

깊이정보 기반의 혼합 가우시안 분포 히스토그램과 Mean Shift Filter를 이용한 깊이정보 맵 부호화 전처리

박성희 유지상

광운대학교 전자공학과

iloveu089@kw.ac.kr jsyoo@kw.ac.kr

Depth Map coding pre-processing using Depth-based Mixed Gaussian Histogram and Mean Shift Filter

Park, Sung-Hee Yoo, Ji-Sang

Dept. Electronic Engineering, Kwangwoon University

요약

본 논문에서는 MPEG의 3차원 비디오 시스템의 표준 깊이정보 맵에 대한 효율적인 부호화를 위하여 전처리 방법을 제안한다. 현재 3차원 비디오 부호화(3DVC)에 대한 표준화가 진행 중에 있지만 아직 깊이정보 맵의 부호화 방법에 대한 표준이 확정되지 않은 상태이다. 제안하는 기법에서는 우선, 입력된 깊이정보 맵에 대하여 원래의 히스토그램 분포를 가우시안 혼합모델(GMM)기반의 EM 군집화 기법에 의한 방법으로 분리 후, 분리된 히스토그램을 기반으로 깊이정보 맵을 여러 개의 영상으로 분리한다. 그 후 분리된 각각의 영상을 배경과 객체에 따라 다른 조건의 mean shift filter로 필터링한다. 결과적으로 영상내의 각 영역 경계는 최대한 살리면서 영역내의 화소 값에 대해서는 평균 연산을 취하여 부호화시 효율을 극대화 하고자 하였다. 실험조건은 1024×768 영상에 대해서 50 프레임으로 H.264/AVC base 프로파일로 부호화를 진행하였다. 최종 실험결과 bit rate는 대략 23% ~ 26% 정도 감소하고 부호화 시간도 다소 줄어드는 것을 확인 할 수 있었다.

1. 서론

방송 통신 기술의 발달에 힘입어 전송되는 디지털 미디어 콘텐츠의 품질이 많이 향상되었다. 또한 흑백 TV부터 최근 고해상도의 3D LED TV 까지 디스플레이의 눈부신 기술발전은 사용자에게 더 높은 품질의 콘텐츠를 제공하는 밑거름이었다. 영화 <아바타>의 흥행이 많은 사람들의 3차원 영상에 대한 눈높이를 높였고, 이는 TV 기술과 3차원 비디오 기술의 발전으로 계속 이어지고 있다. 이 두 가지 기술은 서로 결합되어 실감형 미디어로 서비스 될 수 있으며, 기존의 2차원 비디오보다 실재감과 몰입감을 제공한다.

실감형 미디어의 중요한 요소인 3차원 영상 데이터의 획득, 처리, 전송, 재생 기술은 3차원 비디오 부호화(3DVC)의 표준화 과정을 통하여 점차 구체화 되고 있다. 3차원 비디오 시스템은 양안식(stereoscopic) 입체영상보다 많은 시점의 영상을 이용하여 3차원 비디오 서비스를 제공하는 기술을 말한다. ISO/IEC JCT1 산하의 MPEG(Moving Picture Experts Group)에서는 2002년부터 3차원 비디오 시스템의 중요성을 인식하고, 관련 국제표준을 만들기 위한 노력을 계속해 왔다.[1] 3차원 비디오 부호화(3DVC) 표준은 다수의 카메라로부터 입력 받은 영상 이외에도 사용자가 원하는 다양한 가상 시점의 영상들을 깊이정보 맵을 이용한 시점 보간(view interpolation) 방법을 통하여 무한대까지 생성할 수 있도록 지원한다. 따라서 3차원 비디오 부호화(3DVC) 표준에서는 소수의 시점(view)영상 데이터와 그것의 깊이정보 맵만을 전송하면 되기 때문에, 대역폭과 저장 공간을 절약할 수 있는 장점이 있다.[2]

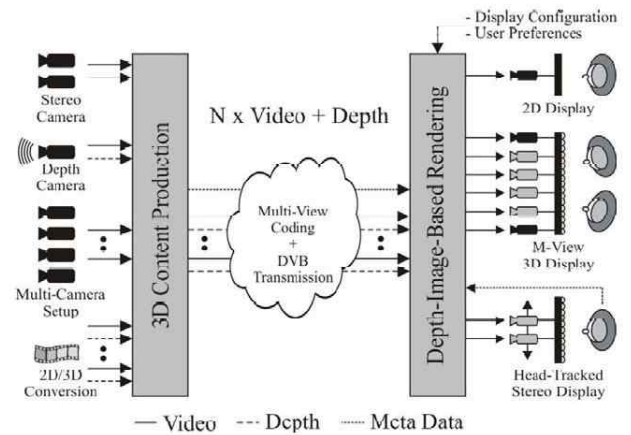


그림 1. 3차원 비디오 시스템의 개념도

(출처: ISO/IEC JTC1/SC29/N9784 Introduction to 3D Video)

