

스테레오 비디오 부호화에서의 시공간적 계위 오버헤드

오세찬⁰ 김길동 박성혁 이한민
한국철도기술연구원 차세대전동차 연구팀
{soh⁰, gdkim, shpark, hanmin}@krii.re.kr

Scalability Overheads Analysis for Stereoscopic Video Coding

Sehchan Oh⁰, Gildong Kim, Sunghyuk Park, Hanmin Lee
Advanced EMU Research Team, Korea Railroad Research Institute

요 약

본 논문에서는 이질적인 수신단말들에 효율적인 스테레오 영상전송을 위한 시공간적 계위를 이용한 부호화 방법을 제시하였으며, 각각의 시간적, 공간적 계위의 오버헤드를 분석하였다. 제시된 부호화기는 유연성있는 스테레오 비디오 서비스를 위해서 하나의 영상에 각각 시간적, 공간적으로 확장 가능한 추가적인 비트열을 정의하였다. 실험을 통하여 공간적 계위 오버헤드는 시간적 계위 오버헤드에 비해 상대적으로 크지만, 스테레오 부호화기의 양안차 예측으로 인해 많이 감소됨을 보였다. 제안된 스테레오 비디오 부호화기는 다양한 디스플레이 장치 및 네트워크 환경을 가지는 유비쿼터스 컴퓨팅 환경을 위한 효율적인 스테레오 비디오 전송 서비스에 활용될 것으로 기대된다.

1. 서 론

인터넷 상의 멀티미디어 통신의 발달로 원격의 사용자들 간에 스테레오 영상과 같은 실감미디어 서비스가 가능해졌다. 하지만, 스테레오 영상과 같은 실감 미디어의 가용성은 전적으로 수신단말의 네트워크 및 디스플레이 환경에 의해 의존한다. 따라서 각각의 수신단말들이 가지는 다양한 네트워크 환경과 처리능력 및 디스플레이 장치들을 고려한 계위를 이용한 스테레오 영상 부호화 방법이 필요하다. 하지만, 계위를 이용한 부호화 방법은 그에 따른 비용이 따른다. 일반적으로 어떠한 유연성을 제공하기 위해 추가적으로 필요한 정보로 인해 전체 전송량이 증가하며, 단일 영상에 대해 기본계층(Base Layer: BL) 비트열과 하나 이상의 향상계층(Enhancement Layer: EL) 비트열을 복원해야 하므로 시간 복잡도가 크다.

현재까지, MPEG-2에서 정의한 여러 계위 (scalability)를 시작으로 스테레오 영상 부호화와 관련된 연구의 대부분이 2차원 기반의 여러 계위들을 3차원 영상 부호화에 적합하도록 변형하고 좌우 영상의 상관관계(양안차)를 이용한 압축효율의 개선에 초점을 맞추고 있다 [1-2]. 특히, 최근에는 양안차 검색의 복잡도를 줄이기 위한 다 해상도 기반의 양안차 예측기법과, 예측의 정확도를 높이기 위해 가변블록 기반의 양안차 예측기법에 대한 연구들이 진행되고 있다 [3]. 하지만 이질적인 네트워크 환경 및 수신측 디스플레이를 고려한 스테레오 비디오 압축기법에 관한 연구는 아직까지 미흡한 실정이다.

본 논문에서는 이질적인 수신 단말들 간에 효율적인 스테레오 비디오 서비스 환경을 제공하도록, 시공간적 계위를 이용한 스테레오 비디오 부호화 방법을 제안한다.

제안된 부호화기는 효율적인 스테레오 비디오의 전송과 재현을 위해서, 스테레오 비디오 영상을 BL과 여러 EL로 나누어 부호화하며, 네트워크 허용대역폭에 따라서 전송할 영상의 시공간적 해상도를 결정한다. 한편, 수신단은 전송 받은 여러 계층의 비트 열을 선택적으로 결합함으로써, 자신의 디스플레이 환경에 맞도록 디스플레이 한다.

실험을 통하여 스테레오 비디오 부호화에서는 다양한 시간적, 공간적 해상도를 지원하기 위해 추가적으로 요구되는 정보량이 스테레오 부호화기의 양안차 예측을 수행함으로써 줄일 수 있음을 보였으며, 각 계층의 효율적인 비트 율 분할 방법을 제시하였다.

