IOP Conf. Series: Materials Science and Engineering **392** (2018) 062167 doi:10.1088/1757-899X/392/6/062167

On the Intelligent Lecture Capture System to Achieve Quality-Enhanced Lecture Recording

Li Xiuhua[1], Zheng Xingchao[2], Yang Weiwei[3] and Zhang Zhe[4]

[1]State Grid Technology College, Jinan, Shandong, China

[2]State Grid Technology College, Jinan, Shandong, China

[3]State Grid Technology College, Jinan, Shandong, China

[4]State Grid Technology College, Jinan, Shandong, China

fallingsnow201007@163.com

Abstract. With the popularity of online learning, the capture of lecture video has become a hot topic where automation is the key challenge. This paper describes an intelligent lecture capture system that takes advantage of computer vision and machine learning technology to achieve automatic and quality-enhanced lecture recording.

1. Introduction

Online learning has become increasingly popular in the past decade, where the lecture video is the essential part [3]. Lecture video facilitates the university students to learn the course without physically attending the lecture and the university staff to distribute and share the education resources conveniently. Most students think it of great help to have lecture videos as resources to understand and review the course content [5]. Despite the great advantage of lecture video, however, there exists a challenge about how to record the lecture video with low cost and high quality.

Nowadays, computer vision and machine learning technology have become more and more mature and shown their great power in many application areas. This proposal describes an intelligent lecture capture system that takes advantage of these two technologies to achieve automatic and quality-enhanced lecture recording.

2. Background / Literature Review

The main technical problem is how to control the camera to track the lecturer. Currently, the mainly used tracking method in lecture capture is still to manually control the motion of the cameras. For example, according to the website of Publishing Technology Center (PTC) in HKUST [2], the motion of the cameras is controlled by staff in PTC, and one person may have to take care of up to four cameras simultaneously to cut down the cost of labour.

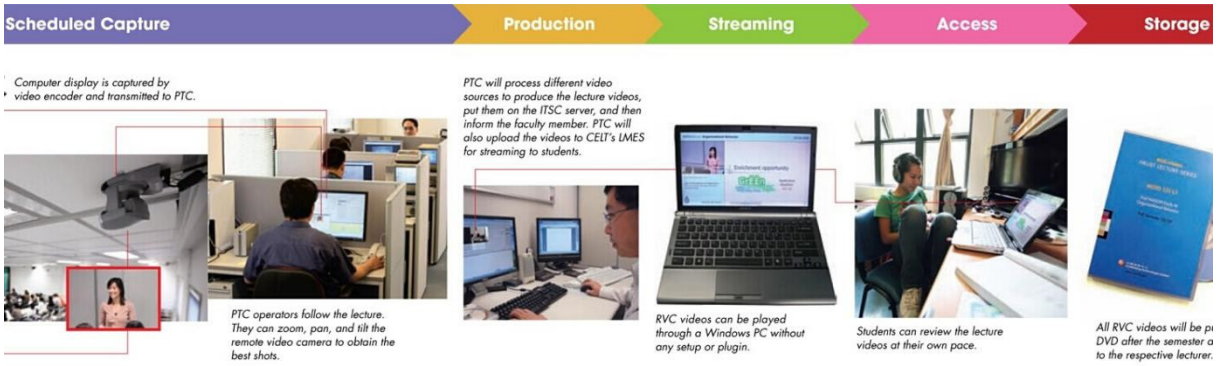

Fig. 1 Current Video Capture Mechanism in HKUST

This method requires little technology investment and the focus of the cameras can be adjusted precisely by the controller. However, the drawbacks are also very obvious. For one thing, salary paid for the staff is a long-term expense. For another, the tracking may be delayed when multiple cameras are controlled by one person at the same time.

Besides this, some other solutions about automatic tracking are also developed or adopted in some universities. One solution is that extra static cameras are used getting the whole picture of the classroom to localize the lecturer [1]. Another is that attention is focused on an effective editing of the lecture videos shot by array of multiple cameras [4]. These solutions are exempt from the assistance of human, but expenses on extra cameras or other equipment are necessary. Furthermore, the quality and precision of the recording cannot be ensured because the focus of the lecturer may not be understood by the computers.

