

Interactive E-book Contents Authoring System based on MPEG-4 LAsER

Injae Lee*, Myungseok Ki*, Seongyong Lim*, Sanghyun Park*, Jihun Cha*, Joongyun Lee**

* Broadcasting&Telecommunications Media Research Deptpartment, ETRI,

161 Gajeong-Dong, Yuseong-Gu, Daejeon, 305-350, Korea

** Service R&D Center, SK Telecom, Korea

{ninja, serdong, seylim, sh_park, jihun}@etri.re.kr, lee.jy@sk.com

Abstract— The current e-book service offers to read books or magazines that are in digital forms. It supports traditional narrative text and image only. It does not support video capabilities, graphic animations and user interaction. The development of handheld devices such as the smart phone, tablet pc is bringing interactivity to e-books. Therefore, the existing e-book contents will suffer not only from lack of multimedia, but the potential for user interaction.

In this paper, we proposed the interactive e-book contents authoring system based on MPEG-4 LAsER and ePub extension. MPEG-4 LAsER is an open standard for rich media service that provides spatio-temporal organization of various multimedia data such as video, audio, text, image, graphics as well as possible interactions and animations of them. The generated e-book contents contain not only multimedia objects (audio, video, text, image and graphic animations) but also user interaction.

Keywords— Interactive, e-book, Authoring, MPEG-4 LAsER, ePub

I. INTRODUCTION

The digital e-book service has grown rapidly in the last few years. Thus, many e-book contents can be downloaded for the iPad, Kindle, Story or any other e-book reader. Most e-book contents are text- and image-based publication in digital form. And current e-book readers support traditional narrative texts and images only.

As the mobile market matures, the numbers of users accessing e-books through mobile devices has continued to grow. And a boom in the tablet PC market will surely result in an increase in e-book consumption. To expend the market for paid content, the new capabilities to enhance the user experience will be required. Rich media services offer enriched contents including audio, video, text, and graphics. And it supports interactivity, animation, and dynamic behaviour over time according to the user's requests. Therefore, it can provide a fresh and active user experience. Several technologies are competing to provide technical format for rich media contents such as MPEG-4 BiFS, LAsER, SVG, FLASH and Silverlight. Compared to the traditional text-based e-book service, the interactive e-book service based on rich media formats can provide attractive features which include dynamic updates of contents and interactivities among

various multimedia combination such as audio, video, graphics and text.

W3C made attempts to define description languages for rich media contents: SVG (Scalable Vector Graphics) [1] and SMIL (Synchronized Multimedia Integration Language) [2]. However, neither SMIL nor SVG supports a streaming method. Therefore, these are inappropriate for fast, dynamic, interactive and interoperable applications. MPEG developed composition and coding technologies for rich media scene descriptions. The first specification developed is MPEG-4 BIFS (Binary Format for Scene) [3]. BIFS was developed to provide virtually all the functionalities rich media can provide to date on a fairly resource abundant environment: PC and the Internet. However, the inherent characteristics of BIFS from VRML (Virtual Reality Meta Language) and its binary encoding structure make it inappropriate applying for services using resource constrained devices. Instead of compromising the performances, MPEG reached to the conclusion that an optimum between feature richness/compression efficiency and device constraints needed to be found, thus decided to create a new standard for rich media in a resource constrained environment. Based on these consideration and decision, MPEG developed a new composition coding technology called LAsER (Lightweight Application Scene Representation) [4]. The main goal of LAsER is to provide a representation of scene data that is well suited for constrained environment, and it is defined to be the rich-media format targeted at constrained devices, such as mobile, embedded and consumer electronics devices. It adopts subsets of SMIL and SVG recommendations and BiFS standard to support use scenarios suitable for rich media streaming and broadcasting.

The interactive e-book contents authoring system is based on the newly developed open standard, MPEG-4 LAsER. It can provide a differentiated e-book service using dynamic and active characteristics. Almost all the e-book readers support ePub(electronic Publication) format. Thus, the proposed authoring system supports the ePub format.

This paper is organized as following. Section II presents the overview of MPEG-4 Part 20 (LAsER) and ePub specification. In section III, the architecture of the interactive e-book

contents authoring system is presented. Section IV shows the implementation results with examples and section V concludes with future work.

