Generation of the Orchestral Media for the Realistic Multimedia Representation

Jae-Kwan Yun', Mi-Kyung Han', Jong-Hyun hug', and Kwang-Roh Parke,

' Meta-Verse&Emotion Technology Research Team, 218 Gajeong-no,
Yuseong-gu, Daejeon, Republic of Korea
[2] Green Computing Research Department, 218 Gajeong-no,
Yuseong-gu, Daejeon, Republic of Korea
{jkyun, mkhan, jangjh, krpark} @etri.re.kr

Abstract. Future digital home is changing into intelligence ubiquitous home, and the future media will be changed into the orchestral media which includes not only visual, audio, text but also effect information related with the specific scene. The orchestral media can be interlocked with user peripheral multiple devises. To play the orchestral media, there needs effect information definition, method for inserting effect information into media. Therefore, in this paper, we explain the SEM (Sensory Effect Metadata), creation of orchestral media and orchestral media authoring tool.

Keywords: Orchestral media, realistic multimedia, SEM.

1 Introduction

Until now, conventional media with visual and sound components has been presented via display devices and speakers. But, nowadays, users want more realistic experiences of multimedia contents with high fidelity. For examples, stereoscopic video, virtual reality, 3DTV, multi-channel audios are typical types of media for realistic experiences. However, these sorts of applications are limited in visual and audio perspectives. For example, special effects can be authored as a separate track in conjunction with an audiovisual content in a synchronized way. While the audiovisual content is being playbacked, a series of special effects can be made by shaking curtains for a sensation of fear effect, by turning on a flashbulb for lightning flash effect, etc. Furthermore, fragrance, flame, fog and scare effects can be made by a scent device, a flame-thrower, a fog generator and a shaking chair, respectively. The orchestral media is the media that includes not only visual, audio, and text but also device control information, synchronization information, other effects descriptions with various useful devices around user [1]. Therefore, the orchestral media can play one media with user around multiple peripheral devices at the same time to give user realistic feelings [2], [3].

Fig. 1 illustrates the concept of the orchestral media service in the service aspect from the media generation to presentation. Media generation process presented in this figure shows the service architecture in which a user films a video at a garden, obtains

environmental information (sensory effect) such as temperature, humidity and strength of the wind to create the orchestral media and plays it at a user terminal. That is, with the orchestral media, it is possible to reproduce the situation when filming was done even after several years. With the orchestral media that has already been made, it defines a device to be connected according to the contents of media and connects various devices which allow more realistic playing of the effect provided by the existing media. To do this, we defined the SEM (Sensory Effect Metadata), the orchestral media which contains audiovisual, and the orchestral media service to present these orchestral media.

This paper is organized as follows. The SEM schema is described in chapter 2. Creation of the orchestral media is proposed in chapter 3. Finally, the conclusion and future work are described in chapter 4.

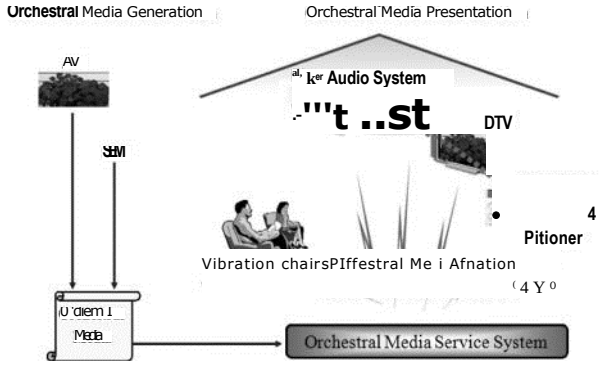


Fig. 1. Concept of the Orchestral Media; From the orchestral media generation to presentation

2 SEM Schema

In this chapter, we will explain about the schema design, syntax & semantics of the SEM.

2.1 Schema Design

The SEM has two main parts those are Effect Property and Effect Variables. Effect Property contains definition of each Sensory Effect applied to the contents. By analyzing Effect Property, the orchestral media service system can match each Sensory Effect to the proper User Device in the user's environment, prepare and initialize User Device before processing media scene. Effect Variables contains the control variables for Sensory Effect synchronized with media stream. Fig. 2 shows the process for the SEM.

