

## MPEG-21 파일 포맷 기술과 동향

조용주, 홍진우, 김진웅(한국전자통신 연구원), 임영권(넷엔티비), 송문섭(픽스트리), 김형중(강원대학교)

### I. 서론

MPEG-21 표준은 다양한 네트워크 환경에서 호환 가능한 멀티미디어 리소스의 사용을 증대하기 위한 프레임워크를 정의하자는 목표로 2000년부터 표준화가 시작되었으며, 현재 MPEG-21 파일 포맷은 part 9으로 정의되어 상기 목적을 충족할 수 있는 새로운 구조의 멀티미디어 파일 포맷을 정의하기 위해 활발하게 진행 중에 있다<sup>1)</sup>.

본 논문에서는 MPEG-21 디지털아이템(Digital Item?DI), 참조(reference)된 정지영상, 동영상, mp4파일 등과 같은 미디어리소스들이 하나의 파일 내에 포함될 수 있는 세부적 방법을 설명한다. 특히, XML기반 텍스트로 제작된 DI를 멀티미디어 파일 포맷화 하기 위해 디지털아이템에서 미디어리소스를 참조하는 방법에 대하여 자세히 설명하고 있다<sup>2)</sup>. 제안한 멀티미디어 파일 포맷의 구조 및 기능은 추후 메타데이터와 미디어리소스의 조합으로 구성된 멀티미디어 콘텐츠의 전달을 위하여 하나의 파일로 제작할 때 참고 모델로서 유용하게 사용될 수 있다.

본 논문은 다음과 같이 구성되어 있다. II장

에서는 MPEG-21 파일 포맷 구조 및 기능에 관하여 자세히 설명하고, III장에서는 파일화 절차, IV장에서는 실험 및 고찰, V장에서는 향후 전망, 그리고 VI장에서는 결론을 맺는다.

### II. MPEG-21 파일 포맷 개요

MPEG-21 파일 포맷은 MPEG-4 파트 12인 ISO Media File Format<sup>3)</sup>의 기반으로 메타데이터인 MPEG-21의 DI를 포함하는 멀티미디어 파일구조를 정의하기 위해 2002년부터 표준화 중에 있다.

#### 1. 디지털아이템 (Digital Item)

디지털아이템(DI)은 MPEG-21 멀티미디어 프레임워크 내에서 표준화된 표현, 식별, 그리고 메타데이터를 지니는 구조화된 객체이다. 이는 효과적인 멀티미디어 제공 서비스를 위해 사용자로 하여금 콘텐츠에 대한 탐색을 가능케 하고, 저작권자의 권리를 이용자에게 알리며, 또한 콘텐츠의 이용에 관련된 정보 전달을 위한 메타데이터를 제공한다.

〈표 1〉 디지털아이템(DI)의 예

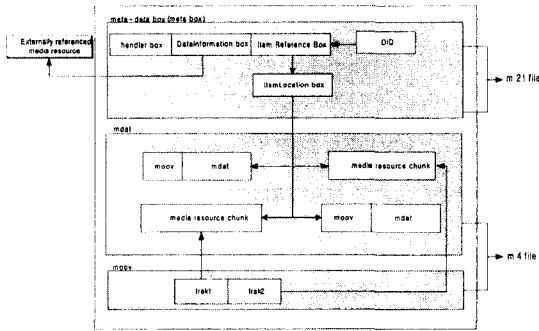
```

<DIDL xmlns="urn:mpeg:mpeg21:2002:01-DIDL-NS"
xmlns:dii="urn:mpeg:mpeg21:2002:01-DII-NS" x
mlns:mpeg7="urn:mpeg:mpeg7:schema:2001"
xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
xsi:schemaLocation="urn:mpeg:mpeg21:2002:01-DIDL-NS
C:\DOCUME~1\dh\1\MP21v3.1\ex\didl.xsd">
<Item>
<Component id="video">
  <Resource ref="weather1.h263" mimeType="video/h263" />
</Component>
<Component id="audio">
  <Resource ref="weather1.g723" mimeType="audio/g723" />
</Component>
<Component id="jpg">
  <Resource ref="fish.JPG" mimeType="image/jpg" />
</Component>
<Component id="mp4">
  <Resource ref="videoaudio.mp4" mimeType="video/mp4" />
</Component>
<Component id="gif">
  <Resource ref="http://www.etri.re.kr/image/logo2.gif" mimeType="image/gif" />
</Component>
<Component id="gif2">
  <Resource ref="http://www.etri.re.kr/image/top_logo.gif" mimeType="image/gif" />
</Component>
</Item>
</DIDL>

