고품질 3DV/FTV 서비스를 위한 DIBR 기반 다시점 비디오 코딩에 관한 연구

*황능주 *남정학 *심동규 **방 건 **정원식 **허남호
*광운대학교 **한국전자통신연구원

*chakdori@kw.ac.kr *qejixfyza@kw.ac.kr *dgsim@kw.ac.kr **gbang@etri.re.kr **wscheong@etri.re.kr **namho@etri.re.kr

A Study of Multiview Video Coding Based on DIBR for High Quality 3DV/FTV Service

*Neung-Joo Hwang *Junghak Nam *Donggyu Sim **Gun Bang **Won-Sik Cheong **Namho Hur
*Kwangwoon University **ETRI

요약

본 논문에서는 최근 국내 및 국제 표준화 작업이 활발히 진행되고 있는 3DV/FTV(Free viewpoint TV)를 위한 코딩에 대하여 알아본다. 먼저 기존의 2D 영상 코딩 방법을 사용하여 Texture와 Depth를 코딩하고, 복원한 영상을 사용하여 가상 시점 영상을 합성하여, 코딩된 영상의 품질과 합성된 영상의 품질 사이의 관계에 대해서 알아본 후, 고품질 3DV/FTV 서비스를 위해 고려해야 할 사항들에 대하여 살펴본다.

1. 서론

최근 3DV/FTV에 관심이 세계적으로 증가하고 있다. 특히 관련 기술을 일반 사용자가 체험할 수 있는 장소와 기회가 매우 제한적이었던, 과거에 비해 최근 디스플레이 장치의 발전 그리고 보급과 맞물려 관심이 한층 고조 되고 있다. 3DV/FTV는 기존의 2D를 넘어서 더욱 향상된 리얼리티와 몰입감를 제공할 수 있고, 사용자는 더 큰 재미와 만족을 느낄 수 있다. 이러한 3DV/FTV와 관련해서는 현재 국제 및 국내 표준이 없는 상태이지만 최근 표준화가 활발히 진행되고 있다. 국제 표준화 기구인 ISO/IEC JTC1/SC29/WG11 MPEG에서는 2010년 4월 draft Call for Proposal이 작성되었으며, 2010년 7월 final Call for Proposal 작성을 앞두고 있다. 국내에서는 EBS, KBS, MBC, SBS등의 방송사, LG전자, 삼성전자등 관련 회사, 연구소, 학교 등이 참여한 가운데 3DTV를 위한 표준화 작업을 활발히 진행하고 있다.

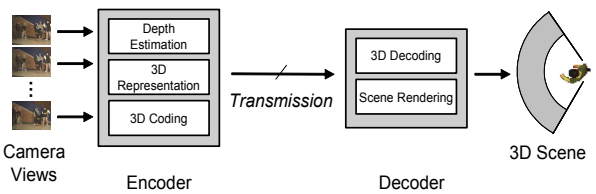

그림 1. 3D 비디오 처리 과정

그림 1은 3D 비디오의 처리 과정을 나타낸다[1]. 송신단에서는 다수의 카메라로부터 다수의 영상을 취득하고, 영상으로부터 Depth를 취득한다. 이렇게 취득한 정보를 코덱을 사용하여 압축하게 된다. 그리고 수신단에서는 수신한 Texture와 Depth 비트스트림으로부터 렌더링을 통하여 목적에 맞는 영상을 사용자에게 제공하게 된다. 특히 FTV는 사용자에게 자유 시점을 제공하는데, 이를 위해 송신단에서 보내 준 시점 보다 더 많은 시점을 제공하기 위해서 Texture와 Depth를 가지고 가상 시점을 합성하게 된다[2]. 이 경우 송신단에서 영상을 코딩할 때에 단순히 2D영상의 개별적인 비트레이트와 품질만을 고려할 것이 아니라, 합성될 시점의 품질을 예측하고, 이것을 고려해야 할 필요가 있다. 이에 본 논문은 Texture와 Depth의 코딩 품질과 합성 시점 영상의 품질 사이의 관계에 대해 알아본 후에, 코딩시에 고려해야할 사항에 대하여 살펴본다.

