

# 입체 디지털 아이템의 MPEG-21 DIA 테스트베드

\*손현식, \*김만배, \*\*송영주, \*홍동희, \*김학수, \*\*홍진우  
\*강원대학교 컴퓨터정보통신공학과  
\*\*한국전자통신연구원 무선방송연구소 방송미디어연구부  
hssohn@kwnu.kangwon.ac.kr, manbae@kangwon.ac.kr

## MPEG-21 DIA Testbed for Stereoscopic Digital Item Adaptation

\*Hyunsik Sohn \*Manbae Kim \*\*Youngjoo Song \*Donghee Hong \*Haksoo Kim \*\*Hong Jinwoo  
Dept. of Computer, Info. and Telecom, Kangwon National University  
Broadcasting Media Research Department, Radio & Broadcasting Research Laboratory, ETRI

### Abstract

MPEG-21의 주요 요소 중 하나인 Digital Item Adaptation(DIA)은 사용자 선호도(User Preferences) 및 터미널 능력(Terminal Capabilities)에 따라 적응 변환된 디지털 아이템(digital item)을 제공한다. 본 논문은 MPEG-21 프레임워크에서 입체 영상 디지털 아이템의 적응을 구현하는 DIA 테스트베드의 구현을 목적으로 한다.

사용자 선호도에는 입체시차 종류, 깊이 범위, 지연 영상 간격 등이 있다. 터미널 능력으로는 코덱, 입체 렌더링 포맷, 디스플레이 타입 등이 있다. 상기 기술자들은 DIA 서버로 전송되어 적응된 입체영상 생성되며, 또한 변환된 입체영상 DI는 사용자의 요구에 따라 MPEG-2 또는 MPEG-4로 인코딩되어 전송된다. 클라이언트는 DIA 서버에서 전송되는 입체영상 DI를 받아 디코딩 한 후 디스플레이 하게 된다. 서버의 DIA는 입체 DI의 실시간 전송을 위하여 RTP와 RTSP를 사용하여 전송한다.

### I. 서론

현존하는 다양한 네트워크 환경 및 단말기들의 연동을 통한 투명하고 통합적인 멀티미디어 자원의 이용을 가능하게 하는 MPEG-21 멀티미디어 프레임워크의 표준화 활동이 진행 중이다[1, 2]. MPEG-21은 멀티미디어 콘텐츠의 체계적인 통합 관리를 가능하게 한다.

MPEG-21의 주요 요소 중 하나인 Digital Item Adaptation(DIA)은 사용자 선호도(User Preferences) 및 터미널 능력(Terminal Capabilities)에 따라 적응 변환된 디지털 아이템(Digital Item: DI)을 제공한다[3]. 리소스 적응(Resource Adaptation)과 기술자 적응(Descriptor Adaptation)으로 구성되어 있는 DIA는 사용자의 멀티미디어 콘텐츠의 종류, 네트워크 상태, 단말기의 사양에 맞추어 신뢰성 있고 투명한 서비스를 제공한다. 이러한 서비스는 사용자 요구에 맞는 디지털 아이템의 적응을 요구한다.

본 논문은 MPEG-21 프레임워크에서 3D 입체영상 DIA를 구현하고자 한다. 이를 위해 StereoscopicVideo Conversion DS 및 StereoscopicVideoDisplay DS (Description Scheme)를 제안한다. 이 DS에 기반하여 2D 동영상의 3D 입체영상 변환(2D-to-3D 동영상 변환) 및 3D 입체동영상의 2D 변환(3D-to-2D 동영상 변환)을 수행한다. DI 리소스는 MPEG-21 DIA에서 정의한 사용자 선호도 및 터미널 능력 등 사용자의 환경에 맞게 적응된다. 사용자 선호도에는 입체시차의 종류, 깊이의 범위, 지연 영상 간격 등의 기술자들이 있다. 터미널 능력으로는 코덱, 렌더링 포맷, 디스플레이 타입 등이 있다. 상기 기술자들은 클라이언트가 서버로 전송되어 적응된 3D 입체영상 DI가 생성되고 변환된 DI는 사용자의 요구에 따라 원하는 MPEG-4 등으로 압축되어 전송된다. 클라이언트에서는 DIA 서버에서 전송되는

입체 DI를 RTP(Realtime Transport Protocol) 및 RTSP(Realtime Streaming Protocol)를 사용하여 실시간으로 단말기에서 디스플레이가 된다[7].

