A Framework of Sensory Information for 4-D Home Theater System

Yongmun Park[1], Bumsuk Choi[1], Hankyu Lee[1], Kyung Sung[2]*
[1]Electronics and Telecommunications Research Institute, Daejeon, Korea
[2]Department of Computer Education, Mokwon University, Korea
Email Address: ympark@etri.re.kr

Abstract: These days the major film makers are producing movies considering 4-D effects. The number of 4-D system equipped theaters is increasing. 4-D is an irresistible new trend in movie industry. Also home devices such as fan, lights, heater, even motion chair are evolved and smart enough to be controlled remotely. Users still prefer to enjoy movies in home by TV and home theater system. 4-D home theater system is a solution for a user who wants to watch movies immersive in private zone. We present a framework for 4-D home theater system based on MPEG-V standard. A complete framework for 4-D home theater system can be supported from authoring of sensory effects to environment description and commanding rendering devices for the sensory effects by MPEG-V and couple of other standards. Now the standard and the framework technology are ready for 4-D contents for users in home.

1 Introduction

Right after 3D contents, 4-D contents knock on the movie industry. More than hundreds of films per a year are produced with 4-D effects now a day. 4-D movies were very special type of entertainment only available at the theme parks such as the Disney World or the Universal Studio, as a form of a short time riding experience. However, the full-time 4D movies are recently released and they brought a great attention. More and more movies are expected to be produced using 4-D effects and accordingly a standard to represent 4-D effects is necessary. The 4-D effects represent the sensory effects such as wind, water spraying and motion.

As the realistic game devices are advanced, those devices can also be used in 4-D contents play especially in home environment because they are small, comparably cheap, and smart enough to process synchronized signals. In this paper, 4-D home theater system based on the MPEG standard, specifically MPEG-V, is proposed with component technology such as description of 4-D contents, decoding and setting in the home network environment, and presentation of 4-D contents.

MPEG-V is initiated in 2007 with the goal of establishing interoperable metadata containing sensory effects information. In the early stage of MPEG-V, it is called RoSE which is the abbreviation of Representation of Sensory Effects. As the scope of RoSE activities became broader to the area of

---

* **Corresponding Author: Kyung Sung, skyys04@mokwon.ac.kr**

virtual world, RoSE was merged with MPEG-V [1] in 2008.

Fig. 1 shows an exemplary service scenario of media service with sensory effects which is described in MPEG-V Part 1 [1]. A video content is authorized with a sensory effects description as defined in Part 3 [3] of MPEG-V standard. This AV content with the sensory effects description is then delivered to user terminal (media processing engine in the Fig. 1) connected with sensory devices. The media processing engine analyses the sensory description written in Part 3. Then the MPEG-V adaptation engine converts the sensory effects description to commands as defined in Part 5 [5] to drive the rendering devices based on the device capability descriptions and the user preferences description specified based on Part 2 of MPEG-V [2].
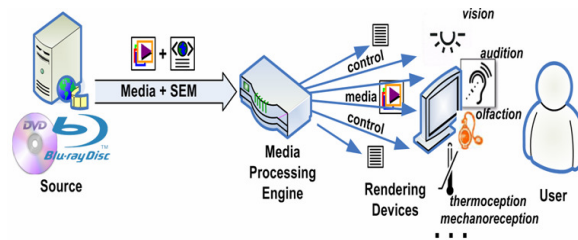

Fig.1 Example of media service with sensory effects

Besides MPEG-V standard supports many other service scenarios such as virtual tour, realistic game, virtual shopping, virtual learning, immersive broadcasting, and so on. In order to support these, MPEG-V standard defines various sensors to collect environmental information such as temperature, humidity, intensity of illumination, motion, position, etc. The virtual world is a core application domain in Part 4 which currently describes only avatar however this Part will be extended as the activities go on.

2 Description of 4-D Content

MPEG-V Part 3 provides means for describing sensory effects such as wind, vibration, temperature, light, etc., that trigger human senses in order to increase the experience of the user. The Sensory Effect Description Language (SEDL) is an XML Schema-based language which defines the skeleton of sensory effect metadata and common attributes which may be used to describe all kinds of sensory effects. Each type of specific sensory effect and its attributes are defined in Sensory Effect Vocabulary (SEV) Schema. A description conforming to SEDL and SEV may be associated with any kind of

multimedia content (e.g., audio, video, image, and flash).