2. 스케러블 스테레오 비디오 복부호화 시스템

2.1 시공간적 계층의 구성

이질적인 수신단말들 사이에 효율적인 스테레오 비디오 서비스를 제공하기 위해서 그림 1과 같이 하나의 BL과 여러 EL을 정의하였다. BL과 EL3은 각각 기본 해상도의 좌우 영상을 표현하기 위한 비트 열이다. BL을 생성하기 위해서 부호화기는 현재 프레임과 이웃한 프레임을 이용한 움직임 보상 예측(Motion Compensation Prediction: MCP)을 수행하여 시간적 중복성을 제거한다. EL3 비트열은 MCP와 더불어 시간적으로 일치하는 BL계층의 복호화된 좌영상과의 양안차 보상 예측(Disparity Compensation Prediction: DCP)을 수행함으로써 생성된다. EL1과 EL4는 BL과 EL3에 추가적으로 완전한 시간적 해상도를 제공하기 위한 비트 열이다. 마찬가지로, EL2와 EL5는 [BL+EL1]과 [EL3+EL4]에 추가적으로 완전한 공간적

해상도를 제공하기 위한 비트 열이다.

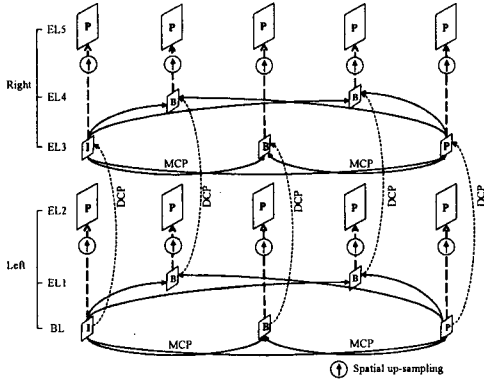


그림 1. 시공간적 계층의 구성

2.2 스케러블 스테레오 비디오 부호화기

제안된 시스템의 스테레오 비디오 부호화기의 구조는 그림 2와 같다. 공간적으로 기본 해상도와 완전한 해상도를 표현하기 위해 입력된 스테레오 영상을 하향 샘플링한다. 공간적 기본 해상도의 좌 영상은 기존의 비계위 MPEG-2 부호화기를 사용하여 부호화 된다. 또한, 여러 시간적 해상도를 표현하기 위해 B 프레임 분할(B-frame Partition)을 거쳐 각각 BL, EL1 비트 열을 생성한다. 공간적 기본 해상도의 우 영상은 현재의 영상과 시간적으로 일치하는 공간적 기본 해상도를 가지는 좌 영상을 참조하여 부호화 된다. 움직임 벡터(Motion Vector: MV) 또는 양안차 벡터(Disparity Vector: DV) 중 전송효율이 높은 쪽을 선택한 후, 그에 따라서 움직임 차이 정보(Dis-

placed Frame Difference: DFD) 혹은 양안 차이정보(Displacement Compensated Difference: DCD)를 부호화하여 전송한다.

해당 이미지에서 하나의 픽셀에 대한 DCD는 수식 (1)과 같이 정의된다.

$$DCD(x, DV) = I_R(x) - I_L(x + DV) \quad (1)$$

공간적 기본 해상도를 가지는 우 영상을 부호화하기 위한 방법은 다음과 같다.

- 우 영상의 I-프레임은 양안차 벡터를 이용하여 예측된 블록을 이용하여 부호화한다.
- 우 영상의 P-프레임은 동일한 시퀀스 내에서 순방향 예측(forward prediction)을 하여 얻어진 MV와 시간적으로 동일한 좌 영상을 참조하여 얻어진 DV를 이용하여 부호화한다. 예측된 P-프레임은 수식 (2)와 같이 표현된다.

