Journal of Korea Multimedia Society Vol. 11, No. 12, DECEMBER 2008(pp. 1765-1774)

Design of Hot-spot generator for inserting content-based interactive advertising on DMB broadcasting

Tae-Jin Park[†], Se-Hyun Park[††], Soon-Bum Lim[†††], Yoon-Chul Choy[††††]

ABSTRACT

Due to recent increase of product placement(aka. PPL) in TV drama and entertainment program, the audience of these programs is exposed to the advertisement in the form of the props and backgrounds without discernment. Advancement in digital data broadcasting technology enabled advertisers to insert detailed information of the product or associated advertisement link into the content of the broadcast program as an interactive advertisement. With the interactive advertisement, the audience can focus on the content of the program and the advertiser can create an advertisement with a rich content through the interaction between the advertisement and the audience. It is possible because the interactive advertisement is dynamic advertisement behavior only activated by the audience's intention. In this research, we develop a Hot-spot generator to insert interactive advertisement in the DMB broadcast program. We also suggest dynamic Hot-spot generation technique that allows the Hot-spot object to follow not only the shape of the advertising object in the program, but also the movement of the object as the program progresses.

Key words: DMB, Digital Data broadcasting, BIFS, interactive advertising, Hot-spot

1. INTRODUCTION

With recent breakthrough of digital media technology, PVR(Personal Video Recorder) device with time shift feature as Tivo does and N-PVR (Networked PVR) services are widely spread. Using the time shift feature, the audience now can enjoy only the broadcast program without having to watch the advertisement[1]. Moreover, IPTV services allow their subscribers to avoid advertisements while enjoying the content at any time they prefer. The advertiser relies heavily on the PPL exposing the specific products as props in the target drama or entertainment program to counteract the effect of these advertisement avoidance, hence the audience is surrounded by the advertisements not knowing they are exposed to the advertisement while they're watching TV programs. The media consortium in Germany addresses the PPL as 'the concealed advertisement'. That is, the consortium regards the PPL as a deception against the audience.

Advancement in digital data broadcasting technology enabled advertisers to insert detailed information of the product or associated advertisement link into the content of the broadcast program as an interactive advertisement. The advertiser can use Hot-spot feature in the interactive data broadcasting service to provide interactive advertise-

※ Corresponding Author : Tae-Jin Park, Address : (120-749) 134 Sinchon-Dong, Seodaemun-Gu, Seoul, Korea, TEL : +82-2-393-7663, FAX : +82-2-393-7663, E-mail : parktj2003@gmail.com

Receipt date : Nov. 5, 2008, Approval date : Dec. 19, 2008

† Department of Computer Science, Yonsei University, Seoul, Korea

†† Department of Computer Science, Yonsei University, Seoul, Korea
(E-mail : i2workshop@gmail.com)

††† Department of Multimedia Science, Sookmyung Women's University, Seoul, Korea
(E-mail : sblim@sookmyung.ac.kr)

†††† Department of Computer Science, Yonsei University, Seoul, Korea
(E-mail : ycchoy@rainbow.yonsei.ac.kr)

※ This work is supported by Seoul R&BD(10581) program.

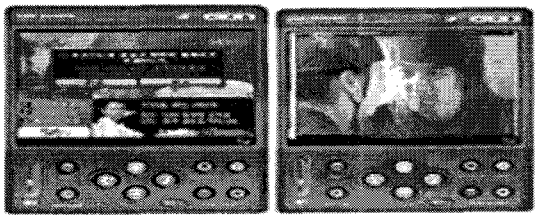
ment. Hot-spot indicates the area on the screen where the audience clicks specific area to get more detailed information or visit other pages from the Internet with touch screen or pointing device[2]. For example, by clicking Hot-spot, the audience is able to search the Internet for additional information of the actor/actress starring in the drama, or buy the cloth the actor/actress put on right away.

This research suggests dynamic Hot-spot generation technique that supports the shape of the advertising object other than simple button or menus, and also follows the movement of the object as the program progresses.

