

# 3DTV를 위한 다시점 동영상 부호화 기법

\*배진우, \*\*송혁, \*\*\*김정수, \*유지상

\*광운대학교 전자공학과, \*\*KETI, \*\*\*한국사이버대학교 컴퓨터정보통신학부

\*bjw8751@image.gwu.ac.kr

## Multi-view Video Codec for 3DTV

\*Jin-Woo Bae \*\*Hyuk Song \*\*\*Jeong-Su Kim \*Ji-Sang Yoo

\*KwangWoon University, \*\*KETI, \*\*\*Korea Cyber University

### 요약

본 논문에서는 3DTV를 위한 다시점 동영상 부호화 기법을 제안한다. 제안하는 다시점 동영상 부호화기는 다수의 카메라 입력에 대하여 기존의 MPEG 기반의 부호화 기법에서 사용된 시·공간적 중복성 제거와 각 시점에 해당하는 영상간의 공간적 중복성을 제거하여 부호화효율을 개선할 수 있다. 각 시점 영상간의 공간적 중복성을 효율적으로 제거하기 위하여 전역 시차 보상(global disparity compensation)된 집적영상(Accumulated Image)를 사용하였다. 또한 기존의 디지털 TV 전송 표준인 MPEG-2를 기반으로 하였기 때문에 기존의 디지털 TV 표준을 크게 수정하지 않고 3DTV를 구현 할 수 있다. 제안하는 기법은 기존의 MPEG-2 기법으로 각 시점에 해당하는 영상을 각각 부호화한(simulcast 기법) 결과와 비교하여 객관적 화질면에서 우수하였으며, 각 시점간의 동기화 문제도 해결 할 수 있었다.

### 1. 서론

고화질의 HDTV의 상용화 이후에 차세대 영상 통신으로 입체감을 제공할 수 있는 3DTV에 대한 관심이 높아지고 있다. 3DTV는 깊이감과 입장감 더 나아가서는 관찰자의 시점에 따라 각각 다른 영상을 보여줄 수 있는 입체, 실감 방송을 의미한다[1,2]. 일반적으로 입체감을 느낄 수 있는 스테레오 영상을 얻기 위해서는 양안에 해당하는 두 대의 카메라가 필요며, 이를 이용하여 획득된 영상을 전송하고 수신단에서 특수한 장비를 이용하여 관찰자는 입체감이 있는 영상을 볼 수 있게 된다. 스테레오 영상은 양안 시차에 의해 입체감은 느낄 수 있지만 고정된 시점에서 영상을 보게 된다는 단점이 있다. 이러한 단점을 보완하기 위해서는 관찰자가 움직이거나 다른 시점에서 보았을 때 해당되는 영상을 제공할 수 있어야 하며, 이를 위하여 두 대 이상의 카메라를 이용하는 다시점 영상 시스템이 필요하다. 하지만 이러한 다시점 영상 시스템은 다수의 카메라를 이용하여 영상을 획득하기 때문에 전송해야 할 데이터 또한 증가하게 된다. 최근 제한된 대역폭 내에서 효율적으로 다시점 영상을 전송할 수 있도록 하기 위한 연구가 많이 진행되고 있다[3,4,5,6].

MPEG(Moving Picture Expert Group)에서는 3차원 비디오와 관련된 규격으로 1996년 MPEG-2 Video 표준에 포함된 MVP(Multi-View profile)를 시초로 3차원 비디오 압축에 대한 표준화에 대한 관심을 가지기 시작하였다. 하지만 MPEG-2 Video MVP는 주 대상이 스테레오 영상에 대한 압축 규격으로, 영상의 깊이 정보 및 영상의 calibration에 대한 정보를 포함하고 있지 않기 때문에 다시점 비디오 압축 기술을 만족시키기에는 크게 부족하다. 뿐만 아니라 시점 변경에 대한 기술이 없어 다시점 비디오를 이용한 상호작용(interactivity)과 같은 다양한 서비스를 제공하지 못하는 한계를 지니고 있다. 이러한 문제점을 해결하기 위하여 MPEG에서는 최근 부각되

