

# 3차원 비디오 부호화 기술의 국제 표준화 동향\*



호요성, 이상범 | 광주과학기술원 실감방송연구센터

## 1. 머리말

최근 디지털 영상처리와 컴퓨터 그래픽 기술이 발전함에 따라, 우리는 현실세계를 재현하고 이를 실감나게 경험할 수 있게 되었다. 하지만, 아직도 3차원 멀티미디어 응용 서비스를 널리 보급하기 위해서는 관련 기술의 개선 및 신기술의 개발이 절실히 요구되고 있다. 다시점 영상을 이용한 3차원 텔레비전(Three-dimensional Television: 3DTV)은 현실 세계를 재구성한 콘텐츠로 현실감있는 느낌을 사용자에게 제공할 수 있기에 차세대 방송 기술로 많은 관심을 끌고 있다[1][2]. 다시점 영상이란 기존의 단일 카메라를 이용하여 평면적인 장면을 획득하는 방법과는 달리, 인접한 여러 대의 카메라를 이용하여 동시에 여러 시점에서 획득한 영상을 말한다. 우리는 다시점 영상을 이용하여 사용자가 원하는 시점에서의 영상을 선택적으로 재생하거나, 인접한 시점의 두 영상 혹은 다시점 영상 전체를 이용하여 3차원 입체 모니터나 다시점 모니터를 통해 사용자에게 실감나는 생생한 3차원 영상을 제공할 수 있다.

3차원 비디오는 기본적으로 한 장면을 여러 시점에서 시청할 수 있는 기능을 제공하므로 다양한 분야에 효과적으로 이용될 수 있다. 예를 들어, 보다 실감나는 3차원 영상을 바탕으로 한 3차원 게임이나, 전시 또는 이벤트에 효율적으로 이용할 수 있으며, 한 사물을 다양한 각도에서 실시간으로 시청할 수 있는 점을 활용하면 보다 생생하고 재미있는 학습용 소프트웨어로도 이용될 수 있다. 의학적으로는 3차원 실감형 가상수술 교육 프로그램을 제작하여 보다 효과적인 교육을 수행할 수 있으며, 입체 영화나 입체형 TV 방송용으로 다양하게 이용할 수도 있다.

3차원 비디오 신호를 획득하기 위해 양안시점(stereoscopic) 카메라, 다시점(multi-view) 카메라, 깊이 카메라, 홀로그래피(holography) 등과 같은 장치들이 개발되었으며, 이를 이용하여 얻은 3차원 비디오 신호를 처리하고 부호화하는 기술들에 대한 연구가 활발히 진행 중이다. 특히, 두 대 이상의 카메라를 이용해 촬영된 다시점 비디오는 최근 3차원 영상처리의 핵심연구 대상으로 주목을 받고 있다[3]. 3차원 비디오는 사용자들에게 자유

\*본 연구는 지식경제부 및 정보통신연구진흥원의 대학IT 연구센터(ITRC)의 지원에 의한 것입니다.

로운 시점과 넓은 시야각을 통해 입체감과 임장감을 제공하는 장점을 지니고 있다. 하지만, 다시점 비디오의 경우 카메라 수에 비례하여 데이터양이 증가하고, 깊이 카메라의 경우 장비가 고가라는 문제점이 있기에 다양한 입체 영상 서비스 개발이 상당히 제한적이다.

멀티미디어 콘텐츠의 압축 부호화에 대한 국제 표준 규격을 제정하는 국제조직인 MPEG(moving picture experts group)에서는 2001년 12월부터 3차원 오디오 비주얼(3-dimensional audio-visual: 3DAV)이란 이름으로 3차원 영상처리에 관한 기술을 표준화하는 작업을 시작했다. MPEG에서는 MPEG-2(ISO/IEC 13818-2)와 MPEG-4(ISO/IEC 14496-2)에서 이미 평행형 다시점 영상 부호화와 3차원 메쉬(mesh) 압축 부호화 기능을 제공했지만, 좀더 구체적인 비디오 부호화의 필요성에 따라 3DAV 부호화 작업을 다시점 비디오 부호화(multi-view video coding: MVC)란 이름으로 바꾸고, 2006년 7월에 열린 MPEG 회의에서부터 MVC에 관한 표준화 작업을 JVT(joint video team)에서 본격적으로 맡아 진행했다.

