

특집논문-04-09-3-05

이질적인 환경을 위한 스케러블 스테레오 영상 부호화

오 세 찬^{a)}, 이 영 호^{a)}, 우 운 택^{a)†}

Scalable Stereoscopic Video Coding for Heterogeneous Environments

Sehchan Oh^{a)}, Youngho Lee^{a)} and Woontack Woo^{a)†}

요 약

본 논문에서는 이질적인 수신 단말을 위해 시공간적 계위(spatio-temporal scalability)를 이용한 새로운 3차원 영상 부호화 방법을 제안한다. 제안된 방법은 좌 영상의 부호화를 위해 MPEG-2 표준을 이용하며, 우 영상의 부호화를 위해서 향상된 호환성(enhanced compatible) 부호화 방법을 이용하여 P-와 B-타입의 영상을 예측한다. 향상된 호환성 부호화 방법은 두 개의 전후방향 움직임으로부터 예측된 매크로 블록과 양안차로부터 예측된 블록을 보간하여 매칭 블록을 예측한다. 유연성 있는 3차원 비디오 서비스를 위해서, 각각의 좌우 영상에 대해 시공간적 확장 가능한 계층을 정의하였다. 실험을 통하여 제안된 3차원 영상 부호화기의 성능을 기존의 방법들과 비교하여 검증하였으며, 양안차 예측의 이점을 계위에 따르는 오버헤드의 관점에서 분석하였다. 제안된 시공간적 계위를 이용한 3차원 영상 부호화기는 다양한 디스플레이 장치 및 네트워크 환경을 가지는 유비쿼터스 컴퓨팅 환경을 위한 효율적인 3차원 비디오 전송 서비스에 활용될 것으로 기대된다.

Abstract

In this paper, we propose a new stereoscopic video coding approach for heterogeneous consumer devices by exploiting the concept of spatio-temporal scalability. The proposed method uses MPEG-2 standard for coding the left or main sequence and an enhanced compatible coding scheme for predicting the P- and B-type of frames of the right or auxiliary sequence. The enhanced compatible coding scheme predicts matching block by interpolating both two forward and backward motion predicted macroblocks and disparity predicted macroblock. To provide flexible stereo video service, we define both a temporally scalable layer and a spatially scalable layer for each eye-view. The experimental results show the efficiency of proposed coding scheme by comparison with already known methods and the advantages of disparity estimation in terms of scalability overhead. According to the experimental results, we expect the proposed functionalities will play a key role in establishing highly flexible stereo video service for ubiquitous computing environment where devices and network connections are heterogeneous.

Keywords : Stereoscopic video coding, Spatio-temporal scalability, Enhanced compatible coding scheme, Flexible stereo video service.

1. 서 론

최근, 인터넷 및 멀티미디어 서비스의 발전과 더불어 멀

리 떨어져 있는 사용자들 사이에 3차원 및 파노라마 영상과 같은 고화질의 실감미디어를 제공하는 협업 환경이 가능해 졌다. 하지만, 3차원 영상과 같은 실감 미디어의 가용성은 전적으로 네트워크 인프라 및 수신 측 디스플레이 환경에 의존한다. 그림 1과 같이, 현재 네트워크 인프라 및 수신 단의 디스플레이 장치들은 상당히 이질적이다. 따라서

a) 컴퓨터공학부 정보통신공학과 U-VR 연구실
GIST U-VR Lab

† 본 연구는 대학 IT 연구센터 육성·지원사업의 연구결과로 수행되었음.

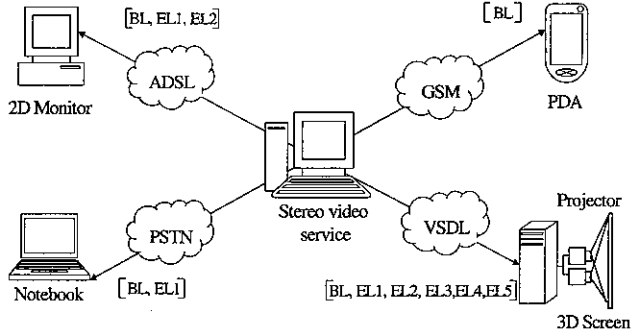


그림 1. 이질적인 네트워크 및 디스플레이 시스템에서의 3차원 비디오 서비스
Fig. 1. Heterogeneous network and display systems with 3D video service

다양한 네트워크 환경과 수신 단말의 처리 능력 및 디스플레이 시스템을 고려한 새로운 3차원 영상 압축기법이 필요하다.

본 논문에서는 이질적인 수신 단말들 간에 효율적인 3차원 영상의 디스플레이 환경을 제공하도록, 시공간적 계위를 이용한 새로운 3차원 영상 부호화 방법을 제안한다^{[1][2]}. 제안된 방법은, 기준 영상(좌 영상)의 부호화를 위해서 MPEG-2 표준을 사용하였다. 또한, 우 영상의 P-픽처와 B-픽처의 효율적인 예측을 위해서 MPEG-2 멀티뷰 프로파일에서 정의한 compatible 방식을 개선한 형태의 DV (Disparity Vector)와 MV(Motion Vector)를 이용한 보간 기법을 사용한다. 각각의 B-픽처(Bidirectionally Predictive-coded Picture)의 매크로블록(MacroBlock: MB)은 I-픽처(Intra-coded Picture), P-픽처(Predictive-coded Picture)로부터 구한 전후 방향의 움직임 예측과 더불어, 시간적으로 동일한 기준 영상으로부터 양안차 예측으로 얻어진 MB를 보간 함으로써 예측된다. 마찬가지로 P-픽처의 경우, 전방향 움직임 예측과 양안차 예측으로 얻어진 MB를 보간 함으로써 예측된다. 이때, 움직임 예측으로부터 얻어진 MB와

양안차 예측으로부터 얻어진 MB는 동일한 가중치를 두어 보간 한다.

