

MPEG-4 SA-DCT 기반의 경계 보존 방법을 이용한 깊이 영상 압축

김동현 서정동 손광훈
연세대학교 전기전자공학부

khsohn@yonsei.ac.kr

Depth Image Compression based on a MPEG-4 SA-DCT for the Edge Preserving Method

Donghyun Kim Jungdong Seo Kwanghoon Sohn
School of Electrical and Electronic Engineering, Yonsei University

요약

멀티미디어 처리 분야의 급속한 발전으로 인해 3차원 TV (3DTV)는 차세대 방송 시스템 시장에서 가장 주목을 받는 제품이 되었다. 3DTV는 사용자가 원하는 시점을 자유롭게 선택할 수 있고, 입체감을 제공하여 사용자가 마치 그 곳에 있는 듯한 효과를 줄 수 있다. 지금까지 입체 영상은 스테레오 영상을 기반으로 하나의 시점에 대한 입체 영상을 제공했지만 최근에는 다시점 영상을 이용하여 다양한 위치에서의 입체 영상을 제공하는 기술이 연구되고 있다.

다시점 영상은 사용자에게 임의 시점의 영상에 대한 시청을 가능케 하여 입체감 있는 화면을 제공할 수 있다. 입체감 있는 영상을 만들기 위해서는 다시점 영상의 시점 간 가상 시점을 생성할 수 있도록 하고 깊이 정보를 포함하고 있는 깊이 영상 (Depth Image)을 획득하여야 한다. 획득된 깊이 영상 데이터와 다시점 비디오 데이터를 동시에 전송하는 다시점 비디오 시스템이 상용화되기 위해서는 방대한 양의 데이터를 효율적으로 압축하는 다시점 비디오 부호화 기술 개발이 필수적이다. 본 논문에서는 기존의 컬러 영상의 효율적인 압축 방법을 제안하던 다시점 비디오 부호화 기술에 국한되지 않고 3차원 영상 화질을 객관적으로 높일 수 있도록 깊이 영상의 효율적 압축 기법에 대한 새로운 방법을 제안한다.

1. 서론

최근 차세대 방송 시스템 시장에서 고품질 TV (HDTV)의 뒤를 이을 제품으로 보다 사실적이고 현장감을 느낄 수 있도록 하는 3차원 TV (3DTV)가 주목 받고 있다. 최근 국내의 다수의 디스플레이 업체들의 3차원 디스플레이, 휴대 단말기에서 사용 가능한 3DTV 서비스 기술 개발 등의 움직임을 비롯해, 유럽을 중심으로 한 3D4YOU 그룹에서는 3D 비디오 콘텐츠에 따른 포맷 기술 개발, 3D@Home을 중심으로 가정에서의 3차원 비디오 서비스 개발 등을 통해 3D 비디오가 주목 받고 있음을 확인할 수 있고, 이에 따라 3DTV의 발전 가능성도 주목 받고 있다.[1]

3DTV의 관심도가 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 영상을 사용하는 Multi-view Video plus Depth (MVD) 데이터를 획득 및 부호화하고, 이를 전송하여 수신단에서 중간 영상을 생성하고 다시점 3D 비디오를 생성하는 것을 목표로 하여 표준화가 진행 중이다.[2] 따라서 2차원 영상을 전송할 때와는 달리 3차원 영상을 전송할 때는 깊이 정보를 포함하고 있는 영상도 전

송해야 하고 전송 시 발생하는 깊이 영상의 왜곡은 컬러 영상의 왜곡에 비해 3차원 영상 화질의 열화에 많은 영향을 끼치므로 깊이 영상을 효율적으로 압축, 전송할 수 있는 기술이 필요하다.

특히 물체의 경계 부분에서의 왜곡은 깊이 영상을 이용하여 가상시점 생성에 영향을 주기 때문에 더욱 심한 화질의 열화를 가져온다. 본 논문에서는 깊이 영상을 압축할 때 경계를 보존하는 효율적인 압축 방법을 제안한다. 깊이 영상을 압축 시 기존의 블록 단위 변환 과정으로 경계 부분이 보존되지 못하는 방법을 블록 단위 내부를 물체와 경계로 분류하여 각각 따로 변환하는 방법을 통해 경계 부분에서 발생하는 왜곡을 최소화 한다.

