

# MPEG-21 DIP 기반 비디오 요약 서비스

류지웅<sup>0</sup> 김문철<sup>0</sup>  
한국정보통신대학원대학교  
{zugoto<sup>0</sup>, mkim}@icu.ac.kr

## Video Summarization Service based on MPEG-21 DIP

Jeewoong Ryu<sup>0</sup> Munchurl Kim<sup>0</sup>  
Information and Communications University

### 요 약

MPEG-21 환경에서의 DI(Digital Item)은 MPEG-21 프레임워크 내에서 표준화된 표현 형식, 식별 체계, 서술 형식을 따르는 구조화된 디지털 객체이며, 유통, 처리의 최소 단위이다. 따라서, 이러한 DI가 MPEG-21 멀티미디어 프레임워크 환경에서 사용자 터미널에 전달되었을 때 어떻게 처리되어야 될 것인지를 규정하는 것은 매우 중요한 과제이며, 이와 관련한 기술이 DIP(Digital Item Processing)이다. 본 논문에서는 DIP의 한 응용 예로서 멀티미디어 콘텐츠를 계층적으로 기술하는 Video Summary의 응용 방안에 대한 연구 결과를 제시하고자 한다.

### 1. 서 론

MPEG-21 멀티미디어 프레임워크는 현재까지 독자적으로 발전한 멀티미디어 콘텐츠의 전달 및 이용을 위한 기반 구조들과 관리 체계 시스템으로 인해, 콘텐츠의 표현 방식에 대한 호환성, 네트워크 전달 방식과 단말간의 호환성 등의 잠재적인 문제를 해결하기 위해 멀티미디어 콘텐츠의 유통을 위한 프레임워크의 주요 요소들과 그들간의 관계를 확립하고 다양한 네트워크 및 장비를 통한 멀티미디어 콘텐츠의 접근, 사용, 가공, 재활용을 가능하게 하는 새로운 규격을 재정하는 것이다. 이러한 MPEG-21 환경에서 DI(Digital Item)는 멀티미디어 콘텐츠와 식별자, 그리고 서술자를 포함하는 유통, 처리의 최소 단위이다. 따라서, 이러한 DI가 MPEG-21 환경에서 사용자 터미널에 전달되었을 때 이미 규정된 방식으로 처리될 수 있는 표준을 정하는 것은 매우 중요한 과제이며, 이와 관련한 기술이 바로 DIP이다. 본 논문에서는 DIP의 한 응용 예로서 MPEG-7 MDS(Multimedia Description Scheme)의 Summarization DS(Description Scheme)를 이용하여 멀티미디어 콘텐츠의 이벤트 별 계층적 구조를 서술하는 Video Summary에 대한 활용 방안에 대해 제시하고자 한다. 또한, 본 논문에서는 DIA(Digital Item Adaptation) 관점에서 사용자의 단말 환경에 따라 적응된 형태를 제공하는 DIP 어플리케이션에 대해 제안하고자 한다.

### 2. DIP

#### 2.1 DIP의 필요성

DI의 경우 그 형태를 규정하는 표준은 이미 MPEG-21 part-2 DID(Digital Item Declaration)에서 정의하고 있지

만 이러한 DI의 Processing 레벨에서의 상호 호환성을 규정하기 위한 표준은 아직 정의되지 않고 있다. 따라서, 이러한 상호 호환성을 위한 DIP는 DI 문서를 받았을 때 DI의 요소들을 사용자에게 일정한 규격에 의해 표현하고 처리할 수 있는 DIM(Digital Item Method)의 목록을 정의하고 표현하는 것이 그 목적이다.

