

향상된 중간 시점 합성 영상을 위한 깊이 영상 부호화

류승철, 서정동, Xingang Liu, 손광훈
연세대학교 전기전자공학과
khsohn@yonsei.ac.kr

Depth Video Coding for Improved Synthesized Intermediate View Video

Seungchul Ryu, Jungdong Seo, Xingang Liu, Kwanghoon Sohn
Department of Electrical and Electronic Engineering, Yonsei University

요 약

본 논문에서는 향상된 중간 시점 합성 영상을 위한 깊이 영상 부호화 방식을 제안한다. 깊이 영상은 실제 영상과 다르게 날카로운 경계를 기준으로 완만한 변화를 가지는 픽셀 값을 가지는 특성이 있다. 따라서 깊이 영상의 부호화에서는 경계 영역을 효율적으로 부호화하는 것이 중요하다. 기존의 다시점 비디오 부호화기 (Multiview Video Coding)가 하나의 프레임 내에서 고정된 양자화 파라미터 값을 사용하는 것에 반해, 제안된 방식에서는 경계 영역을 효율적으로 부호화하기 위해 블록의 특성에 따라 적용적으로 양자화 파라미터를 할당한다. 2 차 미분 영상의 분포에 기반해 각 블록을 경계 블록, 평탄 블록, 일반 블록으로 구분하고 이에 따라 양자화 파라미터를 할당한다. 실험결과로서, 제안하는 방법의 성능이 다시점 비디오 부호화기 참조 소프트웨어 JMVC 8.3 에 비하여 BD-PSNR 이 평균 0.18dB 향상되고, BD-BR 은 평균 4.03% 감소되어 부호화 효율이 우수함을 확인할 수 있었다.

1. 서론

최근 멀티미디어 처리 기술 및 영상 처리 기술의 급격한 발전으로 인해, HD(High-definition)급 이상의 고화질 비디오 서비스가 널리 보급화 되었다. 또한 HD 급 비디오의 다음 세대 비디오 서비스로 3 차원 입체 (Three Dimensional) 비디오 서비스가 주목을 받고 있다. 이와 관련된 연구로써, 여러 시점의 카메라로부터 동시에 획득한 영상을 효율적으로 부호화하는 다시점 비디오 부호화기 (Multiview Video Coding)가 JVT (Joint Video Team of ISO/IEC JTC1/ SC29/WG11 MPEG and ITU-T SG16 Q.6 VCEG)에 의해 표준화가 완료되었다 [1]. 또한 최근 ISO/IEC MPEG 에서 다시점 컬러 비디오 플러스 깊이 정보를 부호화하기 위한 표준 (3DV)을 진행 중이다 [2]. 3DV 는 입력 받은 시점 영상 이외의 다양한 가상 시점 영상을 생성할 수 있는 시스템을 위한 표준으로, 가상 시점 영상은 카메라와 객체간의 거리인 깊이 정보 영상을 이용한 렌더링 방법(Depth Image Based Rendering)을 사용하여 합성된다.

깊이 영상을 부호화하는 방법에 관련한 표준은 현재까지는 따로 존재하지 않기 때문에 컬러 영상의 부호화 방식을 그대로 차용하고 있다 [3]. 하지만 깊이 영상은 컬러 영상과는 다른 특성을 가지기 때문에 기존의 방식은 깊이 영상을 부호화하는데 적합하지 않다. 예를 들어 깊이 영상은 컬러 영상과 다르게 배경 및 객체 내부에서는 픽셀 값의 변화가 완만한 반면, 객체 경계 부분에서는 픽셀 값의 차이가 커서 날카로운 경계를 가진다. 또한, 합성된 가상시점 영상의 화질은 깊이 영상의 정확도에 크게 의존하므로, 이를 효과적으로

부호화하는 방법에 대한 연구가 필요하다.

본 논문의 구성은 다음과 같다. 2 장에서는 본 논문에서 제안하는 깊이 영상 부호화 방법에 대해 자세히 설명하고 3 장에서는 제안된 방법의 성능을 실험을 통해서 확인한다. 마지막으로 4 장에서는 본 논문에 대한 결론을 맺는다.

