

미디어 필터를 이용한 깊이영상 부호화의 인트라 예측모드 설계

이천, 허진, 호요성
광주과학기술원
{leecheon, jinheo, hoyo}@gist.ac.kr

Modified Intra Prediction for Depth Map Coding using Median Filter

Cheon Lee, Jin Heo and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)

요 약

본 논문에서는 깊이영상 부호화 효율을 높이기 위한 새로운 인트라 예측모드를 제안한다. 기존의 H.264/AVC의 인트라 예측모드는 복호된 13개의 인접 화소를 이용하여 9가지 방법으로 부호화할 블록을 채운 후, 가장 효율적인 예측모드를 선택하여 부호화한다. 이때, 인접 화소들의 값을 가중합으로 연산하여 예측 방향에 따라 채운다. 하지만 이 방법은 영상의 특성이 단순한 깊이영상의 부호화에는 비효율적이다. 본 논문에서는 참조할 13개의 인접 화소를 미디어 필터를 이용하여 부호화 오류를 제거하여 예측에 이용한다. 또한 미디어 필터를 이용하여 각 방향으로 값을 채운다. 이와 같은 방법을 이용하여 깊이영상을 부호화한 결과 1.5~5.8 %의 비트량이 감소했다.

I. 서론

3차원 비디오(3D Video)는 여러 대의 카메라로 동시에 획득한 다시점 비디오를 시청자에게 보다 실감나는 영상으로 재현하는 기술이다 [1]. 3차원 TV와 자유시점 TV(free-viewpoint TV)가 그 예이다. 이러한 시스템을 통해 사용자는 보다 실감나는 3차원 입체영상을 감상하거나 시청하는 시점을 자유롭게 선택할 수 있다.

최근 동영상 전문가 그룹인 MPEG(moving picture experts group)에서는 3차원 비디오 시스템에 대한 국제 표준화를 진행하고 있다. 3차원 오디오-비주얼(3D audio-visual)을 시작으로 다시점 비디오 부호화(multiview video coding)의 국제표준안이 제정되었다. 또한 이 다시점 비디오 부호화의 경험을 바탕으로 3차원 비디오 부호화(3D video Coding)에 대한 국제표준화 작업이 진행되고 있다 [2].

현재 연구되고 있는 3차원 비디오 시스템은 제한된 수의 다시점 영상을 이용하여 보다 많은 수의 시점을 생성하기 위하여, 깊이영상을 이용한 중간시점 영상합성 기술을 고려하고 있다. 그러므로 전송단에서 컬러영상과 함께 깊이영상을 동시에 부호화해야 한다.

깊이영상의 부호화는 기존의 부호기를 이용하여 부호화할 수 있다. 하지만 기존의 부호기가 컬러영상을 효율적으로 부호화하도록 설계되었기 때문에, 영상의 특성이 다른 깊이영상의 부호화에는 효율적이지 않다.

깊이영상은 장면내의 각 객체에 대한 3차원 거리를 2차원 영상으로 표현한 영상이다. 그러므로 영상이 비교적 단순하고 객체의 경계 부분에서 값의 불연속이 발생한다. 이러한 특징을 이용하면 깊이영상의 부호화에 효율적인 부호화 방법을 개발할 수 있다.

본 논문에서는 깊이영상의 부호화를 위한 효율적인 인트라 4x4 예측 방법을 제안한다. 기존의 인트라 4x4 예측 방법은 복호된 13개의 인접 화소를 이용하여 9개의 방향으로 값을 채운 후, 가장 효율적인 예측모드를 선택하여 부호화한다 [3]. 이때 인트라 예측에 참조할

복호된 13개의 인접 화소에 부호화 오류가 있다면 예측의 정확도가 떨어지게 된다. 또한, 인접 화소들의 가중합으로 각 방향에 따라 값을 채우므로, 깊이영상의 경계 부분에서는 보간된 값으로 인한 잔여 데이터가 발생하게 된다. 이에 본 논문에서는 미디어 필터를 이용하여 참조할 인접 화소의 오류를 줄이는 방법과, 동 필터를 이용하여 방향성을 가지는 6개 예측모드의 예측효율을 개선한 방법을 제안한다.

