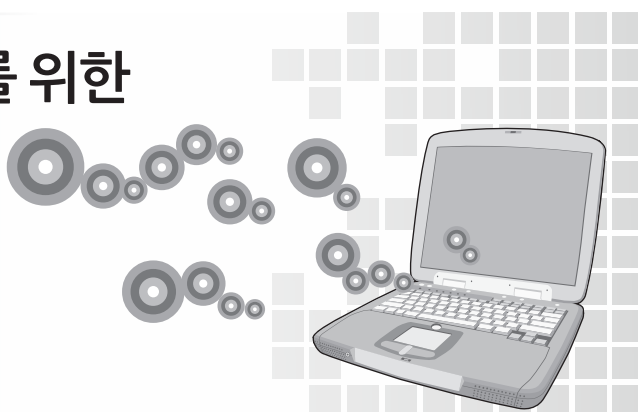


# 3차원 비디오 서비스를 위한 MPEG 표준화 기술

호요성 | 광주과학기술원 실감방송연구센터 교수

강민구 | 광주과학기술원 실감방송연구센터



## 1. 머리말

최근 방송과 인터넷, 개인 멀티미디어 장치의 발달로 고품질의 다양한 영상 콘텐츠에 대한 수요가 빠르게 증가하고 있다. 이러한 수요의 증가는 Full HD나 UHD(ultra high definition)와 같은 고해상도를 지원하는 TV 기술과 3차원 비디오 기술의 발달로 이어지고 있다. 이 두 가지 기술은 서로 결합된 형태의 실감형 멀티미디어로 서비스될 수 있으며, 기존의 2차원 비디오에서 제공하던 시청감을 초월하는 실재감(presence)과 몰입감(immersion)을 제공한다.

특히, 3차원 비디오 기술은 인간이 좌우 두 눈으로부터 서로 다른 시각 정보를 수용하여 3차원 깊이 정보를 처리하는 양안 시각(binocular vision) 시스템의 입체시(stereoscopic vision) 과정을 모방해, 시청자에게 공간감과 깊이감을 제공한다. 여기에는 3차원 영상 데이터를 획득/처리/전송/재생하는 기술이 포함되며, 실감형 방송 서비스를 위한 핵심 기술로서 차세대 디스플레이 시장을 주도할 것으로 전망된다. 3차원 비디오 기술에 대한 관심의 증가는 실감영상에 대한 높은 수요, 디스플레이 시장의 기능성 경쟁시대 진입, 디스플레이 제

작업체들의 수익 창출을 위한 마케팅 전략, 미국 할리우드를 중심으로 한 영화 배급사와 콘텐츠 제작업체들의 적극적인 시장 대응 등에서 비롯된다. 그러나 최근 개봉했던 3차원 영화 <아바타>에서 지적되었던 어지럼증, 멀미, 시각적 피로도 유발 등의 3차원 비디오 기술의 한계와 문제점들이 풀어야 할 과제로 남아 있다. 이를 위해서 3차원 비디오 서비스를 위한 안전 가이드라인 제시와 국제 표준화를 통한 기술의 이질성 완화 및 동질성 확보가 시급하다.

본 고에서는 MVC(multi-view video coding) 기술과 FTV(free viewpoint TV) 기술의 2단계로 진행되고 있는 ISO/IEC/JTC1/SC29/WG11(이하 MPEG)의 표준화 동향과 깊이영상 기반의 비디오 포맷을 이용한 최신 3차원 비디오 응용 기술을 설명한다[1]. 먼저 2장에서는 현재까지 진행된 MPEG 3DV 표준화 현황을 살펴보고, 3장에서는 앞으로 표준화가 진행될 3DVC(3D video coding) 기술의 테스트를 위한 실험 구조와 요소 기술에 대해 알아보겠다. 4장에서는 구체적인 3DVC 기술 제안을 위해 현재까지 논의된 제안요청서(call for proposals) 초안의 요구사항에 대해 기술하고, 5장에서 결론을 맺고자 한다.