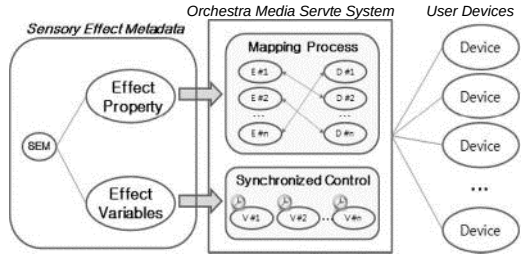


Fig. 2. The Entire Processes for the SEM.

2.2 Syntax & Semantics

The SEM element is the root of the SEM schema. This element has three sub-elements those contains the General Information, the Effect Property, and the Effect Variables. The syntax and semantics of the *SEM* are shown in Fig. 3 and 4.

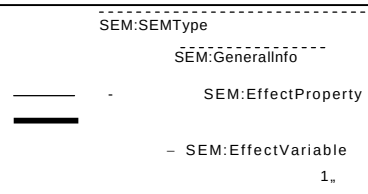
Diagram	
Source	<element name="SEM" type="SEM:SEMType"/> <complexType name="SEMType"> <sequence> <element name="GeneralInfo" type="mpeg7:DescriptionMetadataType" minOccurs="0"/> <element name="SEDescription" type="SEM:SEDescriptionType" maxOccurs="unbounded"/> </sequence> </complexType>

Fig. 3. The XML Schema Syntax of <SEM>.

GeneralInfo	An element containing the information on the metadata creation
EffectProperty	An element contains a list of Sensory Effect and the property of each Sensory Effect applied to the contents
EffectVariable	An element contains a set of Sensory Effect control variables and time information for synchronization with media scene

Fig. 4. The XML Schema Semantics of <SEM>.

EffectProperty contains the information about overall Sensory Effect applied to the contents. EffectlD and Type will be defined for each Sensory Effect (Effect element in the schema) to identify itself and to notify category of the Sensory Effect. Under the Effect element, there is a set of property elements for describing Sensory Effect capabilities through which the orchestral media service system will match each Sensory Effect to User Device. The semantics of attributes are shown in Fig. 5

I. Name	Definition
EffectiD	An attribute containing ID of individual Sensory Effect.
Type	An attribute containing the enumeration set of Sensory Effect type. "VisualEffect": Sensory Effect for visual display such as monitor, TV, wall screen, etc. "SoundEffect": Sensory Effect for sound such as speaker, music instrument, bell, etc. "WindEffect":Sensory Effect for wind such as fan, wind injector, etc. "CoolingEffect": Sensory Effect for temperature such as air conditioner. "HeatingEffect": Sensory Effect for temperature such as heater, fire, etc "DimmmgEffect'. Sensory Effect for light bulb, dimmer, color LED, flash, etc. "FlashEffect": Sensory Effect for flash "ShadingEffect": Sensory Effect for curtain open/close, roll screen up/down, door open/close, etc. "VibrationEffect": Sensory Effect for vibration such as trembling chair, joystick, tickler etc. "DiffusionEffect": Sensory Effect for scent, smog, spray, water fountain, etc. "OtherEffect": Sensory Effect which is not defined or combination of above effect type
Priority	An optional attribute defining priority among the number of Sensory Effects
AltEffect1D	An optional attribute containing ID of alternate Sensory Effect which can be replace current Sensory Effect

Fig. 5. The semantics of attributes of <EffectProperty>.

EffectVariable is the container for various Sensory Effect variables which controls Sensory Effect in detail. Fig. 6 shows attributes of EffectVariable.

In=rtme	Defini
SEfragmentI	An attribute defining ID of the fragment of Sensory Effect
Start	An attribute defining the start time that Sensory Effect will be activated.
duration	An attribute defining the duration time that Sensory Effect will be deactivated.

Fig. 6. The semantics of attributes of <EffectVariable>.

3 Generation of the Orchestral Media

The format of the orchestral media is based on ISO base media file format. In order to satisfy streaming service, we used Timed Metadata Format (TeM) which is a standard method to send metadata in streaming way.

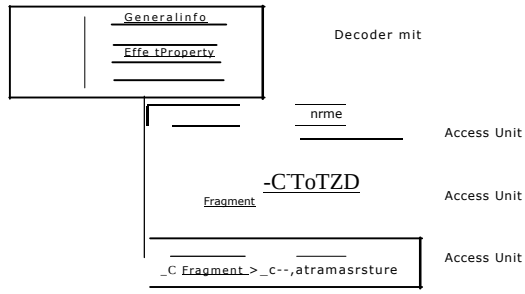


Fig. 7. Fragmentation Strategy of SEM.

