Online AI-Guided Video Extraction for Distance Education with Applications

Hua Gao

Zhengzhou Preschool Education College, Zhengzhou 450000, China

Correspondence should be addressed to Hua Gao; warmerwy@126.com

Received 26 January 2022; Revised 1 March 2022; Accepted 5 March 2022; Published 30 March 2022

Academic Editor: Naeem Jan

Copyright © 2022 Hua Gao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The pervasive applications of artificial intelligence (AI) have made an impact on the human knowledge sharing and education. In practice, a lifelong learning process is becoming an inevitable trend in the future. Although distance education has been developed for a long period, there are still problems such as low learning enthusiasm, low autonomy, and unsatisfactory learning support services. The application of AI to distance education will fundamentally affect the instruction and guidance methods, as well as the instructing content of distance education. Evolutions in distance education have brought significant changes in the instructing mode, learning strategy and support service, evaluation mode, and instructing platform of distance education. Distance education can construct an intelligent education platform to provide learners with personalized learning support. It can also offer flexible, dynamic, and convenient evaluation and on-demand learning resources with guaranteed quality. By leveraging the state-of-the-art data mining technique, we in this work design a set of simple yet effective distance education system, in the context of video conferencing. Our technique depends on the existing video surveillance network and equipment. Our technique can be applied to achieve an effective face-to-face online interaction. Meanwhile, the video surveillance network switch is optimized to enhance the network transmission and ensure the reliability of network operation. Comprehensive experimental results on the two instructing platforms based on embedded system learning and hospital nurse learning have shown the advantage of our method.

1. Introduction

Online conference is an integral part of human social and economic life. Research has shown that approximately 55% of the effectiveness of communication depends on face-to-face visual effects, while 38% depends on speech. While speech is a way for people to acquire the acoustic channel information, face-to-face discussion is the most intuitive way of human communication. The AI advancement has made a breakthrough in online communication technology, such as the invention of telephone and television. In this way, people from far-flung places can transmit voice and documents instantly. In addition to verbal communication, in the past decades, human beings have been investigating the technique to transmit facial expressions and body expressions in real time. This is because the transmission of voice can no longer meet the requirement of modern people communication. Apparently, real-time and interactive images can transmit information more naturally and intuitively [1]. The video conference system was created under the impetus of such requirement. Through various electrical communication transmission media, the static and dynamic images, voice, text, pictures, and other information of characters are delivered to the terminal equipment of each user. In this way, geographically dispersed users can gather together to exchange information through graphics, sound, and other channels. This can increase the understanding of the content of both parties and make people participate in the online meeting as if they were immersed in the same venue. Our unit has many subordinate units and can support faraway distance. In practice, it has been difficult to gather multiple participants together for a meeting, such as academic discussion and exchanges in the same physical location. A relatively simple yet effective way is to leverage the existing special dispatching equipment to listen to the participants’ speech timely. Noticeably, the disadvantage of
this communication scheme is that you cannot understand who the other party is. Moreover, you cannot understand the other party’s body language. To this end, the design and investigation of the distance education system platform is conducted in the form of video conference in this work. An overview of the distance education framework is given in Figure 1.

With the development of cloud computing, big data, and meta-universe, deep learning-guided artificial intelligence has more and more applications in people’s daily life. In 2011 [2], the Google brain recognized cat faces through unsupervised learning. In [3], the image recognition accuracy achieved by Stanford Artificial Intelligence Laboratory surpassed the human eye in the well-known ImageNet image recognition competition. Besides, Microsoft speech recognition system can reduce the English word error rate by 5.2%. Further, in 2019 [4], Google Go beat South Korea’s top player Lee Sedol, which shocked researchers in the AI community. The AI era has substantially arrived, and it has replaced humans in many fields with competitive computing power, super-human precision, and decision-making power. Based on these observations, 2019 is known as the first year of “artificial intelligence,” which has attracted the attention of many people outside the computer industry. Later, Citibank and Oxford University released “TECHNOLOGY AT WORK v2.2” (technology promotes work 2.0); in the report, it was stated that 75% of Ethiopia’s jobs will be replaced by robots. They claimed that in terms of the replacement rate, China ranks first in the world with 79%, and the United States ranks second with a replacement rate of 47%. The research and analysis team of Morgan Stanley in the United States conducted a study on the possibility of 19 occupations being replaced by robots. Chefs are over 87%, and programmers and journalists are 48% and 11%, respectively. Subsequently, the National Development and Reform Commission in China issued the “Internet +” Artificial Intelligence. The three-year action implementation plan emphasizes the importance of artificial intelligence, and the plan involves emerging industries [5–8]. On Sep 2, 2009, China released the first artificial intelligence plan “New Generation Artificial Intelligence Development Plan,” which proposed to accelerate the innovative application of artificial intelligence.

