

## 깊이영상 해상도 조절에 따른 3 차원 비디오 부호화 성능 분석

이도훈, 양윤모, 오병태

한국항공대학교

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

### Analysis of Depth Map Resolution for Coding Performance in 3D Video System

Do Hoon Lee, Yun mo Yang, Byung Tae Oh  
Korea Aerospace University

#### 요 약

본 논문에서는 다시점 깊이영상 기반 3 차원 영상시스템에서 깊이영상의 해상도에 따른 압축성능을 비교 실험한다. 다시점 깊이영상 시스템에서 깊이영상은 중간시점 합성을 위해 사용되며, 깊이영상은 중간영상의 화질에 큰 영향을 미치게 된다. 본 논문에서는 3 차원 영상시스템에서 깊이영상 해상도에 따른 중간영상의 화질변화를 분석하고, dilation filter 로 인한 성능변화를 비교 분석한다.

This paper provides the coding performance comparisons of depth map resolution in 3D video system. In multiview plus depth map system, depth map is used for synthesis view rendering, and affects to synthesis views quality. In the paper, we show the experimental results as depth map resolution in 3D video system, and show performance variation as dilation filter.

#### 1. 서론

차세대 영상 시스템으로서 기존의 2차원 영상에서 벗어나 깊이 감까지 느낄 수 있는 3차원 실감영상에 대한 관심이 점차 커지고 있다. 이중 가장 대중화된 방식으로 양안에 서로 다른 시점의 영상을 보여주는 스테레오스코픽 기반의 방식을 들 수 있다. 이는 현재 제품으로까지 출시되어 시청자가 안방에서까지 편하게 3차원 실감 영상을 시청할 수 있게 되었다. 하지만 이 스테레오스코픽 기반의 방식은 고정된 시점의 영상만 시청할 수 있다는 명백한 한계를 지니고 있다. 이러한 한계를 극복하기 위하여 다수의 시점을 표현할 수 있도록 여러 시점에 따른 영상을 사용하는 다시점 방식 (Multiview)이 새롭게 각광받고

있다. 하지만 다시점 영상 시스템은 여러 시점에 따른 많은 수의 영상을 동시에 처리해야 하므로, 방대한 데이터 크기에 따른 전송 및 압축 등에 대한 문제점이 있다. 이를 해결하기 위해서 깊이영상을 이용한 다시점+깊이영상 (Multiview plus depth map) 방식이 현재 활발하게 연구되고 있다[1-3].

다시점 영상 압축은 시점간 중복성을 제거하는데 초점이 맞추어져 있으며, 이는 2000년대 중반부터 활발하게 연구되어, Multiview Video Coding (MVC) 표준까지 제정되었다. 하지만, 깊이영상의 경우는 기존 영상과는 전혀 다른 성질을 가진 콘텐츠로서, 현재 깊이영상은 KINECT, TOF(Time of frequency) 카메라, 혹은 스테레오매칭 방식 등을 이용하여 획득하는 방법이 많이 사용되고 있다. 하지만 KINECT의 경우 실내에서만 사용할 수 있고, TOF 카메라는 하드웨어상의 문제로 인하여 기존 영상의 해상도에 비해 많이 낮은 해상도의

영상을 얻을 수밖에 없다. 또한 스테레오매칭 방식은 사물의 경계부근에서 매우 흔들림이 심하여 디자이너의 보정을 거치지 않고서 사용하기에는 매우 부적절한 영상이 획득되는 문제점을 갖고 있다. 따라서 그 밖에 다양한 방법을 이용하여 깊이영상을 획득하는 방식에 대한 연구가 활발하게 진행되고 있다[4-5].

본 논문에서는 압축 전 깊이영상의 해상도를 낮추고, 압축 후 해상도를 복원하여 합성하는 방식이 다시점 영상 성능에 어떠한 영향을 미치는지 분석하고 있다. 이를 위해 깊이영상의 폭과 높이를 각각 2배, 4배, 8배로 조절하면서 깊이영상의 해상도에 따른 성능 비교를 해보고자 한다.

본 논문의 구성은 다음과 같다. 2절에서는 본 논문에서 수행한 down- 및 up-sampling 의 방식을 소개하고, 3절에서는 2절에서 소개한 방식을 적용하여 조절한 깊이영상의 해상도에 따른 성능을 실험을 통해서 확인한다. 마지막으로 4절에서는 본 논문에 대한 결론을 맺는다.

