

기준 화면을 이용한 다시점 영상 부호화의 빠른 모드 결정 방법

정충현^o 신광무 정기동

부산대학교 컴퓨터공학과

hobori25@gmail.com, sin@pusan.ac.kr, kdchungm@melon.cs.pusan.ac.kr

A Fast Mode Decision using Anchor Pictures for Multiview Video Coding

Choonghyun Jung^o Kwangmu ShinKidong Chung

Department of Computer Engineering, Pusan National University

요 약

다시점 영상 부호화에서는 시점 간의 공간적 중복성을 이용하여 데이터 중복성을 제거하는 것이 중요하다. 독립적으로 부호화하는 동시 부호화 방법(simulcast)보다 부호화 효율이 더욱 향상하였지만 계산 복잡도가 증가하는 문제가 있다. 본 논문에서는 다시점 영상 부호화기의 계산 복잡도를 감소시키기 위한 빠른 모드 결정 방법을 제안한다. GOP 내의 양 끝에 위치하고 있는 기준 화면의 MAD를 계산하여 영역을 분할하고 영역 맵을 생성한다. 시점 간의 예측을 사용하는 시점의 경우 인접 시점의 기준 화면도 이용하여 영역을 분할한다. 생성된 맵은 비기준 화면의 부호화 시 적용되어 후보 모드를 조기에 판단한다. 이와 같은 방법을 적용한 후의 실험 결과, 화질의 손실이 거의 없으면서 부호화 시간은 평균 58.6% 감소하였고, 비트율은 평균 1.9% 증가하였다.

1. 서 론

영상 관련 기술들의 디지털 방식에서의 전환은 현재 단순한 영상 콘텐츠 형태에서 3D 입체 영상과 같은 실감형 콘텐츠로의 변화를 가속화시키고 있다. 실감형 콘텐츠 응용에는 3D 입체 비디오, 3D 원격 의료, 3D 원격지간 회의, 자유시점 TV, 자유시점 비디오 등이 있다[1][2]. 다시점 영상(Multiview Video)은 실감형 콘텐츠 응용의 기초 기술로서 사용될 수 있다. 다시점 영상은 카메라 수에 비례하는 방대한 양의 데이터를 가지기 때문에 이를 효율적으로 부호화하는 것은 필수적이다. JVT(Joint Video Team)에서 다시점 영상 부호화 표준화[3] 작업을 시작하여 2009년 1월에 완료하였고, 참조 모델인 JMVM(Joint Multiview Video Model)을 개발하였다. JMVM은 기존의 단일시점 영상 부호화인 H.264/AVC를 기반으로 하는 시간적, 공간적, 통계적 중복성을 이용하여 데이터의 중복성을 제거하는 다양한 기술들 뿐만 아니라 여러 대의 카메라를 사용하는 특징을 이용하여 인접한 시점의 데이터 중복성을 제거하는 기술들을 사용한다. 시간적인 예측은 계층적 B 구조와 시점 간 예측은 IBPBP...구조를 적용한다.

다시점 영상 부호화를 위한 참조 소프트웨어인 JMVC 5.0에서 사용하는 모드에는 Skip, Direct 16×16, Direct

8×8, Direct 8×8 FRExt(Fidelity Range Extension), Inter 16×16, Inter 16×8, Inter 8×16, Inter 8×8, Inter 8×8 FRExt, Inter 8×4, Inter 4×8, Inter 4×4, Intra 16×16, Intra 8×8, Intra 4×4, Intra PCM0이 있다. 부호화 시, 각 모드들의 비트율-왜곡 값을 비교하여 최적의 값을 가지는 모드를 선택하고 부호화 한다. 각 모드들 중 Inter 모드들은 움직임 보상 예측을 위하여 설정된 탐색 범위만큼 이전에 부호화된 프레임을 참조하기 때문에 Skip 모드나 Intra 모드에 비해 계산 복잡도가 아주 높다. 인접 시점을 참조 영상으로 사용하는 시점의 경우 변이 보상 예측을 위한 계산 복잡도가 추가적으로 더해진다.

