A Study on Optimized Watermarking Scheme for MPEG-4 Contents Services on Mobile Device

Dae-Jea Cho

Dept. Of Multimedia Engineering, Andong National University, South KOREA
djcho@anu.ac.kr

Abstract

Content for mobile devices, unlike content for PCs, need to be optimized for the features of the mobile device. In recent years, with the progress in mobile telecom technology including Wibro and 3G, various types of mobile devices have become widely available, such as PMP, PDA, or UMPC. This has drawn attention to interactive multimedia services such as IPTV, DMB broadcasting, mobile portals and mobile TVs. However, in order to provide appropriate services for the mobile environment, the process of converting MPEG-4 BIFS to LASeR is required.

This paper proposed optimized watermarking scheme for mobile device. Proposed watermarking algorithm uses z-axis location information of MPEG-4 BIFS and original model is not needed in extraction procedure. This proposed watermarking algorithm has good imperceptibility and robustness against the geometrical attacks such as noise addition, simplification and cropping. And it is proved that this proposed algorithm is good for access control capability by the experiment.

Keywords: Watermark, Mobile security, LASeR, MPEG-4 BIFS, 3D authentication

1. Introduction

Digital content is easily reproduced and since it is almost the same as the original, there is the issue of unauthorized reproduction and distribution. To resolve this issue, there is a need for copyright protection technology and authentication measures that can prevent illegal distribution and usage of digital content. Digital watermarking technology is an essential technology for confirming one's ownership and authentication. Watermarks which are specific forms of data are inserted into the digital content for later extraction, without changing the quality or size of the content. They are not visible or audible by the human eye or ear, either.

Use of the watermark varies in images and videos, as well as in illegal distribution as the current copyright issues, which is recognized in the form of MP3 music files can be applied to. But no matter what the subject of watermarking, several problems remain to be solved. As the size of inserting watermark is large, the quality of original image is getting worse. In other words, it is not easy to find the optimum condition because the size of the watermarks and quality of the images which have inserted watermarks have trade-off relationships, which means as one side becomes advantageous, the other side becomes disadvantageous [1]. Numerous researchers have been studying to insert much information in the original images which have limited sizes.

Thanks to the progress in embedded technology, mobile devices such as smartphones or PDAs are now widely distributed. With the advanced communication technology, mobile 3D games or DMB which are mobile based content services are also made possible. With these trends, mobile telecom companies and broadcasters are producing various 3D contents and offering them to consumers. However, since there is no way to prevent illegal reproduction of mobile content, it is difficult to protect copyrights. In

addition, there is much left to be desired in the field of mobile watermarking research that authenticates mobile content associated with mobile devices.

This paper proposes digital watermarking methods appropriate for mobile contents. Content for mobile devices, unlike content for PCs, need to be optimized for the features of the mobile device. In recent years, with the progress in mobile telecom technology including Wibro and 3G, various types of mobile devices have become widely available, such as PMP, PDA, or UMPC. This has drawn attention to interactive multimedia services such as IPTV, DMB broadcasting, mobile portals and mobile TVs. For interactive multimedia services, in MPEG-4, the multimedia objects such as the image, video and audio are object coded from the early stages of the suggested standard. Then BIFS (Binary Format for Scene), which is a scene description and scene description language used for two dimensional display were suggested. However, since BIFS is suggested based on PC, it is restrained in the mobile environment by the fact that there are limited operation resources and battery. As such, in MPEG-4, as an alternative to BIFS in the mobile environment, MPEG-4 part20 LASeR (Lightweight Application Scene Representation) standard was additionally suggested [2-6].

However, in order to provide appropriate services for the mobile environment, the process of converting MPEG-4 BIFS to LASeR is required. The conversion system from XMT to LASeR parses the XMT-A file that the user put in to generate a DOM-Tree on the memory, and in the XMT-A, converts the nodes that can be converted into LASeR-XML using an XSLT style sheet through an XSLT processor. In the case of three dimensional geometry and associated 3D nodes, associated nodes are searched and the value of the relevant node is converted into a value that can be processed in the LASeR node. Moreover, geometric coordinate conversion and 2D coordinate conversion from 3D coordinate are carried out. In the last stage, to the DOM-Tree generated in the memory is added an alternative node that can be mapped out by the LASeR standard SVG node, and the value is renewed to convert it into an LASeR-XML file [7].

