Construction of Autonomous System of Educational Robot Under the Background of Artificial Intelligence

Taking the Mobile Educational Robots toward Performing Intelligent Invigilation as an Example

Shuya Zhang[1], Jingwen Luo[1*]

[1]*School of Information Science and Technology, Yunnan Normal University, Kunming, Yunnan, China*
{Shuya Zhang, Jingwen Luo}857561945@qq.com, *Corresponding author: by1503117@buaa.edu.cn*

Abstract

At present, it is necessary to improve the ability of robots to perform various educational services; however, in the process of applying robots to educational scenes, the examination scenes are often to be ignored. As The Mobile Educational Robot (MER) performing intelligent invigilation tasks often encounters unknown behaviours of examinees or different examination environments and must face complex uncertainties, this paper comprehensively utilizes various advanced technologies in the field of artificial intelligence (AI) to construct an MER autonomous system that performs intelligent proctoring, in which the three-dimensional simultaneous localization and mapping (3D-SLAM) is the key supporting technology for realizing MER's autonomous navigation and path planning, supplemented by deep learning algorithms, big data, and natural language processing (NLP) technologies to construct the intelligent perception analysis system of the examination room, accordingly, to realize the relevant functions of intelligent invigilation in the examination room, further, comprehensively improving the autonomy of MER in different test rooms and the competence of its services.

Keywords: *mobile educational robot (MER), intelligent invigilation, autonomous navigation, simultaneous localization and mapping (SLAM), intelligent perception analysis*

1 INTRODUCTION

Both "Made in China 2025" [11] and "New Generation of Artificial Intelligence Development Plan" [12] emphasize the importance of developing intelligent education, centring on service robots, actively research new products to meet different situational requirements and promote the standardization and modularized development of robotic technology, and further to expand its field of practical applications. With the rapid development and wide application of artificial intelligence (AI) and robotics, people's demand for intelligent robots with autonomous navigation and autonomous operational capabilities in the field of education has also increased. In future, the development of educational robots will focus on specific application scenes, also, pay more attention to its function and value [4]. In the five educational scenes of teaching, learning, examination, evaluation and management, educational robots are often employed in the service of the other four scenes, but the examination scene is often overlooked. Nowadays, From the examination of basic education to higher education, and then to the vocational qualification examinations in the society, the number of examinees, the times of examination and the degrees of strictness are constantly increasing, additionally, the forms of examinations are more diversified. In many cases, the traditional invigilation mode easily leads to lethargy or boredom in invigilator, so that instances of academic dishonesty cannot be determined in a timely and accurate manner, thereby affecting the seriousness of examinations and the fairness of talent selection.

Kohlberg divided human moral development into "three levels and six stages" in his theory of the Moral Development Theory. Most of the adolescents and adults in China are at the second level, that is, the conventional level. They can understand relevant social behaviours and realize that they should abide by and implement the relevant codes of conduct in line with public opinion. General conscious compliance with norms may be the

B. Fox et al. (Eds.): IC-ICAIE 2022, AHCS 9, pp. 810-816, 2023.
https://doi.org/10.2991/978-94-6463-040-4_123

basic condition for the implementation of robot supervision. Mubin et al., [8] discuss whether robots can prevent students from cheating, and the comparison test concluded: Robots can effectively prevent students from cheating, and proposed that classroom logistics, settings and dynamics need to be considered, so that the robot can autonomously navigate and monitor more students. This viewpoint provides a certain theoretical and practical basis for the research of this paper.

2 RESEARCH STATUS AND CHALLENGING PROBLEMS

In recent years, there have been some relevant cases of intelligent invigilation robots used for invigilation services, e.g. In 2014, an invigilation robot named “Test Assistant” invented by four students of Jiangxi University of Science and Technology, it has the functions required for general invigilator services and runs on batteries. Subsequently, although the "Future Teacher" educational robot that came out in 2015 can assist teachers' management work, it unable to accurately plan its patrolling paths in the face of unknown examination environments, and cannot detect the irregular behaviours of examinees in time. With further development, a robot, AI-MATHS, was developed by Chengdu Zhunxingyunxue Technology, sits for the math test of China's national college entrance exam in Chengdu, June 7, 2017. Furthermore, ATA Intelligent Proctor (AI Proctor), appeared in 2018, is mainly used in classrooms and standardized examination environments to conduct real-time analysis of the examinee’s skeletal behaviour and effectively track abnormal behaviour in the examination room, nevertheless, it incapable of autonomously moving in the examination room to serve examinees.