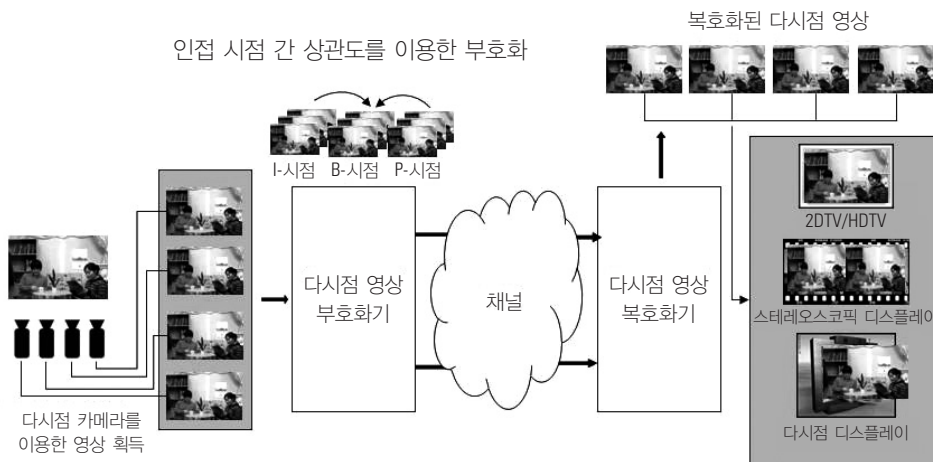
## 2. 3차원 비디오 관련 MPEG 표준화 작업

MPEG에서는 3차원 비디오와 관련해 활발한 표준화 활동을 진행하고 있다. 초기 MPEG에서는 시청자에게 입체감을 제공하기 위한 스테레오 영상의 부호화 기술에 대한 연구를 진행했다. 먼저 MPEG-2(ISO/IEC 13818-2) 표준에서는 좌우 스테레오 영상 간 시차를 보상하기 위한 MVP(multi-view profile) 기술을 개발했고[2], MPEG-4(ISO/IEC 14496-2) 표준의 제2부에서는 이진형상 정보, 변이 정보, 깊이 정보 및 텍스처 정보 등의 다중 부가정보를 포함하는 기술을 만들었다[3]. MPEG-4 제10부에서는 양안식 비디오에 대한 SEI(supplemental enhancement information) 메시지를 정의하는 기술을 포함했다[4]. 이후 새로운 3차원 AV(3-dimensional audio-video) 부호화 기술의 요구에 따라 2001년 12월부터 본격적으로 다시점 비디오 기반의 MVC 기술과 FTV 기술에 대한 연구가 시작되었다[5].

MVC 기술은 2004년 8월부터 ISO와 ITU가 공동으로 설립한 JVT(joint video team)를 중심으로 표준화 작업이 진행되었고, 다시점 영상의 보다 효과적인 부호화를 위한 다양한 알고리즘들이 개발되고 있다[6]. [그림

1]은 MVC 기술과 3차원 비디오 응용 서비스를 위한 시스템의 개념도이다. 먼저, 다시점 카메라를 이용하여 하나의 장면에 대한 여러 시점의 비디오를 촬영하고, 촬영된 영상의 효과적인 전송을 위하여 다시점 비디오 부호화를 수행한다. 수신단에서는 전송된 압축 비트스트림에서 다시점 영상을 다시 복호화하고, 복호화된 다시점 영상을 이용하여 시청자의 요구에 따라 다양한 응용 서비스를 제공한다. 이러한 시스템은 기존의 스테레오 영상과 같은 입체감을 제공할 뿐만 아니라, 시청자가 원하는 시점의 스테레오 영상을 선택하여 제공하는 장점이 있다.

FTV 기술은 사용자가 기존의 다시점 영상 사이의 임의의 시점을 선택하여 시청할 수 있는 기술로 처음 논의되었으나, 2007년 4월에 열린 JVT 회의 이후에 그 범위를 3DTV 응용을 위한 FTV/3DV 표준 시스템 모델로 확장하여 개발되고 있다. FTV/3DV 표준 시스템 모델은 [그림 2]와 같이 표현된다. 먼저, 수신단에서 다시점 색상 영상과 각 색상 영상에 대응하는 깊이 데이터(MVD: multi-view video plus depth)를 획득하고, 이를 부호화해 전송한다. 전송된 MVD 데이터는 복호화 과정을 거쳐 깊이영상을 이용한 가상 시점영상 생성(DIBR: depth



