

# 비디오와 동기화된 물체의 위치정보 표현 data stream 생성 및 해석기 구현

나 희 주\*, 김 정 환\*\*, 정 문 열\*\*\*,  
서강대학교 영상대학원 미디어공학과 디지털방송연구실  
naheejoo@sogang.ac.kr\*

## Generation and Interpretation of data stream for position data of objects synchronized with video

Hee-Joo Na, Jung-Hwan Kim, Moon-Ryul Jung,  
Digital Broadcasting Lab, Dept of Media Technology, Sogang Univ.

### 요 약

본 논문은 디지털 방송 프로그램 진행 중 비디오의 특정 시점에 동기화된 특정 객체의 위치정보를 표현하는 data stream을 생성하고, 그 시점에 해당 위치 정보를 해석하는 해석기에 관한 것이다. 현재의 상용 스트림 생성기는 디지털 방송 표준에서 권고하는 스트림 이벤트의 발생 시각과 셋톱박스에서 디코딩 시에 사용할 참조값을 적절하게 생성하지 못하고 있다. 또한, 셋톱박스에서 동작하는 애플리케이션(Xlet) 역시 STC(System Time Clock), PCR(Program Clock Reference), NPT(Normal Play Time) 등의 시간값을 적절하게 읽어내지 못하고 있다. 더욱이, 현재의 디지털 방송 표준에서는 영상 내 특정 객체를 위해 정보를 제공하는 데에는 한계가 있다. 따라서, 본 논문에서는 다양한 연동형 디지털 방송 프로그램 제작을 위해 비디오의 특정 시점에 동기화된 객체의 위치정보를 표현하는 data stream을 생성하는 방법과, 동기화된 데이터를 처리하는 애플리케이션에 대해서 설명한다.

### I. 서 론

디지털 방송은 아날로그 방송에서 볼 수 없었던 고화질의 서비스를 제공한다는 점에 있어서 기존의 방송과 차이를 나타내지만, 비디오와 오디오 신호만을 제공하는 기존 방송과 달리 수신기에서 구동되는 애플리케이션 프로그램과 이미지 파일이나 텍스트와 같은 데이터도 시청자에게 함께 전달하는 것을 가능하게 함으로써 아날로그 방송과 뚜렷한 차이를 가지게 되었다. 여기서 애플리케이션 프로그램이란 디지털 방송 신호를 수신하여 시청자에게 제공하는 수신기(set-top box)에서 구동되는 프로그램을 말하며, 이를 통하여 방송 서비스

제공자는 시청자에게 보다 다양한 부가 정보를 제공할 수 있다. 통상 이러한 형태의 애플리케이션 프로그램을 엑슬릿(Xlet)이라 부르는데 방송 프로그램과의 연관성에 따라 크게 독립형과 연동형으로 구분할 수 있다. 독립형은 방송 프로그램과는 별도로 구동되는 애플리케이션을 말하고, 연동형은 비디오 내의 특정 시간에 특정 행위를 하도록 구성된 애플리케이션을 지칭한다. 애플리케이션에 대한 상세한 설명과 예시는 문헌 [1], [3]에 기술되어 있다.

초기 디지털 방송에서 제공하는 데이터 방송 서비스는 그 제작의 간편성으로 인하여 방송 프로그램과는 별도로 구성되어 제공되는 독립형 서비스가 대부분이었다.

