

효율적인 보조 정보 생성을 통한 깊이지도 기반의 분산 다시점 비디오 코딩 기법

준회원 유지 환*, 이동 석*, 김 태 준*, 종신회원 유지 상*

Depth Map Based Distributed Multi-view Video Coding Scheme through an Efficient Side Information Generation

Jihwan Yoo*, Dong-seok Lee*, Tae-june Kim* *Associate Members,*
Ji-sang Yoo* *Lifelong Member*

요 약

본 논문에서는 효율적인 보조 정보 생성을 통한 새로운 분산 다시점 비디오 코딩 기법을 제안한다. 분산 비디오 코딩은 원영상과 디코더에서 생성한 보조 정보 간의 오차를 채널 코딩 기법으로 정정한다. 따라서 보조 정보를 정확히 만들수록 분산 비디오 코딩의 성능은 좋아지게 된다. 제안한 기법에서는 깊이지도를 기반으로 하는 다시점 비디오 코딩에 분산 비디오 코딩 기법을 적용한다. 또한 깊이지도를 이용한 3차원 워핑을 통해 인접한 시점의 영상으로부터 보조 정보를 생성하고, 3차원 워핑과 시간 축 상의 인접한 영상을 이용하는 MCTI(motion compensated temporal interpolation)를 효율적으로 혼합하여 사용한다. 실험 결과 제안한 기법으로 생성한 보조 정보는 MCTI와 3차원 워핑을 따로 사용한 방법보다 평균 0.97dB의 PSNR이 향상되었음을 알 수 있었다. 또한 R-D 곡선 상에서 동일 PSNR 대비 평균 8.01%의 비트율이 감소되었다.

Key Words : Distributed Video Coding, Multi-view Video Coding, Side Information, Depth Map, 3D Warping

ABSTRACT

In this paper, we propose a new depth map based distributed multi-view video coding algorithm through an efficient side information generation. A distributed video coding scheme corrects errors between an original image and side information generated at a decoder by using channel coding techniques. Therefore, the more accurate side information is generated, the better performance of distributed video coding scheme is achieved. In the proposed algorithm, a distributed video coding scheme is applied to multi-view video coding based on depth map. Side information is also generated from images of adjacent views through 3D warping by using a depth map and is also combined with MCTI(motion compensated temporal interpolation) which uses images on a temporal axis, and 3D warping. Experimental results show that side information generated by using the proposed algorithm has 0.97dB better average PSNR compared with using MCTI and 3D warping separated. In addition, 8.01% of average bit-rate has been decreased while the same PSNR in R-D curves is kept.

I. 서 론

방송 기술은 지난 수십 년 동안 흑백 TV에서 컬러

※ 본 연구는 지식경제부 및 정보통신산업진흥원의 IT산업원천기술개발사업[2009-F-208-01, 대화형 디지털 홀로그램 통합서비스 시스템의 구현을 위한 신호처리 요소 기술 및 SoC 개발]과 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업(NIPA-2009-(C1090-0902-0017))의 일환으로 수행하였음.

* 광운대학교 전자공학과 디지털 미디어 연구실({uzhwan, dlehdtr86, mirrine}@kw.ac.kr, jsyoo@kw.ac.kr)
논문번호 : KICS2009-07-316, 접수일자 : 2009년 7월 30일, 최종논문접수일자 : 2009년 10월 16일

러 TV, HDTV에 이르기까지 눈부신 발전을 이루어 왔다. 특히 최근에는 세계 각국에서 기존의 아날로그 방송을 중단하고 디지털 방송으로 전환하기 시작했다으며 우리나라도 이러한 추세에 따라 2012년까지 디지털 방송으로 전환할 계획이다. 하지만 이와 같은 발전에도 불구하고 사용자들은 더 현실감 있는 방송을 요구하고 있고, 이에 따라 세계 각국의 연구소 및 기업에서는 실감 방송을 위한 3DTV (3-dimensional television) 및 UHDTV(ultra high definition television)를 연구 중이다.

ISO/IEC 산하의 MPEG(moving picture experts group)에서는 2001년부터 3DAV(3-dimensional audio-visual)란 이름으로 3차원 영상에 관한 표준화를 진행하고 있다^[1]. 특히 2008년 7월에는 다시점 비디오 코딩(multi-view video coding, MVC)을 표준화하였다^[2]. 다시점 비디오 코딩은 시간 축 방향 뿐만 아니라 시점 방향의 인접한 영상들도 참조하기 때문에 인코더의 복잡도가 매우 높다는 단점이 있다.

또한 인코더의 복잡도를 디코더로 옮기는 분산 비디오 코딩(distributed video coding, DVC)에 대한 연구도 진행 중이며, 이를 다시점 비디오 코딩에 적용하려는 시도가 있었다^{[3],[4]}. 분산 비디오 코딩의 인코더는 영상을 인트라 코딩(intra coding)되는 키 영상(key frame)과 패리티 비트(parity bit)만을 전송하는 Wyner-Ziv 영상으로 나눠 부호화한다. 디코더에서는 전송된 키 영상으로부터 Wyner-Ziv 영상과 비슷한 보조 정보(side information)를 생성하고, Wyner-Ziv 영상과 보조 정보 사이의 오차를 전송된 패리티 비트를 이용한 채널 코딩 기법을 통해 정정한다. 기존의 분산 다시점 비디오 코딩에서는 시간 축 상의 움직임 정보를 이용하는 MCTI(motion compensated temporal interpolation)와 인접하는 시점의 영상 간의 호모그래피(homography)를 이용하는 HCII(homography compensated inter-view interpolation)를 적절히 혼합하여 보조 정보를 생성했다. 하지만 HCII 방식은 호모그래피 추정 과정에서 발생하는 대응점의 부정확성과 워핑(warping)할 때의 보간 성능의 한계로 보조 정보의 정확성이 떨어진다.

이 문제를 해결하기 위해서 본 논문에서 깊이 지도(depth map)를 이용한 3차원 워핑(3-dimensional warping)을 통하여 인접하는 시점 영상으로부터 보다 정확한 보조 정보를 생성하는 기법을 제안하고자 한다. 또한 MCTI 기법과 혼합하여 더욱 정확한

보조 정보를 얻기 위해서 먼저, 매크로블록 단위로 Wyner-Ziv 영상의 인접하는 전, 후 영상의 차이를 구한다. 만약 차이 값이 작다면 배경 영역이라고 판단하여 전, 후 영상의 평균값을 가지고 보조 정보를 생성한다. 일반적으로 3차원 워핑은 깊이지도의 경계 영역에서 성능이 많이 떨어지므로 이 영역에서는 MCTI 기법을 사용한다. 마지막으로 MCTI 과정에서 발생하는 움직임 벡터의 크기와 움직임 보상된 매크로블록의 잔여 신호(residual signal) 값이 큰 영역에서는 움직임 예측이 부정확하다고 판단하여 3차원 워핑을 이용한다. 이때 주변 매크로블록의 움직임 크기를 참조하여 부정확한 움직임 예측으로 인한 오차를 방지한다.

