

H.264/AVC에서 다시점 비디오 부호화를 위한 향상된 GoGOP 구조

(An Improved GoGOP Structure for Multi-view Video Coding in H.264/AVC)

신 광 무[†] 이 서 영[†]
(Kwangmu Shin) (Seoyoung Lee)

김 성 민[†] 정 기 동^{**}
(Sungmin Kim) (Kidong Chung)

요 약 근래의 지속적인 멀티미디어 기술의 발전과 콘텐츠를 이용하는 사용자의 향상된 욕구가 부합하여 새로운 실감 미디어 기술이 창출되고 있다. 그 중에서 다시점 비디오는 3차원 TV, 자유 시점 비디오 등의 기반 기술로써 활발하게 연구되고 있다. 하지만 시점 수의 증가에 따른 부호화 시간 증가, 비트율 증가 등이 문제점으로 작용한다.

본 논문에서는 다시점 비디오의 부호화 효율을 높이기 위해 향상된 GoGOP 구조를 제안한다. 즉, Key 프레임 위치 조절, 동적인 I 프레임 수 그리고 동적인 B 프레임 수를 이용하는 방법을 적용하여 향상된 GoGOP 구조를 제안하였다. 실험 결과, 본 논문에서 제안한 기법이 Anchor 구조와 비교하여 유사한 PSNR 값을 가지면서, 비트율이 상당히 감소하였다.

키워드 : 다시점 비디오, GoGOP, 부호화 효율

Abstract Corresponding recent continuous development of multimedia technology with improved desire of using various contents, a new realistic feeling media technology is being created. Of them all, multi-view video is being researched actively as foundation technology of 3D TV, free-view point video etc. But encoding time and bit-rate are increased as view numbers are increased.

In this paper we propose improved GoGOP structure to enhance the coding efficiency of multi-view video by applying methods which are using techniques such as Key frame position adjustment, dynamically changing the number of I frame and B frame. As experimental results, technique proposed in this paper reduces bit-rate having similar PSNR value compared with Anchor structure.

Key words : multi-view video, GoGOP, coding efficiency

1. 서 론

현대 사회에 상당한 영향을 미치고 있는 디지털 미디어는 멀티미디어 기술이 지속적으로 발전함에 따라 디지털 포맷의 형태로 점점 더 많이 양산, 처리, 저장 그리고 전송되고 있다. 이와 같은 현상의 근본적인 이유는 보다 더 복잡하고 다양한 콘텐츠를 요구하는 사용자의 욕구 때문일 것이다. 미디어 포맷의 종류에는 일반적으로 비디오, 오디오, 컴퓨터 그래픽스, 텍스트, 이미지 등이 있는데, 사용자의 욕구와 멀티미디어 기술의 발전이 상호 작용하여 과거에 단일 미디어 포맷에서 실감 미디어로의 접근을 위해 기존 미디어 포맷이 융합되는 형태를 보이고 있다[1].

최근에는 다시점 비디오(Multi-view Video), 양안 시점 시스템(Stereoscopic System), 홀로그래피(Holography) 등의 3차원 비디오 기술이 실감 미디어의 대안으로 연구가 진행되고 있다[2-4].

그 중에서 다시점 비디오는 3차원 TV, 자유 시점 비디오(Free-view Video) 등의 기반 기술로써 활발하게 연구 및 표준화가 진행 중이다.

다시점 비디오는 한 대 이상의 카메라를 통해 촬영된 영상들을 기하학적으로 교정하고 공간적인 합성 등으로 다양한 시점을 사용자에게 제공하는 3차원 영상처리의 한 분야이다.

다시점 비디오의 한 예인 파노라믹(Panoramic) 비디오는 컴퓨터 그래픽스, 영상처리, 컴퓨터 비전, 우주/항공 사진학에서 발전 가능성 있는 연구 주제로 자리 잡고 있으며, 이와 같은 파노라믹 비디오 기술은 비디오 압축, 영상 변화 감지, 항공사진의 해석, 카메라 해상도 확대에서 간단한 영상 편집에 이르기까지 다양한 분야에서 응용되고 있다[5].

· 이 논문은 제34회 추계학술대회에서 'H.264 기반의 다시점 비디오 부호화를 위한 향상된 GoGOP 구조'의 제목으로 발표된 논문을 확장한 것임

[†] 학생회원 : 부산대학교 컴퓨터공학과
sin@pusan.ac.kr
seoyoung@melon.cs.pusan.ac.kr
morethannow@pusan.ac.kr

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

논문접수 : 2007년 12월 7일

심사완료 : 2008년 2월 14일

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 얻고 비용을 지불해야 합니다.

