

3차원 비디오의 합성영상 경계 잡음 제거

이도훈, 양윤모, 오병태

한국항공대학교

dolmengei@kau.ac.kr, yym064@naver.com, byungoh@kau.ac.kr

Boundary artifacts reduction for synthesized views in 3D video

DoHoon Lee, Yoonmo Yang, Byung Tae Oh
Korea Aerospace University

요 약

본 논문에서는 3 차원 영상시스템에서 발생하는 경계 잡음을 Projection onto Convex Sets (POCS) 방법을 활용하여 제거하는 기법에 대해서 소개한다. 3 차원 영상시스템에서의 경계 잡음은 일반적으로 손상된 깊이영상 의하여 가상 시점의 합성 과정에서 발생하는 잡음을 뜻한다. 본 논문에서는 이러한 경계 잡음을 분석하고, 제안 방식을 이용하여 효과적으로 제거할 수 있음을 여러 실험을 통하여 확인하였다.

This paper proposes a boundary artifacts reduction method for synthesized views in 3D video system using Projection onto Convex Sets (POCS). In 3D video system, boundary artifacts are occurred in synthesized views by distorted depth map. In this paper, we analyze boundary artifacts, and show the experimental results by the proposed algorithm on various situations.

1. 서론

다수의 시점에 대한 영상을 통하여 사용자에게 보다 실감적인 영상을 제공하는 3차원 영상 시스템에 대한 관심이 점차 커지고 있다. 다수의 시점에 대한 영상을 표현하기 위하여 많은 3차원 영상 시스템들이 깊이영상을 활용한 합성 방식을 사용하고 있으며, 이 방식을 통하여 다수의 시점에 대한 영상을 표현하기 위해 필요한 방대한 데이터 크기를 효율적으로 감소시켰다. 이러한 방식의 대표적인 예로 다시점+깊이영상 (Multiview plus depth map) 방식이 있으며, 현재 활발하게 연구되고 있다 [1-3].

하지만 깊이영상을 통한 합성은 깊이영상의 상태에 따라 합성영상의 화질이 크게 좌우된다. 따라서 기존의 영상과 다른

성질을 가지는 깊이영상을 효율적으로 처리하기 위하여 많은 연구가 진행되고 있다 [4-11]. 또한 현재 깊이영상은 구조광 카메라, Time-of-flight (ToF) 카메라 등을 이용하여 직접적으로 획득하거나, 혹은 스테레오 영상에 대한 매칭 방식을 이용하여 간접적으로 획득할 수 있고, 각각의 방식마다 장단점이 존재한다. 우선 구조광 카메라의 경우 실내에서만 사용가능하며, ToF 카메라는 하드웨어의 기술적 한계로 인하여 기존 영상 대비 낮은 해상도의 영상을 획득할 수 밖에 없다. 스테레오매칭을 통하여 획득한 깊이영상은 매칭방식의 한계 등으로 인하여 특정 영상의 경우 직접적으로 획득한 영상과 비교하여 정확도가 크게 떨어질 수 있다. 따라서 다양한 방법을 통하여 적절한 깊이영상을 획득하는 연구 또한 다양하게 진행되고 있다 [12-13].

깊이영상을 원본에 가깝게 복원하여 합성영상의 화질을 향상시키는 방식이 아닌, 합성된 영상에서 발생하는 경계

잡음의 성질을 이용하여 이를 제거하는 연구도 다양하게 진행되고 있다. 먼저 경계 잡음이 발생하는 부분에 대하여 한 쪽 시점에서 warping된 영상만으로 대체하는 방식 [14]과 참조 영상을 다른 시점의 참조 영상까지 warping하여 신뢰할 수 있는 정보인지 아닌지를 판단하여 합성을 만드는 방식 [15]이 있으며, 더불어 전경에 대한 깊이영상의 경계가 컬러영상의 경계보다 배경 쪽으로 확장함으로써 경계 잡음을 제거하는 방식 [16] 등이 있다. 본 논문에서는 합성된 영상에서 경계 잡음의 성질을 분석하고, POCS를 활용하여 경계 잡음을 제거하고자 한다.