2. RELATED WORKS

The advent of digital broadcasting era brings interactive advertisement to the stage allowing communications with the audience. Interactive advertisement can be defined as a new method of advertising that the consumer can evaluate, ask for more information like brochures, and catalogs, acquire sample products or discount coupons, or even buy the advertised product at the moment of the advertisement is being watched[3].

There are many forms of interactive advertisement according to the characteristics of the medium or the program genres. Various interactive advertisements can be divided into following 3 classes according to the advertisement entry time-trigger exposure time. CIC(CM in CM) if the trigger is placed within the CM, CIP(CM in Program) if it's in the program, and finally, the trigger can be exposed in the banner window that appears by the action of the remote control[4]. If the classification criterion is the outer form of the trigger the audiences interact, they can be fall into either the simple shape interactive advertisement or the content-based interactive advertisement. In the case of the simple shape interactive advertisement like Figure 1(a), the audience can select the



(a) simple menu shape (b) content-based shape

Fig. 1. Interactive advertisements

Hot-spot of a simple menu type to vote any program of interest, or request for an admission ticket while watching the program. In Figure 1(b), if the audience finds an actress she likes with a beautiful accessory while watching a drama, then she can click the accessory to get more information or purchase it.

In the view of advertisers, CIP and content-based advertisements are more effective since these types do not need separated advertisement time, and the audience can recognize the advertisement more intuitively. However, the development of the authorware-software that supports the dynamic interactive advertisements in the digital broadcasting program that changes every frame-is still insufficient.

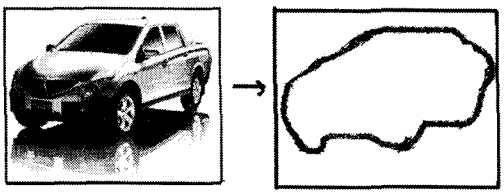
If the interactive advertisements are to be inserted in the digital broadcasting program, the support for the interactive data broadcasting environment is necessary. As of today, T-DMB (Terrestrial - Digital Multimedia Broadcasting) in Korea supports interactive data service on MPEG-4 BIFS(Binary Format for Scene) standard[5,6]. BIFS inherited from VRML(Virtual Reality Modeling Language)[7] standard, a graphic data description language that visualizes 3D Space on the Web, modifying and updating VRML from text based to binary to improve the transmission efficiency[8]. In BIFS, a whole content is segmented into several objects that compose a scene and compressed. The content also includes scene description which describes time & spatial layout of each object. The scene description uses a scene

tree that contains objects as nodes to express the content[9]. Each node describes not only visual feature data, and spatial position data but also required information for the interaction between nodes. In this research, to insert interactive advertisement into the digital broadcasting program, arbitrary area of interest in the broadcasting program is set to the trigger area. Then, the trigger area is described as Hot-spot after it is represented with BIFS Graphics objects and the relevant link is established. In other words, Hot-spot is interactive advertisement trigger that has the same 2D-shape of the target object to be advertised and moves along with the target object.

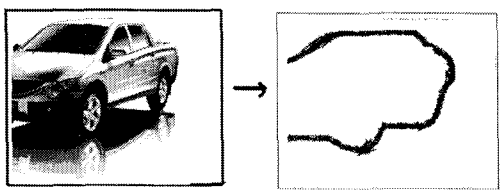
3. SUGGESTED SYSTEM

3.1 General System Flow

Extraction of spatial area of the target advertising object is a prerequisite for inserting a content-based interactive advertisement. As an example, let's say a hero of a drama is driving a car, and the car is the target object to be advertised. Following the plot, the car can move into or goes out of the frame. As the car comes closer to the audience or goes farther, the shape of the Hot-spot should change accordingly.



(a) Appearance of the target object



(b) Target object moves out of the screen

Fig. 2. Content-based Hot-spot extraction

Table 1. Basic Process for Hot-spot

Process	Feature
Shot Detection	Forms a group from similar frames as a scene
Object Extraction	Extracts Contour of object areas
Segment Tracking	Generates object areas using previous selected area by user pen-input
BIFS Encoding	Converts to BIFS code

Table 1 lists general workflow of the suggested system that extracts target objects area information for interactive advertisement insertion in the broadcasting programs and generates Hot-spot using the extracted object area.

