

MPEG-2 MVP를 이용한 스테레오 동영상부호화

°배 태 면, *권 동 현, 한 규 필, 하 영 호
 경북대학교 전자전기공학부
 영남대학교 전자공학과
 702-701 대구광역시 북구 산격동 1370번지
 e-mail: yhha@ee.kyungpook.ac.kr

Stereoscopic Sequence Coding Using MPEG-2 MVP

°Tae-Min Bae, *Dong-hyun Kwon, Kyu-Phil Han, and Yeong-Ho Ha
 School of Electronic Electrical Eng., Kyungpook Nat'l Univ.
 1370 Sankyuk-dong, Puk-gu, Taegu 702-701, Korea.
 * Department of Electronic Eng., Yeungnam Univ.

Abstract

A new stereoscopic codec. structure using MPEG-2 multiview profile is presented in this paper. In the suggested codec., the left image is coded with motion estimation in the base layer and the right image is coded with disparity estimation in the enhancement layer. Since it is possible to calculate rough motion of the right image sequence with disparity and motion of the left image sequence, motion compensation of the enhancement layer is performed without motion estimation. Since the proposed codec. does not perform motion estimation in the enhancement layer encoding, it is simple and reduces the encoding time. We compared the PSNR of encoded image with three different structured codec., and the experimental results show that suggested codec. has comparable with other codecs.

I. 서론

3차원 가상현실에 대한 관심이 증가함에 3차원의 입체감 있는 영상을 볼 수 있는 방법들이 연구되고 있다. 그 가운데 스테레오 영상을 이용한 입체 영상 시스템이 3D TV, 입체영화, 게임등에 적용되고 있다. 그러나 이러한 스테레오 영상은 일반 영상에 비해 정보량이 2배로 증가하므로 영상을 효율적으로 저장하기 위한 입체 영상 부호화기가 필요하다. 스테레오 동영상은 영상열간의 시간적 중복성과 두 스테레오 영상사이의 공간적 중복성을 가지고 있기 때문에 정확한 움직임과 변이를 추정한다면 전송효율을 높일 수 있다.

지금까지 알려진 스테레오 영상 부호화의 방법으로는 객체기반(object-based) 방법[1]과 기존의 MPEG과 같은 블럭기반의 움직임 예측방법을 확장한 방법[2-8]이 연구되어왔다. 객체기반 방법은 두 스테레오 영상에서 동일한 3차원 움직임을 가지는 객체를 추출해내고 부호화기

에서는 그 객체에 관한 움직임 정보를 이용하여 영상을 부호화 하는데, 스테레오 영상열(stereo image sequence)로부터 좌·우 영상 각각의 움직임 정보와 좌·우 영상 사이의 변이(disparity) 정보를 이용하므로 하나의 영상열(mono image sequence)을 이용하는 경우보다 정확한 3차원 움직임을 예측할 수 있다. 그러나 객체 기반의 방법은 높은 부호화율을 얻을 수 있지만 기존의 블럭기반의 부호화기보다 구조가 복잡하고 계산량이 많아 실제 부호화기의 구현이 어렵다.

MPEG-2 표준안에는 스테레오 동영상 부호화를 위한 MVP(Multiview Profile)가 존재하는데, MVP에서는 시간계위(Temporal Scalability)를 이용하여 기저층(base layer)에서는 좌영상에 대한 부호화를 행하고, 상위층(enhancement layer)에서는 우영상에 대한 부호화를 행하는 구조로 되어 있다. 상위층의 부호화시 변이 정보만을 이용하여 스테레오 영상을 압축하는 경우에 부호화기는 간단해 지지만, 변이 정보뿐만 아니라 움직임 정보를 함께 이용하는 경우보다 화질의 저하를 가져온다. 그렇지만 움직임 예측과 변이 예측을 동시에 사용하기 위해서는 부호화기가 복잡해지는 것을 피할 수 없게 된다. 본 논문에서는 좌영상은 움직임 예측에 의해, 우영상은 변이 정보를 예측하여 부호화한다. 이 때 우영상열의 시간적인 중복성 제거를 위해 좌영상의 움직임과 변이 정보로부터 우영상의 움직임을 유도해 낼 수 있다[7]. 이 유도된 움직임 정보를 이용하여 상위층에서의 움직임 예측을 수행하지 않고 움직임 보상이 가능하지만 MPEG-2 표준안을 지키기 위해서는 부호화기를 설계시 움직임 예측모드, field/frame 모드를 고려해야 한다. 본 논문에서는 위의 문제를 해결하여 MPEG-2 MVP 표준안을 지키며 상위계층에서 움직임 예측을 수행하지 않는 구조를 적용하여 부호화에 요구되는 계산량을 줄임과 동시에 구조를 간단히 할 수 있었다.