#### 2.2 DIP 요소들

DIP를 구성하는 요소들에는 DIM, DIME(Digital Item Method Engine), DIBO(Digital Item Base Operation), DIML(Digital Item Method Language) 등이 있으며, MPEG-21의 요구 사항에 따르면 각각의 요소들이 다음과 같은 기능을 담당한다. 첫째로 DIM은 DI를 의도하는 형식 또는 절차대로 처리하기 위해 DIP Engine에 의해 수행되는 일련의 절차적인 동작들로 구성되며, DI 내에 포함되어 DI 요소들을 처리, 식별하는 방법들로 조직화되어 있어야 하며, 발생하는 이벤트에 대해 처리할 수 있어야 한다. 둘째로 DIME는 DIM에서 규정하는 동작들을 수행할 수 있어야 하며, MPEG-21의 다른 요소들(DID, DII(Digital Item Identification), REL(Rights Expression Language), IPMP(Intellectual Property Management and Protection), DIA)과 연동될 수 있어야 한다. 셋째로 DIBO는 다음과 같이 DI 내의 요소들의 display, 규정된 방법에 따라 아이템 또는 리소스를 presentation, 사용자에게 선택 가능한 메소드 또는 리소스를 display, 사용자 입력에 따라 특정 행동을 취하거나 표현 지원, 리소스와 어플리케이션 실행, DI 전체 또는 일부 편집 등의 기능을 수행할 수 있는 것들을 의미한다. 넷째로 DIML은 DIM을 표현하기 위한 언어를 의미하며 다음과 같이 DIBO 또는

DIM을 호출할 수 있어야 하며, 분기나 선택 제어를 지원하여야 하며, 이벤트 또는 예외 처리를 지원하여야 하며, Concurrency 표현 능력 등의 기능을 지원하여야 한다.

2.3 DIP Architecture

다음 그림 1은 DIP의 전체적인 구성을 나타낸다.

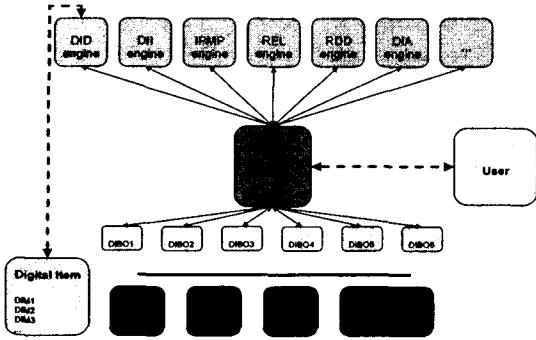


그림 1. DIP Architecture

그림 1과 같이 DIP는 DIME를 중심으로 MPEG-21의 주변 요소들과 연동되어 있다. 왼쪽 정선은 MPEG-21 터미널과 터미널에 도착한 DI의 상호 작용을 나타내며, 오른쪽의 정선은 사용자와 MPEG-21 터미널간의 상호 작용을 나타낸다.

아래 그림 2는 DIP 구성 요소들 간의 제어 흐름을 나타낸다.

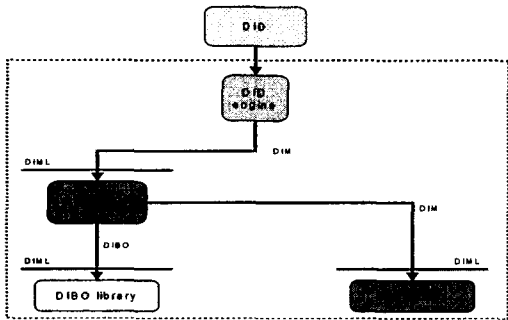


그림 2. DIP 구성 요소들의 제어 흐름

그림 2와 같이 터미널에 DI 문서가 도착하면 먼저 DID Engine이 이를 처리하여 DIM을 인식하여 DIME에 보내고, DIME는 DIBO Library와 DIM Library에 접근하여 필요한 DIBO와 DIM을 사용하거나 서버로부터 제공 받아 동작을 수행하게 된다. 상호간의 제어는 DIML을 이용한다.

3. DIP Test bed

본 논문에서는 MPEG-21 DIP의 한가지 응용 예로써 2002 한일 월드컵 경기 중 한 경기에 대한 Video Summary를 MPEG-7 MDS의 Summarization DS에 의해 서술하고 이를 담고 있는 DI를 활용할 수 있는 어플리케이션을 자바 환경에서 구현하여 본 논문에서 제시하고 있

는 Video Summary DI의 활용 방안에 대한 가능성을 제안하고자 한다. 사용자의 터미널 환경은 일반 PC를 통한 어플리케이션의 전체적인 활용과 PDA 등을 통한 제한적인 활용을 기본으로 한다. 이하 본 논문의 어플리케이션을 DIMlet이라 칭하도록 하며 그림 3은 DIMlet의 사용 시나리오를 나타낸다.