최근 MPEG에서는 자유시점 TV(free-viewpoint television)에 대

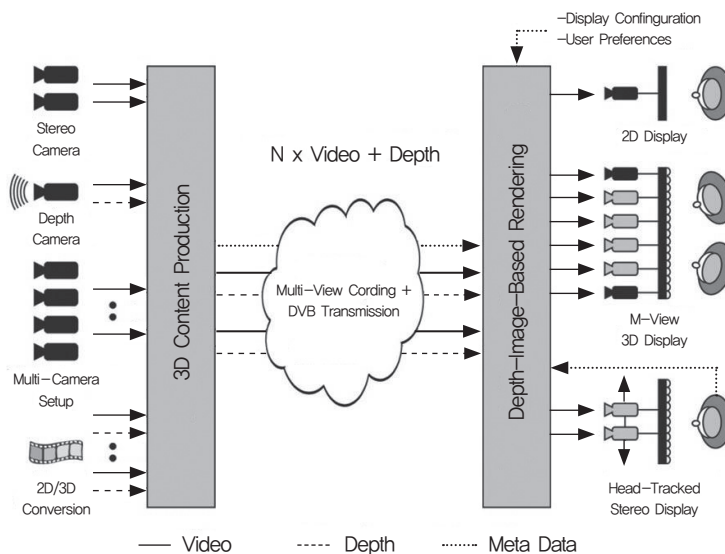
한 표준화 작업을 재개해 다시점 비디오와 깊이 영상을 기반으로 한 3차원 비디오에 대한 표준화 작업을 활발히 진행하고 있다. 본 고에서는 3차원 멀티미디어 서비스를 위한 3차원 비디오 부호화 기술의 국제 표준화 동향 및 표준화기구에서 다루는 주요 기술들을 전체적으로 살펴보고자 한다.

## 2. 3차원 비디오의 기술 동향

### 가. 3차원 비디오 부호화

다시점 비디오 부호화에 대한 표준화 작업이 주로 다시점 영상을 압축하는데 집중했는데, 3차원 비디오 부호화에서는 각 시점의 깊이 영상을 포함한 3차원 비디오의 획득, 처리, 전송 및 재생에 대한 모든 과정을 포함한다. 이는 3DAV를 위한 탐색 실험에서 수행한 자유시점 TV의 내용을 기반으로 다시점의 영상과 함께 다시점의 깊이 영상을 동시에 획득하고 처리하여 전송한다.

[그림 1]에 보인 것과 같이, 다시점의 영상은 지금까지 연구가 진행된 MVC 부호화 방법을 일부 이용한다.



[그림 1] 3차원 비디오의 개념도

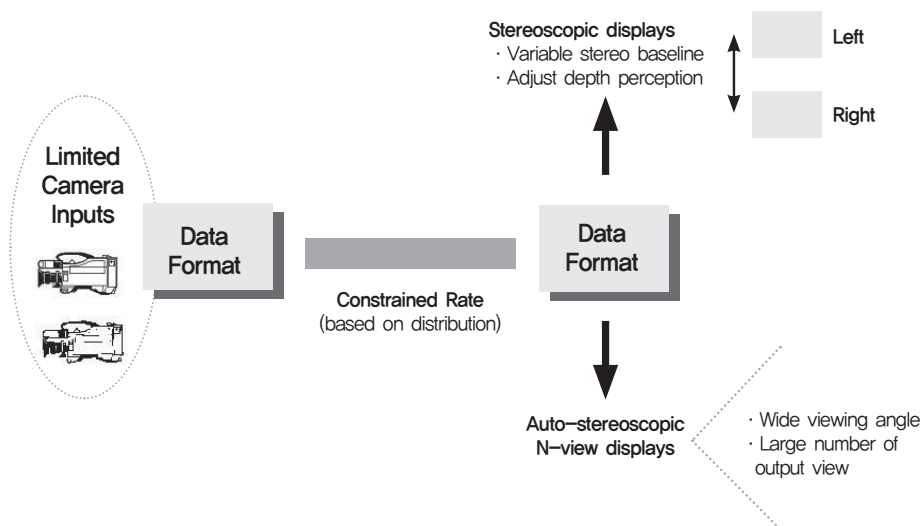
송신단에서 압축하여 수신단으로 전송된 색상 영상과 깊이 영상은 디스플레이 장치에 따라 여러 가지 방법으로 재현될 수 있다. 기본적으로 기존에 이용된 2차원 디스플레이 장치에서 재생할 수 있고, 전송된 다시점 중에서 사용자가 원하는 시점의 2차원 영상을 자유롭게 선택하여 시청할 수 있다. 또한, 다양한 종류의 3차원 디스플레이 장치를 이용하여 재생할 수 있다. 전송된 다시점 영상보다 적은 시점의 3차원 디스플레이 장치로 여러 시점의 입체 영상을 재생하거나, 시청자의 위치를 파악해 시청자의 위치에 맞는 시점의 입체 영상을 재생하는 장치(head-tracked stereo display)를 이용해서 재생할 수도 있다. 이러한 다양한 장치에 따른 여러 시점의 3차원 비디오를 재생하기 위해서는 복원한 다시점 영상보다 많은 수의 시점을 생성해야 하기에 수신단에서 임의의 시점에 대한 시점 합성(view synthesis) 기능을 포함해야 한다.

