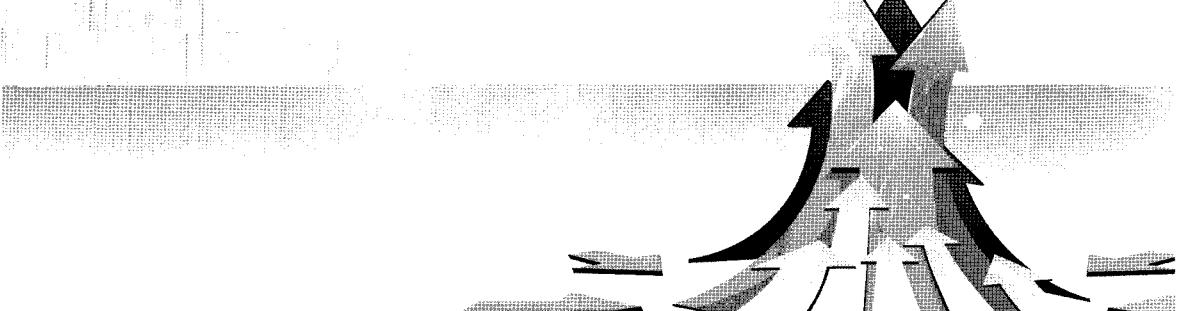


3차원 비디오 부호화 표준 기술

호요성 광주과학기술원 실감방송연구센터

이천 광주과학기술원 실감방송연구센터



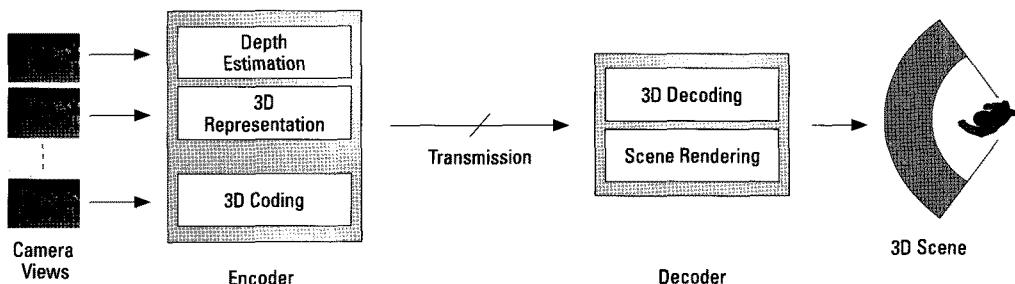
1. 머리말

지난 4월 IT 업계는 LG전자와 삼성전자의 3차원 디스플레이의 성능 공방에 대한 소식으로 뜨거웠다. LG전자와 삼성전자는 각각 독자적인 입체영상 재현 방법이 우수하다고 주장했다. 입체영상의 핵심 원리는 스테레오 카메라로 획득한 두 영상을 좌안(左眼)과 우안(右眼)에 독립적으로 비추어 주는 것이다. 이를 위하여 LG전자는 편광(偏光) 입체안경을 이용하는 방법을 이용하고, 삼성전자는 좌안과 우안을 번갈아 가며 닫았다. 여는 방식의 셔터글라스(shutter-glasses)를 이용하는 방법이다. 3차원 TV 전문가들은 공개 시연회에서 두 회사의 제품을 비교하여 각각의 장점과 단점을 지적했다. 결과적으로 어느 한 회사의 제품이 압도적으로 우세하다는 결론보다는, 각 제품의 특징을 정확하게 인식하게 되는 기회가 되었다. 아이러니 하지만, 이와 같은 두 회사의 갈등으로 인해 소비자들은 3차원 TV에 대한 관심이 더 높아져 3차원 TV의 매출이 더 높아졌다 는 후문이다.

