

# 방송 영상에서 분할된 객체의 정보를 이용한 대화형 방송 시스템 구성 방법

\*박형준

한국디지털케이블연구원

\*hjpark@klabs.re.kr

## A Configuration Method for the Interactive Broadcasting System Using Information of Segmented Objects on the Image

\*Park Hyung-Jun

Korea Digital Cable Laboratories

### 요약

본 논문에서는 사용자가 뉴스나 스포츠 중계 등의 라이브 방송을 제외한 녹화방송 프로그램 및 VoD콘텐츠를 시청하는 중에 화면 상의 상품에 대한 관심이 유발되었을 때, 객체의 좌표정보를 통해 사용자의 입력을 받아 객체의 부가정보를 실시간 검색하고 부가정보 데이터를 생성하는 방법을 통해 사용자에게 해당 객체의 부가 정보를 제공하는 방법을 제안한다.

### 1. 서론

방송이 양방향의 디지털 방송으로 진화되고 대용량의 방송 부가 데이터 통신이 가능해지면서 다양한 대화형 방송 기술의 개발 및 적용이 가능하게 되었다.

이로 인해 방송플랫폼 사업자와 스마트 셋탑박스나 스마트 TV같은 방송 단말 및 세컨드 스크린 단말용 어플리케이션 개발자 등은 대화형 방송 기술의 개발을 통해 세롭고 다양한 서비스의 창출할 수 있게 되었다.

본 논문에서는 사용자가 뉴스나 스포츠 중계 등의 라이브 방송을 제외한 녹화방송 프로그램 및 VoD콘텐츠를 시청하는 중에 화면 상의 상품에 대한 관심이 유발되었을 때, 사용자에게 해당 객체에 대한 부가 정보를 전송해 보여주는 대화형 방송 서비스 제공을 위하여 방송 영상상의 분할된 각 객체들의 부가 정보를 제공하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2 절에서는 대화형 방송에 관련된 주요 연구를 살펴본 후, 3절에서는 분할된 객체의 정보를 이용한 대화형 방송 방법을 제안한다. 4 절에서는 본 논문에서 제안하는 방법을 적용한 시스템 구성을 제시한다. 5절에서는 본 논문에 대한 결론을 맺는다.

### 2. 관련 연구

객체 정보를 이용한 대화형 방송은 다양한 대화형 방송 중 한 방식으로 대화형 방송은 방송 전송 채널을 통해 데이터 캐러셀(data carousel)내에 데이터 방송에 필요한 데이터가 포함되어 가입자 단말에 전송되어 사용자에게 제공되는 지역 대화형 서비스(local interactive service)와 방송프로그램에 관련되거나 관련되지 않은 더 많은 정보를 얻기 위해 원격지의 서버에 접속하는 원격 대화형 서비스

(remote interactive service)로 분류될 수 있다[1]. 지역 대화형 서비스는 사용자가 요청하는 다양한 정보를 미리 수신기에 전송 받기 어려워 원격 대화형 서비스에 대한 관심이 높아지고 있다[2].

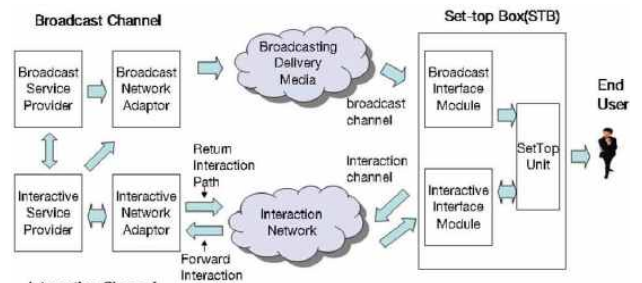


그림1. 대화형 방송 시스템 참조 모델

분할된 객체의 정보를 이용하는 대화형 방송 기술과 관련해 객체 분할을 통해 객체를 저장하는 방법[3], 영상분할을 이용한 영상 검색 시스템[4], 동영상 객체분할 시스템의 구현[5], 시청자가 선택한 객체의 부가정보를 제공하는 대화형 서비스 구현[2]등의 다양한 연구들이 진행되었다.

또한 방송프로그램의 인덱스 정보를 검색하여 이를 UI에 맞게 편집 생성하여 세컨드 스크린 단말에 부가정보를 제공하는 방법[6] 등 부가정보 제공 방법에 대한 연구 개발 및 상용화가 진행되고 있고 객체의 색상정보와 차영상을 이용한 동영상 내 객체 추적 기법[7]등의 연구가 진행되었다.

대화형 방송을 위해서는 기본적으로 메타데이터 관리기술이 필요하다.