본 논문의 구성은 다음과 같다. 1 절의 서론에 이어 II 절에서는 MPEG-21 DIA 테스트베드 개요에 대하여 설명한다. III 절에서는 입체 변환 Description Scheme을 제안한다. IV 절에서는 구현된 전송시스템을 소개하고 마지막 V 절에서는 결론 및 향후 연구 방향에 대해 설명한다.

### II. DIA 테스트베드 개요

그림 1은 DIA 테스트베드의 구조도이다. 이 테스트베드는 크게 DIA 서버와 클라이언트로 구성된다. 클라이언트는 서버의 DI를 요구하고 요구된 DI는 클라이언트의 DIA Negotiation Message에 따라 3D DIA에서 DI를 적응시켜 클라이언트로 전송한다. 전송할 때에는 RTP를 사용하고 전송 제어를 위해 RTSP를 사용한다.

3D DIA는 2D-to-3D 동영상 변환과 3D-to-2D 동영상 변환의 기술자 적응과 리소스 적응으로 구성된다(그림 2). 2D-to-3D 동영상 변환에서는 *ParallaxType*, *DepthRange*, *MaxDelayedFrame*, *RenderingFormat* 등의 기술자를 이용하고 리소스 적응은 2D-to-3D 동영상 변환 모듈을 사용한다. 또한 3차원 DI를 2차원 DI로 변환하는 경우에는 기술자 적응은 *LeftRightInterVideo*의 기술자를 이용하고 리소스 적응은 3D-to-2D 동영상 변환 모듈을 사용한다.

클라이언트의 요구에 따라 3D DIA에서는 2D MPEG-1 DI의 입체 변환작업을 수행한다.

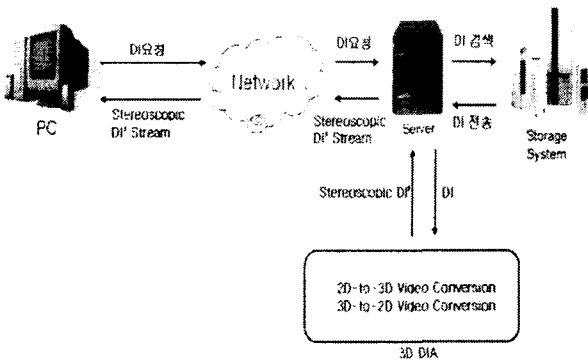


그림 1. 전체 시스템 구조도

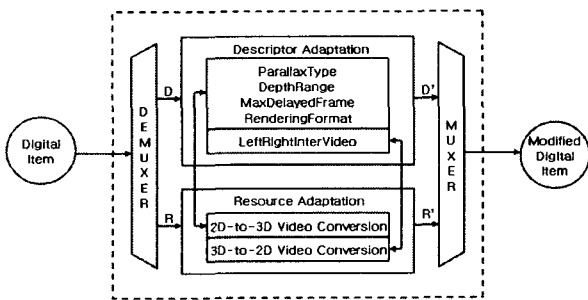


그림 2. DIA 테스트베드의 구조도

먼저 압축 데이터를 디코딩 한 후 사용자의 요구에 따라 입체 변환이 이루어지고 실시간 전송을 위하여 MPEG-2 또는 MPEG-4로 압축되어 RTP/RTSP를 이용하여 클라이언트로 전송된다. 그림 3은 이러한 입체 변환을 구현한 구조도이다.