II. MVP를 이용한 부호화기

MPEG-2의 MVP에서는 시간계위를 이용하여 기저층에서는 움직임 보상 및 DCT기반 알고리즘을 사용하여 부호화하며, 상위층에서는 상위층간의 움직임 보상정보

와 기저층과 상위층간 시각차 정보를 함께 이용해 계층간의 정보를 이용한 부호화를 행하여 전송하려는 정보를 최소화하려고 하고 있다[8].

일반적으로 스테레오 동영상 부호화기는 좌우영상을 부호화 하는 방법에 따라 3가지로 나눌 수 있다[3]. 좌우영상을 독립적으로 부호화하는 simulcast 부호화방법과 좌우영상중 한 쪽영상을 먼저 부호화하고 다른 쪽 영상은 먼저 부호화한 영상을 이용하여 부호화하는 compatible 부호화방법, 그리고 좌우영상을 동시에 부호화 하는 joint부호화방법이 있다. 위 세 가지 방법 가운데 MPEG-2기반의 부호화기의 구조에서는 simulcast 부호화방법과 compatible 부호화방법만이 가능하다. 그러나 simulcast 부호화방법의 경우 좌우영상을 독립적으로 부호화하기 때문에 좌우영상간의 중복성을 제거할 수 없으므로 compatible 부호화방법에 비해 부호화율이 떨어지지만 구현시 병렬화가 가능하고 설계가 간단하다. MPEG-2의 MVP를 기반으로 부호화기를 설계하는 경우, simulcast 부호화방법과 compatible 부호화방법중 어느 방법을 사용하더라도 기저층의 부호화방법은 변하지 않는다. 그러나 상위층의 부호화의 경우 simulcast방법에 의해 부호화하면 움직임정보만을 이용하여 부호화하게 되고 compatible 부호화방법으로 부호화하면 변이정보를 추가하여 부호화할 수 있다. 그림 1은 simulcast방법을 이용한 스테레오 동영상 부호화기의 구조이고 그림 2는 compatible방법을 이용한 스테레오 동영상 부호화기의 구조이다. 그림 1과 2에서와 같이 simulcast부호화방법이 compatible부호화방법보다 부호화기 구조가 복잡하다.

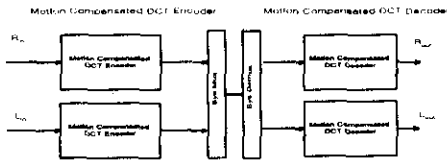


그림 1. simulcast부호화 방법을 이용한 스테레오 동영상 부호화기 구조

Fig. 1. stereoscopic codec using simulcast method

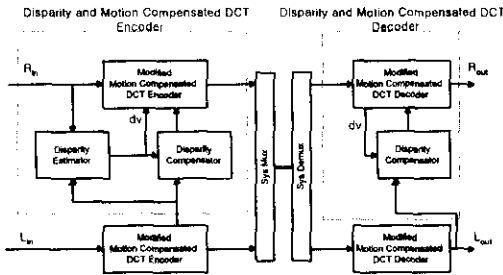


그림 2. compatible부호화 방법을 이용한 스테레오 동영상 부호화기 구조

Fig. 2. stereoscopic codec using compatible method

III. 기저층의 움직임정보와 변이정보를 이용한 상위층의 움직임 정보 추출

평행하게 정렬된 스테레오 카메라에서 실제 물체가 평행이동하는 경우 좌영상에서의 움직임 정보와 우영상의 움직임 정보와의 관계를 유도해 낼수 있다[7].

