

시점 보간법을 이용한 차분 다시점 비디오 부호화 방법

*이상범 *김준엽 *호요성 **최병호

*광주과학기술원 정보통신공학과

**전자부품연구원

*(sblee, jykim77, hoyo)@gist.ac.kr, **bhchoi@keti.re.kr

Differential Multi-view Video Coding using View Interpolation

*Sang-Beom Lee *Jun-Yup Kim *Yo-Sung Ho **Byeong-ho Choi

*Gwangju Institute of Science and Technology (GIST)

**Korea Electronics Technology Institute (KETI)

요약

3차원 비디오는 차세대 정보 통신 서비스 분야의 하나로, 사용자에게 시각적으로 고차원적인 서비스를 제공하는 것을 목적으로 한다. 이 가운데 다시점 비디오는 같은 시간, 여러 시점에서 영상 정보를 획득하여 사용자에게 원하는 시점의 영상 정보를 제공하는 3차원 비디오이며, 현재 방송 관련 연구 기관에서 차세대 실감방송 멀티미디어 서비스 개발을 목적으로 하는 연구가 활발히 진행되고 있다. 최근 MPEG 표준화 그룹에서는 다시점 비디오 부호화 (multi-view video coding, MVC) 방법에 관한 표준화 작업이 진행 중이며, 최신 비디오 압축 표준인 H.264를 이용한 여러 가지 방법들이 제안되었다. 현재 MVC 표준화 작업의 평가 기준이 되는 방법은 각 시점을 H.264로 부호화하는 방법인데, 이는 다시점 비디오 영상의 중요한 특성인 인접 시점들 사이의 공간적 상관도를 전혀 고려하지 않았다. 본 논문에서는 시점 보간법을 이용하여 얻어진 중간 영상과 원영상과의 차분 영상을 부호화하는 알고리즘을 제안하고자 한다. 여기서 시점 보간법이란 좌우 두 시점 영상으로부터 변이값을 얻은 다음, 이를 이용하여 중간 시점 영상을 합성하는 방법을 말한다. 예를 들면, 다시점 비디오의 홀수 번째 시점의 영상은 기존의 방법을 따르고, 짝수 번째 시점의 영상은 이미 부호화된 홀수 번째 시점의 영상을 이용하여 보간적으로 예측하여 원래 영상과 차분 영상을 구하여 부호화한다. 차분 영상은 영상의 복잡도가 많이 감소되어 원영상에 비해 보다 나은 부호화 효율을 보인다. 그러나 합성 영상이 각 장면마다 독립적으로 생성되므로 원영상에 비해 차분 영상의 시간적인 상관도가 줄어들어 I장면의 경우 부호화 효율이 크게 향상되었으나, 시간적인 상관도를 이용하는 P장면과 B장면에서는 오히려 좋지 않은 결과를 보였다.

1. 서론

디지털 기술의 발전과 방송 통신의 융합으로 방송 매체가 다양해짐에 따라, 디지털 기술의 특성을 이용한 방송 관련 부가 서비스들이 새롭게 선보이고 있다. 지금까지의 방송은 TV 화면에 평면적인 화상 및 음성 등의 한정된 정보만을 전송하였다. 하지만 TV의 디지털화 및 방송 네트워크의 지능화와 고도화가 진행되면서 이동 방송, 데이터 방송, 휴대 인터넷 방송, IP TV 및 T-Commerce 등의 다양한 방송 멀티미디어 서비스가 제공되기 시작했다. 현재 방송 중인 TV의 발전 방향은 고화질과 화면 크기의 확대로 가고 있지만 TV 화면 자체가 2차원이기 때문에 TV를 통해서 입체감을 느낄 수는 없다. 다시 말해, 실제 원거리에 떨어져 있는 물체를 TV를 통해 현실감 있게 접하고자 하는 시청자들의 요구 사항을 제대로 만족시키지 못하고 있는 실정이다.

현재 방송 관련 기술자들은 방송 네트워크를 통해 인간의 오감을 이용할 수 있는 실감방송에 관련된 연구를 활발히 진행하고 있다. 그 중에서도 입체적인 화상 표현이 가능하고 시각적으로 실감을 느낄 수 있는 다시점 비디오에 대한 관심이 매우 높다. 시청자에게 보다 입체감

있는 화면을 제공하고 보다 현실감 있는 느낌을 주기 위한 요소 중의 하나가 바로 시각 정보이며, 다시점 비디오는 시청자에게 원하는 시점의 영상을 제공하여 깊이감과 입체감을 제공해 주기 때문이다.