정보과학회논문지: 컴퓨팅의 실제 및 레터 제14권 제3호(2008.5)

표준화 차원에서의 노력으로는 MPEG 3DAV(3 Dimensional Audio Video)가 ISO/IEC SC29 WG11에 설립되었고, 현재는 ISO/IEC와 ITU-T의 공동 스타디 그룹인 JVT(Joint Video Team)에서 SVC(Scalable Video Coding) 참조 모델 기반의 다시점 비디오 부호화(Multi-view Video Coding)에 대한 표준화가 진행 중이다[6].

다시점 비디오는 사용자에게 다양한 시점을 제공해 줌으로써 보다 현실감 있는 콘텐츠의 제공이라는 장점을 지니고 있지만, 시점 수의 증가에 따른 부호화 시간 증가, 비트율(Bit-rate) 증가 그리고 고가의 장비가 요구되는 문제점이 있다.

따라서 최근 표준화 과정에서 시·공간적 예측 구조(CE1: View-temporal Prediction Structures), 조명 보상(CE2: Illumination Compensation), 시점 보간 예측(CE3: View Interpolation Prediction), 변이 벡터 예측(CE4: Disparity Vector Prediction) 그리고 MVC(Multi-view Video Coding)를 위한 디블록킹 필터 적용(CE5: Deblocking Filter Adaptation for MVC) 등의 핵심기술 성능평가 실험(Core Experiments)을 통해 이와 같은 문제점을 해결하고자 하고 있다.

본 논문에서는 이와 같은 주제들 중에서 GoGOP(Group of Group Of Pictures) 구조의 변경을 통한 시·공간적 예측 구조의 성능 향상에 주안점을 두고 있다. GoGOP는 단일 시점(Single View)에서 픽처들의 집합을 나타내는 GOP(Group Of Pictures)를 다시점으로 확장한 개념으로, 시점 수에 따른 GOP의 그룹을 의미한다. 초기의 일반적인 GoGOP 구조로 사용되었던 것이 Anchor 구조인데, 각 시점별로 독립적으로 부호화를 수행하기 때문에 아주 비효율적이다. 따라서 그 이후에 몇 가지 효율적인 GoGOP 구조가 제안되었다. 계층적 B 프레임을 이용한 시·공간적 예측 구조는 그 성능이 입증되어 다시점 비디오의 참조 소프트웨어(Reference Software)인 JMVM에 사용되고 있다[5-8]. 본 논문의 GoGOP 구조에서 적용된 기법은 [8]에서 적용된, 시점의 중간 지점에 Key 프레임을 둬으로써 참조 프레임의 중복 영역을 확장시키는 방법, 시점 수의 변화에 따라 I 프레임 수를 조절하는 방법 그리고 I 프레임 수의 변화에 따라 B 프레임 수를 조절하는 방법이 적용되었다.

본 논문의 구성은 다음과 같다. 2장에서 관련 연구, 3장에서 제안하는 기법, 4장에서 실험 결과, 그리고 마지막으로 5장에서 결론과 향후 과제를 언급한다.

