

특집논문-04-09-3-04

3차원 비디오 압축에서의 다시점 부호화를 위한 적응적 시공간적 예측 부호화

성우철^{a)}, 이영렬^{a)†}

Adaptive Spatio-Temporal Prediction for Multi-view Coding in 3D-Video

Woo Cheol Sung^{a)} and Yung-Lyul Lee^{a)†}

요 약

본 논문에서는 차세대 실감형 3차원 영상 처리와 3차원 TV, 3차원 화상회의등과 같은 3차원 실감 미디어의 부호화에 적합한 H.264 기반의 적응적 시공간적 예측 부호화를 제안한다. 첫 번째로, 두 가지의 IPPP와 IBBP GOP(group of picture) 구조에서 기존의 동시방송(simulcast) 방식과 다르게 same-view와 inter-view 영상의 적응적 시공간적 예측 부호화를 제안한다. 두 번째로, GOP 구조가 IBBP에서 제안된 inter-view 방식의 예측 부호화시 B(bi-predictive) 화면이 MB(macroblock) 모드중 하나인 시간적 직접 방식(temporal direct mode)을 수행하려고 할 때 참조 화면이 inter-view 영상을 참조하는 경우 현재의 시간적 직접 방식을 효율적으로 수행하기 위해서 2차원 inter-view 직접 방식을 제안한다. 제안된 방식은 실험결과에서 기존의 동시방송 방식과 비교되었으며 각각의 GOP 구조 모두 화질 면에서 우수한 결과를 확인할 수 있었다.

Abstract

In this paper, an adaptive spatio-temporal predictive coding based on the H.264 is proposed for 3D immersive media encoding, such as 3D image processing, 3DTV, and 3D videoconferencing. First, we propose a spatio-temporal predictive coding using the same view and inter-view images for the two IPPP, IBBP GOP (group of picture) structures that are different from the conventional simulcast method. Second, an 2D inter-view direct mode for the efficient prediction is proposed when the proposed spatio-temporal prediction uses the IBBP structure. The 2D inter-view direct mode is applied when the temporal direct mode in B(bi-predictive) picture of the H.264 refers to an inter-view image, since the current temporal direct mode in the H.264 standard could not be applied to the inter-view image. The proposed method is compared to the conventional simulcast method in terms of PSNR (peak signal to noise ratio) for the various 3D test video sequences. The proposed method shows better PSNR results than the conventional simulcast mode.

Keywords : H.264, Spatio-temporal prediction, 3DAV, Inter-view coding, Direct mode

I. 서론

최근 3차원 디스플레이 기술의 발전에 따라 사용자들은 기존의 2차원 영상에서는 느끼지 못했던 시각적 효과를 3차

원 영상에서 느낄 수 있게 되었다. 예를 들어 여름의 바닷가를 3차원 영상으로 구성된 매체를 사용자가 접했을 경우 마치 사용자가 현실 세계의 바닷가에 있는 듯한 느낌을 줄 수 있고 3차원 영상에서 자신을 향해 다가오는 자전거를 무의식적으로 피할 만큼 현실과 비슷한 시각적 효과를 줄 수 있다. 더 나아가 3차원 영상은 사용자가 2차원 영상에서처럼 수동적인 입장에서 매체를 접하는 것이 아니라 능동적인 입장에서 매체와 상호작용을 할 수 있다는 것이 커다란 장점

a) 세종대학교 인터넷공학과

Dept. of Internet Engineering, DMS Lab., Sejong University

※ 본 과제는 산자부의 "2004 표준화기술개발사업"에 의하여 지원되었습니다.

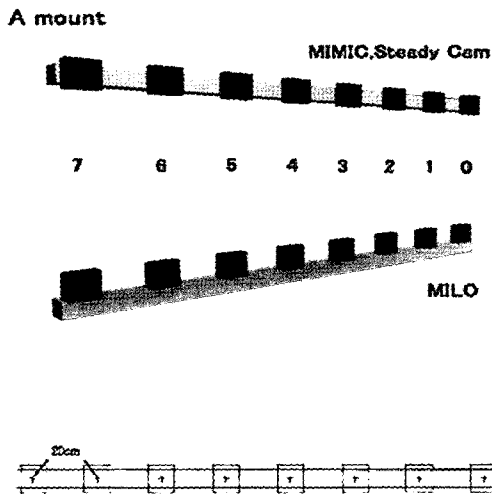


그림 1. 카메라 타입 A
Fig. 1. Camera type A

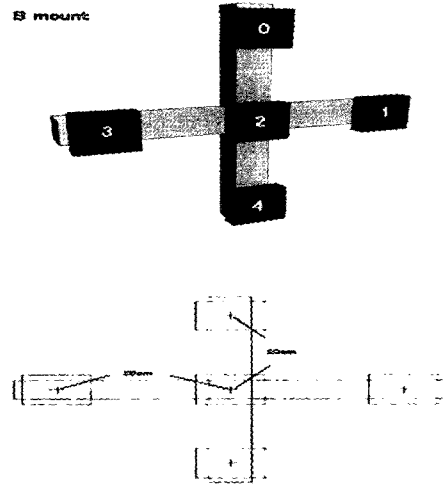


그림 2. 카메라 타입 B
Fig. 2. Camera type B

이다. 상호작용의 한 예로써, 사용자가 영상매체에서 시점을 자유롭게 선택하게 되면 선택되어진 곳이 화면의 중심이 되거나 또는 시점의 물체가 확대 되는 등의 상호작용을 예로 들 수 있다. 이것은 비디오 매체에만 국한되는 것이 아니라 오디오에도 적용될 수 있다. 예를 들어, 시끄러운 콘서트 장소에서의 3차원 오디오 매체라고 하면 처음에는 음악소리가 시끄럽게 들리다가 사용자가 한쪽 구석에서 이야기를 하고 있는 사람들로 시점을 이동하게 되면 콘서트장의 음악소리는 작아지고 시점으로 가리킨 곳의 소리가 상대적으로 커지면서 사람들의 이야기를 들을 수 있는 것이 한 예이다. 최근 2차원 비디오 압축