는 있는 다시점 비디오 및 3차원 오디오/비디오 기술과 관련된 새로운 표준화를 58차 회의부터 3DAV(3 Dimensional Audio Video) [7,8,9]라는 명칭으로 시작하게 되었다. 3DAV 그룹에서는 EE (Exploration experiments)를 통하여 입체 영상과 관련된 응용분야와 표준화 안전에 대하여 회의한 결과 MVC(Multi-view Video Codec)에 대한 표준화가 진행중에 있다[7]. 유럽에서는 COST230(1992~1994) 프로젝트를 시작으로 PANORAMA(1995~1998)와 ATTEST(2002~2004)를 거쳐 3DTV(2004~2009) 프로젝트로 진행 중이며, 특히 3DTV 프로젝트는 기존 연구를 바탕으로 3차원 영상에 대한 획득, 압축/전송 및 표시에 대한 통합화와 표준화를 목표로 진행 되고 있다[7,8]. 국내에서도 연구기관과 각 대학기관을 중심으로 활발히 연구되고 있다.

다시점 영상 시스템에서 입력 영상은 다양한 각도와 거리에 설치된 두 대 이상의 카메라를 이용하여 획득된다. 이러한 카메라의 배치에 따라 관찰자의 시야각이나 입체감에 직접적인 영향을 미칠 뿐 아니라 영상 압축 효율에도 큰 영향을 미친다. 그렇기 때문에 카메라 시스템을 구성할 때 일반적으로 폭주식(convergence) 카메라 모델이나 평행식(parallel) 카메라 모델을 사용하며, 이때 카메라의 내/외부 파라미터(intrinsic/extrinsic parameter) 또한 함께 제공되기도 한다. 모든 카메라 모델에 대하여 독립적으로 우수한 압축 효율을 가질 수 있는 기법에 대한 연구는 아직 미비한 상태이다. 때문에 MPEG에서는 MVC의 표준화에 대한 입력 영상을 일차원 평행식 카메라 모델, 일차원 폭주식 카메라 모델 그리고 2차원 평행식 카메라 모델로 국한하고 있다[7,8]. 현재 진행 중인 표준화에서는 MVC에서 지원되어야 할 몇 가지 요구사항을 정의하고 있는데, 이중 가장 중점으로 생각하는 것은 압축 효율이며, 그 이외에 각 시점간 스케일러빌리티, 저지연성(low delay), 각 시점간 랜덤엑세스(view random access) 및 동기화에 중점을 두고 있다[9].

본 논문에서는 MPEG에서 정의한 요구사항을 최대한 포함할 수 있

으며, 디지털 3DTV에 적합한 다시점 동영상 부호화 기법을 제안한다. 각 시점 영상들을 이용하여 집적영상(Accumulated image)을 생성하고, 생성된 집적영상을 이용하여 각 시점에 해당하는 영상의 시차를 예측하게 된다. 집적영상에는 기준이 되는 영상이 포함되어 있기 때문에 집적영상만 전송할 경우 기존의 2DTV와도 하방 호환성을 가질 수 있다. 또한 집적영상과 예측영상만을 이용하여 시차를 예측하기 때문에 각 시점간 스케일러빌리티를 지원할 수 있으며, 각 시점영상을 각각 부호화하는 simulcast 기법에서 발생할 수 있는 시간 동기화 문제도 해결 가능하다.

본 논문의 II장에서는 제안하는 기법의 전체 구성에 대하여 설명하고 III장에서는 집적영상의 생성 방법과 이를 이용한 부호화 기법을 설명한다. IV장에서는 실험 및 그 결과에 대해 기술하고 V장에서 결론을 내린다.