또한, 제안된 방법은 이질적인 수신 단말들의 효율적인 디스플레이 환경을 제공하도록, 시공간적 계위를 이용하여 3차원 영상의 부호화를 수행한다. 효율적인 3차원 영상의 전송과 재현을 위해서, 하나의 기본계층(Base Layer: BL)과 여러 향상계층(Enhancement Layer: EL)을 제공한다. BL은 입력된 기준영상(좌 영상)에 대한 기본 해상도를 위한 계층이며, EL은 기본 해상도의 우 영상과 더불어 고 해상도의 좌우 영상을 제공하기 위한 추가적인 정보를 제공한다. 일반적으로 계위는 시공간적 해상도의 유연성을 제공하는 반면, 오버헤드를 가지고 있다. 3차원 영상 부호화에서는 양안차 예측을 수행함으로써 2차원 영상에 비해, 이러한 오버헤드를 크게 줄일 수 있다.

본 논문의 구성은 다음과 같다. 2 장에서는 관련연구를 언급한 후, 3 장에서는 제안된 3차원 영상 부호화 방법과 시공간적 계위에 대해 설명한다. 4 장에서는 실험을 통하여 제안된 시스템을 분석하며, 5장에서는 개선 및 보완되어야 할 기술적인 문제점에 대해 간략하게 기술한다.

II. 관련 연구

일반적으로 3차원 영상의 부호화 방법은 그림 2와 같이, 좌우 영상의 상관관계의 정도에 따라서 (1) simulcast stereoscopic coding, (2) compatible stereoscopic coding 그리고 (3) joint stereoscopic coding으로 나뉜다^[3].

압축 효율의 관점에서, 좌우 영상이 서로 참조되어 부호화 되는 joint stereoscopic coding은 상대적으로 가장 높은 압축효율을 갖는다. 하지만, 부호화에 대한 어떠한 독립성도 가지지 않기 때문에 기존 시스템과 호환성 부족이라는

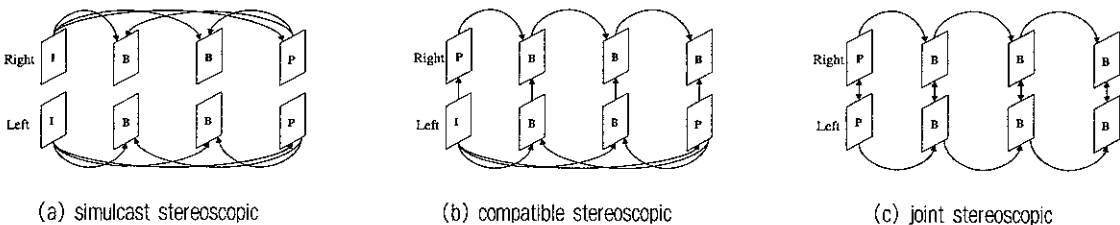


그림 2. 3차원 영상 부호화 방법
Fig. 2. Stereoscopic coding approaches

문제점을 가진다. 반면, compatible stereoscopic coding 기법은 압축 효율과 호환성이 적절히 절충된 부호화 기법이며, 대부분의 3차원 영상 부호화 기법에서 사용된다.

전형적인 3차원 영상 부호화는 각각이 입력 영상에 대한 효율적인 움직임 예측 및 보상과 더불어 좌우 영상 사이에 양안차의 예측 및 보상으로 이루어진다. 기본적으로, 3차원 영상 압축에서 움직임과 양안차 예측 및 보상을 위한 구현은 밝기 값 처리(intensity processing)에 의한 방법과 객체 조작(object handling)에 의한 방법이 있다^[4].

밝기 값 처리에 기반 한 방법 중 가장 일반적인 기법은 MPEG 표준의 움직임 예측에서 사용한 블록 매칭 알고리즘(Block Matching Algorithm: BMA)이다^{[5][6][7]}. BMA는 하나의 프레임을 여러 고정된 크기의 블록으로 나누고 기준 블록과 가장 유사한 블록을 찾는다. 한편, 벡터 탐색의 복잡도와 정확도를 높이기 위한 다 해상도 기반의 양안차 예측 기법과, 가변 블록 기반의 양안차 예측 기법에 대한 연구 등이 진행되고 있다^{[8][9][10]}.

객체 조작에 기반한 움직임 및 양안차 예측 기법은 우선, 처리과정을 거친 장면에서 해당 객체의 특징을 추출하거나 정의한 후, 좌우 영상으로부터 움직임 또는 양안차를 예측한다^[11]. 이러한 기법은 BMA에 비해서 예측의 정교함은 있지만, 계산량이 많아지게 된다.

이들과는 별개로, BMA과 object 기반을 혼합한 방법,

OBDC(Overlapped Block Disparity Compensation), 블록기반 양자화를 위한 비트율-왜곡 최적화 기법(Optimal Rate-Distortion Based Algorithms for Block Dependent Quantization), 움직임/양안차 필드 예측(Motion/Disparity Field Estimation)등의 연구도 있다^{[12][13][14][15]}.

최근, 3차원 비디오에 여러 해상도를 제공하기 위한 연구가 진행되고 있다^[16]. 하지만 이들 기법은 3차원 비디오에 단일 계위만을 결합한 수준에 그쳐서, 표현 가능한 해상도가 제한되어 있다. 따라서 이질적인 네트워크 환경 및 수신측 디스플레이를 고려하여 3차원 비디오에 여러 시공간적 해상도를 제공할 수 있는 3차원 비디오 압축 기법에 관한 연구는 아직까지 미흡한 실정이다.

