



HEVC 3D Extension 표준 기술

I. 서론

차세대 방송 및 멀티미디어 시스템으로서 사실적이고 몰입감을 주는 실감 미디어에 대한 연구가 활발히 진행되고 있다. 보다 실감있는 영상 표현을 위해 2차원의 평면적인 정보를 뛰어넘는 3차원의 입체감을 줄 수 있는 연구가 현재 활발히 진행되고 있으며, 입체감을 제공해 주는 방식에 따라서 양안시점 기술 (Stereoscopy), 다시점 영상 기술 (Multiview), 라이트필드 기술 (Light-field), 그리고 홀로그래프 기술 (Holography) 등이 활발하게 연구 중에 있다^[1-4].

위 언급한 여러 3차원 기술 중에서, 안경 등의 장치를 이용하여 두 눈에 다른 영상을 보여주는 방식의 양안시점 기술은 사업화 및 보급화가 완료되어 있지

만, 보다 사실적이고 실감적인 3차원 입체감을 제공하기에는 부족한 점이 많이 있다. 이러한 단점을 극복하기 위해

다시점 영상 및 깊이영상을 동시에 처리하는 방식인 Multiview plus depth map (MVD)을 기반으로 한 3차원 비디오 표준화가 진행되고 있다.

보다 많은 시점의 영상을 처리하고자 자유시점 방식 (Free-viewpoint TV)에 대한 연구가 진행되었다. 이 과정에서 2006년 77차 MPEG회의에서부터 다시점 방식의 Multiview video coding (MVC) 표준화가 진행되어 완료되었으며^[5], 보다 넓고 다양한 수의 영상을 처리하기 위해 다시점 영상 및 깊이영상을 동시에 처리하는 방식인 Multiview plus depth map (MVD)을 기반으로 한 3차원 비디오 표준화가 2011년부터 현재까지 진행되고 오고 있다^[6-7]. 3차원 비디오 표준화에서도 HEVC의 표준화와 마찬가지로 ISO/IEC 산하 MPEG 그룹과 ITU-T 산하 VCEG 그룹이 공동으로 Joint



오 병 태
한국항공대학교 항공전자
및 정보통신공학부

Collaborative Team on 3D Video Coding Extension Development (JCT-3V)라는 이름으로 표준화를 진행하고 있다^[8].

본고에서는 우선 II장에서 보다 나은 이해를 위해 II장에서 다시점 기반 3차원 비디오 시스템의 소개 및 표준화 현황에 대하여 간략하게 알아본다. 그리고 III장에서는 HEVC기반 3차원 부호화 기술에서 채택하고 있는 여러 압축 기술에 대한 보다 세부적인 접근방법에 대하여 살펴보고, 마지막으로 IV장 결론으로 끝을 맺는다.

II. 3차원 비디오 시스템

3차원 입체감 표현을 하는 방식은 앞서 서론에서 살펴본 것처럼 여러 가지 방식이 존재한다. 하지만 현실적인 제약조건 및 기술적 한계와 현재 기술적 수준을 고려하여 MPEG그룹에서는 다시점 기반의 방식을 차세대 3차원 비디오 시스템으로 선정하고, 이에 대한 표준화를 진행하였다. 우선 기존 H.264/AVC 기반의 다시점 표준화 방식인 MVC 표준을 완료하였으며, 다시점에 깊이영상을 추가한 방식인 MVD 기반 방식에 대한 표준을 2009년부터 지금까지 진행해 오고 있다^[8]. 이번 장에서는 MVD 기반 3차원 비디오 시스템의 간략한 기술 및 이슈들을 소개하고, 이와 관련된 표준화 진행에 대하여 알아보도록 한다.

3차원 비디오 시스템의 궁극적인 목표는 여러 시점의 영상을 동시에 처리하여 시청자들이 보다 사실적인 3차원 입체감을 보여주는 것이다.

1. MVD 기반 3차원 비디오 시스템

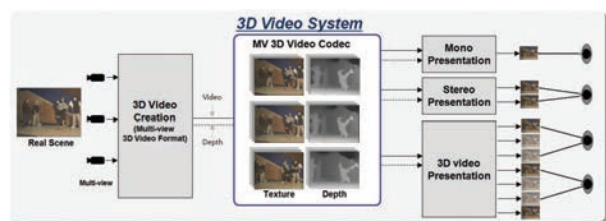
현재 활발하게 연구가 진행 중인 3차원 비디오 시스템의 궁극적인 목표는 여러 시점의 영상을 동시에 처리하여 시청자들이 보다 사실적인 3차원 입체감을 보여주는 것이다. 이와 같은 다시점 방식은 기존 MVC 기반 압축기술을 이용하면 처리할 수 있지만, 실제 수많은 시점(약 100시점 이상)의 영상을 빠르게 처리하고 전송하는 것은 영상의 획득 및 압축 관점에서 큰 어려움이 따르게 된다. 이러한 문제점을 해결하기 위해 제안된 방법이 깊이영상 기반 합성기술을 통해서 처리해