논문의 구성은 2장에서 기존의 H.264를 이용한 깊이 영상 압축 방식[3]과 Platelet 기반의 깊이 영상 부호화 방법[4]에 대해 서술하고, 3장에서는 본 논문에서 제안하는 MPEG-4 SA-DCT (Shape-adaptive Discrete Cosine Transform) 에 기반한 새로운 변환 방법에 대해 설명한다. 4장에서 실험 결과를 보여주고, 5장에서 결론을 맺는다.

2. 깊이 영상 압축 연구 동향

현재 깊이 영상 압축에 관한 연구는 컬러 영상의 압축과는

달리 많은 연구가 진행되어 있지 않다. 본 논문에서는 컬러 영상의 압축 효율이 뛰어난 비디오 부호화 기술로 깊이 영상을 부호화한 결과[3]를 정리하고, 최근에 발표된 Platelet 기반의 깊이 영상 부호화 방법[4]에 대해 언급한다.

가. H.264 기반의 깊이 영상 압축

H.264는 현재 컬러 영상의 압축 효율이 가장 뛰어난 비디오 부호화 표준이다. 대표적으로 H.264/AVC (Advanced Video Coding)는 DMB, DVB-H 등의 모바일 TV와 HD-DVD, Blu-ray 등의 차세대 DVD에 채택되어 사용되고 있고, H.264/MVC (Multi-view Video Coding)는 다시점 비디오의 압축을 위해 채택된 표준으로 향후 몇 년간 H.264는 지속적으로 사용될 것으로 예상된다.

최근 3차원 영상에 대한 관심이 늘어나 다시점 영상의 압축에 대한 연구가 진행되면서 처음에는 H.264/AVC를 이용하여 다시점 영상의 각 시점의 영상을 개별적으로 부호화하는 방법[5]이 제시되었다. H.264/MVC는 H.264/AVC의 단점을 보완하기 위해 Hierarchical B pictures 구조[6]를 이용하여 시점간의 예측 정보와 시간축 상에서의 예측 정보를 이용하여 다시점 컬러 영상을 부호화 하는 표준으로 채택되었다. 그림 1은 본 논문에서 실험할 Microsoft Reseach (MSR)에서 제공하는 Ballet 영상의 다시점 깊이 영상을 H.264/AVC를 이용하여 각 시점의 영상을 개별적으로 부호화한 결과와 H.264/MVC를 이용하여 부호화한 결과를 나타낸다.

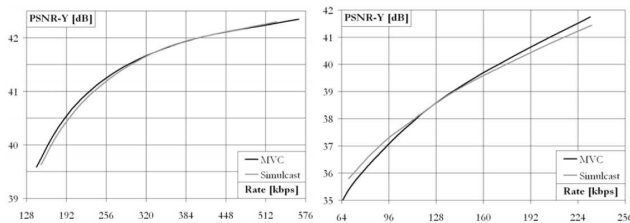


그림 1. Results of coding experiments for multi-view data (left: color video, right: depth video) [3]

그림 1에서 다시점 영상을 부호화할 때 H.264/MVC를 이용하여 부호화한 결과가 H.264/AVC를 이용하여 각 시점의 영상을 개별적으로 부호화한 결과보다 효율적임을 확인할 수 있다. 하지만 깊이 영상이 컬러 영상에 비해 정보가 회색음영의 형태로 단순하고 압축 시 경계의 보존이 잘 이루어지지 않아 동 비트율 대비 PSNR 값이 낮음을 확인할 수 있다.

나. Platelet 기반의 깊이 영상 코딩

3차원 영상에서는 물체 내부의 깊이 정보 왜곡으로 인해 열화 되는 화질 보다 물체 경계 부분의 깊이 정보의 왜곡으로 생기는 열화가 더욱 사용자의 눈에 쉽게 띈다. 이 알고리즘은 깊이 영상의 대부분이 평활한 평면이라는 특성을 기본으로 블록 단위로 전송할 때 한 블록 안에 최대한 한 평면만이 존재하도록 Quadtree 분할 방법[7]을 이용하여 재귀적으로 평면을 분할하였

다. 한 블록 안의 평면은 그림 2와 같이 4가지 형태로 분류할 수 있다.