Proposed watermarking algorithm uses z-axis location information of MPEG-4 BIFS and original model is not needed in extraction procedure. This proposed watermarking algorithm has good imperceptibility and robustness against the geometrical attacks such as noise addition, simplification and cropping attack. And it is proven that this proposed algorithm is good for access control capability by the experiment.

Composition of this paper is as follows. Chapter 2, introduces related works and comparison analysis about MPEG-4 BIFS and LASeR is to be executed, and chapter 3, explains optimized watermarking scheme for a mobile device. In chapter 4, implemented results of suggested algorithm are to be analyzed and compared with other similar research results. Finally in chapter 5, future study subjects are to be described and conclusion is to be made.

2. Related Works

In MPEG-4, Scene Description is to be defined to organize the scene which is composed of multimedia objects. Scene Description defines the expression about visual and auditory objects in time and space and the relationship of time and space between objects. And, it also defines the definition about user events and changeable information by user events. The Scene Description at this moment is called as BIFS(Binary Format For Scene)[7].

MPEG-4 BIFS is an international standard established by MPEG as a composition coding technology to manufacture rich media contents and it is supporting multimedia production using 2D and 3D, consecutive upgrades according to relevant scenes, streaming support in rich media and synchronization between individually different media [8]. In spite of these various technologies for rich media service including streaming,

BIFS is not suitable in mobile environment which has limitations in calculation capability and transmission bandwidth because it was developed based on PC.

In order to overcome these weaknesses and to provide rich media service in mobile environments, ISO/IEC JTC1 WG11 (MPEG) established ISO/IEC 14496-20(MPEG-4 Part 20: LASeR(Lightweight Application Scene Representation) as a new international standard. LASeR is a standard which defines content production and transmission method to provide various rich media service in mobile environments and it is an international standard which will enable various rich media services like web services from existing text or menu-bar services [9]. LASeR is a base technology for scene technology and has limited scene technology using SVG of W3C and can easily convert scene information to other different formats. And, it also can freely express video, audio and graphic data, and transmit by multiplexing as logical stream. Through this, dynamic rich media service can be implemented which supports rich screen expression and compression efficiency in mobile environments.

Figure 1 shows the Composition of LASeR system [9]. In SVG Scene Tree layer, a method is suggested to organize the elements which composes the scenes like media objects and graphic in space-time way. When interaction methods and scenes between the elements composing one scene are changed, a method to reflect the changed parts is suggested. In LASeR Extension layer, clipping function not supported by SVG standards which are base technologies of LASeR standards, limited non-resampling rotation which is based on video and still image, full-screen mode and multiple synchronization references are suggested. And, functions of Longkeypress, Pause & Resume and Text underline are also suggested.
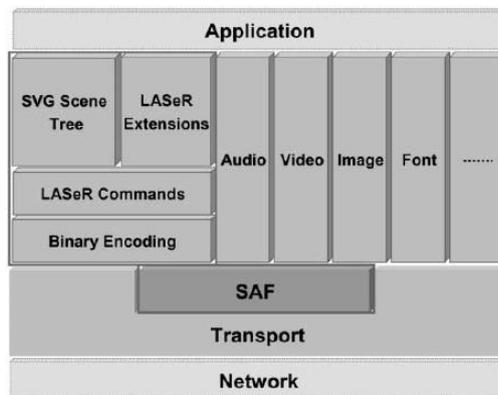
Figure 1. Architecture of LASeR [9]

In dynamic update layer, SVG functions of "Download & Play" and "Progressive Rendering Mode" and "DOM(Document Object Model)" based scene change function are supported in order to serve LASeR contents. Additionally, the function to separate one scene to number of transmission packets and the push function to insert new contents in current scene are additionally supported.

Binary Encoding layer, scene description which is composed of SVG Scene Tree and LASeR Extension is encoded as binary stream. In order to play relevant contents at this moment, parser for XML text processing is required. But, this XML Parser has a problem to require more computation than binary data decoding. Even though it can provide services at "download-and-play" because scene description which is expressed by XML is composed of texts, it has a problem not to provide services through stream. In order to solve this problem, new binary encoding method is provided in LASeR. For this, BiM (Binary format for Meta data) standards are defined in MPEG-7.

