Quantitative Analysis of Classroom Attention Teaching Based on Stochastic Cascade Regression and PnP Method

Bing Gong(✉) and Jing Wei

Information Engineering School, Eurasia University, Xi'an, China
gongbing@eurasia.edu

Abstract. In the process of quantitative analysis of classroom teaching, it is often necessary to evaluate individual indicators related to attention such as "whether the teacher's teaching attracts students, whether it stimulates the enthusiasm for learning" and "whether there is eye contact between teachers and students". Learning attention reflects the students' learning state and learning psychology, and meets the observable, reproducible, and interpretable requirements for quantitative analysis of the classroom teaching process. The project team used random cascade regression to obtain the coordinate information of 11 facial feature points. Used to provide 2D information in PnP solutions. Based on a standard face model obtained by statistical measurement, the mapping relationship between 2D/3D is obtained by solving PnP, and the rotation and translation matrices of the head pose are output. According to the rotation and translation matrix information of the head posture, the student's viewpoint is projected into the teacher's teaching video shot by the rear camera by using the transformation relationship of the spatial coordinates, so as to realize the visual display of learning attention.

Keywords: component: Classroom Teaching · Stochastic Cascade Regression · PnP · Quantitative Analysis

1 Introduction

In traditional primary and secondary classroom teaching, classroom teaching is mainly based on teachers' teaching, supplemented by other forms such as classroom discussions. The main source of teacher-student interaction is not asking and answering questions, but the eye-to-eye interaction between teachers and students [1]. Generally speaking, teachers will pay more attention to high-achieving students, but less attention to students who are good students; students with better grades are more eager to make frequent eye contact with teachers than students who are not interested in the teaching content. In contrast, interested students will pay more attention to the blackboard and the teacher, and in return, the teacher will have more eye contact with interested students. When the teacher asks a question, students who are interested in the question or have mastered the knowledge point will raise their hands, eager to make eye contact with the teacher, and

C. F. Peng et al. (Eds.): EIMT 2022, AHSSEH 3, pp. 596–604, 2023.
https://doi.org/10.2991/978-94-6463-012-1_66

hope to express themselves, while students who are not interested or have not mastered the improvement knowledge point hardly dare to look up at the teacher [2].

At present, with the rapid development of machine vision and image processing technology, the use of digital cameras to record the eye movement process, and the use of image processing methods to analyze the direction of the optical recording method has been widely used, and a wide variety of eye tracker systems have been developed [3]. Including non-wearable fixed eye trackers, wearable portable eye trackers, etc.

However, both types of eye trackers have a certain cost. For example, the price of a portable eye tracker is about 200,000 yuan, while the price of a fixed eye tracker is about 50,000 yuan. Although, in terms of flexibility of use, portable eye trackers can be used for automatic evaluation of classroom teaching, but each student must be equipped with a set of eye trackers, which is difficult to meet the needs of classroom teaching to analyze the attention of 20 to 40 students need [4]. Although the fixed eye tracker is relatively cheap and can also analyze the learning attention of multiple students, it can only analyze the attention of students within 1 to 2 m away from it [7]. Therefore, the fixed eye tracker can be used to analyze the learning attention of students in online education. For the analysis of the learning attention of the classroom teaching process with the nearest distance of more than 3 m, the fixed eye tracker is powerless [9].

In order to carry out long-distance attention analysis of teachers and students in the classroom economically, the paper used a visual analysis method of learning attention based on plug-and-play head pose estimation of a single image, which can better meet the needs of quantitative analysis of attention in classroom teaching process.

2 Attention Visualization Analysis System Based on Front and Rear Cameras

Due to the lack of tools and software to automatically track the gaze of teachers and students in the process of quantitative analysis of classroom teaching, at present, the evaluation teachers enter the class or watch the classroom inspection videos to manually conduct individual evaluations related to attention. Evaluate [10]. At the same time, limited by the professional level and energy of the evaluation teachers, it is difficult for the evaluation teachers to quantitatively analyze and evaluate the attention of the teachers and all the students in the class [9].