III. 시공간적 계위를 이용한 스테레오 비디오 복부호화 시스템

이질적인 디스플레이 환경을 가지는 시스템들 간에 효율적인 3차원 영상 전송을 위하여, 그림 3과 같이 하나의 BL 이외에 추가로 여러 EL을 정의한다. BL, EL3는 각각 좌우 영상의 기본 해상도에 해당하는 비트열이며, EL1과 EL4는 시간적으로 고해상도의 기준영상을 제공하기 위해 추가적인 비트열이다. 또한, EL2와 EL5는 공간적 고해상도를 제

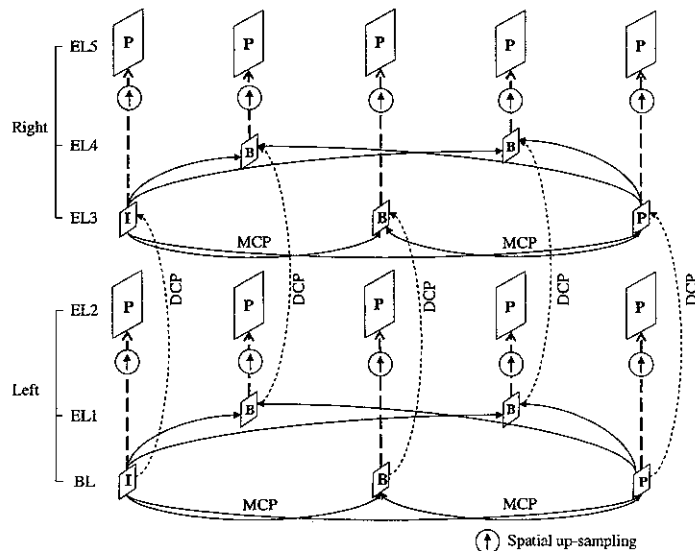


그림 3. 3차원 비디오와 시공간적 계위 구성
Fig. 3. Structure of spatio-temporal scalability with stereoscopic video

공하기 위한 추가적인 비트 열을 표현한다. BL에서는 시간적으로 근접한 자신의 계층에 존재하는 프레임으로부터 움직임보상예측 (Motion Compensation Prediction: MCP)을 기반으로 부호화가 수행되며, EL2와 EL5는 고해상도의 좌우 영상을 제공하기 위한 추가적인 데이터를 다른 예측과 정 없이 인트라 부호화 과정을 수행한다. EL3는 보다 높은 압축 효율을 위해서 시간적으로 일치하는 BL의 영상으로부터 양안차보상예측 (Disparity Compensation Prediction: DCP)과 함께, 시간적으로 근접한 자신의 계층에 존재하는 영상들로부터 MCP를 수행한다. 마찬가지로, EL1과 EL4는 시간적 고해상도의 좌우 영상을 제공하기 위한 추가적인 데이터를 표현한다.

제안된 시스템의 3차원 영상 복부호화는 compatible 방법의 확장된 버전으로, 크게 (1) 3차원 영상 복부호화 (2) 공간적 계위 복부호화와 (3) 시간적 계위 복부호화로 나뉜다.

1. 3차원 영상 복부호화

제안된 시스템의 3차원 비디오 부호화기의 구조는 그림 4(a)와 같다. 3차원 비디오 부호화는 한 쌍의 스테레오 영상이 양안차를 따라 유사성이 있다는 점을 이용한다. 우 영상의 압축을 위해서는 현재의 영상과 시간적으로 일치하는 좌 영상을 참조하여 양안차 벡터를 구한다. 구한 움직임 벡터 또는 양안차 벡터 중 전송효율이 높은 쪽을 선택한 후, 그에 따라서 움직임 차이 정보(Displaced Frame Difference: DFD) 혹은 양안 차이정보 (Displacement Compensated Difference: DCD)를 부호화하여 전송한다. 해당 이미지에 서 하나의 픽셀에 대한 DCD는 수식 (1)과 같이 정의된다.

$$DCD(x, y, dv) = I_R[x][y] - I_L[x + dv][y] \quad (1)$$

위 식에서 IR과 IL은 각각 우 영상과 좌 영상에서의 픽셀 밝기 (intensity) 값이며, dv는 양안차 벡터이다. dv는 수식 (2)와 같이 정의된다.

$$dv(x, y) = \underset{dv \in S}{\text{arg min}} DCD(x, y, dv) \quad (2)$$

위 식에서 S는 윈도우 탐색 영역을 의미한다. 실제 영상에서 양안차 벡터를 구하는 방법은 수식 (3)처럼 각 블록

에 대해 SAD(Sum of Absolute Difference)가장 작은 블록을 매칭 블록으로 정의한다. 양안차 벡터는 SAD 값이 최소가 되는 좌표 (x, y)로 정의한다. 이론적으로 양안차 벡터는 x축 방향의 모션 벡터이나, 실제 카메라의 좌우 렌즈는 정확한 수평을 유지하기 어렵다. 따라서 실제 y는 '0'에 근사한 값을 갖게 된다.

$$dv(x, y) = \underset{dv_{\min}}{\min} \underset{dv_{\max}}{dv} \sum_{i=-S/2}^{S/2} \sum_{j=-S/2}^{S/2} | I_R[x + i][y + j] - I_L[x + i + dv][y + j] | \quad (3)$$

제안된 3차원 비디오 부호화기의 부호화 방법의 특징은 아래와 같다.