Finally, SAF (Simple Aggregation Format) is a transmission standard to transmit comprehensive data binary bit streams of LASeR which are composed of Scene Description and individual media elements. SAF standards enable simple and efficient synchronized transmission by multiplexing LASeR contents elements and subsequently making a single SAF stream [10, 11]. Table 1 shows the characteristics of BIFS and LASeR [7].

Table 1. Features of BIFS and LASeR

	BIFS	LASeR
Standard	MPEG-4 Part1 System	MPEG-4 Part20 LASeR
Target device	P.C	Mobile
Based technology	VRML	SVG
Scene representation	2D/3D	2D
Binary coding	Support	Support
Interactive	Support	Support
Streaming	Support	Support
Specifications	Complex	Simple
Dynamic scenes update	Support	Support

3. Optimized Watermarking Scheme for a Mobile Device

This chapter will explain optimized watermarking scheme for a mobile device. Watermark is inserted using z-axis coordinates on the projection procedures of 3D coordinates to 2D coordinates. When Cylinder, Sphere and other nodes are transformed into LASeR nodes, digital watermark is inserted.

3.1. Watermark Insertion Process

BIFS and LASeR have attributes including location, color, horizontal and vertical length and radius in order to represent the geometry object. Because BIFS and LASeR have different composing methods of these property information, conversion is required that LASeR node can be expressed using extracted property information from BIFS. Figure 2 shows a hierarchical structure of nodes to represent the geometrical object of BIFS [12].
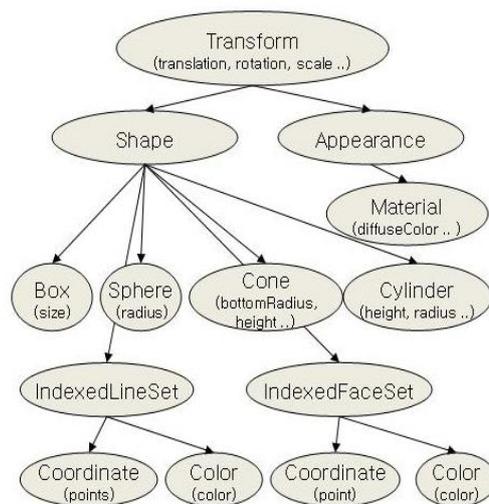


Figure 2. The Hierarchy of BIFS 3D Geometry Nodes

It expressed the node name in the elliptical of figure 2 and the main attribute expressed in parentheses under the node name. Box which is 3D geometry object node of BIFS and Cone, Cylinder, Sphere have individual object's size information properties including horizontal and vertical length, height and radius. And information about object coordinates and rotation and expansion are in Transform which is a higher level node and Material node has color information.

On the other hand, LASeR defines only 2D geometric objects and figure 3 shows the hierarchy of LASeR 2D geometry nodes [12].
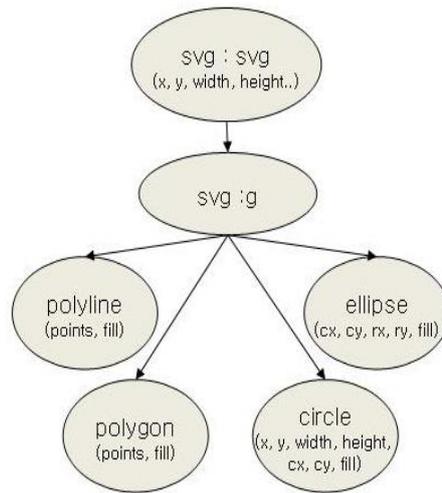
Figure 3. The Hierarchy of LASeR 2D Geometry Nodes