[그림 1] 다시점 비디오 부호화 기술(MVC)과 3차원 비디오 응용을 위한 개념도

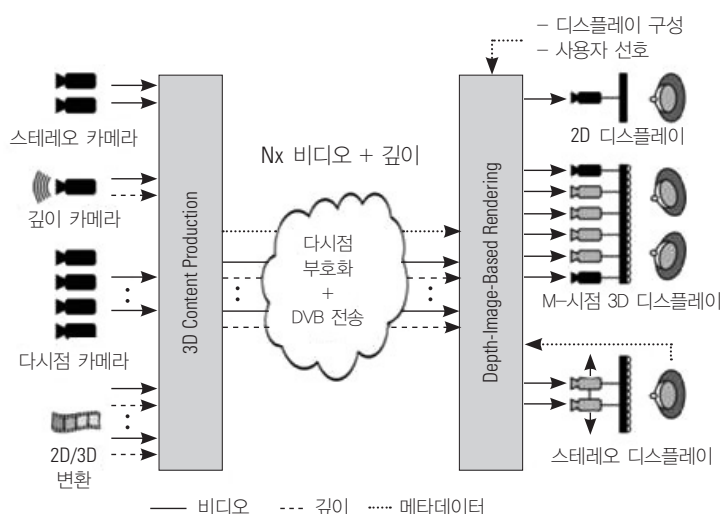
image-based rendering) 기술의 입력으로 사용된다. 최종 생성된 다시점 영상과 가상 시점영상들은 사용자 선호에 따라서 다양한 3차원 비디오 서비스에 제공된다[7]. 이러한 시스템은 제한된 입력영상으로부터 원하는 시점의 영상을 생성할 수 있는 장점이 있으므로 기존의 다시점 영상 기반의 시스템보다 압축과 렌더링 관점에서 효과적이다.

최근에는 3DVC 표준화 작업을 위한 실험환경을 구축하고, 실험영상 수집과 3DV 표준 시스템 모델에 포함된 요소 기술을 개선하는 선행 작업을 진행했다. 2008년 1월 회의에서 처음 3DVC을 위한 실험 영상을 요청했고, 2008년 4월 회의에서 10개의 실험 영상이 기고되었다. 또한 다시점 색상영상으로부터 깊이영상 생성을 위한 소프트웨어(DERS: depth estimation reference software)와 시청자가 원하는 임의 시점의 영상을 생성하기 위한 소프트웨어(VSRS: view synthesis reference software)가 여러 참여기관의 협조를 통해서 배포되었다[8]. 현재까지 반복적인 탐색실험(EE: exploration experiment) 결과에 대한 피드백으로서 기고된 실험영상들이 많이 개선되었으며, 보다 효과적인 알고리즘들

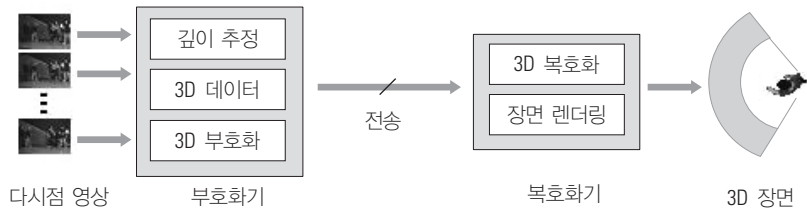
이 DERS와 VSRS에 구현되었다. 그리고 지난 2010년 4월 미팅에서는 6개의 연구기관에서 기고한 19개의 영상이 최종 실험영상으로 통과되었고, 3DVC 실험 구성과 최종 제안요청서에 대한 초안이 작성되었다[9]. 2010년 7월 회의에서는 아직 결정하지 못한 사항들을 논의하고 최종적으로 실험영상을 확정하여 2010년 10월에 최종 제안요청서를 배포할 것으로 보인다.

### 3. 3차원 비디오 부호화 기술의 테스트를 위한 실험 구조

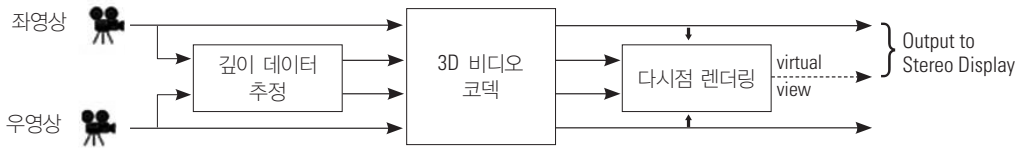
MPEG에서 진행 중인 3차원 비디오 기술은 궁극적으로 고급 양안(advanced stereoscopic) 영상의 처리를 위한 임의 개수의 중간시점 영상을 재구성해, 다시점(auto-stereoscopic) 디스플레이 장치를 지원하는 것을 목표로 한다. 이것을 위한 3DVC 실험 구조는 카메라 두 대 또는 세 대를 이용한 실질적 제작 환경에서 기대할 수 있는 제한적인 수의 영상을 입력으로 받고, 입력 영상으로부터 도출되는 3차원 데이터를 압축 부호화하고, 이것을 다시 복호하여 최종적으로 3차원 디스플레이



