The Application of Artificial Intelligence in Indoor Positioning and Scene Reconstruction of Robots

Dan Ye[1,*] and Mei Gao[2]

[1]The Department of Electronic and Electrical Engineering, Anhui Wenda University of Information Engineering, Hefei, Anhui 231201, China, yedan@mail.wenda.edu.cn
[2]The Department of Electronic and Electrical Engineering, Anhui Wenda University of Information Engineering, Hefei, Anhui 231201, China, gaomei@mail.wenda.edu.cn

Corresponding author: Dan Ye, yedan@mail.wenda.edu.cn

Abstract. The powerful functions of robots in life and production have been proven. It can replace humans to perform some complex and dangerous tasks, and it can also perform related tasks in some dangerous areas. Traditional robot systems rely on GPS systems for information transmission and command sending and receiving. People can monitor the status of the robot's task execution in real time through the computer's host computer system. However, for some relatively closed areas or areas with poor signal, this limits the use of the robot. With the continuous development of sensor technology and visual positioning system, robots have begun to develop towards autonomous positioning and scene reconstruction. This research uses the CNN algorithm of artificial intelligence theory and the GRU method to design an intelligent recognition system, which can also assist the computer system to monitor the task execution process of the robot. It can also monitor the robot's autonomous positioning and scene reconstruction information. The research results found that CNN and GRU technology can assist the robot to identify the characteristics of indoor related objects, and the maximum error rate is only 2.439%. This part of the error is reliable enough for the robot to complete the tasks of autonomous positioning and scene reconstruction.

Keywords: Robot autonomous positioning; Scene reconstruction; Artificial intelligence; Computer-aided technology.
DOI: https://doi.org/10.14733/cadaps.2023.S5.136-145

1 INTRODUCTION

The traditional robot will use the GPS method for positioning. After it is positioned by the GPS, it can carry out the relevant tasks according to the relevant instructions through the system it carries. Robot technology has changed people's life and production methods. It can replace humans to perform some assembly-line engineering operations, and it can also perform some more dangerous tasks [1]. Robot technology has high work efficiency. It does not require rest like manual methods, as long as the robot is regularly maintained and maintained. It can continue to perform some tasks efficiently. Traditional robots can receive instructions through GPS signals to complete tasks related to the path [2]. However, for a relatively closed area, this will limit the robot to receive relevant signals, which will limit the robot to receive positioning commands and motion commands. In this case, the robot is likely to lose the efficiency function of the related work. With the continuous development of vision technology and sensors, vision-based navigation and positioning technology has also undergone many changes. The robot may allow itself to be positioned using the sensors it carries. Once the robot has the function of positioning and scene reconstruction in a relatively closed area, it can be competent for tasks in different terrains or areas. This continues to enrich the use environment of robots. However, the robot's own positioning information or scene effects also need manual assistance to perform tasks.

If the computer-aided system is applied to the robot's positioning and scene reconstruction system, it can check the positioning task completed by the robot and the effect of scene reconstruction in real time, which can better complete the task of the relevant closed area through manual assistance. If only relying on the robot's own positioning and scene reconstruction functions to perform related tasks, it will make it impossible for humans to get relevant feedback. This makes it impossible to know how well or efficiently the robot performs the task. The robot's own positioning technology and scene reconstruction function are generally applied in relatively closed or inaccessible environments, which will reflect the powerful functions of the robot. The application of computer-aided systems in the robot's own positioning and scene reconstruction can be divided into functions such as scene visualization and guidance for reconstruction and positioning [3]. It can send the robot's positioning data or reconstructed scene data to the computer-aided system through related technologies. The three-dimensional effect visualization system built by the computer-aided system can monitor the relevant positioning data and scene data in real time. The relevant data of the robot obtained by the human through the computer-aided system is also the basis for issuing the relevant task instructions.