MPEG-7은 메타데이터 생성과 인덱싱에 대한 기술의 필요성이 제기됨에 따라 멀티미디어 콘텐츠에 대한 구조 정보와 의미 정보를 기

술할 수 있도록 한 국제 표준이다.

그중 MPEG-7 Part5 MDS(Multimedia Description Schemes)는 멀티미디어 콘텐츠의 색인, 검색, 필터링 등을 용이하게 할 수 있도록 정보표현을 구조화하는데 이용되고 있다[8]. MPEG-7 MDS 도구는 콘텐츠 관리에 필요한 정보를 기술하는 콘텐츠 관리(content management)와 콘텐츠의 내용을 기술하는 콘텐츠 기술(content description) 등으로 구성된다. 이 중 콘텐츠 서술은 내용 기반 검색을 위한 동영상 기술 방법으로 동영상 콘텐츠를 시공간적으로 분할하는 Segment DS와 동영상의 내용을 기술하는 Sementic DS로 구성된다. Segment DS는 오디오, 비디오를 시계열상 혹은 공간상에서 분할하고 이러한 Segment들을 Multimedia Segment DS로 합성하고 추상화(abstraction)하여 트리 형태로 구조화하여 기술한다. MPEG-7은 객체 별 정보 기술을 위한 객체분할 방법은 표준화 대상으로 정해 놓고 있지 않으며 관련 응용 개발자의 몫으로 남겨두고 있다[9].

TVAnytime표준은 콘텐츠 검색서비스 제공을 위한 표준으로 Phase1(TVA-1)과 Phase2(TVA-2)의 두 단계로 나누어진다.

TVA-1은 주 방송 프로그램이 단 방향의 방송채널로 전송되고 양 방향 네트워크를 통해서 추가적인 메타데이터를 획득할 수 있는 방송 환경에서, PDR을 중심으로 메타데이터를 이용한 AV 데이터의 탐색(search)-선택(selection)-획득(acquisition)-소비(consumption)의 일련의 콘텐츠 서비스를 가능하게 하는 것을 목표로 한다[10]. TVA-2는 TVA-1을 확장하여 홈네트워크 환경에서 단말간의 콘텐츠 공유(sharing) 및 재분배, 다양한 사용자 환경에 맞는 콘텐츠를 제공하는 타겟팅(targeting) 등의 서비스를 제공하며, 콘텐츠도 TVA-1의AV 뿐만 아니라 다양한 형태의 콘텐츠를 수용한다[11]. TVA-2에서는 또한 패키지(package) 형태의 콘텐츠를 정의한다.

### 3. 분할된 객체의 정보를 이용한 대화형 방송

기존에는 원격 대화형 서비스 방식으로 방송 영상상의 분할된 객체 정보를 이용해 대화형 방송 서비스를 구현하는 연구가 진행 된 적이 있지만 본 논문에서 제안하는 방법처럼 실시간 검색을 통해 부가정보를 생성하고 지역 대화형 서비스와 원격 대화형 서비스를 융합하여 구성하는 방향으로 진행 된 예는 없었다.

제안하는 분할된 객체의 정보를 이용한 대화형 방송 시스템은 방송데이터의 전송과 양방향 데이터 통신이 분리된 환경에서 효과적으로 대화형 방송을 제공하는데 초점을 두고 고안되었다.

이 방법을 간단히 요약하면 프로그램의 인덱스 데이터, 프레임의 인덱스데이터, 프로그램상에서 선정된 객체들의 인덱스 데이터와 각 객체의 화면상 좌표정보(또는 edge정보)를 방송 데이터에 포함시키고 객체의 좌표정보를 토대로 스마트 셋탑박스나 스마트TV와 같은 사용자 방송 단말기 상의 어플리케이션이 객체의 위치를 추적, 객체의 좌표 정보를 갱신하고 이를 이용해 사용자가 데이터 통신 채널을 통해 객체의 정보를 요청하면 단말상의 어플리케이션이 정보검색을 통해 객체의 부가정보를 수신 받아 부가정보를 UI에 맞게 편집하여 방송 화면상에 표시해 주는 방법이다.

분할된 객체 정보를 방송데이터에 포함시키는 작업은 콘텐츠 제작 상에서 이루어진다. 콘텐츠 제작자는 부가 정보를 제공할 객체들을

선정하고 이 객체들이 출연하는 첫 번째 프레임을 표시하고 객체들의 인덱스 정보와 기준 좌표정보를 방송데이터에 포함시킨다.