At present, AI has been utilized in image processing, natural language processing, speech recognition, computer vision and images, intelligent robots, and autonomous driving. It has been pervasively applied in many fields such as medical care and smart home. The “intelligent” attribute of artificial intelligence determines that it will inevitably challenge many conventional occupations that are repetitive, regular, programmed, standardized, and “low-brained.” Gradually, these occupations (such as diagnosis, accounting, and assembly line workers) will be replaced by intelligent machines. Artificial intelligence will inevitably lead to fundamental changes in the structure of social talent demand. Nowadays, the existing education is difficult to adapt to the rapid update and iterative development of information technology. In this way, the future education will pay more attention to lifelong learning and training education.

Distance education that has always undertaken lifelong learning, as we expected, will occupy a larger proportion in the education industry in the future. It will also be the pilot direction and main domain for exploring AI applications in the distance education industry.

Under the AI-guided platform, on the basis of maintaining the dispatching voice transmission system to ensure the voice quality, the lecture sound is sent to the HDMI encoder at the same time. Subsequently, they will be sent to the monitoring server after being encoded. In the case of dispatch system failure, our system provides another strategy for listening to the voice. We leverage massive-scale graph-based data mining techniques to discover videos containing important contents encoded by students and instructors. This will help students and instructors access the key content of online videos more efficiently and effectively. After the platform was constructed, each device was debugged, especially in the refinement of remote camera code stream. Since the existing network transmission bandwidth from the central site to the remote unit is only 10–20 Mbps, the larger the set code stream, the larger the required bandwidth, and the parameter configuration will be reasonably selected regarding these two factors of network bandwidth and resolution. On-site verification was conducted, based on which the online videos were captured and the visual effect was satisfactory. The main venue and branch venues can watch the real-time images within each venue. Then, the educational content was smoothly viewed through the monitoring client, and the real-time sound was clear. The pipeline of our proposed method is elaborated in Figure 2.

The organization of the rest of this article is as follows. We briefly review the related work in Section 2. Section 3 introduces the key components of the proposed method of distance education framework. Section 4 deals with the experimental results of the proposed method with its useful outcomes discussed. Section 5 finally concludes the paper.

2. Related Work

With the popularization of higher education and the expansion of university enrollment, the number of students continues to increase in China in the past decade [3, 5, 6, 9]. Accordingly, the instructing task of professors is becoming more and more challenging. For example, the number of annual instructing classes in mechanics of materials in universities has increased from 24 classes to 30 classes.
Comparatively, the number of instructors who undertake instructing tasks has increased very slightly. Zhu et al. [9] showed the change in the number of classes in the mechanics of materials from 2013 to 2017. The course instructing time is long, and thus instructors are under great pressure to teach. On the other hand, students in many courses have a lot of homework after class. Without the assistance of teaching assistants, teachers will have a heavy task of correcting students’ homework. On the whole, the year-end workload of the instructors of the Engineering Mechanics Center increases every year. With the advancement of education skills and resources, the instructing hours of courses have been significantly reduced. Meanwhile, the instructing content remains basically unchanged. The number of instructing classes is large, while the number of teachers is relatively small. According to the distribution of the final exam results, the instructing effect of many courses is not optimistic. Various factors have to be considered for education reform activities. Online video-assisted instruction no longer simply depends on instructors to convey knowledge to students. Instructors concentrate on understanding and handling students’ problems and guiding students to leverage their knowledge to solve real-world problems. Through the network platform and various learning resources, students independently plan their learning time and learning rhythm after class. Meanwhile, they complete the knowledge learning themselves, based on which classroom becomes a place for interaction between instructors and students and between students and students. These interactions also involve seminars, answering questions, and using knowledge. Through the education reform of online video-assisted instructing, it aims to the so-called “student-centered” educational concept. It focuses on improving students’ experience at universities, making students’ learning more flexible and active, ensuring students’ participation stronger, and cultivating students’ interest in learning and spontaneous learning habit. It can play an exemplary role and function as a guidance for other curriculum reforms.

