A Study on Generating Sensory Effects of MPEG-V Motion for Realistic Content

Sangwook Park, Jae-Kwan Yun, Mikyong Han, and Jong Hyun Jang

Abstract— there is the standard for representing the motion effects in MPEG-V. This paper presents how to use object tracking and motion vector information from realistic content like 4D films and Virtual Reality as motion effects, and how to express the patterns as meaningful motion sensory effects. In this study, the experiment describes the process of analyzing the location and orientation information of each frame extracted from the actual content and recognizing it as a pattern.

Research Keywords— Motion Effect; Sensory Effect Metadata; MPEG-V; Virtual Reality; Augmented Reality

1 INTRODUCTION

Augmented audiovisual data as well as real world effects is a recent trend in film, Virtual Reality (VR), and educational services platforms. The 4D Films give sensory effects to audience watching movies for immerse experience [1]. Designing 4D Effects is a highly sophisticated operation and in-house authoring tools with limited functionality is negative for professional training. Even professional designers go through the process of authoring for a long time each time a new movie comes out, with expertise gained from their experience [2].

Therefore, an algorithm for automatically generating a motion effect and a standard of a method for expressing metadata are required. We briefly introduce the effects representing movement experience in the MPEG-V [3] standard with semantic pattern and the method to utilize it. Following section introduces how to make motion pattern effects based on the standard providing the motion metadata of similar experience to the manually authored effects. The aim of this study is to show the process of generating the motion pattern effects by analyzing the sampled experiment data from 4D media.

2 MOTION EFFECTS

2.1 4D Films and VR

4D film provides an immersive experience with auguendted sensory effects (also called 4D effects) [2], such as motion, light, wind, water, and scent, with a cinema. Realistic VR is similar to the theater, but the tactile sense, by motion, vibration, temperature, and haptics, is more effective for enhanced experience because of the head mounted display with visual maximization.

2.2 Generating 4D Effects

Haptic or motion effects can be extracted automatically from relavant media and cameras. Location by object tracking and vectors by estimating camera motion in the movie matches chairs' movement. Generating sensory effect metadata corresponding to tilting, rotation, zoom-in, zoom-out, camera-movement, and camera-rotation of objects included in the media controls a 4D chair [4]. Extacting information using the color of video scenespoints generate the temperature effect by mapping to cooling and heating fans [1].

2.3 Athoring Tool for Motion Effects

In order to provide the motion effect to the 4D chair fixed to the floor called a rigid body motion, there is an encoder produces the positions in three dimensional space and tilt angles on x, y, and z axis. The six degrees of freedom (6DoF) authoring tool make distance from the center of the chair as surge(forward/backward), heave(up/down), and sway (left/right), and orientation through rotation about normal, lateral, and longitudinal axes, termed yaw, pitch, and roll respectively. As shown in the Fig. 1, motion effects of sampled surge, heave, sway, yaw, pitch, and roll values are generated every frame of the media depeding on the user's intention.

- *S. Park is with Electronics and Telecommunications Research Institute (ETRI), South Korea. E-mail: ssean@etri.re.kr.*
- *J. Yun, M. Han, and J. Jang are with ETRI. E-mail: jkyun, mkhan, and jangjh@etri.re.kr respectively.*

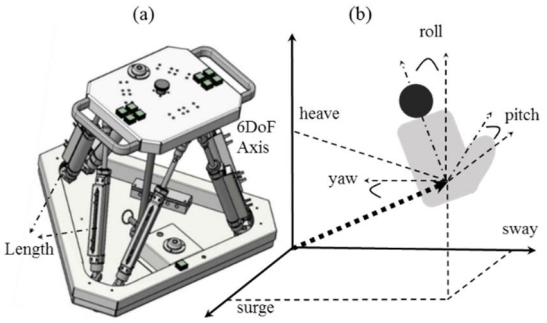

Fig. 1. Motion Authoring Example: (a) is encoder manipulated in experimenter perceives and (b) is 6DoF according to the change in length of the encoder's six axes

3 MPEG-V Sensory Effect Metadata for Motion

In MPEG-V, metadata standardization for representing motion effects is underway, and properties such as TrajectorySamples and MoveToward in the effect definded RigidBodyMotion is discussing to express the motion device properties of the six axes.