### 나. 3차원 비디오에 대한 비전

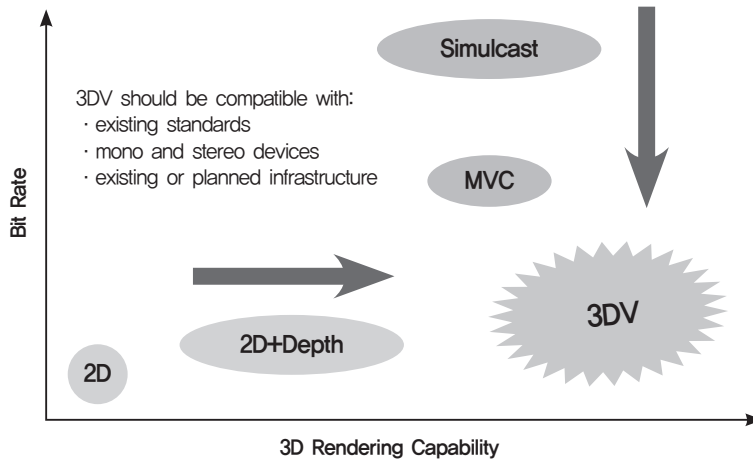
2009년 2월 MPEG 모임에서는 3차원 비디오 포맷에

대한 새로운 비전이 제시됐다[4]. 현재 새롭게 구상 중인 3차원 비디오 포맷은 현재까지 진행된 국제 표준의 기능을 넘어, 스테레오 디스플레이뿐만 아니라 다시점 디스플레이 장치까지도 지원할 수 있는 보다 진보되고 상호 작용이 가능한 기술을 포함한다. [그림 2]는 새롭게 제시된 3차원 비디오 포맷을 도식화한 것이다. [그림 2]에서 알 수 있듯이, 새롭게 제시된 3차원 비디오 포맷은 제한된 수의 카메라를 이용하여 스테레오 디스플레이와 다시점 디스플레이에 모두 사용할 수 있어야 한다.

현재까지 언급한 조건을 충족할 수 있는 3차원 비디오 포맷은 2D+깊이 영상과 다시점 비디오가 있다. [그림 3]은 두 포맷에 대해 비트율과 렌더링 성능을 비교하고 있다. 2D+깊이 영상은 단일 시점의 색상 영상과 이에 대응하는 깊이 영상을 이용하여 스테레오 영상 및 다시점 영상의 생성이 가능한 포맷으로서 현재의 비디오 포맷과 호환이 가능하지만, 시야각(viewing angle)이 좁고 폐색 영역(occlusion)을 처리하기 어렵다. 반면, 다시점 비디오는 다시점 카메라로부터 획득한 여러 시점의 영상을 이용하여 3차원 장면을 생성하는 포맷으로서,



[그림 2] 3차원 비디오 포맷



[그림 3] 2D+깊이 영상 포맷과 다시점 포맷의 비트율과 렌더링 성능에 대한 비교

2D+깊이 영상과는 반대로 넓은 시야각을 제공할 수 있지만, 데이터 양이 카메라 개수에 비례하여 증가하기 때문에 효율적인 부호화 기술이 요구된다.