$$P_{pred}(V_f, V_d) = w_f R_{ref}(V_f) + w_d R_{ref}(V_d) \quad (2)$$

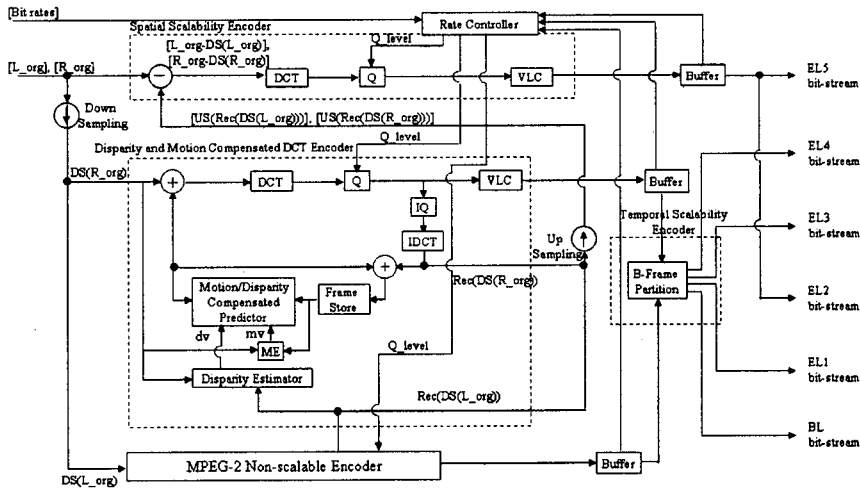
위 식에서 P, v, w는 각각 예측된 P-프레임의 매크로블록(MacroBlock: MB), 벡터, 가중치를 나타낸다. 그리고 인덱스 f, d는 각각 순방향, 양안차를 나타내며, R_{ref}는 복호화 된 프레임의 MB를 나타낸다

- 우 영상의 B-프레임은 동일한 시퀀스 내에서 순방향 움직임 예측 및 역방향 움직임 예측을 하여 얻어진 MB와 양안차 예측을 통해서 얻어진 MB를 이용하여 부호화한다. 예측된 B-프레임은 수식 (3)과 같다.

$$B_{pred}(V_f, V_b, V_d) = w_f R_{ref}(V_f) + w_b R_{ref}(V_b) + w_d R_{ref}(V_d) \quad (3)$$

마찬가지로, B, v, w는 각각 예측된 B-프레임의 MB, 벡터, 가중치를 나타낸다. 마찬가지로, 인덱스 f, b, d는 각각 순방향, 역방향, 양안차를 나타낸다.

완전한 공간적 해상도 표현을 위해 추가적으로 필요한 비트 열(EL2, EL3 비트 열)은 복호화 된 기본 해상도의 영상을 본래 해상도로 상향 샘플링 하여 시간적으로 일치하는 원 영상과의 차이 값을 부호화하여 생성한다.



DCT: Discrete cosine transform VLC: Variable Length Coder IDCT: Inverse discrete cosine transform ME: Motion estimation
 Q: Quantization DS: Down sampling IQ: Inverse quantization US: Up sampling

그림 2. 스케러블 스테레오 비디오 부호화기 구조

3. 실험 및 분석

테스트 영상은 순차방식의 640x480의 해상도를 가지는 컬러 영상이며 YUV 비율은 4:2:0을 가진다. GOP(Group Of Picture)는 16 프레임이며, B-프레임은 GOP의 75%를 차지한다. 그림 3(a), 3(b)와 3(c), 3(d)는 각각 시퀀스 *Laboratory*와 *Ubihome*에 비해 카메라의 뷰의 이동을 비롯한 움직임이 많다.

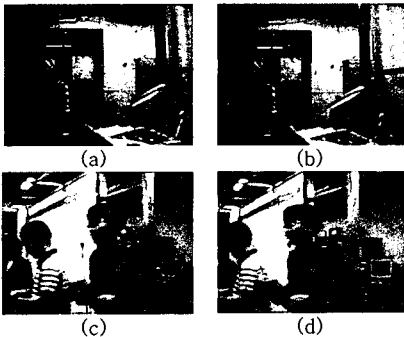


그림 3. 테스트 시퀀스, (a),(b) *Laboratory*, (c)(d) *Ubihome*

표 1은 실험 결과를 비계위 MPEG-2 부호화기와 비교한 값을 보여준다. 공간적 계위에 대한 오버헤드는 제안된 부호화기가 MPEG-2 비계위 부호화기를 통해 얻어진 영상과 같은 화질의 영상을 얻기 위해서 추가적으로 필요한 비트율로 정의 할 수 있다. 제안된 부호화기는 DCP를 이용한 스테레오 비디오 부호화기에 공간적 계위를 결합하여 공간적 계위에 필요한 오버헤드를 줄일 수 있다. 움직임이 비교적 적은 *Laboratory*의 경우 대부분의 블록마다 변화량이 적은 움직임 벡터를 선택하기 때문에 DCP로 얻어지는 압축 효율이 적다. 반면, *Ubihome*의 경우는 블록마다 움직임 변화량이 크기 때문에 MCP 대신, DCP를 이용하여 부호화를 수행한다. 따라서 비계위 부호화기와 비교하였을 때 얻어지는 압축 효율은 상당히 크며, 결과적으로 공간적 계위에 따르는 오버헤드를 크게 줄일 수 있음을 알 수 있다.