The following sections are explaining the effects' samples defined in the standard and relationship between the effects.

3.1 Motion Set - Trajectory Samples

TrajectorySamples is an effect comprised with arranging 6DoF values containing x, y, z positions and pitch, yaw, roll orintations by six matrix in order of time. It has a value of the number of samples per second as the properties of "dim" and "duration". In this case, the motion chair can process the sampling data sequentially in accordance with the number of frames of a given time at the time when the effect starts.

3.2 Motion Patterns

There are patterns MoveToward, Incline, Shake, Wave, Spin, Turn, and Collide that have a certain semantics [4]. MoveToward pattern covers location movement without rotation. Incline pattern covers rotating experience without changing the location. Shake pattern is a movement from one side to opposite side repeatedly. Wave pattern provides the surface expeirance of water from side-up to side-down. Spin pattern means continuous turning without changing the location. Turn pattern gives moving towards a certain direction. Collide pattern descripes reaction after collision against something. Each pattern has attributes to materizlize such as direction, distance, speed, accelation, interval, count and power.

Table 1. Motion Patterns and Corresponding Properties

Effects	Properties - Pattern Semantics of the Samples
MoveToward	Horizontal and Vertical Angles; Distance; Speed - Surge, Heave, ans Sway movement flow reversed
Incline	Pitch, Yaw, and Roll Angles; Speed - Opposite direction of the angular acceleration
Shake	Direction; Distance; Count; Interval - Repetition of MoveToward and +/- of Direction
Wave	Direction; Distance; Count; Start Direction, Interval - Repetition of rolling or pitching of Incline
Spin	Direction; Count; Interval - Repetition of yawing of Incline
Turn	Angle; Distance; Speed - Repetition of Move and Incline pattern
Collide	Horizontal and Vertical Angles; Speed - Repetition of Move and Incline pattern by angle

This table shows an example of the motion patterns based on MPEG-V and properties contains the XML instance respectively.

3.3 Database and Machine Learning

An algorithm changes the sampled motion effect set to a pattern with readability. If the system find changes of the direction on location like as surge, heave, and sway in the sampled motion, it is as a candidate to change into a movetoward pattern. Similarly, if rapid change of yaw, pitch, and roll matches to an incline pattern. As shown in Table 1, if movetoward and incline repeats for a specific period or appear in a specific rule, the algorithm recognizes patterns as Sake, Wave, Spin, Turn, and Collide. Both sample and pattern effects are generated through the encoder and the authoring. The two combined effects verify that they provide the same experience after user testing. A machine learning algorithm extracting MPEG-V motion pattern use the database as a training set.

4 Simulations

This section shows motion effect experiments using the authoring tool for specific videos released on YouTub, analyzing the effects created, and converting the trajectory samples into incline, move toward, shake, and wave patterns. The content collection criteria are those with the moving reference of the background and object. The first list is group of content such as drone shooting, automobile, and dragon-ride fore virtual experience like as VR experience content. The second list is a collection of content that includes moving objects such as dogs, cats, bears, penguins, cars, boats, and trains.

The content using in this experiment verification is the YouTube identification bOJ7lZJXBGM, and it shows the bus slipping to the left and right for about 1,680 mili-seconds from the 3,450 mili-seconds

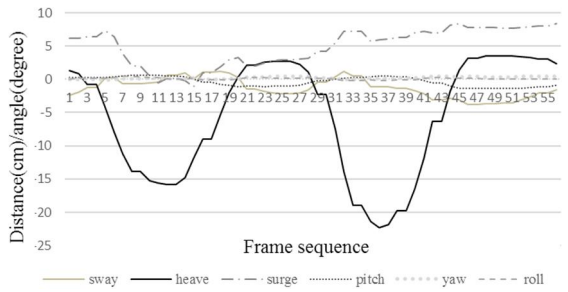

Fig. 2. 6DoF samples relating content streaming by the encoder

point of the video. Fig. 2 shows a graphical representation of the bus object's left/right; front/back; up/down movement, and tilting of the object with 30 fps frame processing through the authoring tool.

Fig. 3 (a) shows an instance accordancing with the MPEG-V standard from the motion coordinates. It is RigidBodyMotion type effect and has the duration as the effect length with pts as the start time. In TrajectorySample, its attribute dim expresses that there are 56 frames in six dimensions of 6DoF.