야 할 시점의 영상의 개수를 줄이는 방법이다. 즉, MVD 기반 방식이란 다수의 N개 시점의 영상을 처리하기 보다는 그보다 적은 K개 시점의 영상만을 처리해서 압축/전송하고 전송되지 못한 (N-K)개 시점 영상은 K개의 전송된 영상을 바탕으로 깊이영상 기반 합성방법 (Depth-image based rendering, DIBR)을 통해 재현하는 방식을 이용하고 있다. <그림 1>은 MVD 시스템의 일련의 과정을 보여주고 있다.

<그림 1>과 같이 MVD 기반 방식에서는 기존 H.264/AVC 혹은 HEVC등에서 다루었던 텍스처 영상과 다른 깊이영상 (depth map)이라는 새로운 형태의 입력이 추가되고, 깊이영상 기반 합성방식이 시스템에 추가됨에 따라, 여러 가지 기술적인 이슈가 많이 제기되었다. 일반적으로 3D 비디오 시스템에서 크게 고려하고 있는 기술적 문제는 1) 효율적인 깊이영상 획득방법, 2) 텍스처 및 깊이영상의 효율적인 압축방법, 3) 효율적인 깊이영상 기반 영상 합성방법이다. 본고에서는 이중에서 표준화 이슈에서 다루는 MVD 데이터의 효율적인 압축방법에 대하여 중점적으로 살펴보도록 하겠다.

2. MVD 포맷의 표준화 진행 현황

3차원 비디오 시스템 표준화를 시작했던 2011년에는 기존 2차원 영상 표준인 H.264/AVC 및 다시점 표준인 MVC가 이미 표준이 완료되어 사용되고 있었으며, 차세대 비디오 표준인 HEVC 표준에 대한 기술개발이 활발하게 진행되고 있었다. 이에 맞추어서 JCT-3V에서는 기존 AVC 및 MVC 뿐 아니라 곧 표준화 완



<그림 1> MVD 기반 3D 비디오 시스템



료를 앞둔 HEVC까지 모두 고려하여 동시에 표준화를 진행하였다. 즉 AVC 및 MVC기반 3차원 비디오 표준 기술 및 HEVC기반 3차원 비디오 표준기술이 동시에 개발된 것이다. 또한 각각 기술에 대한 표준기술을 개발할 때, 단순하게 2차원 영상을 3차원으로 확장한 간단한 방식의 표준도 고려하고, 확장할 때 보다 복잡한 기술을 넣어서 압축효율을 최대한으로 높이는 방식의 표준도 동시에 고려되었다. 이 결과 3차원 비디오 표준은 아래의 4가지 트랙을 동시에 진행하는 것을 목표로 하였다^[8].

- MVC extension for inclusion of depth maps (MVC+D)
- AVC compatible video-plus-depth extension (3D-AVC)
- Multiview HEVC extension (MV-HEVC)
- HEVC 3D extension (3D-HEVC)

위 4가지 트랙 중에서, MVC+D는 기존 MVC에 깊이영상을 추가하도록 확장한 방식이며, MV-HEVC는 기존 HEVC에서 다시점 영상 구조로 확장한 방식이다. 이 2가지 트랙 기술은 기존 기반 시스템에서 syntax 수준의 확장만을 목표로하여 표준화가 각각 2013년 및 2014년 초에 완료되었다. 반면, 3D-AVC는 AVC 기반으로 다시점 및 깊이영상의 효율적인 압축을 위해 보다 복잡한 기술들을 넣고 성능검증까지 완료하여 지난 2013년 10월에 표준화를 완료하였다. 유사하게, 3D-HEVC도 HEVC 기반으로 다시점 및 깊이영상을 위한 압축기술을 추가하여, 현재 MV-HEVC 보다 10~20% 정도 추가 성능향상을 이룬 상태이며, 다가오는 2015년 2월에 표준화 완료를 목표로하고 있다.

3. 3차원 비디오 시스템 압축효율 평가방법

기존의 2차원 영상 시스템과는 다르게 3차원 영상

시스템에서는 텍스처 영상 뿐 아니라 깊이영상이 추가되었기 때문에 압축효율을 평가하는 방법에 있어서 기존과는 많은 차이점이 있다. 2차원 영상에서는 비트량을 동등하게 맞춘 후 주관적 혹은 객관적인 화질을 평가하여 압축효율을 평가했던 반면, 3차원 영상에서는 부호화 하여 전송된 영상 뿐 아니라 복호화 이후 합성을 통해 생성되는 영상까지 화질평가를 해야 하기 때문에 평가방법 선정이 복잡해진다.