- 기준영상(좌 영상)은 MPEG-2 비계위(non-scalable) 부호화기로 부호화한다.
- 우 영상에서의 I-픽처는 시간적으로 동일한 좌 영상으로부터 구한 양안차 벡터를 이용하여 예측된 블록을 이용하여 부호화한다.
- 우 영상에서의 P-픽처는 동일한 시퀀스 내에서 순방향 예측(forward prediction)을 하여 얻어진 움직임 벡터와 시간적으로 동일한 좌 영상을 참조하여 얻어진 양안차 벡터, 그리고 이 둘을 보간하여 얻어진 보간 벡터(interpolated vector)를 이용하여 얻어진 블록의 SSD(Sum of Squared Difference: SSD) 중에서 최소값을 가지는 블록을 이용하여 부호화한다.
- 우 영상에서의 B-픽처는 그림 5에서와 같이 동일한 시퀀스 내에서 순방향 움직임 예측 및 역방향 움직임 예측을 하여 얻어진 MB와 양안차 예측을 통해서 얻어진 MB를 이용하여 부호화한다. 이때, SSD에 따라서 순방향 또는 역방향 인지, 양안차 예측 또는 이 둘 벡터를 보간한 예측인지를 결정한다.

우 영상에서 I-픽처의 경우, 각 MB마다 기준 블록과 가장 유사한 블록을 탐색하기 위해 SAD 뿐만 아니라 SSD를 구한다. 이때 구해진 SSD의 크기와 정해진 문턱값 (threshold value)과 비교하여 인트라 부호화를 수행 할 것인지 양안차를 이용한 부호화를 수행 할 것인지를 판단하게 된다.

P와 B픽처의 경우, 선택된 전후방향 움직임 예측된 프레임 (motion predicted frame)과, 양안차 예측된 프레임 (disparity predicted frame)을 각각 50%의 가중치를 두어 보간된 결과 영상을 이용하여 부호화한다. 예측된 P-픽처는 수식 (4)와 같이 표현 될 수 있다.

$$P_{pred}(v_f, v_d) = w_f R_{rec f}(v_f) + w_d R_{rec d}(v_d) \quad (4)$$

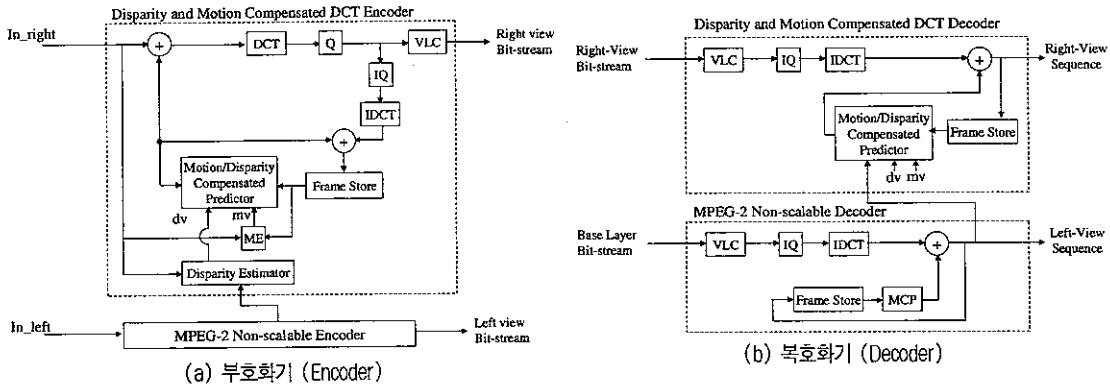


그림 4. 3차원 비디오 복부호화기
Fig. 4. Stereoscopic video codec

위 식에서 P, v, w 는 각각 MB, 벡터, 가중치를 나타낸다. 그리고 인덱스 f, d 는 각각 순방향, 양안차를 나타내며, R_{rec} 는 복호화 된(reconstructed) 픽처의 MB를 나타낸다. B-픽처의 경우, 보간벡터를 이용하여 예측된 픽처는 수식 (5)와 같이 표현 가능하다.

$$B_{pred}(v_f, v_b, v_d) = w_f R_{rec}(v_f) + w_b R_{rec}(v_b) + w_d R_{rec}(V_d) \quad (5)$$

마찬가지로, 위 식에서 B, v, w 는 각각 MB, 벡터, 가중치를 나타낸다. 마찬가지로, 인덱스 f, b, d 는 각각 순방향, 역방향, 양안차를 나타낸다.

그림 4(b)는 3차원 비디오 복호화기 구조를 보여준다. EL3에서 우 영상은 시간적으로 일치하는 BL의 좌 영상을 참조하여 기본 해상도의 우 영상을 복호화한다. 부호화기에서와 마찬가지로 양안차 벡터를 이용하여 복원된 좌 영상으로부터 양안차 보정을 수행하며, 움직임 벡터를 이용하여 시간적으로 이웃하는 우 영상으로부터 움직임 보정을 함께 수행함으로써 우 영상을 복원할 수 있다.

2. 공간적 계위 복부호화

공간적 계위 부호화기는 그림 6(a)와 같이, 기본 해상도를 가지는 기준영상에 고해상도를 지원하기 위한 추가적인 비트 열을 생성한다. BL 부호화기에 입력된 원 영상은 각각 수직 수평 방향으로 하향 샘플링 된 후, 기존의 MPEG-2 부호화기를 이용하여 부호화된다. 한편, 향상계층

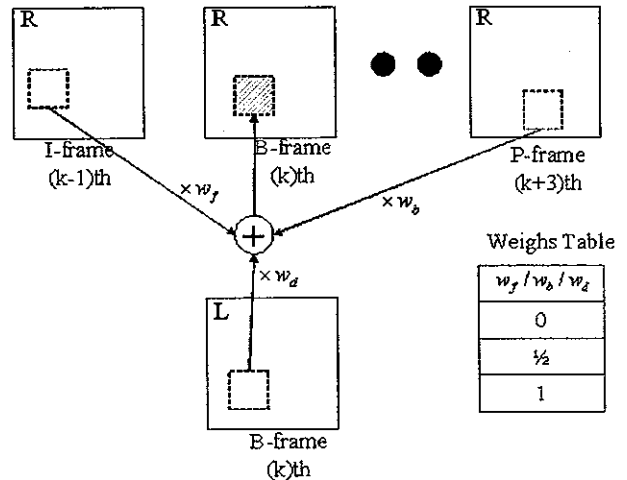


그림 5. 우 영상에서의 B-frame의 예측, $M=4$
Fig. 5. Prediction of B-frame in right sequence, $M=4$

은 원 영상과 복호화 된 BL 영상을 다시 원 영상의 크기로 상향 샘플링 한 영상과의 차분 영상(residual image)을 부호화 한다. 차분 영상은 일반적인 영상의 특성을 가지지 않는다. 따라서 움직임 예측이나 움직임 보상의 수행 없이 인트라 부호화를 수행한다.