본 논문은 다음과 같이 구성된다. II장에서는 기존의 다시점 비디오 코딩 기법에 대해 알아보고, III장에서는 기존의 기법들이 가지는 문제점을 해결하기 위해 제안하는 깊이지도 기반의 3차원 워핑을 통한 보조 정보 생성 기법과 MCTI와의 혼합 기법을 사용한 새로운 깊이지도 기반의 분산 다시점 비디오 코딩 기법에 대해 설명한다. IV장에서 실험 결과를 보이고, V장에서 결론을 기술한다.

II. 기존의 비디오 코딩

2.1 다시점 비디오 코딩

그림 1은 다시점 비디오 코딩의 기본적인 예측 구조(prediction structure)를 나타낸다. 다시점 비디오 코딩은 시간 축 방향으로는 계층적 B 픽처(hierarchical B picture) 구조이며, 시점 간의 예측을 추가하여 압축 효율을 극대화하였다. 따라서 기존의 비디오 코딩 기법에 비해 인코더의 복잡도가

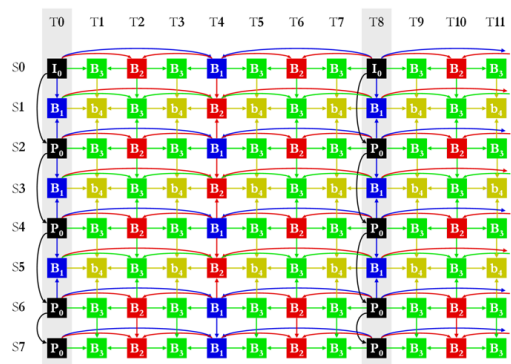


그림 1. 다시점 비디오 코딩의 예측구조
Fig. 1. Prediction structure for multi-view video coding

매우 높다. 인코더의 복잡도를 낮추기 위해서 현재 시점과 인접하는 시점 간의 물체의 움직임 정보가 유사하다는 특성을 이용하여 움직임 예측을 생략하는 움직임 정보 생략 모드(motion skip mode)^[5] 등의 기법들이 제안되었지만 복잡도를 크게 낮추지는 못했다.

2.2 분산 비디오 코딩

앞서 언급하였듯이 분산 비디오 코딩은 인코더의 복잡도를 디코더로 옮기기 위해 도입된 기법이다. 그림 2(a)에 보인 기존의 비디오 코딩 기법은 부호화 하려는 영상 X 와 보조 정보 Y 와의 상관도를 인코더에서 제거하였다. 여기서 보조 정보(side information)는 화면 내 예측(intra prediction)이나 화면 간 예측(inter prediction)을 통해 얻은 예측 영상(predicted image)이며, 인코더와 디코더 모두에서 접근이 가능하다. 반면에 그림 2(b)의 분산 비디오 코딩 기법은 X 와 Y 의 상관도를 디코더에서 제거한다. 여기서 Y 는 디코더에서 X 와의 상관도를 이용하여 생성되는 보조 정보로써 디코더에서만 접근이 가능하다. 여기서 중요한 것은 기존의 비디오 코딩과 분산 비디오 코딩의 성능이 같을 수 있다는 것이다. 이것은 1970년대에 발표된 두 개의 정보이론에 바탕을 두고 있다^{[6],[7]}.

먼저, Slepian-Wolf는 디코더에서 두 데이터의 중복도를 제거하더라도 인코더에서 제거하였을 때와 같은 성능을 낼 수 있다는 것을 보여주었다^[6]. 그림 3은 분산 압축 시에 달성 가능한 Slepian-Wolf 코딩의 비트양의 영역을 나타낸다. 여기서 R_X 와 R_Y 는 각각 두 데이터 X 와 Y 의 비트양을 나타내며,

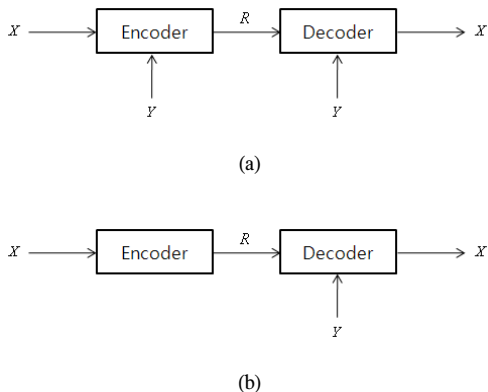


그림 2. 보조 정보의 접근 방법 (a) 기존의 비디오 코딩 (b) 분산 비디오 코딩
Fig. 2. Approach methods for side information (a) Conventional video coding (b) Distributed video coding

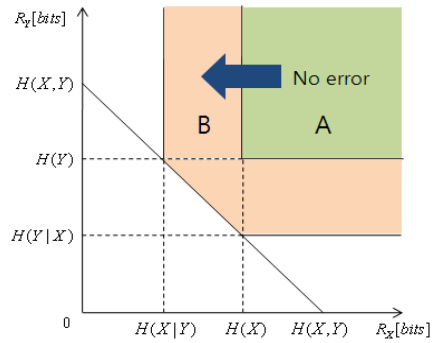


그림 3. Slepian-Wolf : 분산 압축 시 달성 가능한 비트양의 영역
Fig. 3. Slepian-Wolf : achievable rate region for distributed compression

$H(X)$ 와 $H(Y)$ 는 각각 X 와 Y 의 정보량을 정량화한 엔트로피(entropy)이다. 이때 두 데이터 X 와 Y 가 통계적 독립(statistically independent)이라면 발생 비트양은 식 (1)이 나타내는 그림 3의 영역 A 내에 존재하게 된다.

$$\begin{aligned} R_X &\geq H(X) \\ R_Y &\geq H(Y) \end{aligned} \quad (1)$$

하지만 두 데이터 X 와 Y 가 통계적 독립이 아닌 확률적으로 상관도가 있는 관계라면, 발생 비트양의 영역은 식 (2)에 해당하는 그림 3의 영역 B로 확장된다.

$$\begin{aligned} R_X &\geq H(X|Y) \\ R_Y &\geq H(Y|X) \\ R_X + R_Y &\geq H(X, Y) \end{aligned} \quad (2)$$

즉, 두 데이터 X 와 Y 의 상관도가 높을수록 발생하는 비트양은 점차 줄어들게 되며, 결국에는 분산 코딩도 기존 코딩과 같은 압축률을 가질 수 있다는 것을 보여주고 있다.