이러한 여러 복잡한 경우들을 고려하여 현재 JCT-3V에서는 화질평가지 텍스처 영상의 비트량만 고려하는 방법, 혹은 텍스처 영상과 깊이영상의 전체 비트량을 고려하는 방법 2가지로 나누어서 진행하고 있다. 화질 평가를 하는 방법에서도, 압축/전송된 텍스처 영상의 화질만 평가하는 방법, 합성을 통해 생성된 영상의 화질만 평가하는 방법 등 여러 가지 경우를 나누어서 평가하는 방법을 택하고 있다. 이때 주의할 점은 깊이영상의 화질은 고려하지 않는다는 점이다. 이유는 실제 깊이영상은 합성을 위한 부가정보로서 보내는 것이

며, 실제 사용자는 깊이영상 자체를 보게 되지 않기 때문에 깊이영상의 화질 평가는 합성영상의 화질을 통해서만 이루어지도록 정하였다. 위의 여러 가지 기준을 종합하여 3차원 비디오의 압축 성능은 아래의 3가지 기준을 바탕

3D-HEVC에 적용된 기술은 크게 의존 시점 영상(dependent view)에 적용된 압축 기술, 깊이영상에 대한 압축기술, 부호화단계만 적용되는 기술들이 있다.

으로 평가한다^[9].

- Texture 비트량 / Texture 평균 화질
- Texture+depth 비트량 / Texture 평균 화질
- Texture+depth 비트량 / 합성영상 평균 화질

III. 3D-HEVC 주요 기술

앞선 II장에서 MVD 기반의 3차원 비디오 시스템의 표준화 과정 중 HEVC 기반 최적의 압축효율을 목표로 하는 3D-HEVC 기술에 대하여 간략하게 소개하였다. 이번 장에서는 3D-HEVC에 적용된 기술들에 대해서

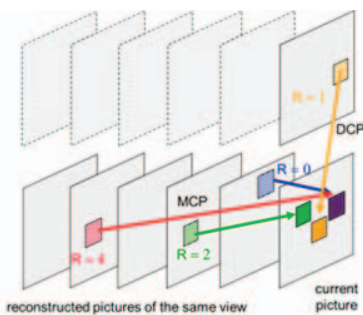
보다 자세하게 알아보도록 한다.

3D-HEVC에 적용된 기술은 크게 다음과 같이 3가지로 나누어 볼 수 있다. 첫 번째는 의존 시점 영상(dependent view)에 적용된 압축 기술이다. 3D 비디오 압축에서는 기존 HEVC와의 호환성을 이유로 독립 시점(base-view)에 대한 추가적인 압축기술은 금지하고 있다. 따라서 3D-HEVC에서 주로 목표로 하고 있는 부분은 의존시점에 대한 압축 기술이다. 두 번째로는 새롭게 추가된 깊이영상에 대한 압축기술이다. 깊이영상이 나타내고 있는 영상의 정보가 기존 텍스처 영상과는 크게 다르기 때문에, 이러한 점을 보완하기 위해서 여러 가지 깊이영상만을 위한 기술들이 소개되었다. 세 번째로 부호화단계만 적용되는 기술들이 있다. 이 기술들은 비표준화(non-normative) 기술로서 표준화 범주 밖의 기술들이지만, 기술 적용여부에 따라 다른 기술들의 영향을 많이 끼칠 수 있는 기술들은 표준화 회의에서 함께 논의되고 있다.

위 언급한 3가지에 대하여 다음의 3개의 절에서 보다 자세히 알아보도록 한다. 본고에서는 지나치게 세부적인 기술적 방식에 대한 언급은 지양하도록 하고, 각각의 새로운 접근방식에 대한 분석과 기존 HEVC방식과의 차이에 중점을 두고 기술하였다.

1. 의존시점 부호화 기술

의존시점 부호화 기술의 가장 기본적인 접근방법은

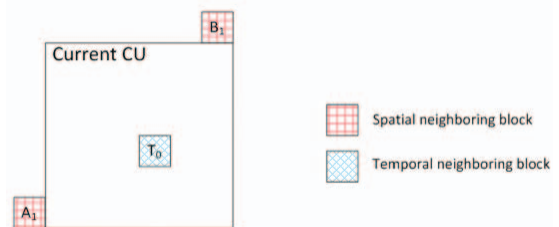


〈그림 2〉 MCP 및 DCP를 이용한 참조

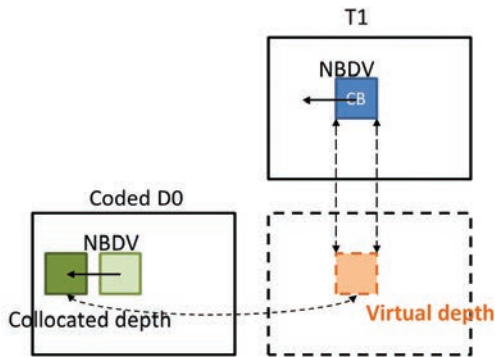
의존시점 부호화 기술의 가장 기본적인 접근방법은 의존시점이 의존하는 독립시점, 혹은 다른 의존시점 영상의 정보를 참조하는 방식이다.

