Development and Application of Audio-triggered Courseware in Remote Music Teaching

Lilian Liu

Hunan University of Humanities, Science and Technology, Loudi, China

Abstract—**This study explored the limitations of contemporary distance video teaching, including unsynchronized audio/video, lack of interaction, and broadband occupancy. The theoretical basis and supporting technology for developing an audio-triggered courseware teaching system were discussed and the corresponding system was designed and developed. The author applied this system to teach a course on Guzheng specialty at the Department of Music and designed a control experiment to verify the effects of this system. Results show that the effect of audio-triggered courseware teaching is significantly superior to that of a traditional distance video teaching system. When students cannot attend classes, using the audio-triggered courseware teaching system may be a preferable option.**

Index Terms—**distance teaching; audio; synchronization; internet**

I. INTRODUCTION

Benefitting from the rapid development of the Internet, remote teaching has been widely applied in the contemporary education arena [1]. In recent years, there are also many universities that have started distance music education in order to enable students to come into contact with the very best instrumentalists and pedagogues [2]. Nowadays, distance music education for advanced students is a developed area, for example at Canada's National Art Centre, IRCAM (Paris), and the Australian National University School of Music [3]. Orman and Whitaker [4] suggested that eye contact was more common in distance teaching compared with face-to-face lessons. In the traditional remote teaching model, the Internet throughput and network transmission delay and other factors at different time change continuously, which leads to the desynchrony between transmission video signal grouping and its corresponding accompanying sound signal grouping, generating labial desynchrony [5]. This limits the popularization and application of remote teaching to certain level, especially for the professional music teaching, which is sensitive to sound signal. This kind of drawback is unacceptable.

Seo et al. [6] developed a scalable extension of H.264, known as scalable video coding (SVC), which had been the main focus of the work of a joint video team. The scalable extension, which was finalized at the end of 2007, proposed an efficient media synchronization mechanism for SVC video transport over IP networks. To support the synchronization of video and audio bit-streams transported over W networks, a Real-time Transport Protocol (RTP)/RTP Control Protocol suite is typically employed. Sang et al [7] showed that the algorithm based on variable-sized matrix segmentation and 3D matrix wide discrete cosine transformation can effectively synchronize video and audio as well as improve compression and synchronization. Casanovas and Cavallaro [8] presented a multimodal method for the automatic synchronization of audio–visual recordings captured by using a set of independent cameras. The proposed method jointly processes data from audio and video channels to estimate inter-camera delays, which are used to align the recordings temporally. An important feature of the study is estimating the confidence level of the results to enable us to reject unreliable recordings automatically for alignment. Tao et al [9] proposed a self-synchronized audio watermarking algorithm based on psychoacoustic model. When watermarking information is hidden, synchronization signals are embedded to enable watermarking to own synchronous ability. According to the psychoacoustic model of human auditory system, when watermarking signals and synchronization signals are embedded in low and medium frequency wavelet coefficient, the embedding strength is self-adaptively controlled by masking threshold. Wang et al [10] utilized USB analog TV box to collect closed-circuit TV programs and encoded them as T-DMB programs to enable real-time broadcast. This paper mainly studied the signal collection method of USB TV box and audio/ video synchronization in program encoding.

The innovation of this study was to apply audio/video synchronization technology in remote teaching; it put forward an audio-triggered courseware-based remote teaching method and constructed an audio-triggered courseware teaching system, which not only solved the current drawbacks of remote teaching, including unsynchronized audio/video, lack of interaction and broadband occupancy, but also owned the characteristics of broadband saving, convenient use and real-time synchronization.

II. THEORETICAL BASIS AND SYSTEM SUPPORT TECHNOLOGY OF AUDIO-TRIGGERED COURSEWARE TEACHING

A. Multiple stimulation theory

The brain stimulation of traditional teaching methods on learners mostly is single while the brain stimulation of audio-triggered courseware teaching method on learners is multidimensional. It can simultaneously mobilize vision, hearing and touch and other senses of learners to generate multiple excitatory areas on the cerebral cortex and the learned contents can quickly enter from sensory memory into deep memory. In this way, it can realize the aims of increasing efficiency and consolidating knowledge.