방법은 MPEG-4 Part 2 비디오 압축, H.264 (또는 MPEG-4 AVC(Advanced Video Coding) 표준) 압축 등 팔목할 만한 압축효율의 증가를 가져왔다. 따라서 앞으로는 3차원 비디오 부호화가 많은 주목을 끌게 될 것이며, MPEG에서는 앞으로 3차원 비디오 부호화 및 전송/수신 렌더링에 관련한 3DAV (3차원 오디오 비디오) 부호화 AHG (AdHoc Group)이 구성되어 2004년 10월에 CFP(Call For Proposal)을 할 예정이다. 또한 일본, 유럽, 미국은 이미 MPEG에서 3차원 비디오 부호화의 중요성을 이야기 하고 있으며 MPEG에 이따 3DAV 전문가 그룹이 형성되어 활발한 연구 활동을 수행하고 있다.

표 1. KDDI 다시점 테스트 영상
Table 1. KDDI multi-view test sequence

번호	영상제목	카메라 타입	총 화면 수	요사
1	Crowd	타입 B	1002장	행단보도를 교차하면서 걷고 있는 사람들을 촬영
2	Race1	타입 A	624장	경주용 자동차들이 경주하는 모습을 위치가 고정된 카메라에서 촬영한 후 카메라를 움직이면서 촬영
3	Race2	타입 A	625장	경주용 자동차에 카메라를 장착하여 촬영
4	Objects1	타입 A	624장	벽 앞쪽의 사람과 테이블을 규칙적으로 밝은 빛을 뿌려주면서 촬영. 동시에 카메라도 움직이면서 촬영.
5	Objects2	타입 A	624장	벽 앞쪽의 사람과 테이블을 규칙적으로 밝은 빛을 뿌려주면서 촬영. 카메라의 위치는 고정되어 있음.
6	Objects3	타입 A	624장	벽 앞쪽의 사람과 테이블을 규칙적으로 밝은 빛을 뿌려주면서 촬영. 동시에 카메라도 움직이면서 촬영.
7	Flamenco1	타입 A	624장	플라멩고 댄서들을 움직이는 카메라로 촬영.
8	Flamenco2	타입 B	1000장	플라멩고 댄서들을 고정된 위치의 카메라로 촬영
9	Golf1	타입 A	623장	카메라가 나무들에서 끝편쪽으로 움직이면서 촬영
10	Golf2	타입 A	624장	골퍼의 모습을 움직이는 카메라로 촬영

본 논문은 다음과 같이 구성되어 있다. II장에서는 본 논문에서 실험되고 있는 다시점 비디오 영상들과 실험 조건에 대해서 설명한다. III장에서는 첫 번째로 GOP 구조가 IPPP에 대해 제안된 적응적 시공간적 방식의 움직임 추정을 설명한다. 두 번째로 현재 H.264에서 사용되는 B 화면에 대한 시간적 직접 방식에 대해서 설명한다. 세 번째로 GOP 구조가 IBBP에 대해 제안된 적응적 시공간적 방식의 움직임 추정을 설명하고, inter-view 예측 부호화를 적용할 경우에 효율적인 시간적 직접 방식을 수행하기 위한 2차원 inter-view 직접 방식을 제안한다. IV장에서는 제안된 방식의 실험결과에 대하여 설명한다. V장에서는 결론을 맺는다.

II. 다시점 비디오 영상과 실험 조건

1. 테스트 영상

본 절에서는 3차원 영상 처리를 위해 2004년 3월 독일 뮌헨에서 열린 68회 MPEG 국제회의에서 KDDI Corp.가 제안한 다시점 비디오 영상들에 대해서 설명하려고 한다. 제안된 비디오 영상들은 총 10가지의 다양한 장면과 움직임을 가진 영상들이다. 제안된 영상들은 3차원 비디오 부호화의 표준화 활동을 증진시키기 위해 효율적으로 쓰여 질 것으로 기대된다. 표 1은 제안된 영상들을 담고 있다.

각각의 영상 크기는 320x240 이고, 영상은 프로그레시브 형태로 되어있다. 포맷은 yuv(4:2:0), 화면율은 초당 30장이 다. 카메라 타입 A는 그림 1에서 보는 것처럼 카메라 간격이 20cm인 8개의 1차원 병렬 형태로 구성된 synchronized DragonflyTM (Point Grey Research) 카메라 구조로 영상을 찍은 것이고 카메라 타입 B는 그림 2에서 보는 것처럼 십자가 구조를 가진 5개의 synchronized DragonflyTM (Point

Grey Research) 카메라로 획득한 영상들이다.

표 1의 다시점 비디오 영상들을 가지고 앞으로 연구해야 할 분야는 크게 다음과 같이 세 가지로 분류해 볼 수 있다. 첫 번째로는 3가지 이상의 다른 카메라에서 얻어진 영상을 가지고 영상의 깊이 정보를 추정하는 분야이다. 두 번째로는 다시점 비디오 영상들에 대한 효율적인 압축방법을 찾는 것이다. 기존의 동시방송 방식과 비교하여 같은 비트율에서 더 나은 화질을 구할 수 있는 방법을 연구해야 할 것이다. 세 번째로는 사용자가 임의의 시점을 가리켰을때 각 카메라의 영상을 효율적으로 보간하여 임의의 영상을 만들어 내는 것이다. 본 논문에서는 세 가지 연구의 커다란 카테고리에서 두 번째인 다시점 비디오 영상들에 대한 효율적인 압축 방법을 제안한다.

