

시점 생성 기법을 이용한 3차원 비디오 부호화의 시점 간 예측 기술

김동현 서정동 함범섭 최진욱 손광훈

연세대학교 전기전자공학부

khsohn@yonsei.ac.kr

View Synthesis for Inter-view Prediction of 3D Video Coding

Donghyun Kim Jungdong Seo Bumseop Ham Jinwook Choi Kwanghoon Sohn

School of Electrical and Electronic Engineering, Yonsei University

요약

멀티미디어 신호 처리 분야의 급속한 발전으로 인해 3차원 TV(3DTV)는 차세대 방송 시스템 시장에서 가장 주목을 받는 제품이 되었다. 3DTV는 사용자에게 입체감을 제공하여 사용자가 마치 그곳에 있는 듯한 효과를 줄 수 있다. 이에 최근 다시점 컬러 영상과 그에 대응되는 깊이 영상을 이용하여 다양한 위치에서의 입체 영상을 제공하는 기술이 활발히 진행되고 있다.

본 논문에서는 주어지는 다시점 컬러 영상과 깊이 영상의 정보를 동시에 이용하여 다시점 영상 부호화의 효율을 높이는데 그 목적이 있다. 이를 위해 시점 생성 기법을 통해 다시점 영상의 부호화 시 참조하는 주변 시점의 영상들로부터 계산되는 변이 벡터 정보를 줄이는 방법을 제안한다. 실험 결과, 기존의 다시점 비디오 부호화에 비해서 약 0.5~1.5 dB정도의 화질 개선이 이뤄졌고, 평균 17.5%의 비트율 감소를 보였다.

1. 서론

차세대 방송 시스템 시장은 기존 HD 보다 2배 이상 향상된 화질의 초고화질 비디오 및 보다 사실적이고 현장감을 느낄 수 있는 3차원 비디오의 형태로 발전하고 있다. 이러한 움직임은 미국 헐리우드를 중심으로 활성화 되고 있는 3D 입체 영화 시장, CES 2010 show에서 삼성전자를 비롯한 다수의 디스플레이 업체들의 3D 디스플레이 시연, 휴대 단말기에서의 3DTV 서비스 기술 개발 및 3D 비디오 콘텐츠의 생성과 포맷에 관한 기술 개발을 위한 유럽의 3D4YOU, 3D@Home을 중심으로하는 가정용 3차원 비디오 서비스 개발, ISO/IEC/JTC1/SC29/WG11(이하 MPEG) 및 JVT에서의 FTV/3DV 표준화 등에서 현실로 나타나고 있다[1].

3차원 비디오에 대한 관심이 늘어남에 따라 현재 다양한 시점의 2차원 컬러 영상과 각 시점의 깊이 정보를 이용해 3차원 영상을 제작하는 방법에 대한 연구도 진행되고 있다. 이에 따라 디스플레이 되는 시점 수 보다 적은 수의 다시점 비디오와 그에 대응되는 깊이 영상(MVD)을 획득 및 부호화하고, 이를 전송하여 수신단에서 중간 영상을 생성하고 다시점 3D 비디오를 생성하는 것을 목표로 하여 표준화가 진행중이다[2]. 3D 비디오는 깊이 영상의 사용으로 인해 기존의 2D 비디오에 비해 방대한 양을 가지므로 이를 효율적으로 부호화 할 수 있는 기술이 요구된다.

현재 다시점 영상을 직접 부호화 하여 3DTV와 FTV

에 대응하는 연구는 비교적 많이 수행되어왔고 관련 표준화도 진행되었으나 일반적인 2D 부호기로 각 시점을 개별 부호화하는 방식에 비해 성능 개선의 폭이 크지 않으며 다시점 영상만을 부호화 하는 방식은 다양한 디스플레이 방식에 대응하기 어렵기 때문에 수신단에서의 활용도가 떨어진다.

본 논문에서는 주어진 컬러 영상과 깊이 영상의 정보를 동시에 이용하여 3차원 비디오 전송을 위한 다시점 영상 부호화의 효율을 높이고자 한다.

본 논문에서는 시점 생성 기법을 이용해 기존의 다시점 영상 부호화 시 참조하는 주변 시점의 영상들로부터 구하는 변이 벡터 정보의 전송량을 줄임으로써 부호화의 효율을 높이는 방법을 제안한다.

논문의 구성은 2장에서 본 논문의 기반이 되는 다시점 영상 부호화 기술과 시점 생성 기술 및 변이 추정 기술에 대해 간략히 서술하고, 3장에서는 제안하는 시점 생성 기법을 이용한 시점 간 예측 기술에 대해 서술한다. 4장에서는 실험 결과를 보여주고, 5장에서 결론을 맺는다.

2. 관련 연구

이 장에서는 본 논문의 기반이 되는 3차원 비디오 전송을 위한 다시점 영상 부호화 기술과 시점 생성 기술에 대해 간략히 서술한다.