이와 같은 흥미로운 사건을 통해 알 수 있듯이, 3차원 영상기술은 이미 우리의 생활에 가까이 다가와 있

다. 앞에서 언급한 3차원 TV뿐만 아니라, 영화 '아바타'의 흥행을 기점으로 3차원 영화가 연이어 흥행에 성공하고 있고, 3차원 모니터, 3차원 노트북, 3차원 휴대폰, 3차원 게임 등이 줄줄이 출시되고 있다. 앞으로는 이러한 3차원 영상 기술을 보다 다양한 분야에서 접할 수 있을 것으로 기대된다[1-3].

현재 주로 이용되고 있는 3차원 영상 기술은 두 대의 카메라로 획득한 영상을 스테레오 입체영상 장치에 재현하는 방법이다. 이 방법은 비교적 간단한 방법으로 입체영상을 제작할 수 있다는 장점이 있지만, 사용자의 위치나 재현장치의 크기와 같은 시청환경에 대한 조정이 어렵다는 단점이 있다. 이에 미래의 3차원 영상 기술은 사용자가 입체영상을 보다 편안하게 볼 수 있는 상호작용 기능을 포함할 것으로 예상된다[4]. [그림 1]과 같이, 이러한 기능을 가진 3차원 비디오 시스템은 국제 표준화 기구인 MPEG(Moving Picture Experts Group)의 3차원 비디오 부호화(3DV: 3D Video coding) 그룹의 주도로 중점적으로 개발되고 있다[5]. 특히 이 시스템은 사용자가 자유롭게 시청 시점을 변경하거나, 다양한 종류의 3차원 재현장치에 재현 가능하도록 다시점의 영상을 지원하는 기능을 가진다. 단 다



[그림 1] 3차원 비디오 시스템의 기본 구조

시점 영상의 이용으로 인하여 채널에 가중되는 부담도 시점 수에 따라 증가하게 되므로, 3차원 비디오의 효율적인 부호화 방법이 개발되어야 한다. MPEG 3DV 그룹에서 이에 해당하는 부호화 방법을 표준화하기 위해 2007년부터 표준화 활동을 진행하고 있다. 본 논문에서는 MPEG 표준화 그룹에서 진행한 3차원 영상 기술을 소개하고, 최근 진행하고 있는 3차원 비디오 부호화 기술의 주요 기능과 요소기술을 살펴본다. 특히, 최근 배포한 제안요청서(Call for Proposal)의 내용을 살펴봄으로 향후 표준화 활동을 소개한다.

2. 3차원 비디오 부호화를 위한 표준화 과정

2.1 3차원 오디오-비주얼(3D Audio-Visual)

MPEG에서는 2001년부터 오랜 기간 동안 3차원 비디오 시스템에 관심을 가져 왔다. 이미 2001년부터 3차원 오디오-비주얼(3D audio-visual, 3DAV) 그룹을 중심으로 다양한 형태의 3차원 영상 서비스 기술을 여러 선형실험(Exploration Experiment)을 수행했다[6]. 구체적으로 EE1에서는 전방향 비디오(omni-directional video)를 이용한 3차원 서비스를 연구했다. 이 연구는 최근 로드뷰(road view)와 같은 서비스로 상용화되어 이용되고 있다. EE2에서는 사용자가 시청하려는 시점을 자유롭게 선택할 수 있는 기능을 가진 자유시점 비디오(Free viewpoint video) 기술을 검토했다. EE3에서는 MPEG-4 MAC(Multiple Auxiliary Component)