2. 관련 연구

그림 1은 Anchor 구조를 나타내는 것으로, 각 시점별로 I 프레임을 각각 둬으로써 독립적으로(Simulcast) 부

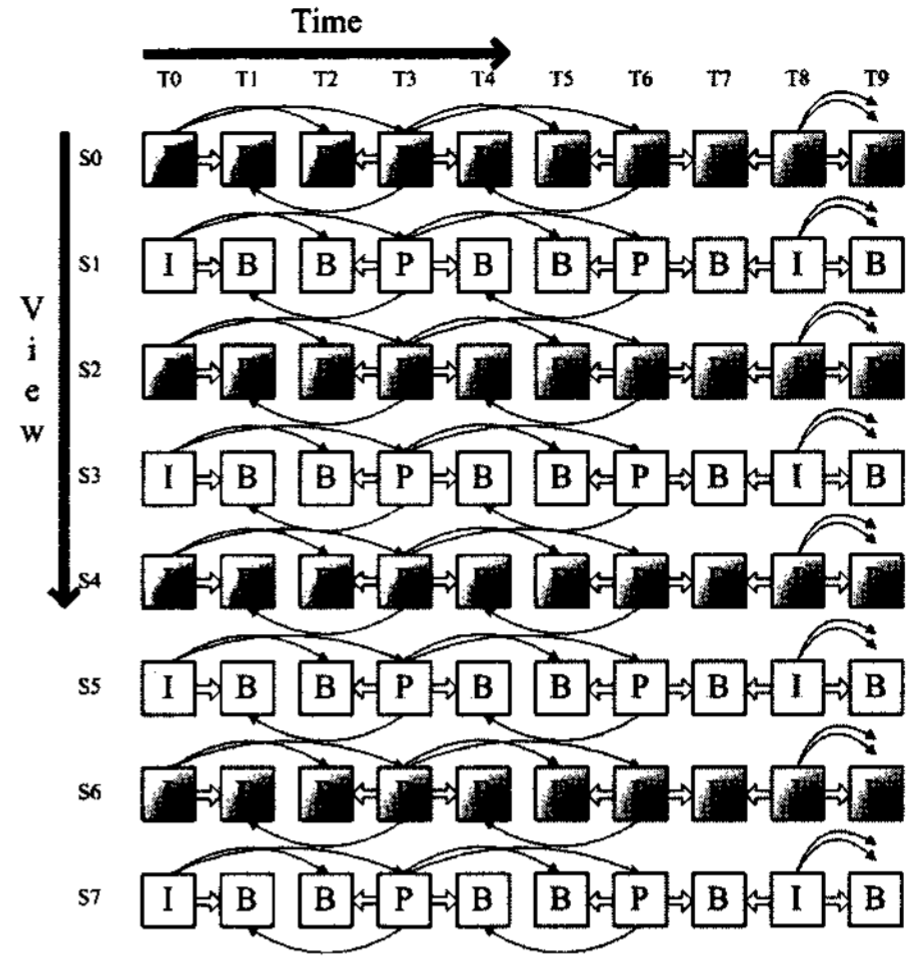


그림 1 Anchor 구조

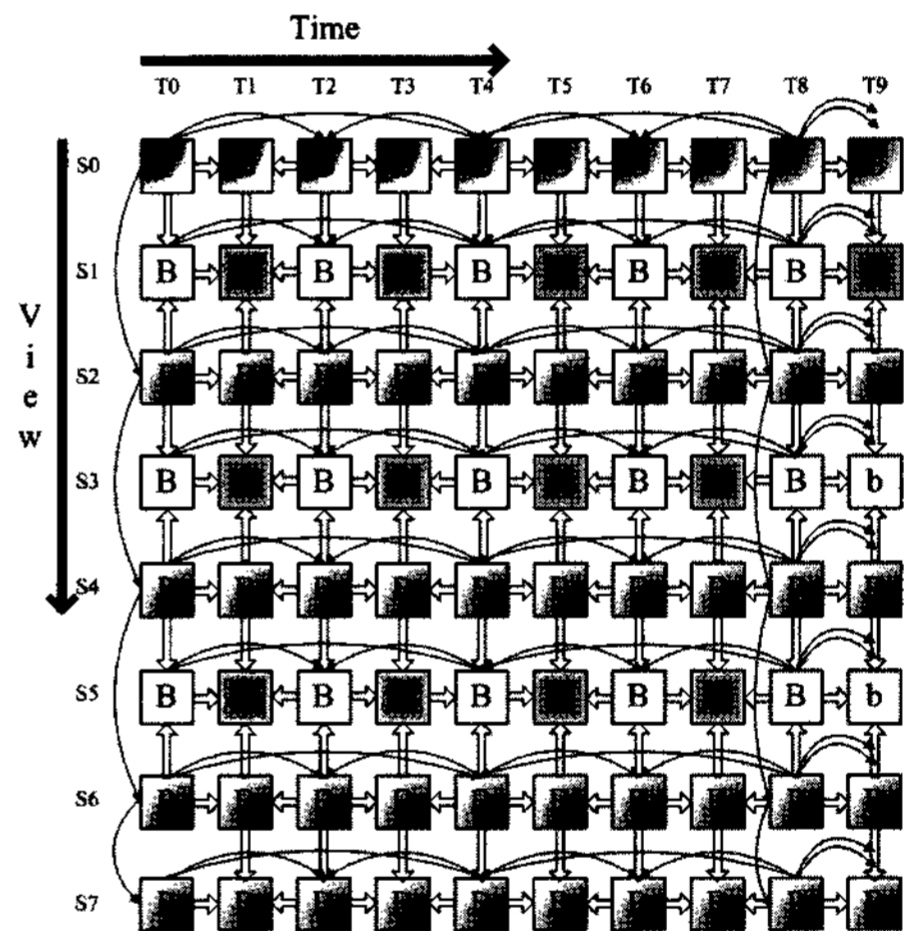


그림 2 계층적 B 프레임을 이용한 시·공간적 예측 구조

호화를 수행한다. 즉, 그림 1에서 시점 S0와 S1을 본다면, 두 시점 간의 참조는 없다. 그리고 (S0/T0)와 (S1/T0) 위치에 I 프레임을 각각 두고, 각 시점 내에서만 시간적으로 참조한다. 이 방법은 프레임 내 부호화를 수행하는 I 프레임 수가 시점의 수에 비례해서 증가하기 때문에 비트율이 상당히 증가한다.

