

# 다시점 비디오와 깊이 정보의 상관도를 이용한 효율적인 다시점 비디오 부호화 기법

\*배병규 \*윤정환 \*김동욱 \*유지상

광운대학교

\*redncap@kw.ac.kr

## An efficient multi-view video coding using correlation between multi-view video and depth map

\*Bae, Byungkyu \*Yun, Junghwan \*Kim, Dongwook \*Yoo, Jisang

Kwangwoon University

### 요약

본 논문에서는 다시점 비디오와 깊이 정보의 상관도를 이용해서 현재 JVT (joint video team)에서 표준화 된 다시점 비디오 부호화 (multi-view video coding : MVC)의 참조 소프트웨어인 JMVM(joint multi-view video model)을 기반으로 하여 효율적인 다시점 비디오 압축 방법을 제안한다. 기존의 일반적인 비디오 부호화 방식은 단일 시점에 대한 비디오 부호화 기술이기 때문에 다시점 비디오 전송을 위해서는 시점 당 각각 전송 채널이 필요하다. 하지만 다시점 비디오 부호화 기법을 이용하게 되면, 단일 전송 채널을 이용하여 전송이 가능하다. 본 논문에서 제안된 방법은 입력된 다시점 입력 영상과 해당 하는 깊이 정보를 이용하여 시점 간의 예측 방법의 효율성을 높였다. 다시점 입력 영상과 깊이 정보의 전역 변이 벡터(global disparity vector : GDV)의 상관도를 이용하였으며, 다시점 영상과 깊이 정보를 동시에 전송해야 할 경우 복잡도를 낮출 수 있고, 약 0.01~0.1dB의 PSNR 이득을 얻을 수 있다.

## 1. 서론

영상 콘텐츠 기술은 사진, 방송이나 영화 분야에서 폭넓게 사용되고 있다. 영상 관련 기술들은 아날로그 방식으로부터 디지털 방식으로 발전하면서, 고해상도의 화질을 지원하는 HDTV(high definition TV)급의 영상 콘텐츠로 발전이 가속화되었고, 현재는 UHDTV(ultra high definition TV)급의 초고해상도를 지원하는 단계까지 발전하고 있다. 그리고 최근 MPEG(moving picture experts group)에서 표준화 의제로 대두되고 있는 FTV(free viewpoint TV) 시스템은 사용자에게 자유로운 시점과 3D 입체 영상을 제공하기 위한 차세대 영상 기술로써 관심이 증대되고 있는 추세이다[1].

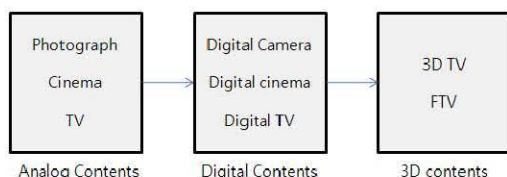


그림 1. 영상 콘텐츠의 발전과 방향

ISO/IEC의 산하의 JTC1/SC29/WG11인 MPEG 그룹에서는 예전부터 3DAV(3-dimensional audio visual)라는 AhG(ad hoc group)이 사용자에게 자유로운 시점 및 넓은 화면을 통한 입체감 제공을 위해 연구를 진행하였다. 기존의 자유로운 시점과 입체감 제공을 위해서는 다시점의 입력 영상이 필요하게 되었다. 이를 위해서는 시점 증가에 비례해서 늘어나는 데이터양 때문에 다양한 서비스를 지원하기 위해서는 전송과 저장, 처리 부분에 많은 어려움이 발생한다. 이에 따라 MPEG-4 3DAV에서는 다시점 비디오 부호화 표준화 작업을 진행하였고, 2006년 7월에 JVT로 다시점 비디오 부호화 표준화 작업이 이전되어 연구가 진행되면서 2008년 7월에 표준화 작업이 완료되었다. 다시점 비디오 부호화 기술은 서로 다른 위치의 카메라들로부터 획득된 영상들에 대해 사용자가 원하는 시점의 장면을 볼 수 있도록 시스템과 사용자 간의 상호작용을 가능하게 하는 미래의 3차원 TV, 자유 시점 비디오, 3차원 보안 시스템 등을 지원하기 위한 새로운 형태의 비디오 부호화 표준이다.