[그림 2] FTV/3DV 표준 시스템 모델



[그림 3] 3D 비디오 프로세스 체인



[그림 4] 두 개 시점의 카메라 구성을 가지는 고급 스테레오스코픽 실험 구조

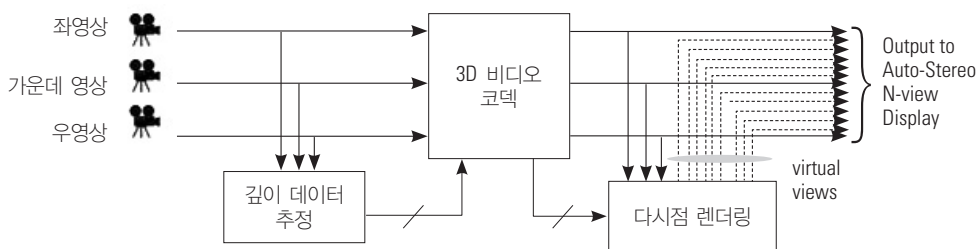
이에 나타내기 위한 중간시점 영상을 생성하는 과정으로 구성된다.

[그림 3]은 3차원 비디오 서비스를 위한 기본 개념도이다. 송신단에서는 실제 3차원 장면에 대한 다시점 색상 영상의 획득, 다시점 깊이 영상의 추정, 그리고 색상 영상과 깊이 영상이 결합된 형태의 3차원 데이터 부호화 등을 수행한다. 수신단에서는 부호화된 3차원 데이터를 복호하고, 복호된 영상들을 이용하여 3차원 장면을 렌더링한다.

MPEG에서는 크게 두 가지 3차원 비디오 응용 서비스를 위한 실험 구조를 가정한다. 먼저 3차원 장면에 대한 깊이감 조절이 가능한 고급 스테레오스코픽 응용 서비스를 위한 실험 구조로는 [그림 4]와 같이 두 개 시점의 카메라 구성을 가정한다. 이 구조에서 각 시점에 대한 깊이 영상은 두 개의 입력영상으로부터 도출되어야 하지만, 현재 MPEG의 부호화 실험에서는 참조 데이터 세트로서 보다 정교하고 고품질의 깊이 영상 추정을 위해서 두 개 시점 이상의 색상 영상을 사용한다. 깊이영상 추정이 끝나면 생성된 스테레오 색상 영상과 깊이 영상에 대한 부호화가 수행된다. 그리고 다시 복

호화 과정을 거쳐서 복호된 출력영상을 이용하여 처음 좌영상과 우영상 시점 사이의 임의의 가상시점 영상을 생성한다. 생성되는 가상시점 영상은 처음 좌영상 또는 우영상과 함께 실제 스테레오스코픽 디스플레이 장치에서 재현된다. 이때 선택되는 가상시점의 위치에 따라서 좌우 영상 간 거리, 즉 양안 시차가 결정되고, 시차의 정도에 따라서 수용되는 깊이감이 결정된다. 이러한 실험 구조에서 제안 기술의 부호화 성능과 3차원 표현 방법의 평가를 위해서는 복호된 영상으로부터 합성된 하나의 가상시점 영상과 복호된 원 시점의 영상이 필요하며, 스테레오스코픽 디스플레이를 이용하여 주관적 화질 평가가 실시된다.

N개 시점의 오토 스테레오스코픽 디스플레이를 지원하기 위해서는 [그림 5]와 같이 3개 시점의 카메라 구성을 실험 구조로 가정한다. 이러한 실험 구조로부터 보다 넓은 범위의 중간시점 영상 생성이 가능하고, 표현할 수 있는 깊이 범위 또한 증가하게 된다. 깊이 영상의 추정은 두 개 시점 카메라 구성과 유사한 방법으로 진행되며, 각 시점에 대응하는 색상 영상과 깊이 영상은 모두 부호화되고 수신단으로 전송되어 복호된다.