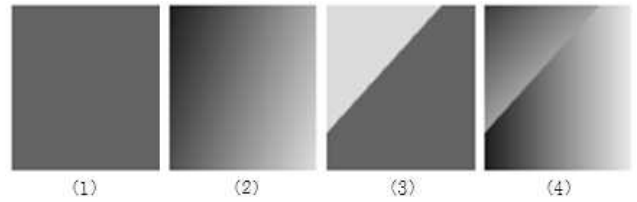


그림 2. Example patterns for quadtree segmented plane[8]

그림 2의 (1)과 (3)은 평면의 방정식이 상수 함수인 군, (2)와 (4)는 평면의 방정식이 선형 함수인 군으로 정의한다. 각각의 평면은 식 1의 함수를 통해 평면의 방정식으로 표현된다.

$$\begin{aligned}
 \hat{f}_1(x,y) &= \alpha_0 \\
 \hat{f}_2(x,y) &= (\beta_0 + \beta_1x + \beta_2y), \text{ for } (x,y) \in S \\
 \hat{f}_3(x,y) &= \begin{cases} \hat{f}_{3A}(x,y) = \gamma_{0A} & \text{for } (x,y) \in A \\ \hat{f}_{3B}(x,y) = \gamma_{0B} & \text{for } (x,y) \in B \end{cases} \\
 \hat{f}_4(x,y) &= \begin{cases} \hat{f}_{4A}(x,y) = \theta_{0A} + \theta_{1A}x + \theta_{2A}y & \text{for } (x,y) \in A \\ \hat{f}_{4B}(x,y) = \theta_{0B} + \theta_{1B}x + \theta_{2B}y & \text{for } (x,y) \in B \end{cases}
 \end{aligned} \tag{1}$$

각 평면은 Rate-Distortion Decision Criterion에 기반하여 선택되며, 이를 통하여 R-D constraint를 만족 시키는 최적의 압축을 목적으로 한다. 이와 같은 방법으로 원본 깊이 영상을 부호화하면 H.264 Intra로 부호화한 것에 비해 형체 내부의 열화는 심해지지만 경계 부분은 더욱 날카롭게 잘라 선명해진다. 새로운 가상 시점을 생성하여 확인하면 H.264에 비해 Platelet 기반의 압축이 경계 부분에서 보다 선명한 영상을 만들어 낸다.[8]

물체 내부의 경우 대부분이 그림 2의 (2), (4)의 예와 같은 평면으로 분류되는데, 그림 2의 (2), (4)의 경우 식 1과 같은 평면의 방정식으로 표현 시 발생하는 에러로 인해 기존의 부호화 결과 보다 효율이 떨어진다. 또한 Platelet 기반의 깊이 영상 부호화 방법은 R-D constraint를 만족할 때까지 Quadtree 분할을 재귀적으로 하기 때문에 복잡도가 증가한다는 단점이 있다. 본 논문에서는 MPEG-4의 SA-DCT[9]에 기반하여 기존의 압축 방법 보다 효과적인 물체의 경계 보존하고 물체 내부에서의 압축 효과를 유지하는 방법을 제안한다.

3. MPEG-4 SA-DCT 기반의 깊이 영상 압축

깊이 영상은 회색음영 형태로 물체와 배경의 구분이 확실하기 때문에 경계 부분이 잘 보존 되어야 한다. 본 논문은 깊이 영상의 경계부분을 보다 효율적으로 보존하며 압축하기 위해서 H.264의 Discrete Cosine Transform (DCT)의 고주파 성분을 제거하여 경계부분을 열화 시키는 단점을 보완하는 방법을 제안한다.

가. MPEG-4 Shaped Adaptive DCT (SA-DCT)

깊이 영상을 압축할 때는 경계 부분의 보존이 중요하다. 회색영으로 표현되는 깊이 영상은 물체와 배경의 밝기 차이가 확실해 경계 부분에서 급격한 색의 변화가 있다. 이런 변화는 주파수 축으로 변환할 경우 고주파 성분을 띄게 된다. 하지만 H.264로 깊이 영상을 부호화 할 경우 색의 변화가 크지 않은 평면 부분에서는 압축 효율이 좋지만 경계 부분에서는 고주파 성분을 제거하는 H.264의 DCT의 특성 때문에 열화가 더욱 심해진다. 이러한 단점을 보완하기 위해 본 논문에서는 MPEG-4에서 사용되는 SA-DCT를 사용하기로 한다.

SA-DCT는 물체 기반의 DCT로써 영상을 8x8 블록으로 나누고, 각 블록을 물체와 배경을 각각 전경과 후경으로 분리해서 각각 따로 주파수 영역으로 변환하여 전송함으로써 물체와 배경 경계의 열화를 최소화 시킬 수 있다. 그림 3은 SA-DCT의 수행 과정이다.