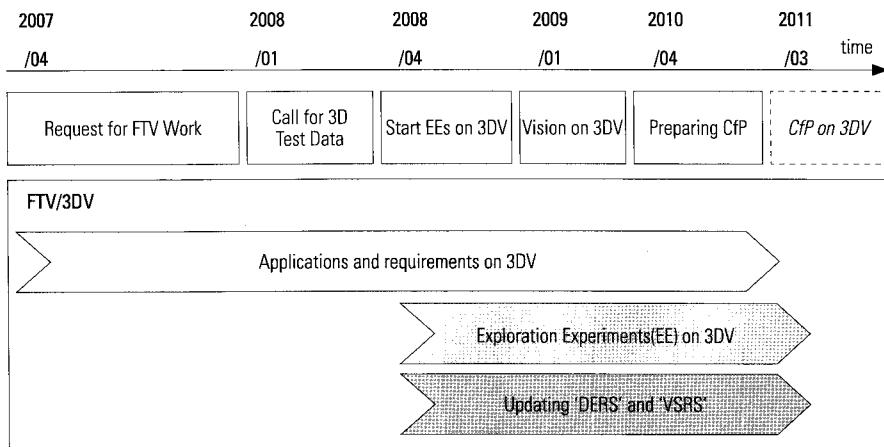
시스템을 기반으로 하는 양안식 비디오 부호화 방법을 실험했다. EE4에서는 깊이 또는 변이 정보를 부호화하기 위한 실험으로, 비디오와 깊이 영상을 부호화하는 방법이 주로 검토되었다.

2.2 다시점 비디오 부호화(Multi-view Video Coding)

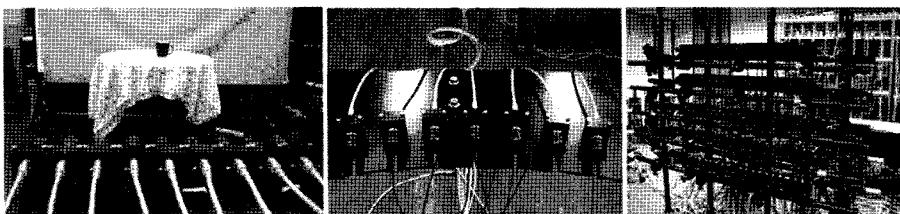
MPEG 표준화 그룹은 3DAV의 탐색실험을 수행하면서 3차원 영상 서비스에 구체적이고 실용적으로 이용될 수 있는 다시점 비디오 부호화(MVC: Multi-view Video Coding)의 필요성을 인지했다. 이에 2004년 8월부터 다시점 비디오 부호화에 관한 표준화 활동을 시작했다. 그리고 2006년 7월부터는 MVC에 관한 표준화 작업을 JVT(Joint Video Team)로 이양하여 수행했다. 다시점 비디오 부호화의 가장 큰 특징은 기존의 H.264/AVC와의 호환성을 유지하면서, 계층적 B화면(Hierarchical B frame) 부호화를 이용한 시점 간 참조 예측 구조이다[7]. 이 외에 시점 간 상관도를 고려하여 부호화 성능을 개선하기 위한 많은 기술들이 개발되었으나, 부호화 구조를 제외한 나머지 기술들은 부호화 성능이 요구조건에 이르지 못해 표준에서 제외되었다[8]. 이렇게 제정된 MVC 표준은 블루레이와 같은 산업 표준에 이용되고 있다.

2.3 3차원 비디오 부호화(3D Video Coding)

2007월부터는 MPEG 3DV 그룹에서 다시점 비디오에서 좀 더 개선된 3차원 비디오 시스템을 정의하고,



[그림 2] 3차원 비디오 부호화를 위한 표준화 과정



[그림 3] 다양한 다시점 카메라 배열

이에 시스템에 필요한 제반 기술을 개발하기 시작했다. [그림 2]는 이와 같은 과정을 간략히 정리한 것이다. 2008년 1월에는 3차원 비디오 부호화의 표준화를 위한 테스트 영상 및 관련 기술을 요청했다. 이에 대한 응답으로 2008년 4월부터 총 10개의 테스트 영상과 깊이영상 추정 소프트웨어와 시점 합성 소프트웨어가 수집되어 본격적인 탐색 실험(exploration experiments)이 수행되었다. 이 과정을 통해 각 테스트 영상의 평가와 시스템 구축에 필요한 제반 기술이 동시에 개발되었다. 그리고 2009년 1월의 스위스 로잔(Lausanne) 회의에서는 3차원 비디오의 비전을 발표하여 3차원 비디오 부호화의 개발 방향을 구체적으로 제시했다. 2008년 4월의 회의부터 탐색실험을 통해 획득한 깊이영상과 제반 참조 소프트웨어를 기반으로 제안 요청서(Call for Proposal)를 준비하기 시작했다. 그리고 지난 2011년 3