공간적 계위 복호화 과정은 그림 6(b)와같이 부호화 과정의 반대이다. EL의 영상을 복호화하기 위해서는 시간적으로 일치하는 BL 영상의 복호화가 선행되어야 한다. 복호화된 BL의 좌(우) 영상은 기본 해상도에 해당하는 디스플레이를 가지는 수신 단에서 그대로 재현이 가능하다. 하지만, 고해상도의 디스플레이 시스템을 갖춘 수신측 디스플레이

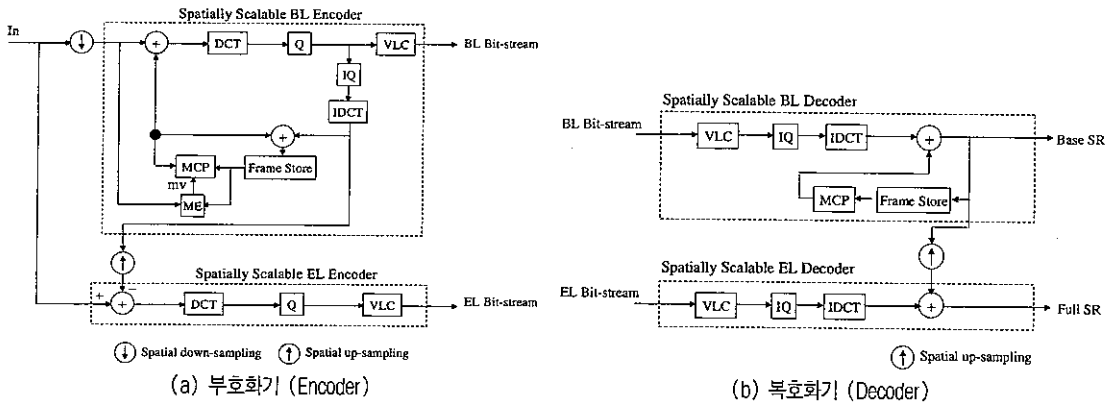


그림 6. 공간적 계위 복부호화기
Fig. 6. Spatial scalability codec

이를 위해서는 원 영상의 크기로 복원한 후, 고해상도를 위한 추가적인 정보인 상향계층의 영상과 결합하여 디스플레이 한다.

3. 시간적 계위 복부호화

시간적 계위 부호화기는 해당 네트워크 허용 대역폭에 맞도록 프레임 레이트(frame rate)를 변화시킴으로써 때로는 부드러운 움직임 혹은, 때로는 거친 움직임을 갖도록 할 수 있다. 일반적으로 시간적 계위에서 2 단계의 서로 다른 해상도 즉, BL과 하나의 EL을 제공하기 위해서는 I-픽처와 P-픽처 사이에 홀수개의 B-픽처가 필요하다^[17]. 시간적 계위 부호화기는 그림 3과 같이, B-픽처를 기본계층과 향상계층으로 분할한다. 분할된 계층의 비트 열은 3차원 비디오 서비스를 담당하는 서버로부터 네트워크 및 수신 단말의 상태에 따라서 다른 시간적 해상도를 가지는 3차원 영상을 제공 받을 수 있다.

시간적 계위 복호화기는 수신 단말기의 네트워크 및 디스플레이 환경에 따라 제공받은 시간적 계위 상향계층의 비트열을 결합하여 최종 디스플레이 되어질 3차원 영상의 시간적 해상도를 결정한다.

IV. 실험 및 비교

실험에 사용된 스테레오 영상은 순차방식의 640×480의 해상도를 가지는 컬러 영상이며, 초당 24프레임의 프레임 레이트를 가진다. 또한 YUV 비율은 4:2:0을 사용하였다. GOP는 16 프레임이며, B-픽처의 개수는 GOP (Group Of Picture)의 75%를 차지한다. 그림 7(a), 7(b)와 7(c), 7(d)는 각각 카메라로부터 얻은 테스트 시퀀스인 Laboratory와 UbiHome 시퀀스 중에서 임의의 프레임에 해당하는 각각의 좌우 영상을 보여준다. 시퀀스 Laboratory는 움직임이 거의 없는 반면, UbiHome는 카메라 뷰의 이동을 비롯하여

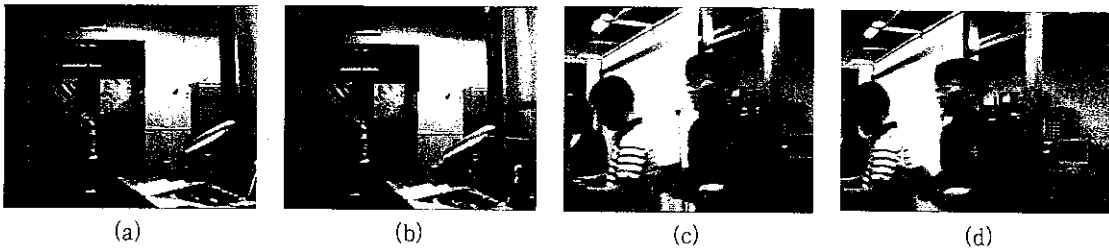


그림 7. 테스트 시퀀스. (a),(b)Laboratory, (c),(d)UbiHome
Fig. 7. Test sequences. (a),(b)Laboratory, (c),(d)UbiHome

