

특집논문-05-10-3-06

# 조명변화에 적응적인 움직임 검색 기법과 2차원 다이렉트 모드를 사용한 다시점 비디오 부호화

이 용 기<sup>a)</sup>, 허 재 호<sup>a)</sup>, 이 영 렬<sup>a)†</sup>

## Multi-View Video Coding Using Illumination Change-Adaptive Motion Estimation and 2D Direct Mode

Yung Ki Lee<sup>a)</sup>, Jae Ho Hur<sup>a)</sup>, and Yung Lyul Lee<sup>a)†</sup>

### 요 약

조명 변화에 적응적인 움직임 탐색 및 보상 방법과 2차원 다이렉트 모드 (2D Direct Mode)를 이용한 MVC (Multi-view Video Coding) 방법을 제안한다. 먼저, 공간적 및 시간적으로 인접한 영상으로부터 휘도 화소 값의 움직임 탐색 및 보상을 수행하기 위한 새로운 SAD (Sum of Absolute Difference) 측정 방법을 제안한다. 조명 변화에 적응적인 움직임 탐색 및 보상은 움직임 벡터의 정확도를 높이고, 비트의 절감을 위하여 새로운 SAD 계산식을 사용한다. 다음으로, 시점 간의 예측 시에 사용될 수 있는 2차원 다이렉트 모드는 MPEG-4 AVC의 시간적 다이렉트 모드 (Temporal Direct Mode)로부터 확장된 방식이다. 본 논문에서 제안된 MVC 방법은 MPEG-4 AVC의 Simulcast 부호화와 비교하여 약 0.8dB의 PSNR (Peak Signal-to-Noise Ratio) 향상을 보였다.

### Abstract

A MVC (Multi-view Video Coding) method, which uses both an illumination change-adaptive ME (Motion Estimation)/MC (Motion Compensation) and a 2D (Dimensional) direct mode, is proposed. Firstly, a new SAD (Sum of Absolute Difference) measure for ME/MC is proposed to compensate the Luma pixel value changes for spatio-temporal motion vector prediction. Illumination change-adaptive (ICA) ME/MC uses the new SAD to improve both MV (Motion Vector) accuracy and bit saving. Secondly, The proposed 2D direct mode that can be used in inter-view prediction is an extended version of the temporal direct mode in MPEG-4 AVC. The proposed MVC method obtains approximately 0.8dB PSNR (Peak Signal-to-Noise Ratio) increment compared with the MPEG-4 AVC simulcast coding.

**Keywords :** Multi-view video coding, illumination, 2D direct mode, MPEG-4 AVC, PSNR

### I. 서 론

a) 세종대학교 인터넷공학과  
Dept. of Internet Engineering, DMS Lab., Sejong University

본 논문은 3차원 TV와 Free View-point video에서 사용할 수 있는 적응적인 공간적 및 시간적 예측 부호화와<sup>[1]</sup>

MPEG-4 AVC를<sup>[2]</sup> 기반으로 하는 MVC를 위한 조명 변화에 적응적인 움직임 탐색 및 보상 방법을 제안한다. 현재 MPEG-4 AVC 표준의<sup>[2]</sup> 시간적 다이렉트 모드는 하나의 시점에 대한 시간적인 예측 방법을 사용하므로, MVC의 수행 시 다른 시점 간의 공간적 예측을 위해서는 사용될 수 없다. 그러나 다른 시점 간의 2차원 다이렉트 모드는 현재 B 프레임의 매크로블록에 대하여 다음 P 프레임의 동일한 위치에 있는 매크로블록이 다른 시점의 (공간적 예측) 프레임을 참조할 때에도 움직임 벡터를 계산할 수 있으므로, IBBP 구조에서 기존의 방법보다 효율적인 공간적 및 시간적 예측 부호 방법이다. 조명 변화에 적응적인 움직임 탐색 및 보상 방법은, 현재 프레임의 블록으로부터 화소들의 평균을 구하여, 각 화소 값과 평균값의 차분을 구하고, 마찬가지로 참조 프레임의 블록 또한 같은 방식으로 차분을 구한 후, 중간 과정의 블록을 생성하고 그 블록을 기반으로 움직임 탐색과 보상을 수행하는 새로운 프레임간 예측 방법이다.