$I_{t+1} - I_t = v_i$, $r_{t+1} - r_t = v_r$, $D_t = r_t - l_t$ 에서 영상의 움직임 벡터를 좌영상의 움직임 벡터와 변이값으로 정리하면

$$v_i = D_t - D_{t+1} + v_r \quad (1)$$

과 같다. 즉 좌영상의 움직임 정보 v_i 를 우영상의 움직임 정보 v_r 와 두 프레임의 변이 정보 D_t 와 D_{t+1} 에서 구할 수 있음을 알 수 있다. 그러나 위의 결과는 화소단위에서는 그대로 적용할 수 있지만 블럭단위 정합에서는 그대로 적용할 수가 없다. 블럭단위 정합을 이용한 움직임 정보의 예측방법은 이차원 평행한 움직임만이 정확하게 예측될 수 있다. 그러므로 실제 물체에서 이차원 평행한 움직임(2D translational motion)만이 있다고 가정하면

$$Z_t = Z_{t+1} \quad (2)$$

이 성립하고 식 (1)은 다음과 같이 간단하게 표현될 수가 있다.

$$v_i = v_r \quad (3)$$

즉 좌영상의 움직임정보를 정합되는 우영상에 그대로 적용할 수가 있다. 이를 바탕으로 우영상의 k 번째 블럭의 움직임 정보를 계산하는 식은 다음과 같다.

$$V_r(k) = V_r(k + \text{SIGN}(D(k)) * \text{ROUND}(\frac{-D(k) + B/2}{B})) \quad (4)$$

이때 ROUND 함수는 부동 소수점을 정수 값으로 변환하는 함수이고 SIGN 함수는 입력값의 +/- 부호를 반환하는 함수이다. 정합 블럭의 크기가 $B \times B$ 일 때 우영상의 k 번째 블럭의 움직임 벡터 $V_r(k)$ 는 k 번째블럭의 변이 정보 $D(k)$ 를 이용하여 좌영상의 $(k + \text{SIGN}(D(k)) \times \text{ROUND}(\frac{-D(k) + B/2}{B}))$ 번째 블럭의 움직임 정보를 이용하여 된다.

이렇게 유도한 움직임 정보를 MPEG-2 MVP를 이용한 부호화기에서 사용하기 위해서는 움직임 예측모드를 고려해야한다. MPEG-2의 경우 움직임 예측방법으로 전방향예측(forward prediction), 후방향예측(backward prediction), 양방향예측(bidirectional prediction)이 있는데 기저층의 움직임 예측정보는 세 가지 모두가 변이정보를 이용하여 상위층의 움직임 정보를 유도할 수 있지만 상위 계층의 경우 변이 정보와 움직임 정보를 모두 사용하기 위해 후방향 예측에 변이 정보를 이용하고 전방향 예측에 움직임 정보를 이용하므로 기저층의 세 가지 움직임 정보를 모두 사용하면서 표준안을 지키기는 어렵다. 또한 MPEG-2에서는 양방향 예측을 위해 영상의 시간적인 순서와는 다르게 재배열하여 I, P, B, B순서로 영상을 부호화하므로 상위계층에서도 기저층과 같이 영상의 재배열을 하지 않고는 하위 계층의 움직임 정보를 이용할 수가 없

다. 이를 해결하기 위해서는 상위계층의 영상을 하위 계층과 같이 재배열하고, 상위계층에서 하위 계층의 움직임 정보중 전방향예측에 의한 움직임정보가 존재하는 매크로블럭에 대해서만 움직임 정보를 유도하는 방법과 하위 계층의 부호화시 양방향예측을 배제하는 방법이 있다. 그림 3은 기저층에서 양방향예측을 배제한 영상배열을 나타낸다. 이러한 배열에서는 상위층의 영상열을 재배열하지 않고도 기저층의 움직임 정보를 사용할 수 있다.