그림 2는 참조 소프트웨어에서 사용하는 계층적 B 프레임을 이용한 시·공간적 예측 구조를 나타낸다. 이 구조는 Anchor 구조와 달리 I 프레임은 첫 번째 시점에만 두고, 시점 간의 참조를 허용한다[6]. 즉, 그림 2에서 (S0/T0) 위치에 I 프레임을 하나만 둔다. 그리고 양방향 예측 부호화를 수행하는 B 프레임인 (S1/T0) 프레임은 다른 시점에 있는 (S0/T0) 프레임과 (S2/T0) 프레임을 참조한다. 따라서 Anchor 구조에 비해 비트율이

감소하는 효과가 있다.

그리고 이와 같은 방법들 이외에도 부호화 효율을 향상시키는 새로운 GoGOP 구조들이 제안되었다.

[7]은 격자 피라미드 GOP 구조 기반의 다시점 비디오 부호화 방법에 대해 기술하였다. 이 방법은 Anchor 구조에서 고려하지 못한 인접 시점 간의 공간적인 상관도를 효과적으로 활용하기 위해 격자 GOP 구조를 제안했고, 각 시점에 대한 효율적인 부호화를 위해 계층적 피라미드 GOP 구조를 이용하였다. 이 방법은 시점 간의 참조를 허용하기는 하지만, I 프레임 수가 여전히 많기 때문에 부호화 효율을 높이는 것에 한계가 있다.

[8]은 다시점 비디오에서 좌영상과 우영상이 대응되는 시점의 상대적인 벡터로 정의되는 전역 변이(Global Disparity)를 이용하여 부호화의 기준 시점을 정하고, 카메라 간의 간격에 따라 B 프레임 수를 조절한다. 그리고 시퀀스의 특성에 따라 다시점 비디오의 부호화 단위인 다시점 GOP 예측 구조를 가변적으로 적용하는 방법을 제안하였다. 이 방법은 GoGOP 단위당 하나의 I 프레임만을 두기 때문에 예측을 위해 참조되는 연결의 수가 증가됨에 따라서 부호화 효율이 떨어지는 단점이 있다.

3. 제안하는 GoGOP 구조

3.1 Key 프레임 위치 조절

그림 3은 제안하는 GoGOP 구조의 1단계를 보여주는 것으로서, [8]에서 Key 프레임을 시점들 사이에 두는 방법을 일부 적용하였다.

본 논문에서는 전역 변이를 이용하여 Key 프레임을 두는 중간 시점을 찾아가는 것이 아니라, 단순히 Key 프레임을 시점들의 중간에 두어 부호화 복잡도를 줄이

