Abstract. This paper describes the enhancement of SMIL 2.0 through the inclusion of a comprehensive set of multimedia processing effects that can be described through XML alike syntax. The enhancements are fully integrated with the time control module in SMIL 2.0 which provides the necessary object synchronization. Such enhancement not only results in better quality Internet presentation, the text based, XML alike syntax also improved the editing environment of the presentation. Furthermore, it improves the reuseability, and partial extraction of the Internet presentation.

1. Introduction

With the rapid development of the Internet, online document in plain form with static objects become unattractive. A number of technologies have been proposed to enhance both visual and audio effects for online documents. Some examples are streaming video by RealNetwork [1] and FLASH by Macromedia [2], etc. However, both approaches do not conform to the text based Internet presentation platform as defined by XML [3]. Furthermore, those commercial products do not support open source architecture, which has imposed subsequent difficulties in editing and maintaining the Internet documents. The non-text based approach also prevents document or objects to be reused by concatenating contents from various sources.

To enable user-friendly authoring of an immersive multimedia presentations such as training courses on the Web, W3C has designed the Synchronized Multimedia Integration Language (SMIL) under the XML standard [4]. The SMIL language is an easy-to-learn HTML-like language. Thus, SMIL presentations can be written using a simple text-editor. Following the conventional HTML concept, the components of the presentation document can be located at various servers and linked together by hyperlinks. The existing SMIL 1.0 standard, however, is very primitive which only support simple manipulation of web objects. The result is far more inferior to the
competing technologies, such as FLASH from macromedia. To enhance the quality of SMIL based presentations, a working draft of a new SMIL 2.0 has been released [5].

In this paper, we will describe the development and implementation of a SMIL 2.0 based browser. In addition, an Enhanced Multimedia Processing Effect Library (EMPEL) has been developed to enhance the multimedia object manipulation capability beyond that specified in the SMIL 2.0 working draft. The new functionalities can be invoked by a new set of directives that follow the XML definitions and fully compatible with existing SMIL specifications.

The EMPEL is implemented with ActiveX control on Microsoft Window platform. The SMIL scene description language together with the EMPEL extensions are used to compose the document. The browser will retrieve the required contents from various locations and compose it at client’s display. As the servers only need to transmit the components and instructions to compose the final presentation at the client side, this approach can also reduce the transmission bandwidth required. The browser also supports various audio and image processing effects on the retrieved objects used for the presentations.

The enhanced effects provided by the EMPEL includes:

1) Audio Effects: Volume control, Fading, Echo, Reverberation, Panning, Surround Sound, Time Stretching, Time Shrinking, etc.

2) Image Filtering Effects: Fading, Brightness Control, Blurring, Embossing, Sharpening, Feathering etc.

3) Text Animations: Curving, Free Rotation, Directional-Scrolling, Shooting, Flashing, Spotlighting, Blinking, Cinema-Scrolling, etc.

4) Video Effects: Successive filtering, Stretching, Shrinking, Background color removal, transitions etc.

5) Scene Object Interactivities: Overlapping, Diffusing, Synchronization etc.

The improved presentation quality with the enhanced effect provided through the EMPEL will be demonstrated.

2. Implementation

As a result of the different nature of audio, image, video and text objects, each category of the multimedia objects is first processed on the client’s machine. Afterward, synchronization and interaction between objects are derived dynamically during execution.

2.1 Audio Processing Effects

The audio effects can be categorized into four main categories: amplitude based effects, time delay effects, waveform distortion effects and frequency response.