다시점 비디오는 똑같은 3차원 장면을 같은 시간에 여러 시점에서 다수의 카메라로 촬영한 것으로, 입체 TV, 자유 시점 TV, 감시 카메라 영상 등 다양한 영상에 응용될 수 있다 [1]. 다시점 비디오는 시청자에게 원하는 시점의 영상을 제공하는 것을 목적으로 한다. 다시점 비디오는 시청자가 원하는 시점에서의 영상을 제공할 수 있다는 장점을 지니는 반면, 카메라의 수에 비례하여 데이터양이 증가하기 때문에 이를 효과적으로 저장하고 전송하기 위해서는 다시점 비디오 정보의 압축 부호화가 필수적이다.

차세대 실감 방송 멀티미디어인 다시점 비디오 정보의 압축 부호화에 대한 필요성이 요구되면서, 국제 표준화 기구인 ISO/IEC JTC1/SC29/WG11 MPEG (Moving Picture Experts Group)은 다시점 비디오 부호화 방법에 대한 표준화 작업을 시작하였다. 최근에 이 부분의 연구가 활발히 진행되면서 다시점 비디오 테스트 시퀀스가 제공되었고, 이에 대한 다양한 압축 방법들이 제안되었다.

지금까지 제안된 다시점 비디오 부호화 방법들은 시공간 상의 상관도를 이용하는 예측 부호화 방법이 대부분이다 [2]. 이러한 방법들을 기존의 2차원 비디오 부호화 방식에서 확장하여, 시간 축 뿐만 아니라 동일 시간대에 있는 각 카메라의 공간적인 상관관계를 고려하여 부호화한다는 공통점을 지닌다.

본 논문에서는 다시점 비디오 부호화 방법들 중 하나로 시점 보간법을 이용한 차분 영상 부호화 방법을 제안한다. 이 알고리즘은 먼저 좌우 스테레오 영상을 획득한 후 이를 이용하여 변이 정보를 추출하고 여기서 얻어진 변이 정보를 중간 영상 합성에 사용하는 시점 보간 단계[3]와 원영상과 중간 영상과의 차분 영상을 부호화하는 차분 영상 부호화 단계로 나누어진다.

2. 시점 보간법

시점 보간법은 크게 변이 측정과 시점 합성으로 나뉜다. 첫 번째로 변이 측정은 입체 영상 쌍 (stereoscopic image pair)에서 대응되는 화소들을 찾는 과정인데, 이 과정은 잡음 영상, 폐색 영역, 반복되는 텍스처, 그리고 깊이 (depth) 불일치 때문에 변이맵을 측정하는 것이 문제가 되어 왔다. 이와 같은 문제점을 해결하기 위해 영역을 기반으로 주변 변이값을 찾는 방법을 사용한다.

두 번째로 시점 합성은 변이 측정 과정에서 얻어낸 변이맵을 이용하여 기준 영상 쌍에 대한 중간 시점의 영상을 복원하는 과정이다. 기준 영상 쌍에 대한 복원된 좌표가 정수가 아니기 때문에 선형 보간 과정을 거쳐서 중간 시점의 화소를 복원한다.

가. 변이 측정

변이는 두 카메라로부터 획득된 영상 화소의 수평 좌표, 즉, x 좌표의 차이로 정의할 수 있다. 이때 좌영상과 우영상 사이의 변이 d 는 두 영상 사이의 새로운 시점의 위치 α 에 의해 나누어지는 두 개의 변이로 나타낼 수 있다. 그림 1에 변이를 측정하는 방법을 나타내었다.

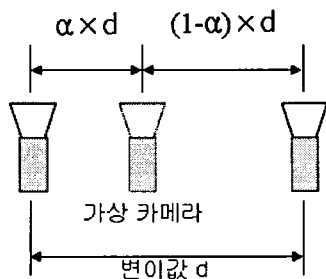


그림 1. 변이 측정 방법

그림 1에서 α 는 $0 \leq \alpha \leq 1$ 의 범위를 갖도록 α 에 의해 정규화된다. 새로운 시점이 좌측 시점에 가까워질수록 α 는 0에 가까워지고, 우측 시점에 가까워질수록 α 는 1에 가까워진다.

(1) 블록 단위 유사도 측정

유사도는 면적이 $A=(2M+1)(2N+1)$ 인 블록 단위로 측정하고 측정을 위한 척도로 MAD (Mean Absolute Difference)를 사용한다. 그림 2는 블록 단위 유사도 측정 방법을 나타낸다.

