논문 2010-47SP-4-11

관심 영역 기반의 픽셀 도메인 분산 비디오 부호

(The Region-of-Interest Based Pixel Domain Distributed Video Coding With Low Decoding Complexity)

정 천 성*, 김 웅 환*, 전 동 산**, 박 현 욱**, 하 정 석**

(Chunsung Jung, Unghwan Kim, Dongsan Jun, Hyunwook Park, and Jeongseok Ha)

요 약

최근 초경량 비디오 부호를 위해 분산 비디오 부호 (Distributed Video Coding)에 대한 연구가 활발하게 진행되고 있다. 분산 비디오 부호는 H.264/AVC와 같은 종래의 비디오 부호 기술보다 부호화 복잡도는 훨씬 낮지만, 복호기 복잡도는 높은 특징이 있다. 본 논문에서는 분산 비디오 부호의 낮은 부호 및 복호 복잡도를 갖는 관심 영역 (Region-Of-Interest, ROI) 기반의 분산 비디오 부호를 제안하였다. 제안하는 분산 비디오 부호화는 기존의 분산 비디오 부호화와 달리 Wyner-Ziv (WZ) 프레임 전체를 WZ 부호화 하지 않고, 움직임이 많은 영역을 ROI로 두어 ROI 영역만 WZ 부/복호화 함으로서 부/복호화의 복잡도 감소 뿐만 아니라 비트율을 절감 하는 장점을 가지고 있다. 실험을 통해서 제안된 ROI기반의 분산 비디오 부호가 Hall Monitor 영상에서 최대 0.95dB 이득과 Salesman 영상에서 최대 1.87dB의 PSNR 성능 향상을 보였다. 또한, 기존의 분산 비디오 부호 구조보다 부호화 복잡도는 최대 73.7%, 복호기 복잡도는 최대 63.3%를 낮추는 것을 확인하였다. 낮은 복호 복잡도를 위해 기존의 Low-Density Parity-Check (LDPC) 복호 알고리즘으로 사용된 Belief Propagation (BP) 알고리즘 대신 수렴 속도가 최대 41.71% 빠른 Layered BP (LBP) 알고리즘을 이용하였다.

Abstract

Recently, distributed video coding (DVC) has been actively studied for low complexity video encoder. The complexity of the encoder in DVC is much simpler than that of traditional video coding schemes such as H.264/AVC, but the complexity of the decoder in DVC increases. In this paper, we propose the Region-Of-Interest (ROI) based DVC with low decoding complexity. The proposed scheme uses the ROI, the region the motion of objects is quickly moving as the input of the Wyner-Ziv (WZ) encoder instead of the whole WZ frame. In this case, the complexity of encoder and decoder is reduced, and the bite rate decreases. Experimental results show that the proposed scheme obtain 0.95 dB as the maximum PSNR gain in Hall Monitor sequence and 1.87 dB in Salesman sequence. Moreover, the complexity of encoder and decoder in the proposed scheme is significantly reduced by 73.7% and 63.3% over the traditional DVC scheme, respectively. In addition, we employ the layered belief propagation (LBP) algorithm whose decoding convergence speed is 1.73 times faster than belief propagation algorithm as the Low-Density Parity-Check (LDPC) decoder for low decoding complexity.

Keywords : Distributed video coding, Wyner-Ziv video coding, Region-of-Interest (ROI), Layered belief propagation algorithm

I. 서 론

현재 비디오 표준 압축기술인 MPEG과 H.26x 기술들은 압축률을 높이기 위해 연속적인 프레임의 상관성을 이용하여 프레임 간의 예측 부호 기술 (Inter-frame predictive coding)을 사용한다. 이 기술에서 사용하는

* 학생회원, ** 정회원, 한국과학기술원 전기 및 전자공학과
(Department of Electrical Engineering, KAIST)
※ 본 연구는 지식경제부 산업원천기술개발사업 '차세대 모바일 영상서비스를 위한 초경량 비디오 부호화 원천기술개발' 과제의 일환으로 수행되었음)
접수일자: 2009년12월2일, 수정완료일: 2010년6월9일

움직임 예측 (Motion Estimation (ME))으로 인해, 일반적으로 부호기의 복잡도가 복호기보다 5배에서 10배 정도 높다. 이런 비대칭 비디오 구조는 한 번 압축하고 여러 번 복호하는 방송용 시스템 (Broadcasting)이나 주문형 비디오 시스템 (Video-On-Demand (VOD))에 적합하다. 그러나 최근 요구되는 비디오 응용 제품인 무선 감시 비디오 센서, 무선 PC 카메라, 휴대용 단말기의 카메라, 일회용 비디오 카메라와 같은 응용 제품들은 메모리와 전력이 제한되어 있고 계산 능력의 한계 등으로 인해 비디오 부호화를 할 때 필요한 높은 복잡도의 알고리즘을 수행하기 어렵다. 따라서 새로운 응용 분야의 제품들은 기존의 비디오 부호 방식인 예측 부호 기술이 아니라, 낮은 복잡도로 비디오 부호화 할 수 있는 새로운 개념의 비디오 부호화 기술이 필요하다. 최근 분산 비디오 부호 (Distributed Video Coding (DVC)) 가 새로운 응용 분야에 적합한 비디오 부호화 기술로 주목을 받고 있다.

지난 수년간 분산 비디오 부호는 1970년대에 발표된 두 개의 정보 이론[1~2]을 근거로 연구되고 있다. 그 중 하나는 1973년 발표된 Slepian-Wolf 정리[1]로 상관성이 있는 소스 X와 Y를 독립적으로 각각 부호화 하여도 결합 복호를 한다면, 두 소스를 결합 부호화 한 것과 같은 압축 효율을 얻을 수 있는 것을 말한다. Slepian- Wolf 부호는 분산 소스 부호를 의미하고, 무손실 압축이다. 다른 하나는 1976년에 발표된 Wyner-Ziv정리[2]로 Slepian-Wolf 부호를 확장하여 복호기에서 보조 정보 (Side Information (SI))를 이용할 때, 손실 압축의 이론적 경계를 제시하였다. 이러한 분산 소스 부호[1~2]를 이용하여 영상 프레임을 독립적으로 부호화하면, 움직임 보상 (Motion Compensation (MC)) 및 움직임 예측 (ME)과 같은 프레임 사이의 통계적 의존성을 계산하는 복잡도 높은 처리 과정 없이 압축이 가능하여 부호기의 복잡도를 낮출 수 있다. 뿐만 아니라 현재 분산 비디오 부호는 MPEG과 같은 기존의 인트라 부호 (Intra-frame coding)와 가까운 성능을 얻을 수 있다[3].

2002년에 분산 비디오 부호의 실제적인 설계 구조인 Wyner-Ziv (WZ) 비디오 부호[3]가 발표되면서 분산 비디오 부호가 현재까지 광범위하게 연구되고 있다. WZ 비디오 부호의 큰 장점은 비디오 부호기에서 복잡하게 수행되던 프레임 간의 예측 부호 기술을 사용하지 않고, 각 프레임을 서로 독립적으로 압축하여 부호기의 복잡도를 낮춘 것이다. 그러나 부호기의 복잡도가 줄어든 만큼 복잡도가 비디오 복호기로 옮겨 복호기의 복잡도는 높아졌다. 즉, 프레임 간의 예측 부호기술이 보조 정보를 생성하는데 사용이 되어 복호기 복잡도가 증가하였고, 더불어 WZ 프레임 복호 복잡도도 추가되어 복호기 복잡도가 높아지게 되었다.