Wyner-Ziv는 Slepian-Wolf의 코딩 이론에 양자화 과정을 추가하여 분산 비디오 코딩의 개념을 손실 압축으로 확장하였다^[7]. 기존 비디오 코딩의 율-왜곡(rate-distortion, R-D) 함수 $R_{X|Y}(d)$ 와 분산 비디오 코딩의 율-왜곡 함수 $R^*(d)$ 의 관계는 일반적으로 $R_{X|Y}(d) < R^*(d)$ 이다. 하지만 데이터가 가우시안 분포(Gaussian distribution)를 갖는 등 어느 특정한 경우에는 $R_{X|Y}(d) = R^*(d)$ 이 될 수 있음을 증명하였다.

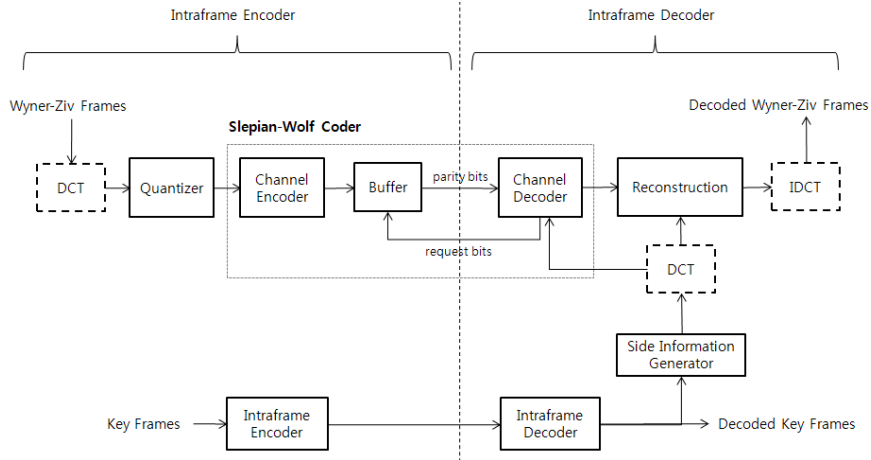


그림 4. 분산 비디오 코딩의 구조
Fig. 4. Structure of distributed video coding

그림 4는 일반적인 분산 비디오 코딩의 구조를 나타낸다. 먼저 원본 영상은 키 영상과 Wyner-Ziv 영상으로 나뉜다. 이때 GOP(groups of picture)는 키 영상과 키 영상 사이의 영상들의 집합이며, 다수의 Wyner-Ziv 영상으로 구성된다^[8]. 키 영상은 기존의 JPEG이나 H.264/AVC의 인트라 코딩 기법을 사용하여 부호화 된다. 반면에 Wyner-Ziv 영상은 양자화 과정을 거친 후 터보 코드(turbo code)^[9]나 LDPC(low-density parity-check) 코드^[10]와 같은 채널 코드를 사용하여 부호화되며 이때 발생된 패리티 비트(parity bit)는 버퍼에 저장된다. 여기서 양자화 과정 전에 주파수 영역으로의 변환 과정이 포함될 수 있고, 일반적으로 주파수 변환 과정이 포함되면 영상의 공간적 중복도(spatial redundancy)가 제거되므로 성능이 더 좋게 된다^[11].

디코더에서는 복호화된 키 영상으로부터 부호화하려는 Wyner-Ziv 영상과 유사한 보조 정보를 생성한다. 이때 디코더는 보조 정보와 Wyner-Ziv 영상의 오차를 가상의 채널 잡음(virtual channel noise)으로 간주하여 전송된 패리티 비트를 이용한 강력한 채널 코딩 기법으로 에러를 정정한다. 만약 복호화에 실패하면 디코더는 복호화가 성공할 때까지 인코더에 추가적인 패리티 비트를 요청하게 된다. 따라서 분산 비디오 코딩의 성능은 정확한 보조 정보의 생성과 채널 코딩의 성능에 달려있다고 할 수 있다.

현재 분산 비디오 코딩에 관한 연구는 미국 Stanford 대학의 Girod^{[8],[11]-[12]}, Berkeley 대학의

Ramchandran^[13]과 유럽의 공동 연구단체인 DISCOVER^[14]에서 주도적으로 진행 중이다.

2.3 분산 다시점 비디오 코딩

다시점 비디오 환경에서는 시간 축의 인접한 영상뿐만 아니라 인접한 시점의 영상에도 접근이 가능하기 때문에 보조 정보(side information)를 더욱 정확하게 만들 수 있다. 일반적으로 분산 다시점 비디오 코딩의 보조 정보를 생성하는 방법에는 다음과 같은 것들이 있다.

첫째는 시간 축의 인접하는 영상 간의 움직임 벡터를 이용하는 MCTI(motion compensated temporal interpolation) 기법이다. 그림 5와 같이 전, 후 영상 k-1과 k+1 번째 프레임으로부터 Wyner-Ziv 영상과 유사한 보조 정보를 생성하는 방법이다. 이를 위해서 먼저, 그림 5(a)와 같이 k-1과 k+1 프레임 간의 움직임 예측을 통해 움직임 벡터를 구하고, 이 값의 절반 값을 이용하여 인접하는 영상들로부터 보조 정보를 보간한다. 하지만 이때 보간된 블록 간에 빈 공간 또는 겹치는 영역이 발생하기 때문에 좌표 A의 움직임 벡터를 블록의 중심 좌표 B로 평행이동시킨다. 그 후에 그림 5(b)와 같이 양방향 움직임 예측을 통해 최종적인 움직임 벡터를 구하게 된다. 마지막으로 k-1과 k+1 프레임의 화소 값의 평균값으로 보간 영상을 생성한다.

보조 정보를 생성하는 다른 방법은 동일 시간대의 인접하는 시점 영상을 이용하는 HCII(homography compensated inter-view interpolation)이다. 그림 6과 같이 영상의 호모그래피(homography)를 이용하

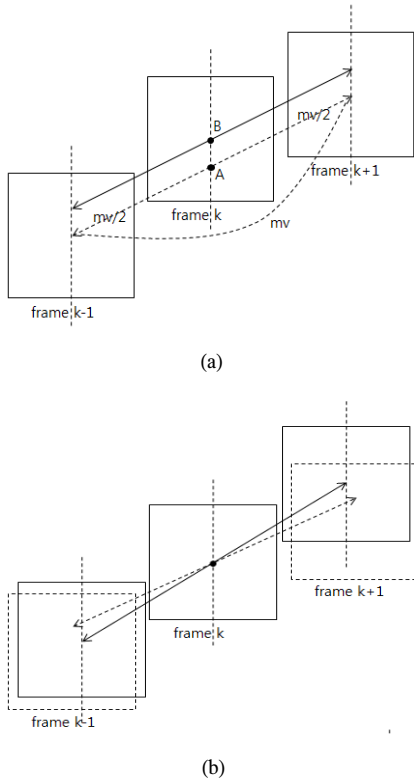


그림 5. MCTI (motion compensated temporal interpolation)
 (a) 움직임 벡터의 선택 (b) 양방향 움직임 추정
 Fig. 5. MCTI (motion compensated temporal interpolation)
 (a) Selection of motion vector (b) Bidirectional motion estimation

여 인접하는 시점 영상과의 변위(displacement)를 구하고 이 값을 이용하여 보조 정보를 생성한다.