Hot-spot designation in continuous video stream requires various procedures. First of all, video processing for segmentation of a whole video stream into the scenes is required. Secondly, image processing for extracting object areas on key frames, and the last step is data service conversion for changing extracted object area into the appropriate DMB BIFS data.

The first step is getting the position of the target object in the whole video stream. The basic unit of dynamic video stream is a frame. Scene change in the video stream is called Cut, and small video unit captured by one camera action is Shot. Segmenting video into shots is called Video Segmentation, and detecting transitional cut where scene change occurs to do video segmentation is called Shot Detection[10]. The content-based Hot-spot generation this research suggests require shot detection to separate the whole video stream into the parts that have the same time feature that is a unit shot.

If the advertisement target object does appear in the segmented shot, the shot is processed to the frames, and the key frames are extracted from the frames of the shot.

The selected key frames hold the object to be advertised, and the objects will move in the unit

shot as the time progresses. Selected key frames go through image processing to extract a target object area. Extracting specific object area from the given image is called Image Segmentation. Image segmentation methods can be categorized into largely two methods: Automatic Segmentation and Semi-automatic Segmentation. Result from automatic segmentation is usually unsatisfactory as the definition of video object happens to be subjective. On the other hand, user's indication of the video object on the first frame or on the frame where it first appears or location data of interest area will result in more satisfactory extraction[11], [12]. For this research, pen-based user input interface is used for the semi-automatic segmentation.

The target object of the advertisement continuously changes it's shape and position as the video stream that contains the object changes dynamically. Thus, the Hot-spot area defined in a key frame differs from the broadcasting program in the process of the time. Setting a key frame for every unit frame in the image segmentation may solve this problem, but it will require too much time and cost. As the movement of the object can occur in the very short time frame leaving the audience without acknowledge the movement at all, it is better to leave the Hot-spot keep the position for a moment giving the audience a chance to acknowledge the Hot-spot and select it. The data rate for information on Hot-spot area is limited due to the bandwidth limitation of the data broadcasting.

Moreover, extracting the target object in every key frame is not an easy job. If the interval between the key frames is too small, the work load of the advertisement producer will increase dramatically, if too large, the Hot-spot area and the object in broadcasting program will differ. Hence, a special technique is required so that generation of key frames with small intervals would not levy too much load on the advertisement producer by helping the input work of the Hot-spot area. This research tries to minimize these work load. By using the prior input stroke, the advertisement producer doesn't have to input the are of the target object area for every key frame.

As the last step, the object area processed through the image processing is converted into DMB BIFS code via data service converter. These BIFS codes are compressed and transmitted to DMB devices. Finally, Hot-spot is triggered by user's interaction while watching DMB TV.

3.2 Design of Hot-spot generator

This section describes suggested system organization and details on the action mechanism. System organization is shown at following Figure 3.

The Hot-spot generator system is composed of three phases: shot detection phase, image segmentation phase, and BIFS conversion phase. Also there are 3 input sources: original video stream input for the interactive advertisement insertion, pen input for setting the area of interest in the key frames, and parameter input to limit the data according to data rate limitations of the data broadcasting. The output is mp4 file that contains interactive advertisement. This mp4 output file is converted to TP(Transport) stream and then finally broadcasted through the DMB devices. As parameter input, the advertisement producer can control adjustment parameter and precision parameter. Advertisement producer not only can control these two parameters to create Hot-spot adaptive to the properties of input video stream,

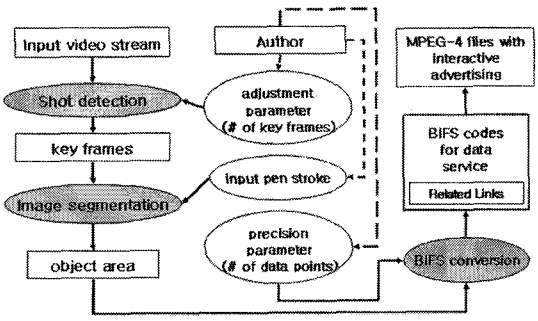
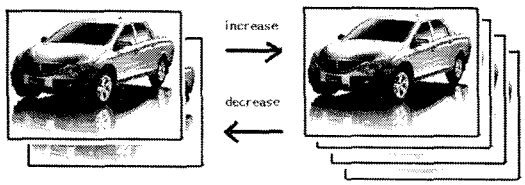