그러나 점차 디지털 방송 서비스가 발전함에 따라 연동형 데이터를 이용한 방송 서비스도 함께 발전하고 있다. 기존의 연동형 애플리케이션은 주로 프로그램의 시작 시점에 맞추어서 구동되는 프로그램 연동형 애플리케이션이 대부분이다. 또한 애플리케이션에서 데이터를 이용함에 있어서 기존의 방송은 특정 시점에 등장하는 특정 객체와는 관련이 적은 데이터를 이용하는 것이 대부분이다. 이는 현재 수신기에서 구동되는 애플리케이션이 각종 데이터 방송 규약의 원형인 DVB-MHP 규약<sup>18)</sup>과 MPEG-2 시스템 규약에서 제공하는 각종 시간값에 대한 정보를 효과적으로 이용하지 못하고 있기 때문이다. 따라서 현재 상용 수신기는 방송 프로그램의 특정 시점에서 등장하는 특정 객체와 동기화된(synchronized) 정보를 제공하거나 활용하는 데 있어서 많은 어려움을 가지고 있고, 이는 디지털 방송이 시청자에게 대화형 서비스를 제공하는 데 있어 중요한 한계점으로 나타나게 된다.

본 논문에서는 방송 프로그램 비디오상의 특정 객체의 화면 내에서의 위치정보와 그 객체가 프로그램에 등장하는 시점에 대한 시간정보를 가지는, 비디오와 동기화된 데이터 스트림을 생성하는 방법과 수신기에서 이러한 정보를 획득하여 시청자에게 알려주기 위한 데이터 스트림 해석기를 제안한다. 본 논문의 이해를 돕기 위하여 비동기, 동기화, 동기화된 데이터의 유형에 대한 설명과 이를 전송하기 위한 데이터 전송 프로토콜, MPEG-2 시스템 타이밍에 대한 설명을 간략하게 설명한다. 이어서 비디오와 동기화된 데이터 스트림의 생성 방법에 대한 설명과 수신기에서 이를 활용하기 위한 해석기에 대해서 기술하고 결론을 맺는다.

## II. 디지털/데이터 방송 개요

본 장에서는 본 논문에서 제안하는 비디오와 동기화된 데이터 스트림의 생성과 수신기에서의 활용을 위한 해석기를 이해하는 데에 필요한 디지털/데이터 방송에서의 데이터의 유형에 대해 살펴보고, 또한 이를 전송하기 위한 프로토콜과 비디오와 동기화된 데이터를 위한 MPEG-2 시스템 타이밍에 대하여 설명한다. 이에 대한 자세한 설명을 원한다면 문헌 [2]를 참고한다.

### 1. Timing Model (Synchronization)

디지털 방송에서 비디오/오디오와 함께 전송되는 데이터의 유형은 방송 프로그램과의 연동성에 따라 비동기, 동기, 동기화된 데이터로 구분하고 있다.

비동기(asynchronous) 데이터는 세 가지 유형 중 방송 프로그램과 시간적인 연관성이 가장 적은 데이터로서 증권정보, 뉴스, 날씨, 교통정보 등과 같은 독립적으로 제공할 수 있는 데이터이다. 현재 데이터 방송 서비스의 대부분을 차지하는 유형이라 할 수 있다.

동기(synchronous) 데이터는 비디오와 오디오에 대한 분석 없이 단순히 방송 프로그램과 동기를 이루는 유형의 데이터로서, 예를 들어 축구경기의 중계방송에 있어서 경기장의 정보나 선발출장 선수에 대한 정보는 비디오와 특정 시점에서 동기화될 필요는 없지만 방송 프로그램이 시작되는 시점에 맞추어 정보를 지니고 있어야 한다.

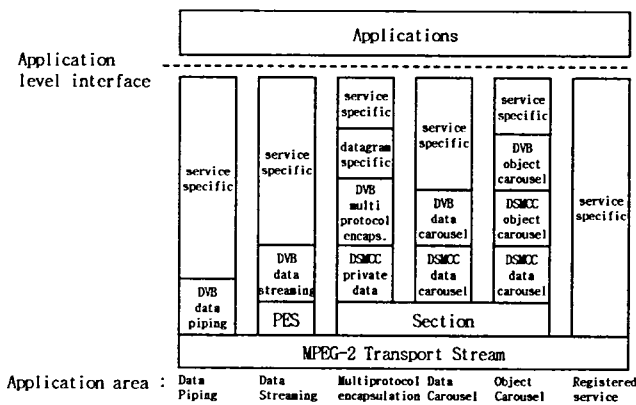
동기화된(synchronized) 데이터는 특정 방송 프로그램 내의 비디오/오디오와 연계되어 특정 시점에서 활성화되어 제공되는 데이터의 유형으로서, 본 논문에서 특정 객체의 출현 여부와 그 위치정보를 제공하는 데에 활용할 데이터이다. 이는 디지털 방송이 추구하는 대화형 서비스를 제공하기 위해 필수적으로 사용해야 할 데이터의 유형이라 할 수 있다.