## 2. 다시점 동영상 부호화를 위한 집적영상 생성

제안하는 다시점 동영상 부호화 기법은 카메라간의 공간적 중복성을 효율적으로 제거하여 높은 부호화 효율을 얻을 수 있으며, 기존의 TV와 하방 호환성 및 각 시점간 시간 동기화를 해결할 수 있다. 제안하는 동영상 부호화기의 전체 흐름도를 그림 1에 보였다. 입력 영상은 평행식과 폭주식 카메라 모델로 제한하였으며, 이를 이용하여 집적영상을 만들어내게 된다. 집적영상은 카메라 모델에 따라 원래 해상도보다 더 커질 수 있으며, 기준 시점 영상 이외에 다른 시점의 영상에 대한 정보를 포함하고 있다.

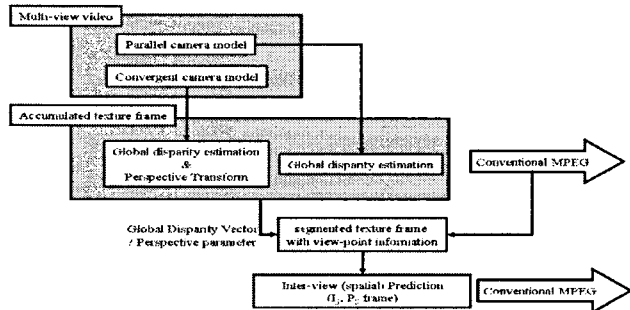


그림 1. 집적영상을 이용한 다시점 동영상 부호화기의 전체 흐름도

그림 1에서 보는바와 같이 다시점 동영상 부호화기는 입력되는 카메라의 구조에 따라 집적영상의 생성 기법을 달리하게 된다. 평행식 카메라 구조인 경우에는 뒤의 배경에 대한 전역 시차만을 보상하여 집적영상을 생성하게 된다. 폭주식 카메라 구조의 경우에는 객체에 초점을 맞추어 뒤의 배경이 달라지므로 전역 시차 보상 후 회전, 뒤틀림, 확대축소가 가능한 원근변환(perspective transform)을 이용하여 배경에 해당하는 공간적 중복성을 제거할 수 있게 된다. 위의 방법으로 생성된 집적영상은 일반적인 동영상 부호화 기법으로 부호화 집적영상에 존재하는 시,공간적 중복성을 제거할 수 있다. 또한 다시점 영상간의 중복성은 집적영상에서 전역 시차 벡터 및 원근변환의 매개변수를 이용하여 분리한 Segmented 영상을 이용하였다.

### 가. 집적영상을 이용한 각 시점별 참조영상 생성

제안하는 다시점 동영상 부호화 기법은 집적영상을 이용하여 각 시점별 참조영상을 생성한 다음 실제로 부호화 할 영상에 대하여 예측하

여 부호화 효율을 높일 수 있다. 평행식 구조를 가지는 다시점 영상에서 집적영상과 분리영상을 이용하여 참조영상을 생성한 결과를 그림 3에 나타내었다. 그림 3의 (d)와 (e)에서 보는 바와 같이 집적영상을 이용하여 생성된 참조영상을 이용하여 예측할 경우 오차성분을 줄일 수 있어 효율적인 부호화를 할 수 있다.

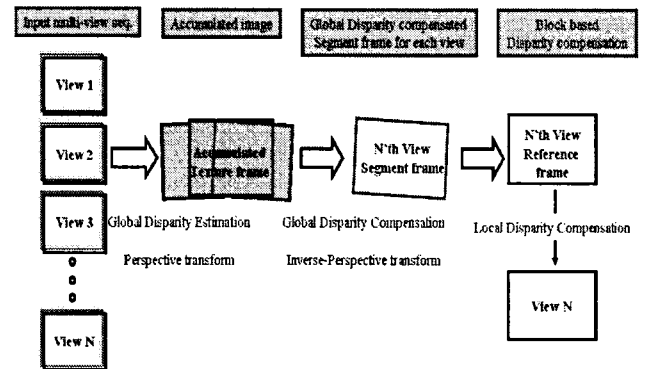
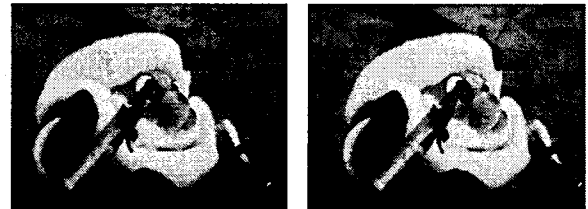