LASeR 2D geometric objects which are ellipse, circle, polyline and polygon include object properties such as size, coordinates and color information. Therefore, in order to express BIFS's 3D nodes as LASeR 2D nodes, first BIFS's Transform node and Material, as well as the value of each geometric object are first extracted and this information is converted to be expressed as an LASeR node. In this paper, watermark is inserted during this conversion process. Figure 4 shows the watermark insertion process.
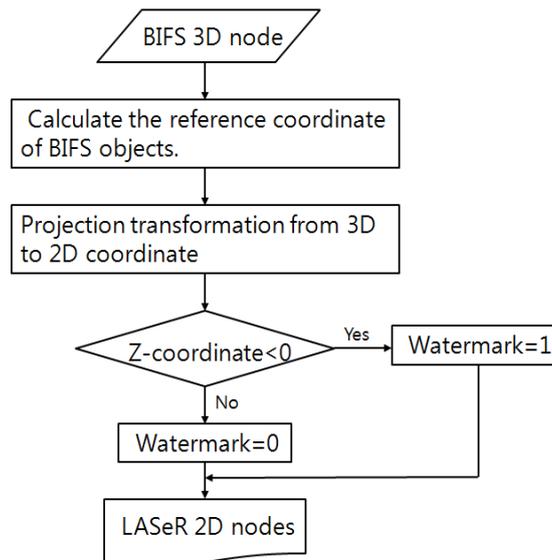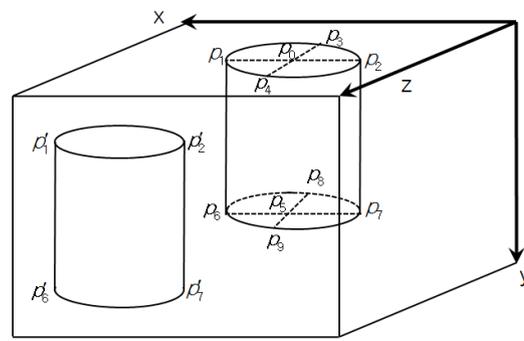

Figure 4. Watermark Insertion Process

First, the property of the BIFS 3D geometric object is extracted. For example, BIFS's Cylinder object has a height property value and a radius property value. Using these property values, the 10 peak coordinates of the Cylinder are calculated. To express them in LASeR, the necessary peak coordinates are defined as reference coordinates. Once the reference coordinates are gained, the location where the actual object will be expressed is calculated using the location movement information, rotation information and expansion information in the Transform node of BIFS. Next, 3D coordinates are projected and converted into 2D coordinates. BIFS defines the center of a scene as ground zero, while LASeR defines the upper left side as the ground zero. Therefore, a coordinate conversion is required.

If it is projected as 2D coordinates at this moment, z-axis coordinates will disappear in original 3D coordinates. Figure 5 shows the procedures to project 3D coordinates to original 2D coordinates. z-axis coordinates P3, P4, P8 and P9 will disappear.

In this paper, watermark is inserted using z-axis coordinates on the projection procedures of 3D coordinates to 2D coordinates. For example, if coordinate value of z-axis is negative, insert '1' as watermark and if it is positive, insert '0'.

Figure 5. The projection conversion process of 3D coordinates to 2D coordinates [13]

3.2. Watermark Insertion Algorithm for Sphere Nodes [14]

Sphere node generates 7 reference coordinate values to express the Sphere as attribute values and radius. Figure 6 shows properties of Sphere node and reference coordinates on Sphere node to express LASeR.
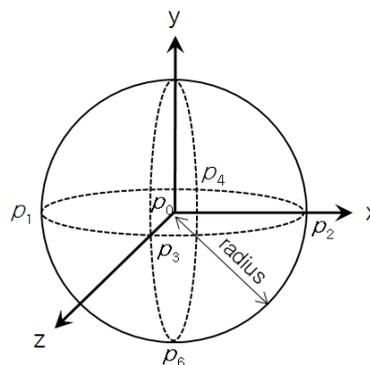
Figure 6. The Reference Coordinates of Sphere Node

Algorithm to insert watermarks in Sphere node is as follows.
Step 1: Extract the properties of 3D geometry objects in BIFS.
Step 2: Calculate 7 reference coordinates which compose Sphere using Radius attribute values of Sphere node.

P0=(0, 0, 0),   P1=(-r, 0, 0), P2=(r, 0, 0),   P3=(0, 0, r), P4=(0, 0, -r),   P5=(0, r, 0), P6=(0, -r, 0)

Step 3: Calculate the locations where physical objects are expressed through location movement, rotation and expansion using location movement information, rotation information and expansion information which are in BIFS Transform node.

Step 4: Project and convert 3D coordinates to 2D coordinates and execute the conversion of coordinate system in acquired 2D coordinates and express as one svg : circle. Insert watermarks as follows using 3D reference coordinates which were used at this moment.

if (3D reference coordinate value of z-axis < 0) then watermark=1
Else watermark=0;
Step 5: Express as LASeR node using converted 2D coordinates on step 4.