월 회의에서는 최종적으로 제안 요청서를 배포했다. 이에 98차 회의부터는 본격적인 3차원 비디오 부호화에 대한 표준화 활동이 시작될 것으로 보인다.

3. 3차원 비디오 부호화를 위한 탐색 실험

MPEG 3DV 표준화 그룹은 본격적인 3차원 비디오 부호화의 기술을 개발하기에 앞서 실험환경을 구축하기 위한 탐색 실험(exploration experiments)을 약 3년간 수행했다. 각 탐색 실험에 대한 내용을 간략히 살펴본다.

3.1 테스트 영상 획득

MPEG에서 개발하고 있는 3차원 비디오 시스템은 [그림 3]과 같이 다시점 카메라를 이용한다. 다시점 카메라를 이용하면 보다 많은 시점의 영상으로 사용자에

게 적합한 두 시점의 영상을 선택하여 재현할 수 있다. 하지만 여러 대의 카메라를 동시에 이용하기 때문에 카메라 사이의 불일치성을 보정해야 한다. 대표적으로 기하학적 불일치성(geometrical mismatches)을 보정하기 위한 영상 정렬(image rectification) 과정과, 카메라 간 색상 불일치성(color mismatches)을 보정하는 색상 보정(color correction) 과정 등을 수행해야 한다[9-10]. 현재 MPEG 3DV 그룹에서 사용하고 있는 테스트 영상은 모두 영상 정렬이 수행되었고, 대부분의 영상이 색상 보정이 완료된 영상이다.

3.2 다시점 비디오를 이용한 깊이영상 생성

깊이영상은 장면을 구성하는 객체의 3차원 거리정보를 나타내는 부가 데이터로서 3차원 비디오 시스템을 구성하는 중요 입력 데이터가 된다. 깊이영상은 적외선 센서를 이용하는 TOF(time-of-flight) 카메라를 이용하거나, 스테레오 정합(stereo matching) 방법과 같은 소프트웨어적인 방법으로 획득할 수 있다. TOF 카메라를 이용하면 실시간으로 비교적 정확한 깊이영상을 획득할 수 있으나, 비용이 많이 들고 낮은 해상도로 인한 제약이 있다. 그렇기 때문에 MPEG 3DV 그룹에서는 깊이추정 참조 소프트웨어(DERS: Depth Estimation Reference Software)를 개발하여 깊이영상을 획득할 수 있도록 했다. 여러 차례의 업그레이드를 통해 현재 DERS 5.1이 배포되고 있다. 현재 표준화 그룹에서 채택한 깊이영상은 DERS를 기본으로 이용하고 일부 수동적인 수정을 통해 얻어진 영상이다[11]. 이렇게 획득한 영상은 향후 3차원 비디오 부호화의 입력 영상으로 이용될 예정이다.

3.3 깊이영상을 이용한 임의시점 합성 방법

3차원 비디오 시스템이 다시점의 영상을 이용하기 때문에 전송에 필요한 채널의 확보가 가장 큰 문제가 된다. 또한 사용자가 직접 시청할 시점을 결정하거나, 보다 많은 수의 시점을 필요로 하는 3차원 재현장치의 경우는 전송된 시점보다 많은 시점의 영상이 필요하게 된

다. 이러한 문제를 해결하기 위한 방법이 깊이영상을 이용한 임의시점 영상 합성 기술이다. 흔히 깊이영상 기반 영상 렌더링(DIBR: Depth-Image-Based Rendering)이라고 알려진 이 기술을 이용하면 임의의 시점의 영상을 자유롭게 획득할 수 있다. MPEG 3DV 그룹에서는 이 DIBR 기술을 기반으로 한 영상합성 참조 소프트웨어(VSRS: View Synthesis Reference Software)를 개발하였다. VSRS 또한 여러 번의 알고리즘 업데이트를 통해 현재 VSRS 3.5가 주로 이용되고 있다[12].