### 2. Data Transport Protocol

현 디지털 방송에서는 비디오, 오디오, 데이터를 수신기로 전송하기 위해 MPEG-2 Systems[4] 규약에서 정의된 MPEG-2 Transport Stream을 사용하고 있다. DVB에서는 MPEG-2 전송 스트림을 구성하는 데이터 전송 프로토콜로 데이터 파이프(data pipe), 데이터 스트리밍(data streaming), 다중프로토콜 캡슐화(multiprotocol encapsulation), 데이터 캐루셀(data carousel), 객체 캐루셀(object carousel)을 제안하고 있다.[5][6]

데이터 파이프(data pipe)란 비동기 전송 메커니즘으로써 MPEG-2 전송 패킷(Transport Packets)의 페이로드(payload)에 직접 실려 전송되는 프로토콜로 전송 메커니즘이 상세히 구현되어 있지는 않다. 데이터 스트리밍(data streaming)은 방송 프로그램과 비동기, 동기, 동기화된 데이터를 전송하는 프로토콜로 MPEG-2

PES(Packetized Elementary Stream) 부분에 데이터를 실어 전송한다. 다중프로토콜 캡슐화(multiprotocol encapsulation)는 DSMCC 형식을 이용하여 MPEG-2 섹션(section) 부분에 데이터를 실어 전송하는 방식으로, 인터넷 프로토콜(IP : Internet Protocol)과 비슷한 메커니즘을 통해 하나의 수신기, 혹은 정해진 수신기 그룹이나 모든 수신기로 전송할 수 있다. 데이터 캐루셀(data carousel)은 방송망에서 데이터 모듈을 주기적으로 전송하여 수신기에서 손실 없이 데이터를 받을 수 있도록 하는 것이고, 객체 캐루셀(object carousel)은 방송 서버의 디렉토리, 파일, 스트림 등의 객체들을 구조화하여 수신기에 전송하는 것을 정의한다. 아래 그림은 각 단계의 프로토콜을 도식화한 것이다.



[그림 1] 데이터 전송 프로토콜<sup>[6]</sup>

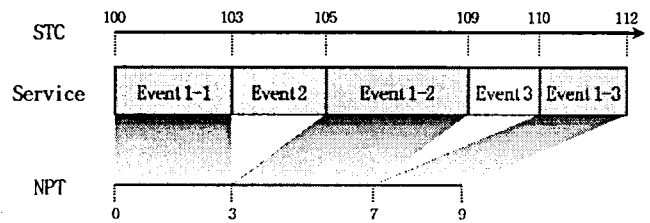
### 3. MPEG-2 System Timing

디지털 방송에서 대화형 서비스를 제공하기 위해서는 현 데이터 방송 전송 규약의 근간이 되는 MPEG-2 System Timing 모델<sup>[4]</sup>은 방송 프로그램의 시간값과 유기적인 관계를 맺고 있기 때문에 중요한 의미를 가진다. 표준에서 정의하는 대표적인 시간값에는 STC(System Time Clock), PCR(Program Clock Reference), PTS(Presentation Time Stamp), DTS(Decoding Time Stamp)와 DVB-MHP 규약에서 정의하고 있는 NPT(Normal Play Time) 등이 있다.