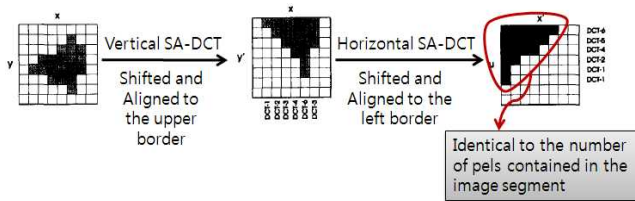


그림 3. Example of SA-DCT processing in an arbitrarily shape image block

그림 3에서 검은 부분은 전경에 해당하고, 하얀 부분은 후경에 해당한다. 우선 전경 부분만 변환하기 위해 세로로 1차원 변환하고, 가로로 1차원 변환한다. SA-DCT의 변환 공식은 식 2와 같다.

$$\begin{aligned}
 \frac{DCT-N(p,k)}{k,p=0 \rightarrow N-1} &= c_0 \cdot \cos \left[p \left(k + \frac{1}{2} \right) \cdot \frac{\pi}{N} \right] \\
 p : pth \text{ DCT basis vector} & \qquad \qquad \text{식 2} \\
 c_0 &= \sqrt{1/2}, \text{ if } p = 0 \\
 c_0 &= 1 \qquad \qquad , \text{ otherwise}
 \end{aligned}$$

H.264의 DCT의 블록 단위는 4x4, 8x8, 16x16으로 유동적이고 정수화 되어 있지만, SA-DCT는 8x8, 16x16으로 제한되어있고 정수화 되어있지 않다. 본 논문에서는 SA-DCT의 효율을 극대화하기 위해서 4x4 정수화 SA-DCT (Integer SA-DCT)[10]를 경계 부분에 적용한다. 평면 부분에서는 H.264의 DCT의 효율과 SA-DCT의 효율 비슷하므로 H.264의 DCT를 사용하기로 한다.

4. 실험 결과

본 논문에서는 깊이 영상에 H.264의 4x4 DCT와 MPEG-4의 4x4 SA-DCT를 적용한 효율을 비교하기 위해서 MSR에서 제공되는 Ballet 영상을 사용한다. Ballet 영상은 1024x768의 크기며 8개의 컬러와 깊이 영상이 제공된다.[11] 그림 4는 8개의

시점으로 구성된 Ballet 영상의 컬러와 깊이 영상의 원본 중 하나이다.



그림 4. Original Ballet Color Image (left), Depth Image (right)

실험은 테스트 영상에 양자화 레벨 변화에 따른 열화 정도와 경계 부분 보존의 효율에 초점을 맞춘다. 본 논문에서는 H.264의 DCT와 MPEG-4의 SA-DCT의 효율을 비교하기 위하여 테스트 영상을 변환한 직후의 결과 영상들을 확인한다. H.264의 DCT와 MPEG-4의 SA-DCT 각 방법에 양자화 레벨을 10, 18, 24, 30, 36, 42에 맞추어 수행한다. 그림 5는 제안된 방법을 사용하여 양자화 레벨을 36으로 하고 얻은 결과이다.



그림 5. Using H.264 DCT (left), Using MPEG-4 SA-DCT (right) (QP=36)

일반적으로 H.264의 DCT는 물체와 배경을 구분하지 않고 블록 전체를 변환하여 물체와 배경의 경계 부분에서 정보 손실이 많아 열화가 심해진다. MPEG-4의 SA-DCT는 물체와 배경을 구분하여 블록을 따로 변환한다. 그림 6은 그림 5의 원으로 표시된 부분을 확대한 것이다.

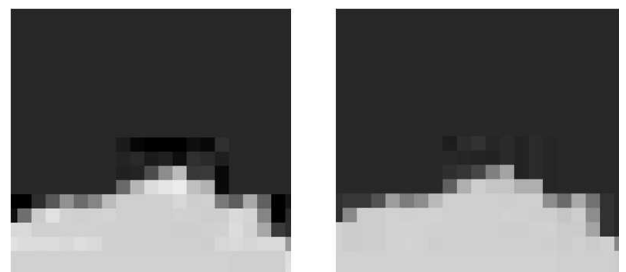


그림 6. Enlarged Image of H.264 DCT (left), Enlarged Image of MPEG-4 SA-DCT (right)

그림 6에서 볼 수 있듯이 SA-DCT로 깊이 영상을 변환한 경

우 DCT로 변환한 경우 보다 경계 부분에서의 Ringing Artifact 가 더 적음을 알 수 있다. 표 1은 각각의 결과 영상의 화질을 PSNR로 비교한 것이다.