분산 비디오 복호기의 복잡도를 줄이기 위해서는 복호기 복잡도를 줄이거나 피드백 채널로 인해 발생되는 복호 지연 문제를 해결해야 한다. 후자의 방법으로 부호기에서 WZ 프레임의 적절한 비트율을 추정하여 압축하는 부호기 제어 방법[12]과 부호기와 복호기에서 비트율 추정이 진행되는 이종 제어 방법[13]이 있다. 그러나 이 방법들은 피드백으로 인해 발생되는 복호 지연 문제만 다루고 있고, 부/복호기의 직접적인 복잡도를 줄이는 알고리즘은 제시하지 않았다. 본 논문에서는 낮은 부/복호 복잡도를 갖는 분산 비디오 부호 설계를 위해 두 가지 알고리즘을 제안한다.

첫째, WZ 프레임에서 Region-Of-Interest (ROI)을 추출하여 선택적 WZ 부호화 알고리즘을 제안한다. WZ 프레임 내에는 영상의 배경과 같이 광도 변화가 적은 영역이 있고, 물체가 움직임으로 광도 변화가 많은 영역이 있다. 따라서 WZ 프레임 전체를 WZ 부호화 하는 것이 아니라, WZ 프레임에서 움직임이 많이 일어났거나 광도 변화가 심한 영역을 ROI로 정의하여, ROI에 대해 WZ 부호화를 선택적으로 할 수 있다. 이를 통해 WZ 프레임의 비트율을 절감하여 Rate-Distortion (RD) 성능을 높일 수 있으며, ROI에 대해 WZ 부/복호화 함으로써 더 낮은 복잡도를 갖게 한다.

둘째, WZ 복호 속도를 개선하는 알고리즘을 제안한다. 기존의 WZ 복호 방법은 Low-Density Parity-Check (LDPC) 복호 알고리즘인 Belief Propagation (BP) 알고리즘을 이용하는 것인데, 기존의 BP 알고리즘을 변형하여 반복 복호 속도를 개선한 Layered BP (LBP) 알고리즘을 적용하여 새로운 WZ 복호기 구조를 제안한다.

본 논문은 총 Ⅳ장으로 구성되며, 제 Ⅰ장에서 새로운 비디오 응용 분야에 적합한 분산 비디오 부호를 기술하고, 본 연구의 배경 및 목적을 설명한다. 제 Ⅱ장에서는 낮은 복잡도를 갖는 분산 비디오 부호 구조를 제안하고, 분산 비디오 부호의 복잡도를 줄이기 위한 알고리즘들을 설명한다. RD 성능 및 복잡도 실험을 통해 기존의 연구 실험 결과와 비교 분석은 제 Ⅲ 에서 소개하고, 제 Ⅳ장에서 결론을 맺는다.

Ⅱ. 본 론

1. 제안하는 분산 비디오 코덱 구조

아래 그림 1은 제안하는 DVC 코덱의 전체적인 구조이다. 제안하는 DVC 코덱은 Aaron[5]이 제안한 픽셀 도메인 DVC 구조와 비슷하다. 그러나 기존의 DVC 코덱은 WZ 프레임 전체를 WZ 부호화 하지만, 제안된 DVC 코덱에서는 움직임이 많은 영역을 ROI로 정의하고, ROI를 WZ 부호화하여 비트율을 절감 및 부/복호화 복잡도를 낮춘다. 그림 1과 같이 제안된 DVC 코덱에서는 점선으로 표시된 ROI 추출하는 과정이 추가되었다.

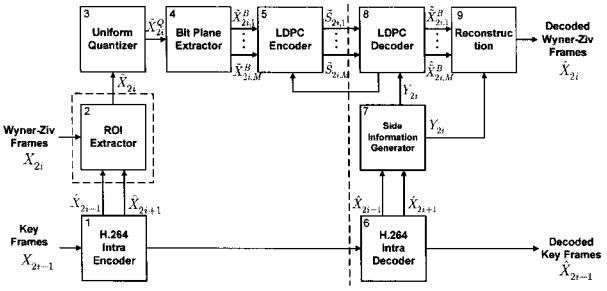


그림 1. 제안된 분산 비디오 코덱 구조
Fig. 1. The structure of proposed distributed video coding.

가. 부호화 과정

먼저 픽셀 도메인 DVC 코덱에서 부호화 하는 과정을 살펴보자. 제안하는 DVC 코덱에서는 기존의 DVC 코덱과 같이 프레임 영상을 홀수 프레임과 짝수 프레임으로 분류하여 각각을 키 프레임 (X_{2i-1})과 WZ 프레임 (X_{2i})으로 나누어 처리한다. 키 프레임과 WZ 프레임은 서로 독립적으로 부호화 되며, 그림 1의 블록 1에서 키 프레임을 H.264/AVC Intra prediction 알고리즘으로 부호화를 한다. 그리고 WZ 프레임을 부호화하기 위해 블록 2의 ROI 추출 과정을 통해 ROI ($\widetilde{X}_{2i}$)를 추출하여 블록 3, 4, 5에서 WZ 부호화 한다. ROI를 WZ 부호화 하는 과정을 살펴보면, 먼저 블록 3에서 M 비트 스칼라 양자화기를 통해 ROI의 각 픽셀 값들을 양자화 하여 양자화 된 심볼 ($\widetilde{X}_{2i}^{Q}$)을 구하게 된다. 그리고 블록 4에서 $\widetilde{X}_{2i}^{Q}$을 비트 스트림으로 변환하여 같은 레벨의 비트(e.g. most significant bit (MSB))들을 추출하여 M개의 비트 평면들로 분류한다. MSB 비트 평면 ($\widetilde{X}_{2i,1}^{B}$)부터 LDPC 부호기(블록 5)에 전달된다. LDPC 부호기는 LDPCA[6] 신드롬 생성기로 WZ 프레임의 M 개의 비트 평면을 각각 누적 신드롬 ($\widetilde{S}_{2i,j}$)으로 압축하고 전송한다.

나. 복호화 과정

다음으로 제안하는 DVC 코덱의 복호화 하는 과정을 살펴보자. 그림 1의 블록 6에서 수행되는 H.264 인트라 복호를 통해 양자화된 키 프레임을 완벽하게 얻게 된다. 보조 정보 생성 블록 (블록 7)에서 복호된 키 프레임 정보인 $\hat{X}_{2i-1}$와 $\hat{X}_{2i+1}$를 이용하여 Ascenso 방법[7]으로 WZ 프레임과 닮은 보조 정보 Y_{2i}를 생성하게 된다. 그리고 생성된 보조 정보 Y_{2i}와 WZ 프레임의 압축된 정보, 즉 블록 5에서 얻어진 누적 신드롬 $\widetilde{S}_{2i,j}$을 이용하여 LDPC 복호기(블록 8)를 통해 WZ 프레임의 양자화 된 비트 스트림 $\hat{\widetilde{X}}_{2i,j}^{B}$을 얻게 된다. 만약 LDPC 복호가 실패하면, WZ 복호기는 WZ 부호기로 여분의 신드롬을 재전송 요청하고, 다시 받은 신드롬을 이용하여 LDPC 복호를 수행하여 LDPC 복호가 성공 될 때까지 이 과정이 반복된다. 복호 과정의 마지막 단계로 양자화된 비트 스트림 $\hat{\widetilde{X}}_{2i,j}^{B}$과 보조 정보 Y_{2i}를 이용하여 재구성 단계(블록 9)를 거쳐 최종적으로 WZ 프레임 $\hat{X}_{2i}$을 복원하게 된다.

