

Proposal of Value-added Service based on Real-time Automatic Content Recognition

Donghwan SHIN*, Jonguk. CHOI*

**MarkAny Inc., 10F Ssanglim Bldg., 286, Toegye-ro, Jung-gu, Seoul, Korea*

dhshin@markany.com, juchoi@marknay.com

Abstract— ACR (Automatic Contents Recognition) technology is a technology for automatically recognizing contents, and mainly recognizes contents using audio signal processing technology. The reason why ACR technology is necessary is to recognize the content currently being broadcast and provide an additional service suitable for it. Because of this necessity, internationally, ACR technology is selected as a standard in ATSC 3.0. In Korea, IBB (Integrated Broadcast and Broadband) service is included as domestic UHD broadcasting technology standard.

In this paper, we describe the standard progress of ACR technology as domestic IBB standard and discuss why ACR technology is needed in actual broadcasting environment and future development direction. In particular, there is a limitation that the standardized part can only be provided by UHD TV. UHD TV Terrestrial broadcasting started in May 2017 based on ATSC 3.0 for the first time in the world. However, considering the domestic situation that more than 90% of users watch using set-top boxes, the number of households receiving UHD broadcasting directly is expected to be less than 10%. In addition, UHD broadcasting can only use UHD TV purchased in 2017 to use additional information service using ACR technology. DTV broadcasting will continue until UHD 100% broadcasting will be available by 2027 because the average TV replacement cycle is about 10 years on average. Accordingly, the time to replace the TV with UHD TV is expected to take more than 10 years. In this paper, the technology and direction to solve these problems are proposed.

Keywords— Audio, ACR, 2nd Screen Service, Watermark, Fingerprint

I. INTRODUCTION

ACR(Automatic Content Recognition) is a technology for automatically recognizing contents, and is widely used for the purpose of recognizing contents currently broadcasted in a broadcasting environment and providing additional information service which is corresponding to the contents.

ACR has mainly used two technologies to the present. Watermarking technology and fingerprinting technology have been mainly used to protect contents since the 2000s.

In particular, fingerprinting technology has been used for recognizing music since the early 2000s.

In the early days, Shazam [1], Sound Hound [2], and Grace Note [3] were known as services that let users know what music they were listening to.

The move to use ACR technology in the broadcast environment began around 2010. Digimarc has attempted a second screen service by inserting an audio watermark into the broadcast signal [4]. Thereafter, attempts to recognize and

service content using fingerprinting technology have been activated.

In Korea, Enswers and Markany have related technologies and are supplying technology to broadcasting stations. Overseas, companies such as Audible Magic, Flingo, and Civolution [5] are putting ACR technology in the foreground, and some of them are pursuing strategic alliances with large US broadcasters.

In Chapter 2, we will study the standard of ACR technology required in UHD broadcasting standard of Korea which is the world first UHD broadcasting in 2017 and discuss its practical operation plan.

In Chapter 3, we present examples of using Korean broadcasting stations using fingerprinting technology.

In Section 4, we introduce a method of providing value-added information service using ACR technology in ambient environment.

II. 1ST SCREEN ACR

Attempts to employ watermarks as ACR technology are steadily being pursued in broadcast standards. A representative example is ACR technology as a standard in ATSC 3.0, the North American terrestrial standard. In Korea, ATSC3.0 is based on Korean UHDTV broadcasting standard, and ACR technology is also included here.

A. Watermark ACR

The ACR standard of UHDTV in Korea was established a little differently from ATSC 3.0. In ATSC 3.0, ACR technology was selected by a limited number of companies. In Korea, however, the standard for watermark technology itself was not specified but standardization was made to expand service opportunities to a large number of companies.

IBB Standard is a standard for using broadcasting network and Broadband network to provide supplementary information service in domestic UHDTV broadcasting. The IBB Standard cites ATSC 3.0 in North America and the HbbTV 2.0 standard in Europe.

Watermark Payload is total 50bits per 1.5 seconds. Each watermark payload structure is shown in table 1.

TABLE 1. WATERMARK PAYLOAD DESCRIPTION

Item	Description	Bits	Comments
Domain-type	1: Large domain 0: Small domain	1	
Server-field (channel-field)	Classification of broadcast channel information or program ID	23	One ID can be assigned per program
Interval-field (time-field)	Counter value(cycle 1.5 sec)	25	0~33,554,431 (50,331,648sec = 388 days)
Query-flag	Dynamic event	1	

The ACR service structure using the watermark according to the IBB standard is shown in figure1.

Watermark insertion is done at the station. The watermark is detected by the WM Client of Smart TV, the IBB Receiver. And Smart TV recognizes the broadcast content using the detected watermark information.