본 논문의 구성은 다음과 같다. 2절에서는 본 논문에서 제거하고자 하는 경계 잡음에 대하여 소개 및 분석하고, 본 논문에서 제안한 POCS를 활용한 경계 잡음 방식을 소개한다. 3절에서는 2절에서 소개한 방식의 성능을 다시점+깊이영상 시스템에서 실험을 통하여 확인한다. 마지막으로 4절에서는 본 논문에 대한 결론을 맺는다.

2. 제안 알고리즘

A. 경계 잡음

3차원 비디오 시스템에서 가상의 시점에 대한 영상은 깊이영상을 통하여 합성된다. 하지만 깊이영상은 앞장에서 언급하였듯이 기존의 컬러영상과 다른 특징을 가지고 있다. 특히, 다시점+깊이영상 시스템에서 깊이영상은 그림 1과 같이 기존의 컬러영상과 같은 방식으로 부호화 및 복호화가 이루어진다. 복호화 된 깊이영상은 압축 에러가 존재하고, 이는 참조 영상의 값이 가상의 시점에서 잘못된 위치로 warping되게 된다. 이러한 현상으로 발생하는 잡음을 경계 잡음이라 하고, 이에 대한 모델은 그림 2와 같다. 실제 가상시점 영상에서 경계 잡음은 그림 3과 같다.

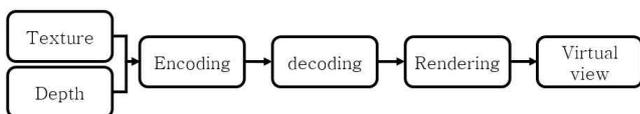


그림 1. 다시점+깊이영상 시스템 구성도.

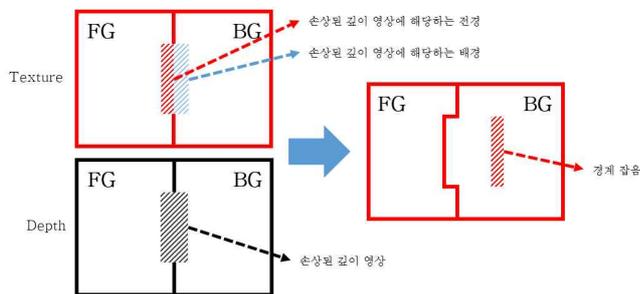


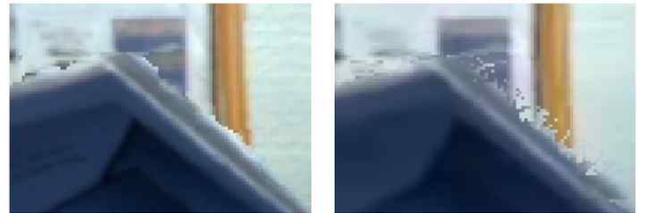
그림 2. 경계 잡음의 발생.



(a) Undo Dancer



(b) Shark



(c) Newspaper

그림 3. 좌영상 : 원본 합성영상, 우영상 : 손상된 깊이영상에 의하여 발생한 경계 잡음이 존재하는 영상.

B. 제안 알고리즘

가상시점 영상의 경계 잡음을 제거하는데 있어서 중요한 점은 전경과 배경의 올바른 정보는 유지한 채 경계 잡음만을 제거해야 한다는 것이다. 기존의 일반적인 필터를 사용할 경우 경계잡음은 제거가 되지만 전경과 배경의 올바른 정보 또한 손상되게 된다. 이러한 문제를 해결하기 위하여 본 논문에서는 POCS 기법을 활용하여 경계 잡음을 제거하고자 한다.