2. 제안하는 분산 비디오 부호기

가. ROI 추출기 (블록 2)

제안하는 DVC 부호기에서는 WZ 프레임 전체를 WZ 부호화 하는 것이 아니라, ROI를 추출하여 그 영역에 대해서 선택적으로 WZ 부호화 한다. 그 이유는 WZ 프레임내에는 영상의 배경과 같이 광도 변화가 적은 영역이 있고, 물체가 움직임으로 광도 변화가 많은 영역이 있기 때문이다. 따라서 WZ 프레임에서 움직임이 많이 일어났거나 광도 변화가 심한 영역을 ROI로 정의하고, ROI에 대해 비트율을 할당하여 WZ 부호화한다. 나머지 영역은 영상에서 배경과 같은 영역이기 때문에 움직임이 거의 없다고 판단할 수 있다. 따라서 ROI 아닌 영역은 WZ 부/복호화를 하지 않고, WZ 프레임의 앞/뒤 키 프레임들의 평균값으로 재구성 할 수 있다. 이를 통해 WZ 프레임의 비트 율을 절감하여 RD 성능을 높일 수 있으며, ROI에 대해 WZ 부/복호화 함으로써 더 낮

은 복잡도를 갖게 한다. 추가적으로 보조 정보도 ROI에 대해서만 생성하여 보조 정보 생성시 필요한 복잡도를 줄일 수 있다.

본 논문에서는 세가지 ROI 추출하는 방법을 제안하였다. 1) 영상 단위 ROI 추출 방법으로 전체 영상에 대해서 Mean Absolute Difference (MAD) 를 이용하여 ROI를 추출하며, 전체 영상 부/복호화를 진행하는 동안 ROI 위치는 고정되어있다. 2) 프레임 단위 ROI 추출 방법으로 WZ 프레임을 압축할 때, WZ 프레임의 이전 키 프레임과 다음 키 프레임의 차이인 Absolute Residual Map (ARM)를 이용하여 ROI를 추출하며, ROI의 위치는 움직임이 많은 영역으로 적응적으로 바뀐다. 3) 마지막 방법은 블록 단위 ROI 추출 방법으로 하나의 프레임을 다수의 블록으로 세분화 하고 각각의 블록에 대하여 블록 Sum Absolute Difference (SAD) 값을 계산한다. 전체 불록들 중에서 평균 SAD값이 큰 불록들만을 취하고 선택된 불록의 모음을 ROI로 이용한다. ROI 크기의 합, 즉 선택된 불록들이 차지하는 영역은 전체 영상 크기의 1/4로 정하였다. 예를 들어, QCIF 영상의 경우 ROI 크기는 88×72 이다. 제안된 방식은 ROI가 프레임의 여러 곳에 나눠져 있는 경우 특히 좋은 성능을 발휘한다. 세 가지 방법 중에 영상 단위의 ROI 추출 방법은 전체 프레임에 대해 ROI를 추출해야 하므로 지연시간이 있을 수 있다. 따라서 프레임 단위 ROI 추출 방법이나, 블록 단위 ROI 추출 방법이 실시간 부/복호화를 위해서는 더 적절한 추출 방법이다. 또한 세 가지 ROI 추출 방법은 각각 MAP, ARM, SAD를 생성하기 위한 추가적인 메모리가 필요하지만, 메모리는 동적으로 할당되기 때문에 복잡도는 크게 증가하지 않는다. 오히려 ROI 연산을 수행하면 비트 평면의 크기가 감소하기 때문에 부호화 복잡도는 최대 73.7%, 복호화 복잡도는 최대 63.3% 감소한다는 것을 실험결과를 통해 보일 것이다.

(1) 영상 단위 ROI 추출

영상 단위 ROI 추출 방법은 WZ 프레임의 부호화 전에 전체 영상에 대해서 ROI를 추출하는 방법을 말한다. 영상 단위 ROI 추출 방법은 다음과 같은 절차로 수행된다.

1 단계: MAD 맵 생성

먼저 수식 (1)과 같이 키 프레임 K_i들의 절대값 차이

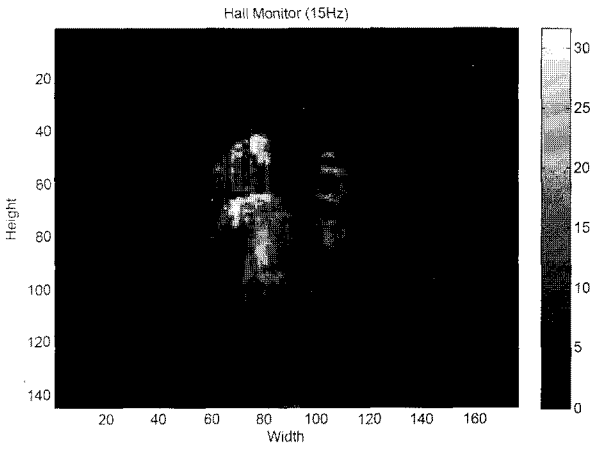


그림 2. MAD 프레임
Fig. 2. The MAD frame.

의 평균값인 MAD 맵 K_{MAD}를 구한다.

$$K_{MAD}(x,y)=\frac{1}{N}\sum_{n=1}^{N}|K_{2n-1}(x,y)-K_{2n+1}(x,y)|,$$
$$(1\le x\le F_h, 1\le y\le F_w) \quad (1)$$

여기서 K_n는 시간 색인이 n번째 프레임을 의미하고, $K_n(x,y)$은 n번째 프레임의 높이 위치 정보인 x, 너비 위치 정보인 y인 (x,y) 위치의 픽셀 값을 의미한다. 그리고 N은 전체 키 프레임 수를 의미하고, F_h와 F_w는 각각 프레임의 높이와 너비를 의미한다.

그림 2는 QCIF 크기의 Hall monitor 영상 100 프레임으로 얻은 MAD 맵이다. Hall monitor 영상은 감시 카메라 영상으로 위치가 고정된 카메라가 사무실 복도를 촬영하는 영상이다. 카메라의 위치가 고정되어 있어 배경은 고정되어 있지만, 복도를 중심으로 사무실 직원들이 움직이고 있다.

Hall monitor 영상의 MAD 맵인 그림 2를 통해 움직임이 많은 복도 중앙 영역을 ROI로 정하는 것이 합당하다. 그러나 이것은 직관적인 결과이고, ROI를 찾는 체계적인 방법이 필요하다. 다음 단계에서 체계적인 방법으로 ROI를 찾는 방법에 대해 소개하겠다.