[그림 5] 세 개 시점의 카메라 구성을 가지는 오토 스테레오스코픽 실험 구조

또 복호된 출력 영상은 N개의 가상시점 영상을 합성하는데 사용된다. 이러한 실험 구조의 성능평가는 두 개 시점의 실험 구조와 유사하게 실시되며, 최근 실험에서는 9개 시점의 디스플레이 장치를 사용하여 주관적 화질 평가를 실시했다.

#### 4. 3차원 비디오 부호화 기술 제안을 위한 절차 및 제약사항

지난 2010년 4월 MPEG 회의에서 내부적으로 3DVC 기술 제안 요청을 위한 초안이 작성되었다[7]. 바야흐로 3차원 비디오 기술 선점을 위한 본격적인 표준화 작업이 시작되었다고 할 수 있다. 추후 회의를 통해서 최종 제안요청서가 작성되고 배포되면 제안요청서의 요구사항에 따라 각 기관에서 3DVC 기술을 제안하게 되고, 제안된 기술들이 전문가들의 주관적 화질 평가를 통해서 비교, 검증되어 표준 기술로서의 채택 유무가 결정된다. 제안된 기술들은 앞 장에서 설명한 두 개 시점의 카메라 구성을 가지는 고급 스테레오스코픽 응용 서비스와 세 개의 카메라 구성을 가지는 오토 스테레오스코픽 응용 서비스의 목적에 부합해야 한다. 제안된 기술의 공정한 성능 평가를 위해서, 공식적인 주관적 화질 평가가 수행되고 그 결과가 공개된다. 하지만 실험결과에 대한 평가에 앞서 제안 기술을 고려하는 어떠한 행동방침도 행해질 수 없다. 제안 기술을 검토

하기 위한 회의는 잠정적으로 2011년 3월에 열릴 예정이다. 기술 제안자는 이 회의에 참석하여 본인의 기술을 설명해 할 필요가 있다. 또한 이 회의에 대한 입력 문서로서 제안 기술을 설명하는 문서가 사전에 등록되어야 한다.

3DVC 기술 제안을 위해서 사용될 참조 영상 세트와 각 영상의 정보를 <표 1>에 정리했다. <표 1>에서 영상의 특성을 나타내기 위해서 영상의 해상도, 초당 프레임 수 그리고 촬영된 영상의 총 프레임 수를 표시했고, 카메라 배치 항목에서 촬영에 사용된 카메라의 수와 카메라 사이의 간격을 보였다. 참조 영상 세트에 포함된 모든 영상은 선형적 카메라 배치를 이용하여 촬영되었고, 영상의 시점 간 기하학적 교정(rectified)이 완료되어 오직 가로방향의 양안 시차를 가진다. 그리고 중간시점 영상의 합성과 렌더링을 위해서 각 시점에 대응하는 카메라 변수가 사용된다.

참조 데이터 영상의 부호화 환경은 역호환성의 만족 여부에 따라 정의되며, 모든 테스트 시나리오(실험 구조 1: 고급 스테레오스코픽 디스플레이, 실험 구조 2: 오토 스테레오스코픽 N-시점 디스플레이)에 대해서 그 결과물을 제출해야 한다. 또 인접한 영상의 시점 간 상관도를 이용하는 다시점 비디오 부호화를 위해서 MVC 부호화기(JMVC 8.0)가 색상영상과 깊이영상에 대해서 독립적으로 사용된다. 이 외에도 제안 기술의 공정한 평가를 위한 제약 사항은 다음과 같다.

### 제안 기술의 부호화 환경에 대한 제약 사항

- 입력영상에 대한 전처리(pre-processing)를 허용하지 않는다.
- 복호화 과정의 한 부분으로서 후처리(post-processing)는 허용된다.
- 양자화 과정은 정적으로 고정해야 하며, 변동이 있을 경우에는 이를 명시해야 한다.
- 부호화 과정에 필요한 최적의 매개변수의 수동 설정을 권장하지 않는다.
- 부호화 성능 평가를 위해서 사용되는 데이터는 엔트로피 부호화 테이블을 위한 트레이닝에 사용되지 않는다.