2. 실험조건

본 실험에서는 1절에서 소개한 다시점 비디오 영상들 중에서 1차원 병렬 형태로 구성된 8개의 카메라로 각각의 영상을 찍은 카메라 타입 A의 영상을 가지고 실험 하였다. 각 등급에 대한 비트율과 초기 QP값들은 표 3에 정리되어 있다. 실험을 위한 참조 코덱으로 JM 7.6이 사용되었으며 실험을 위한 변수들은 표 2에 정리되어 있다.

표 2. JM 7.6 실험조건
Table 2. Experimental condition for JM 7.6

Feature/Tool/Setting	AVC Parameters
비트율 제어	Yes
RD 최적화	Yes
루프필터, CABAC	Yes
검색 영역	±32 for CIF/VGA
참조 화면 수	5장
I 화면 주기	1초
GOP 구조	IPPP, IBBP
IDRIntraEnable	Yes

표 3. 다시점 비디오 부호화를 위한 비트율과 초기 QP값
Table 3. Bit-rates and initial QP for multi-view coding

등급	실험 영상	비트율 [kbps/camera] (initial QP)		
A(easy)	KDDI(8 카메라): flamenco1, golf2	48 kbps (40 KDDI)	64 kbps (37 KDDI)	128 kbps (32 KDDI)
B(moderate)	KDDI(8 카메라): race1, race2	64 kbps (43 race2) (42 race1)	128 kbps (39 race2) (38 race1)	256 kbps (34 race2) (34 race1)
C(difficult)	KDDI(8 카메라): object1	256 kbps (38 KDDI)	384 kbps (36 KDDI)	512 kbps (33 KDDI)

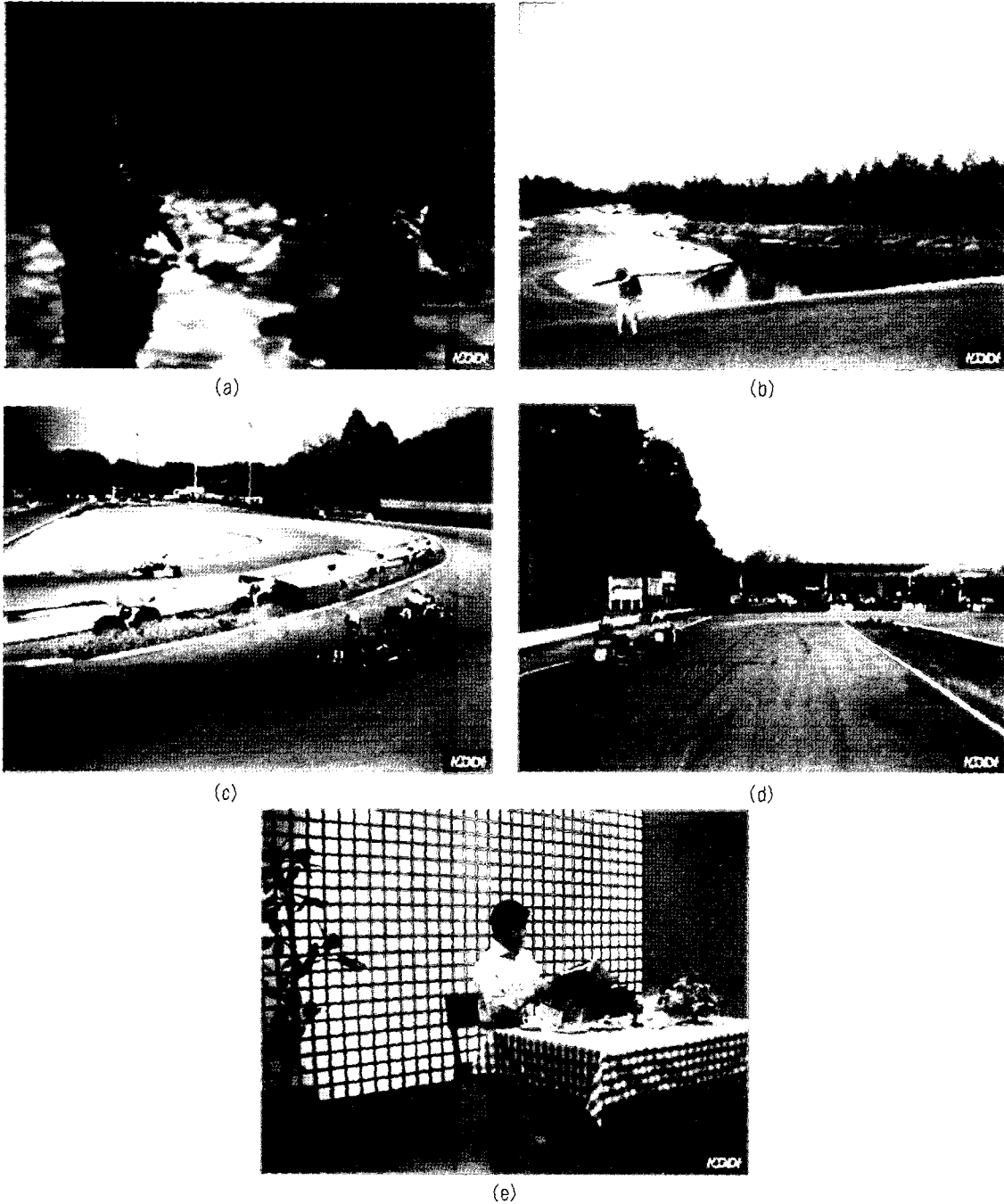


그림 3. 실험을 위한 영상: (a) 플라망고1, (b) 골프2, (c) 경주1, (d) 경주2, (e) 사물1
 Fig. 3. Experiment sequence: (a) flamenco1, (b) golf2, (c) race1, (d) race2, (e) objects1

영상은 각각의 장면과 움직임에 따라서 등급 A(easy), 등급 B(moderate), 등급 C(difficult)의 세 가지 종류로 분류하

고 실험하였다. 실험조건에는 효율적인 비트율 제어를 위해서 각각의 영상에 카메라당 비트율과 QP값을 정해서 실험

하였다. 각각의 영상은 그림 3에 나타내었다. 아래의 영상들 중 다른 영상들에 비해 특이한 점을 가지고 있는 영상인 object1은 12장 마다 한번씩 화면 전체에 굉장히 밝은 빛(Strobe lights)을 뿌려주는 특성이 있어 움직임 추정시 반드시 다수의 참조 화면을 사용해야하는 특이점이 있다.