표 1. 테스트 시퀀스 Laboratory에 대한 프로그램 수행 결과

Table 1. The experimental results for the test sequence Laboratory

	Decoded layers	BL+EL1 (320×240)	BL+EL1+EL2 (640×480)	EL3+EL4 (320×240)	EL3+EL4+EL5 (640×480)
		Average MSE for a GOP	8.22	6.05	10.35
Simulcast stereo coding	Average PSNR for a GOP [dB]	38.98	40.31	37.98	38.86
	Mean PSNR for left and right [dB]	39.53			
Proposed stereo coding	Average MSE for a GOP	8.22	6.05	5.65	6.51
	Average PSNR for a GOP [dB]	38.98	40.31	40.61	39.99
	Mean PSNR for left and right [dB]	40.15			

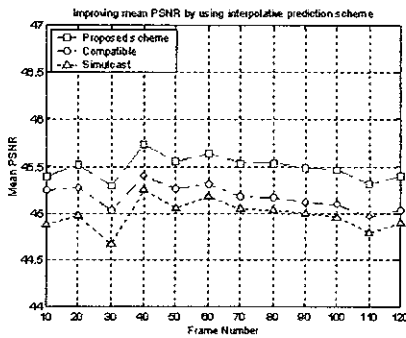
움직임이 많다.

표 1에서는 제안된 스테레오 영상 복부호화기를 이용하여 부호화를 수행한 결과를 보여준다. 좌우 영상에 할당된 목표 비트율은 각각 1 Mbps이며, 좌우 기준 영상 ([BL+EL1]과 [EL3+EL4])의 목표 비트율은 전체 할당된 목표 비트율의 40%이다. 복호화된 기본해상도의 영상을 본래의 해상도로 보간 한 후에 공간적 EL의 추가적인 정보를 더하여 고해상도의 영상을 얻을 수 있다. 실험에 사용된 테스트 시퀀스 Laboratory는 같은 목표 비트율을 가지고 각각 독립적으로 부호화를 수행할지라도 좌우 영상의 차이 때문에, 좌 영상의 PSNR 값이 약 1 dB 정도 높다. 하지만 우 영상 부호화에 양안차 예측 및 보상을 수행하였을 때, 오히려 우 영상의 PSNR 값이 좌 영상 보다 약 1.6 dB 정도 높음을 알 수 있다. 우 영상의 향상계층과 결합하여 생성된 영상 (EL3+EL4+EL5) 또한 독립적으로 부호화 한 경우보다 약 1 dB정도 높은 화질의 고해상도 영상을 얻을

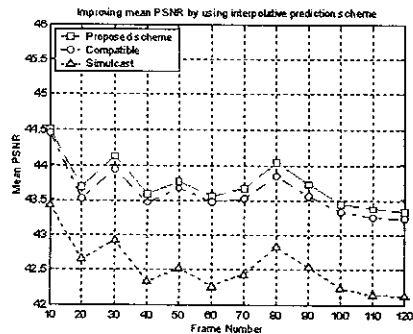
수 있다.

그림 8은 제안된 보간 기법을 이용한 3차원 부호화 방법과 각각의 영상을 독립적으로 부호화하는 simulcast 방식, MPEG-2의 멀티뷰 프로파일에 정의된 compatible 방식의 3차원 영상 부호화 방법의 평균 PSNR값을 비교한 것이다. 그림에서 알 수 있듯이, 테스트 영상 Laboratory와 UbiHome 모두 제안된 보간 기법 방식이 나은 PSNR 값을 가진다.

표 2에서는 실험 결과를 비계위 MPEG-2 부호화기와 비교한 값을 보여준다. 공간적 계위에 대한 오버헤드는 제안된 부호화기가 MPEG-2 비계위 부호화기를 통해 얻어진 영상과 같은 화질의 영상을 얻기 위해서 추가적으로 필요한 비트율로 정의 할 수 있다. 제안된 부호화기는 DCP를 이용한 3차원 비디오 부호화기에 공간적 계위를 결합하여 공간적 계위에 필요한 오버헤드를 줄일 수 있다. 움직임이 비교적 적은 Laboratory의 경우 대부분의 블록마다 변화량



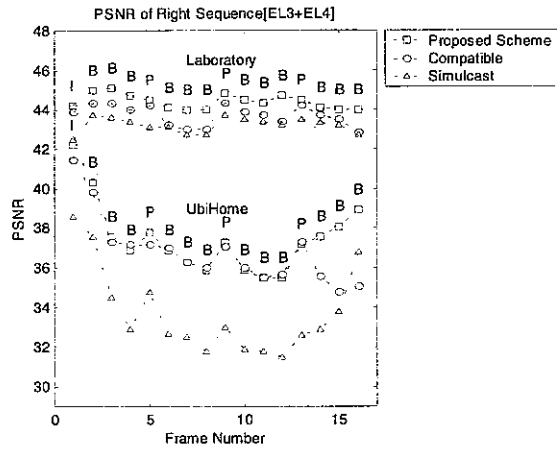
(a) Laboratory



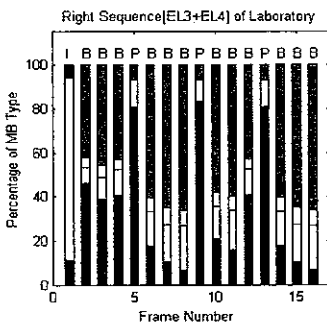
(b) UbiHome

그림 8. 기본해상도의 우 영상에 해당하는 평균 PSNR

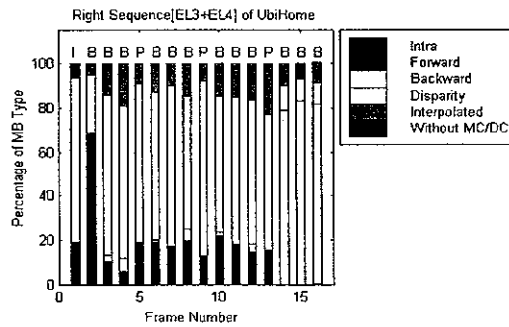
Fig. 8. Average PSNR for the right sequence with lower spatial resolution