의존시점이 의존하는 독립시점, 혹은 다른 의존시점 영상의 정보를 참조하는 방식이다. 기존 독립시점의 부호화 기술에서 가장 압축효율이 높은 방식은 시간적으로 앞뒤에 있는 영상을 참조영상으로 움직임 예측하는 Motion-compensated prediction (MCP) 이다. 의존시점 부호화에서는 이와 유사하게 동일한 access unit 내에서 부호화가 완료된 독립시점 (혹은 다른 의존시점)의 영상을 참조영상 시점간 차이를 예측하는 Disparity-compensated prediction (DCP) 방식을 사용하고 있다. 보다 자세한 시점간 참조방식은 〈그림 2〉에서 보여주고 있다. 이 과정에서 MCP를 위한 참조 영상 뿐 아니라 DCP를 위한 참조영상까지 동시에 고려를 해야 하기 때문에, 참조영상을 어떤 순서로 어떻게 구성하는지에 따라 압축효율에 영향을 미치게 된다. 이에 대한 보다 세부적인 내용은 Test Model 및 Draft text 문서를 참고하기 바란다^[10-11].

MCP에서 motion vector (MV)의 정보량을 줄이기 위해 주변 블록에서 MV 예측하여 motion vector prediction (MVP)값을 사용하는 것과 유사하게, DCP를 수행하는 과정 중 disparity vector (DV)를 예측하는 Derivation of disparity vectors 과정이 필요하다. 3D-AVC의 경우에는 의존시점에서 깊이영상이 텍스처 영상보다 먼저 부호화된다는 점을 이용하여 압축된 깊이영상의 정보를 이용하여 DV를 예측하는 방법을 사용하는 반면, 3D-HEVC의 경우는 항상 텍스처 영상부터 압축하게 되므로, 깊이영상의 정보를 사용할 수 없다. 이 점을 보완한 방식이 다음에 소개할 Disparity vector from neighboring blocks



〈그림 3〉 공간적/시간적 주변 블록의 위치



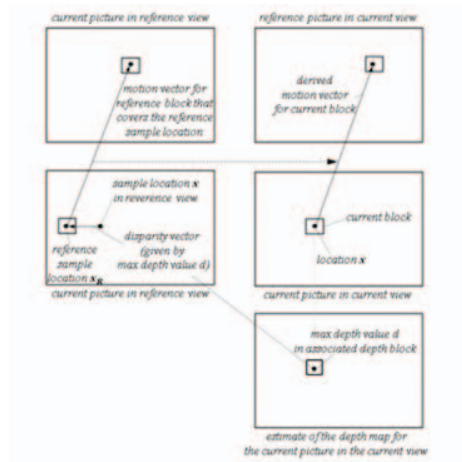
〈그림 4〉 DoNBDV 과정 개략도

(NBDV) 기술이다.

NBDV기술의 DV 예측 접근방법은 현재 부호화 하고자 하는 블록의 공간적 혹은 시간적 주변 블록에서 DV를 찾아오는 방법이다. 현재 3D-HEVC에서는 공간적/시간적 주변 블록으로 〈그림 3〉과 같은 위치의 블록을 탐색한다. 해당 주변 블록이 DCP가 사용되었을 경우 이때 사용했던 부호화된 DV를 참조할 수 있으며, 혹은 MCP가 사용된 주변 블록에서는 MCP가 참조하는 블록을 찾아 여기에서 사용된 DV값을 얻어오는 방식이다. 만약 주변 블록에서 모두 DV를 찾아오지 못할 경우에는 간단하게 zero벡터를 사용하게 된다.

위 언급한 NBDV기술은 깊이영상 없이 DV를 효과적으로 찾아내는 기술이지만, 실제 깊이영상의 정보를 이용하는 것이 아니기 때문에 정확도에 한계가 있게 된다. 이 단점을 극복하기 위해 제시된 기술이 바로 Depth-oriented NBDV (DoNBDV) 기술이다. 〈그림 4〉와 같이 의존시점의 DV를 찾는 과정에서 NBDV를 통해 얻어진 DV를 바로 사용하지 않고, 이 DV를 이용하여 이미 부호화된 다른 시점의 깊이영상의 실제 값을 참조하여 정확도를 높이는 기술이다. 이처럼 DoNBDV에서는 실제 깊이영상의 값을 직접 사용할 수 있게 되므로 의존시점에서 DV 예측 정확도를 보다 향상시킬 수 있다.