II. THE NEW APPROACH : MPEG-4 LAsER AND EPUB

The proposed interactive e-book authoring system supports MPEG-4 LAsER and ePub. The overviews of MPEG-4 LAsER and ePub specification are as follows.

A. LAsER

LAsER is a specification designed for representing rich media services to resource constrained devices. In other words, the main goal of LAsER is to provide a representation of scene data that is well suited for a resource constrained environment [5]. The rich media presentation is a collection of a scene description and media. A scene description specifies the spatio-temporal layout of objects such as a text, graphics, animation, an individual audiovisual content of the form image, video, audio, font data and interactive relationships of them. Binary format is also defined to provide better transport efficiency.

LAsER is based on SVG Tiny which is the W3C Recommendation for mobile vector graphics services. Upon this basis, MPEG added LAsER Extensions; Dynamic Updates, Binary Encoding, Audio/Video support, and Font Information to employ various functionalities required for rich media services in a resource constraint environment. The specification defines a LAsER engine, which is shown in Figure 1, as the viewer for LAsER presentations. A LAsER engine has rich media composition capabilities and these capabilities rely on the usage of an SVG scene tree.

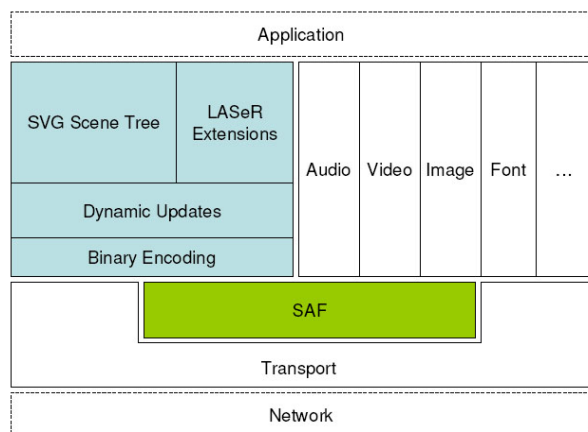


Figure 1. Architecture of a LAsER engine

B. ePub

The ePub format is an open standard for e-books created by the IDPF(International Digital Publishing Forum)[6]. It consists of Open Publication Structure (OPS), Open Packaging Format (OPF) and OEBPS Container Format (OCF). The OPS describes the content markup language such as XHTML, Dasiy DTBook. And it offers the way to include generic XML using an inline XML island. MPEG-4

LAsERML is based on XML and it can be contained in OPS using an inline XML island. The purpose of the OPF is to define the mechanism by which the various components of an OPS publication are tied together and provides additional structure and semantics to the electronic publication. It is accomplished by two XML files with the extensions .opf and .ncx. The opf file houses the ePub book's metadata, file manifest, and linear reading order. And the ncx file contains the hierarchical table of contents for ePub file. An ePub file is a group of files conforming to the OPS/OPF standards that is wrapped in a ZIP file. The OCF specifies how these files should be organized in the ZIP, and defines two additional files: mime type file and container.xml file.

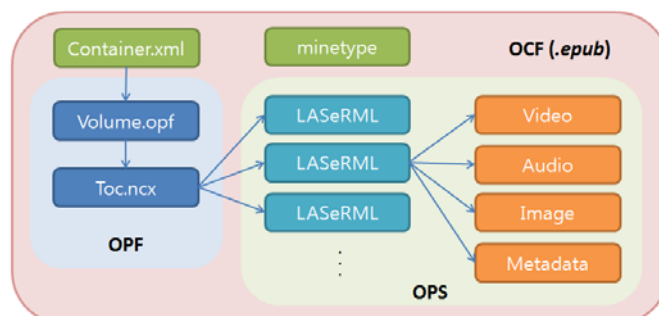


Figure 2. Structure of interactive e-book based on MPEG-4 LAsER

III.ARCHITECTURE OF INTERACTIVE E-BOOK CONTENTS AUTHORIZING SYSTEM

The proposed authoring system enables an author to easily and conveniently produce interactive e-book contents by using MPEG-4 LAsER. It provides powerful graphic user interface and produces dynamic e-book contents in the form of ePub. The architecture of the presented authoring system is shown in Figure 3. It consists of the following parts; User Interface, Authoring Information Editor & Processor, Data Preprocessor.