끝으로 MCTI와 HCII를 혼합하여 보조 정보를 생성하는 방법이다. 움직임이 큰 영역에서는 움직임 벡터가 부정확하기 때문에 MCTI를 사용하여 보간된 블록은 오차가 크다. 따라서 움직임 벡터의 크기를 기준으로 특정 크기를 넘는 영역에서는 MCTI가 아닌 HCII를 사용하여 보조 정보를 생성한다. 일반적으로 다시점 비디오 코딩에서는 시간 축의 인접한 영상들 간의 상관도가 인접한 시점의 영상들 간의 상관도보다 높기 때문에 시간 축 방향의 영상이

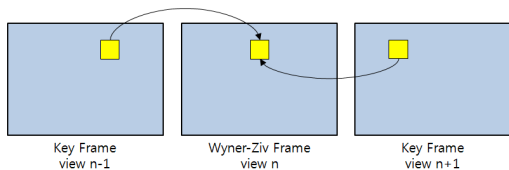


그림 6. HCII(homography compensated inter-view interpolation)
 Fig. 6. HCII(homography compensated inter-view interpolation)

참조영상으로 선택되는 빈도가 높다. 하지만 분산 다시점 비디오 코딩에서는 움직임 예측이 수행되는 키 영상 간의 거리가 멀기 때문에 시간 축 방향의 상관도가 작아지게 된다. 따라서 이러한 문제점을 보완해주는 혼합 방식을 사용하면 보다 정확한 보조 정보를 얻을 수 있다.

III. 제안하는 알고리즘

본 논문에서는 깊이지도 기반의 3차원 워핑을 통하여 보다 정확한 보조 정보를 생성하는 기법과 MCTI(motion compensated temporal interpolation)와의 효율적인 혼합 기법을 제안한다. 제안하는 기법에서 필요한 깊이지도(depth map)와 카메라 파라미터(camera parameter)는 인코더로부터 전송되었다고 가정한다. 그림 7은 본 논문에서 제안하는 깊이지도 기반의 분산 다시점 비디오 코딩의 구조를 보여준다. 일반적인 분산 비디오 코딩 구조를 기반으로 시간 축 및 시점 방향의 영상들 간의 상관도를 이용해서 보조 정보를 생성한다.

그림 8은 일반적인 분산 다시점 비디오 코딩의 영상 구조를 나타낸다. 영상은 인트라 코딩되는 키 영상(I)과 채널 코딩되는 Wyner-Ziv 영상(WZ)이 시간 축 방향과 시점 방향으로 교대로 배치된다. 따라서 Wyner-Ziv 영상의 보조 정보는 인접하는 4개의 키 영상으로부터 생성될 수 있다. 본 논문에서는 인접하는 시점의 키 영상으로부터 3차원 워핑을 이용하여 보조 정보를 생성하는 기법과 시간 축 방향의 인접하는 키 영상을 참조하는 MCTI 기법과의

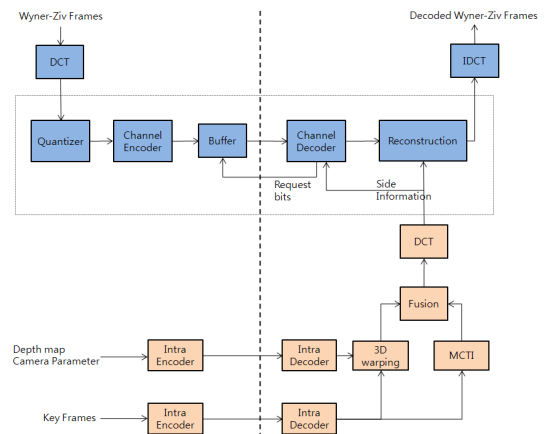


그림 7. 깊이지도 기반의 분산 다시점 비디오 코딩
 Fig. 7. Depth map based distributed multi-view video coding

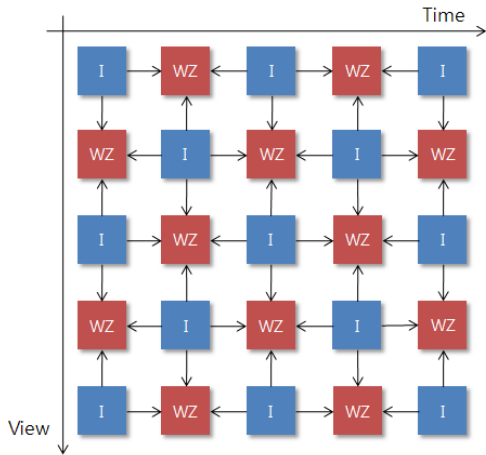


그림 8. 영상 구조
Fig. 8. Frame structure

효율적인 혼합 기법을 제안하고 있다.

3.1 3차원 워핑을 이용한 보조 정보의 생성

Wyner-Ziv 영상의 인접하는 시점 영상의 3차원 워핑을 통해 보조 정보를 생성할 수 있다. 먼저 그림 9(a)에 보인 우 영상의 깊이지도에 소벨 마스크(Sobel mask)를 사용해서 그림 9(b)와 같은 경계를 추출한다. 추출된 깊이지도의 경계는 일반적으로 원 영상의 객체의 경계 부분과 완전히 일치하지 않는다. 따라서 3차원 워핑을 이용한 보조 정보는 깊이

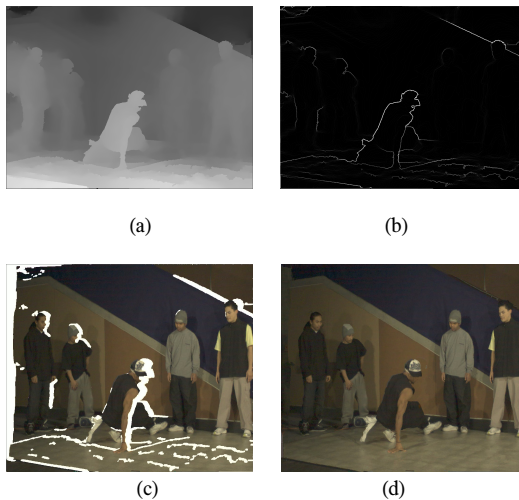


그림 9. 깊이지도와 카메라 파라미터를 이용한 3차원 워핑
(a) 깊이지도 (b) 깊이지도의 경계 (c) 투영된 영상 (d) 보조 정보
Fig. 9. 3D warping using depth map and camera parameters
(a) Depth map (b) Edge of depth map (c) Projected image (d) Side information

영상의 경계 부분에서 오차가 발생한다. 이러한 오차를 방지하기 위해 깊이지도의 경계 영역을 원 영상에서 제거하는 방식을 사용한다¹⁵⁾.