In the process of the robot relying on multi-sensor positioning, it is necessary to identify the characteristics and positions of the relevant scenes, which requires the use of artificial intelligence theory. Some algorithms in artificial intelligence theory can learn the characteristics of related objects to assist the robot system to complete the modeling work. If the relevant theory of artificial intelligence is not used, this will limit the recognition of relevant objects in the scene by the robot, and it will be difficult to establish a relevant model of the scene. Therefore, the combination of artificial intelligence technology and computer-aided system is necessary for the autonomous modeling and scene reconstruction of robots. The intelligent algorithm can judge the attributes and data characteristics of objects in the scene, which can quickly assist in establishing the location information of the indoor scene model. This research mainly uses intelligent algorithms to complete the indoor positioning and scene modeling of robots, which is also an application of closed scene positioning and modeling.

2 RELATED STUDIES

Cao et al. [4] believed that the completeness and accuracy of the robot to complete the 3D scene reconstruction task is more important. An accurate scene reconstruction scheme will capture the complete coverage of the scene where the robot is located. It uses the method of drones and robots to design an indoor pipeline scene reconstruction scheme. The aerial images will automatically complete the process of collecting and fusing indoor scene images. The results of the study also showed the accuracy of this scheme. Hu et al. [5] has also considered that the path planning and scene reconstruction faced by medical service robots are an important part of autonomous navigation. It utilizes 3D point cloud and single-camera techniques to evaluate depth estimation and pose tracking tasks for medical robots. The research results show that point cloud image reconstruction technology can reduce the acquisition of non-uniform data during scene reconstruction. Wu et al. [6] also believes that scene reconstruction is a necessary process in the field of robot vision. It requires the robot to complete the image acquisition at different locations. It uses an improved 3D scene reconstruction method to overcome the difficulty of frame sequence conversion during scene reconstruction. It proposes fast point feature map and iterative closest point method to solve this problem by using depth image matching method. The results of the study found that this method has high value. Tong [7] has also discovered the role of binocular recognition technology in autonomous robot positioning and scene reconstruction. It uses heterogeneous binocular recognition technology and multiple viewpoints to study the scene reconstruction task of robots. This method can obtain relevant scene image information from multiple viewpoints. It combines the related images of soccer robots to analyze the heterogeneous binocular recognition technology. Garcia-Salguero et al. [8] used RGB cameras to estimate the position and orientation information in the environment, which can provide support for the robot to complete the scene reconstruction. It also completes the collection and recognition of scene images based on CNN technology. The findings demonstrate the robustness of this approach in scene re-addition.

This research will use artificial intelligence theory and computer-aided system to solve the function of autonomous positioning technology and scene reconstruction of robots, which solves the shortcomings of traditional robots that are difficult to achieve autonomous positioning and path planning in closed areas. This article analyzes the effects of intelligent algorithms and computer-aided technologies in autonomous robot positioning and scene reconstruction from different perspectives. Section 1 explains the background of robot development and the significance of the application of computer-aided technology and artificial intelligence skills in robots by means of investigation. The related status of robot autonomous localization and scene reconstruction is illustrated in Section 2. The process of CNN and GRU technology for identifying relevant features of indoor scenes and the process of computer-aided systems in autonomous localization and scene reconstruction are described in Section 3. Section 4 illustrates the accuracy of CNN and GRU in recognizing indoor scene features and positioning features through different statistical parameters. Finally, it summarizes the performance in two related ways for robotic systems.