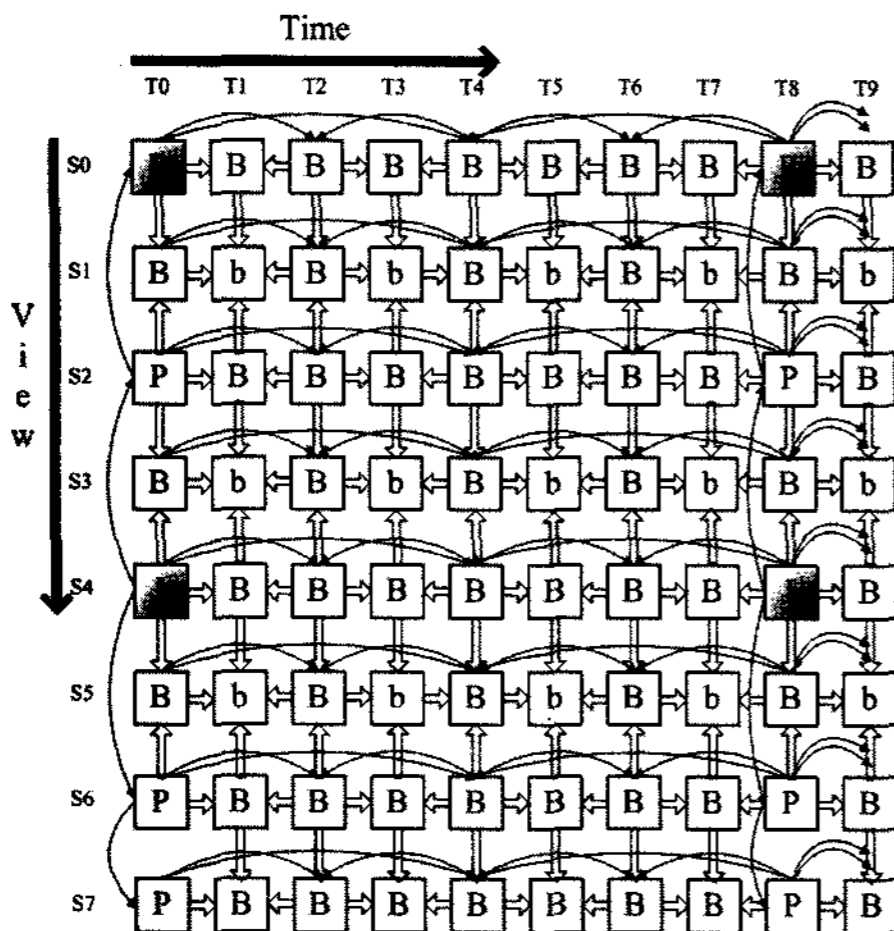


그림 3 제안하는 GoGOP 구조 - 1단계

도록 하였다. Key 프레임을 시점의 중간에 둬으로써 얻을 수 있는 효과는 Key 프레임 주위의 참조 프레임의 중복 영역을 확장시켜 부호화 효율을 증가시키는 것이다. 그림 3은 시점의 수가 8개인 경우로 단순히 4번째 또는 5번째에 Key 프레임을 둘 수 있는데, 본 구조에서는 5번째 위치인 (S4/T0)에 Key 프레임을 두었다.

3.2 동적인 I 프레임 수

그림 4는 제안하는 GoGOP 구조의 2단계를 보여주는 것으로서, I 프레임 수를 시점의 수가 증가함에 따라 동적으로 변화시키는 것이다. 따라서 I 프레임을 하나만 유지하는 것보다 PSNR(Peak Signal to Noise Ratio) 값을 상승시키는 효과가 있다. 하지만 I 프레임 수가 증가한다는 것은 곧, 부호화 효율이 떨어진다는 것을 의미하기 때문에, I 프레임 수 증가에 특정 규칙이 필요하다.

본 논문에서는 I 프레임을 기준으로 예측되는 P 프레임의 연결 수가 2개가 될 때마다 P 프레임을 I 프레임으로 전환시키는 방법을 적용하였다. 그 이유는 실험적으로, P 프레임의 예측 연결 수가 2개가 될 때 비트율의 증가폭이 크게 증가하는 것을 확인하였기 때문이다. 단, P 프레임이 연속적으로 분포하고 있을 때에는 I 프레임으로 전환시키지 않는다. 그림 4에서 (S0/T0) 위치의 P 프레임이 예측 연결 수가 2개가 되어 I 프레임으로 전환되었다. 하지만 (S7/T0) 위치의 P 프레임은 예측 연결 수가 2개이지만 P 프레임이 연속적으로 분포하기 때문에 I 프레임으로 전환되지 않는다.