Fig. 3. Hot-spot Generator system overview

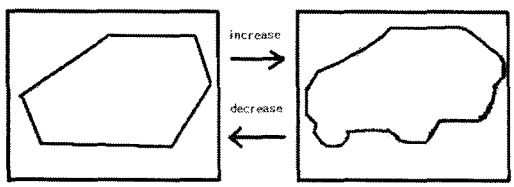
also should control these parameters if the data rate limitation is not met. As an example, Increasing adjustment parameter for a video stream with large movements and frequent scene change like a sports game will make the intervals smaller between the key frames, so the generated Hot-spot and the object in the video stream would match more seamlessly(see Figure 4(a)). On the other hand, for video streams with relatively large area of actor/actress or product on screen but low movement such as TV dramas, increasing precision parameter will increase the points of the polygon that represents the target object, allowing the audience to select complex Hot-spots of their interest with more accuracy(see Figure 4(b)). Conversely, in case of the total data size is bigger than the data rate limitation of broadcasting, the problem can be solved by reducing those two parameters. So it will make the size of the additional information smaller.

Shot detection

This research forms shot indexes by detecting hard cut and gradual transition from the input video stream to extract unit shots. For faster interaction of the system, shot detection algorithm processes I frames only leaving B, P frames intact. Checking whether target objects to be advertised is in the indexed unit shot is done with the unaided eyes of the advertisement producer. Using automatic extraction of moving object and foreground images in the video stream is out of the scope of this research. As the next step, the advertisement producer extracts key frames from selected unit shots. Adjustment parameter is referenced to decide the number of key frames in this step.



(a) adjustment parameter
- change in the key frame intervals



(b) precision parameter
- change in the number of data points

Fig. 4. Application of the adaptive parameter and precision parameter

Image segmentation

Image segmentation on selected key frames is applied for extracting Hot-spot areas. Figure 5 shows the flow chart of the suggested method.

First of all, key frames are automatically segmented into basic segments by Mean Shift algorithm[12]. The advertisement producer draws a simple closed curve around the target object of advertisement on the key frame with pen-based interface. Basic segments that are inside or overlap the simple closed curve are selected as the target advertisement object area. All the other segments are automatically clipped out.

Figure 6(a) shows the key frame with automatically segmented into basic segments. After the advertisement producer draws the simple closed

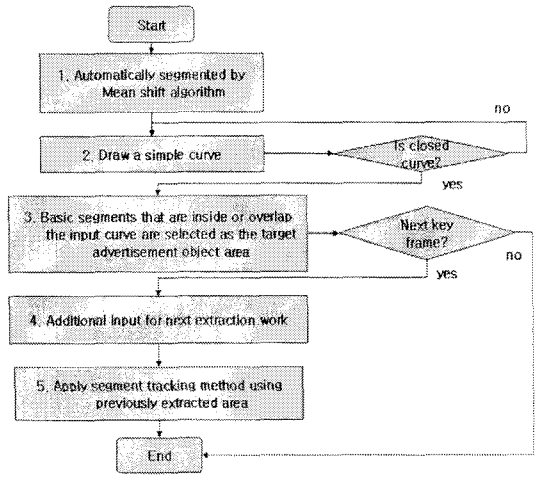
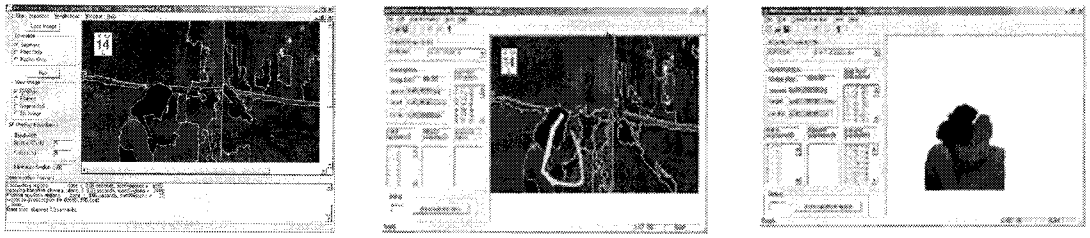


Fig. 5. Suggested object area extraction technique



(a) (b) (c)

Fig. 6. User's pen input and the result of the object extraction

curve as the information of the target object area for advertisement as in Figure 6(b), then Figure 6(c) shows the final result of the extracted target object.