A description conforming to SEDL is referred to as Sensory Effect Metadata (SEM). SEM may contain an optional `autoExtraction` attributes and `DescriptionMetadata` followed by choices of `Declarations`, `GroupOfEffects`, `Effect`, and `ReferenceEffect` elements. The `autoExtraction` attribute is used to signal whether automatic extraction of sensory effect from the media resource is preferable. The `DescriptionMetadata` provides metadata about the SEM itself (e.g., authoring information of the SEM) and aliases for classification schemes (CS) used throughout the whole description. The Declarations element is used to define a set of sensory effects and preparation information for later use in a SEM via an internal reference. A `GroupOfEffects` is used to define combinations of sensory effects which contain at least two single effects which share start time and duration. An Effect is used to describe a single effect with an associated timestamp which can be used for synchronization with the associated media resource. For the timing and synchronization information, the XML Streaming Instructions defined in MPEG-21 Digital Item Adaptation [8] have been adopted. The actual instance of Effect comprises all information pertaining to a single sensory effect such as fade, priority, position, activate, duration, etc. A `ReferenceEffect` may be used to reference the sensory effect declared internally in the Declarations elements.
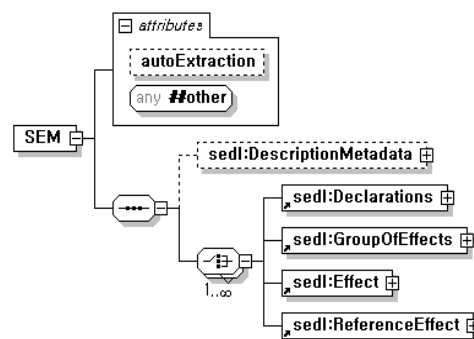

Fig.2 Structure of Sensory Effect Metadata for sensory effect description

The Sensory Effect Vocabulary (SEV) defines an individual type of actual sensory effects. The effects are defined in an abstract way to capture the author's intention, because the sensory effect description should be independent from the end

user's device setting. Currently, the standard defines the following effects:

- Light, colored light, flash light are defined for describing light effects with the intensity in terms of illumination expressed in the unit of Lux. For the color information, a classification scheme (CS) is defined by the standard comprising a comprehensive list of common colors. Furthermore, it is possible to specify the color as a vector of RGB values. The flash light effect extends the basic light effect by the frequency of the flickering in times per second;

- Temperature effect enables describing a temperature effect of heating/cooling with respect to the Celsius scale;

- Wind provides a wind effect where it is possible to define its strength with respect to the Beaufort scale;

- Vibration allows one to describe a vibration effect with its strength according to the Richter magnitude scale;

- For the water sprayer, scent, and fog effect the intensity is provided in terms of ml/h;

- RigidBodyMotion describes the motion chair effect with motion patterns based on 6DoF which is well known factors to express movement in 3 dimensional spaces. There are 7 patterns defined in current standard specification such as MoveToward, Incline, Shake, Wave, Turn, Spin, and Collide;

- PassiveKinestheticMotion, Passibe KinestheticForce, ActiveKinesthetic, and Tatile effects define all haptic effects which are used to describe the feel and the force of a material;

Finally, the color correction provides means to define parameters that may be used to adjust the color information in a media resource to the capabilities of end user devices. Furthermore, it is also possible to define a region of interest where the color correction shall be applied in case this desirable (e.g., black/white movies with one additional color such as red).

3 Sensory Effect Authoring Tool

Ordinary users may have a difficulty in making sensory effect description because of the lack of XML knowledge. We developed a GUI based authoring tool called "RoSE Studio" which provides convenience to users in generating sensory effect description.