차세대 영상 기술에 대한 사용자들과 기업들의 요구도가 증가되면서, 3D 입체 영상과 자유 시점 비디오에 대한 관심이 증대되고, 그에 따라 다시점 비디오 부호화 기술은 요소 기술로써 그 활용 가치의 기대도가 높아지고 있다. 그에 따라 MPEG에서도 2007년 4월 제80차 MPEG회의에서 FTV AHG이 결성되었고, FTV에 대한 요

구 사항이 제안되어 7월 제81차 MPEG 회의에서 채택되었다. 표준화 작업이 완료된 다시점 비디오 부호화 기술은 FTV 시스템과 밀접한 연관이 있기 때문에 FTV AHG내에서도 FTV 시스템 구성을 위해서 다시점 비디오 부호화와 밀접한 관계를 유지하면서 표준화를 진행하고 있다.

FTV 시스템에서 깊이 정보는 필수적인 요소기술이다. FTV 시스템에서 사용되는 다시점 비디오와 깊이 정보의 데이터양은 매우 크다. 그리고 방대한 데이터를 효율적으로 처리할 수 있는 부호화기가 필수적인 요소이다. 본 논문에서는 다시점 비디오와 깊이 정보를 부호화하기 위해서 필요한 처리 과정을 줄여줄 수 있는 방법을 제안한다. 다시점 비디오 부호화 기법의 참조 소프트웨어를 이용하여 다시점 영상과 깊이 정보의 전역 변이 벡터(global disparity vector) 상관성을 이용해서 부호화 효율을 높였다.

본 논문의 II장에서는 다시점 비디오 부호화 기술에서 시점 간 예측에 관해서 간략히 설명하고, III장에서는 다시점 영상과 깊이 정보의 전역 변이 벡터를 이용한 부호화 기법을 제안한다. IV장에서는 기존 다시점 비디오 부호화/복호화와 비교하여 실험 결과가 제시되며, V장에서 결론을 내린다.

## 2. 다시점 비디오 부호화에서의 시점 간 예측

다시점 비디오 부호화에서 시점 간의 예측은 변이 벡터 예측과 움직임 생략 모드를 사용한다. 비디오 부호화에서 B-픽처는 시간적으로 앞뒤에 놓인 두 영상으로부터 부호화되는 영상이다. 따라서 부호화 복잡도는 증가하나, 보다 적은 비트를 이용하여 좋은 화질을 나타낼 수 있다. B-픽처는 이전 참조 영상을 이용하는 경우, 이후 참조 영상만을 이용하는 경우, Direct 모드, 두 참조 영상을 모두 이용하는 경우로 나누어진다. 두 영상을 모두 이용하는 경우에는 한 블록에 대한 움직임 벡터가 두 참조 영상으로부터 사용된다. 그러나 카메라의 위치나 다른 영상의 움직임을 이용한 예측을 통해 이 두 벡터를 보다 적은 비트로 부호화할 수 있다. 그리고 움직임 생략 모드를 사용한 기법은 이웃한 시점 간의 높은 상관도 때문에 움직임 생략 모드는 이전 시점 간의 부호화에 사용된 움직임 정보를 사용하여 부호화 효율을 높일 수 있다. 움직임 생략 모드가 사용되면 현재 매크로 블록의 움직임 정보는 전송되지 않아도 되며, 플래그만 전송함으로써 데이터 양을 줄일 수 있다.