```

디지털아이템 선언(Digital Item Declaration(DID))은 디지털아이템의 구조체를 정의하는 부분으로 디지털아이템 서술 언어(Digital Item Declaration Language)을 이용한다. 이는 DI에 연계된 콘텐츠 및 메타데이터의 상호 관련성을 정의하고 있는 언어로서, XML Schema로 정의되어 있어 XML언어가 지니는 일반적인 특이인 유연성 및 확장성을 지닌다. 표1은 16개의 요소(element)들의 조합으로 제작

된 디지털아이템의 일 예를 보여주고 있다. 상기 디지털아이템에서는 각 미디어리소스의 종류 및 위치정보만을 서술하고 있지만, DIID(Digital Item Identifier), MPEG-7, 저작권 보호 및 관리(Intellectual Property Management and Protection-IPMP) 및 right(Right Expression)등에 대한 정보를 포함할 수 있으며, 추후 추가되는 부분/기능(description)에 대하여 확장 가능하다.



〈그림 1〉 MPEG-21 파일 구조

## 2. MPEG-21 파일 포맷 구조

MPEG-21 파일 포맷은 메타데이터인 DI와 미디어리소스 (예-동영상, 정지영상)들을 효과적이고 체계적으로 하나의 파일 내에 패키징 하기 위한 구조를 갖는다<sup>[3][4]</sup>. 또한, MPEG-21 파일 포맷은 MPEG-4 파일과 상호 양립해야 하는 요구사항을 충족 시키는 구조이다<sup>[5]</sup>. 상기 목적을 위하여 MPEG-21 파일 포맷은 meta-data, mdat, moov 박스로 구성된다. 상기 파일의 구조는 그림 1에서 잘 나타내고 있다<sup>[6]</sup>.

### 가) meta-data box

meta-data 박스는 handler box, Data Information box, Item Reference box, Item Location box, 그리고 DID box로 구성된다. 추후 특별한 목적을 위한 box추가도 가능하다. handler box는 meta-data box의 타입을 정의하는 정보를 포함하며, DataInformation box는 외부에 존재하는 미디어리소스의 위치정보(URI를 이용한 location 정보)를 담고 있는 테이블을 포함한다. Item Reference box는 DI에서 표현된 모든 미디어 리소스의 세부 위치정보(예: mp4 파일 내 특정 ES의 위치정보)를 포함하는 부분으로 Item Location Box에서 정의된 미디어리

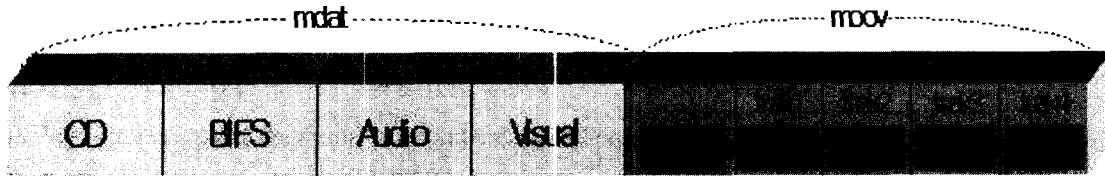
소스의 offset과 length정보와 함께 mp21 파일 내 특정 미디어리소스의 위치정보를 포함한다. 또한, DI에서 중복 정의된 미디어리소스를 Item Location box의 위치정보와 1:1 매핑시키는 기능을 제공한다. 이 논문에서 제시한 미디어리소스의 참조 방법이 표현되는 곳이기도 하다. Item Location box는 파일 디렉토리 기능을 하는 부분으로 파일 내 참조된 미디어리소스의 위치정보인 offset, length 정보를 포함한다. DID box는 XML기반의 디지털아이템을 포함한다.

