Intermediate Data Structure

for MPEG-4 Scene Description

KyungAe Cha, HeeSun Kim, SangWook Kim

Department of Computer Science, Kyungpook National University

Abstract

MPEG-4 content is streaming media that are composed of different types of media objects, organized in a hierarchical fashion. This paper proposes scene composition model for authoring MPEG-4 contents which can support object based interactions. And we have developed MPEG-4 contents authoring tool applied the proposed scene composition model as intermediate data structure. Particularly, for supporting interoperability between multimedia contents, the scene composition model should be used independent of file format. So visual scene composed of media objects on the form of scene composition tree can be transformed variable data format including BIFS, scene description form proposed by MPEG-4 standard and also support the extension of capability.

1. Introduction

It is needed to develop of authoring tool assist the author to compose object based multimedia scenes to author interactive multimedia contents involves both temporal and spatial relations of media items such as MPEG-4 contents. MPEG-4 content is streaming media that are composed of different types of media as individual entities, organized in a hierarchical fashion[1-4]. So the object based user interactions can be supported. And it also contains scene description that specifies spatio-temporal relations among media objects. Scene description must be built in terms of both its structure and the functionality of object composition nodes. So the interactive event can be described as unit of object and scene structure can be changeable any input or output on any node.

There are several systems for authoring MPEG-4 contents. Flavor[5], script paradigm authoring tool provides the interface using script for composition of MPEG-4 contents so it requires detail knowledge of MPEG-4 standard description format on the form of text.

Mpro[6], MPEG-4 contents authoring tool supports synchronization using time-line and object based authoring interface but it also needs special knowledge about BIFS for composing contents. In this paper, to provide the visual configuration of describing interactive multimedia contents we propose scene composition model. And we have developed MPEG-4 contents authoring tool that the visual scene is composed on the form of scene composition tree as intermediate data structure. In section 2, we define scene composition model for MPEG-4 contents for purpose of representing scene description. And we explain MPEG-4 contents authoring tool based on the proposed authoring framework in section 3. In section 4, we show the implementation of MPEG-4 contents authoring tool. Finally we conclude and remark on future plan in section 5.

2. Scene composition model for composing MPEG-4 Scene description

. In this paper we propose scene composition model which can support object based interactions. Scene composition model contains audiovisual object, property objects and their relationships separately so we can treat each object independently. The contents in the form of defined scene composition tree, that is the instance of scene composition model, can be transformed into streaming media followed the specification of MPEG-4 systems because it contains audio-visual object, attribute object and event objects which are used in consisting MPEG-4 scene.

Figure 1 shows an example of scene composition tree which consists of an image object and several 2d geometry objects. These visual objects are connected with property objects which describe their attributes such as their position and their appearance. And the property objects are created or deleted as independent unit of scene tree so the scene composition tree can be extended as followed the process of authoring. If the routing information is specified, the event nodes are created and they are also connected visual object node. So object based interaction information can be described and the scene tree express these scene structure.
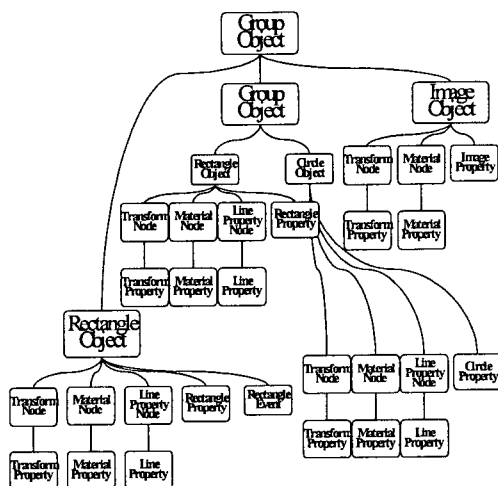

Fig. 1. An example of Scene Composition Tree

To make scene composition model impose the 2d complete nodes in MPEG-4, we classify MPEG-4 audiovisual node into 2dgeometry object, media object such as audio, image and video files. In figure 1, circle object and rectangle object are grouped so it can be treated as one group object. And we also defined the baseobject consists of the common characters of each object, so the variable to be defined repeatedly are avoided. The attribute object class that will become the instance of object attributes are classified as appearance properties and texture of material properties. These effort is to achieve the easy way to make the object description, one of MPEG-4 contents.

3. MPEG-4 Contents Authoring Tool

In this section, MPEG-4 contents authoring tool using proposed scene composition model model as intermediate data structure is explained.

To specify visual scene structure and the relations among media objects, the authoring tool provides abstract authoring concept such as interactive effects and visual configuration of object's attributes including temporal attributes[7-8]. And these information is used for generating scene composition tree which will be transformed into MPEG-4 scene descriptions.

Figure 2 shows the structure of a multimedia authoring tool which consists of a user interface, an event processor that processes the events occurred when a content is authored, scene composition tree generator and the file converter that can generate several types of streams of files by converting the scene composition tree into streaming multimedia.

