

다시점 비디오 부호화에서 시점 간 예측 제한 방법 (Multi-view Video Coding using the Constrained Inter-view Prediction)

천 성 환 [†] 신 광 무 [†]
(Sunghwan Chun) (Kwangmu Shin)

김 기 완 ^{**} 정 기 동 ^{***}
(Kiwan Kim) (Kidong Chung)

요약 본 논문은 다시점 비디오 부호화를 위한 제한적인 시점 간 예측 방법을 제안한다. 다시점 비디오에서는 카메라의 위치 및 각도에 따라 시점 간에 보이지 않는 부분이 발생한다. 이와 같은 특성은 시점 간 예측을 하는 화면에서 매크로블록(Macroblock)이 보이지 않는 부분에 대해서도 양쪽의 참조 화면을 이용함으로써 예측 시간이 증가하는 요인이 된다. 본 논문에서는 이와 같은 점을 해결하기 위해 보이지 않는 부분에 속한 매크로블록의 시점 간 예측을 제한하는 방법을 적용하였다. 실험 결과, 기존의 다시점 비디오 부호화와 비교하여 4% 정도 부호화 시간을 단축시켰고 화질의 저하는 거의 없었다.

키워드 : 다시점 비디오 부호화, 전역 변이 벡터, 제한적 시점 간 예측

Abstract In this paper, we propose a method that uses the constrained inter-view prediction for multi-view video coding. In the multi-view video, there exists occluded area because of the locations and angles of

* 이 논문은 2008 한국컴퓨터종합학술대회에서 '다시점 비디오 부호화에서 시점 간 예측 제한 방법'의 제목으로 발표된 논문을 확장한 것임

[†] 학생회원 : 부산대학교 컴퓨터공학과
rainingman23@gmail.com
sin@pusan.ac.kr

^{**} 정 회 원 : 육군 3사관학교 컴퓨터공학 교수
wan1434@mmaa.or.kr

^{***} 중신회원 : 부산대학교 컴퓨터공학과 교수
kdchung@pusan.ac.kr

논문접수 : 2008년 8월 28일
심사완료 : 2008년 10월 20일

Copyright©2008 한국정보과학회 : 개인 목적이거나 교육 목적인 경우, 이 저작물의 전체 또는 일부에 대한 복사본 혹은 디지털 사본의 제작을 허가합니다. 이 때, 사본은 상업적 수단으로 사용할 수 없으며 첫 페이지에 본 문구와 출처를 반드시 명시해야 합니다. 이 외의 목적으로 복제, 배포, 출판, 전송 등 모든 유형의 사용행위를 하는 경우에 대하여는 사전에 허가를 받고 비용을 지불해야 합니다.

정보과학회논문지: 컴퓨팅의 실제 및 레터 제14권 제8호(2008.11)

cameras. This increases the computational complexity, as it still uses both reference pictures for predicting the area which is not shown in the current frame. In this paper, we propose a method that does not use the inter-view prediction in cases of the occluded macroblocks. Experimental results show that benefits about 4% can be achieved compared with the conventional approaches.

Key words : Multi-view Video Coding, Global Disparity Vector, Restricted Inter-view Prediction

1. 서론

근래 디지털 기술의 발전은 다양한 서비스를 가능하게 하였다. 그 중에서도 TV 방송 분야는, 기존 2D 형태의 평면적인 영상과 음성 정보를 전송하는 것으로부터, 현재는 양안(Stereoscopic) 정보를 이용하여 입체감 있는 3D 입체영상 기술도 선보이고 있다. 또한 촉각, 후각, 청각 등의 정보를 동시에 제공하는 실감 콘텐츠에 대한 연구도 진행되고 있다. 이와 같은 실감 영상은 양안시점 입체영상, 다시점 비디오(Multi-view Video), 홀로그래피(Holography) 등과 같은 기술을 적용함으로써 가능해진다.