### 나) mdat box

Mdat 은 미디어리소스들을 포함하는 box이다. 미디어리소스는 오디오, 비디오, 정지영상, 그리고 MPEG-4 파일(mp4)등 터미널에서 소비 가능한 미디어 콘텐츠를 의미한다. Mdat내에는 모든 포맷의 미디어리소스를 포함할 수 있다는 것이 특징이다. 이 미디어리소스들은 MPEG-21 터미널인 경우, meta-data box내 ItemLocation box 에서, MPEG-4 시스템 디코더인 경우, Moov box의 trak으로부터 참조(address) 받아 사용될 수 있다. MPEG-21 파일(mp21)은 meta-data, mdat, 그리고 moov box로 구성되나, moov와 Mdat box의 조합으로 MPEG-4 파일(mp4)이 구성될 수 있는 특징이 있다. 이는 MPEG-21 파일 포맷 요구 사항인 MPEG-4 backward compatibility를 충족시켜주기 위함이다. mp4인 경우, Mdat내 미디어리소스는 moov내 trak정보를 통해 참조 된다(그림 1, 2)

### 다) moov box

Moov는 Mdat과 함께 mp4를 구성하는 구조체로서, moov의 상세한 구조 및 기능들은



〈그림 2〉 MPEG-4 파일 구조

MPEG-4 systems(14496-1)에 자세히 설명되어 있다. 참고로 moov는 trak들의 집합이며, 각 trak은 미디어리소스의 정보를 담고 있는 “mdia”를 포함하고 있다. 즉, meta-data box에서 미디어리소스를 참조할 수 있는 것과 같이 moov도 trak정보를 이용하여 Mdat내의 미디어리소스를 참조 할 수 있는 것이다. 그림 2는 MPEG-4 파일 포맷의 구조를 나타내고 있다<sup>1)</sup>.

### 3. 미디어리소스 참조 방법

MPEG-21 파일 포맷에서 핵심이 되는 기술로서 파일 내 포함된 메타데이터(DI)에서 특정 미디어리소스를 참조하기 위한 방법을 의미한다. 미디어리소스는 파일 내 또는 외부에 존재가능하며, 본 절에서는 이 두 가지 경우에서 참조할 수 있는 방법을 자세히 기술한다<sup>2)</sup>.

#### 가) 파일 내부 미디어리소스 참조 방법

미디어리소스가 파일 내부에 포함될 경우, item reference box내 세부 위치정보(reference description<sup>1)</sup>라 칭함)와 item location box에 offset, length를 정의하여 특정 미디어리소스를

참조한다.

아래의 예제들은 파일 내부의 미디어리소스 참조 시, item reference box와 location box내 offset, length, 그리고 세부위치 정보 사용 예를 나타내고 있다.

<Resource target= “c:\test1.mp4” mimeType= “video/mp4” > 와 같이 DI에서 특정 미디어리소스(예: test1.mp4)를 참조 했을 때, 이 미디어리소스는 더 이상 PC내 C 디렉토리가 아닌 파일 안에 포함되기 때문에 target값은 <Resource target= “1” mimeType= “video/mp4” > 같이 파일 내에서 참조 가능하게 수정되어야 한다. 여기서 1은 item reference box내 index값을 의미한다. 즉, 첫 번째 item reference index에 “test.mp4” 미디어리소스의 위치 정보(offset와 length 값)를 포함한 item location box의 index값과 미디어리소스 내 특정 부분이 참조될 시, 세부위치 정보가 이 정의였다는 것을 의미이다. 이렇게 Resource element값이 변환된 DI는 meta-data box내 DID box에 포함된다.