B. Theory of cognitive psychology

Attention, feeling, perception, memory and awareness are psychological activities generated from understanding

objective things and they are called cognition process in psychology. Audio-triggered courseware-based teaching shall follow human psychological cognitive development law for achieving good teaching efficiency.

C. Theory of hearing psychology

In human senses, hearing owns unique psychological phenomena, such as auditory break, sound intensity adaptation, auditory fatigue. Auditory break phenomenon refers to the auditory behavior that has an excitement boot threshold; so audio-triggered courseware teaching needs to predict the psychological states of listeners, avoiding auditory break. Sound intensity adaptation phenomenon refers to the emerging moment of audio resources that will generate an excitement on cerebral cortex. When the audio resources continue, the excitement on cerebral cortex will decrease and excitement will transfer to new sound stimulations. Therefore, audio-triggered courseware-based teaching shall focus on the intensity design of audio resources. Auditory fatigue refers to auditory dysfunction, and even hearing loss or tinnitus caused by exposing the auditory organ to one sound. Thus, audio used in teaching shall avoid using audio resources that are easy to cause auditory fatigue.

D. Internet technology

The Internet is global and interconnected by a group of universal protocols and covers the whole world. The Internet can exchange information and remains free from the limitations of space while sharing information. The cost of information exchange is low, which makes virtual classroom communication possible.

With the development of technology, the Internet tends to be mobile, cloud-based and universal, which provides more development options for virtual classroom.

E. Multimedia technology

Multimedia technology refers to the real-time information interaction technology between users' various senses and the computer by comprehensively processing and managing various media information, for example, texts, data, graphics, images, animations, and sounds.

F. Audio compression technology

After audio signal digitalization, data size will be huge, which is not good for network transmission and storage. Thus, in remote education, audio needs to be compressed. At present, the common audio compression encoding standards include MPEG, AC-3, DTS and others.

G. Video frame transmission technology

Video provides an illusion of continuous actions for eyes by playing a series of frame pictures based on the principle of human visual persistence. In videos, each frame is static picture while fast playing of frames continuously gives rise to the illusion of action.

III. AUDIO-TRIGGERED COURSEWARE-BASED REMOTE TEACHING SYSTEM

A. System functions and characteristics

This paper studied and developed an audio-triggered courseware-based remote teaching system. This system can keep synchronization between courseware and audio resources, and at the same time, it can save bandwidth and be easy to use [7].