3.3. Watermark Insertion Algorithm for Cylinder Node

Cylinder node generates 10 reference coordinate values to express the cylinder as attribute values, radius and height. Figure 7 shows the attributes of Cylinder node and necessary reference coordinates to express as LASeR.
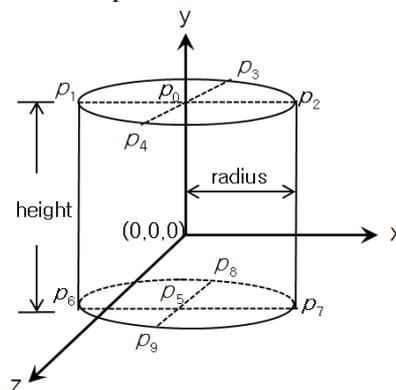
Figure 7. The Reference Coordinates of Cylinder Node

Algorithm to insert watermarks in Cylinder node is as follows.
Step 1: Extract the properties of 3D geometry objects in BIFS.
Step 2: Calculate 10 reference coordinates which compose Cylinder using Radius, Height attribute values of Cylinder node.
P0=(0, h/2, 0),   P1=(-r, h/2, 0), P2=(r, h/2, 0), P3=(0, h/2, -r), P4=(0, h/2, r),
P5=(0, -h/2, 0), P6=(-r, -h/2, 0), P7=(r, -h/2, 0), P8=(0, -h/2, -r),   P9=(0, -h/2, r)
Step 3: Calculate the locations where physical objects are expressed through location movement, rotation and expansion using location movement information, rotation information and expansion information which are in BIFS Transform node.

Step 4: Project and convert 3D coordinates to 2D coordinates and execute the conversion of coordinate system in acquired 2D coordinates and express as two svg : ellipse and one svg : polygon. Insert watermarks as follows using 3D reference coordinates which were used at this moment.

if (3D reference coordinate value of z-axis < 0) then watermark=1
Else watermark=0;
Step 5:
If (y-axis coordinate of P3 = y-axis coordinate of P4) then express as svg : polygon
Else if (x-axis and y-axis coordinate of P1 = x-axis and y-axis coordinate of P6) then express as svg : circle
Step 6: Express as LASeR node using converted 2D coordinates on step 4.

3.4. Watermark Insertion Algorithm for the Other Nodes

The algorithm to insert watermarks in the nodes other than the object(Box, Cylinder, Sphere, Cone node) is as follows. Like in Figure 8, calculate the average by summing the pixel values from $A_{i-1,j-1}$ in the clockwise direction. If $A_{i,j}$ is greater than the average when compared with the average and $A_{i,j}$, $A_{i,j}$ is increased.

When extracting watermarks, original image is not required and we can extract watermark comparing variation of $A_{i,j}$.

if $(A_{i-1,j-1}+A_{i-1,j}+A_{i-1,j+1}+A_{i,j+1}+A_{i+1,j+1}+A_{i+1,j}+A_{i+1,j-1}+A_{i,j-1})/8 < A_{i,j}$
    then $A_{i,j}$++;
    else $A_{i,j}$--;

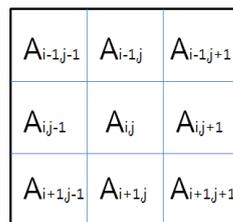| $A_{i-1,j-1}$	$A_{i-1,j}$	$A_{i-1,j+1}$
$A_{i,j-1}$	$A_{i,j}$	$A_{i,j+1}$
$A_{i+1,j-1}$	$A_{i+1,j}$	$A_{i+1,j+1}$

Figure 8. Watermark Mask for the Nodes other than the Object

4. Experimental Results

All the tests were carried out on the P.C. equipped with an Intel core2 2.8 GHz processor and 8GB memory. LASeR based 3D representation system is implemented with JAVA in eclipse3.4 development platform. The proposed method is implemented in C++, and has been tested on several 3D models. Table 2 details information of 3D models and Figure 9 shows four of them: Object1, Object2, Object3 and Object4.