최근에는 다시점 영상 부호화의 계산 복잡도를 감소시키기 위한 빠른 부호화 방법에 대한 연구가 다양하게 진행되고 있다[4][5][6]. 하지만 시점 간의 예측을 사용하는 시점에서의 빠른 부호화 방법만이 제안되었다. 본 논문에서는 다시점 영상 부호화기의 계산 복잡도를 감소시키기 위하여 기준 화면을 이용한 영역 분할 기법을 활용하여 빠른 모드 결정 방법을 제안한다. GOP(Group Of Pictures) 내 두 기준 화면(Anchor Picture)의 대응하는 매크로블록 간의 상관 관계를 계산하여 영역 맵(region map)을 생성하고 비기준 화면의 부호화 시 제안된 방법이 적용되어 후보 모드를 조기(early step)에 판단한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구인 다시점 영상 부호화 구조와 제안한 기법을 위한 모드 분석을 살펴본다. 3장에서는 본 논문에서 제안한

"이 논문은 2009년도 정부(교육과학기술부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임 (No. 2009-0087838)"

방법을 설명하고, 4장에서는 실험을 통해 비트율과 PSNR, 부호화 시간을 중심으로 성능을 평가한다. 5장에서는 결론 및 향후 연구과제에 대해 기술한다.

2. 관련 연구

2.1 다시점 영상 부호화 구조

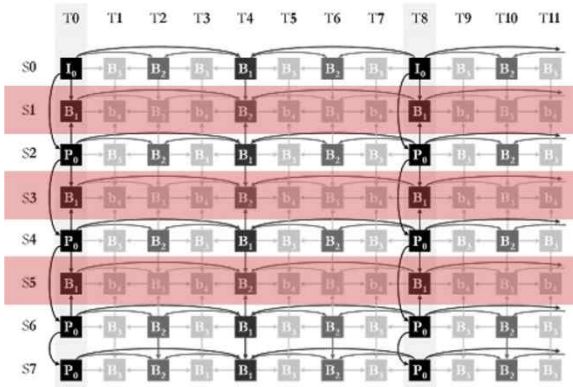


그림 1 다시점 영상 부호화 구조

그림 1은 다시점 영상 부호화 표준에서 사용하는 구조이다. 시간적으로는 계층적 B 구조이며, 시점 간으로는 IBPBP... 구조의 형태이다. S0는 시점 간 예측을 사용하지 않는다. S2, S4, S6은 기준 화면에서 시점 간 예측을 사용한다. S1, S3, S5, S7은 기준 화면뿐 만 아니라 비기준 화면에서도 시점 간 예측을 사용하여 부호화 효율을 높인다.

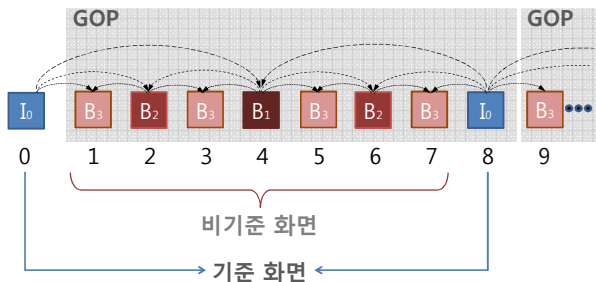


그림 2 계층적 B 예측 구조

그림 2와 같이 계층적 B 구조 예측은 GOP 단위로 수행한다. 하나의 GOP에는 하나의 기준 화면과 여러 비기준 화면들로 구성되어 있다. 부호화 과정은 화면내 예측(intra prediction)을 사용하는 기준 화면을 시작으로 비기준 화면까지 계층적인 양방향 예측(inter prediction)을 사용한다. 따라서 부호화 시 GOP 내 비기준 화면의 매크로블록들의 최적 모드는 두 기준 화면과 연관성이 있다. 본 논문에서는 두 기준 화면 사이의 대응하는 매크로블록들 간의 관계로부터 GOP 내 비기준 화면들의 선택되는 최적 모드와의 관련성을

분석하였다.

2.2 JMVC의 모드 분포 분석

본 논문에서는 참조 소프트웨어인 JMVC 5.0을 사용하여 모드 처리 시간과 모드 결정을 표 1과 같은 파라미터를 적용하여 분석하였다.