B. System design

1) System architecture design

This system adopts B/S architectural pattern, namely, browser/server model, shown in Figure 1.
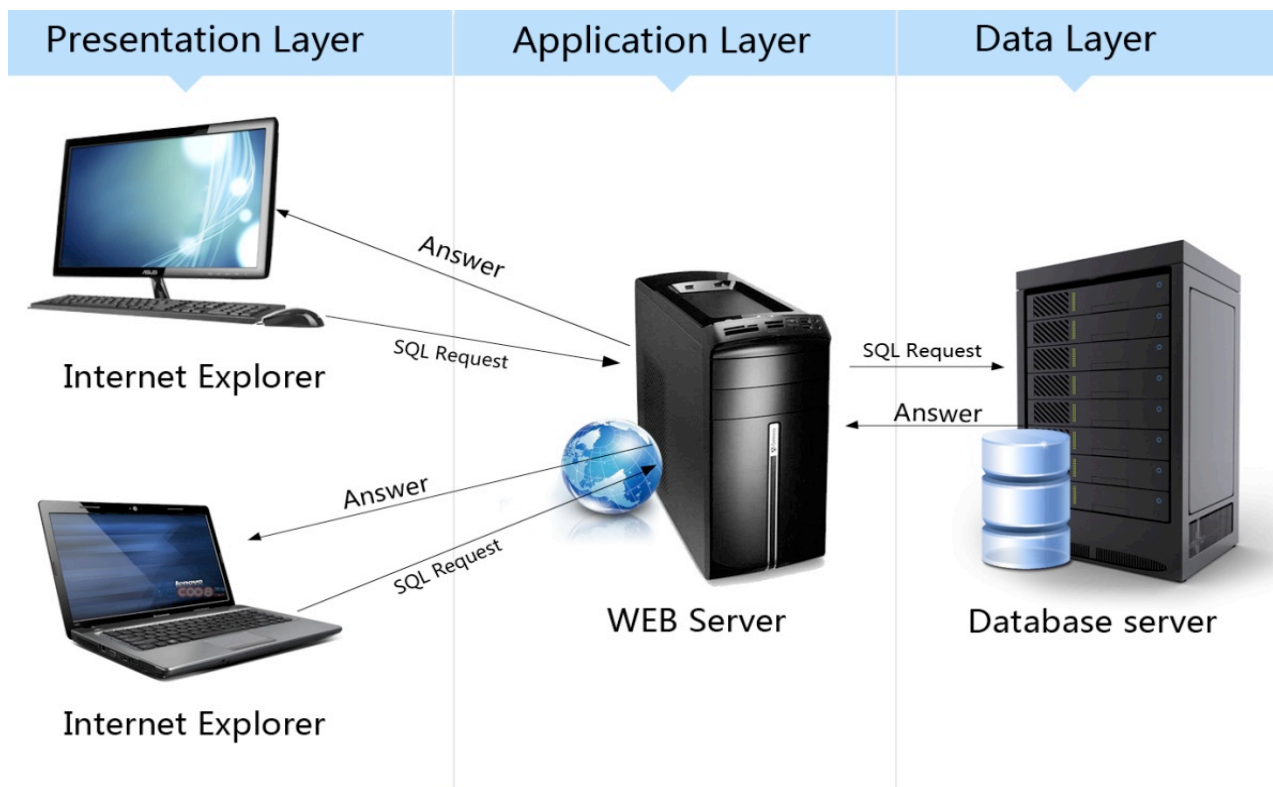

Figure 1.   Schematic diagram of B/S architectural pattern

The system is divided into three-layer structure, namely, database server, WEB server and presentation layer. Data layer is mainly a database server, saving various data required by the system, for example, courseware and audio resources. Application layer is mainly about WEB server to manage each application module in the system. In the presentation layer, students and teachers can visit various applications of remote teaching system through Internet Explorer in a computer.

2) System module design

The system includes student management module, virtual classroom (interface), teaching management module, and learning management module.

Teaching management module consists of virtual teachers and teaching database. Among them, virtual teachers give lessons to students; and teaching database is used to save courseware, audios, exercises, tests, and keys for teaching. At the same time, to guarantee the synchronization between courseware and audio, teaching management module needs to establish the relevance between courseware and audio and save it in the teaching material base.

Student users utilize learning management module to select learning contents and enter into the virtual class-room to have lessons, exercises or tests. Student management module interactively obtains data of other users through database server via the Internet and provides data about other students to virtual classroom, forming an environment of many people studying at the same time.

Student management module includes student database, course-choosing registration base, class-attending historic database, and student grade data. Among them, student database mainly saves personal information and other information of students. Course-choosing registration base mainly records remote teaching courses of students. Class-attending historic database mainly registers students' learning history and attendance. Student grade database mainly records students' exercises, grades, ranking and other information.

Virtual classroom is the presentation interface of the system, facing teachers and students respectively and constructing virtual classroom learning environment.

