Multiple Conference Support Method Using Variable Speed Playing and Video Image Switching

Aki Sugiyama, Kuniaki Suseki and Ken-ichi Okada
Faculty of Science and Technology, Keio University
3-14-1 Hiyoshi, Kohoku-ku, Yokohama, Kanagawa, Japan
(sugiyama, suseki, okada)@mos.ics.keio.ac.jp

Abstract

In this research paper, we propose a new style for conference participation that aims to allow users to participate in two remote video conferences simultaneously. Based on the approach mentioned above, we are able to allow the users to participate in two remote video conferences simultaneously. This method compensates for the users' imperfect memory, which is caused by the switching. Furthermore, from the results of the evaluation experiments, we confirmed that the users are able to watch the two conference videos without significantly decreasing the level of comprehension regarding the contents of conferences.

1. Introduction

In late years, there are many opportunities for people who are non-professional director to shoot and edit video. In addition to camera work, switching is also important for rendition. Kato et al propose an automatic camera switching video images shots automatically between multiples cameras, to produce video images, which do not bore, nor discomfort the viewer due to unnatural switching [1]. Tsumura et al propose gAutomatic Switching Technique of Remote Conference Video Image for Multitasking Workers with Multiple Camerash using time shifting aiming on support on desk work and watching remote conference [4] . Takata et al propose gAgenda based Multiple Work Support for Video Conferencing Participation and Deskworkh [5]. This research aimed for multiple work support of deskwork such as documentation and attending remote conference simultaneously using conference agendas. When we view the difficulty level of multiple work in terms of controllability, the most difficult work is combination of interaction works in terms of the overlapped audio and the grasp of the content and the task switching. Therefore, in this research, we aim for the multiple work support of the combination of interaction work.

2. Proposal

In this research, we set the work environment using a desktop display. The worker carries out the multiple work of alternately watching two conference videos to participate in two remote conferences. As mentioned above, the workers' environment assumed in this research is the multiple work environment in which they participate in two remote conferences simultaneously. The concept of the proposal in this research project is the paradigm shift from single to many. The details regarding the factors for working out this concept are as follows: The compression of the gdensityh of the video by shortening the video play toe,(The time compression of the video). We define this shortened video playing as watching the video while it is being fast-forwarded to shorten the playing time of the video more than that of the original. We define the density of the video as shown in the following equation;

$$Density\ of\ video \quad = \quad \frac{Actual\ time\ of\ video}{Playing\ time}$$

Consequently, by shortening the playing time of the video, the density of the video is increased. Therefore, spare time is created and the video is converted into a video with highly-concentrated contents. Next, we need to study the most appropriate video switching interval in order to propose a new conferencing participation style of to use when alternately watching multiple videos in the same time frame for a short period of time, similar to the technique used to watch multiple videos using alternate switching. Consequently, we actually measure the appropriate value for the switching interval through evaluation experiments and then we introduce the measured value. By doing so, we try to successfully enable multiple videos to be watched. Finally, we introduce an approach that supports video restarting by using overlapping when switching the videos. If the videos are switched automatically after the video switching interval passes, the videos will be broken up during conversation. Consequently, when the video restarts from the position immediately after the break, it becomes difficult to

ICAT 2008
Dec. 1-3, Yokohama, Japan
ISSN: 1345-1278

comprehend the content regarding the conference. To solve this problem, the video is overlapped for a minute time, Δt, and then restarted. By doing this, the problem of imperfectly memorized content is solved and workers can alternately watch multiple videos using shortened video playback for a short period of time without incident. Using the above mentioned techniques, participation in two remote conferences at the same time becomes possible and we can propose the new multiple work style.

3. Pilot Study

3.1. Experiment for appropriate playback speed

The purpose of this experiment is to determine the upper limit of the playback speed that does not impair the level of comprehension of the contents, and the limit of the auditory information-handling ability of the viewers. In the case of video information, it is difficult to obtain the same amount of information in a short amount of time, as opposed to quickly reading through text information. However, the video information is valuable because it contains more information than text. So, we used the audio correction technique called TSM (Time Scale Modification) to preserve the value of the video information when the video playback speed was changed.

