

특집논문(레터)-10-15-2-09

블록기반 부호화할 비트평면 결정을 이용한 깊이정보 맵 부호화

김 경 용^{a)}, 박 광 훈^{a)†}

Depth-map coding using the block-based decision of the bitplane to be encoded

Kyung Yong Kim^{a)} and Gwang Hoon Park^{a)†}

요 약

본 논문에서는 깊이정보 맵의 효율적인 부호화 방법을 제안한다. 블록 기반 적응적 깊이정보 맵 부호화 방법에서는 깊이정보 맵에 대한 비트율 조절을 위해 양자화 변수를 이용하여 부호화할 비트평면의 수를 결정한다. 이러한 양자화 변수에 의한 제약 조건에서 벗어나서 부호화할 비트평면을 블록 단위에서 결정하는 방법을 제안한다. 실험 결과, 제안하는 방법이 블록 기반 적응적 깊이정보 맵 부호화 방법보다 BD-PSNR이 0.25 dB 향상되었고 BD-rate가 3.5% 감소되어 제안하는 방법의 우수함을 확인할 수 있었다.

Abstract

This paper proposes an efficient depth-map coding method. The adaptive block-based depth-map coding method decides the number of bit planes to be encoded according to the quantization parameters to obtain the desired bit rates. So, the depth-map coding using the block-based decision of the bit-plane to be encoded proposes to free from the constraint of the quantization parameters. Simulation results show that the proposed method, in comparison with the adaptive block-based depth-map coding method, improves the average BD-rate savings by 3.5% and the average BD-PSNR gains by 0.25dB.

Keyword : 3D Video, Depth-map Coding, FTV(Free view-point TV), MVC(Multi-view Video Coding)

1. 서 론

실감 방송에 대한 사용자들의 관심이 크게 증가되면서 기존 2차원 TV 방송의 한계를 벗어나 3차원 TV 방송으로의 전환이 곧 다가올 것으로 전망된다. 실제로 일본 BS방송

은 2007년 12월부터 HD급 해상도를 지원하는 3차원 입체 방송을 개시하였다. 또한 미국이나 유럽, 아시아 지역에서도 3차원 TV 방송에 대한 연구 개발 활동이 활발히 진행되고 있다. 이러한 3차원 TV 방송에 대한 비디오 표준으로는 여러 시점의 영상을 부호화하기 위해서 ISO/IEC의 JVT (Joint Video Team of ISO/IEC JTC1/SC29/WG11 MPEG and ITU-T SG16 Q.6 VCEG)에서 제정한 다시점 비디오 부호화(Multi-view Video Coding, H.264/AVC Amendment 4) 표준^[1]을 예로 들 수 있다. 또한 입력된 영상 이외의 가상시점 영상들을 깊이정보 맵을 이용한 시점 보간 방법을 통해 무한대까지 생성할 수 있도록 지원하기 위한 3차

a) 경희대학교 전자정보대학 미디어랩

Media Lab., College of Electronics and Information, Kyung Hee Univ.

† 교신저자 : 박광훈(ghpark@khu.ac.kr)

※ 이 논문은 교육과학기술부 및 한국연구재단의 국가지정연구실 지원 사업(ROA-2005-000-10061-0)과 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 지원 사업(NIPA-2010-(C1090-1011-0001))의 연구 결과로 수행되었음.

· 접수일(2009년12월31일), 수정일(2010년2월2일), 게재확정일(2010년2월10일)

원 비디오 부호화 표준^[2]이 ISO/IEC의 MPEG에서 표준화가 진행 중이다. 3차원 비디오 부호화 표준에서 표준화 대상으로는 입력 받은 영상을 부호화하는 방법과 깊이정보 맵을 부호화하는 방법, 그리고 깊이정보 맵을 이용하여 가상시점 영상을 생성하는 방법이다. 3차원 비디오 부호화 표준에서 입력 영상에 대한 부호화 표준은 다시점 비디오 부호화 표준을 이용할 것으로 예상된다. 하지만 깊이정보 맵 부호화와 관련된 국제표준은 아직 존재하지 않는다. 실제로 3차원 콘텐츠를 이용한 응용 분야에서 깊이정보 맵을 부호화하기 위하여 일반적으로 사용되는 동영상 압축 방법을 그대로 이용하고 있다^[3]. 깊이정보 맵은 일반 영상과 그 특성이 다르기 때문에 깊이정보 맵의 특성에 적합한 부호화 방법이 반드시 필요하다고 할 수 있다.