아직까지 효율적인 DCP를 위한 DV 예측과정에 대한 기술들을 소개하였다. 이 과정은 의존시점의 효율적인 압축에 있어서 중요한 과정임을 직관적으로 알 수 있다. 하지만, 위 과정을 조금만 변형한다면 위에서 언급

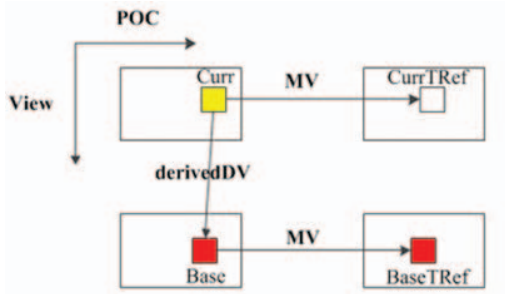


〈그림 5〉 주변 시점을 이용한 움직임 정보 유도과정

한 기술들과 유사하게 MCP를 위한 MV 예측도 향상시킬 수 있다. 개략적으로 설명하자면, 현재 의존시점의 MV를 예측하고자 할 때, 이미 부호화된 다른 시점의 현재 부호화 하고자 하는 블록과 대응되는 블록의 MV 값을 얻어오는 방식이다. 이 과정은 〈그림 5〉에서 보여주고 있다.

〈그림 2〉과 같이 각 블록마다 MCP와 DCP가 혼재되어 있고, 각각을 찾는 방법도 2차원 영상보다 복잡하게 구성되어 있다. 즉, 시간적 참조영상 뿐 아니라, 시점간 참조영상이 더해지면서 예측을 위한 후보군이 추가적으로 늘어나게 되었다. 이들 후보군들 중 최적의 후보를 선택하기 위해 Merge candidate list (MCL)를 효율적으로 구성하는 방법이 필요하다. 이를 위해 수차례 표준화 회의를 걸쳐 보다 효율적인 MCL 구성방법이 연구되어 사용되고 있다. 이 과정은 다소 복잡하여 본고에서는 생략하도록 한다. 자세한 내용은 Test model 및 Draft text 문서를 참고하기 바란다^[10-11].

마지막으로 본고에서 살펴볼 기술은 Advanced residual prediction (ARP)이다. 이 방법은 블록에서 MCP를 통해 1차로 얻어진 잔차신호(residual) 정보를 한 번 더 예측하여 정보량을 줄이는 방법이다. 이때 한 번 더 예측을 하기 위해 다른 시점의 정보를 이용하는 것이 해당 기술의 접근방법이다. 〈그림 6〉과 같이 우선 의존시점의 현재 블록을 MCP를 통해 예측을 한 후, 현재 블록에 대응되는 다른 시점의 블록을 DV를



〈그림 6〉 ARP의 예측과정 유도과정

이용하여 찾고, 해당블록의 MCP 예측 잔차신호를 바탕으로 현재 블록의 잔차신호를 한 번 더 예측하는 방식이다. 위 과정에서 MV와 DV의 순서를 바꾸는 방식 또한 물론 적용 가능하다.

아직까지 살펴본 모든 기술은 현재 3D-HEVC에서 적용하고 있는 다시점 구조를 이용한 방법으로서, 기본적으로 각각의 알고리즘들이 가지고 있는 근본적인 접근방법이 모두 유사함을 알 수 있다. 아직까지 살펴본 기술 외에도, 시점간 luminance 보정을 위한 Illumination compensation (IC) 기술, NBDV등으로 예측된 깊이정보를 이용하여 예측을 진행하는 View synthesis prediction (VSP) 기술, 깊이영상의 mask 정보를 이용하여 해당 블록을 임의의 2개의 영역으로 나누는 Depth-based block partitioning (DBBP) 기술 등이 현재 사용되고 있다.

2. 깊이영상 부호화 기술

본고에서 두 번째로 살펴볼 기술은 깊이영상에 적용된 기술이다. 기존 2차원 영상에서부터 많이 연구되어 온 텍스처 영상과는 다르게, 깊이영상은 영상을 획득하는 방법에서부터 영상이 표현하고 있는 정보까지 기존 텍스처 영상과는 확연히 다른 특징들을 가지고 있다. 〈그림 7〉에서 쉽게 보여지듯이, 깊이영상은 대부분 같은 물체에서는 대부분 유사한 정보를 가지는 homogenous 특징을 가지고 있으며, 전경과 배경, 혹

깊이영상은 대부분 homogenous하며 강한 에지정보를 가지고 있다. 이러한 특징을 위해 Depth Mode Modeling 방법, Segment-wise DC 방법, Motion Parameter Inheritance 기술들이 적용된다.