가상시점의 영상은 두 개의 참조 영상에서 warping된 두 영상의 합성을 통하여 생성된다. 이때 경계 잡음은 그림 4와 같이 대부분 warping된 두 영상 중 한 쪽 영상에서만 나타나는 것을 확인 할 수 있다. 또한 경계 잡음은 충격 잡음과 비슷한 성질을 가지며, 이러한 잡음을 포함한 영역을 주파수 도메인으로 변환할 경우 고주파 성분이 강하게 나타난다. 이러한 성질을 이용하여 본 논문에서 제안하는 알고리즘은 그림 5와 같이 POCS를 활용하여 진행한다. POCS란 몇 개의 constraint set을 설정하고 반복적인 연산을 수행하여 최적의 값을 찾아가는 방식이다 [17-18].

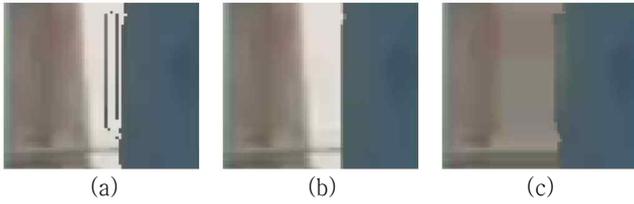


그림 4. (a) 가상 시점의 영상, (b) 우영상을 가상 시점으로 warping한 영상, (c) 좌영상을 가상 시점으로 warping한 영상.

제안 알고리즘에서 사용한 constraint set은 두 개이며, 먼저 첫 번째 set은 warping 된 우영상과 좌영상의 DCT domain에서 AC 성분에 대한 에너지의 집중된 정도를 비교한 후, 집중도가 강한 영상을 선택하고, 수식 (1) 과 같이 가상 시점 영상의 AC 성분을 밀집된 영상의 AC성분으로 투영하였다

$$B_{1AC}(u, v) \rightarrow B_{\text{compact AC}}(u, v) \quad (1)$$

수식 (1)에서 B_1 은 가상 시점의 영상에서 처리하고자 하는 블록에 대하여 DCT를 취한 블록이고, B_{compact} 는 warping 된 우영상과 좌영상 중에서 B_1 에 해당하는 블록에 대해 DCT를 취한 블록 중 AC 성분에 대한 에너지가 밀집된 블록이다.

또한 본 논문에서는 위와 같은 연산을 수행할 경우 전경과 배경의 올바른 정보 또한 손상되는 것을 방지하기 위하여 다른 constraint set을 설정하였다. 이에 따라 두 번째 constraint set을 앞서 언급하였듯이 경계 잡음이 충격 잡음과 유사한 성질을 갖는다는 점에 착안하여 수식 (2)와 같이 원본의 가상 시점의 영상과 median 필터의 영상의 픽셀 값으로 지정하였다.

$$\text{out}(i, j) = [\min(\hat{n}(i, j), f(i, j)), \max(\hat{n}(i, j), f(i, j))] \quad (2)$$

수식 (2)에서 \hat{n} 은 입력된 가상 시점의 영상이고, f 는 median 필터를 취한 영상이며, i 와 j 는 각각 픽셀의 좌표다.

3. 실험 결과

제안방식의 실험 및 평가를 위하여, 본 논문에서는 현재 Joint Collaboration 표준화 활동을 통해 3D video 압축 표준을 제정하고 있는 JCT-3V의 Common Test Condition (CTC)을 따르고, 그곳에서 사용하고 있는 테스트 영상을 이용하여 실험을 하였다. 실험을 위한 reference S/W 로서 영상 압축은 3D-ATM v15.0 [19]을 사용하였고, 비교 평가를 위한 중간영상 생성은 VSRS general mode 를 이용하였다. 보다 자세한 실험영상 세팅 및 실험조건은 CTC 문서를 참조한다 [20].