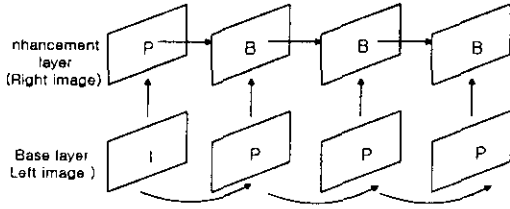


그림 3. 스테레오 영상열간의 관계
Fig. 3. Relation between stereo sequences.

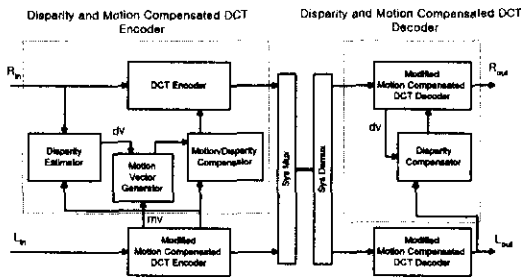


그림 4. 제안한 부호화 방법을 이용한 스테레오 동영상 부호화기 구조
Fig. 4. stereoscopic codec using proposed method

또한 MPEG-2 MVP에 적용하기 위해서는 매크로 블럭의 필드/프레임(field /frame)모드를 고려해야 하는데 기저층에서 움직임정보 예측시 매크로 블럭별로 필드모드 또는 프레임모드 가운데 더 나은 결과를 가지는 모드를 선택하게 되는데, 상위층에서의 변이정보 예측시 블럭의 모드와 기저층의 모드가 다를 수 있다. 상위층의 변이 예측시의 매크로블럭의 모드와 기저층의 움직임 예측시의 모드가 달라도 움직임 정보를 유도하는데는 문제가 없지만 기저층의 블럭모드에 따라 상위층에서는 움직임예측 모드의 선택을 제한해야 한다. 즉 상위층에서는 변이 예측 정보와 움직임 정보를 이용한 양방향 예측이 가능하지만 움직임 정보가 필드 또는 프레임모드의 움직임 정보로 한정되므로 양방향 예측모드 또한 기저층의 블럭모드에 의해 한정시켜야한다. 실제 MPEG-2 MVP를 이용한 부호화 실험시 상위층에서 변이 정보와 움직임 정보를 모두 사용하는 양방향 예측모드가 가장 높은 효율을 보인다[3]. 마지막으로 MPEG-2에서는 반화소(half pel)단위의 움직임 예측을 하지만 기저층의 움직임 정보를 반화소 단위까지 사용하기에는 유도한 움직임 정보의 정확도가 떨어지므로 유도된 움

직임정보는 화소단위로 결정한 후 반화소단위의 예측을 수행하여 움직임 정보를 결정하였다. 그림 4는 MPEG-2 MVP에 제안된 알고리즘을 적용한 부호화기의 구조이다. 상위 계층의 부호화 블럭에는 움직임 예측 블럭을 대신하여 변이 정보와 기저층의 움직임 정보에서 상위층의 움직임 정보를 유도해내는 motion vector generator가 존재한다.