Ⅲ. 제안된 적응적 시공간적 예측 부호화

1. IPPP GOP에서 제안된 예측 부호화

본 절에서는 GOP 구조가 IPPP에 대해서 기존의 동시방송 방식과 제안된 적응적 시공간적 예측 부호화 방식에 대

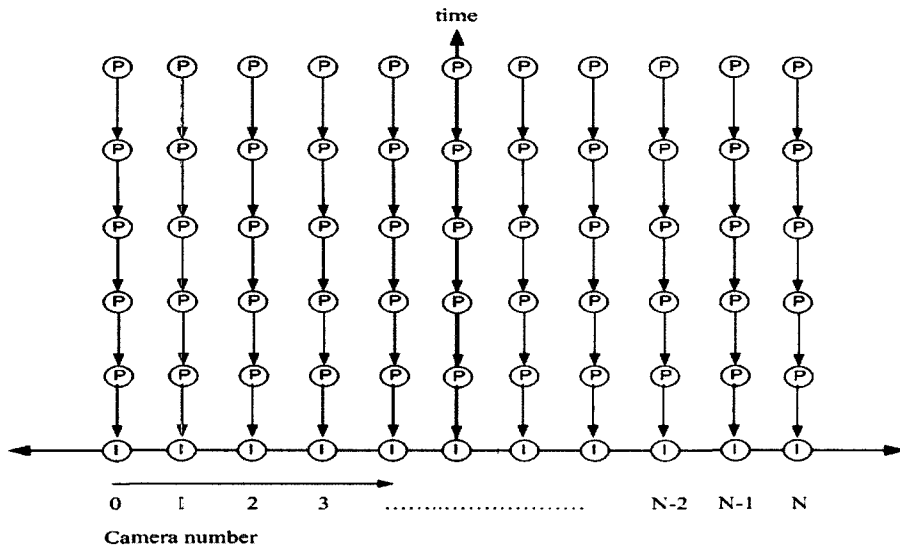


그림 4. 3차원 비디오 부호화를 위한 동시방송 방식
Fig. 4. Simulcast method for 3D-video coding

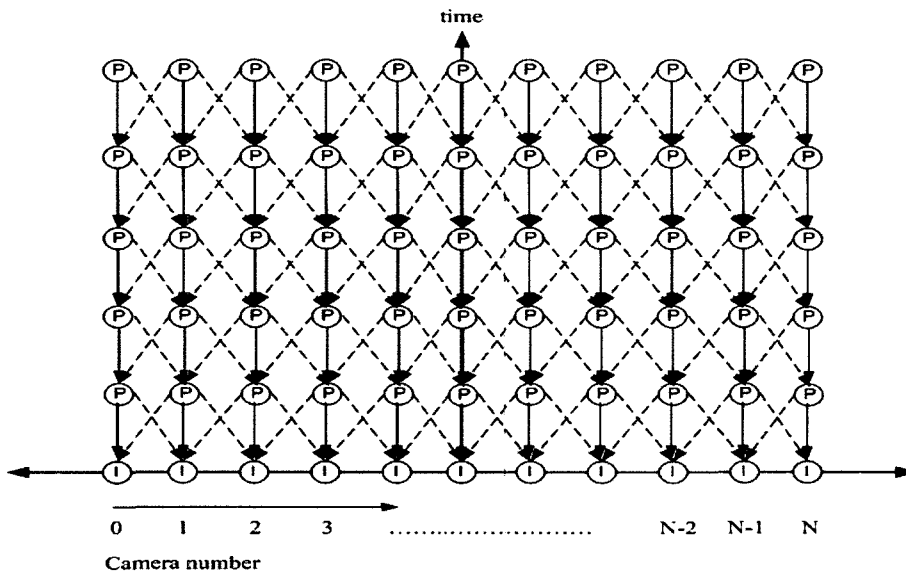


그림 5. 3차원 비디오 부호화를 위해 제안된 적응적 시공간적 예측 부호화 방식
Fig. 5. Proposed adaptive spatio-temporal method for 3D-video coding