As a supplement to higher education, distance education, taking the National Open University as an example, has an annual enrollment of up to 300,000 people and an enrollment of more than 3.66 million people in China [7, 8, 10–12]. The educational levels vary significantly. Even if part-time teachers participating in distance learning provide tutoring to supplement the shortage of in-service instructors [10], it is difficult to meet learners’ requirements timely. Moreover, it is even more difficult to understand students’ learning rhythm and provide recommendations and guidance for students. The implementation of distance education depends to a greater extent on the self-discipline of students. Meanwhile, there is still an insurmountable gap between traditional education and distance education due to the serious shortage of instructors in cultivating and guiding students. The pervasive application of artificial intelligence provides the basic conditions for distance education, especially the training of instructors, which is called Microsoft’s self-driving technique. Currently, Microsoft Translator and Tesla’s unmanned driving have entered people’s field of vision. Meanwhile, Microsoft Xiaobing ChatBot, Baidu Xiaodu, Xiaomi Xiaoi, and Alibaba Tmall Genie have entered the market. Microsoft Translator sneaks into the browser Edge in the form of a plug-in and is pervasively applied in research and meeting sharing. iFLYTEK’s speech recognition can conveniently and accurately convert speech into text, which can facilitate the generation of meeting records and memos. AI has clear advantages in terms of speed, accuracy, load-bearing capacity, repetition consistency, and homework time. For distance education, AI will undertake repetitive tasks such as homework correction, question bank construction, examination, and learning support. AI has become teachers’ favorite assistant [11], which enables more instructors to guide learners in personalized learning. The above platform allows distance education instructors to have more time to identify what kind of people to cultivate. It can also make better use of AI technology to cultivate students. Under the guidance of instructors, students effectively learn knowledge by distance education, acquire skills by solving problems, and develop the habit of independent thinking and collaboration. This can also be utilized to improve students’ psychological diathesis and personal culture. In the AI era, the boundary between instructors and students is becoming more unclear. Every instructor is both an instructor and a learner. As we expected, the future education will be fully learner-centered and we thus have to build learner personalized education. For distance education, we have introduced the AI technology and platform to provide customized education toward each learner. In the AI era [12], “instructing” and “learning” will develop in the direction of “precision education” and “lifelong learning.” Nowadays, we have to face the conversion from learning knowledge to acquiring real-world skills, the conversion from instructor delivering knowledge to guiding students learning, and the conversion from unified instruction to individualized teaching.

3. Our Proposed Method

Our target is to build an education platform by incorporating special equipment for video and teleconferencing. After consideration, we noticed that the equipment and system are expensive and unfeasible. Depending on the existing video surveillance and local network of our unit, we leverage the existing equipment and purchase some additional equipment to design and establish a distance education platform. We notice that the corresponding functions can also be realized, and this scheme is economically efficient. It also exhibits satisfactory practicability. The distance education platform consists of three modules: control
center, local area network, and branch venues. The control center is built at the central site, including three service platforms: video service, dispatching voice service, park broadcast service, and the main venue teaching platform. The video service collects and processes images and videos to provide video services for remote sites. It also dispatches the voice service to ensure voice synchronization between the main venue and each site. Meanwhile, the park broadcast service shares the park broadcast sound between the central site and each remote site. The system architecture of the distance education platform is shown in Figure 3. The local area network relies on the existing video surveillance network to conduct network planning and configuration. The local area network has the feature of simple networking and high transmission rate, based on which the video conferencing system will be easier to achieve good operation results. The branch venues are arranged at remote sites, and the surveillance camera is responsible for collecting the venue images and sending them back to the main venue. Based on this, the monitoring computer provides the main venue images for the branch venues.

