Transformation of MPEG-4 Contents for a PDA Device

Sangwook Kim[1], Kyungdeok Kim[2], and Sookyoung Lee[1]

[1] Department of Computer Science, Kyungpook National University
Daegu, 702-701, Republic of Korea
swkim@cs.knu.ac.kr
[2] Division of Computer and Multimedia Engineering, Uiduk University
Gyeongju, 780-713, Republic of Korea
kdkim@uu.ac.kr

Abstract. We propose a transforming method of MPEG-4 contents in order to present the contents on PDA devices in this paper. The method reconstructs the mp4 file according to transforming a scene tree in MPEG-4 contents when the contents are authored. The method reduces the size of each object in a scene of the content for presenting it efficiently on the small interface of a PDA device, and transforms visual objects in the scene into geometry objects in order to reduce initial loading time and a size of the contents. And an original object is presented on a PDA device when the user clicks the substituted geometry object. The method was applied to a conventional authoring tool, so we could find that the method showed an efficient presentation of the contents on a PDA device.

Keywords: MPEG-4, Transformation, Authoring, Adaptation.

1 Introduction

Conventional MPEG-4 contents are generally used for a video-phone on a communication network. But, in these days, their use is very increased by growth of wireless communication. The MPEG-4 is an international standard for efficient transmission and use of multimedia data and focuses on the content-based encoding that is based on understanding of image contents. Such the content-based encoding splits image contents into object units, transmits the units, and controls and displays split respective units by a user's intention [1, 2, 3].

In MPEG-4 contents, a scene is formed by the split units that are handled individually, and a scene description language, BIFS (BInary Format for Scene), is used to describe temporal and spatial information for scene changes according to user interaction and temporal flow [1]. Users use generally intuitive MPEG-4 authoring tools because authoring contents requires professional information. Most of MPEG-4 contents authoring tools are suitable to desktop computers, and generated MPEG-4 contents from the tools are most suitable to desktop players. It is especially difficult to be presented on a PDA device for MPEG-4 contents due to a small size of a PDA screen. So conventional MPEG-4 authoring

K. Aizawa, Y. Nakamura, and S. Satoh (Eds.): PCM 2004, LNCS 3332, pp. 762–769, 2004.

tools need a transformation method in order to play MPEG-4 contents on a PDA device [6, 9, 10].

In this paper, we describe the method that scales MPEG-4 scene tree in authoring for presentation of MPEG-4 contents on a PDA device. The method considers a displaying screen size on a PDA device and a file size of transmitted contents to a PDA device. A scene of MEPG-4 contents can include lots of media objects (visual and aural objects). The more the MPEG-4 contents include media objects, the more a file size of MPEG-4 contents is increased. And such MPEG-4 contents are difficult to transmit and play on mobile environment. So the suggested method supports that visual objects in a scene are represented by geometry objects and an initial transmission file has only basic information about media streams. The geometry object and reducing initial file size efficiently support to play the contents on a PDA device.

This paper is organized as follows. Section 2 introduces related researches. Section 3 describes a transformation method for a PDA device. Section 4 explains implementation and comparison between scaled contents and not scaled contents. Finally, section 5 describes some concluding remarks and presents future plans.

2 Related Works

There are lots of researches on MPEG-4 contents, but researches on authoring and a transformation of contents for a PDA device are not activated yet. There are researches about authoring tools and models of contents as follows.