제안된 방법은 여러 다 시점 테스트 영상들에 대한 객관적인 평가를 MPEG-4 AVC<sup>[2]</sup> Simulcast 부호화 방법과 비교하였는데 같은 화질 (PSNR)에서 필요로 하는 비트는 제안된 방법이 MPEG-4 AVC<sup>[2]</sup> Simulcast 부호화 방법을 사용했을 때 보다 최대 43%를 절약할 수 있었다.

본 논문은 다음과 같이 구성되어 있다. 제 2장에서, MVC를 수행할 때 참조 프레임의 배열 순서를 정의하는 방식에 대하여 소개하고, 제 3장과 4장에서는, 제안된 방법의 세부적인 알고리즘을 서술한다. 마지막으로 제 5장과 6장에서 각각 기존의 MPEG-4 AVC 표준과 본 논문에서 제안하는 방법과의 비교 실험 결과와 결론으로 본 논문의 끝을 맺는다.

## II. 참조 색인 순서 (Reference Index Order)

본 논문에서 사용된 입력 영상은 카메라의 배열 형태에

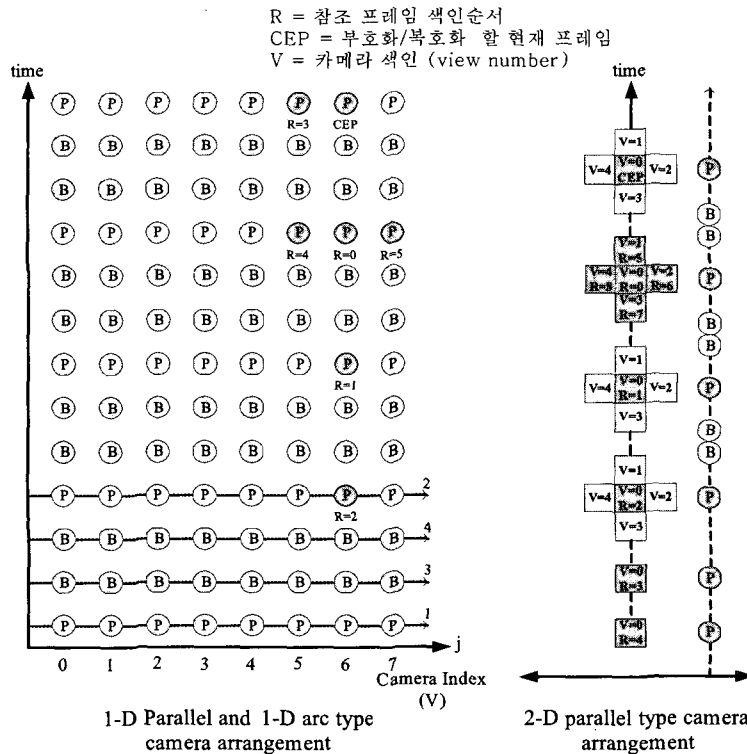


그림 1. 카메라 배열에 따른 참조 프레임 순서 (a) 1차원 평행배열영상 및 1차원 호배열영상 (b) 2차원 평행배열영상

Fig. 1. Reference picture ordering of (a) 1-D parallel and 1-D arc camera arrangement and (b) 2-D parallel camera arrangement.

따라 1차원 평행배열영상, 2차원 평행배열영상, 1차원 호배열영상으로 3가지 형태가 있다<sup>[3]</sup>. 각 세 가지 영상 형태의 참조 프레임 순서는 그림 1의 (a), (b)와 같이 정의 되어 있다. 그림 1 (a)에서처럼 1차원 평행배열영상과 1차원 호배열영상의 각 시점 간의 부호화/복호화 순서는 화살표 방향과 같다. 그림 1에서 R은 참조 프레임 색인 순서이고, CEP는 부호화/복호화를 할 현재 프레임이다. 그리고 그림 1 (b)에서 V는 카메라 색인을 나타낸다.