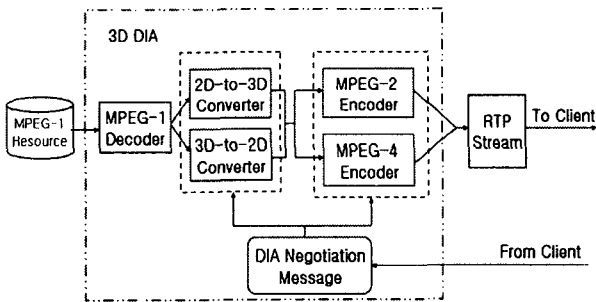


그림 3. 2D-to-3D 동영상 변환

### III. 입체 변환 Description Scheme

입체 변환 DS는 StereoscopicVideo Conversion DS와 StereoscopicVideoDisplay DS로 구성되어 있다[3, 4, 5].

#### 1. StereoscopicVideoConversion DS

그림 4는 StereoscopicVideoConversion DS의 Schema Diagram이다. 이 DS는 DIA/UserCharacteristics/PresentationPreference/Display/ 노드 밑에 위치하고 있다. From2DTo3D Stereoscopic은 2D동영상이 3D 입체 동영상으로 변환을 의미하고 3개의 ParallaxType, DepthRange, MaxDelayedFrame 요소를 가지고 있다. ParallaxType은 양의

시차(positive parallax)와 음의시차(negative parallax)로 구성되며 이는 시차의 유형을 의미한다. 음의 시차에서 입체의 깊이는 모니터 화면과 사람의 눈 사이에서 인식된다. 반대로 양의 시차에서 입체감은 모니터 화면 뒤에서 인식된다.

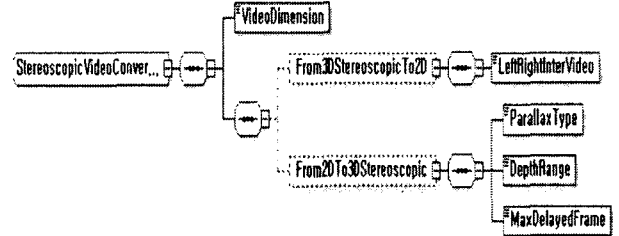


그림 4. StereoscopicVideoConversion DS

DepthRange는 사용자에 의해 인식되고 있는 3D깊이의 범위를 나타낸다. 예를 들면, 음의 시차에서 인식된 3D의 깊이는 증가 혹은 감소될 수 있다. 좌영상 또는 우영상을 적당하게 이동시킴으로써 3D 깊이 범위는 사용자의 선호에 따라 변할 수 있다. 입체 영상을 만들기 위해 지연(이전) 영상을 사용할 수 있다. 이때 영상 시퀀스를 {...I<sub>k-3</sub>, I<sub>k-2</sub>, I<sub>k-1</sub>, I<sub>k</sub>...}라고 가정하자. I<sub>k</sub>는 현재 프레임이다. 이전 프레임 중 하나인 I<sub>k-1</sub> (>1)을 선택한다. 그러면 입체영상은 I<sub>k</sub>와 I<sub>k-1</sub>로 구성된다.

MaxDelayedFrame은 현재 영상과 지연영상을 이용하여 입체 영상을 만들 때 현재 영상으로부터 몇 번째 지연영상을 사용하여 만들 것인지에 대한 값이다. 이 값이 크면 클수록 사용자는 더 많은 깊이를 느낀다.

From3DStereoscopicTo2D는 3D 입체 동영상이 2D 동영상으로 변환됨을 나타낸다. LeftRightInterVideo는 3D 입체 동영상이 2D 입체 동영상으로 변환될 때 필요하다. 입체 영상이 보통 좌영상 I<sub>L</sub>과 우영상 I<sub>R</sub>로 구성되므로 사용자는 원하는 어느 하나의 영상을 선택할 수 있다. 따라서 I<sub>L</sub>과 I<sub>R</sub>중 하나의 동영상을 선택할 수 있다. 또한 영상 처리 기술을 이용하여 새로운 합성 영상을 I<sub>L</sub>과 I<sub>R</sub>로부터 얻을 수 있다.