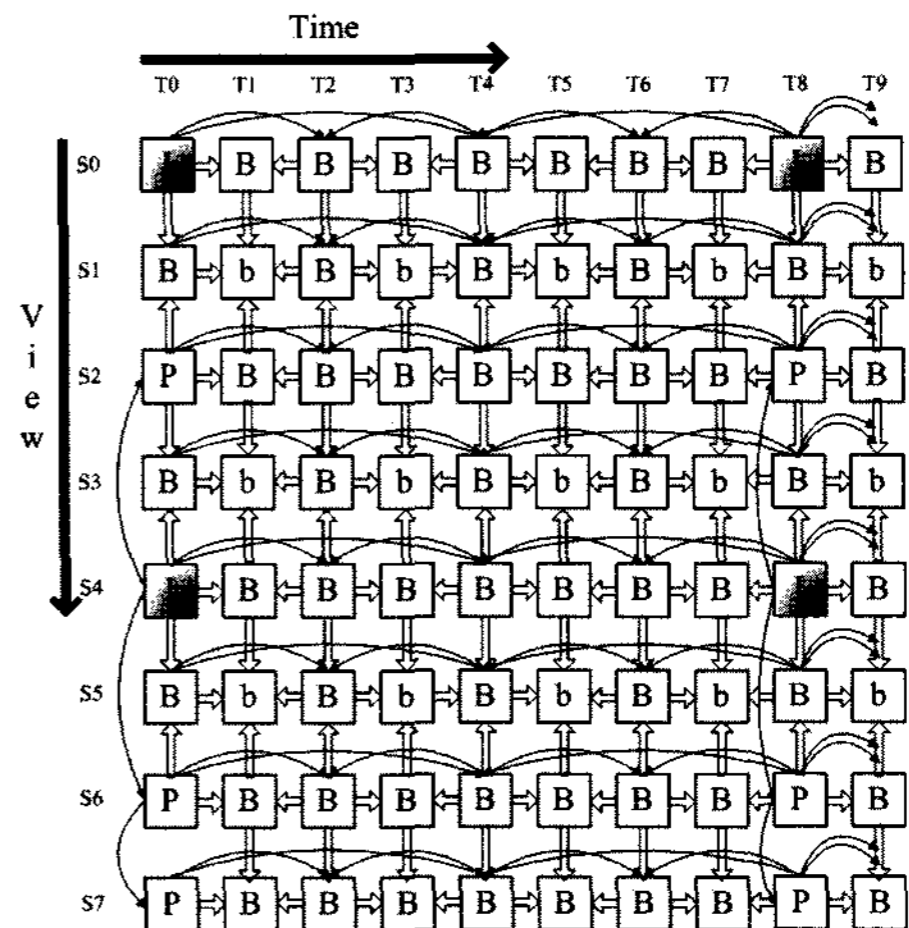


그림 4 제안하는 GoGOP 구조 - 2단계

3.3 동적인 B 프레임 수

그림 5는 제안하는 GoGOP 구조의 최종 단계를 보여주는 것으로서, 1단계와 2단계를 거친 후에 I 프레임 사이의 P 프레임을 B 프레임으로 전환시킨다. 따라서 하

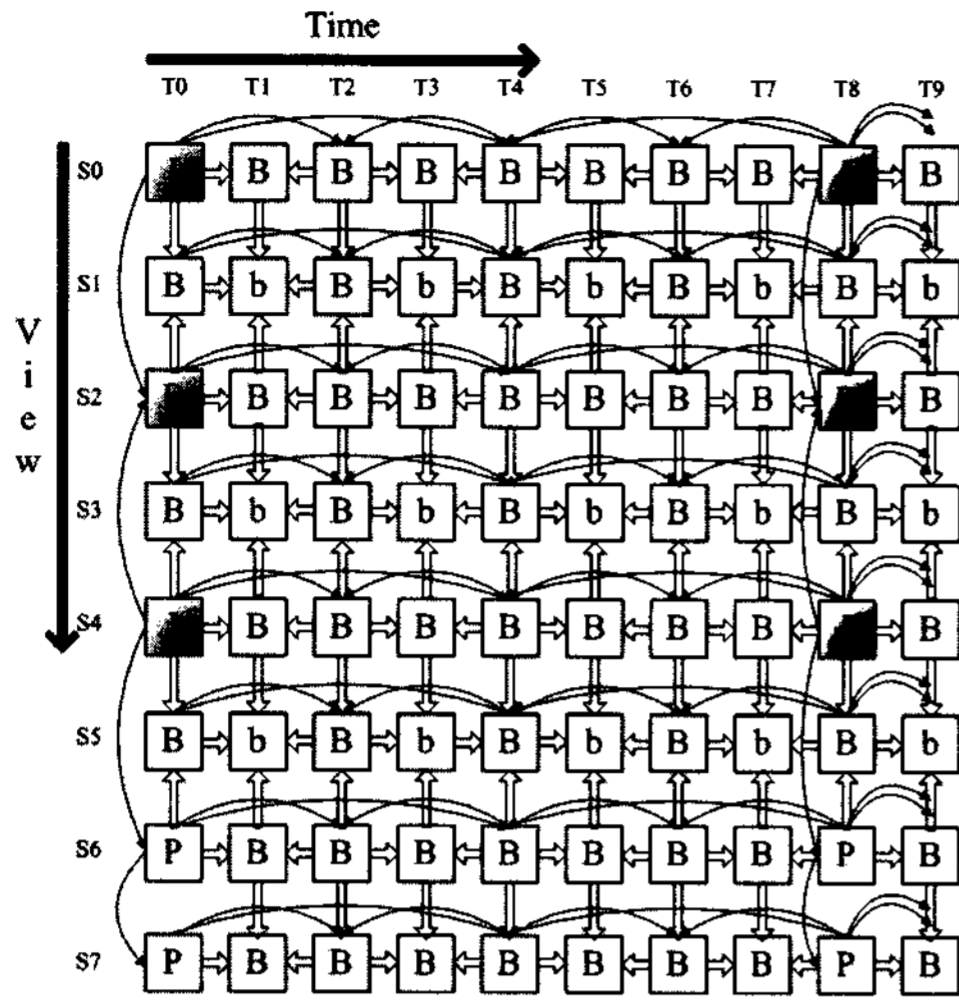


그림 5 제안하는 GoGOP 구조 - 3단계(최종 단계)

나의 I 프레임에서 예측되던 P 프레임보다 부호화 효율이 증가한다.

그림 5에서 (S0/T0) 위치의 P 프레임이 I 프레임으로 전환됨에 따라, (S2/T0) 위치의 P 프레임을 B 프레임으로 전환한다.

4. 실험 결과

제안하는 기법의 성능을 검증하기 위해 표 1의 부호화 환경 설정으로 표 2의 실험 시퀀스를 표 3의 양자화 파라미터(Quantization Parameters)에 따라 변화시키면서 Anchor 구조와 비교하였다. 양자화 파라미터는 부호화 효율에 영향을 미치는 중요한 요소로서 그 값이 낮을수록 고비트율을 나타낸다.