### III. 2차원 다이렉트 모드를 이용한 공간적 및 시간적 예측 부호화

MVC를 위하여 공간적 및 시간적인 예측이 수행되는데, 공간적 예측 부호화를 적용할 경우, 이미 부호화 된 다른 시점의 참조가 가능한 P 프레임을 공간적 예측을 위한 프레임으로 선택한다.

#### 1. MPEG-4 AVC의 시간적 다이렉트 모드

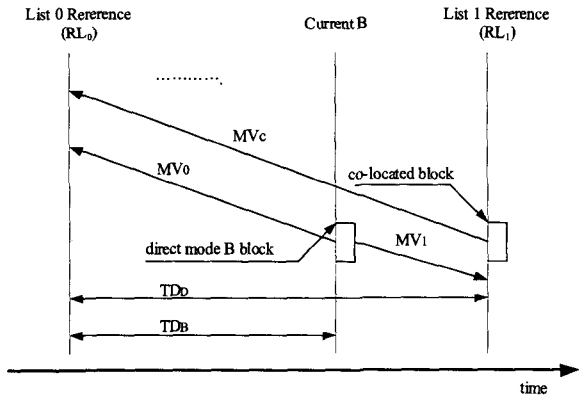


그림 2. MPEG-4 AVC에서 다이렉트 모드 : 레퍼런스 프레임  $RL_0, RL_1$ 을 가리키는 전 방향 움직임 벡터  $MV_0$ 와 후 방향 움직임 벡터  $MV_1$   
 Fig. 2. A direct mode block has the forward and backward motion vectors  $MV_0$  and  $MV_1$  pointing to two reference pictures  $RL_0$  and  $RL_1$  in MPEG-4 AVC.

시간적 다이렉트 모드 (Temporal Direct Mode)는 현재 B 프레임의 특정 블록을 부호화하기 위하여 이후의 프레임  $RL_1$  (List 1 Reference) 내의 동일 위치에 있는 블록의 움직

임 벡터  $MV_C$ 로부터 현재 블록의 전 방향과 후 방향 움직임 벡터 쌍 ( $MV_0, MV_1$ )을 유도한다. 예측 신호는 전 방향과 후방향의 참조 프레임 ( $RL_0, RL_1$ )로부터 움직임 벡터가 가리키는 위치의 두 개의 블록을 선형 조합(Linear Combination)하여 생성한다. 시간적 다이렉트 모드의 방법은 그림 2에서 나타내고, 다이렉트 모드 블록의 전 방향과 후 방향 움직임 벡터의 유도 과정은 식 (1)과 같다.

$$\begin{aligned}
 MV_0 &= \frac{TD_B}{TD_D} MV_C \\
 MV_1 &= \frac{TD_B - TD_D}{TD_D} MV_C
 \end{aligned}
 \tag{1}$$

여기서  $MV_C$ 는  $RL_1$  내의 동일 위치에 있는 블록의 움직임 벡터이다.  $MV_0$ 는 전 방향 움직임 벡터이고,  $MV_1$ 은 후 방향 움직임 벡터를 의미한다<sup>[4]</sup>.

#### 2. 제안된 2차원 다이렉트 모드

일반적으로 MVC에서는 부호화 효율을 높이기 위해서 공간적 및 시간적 예측을 수행한다. 한 시점의 부호화를 위한 시간적 다이렉트 모드는 여러 개의 시점을 가지는 MVC 환경에서 다른 시점 간의 예측을 필요로 하는 경우에 적용될 수 없기 때문에 MVC 환경에 적합한 새로운 다이렉트 모드가 필요하다. 따라서 다른 시점 간의 공간적 예측을 위한 2차원 다이렉트 모드를 제안한다. 그림 3은 2차원 다이렉트 모드의 방법을 나타낸다. 가로축의  $i$ 는 카메라 색인을 의미하고, 세로축은 시간의 경과를 의미한다. 그림 3에서  $P(i,t)$ 는 카메라 색인  $i$ 와 시간  $t$ 에 존재하는 P 프레임을 나타내고,  $B(i,t-1)$ 은 카메라 색인  $i$ 와 시간  $t-1$ 에 존재하는 B 프레임을 의미한다.