II. 개선한 인트라 4x4 예측모드

2.1 H.264/AVC의 인트라 4x4 예측 모드

기존의 H.264/AVC의 인트라 4x4 예측모드는 13개의 인접 화소를 이용하여 9가지의 방향으로 부호화할 블록의 정보를 예측하는 방법이다. 수직, 수평, DC 모드를 제외한 나머지 6개 모드는 각 방향에 따라 참조화소를 가중합 연산을 이용하여 값을 채운다. 이렇게 예측된 인트라 모드들 중에서 비트율-왜곡 비용이 가장 작은 모드를 선택하여 부호화한다.

2.2 미디어 필터를 이용한 참조화소 결정

인트라 4x4 예측에 이용되는 인접 화소들은 부호화 과정에서 양자화 오류가 포함된 정보이다. 이 양자화 오류는 깊이영상에서 객체의 경계 부분에서 강하게 나타나는데, 이러한 오류를 포함한 정보는 인트라 예측의 정확도를 낮추어 부호화 효율을 떨어뜨린다. 제안하는 방법은 미디어 필터를 이용하여 양자화 오류를 제거하는 것이다. 깊이영상의 경계 부근에서 발생하는 양자화 오류는 마치 잡음처럼 발생하므로, 잡음 제거에 효과적인 미디어 필터를 이용한다. 식 (1)은 일반적인 미디어 필터를 표현한 것이다. 이와 같은 방법으로 참조할 화소를 결정하면 객체의 경계 부근의 인트라 예측의 정확도가 향상한다.

$$\text{MEDIAN}(x,y,z) = x + y + z - \text{Min}(x,y,z) - \text{Max}(x,y,z) \quad (1)$$

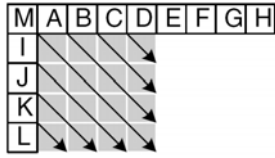


그림 1. 인트라 4x4 예측 모드 4번

2.3 미디언 필터를 이용한 인트라 4x4 예측모드

기존의 인트라 4x4 예측 모드는 인접한 13개의 화소를 이용하여 각 방향에 따라 블록의 값을 채운다. 예를 들어, 그림 1과 같은 인트라 4x4 예측 모드는 오른쪽 대각선 아래 방향으로 값을 채운다. 주목할 것은, 3개의 화소값의 가중합을 각 방향으로 채운다. 식(2)는 현재 블록의 (0,0) 위치의 값을 채울 때 이용하는 방법이다.

$$P(0,0) = \text{ROUND}((M + 2A + I) / 4) \quad (2)$$

위의 방법은 3개 화소의 가중합을 이용하므로 보간된 값으로 대상 화소를 채운다. 하지만, 깊이영상은 객체의 경계 부근에서 값의 차이가 뚜렷하므로 보간한 중간값을 이용하는 것보다 실제 값에 가장 가까운 값을 이용하는 것이 효율적이다. 그러므로 식(3)과 같이 미디언 필터를 이용하여 현재 블록의 (0,0) 위치의 값을 채운다. 이와 같은 방법을 이용하여 나머지 15개의 화소값을 채운다.

$$P(0,0) = \text{MEDIAN}(M, A, I) \quad (3)$$

수직모드(vertical)와 수평모드(horizontal)는 인접 화소의 값을 가중합 없이 바로 이용하고, DC모드는 인접 8개의 화소값의 평균을 이용하므로 변경없이 그대로 사용한다. 3개 모드를 제외한 6개 예측모드를 미디언 필터를 사용하여 변경한다.

III. 실험 결과 및 분석

제안하는 방법의 성능을 평가하기 위해, 깊이맵이 제공되는 다시점 비디오 영상인 "Breakdancers" 와 "Ballet", MPEG의 3차원 비디오 부호화 그룹에서 테스트 영상으로 사용 중인 "Book Arrival"와 "Mobile" 영상을 사용하였다. "Book Arrival" 영상은 깊이영상이 제공되지 않으므로, MPEG에서 제공하는 깊이맵 추정 소프트웨어 (DERS 4.0 [4], depth estimation reference software)를 이용하여 얻은 깊이영상을 이용했다. 반면 "Mobile" 영상은 제공되는 깊이영상을 이용했다.