2 단계: SAD 분포도 생성

ROI를 구하기 위해 1 단계 과정에서 얻은 MAD 맵 K_{MAD}을 기반으로 먼저 수직 SAD 분포도와 수평 SAD 분포도를 생성한다. 그림 2인 K_{MAD}를 수직 방향으로 더하여 수식 (3.a)과 같이 F_w길이인 벡터 K_{SAD}^{V}을 만들고, 수식 (3.b)와 같이 수평 방향으로 더하여 F_h

길이인 벡터 K_{SAD}^{H}을 만든다.

$$K_{SAD}^{V}(y)=\sum_{i=1}^{F_h}K_{MAD}(i,y) \tag{3.a}$$

$$K_{SAD}^{H}(x)=\sum_{j=1}^{F_w}K_{MAD}(x,j) \tag{3.b}$$

3 단계: ROI 위치 찾기

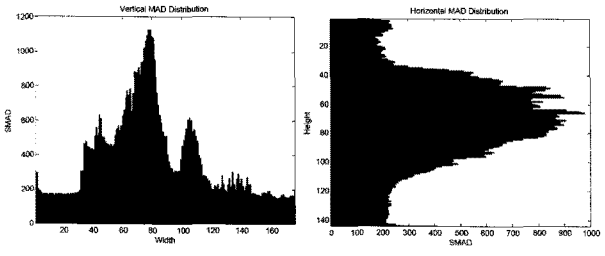
그림 3.(a)는 1 단계 과정에서 예시로 MAD 맵을 생성한 Hall monitor 영상의 MAD 맵 그림 2를 기반으로 수직 SAD 분포도를 생성한 것이고 그림 3.(b)는 수평 SAD 분포도를 생성한 것이다.

마지막 단계는 수직 SAD 분포도 K_{SAD}^{V}와 수평 SAD 분포도 K_{SAD}^{H}를 기반으로 ROI의 위치 정보인 (x_{ROI},y_{ROI})를 찾는 것이다. ROI의 높이 위치를 찾기 위해 수식 (4.a)와 같이 수평 SAD 분포도에서 ROI의 높이 크기의 창에 속하는 값들의 합이 제일 큰 창의 위치를 찾는다.

$$x_{ROI}=\arg\max_x \sum_{i=x}^{x+R_h-1}K_{SAD}^{H}(i),\quad (1\le i\le F_h-R_h+1) \tag{4.a}$$

$$y_{ROI}=\arg\max_y \sum_{j=y}^{y+R_w-1}K_{SAD}^{V}(j),\quad (1\le j\le F_w-R_w+1) \tag{4.b}$$

여기서 R_h와 R_w는 각각 ROI의 높이와 너비를 의미한다. 같은 방법으로 ROI의 너비 위치를 찾기 위해 수식 (4.b)와 같이 수직 SAD 분포도에서 ROI의 너비 크기의 창에 속하는 값들의 합이 제일 큰 창의 위치를 찾는다.



(a) (b)

그림 3. (a) 수직 SAD 분포도, (b) 수평 SAD 분포도
Fig. 3. (a) Vertical SAD Distribution, (b) Horizontal SAD Distribution.

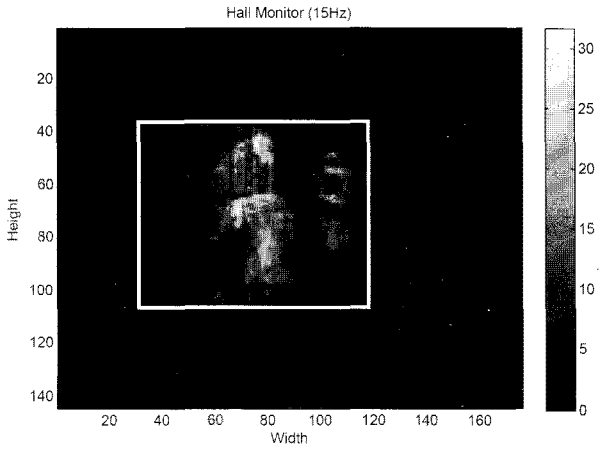


그림 4. 영상 단위 ROI 추출
Fig. 4. ROI extraction at sequence level.

그림 4는 Hall monitor 영상에 대해서 ROI 추출 알고리즘으로 구한 ROI 이다.

특히, 제안된 방식은 키 프레임은 부호기와 복호기가 동시에 가지는 정보인 키 프레임만을 이용하여 ROI를 추출하는 특징을 가지고 있다. 이 때, 부호기는 H.264/AVC Intra coding이 끝난 후 주변 영역에 대한 공간적 예측 (Spatial Prediction)을 위해 생성된 복원된 영상을 기반으로 키 프레임을 만들기 때문에 복호기와 동일한 키 프레임을 갖게 된다. 따라서 부호기가 ROI정보를 복호기에 알려주지 않아도 복호기가 독립적으로 ROI를 추정하여 복호에 이용할 수 있는 장점을 가지게 된다.

(2) 프레임 단위 ROI 추출

영상 단위 ROI 추출 방법은 전체 영상 부호화 하기 전 한 번만 수행되지만, 프레임 단위 ROI 추출 방법은 WZ 프레임을 부호화 하기 전에 매번 수행된다. 프레임 단위 ROI 추출 방법 다음과 같은 절차로 수행된다.

1 단계: Absolute Residual 맵 생성

먼저 수식 (5)와 같이 $2i$번째 WZ 프레임을 부호화 한다고 가정하면, 이전 키 프레임과 다음 키 프레임들의 절대값 차이인 Absolute Residual Map (ARM)를 생성한다.

$$K_{2n,ARM}(x,y)=|K_{2n-1}(x,y)-K_{2n+1}(x,y)| \tag{5}$$

여기서 K_n는 시간 색인인 n번째 프레임을 의미하고, $K_n(x,y)$은 n번째 프레임의 높이 위치 정보인 x, 너비 위치 정보인 y인 (x,y)위치의 픽셀 값을 의미한다.

2 단계: SAD 분포도 생성

ROI의 영역은 앞서 소개한 영상 단위 ROI 추출 방법 (1)과 같은 방법으로 수직 SAD 분포도와 수평 SAD 분포도를 생성한다.

$$K^{V}_{2n,SAD}(y)=\sum_{i=1}^{F_h}K_{2n,ARM}(i,y) \tag{6.a}$$

$$K^{H}_{2n,SAD}(x)=\sum_{j=1}^{F_w}K_{2n,ARM}(x,j) \tag{6.b}$$

3 단계: ROI 위치 찾기

마지막 단계는 수직 SAD 분포도 $K^{V}_{2n,SAD}$ 와 수평 SAD 분포도 $K^{H}_{2n,SAD}$를 기반으로 영상 단위 ROI 추출 방법 (1)의 3단계와 같은 방법으로 수식 (7.a)와 (7.b)를 이용하여 ROI 창에 속하는 값들의 합이 최대가 되는 위치인 ROI의 위치 정보인 $(x_{2n,ROI},y_{2n,ROI})$를 찾는다.

$$x_{2n,ROI}=\arg\max_x\sum_{i=x}^{x+R_h-1}K^{H}_{2n,SAD}(i),\quad (1\le i\le F_h-R_h+1) \tag{7.a}$$

$$y_{2n,ROI}=\arg\max_y\sum_{j=y}^{y+R_w-1}K^{V}_{2n,SAD}(j),\quad (1\le j\le F_w-R_w+1) \tag{7.b}$$