## 2. Up/Down – sampling

3차원 비디오 시스템에서 깊이영상의 해상도 조절은 압축 전 down-sampling 한 후, 복호화된 깊이영상이 중간영상 합성을 위한 입력으로 들어가기 전 up-sampling을 통하여 이루어진다. 특히 3D-AVC기반 3차원 시스템에서는 깊이영상의 해상도를 원본의 1/4로 축소하였을 때의 성능이 좋다는 것은 실험적으로 확인되었다 [6]. 이와 같은 결과를 바탕으로 본 논문에서는 깊이영상의 해상도에 따라 3차원 시스템의 압축이 어떤 효율을 가지게 되는지 실험을 진행하였다. 이때, 각각의 down- 및 up-sampling은 아래와 같은 방식을 사용하였다.

### A. Down-sampling

현재 3D-AVC 기반 3 차원 비디오 시스템에서 깊이영상의 down-sampling 은 기존 Multiview video coding (MVC)의 필터를 사용하고 있다. MVC 의 Reference S/W 인 JSVM 에서는 다양한 down-sampling 방식을 제공하고 있는데, 이중 dyadic down-sampling (MPEG-4 down-sampling filter)방식이 사용되고 있다. 보다 자세한 필터계수는 아래의 수식 (1)과 같다.

$$[2, 0, -4, -3, 5, 19, 26, 19, 5, -3, -4, 0, 2] \quad (1)$$

### B. Up-sampling

Up-sampling 은 영상처리 분야에서 꾸준히 연구되고 있는 분야로서, nearest neighborhood interpolation, bilinear interpolation, bicubic interpolation, new edge directed interpolation (NEDI), soft-decision interpolation (SAI)등 다양하게 존재한다. 3D-AVC 에서는 위 방식 중 속도면에서 탁월한 bilinear interpolation 을 택하여 압축시 입력된 깊이영상의 해상도가 컬러영상보다 작은 경우 up-sampling 을 수행한다.

### C. Dilation filter

압축된 깊이영상은 경계부근에서 artifact 가 발생하는데, 이는 합성과정에서 악영향을 미친다. Dilation filter 는 3D-AVC 에 내장되어있는 post filter 로 [7], 이러한 문제점을 해결하기 위하여 주위 픽셀들의 값 중에서 최대값을 취함으로써 foreground 를 background 로 확장하여 경계 밖에서 artifact 가 일어나도록 하는 방식을 취한 filter 다.

## 3. 실험 결과

제안한 실험 및 평가를 위하여, 본 논문에서는 현재 Joint Collaboration 표준화 활동을 통해 3D video 압축 표준을 제정하고 있는 JCT-3V 의 Common Test Condition (CTC)을 따르고, 그곳에서 사용하고 있는 테스트 영상을 이용하여 실험을 하였다. 실험을 위한 reference S/W 로서 영상 압축은 현재 개발 진행중인 3D-ATM v9.0 [8]을 사용하였고, 비교 평가를 위한 중간영상 생성은 1D fast VSRS [9]를 이용하였다. 보다 자세한 실험영상 세팅 및 실험조건은 CTC 문서를 참조한다 [10].

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

표 I. 깊이영상의 높이 변화에 따른 성능 비교.

	2x	4x	2x+dilation	4x+dilation
Poznan Hall2	-1.5	-5.78	-0.62	-5.58
Poznan Street	-5.75	-8.42	-6.86	-10.83
Undo Dancer	4.38	7.09	1.12	1.83
GT Fly	-5.45	-5.47	-9.17	-8.67
Kendo	-11.18	-19.04	-9.41	-18.33
Balloons	-3.93	-5.19	1.6	-4.46
Newspaper	-4.17	-3.38	-0.64	-3.03
<b>Avg.</b>	<b>-3.94%</b>	<b>-5.74%</b>	<b>-3.43%</b>	<b>-7.01%</b>

표 II. 깊이영상의 넓이 변화에 따른 성능 비교

	2x	4x	2x+dilation	4x+dilation
Poznan Hall2	3.91	2.23	2.81	0.7
Poznan Street	-5.18	-7.13	-7.68	-10.28
Undo Dancer	38.55	81.95	10.31	52.01
GT Fly	-4.78	-3.72	-8.94	-7.16
Kendo	-9.97	-14.88	-9.29	-16.02
Balloons	3.72	10.19	2.84	4.57
Newspaper	-1.09	8.93	-1.21	2.73
<b>Avg.</b>	<b>3.59%</b>	<b>11.08%</b>	<b>-1.59%</b>	<b>3.79%</b>