The service platforms are built at the central site to provide effective services for the central site and remote users. The video surveillance server is the core of the video surveillance local area network, which collects, processes, and stores remote backup information. There is a video surveillance server in the central site, which supports TCP/IP, multicast, and other network protocols. Afterward, the ONVIF network video standard specification is operated, which can be remotely configured and managed through a Web browser. It can be operated 24 hours each day, based on which the operation is stable and reliable. Online users can use the monitoring client to view the required real-time images and playback images according to the permission level. In our framework, we install one camera and one computer in the study room of the remote site. The video and audio signals of the conference room can be collected through the camera and uploaded to the server. Simultaneously, the computer can access the server through the client, watch the required image source, and access the server through FTP to download and upload the required data.

In our framework, we incorporate an IP network campus broadcast system audio acquisition host at the central site to collect audio signals. We also encode them into digital signals and transmit them to remote terminals via the network. We break through the distance limitations and realize the sharing of campus broadcast signals. The park broadcasting terminal is installed in the broadcasting room of the remote unit, and the audio cable can be connected to the outdoor sound to listen to the broadcasting sound of the server. A certain type of dispatching host that has been eliminated and is idle is connected to the video surveillance network to realize the separation of conference dispatching and dedicated dispatching services without affecting each other. The dispatching unit in the study room of remote units is also changed from dedicated data transmission network access to video surveillance network access to reduce the probability of network crossover. The video surveillance local area network is an internal independent network, which is physically isolated from other data transmission networks such as the Internet. It has been extended to each remote site. Its network architecture is displayed in detail in Figure 3.

Next, we construct the main venue teaching platform. The main venue has an LED screen, multiple computers, and a set of auxiliary audio equipment. On this basis, an instructing platform is built. Two computers are leveraged in the main venue as the instructing computer and the video-on-demand computer, respectively. The output signal of the teaching computer is divided into two parts, one of which is sent to the LED screen for display. Meanwhile, the other is sent to the HDMI signal encoder, which is subsequently sent to the monitoring server after the encoding process. The video-on-demand computer can view the images of the branch venue and the instructing computer in real time by accessing the server. It can also display multiple screens. The output signal of the monitoring computer is divided into two parts: one is sent to the LED screen for display, while the other is sent to the encoder and further to the Dahua monitoring server after being encoded. When sending the signal source to the LED screen, it is feasible to switch to choose the main teaching screen or the monitoring screen. The schematic diagram of device connection is shown in the next section. In our work, we aim to search for the $k$ most frequently occurring patterns for distance learning, that is,

$$T(\text{item}) = \min\{\text{count}, 2^{l-1}\},$$

(1)

where $T(\text{item})$ denotes the load function and count indicates the number of occurrence of each item.

Our constructed network configuration management is a very important part of network management. Its purpose is to improve the running state of the network and realize a specific function or optimize the network performance [13]. The video surveillance network is a small internal local area network connecting four areas, based on which the network scale and number of hosts are small. According to the IP address and subnet mask addressing specification, in order to ensure that the IP addresses in the network are not in conflict with each other, the address resources will be completely utilized. Meanwhile, we will take into account the convenience of the administrator for terminal and address management, and the IP addresses are continuously divided.