표 1 실험 환경

프레임 수	121
GOP 크기	12
탐색 범위	32
실험 영상	Ballroom, Exit, Vassar

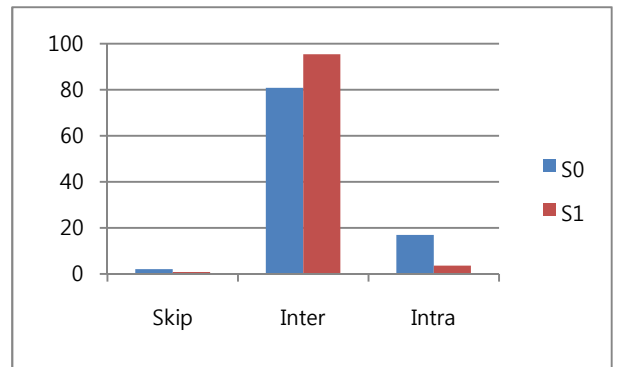


그림 3 시점 간 예측 사용 여부에 따른 모드 처리 시간 비율

그림 3에서 S0와 S1는 시점을 나타낸 것이고, 시점에 따른 모드 처리 시간 비율을 나타낸다. 대부분의 처리 시간은 Inter 모드가 차지하며 시점 간 예측을 하는 경우 그 비율이 더욱 커진다. 본 논문에서는 Inter 모드를 생략하여 부호화 복잡도를 줄이고자 한다.

두 기준 화면 사이의 대응하는 매크로블록 간의 관계와 GOP 내 비기준 화면의 선택되는 최적 모드와의 관련성을 분석하기 위하여 부호화 시, GOP 내 비기준 화면들의 최적 모드들 중 Inter 모드가 한번이라도 선택되지 않은 것을 기준으로 하여 비인터 모드 영역과 인터 모드 영역을 나누었다. 그리고 두 기준 화면의 대응하는 매크로블록들 간의 MAD를 수식 (1) 과 같이 계산하여 각각 비인터 모드 영역과 인터 모드 영역의 MAD(Mean of Absoulte Difference)의 평균을 계산하였다.

$$MAD(x,y) = \frac{\sum_i \sum_j |img1(i,j) - img2(i,j)|}{I \times J} \quad (1)$$

수식 (1)은 img1과 img2는 두 기준 화면의 대응하는 위치 x, y 의 매크로블록들 간의 절대 차이의 평균을

나타낸다.

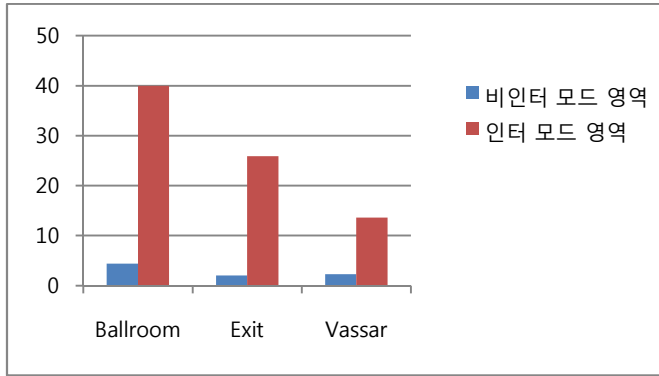


그림 4 두 영역에 따른 MAD의 평균

그림 4는 두 영역의 MAD의 평균을 나타낸 것이며, 실험 영상 마다 두 영역의 MAD의 평균 값이 차이가 나는 것을 알 수 있다. 따라서 본 논문에서는 부호화 복잡도를 감소시키기 위하여 비기준 화면의 빠른 모드 결정을 두 기준 화면의 대응하는 매크로블록의 MAD를 이용하여 영역을 분할하는 기법을 제안하였다.

3. 제안하는 빠른 모드 결정 방법

본 논문에서는 JMVM의 계층적 B 예측 구조의 특징을 이용하여 두 기준 화면의 MAD를 계산하여 영역을 분할하고 분할된 영역에서 빠른 모드 결정을 수행한다.

이웃 시점의 기준 화면들도 사용하여 동일하게 처리한다. 수식 (2)의 TH는 영역 분할을 위한 임계치를 나타내며 전체 MAD의 평균으로 설정한다. 계산된 임계치를 적용하여 수식 (3)과 같이 움직임이 적은 영역과 움직임이 많은 영역으로 분할한 움직임 영역 맵을 생성한다.

$$TH = \frac{\sum_x^X \sum_y^Y MAD(x,y)}{X \times Y} \quad (2)$$

$$B(x,y) = \begin{cases} 1 & TH > MAD(x,y) \\ 0 & otherwise \end{cases} \quad (3)$$