$B(i,t-1)$  프레임의 현재 블록( $16 \times 16$  또는  $8 \times 8$  블록)이 2차원 다이렉트 모드인 경우, 현재 프레임 이후의  $P(i,t)$  프레임 내의 동일 위치에 있는 블록이  $P(i-1,t-3)$  프레임을 참조하는 움직임 벡터를 사용하여 전 방향과 후 방향의 움직임 벡터를 유도한다. 위의 과정으로 얻어진 움직임 벡터가 가리키는 예측 신호는 전방향일 경우  $P(i-1,t-3)$ 와  $P(i,t-3)$

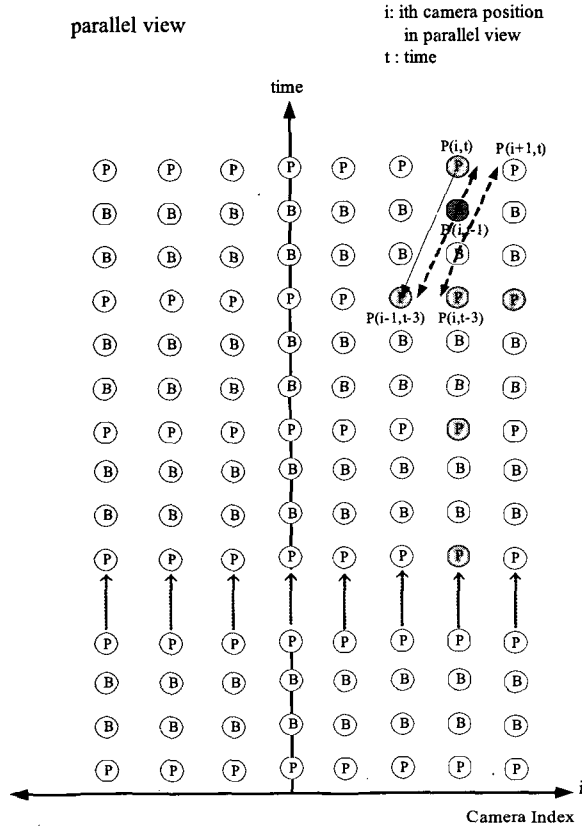


그림 3. 2차원 다이렉트 모드의 예  
Fig. 3. An example of 2D direct mode.

프레임, 후방향일 경우  $P(i,t)$ 와  $P(i+1,t)$  프레임 내의 참조 블록을 이용하여 보간하는 방법으로 얻을 수 있다.

이 모드를 사용하는 블록의 두 개의 전 방향 움직임 벡터 ( $MV_{F1}$ ,  $MV_{F2}$ )는 현재 프레임 이후의  $P(i,j)$  프레임 내의 동일 위치에 있는 블록이 가리키는 움직임 벡터를 이용하여 유도된다. 전 방향 예측 블록  $R_F$ 는 두 개의 예측 블록  $R_{F1}$ 과  $R_{F2}$ 를 이용하여 생성될 수 있다. 위의 방법은 식 (2)와 같다.

$$\begin{aligned} MV_{F1} &= MV_C \frac{TD_B}{TD_D} \Big|_{to P(i-1,t-3)} \\ MV_{F2} &= MV_C \frac{TD_B}{TD_D} \Big|_{to P(i,t-3)} \\ R_F &= (3R_{F1} + R_{F2} + 2) / 4 \end{aligned} \quad (2)$$

$MV_{F1}$ 과  $MV_{F2}$ 는 각각  $P(i-1,t-3)$ 와  $P(i,t-3)$  프레임 내의 참조 블록  $R_{F1}$ 과  $R_{F2}$ 를 가리킨다.