측정 요소는 부호화 효율을 나타내는 비트율, 비디오의 화질을 객관적으로 나타내는 PSNR을 이용하였다. 비트율의 측정 단위는 kbps이고, PSNR의 측정 단위는 dB이다. 사용된 참조 소프트웨어는 H.264 참조 소프트웨어인 JM 기반으로 만들어진 JMVM 4.0 이다[9].

표 1은 실험 시퀀스의 부호화 환경 설정을 나타낸 것이다.

표 2는 실험 시퀀스의 특성을 항목별로 나타낸 것으로서[10], 카메라의 대수(시점 수) 및 카메라 배열이 일정한 시퀀스에 한해서 선택하였다.

표 1 부호화 환경 설정

부호화 환경 설정	값
부호화 프레임수	30
참조 프레임수	2
탐색 범위	32
GOP 크기	8

표 2 실험 시퀀스 특성

실험 시퀀스	크기	프레임율	카메라 대수 (시점 수)	카메라 간격	카메라 배열
Race1	640×480	30fps	8	20cm	1D/parallel
Ballroom	640×480	25fps	8	20cm	1D/parallel
Exit	640×480	25fps	8	20cm	1D/parallel

표 3 기본 양자화 파라미터(Basic QP)

실험 시퀀스	고비트율	저비트율
Race1	24	28
Ballroom	29	34
Exit	26	31

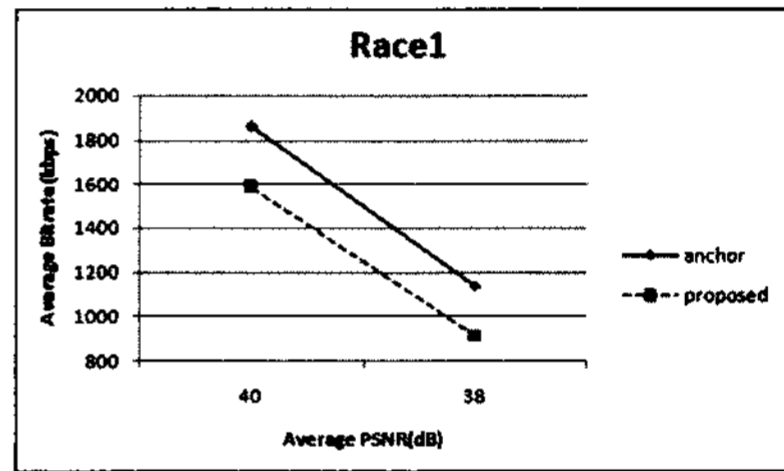


그림 6 Race1 RD(Rate-Distortion) 성능 비교

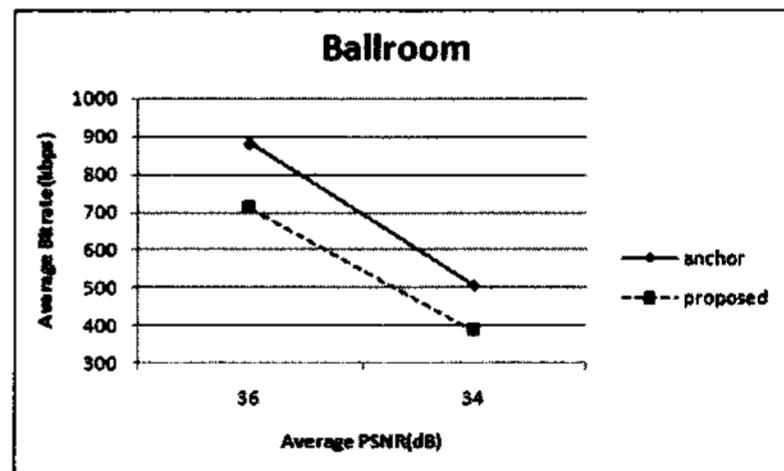


그림 7 Ballroom RD 성능 비교

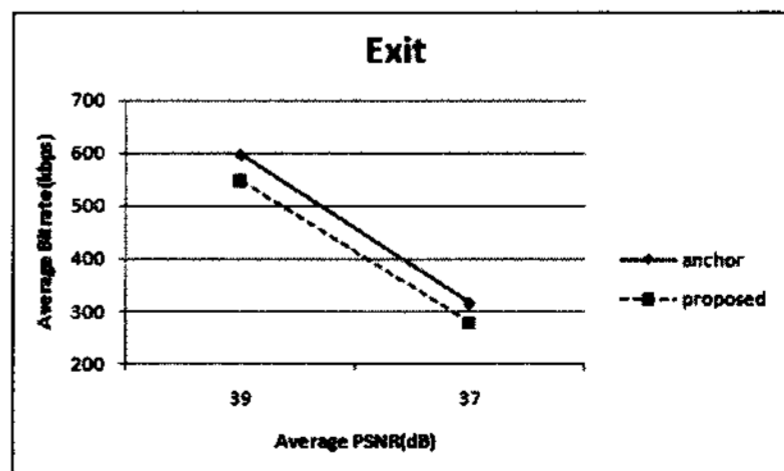


그림 8 Exit RD 성능 비교

표 3은 실험 시퀀스의 부호화를 위해 필요한 기본 양자화 파라미터의 값을 나타낸다[5].