다음으로 경계가 제거된 2차원 영상을 깊이지도와 카메라 파라미터를 이용하여 식 (3)을 통해 3차원 좌표로 변환한다. 식 (3)은 일반적인 핀홀 카메라(pinhole camera)의 기하학적 구조를 이용하여 얻은 2차원 좌표와 3차원 좌표의 관계를 나타낸다.

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K[R|T] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (3)$$

여기서 x 와 y 는 영상의 2차원 좌표이고, X , Y , Z 는 실세계의 3차원 좌표를 나타낸다. K 는 카메라의 내부 파라미터(intrinsic parameter)를 나타내는 3x3 행렬이다. R 은 카메라의 회전(rotation)을 나타내는 3x3 행렬의 외부 파라미터(extrinsic parameter)이며, T 는 카메라의 이동(translation)을 나타내는 3x1 행렬의 외부 파라미터이다. 이때 Z 값은 식 (4)를 통해 계산된다.

$$Z(i,j) = 1 / \left(\frac{P(i,j)}{255} \times \left(\frac{1}{MinZ} - \frac{1}{MaxZ} \right) + \frac{1}{MaxZ} \right) \quad (4)$$

여기서 $P(i,j)$ 는 각각의 화소에 대응되는 깊이지도의 값이며, $MinZ$ 와 $MaxZ$ 를 통해 Z 의 최소값과 최대값을 정의한다.

변환된 3차원 좌표는 다시 식 (3)을 이용하여 그림 9(c)와 같이 원하는 시점의 2차원 평면으로 재투영된다. 이때, 카메라 파라미터는 Wyner-Ziv 영상이 속하는 시점의 카메라 파라미터 값이 된다. 하지만 그림 9(c)에서 확인할 수 있듯이 재투영된 영상에는 홀(hole)과 가려진 영역(occluded region)이 존재하게 된다. 홀은 3차원 실수 좌표가 2차원 정수 좌표로 변환되는 과정에서 발생하는 것이고, 가려진 영역은 시점이 바뀔 때 보이지 않던 부분이 나타나면서 발생하는 영역이다. 홀은 인접한 화소에 선형 보간법을 적용하여 채울 수 있고, 가려진 영역은 좌영상의 정보를 가지고 채울 수 있다¹⁵⁾. 홀과 가려진 영역을 채운 뒤 미디언 필터를 적용함으로써 그림 9(d)와 같은 최종적인 보조 정보를 얻을 수 있다.

3.2 효율적인 보조 정보의 생성 기법

기존의 다시점 비디오 코딩에서는 부호화 하려는

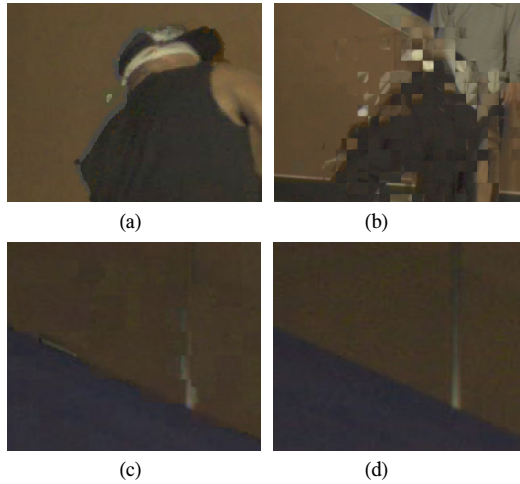


그림 10. 보조 정보의 문제점
 (a) 3차원 워핑에서 깊이지도의 경계 영역 (b) MCTI에서 움직임이 큰 영역 (c) 배경 오차의 처리 전 (d) 배경 오차의 처리 후
 Fig. 10. Problems of side information
 (a) Edge region of depth image in 3D warping (b) Large motion region in MCTI (c) Before processing background error (d) After processing background error

영상과 인접하는 참조 영상에 모두 접근할 수 있었기 때문에 각각의 참조 영상에 대해서 율-왜곡 (rate-distortion) 값을 구해 최소값을 갖는 모드를 선택하였다. 하지만 분산 다시점 비디오 코딩에서는 Wyner-Ziv 영상이 전송되지 않고, 패리티 비트만이 전송되기 때문에 이와 같은 방법은 불가능하다. 따라서 본 논문에서는 전, 후 영상 간의 차이 값, 깊이지도의 경계 정보, 움직임 벡터의 크기와 움직임 보상된 매크로블록의 잔여 신호(residual signal)의 크기를 이용하여 보조 정보를 효율적으로 생성하는 기법을 제안한다.

일반적으로 3차원 워핑을 이용하여 생성한 보조 정보는 그림 10(a)와 같이 깊이지도의 경계 영역에서 오차가 발생한다. 또한 MCTI를 이용하여 생성된 보조 정보는 그림 10(b)와 같이 움직임이 큰 영역에서 부정확한 값을 갖는다. 따라서 깊이지도의 경계에서는 MCTI를 이용한 보조 정보를 사용하고, 움직임이 큰 영역에서는 3차원 워핑을 이용한 보조 정보를 혼합하여 사용할 수 있다면 더욱 정확한 보조 정보를 생성할 수 있다. 즉, 깊이지도의 경계 영역에 해당하는 매크로블록에서는 MCTI를 이용한 보조 정보를 사용하고, 움직임 벡터의 크기 및 움직임 보상된 매크로블록의 잔여 신호의 크기가 각각 임계치 Th_2 , Th_3 를 넘는다면 3차원 워핑을 통해 생성된 보조 정보를 이용하는 것이다.