두 개의 후 방향 움직임 벡터 ( $MV_{B1}$ ,  $MV_{B2}$ )와 예측신호 또한 위와 유사한 방식으로 유도될 수 있으며, 식 (3)과 같다.

$$\begin{aligned} MV_{B1} &= MV_C \frac{TD_B - TD_D}{TD_D} \Big|_{to P(i,t)} \\ MV_{B2} &= MV_C \frac{TD_B - TD_D}{TD_D} \Big|_{to P(i+1,t)} \\ R_B &= (3R_{B1} + R_{B2} + 2) / 4 \end{aligned} \quad (3)$$

$MV_{B1}$ 과  $MV_{B2}$ 는 각각  $P(i,t)$ 와  $P(i+1,t)$  프레임 내의 참조 블록  $R_{B1}$ 과  $R_{B2}$ 를 가리킨다.

실험 결과에서 예측을 위한 기존의 MPEG-4 AVC<sup>[2]</sup> 참조 프레임은 시간적 예측에만 사용되는데, 2차원 다이렉트 모드를 사용하면 차분값과 2차원 다이렉트 모드라는 것만 부호화되기 때문에 부호화 효율이 향상되는 것을 볼 수 있다.

$B(i,t-1)$  프레임의 현재 블록(16X16 또는 8X8 블록)이 2차원 다이렉트 모드인 경우, 현재 프레임 이후의  $P(i,t)$  프레임 내의 동일 위치에 있는 블록이  $P(i-2,t-3)$  프레임을 참조하는 경우에도 식(2)와 (3)의 가중치를 바꿔 위와 유사한 방법으로 유도 해낼 수 있다.

#### IV. 조명 변화에 적응적인 움직임 탐색 및 보상

MVC에서는 일반적으로 어느 한 시점과 다른 시점 간의 조명 변화가 존재한다. 따라서 본 논문은 매크로블록 기반의 조명 변화에 적응적인 움직임 탐색 및 보상을 제안한다. 현재 프레임을  $f(i,j)$ 로, 참조 프레임을  $r(i,j)$ 로 정의하고,  $i$ 와  $j$ 는 각각 공간 좌표의 가로축 과 세로축에 대응된다. 가변 크기의 SXT (16X16, 16X8, 8X16, 8X8, 8X4, 4X8, 4X4) 블록의 움직임 탐색을 위한 기존의 SAD 계산은 식 (4)와 같다.

$$SAD(x, y) = \sum_{i=0}^{S-1} \sum_{j=0}^{T-1} |f(i, j) - r(i+x, j+y)| \quad (4)$$

(x, y)는 움직임 벡터를 나타낸다. 조명 변화를 보상하기 위해서 식 (5)와 같이 새로운 SAD 계산식을 정의한다.

$$NewSAD(x, y) = \sum_{i=0}^{U-1} \sum_{j=0}^{V-1} |(f(i, j) - M_{curr}) - (r(i+x, j+y) - M_{ref})| \quad (5)$$

$M_{curr}$ 과  $M_{ref}$ 는 각각 현재 블록과 참조 블록의 화소들에 대한 평균값이다. 실험을 통하여  $U$ 와  $V$ 를 16 또는 8로 결정하였다.  $UXV$  블록의 차분 신호들은 MPEG-4 AVC 표준안에<sup>[2][5]</sup> 정의된 정수  $4 \times 4$  DCT (Discrete Cosine Transform)와 양자화를 통하여 부호화 된다. 이 방법을 위하여, 우선적으로 MB\_type이 결정되며, 제안된 방법을 사용하는 경우와 기존의 방법을 사용하는 경우를 표시하기 위하여 SXT 블록에 대하여 최소 1 비트에서 최대 4 비트까지의 플래그 비트가 추가된다. 또한, 현재 블록이 제안된 방식을 사용할 경우 8 비트의  $M_{curr}$ 이 추가된다.