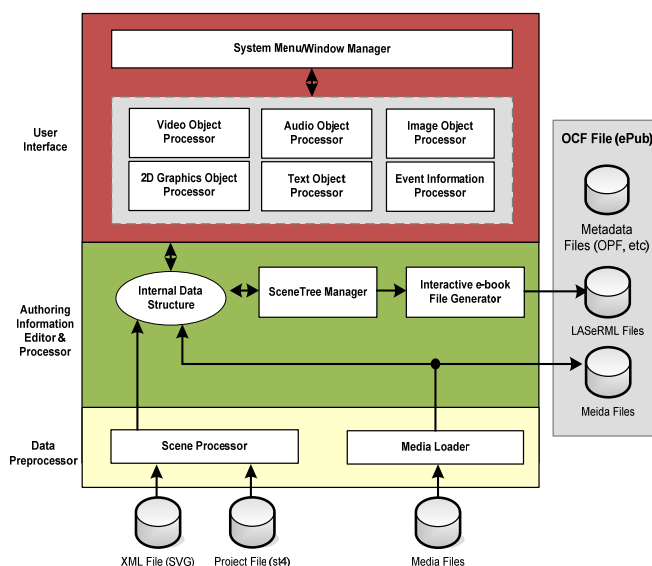


Figure 3. Architecture of the presented authoring system

A. User Interface

The presented e-book authoring system is furnished with a powerful graphic user interface to help an author conveniently create and edit interactive e-book contents. And it provides the spatio-temporal authoring environment. Figure 4 shows the graphic user interface in the presented authoring system. It consists of menu and tool bars, canvas, pageview, event, action, property and timeline windows.

The menu and tool bars provides fundamental functions for the authoring system such as open, save, publish, import, etc. In the canvas window, e-book contents can be created and edited by inserting, deleting and modifying audio-visual objects. It supports multimedia data such as MP4, MP3, JPEG, PNG and graphic objects such as circle, rectangle, text, line. The pageview window offers the ability to show the spatial authoring results about e-book pages and access to each page. Many event properties such as show, hide, animation can be realized in the event&action window. The attributes of individual objects are shown and can be edited in the property window. The temporal attributes of individual audio-visual objects are displayed as a track in the time line windows.

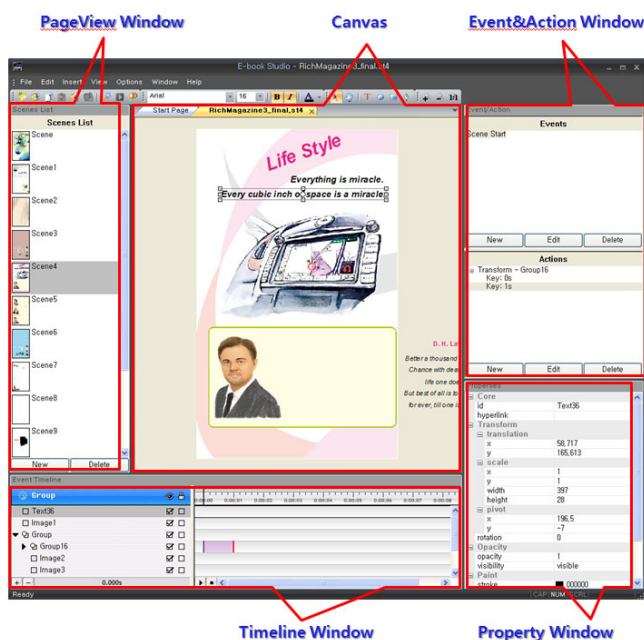


Figure 4. GUI of the presented interactive e-book authoring system

B. Authoring Information Editor & Processor

Authoring Information Editor & Processor provides the interfaces of easy access to all the functionalities of the presented authoring system using Data access APIs. Data access APIs contains the interfaces that construct and edit an internal data structure such as the uDOM(micro Document Object Model). The uDOM is based on XMT DOM, and thus has a tree structure with a variety of elements for e-book description and reference to media object. It also allows hierarchical access to the scene information.