깊이정보 맵(Depth-map)이란, 카메라와 객체 간의 실제 거리를 정수단위로 표현한 것으로써, 카메라와 가까울수록 큰 값(밝은 값)으로 표현된다. 실제 영상과는 다르게 깊이정보 맵은 상당히 완만하고 단조로운 특성을 나타낸다. 이러한 깊이정보 맵의 단조로운 특성을 이용한 무손실 부호화 방법으로 비트평면 기반 무손실 깊이정보 맵 부호화 방법^[4]이 제안되었으며, 또한 손실 부호화 방법으로 H.264/AVC^[5]를 이용한 것보다 높은 부호화 효율을 갖는 블록 기반 적응적 깊이정보 맵 부호화 방법^[6]이 제안되었다.

II. 제안하는 깊이정보 맵 부호화 방법

블록 기반 적응적 깊이정보 맵 부호화 방법^[6]은 양자화 변수의 값에 따라 부호화되는 비트평면의 수가 결정된다. 예를 들어, 입력 받은 양자화 변수 값이 '33' 이상일 경우에는 MSB, MSB-1, MSB-2, MSB-3, MSB-4 비트평면까지만 부호화를 수행한다. 이렇게 양자화 변수에 따라 부호화할 비트평면의 수가 결정되기 때문에, 한 화면의 전체 블록에서 부호화되는 비트평면의 수는 일정하다. 가상시점 영상의 화질은 깊이정보 맵의 배경부분보다 객체 경계부분의 화질에 따라 결정되므로, 더 나은 가상시점 영상의 화질을 얻기 위해서는 배경부분보다 객체 경계부분에 더 많은 비트평면들을 부호화하여야 한다. 본 논문에서는 양자화 변

수에 따라 부호화할 비트평면의 수를 결정하는 기존 방법의 제약조건에서 벗어나고, 깊이정보 맵의 객체 경계부분에 더 높은 중요도를 부여하기 위한 방법을 제안한다.

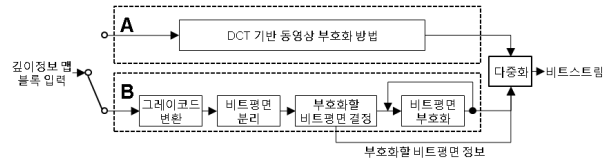


그림 1. 제안하는 방법의 부호화 블록도
Fig. 1. Encoder block diagram of proposed method

그림 1은 본 논문에서 제안하는 부호화할 비트평면의 수를 블록단위에서 적응적으로 결정하는 방법을 이용한 비트평면 단위 부호화와 DCT 기반 동영상 부호화를 적응적으로 선택하는 방법의 부호화 블록도를 나타낸다. 그림 1(A)의 DCT 기반 동영상 부호화 방법은 일반 영상의 부호화에 사용되는 H.264/AVC를 이용한다. 그림 1(B)의 비트평면 단위 부호화 방법은 다음과 같은 4단계를 거쳐 부호화가 진행된다.