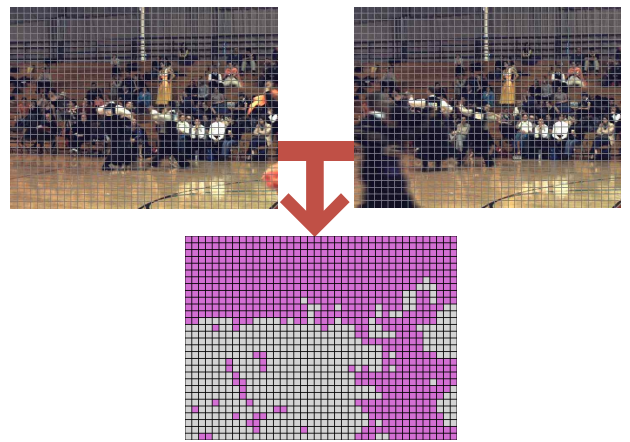


그림 6 기준 화면을 이용하여 생성한 움직임 영역 맵

그림 6은 'Ballroom' 영상에서 생성된 영역 맵을 나타낸 것이다. 영역 맵은 비기준 화면의 부호화 시 빠른 모드 결정을 위하여 매크로블록 수준에서 그림 7과 같이 적용된다.

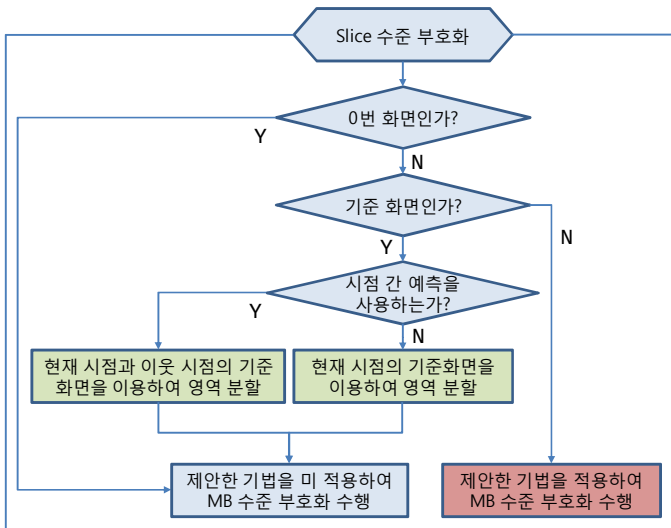


그림 5 슬라이스 수준에서의 제안하는 방법의 알고리즘

그림 5는 부호화 시 순차적으로 기준 화면과 비기준 화면에 따른 제안한 슬라이스(slice) 수준의 순서도를 나타낸 것이다. 부호화 하려는 화면이 기준 화면일 경우, 두 기준 화면을 이용하여 각 화면에 대응하는 매크로블록들의 MAD를 계산하고 영역 분할을 위한 임계치를 수식 (2)와 같이 계산한다. 만약 시점 간 예측을 사용하는 시점일 경우 현재 시점뿐 만 아니라

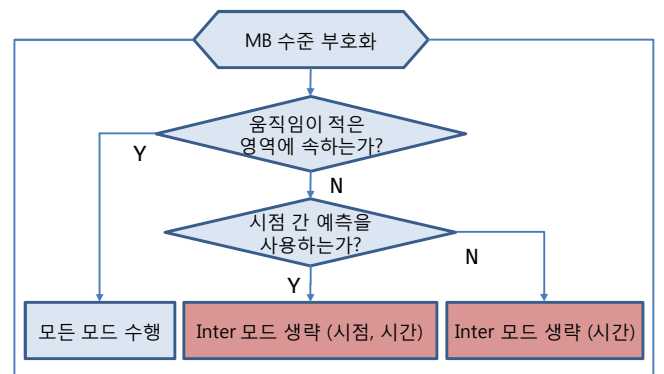


그림 7 매크로블록 수준에서의 제안하는 기법의 알고리즘

현재 매크로블록을 영역 맵과 비교하여 움직임이 적은 영역의 경우에 Inter 모드를 생략한다. 시점 간 예측을 사용하는 시점의 경우 시간뿐 만 아니라 시점간의 Inter 모드를 생략하여 부호화 복잡도를 감소시킨다.

4. 실험 및 성능평가

본 논문에서는 제안한 방법의 성능을 검증하기 위하여 다시점 영상 부호화 참조 소프트웨어인 JMVC 5.0을 기반으로 실험을 수행하였다. 실험 영상으로는 ‘Ballroom’, ‘Exit’, ‘Vassar’를 사용하였다. 실험 환경은 표 2와 같다.