특히 다시점 비디오는 기존의 단일 시점 비디오와는 달리, 여러 대의 카메라를 다양한 방식으로 배열하여 하나의 장면을 촬영 후 얻은 영상들을 통해 사용자들에게 다양한 시점의 화면을 제공한다. 이 기술을 통해 입체 TV, 자유 시점 TV[1,2] 등 다양한 영상에 응용될 수 있다. 이와 같은 다시점 비디오 기술은 사용자에게 자유로운 시점 및 영상을 입체감 있게 제공할 수 있지만, 여러 대의 카메라에서 획득한 정보의 양은 카메라의 수에 따라 비례적으로 증가하는 문제점이 있다. 따라서 효율적인 다시점 비디오 부호화가 필수적이다. 이와 같은 필요성이 요구됨에 따라, 국제 표준화 기구인 ISO/IEC JTC1/SC29/WG11 MPEG은 다시점 비디오 부호화(Multi-view Video Coding) 방법에 대한 표준화 작업을 시작하였고, 현재 JVT(Joint Video Team)에서 H.264를 기반으로 다시점 비디오 부호화의 표준화[3]를 진행 중에 있다. 다시점 비디오 부호화는 기존의 단일 시점 비디오 부호화에서 시간적, 공간적, 통계적 중복성을 이용하여 데이터의 중복을 제거하는 것 외에 다시점 비디오 부호화의 특징인 시점 간 상관도에 따른 중복성을 고려한다. 이와 같은 특징으로 인해 다시점 비디오 부호화의 예측 복잡도는 시점 간 예측을 포함하기 때문에 단일 시점 비디오 부호화의 예측 복잡도 보다 더욱 증가하게 된다. 본 논문에서는 다시점 비디오 부호화가 가지는 특징인 카메라 간의 거리 차이에 따른 객체 또는 전경의 보이지 않는 부분에 대한 시점 간 예측을 제한함으로써 예측 복잡도를 줄이는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련 연구에 대하여 살펴보고, 3장에서는 제안하는 방법을 이용하여 제한적인 시점 간 예측 방법을 설명한다. 4장에서는 제안 방법을 이용한 다시점 비디오 부호화의 성능을 분석한다. 마지막으로 5장의 결론 및 향후 과제로 글을 맺는다.

2. 관련 연구

본 장에서는 다시점 비디오 부호화에서 사용하는 참조 부호화 모델의 시·공간적 예측 구조에 대해서 간단히 소개한다. 또한, 본 논문에서 이용하는 전역 변이의 개념에 대해 설명하고, 기존 연구에서 전역 변이를 구하는 방법에 대해 소개한다.

2.1 다시점 비디오 부호화 구조

다시점 비디오 부호화의 예측 구조[4,5]는 그림 1과 같다. 시간적으로는 계층적 B 구조 예측을 이용하여 시간적 중복성을 제거하고, 시점 간 예측은 IBPBP... 구조를 이용하여 시점 간의 공간적 중복성을 제거하는 예측 구조를 사용한다. 여기서 시점 간 예측은 인접한 두 개의 시점을 이용하기 때문에 그림 1의 S1, S3, S5...에서 시점 간 예측이 사용된다.

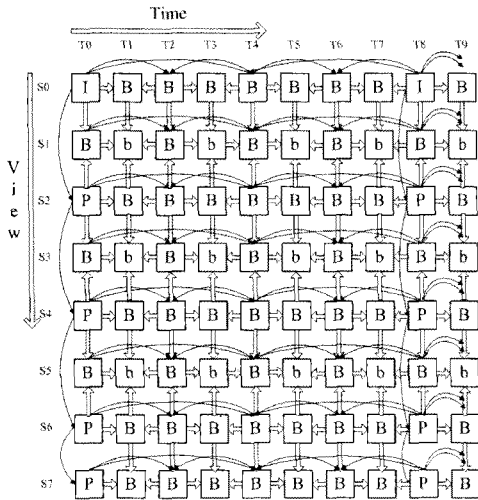


그림 1 다시점 비디오 부호화 구조

2.2 전역 변이 벡터

다시점 비디오는 여러 대의 카메라를 평행 또는 수렴으로 배치하여 객체 또는 전경을 촬영하여 영상을 획득하기 때문에 카메라간의 거리 차이가 발생한다. 즉 평행하는 카메라간의 위치 차이로 인해 인접하는 시점 간에 보이지 않는 부분이 발생한다.