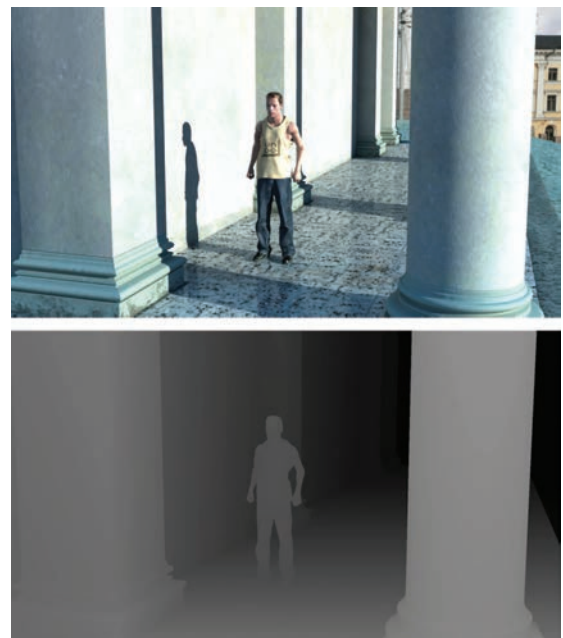
은 물체와 물체 사이에 강한 에지정보를 가지고 있는 특징이 있다. 특히 이 에지정보를 손실없이 잘 부호화하는 것이 합성에 중요한 역할을 하기 때문에 이 부분을 위한 여러 가지 새로운 압축 기술들이 연구되어 오고 있다. 이번 절에서는 깊이영상 기술 중 중요한 기술에 대하여 소개하도록 한다.

기존 HEVC에서도 효율적인 Intra 압축을 위해서 여러 가지 예측 방법을 제안하고 있다. 하지만, 깊이영상 압축에서는 이보다 더 다양한 경우의 수를 고려하는 Depth mode modeling (DMM) 기술을 적용하고 있다. DMM 기술은 초기에 DMM1 ~ DMM4의 4가지

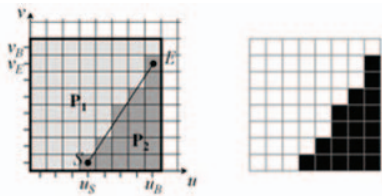
기술이 제안되었으나, 표준화 과정을 거치면서 현재는 DMM1과 DMM4의 2가지 기술만 적용되고 있다.

DMM1 기술은 직선으로 분리할 수 있는 모든 경우의 수를 탐색해 보고 최적의 블록 분할을 찾는 방식이다. 위 과정을 위해서는

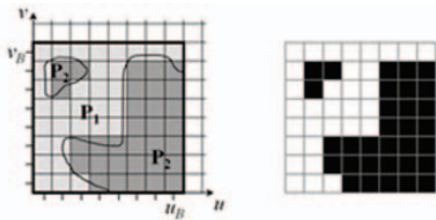
〈그림 8〉과 같이 해당 블록에서 직선이 시작하는 위치 및 끝나는 위치 등의 정보가 추가로 압축/전송해야 한



〈그림 7〉 Undo Dancer 영상의 텍스처 (위) 및 깊이영상 (아래)



〈그림 8〉 DMM1의 블록 분할 방법



〈그림 9〉 DMM4의 블록 분할 방법

다. 즉 DMM1에서는 기존 방법에 비해 부가정보를 늘려 최대한 잔차신호를 줄이는 방법이다. 이 과정에서 블록 분할의 모든 경우의 수를 전부 탐색하는 과정은 시간적 복잡도가 너무 커서, 이를 줄이는 여러 가지 기술도 추가로 제안되었다.

DMM4는 분할을 위해 직선을 사용하는 것이 아니라, 〈그림 9〉와 같이 현재 블록에 상응되는 텍스처 블록을 우선 분할하고, 분할된 경계를 이용하여 깊이영상 블록을 분할하는 방식이다. 이 경우는 복호화 단에서 텍스처 블록의 경계를 찾아야 하는 복호화의 복잡도 증가를 가져오지만, 부가정보의 양을 줄일 수 있으며 보다 다양한 형태의 분할이 가능해져 압축효율의 향상을 기대할 수 있다.

DMM1 및 DMM4에서 블록 분할이 완료되면 각각의 영역은 DC 모드와 같이 Constant partition value (CPV)라 불리는 1개의 대푯값으로 예측된다. 이 경우 해당 예측값은 블록의 주변값을 이용하여 예측되는데, 실제 평균값과 예측값이 상이할 경우를 대비하여 그 차분값 (ΔCPV)을 추가로 부호화하여 전송하는 방법도 채택하고 있다.