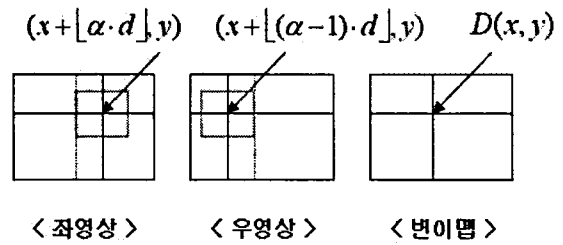


그림 2. 블록 단위 유사도 측정 방법

유사도에 대한 식은 다음과 같이 나타낼 수 있다. 블록 내에서 영상을 벗어나는 부분이 생길 경우, 영상 내부에 포함된 면적 내에서만 MAD를 구한다.

$$E_{block}(x, y, d, \alpha) = \frac{1}{A} \sum_{j=-M}^M \sum_{i=-N}^N \Psi(x+i, y+j, d, \alpha) \quad (1)$$

$$\Psi(x, y, d, \alpha) = |I_L(x + \lfloor \alpha \cdot d \rfloor, y) - I_R(x + \lfloor (\alpha - 1) \cdot d \rfloor, y)| \quad (2)$$

(2) 변이맵의 공간적 상관도 이용

폐색 영역, 잡음, 그리고 반복되는 텍스처 때문에 인접 화소에 대해서만 변이를 찾는 것은 좋은 방법이라 할 수 없다. 따라서 변이맵을 더욱 정확하게 구하기 위해서는 주변 변이값들과의 유사성을 이용한다. 다음 에너지 함수는 변이맵 D 내부에서 이웃 간 변이값들과의 유사도를 측정한다.

$$E_{reg}(x, y, d) = \frac{1}{4} [|D(x-1, y) - d| + |D(x-1, y-1) - d| + |D(x, y-1) - d| + |D(x+1, y-1) - d|] \quad (3)$$

(3) 변이맵의 시간적 상관도 이용

영상의 대부분 영역에서는 시간적인 변화가 크지 않으므로, 변이값을 측정할 때 이전 장면에서 동일한 위치에서의 변이값을 이용하면 변이값의 정확도를 높일 수 있다. 시간적 상관도를 이용한 에너지 함수는 다음과 같다.

$$E_{temp}(x, y, d, t) = |D(x, y, t) - d| \quad (4)$$

최종적으로 위의 모든 에너지 함수들의 합을 최소화하는 d 를 $D(x, y)$ 로 결정한다.

나. 시점 합성

α 위치에서의 새로운 시점 합성은 기준 영상 쌍의 선형 보간 과정과 대응하는 화소의 혼합 (blending) 과정으로 이루어진다. 혼합 과정에서 기준 영상 화소의 복원된 좌표가 정수값을 가지지 않기 때문에 선형 보간을 통하여 복원된 좌표들을 정수값으로 만들고, 선형 보간이 이루어진 좌표와 α 를 이용하여 기준이 되는 좌우 영상에서의 대응하는 화소값을 이용하여 혼합한다. 복원된 좌표가 영상의 범위를 벗어나는 경우에는 해당 영상을 무시하고 나머지 영상에 대해서만 복원한다.

표 2. 'Rena_p00046' 영상에 대한 부호화 효율 비교

QP	Anchor		제안하는 알고리즘	
	비트율 (kbps)	PSNR (dB)	비트율 (kbps)	PSNR (dB)
28	712.50	42.59	399.16	42.10
32	411.58	40.01	193.61	39.68
36	235.56	37.54	96.80	37.35
40	141.41	35.19	54.35	35.14

표 2에서 알 수 있듯이, 제안한 알고리즘에서 PSNR은 0.1~0.5 dB 떨어진 반면, 비트율은 절반 이상으로 줄어들었다. 그림 5와 그림 6은 'Anchor'와 제안하는 알고리즘에 대한 비트율-왜곡 곡선을 비교하여 나타내고 있다.