그림 2는 인접한 시점 간에 보이지 않는 부분을 나타

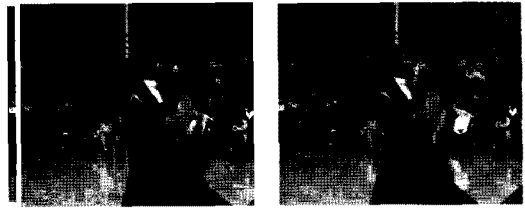


그림 2 ballroom의 시점 0과 시점 1의 전역 변이

내고 있다. 이와 같이 보이지 않는 부분에 대한 화면의 전체적인 움직임 벡터를 전역 변이 벡터(Global Disparity Vector)라고 한다.

전역 변이 벡터를 계산하는 여러 가지 방법[6] 중에 움직임 생략 모드(Motion Skip Mode)에서는 기준 화면에서 픽셀 단위로 MAD(Mean Absolute Difference)를 적용하여 기준 화면의 전역 변이 벡터를 구하게 되고[7,8] 이것을 비 기준 화면의 전역 변이 벡터로 이용한다.

2.2.1 기준 화면 전역 변이 벡터 계산

움직임 정보 생략 모드란 인접하는 시점의 영상들 간의 시간적 상관도가 유사하다는 특징을 고려하여, 이전에 부호화된 인접 시점의 움직임 정보를 공유하는 방법을 통해 부호화 효율을 높이는 방법이다.

전역 변이를 계산하기 위해서 적용하는 기법인 MAD 방법은 그림 3에서 보는 바와 같이 인접한 시점 간에 이미지가 겹쳤을 때의 가장 최소값인 (g_x, g_y) 를 찾아내는 방법이다. 하지만 이 방법은 매크로블록안의 픽셀단위로 가장 최소값을 찾아내어 전역 변이를 구하는 방법이기 때문에 복잡도가 높다는 단점이 있다.

MAD를 이용한 식은 다음과 같이 정의 된다.

$$(g_x, g_y)_{MAD} = \min[\frac{1}{R} \sum_{i,j \in R} |img0(i,j) - img1(i-x, j-y)|] \tag{1}$$

$img0$ 와 $img1$ 은 그림 3에서 전역 변이를 계산하기 위한 두 장의 이미지이고 R은 겹쳐진 영역의 픽셀들의 수

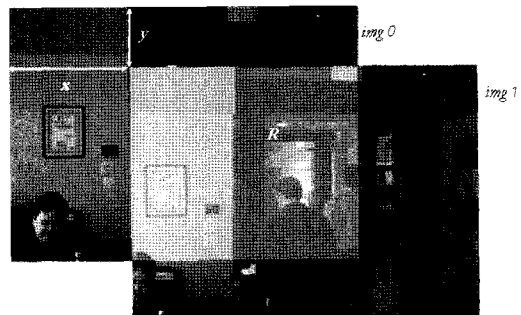


그림 3 전역 변이 계산(MAD 방법)

를 나타낸다. 가장 작은 차이의 절대값을 가지는 x, y를 화면의 전역 변이 벡터로 구한다.

2.2.2 비 기준 화면 전역 변이 벡터 계산

그림 4는 중앙에 있는 비 기준 화면(Non-anchor Frame)에서 전역 변이를 구하기 위해 양쪽의 기준 화면에서 구한 전역 변이 정보를 이용하는 것을 나타낸다. 비 기준 화면의 전역 변이를 구하는 식은 아래 식 (2)와 같다. 현재 부호화 하는 매크로블록의 전역 변이는 각 화면의 POC(Picture Order Count)와 해당 비 기준 화면이 속한 GOP(Group of Picture)의 양 끝단에 위치한 기준 화면의 전역 변이 값을 통해 유도한다. 여기서 GDV_{cur} 은 비 기준 화면에서 구하려는 현재 매크로블록의 전역 변이를 나타내고, GDV_A 와 GDV_B 는 기준 화면에서 구한 전역 변이 값이다. 또한, POC는 각 화면의 표시 순서를 나타내는 값이다.