IV. 실험 및 결과

ITU-R BT.-601 720x480 크기인 4:2:0 컬러 포맷의 'flowerpot', 'street organ', 'trapeze', 'starsailer' 스테레오 영상들이 실험에 사용되었다. 실험에 사용된 부호화기는 기저층은 영상의 재배열을 피하기 위해 I, P, P 구조가 되도록 하였고, 상위층은 변이 정보만을 이용하는 경우, 움직임 정보와 변이 정보를 모두 이용하는 경우, 그리고 기저층의 움직임 정보와 변이정보를 이용하여 상위층의 움직임 정보를 유도하는 제안한 방법을 적용한 3가지 부호화기를 사용하여 부호화 결과 영상의 PSNR(Peak Signal to Noise Ratio)을 비교하였다. 이때 GOP(Group of Picture)의 길이 M은 15로 I, P영상간의 거리 N은 3으로 두었으며 기저층과 상위층의 비트율은 각각 4Mbps로 설정하였다. 그림 5, 6, 7은 각 30프레임에 대한 부호화된 우영상의 Y신호에 대한 PSNR을 나타낸 것이고, 표 1은 부호화된 우영상열들의 Y 신호의 평균PSNR이다. 이 때 상위층을 부호화시 Case1은 변이 예측만을 사용하여 경우, Case2는 변이 예측과 움직임 예측을 모두 사용하는 경우, 그리고 Case3는 제안한 방법에 의한 경우이다. 'street organ'영상을 제외한 나머지 세 가지 영상에 대한 부호화 결과에서 제안한 방법에 의한 부호화방법이 직접 움직임을 예측하여 부호화한 결과에 근접하면서 변이만을 예측하여 부호화한 경우보다 나은 결과를 보였다. 'flowerpot', 'trapeze', 'starsailer'영상들은 움직이지 않는 넓은 배경과 움직임을 가지는 작은 영역을 가지는 공통점을 가지고 있는데 움직임이 없는 배경의 경우 기저층에서도 움직임이 없으므로 항상 변이가 존재하는 변이에 의한 영상예측보다 나은 예측을 할 수 있으며 움직이는 작은 영상에 대해서도 어느 정도 타당한 움직임 예측을 하는 것을 추측할 수 있다. 이에 비해 걸어가는 사람이 화면의 대부분을 차지하는 'street organ'영상의 경우 변이정보를 이용하는 부호화기보다는 좀더 나은 결과를 보이지만 움직임 예측을 함께 수행하는 부호화기보다는 낮은 예측능력을 보이는 것을 알 수 있다.

표 1. 부호화된 우영상 Y신호의 PSNR
Table 1. Luminance PSNR(dB) of right image

sequence	Case 1	Case 2	Case 3
street organ	33.8	34.7	34.1
starsailer	31.6	33.5	33.5
Trapeze	31.1	33.6	32.7
flowerpot	34.9	37.5	37.4

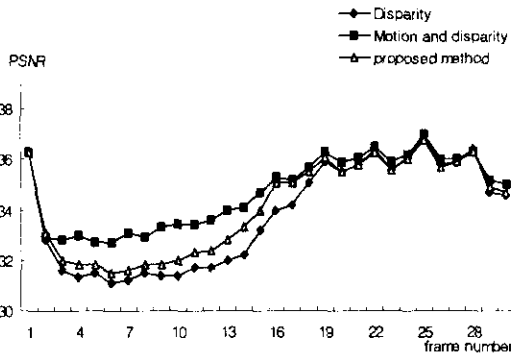


그림 5. 부호화된 'street organ' 우영상의 PSNR
Fig. 5. PSNR of encoded 'street organ' right sequence.

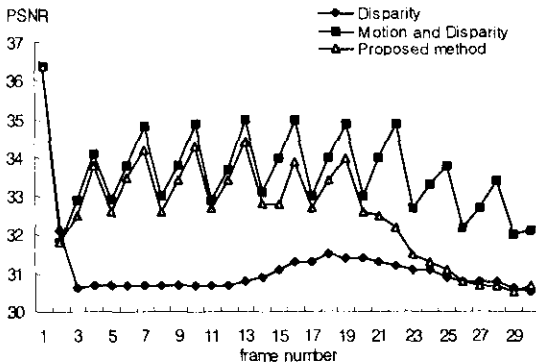


그림 6. 부호화된 'Trapeze' 우영상의 PSNR
Fig. 6. PSNR of encoded 'Trapeze' right sequence.