### 3. 3차원 비디오의 주요 기술

#### 가. 다시점 깊이 영상 추정 기술

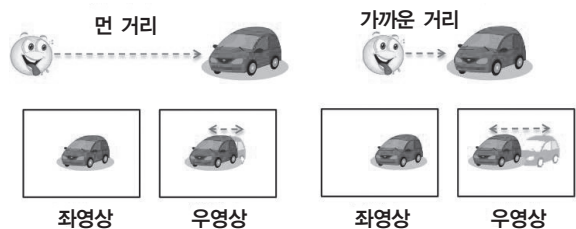
이상적으로 3차원 비디오의 자유시점 기능을 구현하기 위해서는 여러 시점의 영상을 이용해 시점 사이에 존재하는 가상의 중간시점 영상을 합성하는 중간시점 영상 합성 기술이 필수적이다. 중간시점 영상을 합성하기 위해서는 깊이 영상이 필요한데, 깊이 영상이란 영상 내에 존재하는 객체의 3차원 거리 정보를 나타내는 영상으로, 깊이 영상의 화소값은 해당 화소의 깊이 정보를 알려준다. 깊이 영상의 정확도는 합성된 중간시점 영상의 화질을 좌우하기 때문에 정확한 깊이 영상을 생성하는 것이 매우 중요하다.

두 대의 카메라가 있다고 가정을 하자. 또한, 이 두 대의 카메라는 정확하게 수평이 맞도록 정렬이 되어 있고, 또한 카메라의 방향이 정확하게 평행을 이룬다

고 가정한다. 이러한 카메라 구성에서 어느 한 영상에 존재하는 특정한 화소가 다른 시점의 영상에도 존재할 때, 이 두 시점에서의 좌표의 수평 차이를 변이(disparity)라고 한다.

[그림 4]는 변이와 깊이의 상관 관계를 나타내고 있다. [그림 4]에서 알 수 있듯이, 만약 객체가 카메라 가까이 있을 경우, 변이는 큰 값을 가지게 된다. 반대로, 객체가 카메라에서 멀어질수록 변이는 작은 값을 가지게 된다. 이러한 변이의 특성을 이용하여 우리는 장면의 깊이 정보를 얻어낼 수 있다.

2008년 1월 MPEG 내의 3차원 비디오 부호화 그룹에서는 다시점 비디오와 다시점 깊이 영상의 중요성을



[그림 4] 변이와 깊이의 상관 관계

인지하여 3차원 비디오 기술에 대한 성능을 평가하기 위한 테스트 영상을 요청했다[5]. 테스트 영상에는 3차원 비디오뿐만 아니라, 각 시점에 대응하는 깊이 영상도 포함된다. 이에 대한 응답으로 일본의 나고야대학에서 다시점 영상의 시점 간 상관 관계를 이용하여 깊이를 추정하는 소프트웨어를 구현했고 이를 배포하였다[6][7]. 이 기술은 가운데 시점 영상과 좌우 시점 영상의 화소값을 이용하여 대응하는 화소의 위치를 추정한 후 깊이 정보를 획득하는 방법을 이용한다.

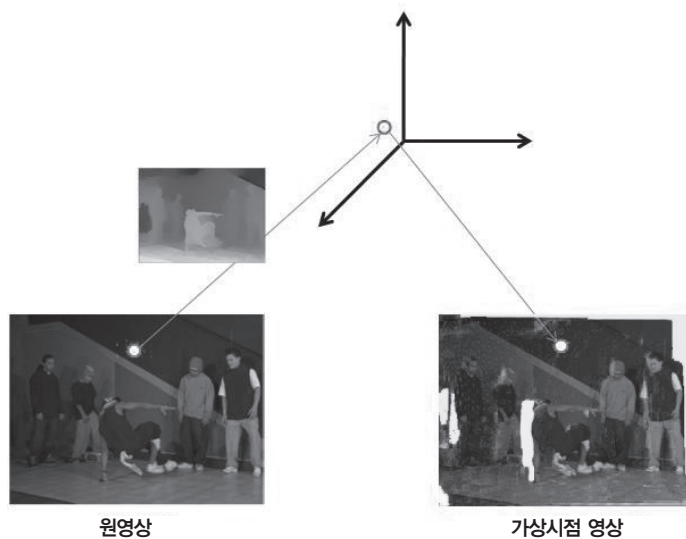
#### 나. 가상시점 영상 합성 기술

3차원 비디오의 자유시점 기능을 구현하기 위해서는 사용자가 원하는 위치의 모든 시점에 대한 정보를 가지고 있어야 한다. 하지만 카메라의 부피, 비용 등의 현실적인 문제가 존재하기에, 한정된 대수의 카메라를 이용하여 모든 시점에서의 영상을 획득하는 데에는 한계가 있다. 이러한 문제점은 여러 시점의 영상을 이용하여 시점 사이에 존재하는 가상의 중간시점 영상을 합성하

여 해결할 수 있다. 자유시점 기능을 가능케 하는 이러한 기술을 중간시점 영상 합성 기술이라고 한다.