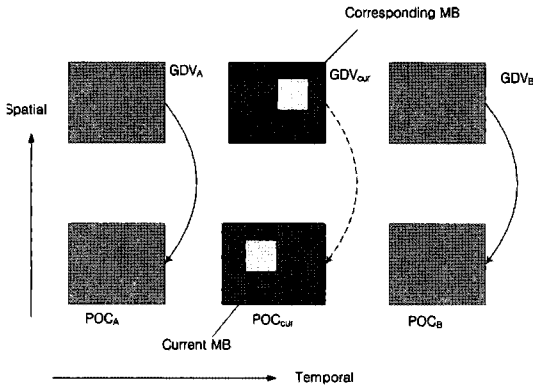


그림 4 전역 변이 계산

$$GDV_{cur} = GDV_A + \left[\frac{POC_{cur} - POC_A}{POC_B - POC_A} \times (GDV_B - GDV_A) \right] \quad (2)$$

3. 제안하는 제한적 시점 간 예측 구조

본 논문에서 제안하는 제한적 시점 간 예측은 먼저 인접하는 시점 간에 보이지 않는 부분, 즉 화면 간의 차이를 전역 변이를 통해서 구한다. 다음, 구해진 부분에 속한 매크로블록들이 기존의 시점 간 예측 구조에서 양쪽의 참조 화면 모두를 이용하는 것 대신, 현재 시점이 가지고 있지 않는 부분에 대한 예측을 제거한다.

그림 5는 B 화면으로 부호화 되어야 하는 화면이고, 그림 6은 두 개의 참조화면이다. 그림 5의 오른쪽 직사각형 부분은 그림 6의 왼쪽 참조화면에는 존재하지 않는 부분이고 왼쪽의 직사각형 부분은 그림 6의 오른쪽 참조화면에는 존재하지 않는 부분이다. 그렇기 때문에 왼쪽 부분에 속한 매크로블록과 오른쪽 부분에 속한 매크로블록은 양쪽의 참조 화면을 모두 예측할 필요 없이 같은 부분이 있는 참조화면만 예측을 하는 것이 효율적이다.



그림 5 시점 간 예측을 이용하는 화면

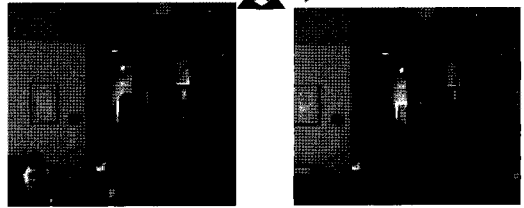


그림 6 그림 5의 두 참조 화면

크로 블록은 양쪽의 참조 화면을 모두 예측할 필요 없이 같은 부분이 있는 참조화면만 예측을 하는 것이 효율적이다.

전체적인 과정을 살펴보면 먼저 시점 간 예측을 하려는 화면의 전역 변이를 구해야 한다. 본 논문에서는 전역 변이를 구하기 위해 기준 화면에서 16x16 블록 크기의 시점 간 예측이 선택된 매크로블록의 변이를 구하고, 16x16이 선택된 횡수로 평균한 값을 전역 변이로 구하였다. 이것은 전체적인 화면의 변위를 구하는 것이기 때문에 움직임이 거의 없는 16x16 시점 간 예측 모드가 선택되었을 때의 변이를 측정하였다. 식은 다음과 같이 정의된다.

$$(g_x, g_y)_L = \frac{1}{C} \sum_{i,j \in R} \{img0(i,j) - img1(i-x, j-y)\} \quad (3)$$

$$(g_x, g_y)_R = \frac{1}{C} \sum_{i,j \in R} \{img0(i,j) - img1(i-x, j-y)\} \quad (4)$$