VSRS에서 주로 이용하고 있는 시점 이동 방법은 크게 3차원 영상 워핑(3D image warping) 방법과 변위를 이용한 시점 이동이 있다. 3차원 워핑(warping) 기술이란 깊이값을 이용하여 원래 영상의 화소들을 실제 공간 좌표계로 반환하고, 이를 다시 가상 시점의 영상으로 재투영하는 과정을 의미한다. 그리고 변위 이동은 1차원 평행형 구조에서 변위를 이용하여 목표 시점으로 시점을 이동하는 방법이다.

3차원 영상 워핑이나 변위 이동을 이용하여 시점을 이동하면 원영상에서는 보이지 않았던 영역이 가상 시점에서 보이게 된다. 이런 영역은 원 영상에서 참고할 수 있는 정보가 없기 때문에 빈 영역(hole region)으로 남게 된다. 따라서 빈 영역을 처리하는 방법이 합성된 가상시점 영상의 품질을 결정한다. 빈 영역을 다른 참조 시점에서 참조 가능하면 그 정보를 참조하여 채우게 된다. 하지만 두 참조 시점 모두에서 참조할 수 없을 경우, 이 영역은 공통적인 빈공간이 되므로 인페인팅(inpainting)이나 주변의 깊이값을 고려하여 효과적으로 채운다[13-14].

4. 3차원 비디오 부호화를 위한 제안요청서 (Call for Proposal for 3DV)

지난 2011년 3월에 열린 제96차 MPEG 회의에서는 3차원 비디오 부호화를 위한 제안 요청서가 배포되었다 [15]. 이로 인해 세계 각국의 주요 연구기관에서 3차원

비디오 부호화에 대한 다양한 기술을 제안할 것으로 예상된다. 본 장에서는 제안요청서의 내용을 바탕으로 중요한 사항을 살펴본다.

4.1 목적과 절차

3차원 비디오 부호화의 목적은 고화질의 합성영상 을 위한 데이터 형식(data format)을 정의하고, 이에 대한 부호화 기술을 개발하는 것이다. 새로 정의될 3차원 비디오의 데이터 형식은 기준의 양안식 입체영상 재현장치에 이용가능하고, 다시점의 입체영상 재현장치에도 이용이 가능한 형식이다. 3차원 비디오 부호화 방법은 이 데이터 형식으로 생성한 데이터를 효율적으로 부호화할 수 있는 기술을 말한다. 이와 더불어, 3차원 비디오 부호화에 필요한 제반 기술을 고려하고, 필요할 경우 표준안에 포함할 수 있다.

3차원 비디오 부호화의 표준화 일정은 <표 1>과 같다. 2011년 4월 15일에 최종 제안요청서가 배포되고, 7월 20일까지 제안하는 기술에 대한 등록이 이루어진다. 그리고 10월 1일부터 제안된 기술에 대한 주관적 화질평가가 MPEG의 화질 평가기관의 주관으로 이루어질 예정이다. 이 평가 결과를 바탕으로 이어 열리는 제98차 MEPEG 회의에서 제안된 기술들이 평가될 예정이다.