MP4 파일은 여러 종류의 미디어리소스(Elementary Stream(ES)라 칭함)의 조합으로 생성될 수 있기 때문에, 위의 mp4 파일 내 특정

1) Reference description은 URI에서 사용하는 #, ?, &와 같은 기호를 이용하여 mp4 파일 내, 특정 ES의 위치를 표시하는 string 값을 의미한다. #, ?, &은 각각 Segment, query 그리고 AND를 의미한다.

ES를 참조할 수 있는 기능도 지원 가능해야 한다. 이 경우에는 item location box에서 offset, length 값 이외에 세부위치정보인 reference description을 item reference box에 정의하여 상기 미디어리소스의 특정 ES를 참조할 수 있다. 예로, test.mp4내 첫 번째 ES를 참조할 시, reference description은 "#moov/trak[1]"과 같이 정의된다. 만일 test.mp4 파일 내, odid와 esid값이 0x10과 0x11인 ES를 참조할 때에는, item reference box내에 "?odid=0x10&esid=0x11"와 같이 reference description을 item location box index값("test.mp4"파일의 위치정보를 기술한 부분)과 함께 정의하게 된다.

#### 나) 파일 외부 미디어리소스 참조 방법

파일 외부 미디어리소스를 참조할 때에는, data information box에서 이 미디어리소스의 위치정보인 URI값을 정의하고, Item location box와 item reference box에서는 파일의 인 offset, length, 그리고 세부 위치정보인 reference description을 각각 정의한다.

다음은 파일 외부의 미디어리소스 참조 시, item location과 data information box의 내용을 나타내고 있다.

<Resource target="http://www.etri.re.kr/test2.mp4" mimeType="video/mp4"> 와 같이 DI에서 외부 미디어리소스(예: test2.mp4)를 참조 하였을 때, <Resource target="2" mimeType="video/mp4"> 같이 수정된 DI는 DID box에 포함되며, item location box내 data?reference?index를 통해 data information box내 index를 참조하게 된다. 그리고 data information box에서는 외부 참조된 미디어리소스의 위치정보인 URI값인 "http://www.etri.

re.kr/test2.mp4"이 정의된다.

또한, 외부 참조된 mp21 파일 내 mp4 파일의 특정 ES를 참조할 경우, item location box에 mp4 파일의 offset, length 정보를, item reference box에 reference description(예: #moov/trak[1])을 정의하여 이 ES의 세부 위치 정보를, 그리고 data information box에서는 mp4 파일의 위치정보인 URI값(www.etri.re.kr/test.mp21)을 정의한다.

### III. MP21 파일화 절차

MP21 파일화 절차는 mp4 파일화 단계와 mp21 파일화 단계로 구분화된다.

#### 1. MP4 파일화 과정

MP4 파일화 과정은 텍스트로 저작된 IOD(Initial Object Description), OD(Object Description), SD(Scene Description), 그리고 미디어리소스를 입력 받아 mp4 파일을 생성하는 과정이다<sup>7)</sup>. 이 과정에서는 MPEG-4 표준에서 제공하는 reference S/W인 mp4 부호화 모듈을 이용한다. 생성된 mp4 파일은 그림 2와 같이 mdat과 moov box로 구성되는 구조를 갖는데, 이 box들은 그림 1에서 나타내는 것과 같이 mp21 파일을 구성하는 box이기도 하다.

#### 2. MP21 파일화 과정

MP21 파일화 과정은 입력된 DI를 이용하여 meta-data box를 생성하고, 상기 mp4 파일화 과정을 통해 나온 mdat과 moov box와 함께 완전한 mp21파일을 생성하는 과정이다.

Meta-data box는 DI에서 정의된 미디어리소스(mp4 파일화 과정을 통해 나온 MDAT box내 미디어리소스와 동일)들을 본 논문에서 제안한 미디어리소스 참조 방법을 이용하여 생성 된다.

세부적인 절차는 다음과 같다.

우선, 입력된 DI는 특정 미디어리소스 참조 정보가 파일 구조에 맞도록 3.1, 3.2장에서 기술한 방법과 같이 변환되어 DID box에 포함된다. 다음으로는, meta-data box를 구성하는 item location, item reference와 data information box를 앞장에서 기술한 정보들을 이용하여 생성한다. 이로써, mp21 파일을 구성하는 meta-data, moov, mdat box가 생성된다.