In our distance education system, there are four slices, and the address segment uses 192.168.1.X-192.168.4.X, respectively. There are three types of services, each of which can be divided into different subnets. The number of monitoring terminal hosts is fewer than 120, and the 25 bit subnet mask is utilized, that is, 192.168.X.0/25. The number of hosts for scheduling services and broadcasting services is in single digits. Based on this, the 27 bit subnet mask is leveraged, that is, 192.168.X.128/27 and 192.168.X.160/27. The undivided subnet part of each segment can be utilized as the requirement for the subsequent expansion services. The VLAN division is constructed based on ports. Virtual local area network (VLAN) [4] is an Ethernet technology that logically divides a physical LAN into multiple virtual LANs. It can achieve the purpose of reducing the broadcast domain.
VLAN is a logical segmentation of network users connected to the second-layer switch port. Also, it is a network segmentation based on user requirements without being restricted by the physical location of network users. By dividing the different VLANs, during the information exchange, the broadcast information can only be replicated by the same VLAN. Also, it cannot span across different VLANs, which can block broadcasts more conveniently and effectively. It can also improve the performance during the switch process. By leveraging the VLAN technique, we combine the data link layer and network layer switching equipment which can build a safe and reliable network. There are three main schemes to divide the VLAN based on switched Ethernet: (1) based on port, (2) based on MAC address, and (3) based on IP address. The port-based VLAN division method takes the switch port as the unit, a VLAN corresponds to a group of ports of the switch, and the communication within the VLAN is conducted within the range of this group of designated ports. Meanwhile, the connection with other ports is restricted, and its advantage is that the configuration is simple. In this way, the network is built upon the port-based VLAN division.

### 4. Experimental Results and Analysis

**Experiment 1.** At the end of the semester, instructors should evaluate the courses in the form of self-evaluation reports. They will also conduct data analysis and evaluation of students’ learning effects. If there are parallel ordinary classes, the instructors should conduct comparative evaluations. The evaluation content includes instructing content, instructing methods and means, student assessment and evaluation, and student learning effect. It is necessary to realize the analysis and reflection on the implementation of the course.

In this experiment, we analyze the online video-assisted teaching effect. The Mechanics Center counted the number of teaching classes using various teaching methods in 2016 and 2017. We can also calculate statistics based on the results of each instructing class in each semester to analyze the instructing effects of multiple instructing schemes. Table 1 elaborates the number of classes with different teaching methods from the 2016 and 2017 material mechanics teaching classes. All teaching classes utilize the same final exam papers. According to the statistical results of the final exam results of material mechanics in the past two years, the instructing effect of different teaching methods is evaluated and analyzed in each semester. Table 2 shows the average grades of classes with different teaching methods in 2016 and 2017. From the average of the final exam grades, the grades of the classes taught in the traditional classroom are the lowest, followed by the classes taught by the flipped classroom. Classes that combine traditional classroom instruction and online video teaching have the highest grades.

Questionnaires were distributed to investigate students’ appraisal of different instructing strategies and the relevant opinions/suggestions. The results have shown that a vast majority of students will use online video for learning. They believed that online video has a positive effect on preview and review after class. Different videos can attract students to different degrees. Videos containing engineering application backgrounds and life examples are more attractive to students and can improve students’ motivation of learning material mechanics. Online video-assisted teaching helps to improve students’ interest in learning and improve their participation in learning material mechanics. The length of each class in traditional classroom teaching is generally 45 minutes, which makes it difficult for students to concentrate for a long time. However, online video-assisted teaching allows students to study in sections, making it easier for students to concentrate on learning a certain knowledge point. In video-assisted learning, the duration of each video should be controlled within 15 minutes, and the video should be recreated according to the knowledge points. Students are required to complete the relevant exercises or homework before learning the next knowledge point. This can be utilized to improve the learning progress of students using video-based distance education. In addition, a few
problems were also observed during the online video-assisted teaching. How to ensure that students have a higher degree of participation in distance education to enhance student learning, how to ensure that students have learned all the video content, and how to make the video more attractive to students are issues that instructors have to tackle when taking educational videos.