QP	10	18	24	30	36	42
DCT	40.26	40.26	40.21	39.93	38.90	36.36
SA-DCT	50.62	49.44	47.68	44.74	41.71	36.26

표 1. QP 변화에 따른 DCT와 SA-DCT의 PSNR 변화

양자화 값이 커질수록 SA-DCT로 깊이 영상을 변환한 경우 물체 경계 부분은 보존 되지만 물체 내부는 DCT에 비해 열화가 심해지므로 표 1과 같은 결과가 발생된다.

5. 결론

본 논문에서 제안한 알고리즘을 실험한 결과, H.264의 일반 적인 DCT보다 MPEG-4의 SA-DCT를 기반으로 하여 깊이 영상을 변환했을 때, 경계 부분에서의 열화를 줄여 비교적 영상의 경계 부분을 보존하는 효과를 얻을 수 있다. 본 논문에서의 실험 환경이 DCT와 SA-DCT의 단순한 효율 비교였기 때문에 비트 율 제어에 따라 달라지는 결과를 확인할 수 없었다. 하지만 깊이 영상의 경계 부분에서 SA-DCT의 보존 효율을 확인하였기 때문에 추후에 실제의 H.264의 참조 소프트웨어에 SA-DCT를 혼합하여 비트율에 따른 결과를 확인하고자 한다.

또한, SA-DCT는 전경과 후경을 분리하기 위해 Image Segment 기술이 필요하다. 하지만 Image Segment 기술에 따라 압축효율의 차이를 가져올 수 있으므로 MPEG-4에서는 전경은 백색, 후경은 흑색으로 이루어진 Segment 파일을 사용하여 영상을 분리한다. 따라서 깊이 영상뿐만 아니라 이러한 파일을 추가적으로 사용함에 따라 전송 용량이 증가한다는 단점이 있다. 추후에 본인은 전송 시 Segment 파일이 필요하지 않도록 기존의 H.264에 추가적으로 Image Segment 기술을 적용할 예정이다.

6. Reference

[1] 엄기문, 방건, 허남호, 김진웅, "3D 비디오 MPEG 표준화 동향", 전자통신동향분석 제24권 제3호, 2009년 6월
 [2] ISO/IEC JTC1/SC29/WG11, "Call for Contribution on 3D Video Test Material(Update)", N9595, Jan. 2008.
 [3] P. Merkle, A. Smolic, K. Müller, and T. Wiegand, "Multi-view Video Plus Depth Representation and Coding", *Proc. IEEE International Conference on Image Processing*, San Antonio, TX, USA, Sept. 2007.
 [4] P. Merkle, Y. Morvan, A. Smolic, D. Farin, K. Müller, P.H.N. de With, and T. Wiegand, "The Effectss of multiview depth video compression on multiview rendering", *Signal processing: Image communication*,

Vol. 24, Issues 1-2, Jan. 2009, pp. 73-88.

[5] H. Schwarz, D. Marpe, and T. Wiegand, "Analysis of hierarchical B pictures and MCTF", *ICME 2006, IEEE International Conference on Multimedia and Expo*, Toronto, Ontario, Canada, July. 2006.
 [6] T. Wiegand, G.J. Sullivan, G. Bjøntegaard, and A. Lutha, "Overview of the H.264/AVC video coding standard", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 13, No. 7, p. 560, July. 2003.
 [7] R. Shukla, P. L. Dragotti, M. N. Do, and M. Vetterli, "Rate-distortion optimized tree structured compression algorithms for piecewise smooth images", *IEEE Trans. Image Proc.*, vol. 14, pp. 343 - 59, Mar. 2005.
 [8] Y. Morvana, P.H.N. de Witha, D. Farina, "Platelet-Based Coding of Depth Maps for the Transmission of Multiview Images", *Proceedings of SPIE, Stereoscopic Displays and Applications*, vol. 6055 p. 93-100, Jan. 2006, San Jose (CA), USA
 [9] Sikora, T, and B. Makai, "Shape-adaptive DCT for generic coding of video", *IEEE Trans. Circuits Syst. Video Technology.*, vol. 5, no. 1, pp. 59-62, 1995.
 [10] X. Li, E. Edirisinghe, and H. Bez, "Shape Adaptive Integer Transform for coding arbitrarily shaped objects in H.264/AVC", *Proc. SPIE 2006*, vol. 6077, 60770C, Jan. 2006.
 [11] <http://research.microsoft.com>