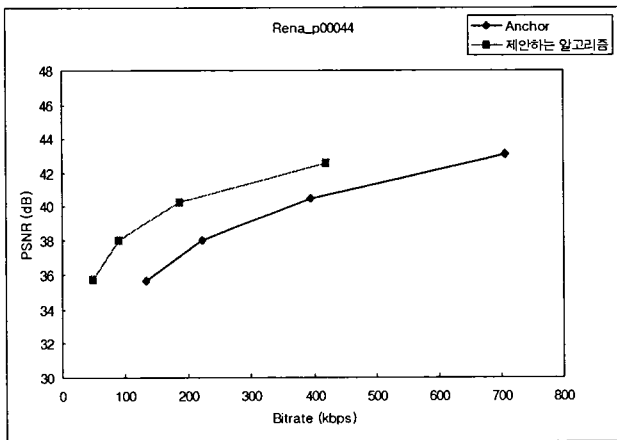


그림 5. Rena_p00044 영상에 대한 비트율-왜곡 곡선

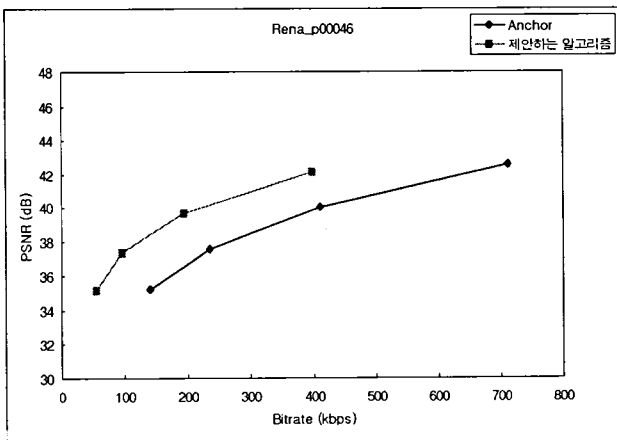


그림 6. Rena_p00046 영상에 대한 비트율-왜곡 곡선

그림 5와 그림 6에서 알 수 있듯이, 제안하는 알고리즘의 비트율이 'Anchor'에 비해 크게 줄어든 것을 알 수 있다. 하지만 QP가 증가함에 따라 비트율이 감소하는 정도가 다소 줄어드는 경향을 보이기도 했다. 다시 말해, 화질이 떨어질수록 비트율의 이득 또한 떨어진다는 사실을 알 수 있었다.

5. 결론

본 논문에서는 시점 보간법을 이용한 다시점 비디오 부호화 방법을 제안하였다. 다시점 비디오 부호화를 위해 사용된 시점 보간법에서

는 화소 단위로 변이맵을 생성하는 경우 발생하는 문제점인 잡음 영상, 폐색 영역, 반복되는 텍스처, 그리고 깊이 불일치를 해결하기 위해 영역을 기반으로 주변 변이값을 찾는 방법을 이용하였다. 본 논문에서 제안한 다시점 비디오 부호화 방법은 홀수 번째 시점 영상들은 'Anchor'와 동일하게 부호화하고, 짝수 번째 시점 영상들은 홀수 번째 시점 영상들을 이용하여 시점 보간한 중간 영상과 짝수 번째 시점의 원영상과의 차분 영상을 부호화하는 방법을 사용하였다. 제안한 알고리즘이 높은 부호화 효율을 내기 위해서는 시점 보간법을 이용하여 얻은 영상과 원영상과의 차이가 적어야 하고 이를 위해서는 카메라 간격이 조밀해야 한다. 컴퓨터 모의실험을 통해서 제안하는 다시점 비디오 부호화 방법이 'Anchor' 부호화 방법에 비해 PSNR은 0.1~0.5 dB 떨어진 반면, 비트율이 크게 줄어든 것을 확인하였다.

감사의 글

본 연구는 광주과학기술원(GIST) 실감방송연구센터(RBRC)를 통한 정보통신부 및 정보통신연구진흥원의 대학IT연구센터(ITRC) 지원 사업, 교육인적자원부 두뇌한국21 (BK21) 정보기술사업, 그리고 전자부품연구원(KETI)의 지원에 의한 것입니다.

참고 문헌

- [1] ISO/IEC JTC1/SC29/WG11 M11259, "FTV (Free Viewpoint Television): Achievements and Challenge," October 2004.
- [2] ISO/IEC JTC1/SC29/WG11 N6909, "Survey of Algorithms used for Multi-view Video Coding (MVC)," January 2005.
- [3] M. Droege, T. Fujii and M. Tanimoto, "Ray-Space Interpolation based on Filtering in Disparity Domain," Proc. 3D Conference 2004, pp. 213-216, Tokyo, June 2004.
- [4] Tamas Frajka and Kenneth Zeger, "Residual Image Coding for Stereo Image Compression," International Conference on Image Processing (ICIP) Rochester, New York, vol. 2, pp. 217-220, October 2002.