Adaptive Synchronization Framework for Navigated Hypermedia Document Presentation

Kuo-Yu Liu and Herng-Yow Chen
Department of Computer Science and Information Engineering, National Chi-Nan University
1st University Rd., Pu-Li, Nan-Tou County, Taiwan 545, ROC
caspar@mclab.csie.ncnu.edu.tw  hychen@csie.ncnu.edu.tw

Abstract

Synchronization anomalies, or so-called out-of-sync problems, may occur in a composite multimedia document presentation especially when a presentation tool receives the unexpected events that destroy inter-media relations in the document. This paper proposes an adaptive synchronization framework (ASF) to solve this problem by adopting resynchronization procedure. In contrast to the existing approaches focusing on temporal synchronization only, spatial and content relations for a more general presentation are also considered and incorporated into the proposed synchronization core. The experimental results show that the proposed framework can adapt a document to an unsupervised presentation environment and the performance is satisfactory (85%) under some synchronization-killer operations.

1. Introduction

With the advance in multimedia and web technologies, many hypermedia applications such as education-on-demand systems have successfully incorporated different types of multimedia objects and HTML pages into a logically composite document, and allowing a playback system to present them in a synchronized way [1, 2]. Such systems require an algorithm that can dispatch the corresponding media objects to appropriate physical or logical devices (e.g., audio device and windows of screen) under a certain constraint on either temporal or spatial domain. Information about dispatching rule, called synchronization relation, is necessary for a document presentation and may be related to intra-media synchronization or inter-media synchronization [3].

All media objects involved in a composite multimedia document should be able to be presented synchronously under the synchronization control of the presentation tool; nevertheless, the document may still suffer out-of-sync anomalies. As shown in Figure 1, out-of-sync problem occurs in spatial domain when a user narrows the displaying window in width, browser itself maintains the intra-media synchronization only (i.e., atomic property of word “synchronization”) but it fails to guarantee the inter-media spatial synchronization between the word “synchronization” and the ellipse mark over it. The result is therefore unsatisfactory and need to be corrected by another resynchronization process.

Several techniques have been proposed for solving the out-of-sync problem described above. The techniques can be classified into two classes: group-based and constraint-based. The former means that all media objects are bound together during the presentation (e.g., SVG and SMIL), and the latter is to define several rules for adapting the presentation to different display environments [4]. However, both techniques require the development of advanced player and the users’ interactions are limited. In this paper, the general browser is used as a presentation tool to render the hypermedia document, and we proposed an Adaptive Synchronization Framework for recovering the out-of-sync problem when unexpected user interactions are involved.

2. The WSML system

The Web-based Synchronized Multimedia Lecture system integrates audiovisual lectures, HTML pages, and navigation events to provide synchronized presentations [5]. A teacher’s oral guidance along with
several navigation events can be captured by a synchronization recorder. The navigation events, such as pen strokes, highlight, dynamic annotation, mouse track (tele-pointer) and scrolling, are guided media. Those media objects and navigated events will be presented dynamically in a browser by using state-of-the-art dynamic HTML techniques.

Figure 2 shows an example of the synchronized presentation in the WSML Browser. Two short terms, P.E. and U.E., indicate the pre-orchestrated events and unexpected user interaction events, respectively.

Figure 2. An example of synchronized presentation with unexpected interactions

- **T1**: The AV and HTML URL1 are loaded at T1. The HTML page is then rendered by the browser and the AV player starts playing the AV lecture at the same time. The mouse track events would exist all the time until the teacher stops recording in the recording stage. Note that the navigation events and the corresponding audio clip should never be presented unless the associated HTML page (namely, a base page) has been retrieved into the browser. If the HTML base slide cannot be loaded successfully, both navigation events rendering and audio clip playing are meaningless (content synchronization).

- **T2**: A scrolling event is triggered to move down the scroll bar so that the content which is originally out of screen could be displayed. The event here is P.E.

- **T3**: User enlarges the font size of document and the hypertext is subsequently rearranged by the browser. The event here is U.E.

- **T4**: A scrolling event is triggered to show the content which is out of screen. The event here is P.E.

- **T5**: A highlight event over a keyword is invoked at T5. The event here is P.E.

- **T6**: A pen stroke event is driven again to show a continuous drawing effect. Resynchronization is needed here because the positions of hypertext are rearranged. The event here is P.E.

- **T7**: At that time, the user narrows the browsing window and the document layout is changed again. The event here is U.E.

- **T8**: A pen stroke event is driven again to show a continuous drawing effect. The events here are P.E. They are rendered according to the correct coordinates computed by the resynchronization mechanism after the document layout is changed.

In the presentation stage of the example given above, navigation events are triggered dynamically to enrich multimedia presentation on the web. During audio/video playback, navigation events (P.E.) will be presented at appropriate time and spatial positions. In addition, resynchronization process is invoked when synchronization relations are corrupted by destructive user interactions (U.E.)

3. Adaptive synchronization framework

The synchronized hypermedia document can be presented synchronously if there is no destructive event. However, several events, such as slow/fast playback of audio clip, window resizing, or disabled/enabled partial content, will cause the presentation out-of-sync unexpectedly. In order to restore the adaptive synchronous presentation, the resynchronization framework is proposed to handle those out-of-sync problems. The framework, as shown in Figure 3, mainly comprises three mechanisms: synchronization, resynchronization and smoothing mechanism.