2-1. 3차원 비디오 전송을 위한 다시점 영상 부호화 기술

현재 MPEG FTV/3DV에서는 여러 MVD 부호화에 대한 실험이 진행되고 있으며, 곧 이를 이용한 MVD 데이터 부호화에 대한 제안 기술 수집(CfP)을 공지할 예정이다. FTV/3DV에서 표준화하려고 하는 것은 3D 디스플레이에서 동시에 N시점 비디오를 시청자에게 보여주는 다시점 3D 비디오를 대상으로 하고 있다. 이런 시스템의 경우 효율성을 위하여 디스플레이 되는 시점 수 보다 적은 수인 K개의 다시점 컬러 영상과 깊이 영상을 전송한다. FTV/3DV에서는 여러 실험을 통해 3개의 컬러 영상과 2개의 깊이 영상을 전송하기로 한다. 3개의 컬러 영상은 그림 1과 같이 Multi-view Coding(MVC)과 같은 예측 구조로 부호화를 진행한다. 부호화는 우선 시점 0 (이하 I 시점)을 부호화 하고, 부호화 된 I 시점을 참조하여 시점 2 (이하 P 시점)을 부호화 한 후, 부호화된 I 시점과 P 시점을 참조하여 시점 1 (이하 B 시점)을 부호화하는 순서로 진행된다.

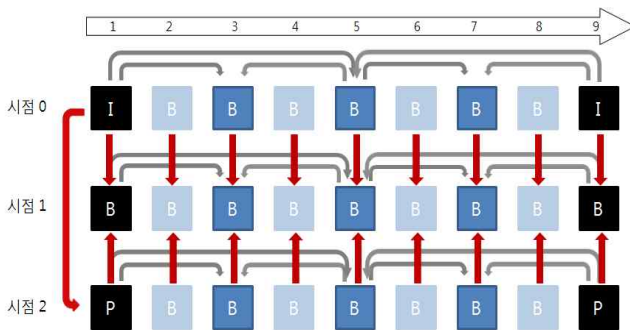


그림 1. MPEG FTV/3DV에서 사용하는 MVC기반의 다시점 영상의 부호화 구조

2-2. 시점 생성 기술 및 변이 추정 기술

현재까지 다양한 시점 생성 기술이 제안되었다. 일반적으로 시점 생성 기술은 기본적으로 두 개 이상의 시점의 위치와 획득된 영상의 기하학적인 관계를 이용하여 소수의 한정된 시점 영상을 다수의 중간 시점의 영상으로 합성하는 방법[3]을 사용한다. 시점의 위치와 영상들 간의 기하학적 관계는 행렬의 형태로 표현되어 행렬의 회전과 평행 이동 연산을 통해 시점의 변화를 표현할 수 있다. 따라서 카메라의 위치와 영상간의 기하관계를 통해 가상 시점의 영상을 합성할 수 있다. 본 논문에서는 컬러 영상과 해당 시점 깊이 영상으로부터 3차원 투영에 기반한 두 시점 사이의 중간 영상을 생성하는 기술을 적용한다[4]. 이 방법은 카메라 외부 및 내부 파라미터와 식 (1)에 표현된 공간 좌표계와 카메라 좌표계 변환식을 이용하여 생성하고자 하는 시점의 주변 카메라 영상 및 깊이 영상으로부터 생성하고자 하는 시점 위치로 먼저 양쪽 시점의 깊이 영상을 투영한다.

$$sC_p(i) = A(i)R^{-1}(i) [W_p(i) - t(i)] \quad \text{식 (1)}$$

$C_p(i)$: 카메라좌표계의 위치
 $W_p(i)$: 공간좌표계의 위치
 $A(i)$: 내부 파라미터
 $R^{-1}(i), t(i)$: 회전, 이동 파라미터
 s : 깊이 정보

이 때 가려짐 영역에서 생기는 작은 홀은 이중 선형 (Bilinear) 필터를 이용하여 채워준다.

다음으로, 투영된 깊이 영상들을 이용하여 주변 두 시점의 각 화소를 가상 시점 위치로 3차원 투영하고 두 시점 영상에 의해 투영되지 않은 위치의 컬러 정보는 주변 화소의 컬러 정보를 이용하여 보간 기법을 통해 채워준다. 만약 두 시점에서의 컬러 정보가 동일 위치로 투영될 경우 두 시점에서 가까운 시점의 컬러 정보로 채워주는 기법을 사용한다.

본 논문에서는 MPEG FTV/3DV의 실험 환경과 동일한 3개의 다시점 컬러 영상과 2개의 깊이 영상을 사용하므로 가운데 시점인 B 시점의 깊이 영상을 필요로 한다. 이 과정에서 많은 계산량이 문제가 되는데, 본 논문에서는 기존 영상의 비용함수를 양쪽으로 전달함으로써 계산량을 줄이는 방법[5]을 선택한다.