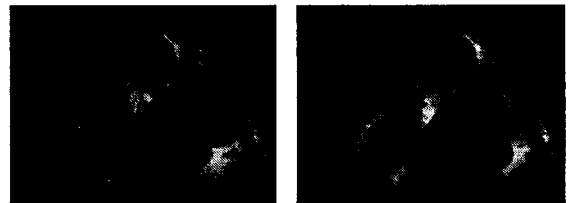
그림 2. 집적영상을 이용한 시차 예측



(a) 좌측 원영상 (b) 집적영상에서 분리된 참조영상(좌영상)



(c) 우측 원영상



(d) (a)와 (c)의 차영상 (e) (a)와 (b)의 차영상

그림 3. 집적영상을 이용한 참조 영상 생성 (평행식 카메라 구조)

하지만 폭주식 카메라 구조를 가지는 다시점 영상의 경우에는 배경을 기준으로 원근변환을 수행하기 때문에 배경에 대한 시차는 줄일 수 있지만 카메라 구조의 특성상 객체에 대한 시차는 증가하였다. 폭주식 구조를 가지는 다시점 영상에서 집적영상을 이용하여 생성한 참조영상과 원 영상의 시차를 그림 4에 나타내었다. 그림 4의 (a)는 집적영상을 이용하여 생성한 참조영상과 좌영상의 차이를 나타낸 영상이며, (b)는 우영상과 좌영상의 차이를 나타낸 영상이다. 그림 4의 (a)에서 보는 바와 같이 배경에 대한 차성분은 줄일 수 있으나 객체에 대한 차성분은 증가하게 된다. 이는 원근변환을 적용할 때 단순히 배경에 기준을 둔 이차원 변환을 적용하였기 때문이다. 객체에 대한 차성분을 효율적으로



(a) 좌영상과 우영상의 중첩영상 (b) 생성된 참조영상과 좌영상의 중첩영상  
그림 4. 집적영상을 이용한 참조 영상 생성 (폭주식 카메라 구조)

로 감소시키기 위해서는 카메라와 객체 사이의 깊이정보를 고려한 3차원 원근변환이 필요 하다.

### 3. 집적영상을 이용한 다시점 동영상 부호화 기법

제안하는 다시점 동영상 부호화 기법은 크게 각 프레임간의 시간적 중복성을 줄여주는 부분과 각 카메라 간의 공간적 중복성을 줄여주는 부분으로 구성된다. 그림 5은 집적영상을 이용한 다시점 동영상 부호화 기법을 나타낸 것이다.