First, the work of design and development of MPEG-4 contents authoring system [1] provides an intuitive authoring for users on windows environments. The system supports to compose a visual and aural scene using video objects, audio objects, image objects, and text objects, and set up intuitively a scene change by a mouse click. And the system defines a scene tree for the representation of MPEG-4 contents that is consisted of object units in a scene. The scene tree is used inner data structure for authoring in the system. A scene that is represented by using a scene tree generates a scene descriptor in the form of text, and then finally MPEG-4 contents are created after the scene goes through encoding phase. The work of design and implementation of a visual MPEG-4 scene-authoring tool [4] is developed on windows environment, and the tool can store user's making scene in the data form of the server storing MPEG-4 contents. Also the authoring tool supports BIFS commands and JAVA scripts in order to user's interaction. The work of MPEG-4 authoring tool for the composition of 3D audiovisual scenes [5] generates MPEG-4 contents using 3D APIs of the OpenGL library. The tool generates 3D contents and media objects such as a box, a sphere, a cone, video, audio, etc., and then transforms the objects into the BIFS commands. Finally, the tool generates an mp4 file using the generated BIFS file and provides a preview function for a 3D MPEG-4 scene. The work of content model for mobile adaptation of multimedia information [6] proposes a new abstract model to represent and adapt multimedia to hybrid environments. The model includes a layered mapping of semantic and physical entities and is

combined under the taxonomy of multimedia adaptation to optimize end-to-end service. The adaptation taxonomy consists of two parts; semantic adaptation and physical adaptation. The semantic adaptation is based on users' and service providers' choice and it is affected by the semantic content of a presentation. The physical adaptation is based on physical QoS and the characteristics of media objects consisting multimedia contents. The work of adapting multimedia internet content for universal access [7] presents a system adapting multi-media web documents to optimally match the capabilities of the client device requesting it. The system shows that multimedia contents are adapted by using two components; a representation scheme that provides a multimodal and multi-resolution representation hierarchy for multimedia and a customizer that selects the best content representation to meet the client capabilities while delivering the most value. The scalability for adaptive MPEG-4 contents describes an adapting technique for MPEG-4 contents on various service environments. The adapting technique includes adaptations of each media object in a scene and streaming service using network bandwidth. But the technique doesn't support to author MEPG-4 contents with interactive capability for mobile devices such as a PDA device, and it introduce capability which MPEG-4 contents is reconstructed using media objects in a scene.

Most of above-described researches are difficult to suitably support contents authoring for mobile devices. Conventional MEPG-4 contents need an adaptation in order to be used on mobile devices. So, in this paper, we propose a transformation method that uses characteristics of a mobile device in authoring contents.

3 Transformation of MPEG-4 Contents

Conventional MPEG-4 contents are difficult to present on a PDA device because they are optimized at desktop computers. Specially, they have a problem on a displayer due to difference on a screen size of a PDA device and conventional devices. So, in this paper, we consider two sides on authoring MPEG-4 contents. First, we consider reduction of a physical size of each object in each scene because of a small screen size of a PDA device. Second, we consider a reduction of transmission quantity on content's stream for fast initial presentation on a PDA device.

The figure 1 presents an authoring course of MPEG-4 contents for a PDA device. The authoring course includes 5 steps as following; arrangement of media objects, and setting up of object's attributes, management of scene description information, and scaling objects, adjustment of scene tree, encoding of BIFS, and generation of an mp4 file. The first step (arrangement of media objects, and setting up of an object's attributes) supports to compose a scene using geometry objects which represent media objects. The object's attributes which set up visual characteristics at each media object are established by a user's a mouse click or simply textual input on a dialog box. The second step (management of scene description information, and transforming objects) supports to generate a scene