2. 제안된 깊이 영상 부호화 방법

깊이 영상은 카메라(월드 좌표의 원점)와 객체 간의 거리를 나타내는 영상이므로 (그림-1.b), 배경 및 객체 내부에서는 실제 영상에 비해 픽셀 값의 변화가 완만하다. 반면에 객체 경계 부분에서는 픽셀 값의 차이가 커서 아주

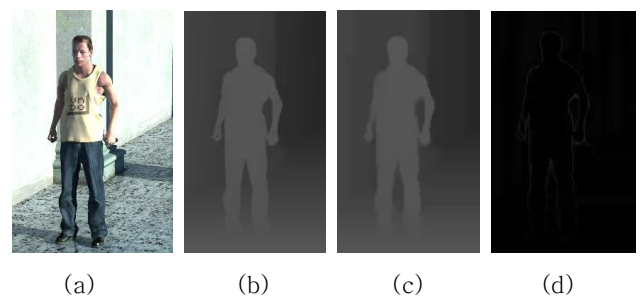


그림 1. 실제 영상과 깊이 영상: a) 실제 영상, b) 원본 깊이 영상, c) 부호화된 깊이 영상 (QP 42), d) b)와 c)의 차이 영상

날카로운 경계를 가지는 특성을 보인다. 이러한 특성 때문에 객체의 경계 부분에서 부호화 오차가 다른 부분에 비해 크게 나타나게 된다 (그림-1.d). 일반적인 자연 영상의 압축과는 다르게 깊이 영상 압축의 주 목적은 중간시점 합성 영상의 화질을 높이는 것이다. 즉, 깊이 영상 자체의 화질을 높이는 것보다 이 깊이 영상을 이용하여 합성한 중간시점 영상의 화질을 높이는 것이 더 중요하다. 깊이 영상을 기반으로 합성된 중간시점 영상의 화질은 일반적으로 깊이 영상의 경계영역의 부호화 오차에 큰 영향을 받는다[4].

본 논문에서는 이러한 깊이 영상의 특성을 고려한 깊이 영상 부호화를 위한 양자화 파라미터(Quantization Parameter) 할당 방법을 제안한다. 깊이 영상을 부호화하기 위해 일반적으로 사용되는 방식은 다시점 영상 부호화기(Multiview Video Coding) 이다. 기존의 다시점 영상 부호화기에서는 한 프레임 내에서 고정된 양자화 파라미터(QP)를 사용한다. 이 양자화 파라미터는 식-(1)과 같이 라그랑지 비용함수에 사용되며, 라그랑지 비용함수는 최종적으로 비트율-왜곡 최적화된 블록 모드를 결정하는데 사용된다.

$$J(Q) = D(Q) + \lambda R(Q) \quad (1)$$

Q는 양자화 파라미터이며, J(Q)는 블록을 Q로 양자화 하였을 때 얻어지는 비용함수, λ는 라그랑지 승수, D(Q)와 R(Q)는 각각 블록을 Q로 양자화 하였을 때의 비트율과 왜곡을 의미한다. Q의 값이 높을수록 양자화 간격이 커지게 되고, 따라서 비트율은 낮아지며 왜곡율은 반대로 높아지게 된다.

제안된 부호화 방법에서는 객체의 경계영역에서의 화질을 높이기 위해, 2 차 미분 영상 정보를 이용하여 적응적으로 양자화 파라미터를 선택한다. 2 차 미분 영상을 얻기 위해 라플라시안 필터(Laplacian Filter)를 이용하였다. 다시 말해, 객체의 경계영역에서는 상대적으로 낮은 양자화 파라미터를 평탄한 영역에서는 상대적으로 높은 양자화 파라미터를 사용한다. 제안된 방식의 알고리즘은 다음과 같다.

- 1) 라플라시안 필터를 이용하여 깊이 영상의 2 차 미분 영상을 획득한다.