Table 2. Information of 3D Models

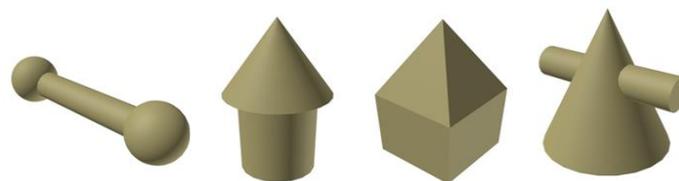| Model	Number of reference coordinates	Number of polygons
Object1	24	3
Object2	16	4
Object3	13	9
Object4	16	5

Figure 9. Original 3D Models: Object1, Object2, Object3, Object4

In this paper, proposed method was verified by distortion, robustness and imperceptibility evaluation. The objective distortion between the normalized cover and stego 3D models are measured by MRMS (maximum root mean square error) [15].

The robustness indicates how resistant the watermark is against various routine operations on the watermarked content. The robustness is evaluated in terms of the

BER(bit error rate) of the extracted watermark bit sequence, as well as the correlation coefficient[16] between the extracted watermark bit string w_n' and the originally inserted one w_n as given by the following equation:

$$\text{Corr} = \frac{\sum_{n=1}^{N=1}(w_n' - \overline{w'})(w_n - \overline{w})}{\sqrt{\sum_{n=1}^{N=1}(w_n' - \overline{w'})^2 \sum_{n=1}^{N-1}(w_n - \overline{w})^2}} \tag{1}$$

where $\overline{w'}$ and $\overline{w}$ indicate, respectively, the averages of the watermark bitstrings w_n' and w_n.

The imperceptibility of a watermarking algorithm refers to the perceptual similarity between the original and watermarked images. It is useful to measure the extent of distortion that the watermarking introduces to the work for the quality appraisal. PSNR is measured to compare visual quality between the cover image and watermark embedded image[17].

$$PSNR(I, I_w) = 10 \, log_{10} \frac{L_{max}^2}{MSE(I,I_w)} \tag{2}$$

where I is the original image and I_w is the watermarked image and L_{max} is the maximum value of luminance level(0-255). The mean squared error is given as

$$MSE(I, I_w) = \frac{1}{m \times n} \sum_i^{m-1} \sum_j^{n-1} |I(i,j) - I_w(i,j)|^2 \tag{3}$$

where m and n denote the width and height of the cover image. Distortion is generally indiscernible to the human eye when PSNR exceeds 30 dB. The average PSNR of four watermark embedded models is over 48dB which demonstrates that our watermarking scheme has good imperceptibility.

To evaluate the robustness of proposed scheme, the watermark embedded models are tested against attacks such as noise addition, simplification and cropping. The BER due to mesh simplification, adding noise and cropping attack is shown in Figure 10. Noise addition attack aims to simulate the artifacts introduced during 3D mesh generation and the errors induced during data transmission. In this experiment, pseudo-random noises are added.
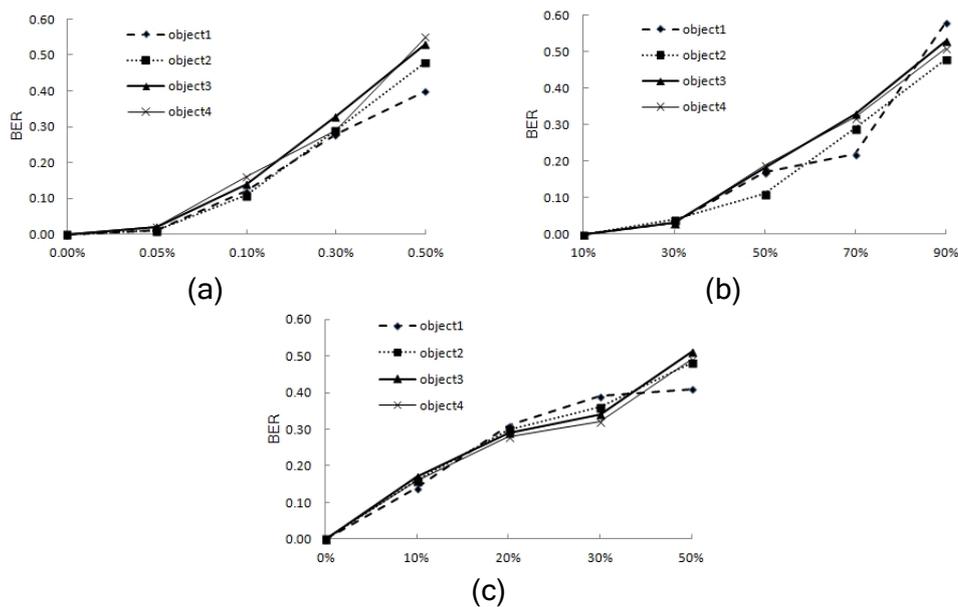
Figure 10. Experimental Results after Various Attacks; (a) Noise Addition, (b) Simplification and (c) Cropping