As seen in the graph, the heave distances at the 12, 24, and 36th frames show peak, which produces MoveToward pattern in Fig. 3 (b). By using the transformation, formula of the surface coordinate system and the spherical coordinate system changes location to direction. Likewise, the inflection point of the pitch produces Incline pattern as shown as Fig. 3 (c). Shake replaces MoveToward repetitions in the opposite direction more simply. If a significant change of the surge in the same period and the rotation in the opposite cycle occur, Wave pattern appears. Each representation is in Fig. 3 (d) and (e) respectively. Fig. 3 shows deleting repeated syntax and omiting namespace in all effects.

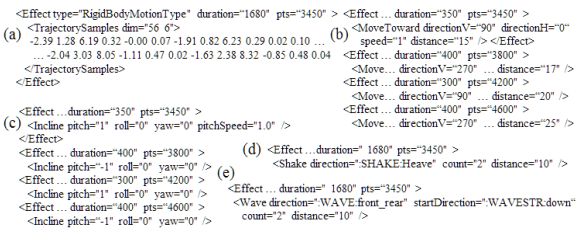

Fig. 3. Motion Effect Examples in MPEG-V: (a) is Trajectory Samples from the Authroing Tool, (b) is MoveToward set, (c) is Incline set, (d) is Shake pattern without rotation, and (d) is Wave pattern with both movement of location and orentation rotation

5 Conclusions and Future Works

This paper proposes a study to provide motion effects to the content. There is the standard to represente the motion effects and the MPEG-V metadata covers tracking information extracted from content automatically and patterns of 4D motion. After our research make a database of the metadata instances corresponding to the content, and we will measure the performance of algorithms that convert between two standards with artificial intelligences.

Acknowledgment

This work was supported by 'The Cross-Ministry Giga KOREA Project', grant funded by the Korea goverment (MSIT) (No.GK17P0100, Development of Tele-Experience Service SW Platform based on Giga Media)

References

[1] S. Kim, S. Yang, C. Ahn, and Y. Joo, "Sensorial Information Extraction and Mapping to Generate Temperature Sensory Effects," ETRI J., vol. 36, no. 2, pp. 224-231, Apr. 2014.

[2] J. Lee, B .Han, and S.Choi, "Interactive motion effects design for a moving object in 4D films," Proc. 22nd ACM Conf. Virtual Reality Software and Technology (VRST '16), pp. 219-228, Nov. 2016. Doi:10.1145/2993369.2993389.

[3] ISO/IEC 23005-3:2016, Information Technology – Media Context and Control – Part 3: Sensory Information, 2016.

[4] H. Oh and J. Huh, "Auto Generation System of MPEG-V Motion Sensory Effects based on Media Scene," Proc. 2017 IEEE Int. Conf. Consumer Electronics (ICCE '17), pp. 160-163, Jan. 2017. doi: 10.1109/ICCE.2017.7889269.

S. Park received his B.S. and M.S. degree from Kyungpook National University in 2001, 2004 respectively and joined the Electronics and Telecommunications Research Institute (ETRI), South Korea in 2004. He is currently working in realistic content service technology as a senior reseacher of the Real and Emotional Sense Section has been engaged in projects involving content server application and multimedia streaming services.

J. Yun received his B.S., M.S. and Ph.D. degrees in Computer Science at Konkuk University in Seoul, Korea, in 1997, 1999, and 2003, respectively. He is currently a senior researcher of ETRI and interests include geographic information systems (GIS), Telematics, mobile-GIS, digital home computing, and realsensemedia representation technology.

M. Han received her M.S. degree in computing engineering from the School of Electronics and Information, Kyung Hee University, South Korea, in 1993. She joined ETRI in 1993 and was a visiting professor at the School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, USA from 2012 to 2013. Her major research interests include multi-media service, emotional service, and immersive media service platforms.

J. Jang received his Ph.D. degree in computer science and engineering from Hankuk University of Foreign Studies, Soth Korea, in

2004. Since 1994, he has been with ETRI, where he is currently a leader of the Giga Service Research Group. His research interests are real-time middleware for telecommunication systems, giga media service and home networking systems, and real-sense media services.