunities and limitations. This may require some training and also become one of the aspects of the education of any member of society in the near future. In all cases, the robot would remain the executor of human will, other than slave, but a friendly partner.

## Competing Interests

The author declare that he has no competing interests.

## References

1. Zadeh LA (1997) Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic. *Fuzzy sets and systems* 90: 111-127.
2. Kandrashina EYu, Litvinceva LV, Pospelov DA (1989) Knowledge representation of time and space in intelligence systems, Nauka, Moscow.
3. Gradetsky I, Ermolov I, Knyazkov M, Semenov E, Sukhanov A (2018) The Dynamic Model of Operator-Exoskeleton, Proceedings of the 3rd International Conference on Interactive Collaborative Robotics (ICR), Switzerland AG, Printforce, Netherland, pp. 52-59.
4. Yuschenko AS (2015) Dialogue Mode of Robot Control on the Base of Fuzzy Logic. *Transaction of the Conference "Extreme Robotics"(ER)*, S-Pb, Politechnika service, p.p. 143-146.
5. Konyshv DV, Vorotnikov SA, Yuschenko AS, Zhonin AA (2016) Mimic Recognition and Reproduction in Bilateral Human-Robot Speech Communication. *Proceedings of the First International Conference ICR*, 2016, Budapest, Hungary, Springer, LNAI 9812, pp. 133-142.

6. Rogalla O, Ehrenmann M, Zöllner R, Becher R, Dillmann R (2002) Using gesture and speech control for commanding a robot assistant. Proceedings of the 11th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN), Berlin, Germany, p.p. 454-459.
7. Sutton RS, Barto AG (2018) Reinforcement Learning: An introduction. Cambridge, MIT Press, 548 p.
8. Shuai Y, Yuschenko AS (2019) Dialogue control system for a robot based on the theory of finite automata. *J Mechatronics, Automation, Management* 20: 686-695.
9. Gruber I, Ryumin D, Hruz M, Karpov A (2018) Sign Language Numerical Gestures Recognition Using Convolutional Neural Network. Proceedings of the 3-rd International Conference ICR, Springer, LNAI 11097, Netherland, pp. 70-77.
10. Ekman P, Rosenberg EL (1997) What the face reveals: basic and applied studies of spontaneous expression using the Facial Action Coding System (FACS). New York, Oxford Univ. Press.
11. Cohn JF, Zlochower AJ, Lien J, Kanade T (1999) Automated face analysis feature point tracking has high concurrent validity with FACS manual coding. *J Psychophysiology* 36: 35-43.
12. Yin S, Yuschenko A (2019) Application of Convolutional Neural Network to Organize the Work of Collaborative Robot as a Surgeon Assistant. In: Ronzhin A, Rigoll G, Meshcheryakov R (eds) *Interactive Collaborative Robotics. ICR 2019. Lecture Notes in Computer Science*, vol 11659. Springer, Cham.
13. Yin S, Yuschenko AS (2020) Object Recognition of the Robotic System with Using a Parallel Convolutional Neural Network. In: Kravets A (eds) *Robotics: Industry 4.0 Issues & New Intelligent Control Paradigms. Studies in Systems, Decision and Control*, vol 272. Springer, Cham.
14. Pettersen T, Pretlove J, Skourup C, Engedal T, Lokstad T (2003) Augmented reality for programming industrial robots," *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality*, 2003. Proceedings, pp 319-320.
15. Akan B, Ameri A, Curuklu B, Asplund L (2011) Intuitive industrial robot programming through incremental multimodal language and augmented reality," *2011 IEEE International Conference on Robotics and Automation*, 2011, pp. 3934-3939.
16. Schwandt A, Yuschenko A (2019) Robot Manipulator Programming Interface based on Augmeed Reality. *International Journal of Recent Technology and Engineering (IJRTE)* 8: 819-824.
17. Wu PC, Wang R, Kin K, Twigg, Han S, Yang MH, et al. (2017) Accurate 6Dof Tracking of a Passive Stylus. *Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*, pp. 365-374.
18. Yuschenko A (2019) Ergonomic problems of collaborative robotics. *Robotics and Technical Cybernetics* 7: 85-93.