3. 시점 생성 기법을 이용한 시점 간 예측 기술

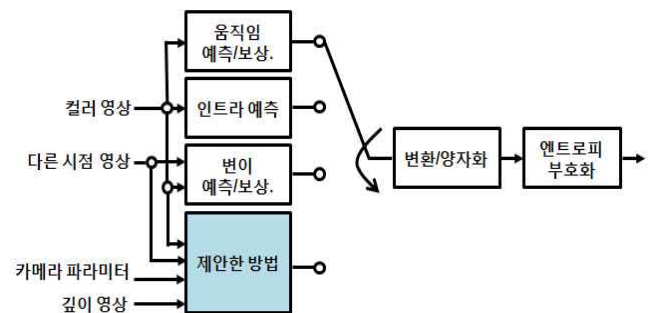


그림 2. 시점 생성 기법을 이용한 시점 간 예측 기술을 적용한 부호화 구조 블록 다이어그램

본 논문에서는 I 시점과 P 시점에 시점 생성 기법을 적용하여 새로 생성된 B 시점 영상을 참조하여 기존의 I 시점과 P 시점에서 구하는 변이 벡터의 전송량을 줄여 부호화의 효율을 높이는 방법을 제안한다. 그림 2는 본 논문에서 제안한 시스템의 개괄적인 블록 다이어그램으로 나타낸 것이다.

기존의 다시점 부호화 기술은 예측하고자 하는 매크로 블록의 위치에서 시점의 위치가 다른 I 시점이나 P 시점에서 변이 벡터를 계산하고, 움직임을 예측하여 얻은 정보를 부호화 하여 전송하였다. 그러나 본 논문에서 제안한 방법은 변이 벡터를 계산하고, 움직임을 예측할 필요 없이 현재 예측하고자 하는 매크로 블록의 위치에 해당하는 시점 생성 기법 기술을 통해 현재의 B 시점과 동일한 시점

으로 생성된 B 시점의 정보를 부호화 하여 전송한다. 이 방법은 움직임이 거의 없는 배경 영역에서보다 주로 움직임이 많은 물체의 경계 부분에서 많이 선택된다. 또한 기존의 다시점 영상을 부호화하던 기술과는 달리 변이 벡터를 전송하는 과정이 생략되므로 비트율 측면에서 이득을 얻을 수 있다.

4. 실험결과

본 논문에서는 기존의 다시점 영상 부호화 기술과 시점 생성 기법 기술을 이용해 시점 간 예측하는 다시점 영상 부호화 효율을 비교하기 위해 B 시점의 깊이 영상의 유무에 따라 두 가지의 데이터를 사용하기로 한다. B 시점의 깊이 영상이 있는 경우 변이 추정 기술이 필요 없으므로 변이 추정 기술에서 발생할 수 있는 오류를 무시할 수 있다.

본 논문에서 1024×768 크기로 8개 시점의 컬러영상과 깊이 영상 모두 제공되는 MSR(Microsoft Research)의 Breakdancer 영상[6]을 가운데 시점인 B 시점의 깊이 영상이 있는 데이터로 사용하기로 한다. B 시점의 영상이 제공되지 않는 경우는 GIST에서 제공한 MPEG test material for 3D video (Newspaper)를 실험 영상으로 사용하기로 한다.

	Breakdancer	Newspaper
Frame rate	15 Hz	16.67 Hz
QP	27, 32, 37, 40	28, 32, 36, 40
GOP Size	15	12
GOP Structure	hier B	hier B
Search Range	±32	±32
Motion skip mode	On	On

표 1. 다시점 영상 부호화를 위한 실험 환경

본 논문에서는 JMVC 5.0.5 참조 소프트웨어를 표 1에 기재되어 있는 실험 환경에 따라 실험을 진행하였다. 그림 3은 Breakdancer 영상과 Newspaper 영상에 각각 일반적인 다시점 영상 부호화 기술과 제안된 다시점 영상 부호화 방법을 비교한 것이다.

그림 3(a)는 깊이 영상이 제공되는 경우의 실험이고, 그림 3(b)는 깊이 영상이 제공되지 않아 변이 추정 기술을 통해 B 시점의 깊이 영상을 예측한 후 시점을 생성해 실험한 결과이다. 그림을 통해 볼 수 있듯이, 제안한 방법이 저 비트율에서는 기존의 다시점 영상 부호화 기술에 비해 PSNR이 약 0.5~1.5dB의 효율을 보이고, 약 17.5% 정도의 비트율 감소를 보이고 있다. 하지만 고 비트율로 갈수록 다시점 영상 부호화 효율이 점점 기존의 다시점 영상 부호화 기술의 효율과 비슷해지는 것을 확인할 수 있다.