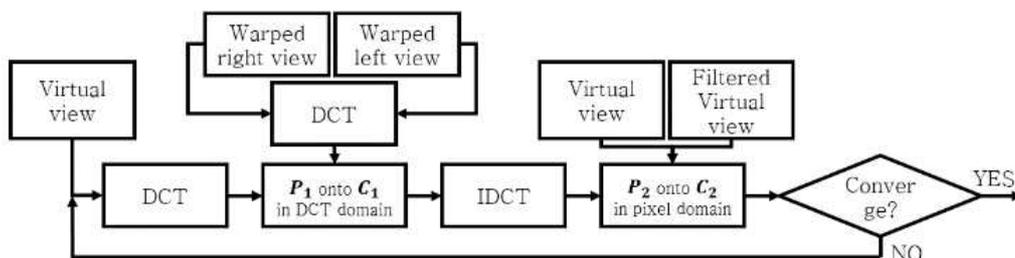


그림 5. 제안 알고리즘

본 논문에서는 3D-AVC 기반에서 깊이영상의 해상도에 따른 성능을 객관적으로 비교하기 위해 PSNR 및 비트량과 PSNR에 기반한 BDBR을 이용하였다.

표 I. BDBR 관점에서의 성능 평가.

Sequence	Proposed algorithm
Poznan Hall2	0.02
Poznan Street	-0.96
Undo Dancer	-9.84
GT Fly	0.20
Kendo	-0.28
Balloons	-4.54
Newspaper	-3.15
Shark	-9.18
Avg.	-3.47

표 II. PSNR 관점에서의 성능 평가

Sequence	R1	R2	R3	R4
Poznan Hall2	0.000	0.000	0.000	-0.003
Poznan Street	0.007	0.019	0.034	0.060
Undo Dancer	0.146	0.198	0.294	0.505
GT Fly	0.008	0.006	-0.008	-0.039
Kendo	0.004	0.007	0.012	0.029
Balloons	0.061	0.142	0.203	0.305
Newspaper	0.056	0.091	0.120	0.162
Shark	0.031	0.127	0.654	1.580
Avg.	0.039	0.074	0.164	0.325

위 결과를 분석해보면, Undo Dancer 와 Shark 와 같이 압축 후에도 깊이영상이 컬러영상의 정보와 비교적 매칭이 잘 되는 영상에서 많은 성능 향상이 되었음을 확인 할 수 있었다. 이러한 이유는 깊이영상이 압축으로 인하여 크게 손상될 경우 정확한 가상시점에서의 warping 된 영상을 얻기 힘들며, 이로 인하여 POCS 의 projection 될 블록이 잘못된 블록일 확률이 커지기 때문이다. 이는 표 II 의 rate 별 PSNR 관점에서의 성능 변화로도 확인 할 수 있는데, 낮은 QP 로 압축했을 경우 깊이영상의 손상이 비교적 적게 발생하고 이에 따라 비교적 정확한 projection 시킬 블록을 획득할 수 있기 때문에 본 논문의 알고리즘에 대한 성능 향상이 큰 것을 확인할 수 있다.



(a) Undo Dancer



(b) Shark



(c) Newspaper

그림 6. 좌영상 : 원본 합성영상, 가운데영상 : 손상된 깊이영상에 의하여 발생한 경계 잡음이 존재하는 영상, 우영상 : 제안 알고리즘 후 영상.

4. 결론

본 논문은 POCS 를 활용하여 깊이영상의 손상으로 인하여 가상의 시점에 발생하는 경계 잡음을 제거하는 기법을 소개하였다. 실험 결과 제안 방식이 손상된 깊이영상으로 인하여 발생하는 많은 경계 잡음을 효과적으로 제거하였음을 확인할 수 있었다. 추후 연구에서는 깊이영상 기반의 영상 합성 과정에서 발생하는 경계잡음을 제거하기 위한 보다 효율적인 접근 방법에 대해 연구를 진행할 예정이다.

감사의 글

이 논문은 2013 년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (NRF-2013R1A1A1057779).

5. 참조문헌

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

[2] A. Smolic, K. Mueller, N. Stefanoski, J. Ostermann, A. Gotchev, G. B. Akar, G. Triantafyllidis, and A. Koz, "Coding algorithms for 3DTV a survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 11, pp. 1606-1621, Nov. 2007.