위 결과를 분석해보면, Undo Dancer 는 다른 sequence 와 비교하여 ground-truth 에 가까운 특징을 가진다. 따라서 down- 및 up-sampling 과정을 통하여 정보 손실이 발생할 때, 다른 sequence 에 비하여 성능이 많이 떨어짐을 확인할

수 있었다. 또한, 아래 그림 1 과 그림 2 에서 확인할 수 있듯이 수평방향의 경계정보가 수직방향의 경계정보보다 월등히 많아서 수평방향으로 down- 및 up-sampling 을 진행 할 때, 더욱 성능이 하락함을 확인할 수 있었다. 또한, 현재 MVD-system 의 특성상 수평방향으로 합성이 진행되므로, 전반적으로 모든 sequence 에 걸쳐 수직방향으로 down- 및 up-sampling 을 진행하는 경우가 수평방향으로 down- 및 up-sampling 된 경우보다 성능이 평균 10%이상 좋은 것을 확인할 수 있었다.

그러나 앞 절에서 언급한 dilation filter 에 의한 성능 향상은 수평방향으로 down- 및 up-sampling 이 수행된 경우가 수직방향으로 down- 및 up-sampling 이 수행된 경우보다 크게 향상됨을 확인할 수 있었는데, 이는 앞서 언급하였듯이 다시점 깊이영상 기반 3 차원 영상 시스템은 수평방향으로 합성이 진행이 되어, dilation filter 가 수평방향에서 down- 및 up-sampling 과정에서 발생하는 정보손실을 수직방향에서 발생하는 정보손실 보다 효율적으로 보완하기 때문이다.

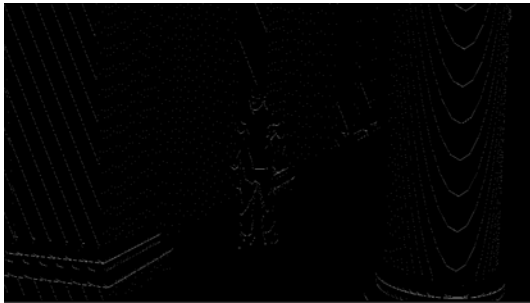


그림 1. Undo Dancer 의 수직방향의 경계

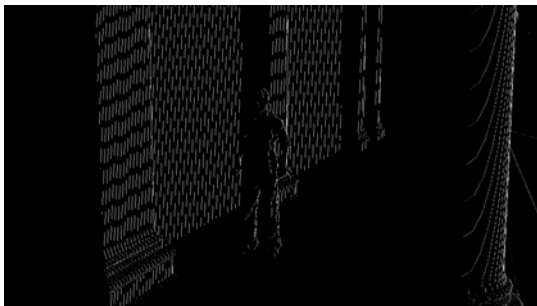


그림 2. Undo Dancer 의 수평방향의 경계

#### 4. 결론

본 논문은 깊이영상의 다양한 해상도에 따른 3 차원 영상시스템의 압축 성능을 비교하였다. 성능 비교는 3D-AVC 에서 down- 및 up-sampling 과정을 통하여 수행하였다. 실험결과 수평방향보다 수직방향에 대하여 down- 및 up-sampling 을 수행하는 것이 성능이 우수함을 확인할 수 있었다. 추후 연구에서는 깊이영상의 특징에 맞춰 다양한 접근을 통하여 보다 압축 성능이 좋은 알고리즘에 대한 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] 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.
- [5] 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.
- [6] "Common test conditions for 3DV experimentation," ISO/IEC JTC1/SC29/WG11 MPEG2012/N12560, Feb. 2012.
- [7] 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, Jul. 2012
- [8]
- [9] *3D-AVC Test Model 5*, ITU-T SG16 WP3 ISO/IEC JTC1/SC29/WG11, Doc. JCT3V-C1003, Geneva, Jan. 2013.
- [10] *3D-HEVC Test Model 3*, ITU-T SG16 WP3 ISO/IEC JTC1/SC29/WG11, Doc. JCT3V-C1005, Geneva, Jan. 2013.
- [11] *Common test condition of 3DV core experiments*, ITU-T SG16 WP3 ISO/IEC JTC1/SC29/WG11, Doc. JCT3V-C1100, Geneva, Jan. 2013.