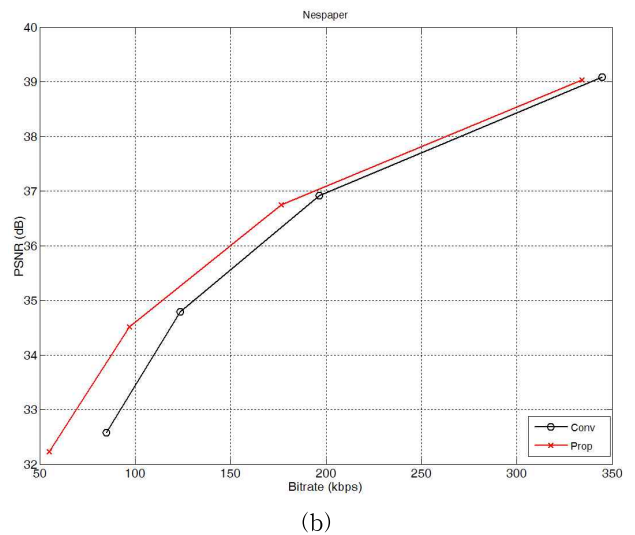
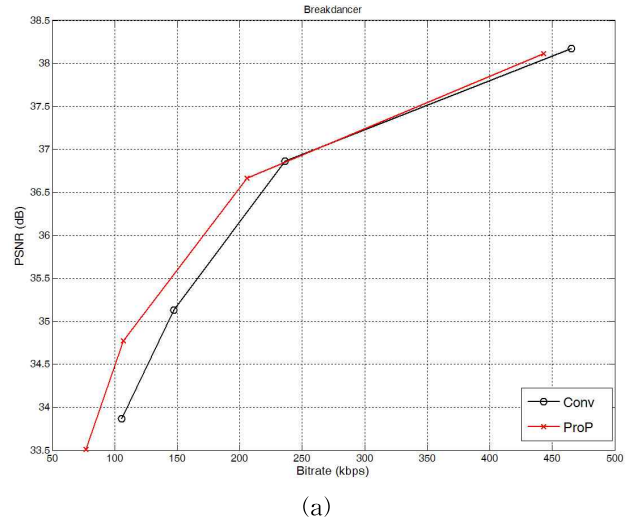


그림 3. 실험 영상에 대한 부호화 R-D 곡선 결과 (a) Breakdancer 영상, (b) Newspaper 영상

5. 결론

본 논문에서는 다시점 영상을 부호화할 때 시점 생성 기법을 통해 생성된 시점을 이용하는 시점 간 예측 기술을 제안하였다. 이는 B 시점 영상을 부호화할 때 그 시점에 맞게 시점을 생성하여 기존에 I 시점의 영상과 P 시점의 영상에서 변이 벡터를 전송하는 과정이 줄어든다. 왜곡이 많이 발생하는 저 비트율에서는 변이 벡터 계산과정이 생략되는 경우가 많아 동일 비트율 대비 PSNR 측정 결과가 약 0.5~1.5dB 정도 이뤄지는 것을 확인할 수 있다.

그러나 본 논문에서 제안한 알고리즘은 시점 생성 기술의 성능에 따라 몇 가지의 한계점을 가지고 있으며 그것들은 향후 과제로 남겨두었다. 우선, 시점 생성 시 가려짐 영역 부분의 깊이 값을 알 수 없어 고 비트율로 갈수록 생성된 시점 영상 참조 비율이 줄어들어 기존의 다시점 영상 부호화 기술의 효율과 거의 동일해지는 것을 실험 결과를 통해 확인할 수 있다. 또한 변이 추정 기술에

의해 생성되는 깊이 영상의 품질에 의존적인 경향이 있기 때문에 올바른 시점 생성을 위해서는 정확한 깊이 영상을 사용하여야 하는 한계점이 있으며, 변이 추정 기술의 수행 과정에 의해 전체 수행 시간이 증가하는 부분을 해결해야 할 것이다.

6. 참고문헌

- [1] 한국방송기술인연합회, "HD Display의 기술발전," 방송기술저널, 제 72호, 2009년 3월, p.4.
- [2] ISO/IEC JTC1/SC29/WG11, "Call for Contribution on 3D Video Test Material(Update)", N9595, Jan. 2008.
- [3] D. Scharstein, "View Synthesis Using Stereo Vision," Springer Press, 1999.
- [4] ISO/IEC JTC1/SC29/WG11, "View Synthesis Method without Blending," M16091, Feb. 2009.
- [5] D. Min, D. Kim, S. Yun and K. Sohn, "2D/3D freeview video generation for 3DTV System," Signal Processing: Image Communication, vol.24, no. 1-2, pp 31-48, January 2009.
- [6] <http://research.microsoft.com>