-1	-1	-1
-1	8	-1
-1	-1	-1

<라플라시안 필터>

- 2) 각 블록의 평균 라플라시안 필터링 값을 AL로 정의하고 이를 이용하여 각 블록을 경계 블록, 평탄 블록, 일반 블록으로 분류한다.

$$AL = \sum_x \sum_y L(x, y) \quad (2)$$

$$\begin{aligned} & \text{if } AL > Thd_H \\ & \quad QP = Q + \Delta \\ & \text{else if } AL < Thd_L \\ & \quad QP = Q - \Delta \\ & \text{else} \\ & \quad QP = Q \end{aligned}$$

그림 2. 적응적 양자화 파라미터 선택 알고리즘: QP는 현재 블록의 양자화에 사용되는 파라미터이며, Q는 현재 프레임의 양자화 파라미터를 의미한다.

L(x,y)는 현재 블록의 (x,y)의 라플라시안 필터링 값을 의미하며, Thd_H, Thd_L은 블록의 타입을 결정하기 위한 임계 값이다.

- 3) 블록의 타입에 따라 현재 블록의 양자화 파라미터를 선택한다. 즉, 경계 블록에서는 Q-Δ, 평탄 블록에서는 Q+Δ, 일반 블록에서는 Q의 양자화 파라미터를 사용하며, 양자화 파라미터 선택 알고리즘을 그림 2에 나타내었다.

3. 실험 결과 및 분석

3.1 실험 조건

제안된 깊이 영상 부호화 방법의 성능을 검증하기 위해 2 개의 테스트 영상('Mobile' [5], 'Undo_Dancer' [6])을 사용하였다 (그림 3). 제안된 방식은 다시점 영상 부호화기 참조 소프트웨어 JMVC 8.3에 구현되었다. 각 영상은 [5]의 실험 조건에 따라서 부호화 되었으며 블록의 타입을 결정하기 위한 임계 값 Thd_H와 Thd_L은 라플라시안 필터링 된 영상의 히스토그램 분석을 통해 결정되었고, Δ 값으로 2를 사용하였다. 각 테스트영상의 실험 조건을 표 1에 나타내었다.

3.2 제안된 방법의 성능 평가

제안된 부호화 방법의 성능을 평가하기 위하여 JMVC 8.4와 비교 분석을 수행하였으며, 성능 평가 방법으로 BD-PSNR과 BD-BR을 사용하였다[7]. BD-PSNR은 동일한 비트율에서의 PSNR의 증가분을, BD-BR은 동일한 PSNR에서의 비트율 감소분을 의미하며, BD-PSNR이 높을수록 혹은 BD-BR이 낮을수록 높은 성능을 나타낸다.

표 1. 실험 조건

Sequence	Mobile	Undo_Dancer
GOP	15	12
탐색 범위	± 64	± 64
총 프레임 수	76 (5 GOP)	61 (5 GOP)
해상도	720*540	1920*1088
시점	3-5-7	1-2-3
Thd_H	60	40
Thd_L	5	2

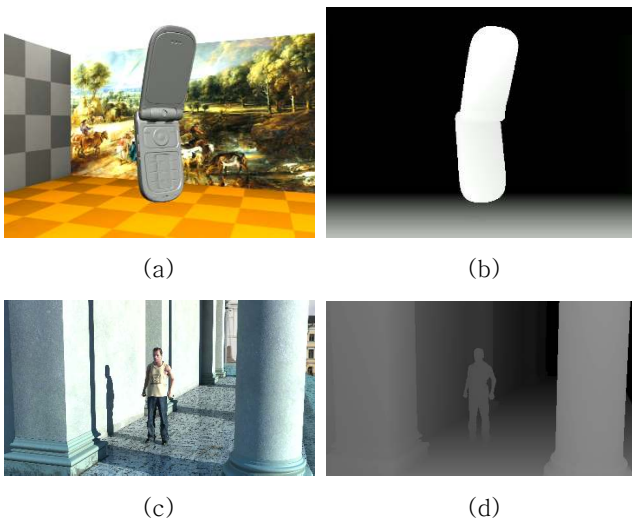


그림 3. 테스트 영상: a) 'Mobile' 컬러 영상, b) 'Mobile' 깊이 영상, c) 'Undo_Dancer' 컬러 영상, d) 'Undo_Dancer' 깊이 영상.