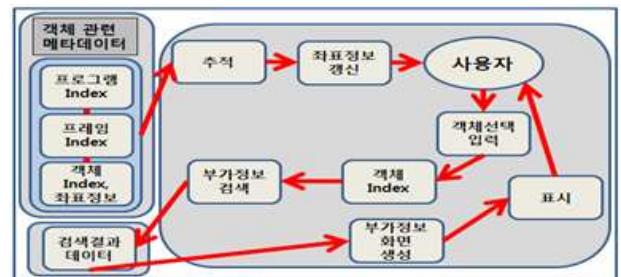


그림2. 제안하는 방법의 operation flow

사용자 단말기 상의 어플리케이션은 객체의 최초 출연 프레임 인덱스 정보와 기준좌표정보를 이용해 객체를 추적하여 객체의 좌표정보를 갱신한다. 객체추적은 비디오 스트림을 프레임 단위로 분할하여 기준좌표를 통해 객체의 영역을 지정하고 스트림 진행에 따른 객체의 위치변화를 추적하여 각 객체의 위치, 시각, 모양과 해당 프레임 간의 동기를 맞추는 기술이 필요하다.[2]

객체에 대한 부가정보는 방송플랫폼 사업자가 제작하여 제공하는 방향으로 진행될 수 있고 콘텐츠 제작자가 제작하여 콘텐츠와 함께 제공하는 방향으로 진행될 수도 있다. 또한 사전 제작을 하지 않고 실시간 검색을 통해 획득하는 방향으로 진행될 수도 있는데 본 논문에서는 이와 같이 실시간으로 홈쇼핑 및 인터넷 쇼핑정보, youtube, SNS 등을 해당 검색API를 활용하여 사용자 단말기의 어플리케이션을 통해 검색해 검색된 정보를 편집 생성하여 보여주는 방법으로 구상하였다. 부가정보의 구성방법에 따라 새로운 광고 수익이 발생할 수 있을 것이고 구매루트 제공을 통한 수수료 수익의 발생도 예상할 수 있다.

제안하는 방법을 사용자의 객체 입력 시점에서 보면

첫 번째로 객체의 좌표데이터가 존재한다. 이 좌표데이터는 방송 데이터 상의 프로그램인덱스 데이터, 객체가 최초로 출연하는 프레임 상의 기준 좌표 데이터를 추적하여 갱신된다. 이 좌표에 커서를 위치시켜 부가정보를 요청하면 해당 좌표의 객체에 대한 인덱스 데이터를 통해 해당 객체에 맞는 부가데이터가 검색되고 이 데이터가 수신되어 구현된 인터페이스에 맞게 편집 생성돼 시청화면상에 표시되는 구조로 데이터가 구성된다.

자동 영상 분할 기술과 영상 인식기술을 이용하면 플랫폼 사업자가 콘텐츠 제작단계와 상관 없이 분할된 객체에 대한 부가정보를 제공하는 기능을 제공할 수 있고 방송과 직접적 관련이 없는 주제가 사용자 방송 단말기 상이나 세컨드 스크린 단말 상의 어플리케이션 개발을 통해 해당 기능을 제공하는 것도 가능할 것이다. 또한 객체의 좌표 데이터는 사용자가 화면상의 커서를 통해 객체를 선택하여 입력하기 위해 필요하므로 음성입력을 통해 입력이 가능할 시 필요하지 않을 수 있으나 음성입력이 불가능한 상황이 발생 가능하므로 분할된 객체의 좌표데이터는 포함되어야 한다.

방송 전송과 양방향 데이터의 송수신이 분리된 환경에서의 객체의 부가정보제공은 동작속도가 늦어지는 단점을 야기할 수 있으나 방송전송과 데이터 송수신 채널의 분리로 인해 부가데이터의 송신이 방송의 전송에 영향을 주지 않으므로 향후 방송데이터의 용량이 커져도

방송의 QoS를 유지하면서 대용량의 데이터 송신이 가능하고 PIP와 같은 기능을 제공해 상품 광고 및 상품 리뷰 영상 등의 확인이 가능할 것으로 예상된다.

#### 4. 시스템 구성

제안 하는 방법의 시스템은 기존의 시스템 구성에 필요 장비 및 소프트웨어가 추가되는 형태로 구성된다. 제안하는 방법의 서비스를 제공하는 시스템을 콘텐츠 제작 시스템부분과 양방향 방송 시스템부분으로 나누어 볼 수 있다.

콘텐츠 제작 시스템부분에서는 메타데이터 저작 툴로 생성된 AV 콘텐츠를 멀티플렉서를 통해 하나의 콘텐츠로 생성한다.