현재 참조 소프트웨어는 전역 변이 벡터를 기본으로 현재 시점에 인접한 주변 시점들의 매크로블록 모드, 참조 영상 색인, 움직임 벡터 등의 움직임 정보를 사용하여 부호화 효율을 높이는 방법을 사용한다. 여기서, 전역 변이 벡터가 슬라이스 헤더에서 전송되기 때문에, 전역 변이 벡터를 움직임 생략 모드뿐만 아니라 움직임/변이 예측 방법에서도 고려할 수 있다. 이는 현재 부호화하는 영상이 다른 시점의 영상을 참조하고, 부호화하는 블록의 참조 영상 인덱스 값이 주변 블록들의 참조 영상 인덱스 값들과 같지 않을 경우, 참조 영상 인덱스 값이 같지 않은 주변 블록들의 움직임/변이 벡터를 전역 변이 벡터로 가져다 사용할

수가 있다[4].

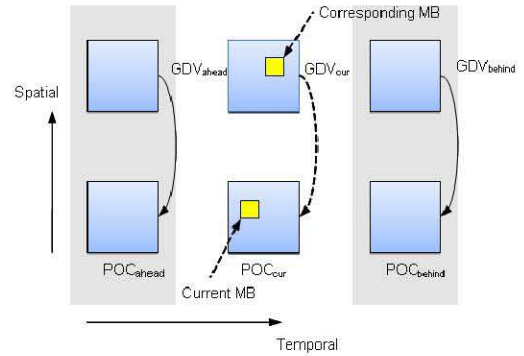


그림 2. 움직임 생략 모드를 위한 전역 변이 벡터

그림 2에서 참조 영상 사이에서 복호화되는 영상들을 위한 전역 변이 벡터는 현재 시점과 주변 시점 영상 사이의 전역 변이 벡터값을 고려하며, 식 (1)에 의해 구해진다[2][4].

$$GDV_{cur} = GDV_{ahead} + \left[ \frac{POC_{cur} - POC_{ahead}}{POC_{behind} - POC_{ahead}} \times (GDV_{behind} - GDV_{ahead}) \right] \quad (1)$$

$GDV_{cur}$ 는 현재 프레임의 전역 변이 벡터값이며,  $GDV_{ahead}$ ,  $GDV_{behind}$ 은 각각 복호된 과거와 미래 프레임의 전역 변이 벡터값을 의미한다.  $POC_{cur}$ ,  $POC_{ahead}$ ,  $POC_{behind}$ 는 프레임의 순서를 의미하는 POC(picture of count)의 각각 현재, 과거, 미래의 프레임임을 의미한다.

## 3. 제안하는 방법

다시점 영상과 깊이 정보는 경계정보의 거의 동일하기 때문에 시점 간의 변이값이 비슷하게 발생하게 된다. 그림 3에서 위쪽의 두 그림은 카메라에서 획득된 3번,4번 시점의 영상이다. 그리고 아래쪽 두 그림은 각각 위의 그림에 해당하는 깊이 정보이다. 그림 3에서 선의 왼쪽 부분은 영상이 카메라로 획득될 때 있어서 카메라 위치에 따른 영상차이다. 이 때 발생하게 되는 영상 차는 전역 변이 벡터 크기로 나타낼 수 있고, 실제로 영상 간의 상관도 기반으로 추정할 수 있다. 또한 카메라에 의해 획득된 영상들에 대한 깊이 정보를 구하면, 원 영상의 경계 부분과 깊이 정보의 경계 부분이 상관도를 비교했을 때 원영상과 유사하게 발생된다. 실제로 그림 4에서 3번과 4번 시점의 원 영상과 깊이