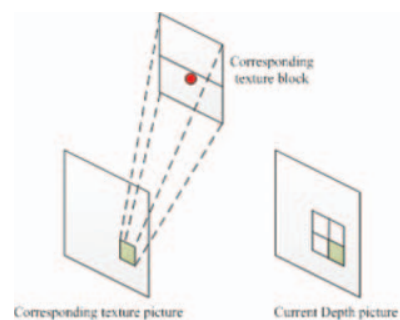
DMM과 더불어 또 한가지 중요한 깊이영상 압축기술

깊이영상의 왜곡을 측정할 때 기존과 달리 왜곡된 깊이영상을 바탕으로 합성영상을 생성해보고, 이를 왜곡없이 합성된 영상과 비교하는 View Synthesis Optimization (VSO) 방식을 적용한다.

은 Segment-wise DC (SDC) 압축 방법이다. SDC 방법에서는 잔차신호를 압축할 때, 기존 방식대로 transform 및 quantization을 진행하지 않고, 간단하게 각 예측 분할 영역을 하나의 DC값으로 대체한 후, Depth look-up table (DLT)을 사용하여 quantization 없이 적은 비트로 압축을 할 수 있다. 비교하자면, DMM이 깊이블록의 예측에 관한 기술이었다면, SDC는 잔차신호를 압축하는데 깊이영상의 특징을 이용한 방법이라고 생각할 수 있다. 초기 SDC 기술은 Intra mode에서만 적용되었지만, 현재는 Simplified inter-mode depth coding (SIDC)이라는 이름으로 inter mode에까지 유사한 개념의 기술이 적용되고 있다.

Motion parameter inheritance (MPI) 기술은 깊이영상의 MV를 찾을 때 적용되는 기술이다. 앞서 III-1절에서 텍스처 영상의 MV를 효율적으로 찾는 방법에 대하여 살펴보았는데, 이 MPI 기술은 깊이영상에서 효율적으로 MV를 찾는 기술이다. 세부적인 접근방법을 살펴보면, 해당 깊이영상 블록의 MV를 찾을 때, 이와 상응되는 텍스처 영상의 MV를 재활용 하는 방법이다. 이때, 해당 깊이영상의 PU와 상응하는

텍스처 영상의 PU 크기가 다를 수 있는 경우, 〈그림 10〉과 같이 중간 위치에 있는 MV값을 참조한다. MPI를 통해 얻어진 MV는 merge list에 첫 번째 위치로 할당되어, 다른 경쟁 MV 후보군과 경쟁하여 선택된다.



〈그림 10〉 MPI의 MV정보 참조 방법

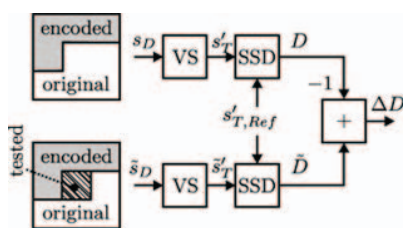
위 소개한 기술 이외에도 의존시점 깊이영상 블록에서 DV를 찾을 때 깊이영상 정보를 이용하는 Disparity derived depth (DDD) 기술, 깊이영상의 블록 분할 시 텍스처 영상의 분할정보를 이용하는 Depth quadtree prediction (DQP) 기술 등이 현재 사용되고 있다.

3. Encoder 제어 기술

III-1 절 및 III-2 절에서 언급한 기술은 모두 encoder 및 decoder 단에 동시에 적용되는 표준화 기술인 반면, 이번 절에서는 encoder에서만 적용되는 기술에 대하여 살펴보도록 한다.

HEVC 등에서 사용되는 가장 중요한 encoder 제어 기술은 Rate-distortion (R-D) 비용 함수를 통한 모드 결정방법이다. 이 과정은 여러 가지 후보 압축기술 중 한 가지를 선택할 때 적용되는 방법으로서, 압축효율에 아주 큰 영향을 미치게 된다. 3D-HEVC에서도 마찬가지로 R-D 비용함수를 사용하게 되는데, 이때 깊이영상의 경우 새롭게 distortion을 정하여 사용하는 방식을 적용하고 있다.

앞서 언급하였듯이, MVD 시스템에서 깊이영상은 새로운 시점의 영상을 합성하기 위하여 사용되는 부가정보로서, 실제 시청자에게는 보이지 않는 정보이다. 이러한 점 때문에, 깊이영상의 왜곡을 측정할 때, 기존의 MSE와 같이 각각의 픽셀 luminance의 차이를 이용한 왜곡함수 결정은 큰 의미가 없게 된다. 이와는 달리, 왜곡된 깊이영상을 바탕으로 합성영상을 생성해보고, 이를 왜곡없이 합성된 영상과 비교하여 distortion을 측정하는 방식이 바로 View Synthesis Optimization (VSO) 방식이다. VSO 방식의 개략도는 <그림 11>과 같다.