STC란 인코더(방송국 서버)와 디코더(수신기) 모두 지니고 있는 시간으로, 인코딩, 디코딩 할 시에 방송국 서버와 수신기가 참조하게 되는 절대적인 시간값이다. 디코더의 STC는 MPEG-2 TS 패킷의 adaptation field로

전송되는 PCR을 이용하여 보정하게 된다. DTS는 access unit을 디코딩해야 하는 시간이고 PTS는 디코딩한 unit을 화면에 보여주는 시간이다.

그리고 NPT는 객체 캐루셀로 전송되는 시간값으로서, DVB에서 'event' 라고 부르는 단위 프로그램 내에서 증가하는 시간이다. NPT는 단위 프로그램(이벤트)에 종속적인 시간이므로 진행 중인 단위 프로그램에 광고 등과 같은 다른 프로그램이 삽입되는 경우에도 변하지 않는다. 따라서 본 논문에서 제시할 비디오와 동기화된 데이터 스트림을 처리함에 있어서도 언제나 해당 프로그램 내의 동일한 위치를 가리키는 NPT 값의 사용이 적절하다.<sup>[7]</sup> 아래 그림은 전역시간인 STC와 국지시간인 NPT의 차이를 보여준다.



[그림 2] STC와 NPT의 차이<sup>[7]</sup>

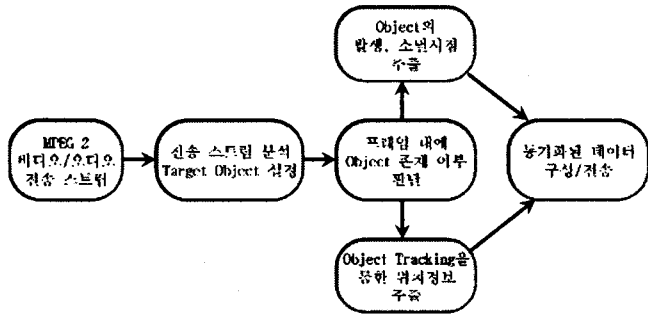
### III. 비디오와 동기화된 data stream 생성

본 논문에서는 단위 프로그램 내에서 재생되고 있는 동영상 프레임에 할당된 시간값을 통하여 프로그램 내의 특정 객체의 존재 여부와 위치, 표시 시간 등의 부가정보를 생성하여 수신기에서 활용할 수 있도록 구조화시키는 방법을 제안한다.

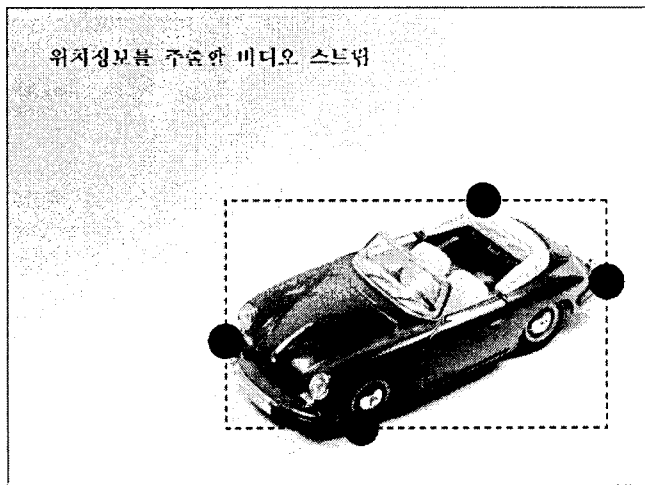
#### 1. Target Object 선정, 대상 스트림 추출

[그림 3]은 비디오/오디오로 이루어진 전송 스트림에 특정 객체에 해당하는 동기화된 데이터를 구성하여 수신기에 전송하는 과정을 도식화한 것이다. 부가정보를 삽입하고자 하는 방송 프로그램을 선택하면 첫 번째로 데이터의 구성에 필요한 정보를 추출하기 위해 비디오/오디오 만으로 이루어진 전송 스트림을 분석하여 타겟으로 설정할 객체를 선정하는 작업을 거친다. 부가정보를 삽입하고자 하는 객체가 선정되면, 그

다음으로 단위 프로그램의 비디오를 프레임 별로 순차적인 번호를 부여해서 분석하고 각 프레임 내에 target object가 존재하는지 여부를 판단하여 object가 존재하는 프레임을 추출해내고 포인터를 부여하여 각 object 별로 구분한다.