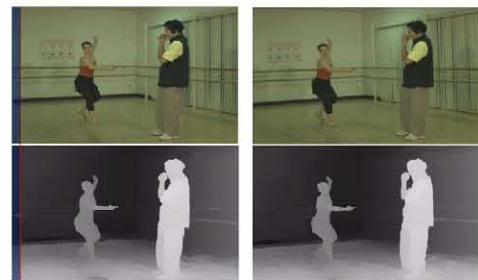


그림 3. 시점 간의 카메라 위치에 따른 영상차

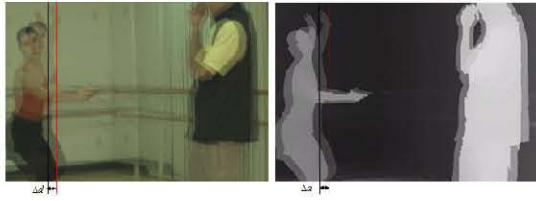


그림 4. 원 영상과 깊이 정보와의 전역 변이 벡터의 유사성

정보가  $\Delta d$  만큼의 차이 값이 동일 위치선 상에서 발생한다는 것을 알 수 있다. 따라서 변이 벡터는 3번과 4번 영상의 원 영상과 깊이 정보의 영상에서 값이 거의 동일하게 구해진다.

다시점 비디오 부호화에서는 변이 벡터를 움직임 생략 모드에 의해서 부호화하는 방법이 사용된다. 움직임 생략 모드를 위해서 전역 변이 벡터를 구해야 하며, 각각의 세밀한 움직임 벡터를 위해 RDV(region of disparity vector)를 사용한다[5]. 전역 변이 벡터를 유도하는 과정은 식 (2)과 같으며, 식 (3)을 통해서 시점 간의 매칭되는 매크로 블록을 탐색하게 된다. 참조 소프트웨어에서 시점 간의 전역 변이 벡터 값은 GDV라는 클래스의 멤버 변수인  $gdvL0$ 과  $gdvL1$ 에 저장되어 슬라이스 헤더 정보에 저장된다[3]. 그리고 변이 정보(disparity)를 구할 때 GDV의 정보만큼 이동하여 8-pel의 정확도로 주변 픽셀을 탐색하게 된다. 제안하는 방법은 깊이 정보에서 전역 변이 벡터를 구해서 GDV의 정보를 원영상에 적용시키고,  $gdvL0$ 과  $gdvL1$ 의 값을 갱신하여 전역 변이 벡터의 값을 적용시킨다. 따라서 카메라가 간격이 좁아 겹치는 영역이 많을수록 시점간의 탐색 범위를 줄일 수 있기 때문에 전역 변이 벡터값에 대한 연산을 줄일 수 있다.

$$GDV = (x, y) = \arg \min_{-SR \leq x, y \leq SR} \{MAD(8 * x, 8 * y)\} \quad (2)$$

$$MAD(x, y) = \frac{1}{(h-y)(w-x)} \sum_{i=0}^{w-x-1} \sum_{j=0}^{h-y-1} |I_r(i+x, j+y) - I_c(i, j)| \quad (3)$$

$SR$  : the search range with 8-pel accuracy  
 $I_c$  : current picture     $I_r$  : inter-view reference picture  
 $w$  : the width of the picture     $h$  : the height of the picture  
 $i$  : horizontal coordinate     $j$  : vertical coordinate

#### 4. 실험 조건 및 결과

본 실험에서는 테스트 영상들을 이용하여 기존의 참조 소프트웨어(joint multi view video model : JMVM) 와 제안한 알고리즘의 PSNR, 평균비트율, 수행시간의 변화량을 측정하여 비교하였다.

표 1은 참조 소프트웨어의 실험환경을 나타내었고, 탐색 범위는  $\pm 64$ 로 전탐색을 실시하였다.

표 2는 실험 영상들에 대한 정보를 나타내었으며, 표준의 Basis QP값을 사용하여 실험을 수행하였다.

표 1. 실험 환경

CPU	Core 2 Duo E6600
메모리	2 GBytes
O S	Window XP
참조 소프트웨어	JMVM7.2
Search range	$\pm 64$

표 2. 실험 영상

영상	영상 크기	프레임	Basis QP
Breakdancers	1024x768	100	22, 27, 32, 37
Ballet			

표 3은 실험한 영상의 모든 시점의 평균값의 결과를 정리한 것으로 PSNR, 평균비트율, 부호화에 걸린 시간의 변화량은 각각 식 (4), (5), (6)에 정의하였다.