### V. 실험 결과

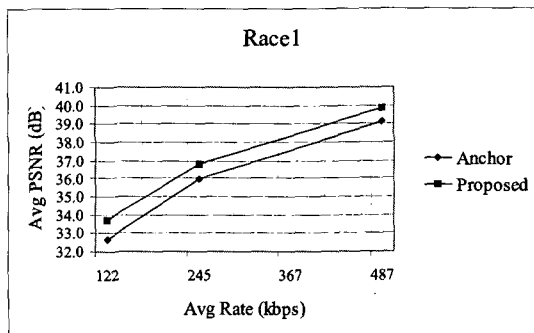
실험은 시간적 및 공간적 참조 프레임과 매크로블록 기반의 조명 변화에 적응적인 움직임 탐색 및 보상기법을 카메라 간격이 좁은 영상에 대해서만 수행하였다. 그리고 약간의

화질열화현상이 있지만, 헥사곤 탐색 기법 (Unsymmetrical-cross Multi-Hexagon-grid Search)<sup>[6]</sup> 이라 불리는 JM7.6에<sup>[7]</sup> 제안된 빠른 움직임 탐색기법을 제안하는 방법과 함께 사용하였다. MVC 성능 평가를 위해 MPEG-4 AVC Simulcast 방법으로 부호화된 비트스트림과 제안된 방법을 비교하였다. 실험에 쓰인 영상은 모두 SIF (320X240) 이며, IBBPBBP...구조로 부호화 하였고, 이외에 코딩 조건은 참고문헌의<sup>[3]</sup> Annex A에 자세히 기술되어 있다.

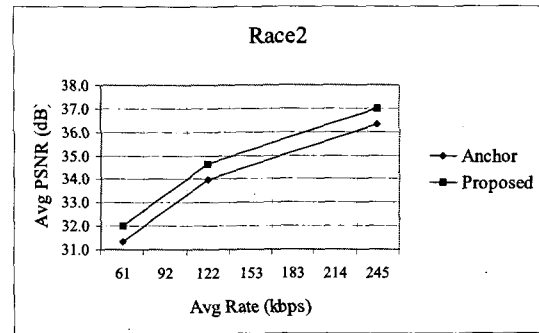
모두 10개의 참조 프레임 중에 2차원 평행배열영상에서는 5개의 참조 프레임이 공간적 참조 프레임으로 사용되었고, 1차원 평행배열영상과 1차원 호배열영상에서는 6개의 참조 프레임 가운데 3개의 프레임이 공간적 예측 프레임으로 사용되었다. 더 많은 공간적 예측 영상을 사용할 수 있지만 이 실험에서는 그 수를 제한하였고, 참조 영상의 수가 위와 같이 줄어도 PSNR의 손실은 없었다.

제안된 2차원 다이렉트 모드는 공간적 예측 프레임에만 적용되어 사용되었고, 매크로블록 기반의 조명 변화에 적응적인 움직임 탐색 및 보상기법은 모든 참조 프레임에 대해 사용되었다.

그림 4는 "Race1", "Race2", "Flamenco1", "Golf2", "Flamenco2", "Crowd", "Aquarium" 영상에 대한 울-왜곡 곡선을 나타낸다. "Aquarium"은 10Hz의 프레임율을 가지며 그 외의 영상은 30Hz의 프레임율을 갖는다. 제안한 방법을 통하여 그림 4(a)-(g)에서 보이는 바와 같이 MPEG-4 AVC Simulcast 방법의 PSNR과 비교하여 각각 0.7dB, 0.9dB, 0.8dB, 2.2dB, 0.8dB, 0.2dB, 0.4dB의 향상을 보였다.