<그림 11> VSO 방식의 개략도

VSO 방식은 encoder 내부에서 실제 합성을 진행하여 distortion을 예측하고 있어 그 예측값이 상당히 정확하다는 장점이 있지만, R-D 과정에서 항상 블록 레벨의 합성을 진행해야 하기 때문에 높은 복잡도의 연산량을 가지게 된다. 이러한 점을 극복하기 위해, 보다 빠른 방법으로 distortion을 예측하는 방식이 제안되어 실제 렌더링을 진행하지 않고 모델을 이용한 방법으로 합성영상의 distortion을 예측하는 방식이 제안되었다. 이 방식은 VSO 방식 중 TU의 split 모드 결정 과정을 대체하여 복잡도를 크게 감소시켰다.

또한 초기에는 깊이영상 압축 시 기존 사용하던 distortion 측정방식인 Sum of squared differences (SSD)를 모두 VSO기반 distortion으로 대체하여 사용하였지만, 위 과정으로 인해서 압축된 깊이영상이 너무 심하게 왜곡되는 단점이 존재하였다. 따라서, 이후 조정을 통해 VSO와 SSD기반 distortion을 일정 비율로 섞어서 고려하는 방식이 현재 사용되고 있다.

4. 주요 기술 성능 평가

JCT-3V에서는 3D-HEVC에 채택된 각각의 기술들에 대하여 별도의 Ad-Hoc Group (AHG)를 두어 표준화 진행에 따른 성능을 꾸준히 평가 관리하고 있다. 현재 소개된 각각의 기술들에 대한 성능 및 3D-HEVC에 대한 전체적인 성능은 AHG에서 기고한 성능 평가 기고서를 통해 확인할 수 있다^[12].

IV. 결론

지금까지 미래 비디오 표준기술 중 3차원 영상을 위한 깊이영상 기반 다시점 시스템에 대한 연구를 소개하고, 이를 위한 표준기술로서 HEVC 확장 기술인 3D-HEVC의 세부 기술에 대한 소개 및 접근방법에 대해 간략하게 살펴보았다. 현재 3차원 비디오에 대한 표준화가 크게 4가지 트랙으로 동시에 진행되었으며, 각각의 기술에 대한 표준화가 이미 완료되었거나 1년 내 완료가 계획되어 있다. 이 기술들은 각각의 용도에 맞게 디지털 방송, IPTV 등의 실감미디어 뿐 아니라 영



화, 박물관, 게임 등의 각종 엔터테인먼트 등에서 널리 사용될 것으로 예상된다.

참 고 문 헌

[1] B. Lee, "Three-dimensional displays, past and present", Phys. Today, 2013.

[2] C. Zhang, "Multiview imaging and 3DTV", IEEE Signal Processing Magazine, 2007.

[3] N. S. Holliman, N. A. Dodgson, G. E. Favalora, L. Pockett, "Three-Dimensional Displays: A Review and Applications Analysis", IEEE Trans. on Broadcasting, 2011.

[4] H. Urey, K. V. Chellappan, E. Erden, P. Surman, "State of the Art in Stereoscopic and Autostereoscopic Displays", Proceedings of the IEEE, 2011.

[5] A. Vetro, T. Wiegand, G. J. Sullivan, "Overview of the stereo and multiview video coding extensions of the H. 264/MPEG-4 AVC standard", Proceedings of the IEEE, 2011.

[6] P. Merkle, A. Smolic, K. Muller, and T. Wiegand, "Multiview video plus depth representation and coding," in Proc. IEEE Int. Conf. on Image Processing, 2007.

[7] 호요성, 이천, "다시점 3차원 영상 처리 및 부호화 기술", 진샘미디어, 2013

[8] J.-R. Ohm, "Overview of 3D video coding standardization", Three Dimensional Systems and Applications (3DSA), 2013.

[9] Common Test Conditions of 3DV Core Experiments, ITU-T SG16 WP3 ISO/IEC JTC1/SC29/WG11, Doc. JCT3V-G1100.

[10] Test model 9 of 3D-HEVC and MV-HEVC, ITU-T SG16 WP3 ISO/IEC JTC1/SC29/WG11, Doc. JCT3V-I1003.

[11] 3D-HEVC Draft Text 5, ITU-T SG16 WP3 ISO/IEC JTC1/SC29/WG11, Doc. JCT3V-I1001.

[12] AHG5 : Performance evaluation for 3D-HEVC and its coding tools, ITU-T SG16 WP3 ISO/IEC JTC1/SC29/WG11, Doc. JCT3V-I0115.



오 병 태

- 2003년 연세대학교, 공학사 (전기전자공학)
- 2009년 University of Southern California, Ph.D. (Electrical Engineering)
- 2009년~2013년 삼성 종합기술원 전문연구원
- 2013년~현재 한국항공대학교 조교수

〈관심분야〉
영상처리, 멀티미디어 압축, 3차원 비디오 시스템