프레임 단위 ROI 추출 방법은 영상의 움직임에 따라서 적응적으로 ROI 영역이 바뀌게 된다. 그림 5는

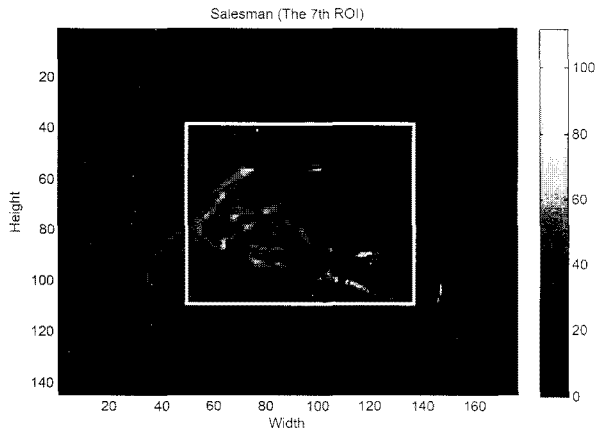


그림 5. Salesman 영상의 7번째 프레임 단위 ROI 추출
Fig. 5. ROI extraction at frame level in 7th salesman sequence.

Salesman 영상에서 프레임 단위 ROI 추출 방법으로 ROI를 추출한 7번째 ROI 영상이다.

(3) 블록 단위 ROI 추출

앞에서 소개한 두 ROI 추출 방법은 ROI 크기는 직사각형으로 고정되어 있다. 그러나 움직임이 큰 부분이 특정영역에 몰려있는 경우가 아니면, (1)과 (2) 방법처럼 하나의 큰 ROI을 추출하는 것은 한계가 있다. 따라서 다양한 영상에 적합한 ROI를 추출하기 위해서는 여러 개의 작은 ROI를 추출하거나 ROI 크기도 바꿔어야 할 필요가 있다. 이 문제를 해결하기 위해 프레임을 작은 크기의 블록으로 나누어 움직임이 많은 블록을 선택하고 선택된 블록들은 제안된 블록 단위 ROI를 구성한다. 이때 선택된 블록들의 개수는 선택된 블록의 전체 면적이 영상단위 및 프레임 단위 ROI 추출 시에 사용한 ROI의 크기와 같게 정한다. 블록 단위 ROI 추출 방법은 다음과 같은 절차로 수행된다.

1 단계: 블록 SAD 계산

블록 단위 ROI 추출 방법은 프레임 단위 ROI 추출 방법과 같이 WZ 프레임을 부호화하기 전에 매번 수행되며, 1단계는 WZ 프레임 (K_{2n})의 앞 키 프레임 (K_{2n-1})과 뒤 키 프레임 (K_{2n+1})의 차이를 절대값을 취하여 블록 단위로 더한다. 이를 수식으로 표현하면 다음 식 (8)과 같다.

$$B_{2n}(i,j)=\sum_{y=y_s}^{y_f}\sum_{x=x_s}^{x_f}|K_{2n-1}(x,y)-K_{2n+1}(x,y)|,\quad (x_s=m(i-1)+1,x_f=mi),\ (y_s=n(j-1)+1,y_f=nj) \tag{8}$$

여기서 B_{2n}은 시간 색인이 $2n$번째 프레임을 의미하고, i와 j는 각각 블록 SAD의 높이 위치와 너비 위치 정보를 의미하고, 그 범위는 각각 $(1\le i\le F_h/m)$, $(1\le j\le F_w/n)$이다.

2 단계: ROI 블록 선택

이 단계에서는 이전 단계에서 구한 블록 SAD 값을 이용하여 ROI 만큼의 블록을 선택하는 단계이다. 움직임이 많은 영역이 ROI 이므로 블록 중에서 SAD 값이 큰 블록을 ROI 만큼 선택한다. 그림 6의 Mother 영상의 경우 움직이는 주요 객체가 2개로 나뉘어져 있고,

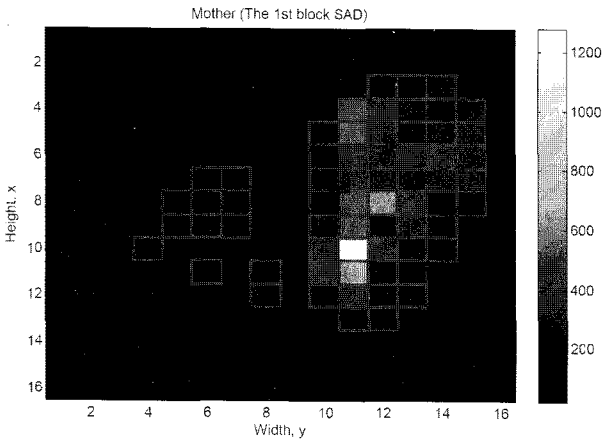


그림 6. Mother 영상의 1번째 블록 단위 ROI 추출
Fig. 6. ROI extraction at block level in 1th salesman sequence.

객체의 크기도 서로 다르기 때문에 프레임 단위 추출 방법 (2)은 적합하지 않다. 그림 6은 Mother 영상에서 블록 단위로 ROI를 추출한 1번째의 ROI 영상이다. 이 방법을 통해 효율적으로 움직임 영역을 모두 ROI로 추출이 가능하다.

3. 제안하는 분산 비디오 복호기

가. 제안하는 LDPC 복호기 (블록 8)

제안된 WZ 복호기는 계층적 신뢰 전파복호(Layered Belief Propagation)를 적용한 LDPC 복호기를 사용한다. Belief Propagation (BP) 복호기는 일반적으로 LDPC부호의 복호에 널리 쓰이지만, 수렴 속도(convergence rate)측면에서는 효율적이지 못하다. 즉, 계층적 신뢰 전파 복호의 주요한 장점은 BP 복호에 비해 수렴속도가 빠르다는 것이다. 계층적 신뢰 전파 복호기는 약 절반의 하드웨어 연산으로도 BP 복호기와 동일한 복호 수율(decoding throughput)을 낼 수 있으며, BP 복호기에 비해 더 적은 메모리로 효율적인 복호가 가능하다[10]. 계층적 신뢰 전파 LDPC 복호기의 복호 절차는 다음과 같다.

복호기는 부호기부터 M개의 비트 평면에 해당되는 M개의 WZ 프레임 신드롬을 받는다. 또한 복호기의 키 프레임들에 의해 생성된 보조 정보를 통해 WZ 프레임 비트 스트림에 대응되는 연판정 (soft decision)값*을 추출한다. 연판정 값을 LDPC 복호기의 가변노드 초기 정보로 입력하고, 계층적 신뢰 전파 복호를 수행한다.

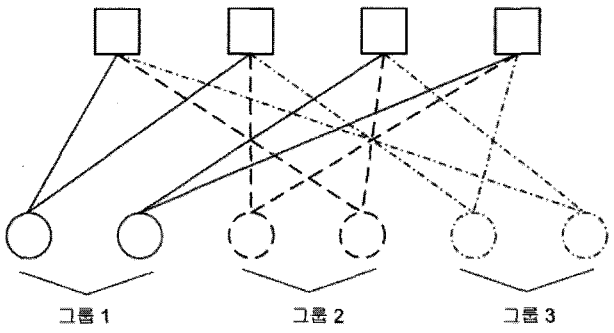


그림 7. 인수 그래프를 3개의 그룹으로 나눔
Fig. 7. Factor graph separated with 3 groups.