본 논문은 다음과 같은 순서로 구성된다. 2장에서는 다양한 코딩 조건에서 합성 영상의 품질을 측정하고, 3장에서는 2장에서의 측정 결과를 분석한다. 4장에서는 결론과 향후 연구 과제를 기술한다.

2. 합성 영상의 품질 측정

표 1. 코딩 조건

영상	Lovebird1 6, 8
영상 크기	1024x768
사용한 프레임 수	200
GOP 크기	15
인트라 주기	15

본 논문에서는 영상의 품질 평가를 위하여 표 1의 조건과 더불어 4가지의 타겟 비트레이트(375, 775, 1255, 1750 kbps)를 정하고, 토탈 비트레이트가 타겟 비트레이트를 만족하는 Texture와 Depth의 비트스트림 세트를 구성하였다. 토탈 비트레이트는 아래와 같이 나타낼 수 있다.

Texture bitrate = texture_bitrate(view6)+texture_bitrate(view8)
Depth bitrate = depth_bitrate(view6)+depth_bitrate(view8)
Total bitrate = texture_bitrate + depth_ bitrate

그림 2는 각 타겟 비트레이트를 만족하는 비트스트림 세트를 나타낸다. 각 비트스트림 세트로부터 영상을 복원하고, 복원한 영상을 가지

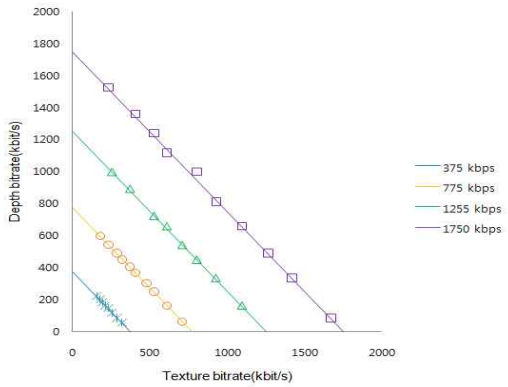

그림 2. 타겟 비트레이트를 만족하는 비트스트림 세트

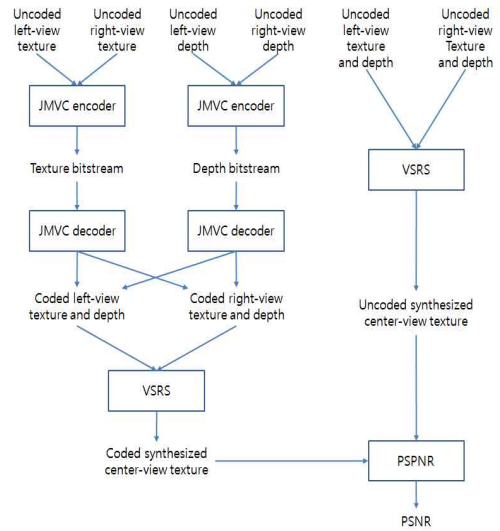

그림 3. 본 논문에서 사용한 가상 시점 영상 품질 측정 방법

고 가상 시점 영상(view7)을 합성하였다. 그리고 나서 2D이미지의 코딩 품질과 합성된 가상 시점 영상의 품질을 비교하였다. 가상 시점 영상의 품질 측정에 관해서는 여러 가지 방법이 있을 수 있으나, 본 논문에서는현재 MPEG EE과 동일한 방법을 사용하였으며, 그 과정이그림 3에 나타나 있다[3].

그림 2는 각 타겟 비트레이트를 만족하는 비트스트림 세트를 나타낸다. 각 비트스트림 세트로부터 영상을 복원하고, 복원한 영상을 가지고 가상 시점 영상(view7)을 합성하였다. 그리고 나서 2D이미지의 코딩 품질과 합성된 가상 시점 영상의 품질을 비교하였다. 가상 시점 영상의 품질 측정에 관해서는 여러 가지 방법이 있을 수 있으나, 본 논문에서는현재 MPEG EE과 동일한 방법을 사용하였으며, 그 과정이 그림 3에 나타나 있다[3].