The system proposed by this report adopts gesture recognition technology to get gesture information of the lecturer, and uses machine learning to analyze the timing of zoom and motion of the cameras. It can achieve the automation recording of high quality at low cost

3. Technical Description

3.1. Video sources

The intelligent lecture capture system will take three video inputs as sources, which are one camera tracking and recording the lecturer, another one focusing on the whiteboard, and one video stream from the desktop that records what is displayed by the projector. The three videos will be integrated together to provide thorough and quality-ensured lecture content for students.

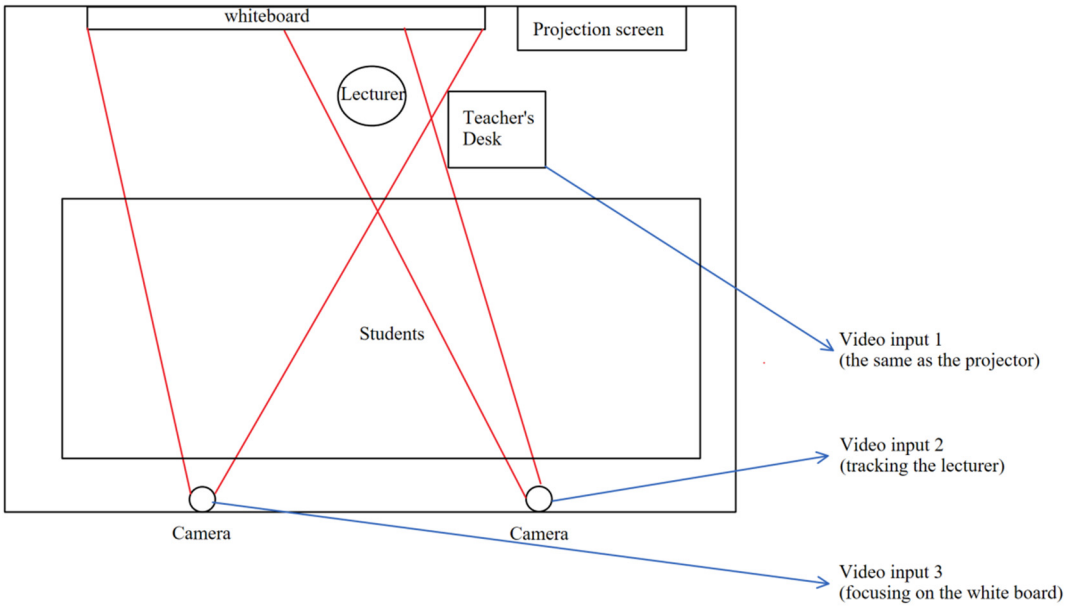

Fig. 2 Video Inputs of the System

3.2. Gesture recognition and tracking

The lecturers will be recognized and tracked by the cameras automatically, which is realized by cooperation of four components: the camera array, the integrated video analyzer, the camera motion controller, and the backend server.

