

특집논문-08-13-6-11

복호 신뢰도에 기반하여 점진적으로 보조정보를 향상시키는 변환영역 Wyner-Ziv 부호화 방법

고 봉 혁^{a)}, 심 혁 재^{a)}, 전 병 우^{a)†}

Transform domain Wyner-Ziv video coding with successively improving side information based on decoding reliability

Bonghyuck Ko^{a)}, Hiuk Jae Shim^{a)}, and Byeungwoo Jeon^{a)†}

요 약

센서네트워크와 같은 자원한정적인 상황에서의 비디오 부호화가 중요해짐에 따라 경량화 비디오 부호화 문제의 솔루션으로 분산비디오 부호화 기술(DVC: Distributed Video Coding)이 폭넓게 연구되고 있다. 이의 대표적인 기술인 Wyner-Ziv 부호화 기술은, 복호화기에서 영상간 유사도를 이용하여 부호화하는 현재 프레임에 대한 보조정보를 생성하고, 발생하는 잡음을 채널코드를 이용하여 제거함으로써 영상을 복원한다. 따라서 복호화기에서 보조정보를 정교하게 만들수록, 즉 제거되어야 할 잡음이 적게 생성될수록 전체적인 시스템의 부호화 효율이 좋아진다. 하지만 실제로 복호화기가 부호화하는 현재 프레임에 대한 정보 없이 정확도 높은 보조정보를 만드는 것은 상당히 어려운 일이므로, 복호되는 현재 프레임의 정보 및 그 신뢰도에 기반하여 점진적으로 보조정보를 향상시키는 방법이 개발되었다. 하지만 복호결과를 이용하여 보조정보를 향상시키기 위해서는 복호결과 정보의 에러율에 못지않게 복호결과로부터 유효한 정보의 양 또한 중요하다. 따라서 본 논문에서는 변환영역의 Wyner-Ziv 부호화 기술에서 점진적으로 복호되는 현재 프레임에 대한 정보의 에러율 및 유효한 정보 양을 고려한 신뢰도에 기반하여 점진적으로 보조정보를 향상시키는 부호화 방법을 제안한다. 실험결과는 제안 방법이 점진적 보조정보 향상방법을 사용하지 않는 이전 변환영역 Wyner-Ziv 부호화 방법에 비해 최대 약 1.7 dB에 이르는 성능향상이 있음을 보였다.

Abstract

As a video encoding in resource constrained environments such as sensor networks has become an important issue, DVC(Distributed Video Coding) has been intensively investigated as a solution for light weighted video encoding problem. Known as one of the representative schemes of DVC, the Wyner-Ziv coding generates side information of current frame only at decoder, using correlation among frames, and reconstructs video through noise elimination on the side information using channel code. Accordingly, the better quality of side information brings less channel noise, thus attains better coding performance of the Wyner-Ziv coder. However, since it is hard for decoder to generate an accurate side information without any information of original frame, a method to successively improve side information using successively decoded original frame, based on decoding reliability, was previously developed. However, to improve side information from decoding results, not only an error rate of the decoding result as a reliability, but also the amount of reliable information from the decoding result is important. Therefore, we propose TDWZ(Transform-domain Wyner-Ziv coding) with successively improving side information based on decoding reliability considering not only an error rate but also the amount of reliable information of the decoding results. Our experiment shows the proposed method gains average PSNR up to 1.7 dB over the previous TDWZ, that is without successive side information improvement.

Keywords : DSC, DVC, Wyner-Ziv coding, Decoding reliability

1. 서론

최근 각종 멀티미디어 기술의 급속한 발전으로 디지털 콘텐츠 시장의 규모가 급격히 성장하였다. 그리고 소비자들의 다양한 미디어에 대한 요구에 맞춰 사용자 제작 콘텐츠(UCC: User Creative Contents)로 불리는 개인 제작 비디오 콘텐츠 및 관련 시장 규모가 급증하였다. 여기에 이종망들과는 독립적으로 운영되었던 종래의 방송망이 IPTV(Internet Protocol Television)의 발달로 인터넷망에 통합되고 있어 향후의 방송환경에서는 사용자 제작 콘텐츠의 저변이 더욱 확대될 것으로 예상되고 있다. 이러한 사용자 제작 콘텐츠는 비전문성 및 편의성 등의 이유로 고가의 전문적인 기기대신 저가의 모바일 캠코더가 선호되고 있다. 하지만 이러한 모바일 기기에서는 사용 가능한 에너지와 컴퓨팅 자원이 한정적이기 때문에 높은 압축 성능을 보이는 저 복잡도의 부호화기술의 필요성이 대두된다. 종래의 MPEG-1/2/4나 H.26x와 같은 표준 기술은 기본적으로 부호화기에서 영상에 존재하는 중복성(Redundancy)를 제거하는 일련의 처리과정을 수행하기 때문에 자원 한정적인 상황에서는 압축 성능을 떨어뜨려 부호화 복잡도를 낮출 수밖에 없고, 따라서 경량화 부호화에는 적합하지 못하다는 문제점이 지적되어 왔다.