그림 1은 3차원 비디오 시스템의 개념도이다. 다시점 카메라를 이용하여 여러 시점의 비디오를 획득하고, 획득된 영상의 효과적인 전송을 위하여 다시점 비디오 부호화를 수행한다. 수신단에서는 전송된 부호화된 비트스트림을 다시 다시점 영상으로 복호화하고, 복호된 다시점 영상을 이용하여 사용자가 원하는 시점의 영상을 새롭게 생성한다. 3차원 비디오 부호화(3DVC) 표준에서 표준화 대상은 입력 받은 다시점 영상을 부호화하는 방법과 깊이정보 맵을 부호화하는 방법, 그리고 깊이정보 맵을 이용하여 가상시점 영상을 생성하는 방법이다.[3] 입력된 다

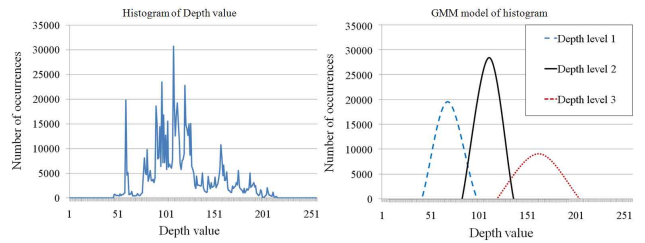
시점 영상에 대한 부호화는 이미 표준화 작업이 끝난 다시점 비디오 부호화(MVC: multi-view video coding) 방법을 이용할 수 있다. 깊이정보 맵의 부호화는 H.264/AVC나 현재 표준화가 진행 중인 High Efficiency Video Coding(HEVC)방법이 예상된다.[4] 제한된 채널을 통해 다시점 영상과 각 영상에 대응하는 깊이 데이터(MVD: multi-view video plus depth)를 전송하기 때문에 전달되어야 하는 영상 데이터의 양이 커지게 된다. 그러므로 보다 효율적인 비디오 부호화 기술이 요구된다. 즉, 가상시점 영상의 합성을 위해서, 다시점 영상뿐만 아니라 깊이 정보 맵의 고효율 부호화 방법이 반드시 필요하다고 할 수 있다.

본 논문에서는 깊이정보 맵의 특성을 고려한 깊이정보 맵 부호화의 전처리 과정을 제안한다. 깊이정보 맵(Depth-map)이란, 카메라와 객체 간의 실제 거리를 정수단위로 표현한 것으로써, 카메라와 가까울수록 큰 값(밝은 값)으로 표현된다. 따라서 영상내의 객체나 배경은 카메라와 거리에 따라 어느 정도 군집화(clustering)되어 있는 것이 특징이다. 제안된 기법에서는 이러한 특성에 따라 깊이정보 맵을 분리하여 각각 필터링하게 된다.

2. 제안하는 전처리 방법

그림 2는 깊이정보 맵 부호화를 위한 전처리 과정의 흐름도이다. 먼저 깊이정보 맵을 객체와 배경의 깊이에 따라 여러 개의 영상으로 분리한다. 분리하는 방법으로 가우시안 혼합 모델(GMM: Gaussian Mixture Model)기반의 EM(Expectation Maximization)군집화 기법을 적용한다. 그림 3은 깊이정보 맵의 히스토그램 분포를 서로 다른 평균값과 분산을 가지는 여러 개의 혼합된 가우시안 모델의 조합으로 나타낸 결과이다. 두 개의 가우시안 모델의 겹쳐진 부분에 대해서는 군집화 과정이 필요함을 알 수 있다.[5] 0부터 255사이의 모든 깊이정보 맵의 값이 여러 가우시안 모델가운데 어느 한 모델에 해당되는 지를 좀 정확하게 추정하기 위해 각 모델에 포함될 확률을 반복적으로 계산하고 비교하는 EM군집화 기법을 사용한다. 따라서 가우시안 모델들이 서로 겹쳐진 부분의 깊이정보가 어느 모델에 해당하는지 알 수 있다. 군집화된 결과를 바탕으로 영상을 분리하면 그림 4와 같다. 세 개로 분리된 깊이정보 맵 Depth level1 부터 Depth level3 은 그림 3.(b)의 가우시안 모델 Depth

level1 부터 Depth level3을 따르게 된다.