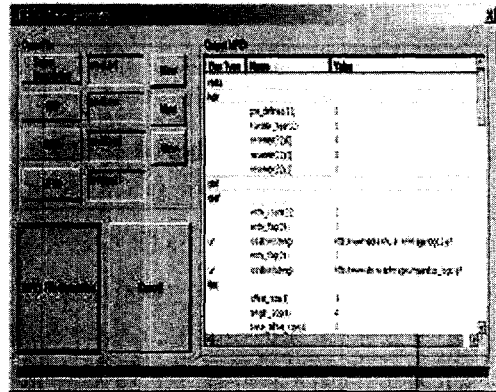
### IV. 실험 및 결과

#### 1. 실험

제안한 파일 구조 및 미디어리소스의 참조기능을 검증하기 위해 파일 부호화와 복호화를 개발하여 실험하였다.

사용된 모듈의 규격으로는 1.0GHz 펜티엄 CPU, 윈도우2000을 사용하였다.

또한, MPEG-21 파일 포맷 모듈에서의 테스트콘텐츠로서, 비디오(H.263), 오디오(G.723), 정지영상(JPG), mp4 파일과 같은 미디어리소스, MPEG-21 콘텐츠인 xml기반 디지털아이템, mp4파일화를 위해 IOD/OD, SD를 제작하여 사용하였다. 실험용 DI는 상기 4개의 미디어리소스를 포함하여 제작하였으며, OD, SD에서는 mp4를 제외한 비디오, 오디오, 정지영상만을 포함하여 제작되었다. 현재 mp4 부호화 모듈에서는 mp4 파일을 하나의 미디어리소스로 포함할 수 있는 기능이 지원되지 않기 때문이다. 따라



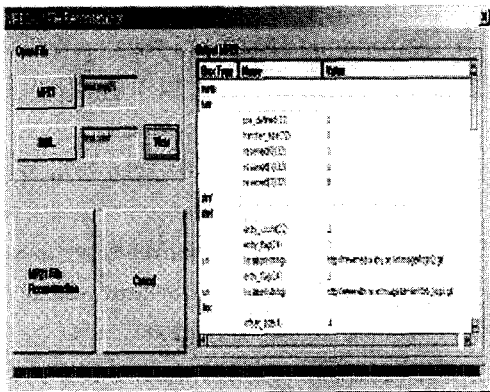
〈그림 3〉 mp21 file 부호화 모듈

서, MDAT내 포함함으로써, meta-data box에서 참조 가능케 하였다.

아래의 그림 3, 4는 본 논문에서 제안한 mp21 파일화 모듈에서 부호화 및 복호화 동작화면의 예를 보여주고 있다. 이 실험에서는 부호화 시, 제안한 구조(파일 포맷 구조 및 미디어리소스 참조 방법)로서 mp21 파일이 생성되는지에 대한 여부, 생성된 mp21 파일이 mp4 시스템 디코더에서 소비 가능한지에 대한 여부, 그리고 mp21 터미널에서 소비 가능한지에 대한 여부를 검증하게 된다.

그림 3은 부호화 화면으로서, mp21파일 생성의 예를 보여주고 있다. IOD, OD, SD, DI를 입력 받아 앞장에서 정의한 방법 및 절차를 통해 mp21파일을 생성하게 된다. 오른쪽 창은 부호화가 성공적으로 되었을 경우 파일을 구성하는 변수 정보를 표시하고 있어 파일의 구조를 쉽게 알아 볼 수 있다.

그림 4는 복호화 화면으로서, mp21 파일의 복호화 예를 보여주고 있다. 부호화 모듈과 마찬가지로 입력된 mp21파일의 구성 정보는 오른쪽 창에 표시가 되어 파일의 구조를 이해할 수 있고 또한 복호화 시, 검증의 역할도 할 수 있다.



〈그림 4〉 mp21 file 복호화 모듈

Mp 21파일을 입력으로 받아 이 파일에 포함된 DI와 MDAT내 각 미디어리소스를 분리하여 정의된 파일 이름으로 저장한다.