양방향 방송 시스템에는 스마트 셋톱박스나 스마트TV의 어플리케이션 상에서 방송 데이터로 수신된 객체의 프레임 인덱스와 기준좌표를 이용하여 객체를 추적하는 소프트웨어 모듈과 검색기능을 통해 부가정보데이터를 제공하는 검색API서버와 통신 하고 부가정보를 수신 받아 수신 받은 데이터를 UI에 맞춰 편집 생성하여 화면상에 표시 하는 소프트웨어 모듈이 장착되는 형태이다.

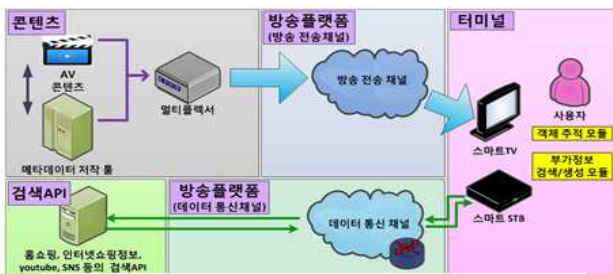


그림3. 제안하는 방법을 적용한 실시간 방송 시스템 구성

#### 5. 결론

양방향 통신이 가능한 디지털 방송 기술이 발전하면서 다양한 방식의 대화형방송 서비스가 가능해졌고 대화형 방송 서비스에 대한 사용자들의 요구가 증가하고 있다.

본 논문에서는 대화형 방송의 한 방식으로 분할된 객체의 정보를 이용한 대화형 방송 시스템의 구성 방법을 제안하였다.

이 방법을 통해서 생성된 방송콘텐츠는VoD로 제작되어도 동일한 방법으로 부가정보 서비스를 제공할 수 있는 장점이 있고 자동 영상 분할 작업 없이 실시간 검색을 통해 부가정보를 생성해 제공하므로 부가정보의 갱신작업이 필요하지 않아 서비스의 구현 면에서 현실적이고 유지 면에서 유리하다.

본 논문에서 제안하는 방법을 구현하기 위해서는 방송데이터에 메타데이터를 포함시키는 기술의 도입이 진행되어야 할 것이고 객체 추적 및 부가정보 검색 API를 이용해 획득한 데이터를 실시간으로 재구성하는 어플리케이션의 개발이 이루어 져야 할 것이다.

향후에는 기존 데이터를 이용한 객체 추적 기술, 실시간 검색을 통해 획득한 부가데이터를 편집 생성하는 기술을 개발 및 적용하여 제안하는 방법의 시스템을 구현하는 연구를 진행 할 것이다.

#### 참 고 문 헌

- [1] ATSCStandard, A/98, ATSC Interaction Channel Protocols, 2004. 2
- [2] 천왕성, 이상홍 “동영상 내 시청자 선택의미 객체의 부가정보를 제공하는 대화형 방송 서비스 구현,” 통신공학회논문지 (2008년 하계), pp.1874-1877, 2008. 7
- [3] 한대영, 황부현, “대화형 멀티미디어 방송을 위한 객체 저장 방법,” 한국콘텐츠학회논문지, 제9권, 제2호, pp.51-59, 2009.2
- [4] 이희경, 호요성, “영상분할과 특징점 추출을 이용한 영역기반 영상 검색 시스템,” 방송공학회논문지 제7권, 제3호, pp.262-270, 2002.9
- [5] 김장희, 강대성, “장면 전환 기법을 이용한 동영상검색 시스템 설계,” 전자공학회논문지, 제44권 SP편, 제3호, pp.8-15, 2007. 5
- [6] www.viaccess-orca.com
- [7] 고민구, 유지상 “객체의 색상 정보와 차영상을 이용한 동영상 내 객체 추적 기법,” 방송공학회추계학술대회(2010년 추계), pp.180-183, 2010
- [8] 윤경로, “맞춤형 방송 기술 및 동향,” 전자공학회지, 제33권, 제 6호, pp.39-48, 2006. 6
- [9] ISO/IEC “MPEG-7 Overview(version 10),”N6862, 2004. 10
- [10] 강정원, 이희경, 김재근, “맞춤형 콘텐츠 서비스를 위한 TV-Anytime 기반 콘텐츠 패키징,” 방송공학회논문지, 제9 권, 제4호, pp.322-333, 2004. 9
- [11] The TV-Anytime Forum, Requirement Series: RQ001v2.0 on Phase 2 Business Model: TV140, 2003.
- [12] Bumsuk Choi, Jeonghak Kim, “A Metadata Design for Augmented Broadcasting and Testbed System Implementation,” ETRI Journal, Volume 35, Number 2, pp. 292-300, 2013. 4