- (1 단계) 입력된 깊이정보 맵 블록을 그레이코드로 변환한다.
- (2 단계) 그레이코드로 변환된 n -비트의 깊이정보 맵 블록을 n 개의 비트평면 블록으로 분리한다.
- (3 단계) 입력된 n 개의 비트평면들 중에서 부호화할 비트평면들을 적응적으로 결정한다. 이때 이전 그레이코드 변환 단계에서 인접한 비트평면을 통해 부호화가 수행된 그레이코드의 비트평면 간 의존성 때문에 MSB 비트평면부터 LSB 비트평면의 순서로 차례대로 부호화할 비트평면의 수를 증가시키면서 부호화를 수행하여야 한다. 따라서 부호화할 비트평면을 결정하는 방법은 부호화할 비트평면의 수를 결정하는 것과 같다. 부호화할 비트평면의 수를 결정하는 순서도는 그림 2에 도시하였다. 여기서 결정된 부호화할 비트평면의 수에 대한 정보는 양자화 변수에 따라 가변 길이 부호화되어 비트스트림에 포함된다.
- (4 단계) 상위 단계에서 선택된 비트평면들은 각각 부호화되며, 이때 사용되는 비트평면 부호화 방법은 국제 동영상 표준인 MPEG-4 Part-2 Visual(ISO/IEC 14496-2)^[8]의 이전 영상 부호화 방법^[8,9]을 사용한다.

그림 2는 부호화할 비트평면의 수를 결정하는 순서도를 나타내며, 알고리즘의 자세한 진행과정은 다음과 같다.

- (1 단계) 부호화할 비트평면의 수('N')를 '4'로 설정한다. 5개 미만일 경우는 깊이정보 맵의 왜곡이 심하므로 복잡도를 감소시키기 위해 부호화를 수행하지 않는다.
- (2 단계) 부호화할 비트평면의 수('N')를 '1' 증가시킨다.
- (3 단계) 부호화할 비트평면의 수('N')에 따라 MSB 비트평면부터 LSB 비트평면까지 차례대로 부호화 모드를 설정한다. 예를 들어 'N'이 '5'일 경우는 MSB, MSB-1, MSB-2, MSB-3, MSB-4 비트평면만 부호화하는 모드이고, N이 'm'일 경우에는 MSB와 MSB-1, MSB-2, MSB-3, ..., MSB-(m-1) 비트평면을 부호화하는 모드이다.
- (4 단계) 해당 부호화 모드로 부호화를 수행한 후 H.264/AVC에서 사용하는 비트율-왜곡 최적화 방법을 이용하여 해당 부호화 모드의 비용($Cost_i$) 값을 계산한다.
- (5 단계) 만약 해당 부호화 모드가 모든 부호화 모드 중 비용($Cost_i$) 값이 최소라면, 6 단계로 분기한다. 그렇지 않다면, 2 단계로 분기한다.
- (6 단계) 가장 최소의 비용($Cost_i$) 값을 가진 부호화 모드를 통해 부호화할 비트평면의 수를 결정한다.

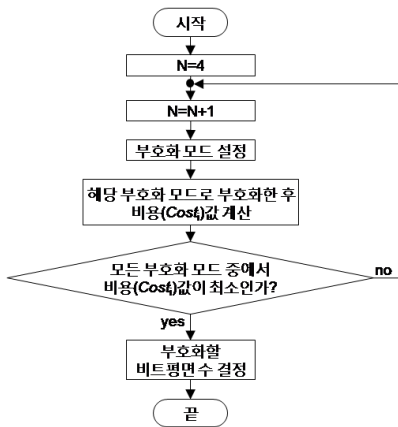


그림 2. 부호화할 비트평면 결정 방법의 순서도
 Fig. 2. Block diagram of decision method of the bitplane to be encoded

본 논문에서 제안하는 방법의 복호화 블록도는 그림 3과 같다. 그림 3에서 역다중화에서는 부호화된 모드 정보를 복호화하며, 해당하는 모드로 복호화가 진행된다. DCT 기반 동영상 복호화 모드일 경우에는 그림 3(A)가 수행되며, 비트평면 단위 복호화 모드일 경우에는 (그림 3(B))가 수행된다. 비트평면 단위 복호화에서는 복호화할 비트평면 정보

를 통해 부호화 과정에서 정한 비트평면 수만큼 반복해서 비트평면 복호화를 수행하고 각각의 비트평면 영상을 결합하여 그레이코드 역변환 과정을 통해 재구성된 깊이정보 맵 블록을 복원한다.