The proposed interactive e-book authoring system generates not only the OPS file which is composed of LAsERML files, but also the OPF file which describes the structure of an ePub such as opf, ncx. And the final interactive e-book content is the .epub file which bundles files (OPS, OPF, multimedia files) together according to OCF.

C. Data Preprocessor

Data Preprocessor supports the collection of the codecs for decoding audio-visual objects such as MP4 for visual sequences, MP3 for audio sequences, JPEG, PNG for still images. Since the Media Library separately manages media codecs from other parts of the presented authoring system, it is much easier to extend the scope of the Media Library by adding new media codecs. And it can manage the external SVG files. The processed SVG files can be used as an independent object in the e-book contents

IV. IMPLEMENTATION

In this paper, the interactive e-book contents authoring system which includes MPEG-4 LAsER and ePub specification is presented. The interactive e-book contents are shown in figure 5, 6. They are interactive e-book contents including various media, graphics and events. The e-book player for authorized e-book contents consumption has been implemented as an application for the android smartphone such as Galaxy S.

Figure 5 shows an interactive film magazine including audio, video, images, graphics and text. This content can be offered additional information such as the synopsis, cast, director, trailer, photo gallery, advertisement. Thus, a user can select the desired information. A user can control the trailer video such as play, pause in the interactive magazine as shown in figure 5(a). When a user selects the director or characters menu, the detailed information is displayed on the screen (illustrated in figure 5(b)). Figure 5(c) shows the interactive photo gallery. If a user selects a specific photo, the selected picture is animated and appeared in the center of the page. To arouse a user's interest, the animated graphics and the advertising video can be served as shown in figure 5(d).

The interactive e-book contents based on MPEG-4 LAsER can be offered fresh and active graphic animation like FLASH. Figure 6 shows animated graphical features. The text objects and images as well as graphics are animated. And it has the scalability of vector graphics. Therefore, it can be rendered to fit the size of the rendering plane.

As shown in figure 5, 6, interactive e-book contents are easily produced by temporal and spatial composition of objects, and by animation and events managements using GUI.



(a)



(b)



(a)



(b)



(c)



(d)

Figure 5. An example of the interactive film magazine



(c)



(d)

Figure 6. An example of the interactive magazine

V. CONCLUSION

The interactive e-book service is one of new paradigms for e-book services. LAsER is a newly introduced open standard for rich media service that provides interactive functionalities in a lightweight contents format. Providing interactive e-book contents make user behaviors from passive reading to active playing on the contents. A user's interest will be aroused by dynamic multimedia combination and animation. By interacting directly on the contents, users can acquire additional information easier.

The presented e-book authoring system provided a noble approach to produce interactive e-book contents by using MPEG-4 LAsER and ePub. Users can enjoy more versatile value added services compared to traditional text-based services with interactive and dynamic features.

By the rapid growth of the mobile and tablet PC market, the numbers of users accessing the e-book through handheld devices has continued to grow. Recently introduced handheld devices have interesting features such as sensors. The easier way to create interactive e-book contents using various sensors such as accelerometers, tilt, proximity, and activity sensors needs to be further supported.

ACKNOWLEDGMENT

This work was supported by SK telecom under the title of "Next-Generation Rich Media Technology".

REFERENCES

- [1] W3C, SVG Tiny1.1 Specification. <http://www.w3.org/TR/SVG11/>.
- [2] W3C, SMIL 2.0 Specification. <http://www.w3c.org/TR/smil20/>.
- [3] ISO/IEC 14496-1, Information Technology - Coding of Audio-Visual Objects - Part 1: Systems-First Edition, Dec. 1999.
- [4] ISO/IEC 14496-20, Information Technology - Coding of audio-visual objects, Part 20: lightweight Application Scene Representation (LAsER) and Simple Aggregation Format (SAF), June 2006.
- [5] Jean-Claude Duford, and Olivier Avaro, "LAsER: the MPEG standard for rich media services", *IEEE multimedia*, vol. 12, pp. 60 - 68, Dec. 2005.
- [6] <http://en.wikipedia.org/wiki/EPUB>