## 2. 결과 및 고찰

상기 mp21 파일화 모듈을 통한 실험 결과, mp21 파일은 그림 1에서 나타내고 있는 구조와 같이 생성되는 것을 확인하였으며, mp4 시스템 디코더 (IM?1 2D v5.0) 및 mp21 터미널 (자체개발)에서도 소비 가능함을 검증할 수 있었다. 즉, 제안한 mp21 파일 구조 및 미디어리소스 참조 방법은 여러 형태의 미디어리소스를 체계적이고 효율적으로 구성할 수 있음을 검증할 수 있었다. 단, 앞서 말한 바와 같이 mp4 파일화 과정에서 mp4 파일을 미디어리소스로 포함할 수 있는 기능을 지원하는 mp4 부호화 모듈의 개발이 선행되어야 할 것이다. 또한, mp21 파일 포맷은 현재 진행중인 표준화이므로, 지속적인 기능 업그레이드로 있어야 할 것이다.

## V. MPEG-21 파일 포맷 향후 전망

앞서 설명한 것과 같이, MPEG-4 파트 12?ISO Base Media File Format 표준화는 MPEG-21 파일 포맷의 모체로서, 현재 Proposed Draft Amendment(PDAM) 상황이며, 2004년 7월에 FDIS(Final Draft International Standard)를 거쳐 년 말에 International Standard(IS)를 목표로 표준화가 추진 중에 있다[8]. 반면, MPEG-21 File Format은 2004년 3월 회의에서 CD(Committee Draft)에 이어, 2005년 7월에 FDIS를 통해 당해 상반기에 IS를 목표를 표준화가 추진 중에 있다.

기술적으로는, MPEG-21 MDS내 Media Locator분과에서 진행중인 표준기술(MPEG-7 media description 포함)이 본문에서 정의한 미디어리소스 참조방법과 함께 MPEG-21 파일 포맷에서 포함되어 표준기술이 결정될 것으로 판단된다.

향후, MPEG-21 파일 포맷은 ISO Base Media File Format과 함께 다양한 멀티미디어를 포괄적으로 포함할 수 있는 저장용 파일 형태로서 MPEG-21 기반 멀티미디어 프레임워크 내 구성 시스템(서버 및 터미널)에서 널리 사용될 것으로 예상된다.

## VI. 결론

본 논문에서는 디지털 방송, 무선통신, 인터넷 환경에서의 다양한 형태의 콘텐츠들을 효과적이고 체계적으로 조합할 수 있는 기능 및 효과적인 저장 기능을 제공할 수 있는 MPEG-21 파일의 구조, 파일 내/외에 존재하는 미디어리소스를 참조 하는 방법, 그리고 향후 전망에 대해

서 설명하였다.

MPEG-21 파일 포맷은 XML기반 DI, mp4관련 데이터 및 여러 종류의 미디어리소스들을 하나의 파일 내에 체계적으로 구성함으로써 이 파일 내/외에 존재하는 미디어리소스들을 여러 종류의 터미널(예: mp4 시스템디코더, mp21 터미널)에서 상호운용성 기반 하에 소비 가능하게 하며, 각 미디어리소스의 random access기능을 지원 하며, 그리고 저장의 목적으로 사용하기에 적합하다는 것을 실험을 통해 입증할 수 있었다. 이는 다양한 메타데이터와 미디어리소스를 효과적으로 조합하여 다양한 정보를 하나의 파일로 구성할 수 있음에 의미를 가질 수 있다<sup>[6]</sup>. 더불어, 향후 본 논문에서 기술한 멀티미디어 파일 포맷은 MPEG-21 File Format 표준화 방향에 맞추어 수정 보완된 후 표준 안으로 제안함으로써 표준화에 기여할 수 있을 것으로 판단 된다.