<표 1> 3차원 비디오 부호화의 표준화 일정

일정	내용
2011년 4월 15일	최종 제안요청서 배포 및 기술 등록 시작
2011년 7월 20일	기술 등록 마감
2011년 8월 15일	기술 등록 취소 마감
2011년 10월 1일	주관적 화질 평가
2011년 11월 21일	제안서 등록 시작
2011년 11월 26일	제안된 기술 평가(98차 회의)

4.2 테스트 영상

3차원 비디오 부호화의 표준화 작업을 진행하기 위해서는 부호화 실험의 방법 및 환경이 정의되어야 한다. 3차원 비디오의 입력 데이터는 YUV4:2:0 8비트 형

식의 컬러 비디오와 깊이 비디오로 이루어진다. 테스트 영상은 영상의 크기 별로 클래스 A와 클래스 C로 구분한다. 클래스 A는 풀HD(Full HD)급 영상으로 구성된 영상이나, AVC 매크로블록의 크기가 16x16인 것을 고려하여 영상의 사이즈를 1920x1088로 확장하여 영상을 제공했다. 그리고 ‘Undo_Dancer’와 ‘GT_Fly’ 영상은 컴퓨터 그래픽스(computer graphics)를 이용한 영상으로, 정확한 깊이영상이 제공되고 있다. 클래스 C는 XGA급(1024x768) 해상도로 이루어진 영상으로, 모두 다시점 카메라로 직접 획득한 영상들이다. 각 테스트 영상의 깊이영상은 EE1(depth estimation)을 통해 결정된 데이터를 이용한다. <표 2>는 각 테스트 영상에 대한 정보를 정리한 것이고, [그림 4]는 각 테스트 영상을 보인 것이다.

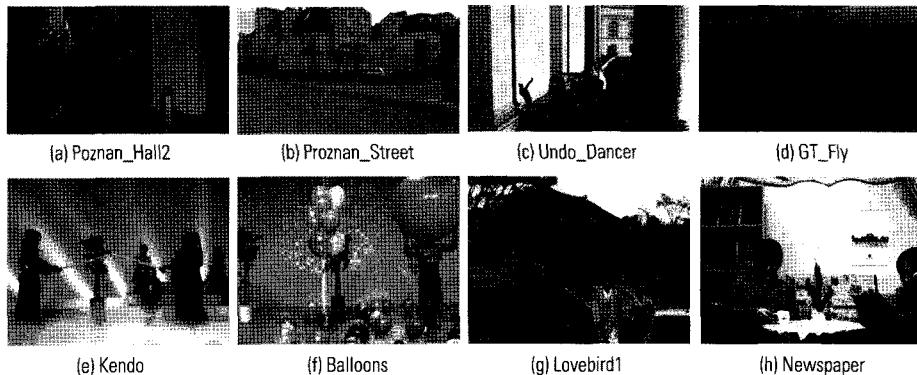
테스트 시나리오는 2-시점 카메라 구조와 3-시점 카메라 구조로 구분된다. 2-시점 카메라 구조의 실험환경은 일반적인 스테레오 재현장치에 이용 가능한 시나리오이고, 3-시점은 다시점 입체 재현장치를 고려한 시나리오이다. 각 시나리오에 따라 부호화 할 시점 수가 다르기 때문에, 부호화 과정에서 사용하는 예측구조도 다르다. 예를 들어, 2-시점 카메라 구조에서는 부호화 할 시점이 2개이므로 시점 방향으로 I-P 구조를 이용한다. 반면에 3-시점 카메라 구조는 부호화할 시점이 3개이므로 시점 방향으로 I-B-P 구조를 이용한다.