Based on the above analysis of the current research status of invigilation robots, it can be easily determined that intelligent invigilation robots have been applied within a certain range, but they have not been popularized. The reasons can be summed up as three challenges: 1) poor autonomous and lack of the autonomous navigation capability in the true sense, the invigilator robot is either fixed in the examination room, or moves according to the pre-specified path or needs to store the environment map in advance to complete the path planning and obstacle avoidance; 2) poor environmental adaptability, the monitoring robot can only work in a specific structured environment, and unable to successfully perform the monitoring task in a dynamic environment or an environment with unknown interference; 3) Single function and poor interactivity, although a certain level of recognition accuracy has been achieved, invigilator robots can only perform certain simple operations, resulting in the low intelligence of the robots and the week competence of the invigilating services.

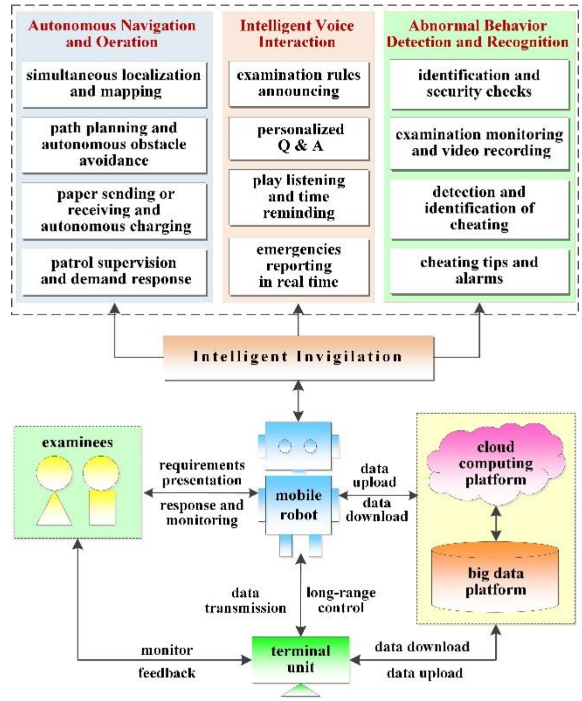

Figure 1: Architecture of autonomous system for mobile education robot towards intelligent invigilation.

3 AUTONOMOUS SYSTEM CONSTRUCTION

In order to improve the applicability of invigilator robot in the examination room environment, it should develop towards intelligence and autonomy. This paper will meet the application need for mobile educational robots (MER) to perform intelligent invigilation, the 3D Simultaneous Localization and Mapping (3D-SLAM) is employed as the key technology of autonomous navigation to help the MER complete real-time positioning and map construction. Then, based on the global path planning, the local path planning is implemented to obtain the reasonable motion path of the invigilator robot, and further performs the tasks of patrolling, invigilating and sending or receiving test papers according to the path. Meanwhile, combined with deep learning, big data and natural language processing technologies to realize the functions of identity verification, cheating monitoring and voice question and answer, so as to ensure that the MER implementing intelligent invigilation can maintain good order in any indoor examination room, achieve all-round strict invigilation and fair invigilation, also, realize the intelligent examination room without invigilation. The full scheme of autonomous system constructed in this paper is illustrated in Fig. 1, and the specific descriptions are as follows:

3.1 Autonomous Navigation and Autonomous Operation

3.1.1 Simultaneous Localization and Mapping

The autonomous navigation and autonomous operation modules are the key modules for MER to realize real autonomy. In the process of performing invigilation, the MER needs to locate the target position to move autonomously in the examination room, which requires the robot to be able to perceive the information of the surrounding environment in real-time, and to effectively determine its own location, so as to facilitate further path planning, obstacle avoidance and autonomous operation for specific invigilation tasks. Therefore, this paper adopts the 3D-SLAM technology based on visual perception for the modules: by first obtaining RGB-D visual information on the surrounding environment through a depth camera, and then extracting ORB (Oriented FAST and Rotated Brief) feature points, as well as utilizing key-frame extraction technology, the robot is able to acquire environmental information through its sensors, and ultimately realize real-time autonomous localization and construction of a working map of its environment. The algorithm is based on the ORB-SLAM2 [7] architecture, with the advantage of its operation not being restricted by the workplace, with high positioning accuracy and strong robustness. The system perceives and obtains the surrounding environment information through the depth camera. The whole framework mainly includes three threads: feature tracking, local mapping, and loop detection, as shown in Fig. 2.