- **The synchronization mechanism** is to load the composite hypermedia document that consists of different types of media objects and synchronization information. The cooperation of temporal, spatial and content synchronization guarantees the synchronized presentation.
The resynchronization mechanism is to recompute the synchronization information if relations are changed by user events. The principle of the resynchronization mechanism can be summarized as the three procedures: (1) checking whether or not the relations have been changed, (2) updating the respective information by re-computing values, and (3) re-applying the updated information and the media object to the synchronization mechanism.

The purpose of smoothing mechanism is for human perception considerations. The jitter-defects of continuous media usually result in perceptive annoyances due to the performance limitation in synchronization mechanism. Thus, smoothing mechanism can be adopted to conceal the annoyed presentation defects.

### 4. Human perception consideration

Human beings usually have different perception levels of “out-of-sync” in different applications [6]. The general approach, for example, dealing with intermedia synchronization usually result in perceptive annoyances due to the performance limitation in synchronization mechanism. Thus, smoothing mechanism can be adopted to conceal the annoyed presentation defects.

![Figure 4. Mouse track of a teacher's recitation](image)

Mouse track (tele-pointer) is captured in the recording stage continuously in our system. Although the spatial out-of-sync problem of mouse track can be recovered very soon by the proposed resynchronization mechanism, jitters may still happen. Figure 4 (a) shows the presentation under normal displaying environment. Once a user resizes the window, the objects in displaying window will be rearranged by browser automatically. At that time, the resynchronization mechanism is triggered to update new spatial information of tele-pointer. However, as shown in Figure 4 (b), the presentation sequence of tele-pointer, 12, 13, 14 and 15 will annoy the users due to spatial jitter.

Two policies: “drop_then_replace” and “insert_if_far,” are adapted to smooth the jitter and then get better presentation effect. Figure 5 shows the relationships of tele-pointers of Figure 4 from time and distance viewpoints. Each tele-pointer will keep the attributes: time duration itself and the distance to the preceding one. We use the time duration and distance between two pointers to specify which policy should be applied.

![Figure 5. Mouse track mapping to time_distance relationships of Figure 4.](image)

The following rules can be used to identify the appropriate policy. \( P_i \) means the \( i \)th tele-pointer in the presentation sequence.

1. If (time duration of \( P_i \) > time duration threshold \( TT \)) \&\& (the distance of \( P_{i+1} \) > distance threshold \( DT \)), the “insert_if_far” policy is adopted.
2. If (time duration of \( P_i \) < time duration threshold \( TT \)) \&\& (the distance of \( P_{i+1} \) > distance threshold \( DT \)), the “drop_then_replace” policy is used.

In “insert_if_far” policy, we should first determine the number of tele-pointers to insert (NC).

\[
NC = \frac{\text{Distance of } P_{i+1}}{\text{Distance threshold}} \times \frac{\text{Duration time of } P_i}{\text{Duration time of } P_{i+1}}
\]

Then, we compute the \( NC_i \) and \( NC_{i+1} \), to specify how many extra tele-pointers will occupy time duration of \( P_i \) and \( P_{i+1} \), respectively.

The end time of \( P_i \) is (timestamp of \( P_i \) + (duration of \( P_i \) – \( NC_i \) * sampling rate (20ms))) and the start time of \( P_{i+1} \) is (timestamp of \( P_{i+1} \) + \( NC_{i+1} \) * sampling rate). The
duration time of $P_i$ and $P_{i+1}$ can be derived as 180 ms and 360 ms.

Figure 6. An example of “insert_if_far” policy
The “drop_then_replace” policy is to change the position of dropped tele-pointer to another position closer to the previous pointer. As shown in Figure 5 (b), if time duration and distance threshold can be determined properly, the jitters can be concealed and the movements of tele-pointer become smooth.

5. Experimental results

The purpose of smoothing mechanism within WSML is to smooth spatial noise that is caused by playback under an unexpected user interaction. We perform a subjective experiment to evaluate the result of the smoothing mechanism for tele-pointer movements in a synchronized hypermedia document. The experimental scenario on a synchronized hypermedia document is that a teacher explains a paragraph and moves the tele-pointer to point words that he is talking about. We invited several students who were selected as fairly as possible across all grades and both sexes to evaluate the tele-pointer movement. According to their perceptions, the mean opinion scores (MOS) were obtained to be subjective performance measures of the smoothing mechanism.

Figure 7. Human perception acceptance
The subjective experimental results show that people feel acceptable if threshold of time duration lies between 250ms and 350ms, and distance threshold lies between 150 and 300 pixels. Figure 7 illustrates the acceptable ratio of tele-pointer movement for human perception. The X-axis denotes the different time duration threshold settings, Y-axis presents another threshold value, distance, and Z-axis means the percentage of test candidates who feel comfortable with tele-pointer movement in different threshold values.

If the threshold of distance or duration time is lower, the tele-pointer points to a wrong word that a teacher is talking about. If the thresholds of both are higher, the jitters occur and annoy the users. In both cases, users can’t concentrate on learning materials and get poor learning effects.

6. Conclusion

This paper has presented the effects of various synchronization anomalies and resynchronization strategies according to temporal and spatial synchronization viewpoints. The robust, adaptive synchronization framework which is for presenting the composite hypermedia document is proposed to ensure the synchronized presentation. Human perceptions on spatial synchronization have also been discussed. The experimental results show that the resynchronization performance is satisfactory (85%) under some synchronization-killer operations. The resynchronize techniques presented in this research may be beneficial to the development of QoS-aware multimedia applications. An English overview page with demonstrations can be accessed at http://english.csie.ncnu.edu.tw/demo/index.html.

7. References