Experiment 2. A total of 100 patients with type two diabetes who visited the community health service center from Shanghai were selected [14–16]. The selection criteria are given as follows: there should be type 2 diabetes criteria; aged between 50 and 80; gender is not limited; glycosylated hemoglobin >6.5%; vision and hearing are guaranteed to accept online courses; and able to use the Internet or mobile smart devices. Exclusion criteria are given as follows: type 1 diabetes mellitus; type 2 diabetes mellitus with severe chronic complications of diabetes; cognitive impairment; unable to take care of themselves; and education level below junior high school. We do not use smart mobile devices. The patients were randomly divided into a control group and an experimental group (online video micro-classroom group), with 50 cases in each group. Herein, methods of all patients were diagnosed and treated according to the clinical pathway of diabetes. Also, no special intervention was required for the study.

When the fasting blood glucose was >10 mmol/L, the patients were referred to the endocrinology department of a tertiary hospital for treatment. The control group did not receive diabetes self-management education except for routine outpatient consultation and communication. When the experimental group was enrolled, they were given a brochure of “Diabetes Self-Management Video Course.” This course contains 20 micro-course video URLs and can be scanned by the QR code. The video course was designed and produced by the Department of Endocrinology from Tongji Hospital. It consisted of 20 micro-course videos, each of which lasts 5–10 minutes long. The key topics include basics of diabetes, diet therapy, exercise therapy, day-to-day care and problem solving, and medication for patients with diabetes. Doctors in the receiving community instruct patients on how to obtain video resources. The doctors called back once in every two weeks to learn about video viewing and answer patients’ questions. Before the study and 6 months after the study, the height, weight, and blood pressure data of the two groups of patients were collected. Meanwhile, the fasting venous blood was collected, and the fasting blood glucose, glycosylated hemoglobin, total cholesterol, triglyceride, low-density lipoprotein, and high-density lipoprotein were measured.

For the protein detection, after six months of the study, the patients in the control group and the experimental group were evaluated for their knowledge and abilities [17–19]. Meanwhile, the evaluation was completed by community general practitioners in the form of questionnaires. The content is divided into two parts, including basic knowledge of diabetes and problem-solving ability, each of which has 20 multiple-choice questions. Each question contains five key points. The problem-solving ability is case or scenario questions. This study is a community outpatient study because self-monitoring of blood glucose is uncommon in patients with type two diabetes (20.6% of patients in this study had their own blood glucose meters). Also, it is difficult to accurately collect data such as daily blood glucose assessment, blood glucose fluctuation, and confirmation of hypoglycemia events. It does not include pre and poststudy assessments.

Based on the above analysis, statistics obtained from the SPSS 20.0 software were utilized for an in-depth data analysis. Measurement data were calculated in the form of \( x \pm s \). The independent sample evaluation was utilized for comparison between the two groups. Also, the test was utilized for comparison of enumeration data between groups. The baseline indicators of the two groups are shown in Table 3. There was no significant difference in the metabolic indicators between the two groups. There was no difference in the baseline body weight, body mass index, and blood pressure between the two groups. The evaluation after six months of study showed that the glycated hemoglobin and fasting blood glucose of the experimental group were significantly lower than those at the time of entry. Simultaneously, the fasting blood glucose and glycated hemoglobin of the control group decreased slightly. Noticeably, the improvement of glycated hemoglobin and fasting blood glucose of the experimental group was higher than that of the experimental group. Simultaneously, the control group was more significant \((P < 0.001)\).
5. Conclusions

This paper introduces a novel distance education framework in the course instruction reform by leveraging online video mining based on AI. The algorithm proposed by us can play a greater role in the talent training process in universities by virtue of its powerful simulation and online debugging capabilities. Our proposed framework can greatly promote the timeliness of the instructing quality. In order to evaluate our designed framework, we apply the structured diabetes self-management video course education to help the elderly diabetic patients treated in community health service centers. We observed that our method can enhance patients’ blood sugar control, diabetes knowledge, and problem-solving ability. We also noticed that systematic diabetes education is a feasible way of health education, and it has some promotion value. Meanwhile, the distance education of single-chip microcomputer also confirms the competitiveness of our distance education framework.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

References