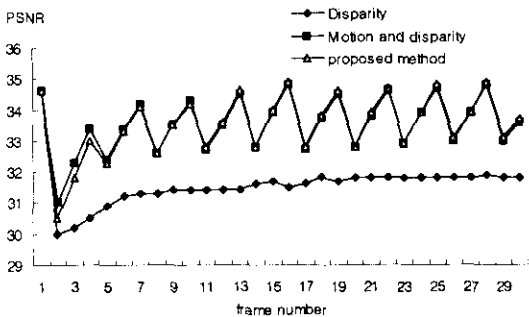


그림 7 부호화된. 'flowerpot' 우영상의 PSNR
Fig. 7. PSNR of coded 'flowerpot' right sequence.

V. 결론

본 논문에서는 스테레오 동영상 부호화에 MPEG-2 MVP를 이용한 구조를 제안하였다. 제안한 방법은 스테

레오 영상 부호화시 변이 정보를 이용하여 좌·우 영상 사이의 공간적 상관성을 이용하였으며 변이 정보만을 예측하면서도 움직임 정보를 스테레오 카메라의 기하학적 구조에서 계산함으로써 시간적인 상관성도 고려하여 변이만을 사용하는 경우보다 높은 PSNR을 얻을 수 있었다. 상위계층의 부호화시 움직임 예측을 생략하여 구조가 간단해진 부호화기를 구현하였으며 빠른 부호화를 실현하였다. 실영상에 대한 제안한 방법의 실험 결과는 변이 예측만을 한 부호화기보다 객관적 성능의 향상을 보여주었으며 움직임과 변이를 모두 예측하여 부호화한 경우에 근접하는 결과를 보여주었다. 제안된 부호화기는 기저층의 움직임 정보를 이용하기 위해 기저층에서 전방향 예측만을 수행하여 양방향 예측을 이용하여 단일 영상을 부호화하는 MPEG기반의 부호화기가 움직임 예측을 수행하는 시간에 기저층의 움직임 예측과 상위층의 변이 예측이 가능하지만 결과적으로 기저층의 부호화율이 떨어지는 단점이 있다.

참고문헌

- [1] Dimitros Tzovaras, Nikos Grammalidis, and Michel G. Strintzis, "Object-Based Coding of Stereo Image Sequences Using Joint 3-D Motion/Disparity Compensation" *IEEE Trans. on C.S.V.T* vol. 7, pp. 312-327, 1997.
- [2] Jin Liu and Robert Skerjanc, "Stereo and motion correspondence in a sequence of stereo images," *Signal Processing Image Communication*, vol. 5, pp. 305-318, 1993.
- [3] A. Puri, R.V. Kollarits, and B.G. Haskell, "Basics of stereoscopic video, new compression results with MPEG-2 and a proposal for MPEG-4," *Signal Processing Image Communication*, vol. 10, pp. 201-233, 1997.
- [4] R. Horst, "A digital codec for 3D-TV transmission," *Signal Processing of HDTV, IV*, pp. 489-495, 1993.
- [5] Y. J. Song, "Improved disparity estimation algorithm with MPEG-2's scalability for stereoscopic sequence," *IEEE Trans. on Consumer Electronics*, vol. 42, no. 3, pp. 306-311, 1996.
- [6] 권동현, 장현식, 김진웅, 이명호, 김영식 "MPEG-2 MVP를 이용한 입체 영상 부호화기" *대한전자공학회 추계종합학술대회 논문집*, 제19권, B편, 제2호, pp. 1191-1194, 1996.
- [7] 배태면, 한규필, 김용석, 이호영, 하영호 "다해상도 블러 정합기법을 이용한 스테레오 동영상의 부호화" *한국통신공학회 하계종합학술대회 논문집*, vol. 16, pp. 420-423, 1997.
- [8] ISO/IEC 13818-2, "Information technology - Generic Coding of Moving Pictures and Associated Audio Information : Video," April, 1996.