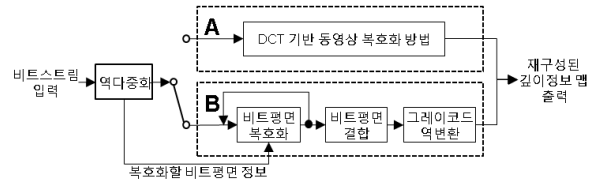


그림 3. 제안하는 방법의 복호화 블록도
 Fig. 3. Decoder block diagram of proposed method

III. 제안하는 방법의 성능 평가

본 논문에서는 제안하는 방법의 성능을 평가하기 위해서 H.264/AVC의 참조 소프트웨어인 JM(Joint Model) 13.2^[10]에 실제 구현하여 블록 기반 적응적 깊이정보 맵 부호화 방법^[6]과 제안하는 방법의 비교를 수행하였다. 실험에 사용한 영상은 Microsoft에서 제공한 "Ballet", "Breakdancers" 영상의 깊이정보 맵이며, 해상도는 1024x768이고 실험에 사용된 양자화 변수는 22, 27, 32, 37이고 예측구조는 I-P-P를 사용하였으며, 엔트로피 부호화 방법은 CAVLC (Context-based Adaptive Variable Length Coding)^[5]와 CAE(Context-based Arithmetic Encoding)^[9]를 사용하였다.

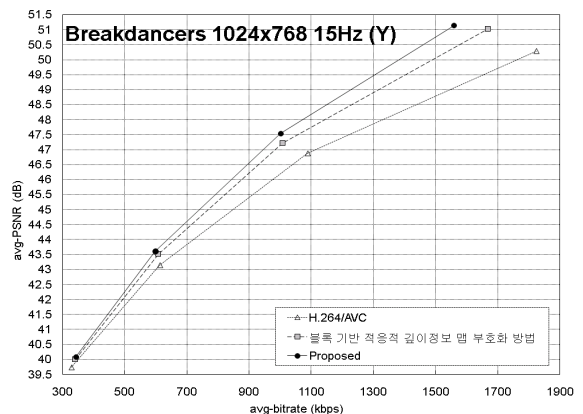


그림 4. "Breakdancers" 영상의 깊이정보 맵 실험 결과 그래프
 Fig. 4. Rate-distortion curves for the "Breakdancers" depth-map sequence

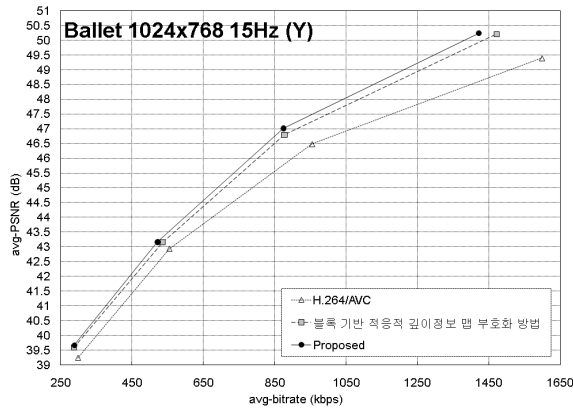


그림 5. "Ballet" 영상의 깊이정보 맵 실험 결과 그래프
Fig. 5. Rate-distortion curves for "Ballet" depth-map sequence

그림 4와 그림 5는 H.264/AVC와 블록 기반 적응적 깊이 정보 맵 부호화 방법^[6]과 본 논문에서 제안하는 방법 ("Proposed")에 대한 비트율 대비 PSNR(Peak Signal-to-Noise Ratio)의 R-D(Rate-distortion) 곡선을 나타낸다. 그래프 결과를 통해 제안하는 방법의 결과가 기존 방법의 결과보다 성능이 향상된 것을 확인할 수 있다. 표 1은 블록 기반 적응적 깊이정보 맵 부호화 방법과 제안하는 방법 간의 PSNR(dB) 및 비트율 결과를 비교한 것이다. 비교 결과, "Ballet"과 "Breakdancers" 깊이정보 맵 모두 저비트율에서 큰 변화가 없었지만, 고비트율로 갈수록 비트율이 줄거나 혹은 화질이 향상되었다.