This semi-automatic segmentation method holds advertisement producer's intention better than full-automatic method and also extracts aimed Hot-spot area much quickly using pen-based interface. Moreover, using extracted area of prior the key frame, the advertisement producer's effort is minimized to add only the part changed in the continuous key frames.

BIFS conversion

Extracted object area is represented as 2D coordinates sets, and described as Hot-spot by adding the associated links. Figure 7 shows a structure and an example of the key frame list processed on a unit shot contains advertising object. Total number of key frames is 41, the number of the first key frame is 0, and the second one is 50. If target advertising object area is extracted for key frame

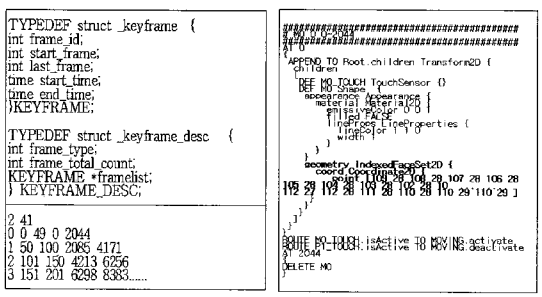


Fig. 7. Polygonal representation of object area and BIFS code

number 0, then a coordinates set value needs to be inserted into the part of bold text in Figure 7. geometry IndexedFaceSet2D element represent the area of Hot-spot and BIFS recognizes the first value as the x-axis coordinate, second value as the y-axis coordinate in the 2D coordinates set, and processes all continuing values in this manner[13].

After BIFS conversion, BIFS codes of each key frame are combined into one MPEG-4 BIFS data file. This data file is multiplexed with audio and video data to create a new mp4 file. This newly created mp4 file is transferred to DMB device via digital broadcasting transfer device, then whether the intended advertisement area is working as the Hot-spot can be verified.

3.3 Implementation of the system and Working study

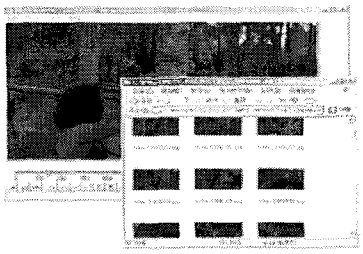
This section describes the implementation of the suggested Hot-spot generator and how it works. Microsoft Visual Studio 2005 on Windows XP was used to develop the Hot-spot generator. The system has largely 3 executable modules. The first one is ShotDetector which detects shots from the input video stream and decides the key frames. The second module is ContourTracer that extracts object area from the key frames. The last one, BTMaker, is in charge of the DMB BIFS data conversion. For validating the concordance of result Hot-spots and the video source, we used the Osmo4 Player which is provided by the open source project-GPAC[14]. The result mp4 file was played and tested on the

Osmo4 Player.