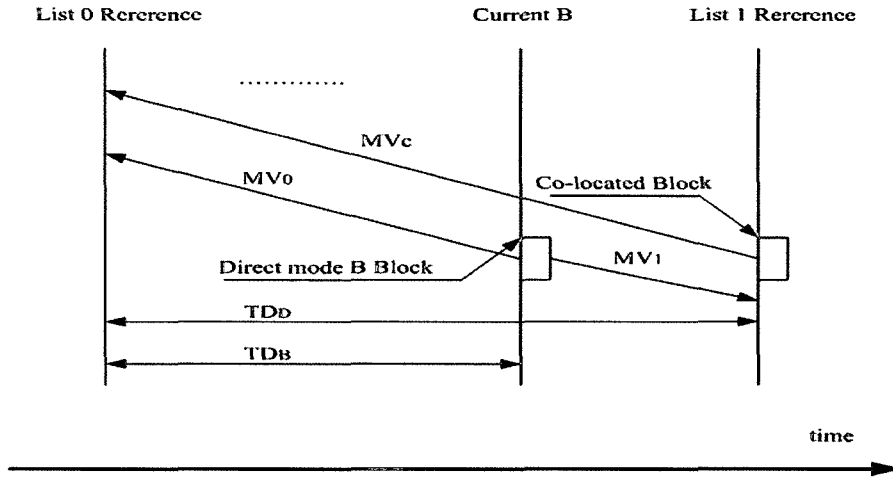


그림 6. B 화면의 시간적 직접 방식
Fig 6. Temporal direct mode in B picture

해서 설명한다. 본 실험에서 GOP 구조가 IPPP에 대한 영상에 대해서 스킵 2장을 주고 10Hz로 압축하였다. I 화면의 주기는 1초로 설정했다. 즉 I 화면은 30장마다 한번씩 부호화되는 구조이다. 기존의 동시방송 방식은 그림 4처럼 각각의 카메라에 대한 영상들이 다른 카메라의 영상들과는 독립적으로 부호화 된다. 동시방송 방식은 움직임 추정시 same-view 영상들에 대해서 기존의 시간적 예측만을 사용하고 inter-view 영상간의 공간적 예측은 사용하지 않는 구조이다. 이와는 달리 본 논문에서 제안된 적응적 시공간적 예측 부호화 방법은 그림 5처럼 각각의 카메라에 대한 영상들이 same-view의 영상들과 움직임 추정을 위해 기존의 시간적 예측을 사용하고 여기에 inter-view의 영상에 대하여 공간적 예측을 수행한다.

예를 들어 동시방송 방식이 시간적으로 앞에 있는 5장의 참조 화면을 사용해서 움직임 추정을 수행한다면, 제안된 방식은 현재 압축하고 있는 영상이 시간 축으로 t 번째이고 카메라 번호가 N이라고 할 때 카메라 번호가 N인 것을 기준으로 하여 시간 축에서 t-1, t-2, t-3 에서의 영상을 참조 화면으로 사용하고 여기에 시간 축 t-1에서 카메라 번호가 N-1, N+1의 영상에 대해서 공간적 예측을 추가로 해서 움직임 추정을 하는 구조이다. 제안된 방식은 기존의 동시방송 방식에서처럼 총 5장의 참조 화면을 갖게 된다.

2. H.264의 시간적 직접 방식

현재 JVT 표준의 B 화면은 list0(forward), list1(bac-

kward), 양방향 예측(bi-predictive), 직접(direct) 방식, 그리고 인트라(intra)방식 등 5가지의 예측방식을 가지고 있다^[3]. 이중 움직임 벡터 정보를 디코더에 전송하지 않음으로써 비트율을 감소시키는 장점이 있는 직접 방식은 슬라이스 헤더에서 전송되는 direct_spatial_mv_pred_flag 파라미터에 의해 공간적(direct_spatial_mv_pred_flag 값이 1일 경우) 또는 시간적(direct_spatial_mv_pred_flag 값이 0일 경우)으로 움직임 벡터를 유도하게 된다. 본 논문에서는 MPEG-4에서 그 성능을 확인하고 현재 사용되고 있는 시간적 직접 방식을 실험에 사용하였다. 시간적 직접 방식은 움직임 벡터가 인접한 두 화면 사이에 움직임이 일정하게 유지된다는 시간적 중복성을 이용하여 list1의 참조 화면중 참조 인덱스가 0인 화면에 있는 co-located 블록의 list0 방향의 모션벡터를 시간적 거리로 스케일링하여 list0과 list1에 대한 모션 벡터를 유도한다.

그림 6은 JVT 표준에서 사용되고 있는 시간적 직접 방식의 구조를 도식화 시켜놓은 것이다. 현재 압축될 블록의 list0의 움직임 벡터와 list1의 움직임 벡터를 구하기 위해서 list1 방향의 참조 인덱스가 0인 co-located 블록의 움직임 벡터를 스케일링하게 된다. 구체적인 식은 아래의 식 (1)에 나타내었다.

$$\begin{aligned}
 M V_0 &= \frac{T D_B}{T D_D} M V_C \\
 M V_1 &= \frac{T D_B - T D_D}{T D_D} M V_C
 \end{aligned}
 \tag{1}$$

식 (1)에서 MVc는 현재 압축할 블록의 list1 방향의 참조 인덱스가 0인 화면의 co-located 블록의 움직임 벡터를 나타내고, TDD는 그림 6의 list0 참조 화면과 list1 참조 화면간의 거리를 나타낸다. TDB는 현재 압축할 블록의 화면과 co-located 블록이 list0 방향에서 참조하고 있는 화면간의 거리를 나타낸다. MV0 과 MV1 은 각각 현재 블록의 list0과 list1방향의 움직임 벡터를 나타낸다. 시간적 직접 방식은 다른 방식과는 달리 움직임 벡터에 대한 정보를 디코더에 전송하지 않으므로 비트율을 감소시키는 장점이 있다.

다음 절에서는 첫 번째로, GOP 구조가 IBBP에 대해 제안된 적응적 시공간적 예측 부호화를 설명하고, 두 번째로 제안된 방법에서 시간적 직접 방식을 쓰는 경우 비트율 감소의 장점을 그대로 유지하면서 효율적인 예측 부호화를 할 수 있는 새로운 방식의 2차원 inter-view 직접 방식을 제안한다.

3. IBBP에서 제안된 적응적 시공간적 예측 부호화와 2차원 inter-view 직접 방식

3.1 IBBP GOP에서 제안된 예측 부호화

본 절에서는 GOP 구조가 IBBP에 대해서 제안된 적응적

시공간적 예측 부호화 방식에 대해서 설명한다. 본 실험에서 IBBP에 대한 화면들에 대해서 I 화면의 주기를 1초로 설정했다. 즉 I 화면은 30장마다 한번씩 부호화되는 구조이다. 그림 7은 IBBP에 대한 적응적 시공간적 예측 부호화 방식을 그림으로 도식화 시켜 놓은 것이다. 제안된 방식은 현재 압축하고 있는 영상이 시간 축으로 t 번째 이고 카메라 번호가 N이라고 할 때 카메라 번호가 N인 것을 기준으로 하여 시간 축에서 list0 방향의 t-1,t-2,t-3 번의 영상들과 list1 방향의 t+1번의 영상을 참조 화면으로 사용하고 여기에 시간 축 t-1 에서 카메라 번호가 N-1, N+1의 영상에 대해서 공간적 예측을 추가로 해서 움직임 추정을 하는 구조이다. 제안된 방식은 list1 방향의 참조 화면을 포함해서 총 6장의 참조 화면을 참조하게 되므로 기존의 동시방송 방식과 같은 참조 화면 수를 갖는다.