4.3 3차원 비디오 부호화의 실험 조건

MPEG 3DV 그룹에서는 다양한 방법의 3차원 비디오 부호화의 기술을 장려하고 있으나, 비디오 부호화의 성능만을 고려하기 위하여 몇 가지 제한사항을 기술했다. 먼저 입력 영상에 대한 선처리(pre-processing)를 금하고 있다. 하지만, 복호단에서 수행하는 후처리(post-processing)는 가능하도록 했다. 단 후처리 기술이 포함되었을 때, 그 기술에 대한 자세한 기술이 동반되어야 한다. 부호화 과정에서 이용하는 양자화 설정(table)을 고정하도록 했다. 이로써 각 영상별로 최적

〈표 2〉 3차원 비디오 부호화에 이용되는 테스트 영상

클래스	Seq. ID	테스트 영상	2-시점	3-시점	제공 기관	영상 크기
A	S01	Poznan_Hall2	7-6	7-6-5	Poznan 대학	1920x1088
	S02	Poznan_Street	4-3	5-4-3	Poznan 대학	
	S03	Undo_Dancer	2-5	1-5-9	Nokia	
	S04	GT_Fly	5-2	9-5-1	Nokia	
C	S05	Kendo	3-5	1-3-5	Nagoya 대학	1024x768
	S06	Balloons	3-5	1-3-5	Nagoya 대학	
	S07	Lovebird1	6-8	4-6-8	MPEG_KOREA	
	S08	Newspaper	4-6	2-4-6	광주과학기술원	



[그림 4] 3차원 비디오 부호화에 이용되는 테스트 영상

화된 양자화 방법을 이용하는 것을 봉쇄했다. 이와 더불어 각 영상별로 수동으로 설정된 부호화 파라미터를 이용하는 것을 금지했다.

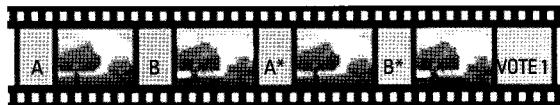
3차원 비디오 부호화는 두 가지의 부호화 표준에 호환되는 부호화 방법을 개발할 예정이다. 먼저 기존의 AVC(Advanced Video Coding) 부호화 표준과의 호환성을 유지하기 위한 'AVC-Compatible' 모드를 개발할 예정이다. 그리고 향후 JCT-VC(Joint Collaborative Team on Video Coding) 그룹에서 연구 중인 HEVC(High Efficiency Video Coding)와의 호환성을 가지는 'HEVC-Compatible & Unconstrained' 모드를 개발할 예정이다. 또한 이 모드에서는 기존 부호화 방법에 구애되지 않은 새로운 부호화 방법에 대한 가능성도 열어두고 있다. 다만, 각 기술에 대한 상세한 기술이 동반되어야 한다.

4.4 주관적 화질 평가

3차원 비디오 부호화 기술은 입력영상을 효율적으로 부호화해야 한다. 이 때, 부호화 성능을 비교하기 위한 주관적 화질 평가가 동시에 이루어질 예정이다. 주관적 화질평가는 MPEG 내의 화질평가 기관의 주관으로 수행될 예정이다. [그림 5]는 주관적 화질평가에서 사용할 화질테스트 방법을 설명한 것이다. A와 A*는 원본 영상이고, B와 B*는 제안한 방법으로 복원한 영상이다. 각각을 순서에 따라 보여주고 마지막에 평가하는 방법을 이용한다. 각 영상에 대한 평가는 0점(가장 나쁨)부터 10점(가장 좋음)으로 이루어진 등급으로 평가된다.

이와 같은 방법으로 평가된 결과는 제 98차 MPEG 회의에서 종합적으로 평가될 예정이다. 주관적 화질평가가 3차원 재현장치의 태입과 실험환경, 평가자에 따라 결과가 변동될 수 있는 여지가 있으므로 보다 객관

적인 평가 지표가 필요할 것으로 예상되므로, 향후 표준화 과정에서 이에 대한 보완이 이루어질 것으로 예상된다.