식 (3), (4)에서 (g_x, g_y) 는 전역 변이를 나타내고, L, R은 시점 간 예측에서 중간 시점을 기준으로 좌측 시점, 우측 시점을 나타낸다. 즉, 현재 시점을 기준으로 좌측 참조 화면과의 전역 변이를 구하고 우측 참조 화면과의 전역 변이를 구하여 현재 시점의 전역 변이로 이용하는 것이다. 양쪽의 전역 변이를 따로 구하여야만 왼쪽 참조 화면의 보이지 않는 부분과 오른쪽 참조 화면의 보이지 않는 부분에 대해 보다 정확한 전역 변이 벡터를 기준 화면에서 구할 수 있다. C는 기준 화면에서 16x16 모드가 선택된 횟수이다.

비 기준 화면에서는, 기준 화면에서 구한 전역 변이를 그대로 이용하여 참조 화면의 예측을 제한하는 매크로블록을 설정하게 된다. 시점 간 예측을 이용하는 화면에

서 매크로블록을 제한하는 것을 살펴보면, 기준 화면에서 구한 2개의 수평, 수직의 전역 변이 벡터를 이용하여 매크로블록을 구분한다. 이와 같이 제한된 매크로블록들은 양쪽의 참조 화면을 이용하는 것이 아니고 예측되지 않는 참조 화면 1개만 선택해서 예측한다. 즉, 보이지 않는 부분에 대한 시점 간 예측을 제한함으로써 예측 시간을 단축시키는 방법이다.

4. 실험 결과 및 분석

본 논문에서 제안한 알고리즘의 성능을 검증하기 위해 참조 소프트웨어인 JMVM7을 기반으로 실험을 수행하였다. GOP 구조는 시간적으로 IBBB 구조를 사용하는 계층적 B 구조를 이용하였고, 시점 간으로는 기준 화면이 인트라(Intra) 모드만 선택하는 I 구조를 사용하였다. 그리고 기본적인 QP 값은 37을 적용하였다. 실험에서 각 시점 간의 거리 차가 일정하기 때문에 사용한 시점은 3가지이다.

표 1은 실험에서 사용된 영상[3]을 나타내며, 모두 다시점 비디오 부호화 표준화에서 사용하기 위해 제공된 영상들이다. 각 영상은 카메라 개수만큼의 시퀀스를 가진다.

움직임이 많은 Ballroom 영상과 움직임이 적은 Exit 영상을 통해 실험을 하였고, 표 2는 위의 실험 영상에 다양한 탐색 범위를 통해 기존의 JMVM과 제안한 기법을 적용한 것으로 나누어 측정된 부호화 시간을 나타낸다.

표 2를 보면 Ballroom과 Exit 영상에서 기존의 JMVM보다 본 논문에서 제안한 기법의 시점 간 부호화 시간이 줄어든 것을 알 수 있다. 이것은 전역 변이 벡터만큼 제한되어진 매크로블록들이 시점 간 예측을 한쪽 만 함으로써 시간이 줄어들었기 때문이다. 또한 탐색 범위가 커질수록 부호화 시간이 줄어드는 것을 알 수 있다.

표 1 실험 영상의 종류 및 특성

영상 종류	영상 속성	카메라 배열	각 카메라당 화면수
Ballroom	640×480 25fps	8 cameras, 1-D parallel	250
Exit	640×480 25fps	8 cameras, 1-D parallel	250

표 2 부호화 시간 비교

영상 종류	탐색범위	32	64	96
		JMVM(sec)	8725	14729
Ballroom	Proposed(sec)	8512	14207	22787
	JMVM(sec)	8140	13885	20988
Exit	Proposed(sec)	8035	13354	20036