$$\Delta YPSNR = PSNR_{reference} - PSNR_{proposed} \quad (4)$$

$$\Delta Bitrate = \frac{Bitrate_{reference} - Bitrate_{proposed}}{Bitrate_{reference}} \times 100\% \quad (5)$$

$$\Delta T = \frac{Time_{reference} - Time_{proposed}}{Time_{reference}} \times 100\% \quad (6)$$

표 3. 실험 결과

	QP	$\Delta YPSNR$ (dB)	$\Delta Bitrate$	$\Delta T$
Breakdancer	22	0.04	3.04%	10.3%
	27	0.08	3.91%	10.9%
	32	0.07	1.18%	10.8%
	37	0.02	1.15%	10.3%
평균		0.05	2.41%	10.6%
Ballet	22	0.06	1.23%	10.4%
	27	0.09	1.22%	10.3%
	32	0.12	0.76%	10.9%
	37	0.11	0.92%	11%
평균		0.09	1.03%	10.6%

제안한 방법이 Anchor에 비해 통해 수행시간이 10%정도 감소되었다는 것을 표 3을 통해 알 수 있고, 그림5와 6을 통해 이전 알고리즘에 비해 PSNR이 약0.01~0.1dB 정도의 향상된 것을 확인할 수 있다.

#### 5. 결론

본 논문에서는 다시점 비디오와 깊이 정보의 상관도를 이용한 시점 간의 예측 알고리즘을 제안하였다. 기존의 전 탐색 알고리즘과 비교해 본 논문에서 제안한 알고리즘은 PSNR과 비트는 거의 유지하면서, 비디오 부호화 시간을 약 10%정도 단축시킬 수 있었

다. 제안하는 알고리즘은 영상의 크기가 클수록, 영상 간의 카메라 위치가 가까울수록 연산이 추가로 줄어들기 때문에 해상도가 높은 HDTV급 화질에 활용하면 높은 효율을 기대할 수도 있다.

### 감사의 글

“본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 성·지원사업의 연구결과로 수행되었음.” (IITA-2008-C1090-0801-0017)

### 참 고 문 헌

- [1] ISO/IEC JTC1/SC29/WG11, "Proposal on Requirements for FTV", N14417, April 2007
- [2] ISO/IEC JTC1/SC29/WG11, "Joint Multi-view Video Model(JMVM) 8.0", N9762, April 2008
- [3] ISO/IEC JTC1/SC29/WG11, "Study of Text of ISO/IEC 14496-10:200X/FPDAM 1 Multi-view Video Coding", JVT-U211, April 2008.
- [4] ISO/IEC JTC1/SC29/WG11, "MVC Motion Skip Mode", JVT-W081, April 2007.
- [5] ISO/IEC JTC1/SC29/WG11, "Regional Disparity Est/Comp for MVC", JVT-U047, Oct. 2006.
- [6] C.L. Zitnick, S.B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High-Quality Video View Interpolation Using a Layered Representation," ACM Trans. Graphics, vol. 23, no. 3, July 2004.
- [7] ISO/IEC JTC1/SC29/WG11, "Common Test Conditions for Multiview Video Coding", JVT-U211, Oct. 2006.
- [8] G. Bjontegaard, "Calculation of average PSNR differences between RD-curves," ITU-T Q6/SG16, Doc. VCEG-M33, Apr. 2001.
- [9] Lain E.G, Richardson, "H.264 and MPEG-4 Video Compression: Video Coding for Next-generation Multimedia", Jhon Wiley and Sons, 2003.
- [10] 호요성, 김승환, "H.264/AVC 표준의 소스 코드 분석" 2007 두양사
- [11] Y.-S. Ho, K.-J. Oh, C. Lee, P.-K. Park, and B. H. Choi, "Global Disparity Compensation for Multi-view Video Coding," ISO/IEC JTC1/SC29/WG11, Doc. M13581, Jul. 2006.
- [12] ISO/IEC JTC1/SC29/WG11, "Fine Motion Matching for Motion Skip Mode in MVC", JVT-Z021, Jan. 2008.