The feature tracking thread realizes the real-time positioning of the visual sensor (depth camera), and tracks each frame of image collected by the depth camera. The local mapping thread completes the local map construction, filters out the recently generated map points and generates new map points, and then uses the local Bundle Adjustment (BA) [13] for optimization, and finally filters out redundant keyframes to obtain a more precise pose and location of feature points. The loop closing thread performs loop closure detection by evaluating the similarity between the image and the previous image, and then corrects the pose of the robot using the repeatedly observed information, optimizes only if multiple key frames are similar, and finally updates the map. Through the above steps, a concise and efficient environment 3D point cloud map can be constructed. In order to construct a 3D navigation map, a map form based on voxel representation is further adopted, that is, an OctoMap is established by using a point cloud map [1].

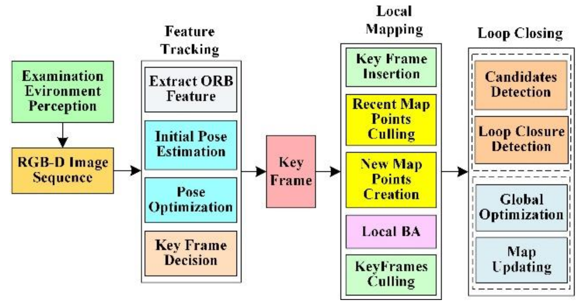

Figure 2: Systematic framework for the 3D SLAM algorithm based on ORB feature points.

3.1.2 Path Planning and Autonomous Obstacle Avoidance

On the basis of completing simultaneous localization and mapping, the MER needs to further implement path planning and autonomous obstacle avoidance, i.e., it needs to have the navigation ability for some target points (each examinee). Specifically, one is test papers sending and collecting, and auto-charging. The difference between the process of collecting and the sending is that the examinee needs to place the back of the test paper and answer card upward, when the robot recognizes that all the test papers are collected, the process is ended. When the power of the robot is lower than the threshold, navigate to the charging area according to the independently planned path, and complete the accurate docking of the two charging ports under the guidance of infrared signal recognition. The other is intelligent patrol inspection and demand response. Similar to invigilator, the robot can move and patrol in the examination room from time to time and randomly plan the route, meanwhile, it can accurately locate and move to the corresponding candidate's location to provide services, such as sending draft paper and 2B-pencils et al., or collecting test materials from candidates separately.

The OctoMap constructed in the 3D-SLAM stage is projected into a two-dimensional grid map, and then the path planning and autonomous obstacle avoidance of MER can be realized according to this grid map. In our system, path planning includes two modules: global path planning and local path planning. The former includes the front-end and back-end. The front-end completes "Path Finding", and uses the D* algorithm [9] [10] based on dynamic heuristic path search to search in the grid map to obtain a preliminary path. The paths obtained through the front-end search generally cannot meet the requirements of robot navigation in terms of stability, smoothness and other performance indicators. Therefore, the back-end is also needed to complete "Trajectory Optimization", i.e., the path optimization is converted into a Minimum Snap [6] trajectory planning problem by combining the objective function and the surrounding environment constraints. Considering that there will be many emergencies in the actual human-machine coexistence environment, such as road blocking when examinees submit their papers in advance and leave the examination room, it is necessary to make local path planning at another level, so that the robot can safely, reliably, and efficiently bypass the obstacles to continue to move. The local path planning of this system adopts a relatively perfected online obstacle avoidance strategy, namely, Dynamic Window Approach (DWA) [2]. According to the motion state model of the MER, the robot motion mode is divided into straight walking without angular velocity and arc walking with angular velocity. In the implementation process, DWA needs to comprehensively consider the factors such as feasible arc trajectory, allowable speed and environmental constraints to form a series of selectable speed windows, then traverses these speed windows, and calculates the evaluation function values corresponding to these speed windows by constructing an evaluation function. Finally, the speed with the maximum evaluation function value is selected as the speed of the robot.