3 A SCHEME OF ARTIFICIAL INTELLIGENCE AND COMPUTER-AIDED THEORY IN ROBOT AUTONOMOUS POSITIONING AND SCENE RECONSTRUCTION

3.1 The scheme of Intelligent Positioning and Scene Reconstruction and the Introduction of CNN

Taking into account the relationship between the autonomous positioning of the robot and the related features of scene reconstruction, it uses the convolutional neural network CNN and GRU methods in artificial intelligence technology to identify and evaluate the positioning information and object features in the scene. The positioning and reconstruction functions of indoor scenes will be designed with information such as indoor coordinates, the location of scene objects, and the shape of objects [9]. The coordinate data of indoor objects is the key for the robot to identify the positioning system. It needs to use its own identification and modeling system to update and establish the position information. The position of indoor objects not only has a certain spatial relationship, but also changes in a relatively small area over time. The shape information of indoor research objects also has a strong relationship with time [10]. Therefore, it not only needs to apply CNN technology to autonomous positioning and intelligent scene recognition of robots, but also needs to consider the application of GRU in robot systems.

Through the above analysis, we can find that the coordinate information of the indoor scene, the relative position information of the object and the shape information of the object are the key factors for the robot to locate and reconstruct the scene. First, it needs to take into account the spatial characteristics and complex relationships of these related information. The sensor technology carried by the robot itself cannot complete the function of recognition and feedback, it can only complete the function of recognition and visual display. Therefore, it is necessary to use the CNN technology in the intelligent algorithm to complete the identification and evaluation of spatial information. CNN is a relatively mature and simple spatial feature recognition algorithm. For the autonomous positioning and scene reconstruction tasks of robots, CNN technology is sufficient for this task. Many successful cases also verify the expressiveness of CNN in spatial information recognition. CNN can not only handle the complex relationship of data, but the amount of data it can handle is allowed to be large, which is due to the role of the weight sharing mechanism of CNN.

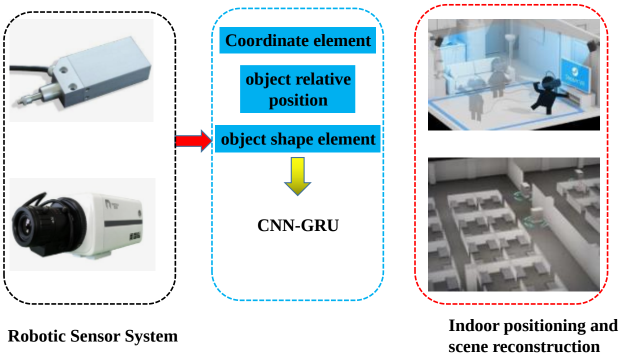

Figure 1: Schemes of CNN-LSTM and computer-aided methods in robot autonomous localization and scene reconstruction.

Figure 1 introduces the related theory of CNN-GRU and the role of computer-aided technology in robot indoor positioning and scene reconstruction. Researchers first need to train relevant theoretical algorithms of artificial intelligence, which requires collecting relevant data sets as support for feature learning. At the same time, it also ensures the accuracy of scene object features. The computer-aided system carried by the robot will learn from the reconstructed scene features. Once it learns the relationship between the features and the scene data, it can identify the features of related objects from unfamiliar scenes. From Figure 1, it can be seen that the large-scale data of robot positioning and scene reconstruction will go through the CNN and GRU algorithms. During the testing and application stage, the data in the test set will not go through the CNN or GRU algorithms, and these data will be directly related to the trained ones. Weights are calculated and evaluated accordingly. When obtaining the optimal weight distribution for robot positioning and scene reconstruction, it needs to go through a large number of iterations and training processes. Once the training phase of the CNN and GRU is completed, this AI-assisted computer system is an efficient recognition scheme.

3.2 Indoor Positioning of Robots and Identification of Time Factors for Scene Reconstruction

For indoor positioning and scene reconstruction tasks, indoor coordinate information, the relative position of objects in the scene, and the shape of scene objects are all related to time. At the same time, in the process of the robot performing tasks, the three elements will change with time, which requires full consideration of the time relationship between the relative position elements of the scene objects and the shape elements of the scene objects. If the temporal characteristics of the relevant elements of the scene are not considered, the intelligent strategy will map the existing temporal relationship, which will cause certain inaccuracies.