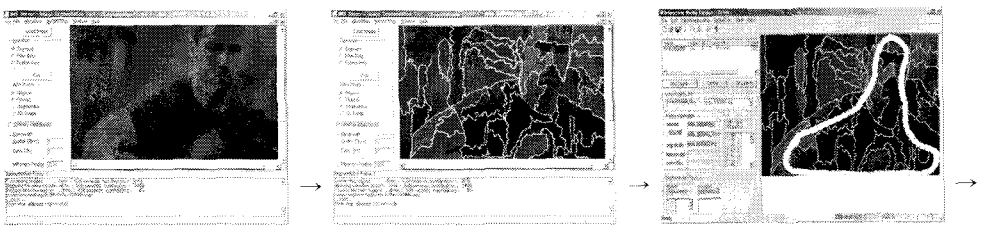
Following Figure 8(a) demonstrates video segmentation on input video content using ShotDetector executable. In Figure 8(b), object areas are extracted from selected key frames and polygonal regions are calculated using ContourTracer executable. After that, extracted object areas are converted into BIFS code by BTMaker executable. The number of key frames to be generated and the size of the additional information automatically reference the two parameter inputs. In Figure 8(c), the link is added to the final advertisement area before converted to the DMB BIFS code by BTMaker. Figure 8(d) shows the user's validation of interactive advertisement using Osmo4 Player. The result of the experiment on DMB drama video stream with less movement is on the left, and the result of a sport game video stream with lots of movement is on the right side.

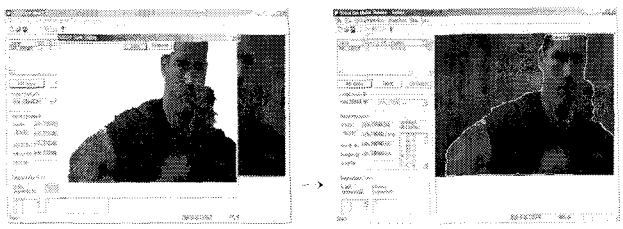
4. EXPERIMENT AND THE RESULT

This section introduces the environment for the

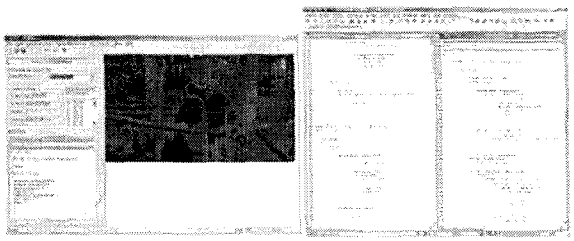


(a) Unit shot segmentation by ShotDetector





(b) Advertising object edge extraction by ContourTracer



(c) Generated BIFS text code

(d) Test of Hot-spot selection

Fig. 8. Interactive advertisement insertion and broadcasting in DMB

Table 2. Plans for experimentation

(a) Plan 1

Goal	Measure the precision of object extraction for the given time
Content	1. Drama program (less keyframes) 2. Sports program (many keyframes)
Method	Calculate the error rate between tester extracted area and accurately extracted area

(b) Plan 2

Goal	Accuracy of the content-based Hot-spot
Content	Video contents with phases that have different key frame extract intervals
Method	Percentage of the successful/failed cases of users' Hot-spot area selection using phases.

experiment and analyzes the result of the experiment. Table 2 shows plans for the experimentation.

In the first experiment, testers are classified into two groups: group A manually extracts the target object using a drawing tool, and group B extracts the target object using our system. The Drama program used as the first experiment content was set 16 key frames, and within the key frames testers are required to extract the face of the actor. For the second content, a basketball game program was used, and the testers are asked to extract the body of the player within 30 key frames. The testers were given 3 minutes to perform object extraction using either the drawing tool or our system. Following Equation (1) and (2) are used to estimate the error rate of the object that the test subjects generated.

$$Eu = (\frac{Ru}{Ra}) \times 100,\ Eo = (\frac{Ro}{Ra}) \times 100 \quad (1)$$

$$Match = (\frac{Ra - Ru}{Ra}) \times 100,\ Error = Eu + Eo \quad (2)$$

Ra is the precisely cropped area of target object in the selected key frame by using the graphics software Adobe Photoshop CS3 Extended. Ru is the un-extracted area of the target object area, and Ro is incorrectly extracted area. For every area, pixel numbers were counted and represented as Match value, and the precision is estimated from the calculated error rate.

If all the objects are correctly extracted in the given experiment, then the error rate would be close to 0. Figure 9 shows the result of the change in the given time for object extraction. For the subject group A, who used graphics tool for extraction, error rate increased drastically as the number of frames increased or given time reduced. On the other hand, the error rate of the subject group B was not susceptible to the change of the content or the given time. In other words, the suggested technique gets more useful for the object extraction in the programs that have many scene changes or very short key frame intervals.