3.2 2차원 inter-view 직접 방식

그림 8에서 현재 N 번째 카메라 상에서 예측된 화면인 P1과 P2로 이미 압축된 영상들에 대해서 그 중간의 2개의 영상 B1과 B2를 시간적 직접 방식으로 예측 부호화할 경우 co-located 블록의 참조 인덱스에 따라서 세 가지 경우의 예

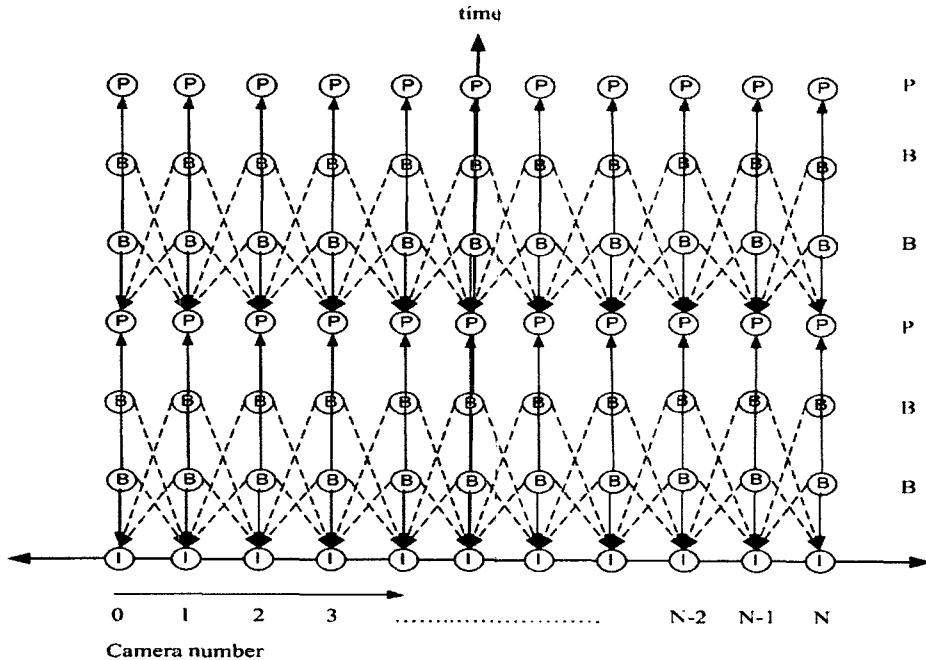


그림 7. 3차원 비디오 부호화를 위해 제안된 적응적 시공간적 예측 부호화 방식
 Fig. 7. Proposed adaptive spatio-temporal method for 3D-video coding

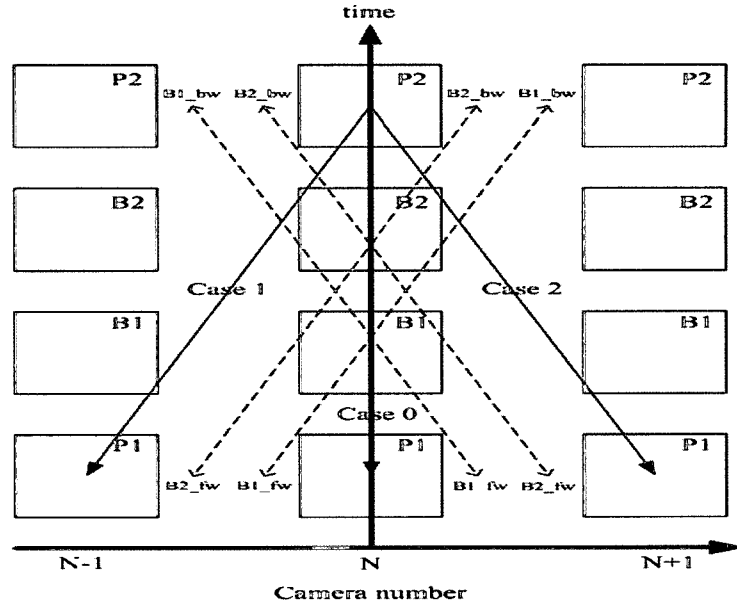


그림 8. 제안된 2차원 inter-view 직접 방식
Fig. 8. Proposed 2D inter-view direct mode

측 부호화를 수행할 수 있다. Case 0인 경우는 co-located 블록의 참조 인덱스가 같은 카메라를 가리키는 경우인데, 이 경우에는 기존의 시간적 직접 방식을 사용하여 예측 부호화를 수행한다. Case 1과 Case 2 는 각각의 co-located 블록의 참조 인덱스가 다른 카메라를 가리키는 경우인데 이때 효율적인 시간적 직접 방식을 수행하기 위해 본 논문에서는 2차원 inter-view 직접 방식을 제안한다.