For the identification and evaluation of time-related factors, most researchers will use LSTM technology to process, which is an earlier and more mature algorithm for processing time information. However, the four-gate structure of LSTM will increase the amount of parameter computation, which puts forward higher demands on the computing power and storage capacity of the computer. These data are continuously collected when the robot performs indoor positioning and scene reconstruction tasks. The amount of data in this process is huge, which will show the update and calculation of the data in the robot's computer system. The GRU method can allow more datasets to be added than the LSTM technique while maintaining accuracy. This is because the gate structure of GRU has been reduced to 2, but it has been proved that the evaluation accuracy of GRU can be guaranteed.

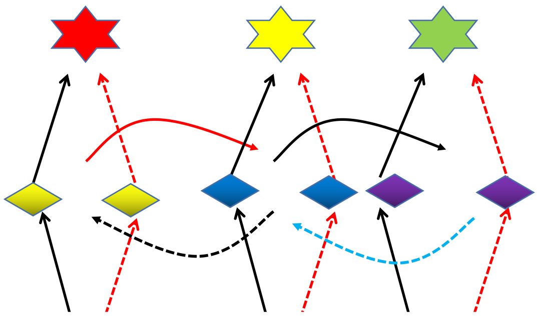

Figure 2: Extraction process of indoor positioning and scene reconstruction time factor.

Figure 2 shows the difference between the LSTM structure and the GRU structure in identifying elements related to robot indoor localization and scene reconstruction. The LSTM strategy is a more detailed strategy, which isolates the filtering of historical information from the input of new time information. The GRU theoretical strategy combines the functions of the historical information filtering of the LSTM strategy and the input of new time information through the reset gate. Such a reset gate performing two tasks already saves a lot of time, and it also reduces the repetition of parameter calculations. In addition, the weight update of LSTM and the output of combined time information are also divided into two different gate structures, while GRU is the output function of the weight update process and combined time information through an update gate. Compared with the four-gate structure, the two-gate structure can reduce a large number of parameters involved in the related calculation.

3.3 A Note on Computer-Aided Technology and Intelligent Strategies

The computer aided system is a necessary equipment for the robot to complete the positioning and scene reconstruction. It can not only display the information that has been positioned and the scene information that has been reconstructed, but it can also use manual intervention to correct relatively inaccurate information. Computer-aided systems also carry relevant intelligent strategies, which are also a key part of processing data collected by robotic multi-sensor systems. When the information collected by the sensor system carried by the robot will be aggregated into the computer-aided system, the computer-aided technology will use data processing strategies and intelligent strategies to learn the relationship between these positioning elements and scene reconstruction elements. Therefore, the computer-aided system plays the role of visualization and calculation in the robot indoor positioning and scene reconstruction system designed in this study.

The intelligent strategy CNN-GRU is the core part of the robot to complete the indoor positioning and scene reconstruction, and the data required for the intelligence roughly comes from the multi-sensor system. These sensors include sensors for strategic distance, cameras for capturing images, and sensors for more measurements. These sensor systems will be combined, and the measured indoor-related data will be fed into a computer-aided system.

4 RESULTS AND DISCUSSION SECTION

This research will use the CNN and GRU theory in the field of artificial intelligence to evaluate the three indicators of the coordinate position information of the robot indoor positioning and scene reconstruction, the position information of the scene objects, and the shape information of the scene objects. The coordinate position information is the main data basis for the robot to complete the indoor positioning function. If the CNN and GRU technology can effectively evaluate the coordinate position information of the environment where the robot is located, it can accurately obtain the indoor coordinate information, which is beneficial for the robot to complete the positioning information. The position information of scene objects and the shape information of scene objects are necessary factors for robots to complete scene reconstruction. Only when CNN and GRU technology can effectively complete the evaluation of the position and shape of scene objects, it can accurately complete the task of indoor scene reconstruction. It can be seen from Figure 3 that the errors of the three factors of indoor positioning and reconstruction of the robot can meet the requirements. The evaluation error of the coordinate position information is the smallest, which is only 1.783%, and the error has not exceeded 2%. The errors of the position information of the scene objects and the shape information of the scene objects are only 2.191% and 2.439%.