The watermark is often embedded in the original complex model, and then the 3D mesh is simplified so as to adapt to the capacity of the available resources. This simplification operation is the most important one of various connectivity attacks. After cropping attack, one part of the watermarked 3D model is cut off and thus lost. This attack happens when we create a new model by combining parts extracted from several other objects.

The correlation coefficient affected by noise addition, simplification and cropping attack is shown in Figure 11.

(a)

(b)

(c)
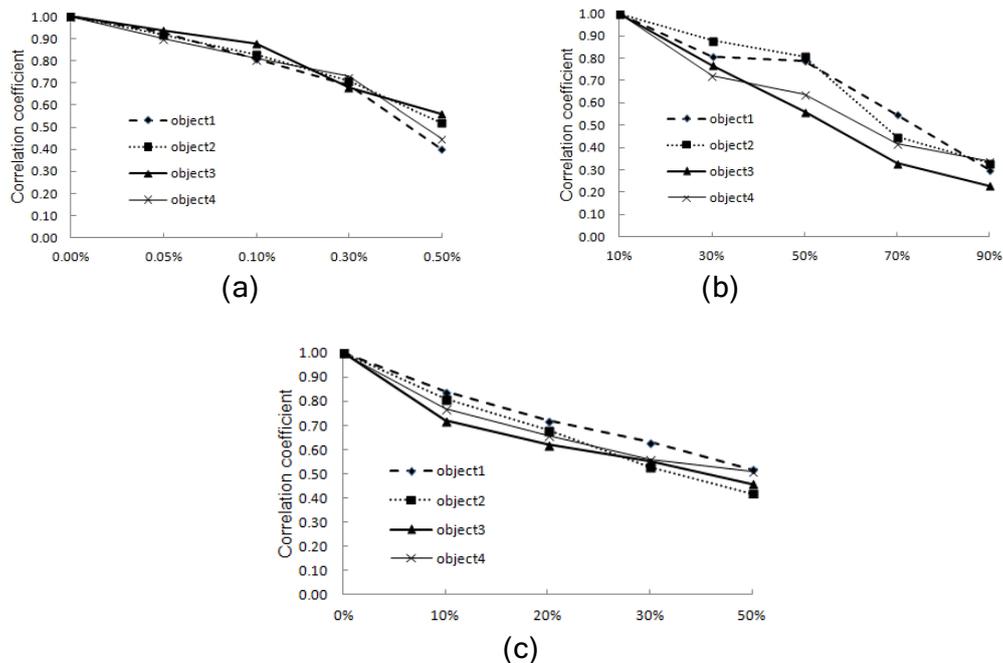
Figure 11. Correlation coefficient affected by noise addition, simplification and cropping. (a) Noise addition, (b) Simplification and (c) Cropping

From Figure 10 and Figure 11, we can see that for the four models, the BER is endurable and the correlation coefficient values are large enough to detect the watermark after various attacking operations on the watermarked model such as mesh noise addition, simplification and copping attack.

5. Conclusion

Content for mobile devices, unlike content for PCs, need to be optimized for the features of the mobile device. In recent years, with the progress in mobile telecom technology including Wibro and 3G, various types of mobile devices have become widely available, such as PMP, PDA, or UMPC. This has drawn attention to interactive multimedia services such as IPTV, DMB broadcasting, mobile portals and mobile TVs. However, in order to provide appropriate services for the mobile environment, the process of converting MPEG-4 BIFS to LASeR is required. The conversion system from XMT to LASeR parses the XMT-A file that the user put in to generate a DOM-Tree on the memory, and in the XMT-A, converts the nodes that can be converted into LASeR-XML using an XSLT style sheet through an XSLT processor. In the case of three dimensional geometry and associated 3D nodes, associated nodes are searched and the value of the relevant node is converted into a value that can be processed in the LASeR node. Moreover, geometric coordinate conversion and 2D coordinate conversion from 3D

coordinate are carried out. In the last stage, to the DOM-Tree generated in the memory is added an alternative node that can be mapped out by the LASeR standard SVG node, and the value is renewed to convert it into a LASeR-XML file.