예를 들어, N번째 카메라의 B2 화면의 Case 1에 대해서 살펴보자. Case 1은 B2에 대한 co-located 블록의 참조 인덱스가 N-1번째 카메라, 즉 inter-view를 가리키고 있다. 다수개의 카메라에 대한 영상을 카메라 번호와 화면 (Camera numbers, Frame)을 가지고 2차원으로 나타내는 경우 B2 화면의 list0에 대한 예측 블록은 (N-1, P1), (N, P1) 화면 사이에 위치하는 것을 유추할 수 있고, 공간적으로 (N-1, P1)에 더 가깝고 유사하다는 것을 알 수 있다. 또한 같은 방식으로 list1에 대한 예측 블록은 (N, P2), (N+1, P2) 화면 사이에 위치하는 것을 유추할 수 있고, 공간적으로 (N, P2)에 더 가깝고 유사하다는 것을 알 수 있다. 제안된 방식은 B2 화면의 현재 예측 부호화를 수행하는 블록의 list0의 예측블록을 위해 B2_fw를 구하고 list1의 예측을 위해 B2_bw를 구한다. 각 방향의 B2_fw 와 B2_bw는 아래의 식 (2)와 같다.

$$MV_{F1} = MV_C \frac{TD_B}{TD_{D(N-1,P1)}}, MV_{F2} = MV_C \frac{TD_B}{TD_{D(N,P1)}} \\ B2_fw = (3R_{F1} + R_{F2} + 2)/4 \tag{2}$$

$$MV_{B1} = MV_C \frac{TD_B - TD_D}{TD_D}_{(N,P2)}, MV_{B2} = MV_C \frac{TD_B - TD_D}{TD_D}_{(N+1,P2)} \\ B2_bw = (3R_{B1} + R_{B2} + 2)/4$$

식 (2)에서 MVF1과 MVF2는 현재 예측 부호화를 하는 블록의 co-located 블록의 움직임 벡터를 list0방향에 대해 스케일링한 움직임 벡터이다. RF1과 RF2는 각각 (N-1,P1)과 (N,P1)의 참조 화면에서 MVF1과 MVF2의 움직임 벡터가 가리키는 곳의 블록들이다. B2_fw는 RF1과 RF2로 예측된 블록들을 제안된 2차원 inter-view 직접 방식을 사용하여 공간적 유사성을 계산하여 가중치를 적용한 예측 블록이다. 2차원 inter-view 직접 방식을 위해서 사용되어진 가중치 요소는 (3.1)을 사용하였다. B2_fw와 같이 B2_bw도 2차원 inter-view 직접 방식을 사용하여 구할 수 있다. B2와 마찬가지로 B1에 대해서도 공간적 유사성을 계산하여 적용적으로 가중치를 줄 수 있으며 Case 2의 경우에도 같은 방식으로 예측 할 수 있다. 제안된 방식은 식 (2)에서의 같이 현재

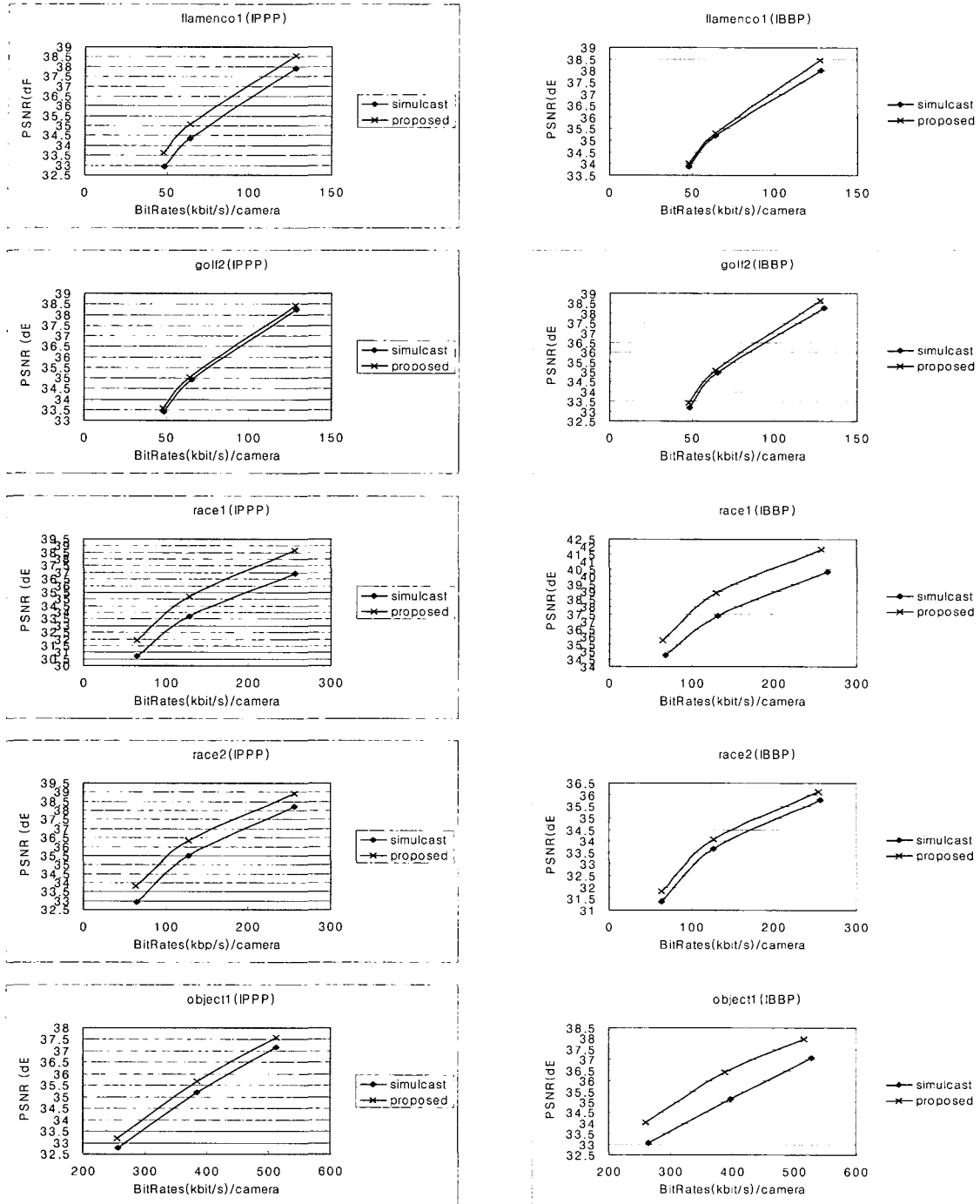


그림 9. 실험 결과
Fig. 9. Experimental result

co-located 블록의 참조 인덱스가 inter-view를 가리키는 경우 공간적인 유사성을 구해서 유사성이 더 높은 참조 화면의 블록에 대해 가중치를 주는 방식이다. 제안된 방식은 inter-view 예측 부호화에서 시간적 직접 방식을 수행 할 경우 현재 예측되는 블록을 위해서 좀더 효율적으로 예측 부호화를 수행한다.