(a) PSNR comparison



(b) Percentage of MB types for Laboratory

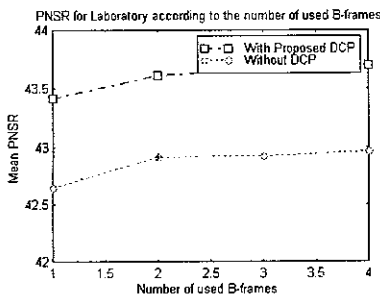


(c) Percentage of MB types for UbiHome

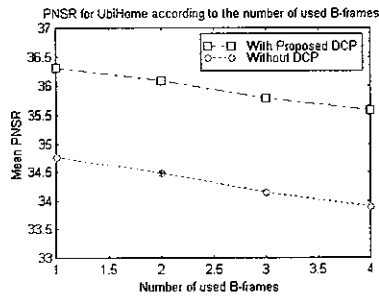
그림 9. 움직임 정도에 따른 양안차 예측 정도와 부호화 효율성
Fig. 9. The Number of Disparity Estimated Macroblocks According to the Motion Rate

이 적은 움직임 벡터를 선택하기 때문에 DCP로 얻어지는 압축 효율이 적다. 반면, UbiHome 경우는 블록마다 움직임

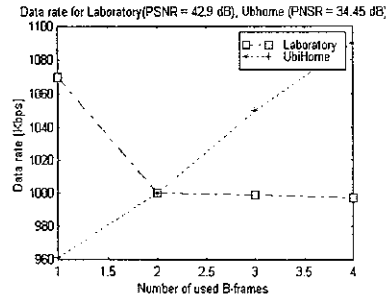
변화량이 크기 때문에 MCP 대신, DCP를 이용하여 부호화를 수행하게 된다. 따라서 비계위 부호화기와 비교하였을



(a) PSNR for Laboratory



(b) PSNR for UbiHome



(c) Data rates for fixed PSNR

그림 10. 시간적 계위에 따른 오버헤드
Fig. 10. Temporal scalability overhead

표 2. 공간적 계위 오버헤드

Table 2. Spatial scalability overhead.

	Test sequences	Laboratory		UbiHome	
		Left	Right	Left	Right
Non-scalable MPEG-2 encoder	Bit-rate [Mb]	1.0	1.0	1.0	1.0
	Average PSNR for luminance [dB]	41.03	40.25	36.50	36.33
	Mean PSNR for left and right [dB]	40.62		36.42	
Proposed scalable stereoscopic encoder	Bit-rate [Mb]	1.16	1.16	1.045	1.045
	Average PSNR for luminance [dB]	40.89	40.28	36.83	36.07
	Base layer bit-rate [Mb]	0.46	0.46	0.41	0.41
	Base layer bit-rate as percent of total bit-rate [%]	39.7	39.7	39.2	39.2
	Spatial scalability overhead [%]	16		4.5	
	Mean PSNR for left and right [dB]	40.63		36.44	

때, 얻어지는 압축 효율은 상당히 크며, 결과적으로 공간적 계위에 따르는 오버헤드를 줄일 수 있음을 알 수 있다.

그림 9는 공간적 계위 BL에 해당하는 우 영상에서 움직임 변화에 따르는 DCP의 정도와 압축 효율성을 보여준다. 그림 9(a)는 GOP에 해당하는 영상을 각각 simulcast 방식, compatible 방식, 제안된 방식을 이용하여 수행한 경우에 대한 PSNR 값을 보여준다. 또한, 그림 9(b)와 9(c)는 제안된 방식에서 각 픽처의 블록이 어떠한 형태로 부호화를 수행하였는지 퍼센트로 보여준다. 움직임 변화량이 큰 UbiHome의 경우에는 대부분의 블록이 DCP 통해서 부호화되었으며 얻어지는 압축효율도 양안차 벡터를 참조한 블록의 수에 비례한다.

그림 10은 시간적 계위에 따르는 오버헤드를 보여준다. 전체 해상도와 1/2 해상도를 지원하기 위해서는 I 픽처 또는 P 픽처 사이에 홀수개의 B 픽처가 존재해야한다. 하지만, 그림 10(a)와 10(b)에서 볼 수 있듯이 일반적으로 B 픽처의 개수가 2개인 경우 전체적으로 안정된 성능을 보인다. 그림 10(c)에서는 각각의 시퀀스에 대해 같은 화질의 영상을 얻기 위해서 필요한 data rate를 B 픽처의 개수에 따른 변화를 보인다. 두 시퀀스를 고려해 볼 때 B 픽처가 2개인 경우 가장 적은 data rate를 가지며 B 픽처가 3개인 경우에는 UbiHome에서만 약 5%의 오버헤드를 가진다. 하지만, 그림 10(a), 10(b)에서 볼 수 있듯이, 제안된 부호화기의 DCP를 이용할 경우 시간적 계위에 따르는 오버헤드는 제거될 수 있다.