The first step is to fragment the SEM and encapsulate each fragmentation with TeM message. Fig. 7 shows the fragmentation strategy of SEM. The whole sub-elements of General Information and Effect Property, and the attributes of Effect Variable will be encapsulated with Decoderinit message which should be transported first hand. Fig. 8 shows the example of Decoderinit message.

```
<DecoderInit..>
    <SchemaReference...1>
    <InitialDescription>
        <FragmentUpdateUnit>
            <FUCommand>addNode</FUCommand>
            <FUContext>/</FUContext>
            <FUPayload>
                <SEM>
                    <GeneralInfo>... </GeneralInfo>
                    <EffectProperty>
                        <Effect EffectlD="1" ...> ... </Effect>
                        <Effect EffectlD="2" ...> ... </Effect>
                    </EffectProperty >
                        <EffectVariable RefEffectlD="1" .../>
                        <EffectVariable RefEffectID="2" .../>
                 </SEM>
            </FUPayload>
        </FragmentUpdateUnit>
    </InitialDescription>
</Decoderinit>
```

Fig. 8. The example of Decoderinit Message.

Then each fragment containing start time and duration information and actual commands to control user devices will be encapsulated by AccessUnit message those should be transported synchronized with audio/visual packets. Fig. 9 shows the example of AccessUnit messages.

```
<AccessUnit>
        <FragmentUpdateUnit>
            <FUCommand>addNode</FUCommand>
            <FUContext>/EffectVariable</FUContext>
            <FUPayload><Fragment start="20" duration="10" ...>...
            </ Fragment ></FUPayload>
        </FragmentUpdateUnit>
        <FragmentUpdateUnit>
            <FUCommand>addNode</FUCommand>
            <FUContext>EffectVariable</FUContext>
            <FUPayload>< Fragment start="20" duratiori="10" ...>...
            </ Fragment ></FUPayload>
        </FragmentUpdateUnit>
</AccessUnit>
```

Fig. 9. The example of AccesUnit messages.

The second step is to embed TeM access units into ISO media file format as a metadata track. Embedding the SEM in ISO base media file format follows the standard specification of MPEG-4 Part 12 Amendment 1. Since the SEM is a kind of timed metadata, we inserted the SEM in 'meta' box defined in ISO file format. The handler_type of Handler Reference Box is 'meta'. Fig. 10 shows the sample description box for the SEM and fig. 11 shows authoring tools which can insert, delete, edit, and publish the orchestral media.

```
class MetaDataSampleEntry(codingname) extends SampleEntry (codingname)
  class XMLMetaDataSampleEntry() extends MetaDataSampleEntry ()meta') {
  string content_encoding; // optional
    string namespace;
    string schema_location; // optional
    BitRateBox 0; // optional }
```

Fig. 10. The sample description box for the SEM.

Fig. 11. Authoring tools for creating the orchestral media.

5 **Conclusion.**

The Orchestral media service can maximize the reproduction effects via any device, and at any time and any place and restructure the media created in this way to make a media customized for and by users as interlocking the device depending on the environment of a user's peripheral devices. The orchestral media service can be used in the field of real sensing digital cinema, device cooperated education, user created contents (UCC). In this paper, we described SEM, concept of the orchestral media service, creation of the orchestral media and authoring tool. In the future, we have to study about binary encoding of the sensory information to reduce amount of the SEM.

Acknowledgments. This work was supported by the Industrial Strategic Technology Development Program funded by the Ministry of Knowledge Economy(MKE, Korea). [10039171, Development of Convergence Media Service Platform for Realistic Experience in Smart Home Environment]

References

1. Yun, J.K., Shin, H.S., Lee, H.R., K.R. Park.: Development of the Multiple Devices Control Technology for Ubiquitous Home Media Service System. In : International Conference on Ubiquitous Information Technology & Applications, pp.404--408. (2007)
2. Choi, B.S., Joo, S.H., Lee, H.R., Park, K.R.: Metadata structure for Media and Device Interlocking, and the Method for Mapping It in the Media Format. In : Advances in Information Sciences and Services, pp186--190. (2007)
3. Elting, C.: Orchestrating Output Devices - Planning Multimedia Presentations for Home Entertainment with Ambient Intelligence. In : Joint sOc-EUSAI conference. (2005)