계층적 신뢰 전파 복호의 주요한 아이디어는 BP 복호 과정에서 갱신된 메시지 정보를 1회 반복이 완료되기 전에 미리 사용하는 것이다. 계층적 신뢰 전파 복호는 인수 그래프 (Factor graph)를 여러 개의 그룹으로 나누고, 각각의 그룹에 대해서 BP 복호를 수행한다. 그림 7은 인수 그래프를 3 개의 그룹으로 나눈 것이며, 각각의 그룹은 2개의 가변노드 (variable node)와 가변노드에 연결된 선(edge), 그리고 선에 연결된 검사노드 (check node)로 이루어져 있다. 첫 번째 그룹에서 BP 복호가 완료되면 두 번째 그룹이 BP 복호를 수행하며, 이러한 과정은 마지막 그룹까지 유지되어 전체 인수 그래프에 대한 1회 BP 복호를 수행한다. 이 때 앞쪽의 그룹에서 갱신된 정보를 뒤쪽의 그룹에서 사용할 수 있으므로 BP에 비해 수렴속도가 빨라지게 된다. 계층적 신뢰 전파 복호의 그룹 분류는 가변노드 기준으로 그룹을 나누는 방법과 검사노드를 기준으로 나누는 방법이 있으며 본 논문의 복호기는 첫 번째 방법인 가변 노드 기준을 사용한다.

계층적 신뢰전파 복호을 이용한 복호 시에는 각각의 그룹에 대해서 아래의 절차를 수행한다. N개의 가변노드를 L개의 그룹으로 나눈 인수 그래프에서 각각의 그룹에 속한 가변노드의 개수는 $N/L = N_L$이 되며, $1 \le l \le L$ 에 속하는 l번째 그룹에 대한 복호 과정[11]은 다음과 같다.

1) 검사노드에서 가변노드로 메시지를 전파하는 단계로써, $lN_l + 1 \le n \le (l+1)N_l$ 에 속하는 가변노드 n과 $m \in M(n)$을 만족하는 검사노드 m에 대하여 (9), (10)의 연산을 수행한다.

* 본 논문에서는 연판정 값으로 log-likelihood ratio를 사용한다.

$$k_{mn}^{(i)} = \prod_{\substack{n' \in N(m)|n \\ n' \le lN_l}} \tanh\left(\frac{q_{mn'}^{(i)}}{2}\right) \times \prod_{\substack{n' \in N(m)|n \\ n' \ge lN_l+1}} \tanh\left(\frac{q_{mn'}^{(i-1)}}{2}\right) \quad (9)$$

$$\gamma_{mn}^{(i)} = sign(m) \times \log \frac{1+k_{mn}^{(i)}}{1-k_{mn}^{(i)}} \quad (10)$$

2) 가변노드에서 검사노드로 메시지를 전파하는 단계로써, $lN_l+1 \le n \le (l+1)N_l$ 에 속하는 가변노드 n과 $m \in M(n)$을 만족하는 검사노드 m에 대하여 (11), (12)의 연산을 수행한다.

$$q_{mn}^{(i)} = F_n + \sum_{m' \in M(n)|m} \gamma_{m'n}^{(i)} \quad (11)$$

$$q_n^{(i)} = F_n + \sum_{m \in M(n)} \gamma_{mn}^{(i)} \quad (12)$$

위의 식에서 사용한 기호는 다음과 같이 정의된다.

$\gamma_{mn}^{(i)}$: i번째 복호에서 검사노드 m에서 가변노드 n으로 전파하는 Log-Likelihood Ratio (LLR) 값

$q_{mn}^{(i)}$: i번째 복호에서 가변노드 n에서 검사노드 m으로 전파하는 LLR값

$M(n)$: 가변노드 n에 연결된 검사노드의 집합

$M(n)|m$: 검사노드 m을 제외하고 가변노드 n에 연결된 검사노드의 집합

$N(m)$: 검사노드 m에 연결된 가변노드의 집합

$N(m)|n$: 가변노드 n을 제외하고 검사노드 m에 연결된 가변노드의 집합

F_n: 채널로부터 받은 가변노드의 초기 값

$sign(m)$: 신드롬으로부터 받은 검사노드 m의 부호로써 검사노드 m의 신드롬 값이 0 이면 $sign(m)=1$, 신드롬 값이 1이면 $sign(m)=-1$

$1 \le l \le L$에 속하는 모든 그룹에 대해 위의 복호과정을 수행하면 1회 복호가 완료된다. 위 복호 과정을 최대 반복 횟수에 도달할 때 까지 반복 복호한다. 여기서 사용되는 최대 반복 횟수는 50회이다. 그리고 매 회 복호가 완료될 때마다, 가변노드의 LLR 값을 경판정(hard decision)하여 패리티 검사를 통해 신드롬 값과 동일한지 확인하고, 동일할 경우 복호가 성공하였으므로 반복 복호를 정지한다. 따라서 수렴 속도가 빠른 복호 방식을 사용하는 경우 높은 확률로 최대 반복 횟수에 도달하기 전에 복호를 성공적으로 마치게 된다.

위의 복호 과정을 M개의 비트 평면에 대한 M개의 신드롬과 연판정 값에 대해 수행하면, 한 개의 WZ 프레임에 대한 복호가 완료된다.

Ⅲ. 실 험

1. 실험 환경

제안하는 관심 영역 기반의 분산 비디오 부호의 성능 평가를 위해 RD 성능과 코덱의 복잡도를 중점으로 실험하였다. 본 논문에서 RD 성능 평가를 위해 사용된 QCIF 영상은 15Hz Hall Monitor 영상[14], 30Hz Salesman 영상[15]이고, 100 프레임을 이용하였다.

실험은 제안하는 세 가지 ROI 추출 방법과 DISCOVER 코덱[8~9]과 비교로 이루어졌다. 비트율은 키 프레임과 WZ 프레임 각각 비트율을 합하여 계산되었으며, PSNR은 원본 영상과 복원된 영상간의 휘도 성분으로 계산된 값이다. 실험에 사용된 시뮬레이터 사양은 Intel(R) Core(TM)2 Quad CPU Q9650 3.00GHz이며, 메모리는 2.00 GB이다. 그리고 실험에 사용된 프로그램은 C로 구현하였으며 Windows XP 운영 체제에서 실행되었다. 위 프로그램은 CPU 한 개에서 실행되기 때문에, 싱글 코어 환경에서도 동일한 실험이 가능하다.

2. 실험 결과

가. RD 성능

그림 8, 9, 10의 RD 성능 그래프의 4개의 포인트는 각각 2, 3, 4, 5 비트 양자화기를 사용한 것이다. 그림 8

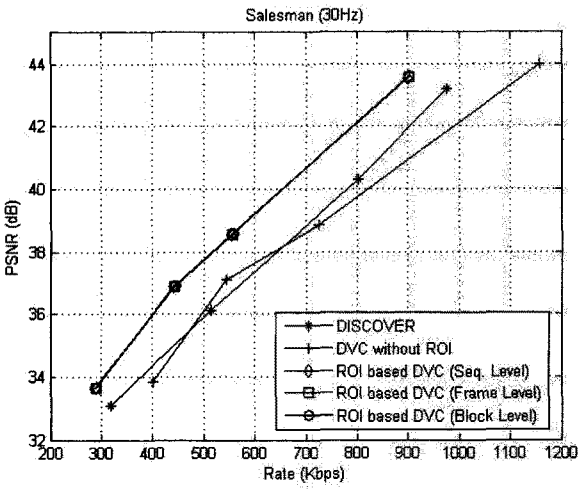


그림 8. Salesman 영상(QCIF at 30Hz)의 RD성능
Fig. 8. RD performance for Salesman (QCIF at 30Hz).