표 2 와 표 3 은 'Mobile' 영상과 'Undo_Dancer' 영상의 실험 결과를 보여준다. 깊이 영상의 PSNR 과 비트율은 좌 영상과 우 영상의 결과를 평균한 값이며, 합성영상의 PSNR 은 실제 중간 시점 영상과 깊이 영상을 이용하여 합성한 중간 시점 영상의 차이를 측정된 값이다. 표에서 확인 할 수 있듯이 JMVC 8.3 을 이용하여 얻은 합성영상보다 제안된 방법으로 얻은 합성영상이 좋은 화질을 보임을 확인 할 수 있다. 표 4 에 BD-PSNR 과 BD-BR 을 나타내었다. 'Mobile' 영상의 경우 0.3dB 의 BD-PSNR 과 5.74%의 BD-BR, 'Undo_Dancer' 영상의 경우 0.06dB 의 BD-PSNR 과 2.32%의 BD-BR 의 부호화 효율이 증가하였다.

4. 결론

본 논문에서는 효율적인 깊이 영상 부호화를 위한 양자화 파라미터 할당 방법을 제안하였다. 깊이 영상은 경계영역이 매우 날카로운 반면, 배경이나 객체 내부에서는 평탄한 특성을 가진다. 이를 바탕으로 제안된 방식에서는 경계 블록, 평탄 블록, 일반 블록으로 각 블록을 구분하고 이에 따라 양자화 파라미터를 할당한다. 실험결과, 제안하는 방법은 다시점 비디오 부호화기 참조 소프트웨어 JMVC 8.3 에 비하여 평균 0.18dB 향상된 BD-PSNR 과 평균 4.03% 감소된 BD-BR 을 보여 부호화 성능이 뛰어난을 확인할 수 있었다.

표 2. 'Mobile' 실험 결과

QP	깊이 영상 PSNR (dB)		깊이 영상 비트율 (kbit/s)		합성 영상 PSNR (dB)	
	JMVC 8.3	제안 방법	JMVC 8.3	제안 방법	JMVC 8.3	제안 방법
32	46.55	46.57	173.59	181.68	41.38	41.46
37	43.28	42.42	128.51	132.21	40.65	40.72
42	37.28	39.87	96.95	97.51	38.59	39.51
47	34.30	34.80	73.04	73.92	37.10	37.17

표 3. 'Undo_Dancer' 실험 결과

QP	깊이 영상 PSNR (dB)		깊이 영상 비트율 (kbit/s)		합성 영상 PSNR (dB)	
	JMVC 8.3	제안 방법	JMVC 8.3	제안 방법	JMVC 8.3	제안 방법
32	46.74	46.76	345.09	348.71	39.20	39.23
37	43.87	43.68	244.70	243.10	38.79	38.80
42	40.87	40.18	173.70	171.96	37.74	37.80
47	36.95	36.68	113.22	115.47	36.36	36.46

표 4. BD-PSNR 과 BD-BR 결과

	Mobile	Undo_Dancer
BD-PSNR (dB)	0.30	0.06
BD-BR (%)	-5.74	-2.32

참고문헌

[1] ISO/IEC JTC1/SC29/WG11, "Text of ISO/IEC 14496-10:200X/FDAM 1 Multi-view Video Coding," N9978, Hannover, Germany, Jul. 2008.

[2] ISO/IEC JTC1/SC29/WG11, "Applications and Requirements of 3D Video Coding," N10358, Lausanne, Switzerland, Feb. 2009.

[3] A. Smolic, K. Mueller, N. Stefanoski, J. Ostermann, A. Gotchev, G.B. Akar, G.A. Triantafyllidis and A.Koz: "Coding Algorithms for 3DTV - A Survey," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol 7, Issue 11, pp. 1606-1621, Nov. 2007.

[4] P. Merkle, A. Smolic, K. Muller, "Multi-view video plus depth representation and coding", *Proc. IEEE International Conference on Image Processing*, San Antonio, Texas, Sep. 2007.

[5] ISO/IEC JTC1/SC29/WG11, "Description of Exploration Experiments in 3D Video Coding," N11274, Dresden, Germany, Apr. 2010.

[6] ISO/IEC JTC1/SC29/WG11, "Call for Proposals on 3D Video Coding Technology," N12036, Geneva, Switzerland, Mar. 2011.

[7] G. Bjontegaard, "Calculation of Average PSNR Differences between RD-curves," ITU-T Video Coding Experts Group document VCEG-M33, Mar. 2001.