(a) 히스토그램 분포 (b) 가우시안 혼합 모델
그림 3. 깊이 영상의 히스토그램 분포와 가우시안 혼합 모델링

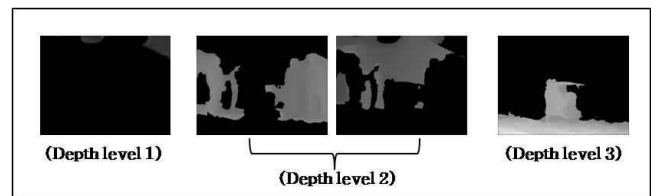


그림 4. EM군집화 기법에 의해 분리된 영상

이렇게 분리된 영상에 각각 mean-shift filter를 적용한다. mean-shift filter를 통과한 영상들은 영역간의 경계는 보존되면서 영역 내의 화소 값들은 평균화 된다. 혼합 가우시안 모델에 의해 나누어진 각 영상들은 깊이 정보에 따라 배경과 객체들로 나누어지는데, 본 논문에서 제안한 기법은 깊이정보를 고려하여 분리된 각 영상에 mean shift filter를 다르게 적용한다는 것이다. 즉, 멀리 있는 배경이나 객체는 평균화 연산을 많이 하고, 가까이 있는 객체는 평균화 연산을 적게 하여 왜곡을 최소화 한다. 평균화 연산을 통하여, 깊이정보 맵에 존재하는 잡음과 같은 고주파 성분을 제거하여 부호화 효율을 극대화할 수 있다. 분리된 각 영상을 필터링한 후 다시 결합하는 과정을 거친 결과는 그림 5와 같다. 배경과 멀리 있는 객체에 대해서 블러링 효과를 확인 할 수 있다. 반면 가까이 있는 출구는 사람은 거의 변화가 없음을 알 수 있다.[6][7]