3) Workflow design of system server

System server mainly provides courseware and the corresponding audio and other materials, which requires the synchronization between courseware and audio. The schematic diagram of system server side process is shown in Figure 2.
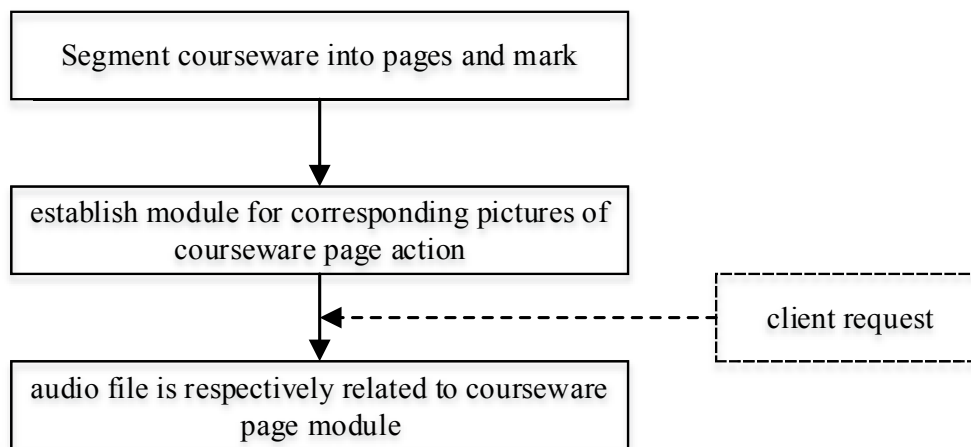

Figure 2. Schematic diagram of system server side process

The system server segments the courseware into pages and each page establishes a globally unique SildeID mark. At the same time, frame module is established for the corresponding pictures of courseware action as well as FrameID mark. To guarantee the synchronization between frame module and audio, the server establishes the relevance between audio files and SildeID and FrameID. If the server receives the page request from a client side, it will output audio files related to the courseware page module. If the server receives the client request of page action, it will output frame module and related audios.

4) Workflow design of system client side

If the system client side receives customer's operation command from users, different service requests will be made according to different user commands. The schematic diagram of system client process is shown in Figure 3.

Client side judges user request types according to customer's operation command. If user request is a page, client side and server interact internally to request the server side to transmit audio files related to the page.

After receiving files, page and related audio files are synchronously outputted to users. If user request is page action, client side will request the server side to output frame modules and related audio files. After receiving files, frame action and related audio files are synchronously outputted to users.

At the server side, the relevance between page and audio and the relevance between frame and audio are established. After the corresponding request, files are received by the server and are synchronously outputted to users, which guarantee the synchronization between playing the courseware and accompanying sound, thereby avoiding time delay.