#### 2. StereoscopicVideoDisplay DS

그림 5는 DIA/terminalCapabilities/노드 트리 밑에 있는 StereoscopicVideoDisplay DS를 나타낸다. 2개의 하위 요소로 DisplayDevice와 RenderingFormat이 있다. DisplayDevice는 3D 입체를 보여주는 디스플레이 장치에 관한 특성이다. 단안 디스플레이 장치나 3D 입체 디스플레이 장치가 될 수 있는데 단안 디스플레이 장치는 3D를 보여주는 특성이 없는 모니터를 의미한다. 그러므로 단안 디스플레이 장치에서 3D입체 동영상을 보여주기 위해서는 2D 동영상을 3D입체 동영상으로 변환하는 작업이 필요하다. RenderingFormat은 2D 동영상을 3D 입체 동영상으로 변환하는 방법을 나타낸다. 이 값들은 다음 중 하나의 값을 가질 수 있다: Interlaced, Sync-Double, Page-Flipping, Anaglyph. 여기서 Anaglyph은 Red-Blue, Red-Cyan.

Blue-Yellow 등으로 구성 된다.

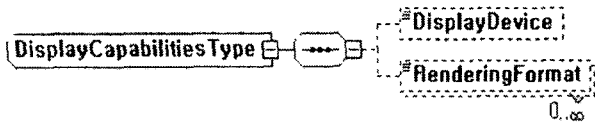


그림 5. StereoscopicVideoDisplay DS

#### IV. 전송 시스템

DIA 테스트베드에서 입체 동영상 디지털 아이템의 실시간 전송을 위하여 RTP와 RTSP가 사용한다.

RTP(Real-time Transport Protocol)는 종단간의 네트워크상에서 음성, 영상 또는 시뮬레이션 데이터 등 실시간 전송이 필요한 응용에 대하여 편리한 전송기능을 제공하는 프로토콜이다. 기존의 TCP(Transmission Control Protocol)는 신뢰성을 너무 강조한 나머지 실시간 전송은 불구하고 일반적인 데이터 전송조차도 느린 속도를 갖는다. 반면, UDP(User Datagram Protocol)는 실시간 전송은 가능하지만 비신뢰적이다. 이러한 이유로 조금은 비신뢰적이지만, 실시간 전송을 가능케 할 수 있는 RTP가 널리 사용되고 있다.

RTP는 UDP와 IP 멀티캐스트 상에서 오디오와 비디오 데이터의 실시간 전송을 제공하는 프로토콜이다. UDP와 RTP 프로토콜 모두 전송 프로토콜 기능의 일부분을 제공하지만, RTP는 다른 유용한 망 또는 전송 계층 프로토콜과 함께 쓰이곤 한다.

RTSP는 On Demand 형식으로 실시간 미디어 전송을 행하는 애플리케이션 계층의 프로토콜을 말한다. RTSP는 인터넷상에서 스트리밍 데이터를 제어하는 방법에 대한 표준안으로 스트리밍 기술이 사용하는 표준 프로토콜은 RTSP(Real Time Streaming Protocol)이다. 현재 테스트베드에서는 DESCRIBE, SETUP, PLAY, TEARDOWN 등 네개의 RTSP 메시지를 사용하고 있다[8].

RTSP 모듈이 각각의 RTSP 메시지의 수신과정에서의 동작은 다음과 같다.

**DESCRIBE:** 이 메시지를 수신하면 서버는 클라이언트와의 연결을 설정하고 클라이언트로부터의 RTSP URL에 근거하여 클라이언트가 원하는 입체 디지털 아이템이 서버에 존재하는지 검사하고 있을 경우 RTSP OK 메시지와 함께 SDP parser를 이용해 SDP를 뽑아낸 후 클라이언트에게 전달한다.

**SETUP:** 이 메시지를 수신 받으면 서버는 응답메시지와 함께 RTP로 입체 디지털 아이템을 전송할 서버와 클라이언트 포트번호를 설정한다.

**PLAY:** 이 메시지를 수신하면 해당 입체 디지털 아이템을 RTP packetize하여 SETUP과정에서 설정된 포트번호로 packet을 송신한다.

**TEARDOWN:** 이 메시지를 수신하게 되면 서버는 해당 클라이언트와 접속을 종료하고 할당된 자원을 해제한다. 또한 새로운 클라이언트의 접속을 위해 서버를 대기상태로 만든다.

그림 6은 RTP와 RTSP를 이용한 입체 디지털 아이템의 상기 전송 과정을 설명한다.