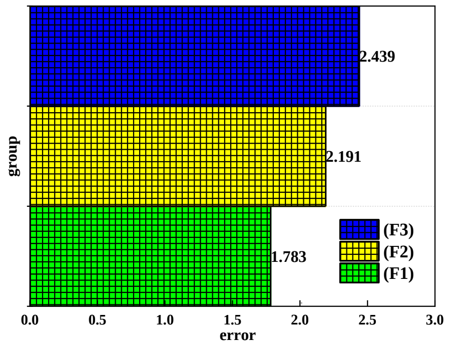

Figure 3: Evaluation errors of factors related to robot indoor positioning and scene reconstruction.

Although robot indoor positioning and scene reconstruction is a continuous system, it is mainly divided into two processes: positioning and scene reconstruction. First, it uses the evaluation effect curve to demonstrate the performance of CNN and GRU related technologies in evaluating the coordinate information of the robot room. It also compares the difference between the CNN--GRU system and a single CNN system in evaluating the coordinate information for indoor positioning of robots. Figure 4 shows the evaluation curve of the coordinate information obtained by using two different artificial intelligence strategies and the evaluation curve of the actual coordinate information. It can be seen from Figure 4 that there are many peaks and troughs in the coordinate information of the indoor positioning of the robot, which means that the coordinate information of the indoor positioning of the robot is relatively complex, which is mainly due to the indoor layout, which may be a kind of complex Shaped interior layout. However, both the CNN-GRU strategy and the CNN strategy can successfully evaluate the changing trend of coordinate information. The trend of the coordinate information evaluated by the CNN-GRU strategy is closer to the actual coordinate information curve, which shows that the CNN-GRU method is more suitable for evaluating the coordinate and position information factors for indoor positioning of robots.

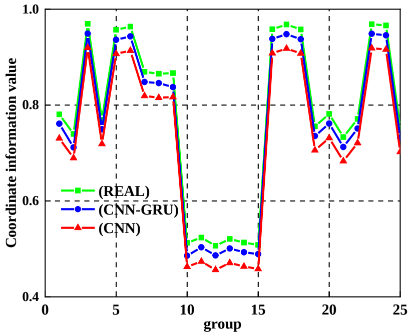

Figure 4: Evaluation efficiency of coordinate location information for robot indoor positioning using two intelligent strategies.

To complete the indoor scene reconstruction task, the robot needs to identify the relative position information of indoor objects and the shape information of the objects. Only by accurately knowing the actual position information of different indoor objects and the location of the occupied area, can the scene occupied by different objects be efficiently reconstructed, which is also the first step of the indoor scene reconstruction task. Combining the above studies, it only utilizes the more accurate CNN-GRU strategy when exploring the relative position information of objects reconstructed in robotic scenes. Figure 5 shows the prediction error distribution of the relative position information of indoor objects. Figure 5 shows that the CNN-GRU theory has successfully evaluated and identified different indoor locations, the relative position information of different objects, and the location of the area occupied by the objects. The errors of the position information of some objects selected in this study are mostly distributed between 1.8% and 2.4%. This is a considerable error distribution area relative to the acceptable error range of 5%. There are also a small number of objects whose relative position error is less than 1.5%, which may be a relatively fixed object for a long time.

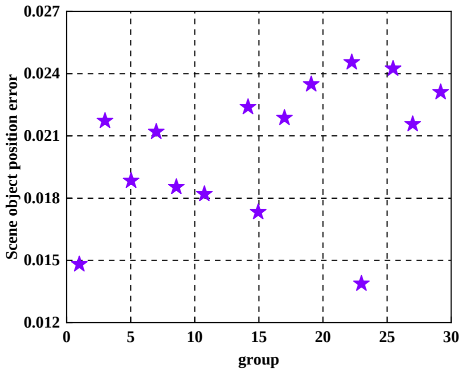

Figure 5: Error distribution of evaluation effect of relative position information of indoor objects.