[그림 3] 특정 Object에 해당하는 동기화된 데이터 구성

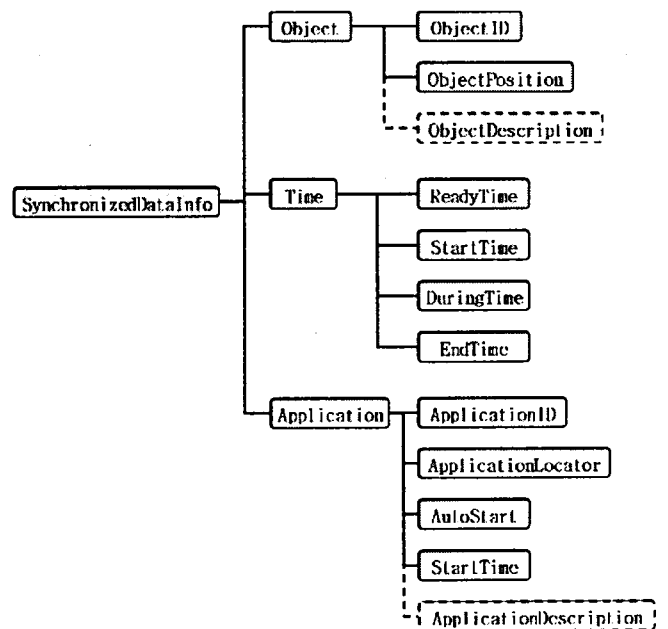


[그림 4] 위치정보 데이터를 구성하는 object의 바깥쪽 네 점

## 2. 특정 객체와 동기화된 정보의 구성

추출한 프레임에서 정보를 구성하는 것은 두 가지로 진행되는데, 첫 번째는 본 연구팀에서 이미 실험을 통하여 검증한 바 있는 NPT를 이용한 동기화, NPT 재구성 기법을 활용하여 *currentNPT*를 구하여<sup>[7]</sup> 추출한 첫 번째 프레임과 마지막 프레임의 시간 정보를 얻어 target object의 발생 시점과 소멸 시점을 구성하는 것이다. 그리고 두 번째, 파티클 필터(particle filter)와 CONDENSATION 기법을 통하여 target object가 존재하는 부분의 비디오 스트림을 분석하여 object tracking<sup>[9]</sup>을 통해 발생한 object의 위치정보를 구성한다. 또한

위치정보에 대한 데이터는 일정한 주기로 target object의 위/아래, 좌/우의 가장 바깥쪽의 좌표값 네 개를 스캔하여 구성한다.[그림 4] 데이터를 바깥쪽의 좌표 네 개로 구성하는 이유는 위치정보를 단순화하여 전송함으로써 해석기에서 target object에 대한 추가정보를 분석할 때 과도한 부하를 막아 수신기 내부에서 발생할 수 있는 지연시간을 줄이기 위함이다. 이렇게 얻어진 기본정보를 포함하는 동기화된 데이터를 효과적으로 처리하기 위하여 단위 프로그램에 대한 전반적인 정보와 재구성한 특정 시점에 대한 정보[그림 5]를 구성하여 전송한다. NPT를 이용한 동기화와 재생되는 비디오의 오브젝트 트래킹에 대한 자세한 설명은 문헌 [7], [9]를 참고하기 바란다.