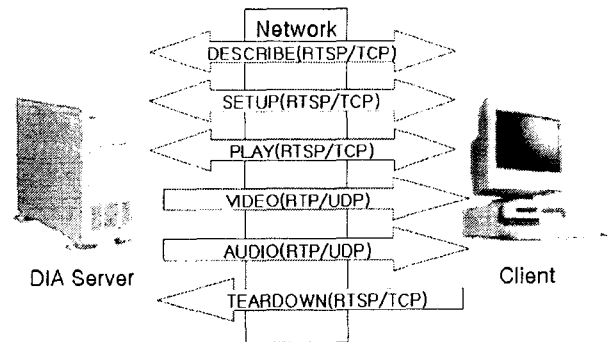


그림 6. 입체 DIA의 전송 과정

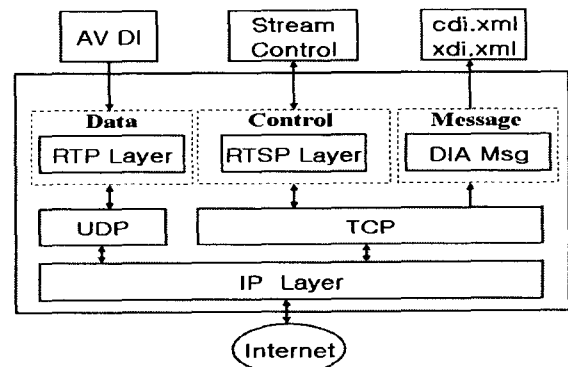


그림 7. 3D DIA의 Protocol Stack

그림 7은 3D DIA의 Protocol Stack을 나타내고 있다. DI의 리소스는 RTP Layer를 통해 UDP로 전송되고 리소스의 스트리밍을 제어하기 위한 RTSP 메시지는 RTSP Layer를 통해 TCP로 전송된다. 또한 디지털 아이템의 적응을 위한 DIA Negotiation message는 cdi.xml과 xdi.xml의 형태로 만들어져 TCP로 전송된다.

#### V. 테스트베드 구현

테스트베드는 크게 3D DIA와 디지털 아이템 전송 모듈로 구성되어 있다.

3D DIA는 MPEG-1 동영상을 디코딩할 수 있는 디코더와 2D-to-3D converter, 3D-to-2D converter, MPEG-2 encoder/MPEG-4 encoder, DIA Negotiation Message parser로 구성되어 있으며 디지털 아이템 전송 모듈은 RTP/RTSP 모듈로 구성되어 있다. 모듈들은 C/C++로 구현되었으며 MPEG-4 RTP packetize를 위한 모듈은 Darwin Streaming Server를 기반으로 해서 작성되었고 클라이언트는 MPEG-4IP의 wmp4player를 기반으로 제작되었다.

그림 8은 클라이언트에서 서버로 보낼 DIA Negotiation Message를 설정하는 화면이다[6]. 2D-to-3D Stereoscopic Conversion 선택 화면으로서 선택할 수 있는 항목은 StereoscopicVideoConversion DS인 ParallaxType, DepthRange,

MaxDelayed Frame가 있고 StereoscopicVideoDisplay DS인 DisplayDevice, RenderingFormat이 있다. 그밖에 MotionType과 VideoFormat이 있다.

그림9는 StereoscopicVideoDisplay DS의 RenderingFormat 설정에 따른 결과 화면이다. 왼쪽은 Anaglyph(Red-Blue) 영상이고 오른쪽은 Interlaced 영상이다. 그림 9는 StereoscopicVideoConversion DS의 Parallax Type의 설정에 따른 결과 영상이다. 왼쪽 영상은 Negative parallax로서 객체가 모니터 스크린과 관찰자 사이에 있는 음의 시차를 갖게 되고 오른쪽 영상은 Positive parallax로서 객체가 모니터 후면에 보이는 것처럼 느껴지는 양의 시차를 가지게 된다.