When the robot completes the indoor scene reconstruction task, it needs to know the shape information of each object in the room, which can assist the computer-aided system of the robot to complete the reconstruction task of the shape information such as the outline and size of the indoor object. Figure 6 demonstrates the performance of the CNN-GRU strategy in evaluating indoor object shapes in the form of predicted distribution curves. The shape information of an indoor object includes information such as the size and outline of the object. It can be seen from Figure 6 that the CNN-GRU strategy can accurately reconstruct and identify the shape information of objects with more fluctuations, which lays a good foundation for the robot to complete the reconstruction of indoor scenes. However, it can also be found from Figure 6 that the values of object shapes in indoor scenes are quite different, which indicates that the selected test set may be a relatively dense indoor area, where there will be more changes in object shape information. This further demonstrates the reliability of CNN-GRU.

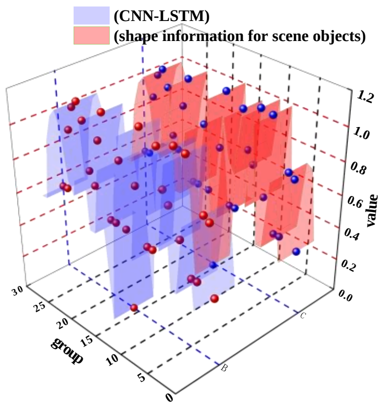

Figure 6: Evaluation effect of object shape information in indoor scenes.

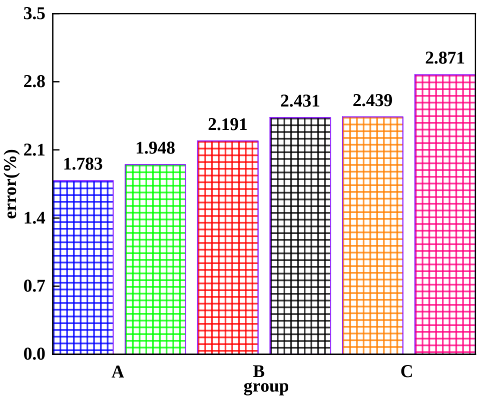

Figure 7: Global performance differences in robotic indoor localization and reconstruction utilizing different intelligence strategies.

Figure 7 shows the effect of two intelligent strategies in evaluating factors related to robot indoor positioning and scene reconstruction in the form of global error distribution. These two intelligent strategies are mainly the CNN-GRU and CNN strategies mentioned above. Here is the strength of the temporal relationship between the three factors used to illustrate indoor localization and scene reconstruction. From Figure 7, it can be clearly seen that the error between indoor positioning and scene reconstruction evaluated by the CNN strategy is high. This shows that the three elements of robot indoor positioning and scene reconstruction are closely related to time. For the coordinate location elements of indoor positioning, the evaluation error has increased from 1.783% to 1.948%. However, the global error of object shape elements for indoor scene reconstruction increases from 2.439% to 2.871%. This fully reveals that the relationship between object shape information and time for indoor scene reconstruction is high.

5 CONCLUSION

Robots have been able to replace people to perform tasks that are more difficult and risky, and it has also demonstrated high work efficiency. In general, after the computer system issues a work order to the robot system, it will execute it autonomously. For robotic systems, localization, path planning, and scene recognition are key factors in performing tasks. Traditional robots rely on GPS for positioning. In this positioning method, people can control the robot in real time and control the robot to perform different action tasks. However, for relatively closed areas, the GPS navigation method loses its corresponding effect, which requires the robot to locate and reconstruct the scene autonomously. As a result of this research, the CNN-GRU intelligent strategy and computer-aided system in the field of artificial intelligence have designed and identified the coordinate position information elements, scene object position elements and scene object shape elements for robot indoor positioning and scene reconstruction. The research results found that the CNN-GRU strategy has higher performance than the CNN strategy in terms of identifying the three elements of localization and scene reconstruction. Through investigation, it can also be found that the shape elements of objects in indoor scenes are closely related to time elements, which is information that cannot be ignored.