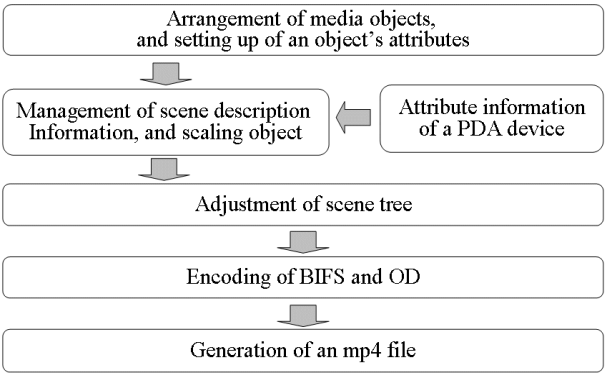

Fig. 1. An MPEG-4 authoring course for a PDA device

tree and scale a media object. The generated scene tree is used to manage scene description information, and each node in the scene tree has information on a temporal and spatial relation among objects as following; a geometry object, an image object, and a video or audio object. Whenever an object is generated in an authoring space, an object node is added in the scene tree. And an attribute object is added as a subordinate node of the object node in the scene tree on setting up temporal and spatial information. The subordinate node has information about form, location, playback time, etc. of a media object. Such information of object's attributes is represented as an object descriptor(OD) in MPEG-4 contents. The third step (adjustment of a scene tree) supports to modify information of a node in the scene tree and physical size of a media object in order to generate MPEG-4 contents for a PDA device. The fourth step (encoding of BIFS and OD) supports to generate BIFS and OD contents from information of each node in the adjusted scene tree using rules generating BIFS and OD. The BIFS is represented as a textual form, and then they are encoded as a binary form. The fifth step (generation of an mp4 file) supports to generate an mp4 file from multi-flexing an encoded binary data (BIFS and OD).

The figure 2 presents a change of a scene tree using the suggested transforming technique. The upper part and lower part in the figure presents an original scene tree and a scaled scene tree respectively. In the figure, the part showed dotted line presents an image or video object, and the part is transformed into a geometry object by the suggested transforming technique. In the figure 2, the GO1 object or GO2 object means a visual object as followings; rectangle, triangle, line, etc. In a scaled scene tree, the geometry object has subordinate nodes that have information in order to activate a corresponding media object in an original scene tree. So, when contents are played, an original media object described in an original scene tree is replayed if user clicks the generated geometry object in the scaled scene.

In order to generate contents for a PDA device, information of media objects to be scaled is extracted from a scene tree. The information is made as a form of

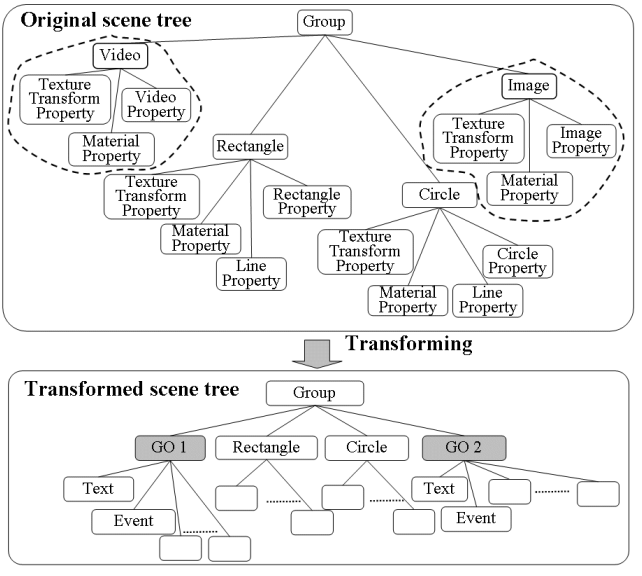

Fig. 2. Change of a scene tree by the transforming technique (A rounded rectangular in a tree presents a node on a scene tree)

linked list, and is used to scale and transform a media object. The transforming a media object transforms the media object into a geometry object, and generates an event which supports to play an original media object instead of showing the geometry object. The event is added as a subordinate node of a corresponding node in a scaled scene tree.