3. 실험 결과

그림 4는 각 타겟 비트레이트를 만족하는 Texture와 Depth의 비트스트림 조합의 품질을 2절에서 설명한 방법에 의해 측정한 결과(붉은 선)와, 단순히 Texture와 Depth의 평균 PSNR(검은 선)이 함께 나타낸다. 실험 결과는 Texture와 Depth의 PSNR을 단순히 평균 낸 결과는 Depth에 할당된 비트레이트가 많을수록 높은 PSNR 값을 나타낸다. 이와는 반대로 모든 타겟 비트레이트에서 Texture에 할당된 비트(d)의 경우 Texture에 할당된 비트가 가장 많은 경우 오히려적게 할당된 경우보다 합성된 가상 시점 영상의 품질이 약간 떨어짐을 볼 수 있다. 이 경우는 Texture에 비하여 Depth의 비트레이트가 1/40

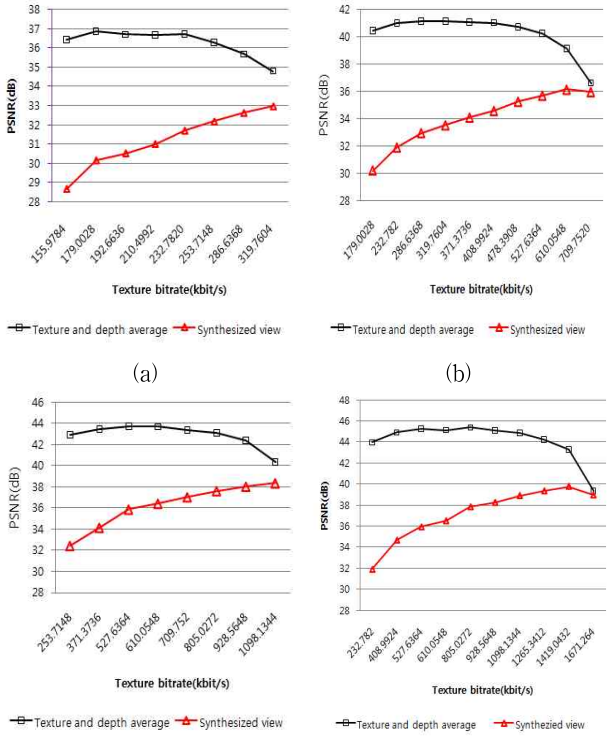

그림4. 각 타겟 비트레이트를 만족하는 합성 영상의 품질과 평균 PSNR (a) 375 kbps (b) 775 kbps (c) 1255 kbps (d) 1750 kbps

이하인 매우 극단적인 경우이다.

4. 결론

본 논문에서는 제한된 환경에서 고품질 3DV/FTV를 위해서는 기존의 2D영상을 코딩 할 때처럼 2D영상의 비트레이트와 품질만을 고려하는 것이 부적합 하다는 것을 실험으로 보였다. 고품질 3DV/FTV를 위해서는 단순히 2D영상이 아니라 타겟 비트레이트를 만족하는 Texture와 Depth의 최적의 조합이 중요한데, Texture에 비트레이트를 더 많이 할당 할 때 더 나은 합성 시점 영상의 품질을 얻을 수 있음을 실험을 통해서 알 수 있었다. 하지만 Texture에 무한히 많은 비트레이트를 할당 하는 것이 아니라 적절한 범위 내에서 이루어 져야 할 것이다[4].

감사의 글

본 연구는 서울시 산학연협력사업(10560)의 일환으로 수행하였음.

본 연구는 지식경제부, 방송통신위원회 및 한국산업기술평가관리원의 산업원천기술개발사업(정보통신)의 일환으로 수행하였음.[KI001932, 차세대 DTV 핵심기술 개발]

참고문헌

[1] “Draft Report on Experimental Framework for 3D Video,” N11201, ISO/IEC JTC/1SC29/WG11 Kyoto, Japan, January 2010.

[2] “Applications and Requirements on 3D Video Coding,” N11061, ISO/IEC JTC/1SC29/WG11 Kyoto, Japan, January 2010.

[3] “Description of Exploration Expreiments in 3D Video Coding,” N11095, ISO/IEC JTC/1SC29/WG11 Kyoto, Japan, January 2010.

[4] “3DV/FTV EE4 result on "Lovebird1" and "Kendo" sequences,” M17459, ISO/IEC JTC/1SC29/WG11 Dresden, Germany, April 2010.