[3] P. Merkle, A. Smolic, K. Muller, and T. Wiegand, "Multiview video plus depth representation and coding," in *Proc. IEEE Int. Conf. on Image Processing*, 2007, pp. 201-204.

[4] Zhao, Yin, et al. "Boundary artifact reduction in view synthesis of 3D video: from perspective of texture-depth alignment." *IEEE Trans. Broadcasting*, vol. 57, no. 2, pp. 510-522, 2011.

[5] 3D-CE3.a results on dilation filter for depth post processing, ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11, Doc. JCT2-A0038, Stockholm, 2012.

[6] K.J. Oh, S. Yea, A. Vetro, Y.S. Ho, "Depth reconstruction filter and down/up sampling for depth coding in 3-D video," *IEEE Signal Processing. Lett.*, vol. 16, no. 9, pp. 747-750, 2009.

[7] K.J. Oh, A. Vetro, Y.S. Ho, "Depth coding using a boundary reconstruction filter for 3D video system," *IEEE Trans. Circuits System. Video Technology*, vol. 21, no. 3, pp. 350-359, 2011.

[8] J. Kopf, M. Cohen, D. Lischinski, M. Uyttendaele, "Joint bilateral upsampling", *Proc. ACM Siggraph*, vol. 26, no. 3, pp. 96, 2007.

[9] D. De Silva, W. Fernando, H. Kodikaraarachchi, S. Worrall, A. Kondoz, "Adaptive sharpening of depth maps for 3D-TV," *Electronics Lett.*, vol. 46, no. 23, pp. 1546-1548, 2010.

[10] D. Min, J. Lu, M. N. Do, "Depth video enhancement based on weighted mode filtering," *IEEE Trans. Image Processing*, vol. 21, no. 3, pp. 1176-1190, 2012.

[11] Y.S. Kang, S.B. LEE, Y.S. Ho, "Depth map upsampling using depth local features," *Electronics Lett.*, vol. 50, no. 3, pp. 170-171, 2014

[12] Q. Yang, L. Wang, R. Yang, H. Stewenius, and D. Nister, "Stereo Matching with Color-Weighted Correlation, Hierarchical Belief Propagation and Occlusion Handling," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, June 2006.

[13] J. Sun, Y. Li, S. Kang, and H.Y. Shum, "Symmetric Stereo Matching for Occlusion Handling," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, vol. 2, pp. 399-406, June 2005.

[14] Lee, Cheon, and Yo-Sung Ho. "Boundary filtering on synthesized views of 3D video." *Future Generation Communication and Networking Symposia*, 2008. FGCNS'08. Second International Conference on. Vol. 3. IEEE, 2008.

[15] Yang, Lu, et al. "Error supression in view synthesis using reliability reasoning for FTV." *3DTV-Conference: The True Vision-Capture, Transmission and Display of 3D Video (3DTV-CON)*, 2010. IEEE, 2010.

[16] Zhao, Yin, et al. "Boundary artifact reduction in view synthesis of 3D video: from perspective of texture-depth alignment." *Broadcasting, IEEE Transactions on* 57.2 (2011): 510-522.

[17] Y. Kim, C.S. Park, and S.J. Ko. "Fast POCS based post-processing technique for HDTV." *Consumer Electronics, IEEE Transactions on* 49.4 (2003): 1438-1447.

[18] G.R. Kwon, H.K. Kim, Y. Kim, S.J. Ko. "An efficient POCS-based post-processing technique using wavelet transform in HDTV." *Consumer Electronics, IEEE Transactions on* 51.4 (2005): 1283-1290.

[19] *3D-AVC Test Model 5*, ITU-T SG16 WP3 ISO/IEC JTC1/SC29/WG11, Doc. JCT3V-C1003, Geneva, Jan. 2013.

[20] *Common test condition of 3DV core experiments*, ITU-T SG16 WP3 ISO/IEC JTC1/SC29/WG11, Doc. JCT3V-C1100, Geneva, Jan. 2013.