3.2 Abnormal Behaviour Detection and Recognition

This module constitutes the intelligent perception analysis system of the invigilator robot, which adopts deep learning-based object recognition technology and big data technology. Since the Convolutional Neural Network (CNN) is similar to the human brain in image processing, it is often used in image recognition, such as identifying the abnormal behaviour among examinees in exams [3].

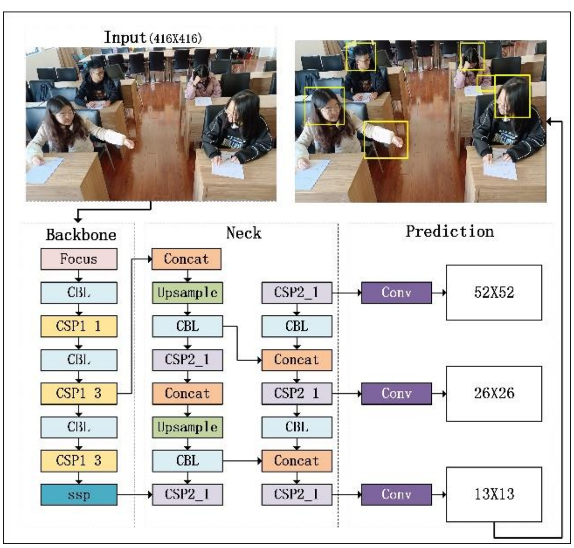

Figure 3: A structural network illustration of YOLOv5 as applied in an examination environment.

The key of the MER executing intelligent invigilation to judge whether the examinee cheats or not is to accurately identify the examinee's behaviour and the objects they touched. The effective method to solve this kind of problem is to employ neural network, among which the YoloV5 algorithm series is widely used. In this paper, it is applied to the recognition of abnormal behaviour of examinees, as given by Fig. 3, including Input, Backbone, Neck, and Prediction. The Input part uses Mosaic data enhancement to input 416×416 size images, the Backbone employs the structures of Focus and CSP, and to balance speed and accuracy. The Neck part adopts the structure of "FPN+PAN" to strengthen the ability of network feature fusion. The Prediction end utilizes GIoU (Generalized Intersection over Union) as the loss function, and outputs 13×13, 26×26 and 54×54 grids to detect targets of different sizes. Finally, non-maximum suppression (NMS) is used to screen multiple target frames so that the occluded objects can also be accurately identified.

Generally, behaviour or object recognition is supported by big data technology. The examinee's behaviour information collected in the huge amount of data in the examination room surveillance video provides rich and diverse comparative data for abnormal behaviour recognition, after obtaining the data stream in these videos, the examinee's behaviour of examination room is divided into the following 7 categories: 1) Thinking: behaviours such as supporting the jaw or the forehead, looking straight ahead, bowing their heads or pondering; 2) Answering: behaviours such as writing or filling in answer sheet; 3) Raising hands: one hand over head; 4) Peeping: behaviours such as glancing left and right, or the significant angling or contortion of body or head; 5) Passing objects: i.e., limbs exceeding prescribed areas; 6) Leaving the seat: i.e., the candidate leaving specified areas; 7) using electronic devices, such as mobile phones or electronic watches. After cleaning the obtained data, the video data set is broken down into multiple images and labelled according to the category and the corresponding model is then obtained by training in the convolutional neural network. Finally, the obtained model is applied to the monitoring video of the examination room to further analyse the behavioural status of examinees in real time. Among these seven types of behaviours, the first two types are normal behaviours, and the last five types are abnormal behaviours, which will trigger the action of MER. Noticeably, the third type does not constitute academic misconduct but is recognized as a signal for MER attention. The third, fourth, fifth, and sixth types are detected through the range of examinees' activities, and the seventh type is detected by identifying abnormal items. Once the abnormality is found, MER will immediately give an early warning to the background administrator; move to the examinee in question, and issue a warning to them of exam protocols. In the meantime, the whole process will be recorded for backup evidence. In addition, when the examinees enter the examination room, the examinees' ID cards and admission tickets are verified by face recognition and safety detection is carried out. If an abnormality or contraband is detected, an early warning will be issued, and a background administrator will arrive to intervene.