중간시점 영상을 합성하기 위해서는 3차원 워핑 기술을 주로 이용한다[8]. 3차원 워핑 기술이란 깊이값을 이용하여 원영상의 화소들을 실제 공간 좌표계로 변환하고, 이를 다시 가상 시점의 영상으로 재투영하는 과정을 의미한다. 그러나 재투영하는 과정이 정확하게 1:1 사상이 아니기 때문에 기하학적인 차이로 인해 빈 영역(holes)들이 발생하게 된다. 따라서, 빈 영역들을 처리하는 방법이 중간시점 영상 합성 기술의 핵심이 된다. [그림 5]는 중간시점 영상을 합성하는 과정을 도식화한 것이다. [그림 5]의 가상시점 영상에서 흰색으로 된 부분이 기하학적인 차이로 인해 발생하는 빈 영역을 나타낸다.

일반적으로 다시점 영상을 이용하는 중간시점 합성 과정은 가상 시점을 기준으로 가장 가까운 좌우영상을 이용해 빈 영역을 채운다. [그림 6]의 (a)와 [그림 6]의 (b)에서 알 수 있듯이, 좌영상만을 이용해 시점 합성을 할



[그림 5] 중간시점 영상 합성 방법

경우에는 빈 영역이 객체의 왼쪽 영역에 존재하는 것을 알 수 있고, 우영상만을 이용할 경우에는 빈 영역이 객체의 오른쪽 영역에 존재하는 것을 알 수 있다. 따라서 한 쪽 시점의 결과를 이용해 다른 시점의 결과에 존재하는 빈 영역을 효과적으로 제거할 수 있다. [그림 6]의 (c)는 두 시점에 대한 최종적인 합성 영상 결과이다. 두 시점의 영상을 이용하면 빈 영역을 거의 채울 수 있다.

#### 4. 3차원 비디오의 국제 표준화 동향

2008년 1월 MPEG 미팅에서는 3차원 비디오 부호화 표준화를 위한 테스트 영상을 요청했다(call for test material: CFT) [5]. 이에 대한 응답으로 2008년 4월 미팅에서 4개 연구 기관에서 총 10개의 테스트 영상이 기고됐다[9]~[12]. 각 영상의 깊이 영상을 생성하는 소프트웨어와 중간시점의 영상을 합성하는 소프트웨어를 일본의 Nagoya 대학에서 제공하여 깊이 영상과 시점 합성의 정확도를 테스트했다[13]. 표준화 작업이 거듭될수록 소프트웨어의 성능은 향상되어 왔으나, 2009년 2월 MPEG 모임에서도 깊이 영상이 부호화에 적합할 수준의 정확도를 보이지 못하고 있다. 이를 위한 대안으로 표준화를 위한 테스트 영상으로서 신뢰할 만한 고화질의 깊이 영상을 다시 요청했다[14]. 이에 대한 응답으로 2009년 4월 MPEG

회의에서는 다양한 깊이 영상추정 기술들이 제안됐다. 제안된 기술들 중에서 자동으로 깊이 영상을 추정하는 기술들은 여전히 한계를 보였지만, 수작업을 거친 깊이 영상은 신뢰할 만한 화질을 제공했다[15].

2008년 1월 MPEG 회의에서는 3차원 비디오 테스트 영상과 함께 중간시점 영상의 합성 기술도 요청했다. 하지만, 앞서 언급한 깊이 영상의 화질 저하 문제가 합성된 중간시점 영상의 화질도 저하시키기에 이를 보완하기 위한 다양한 기술들이 제안됐다. 따라서 다양한 기술들이 통합된 작업 환경에서 실험하기 위해, 2009년 2월 MPEG 회의에서는 다양한 중간시점 합성 기술을 통합하자는 의견이 제시됐다[16]. 2009년 4월 MPEG 모임에서는 중간시점 합성 소프트웨어를 통합하는 작업이 진행 완료됐고, 이를 이용한 검증 실험을 진행했다.

3차원 비디오 부호화에 대한 표준은 깊이 영상 부호화를 포함한 다시점 영상의 부호화에 초점이 맞춰져 있지만, 앞서 언급한 깊이 영상의 화질 저하 문제로 인해 부호화에 대한 표준을 시작하지 못하고 있는 상황이다. 2009년 4월 MPEG 회의에서는 그 동안 진행한 탐색 실험(Exploration Experiment)을 통해 신뢰할 만한 깊이 영상을 제공하는 테스트 영상을 선별했고, 이를 이용한 깊이 영상 부호화 탐색 실험을 수행할 예정이다[17].