In the experiment 2, the subjects were tested for the success rate of Hot-spot selection with the varying key frame intervals in a drama and a sport

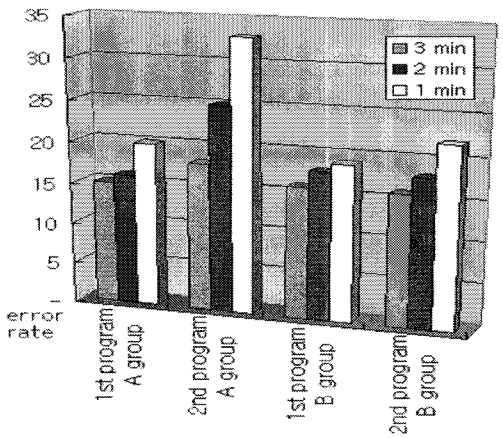


Fig. 9. The result of the precision test of object extraction

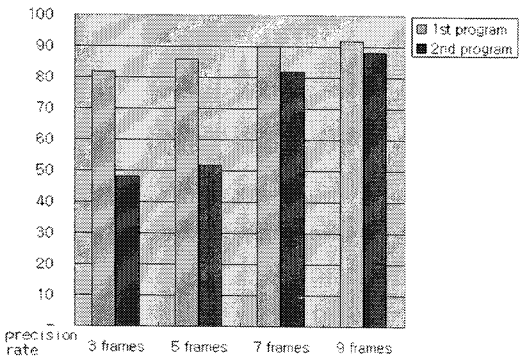


Fig. 10. The result of the precision test of Hot-spot selection

game. Experiment contents were saved with 3, 5, 7, and 9 key frames per 10 seconds, and 10 subjects were tested 10 rounds per each testers. As Figure 10 shows the result, the drama program was not susceptible to the change in the number of key frames, but testers found it hard to click the target advertisement in the sport game with lots of movement. Therefore, for the video content with lots of movement such as sport game program, it is advisable to use the suggested Hot-spot generator to detect a relatively simple form of Hot-spot area for as many numbers of key frames and then broadcast it as a interactive advertisement.

5. CONCLUSION AND FUTURE WORKS

Advancement in digital data broadcasting technology enabled advertisers to insert detailed information of the product or associated advertisement link into the content of the broadcast program as an interactive advertisement. With the interactive advertisement, the audience can focus on the content of the program since this type of advertisement is activated only by the audience's active behavior. The advertisers can use this as a new channel of selling their product, or as a marketing tool for more direct consumer analysis. This research suggests a system for easy generation of dynamic Hot-spot that follows the target object of advertisement.

The proposed system is evaluated using GPAC MPEG-4 player, and has proved it's usefulness using two DMB broadcasting programs that have different video environmental features.

Future works include the technique for generation of satisfactory Hot-spot without the advertisement producer's manual parameter input by automatically detecting features of the input video stream. Intelligent system for handling the data rate of additional information as the number of Hot-spots grows is also in the list of the problem to solve. As for the future study, it is desirable to include the target data rate as a new parameter, and integrate the parameter with existing ones in the system.

REFERENCES

[1] Y.J.Lee, "Audience Attitudes towards Product Placement in TV Drama," *Journal Of Communication Science*, Vol.6, No.2, pp. 323-355, 2006.

[2] J.G.Kim, S.J.Choi, and S.H.Lee, "Research for Application of Interactive Data Broadcasting Service in DMB," *Korea Society Broadcast Engineers Magazine*, Vol.11, No.4, pp. 104-117, Dec, 2006.

[3] S.H.Lee and J.Y.Tak, "An Experimental Study on the Effectiveness of Interactive TV Advertising," *KOBACO*, No.71, pp. 177-196, 2006.

[4] H.S.Yoon, "U-Cast based interactive broadcasting," *TELECOM*, Vol.23, No.2, pp. 67-73, 2007.

[5] ISO/IEC 14496-1, Information technology - Generic coding of audio-visual objects, Part 1: Systems, International Standard, 2001.