During the implementation of intelligent invigilation, the MER tracks the abnormal behaviours of candidates through classification and comparison of candidates' behaviours according to the computer vision perception technology of CNN classification model, thereby, to judge whether candidates cheat, simultaneously, the MER verifies the identity of candidates by combining face recognition technology, and thus to build a series of artificial intelligence perception analysis systems to provide support for examination room inspection and monitoring services.

3.3 Intelligent Voice Interaction

Natural language processing (NLP) technology is one of the most common human-computer interaction methods, while interaction is an important symbol of robot intelligence [5]. Before MER's Q & A, developers need to collect data through carrying out multi-channel research to build an Q & A database, in which the common problems raised by most examinees as the main data element and the individual problems as the auxiliary data element, as shown in Fig. 4.

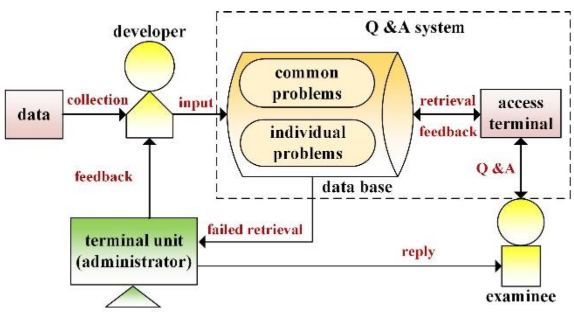

Figure 4: Working model of the Q & A system.

After perceiving the signal of examinee's question, the access terminal develops semantic understanding and analysis for the question, and then retrieves the matching information of answers in the established Q & A database, also, translates the information into a natural language acceptable to the examinee to interact with it. When the retrieval fails, the question is sent to the background administrator through the terminal unit, and answered by administrator. Finally, the developer enters it into the Q & A database. NLP technology enables the MER to be accurate, real-time, and flexible in its performance, and makes MER own two functions: one is to monitor the examination process and the other is to support the personalized inquiries of examinees.

The intelligent voice interaction module enables the MER to make announcements as well as answer questions in the examination room. That is, announce the examination rules and other requisite content before the commencement of exams, and broadcast listening comprehension in the language exam, e.g. CET4&6, TOEFL, IELTS, etc. When examinees have some questions, they can guide the MER to move to their position by raising their hands to put forward needs, the robot will answer accordingly. If human intervention is required, MER will notify the background administrator immediately. For common problems, MER answers in a low voice first, and then turn up the volume to inform other examinees. For example, if the examinee asks, "My answer for this question has been written into the space for another question", MER will first answer the examinee's question in a low voice, "you can mark the correct question number in front". Then it will turn its volume up to announce, "Everyone, please pay attention to the question number on the answer sheet before writing." It follows that, the intelligent voice interaction module can enhance the intelligence of MER performing intelligent invigilation tasks and it is one of the important modules for candidates to interact with robots in the exam, which effectively avoids these cases, e.g. proctors whisper to answer candidates' questions or reveal answers, and thus eliminate the doubt of collusion between invigilators and candidates.

3.4 Environment Description and Intelligent Inspection

The traditional examination scene consists of four basic elements: invigilator, examinee, surveillance camera and storage devices, while the intelligent examination scene using MER includes intelligent invigilation mobile robot, examinee, terminal equipment, cloud computing and big data platform, these parts together form a closed-loop service support system. The MER equipped with surveillance cameras, and storage devices, and possess the ability of invigilator, which is supervised by the background administrator. In this scene, MER not merely have the abilities of sending or receiving test papers, invigilating, patrolling and auto-charging, but can complete the all-round invigilation that invigilators cannot do, which makes the invigilator abandon repetitive work and become a general supervisor.