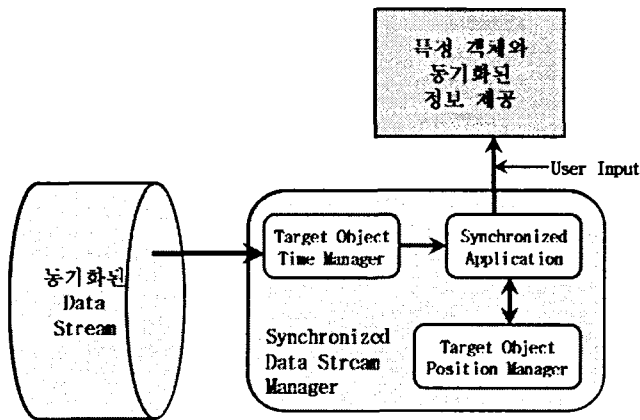
- ObjectID : target object의 고유 ID
- ObjectPosition : object의 위치정보
- ObjectDescription : object의 상세정보
- ReadyTime : 동기화된 데이터 스트림 준비시간
- StartTime : 동기화된 데이터 스트림 시작시간
- DuringTime : 동기화된 데이터 스트림 지속시간
- EndTime : 동기화된 데이터 스트림 종료시간
- ApplicationID : 구동되는 애플리케이션의 고유 ID
- ApplicationLocator : 애플리케이션의 target 정보
- AutoStart : 자동실행 flag
- StartTime : 애플리케이션 구동시간
- ApplicationDescription : 애플리케이션의 상세정보

[그림 5] 특정 시점에 동기화된 data stream의 계층구조

## IV. 동기화된 data stream 해석기

### 1. 해석기의 구성

수신기에서 구동될 해석기는 비디오와 동기화된 데이터 스트림을 이용하여 방송 프로그램에서 target object가 발생할 특정 시점에 도달하였을 때 시청자의 요청이 있을 경우 데이터 스트림에 포함된 부가정보에 대한 처리를 관장한다. 아래 그림은 동기화된 데이터 스트림을 처리하는 해석기에 대한 구조를 나타낸 것이다.



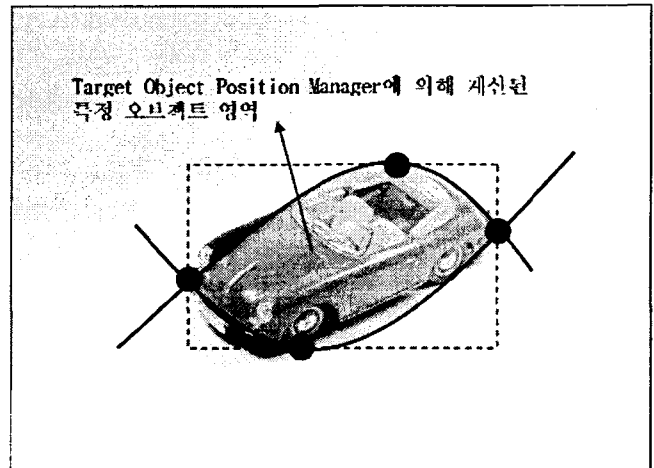
[그림 6] 동기화된 데이터 스트림을 처리하는 해석기 구성

### 2. 해석기 내부의 데이터 스트림 처리 과정

동기화된 데이터 스트림이 Synchronized Data Stream Manager의 Target Object Time Manager로 전송되면 Time Manager는 특정 객체의 발생시점과 지속시간, 소멸시간을 해석하고, 방송 스트림에 포함되어 있는 NPT 시간값을 분석/비교하여 방송 프로그램이 발생시점에 도달하면 Synchronized Application을 구동시킨다. 이 Application은 시청자에게 화면상으로 시청 중인 방송 프로그램에 부가정보가 포함되어 있음을 알리고 시청자의 input을 기다린다. 시청자의 input은 시청 중인 화면 상에 마우스 포인터와 같은 커서를 출현시켜 리모컨의 방향 버튼을 이용하여 시청자가 커서를 움직임으로 해서 획득할 수 있다. 만약 시청자가 움직이는 커서가 특정 객체의 위치영역에 포함되면 이 객체의 부가정보가 있음을 시각적으로 알리고, 시청자가 확인 버튼을 눌렀을 경우 Application이 획득한 부가정보를 화면에 나타내어 준다.