표 2 실험 환경

프레임 수	GOP 크기	탐색 범위	FPS
121	12	32	25.0
해상도	QP	Symbol mode	시점
640×480	28, 32	CABAC	S0, S1, S2

제안한 방법의 성능 평가를 위해 $\Delta PSNR$, $\Delta Bitrate$, $\Delta Time$ 은 아래 수식과 같이 계산하였다.

$$\Delta PSNR = PSNR_{JMVC} - PSNR_{proposal} \quad (4)$$

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

$$\Delta Time = \frac{Time_{proposal} - Time_{JMVC}}{Time_{JMVC}} \times 100\% \quad (6)$$

표 3 QP28에서의 성능 비교

Qp28	JMVC vs proposal		
	$\Delta PSNR(dB)$	$\Delta Bitrate(\%)$	$\Delta Time(\%)$
Ballroom	0.1	4.06	-63.6
Exit	0.05	1.6	-59.6
Vassar	0.08	0.7	-46.3

표 4 QP32에서의 성능 비교

Qp32	JMVC vs proposal		
	$\Delta PSNR(dB)$	$\Delta Bitrate(\%)$	$\Delta Time(\%)$
Ballroom	0.08	3.8	-68.3
Exit	0.04	1.2	-68.2
Vassar	0.06	0.5	-46.6

제안한 방법은 시점간의 예측을 사용하는 시점뿐 만 아니라 시점간의 예측을 사용하지 않은 시점도 적용하였고, 표 3과 표 4는 QP 값에 따른 성능비교를 나타낸 것이다. 화질의 차이는 거의 없었으며, 부호화 시간은 평균 58.6% 감소하였다. 비트율은 평균 1.9% 증가하였다. ‘Ballroom’ 영상의 특징은 ‘Exit’와

‘Vassar’의 특징과 달리 움직임이 많기 때문에 두 기준 화면을 이용하여 생성한 영역 맵이 실제 비기준 화면의 영상과 일치하는 부분이 적어서 비트율의 증가가 조금 더 크게 나타났다.

5. 결론 및 향후 연구과제

본 논문에서는 다시점 영상 부호화 시 계산 복잡도를 줄이기 위하여 기준 화면의 MAD를 활용한 빠른 모드 결정 방법을 제안한다. 부호화 시 두 기준 화면의 대응하는 매크로블록간의 MAD를 계산하여 영역을 구분하고 영역 맵을 생성하였다. 비기준 화면의 부호화 시 움직임이 적은 영역에 대해서 빠른 모드 결정 방법을 적용하였다. 시점간의 예측을 사용하는 시점의 경우 인접 시점의 기준 화면도 변이맵 생성에 활용하여 예측의 정확성을 높였다. 성능을 평가하기 위하여 JMVC 5.0 참조 소프트웨어를 이용하였고 본 논문에서 제안한 방법과 비교 분석하였다. 실험결과에서 JMVC 5.0에 비해 화질의 저하가 거의 없으면서 적은 비트율 상승이 있었고 부호화 시간을 감소 시켰다.

향후 연구로서, GOP 양 끝에 위치하고 있는 기준 화면을 활용하기 때문에 GOP 크기에 따른 실험이 필요하며, 영역 분할을 위한 임계치 설정 시 누적 히스토그램 방식을 이용하면 효과적으로 영역을 분할하여 정확성 향상이 기대된다.

참고문헌

[1] ISO/IEC JTC1/SC29/WG11 M11259, "FTV (Free Viewpoint Television): achievements and Challenge," October 2004.
 [2] A. Smolic and D. McCutchen, "3 DAV exploration of video-based rendering technology in MPEG," IEEE Transactions on Circuits System Video Technology, vol. 14, no. 3, pp.348-356, March 2004.
 [3] ISO/IEC JTC1/SC29/WG11 W8019, "Description of Core Experiments in MVC," April 2006.
 [4] Xiaozhong Xu, Yun He, "Fast disparity motion estimation in MVC based on Range Prediction" ICIP 2008, IEEE International Conference on Image Processing, California, USA, October 2008.
 [5] Jianbo Lu, Hua Cai, Jian-Guang Lou and Juang Li, "An Epipolar Geometry-based Fast Disparity Estimation Algorithm for Multi-view Image and Video Coding," IEEE Trans. On Circuits and Systems for Video Technologies, vol. 17, no.6, pp. 737-750, June 2007.
 [6] Gangyi Jiang, Mei Yu, Feng Shao, You Yang and Haitao Dong, "Fast Multi-view Disparity Estimation for Multi-view Video Systems," LNCS 4179, ACIVS 2006, pp. 493-500, 2006.