Fig. 3.1 shows the main user interface of the RoSE Studio. The RoSE Studio has three parts of functional windows. Upper-left window shows playback of movie clip in order for the user to find out proper scenes by controlling progress bar or control buttons. Upper-right grid window is to modify or insert values for General Information which includes copyright information and issued date. The Effect Property contains definition of each sensory effect applied to the contents which includes position or direction of the sensory effects. The Effect Variables contains the control variables for Sensory Effect synchronized with media stream for example wind speed or color of light. Fig. 3.2 shows the example of sensory effect variable input interface for motion effect.
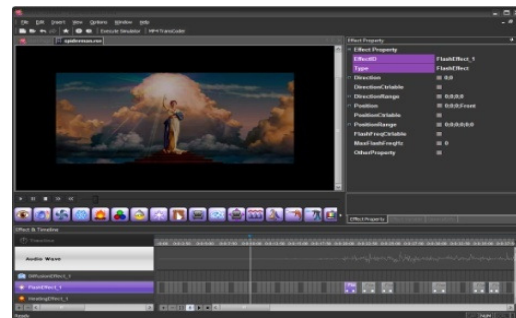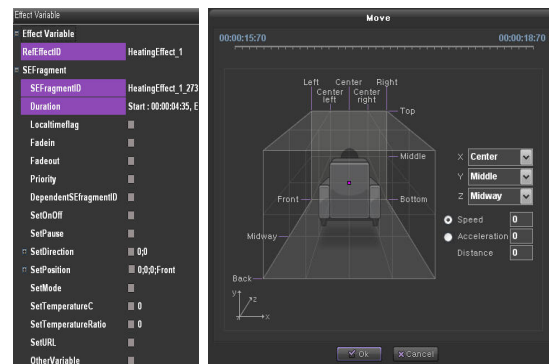

Fig.3.1 RoSE Studio

Fig.3.2 Sensory effect variable for motion effect

Lower-wide window shows timeline according to the length of movie clip. Under the timeline, the sensory effect types and effect fragment will be added as shown in the Fig. 3.3. The effect fragment is a smallest unit to identify individual control of effect.
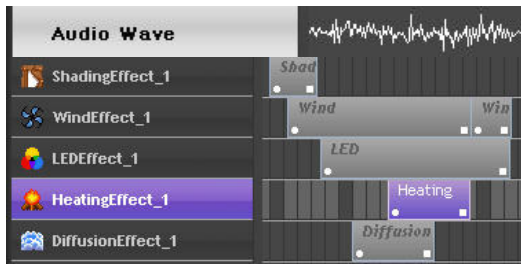
Fig.3.3 Sensory effect layer in timeline

Fig. 3.4 shows the effect icons which user can add a new type of sensory effect by clicking effect icon button. Added effect type will be shown in the left part of the timeline window.

Fig.3.4 Sensory effect icon buttons

After adding a new type of effect, user can create an effect fragment by clicking right mouse button on the selected effect type. Added effect fragment will be located at the same row of the effect type. User can change the start time and end time of the effect fragment by moving it or dragging it with holding the corner of the rectangle of effect fragment. For each effect fragment, user can edit the Effect Variables in the grid window.

Fig.3.5 Publishing sensory effect description

After finishing editing sensory effect description, user can generate a sensory effect description file by executing "Publish" in the main menu. Fig. 3.5 shows the window for publishing a sensory effect description.

4 Decoding and Setting in the Home Network Environment

This section describes the concept of home-network environment for receiving and reproducing 4-D contents.

Focusing on a home server, home network environment consists not only of the multi-media devices such as video and audio equipment, but also of actuators which can reproduce the real-sense effects (4-D information) in the 4-D contents.

In addition, various sensors should be elements of the home network for an interface with the home devices using the user's movements and the adaptive services optimized to the home environments. In order to represent the various 4-D information effectively, description of device capability and user preferences for the specific effects are needed. In MPEG-V Part 2, these effects related information is standardized.

A. Home Network Environment with Sensors and Actuators

The devices in home network environment can be grouped into 3 specific categories, such as Multimedia devices, Actuator and Sensor. Home server provides various functionalities such as reception and reproduction of the 4-D contents, interfacing between devices and sensors, and synchronization between contents and devices.