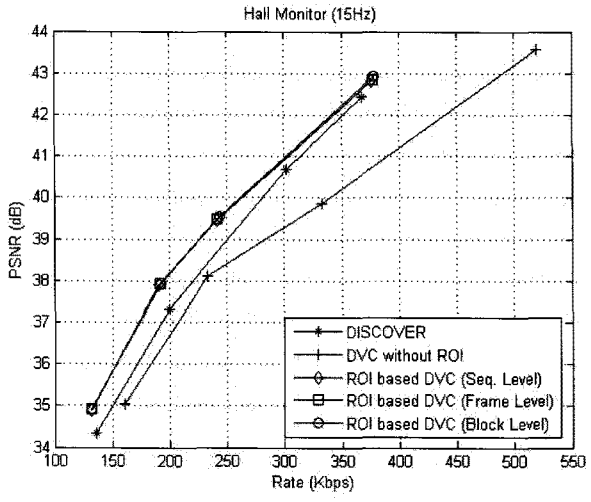


그림 9. Hall Monitor 영상(QCIF at 15Hz)의 RD성능
Fig. 9. RD performance for Hall Monitor. (QCIF at 15Hz)

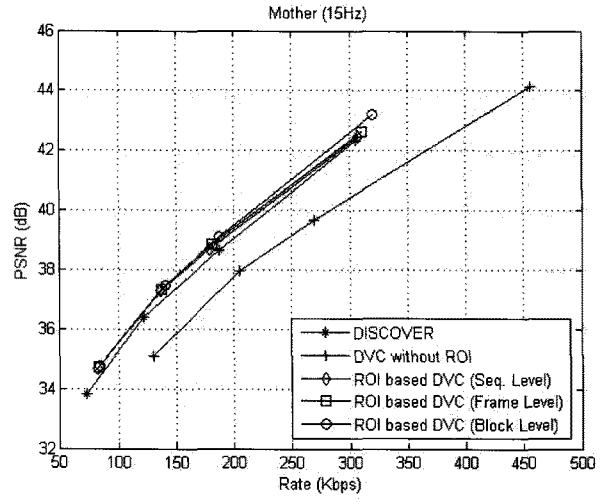


그림 10. Mother 영상(QCIF at 15Hz)의 RD성능
Fig. 10. RD performance for Mother (QCIF at 15Hz).

은 Salesman 영상의 RD 성능 그래프이다. 제안된 관심 영역 기반의 분산 비디오 부호가 기존의 DISCOVER RD 성능보다 PSNR이 최대 1.87dB 향상된 성능을 보인다. 이를 통해 WZ 프레임을 ROI와 ROI 아닌 영역으로 나누어 ROI만 WZ 부/복호화 하여 비트율 절감이 되었음을 확인할 수 있다. 그림 9는 Hall Monitor 영상의 RD 성능 그래프로 그림을 통해 종래의 분산 비디오 부호인 DISCOVER 보다 PSNR이 최대 0.95dB 향상된 성능을 보여준다. 그림 10은 Mother 영상의 RD 성능그래프이며 DISCOVER 보다 PSNR이 최대 0.45dB 향상된 성능을 보인다. Mother영상의 경우 움직임 영역이 분산되어 있기 때문에 블록 단위 ROI추출방법이 영상 단위 ROI추출방법이나 프레임 단위 ROI 추출방법보다 최대 0.3dB정도 더 좋은 성능을 보인다.

나. 제안하는 분산 비디오 부호기 복잡도

표 1은 제안하는 분산 비디오 부호기의 복잡도를 100 프레임의 Hall Monitor 영상에서 WZ 프레임을 부호화 하는데 걸린 전체 시간을 측정한 실험 결과이다. 분산 비디오 부호기의 복잡도가 낮음에도 불구하고 ROI 기반의 분산 비디오 부호기의 복잡도는 최대 73.7%의 부호 시간을 절감하여 초경량 비디오 부호기에 적합한 구조임을 알 수 있다.

표 1. ROI 추출 방법에 따른 부호기 복잡도 (ms) 비교
Table 1. Encoding complexity measured in terms of encoding time (ms) for the Hall Monitor sequence.

M (bits)		5	4	3	2
w/o ROI	LDPC Encoding +Quantization	175	108	91	63
	ROI Extraction	0	0	0	0
	Total	175	108	91	63
영상 단위 ROI	LDPC Encoding +Quantization	32	31	31	32
	ROI Extraction	16	16	16	16
	Total	48	47	47	48
프레임 단위 ROI	LDPC Encoding +Quantization	31	32	32	31
	ROI Extraction	15	15	15	16
	Total	46	47	47	47
블록 단위 ROI	LDPC Encoding +Quantization	64	32	31	31
	ROI Extraction	16	16	16	16
	Total	80	48	47	47
Diff.	w/o ROI - (1)	127	61	44	15
	(%)	72.5%	56.4%	48.3%	23.8%
	w/o ROI - (2)	129	61	44	16
	(%)	73.7%	56.4%	48.3%	25.3%
	w/o ROI - (3)	95	60	44	16
	(%)	54.2%	55.5%	48.3%	25.3%

다. 제안하는 분산 비디오 복호기 복잡도

표 2 는 제안하는 분산 비디오 복호기의 평균 WZ 프레임 복호 시간을 측정한 결과이다. 복호화 과정에는 보조 정보 생성 시간과 LDPC BP 복호 성공한 시간이 포함되어 있다. 표 2를 통해서 ROI를 적용하였을 때 최대 63.3%의 복호 시간 절감을 얻을 수 있다.

표 2. LBP 알고리즘을 사용 했을 때 ROI 추출 방법에 따른 복호기 복잡도 (s) 비교

Table 2. Decoding complexity measured in terms of decoding time (sec) for the Hall Monitor sequence using LBP.

M (bits)		5	4	3	2
w/o ROI		2.629	2.053	1.353	0.942
영상 단위 ROI (1)		0.977	0.81	0.65	0.523
프레임 단위 ROI (2)		0.963	0.829	0.668	0.55
블록 단위 ROI (3)		0.991	0.791	0.655	0.536
Diff.	w/o ROI - (1)	1.652	1.243	0.703	0.419
	(%)	62.8%	60.5%	51.9%	44.4%
	w/o ROI - (2)	1.666	1.224	0.685	0.392
	(%)	63.3%	59.6%	50.6%	41.6%
	w/o ROI - (3)	1.638	1.262	0.698	0.406
	(%)	62.3%	61.4%	51.5%	43.1%

라. LDPC 복호기 복잡도

표 3은 LDPC 복호 복잡도를 측정하는 실험으로 Hall Monitor 영상에서 ROI를 적용하지 않은 분산 비디오 부호와 프레임 단위의 ROI 추출 방법을 적용한 분산 비디오 부호인 두 가지 경우에 대해 실험하였고, 복호를 성공한 경우 기존 복호 알고리즘인 BP와 제안하는 LBP 알고리즘의 평균 WZ 프레임 복호 시간(s)으로 비교한 것이다. LBP 알고리즘으로 LDPC 복호를 한다면 동일한 RD 성능을 유지하면서 최대 41.71%의 복호 시간을 절감할 수 있다.