[6] ISO/IEC 14496-11, Information technology - Coding of audio-visual objects, Part 11: Scene description and application engine, International Standard, 2005.

[7] ISO/IEC 14772-1, Information technology -

Computer graphics and image processing - The Virtual Reality Modeling Language - Part 1: Functional specification and UTF-8 encoding, International Standard, 1998.

[8] Sang-Hun Kim, Chun-Sub Kwak, and Man-Sik Kim, "The Development of Real-time Video Associated Data Service System for T-DMB," *Journal of Broadcast Engineering*, Vol.10, No.4, pp. 474-487, Dec, 2005.

[9] J.S.Choi, I.S.Lee, and Y.M.Kim, "Wibro-DMB integration Service Technique," *TTA Journal*, No.105, pp. 116-123, June, 2006.

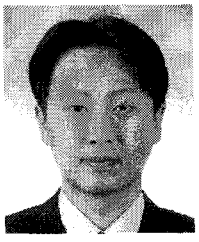
[10] Zhang and Yu-Jin, *Advances in Image And Video Segmentation*, IRM Press, Pennsylvania, USA, May, 2006.

[11] John C. Russ, *The Image Processing Handbook*, CRC Press, 2006.

[12] D. Comaniciu and P.Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Patterns Analysis and Machine Intelligence*, Vol.24, No.5, pp. 603-619, May, 2002.

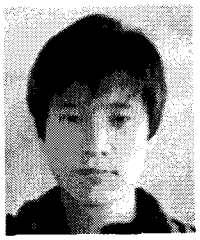
[13] Shin Kong, Tae-Jin Park, Soon-Bum Lim, and Yoon-Chul Choy, "Development of Authoring Tool for Mobile DMB Contents and Templates with 3D Objects," *KIISE Fall Conferenec*, Vol.34, No.2(B), pp. 178-183, Oct, 2007.

[14] GPAC, "http://gpac.sourceforge.net/".



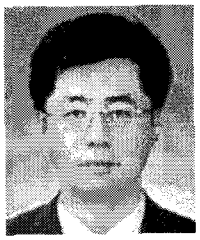
Tae-Jin Park

received the BS and MS degrees in Computer Science from Yonsei University, Seoul, Korea, in 1993 and 1995, respectively. Previously, he worked as a research engineer at LG Digital Media Laboratory and LG DTV Laboratory in 2000 and 2005. Currently, he is working towards the PhD degrees in the Department of Computer Science at Yonsei University. His research interests include DTV system middle-ware, DTV data broadcasting, DTV information interface, multimedia data presentation and DMB.



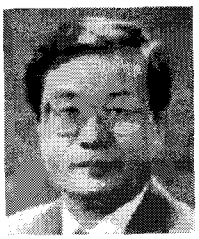
Se-Hyun Park

received the BS degree in Business Management from Korea University, Seoul, Korea in 2006. He is a research member of Multimedia/Graphics Lab in Yonsei University. His research interests include multi-touch interface, video visualization, and video editing. He is awarded outstanding paper prizes from the Korea Multimedia Society in 2008.



Soon-Bum Lim

received his MS and PhD degrees in Computer Science from Korea Advanced Institute of Science and Technology, Daejun, Korea, in 1983 and 1992, respectively. He was an assistant professor at Kunkook University from 1992 to 1997. He has been with the Department of Multimedia Science at Sookmyung Women's University, Seoul, Korea, where he is currently a professor. His research interests include computer graphics, multimedia application, electronic publication, DMB data broadcasting and mobile interface.



Yoon-Chul Choy

received his MS degree from the University of Pittsburgh in 1975 and the MS and PhD degrees in IE & OR from the University of California, Berkeley, U.S.A, in 1976 and 1979, respectively. He was at Lockheed and Rockwell as an international researcher in 1979 and 1982. He was a visiting professor at the University of Massachusetts, U.S.A in 1990 and 1991 and Keio University, Japan, from 2002 and 2003. He has been with the Department of Computer Science at Yonsei University, Seoul, Korea, where he is currently a professor since 1984. His research interests include computer graphics, multimedia, non-photo realistic rendering and sketch-based interface.