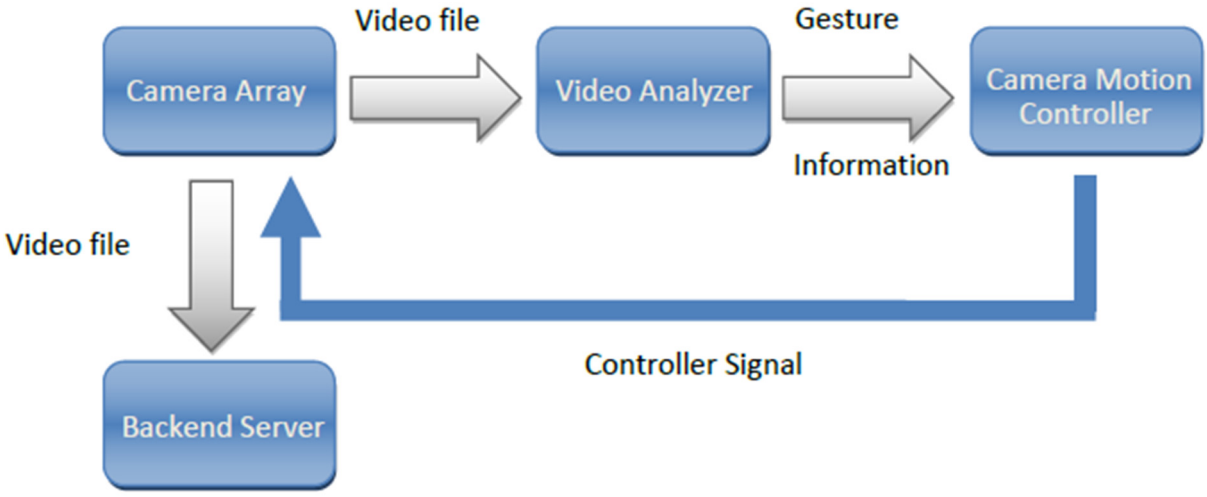

Fig. 3 Mechanism of the Tracking System

First, the lecture will recorded by camera the, and video files generated will be sent to the video analyzer for the lecturer detection and the backend server for storage. The video analyzer will perform gesture recognition on the lecturer, which can outline the position and the posture of body and limbs of the lecturer.

The following step is that the camera motion controller analyzes the indication of the lecturer based on the gesture information and directs the cameras to move appropriately. For instance, if the lecturer is pointing at somewhere on the whiteboard or the projection screen, the controller will give out signal of zooming in to give a clearer view on the content. If the lecturer is going to walk out of the camera view, the camera will be directed to move in the same direction to catch up with the lecturer.

Finally, the cameras will receive signals from the controller and move correspondingly to achieve the automatic lecturer.-tracking

3.3. *Machine learning on the indication of the gesture*

In order to move the cameras properly, besides some preset reactions mode, machine learning technology can be adopted to train the software. By learning the huge amount of lecture videos recorded manually previously, the system can figure out the relationship between the some specific gestures and the regular response of human controller. This method can help save lots of cost on analyzing and programming of the camera's response pattern and achieve better judgement on the lecturer action.

4. Feasibility / Benefits

Thanks to the rapid development in the gesture recognition, machine learning and related fields in the recent years, the related technical tools in this system are already very mature and easy to integrate [6]. Besides, the expenses on development or installment of this system are acceptable. Consequently, this innovation has both technical and economic feasibility.

This system has several major advantages. First and foremost, it can help cut down a considerable amount of cost, especially in cities like Hong Kong where the cost of labour is relatively high. Moreover, with the lecture video of higher quality, students can learn the course content more effectively and thus achieve better academic performance. Last but not least, due to the low dependence on the hardware equipment and adoption of the machine learning technology, the intelligent lecture capture system can be installed in almost any classroom with acceptable cost.

The only possible limitation in the intelligent lecture capture system lies in the requirement of the server for huge video storage and the internet bandwidth for transmission. Nevertheless, as the internet technology develops so rapidly, such problems will not be major concern in the near future.

5. Conclusion

The capture of lecture video has become a hot topic in online education, where automation is the key challenge. As proposed in this report, integrated with the gesture recognition and machine learning technology, the intelligent lecture capture system provides a thorough and feasible solution to this problem.

Thanks to the rapid development in related technology in recent years, the implementation of the system has become totally practical. The deployment of this system will greatly help the university to cut down the cost, assist the lecturer to review or share their teaching, and provide a convenient way for the students to get better education.

References

[1] Zhang, Y. Rui, L.-W. He, and M. Wallick, "Hybrid Speaker Tracking in An Automated Lecture Room" 2005 IEEE International Conference on Multimedia and Expo. ICME 2005, 2005, p. 4 pp.

[2] "Video Capture Mechanism", Rvc.ust.hk, 2017. [Online]. Available: https://rvc.ust.hk/how_vcm.html . [Accessed: 17- Apr- 2017].

[3] M. Ketterl and K. Morisse. User generated web lecture snippets to support a blended learning approach. In G. Siemens and C. Fulford, editors, Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009, pages 2886–2893, Honolulu, HI, USA, June 2009. AACE.

[4] " Capturing & Managing Video Content for Offsite Delivery | AVerMedia", Avermedia.com, 2017. [Online]. Available: http://www.avermedia.com/professional/industries/training_education/solution/classroom_recording_system . [Accessed: 17- Apr- 2017]

[5] "Student Survey in Fall Semester 2009", Rvc.ust.hk, 2017. [Online]. Available: https://rvc.ust.hk/findings_survey_student.html. [Accessed: 17- Apr- 2017].

[6] H. Badi, "Recent methods in vision-based hand gesture recognition", International Journal of Data Science and Analytics, vol. 1, no. 2, pp. 77-87, 2016.