(a)



(b)

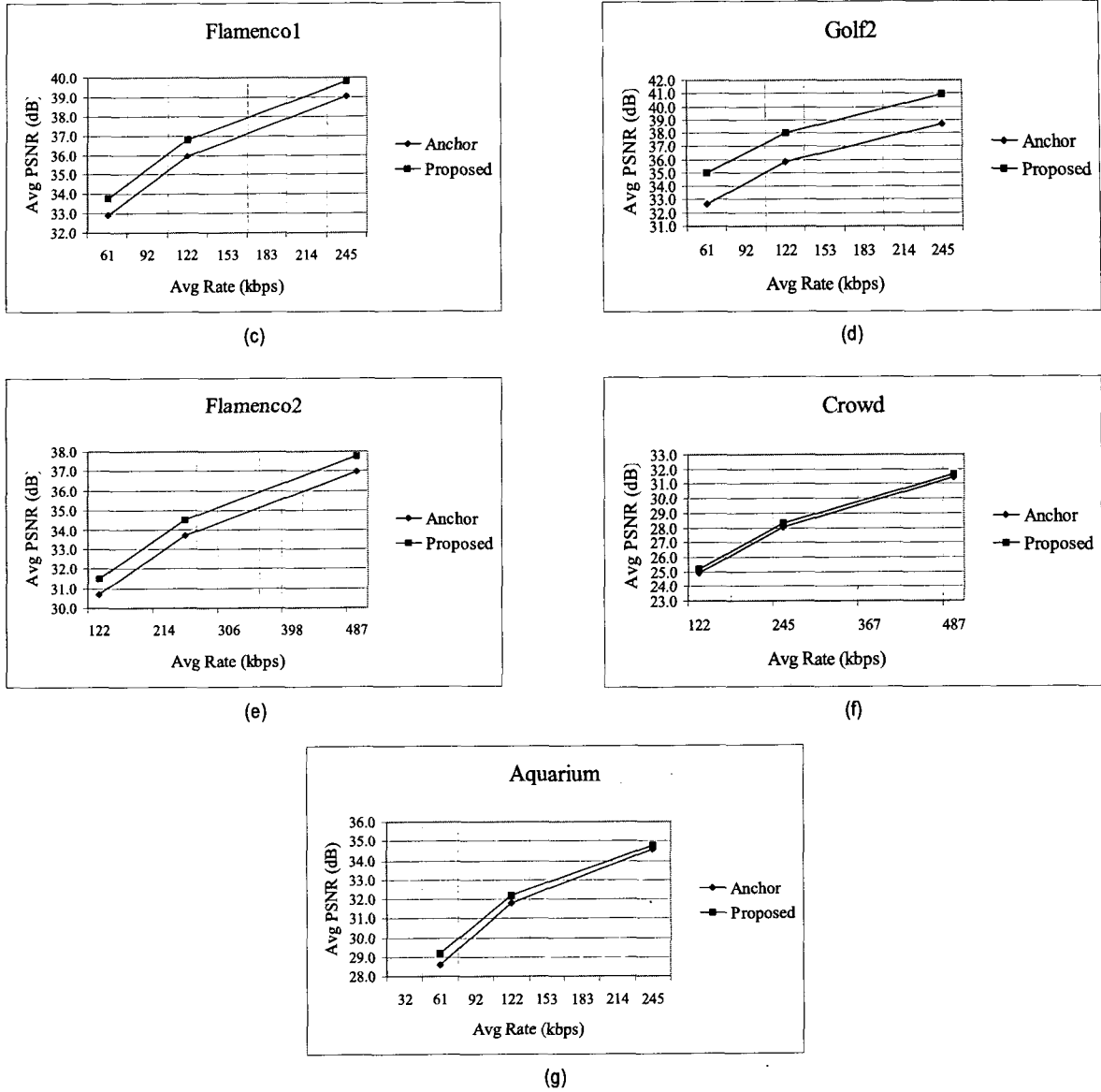


그림 4. 제안된 방법과 MPEG-4 AVC simulcast와의 비교

Fig. 4. Comparison of PSNR between the MPEG-4 AVC simulcast and proposed method: (a) race1: 1-D parallel data set, (b) race2: 1-D parallel data set, (c) flamenco1: 1-D parallel data set, (d) golf2: 1-D parallel data set, (e) flamenco2: 2-D parallel data set, (f) crowd: 2-D parallel data set, and (g) aquarium: 1-D arc data set.