### 3. 특정 객체의 위치정보 판독

Application은 시청자의 input이 입력될 때마다 Target Object Position Manager를 호출하여 시청자가 움직인 커서의 위치정보를 넘겨주고, Position Manager가 커서와 특정 객체의 위치정보를 비교하여 커서가 특정 객체 내에 존재하면 Application에게 알려주어 Application이 화면에 부가정보를 표시하도록 한다. Position Manager이 데이터 스트림에 포함된 특정 객체의 위치정보를 분석하는 방법은 동기화된 데이터 스트림이 가지고 있는 특정 오브젝트의 바깥쪽 네 점을 point\_north, point\_south, point\_east, point\_west로 구분하여 동, 서, 북쪽의 포인트로 북쪽이 최고점이 되는 곡선을 만들고 동, 서, 남쪽의 포인트로 남쪽이 최저점이 되는 곡선을 만들어 커서의 위치가 두 곡선이 만드는 영역 내에 포함되는지를 판별하는 것이다.[그림 7] 이러한 위치 판독은 비록 완벽하지는 않지만 데이터 스트림으로 전송되는 정보의 양을 줄일 수 있다는 점에서 효과적이라 할 수 있다.



[그림 7] 특정 오브젝트의 영역 판별

## V. 결 론

현재 디지털 방송 서비스는 매체 통합 등의 다양한 이슈를 가지고 있으며, 진정한 데이터 방송 서비스를 구현하기 위해서 시청자에게 프로그램과 동기화된 정보를 제공하는 것도 그 중 하나라고 할 수 있다. 프로그램의 특정 객체와 연관된 부가정보를 제공함으로써 시청자가

자신의 선택에 따른 정보를 선별적으로 제공받을 수 있어 넘쳐나는 정보 속에서 사용자가 직접 필터링을 할 수 있다는 점에 있어서는 데이터 방송 서비스가 지향하는 대화형 서비스의 제공과도 무관하지 않다고 할 수 있다. 또한 현재까지 제한적으로 이루어졌던 연동형 서비스의 질을 한 단계 높이는 방법으로 큰 보탬이 되리라 기대된다.

## 참 고 문 헌

- [1] Steven Morris, Anthony Smith-Chaigneau.  
“Interactive TV Standards.” Focal Press. 2005.
- [2] Richard Chernock. “Data Broadcasting.” McGraw-Hill. 2001.
- [3] 정문열, 백두원. “연동형 데이터 방송 애플리케이션의 구조”. 방송공학회논문지. 제9권 제1호. 2004.
- [4] ISO/IEC 13818-1 Generic Coding of Moving Picture and Associated Audio : Systems. 2000.
- [5] ISO/IEC 13818-6 Generic Coding of Moving Picture and Associated Audio : Digital Storage Media Command and Control. 1996.
- [6] TR 101 202 Digital Video Broadcasting (DVB) ; Implementation guidelines for Data Broadcasting. V1.1.1. 1999.
- [7] 정문열, 김용한, 백두원. “동기화된 데이터방송을 위한 근사적인 NPT 재구성 기법”. 방송공학회논문지. 제9권 제1호. 2004.
- [8] ETSI TS 102 812 Digital Video Broadcasting (DVB) ; Multimedia Home Platform (MHP) Specification 1.1. V1.1.1. 2001.
- [9] 최규형. “합성템플릿 매칭을 이용한 축구선수와 공의추적”. 서강대학교 학위논문. 2002.

---

\* 서강대학교 영상대학원 미디어공학과 석사과정  
 \*\* 서강대학교 영상대학원 미디어공학과 석사과정  
 \*\*\* 서강대학교 영상대학원 미디어공학과 정교수