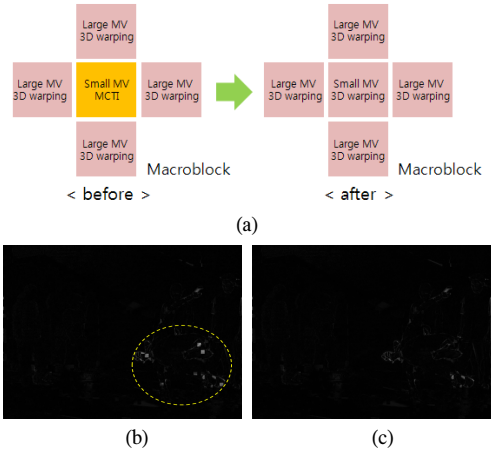


그림 11. 부정확한 움직임 영역의 처리
 (a) 부정확한 움직임 영역 (b) 처리 전 차영상 (c) 처리 후 차영상
 Fig. 11. Processing of inaccurate motion region
 (a) Inaccurate motion region (b) Difference image before process (c) Difference image after process

하지만 움직임 예측은 종종 부정확하기 때문에 그림 10(c)에서와 같이 움직임이 없는 배경 영상에서 오차가 발생하는 것을 확인할 수 있다. 따라서 본 논문에서는 이러한 문제점을 해결하기 위해서 Wyner-Ziv 영상의 전, 후 영상의 차이 값을 구하고 임계치 Th_1 과 비교한다. 만약 차이 값이 Th_1 보다 작다면 움직임이 없는 배경 영상이라 판단하여 전, 후 영상의 평균값으로 보조 정보를 생성한다. 그림 10(d)에서 배경 영상에 대한 오차가 정정된 것을 확인할 수 있다.

또한 움직임 벡터의 값을 이용할 때는 주변 매크로 블록의 움직임 벡터의 값도 고려하여야 한다. 예를 들어, 그림 11(a)와 같이 주변 매크로 블록의 움직임 벡터는 큰 값을 갖지만 중앙 매크로 블록의 움직임 벡터는 작은 값을 갖는 경우이다. 이때는 중앙 매크로 블록의 움직임 예측 과정이 부정확했을 확률이 크다. 이런 경우 주변 매크로 블록의 움직임 벡터의 값을 고려하지 않고 보조 정보를 선택한다면 그림 11(b)의 노란 원 영역과 같이 오차가 발생하는 것을 확인할 수 있다. 이런 오차를 보정하기 위해서는 보조 정보를 선택할 때 주변 매크로 블록의 움직임 벡터의 값도 고려해서 그림 11(a)와 같이 중앙 매크로 블록의 움직임 벡터가 작은 크기를 갖더라도 MCTI가 아닌 3D 워핑을 이용한 보조 정보를 사용하도록 했다. 그림 11(c)에서 오차가 정정된 것을 확인할 수 있다.

그림 12는 본 논문에서 제안하는 효율적인 보조

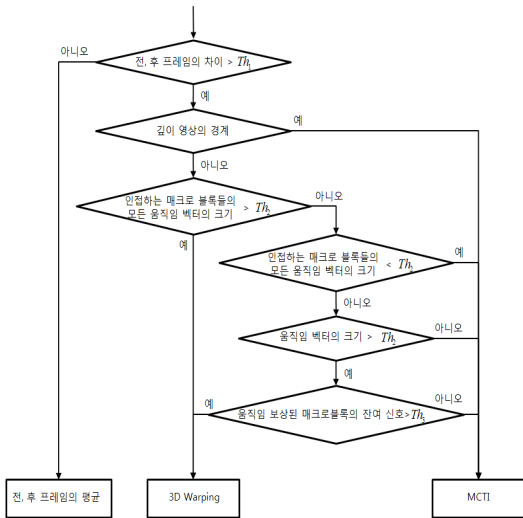


그림 12. 보조 정보 생성 기법
Fig. 12. Side information fusion scheme

정보 생성 기법에 대한 순서도를 보여주고 있다.

IV. 실험 결과

본 논문에서는 실험 영상으로 MSR(microsoft research)에서 제공하는 8시점의 Breakdancers와 Ballet 영상을 사용하였다. 영상은 각 시점 당 100 프레임으로 구성되어 있으며, 본 실험을 위해서 각 시점

당 30 프레임의 영상을 사용하였다. 실험 영상의 해상도는 XVGA(1024x768)이며 초당 15 프레임으로 구성되어 있다. 또한 주파수 변환은 8x8 DCT를 사용하였고, 양자화 계수는 JPEG Annex K^[16]의 양자화 표에 Q = 0.5, 1, 2, 4의 가중치를 곱하여 사용하였다. 그리고 본 논문에서는 채널 코드로 Varodayan 등이 제안한 LDPCA(low-density parity-check accumulate) 코드를 사용하였다^[17]. LDPCA는 기존의 고정 비트율을 가지는 LDPC(low-density parity-check)에 비해 비트율이 적응적인 장점을 가지고 있다.

그림 13(a), (b), (c)는 각각 실험 영상 Breakdancers와 Ballet에 대하여 MCTI, 3차원 워핑, 그리고 제안하는 새로운 혼합 기법을 사용하여 생성한 보조 정보이고, 그림 14(a), (b), (c)는 두 실험 영상에 대하여 원본 영상과 보조 정보와의 오차를 나타낸 차영상이다. MCTI의 차영상인 그림 14(a)에서 움직임이 큰 영역의 오차가 큰 것을 확인할 수 있으며, 3차원 워핑의 차영상인 그림 14(b)에서는 객체의 경계 부분의 오차가 큰 것을 확인할 수 있다. 또한 제안하는 효율적인 혼합 기법의 차영상인 그림 14(c)에서는 MCTI와 3차원 워핑에서 나타났던 오차가 크게 준 것을 확인할 수 있다.

표 1은 Breakdancers와 Ballet 영상의 원본 영상과 보조 정보를 비교한 PSNR을 나타낸다. 본 논문에서 제안하는 MCTI와 3차원 워핑을 혼합하여 보