In this experiment, we had 16 subjects watch the conference video played at six different playback speeds: 2.0×, 2.2×, 2.4×, 2.6×, 2.8×, 3.0× speed (for example 2.0× defines playing at double speed). By doing so, we researched the appropriate playback speed value. The video image is a passive conveyance conference and the topic of the conference is freshman staffing. The subjects ranked the playback speeds that they considered appropriate. Specifically, they ranked the speeds in order of playback speed from best to worst, assigning 5, 4, 3, 2, 1, or 0 points. Consequently, the average of this result is used in the evaluation. Table 1 shows the results of this evaluation. The value of 2.4× playback speed was the highest in this result. However, in consideration of safety, the appropriate playback speed value was set to 2.2× speed for this research project.

Table 1. Appropriate Playback speed Evaluation

Playback speed	2.0×	2.2×	2.4×	2.6×	2.8×	3.0×
Score	2.3	3.4	4.1	3.1	1.6	0.4

3.2. Experiment of switching conference video

Next, we research the appropriate interval at which to switch the video image. We set the interval to switch the video image at: 5, 10, 20, 40, 60, and greater than 60 seconds. The conference video was played at the 2.2× speed that was obtained in the previously mentioned experiment. The topics of the conference in this experiment were travel planning and freshman education. The subjects alternately watched these two conference videos at the previously mentioned switching interval, and we researched the appropriate switching interval. The evaluation method used was the same as the previously mentioned experiment. The results are shown in Table 2.

Table 2. Switching Interval Evaluation

Switching interval (sec)	5	10	20	40	60	greater than 60
Score	0	1	2.6	4.7	4.1	2.5

3.3. Experiment to test the usability

We performed the following experiment which attempts to examine the amount of influence that variable speed playing has on the userfs level of comprehension of the video contents. In this experiment, two videos played at 2.2× the normal speed and 1.0× the normal speed were used. We examined the level of comprehension for each video and compared these results. We define the level of comprehension of the conference as the number of keywords answered correctly out of the total keywords. The topics of the conferences in this experiment were “comfortable death” and “travel planning”. Each topic had two minutes of content. The evaluation item is the level of comprehension regarding the conference. Examples of the questions asked are: gWhat is the problem with the concept of comfortable death?h and gWhere was the travel destination. Table 3 shows the results of this experiment.

Table 3. Usability of Variable Speed Playing Evaluation

Playback speed	Comfortable death	Travel planning
1.0×	65.1	82.5
2.2×	63	79.1

The previously mentioned results of the pilot study can be summarized as follows:

Appropriate playback speed experiment: A playing speed of 2.2× the normal speed is appropriate Experiment of switching conference video: 40 second-interval of switching is appropriate Experiment for usability of variable speed playing: the playback speed do not influence the level of comprehension

However, a problem occurs when the video image is switched at an interval of 40 seconds. In next section, we will describe this problem and how it can be resolved.

4. Problem in switching video

4.1. Timing of switching video

Specifically, when the user participates in Conference A and Conference B simultaneously, the video is switched based only on time without regarding the content of the conversation. Consequently, when the user starts to watch the remainder of the video from Conference A, it becomes difficult for the user remember the content of the conference smoothly because the video was broken up in the middle of the conversation.

4.2. Video Restarting Support

To solve the previously mentioned problem, the video is overlapped for a minute time of Δt seconds and is restarted to play as shown in Figure 1. We define absolute time as the time in the real world, and the relative time as the time relative to the video content in Figure 1. Consequently, the method allows the conference content to be restarted. This solves the problem of ambiguity of the memorized content in terms of human memory. If the users watch the videos alternately for a short time using our proposed method, the problem caused by switching the video is does not occur.

4.3. Calculation of overlap time

In this section, we will describe the method used to calculate Δt, with the goal to make it possible to finish watching two videos alternately in the same time frame as watching a single video. We set the overlap time, the shortened playback speed, the time of the video, and the video switching interval to Δt(sec), k, T(sec) (we assume that the times of the two videos are the same for ease of calculation) and 40 (sec) respectively.