표 1. "Mobile" 5번 시점의 깊이영상 부호화 결과

QP	PSNR (dB)		비트율 (kbit/s)	
	기존의 방법	제안한 방법	기존의 방법	제안한 방법
17	54.13	53.93	500.16	468.87
22	52.58	52.49	394.00	366.31
27	51.11	51.07	310.96	288.60
32	48.95	49.02	242.46	229.27

제안하는 방법은 H.264/AVC의 참조 소프트웨어인 Jm14.2에 구현했다. 인트라 모드만을 이용하는 I

프레임을 총 100 프레임 부호화했다. 사용한 양자화 계수는 17, 22, 27, 32 이다. 표 1은 "Mobile" 영상의 부호화 결과를 나타낸 것이다. 화질의 변화는 적은 반면 비트량이 상당히 줄어든 것을 확인할 수 있다. 그림 2는 표 1에 대한 결과를 비트율-왜곡 곡선으로 표시한 것이다. 전반적으로 높은 성능 향상을 확인할 수 있다.

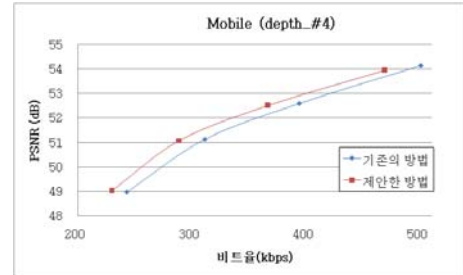


그림 2. "Mobile" 5번 시점의 비트율-왜곡 곡선

표 2는 나머지 테스트 영상에 대한 종합적인 실험 결과를 나타낸 것이다. 전반적으로 0.1~0.4 dB 이상의 성능 화질 개선이 있거나, 1.5~5.8%의 비트량 감소가 발생함을 확인했다. 여기서 특히 주목할 점은 제안하는 방법이 높은 비트율 환경에서 두드러진 성능 개선을 보였다는 것이다. 이는 높은 비트율 환경에서 인접 화소의 신뢰도가 보장되기 때문인 것으로 보인다.

표 2. 제안한 방법의 성능향상

테스트 영상	시점	BDPSNR (dB)	BDBR (%)
Breakdancers	4	0.159	-2.45
Ballet	4	0.176	-2.57
Mobile	5	0.422	-5.80
Book_arrival	4	0.116	-1.59

IV. 결론

본 논문에서는 깊이영상의 부호화 성능을 개선하기 위하여 인트라 4x4 예측 방법을 미디언 필터를 이용하여 설계했다. 기존의 인트라 예측모드는 인접 화소의 값을 가중합을 이용하였으나, 제안하는 방법은 미디언 필터를 이용하여 불필요한 중간값의 생성을 방지했다. 또한 인접 화소의 값 중에서 부호화 오류로 인한 영향을 줄이기 위하여, 인접 화소의 값도 미디언 필터링을 수행했다. 이러한 방법을 적용하여 실험한 결과 평균 3.1%의 비트량이 감소했다.

감사의 글

본 연구는 삼성전자의 지원에 의한 것입니다.

참고문헌

- [1] A. Smolic, K. Müller, P. Merkle, C. Fehn, P. Kauff, P. Eisert, T. Wiegand, "3D Video and Free Viewpoint Video - Technologies, Applications and MPEG Standards," IEEE ICME, July 2006.
- [2] ISO/IEC JTC1/SC29/WG11 "Call for Contributions on 3D Video Test Material," N9595, Jan. 2008.
- [3] ITU-T Rec. H.264 / ISO/IEC 11496-10, "Advanced Video Coding", Final Committee Draft, Do. JVT-F100, Dec. 2002.
- [4] ISO/IEC JTC1/SC29/WG11 "Reference Software of Depth Estimation and View Synthesis for FTV/3DV," M15836, Oct. 2008.