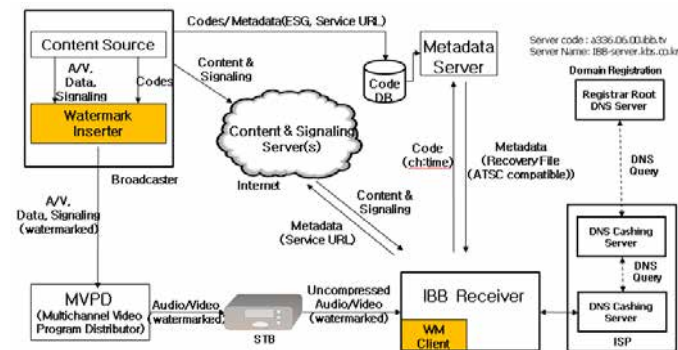


Figure 1. Watermark ACR Service Structure

The detailed operation flow is shown in figure 2.

- 1) Smart TV detects watermark. Using the detected watermark code, the metadata server address is retrieved by DNS query.
- 2) The broadcast channel information and the time code extracted from the watermark detection code are transmitted to the metadata server.
- 3) In metadata server, ESG information and time information can be used to know broadcasting program being broadcast at extraction time. Based on this, the program code and corresponding recovery file are provided to IBB Receiver (Smart TV). The Recovery File contains a list of services to be performed over time for the program.
- 4) The IBB receiver analyzes the recovery file and deliver the service URL to the Contents & Signal server in the internet network.
- 5) Contents and Signal server provides contents and signals to the IBB Receiver.

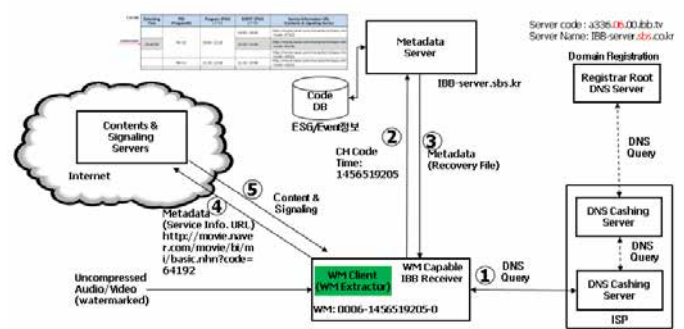


Figure 2. Watermark Interactive Processing Flow

B. Fingerprint ACR

Another technique for identifying content is Fingerprint technology. Fig. 3 shows the Fingerprint ACR service structure.

Before contents are broadcasted from a broadcasting station, the FP (fingerprint) is extracted and registered in a FP (fingerprint) server to construct an FP server. In the Smart TV, which is the IBB receiver, the FP Client extracts the FP for the broadcast content and inquires the FP server to recognize the broadcast content.

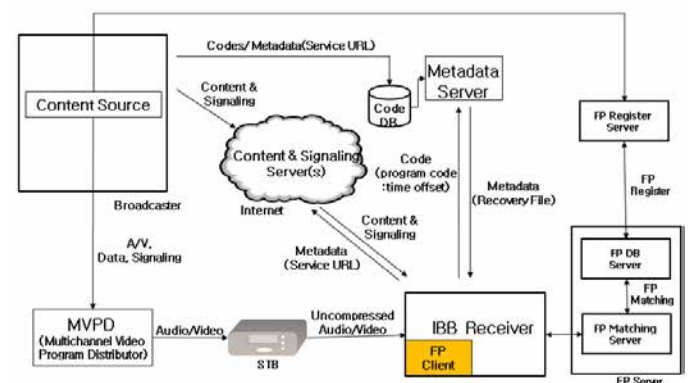


Figure 3. Fingerprint ACR Service Structure

The detailed fingerprint ACR operation flow is shown in figure 4.

- 1) Fingerprint query to retrieve program code and time offset value.
- 2) Transfer the program code and time offset value to the metadata server.
- 3) The metadata server transfers the recovery file corresponding to the program code to the IBB receiver.
- 4) The IBB receiver analyzes the input recovery file and transmits the service information corresponding to the time offset value to the content signaling server using the Internet network
- 5) In the content signaling server, the content and signal corresponding to the service information are transmitted to the IBB receiver and the additional information is displayed on the smart TV.

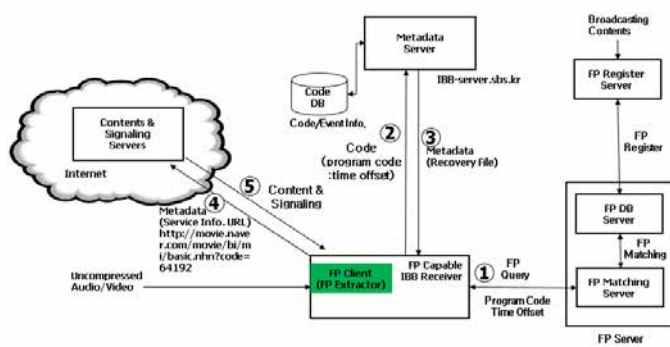


Figure 4. Fingerprint Interactive Processing Flow

III. 2ND SCREEN ACR

In order to perform additional information service on the second screen equipment (smart terminal) according to the IBB standard described in the above, the TV can be provided only to a home having a TV manufactured in 2017.