Furthermore, considering that the MER is likely to interfere with the candidates if it keeps moving during the invigilation process, thus, a certain time interval is set during the whole invigilation process to make the robot invigilate alternately in the static and mobile working modes. In the static mode, the robot is located in the middle of the lecture platform, which is conducive to monitoring every candidate; in the mobile mode, the robot patrols along the random route and can monitor the candidates in all directions. When the robot detects that the candidate has abnormal behaviour, it will divide the behaviour into two cases: raising hands and cheating. In case of raising hands, the robot will acquire the candidate's position and move to the side to receive questions, and feedback information or respond to needs in the form of natural language; in case of cheating, the robot will immediately send an early warning to the background administrator, move to the location of the cheating candidate, prompt the candidate or retract the paper according to the severity of cheating, and store the process data, e.g. monitoring video, Q & A process, etc.

4 CONCLUSION AND PROSPECT

In this paper, a MER autonomous system for performing intelligent invigilation tasks is constructed by comprehensively using various kinds of intelligent technologies in the field of AI. The system enables the MER to have strong autonomous ability in the whole invigilation process, so that it can complete the invigilation task intelligently and efficiently in an unknown examination environment.

Using the autonomy and interactive function of the robots to monitor the situation of the examination room cannot merely save human resources, make the examination organization more flexible and convenient, but improve the efficiency while ensuring the scientific and fair of invigilation process, and thus the rights and interests of candidates are protected to a certain extent.

As part of future work, we intend to link the MER performing intelligent supervision with the moral cultivation of students, and formulate relevant teaching strategies for the moral development of students of all ages, so as to achieve the dual effect of moral education and serving. In addition, to promote the autonomy and applicability of future educational robots in different educational environments, we plane to improve the algorithm to enable robots to move autonomously in complex dynamic environments with interference, such as invigilating of outdoor sports, etc.

ACKNOWLEDGEMENTS

This work was supported by the National Nature Science Foundation of China (Grant No. 62063036) and Research Foundation for Doctor of Yunnan Normal University (Grant No. 01000205020503115)

REFERENCES

[1] Meagher, D. (1982). Geometric modeling using octree encoding. J. Computer graphics and image processing. 19(2), 129-147.

[2] Fox, D., Burgard, W., & S. Thrun (1997). The dynamic window approach to collision avoidance. J. IEEE Robotics & Automation Magazine, 4(1), 23-33.

[3] Liu, Y., Li, Q., & C. B. Yu (2017). Application of deep learning technology education: status quo and prospect. J. Open education research. 23(05), 113 – 120.

[4] Li, X., Li, J.J., Gao, B.J., & Y. Y. Li (2020). Research status and development suggestions of educational robots - Based on literature analysis of the Web of Science core database from 2014 to 2018. J. Modern educational technology. 30(01), 5-11.

[5] Li, Y.Y., Li, X., & J. X. Wang (2020). Design and key problem analysis of educational robot. J. Modern educational technology. 30(01), 12-17.

[6] Mellinger, D., & Kumar, V., 2011. Minimum snap trajectory generation and control for quadrotors. In 2011 IEEE international conference on robotics and automation. IEEE.

[7] Mur-Artal, R. & J. D. Tardós (2017). Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. J. IEEE transactions on robotics. 33(5), 1255-1262.

[8] Mubin, O., Cappuccio, M., Alnajjar, F., Ahmad, M. I., & S. Shahid (2020). Can a robot invigilator prevent cheating?. J. AI & SOCIETY. 35(4), 981-989.

[9] Stentz, A., 1995. The focussed d^* algorithm for real-time replanning. In IJCAI.

[10] Stentz, A., 1997. Optimal and efficient path planning for partially known environments. In Intelligent unmanned ground vehicles. Springer.

[11] State Council. (2015). Circular of the State Council on the Issuance of Made in China 2025. EB/OL. <http://www.gov.cn/zhengce/content/2015-05/19/content_9784.htm >

[12] State Council. (2017). Circular of the State Council on Printing and Issuing a New Generation of Artificial Intelligence Development Plan. EB/OL. http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm

[13] Triggs, B., McLauchlan, P. F., Hartley, R. I., & Fitzgibbon, A. W., 1999. Bundle adjustment—a modern synthesis. In International workshop on vision algorithms. Springer.