This paper proposed optimized watermarking scheme for mobile device. In this paper, suggested method is that watermark is inserted in the conversion process of 3D graphic nodes in MPEG-4 BIFS into LASeR 2D graphic nodes. Proposed watermarking algorithm uses z-axis location information of MPEG-4 BIFS and original model is not needed in extraction procedure. This proposed watermarking algorithm has good imperceptibility and robustness against the geometrical attacks such as noise addition, simplification and cropping attack. Experiments show that the proposed scheme has good imperceptibility and is able to resist against common attacks such as noise addition, simplification and cropping attack. The future research is to develop the algorithm for embedding and extracting a watermark in the rendering information, when rendering 3D mesh models.

Acknowledgements

References

[1] Kai Wang, Guillaume Lavoue, Florence Denis and Atilla Baskurt, Robust and Blind Mesh Watermarking based on Volume Moments. Journal of Computers &Graphics **(2011)**, Vol. 35, pp. 1-19.

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

[3] ISO/IEC FDIS 14496-11:2005(E) Information technology : Coding of audio-visual objects - Part 11 : Scene description and application engine.

[4] IBM Toolkit for MPEG-4, http://www .alphaworks.ibm.com/tech/tk4mpeg4

[5] MPEG-4 Tools, http://gpac.sourceforge.net/ index.php

[6] Qonita M. Shahab, A Study on Transcoding of MPEG-4 BIFS scenes to MPEG-4 LASeR Scenes in MPEG-21 Digital Item Adaptation Framework, School of Engineering Information Center Univ, **(2006)**

[7] I. G. Ko, A Converting Method of MPEG-4 XMT-A 3D Graphics to MPEG-4 LASeR for Mobile Environment. Md. Thesis of Andong Nat. University **(2008)**

[8] ISO/IEC 14496-1:2004 Information technology : Coding of audio-visual objects - Part 1:Systems.

[9] ISO/IEC 14496-20:2006(E) Information technology : Coding of audio-visual objects - Part 20: Lightweight Application Scene Representation (LASeR) and Simple Aggregation Format(SAF).

[10] K. H. Kim, Lightweight Application Scene Representation: LASeR. Journal of IEIE **(2006)**, Vol. 33, No. 9, pp. 984-990.

[11] Jean-Claude Duford, Olivier Avaro and Cyril Concolate, LASeR : the MPEG Standard for Rich Media Services, IEEE Multimedia.

[12] H. S. Kim and D. J. Cho, 3D Object Representation Method based on MPEG-4 LASeR. Journal of KIIT, **(2009)**, No. 7, Vol. 5, pp. 281-289.

[13] D. J. Cho, Watermarking Scheme of MPEG-4 LASeR Object for Mobile Device. International Journal of Security and its Application, **(2015)**, Vol. 9, No. 1, pp. 305-312.

[14] D. J. Cho, Watermarking Scheme of Sphere Object Based on MPEG-4 LASeR. Proceedings of Advanced Science and Technology Letters, **(2015)**, Vol. 109, pp. 1-4.

[15] Cox I. J., Miller ML, Bloom JA, Fridrich J. and Kalker T, Digital Watermarking and Steganography, 2nd ed. Morgan Kaufmann Publishers Inc. **(2007)**

[16] G. Taubin, Geometric Signal Processing on Polygonal Meshes. Proceedings of the Eurographics State of the Art Reports, **(2000)**, pp. 81-96.

[17] Kai WANG, Guillaume Lavou´e, Florence Denis, Atilla Baskurt, Xiyan He, A Benchmark for 3D Mesh Watermarking. Proceedings of 2010 Shape Modeling International Conference, IEEE computer Society, **(2010)**, pp. 231-235.

Author

Dae-Jea Cho, He received the Ph.D. degree in computer engineering from Kyungpook National University, Korea, 2001. He is professor at Andong National University, Korea. His areas of interest include digital watermarking and multimedia contents authentication.