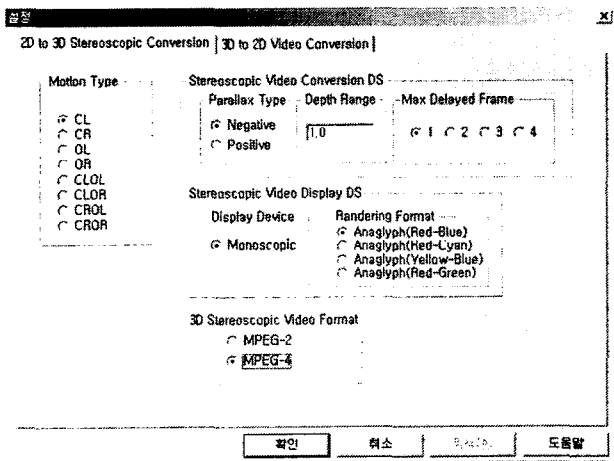


그림 8. 2D-to-3D Stereoscopic Video Conversion 선택화면

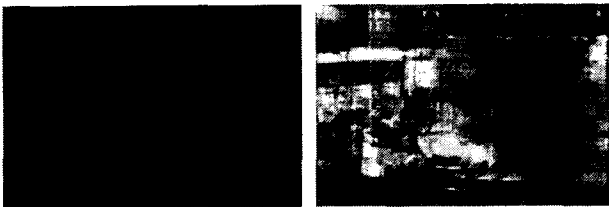


그림 9. RenderingFormat: Anaglyph 및 Interlaced

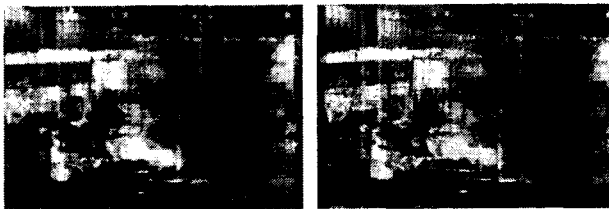


그림 10. Negative parallax, Positive parallax

그림 11은 DepthRange의 변화에 따른 결과 영상이다. ParallaxType은 Negative, RenderingFormat은 Interlaced, DepthRange는 각각 0.05, 0.35, 0.65, 0.95로 설정하였다. 마찬가지로 그림 12는 그림 11과 같은 DepthRange를 준 Anaglyph 영상이다.

그림 13은 MaxDelayedFrame의 변화에 따른 결과 영상이다. 값이 커짐에 따라 입체감도 커짐을 알 수 있다.