표 1. 공간적 계위 오버헤드

Sequence		<i>Laboratory</i>		<i>UbiHome</i>	
		Left	Right	Left	Right
Single Layer Encoder	Bit-rate [Mbps]	1.0	1.0	1.0	1.0
	Average PSNR for luminance [dB]	41.03	40.25	36.50	36.33
	Mean PSNR [dB]	40.62		36.42	
Proposed Scalable Stereo Coder	Bit-rate [Mbps]	1.14	1.14	1.045	1.045
	Average PSNR for luminance [dB]	40.81	40.43	36.83	36.07
	Base layer bit-rate [Mb]	0.49	0.49	0.41	0.41
	Base layer bits-rate as percent of total bit-rate [%]	43	43	39.2	39.2
	Spatial scalability overhead [%d]	14		4.5	
	Mean PSNR [dB]	40.61		36.44	

그림 4는 시간적 계위 오버헤드를 보여준다. 완전한 해상도와 기본 해상도 즉, 2단계 시간적해상도를 지원하기

위해서 I-프레임과 P-프레임 또는 두 개의 P-프레임 사이에 반드시 출수개의 B-프레임이 존재해야 한다 [4]. 하지만, 그림 4(a)와 4(b)에서 볼 수 있듯이 일반적으로 B-프레임의 개수가 2개인 경우 전체적으로 안정된 성능을 보인다. 그림 4(c)는 각각의 시퀀스에 대해 같은 화질의 영상을 얻기 위해서 필요한 bit rate를 B-프레임의 개수에 따른 변화를 보인다. 두 시퀀스를 고려해 볼 때 B-프레임의 개수가 2개인 경우 가장 적은 bit rate를 가지며, 3개인 경우 *Ubihome*은 약 5%의 오버헤드를 가진다. 하지만, 양안차 예측을 통해 이러한 오버헤드는 제거될 수 있다.

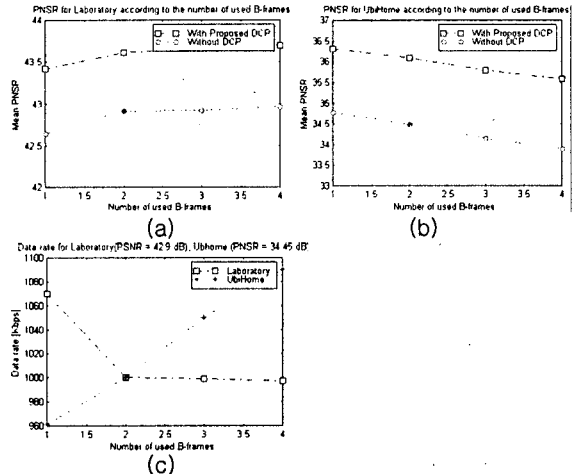


그림 4. 시간적 계위 오버헤드

4. 결론

본 논문은 이질적인 환경에서의 스테레오 비디오 전송을 위한 부호화 방법을 제안하였다. 또한 실험을 통하여 시공간적 계위에 따르는 오버헤드를 분석하였다. 제안된 스테레오 비디오 부호화 방법은 다양한 디스플레이 장치 및 네트워크 환경을 가지는 유비쿼터스 컴퓨팅 환경을 위한 효율적인 스테레오 비디오 전송 서비스에 활용될 것으로 기대된다.

참고 문헌

[1] B. Haskell, A. Puri, A. N. Netravali, "Digital Video: An Introduction to MPEG-2", New York: Chapman & Hall, 1997.
 [2] Domanski, A. Luczak, S. Mackowiak, "On Improving MPEG Spatial Scalability", ICIP Proc. 2000.
 [3] S.Sethuraman, M.W. Siegel, A.G. Jordan, "A multi-resolutional region based segmentation scheme for stereoscopic image compression", Proc. of the IS&T/SPIE Symp. on Electronic Imaging, Digital Video Compression- Algorithms and Technologies 1995, Vol. 2419, 1995.
 [4] M. Narroschke, "Functionalities and Costs of Scalable Video Coding for Streaming Services", Proc. of 36th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, California, Nov. 2002.