표 3. BP 알고리즘과 LBP 알고리즘의 복호시간(s) 비교

Table 3. LDPC decoding time (sec) comparison between BP and LBP for the Hall Monitor sequence.

M (bits)		5	4	3	2
w/o ROI	BP	3.25	2.37	1.52	0.97
	LBP	2.07	1.59	0.95	0.56
	Diff.	1.18	0.78	0.57	0.41
	(%)	36.33%	33.03%	37.55%	41.71%
프레임 단위 ROI	BP	0.87	0.66	0.46	0.26
	LBP	0.56	0.45	0.31	0.20
	Diff.	0.31	0.21	0.15	0.06
	(%)	35.02%	31.29%	32.41%	24.96%

IV. 결 론

본 논문에서 제안하는 관심 영역 기반의 분산 비디오 부호는 기존의 분산 비디오 부호와 달리 WZ 프레임 전체를 WZ 부호화 하지 않고, 움직임이 많은 영역을 ROI로 두어 ROI 영역만 WZ 부/복호화 하여 비트율을 절감하여 RD 성능에서 Hall Monitor 영상에서는 PSNR이 최대 0.95dB, Salesman 영상에서는 최대 1.87dB의 성능 향상이 있었다. 또한 부호화 복잡도를 최대 73.7%, 복호화 복잡도를 최대 63.3% 낮추었다. 제안된 복호기에 사용된 LDPC 복호 알고리즘은 낮은 복호 복잡도를 위해 기존의 LDPC 복호 알고리즘으로 사용된 BP 알고리즘 대신 최대 41.71% 속도가 빠른 LBP 알고리즘을 이용하였다. 관심 영역 기반의 분산 비디오 부호는 RD 성능뿐 아니라 부/복호기의 복잡도도 절감하여 새로운 분산 비디오 부호 기술의 핵심 요소 기술로 기대된다.

참 고 문 헌

[1] D. Slepian and J. Wolf, "Noiseless coding of correlated information sources," *Information Theory, IEEE Transactions on*, vol. 19, no. 4, pp. 471-480, Jul. 1973.

[2] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *Information Theory, IEEE Transactions on*, vol. 22, no. 1, pp. 1-10, Jan. 1976.

[3] B. Girod, A. Aaron, S. Rane, and D. Rebollo-Monedero, "Distributed video coding," *Proceedings of the IEEE*, vol. 93, no. 1, pp. 71-83, Jan. 2005.

[4] A. Wyner, "Recent results in the shannon theory," *Information Theory, IEEE Transactions on*, vol. 20, no. 1, pp. 2-10, Jan. 1974.

[5] A. Aaron, S. Rane, R. Zhang, and B. Girod, "Wyner-ziv coding for video: applications to compression and error resilience," in *Data Compression Conference, 2003. Proceedings. DCC 2003*, pp. 93-102, March. 2003.

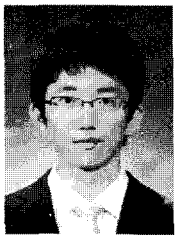
[6] D. Varodayan, A. Aaron and B. Girod, "Rate-Adaptive Codes for Distributed Source Coding," *EURASIP Signal Processing Journal, Special Section on Distributed Source Coding*, vol. 86, no. 11, Nov. 2006.

[7] J. Ascenso, C. Brites and F. Pereira, "Improving Frame Interpolation with Spatial Motion Smoothing for Pixel Domain Distributed Video Coding," *5th EURASIP Conference on Speech and Image Processing*, Multimedia Communications and Services, Smolenice, Slovak Republic, Jul. 2005.

[8] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D.

Kubasov, and M. Ouaret, "The DISCOVER codec: Architecture, Techniques and Evaluation," in *Proc. of PCS 2007*, Lisbon, Portugal, Nov. 2007.
[9] http://www.discoverdvc.org/
[10] E. Sharon, S. Litsyn and J. Goldberger, "Efficient Serial Message-Passing Schedules for LDPC Decoding," *Information Theory, IEEE Transactions on*, vol. 53, no. 11, Nov. 2007.
[11] J. Zhang and M. P. C. Fossorier, "Shuffled Iterative Decoding," *Communications, IEEE Transations on*, vol. 53, no. 1, pp. 209-213, Feb. 2005.
[12] M. Morbee, A. Roca, J. Prades-Nebot, A. Pizurica, and W. Philips, "Reduced decoder complexity and latency in pixel-domain Wyner-Ziv video coders," in *Springer Journal on Signal, Image and Video Processing (SIViP)*, Vol. 2, No. 2, pp. 129-140, June 2008.
[13] J. D. Areia, J. Ascenso, C. Brites, F. Pereira, "Low Complexity Hybrid Rate Control for Lower Complexity Wyner-Ziv Video Decoding." *16th European Signal Processing Conference*, Lausanne, Switzerland, Aug. 2008.
[14] http://amalia.img.lx.it.pt/~discover/test_conditions.html
[15] http://trace.eas.asu.edu/yuv/index.html

저 자 소 개



정 천 성(학생회원)
2008년 한동대학교 전산전자 공학부 학사 졸업
2010년 한국과학기술원 전기 및 전자공학과 석사 졸업
2010년~현재 LG전자 멀티미디어 연구소 연구원
<주관심분야 : 통신, 채널부호>

전 동 산(정회원)
2002년 부산대학교 전자공학과 학사 졸업
2004년 한국과학기술원 전기 및 전자공학과 석사 졸업
2004년~현재 한국전자통신연구원 전파방송연구단 연구원
2007년~현재 한국과학기술원 박사과정
<주관심분야 : 영상처리, 영상압축, 디지털 방송 시스템>

하 정 석(정회원)-교신저자
1992년 경북대학교 전자공학과 학사 졸업
1994년 포항공과대학교 전자전기 석사 졸업
2003년 Georgia Tech 박사 졸업
2004년~2009년 한국정보통신대학교 조교수.
2009년~현재 한국과학기술원 조교수.
<주관심분야 : 통신, 채널부호, 물리계층보안>

김 웅 환(학생회원)
2008년 홍익대학교 전자전기 공학부 학사 졸업
2009년~현재 한국과학기술원 전기 및 전자공학과 석사과정
<주관심분야 : 통신, 채널부호>

박 현 욱(정회원)
1981년 서울대학교 전기공학과 학사 졸업
1983년 한국과학기술원 전기 및 전자공학과 석사 졸업
1988년 한국과학기술원 전기 및 전자공학과 박사 졸업
1989년~1992년 University of Washington 연구원
1992년~1992년 삼성전자 수석연구원
1993년~현재 한국과학기술원 전자전산학과 전기 및 전자공학전공 교수
<주관심분야 : 영상처리, 영상압축, 의료영상시스템, 멀티미디어 시스템>