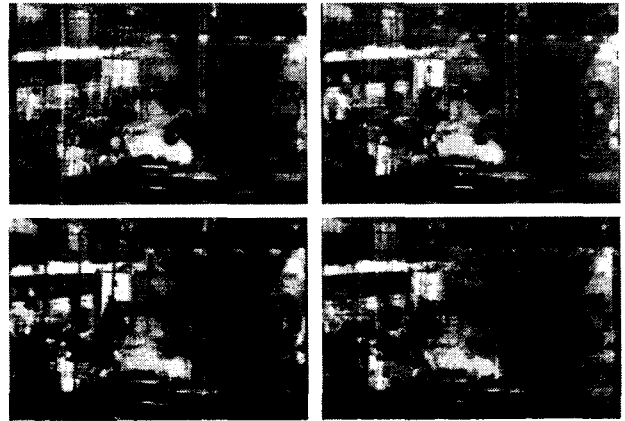


그림 11. Negative parallax, DepthRange 0.05, 0.35, 0.65, 0.95

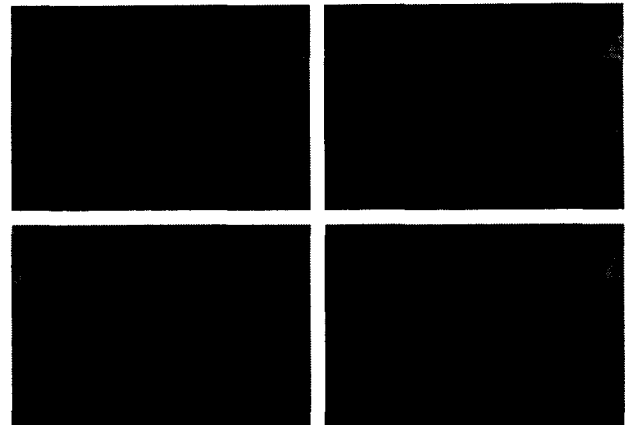


그림 12. Negative parallax, DepthRange 0.05, 0.35, 0.65, 0.95

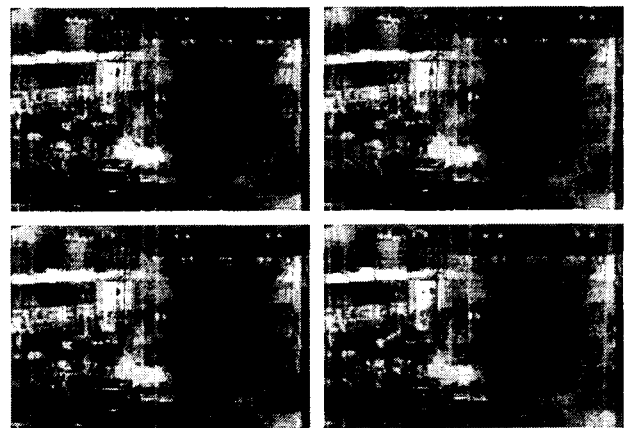


그림 13. MaxDelayedFrame = 1, 2, 3, 4

그림 14는 3D-to-2D Stereo Video Conversion을 위한 화면이다. 이 화면에서는 서버의 입체 동영상을 DIA에서 2D로 변환하는 과정에서 사용자 선호도를 선택할 수 있다.

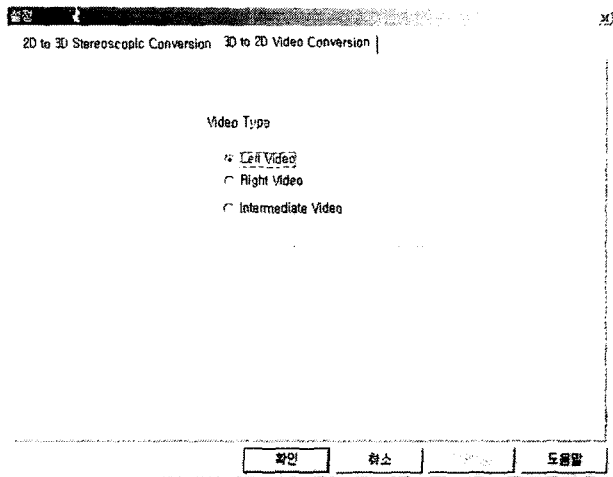


그림 14. 3D-to-2D Stereo Video Conversion 선택 화면

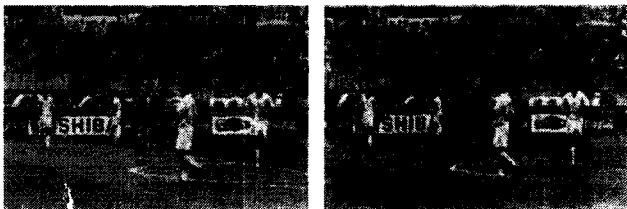


그림 15. 3D-to-2D 변환된 좌영상과 우영상

그림 15의 영상중 왼쪽 영상은 그림 14에서 Left Video를 선택하였을 때의 결과 영상이며 오른쪽 영상은 Right Video를 선택하였을 때의 결과 영상이다. 3D-to-2D 비디오 변환에서 Interlaced 영상의 경우 짝수 라인은 입체영상으로 합성되기 전의 좌영상의 값을 가지고 있고 홀수 라인은 우영상의 값을 가지고 있기 때문에 입체 영상에서 좌영상과 우영상을 추출해 낼 수 있다. 그러나 추출되어진 영상은 가로의 크기는 원래 영상과 같지만 세로는 1/2로 줄어들기 때문에 영상보간법이 필요하다.