#### ===== 참고문헌 =====

- [1] WD of 21000-1: MPEG-21 Vision, Technologies and Strategy, ISO/IEC JTC1/SC29/WG11/N6042, Oct 2003, Brisbane
- [2] Yongju Cho, Youngkwon Lim, Moonup Song, Hyungjoong Kim, "Suggestion to MPEG-21 File Format," ISO/IEC JTC1/SC29/WG11/M10077, Oct 2003.
- [3] Proposed MPEG-21 Container File Format ISO/IEC 21000-9 1.0, ISO/IEC JTC1/SC29/WG11/N5095, July 2002, Klagenfurt
- [4] Storage of untimed meta-data in ISO media files, ISO/IEC JTC1/SC29/WG11/W5991, Oct 2003, Brisbane
- [5] Draft Requirements for Digital Item Processing for

Applications and MP21 File Format, ISO/IEC JTC1/SC29/WG11/N4802, May 2002, Fairfax

- [6] 조용주, 홍진우, 김형중, 임영권, 김진웅, "MPEG-21 파일 포맷 - 미디어리소스 어드레싱 방법 및 보완," 한국정보통신설비학회 하계 학술대회, 2003.
- [7] ISO/IEC 14496:1, "Information technology- Coding of moving pictures and audio: Systems", Study of Draft Technical Corrigendum 1, 2000.
- [8] ISO/IEC 14496-12 & 15444-12, "ISO Base Media File Format, Amendment 1", Proposed Draft Amendment 12, 2003.



저자소개



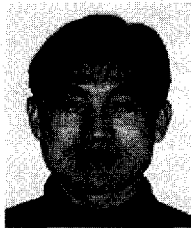
조용주

1997년 Iowa State University 컴퓨터공학과 학사  
 1999년 Iowa State University 전기과 석사  
 2000년 Iowa State University 전기과 박사수로  
 2001년-현재 한국전자통신연구원 근무  
 주관심분야 신호처리, 데이터방송, MPEG-2/4/21,  
 디지털 콘텐츠 보호 및 관리



임영권

1994년 한국항공대학교 항공전자공학과 학사  
 1996년 한국항공대학교 대학원 항공전자공학과 석사  
 1996년-2000년 한국전자통신연구원 근무  
 2000년-현재 ㈜넷앤티비 근무  
 주관심분야 통신신호처리, MPEG-2/4/21, 대화형 리  
 치미디어, 디지털 콘텐츠 보호 및 유통



송문섭

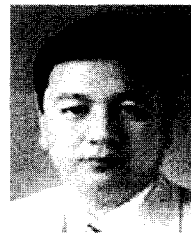
1992년 포항공과대학교 수학과 학사  
 1994년 포항공과대학교 수학과 석사  
 2002년-현재 ㈜픽스트리 근무  
 주관심분야 비디오 신호처리, 그래픽, MPEG-  
 2/4/7/21, 디지털 콘텐츠 관리

저자소개



김형중

1978년 서울대학교 전자공학과 학사  
 1986년 서울대학교 전자공학과 석사  
 1989년 서울대학교 전자공학과 박사  
 1989년-현재 강원대학교 교수  
 주관심분야 parallel computing, multimedia  
 computing, and IPMP(Intellectual Property  
 Management and Protection)



홍진우

1978년-1982년 광운대학교 응용전자공학과 공학사  
 1982년-1984년 광운대학교 대학원 전자공학과 공학석사  
 1990년-1993년 광운대학교 대학원 전자계산기공학과 공학박사  
 1998년-1999년 독일 프라운호퍼연구소(파견연구원)  
 1984년-현재 한국전자통신연구원  
 주관심분야 오디오 신호처리 및 부호화, 디지털 콘텐  
 츠 보호 및 관리, 디지털 오디오 방송



김진웅

1981년 서울대학교 전자공학과 학사,  
 1983년 서울대학교 전자공학과 석사  
 1993년 Texas A&M University 전기공학과 박사  
 1993년-현재 한국전자통신연구원 근무 (책임연  
 구원/방송미디어연구부장)  
 주관심분야 디지털 VLSI 신호처리, 영상 압축, 영상 통  
 신, 멀티미디어 데이터방송, MPEG-4/7/21, 콘텐츠 보호