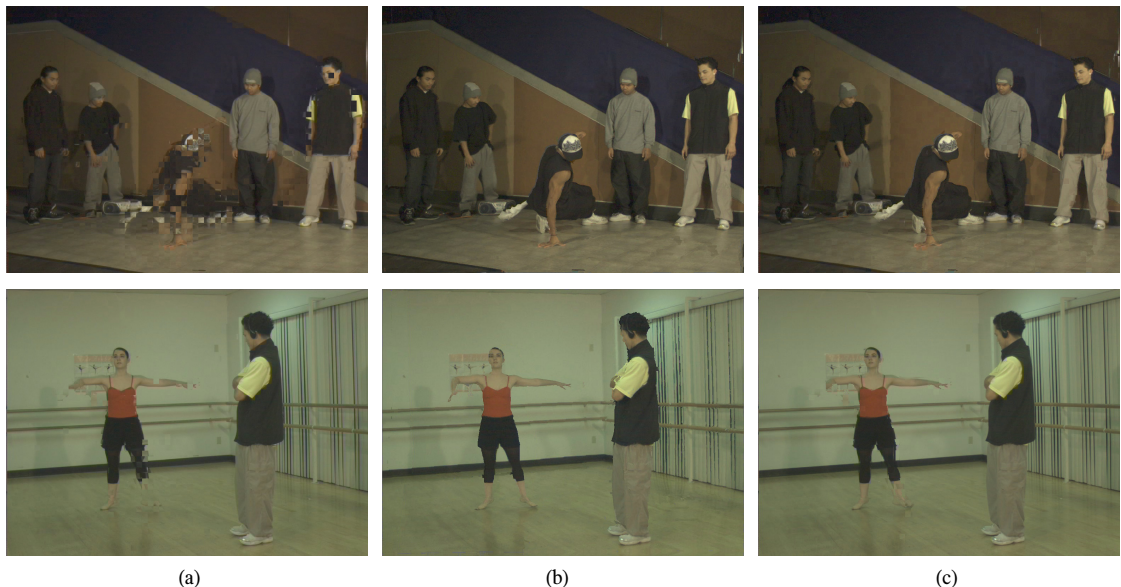


그림 13. Breakdancers와 Ballet의 보조 정보 (a) MCTI (b) 3차원 워핑 (c) 제안하는 혼합 방법
Fig. 13. Side information for Breakdancers and Ballet (a) MCTI (b) 3D warping (c) Proposed method

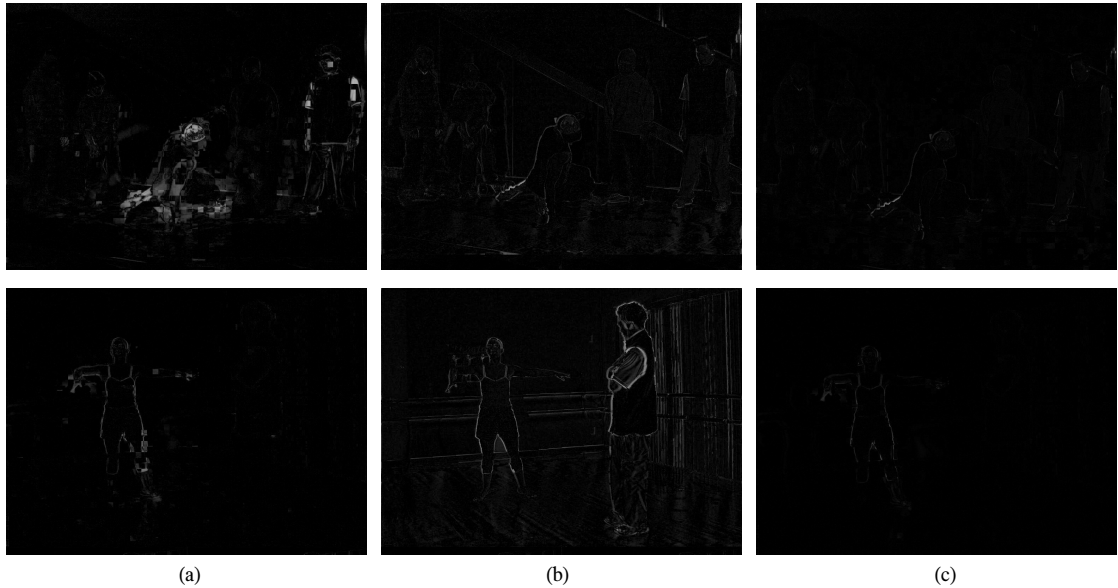


그림 14. Breakdancers와 Ballet의 차영상 (a) MCTI (b) 3차원 워핑 (c) 제안하는 혼합 방법
 Fig. 14. Difference images for Breakdancers and Ballet (a) MCTI (b) 3D warping (c) Proposed method

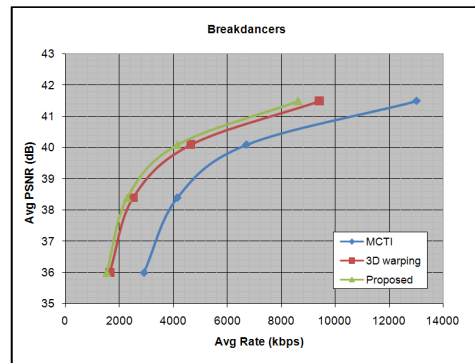
표 1. 평균 PSNR
 Table 1. Average PSNR

Sequence	Q(scaling factor)	PSNR(dB)		
		MCTI	3D warping	Proposed
Breakdancers	Q=0.5	26.82	33.23	33.41
	Q=1	26.85	33.34	33.54
	Q=2	26.81	33.06	33.13
	Q=4	26.68	32.00	32.13
	Average	26.79	32.91	33.05 (+0.14)
Ballet	Q=0.5	33.03	29.14	34.93
	Q=1	33.00	29.15	34.93
	Q=2	32.94	29.02	34.74
	Q=4	32.62	28.55	34.21
	Average	32.90	28.97	34.70 (+1.80)

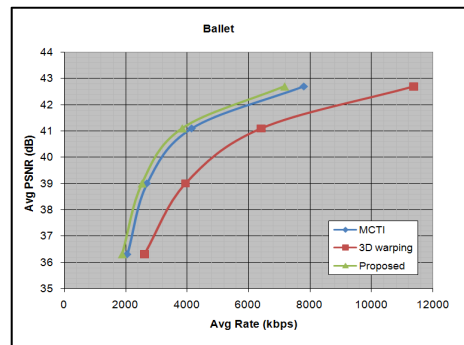
조 정보를 생성하는 기법이 각각을 따로 사용하는 기법보다 0.14~1.8dB 정도의 PSNR이 향상되는 것을 확인할 수 있다. 여기서 MCTI와 3차원 워핑의 성능이 실험 영상에 따라 달라지는 것은 실험 영상의 특성 때문이다. Breakdancers 영상은 영상 내 객체의 움직임이 매우 빨라 시간 축 방향의 정보를 이용하는 MCTI의 성능이 상대적으로 떨어지게 된다. 반면에 Ballet 영상은 카메라와 객체 사이의 거리가 가까워 시점의 이동에 따른 가려진 영역이 상대적으로 넓게 되고, 따라서 3차원 워핑의 성능이 떨어지게 된다.