6 ACKNOWLEDGEMENTS

This work was supported by the Academic funding project of top-notch in Disciplines or Majors in Universities in 2021 (No. gxbjZD2021092); The Academic funding project of top-notch in Disciplines or Majors in Universities in 2020 (No. gxbjZD2020099); The research project of Anhui Wenda University of Information Engineering in 2018: Research on Speech Signal PCM Coding Technology based on Compressed sensing (No. XZR2018A13).

Dan Ye, https://orcid.org/0000-0003-0360-6596
Mei Gao, https://orcid.org/0000-0002-7965-698X

REFERENCES

[1] Reyes-Acosta, A.-V.; Lopez-Juarez, I.; Osorio-Comparan, R.: 3D pipe reconstruction employing video information from mobile robots, Applied soft computing, 75(1), 2019, 562-574. https://doi.org/10.1016/j.asoc.2018.11.016

[2] Wang, L.; Li, R.-F.; Shi, H.-Z.: Multi-Channel Convolutional Neural Network Based 3D Object Detection for Indoor Robot Environmental Perception, Sensors, 19(4), 2019, 893. https://doi.org/10.3390/s19040893

[3] Gong, L.; Wang, W.-J.; Wang, T.: Robotic harvesting of the occluded fruits with a precise shape and position reconstruction approach, Journal of field robotics, 39(1), 2022, 69-84. https://doi.org/10.1002/rob.22041

[4] Cao, X.; Zhu, L.-J.; Cui, H.-N.: Complete and Accurate Indoor Scene Capturing and Reconstruction Using a Drone and a Robot, IEEE sensors journal, 21(10), 2021, 11858-11869. https://doi.org/10.1109/JSEN.2020.3024702
[5] Hu, Y.-Z.; Wang, Y.-J.; Wang, S.: Fusion Key Frame Image Confidence Assessment of the Medical Service Robot Whole Scene Reconstruction, Journal of imaging science and technology, 65(3), 2018, 030409. https://doi.org/10.2352/J.ImagingSci.Technol.2021.65.3.030409
[6] Wu, P.; Li, W.; Yan, M.: 3D scene reconstruction based on improved ICP algorithm, Microprocessors and microsystems, 75(6), 2020, 103064. https://doi.org/10.1016/j.micpro.2020.103064
[7] Tong, C.-Y.: Three-dimensional reconstruction of the dribble track of soccer robot based on heterogeneous binocular vision, Journal of ambient intelligence and humanized computing, 11(12), 2020, 6361-6372. https://doi.org/10.1007/s12652-020-02039-2.
[8] Garcia-Salguero, M.; Gonzalez-Jimenez, J.; Moreno, F.-A.: Human 3D Pose Estimation with a Tilting Camera for Social Mobile Robot Interaction, Sensors, 19(22), 2019, 4943. https://doi.org/10.3390/s19224943
[9] Farsoni, S.; Rizzi, J.; Ufondu, G.-N.: Planning Collision-Free Robot Motions in a Human-Robot Shared Workspace via Mixed Reality and Sensor-Fusion Skeleton Tracking, Electronics, 11(15), 2022, 2407. https://doi.org/10.3390/electronics11152407.
[10] Esper, I.; Smolkin, O.; Manko, M.: Evaluation of RGB-D Multi-Camera Pose Estimation for 3D Reconstruction, Applied sciences-basel, 12(9), 2022, 4234. https://doi.org/10.3390/app12094134