Two video overlaps are caused per switch set. The total overlap time can be expressed as the number of switches $\times$ $2\Delta t$. (T/40 - 1) switching is caused per video, so the total overlap time is $2\Delta t \times$ (T/40 - 1). Consequently, the following equation must be satisfied to finish watching two videos in the same time that it is required to finish watching one video:

$$\frac{2 \times [\frac{T}{40} - 1] \times \Delta t + 2T}{T} = \frac{T}{1}$$

Therefore, Δt must satisfy the following equation:

$$\Delta t = \frac{20T \times (k-2)}{T - 40}$$

The graph depicting the relationship between the overlap time and the time of the video is made from the mathematical formula of Δt. Here, T is considerably larger than the 40-second switching interval. Therefore, there is no problem with the calculation of Δt when adhering to the assumption that T is infinite. The result of the calculation is

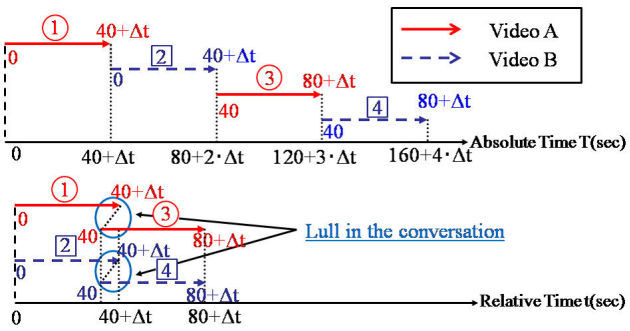

Figure 1. Restarting Support

that the value of Δt is 20(k-2) seconds. When this value is set to Δt of the overlap time, the viewers can finish watching in a shorter time than when watching at 1.0$\times$ normal speed.

5. Implementation

5.1. Algorithm of video outputting

Each part of the configuration diagram is described below. Two videos are stored temporarily in the short term video storage device. The stored videos are input into the variable compression device. The time-compressed video and audio are stored temporarily in the compressed video storage section . The videos that have to be output are selected from the stored videos through the video selecting section. The video and audio are output at the appropriate switching intervals through the video output section. The overlap handling section operates for the videos stored since two round and the videos and audios overlapped in switching are alternately output. The users can shift the switching interval of the video and the playback speed in the user operating section. The values are incorporated into the system through the control section.

Based on the previously mentioned system configuration, input Video A and Video B are actually output as shown in Figure 2. First, one cycle of Video A and Video B as the input videos is stored. Next, the stored video is compressed. Finally, the compressed video is output.

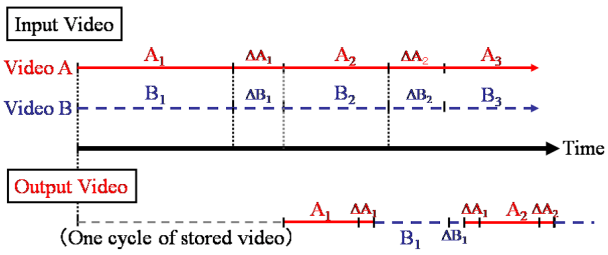

Figure 2. Relation between Input Video and Output Video

Here, we define the overlap time of $A_n(B_n)$ as ΔA_n (ΔB_n). Furthermore, we define the compressed $A_n(B_n)$ and $\Delta A_n(\Delta B_n)$ as $A'_n(B'_n)$ and $\Delta A'_n(\Delta B'_n)$ respectively. One cycle is defined as (A'_1 $\Delta A'_1$) only in the first round and defined as ($\Delta A'_n$ A'_{n+1} $\Delta A'_{n+1}$) (n 2) from the second round on. For example, ($\Delta A'_1$ A'_2 $\Delta A'_2$) in the second round and ($\Delta A'_2$ A'_3 $\Delta A'_3$) in the third round.

When we set the video switching intervals and Δt of the overlap time as 40 seconds and 4 seconds respectively at 2.2$\times$ playback speed, the time for $A'_1(B'_1)$ is set as 40 seconds and $\Delta A'_1(\Delta B'_1)$ is set as 4 seconds.