The figure 3 presents a transforming process of a scene tree. In this process, a screen size for playback is decided by using location values of media objects in each scene of contents. The minimum screen size for playback on a PDA device is computed by leftmost, rightmost, uppermost, and lowermost location of each media object in a scene. Also, the size is used to compute a reduction rate in contrast with a screen size for playback of original contents. The reduction rate is applied to reduce an external size of each geometry object (line, circle, rectangular, etc.) in a scene, and adjusts a coordinates' value of the geometry object, too. The rest of a text, an image, an audio, a video, and an event object are scaled as followings; In the case of a text object, a big reduction can make user not seeing itself on a PDA device. So, in order to avoid this problem, a reduction rate for exceeding a certain font size should be considered. In contrary of a text object, an image object dose not reduced, but if a size of an image object exceeds a screen size of a PDA device, a corresponding node in a scene tree is removed in the transforming technique. So, when an external size of an image object is simply smaller than a screen size of a PDA device, the image object is play backed. In the case of a video object, a size of the object (QCIF(144x176) or CIF(288x176)) is mostly smaller than a screen size of a PDA device. However,

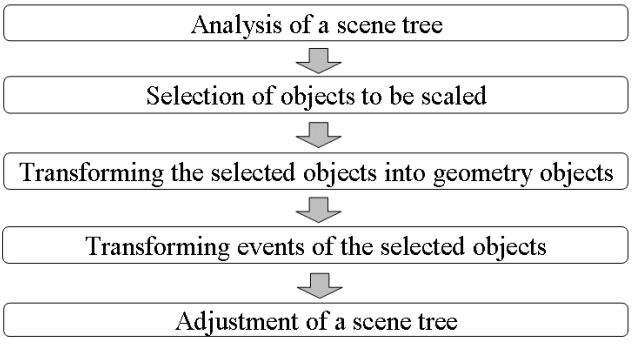

Fig. 3. Transforming process of a scene tree

lacking resources on a media player in a PDA device can bring about quality deterioration on playback. This problem is solved by encoding scalability, but is not considered in this paper. In the case of an image or a video object with temporal information, if the objects should be played back at same location on a PDA screen in regular sequence, a firstly played object is substituted for a geometry object. The geometry object indicates a text 'I'(image) or 'V'(video) according to a really played object. Activation of the geometry object by a users' click ignites playback of connected media objects with time.

4 Implementation and Results

In this paper, the proposed transforming method is applied to a conventional authoring tool [8] for presentation of MPEG-4 contents on a PDA device (screen size: 288x352). Authored MPEG-4 contents using the proposed transforming method indicate geometry objects instead of visual objects on a PDA device. The geometry objects ignite playback of a connected real media object by users' click.

The figure 4 presents MPEG-4 contents on a PC and a PDA respectively. In the figure, left side figure present a complete scene of MPEG-4 contents on a PC, and right side figure present a partial scene of MPEG-4 contents on a PDA due to the suggested transforming technique. However, we could find that the method showed an efficient presentation of the MPEG-4 contents on a PDA device.

The figure 5 presents a comparison between pre-application and post-application of the transforming method in authoring MPEG-4 contents. In the experiment, we compared a file size of an mp4 and BIFS according to increase in number of an image object or a geometry object; in the figure 4, the GO is shorts for a geometry object, and the IO is shorts for an image object. The experiment showed that a file size of BIFS is not enough of a difference whether a geometry object is included or an image object is included in contents. However, a size