As shown in Fig. 1, by analyzing the posture of the teacher's head (back of the head), the teacher is looking at the students on the right side of the center to teach. Guide, listen to the class with one's own computer as the carrier, and this is a very common phenomenon in many classroom teaching at this time. Combined with the gaze tracking of other students, it can be seen that the teacher's teaching still attracts the students' attention. At the same time, combined with the gaze tracking in the video stream, it can be seen that there is also eye contact between teachers and students. All in all, the teacher's teaching can stimulate students' enthusiasm for learning [11].

Due to the objective and inestimable characteristics of the classroom scene, such as the cluttered background and the large venue, and due to the limitation that the current camera equipment uses high-definition images, it is still not enough to meet the requirements of individual eye tracking, so it is necessary to directly detect and estimate

Fig. 1. A classroom teaching scene in university education

through machine vision. It is more difficult to get out the students' line of sight, but in the teaching scene, the student's head posture can basically be used to replace the general direction of the student's line of sight, so the visual analysis of attention for the quantification of the classroom teaching process is based on the student's head posture of [5].

Figure 2 shows the smart classroom laboratory of Xi'an Eurasia University. The teacher teaches about 10 students in the classroom in front of a blackboard with a width of about 4 m. In order to monitor the students' learning status process and evaluate the students' learning effect and learning status, two high-definition surveillance cameras are usually installed above the center of the blackboard at point F or above the upper left corner of the blackboard; in order to analyze and summarize teachers' teaching skills, and To monitor students facing the back in group teaching, the project team also installed three high-definition cameras at point B above the middle of the back wall (opposite the blackboard). The teaching process of the staff is monitored, and two high-definition cameras are installed on each of the two walls. There are 9 high-definition cameras in the whole classroom.

Based on the video image data captured by the front and rear cameras shown in Fig. 2, we mainly use the following six functional modules for visual analysis of attention in the classroom teaching process.

(1) Data (video frame) acquisition: The classroom teaching video is collected through the 1080p high-definition PTZ camera based on the H.264 high-definition compression algorithm, and the video frames are separated to obtain image samples for attention analysis.
(2) Camera calibration: In order to improve the accuracy of head gesture recognition, it is necessary to use a convenient and accurate calibration method to calibrate the camera parameters.

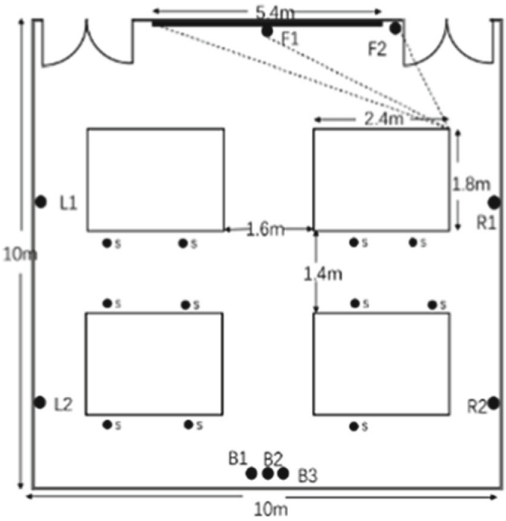


Fig. 2. The layout of a standard classroom and the installation positions of the two cameras in the front and rear

(3) Face detection: Detect faces from discrete video frames using publicly available face detection algorithms.
(4) Face feature point detection: This section uses random cascade regression to obtain the coordinate information of 11 face feature points, which is used to provide two-dimensional information in the PnP solution.
(5) Head pose estimation: Based on a standard face model obtained by a statistical measurement, the mapping relationship between 2D/3D is obtained by solving PnP, and the rotation and translation matrix of the head pose is output [8].
(6) Student's viewpoint positioning: According to the rotation and translation matrix information of the head posture, the student's viewpoint is projected into the teacher's teaching video shot by the rear camera by using the transformation relationship of spatial coordinates to realize the visual display of learning attention.