It is expected that it will take more than 10 years to replace UHD TV because the TV replacement cycle is 10 years and the existing digital TV broadcasting service and UHD TV broadcasting service are provided concurrently by 2027. Until then, normal second-screen service becomes a very difficult environment.

In order to overcome this difficult situation and to implement the second screen service in the current situation, we will explain how to provide second screen service of UHD TV broadcasting by using audio watermark and fingerprint technology recognizable in ambient environment.

Figure 5 shows the structure that enables CS (Companion Screen) service using watermark technology that can be recognized in ambient environment.

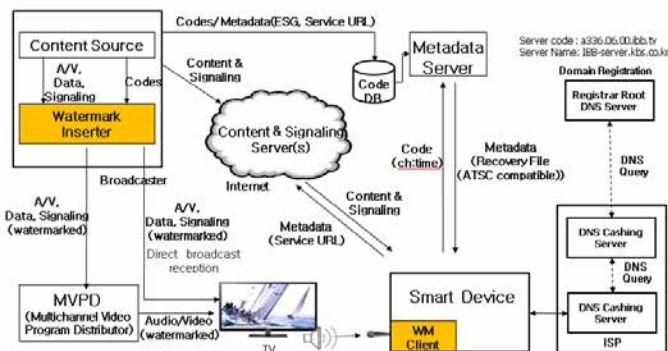


Figure 5. Watermark 2nd Screen ACR Structure

Figure 6 shows a structure for recognizing contents in ambient environment using fingerprint and providing CS service for smart terminal.

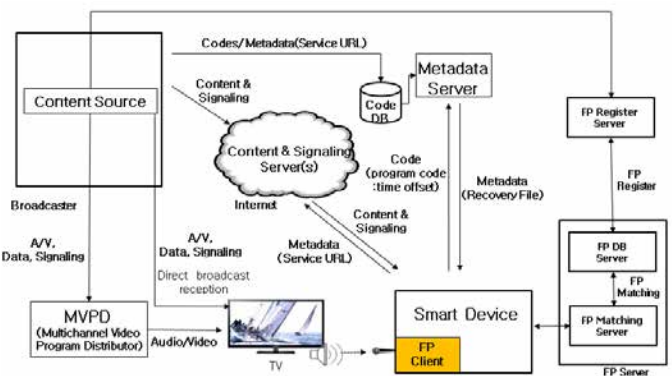


Figure 6. Fingerprint 2nd Screen ACR Structure

IV. CONCLUSIONS

This paper describes the implementation of CS service (Second Screen Service) in UHD broadcasting based on ATSC3.0 for the first time in the world in Korea. In reality, if the reception rate of terrestrial broadcasting is low as 5% and there are not many UHD TV supply assumptions, the spread of CS service is expected to be difficult.

As a solution to this problem, we proposed a method for CS service in digital TV households instead of UHD TVs using watermarking technology and fingerprint technology which can be recognized in ambient environment.

ACKNOWLEDGMENT

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No. 2015-0-00219, Development of smart broadcast service platform based on semantic cluster to build an open-media ecosystem)

REFERENCES

- [1] A. Wang. “The Shazam music recognition service.” *Communications of the ACM*, 49(8):44–48, Aug 2006.
- [2] <https://www.soundhound.com/soundhound>
- [3] J. Haitsma, T. Kalker, and J. Oostveen, “An efficient database search strategy for audio fingerprinting.” In *IEEE Workshop on Multimedia Signal Processing*, pp. 178–181, St. Thomas, USVI, 2002.
- [4] “Nielsen and Digimarc’s Media-Sync Powers Innovative ‘Greys Anatomy’ Enhanced Television Viewing App.,” Feb., 2011.
- [5] Civation (2012). ACR Creating content-aware ecosystems, Civation White Paper, Sep. 2012.



machine learning.



Donghwan Shin received the MS degree and Ph.D. in electronics engineering from University of Seoul, Korea. From 1992 to 1994, he was a member of the LG Electronics Inc. He worked as a senior researcher in the Korea Sports Science Institute from 1996 to 2000. He has been currently a chief manager of MarkAny Inc. from 2000. His research interests are in the areas of copyright protection, fingerprinting, watermarking and

Jonguk Choi received the Ph.D. in artificial intelligence from University of South Carolina, USA in 1989. From 1989 to 1992, he was the director of Artificial Intelligence Division. He has been currently a professor of Sangmyung Univ. from 1992. His main interests are in the areas of digital watermarking, DRM and artificial intelligence