그림 15는 Breakdancers와 Ballet 영상의 R-D

곡선(rate-distortion curve)을 나타낸다. Breakdancers 영상은 제안한 방법이 3차원 워핑을 이용한 방식보



(a)



(b)

그림 15. R-D 곡선 (a) Breakdancers (b) Ballet
 Fig. 15. R-D curve (a) Breakdancers (b) Ballet

다 동일 PSNR 대비 평균 8.62%의 비트 감소 효과가 있었다. 또한 Ballet 영상은 제안한 방법이 MCTI를 이용한 방식보다 동일 PSNR 대비 평균 7.53%의 비트 감소 효과가 있었다.

V. 결 론

본 논문에서는 효율적인 보조 정보를 이용한 깊이지도 기반의 새로운 분산 다시점 비디오 코딩 기법을 제안하였다. 제안된 기법에서는 분산 비디오 코딩의 성능을 향상시키기 위하여 보조 정보를 더욱 정확하게 생성하는 방법을 제안하고 있다. 즉 영상의 깊이지도를 이용한 3차원 위평을 통해 보조 정보를 생성하고, 3차원 위평과 MCTI를 적절히 혼합함으로써 코딩 성능을 향상시킬 수 있었다. 또한 MCTI와 3차원 위평의 보조 정보를 혼합하기 위해서 전, 후 영상 간의 차이 값, 깊이지도의 경계 정보, 움직임 벡터의 크기와 움직임 보상된 매크로 블록의 잔여 신호의 크기를 이용하였으며, 주변 매크로블록의 움직임 벡터의 크기를 고려하여 부정확한 움직임 예측으로 인한 오차를 줄일 수 있었다.

실험 결과에서 알 수 있듯이 제안된 기법을 사용한 보조 정보는 MCTI, 3차원 위평을 각각 사용하였을 때보다 평균 0.97dB의 PSNR이 향상되었고 또한, R-D 곡선 상에서 동일 PSNR 대비 평균 8.01%의 비트 감소율도 보임으로써 매우 효율적이라는 것을 확인할 수 있었다.

참 고 문 헌

- [1] ISO/IEC JTC1/SC29/WG11, "Requirements for Standardization of 3D Video," m8107, Jeju Island, Korea, March 2002.
- [2] ISO/IEC JTC1/SC29/WG11, "Text of ISO/IEC 14496-10:200X/FDAM 1 Multi-view Video Coding," N9978, Hannover, Germany, July 2008.
- [3] C. Guillemot, F. Pereira, L. Torres, T. Ebrahimi, R. Leonardi, and J. Ostermann, "Distributed monoview and multiview video coding," *IEEE Signal Processing Magazine*, vol.24, no.5, pp.67-76, September 2007.
- [4] F. Dufaux, M. Oualet, and T. Ebrahimi, "Recent Advances in Multi-view Distributed Video

- Coding," *SPIE Mobile Multimedia/Image Processing for Military and Security Applications*, Orlando, FL, April 2007.
- [5] H.-S. Koo, Y.-J. Jeon, B.-M. Jeon, "MVC Motion Skip Mode," ITU-T and ISO/IEC JTC1, JVT-W081, San Jose, California, USA, April 2007.
- [6] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *IEEE Trans. Inform. Theory* 19 (1973) 471-480.
- [7] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the receiver," *IEEE Trans. Inform. Theory* 22 (1976) 1-11.
- [8] B. Girod, A. Aaron, S. Rane and D. Rebollo Monedero, "Distributed video coding," in *Proc. IEEE*, vol. 93, pp.447-460, January. 2005.
- [9] J. Garcia-Frias, "Compression of correlated binary sources using Turbo codes," *Communications Letters, IEEE* Vol.5, Issue 10, October, 2001.
- [10] A. Liveris, Z. Xiong, and C. Georgiades, "Compression of binary sources with side information at the decoder using LDPC codes," *IEEE Commun. Lett.*, vol.6, no.10, pp.440-442, October. 2002.
- [11] A. Aaron, S. Rane, E. Setton and B. Girod, "Transform-domain Wyner-Ziv codec for video," in: *SPIE Visual Communications and Image Processing Conference*, vol.5308, pp.520-528, San Jose, CA, 2004.
- [12] A. Aaron, R. Zhang and B. Girod, "Wyner-Ziv coding of motion video," in: *Proceedings of Asilomar Conference on Signals and Systems*, Pacific Grove, CA, November 2002.
- [13] R. Puri and K. Ramchandran, "PRISM: A new robust video coding architecture based on distributed compression principles," in *Proc. Allerton Conference on Communication, Control, and Computing*, Allerton, IL, October 2002.
- [14] <http://www.discoverdvc.org>
- [15] 박세환, 송혁, 장은영, 허남호, 김진웅, 김진수, 이상훈, 유지상, "다시점 동영상에서 임의시점영상 생성을 위한 가려진 영역 보상기법," *한국통신학*

회논문지 제 33권, 제 12호, pp.1029-1038, 12월 2008.

- [16] ITU-T, I. JTC1, "Digital compression and coding of continuous-tone still images," ISO/IEC 10918-1 - ITU-T Recommendation T.81 (JPEG).
- [17] D. Varodayan, A. Aaron, B. Girod, "Rate-adaptive distributed source coding using low-density parity-check codes," Signals Systems and Computers Conference Record of the Thirty-Ninth Asilomar Conference, November 2005.

유 지 환 (Jihwan Yoo)

준회원



2005년 2월 광운대학교 전자공학부 졸업
 2008년 3월~현재 광운대학교 전자공학과 석사과정
 <관심분야> Video coding, Free-view point system, GPGPU

이 동 석 (Dong-seok Lee)

준회원



2009년 2월 광운대학교 전자공학과 졸업
 2009년 3월~현재 광운대학교 전자공학과 석사과정
 <관심분야> Video Coding, Image Processing

김 태 준 (Tae-june Kim)

준회원



2008년 2월 광운대학교 전자공학과 졸업
 2008년 3월~현재 광운대학교 전자공학과 석사과정
 <관심분야> Stereo Matching, Computer Vision, 동영상 코덱

유 지 상 (Ji-sang Yoo)

종신회원



1985년 2월 서울대학교 전자공학과 졸업(공학사)
 1987년 2월 서울대학교 대학원 전자공학과 졸업(공학석사)
 1993년 5월 Purdue 대학교 전기공학과 졸업(Ph.D.)
 1993년 9월~1994년 8월 현대 전자산업(주) 산전연구소 선임연구원
 1994년 9월~1997년 8월 한림대학교 전자공학과 조교수
 1997년 9월~2001년 8월 광운대학교 전자공학과 조교수
 2001년 9월~2006년 8월 광운대학교 전자공학과 부교수
 2006년 9월~현재 광운대학교 전자공학과 정교수
 <관심분야> 웨이블릿 기반 영상처리, 영상압축, 영상인식, 비선형 신호처리