Furthermore, the users watch Video A' again after watching each A'_1 and B'_1. Then, Video A' is not restarted from A'_2 of the continuation of Video A', but from $\Delta A'_1$ because the

part of overlap is incorporated.
Here, the time of one round is set to 44 seconds (40 seconds as the switching interval + 4 seconds as the overlap time), so $\Delta A'_1$ A'_2 $\Delta A'_2$ is 44 seconds and A'_2 is 36 seconds. Therefore, $A'n$ is 36 seconds after n is 2. The output video is generated in the same way as mentioned above.

6. Evaluation

6.1. Evaluation items

We performed experiments to test the level of comprehension and to evaluate the approach of the new conferencing participation style with this proposal. Specifically, we set the evaluation criteria as the accuracy rate. The accuracy rate is the number of questions answered correctly out of the total number of questions. This shows the degree of comprehension of the conference. The video used in this experiment was five minutes long and the content of the video included a free discussion among four participants.

6.2. Comparative approach

To evaluate the usability of our proposal, we defined and used the following comparison methods. Sequential conferencing style (Our proposal) is alternately watching two conference videos for a short time using shortened video playing. Multi conferencing style is imultaneously watching two conference videos. Serial conferencing style is Twatching one conference video after watching another conference video Next, we explain the experimental procedure. The subjects of this experiment included 16 college students. The subjects are divided into one group of six, and one group of five, and one group of five. Each group watches the conference videos in the sequential conferencing style, the multi conferencing style, and in the serial conferencing style respectively.

6.3. Result of the experiments

Table 4 shows the results of the sequential conferencing style, multi conferencing style, and the serial conferencing style, based on our proposal. The accuracy rate of the multi conferencing style was 45%, which was much lower than the other styles. This low rate is because the audio of the two conferences negate each other. The accuracy rate of the serial conferencing style was 75%, which is the highest among the three styles because the subjects can watch each of the conferences accurately. However, there is a problem of decreasing the conference density when using the serial conferencing style because the time until the viewers can finish watching the two conferences is twice as long as the other styles. On the other hand, the accuracy rate of the sequential conferencing style based on our proposal was 66%, which was about the same result of the serial conferencing style. Furthermore, the conference achieved a high density because the subjects can watch two conference videos in the same time that would be required to watch one conference. From the result of this experiment, it was confirmed that the sequential conferencing style of our proposal is useful as the new style of participating conference.

Table 4. Level of Comprehension about Content of Conference

Conferencing style	Sequential	Multi	Serial
Accuracy rate	66	45	75

7. Conclusions

We have proposed a new conferencing participation style with the goal of allowing viewers to participate in two remote video conferences simultaneously. In this research project, the users alternately watched the videos of the remote conferences for a short period of time using shortened video playback and video switching. By doing so, we attempted to create a method that would support multiple conferences so that the users could watch two real-time contents.

8. Acknowledgments

This work is supported in part by a Grant-in-Aid for the Global Center of Excellence for High-Level Global Cooperation for Leading-Edge Platform on Access Spaces from the Ministry of Education, Culture, Sport, Science, and Technology in Japan.

References

[1] J. Kato, T. Sumiya, A. Inoue, H. Shigeno, K. Okada: Automatic Camera Switching Method using Time Shifting, Journal of Information Processing Society of Japan (in Japanese, in press), Vol.47,No.3,March 2006D 1

[2] Blair Maclntyre. Support for multitasking and background awareness using interactive peripheral displays. Proc. Of ACM, 2001.

[3] Andrew Dahley. Water lamp and pinwheels: ambient projection of digital information into architectural space. Proceedings of CHI98, 1998.

[4] K. Tsumura, T. Sumiya, I. Takata, H. Shigeno, K. Okada: Automatic Switching Technique of Remote Conference Video Image for Multitasking Workers with Multiple Cameras IPSJ SIGNotes Groupware and Network services (in Japanese), GN-58,pp.155–160, 2006. 1

[5] Itaru Takata, Kousuke Tsumura, Hisashi Anzai, Hironori Egi, Kenichi Okada. Agenda based Multiple Work Support for Video Conferencing Participation and Deskwork, In Proc. of The Second International Conference on Collaboration Technologies, pp. 21-26, July 2006. 1

[6] Mary Czerwinski, Eric Horvitz and Susan Wilhite, A Diary Study of Task Switching and Interruption,In Proceedings of CHI , pp. 175-182, 2004.