3 Quantitative Analysis of Classroom Teaching Process Based on Visual Analysis of Attention

In order to conduct a visual analysis of attention in the classroom teaching process, in Fig. 3(a), the subject sits 3 m in front of the blackboard and looks at the mathematician who is writing the edition. A front camera is installed just above the blackboard. Monitor the subject's learning status. In Fig. 3(b), the rear camera is installed behind the subject to monitor the teacher's teaching status. In order to ensure the test accuracy, the front and rear cameras need to be calibrated during the installation process, so that the hardware part of the attention visual analysis system is built [6].

According to the processing flow shown in Fig. 4, further develop the software part of the attention visualization analysis system. As shown in Fig. 4, based on the video frame

(a) Installation position of the front camera

(b) The installation position of the rear camera and the visualization of the subject's head pose

Fig. 3. The experimental setup and display effect of the visual analysis system of students' learning attention

image captured by the front camera, the head posture of the tested students is estimated, and the three-dimensional angle information and three-dimensional position coordinates of the tested students' head posture are calculated (a total of 6 dimensional information), for the convenience of display, six-dimensional information is marked on the nose tip, x-axis, y-axis, z-axis, as shown in Fig. 3. Using the derived head pose-based student attention visualization method, the physical position of the student's attention point can be calculated and marked on the video frame image captured by the rear camera. The light in Fig. 4(a). The point is the gaze point of the student's gaze.

The above is the implementation process of the entire attention evaluation system. According to the relevant process, the project team simulated the scene of the attention visualization analysis system and analyzed the results.

As shown in Fig. 4(a), the students are in a focused learning state, their eyes are focused on the teacher, and they are eager to make eye contact with the teacher. The student's viewpoint based on the head posture is densely distributed near the teacher's head. It can also be seen from the head posture of the students that their eyes will be focused on the teacher.

As shown in Fig. 4(b), the students are in a state of paying attention to learning. From the head posture of the students, it can be seen that they are looking at an area of the blackboard, and the students' viewpoints based on the head posture are densely distributed in this area.

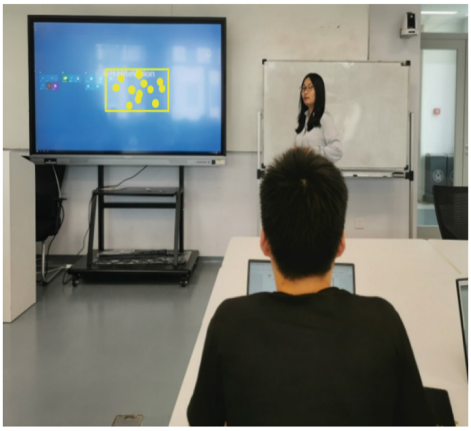
As shown in Fig. 4(c), the student is in a state of ignoring learning. From the student's head posture, it can be seen that he is looking at a certain area in front of the desk, and the student's viewpoints based on the head posture are densely distributed in this area.

From the three visualization results in Fig. 4, it can be seen that the attention analysis system can more accurately give the students' viewpoint distribution, can more accurately analyze the students' learning attention, and provide more accurate quantitative analysis of the classroom teaching process. A clear research basis.

Through the above data analysis of the whole 15-min classroom teaching, the project team has initially achieved some results under the premise of incomplete data.

(1) Teachers should flexibly prefabricate some measures to accelerate the emotional involvement of students in the classroom, so as to improve the concentration of all students. For example, the use of pre-class questions, knowledge consolidation, classroom random questions and other means. During the course of the class, measures such as teacher-student interaction and other measures to ease the classroom mood should also be appropriately increased, so as to relieve the fatigue of the students in class and relieve the students' emotional disengagement before get out of class.
(2) The research results can conduct automatic, full-process, non-perceptual teaching effect evaluation for the classroom teaching process, and can assist teachers in decision-making on teaching strategies. If at a certain moment the students' emotional involvement, participation, attention and difficulty in the classroom all decrease, it means that the students' listening status needs to be adjusted in time. At the same time, teachers should be more vigilant, adjust teaching strategies in a timely manner, activate classroom atmosphere, improve teaching efficiency, and improve teaching effects.
(3) According to the data of individual students' classroom emotional involvement, combined with the progress of classroom teaching, highly targeted learning strategies can be provided for students with learning difficulties. Aiming at the basic learning ability of some students in our school, their study habits leading to the loss of most of their emotional input in classroom teaching is unconscious. Therefore, through the analysis of personal emotional data after class, it is possible to put forward personalized learning strategies and suggestions, change learning habits, optimize learning effect, and improve learning level.