〈표 1〉 실험영상 명세서

제공기관	영상 이름	영상 특성	카메라 배치
Nagoya	Pantomime	1280x960, 30fps, 300 frames	80 cameras with 5cm spacing
	Champagne_tower		
	Dog		
	Kendo	1024x768, 30fps, 300 frames	7 cameras with 5cm spacing, moving camera array
	Balloons		
HHI	Book Arrival	1024x768, 16,67fps, 300 frames	16 cameras with 6,5cm spacing
	Leaving Laptop		
	Door Flowers		
	Alt Moabit		
ETRI	Lovebird1	1024x768, 30fps, 300 frames	12 cameras with 3,5cm spacing
	Lovebird2		
GIST	Newspaper	1024x768, 30fps, 300 frames	9 cameras with 5cm spacing
	Café	1920x1080, 30fps, 200 frames	5 cameras with 6.5cm spacing
Philips	Mobile	720x540, 30fps, 200 frames	3 cameras with 5cm spacing
	Beergarden	1920x1080, 25fps, 150 frames	2 cameras
Poznan	Poznan_Hall1	1920x1088, 25fps, 200 frames	9 cameras with 13.75cm spacing, moving camera array
	Poznan_Hall2		
	Poznan_Street	1920x1088, 25fps, 600 frames	9 cameras with 13.75cm spacing
	Poznan_CarPark		

## 5. 맺음말

지금까지 ISO/IEC/JTC1/SC29/WG11에서 진행하는 3차원 비디오 서비스를 위한 표준화 활동을 살펴보았다. 처음 스테레오 영상과 관련한 기술에서부터 MVC, FTV 등의 다시점 영상 기반의 기술, 그리고 깊이 데이터를

이용한 3차원 비디오 기술에 이르기까지 다양한 기술들이 MPEG 회의를 통해서 논의되고 발전되어 왔다. 최근 MPEG에서는 본격적인 3DVC 기술의 표준화 작업에 앞서 실험영상을 수집하고, 수집된 영상에 대한 다양한 탐색실험을 수행함으로써, 수집된 실험영상 및 관련 요소 기술을 개선하고 발전시키기 위해 노력해 왔다. 이

리한 충분한 준비를 통해 현재 본격적인 3DVC 표준화 작업을 진행할 수 있는 환경이 구축되었고, 2010년 4월 회의에서 3DVC 기술에 대한 제안기술 요청서의 초안이 작성되었다. 앞으로 한 두 번 정도의 회의를 통해 내용이 완성되면 최종 기술제안요청서가 배포되고, 본격적인 기술 경쟁이 이뤄질 것이다. 따라서 차세대 실감 방송 서비스의 주요 기술인 3차원 비디오 부호화 방식의 국제적 선점을 위해 국내 여러 기관들의 적극적인 MPEG 표준화 활동의 참여와 협력이 요구된다.

### [참고문헌]

- [1] 호요성, 이천, “3차원 비디오 부호화를 위한 국제 표준화 기술,” 한국방송공학회 학술지, 제14권, pp. 031-044, 2009. 06.
- [2] ISO/IEC JTC1/SC29/WG11, “Work Plan for Progression of Multi-view Profile,” N1081, Dallas, USA, Nov. 1995.
- [3] ISO/IEC JTC1/SC29/WG11, “Information Technology-coding of Audio-visual Objects Part 2: Visual,” N4350, Sydney, Australia, July 2001.
- [4] ISO/IEC JTC1/SC29/WG11, “Editors’ Text for ISO/IEC 14496-10:2005 (AVC 3rd Edition),” N7081, Busan, Korea, April 2005.
- [5] 엄기문, 방건, 허남호, 김진웅, “3D 비디오 MPEG 표준화 동향,” 전자통신동향분석, 제24권 제3호, 2009. 06.
- [6] 호요성, 오관정, “3차원 비디오 부호화 기술,” 한국통신학회지 정보와 통신, 제27권, 제3호, pp. 029-035, 2010. 03.
- [7] ISO/IEC JTC1/SC29/WG11, “Call for Contribution on 3D Video Test Material (Update),” N9595, Antalya, Turkey, Jan. 2008.
- [8] 호요성, 이천, “자유시점 TV와 3차원 비디오 국제 표준화 동향,” TTA Journal, 제116권, pp. 086-094, 2008. 04.
- [9] ISO/IEC JTC1/SC29/WG11, “Draft Call for Proposals on 3D Video Coding Technology,” N11272, Dresden, Germany, April 2010.

\*본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행되었음[NIPA-2010-(C1090-1011-0003)].

**TTA**