IV. 결론

본 논문에서는 비트율-왜곡 최적화 방법을 이용하여 부호화할 비트평면의 수를 블록 단위로 결정하게 함으로써 깊이정보 맵의 객체 경계부분에 더 높은 중요도를 부여하는 방법을 제안하였다. 실험결과, "Ballet" 깊이정보 맵의 경우에는 기존 방법에 비해 제안하는 방법이 평균적으로 약 0.2 dB 이상 높았으며, 평균적으로 2.97%의 비트율 절감을 얻을 수 있었고, "Breakdancers" 깊이정보 맵의 경우에는 기존 방법에 비해 제안하는 방법이 평균적으로 약 0.28 dB 이상 높았으며, 평균적으로 3.89%의 비트율 절감을 얻을 수 있었다.

표 1. 제안하는 방법의 부호화 효율 비교
Table 1. Coding efficiency comparisons of the proposed method

QP	Breakdancers				Ballet			
	블록 기반 적응적 깊이정보 맵 부호화 방법[6]		제안하는 방법		블록 기반 적응적 깊이정보 맵 부호화 방법[6]		제안하는 방법	
	Rate (kbps)	PSNR (dB)	Rate (kbps)	PSNR (dB)	Rate (kbps)	PSNR (dB)	Rate (kbps)	PSNR (dB)
37	341.20	40.00	343.71	40.08	287.08	39.59	288.82	39.68
32	608.34	43.51	598.86	43.60	535.96	43.15	521.81	43.17
27	1007.76	47.21	1001.98	47.53	877.39	46.80	873.84	47.02
22	1667.15	51.02	1557.99	51.13	1472.12	50.21	1421.19	50.24

참고 문헌

- [1] ISO/IEC JTC1/SC29/WG11, "Text of ISO/IEC 14496-10:200X/FDAM 1 Multi-view Video Coding," N9978, Hannover, Germany, July 2008.
- [2] ISO/IEC JTC1/SC29/WG11, "Vision on 3D Video," N10357, Lausanne, Switzerland, February 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, November 2007.
- [4] 김경용, 박광훈, 서덕영, "비트평면 기반 무손실 깊이정보 맵 부호화 방법," 방송공학회논문지, 제14권 제 5호, pp.551 ~ 560, September 2009.
- [5] ITU-T Recommendation H.264/AVC and ISO/IEC 14496-10 (MPEG-4 Part 10 AVC), "Advanced Video Coding for Generic Audiovisual Services," Version 1: March 2003, Version 2: May 2004, Version 3: March 2005, Version 4: September 2005, Version 5 and Version 6: June 2006, Version 7: April 2007, Version 8: July 2007.
- [6] 김경용, 박광훈, 서덕영, "적응적 블록기반 깊이정보 맵 부호화 방법," 방송공학회논문지, 제14권 제 5호, pp.601 ~ 615, September 2009
- [7] T. Wiegand, H. Schwarz, A. Joch, F. Kossentini, and G.J. Sullivan, "Rate-Constrained Coder Control and Comparison of Video Coding Standards," IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13, No. 7, pp. 688-703, July 2003.
- [8] ISO/IEC 14496-2 (MPEG-4 Visual), "Coding of Audio-Visual Objects - Part 2: Visual," Version 1: April 1999, Version 2: February 2000, Version 3: May 2004.
- [9] N. Brady, and F. Bossen, "Shape compression of moving objects using context-based arithmetic encoding," Signal Processing: Image Communication, Volume 15, pp. 601-617(17), Number 7, May 2000.
- [10] F. Heinrich-Hertz-Institut, "H.264/AVC Reference Software Version JM13.2," <http://iphome.hhi.de/suehring/tml>, May 2008.