(a) Focus: Students focus on the teacher and make eye contact with the teacher



(b) Attention: Students focus on the area inside the blackboard

(c) Ignoring: Students look away from the blackboard area

Fig. 4. Visual analysis of attention of three typical students' learning states

4 Conclusions

To sum up, quantitative analysis of attention in the classroom teaching process requires a long-distance attention analysis of teachers and many students in the classroom to evaluate the attention distribution of teachers and students. By analyzing the attention of the tested students, it is possible to evaluate whether a certain student or all students pay attention to the teacher's teaching content, indirectly whether the teacher's teaching stimulates students' enthusiasm for learning, and whether the content can catch the students' attention.

In the current educational environment, learners can only understand their learning effects through academic performance evaluations or other written evaluations given by teachers, lacking in-depth understanding of their own physiology or psychology, and unable to correctly understand their own learning. Learning habits after points of interest. It is difficult for the tested students to clearly and intuitively understand their own learning situation during the learning process. Therefore, when learning problems occur, they often cannot accurately focus on the root cause of the problem, or cannot accurately understand the root cause of the problem. For example, when some of the tested students have unsatisfactory academic performance due to lack of concentration, the individual tested students may not be aware of the problem of distraction. Therefore, the test students need tools to help them improve their self-assessment, and use data facts to help the test students understand their personal study habits. Through the inquiry of personal data, establish self-awareness and evaluation of personal learning attitude and other aspects, understand their own learning situation, conduct self-reflection, and actively adjust during the class process, form good study habits, and improve learning efficiency.

When the tested students are studying in the classroom, the computer can monitor the learners' current head posture in real time through technical means, and analyze the data of their attention. Complete the attention analysis of the student based on the attention situation data. At the same time, the collection of data should be completed without affecting classroom teaching, so as to ensure the non-perceptual, full-process, and fully automatic implementation of the analysis process.

Acknowledgements. In this paper, the research was sponsored by the Social science fund of Shaanxi Province (Project No. 2019Q019), China Higher Education Association Special Project (Project No. 2020XXHYB13) and Shaanxi Provincial Department of Science and Technology Key R&D Program (Project No. 2022GY-317).

References

1. Dolecki, M., P. Karczmarek, A. Kiersztyn, and W. Pedrycz. 2016. Face Recognition by Humans Performed on Basis of Linguistic Descriptors and Neural Networks. New York: IEEE Press.
2. Feng, Z.H., P. Huber, J. Kittler, et al. 2015. Random cascaded - regression copse for robust facial landmark detection. *IEEE Signal Processing Letters.*
3. Fu, Z.P., Y.N. Zhang, and H.Y. Hou. 2014. *Survey of Deep Learning in Face Recognition.* New York: IEEE International Conference on Orange Technologies.

4. Hang, Z. 2000. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*
5. Huber, P., G. Hu, R. Tena, et al. 2016. A multiresolution 3D morphable face model and fitting framework. In *Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications.*
6. Mello, S.K. 2013. A selective meta-analysis on the relative incidence of discrete affective states during learning with technology. *Journal of Educational Psychology.*
7. Ren, S., K. He, R. Girshick, et al. 2015. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis & Machine Intelligence.*
8. Good, Thomas L., and Jere E. Brophy. 1970. Teacher-child dyadic interactions: A new method of classroom observation. *Journal of School Psychology.*
9. Xing, Beibei, Xianmin Yang, and Qinsheng Li. 2016. Education big data comes from collection technology. *Modern Education Technology.*
10. Zhan, Licai, Fengyan Chen, and Shimin Meng. 2016. Technical support, dynamic structure and mechanism of online campus management under education big data research paradigm based on system dynamic theory. *Journal of Distance Education.*
11. Zhan, Licai, Fengyan Chen, and Shimin Meng. 2016. Technical support, dynamic structure and mechanism of online campus management under education big data - research paradigm based on system dynamic theory. *Journal of Distance Education.*