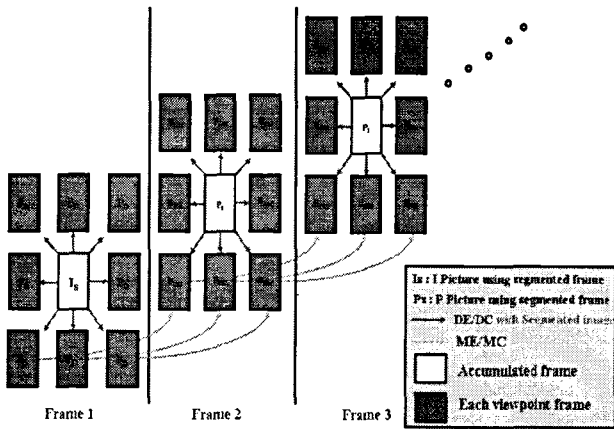


그림 5. 집적영상을 이용한 다시점 동영상 부호화 기법

그림 5는 3X3의 이차원 카메라 구조를 가지는 다시점 영상을 부호화 하는 예이다.  $I_s$ 와  $P_s$ 는 기존의 동영상 부호화기에서 사용되는 I와 P 픽처를 의미한다. 집적영상은 중앙영상을 기준 프레임으로 다른 시점 영상을 변환하여 생성하게 된다. 생성된 집적영상을 부호화하는 과정에서 움직임 예측을 통한 프레임간의 시간적 중복성을 제거할 수 있다.  $P_D$ 는 집적영상에서 분리된 참조영상을 이용하여 각 시점에 해당하는 영상과의 공간적 중복성을 제거하는 시차예측 영상을 의미한다.  $B_{DM}$ 은 이전의  $P_D$ 영상을 이용하여 움직임 예측과  $P_s$ 를 이용한 시차 예측을 동시에 수행하는 구조로써 기존의 동영상 부호화기에서 양방향 예측을 하는 B픽처와 유사한 개념을 가지게 된다.

이와 같은 구조를 이용하여 최종적으로 집적영상에 대한 부호화 시퀀스와 각 시점별 시차벡터와 차영상을 전송하게 되고 수신단에서는 두개의 시퀀스를 이용하여 복호화 할 수 있다. 이때 집적영상의 시퀀스에 동기를 맞추어 각 시점별 시차벡터와 차영상을 전송함으로써 별도의 다중화기 없이 각 시점별 동기화 문제를 해결 할 수 있다.

### IV. 실험 및 결과

제안한 다시점 동영상 부호화기의 성능을 평가하기 위하여 3X3의 평행식 카메라 구조를 가진 Santa영상을 부호화 하였다. Santa 영상은 각 카메라 간의 캘리브레이션이 아주 잘 된 영상으로 각 시점에 해당하는 영상의 전역 시차가 정확하게 일치한다. 실험영상의 각 시점에 대하여 각각 부호화 한 방법(simulcast)과 제안한 기법의 부호화 효율에 대하여 비교하였다. 동영상 부호화는 기존의 디지털 방송 표준인 MPEG-2를 사용하였으며, 이때 GOP는 6이고, 예측 모드는 I, P 픽처를 사용하였다. Q-factor는 고정하여 가변비트율(VBR)방식으로 비교하였다. 표 1은 두 기법으로 부호화된 데이터의 크기를 비교한 것이다. 표에서 보는바와 같이 9개의 영상을 각각 독립적으로 부호화 한 결과와 제안한 기법으로 부호화 한 결과를 비교해 볼때 제안한 기법이 독립적 부호화 기법에 비해 약 10% 정도 데이터량이 줄어든 것을 확인할 수 있다.

표 1. 압축 영상의 데이터 크기 (단위 : byte)

	0.0	0.1	0.2	1.0	1.1	1.2	2.0	2.1	2.2
simulcast	331,430	324,804	309,653	327,910	334,499	319,709	329,741	335,885	327,406
Proposed	302,911	302,611	288,676	299,828	312,037	300,488	301,242	311,765	306,829

그림 6과 그림 7은 제안한 기법과 독립 부호화 한 기법의 프레임별 PSNR을 나타낸다. 시간적 중복성만 제거 가능한 독립적 부호화 방식의 경우에는 각 프레임 별 PSNR이 고르지 못한 것을 확인할 수 있다. 하지만 제안한 기법의 경우에는 집적영상을 이용하여 참조영상을 만들어내고, 해당되는 참조영상에서 시차를 예측하기 때문에 각 프레임에 대한 PSNR이 고르게 분포되는 것을 확인할 수 있다. 평균 PSNR은 제안한 기법이 독립적 부호화 기법에 비하여 거의 유사하였으며, 주관적 화질 측면에서도 거의 차이를 느낄 수 없었다.