(a) 좌영상만 이용한 합성영상



(b) 우영상만 이용한 합성영상



(c) 양쪽시점 모두 이용한 합성영상

[그림 6] 중간시점 영상 합성 결과

## 5. 맺음말

본 고에서는 3차원 비디오 부호화의 국제 표준화 동향 및 주요 기술들을 살펴보았다. 현재까지 진행된 표준화 작업 중 가장 중점적으로 논의될 사항은 깊이 영상과 색상 영상을 포함한 모든 3차원 비디오 데이터를 부호화하는 것이다. 또한, 이미 표준화 작업이 완료된 다시점 비디오 부호화를 기반으로 깊이 영상을 포함하는 보다 진보된 형태의 부호화 방법이 논의될 것으로 전망된다. 현재 테스트 영상의 문제점으로 인해 3DV 표준화 작업에 난항을 겪고 있지만, 3차원 비디오 부호화에 대한 표준 제정이 빠른 시일 내에 이루어질 것으로 기대된다. 그렇게 된다면 3차원 방송과 입체영상 산업 역시 그 규모가 급격히 증가할 것이다. 차세대 방송 기술의 핵심이 될 3차원 비디오에 대한 국제 표준은 미래의 영상 산업을 국제적으로 주도할 수 있는 좋은 기회이다. 그러므로, 3차원 비디오 부호화에 관한 국내 학계와 산업계의 지속적인 연구와 정부 차원의 적극적인 지원이 뒷받침 된다면, 우리나라가 차세대 방송 기술을 선점할 수 있는 좋은 기회가 될 것이다.

### [참고문헌]

- [1] 호요성, 이천, “자유시점 TV와 3차원 비디오 국제 표준화 동향,” TTA Journal, vol. 116, pp. 86–94, 2008.
- [2] 호요성, 오관정, “다시점 비디오 부호화,” TTA Journal, vol. 115, pp. 93–100, 2008.
- [3] A. Smolic, K. Müller, P. Merkle, C. Fehn, P. Kauff, P. Eisert, and T. Wiegand, “3D Video and Free Viewpoint Video – Technologies, Applications and MPEG Standards,” in Proc. of IEEE International Conference on Multimedia and Expo, pp. 2161–2164, July 2006.
- [4] ISO/IEC JTC1/SC29/WG11, “Vision on 3D Video,” N10357, Lausanne, Switzerland, Feb. 2009.
- [5] ISO/IEC JTC1/SC29/WG11 “Call for Contributions on 3D Video Test Material,” N9595, Jan. 2008.
- [6] ISO/IEC JTC1/SC29/WG11 “Reference Software of Depth Estimation and View Synthesis for FTV/3DV,” M15836, Oct. 2008.
- [7] ISO/IEC JTC1/SC29/WG11 “Experimental Results on Improved Temporal Consistency Enhancement,” M16063, Feb. 2008.
- [8] W. R. Mark, L. McMillan, G. Bishop, “Post-Rendering 3D Warping,” in Proc. of Symposium on Interactive 3D Graphics, pp. 7–16, April 1997.
- [9] ISO/IEC JTC1/SC29/WG11 “Contribution for 3D Video Test Material of Outdoor Scene,” M15371, April 2008.
- [10] ISO/IEC JTC1/SC29/WG11 “1D Parallel Test Sequences for MPEG-FTV,” M15378, April 2008.
- [11] ISO/IEC JTC1/SC29/WG11 “HHI Test Material for 3D Video,” M15413, April 2008.
- [12] ISO/IEC JTC1/SC29/WG11 “Multiview Video Test Sequence and Camera Parameters,” M15419, April 2008.
- [13] SO/IEC JTC1/SC29/WG11 “Reference Softwares for Depth Estimation and View Synthesis,” M15377, April 2008.
- [14] SO/IEC JTC1/SC29/WG11 “Call for 3D Test Material: Depth Maps & Supplementary Information,” N10359, Feb. 2009.
- [15] ISO/IEC JTC1/SC29/WG11 “Semi-automatic Depth Estimation for FTV,” M16391, April 2009.
- [16] ISO/IEC JTC1/SC29/WG11 “Description of Exploration Experiments in 3D Video Coding,” N10360, Feb. 2009.
- [17] ISO/IEC JTC1/SC29/WG11 “Description of Exploration Experiments in 3D Video Coding,” N10552, April 2009.

**TTA**