그림 6~8은 각 실험 결과를 보여주는 그래프로서 Anchor 구조와 비교하여 유사한 PSNR 값을 가지면서

표 4 Anchor 구조와의 비트율 차이

실험 시퀀스	QP 값	비트율 차이(%)
Race1	24	-14.9
	28	-19.8
Ballroom	29	-19.6
	34	-23.7
Exit	26	-8.4
	31	-11.5

낮은 비트율을 보여준다.

표 4는 그림 6~8의 실험 결과를 정리한 것으로서, Anchor 구조와 비교하여 비트율이 상당히 감소했다는 것을 수치적으로 보여준다. Exit을 제외한 나머지 시퀀스들의 평균 비트율 감소는 약 20% 수준으로 상당히 크다.

5. 결론 및 향후 과제

본 논문은 다시점 비디오 부호화의 압축 효율을 높이기 위해 GoGOP 구조의 변화를 시도하였다. 즉, Key 프레임 위치 조절, 동적인 I 프레임 수 그리고 동적인 B 프레임 수를 이용하는 방법을 적용하여 향상된 GoGOP 구조를 제안하였다.

실험 결과, 본 논문에서 제안한 기법이 Anchor 구조와 비교하여 유사한 PSNR 값을 가지고, Exit을 제외한 나머지 시퀀스들의 평균 비트율 감소는 약 20% 수준으로 상당히 크다. 이것은 그 성능이 입증된 계층적 B 프레임을 이용한 시·공간적 예측 구조와 충분히 비교될 수 있는 결과로 볼 수 있다.

향후 과제로는 실제로 참조 소프트웨어에서 사용되고 있는 계층적 B 프레임을 이용한 시·공간적 예측 구조와 비교 실험하여 그 성능을 확인하는 것이다.

참고 문헌

- [1] Aljoscha Smolic, Karsten Mueller, Philipp Merkle, Christoph Fehn, Peter Kauff, Peter Eisert, Thomas Wiegand, "3D VIDEO AND FREE VIEWPOINT VIDEO - TECHNOLOGIES, APPLICATIONS AND MPEG STANDARDS," Proceedings of the IEEE International Conference on Multimedia and Expo, pp. 2161-2164, July 2006.
- [2] A. Puri, R.V. Kollarits, B.G. Haskell, "Basics of stereoscopic video, new compression results with MPEG-2 and a proposal for MPEG-4," Transactions of the Elsevier(Signal Processing: Image Communication), Vol. 10, No. 1, pp. 201-234, July 1997.
- [3] A.B.B. Adikari, W.A.C. Fernando, H. Kodikara Arachchi, K.K. Loo, "A H.264 COMPLIANT STEREOSCOPIC VIDEO CODEC," Proceedings of

the IEEE Electrical and Computer Engineering, pp. 1614-1617, May 2005.

- [4] Jens-Rainer Ohm, "Stereo/Multiview Video Encoding Using the MPEG Family of Standards".
- [5] 호요성, 오관정, "MPEG 다시점 비디오 부호화", TTA 저널, 제105호, pp. 132-140, 2006.
- [6] P. Merkle, K. Muller, A. Smolic, T. Wiegand, "EFFICIENT COMPRESSION OF MULTI-VIEW VIDEO EXPLOITING INTER-VIEW DEPENDENCIES BASED ON H.264/MPEG4-AVC," Proceedings of the IEEE Multimedia and Expo, pp. 1717-1720, July 2006.
- [7] 오관정, 오한, 호요성, 최병호, "격자 피라미드 GOP 구조 기반의 다시점 비디오 부호화 방법", 한국방송공학 학회 학술대회, pp. 25-28, 2005.11.
- [8] 윤재원, 서정동, 김용태, 박창섭, 손광훈, "다시점 동영상 부호화를 위한 가변형 다시점 GOP 예측 구조", 한국방송공학 학회 논문지, 제11권, 제4호, pp. 420-430, 2006.12.
- [9] ISO/IEC JTC1/SC29/WG11 N8968, "Joint Multi-view Video Model(JMVM) 4 Reference Software," San Jose, USA, Apr. 2007.
- [10] ISO/IEC JTC1/SC29/WG11 N7567, "Updated Call for proposals on Multi-view Video Coding," Nice, France, Oct. 2005.