Sufficient system support for an effective dialog between the user and the system is critical for authoring interactive multimedia contents. In terms of satisfying these requirement, authoring tool provides a high-level authoring technique that can specify the interactive events and specific effects conceptually, as well as a low-level authoring technique that needs more detail specification. For example, when a user authors a play and a stop time of a video object,

it provides not only the low-level authoring based on a frame unit of a video object but also the high-level authoring based on a second unit. The detail description method can be used for specific information of various type of description format such as HTML, ASF and so on.
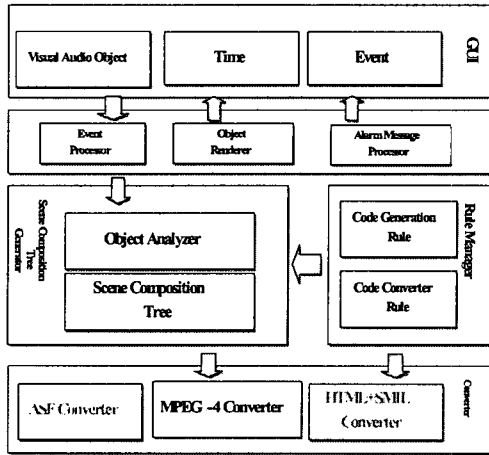

Fig.2. A structure of Multimedia Contents Athoring Tool

The created visual scene is formed as scene composition tree by scene composition tree generator. The scene composition tree extended as followed the process of authoring so it expresses dynamic multimedia scene structure. The attributes of primitive media objects in the scene composition tree are extracted from drawing attributes and they are used for generating BIFS and Object descriptor. Scene composition tree has enough information needed to generate BIFS of MPEG-4 because we have designed scene composition model based on MPEG-4 audio-visual scene. First, BIFS is generated on the text form MPEG-4 scene description and then it is encoded for generating binary stream format. So the visual scene is transformed into the MPEG-4 streaming media that is based on MPEG-4 standard by converter.

4. Development

The Proposed MPEG-4 authoring tool is developed using

Microsoft Visual C++ under the Windows 98/2000/NT operating system. Figure 3 represents an example of authoring a multimedia content. User authors contents for internet TV using various types of media objects such as image objects, text objects, video clips, and audio clips. And a user authors the contents having event and command information interactively using route information dialog to provide a real-time scene change. When a generated content is presented, a user can see the audiovisual scene changed dynamically.
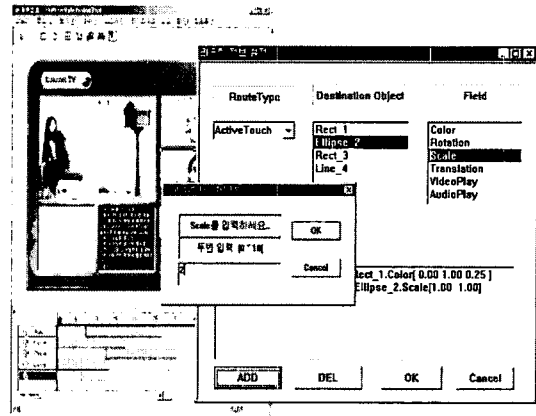

Fig.3. An example of Authoring Multimedia Scene

Figure 4 shows a portion of the text form BIFS corresponding MPEG-4 scene in figure 3. MPEG-4 authoring tool we have developed can generate text form BIFS automatically.

5. Conclusions

To provide the features of interactive multimedia contents, authoring tool must be able to compose the multimedia contents using many types of media, support the user interaction based on each object, and generate the stream supported real-time streaming. To support these features, this paper proposed the scene composition model for representing MPEG-4 scene description. Developed MPEG-4 contents authoring tool provides the high-level authoring

environments such as templates and effects to support that a user authors easily multimedia contents having the user interaction based on object. The direction of future research is the study to generate many types of multimedia streams form proposed scene composition tree.

```
Group {
    children [
    DEF Transform2D1000 Transform2D {
        translation -9.00 183.00
        scale 1.00 1.00
        rotationAngle 0.00
        children [
            Shape {
                appearance Appearance {
                    texture ImageTexture {
                        url 1
                        repeatS TRUE
                        repeatT TRUE
                    }
                }
                geometry Bitmap {
                }
            }

            ....

    DEF Condi1 Conditional {
    buffer {
        REPLACE Switch1006.whichChoice BY 0
    }
    DEF Condi2 Conditional {
    buffer {
        REPLACE Switch1005.whichChoice BY 1
                        ....
            . ]
    }

ROUTE TouchS1007.isActive TO Condi1.activate
ROUTE TouchS1007.isActive TO Condi2.activate
    ....

UPDATE OD [
    {
    objectDescriptorID 1 muxScript golf.scr
                ....
]
```

Fig.4   BIFS of an MPEG-4 audiovisual scene in Fig.3

References

[1] ISO/IEC FCD 14496-1, Part 1:Systems, ISO/IEC JTC1/SC29/WG11 N2201, Approved at the 43rd Meeting, 1998.

[2] ISO/IEC CD 14496-6, Part 6: Delivery Multimedia Integration Framework, ISO/IEC JCT 1/SC 29 N1906, 1997.

[3] http://www.cselt.it/mpeg/

[4] A. Puri and A. Eleftheriadis "MPEG-4 : An object-based multimedia coding standard supporting mobile application," Mobile Networks and Application, p5-32, 1998.

[5] A. Eleftheriadis, "Flavor: A language for media representation," Proc. of '97 ACM '97, pp. 1-9, 1997.

[6] Souhila Boughoufalah, Jean-Claude Dufourd, Frederic Bouihaguest, "MPEG-Pro, an Authoring System for MPEG-4 with Temporal Constraints and Template Guided Editing," Proc. of '2000 IEEE Multimedia Multimedia and Expo, pp. 175-178, 2000.

[7] E. Bertion and E.Ferrari, "Temporal synchronization models for multimedia data," IEEE Trans. Knowledge and Data Eng., vol. 10, No. 4, July/Aug. 1998.

[8] Thomas D.C. Little and Arif Ghafoor, "Interval-Based Multimedia Data," IEEE Trans. On Knowledge and Data Engineering, vol.5, No.4, Aug, 1993.