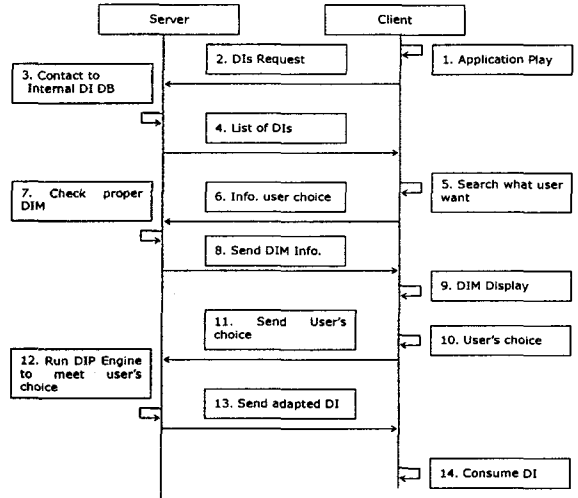


그림 3. DIMlet의 전체적인 흐름도

DIMlet은 서버, 사용자, 그리고 서로 간의 정보 전달을 위한 CDI(Content DI), XDI(Context DI)로 구성되며, 서버는 DI 목록, DIM 목록, 그리고 DI 처리들을 담당하며 CDI는 실제 멀티미디어 콘텐츠에 대한 정보를, XDI는 서버와 사용자간의 상호 작용 정보를 전달한다.

3.1 DIMlet

먼저 DIMlet을 실행하기 전에 전처리 과정으로서 사용자는 그림 4의 좌와 같이 자신의 터미널 환경에 대해 서술하고 있는 DIA 파일을 선택하여 서버에 전달하게 되며 서버는 그림 4의 중, 우와 같이 사용자의 터미널 환경에 적합한 DIMlet Interface를 제공하게 된다.

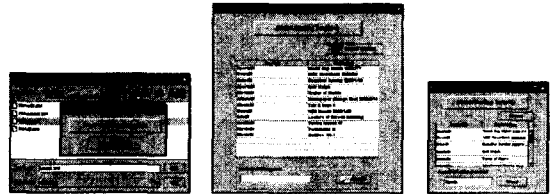


그림 4. DIMlet(좌: DIA 정보 제공 화면, 중: 터미널 초기 화면 PC, 우: 터미널 초기 화면 PDA)

사용자가 서버에 DI 목록이나 자신이 원하는 DI의 검색 쿼리를 제공하면 서버는 자신의 DB에서 DI 목록이나 검색 쿼리에 해당하는 DI들을 사용자에게 제공한다. 이때, 서버는 자신의 내부에 모든 DI 문서와 멀티미디어 콘텐츠를 가지고 있을 필요는 없으며, 단지 접근할 수 있는 경로 정보만을 가지고 있으면 된다. 사용자가 자신이 원하는

DI를 선택하면 서버는 선택 DI에 해당하는 DIM 목록을 사용자에게 표현하게 된다. 본 논문에서 사용하는 DIM 목록은 표 1과 같으며 DIMlet 화면은 그림 5의 우 상과 같다.

표 1. DIMlet의 DIM 목록들

Whole Video		Display whole video
Synopsis / Text		Web text summary / general text summary
Summary	Keyframe (slideshow)	Display keyframe image (slide show)
	Keyframe (Treeview)	Display keyframe image (tree view)
	Highlight (30 sec)	Display highlight video (30 seconds)
	Highlight (1 min)	Display highlight video (1 minute)
	Event summary	Event based summary (tree view, text, video)

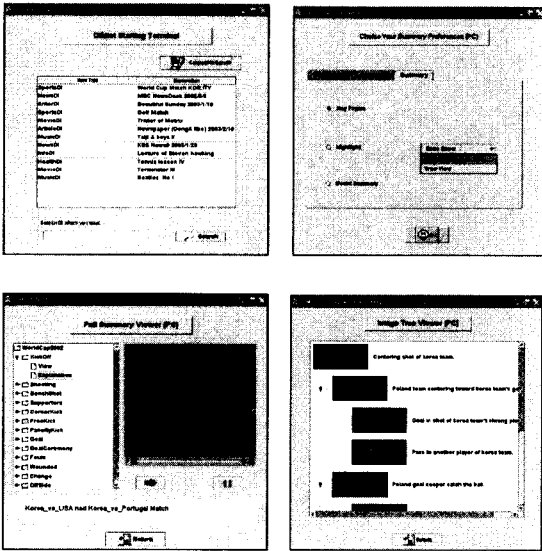


그림 5. DIMlet (좌상: DI 목록 표시, 우상: DIM 목록, 좌하: Event Summary, 우하: Keyframe 트리뷰)