C. System implementation

Through the above mentioned system support technology and according to relevant designs, program development is done to realize the system functions. Figure 4 and Figure 5 respectively show the relevance between courseware page and audio frequency, and courseware animation and audio frequency with the aid of the system.
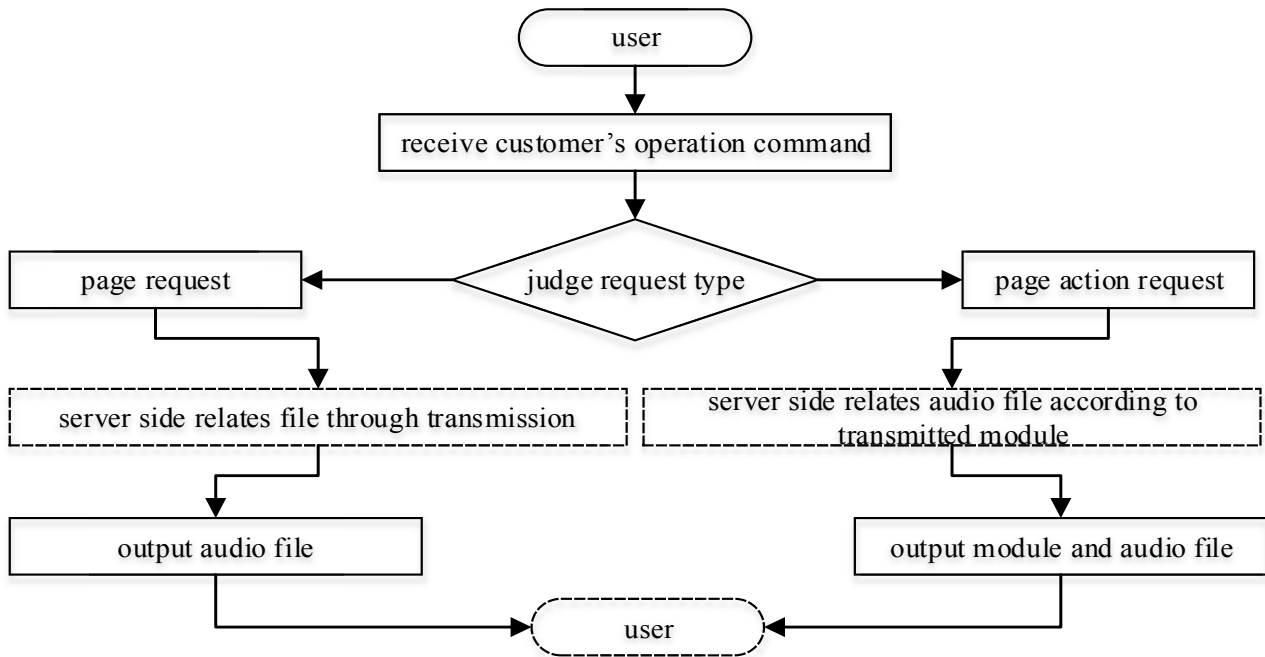
Figure 3.   Schematic diagram of system client process
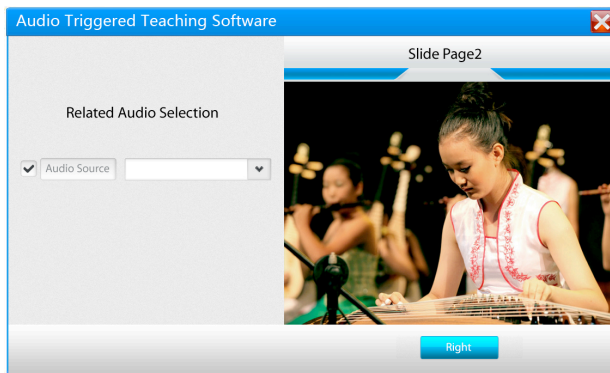

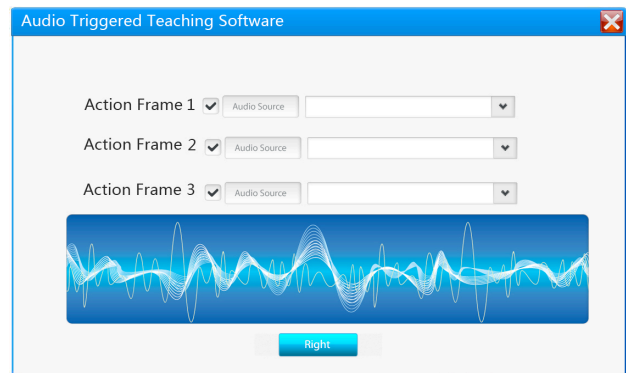Figure 4.   Establishment of relevance between courseware page and audio frequency

Figure 5.   Establishment of relevance between courseware animation and audio frequency

IV.   EFFECT VERIFICATION

To verify the availability of audio-triggered courseware-based teaching system, the author applied the system in a course of Guzheng specialty in the Department of Music on a pilot basis. At the same time, the author designed a control experiment for comparison to verify the effects of this system. It respectively selected three groups of students to take part in this course. The first group was real classroom group, where students attended classes in real classrooms at school. The second group was remote class attending group. Students attended classes of remote video through traditional remote education method. The third group was remote courseware group, which studied through the audio-triggered courseware-based teaching system designed and implemented in this paper. The author respectively assessed the students' Guzheng performance level through four control dimensions, intonation, tone, volume and rhythm. Each dimension carried 25 points and the total score was 100 points.