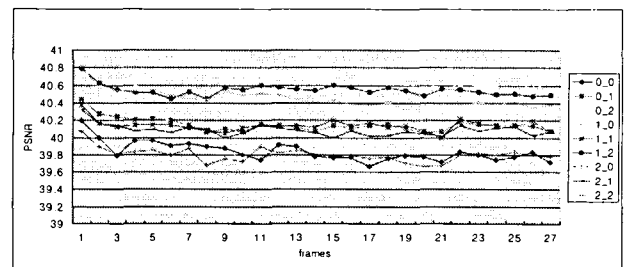


그림 6. 제안된 부호화기법을 이용한 PSNR 결과

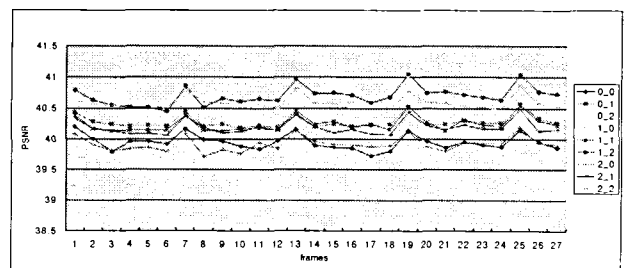


그림 7. 독립적 부호화기법을 이용한 PSNR 결과

또한 각 시점에 대하여 독립적으로 부호화 할 경우 프레임간의 동기화를 위한 별도의 다중화기와 역 다중화기가 필요한 반면 제안하는 기법은 별도의 다중화기 없이 동기화 문제를 해결할 수 있었으며, 집적영상만 전송하여 중앙영상을 볼 수 있으므로 기존의 TV시스템과도 호환이 가능하다. 하지만 제안한 기법에서 폭주식 카메라 구조를 가지는 실험영상의 경우에는 깊이정보를 고려하지 않은 원근변환으로 인하여 객체의 집적영상에서 분리해낸 참조영상과 원 영상의 객체 부분에서 시차가 커지는 현상을 확인할 수 있었다. 이는 부호화 효율의 저하시키는 원인이 된다. 결론적으로 폭주식 카메라 구조를 가지는 다시점 영상의 경우에는 깊이정보를 고려한다면 개선된 결과를 얻을 수 있을 것으로 기대된다.

\*본 연구는 대학 IT연구센터 육성\*지원 사업의 연구결과로 수행되었음.

#### <참고 문헌>

- [1] A. Redert, E. Jendriks, J. Biemond "3-D scene reconstruction with viewpoint adaptation on stereo displays", *IEEE Trans. Circuits and systems for video tech.*, vol.10, pp.550-562, 2000
- [2] Puri, R.V. Kollarits, B.G. Haskell, "Basics of stereoscopic video, new compression results with MPEG-2 and a proposal for MPEG-4", *signal processing : Image comm.*, vol.10, pp.201-234, 1997
- [3] Veksler, O."Semi-dense stereo correspondence with dense features", *Stereo and Multi-Baseline Vision, 2001. (SMBV 2001). Proceedings. IEEE Workshop on* , 9-10 Dec. 2001 Pages:149 - 157
- [4] A. Redert, et al., "ATTEST : Advanced Three-Dimensional Television System Technologies," *3D Data Processing Visualization and Transmission, 2002. First International Symposium*, pp. 313-319, Jun, 2002
- [5] ATTEST : Advanced Three-Dimensional Television System Technologies, <http://eigg.res.cse.dmu.ac.uk/attest/> 2002 - 2004
- [6]MPEG, <http://mpeg.telecomitalialab.com/>
- [7]"Applications and Requirements for 3DAV," MPEG N5539, March 2003.
- [8]"Description of Exploration Experiments in 3DAV," MPEG N5557, March 2003.
- [9]"Report on Status of 3DAV Work Exploration," MPEG N5558, March 2003.