The multimedia device includes video display, audio equipment, text viewer, etc. Actuator is a device that reproduces various real-sense effects, which include fan for the wind effect, heater for the temperature effect, LED for the lighting effects. To collect the user environment information, there exist several sensors in home network such as motion sensors, ambient light sensors, temperature sensors and position sensors. Aggregated information can be used in various applications such as user interface and adaptive home services. Fig. 4.1 described structure of home network system.
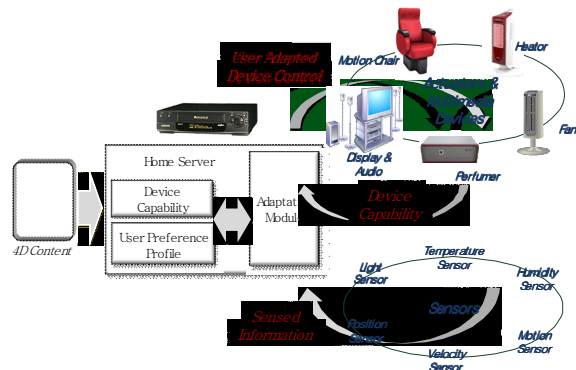

Fig.4.1 Example layout of the receiver and accompanying devices for 4-D reception in home-network environment

B. Service Discovery in Home Network Environment with Sensors and Actuators

In our proposed method, we need to find devices capable to handle MPEG-V functions in a home network. For user convenience, finding and recognizing the devices in the given home network should be done automatically. There exist several technologies for service discovery such as SSDP (Simple Service Discovery Protocol) [9] or Bonjour [10]. Both technologies provide ways to find devices and get their capabilities on a network automatically, but the latter is a proprietary technology from Apple, while the former is from IETF draft. Also SSDP is the basis of Universal Plug-and-Play [11], which is one of the mostly used home network protocols. To find a device and its capability on a home network regarding device compatibility, we define a modified SSDP approach from UPnP.

In UPnP, a device is a logical entity which can provide UPnP functionalities on UPnP network, and a control point is to control those devices on UPnP network. SSDP provides a simple way to find these control points and devices through HTTP protocol. When one device is plugged into an UPnP network, it must advertise its presence and it is done with a SSDP message with NOTIFY header as shown in Fig. 4.2 (a). This SSDP message from a device has URI for the device information or the service description. Control points which receive this NOTIFY SSDP message can request further information of the device using XML form through standard HTTP protocol. For a control point, it can search UPnP devices with an SSDP message with M-SEARCH header as shown in Fig. 4.2 (b). For MPEG-V devices, 2 new terms has been defined for MPEG-V capable entities, a MPEG-V Device and MPEG-V Control Point. And some variables in NTS header have been modified.

Typically, NOTIFY message for UPnP device is as follows,

```
NOTIFY * HTTP/ 1.1
HOST: 293.255.255.2500:1900
CACHE−CONTROL:
LOCATION:
NT:
NTS: ssdp:alive
SERVER:
USN:
BOOTID.UPNP.ORG:
SERCHPORT.UPNP.ORG:
```
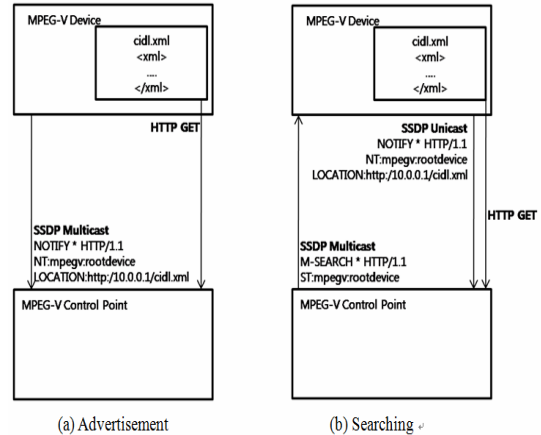

Fig.4.2 Advertisement and searching for MPEG-V devices

For UPnP, NT field means Notification Type, and it shows what kind of device or service is described in that SSDP message. Since we're going to describe the capability of device in UPnP device or service profile, and we don't want to make the discovery procedure complex, only one value for NT is defined – `mpegv:rootdevice`. The Location field stores the URI location of device information. For UPnP, the designated URI on Location field contains a device or service profile. Instead of UPnP device/service profile, CIDL description is located at the URI. MPEG-V application can get the CIDL description from that URI using standard HTTP protocol.

For M-SEARCH message, there is a ST field which describes search targets. When control point issues an M-SEARCH message, search range is described in ST field. In this way, UPnP control point can restrict its searching range from service, device or all the devices in home network (ssdp:all). Instead of these values, `mpegv:rootdevice` is used for MPEG-V applications. MPEG-V Control Point issues M-SEARCH message with `ST:mpegv:rootdevice`, and MPEG-V devices will respond with its own NOTIFY message.