The average test scores of the three groups are shown in Table 1.

TABLE I.
AVERAGE SCORE OF EXPERIMENTAL RESULTS

Group	Intonation	Tone	Volume	Rhythm control	Total
Real classroom group	24	24	22	23	93
Remote class attending group	21	22	20	18	81
Remote courseware group	23	23	22	23	91

V. CONCLUSION

The score of remote class attending group was lower than the other two groups, especially, in rhythm control. The reason may be that, it was easy to have time delay between video and audio in traditional remote video teaching, and therefore the learning experience was poor. The score of students in real classroom group differed slightly from students in remote courseware group. However, audio-triggered courseware teaching, namely, distance teaching, lacked learning atmosphere compared with real teaching environment; consequently its teaching efficiency was slightly lower than that of real classroom group. On the whole, the effect of audio-triggered courseware teaching was much superior to traditional distance video teaching system. When students cannot normally attend classes in classroom, audio-triggered courseware teaching system may serve as a good alternative.

REFERENCES

[1] Barratt R.S., "Meeting lifelong learning needs by distance teaching: clean technology," *Journal of Cleaner Production*, vol. 14, no. 9-11, pp. Pages 906-915, March 2006.

[2] Herbert D., "Five challenges and solutions in online music teacher education," *Research and Issues in Music Education,* vol. 5, no. 1, pp. 1-10, September 2007.

[3] Brandstrom S., Wiklund C., & Lundstrom E., "Developing distance music education in arctic scandinavia: electric guitar teaching and master classes," *Music Education Research*, vol. 14, no. 4, 448-456, December 2012. http://dx.doi.org/10.1080/14613808.2012.703173

[4] Orman E.K., Whitaker J.A., "Time usage during face-to-face and synchronous distance lessons," *American Journal of Distance Education*, vol. 24, no. 2, pp. 92-103, June 2010. http://dx.doi.org/10.1080/08923641003666854

[5] Messaoudi T., Bodin F., Diaz J.J.H., Ichihara S., Fikry T., Lacreuse I., et al., "Evaluation of a new eLearning platform for distance teaching of microsurgery," *Chirurgie de la Main*, vol. 34, no. 3, pp. 109-112, June 2015. http://dx.doi.org/10.1016/j.main.2015.02.002

[6] Seo K.D., Jung S.H., & Kim J.S., "Efficient media synchronization mechanism for svc video transport over ip networks," *Etri Journal*. vol. 30, no. 3, pp. 441-450, June 2008. http://dx.doi.org/10.4218/etrij.08.0107.0270

[7] Sang A.J., Gabbouj M., Chen H.X., "Video and Audio Synchronization Coding Based on Three Dimensional Matrix in Ubiquitous Computing," *Lecture Notes in Electrical Engineering*, vol. 128, pp. 87-93. 2012. http://dx.doi.org/10.1007/978-3-642-25792-6_14

[8] Casanovas A.L., & Cavallaro A., "Audio-visual events for multi-camera synchronization," *Multimedia Tools and Applications*, vol.74, no. 4, pp. 1317-1340, February 2015. http://dx.doi.org/10.1007/s11042-014-1872-y

[9] Tao Z., Zhao H.M., Gu J.H., Tan X.D., "self-synchronizing audio digital watermark based on psychoacoustic model," *Microelectronics & Computer*. vol. 23, no. 4, pp. 27-30, April 2006.

[10] Wang G.Y., Zhang Y.M., Zhang H.S., "audio-video acquisition used to relay T-DMB television program and synchronous design," *Electronic World*, vol. 5, pp. 144-146, March 2013.

AUTHOR

Lilian Liu (Corresponding author) is a lecturer in Music Dance Department, Hunan University of Humanities, Science and Technology, Loudi 417000, China. Her research interests include Guzheng teaching and offline education (liulilian2015@163.com).