IV. 실험 결과

본 실험은 1차원 병렬 형태로 구성된 8개의 카메라로 영상을 구성하였을 때 기존의 동시방송 방식과 제안된 방식을 동일한 비트율에서 화질을 가지고 비교하였다. 그림 9에서 왼쪽의 실험결과는 GOP 구조가 IPPP이고 각 카메라 당 100장을 실험하여 총 800장의 화면을 부호화한 실험 결과이고 오른쪽의 것이 GOP구조 IBBP에서 각 카메라 당 88장으로 실험하여 총 704장을 부호화한 실험결과이다. 좌표의 X축은 카메라 당 비트율을 나타내며 Y축은 화질(PSNR)을 나타낸다. 각각의 영상은 등급 A(easy), 등급 B(moderate), 등급 C(difficult)의 세 가지 부류로 나뉘어서 실험 되었으며 등급 C(difficult) 영상인 object 1은 영상의 특성상 5장의 전방향 참조 화면이 필요하므로 제안된 방법에서 object 1에 대한 전방향 참조 화면은 총 7장이 사용된다. 위의 실험 결과에서 알 수 있듯이 IPPP구조인 경우 동일 비트율에서 각각의 등급에 대한 영상 모두에서 기존의 동시방송 방식보다 제안한 방식이 0.2dB에서 1.5dB의 화질 향상이 있었다. IBBP구조인 경우 0.2dB에서 1.3dB의 화질 향상이 있었다.

3차원 비디오 부호화를 위해서 제안된 방식을 사용할 경우 현재 화면이 참조하고 있는 카메라에 대한 인덱스를 디코더에 보내주는 여분의 비트가 필요하다. 이 비트는 전체 비트율에 크게 영향을 주지 않을 것으로 생각 된다. 위의 실험 결과에서 제안한 방식이 기존의 동시방송 방식보다 동일 비트율에서의 화질 면에서 우수한 성능을 나타내고 있음을 확인할 수 있었다.

V. 결론

본 논문은 차세대 3차원 비디오 부호화에서 효율적인 영상 압축 방법을 제안하였다. 제안된 방식은 기존의 동시방송 방식과 같은 수의 참조 영상을 가지면서 동일 비트율에서 더 좋은 화질을 효과적으로 얻기 위해 적응적으로 카메라간의 공간적 움직임 추정과 시간적 움직임 추정 방법을 이용하였다. 또한 제안된 방법에서 B 화면의 움직임 추정을 위한 시간적 직접 방식을 위해 2차원 inter-view 직접 방식을 제안하여 효율적으로 화질을 향상하였다. 앞으로는 3DAV가 관련 산업 분야에서 많은 주목을 받을 것이며, 수요 또한 늘어날 것으로 예상된다. 이에 따라 새로운 3차원 영상 처리와 응용들을 연구하고 개발해야 할 것이다.

참고 문헌

- [1] F. Isgro, E. Trucco, P. Kauff, and O. Schreer, "Three-dimensional image processing in the future of immersive media," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 3, pp. 288- 303, Mar. 2004.
- [2] A. Smolic and D. McCutchen, "3DAV exploration of video-based rendering technology in MPEG," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 3, pp. 348-356, Mar. 2004.
- [3] O. Grau, T. Pullen, and G. A. Thomas, "A Combined Studio Production System for 3-D Capturing of Live Action and Immersive Actor Feedback," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 14, no. 3, pp. 370- 380, Mar.2004.
- [4] "Study of Final Committee Draft(SFCD) of Joint Video Specification(ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)", Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Mar. 2003
- [5] M. Ziegler et al., "Evolution of stereoscope and three-dimensional video," *Signal Processing: Image Commun.*, vol. 14, pp. 173-194, 1988
- [6] <http://bs.hhi.de/~suehring/tml/download/jm76.zip>
- [7] A. Smolic, Y. Guo, J. Guether, and T. Selinger, "Demonstration of Streaming of MPEG-4 3-D Scenes with Live Video," *ISO/IEC/JTC1/SC29/WG 11, Pattaya, Thailand, Doc. M7811.2001*

 저 자 소 개



성 우 철

- 2002년 2월 : 세종대학교 전산학과 졸업
- 2002년 3월~현재 : 세종대학교 인터넷공학과 석사과정
- 주관심분야 : 영상처리, 영상압축, 3차원 비디오 부호화



이 영 렬

- 1985년 2월 : 서강대학교 전자공학과 졸업
- 1987년 2월 : 서강대학교 전자공학과 석사
- 1999년 2월 : 한국 과학 기술원 전기 및 전자공학과 박사
- 2001년 8월 : 삼성전자 중앙연구소 DMS Lab. 수석연구원
- 2001년 9월~현재 : 세종대학교 컴퓨터공학부 인터넷공학과 부교수
- 주관심분야 : 영상처리(압축, 복원), 영상전송, 멀티미디어 시스템, 3차원 비디오 압축복원, 트랜스코딩, 스케일러블 비디오 압축복원