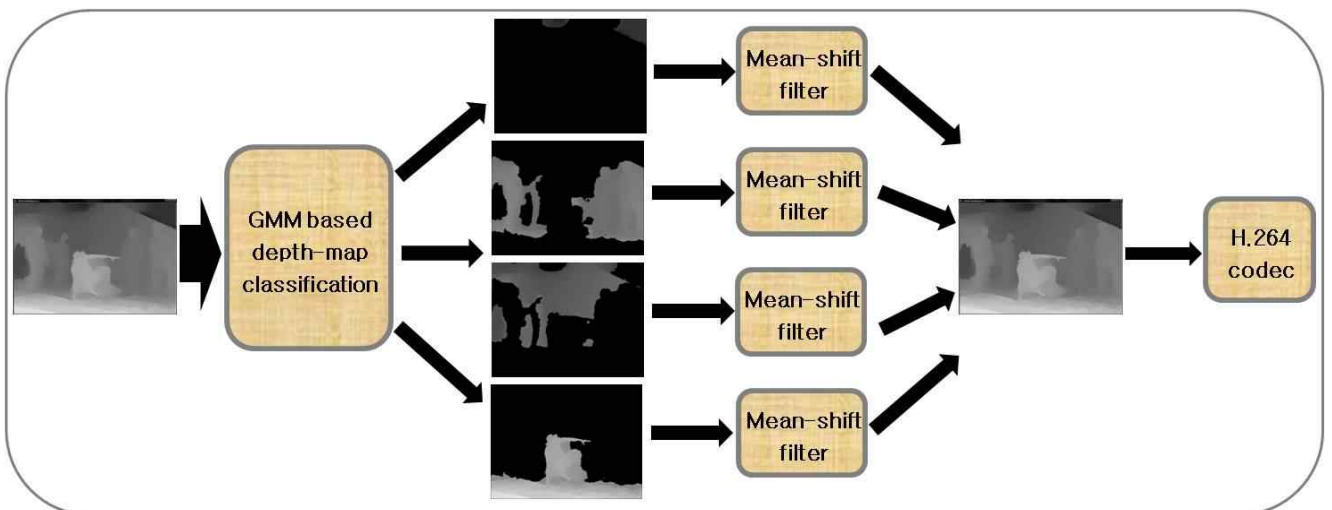


그림 2. 깊이정보 맵을 부호화하기 위한 전처리 과정의 흐름도

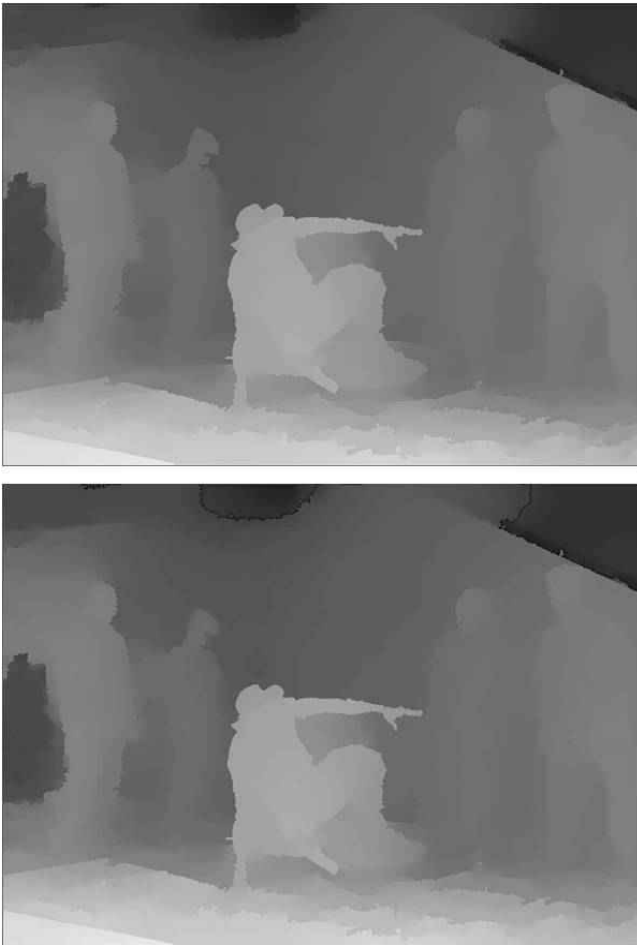


그림 5.
원래의 깊이정보 맵(위)
Mean shift filter 처리 후 다시 결합한 영상(아래)

3. 실험 결과

본 논문에서 제안하는 방법에 대한 실험 조건은 다음과 같다. 해상도 (1027×768)의 breakdancers와 lovebird1 영상을 각각 50 프레임으로 H.264/AVC BASE 프로파일로 로 부호화를 하였다. 결과는 표1, 표2와 같다.

표. 1. Breakdancer

	Original	Proposed (전처리 과정 적용)
Bit-rate(kbit/s)	4305	3283
Encoding time (fps)	0.61	0.49

표. 2. Lovebird1

	Original	Proposed (전처리 과정 적용)
Bit-rate(kbit/s)	3117	2306
Encoding time (fps)	0.53	0.44

표. 1 의 실험 결과를 보면, bit-rate 는 23% ~ 26% 정도 감소한 결과를 보이고, 한 프레임 당 encoding time도 다소 감소하는 것을 확인할 수 있었다.

4. 결론

본 논문에서는 3차원 비디오 표준의 일부분인 깊이정보 맵의 부호화 수행 시 효과적인 부호화를 위한 전처리과정에 대한 연구를 수행하였다. 결과적으로 부호화 성과와 시간이 다소 향상되는 결과를 확인할 수 있었다. 향후, 본 논문에서 제안한 전처리 방법을 적용한 깊이정보 맵을 가지고 다시 가상시점의 영상을 합성을 하는 과정과 깊이정보를 전 처리하는 과정에서 화소 값에 대한 왜곡을 최소화하는 연구가 이루어져야 할 것이다.

감사의 글

※본 연구는 한국 산업기술 평가관리원(KEIT)의 IT산업 원천 기술개발 사업의 일환으로 수행하였음. [KI002058, 대화형 디지털 홀로그램 통합서비스 시스템의 구현을 위한 신호처리 요소 기술 및 SoC 개발]

참 고 문 헌

- [1] ISO/IEC JTC1/SC29/WG11, "Vision on 3D Video," N10357, Lausanne, Switzerland, February 2009.
- [2] 호요성, 이천, "3차원 비디오 부호화를 위한 국제 표준화 기술" 한국 방송공학회 학술지, 제 14권, pp. 031-044, 2009. 06.
- [3] 엄기문, 방건, 허남호, 김진웅, "3D 비디오 MPEG 표준화 동향." 전자통신동향분석, 제 24권 제3호, 2009 06.
- [4] 장세윤, 최진수, 김동형, 정원식, 문경애, 홍진우 "H.264 이후의 차세대 비디오 부호화 표준화 동향 및 전망" 전자통신동향분석 제 23권 제 1호 2008 2월
- [5] 이동석, 유지상, "깊이정보에 따른 레이얼별 히스토그램 매칭을 이용한 조명 불일치 보상 기법", 2010년 8월 한국통신학회, 제 35권 제 8호 pp. 651-660.
- [6] Dorin Comaniciu, Peter Meer, "Mean Shift Analysis and Applications", Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on, vol.2 1197 - 1203
- [7] Dorin Comaniciu, Peter Meer, "Mean Shift: A Robust Approach Toward Feature Space Analysis", IEEE Trans. , vol 24, no 5, May 2002.