사용자가 현재의 Video Summary DI를 표현하는 DIM 가운데 원하는 형태를 선택하여 서버에 이 정보를 전달하게 되면 서버는 해당되는 DIMlet Interface 또는 플레이어를 사용자에게 제공하게 된다. 그림 5의 좌하는 사용자가 이벤트별 Summary를 선택했을 경우의 DIMlet 화면이며, 우하는 사용자가 키프레임 이미지를 트리뷰 형태로 보기를 원했을 경우의 DIMlet 화면이다.

3.2 DIBO & DlxO

현재 DIBO는 DI를 처리함에 있어 가장 기본이 되는 기능들을 담당하기 위해 표2, 표3과 같이 리소스들을 처리하는 DIBO와 DID 처리를 위한 DOM 관련 DIBO가 있다.

표 2. 리소스 처리를 위한 DIBO

Play()	Copy()	Wait()	Stop()	Pause()
Resume()	Display()	Delete()	Store()	Etc.

표 3. DOM관련 DIBO

GetDIDLNode()	CreateDIDLNode()	RemoveDIDLNode()
CopyDIDLNode()	PresentChoice()	Etc.

또한, 이러한 DIBO의 기능을 확장하여 각각의 어플리케이션에서 특정 목적을 위해 사용할 수 있는 DlxO(Digital Item eXtension Operation)를 규정하였다. DlxO의 예로서 이미지를 라벨위에 순차적으로 보여주는 기능을 하는 IMGSlide(), 텍스트를 화면에 한자 한자씩 타이핑하듯이 표시하는 기능을 하는 TypeText(), 이미지를 계층 구조로 표현하는 기능을 하는 IMGTree() 등이 될 수 있다.

MPEG-21 DIP의 DIM 구성 요소들은 그림 6과 같이 M3W API라는 미들웨어 상에서 DIM을 기술하는 언어는 ECMAScript이며, DIBO와 DlxO를 구현하기 위해서는 어떠한 언어도 사용될 수 있으며 DIBO의 형태는 표준으로 정해져 있지만 DlxO의 형태는 제작자의 의도에 따라 제작될 수 있다.

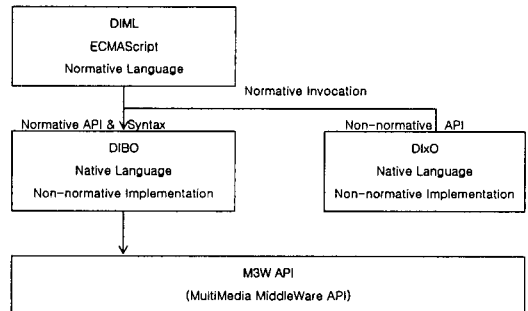


그림 6. DIM 구성 요소

4. 결론

본 논문에서는 자바 환경에서 Video Summary를 DI로써 이용한 DIP 어플리케이션의 활용을 제시함으로써 실제적인 DIM과 DlxO의 활용 예를 제시했으며, 사용자의 다양한 터미널 환경을 고려하기 위해 MPEG-21 DIA를 활용한 DIP 어플리케이션의 적응 환경 등을 제시함으로써 DIP와 DIA의 연동 가능성과 그 방향을 제시하는 좋은 예제가 되리라 생각한다.

참조

- [1] MPEG-21 Digital Item Processing Working Draft (WD) v.1, ISO/IEC JTC1/SC29/WG11/N5621, March 2003, Pattaya
- [2] Digital Item Basic Operation Functions and DIP Architecture, ISO/IEC JTC1/SC29/WG11/m9406, March 2003, Pattaya
- [3] MPEG-21 Digital Item Adaptation Draft v0.6, ISO/IEC JTC1/SC29/WG11/N5353, December 2002, Awaji