그림 8 및 14에서 선택 되어진 항목들은 내부적으로 III절에 제시한 Description Scheme인 Stereoscopic VideoConversion DS와 StereoscopicVideoDisplay DS로 재구성되고 적용된 cdi.xml 및 xdi.xml로 저장되어 서버로 전송된다. 본 논문에서는 디지털 아이템의 Description Scheme은 xdi.xml에 저장되어 3D DIA로 전송되고 콘텐츠의 정보는 cdi.xml에 저장되어 3D DIA로 전송된다. 이 데이터를 받은 3D DIA는 XML parser를 이용해 정보를 추출한 후 해당 DI를 적용한다.

그림 16는 그림 8과 14에서 설정한 사용자 선호도에 따라 적용된 MPEG-4 DI의 디스플레이 화면이다. 서버는 사용자의 선호도에 따라 리소스와 기술자가 적용된 입체 DI를 RTP와 RTSP를 이용하여 전송하고 클라이언트에서 디스플레이 된다.

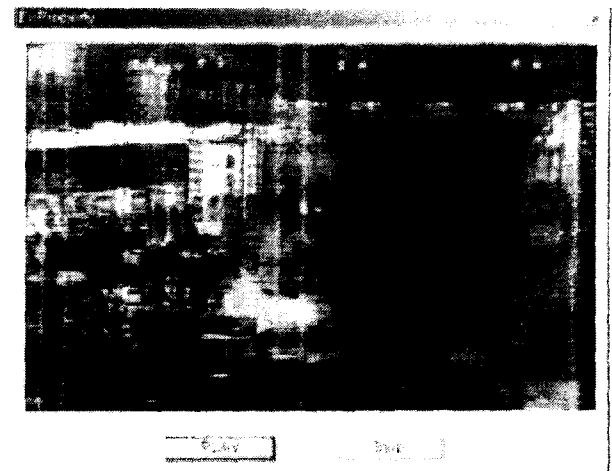


그림 16. 클라이언트에서 3D DI가 Display되는 화면

## VI. 결론

본 논문에서는 MPEG-21 Digital Item Adaption에 기반을 둔 Stereoscopic Video Conversion DS와 Stereoscopic Video Display DS를 사용하여 구동될 수 있는 MPEG-21 3D DIA를 구현하였다. 실시간 영상 전송을 위한 RTP, 제어를 위한 RTSP를 사용하여 사용자의 선호도에 따른 입체영상 디지털 아이템의 스트리밍 서비스가 가능하다.

## 참고문헌

- [1] "MPEG-21 Overview V.4", ISO/IEC JTC1/SC29/WG11 N4801, May 2002
- [2] ISO/IEC Draft of Technical Reports 21000-1, "Part1: Vision, Technologies and Strategy", MPEG/N4333, July 2001.
- [3] MPEG-21 Digital Item Adaptation CD 21000-7 ISO/IEC JTC 1/SC 29/WG 11/N5612 March 2003, Pattaya, Thailand
- [4] Man Bae Kim, Jeho Nam, Woonhak Baek, Jungwha Son, Jinwoo Hong, "MPEG-21 DIA Description for 3D Stereoscopic Video Conversion", IWAIT January 21-22, 2003, Nagasaki, Japan
- [5] MPEG-21 DIA Description Tools for 3D Stereoscopic Video Conversion, ISO/IEC/JTC1/ SC29/WG11, M8613, Klagenfurt, Austria, July 2002
- [6] 손정화, 손현식, 권혁민, 조영란, 김만배, "웹 환경에서의 MPEG-21 테스트베드의 구현", 정보통신설비 제 1권 제2호, 2002년.
- [7] IETF RFC(1889) RTP: A Transport Protocol for Real-Time Applications.
- [8] IETF RFC(2326) RTSP: Real Time Streaming Protocol (RTSP)