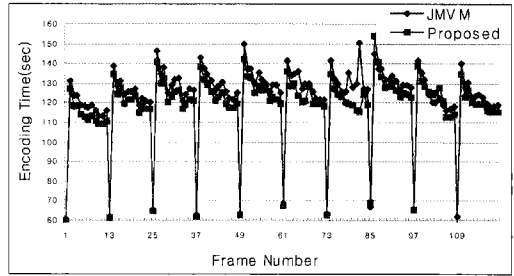


그림 7 Ballroom 영상에서 화면 간 부호화 시간 비교

그림 6은 Ballroom 영상에서 화면에 따라 JMVM과 제안한 기법의부호화 시간을 측정하여 비교한 것이다. 기존의 JMVM보다 화면마다 제안한 방법이 5~8초 정도 시점 간 예측 시간을 단축시킨다. 그림 6의 기준 화면의 부호화 시간이 비 기준 화면의 부호화 시간과 비교해서 반으로 줄어드는 결과를 보이는데 이것은 기준 화면에서 시간적 예측을 하지 않고 시점 간 예측 시간만 포함하기 때문이다.

표 3과 4는 탐색 범위에 따라 PSNR 값과 Bitrate를 나타낸 것이다. JMVM과 제안한 기법을 비교하면 PSNR 값과 Bitrate는 거의 비슷한 것을 알 수 있다. 이것은 예측할 필요가 없는 부분을 제한하기 때문이다. Ballroom 영상과 마찬가지로 Exit 영상도 비슷하게 부호화 시간이 감소하는 것을 볼 수 있다.

표 3 Ballroom 영상에서 PSNR과 Bitrate 비교

탐색 범위	PSNR(dB)		Bitrates(bits/sec)	
	JMVM	Proposed	JMVM	Proposed
32	31.8482	31.8452	190.1868	191.6843
64	31.8642	31.8602	183.9471	185.9355
96	31.8650	31.8590	183.8542	185.3636

표 4 Exit 영상에서 PSNR과 Bitrates 비교

탐색 범위	PSNR(dB)		Bitrate(bits/sec)	
	JMVM	Proposed	JMVM	Proposed
32	34.7129	34.7101	77.2926	77.5041
64	34.7263	34.7190	75.8793	76.5851
96	34.7289	34.7196	76.0612	76.4579

5. 결론

본 논문에서는 다시점 비디오 부호화 시 카메라 간의 거리에 따른 시점 간 예측에서 발생하는 보이지 않는 부분에 대한 매크로블록의 제한적인 시점 간 예측을 제안하였다. 기존의 방법은 시점 간 예측을 할 때 보이지 않는 부분이 있음에도 불구하고 양쪽의 참조 화면을 모두 예측함으로써 예측 시간이 오래 걸리는 문제점이 있

었지만 제안한 기법을 적용함으로써 예측 시간을 단축시킬 수 있었고 화질의 저하도 거의 없었다.

향후 과제로는 더욱 정확한 전역 변이 벡터를 찾아내는 것과 이것을 통해 정확한 예측을 하는 연구가 필요하다. 또한 실험적으로 정확한 통계를 위해 위의 2가지 영상 뿐 아니라 다양한 영상 환경에서의 실험도 필요하다.

참 고 문 헌

- [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 Trans. Circuits Syst. Video Technol., 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] ISO/IEC JTC1/SC29/WG11, "Description of Core Experiments in MVC," MPEG2006/W7798, January 2006.
- [5] 신광무, 이서영, 김성민, 정기동, "H.264 기반의 다시점 비디오 부호화를 위한 GoGop구조", 한국정보과학회, 가을 학술발표논문집(A), 제34권 제 2호, pp. 383-387, 2007.
- [6] ISO/IEC JTC1/SC29/WG11, "Regional Disparity Est/Comp for MVC" October 2006.
- [7] Kwan-Jung Oh, Yo-Sung Ho, "Global Disparity Compensation for Multi-view Video Coding," 방송공학회논문지, 제12권 제6호, 2007.
- [8] 효요성, 오관정, "다시점 비디오 부호화 기술 동향", 전자공학회지, 제34권 제8호, pp. 916-917, 2007.