C. Environment recognition and setting with Part 2

In order to reproduce various information in the 4-D contents effectively, home server must store performance information of actuator, sensor and user preference information for each effect. Home server collects such information in the form of standardization and provides the appropriate service. In less than no time, MPEG-V Part2 will be described in detail.

MPEG-V Part2 specifies syntax and semantics of the tools required to provide interoperability in controlling devices in real as well as virtual world.
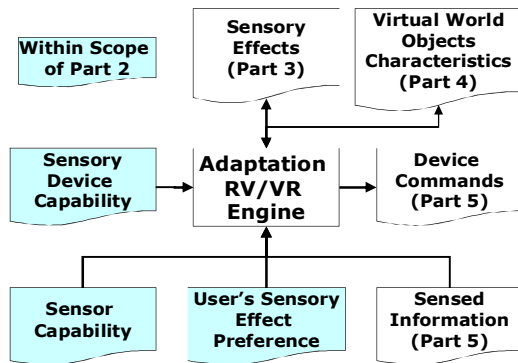

Fig.4.3 Scope of MPEG-V Part2 Control Information

In Fig. 4.3, the scope of this Part of the standard with tools defined in this Part is shown. The adaptation engine (RV or VR engine), which is not within the scope of standardization, takes five inputs (sensory effects (SE), user's sensory effect preferences (USEP), sensory devices capabilities (SDC), sensor capability (SC), and sensed information (SI)) and outputs sensory devices commands (SDC) and/or sensed information (SI) to control the devices in real world or virtual world object. The scope of this Part of the standard covers the interfaces between the adaptation engine and the capability descriptions of actuators/sensors in the real world and the user's sensory preference information, which characterizes devices and users, so that appropriate information to control devices (actuators and sensors) can be generated. In other words, user's sensory preferences, sensory device capabilities, and sensor capabilities are within the scope of this Part of the standard.

The next is the example of fan device capability information schema. It is the description of a wind device capability with the following semantics. The wind device identifier is "fan01". The maximum wind speed of the wind device (possibly a fan) is 30 meter per second. This specified device can support 5 levels in controlling the wind speed. This device takes 10 milliseconds to start and 10 milliseconds to reach the target intensity.

```
<cidl:SensoryDeviceCapability
  xsi:type="dcdv:WindCapabilityType"
  id="fan01" zerothOrderDelayTime="10"
  firstOrderDelayTime="10"
  maxWindSpeed="30" numOfLevels="5" />
```

5 Presentation of 4-D Contents with Part 5

As shown in the Fig. 3.4, the adaptation module produces device commands following the specification of MPEG-V Part 5 based on the available devices and capabilities of the available devices as well as the user's preferences on sensory effects. For example, when a wind effect is presented and there are no fans available but an air-conditioner is available, the adaptation module should issue a command to the air-conditioner to be turned on with the current temperature.

A device command specifies the device to be used to render a certain sensory effect with relative intensity of the effect in percentage of the maximum capable intensity of the effect that the selected device can provide with the timing information to activate the effect. Fig. 5.1 shows the structure (schema) of the wind device command.
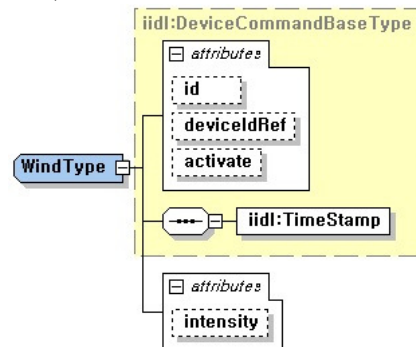

Fig.5.1 Schema diagram of WindType device command

This wind type device command has attributes to represent the id for identifying each individual command, deviceIdRef to uniquely identify a specific device, activate flag to signal turning on and off, and the TimeStamp to carry the timing information regarding the presentation time of the effect. Also, the structure represents an intensity attribute, which takes a value between 0 percentiles to 100 percentile of its maximum capable intensity.