[그림 5] 주관적 화질평가 방법: 주관적 화질평가를 위한 영상 배열

5. 맷음말

3차원 영상 서비스가 시작되면 시청자는 보다 입체감 있고 실감나는 영상을 체험할 수 있을 것이다. 이러한 3차원 비디오 시대를 이끌 핵심 기술이 바로 3차원 비디오 부호화 기술이다. 본 고에서는 MPEG의 3D 비디오 부호화 그룹에서 진행하고 있는 3차원 비디오 부호화의 기술 표준화 동향을 살펴보았다. 이 표준화 활동에서 개발하고 있는 3차원 비디오 시스템은 사용자가 시청시점을 직접 선택하거나, 보다 편한 입체영상 시청을 위한 미래형 3차원 재현장치에 활용할 수 있는 기능을 포함한다. 특히 다시점 기반의 3차원 비디오 부호화 방법을 개발하고 있다. 주목할 점은 지난 2011년 3월에 열린 MPEG 표준화 회의에서 3차원 비디오 부호화 기술에 대한 최종 제안 요청서(Call for Proposal)를 배포한 것이다. 이로 인해 3차원 비디오 부호화의 국제 표준화가 향후 몇 년 안에 제정될 것으로 예상되며, 이렇게 제정된 압축 표준이 3차원 비디오 산업을 이끌 견인차 역할을 할 것으로 기대된다.

※ 본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행되었음(NIPA-2011-C1090-1111-0003)).

[참고문헌]

- [1] A. Kubota, A. Smolic, M. Magnor, T. Chen, and M. Tanimoto, 'Multi-View Imaging and 3DTV – Special Issue Overview and Introduction,' IEEE Signal Processing Magazine, vol.24(6), Nov. 2007.
- [2] A. Smolić, K. Mueller, P. Merkle, C. Fehn, P. Kauff, P. Eisert, and T. Wiegand, '3D Video and Free Viewpoint Video – Technologies, Applications and MPEG Standards,' IEEE ICME, pp. 2161–2164, 2006.
- [3] 호요성, 이천, '3차원 비디오 부호화를 위한 국제 표준화 기술,' 한국 방송공학회 학술지, 제14권, pp. 31–44, 2009. 06.
- [4] A. Smolić and P. Kauff, 'Interactive 3-D Video Representation and Coding Technologies,' Proceedings of the IEEE, Special Issue on Advances in Video Coding and Delivery, vol. 93, pp. 98–110, Jan. 2005.
- [5] ISO/IEC JTC1/SC29/WG11, 'Introduction to 3D Video,' MPEG document N9784; May 2008.
- [6] ISO/IEC JTC1/SC29/WG11, 'Report on 3DAV Exploration,' MPEG document N5878, July 2003.
- [7] ISO/IEC JTC1/SC29/WG11, 'Description of Core Experiments in MVC,' MPEG document N8089, July 2006.
- [8] ISO/IEC JTC1/SC29/WG11, 'Study Text of ISO/IEC 14496-10:200x/FPDAM 1 Constrained Baseline Profile and supplementary enhancement information,' MPEG document N10357, April 2009.
- [9] Y.S. Kang and Y.S. Ho, 'Geometrical Compensation Algorithm of Multiview Image for Arc Multi-camera Arrays,' LNCS, vol. 5353, pp. 543–552, 2008.
- [10] J.I. Jung and Y.S. Ho, 'Color Correction Method Using Gray Gradient Bar For Multi-view Camera System,' International Workshop on Advanced Image Technology, Jan. 2009.
- [11] DERS: http://wg11.sc29.org/svn/repos/MPEG-4/test/trunk/3D/depth_estimation/DERS/DERS
- [12] VSRS: http://wg11.sc29.org/svn/repos/MPEG-4/test/trunk/3D/view_synthetis/VSRS
- [13] A. Telea, 'An Image Inpainting Technique based on The Fast Marching Method,' Journal Graphics Tools, vol. 9, pp. 25–36, May 2004.
- [14] ISO/IEC JTC1/SC29/WG11, 'Common-hole Filling for Boundary Noise Removal in VSRS,' MPEG document M18514, Oct. 2010.
- [15] ISO/IEC JTC1/SC29/WG11, 'Call for Proposals on 3D Video Coding Technology,' MPEG document N12036, March 2011.