## VI. 결론

본 논문에서는 주어진 화질에서 최고 43%의 비트를 절약할 수 있는 MVC 방법을 제안하였다. 이 방법과 더불어

참고문헌 [8]에서 제안하는 GOP구조를 사용하면 효율은 더욱 높아질 것이다. 이를 미래의 TV라고 불리는 3차원 TV나 Free View-point video 같은 어플리케이션에서 사용된다면 같은 네트워크 상황에서 더 뛰어난 화질의 서비스

를 제공할 수 있고, 아니면 같은 화질의 서비스를 네트워크 대역폭이 원만하지 않은 곳에도 제공할 수 있는 장점이 있고 현재 MPEG-4 AVC에 쉽게 확장될 수 있어서 많은 사용이 예상된다.

**참 고 문 헌**

[1] Yung-Lyul Lee and Woo-Chul Sung, "Multi-view video coding using 2D direct mode," ISO/IEC JTC1/SC29 WG11 m11266, Palma de Mallorca, Spain, October 2004.

[2] T. Wiegand, Final draft international standard for joint video specification H.264, Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, JVT-G050, March 2003.

[3] "Call for Evidence on Multi-View Coding," ISO/IEC JTC1/SC29/WG11 N6720, October 2004.

[4] M. Flierl and B. Girod, "Generalized B pictures and the draft JVT/H.264 video compression standard," IEEE Trans. Circuits Syst. Video Technol., vol. 13, pp. 587597, July 2003.

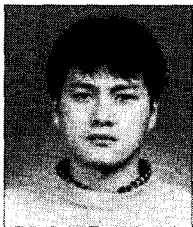
[5] H. Malvar, A. Hallapuro, M. Karczewicz, and L. Kerofsky, "Low-Complexity transform and quantization in H.264/AVC," IEEE Trans. Circuits Syst. Video Technol., vol. 13, pp. 598603, July 2003.

[6] Zhibo Chen, Peng Zhou, Yun He, "Fast Motion estimation for JVT," JVT-G016.doc, Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, 7th meeting, Pattaya, TH, 5-13 March 2003.

[7] [http://bs.hhi.de/~suehring/tml/download/old\\_jm/jm76.zip](http://bs.hhi.de/~suehring/tml/download/old_jm/jm76.zip)

[8] "Survey of algorithms used for Multi-view Video Coding (MVC)," ISO/IEC JTC1/SC29/WG11, N6909, HongKong, China, January 2005.

**저 자 소 개**



**이 용 기**

- 2002년 2월 : 세종대학교 전산과학과 학사과정 졸업
- 2004년 8월 : 세종대학교 컴퓨터공학과 석사과정 졸업
- 2004년 9월 ~ 현재 : 세종대학교 인터넷공학과 박사과정 재학
- 주관심분야 : Multimedia(MPEG) Systems, H.264, Image Processing, 3-Dimensional Audio Vid대, Transcoding



**허 재 호**

- 2005년 2월 : 세종대학교 컴퓨터공학과 학사과정 졸업
- 2005년 3월 ~ 현재 : 세종대학교 인터넷공학과 석박사통합과정 재학
- 주관심분야 : Image Processing, MPEG Video Compression, Transcoding, 3-Dimensional Audio Vid대



**이 영 렬**

- 1985년 2월 : 서강대학교 전자공학과 학사과정 졸업
- 1987년 2월 : 서강대학교 전자공학과 석사과정 졸업
- 1999년 2월 : 한국과학기술원 전기·전자공학과 박사과정 졸업
- 1987년 1월 ~ 1994년 2월 : 삼성전자 R&D 센터 Digital Media Lab.
- 1999년 3월 ~ 2001년 8월 : 삼성전자 R&D 센터 Digital Media Lab. 수석연구원
- 2001년 9월 ~ 현재 : 세종대학교 전자정보대학 컴퓨터공학부 인터넷공학과 부교수
- 주관심분야 : Multi-view video coding, MPEG Video Compression, Multimedia(MPEG) Systems, H.264, Image Compression, Image Processing, Multimedia Data(A/V/System) over IP, Embeded System for A/V Transport, Transcoding