Fig. 5.2 illustrates a sample command of wind device, where a wind device with device id of fd002 is to be activated at 361000 clock ticks at 100 clock ticks per second, meaning 3610 seconds after the initiation, with intensity of 50% of maximum output.

```
<iidl:DeviceCommand
  xsi:type="dcv:WindType" id="f001"
  deviceIdRef="fd001"
  activate="true" intensity="50">
<iidl:TimeStamp
  xsi:type="mpegvct:ClockTickTimeType"
  timescale="100"  pts="361000"/>
</iidl:DeviceCommand>
```

Fig.5.2 Sample command for wind device

6 Conclusions

Even though the MPEG-V has much larger scope, which includes defining standardized interface between virtual worlds, 4-D home theater systems is also one of the key applications of MPEG-V. In this paper, we have presented a framework for 4-D home theater system based on MPEG-V standard.

The authors of the 4-D content construct the sensory effect metadata using MPEG-V Part 3, in synchronization to the content. The authored sensory effect metadata and AV content are delivered to the user terminal. Then the user terminal analyzes the sensory effect metadata through Part 3. The devices around the user environment are detected using SSDP and the device capabilities are written using MPEG-V Part 2. The adaptation engine, then, generates a sequence of adequate device commands using MPEG-V Part 5, so that the devices can generate the sensory effects properly, creating 4-D effects.

Even though this framework is still at a very early stage of development, this framework provides an almost complete guideline to start building 4-D home theater system.

Acknowledgements

References

[1] ISO/IEC FDIS 23005-1 Information technology -- Media context and control -- Part 1: Architecture, August 2010, ISO Publication

[2] ISO/IEC FDIS 23005-2 Information technology -- Media context and control -- Part 2: Control Information, August 2010, ISO Publication

[3] ISO/IEC FDIS 23005-3 Information technology -- Media context and control -- Part 3: Sensory Information, August 2010, ISO Publication

[4] ISO/IEC FDIS 23005-4 Information technology -- Media context and control -- Part 4: Virtual World Object Characteristics, August 2010, ISO Publication

[5] ISO/IEC FDIS 23005-5 Information technology -- Media context and control -- Part 5: Data Format for Interaction Devices, August 2010, ISO Publication

[6] ISO/IEC FDIS 23005-6 Information technology -- Media context and control -- Part 6: Common Types and Tools, August 2010, ISO Publication

[7] ISO/IEC CD 23005-7 Information technology -- Media context and control -- Part 7: Reference Software and Conformance, August 2010, ISO Publication

[8] ISO/IEC 21000-7:2007 Information technology -- Multimedia framework (MPEG-21) -- Part 7: Digital Item Adaptation, November 2007, ISO Publication

[9] Goland, Yaron et al., "Simple Service Discovery Protocol/1.0", IETF Internet-Draft draft-cai-ssdp-v1-0.txt, October 1999.

[10] http://apple.com/bonjour

[11] http://www.upnp.org

[12] Shin Jin Kang, Soo Kyun Kim, Sung Hyun Cho, Procedural Game Level Generation with Interactive Evolutionary Computation, *Journal of Future Game Technology*, 1, 1(2011) 9-20.

[13] Kang-Hee Lee, Ubiquitous Robotic Game incorporating Hardware and Software, *Journal of Future Game Technology*, 1, 1(2011) 35-42.

Yongmun Park received his Ph.D degree in Computer Science from Hannam University in 2002. Currently he is a principal member in Electronics and Telecommunications Research Institute. His research interests include Content-based Advertising, Smart TV and Real-sense Media Convergence. Contact him at ympark@etri.re.kr.

Bumsuk Choi received his Master degree in Computer Science from Chungnam National University in 2001. Currently he is a senior member in Electronics and Telecommunications Research Institute. His research interests include 3D sound, Digital Rights Management system, Augmented Reality, and Real-sense Media Convergence. Contact him at bschoi@etri.re.kr.

Hankyu Lee received his Master degree in Computer Science from Kyungpook National University in 2006. Currently he is a principal member in Electronics and Telecommunications Research Institute. His research interests include Content-based Advertising, Smart TV, Customized Broadcasting and Standards. Contact him at hkl@etri.re.kr.

Kyung Sung is received a MS. degree in computer engineering Kyunghee University 1993 and Ph.D. degree in computer engineering from Hannam University 2003. In 2004, he joined the faculty of Mokwon University where he is currently a professor in Department of Computer Education. His research interests Information Security, Neural Network, Web Programming.