Fig. 4. Presentation of MPEG-4 contents on a PC and a PDA

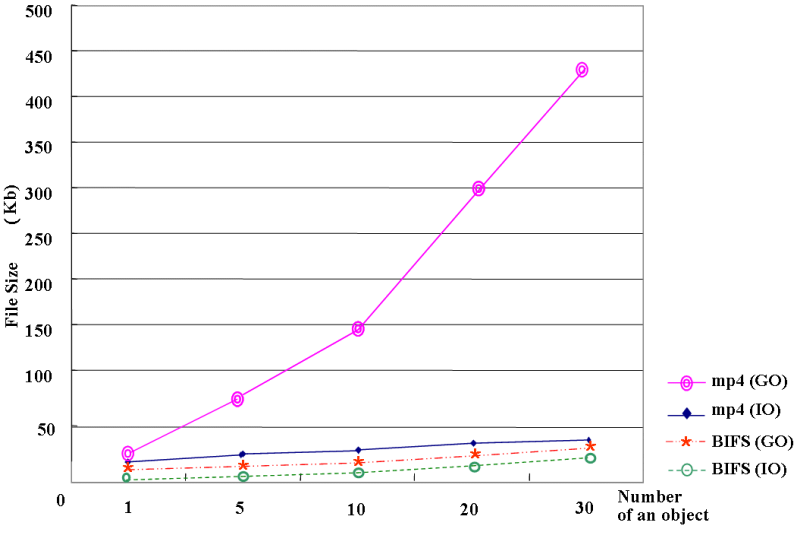

Fig. 5. Comparison between pre-application and post-application of the transforming method (GO: geometry object, IO: image object)

of an mp4 file has all the difference between using an image object and using a geometry object according to increase in number of an object.

In view of the result, scaled contents have a very small file size due to not transmitting a media object. Namely, only BIFS of scaled contents is enough of construction for whole scene and efficient playback on a PDA device. Generally, MPEG-4 contents with media objects are multiplexed with BIFS, OD, and media objects for generation of an mp4 file, which are transmitted to a player device. So the more contents have media objects, the more an mp4 file size is increased. But, a scene of scaled MPEG-4 contents uses geometry objects instead of visual objects; a PDA device receives only BIFS information for initial scene generation, and a user can play back an original visual object according to a user's click.

Such BIFS has not a large size due to not including a visual object, and it can reduce usage on resources of a PDA device and network bandwidth.

5 Conclusions

In this paper, we proposed a transforming method in order to present MPEG-4 contents on a PDA device. The transforming method selects visual objects to be scaled in a scene tree and transforms the selected objects into geometry objects or smaller objects to fit on a PDA screen for automatic generation of an mp4 file. Also the transforming method has hardly information lose of original contents. It is easily applicable to a conventional authoring tool, supports user's authoring contents with user not having information on BIFS, and supports to directly author contents for a PDA device on desktop environment.

Future works are development of an adjustment technique on a scene with events and an encoding technique for scalability on visual objects.

Acknowledgement. This work was supported by Korea Research Foundation Grant (KRF–2003–002–D00304).

References

1. Cha, K., Kim, H., Kim, S.: The Design and Development of MPEG-4 Contents Authoring System. J. of Korea Information Science Society, Vol. 7. (2001) 309–311
2. WG11(MPEG), MPEG-4 Overview (V.16 La Baule Version) document. ISO/IEC JTC1/SC29/WG11 N3747 (2000)
3. WG11(MPEG), MPEG-4 Overview (V.18 Singapore Version) document. ISO/IEC JTC1/SC29/WG11 N4030 (2001)
4. Shieh, M., Perngand, K., Chen, W.: The Design and Implementation of A Visual MPEG-4 Scene-Authoring Tool. Proc. of Workshop and Exhibition on MPEG-4 (2001) 21–24
5. Daras, P., Kompatsiaris, I., Strintzis, M.: MPEG-4 Authoring Tool for The Composition of 3D Audiovisual Scenes. Proc. of Second Int. Workshop on Digital and Computational Video (2001) 110–117
6. Metso, M., Koivisto, A., Sauvola , J.: Content Model for Mobile Adaptation of Multimedia Information. Proc. of IEEE 3rd Workshop on Multimedia Signal. (1999) 39–44
7. Mohan, R., Smith, J., Li, C.: Adapting Multimedia Internet Content for Universal Access. IEEE Transactions on Multimedia, Vol. 1. (1999) 104–114
8. Cha, K., Kim, S.: MPEG-4 STUDIO: An Object-Based Authoring System for MPEG-4 Contents. Multimedia Tools and Applications. (2003)
9. Lee, S., Cha, K., Kim, S.: An MPEG-4 Contents Authoring for Mobile Devices. Proc. of 2003 HCI (2003) 402–405
10. Cha, K.: A Scalability for Adaptive MPEG-4 Contents. Dissertation for The Degree of Doctor of Philosophy, Kyungpook National Univ., Korea (2003)