V. 결론 및 추후연구

본 논문에서 제안한 시스템은 이질적인 네트워크 환경

및 디스플레이 환경을 가지는 수신 단말들 사이에 효율적인 3차원 영상 서비스를 위해서 시공간적 계위를 이용한 3차원 영상 부호화 방법을 제안하였다. 또한, 다양한 실험을 통해서 제안된 시스템의 효율성을 보였으며, 구현된 부호화 및 복호화 시스템은 스테레오 영상에 대해서 여러 시간적, 공간적 해상도를 제공하여 이질적인 수신단말들 사이에 여러 형태의 3차원 영상 서비스가 가능하도록 하였다. 또한, 실험을 통하여 제안된 3차원 영상 부호화기의 성능을 기존의 방법들과 비교하여 검증하였으며, 양안차 예측의 이점을 계위에 따르는 오버헤드의 관점에서 분석하였다. 제안된 스테레오 영상 부호화 방법은 유비쿼터스 컴퓨팅 환경에서 3차원 영상 서비스의 한계를 극복하는 데 기여할 수 있을 것으로 판단된다. 앞으로 공간적 계위 부호화기에 입력 영상인 차분 값을 위한 새로운 양자화기 개발과 더불어 보다 효율적인 양안차 예측의 개선이 이루어진다면 향상된 압축 효율을 기대할 수 있다.

참고문헌

- [1] M.Domanski, S.Mackowiak, "Modified MPEG-2 video coders with efficient multi-layer scalability," Proc. of ICIP, vol. 2, pp. 1033-36, Oct. 2001.
- [2] M.Naroschke, "Functionalities and Costs of Scalable Video Coding for Streaming Services", Proc. of Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, Nov. 2002.
- [3] B. Haskell, A. Puri, A. N. Netravali, "Digital Video: An Introduction to MPEG-2," New York: Chapman & Hall, 1997.
- [4] R.Shukla, H.Radha, "Disparity Dependent Segmentation based Stereo Image Coding,"Proc. of ICIP, Sept. 2003.
- [5] A.Puri, R.V.Kollarits, and B.G.Haskell, "Basics of Stereoscopic Video, New Compression Results with MPEG-2 and a Proposal for

- MPEG-4," Proc. of Image Communications, vol. 10, pp. 201-234, 1997.
- [6] W.Woo, A.Ortega, "Overlapped Block Disparity Compensation with Adaptive Windows for Stereo Image Coding," Proc. of IEEE Tr. Circuits Syst. Video Tech. vol. 10, pp. 194-200, Mar. 2000.
- [7] A.Mancini, J.Konrad, "Robust Quadtree-Based Disparity Estimation for the Reconstruction of Intermediate Stereoscopic Images," Proc. of SPIE Stereoscopic Displays and Virtual Reality Systems, vol. 3295, pp. 53-64, Jan. 1998.
- [8] M.W.Siegel, S.Sethuraman, J.S.McVeigh, and A.G.Jordan, "Compression and interpolation of 3D-stereoscopic and multi-view video," Proc. of SPIE, vol. 3012, pp. 227-238, Feb. 1997.
- [9] S.Sethuraman, M.W.Siegel, A.G.Jordan, "A Multiresolution Framework For Stereoscopic Image Sequence Compression," Proc. of ICIP, vol. 2, pp. 361-365, Nov. 1994.
- [10] S.Sethuraman, A.G.Jordan, M.W.Siegel, "Multiresolution Based Hierarchical Disparity Estimation for Stereo Image Pair Compression," Proc. of the Symposium on Application of subbands and wavelets, 1994.
- [11] D.Tzovaras, N.Grammalidis, and M.G.Strintzis, "Object-Based Coding of Stereo Image Sequences using Joint 3-D Motion/Disparity Compensation," Proc. of IEEE Tr. on Video Technology, vol. 7, no. 2, pp. 312-328, April 1997.
- [12] J.Jiang and E.A.Edirisinghe, "A Hybrid Scheme for Low Bit-rate Coding of Stereo Images," Proc. of IEEE Tr. Image, vol. 11, no. 2, pp. 123-134, Feb. 2002.
- [13] W.Woo, A.Ortega, "Overlapped Block Disparity Compensation with Adaptive Windows for Stereo Image Coding," Proc. of IEEE Tr. Circuits Syst. Video Tech. vol. 10, pp. 194-200, Mar. 2000.
- [14] W.Woo, A.Ortega, "Optimal Blockwise Dependent Quantization for Stereo Image Coding," Proc. of IEEE Tr. CSVT, vol. 9, pp. 861-867, Sept. 1999.
- [15] D.Tzovaras, M.G.Strintzis, "Motion and Disparity Field Estimation using Rate-distortion Optimization," Proc. of IEEE Tr. on Video Technology, vol. 8, pp. 171-180, April 1998.
- [16] Woon-Tack Woo, "http://vr.gist.ac.kr/~3D/," 2001.
- [17] M.Naroschke, "Functionalities and Costs of Scalable Video Coding for Streaming Services," Proc. of Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, Nov. 2002.

저 자 소 개



오 세 찬

- 2002년 : 강원대학교 정보통신공학과(학사)
- 2002년~현재 : 광주과학기술원 정보통신공학과 석사과정
- 주관심분야 : 신호처리, 3차원 비디오, HCI, 컴퓨터 비전, 혼합현실



이 영 호

- 1999년 : 한국과학기술원 수학과 (학사)
- 2001년 : 광주과학기술원 정보통신공학과 (석사)
- 2002년~현재 : 광주과학기술원 정보통신공학과 박사과정
- 주관심분야 : 네트워크 가상환경, 유비쿼터스 컴퓨팅, HCI 등

저 자 소 개



우 은 택

- 1989년 : 경북대학교 전자공학과 (학사).
- 1991년 : 포항공과대학교 전자전기공학과(석사).
- 1998년 : University of Southern California 전기공학과 (박사).
- 1991년~1992년 : 삼성종합기술원 연구원.
- 1999년~2001년 : ATR, Japan 초빙 연구원.
- 2001년~현재 : 현재 광주과학기술원 정보통신공학과 조교수.
- 주관심분야 : 유비쿼터스 컴퓨팅, 컨텍스트 인식, 혼합현실, 3차원시각, HCI 등