그리고 근래에 다시점 비디오(Multiview Video)가 차세대 방송기술 중 하나로 각광받고 있다. 이 기술은 여러 대의 카메라로 한 장면에 대한 다시점의 영상을 촬영하여 방송 콘텐츠를 제작한다. 따라서 사용자는 TV를 통하여 어떠한 장면이라도 원하는 시청 시점을 자유롭게 조정할 수 있다. 하지만 다수의 카메라로부터 얻어지는 영상에 대한 데이터의 양의 방대하므로 다시점 비디오에서는 효율적인 압축 기술이 무엇보다 중요하다. 다시점 비디오는 다수의 카메라를 이용하여 동시에 촬영하기 때문에 시점간 상관도(Correlation)가 높다. 따라서 다시점 부호화 기술(MVC: Multiview Video Coding)에서는 이를 효율적으로 이용하

여 압축성을 높힐 수 있다. 하지만 영상 간에 존재하는 중복성(Redundancy)을 제거하는 일련의 처리절차를 부호화기에서 수행하려면 각 카메라로부터 획득한 영상들을 연계하여 부호화해야 한다. 다시점에서 촬영한 영상은 카메라 간에 서로 공유하기가 실제로는 불가능하므로 이를 위해서는 각각의 카메라들이 다시점 부호화기에 획득한 영상을 전송하여야 한다. 따라서 다시점 부호화기는 다수의 카메라로부터 원본영상을 전송받아야 하므로 많은 양의 메모리 공간과 컴퓨팅 자원을 갖추어야 한다. 이를 피하기 위해 압축기능을 내장한 카메라를 사용하여 각 카메라에서 다시점 부호화기로 전송하는 데이터의 양을 줄일 수도 있지만 이러한 카메라는 가격이 비싸고 다시점 영상은 수많은 카메라를 필요로 하므로 시스템을 구성하는데 따른 가격의 증가는 피할 수 없다. 또한 연계부호화(Joint Encoding)방식은 시청 시점의 수로 비트율을 조절한다면 시청 시점의 수에 따라 부호화를 각각 수행하여야 하는 문제점이 있다. 분산화된 형태로 각 카메라에서 획득한 영상을 독립적으로 압축하여 전송하는 방식의 경우, 부호화기에서 영상간 중복성을 제거하는 일련의 복잡한 처리절차를 수행하지 않기 때문에 경량화 부호화가 가능하고, 또한 부호화기에서 단순히 전송받는 시점 및 시점의 개수를 조정하여 시청 시점 영상의 화질 열화없이 전송률을 줄일 수도 있다.

따라서 종래의 표준 압축기술에 기반하여 영상 및 영상 간에 존재하는 중복성을 제거하는 일련의 처리절차를 부호화기에서 수행하였을때 발생하는 문제들의 주요 해결방안으로 분산 소스 부호화 기술(DSC : Distributed Source Coding)이 관심을 받아왔다. 이 기술은 Slepian-Wolf이론과 Wyner-Ziv이론에 기반하여, 분산된 여러 소스들을 소스 데이터간의 중복성(redundancy)을 제거하는 일련의 처리 없이 독립적으로 부호화하는 문제를 다루며 이론적으로 분산소스 부호화 기술이 소스간 유사도를 이용하는 기존의 부호화 기술과 동일한 정도까지의 유효-왜곡 성능을 얻을 수 있음을 증명하였다^{[1],[2]}.

비디오 압축기술의 구현 관점에서 이 이론들은, 기존의 비디오 압축 기술이 프레임 간 중복성을 줄이기 위한 모든 처리절차를 부호화기에서 행하여 부호화기의 복잡도가 부호화기의 복잡도보다 상당히 높은 구조로 이루어져 있었으

a) 성균관대학교 정보통신공학부
 School of Information & Communication Eng., Sungkyunkwan Univ.
 † 교신저자 : 전병우(bjeon@skku.edu)
 ※ 이 논문은 과학기술부의 재원으로 한국과학재단 국가지정 연구실사업으로 수행된 연구임(R0A-2006-10286-0(2008))

나, 부호화 이득에 별다른 손실 없이 부호화기에서 행하였던 모든 신호처리절차를 복호화기 쪽으로 옮길 수 있음을 시사한다. 이에 분산 비디오 부호화 기술(DVC : Distributed Video Coding)은 각 소스들을 서로 독립적으로 부호화하며 기존 기술과 같은 복잡한 처리절차를 부호화기에서 행하지 않기 때문에 비디오 압축기술의 경량화 부호화기 문제 해결방안으로 각광받고 있다.

DVC 기술 중 현재 가장 잘 알려져 있는 기술은 Wyner-Ziv 부호화 기술이다^[3]. 이 기술에서는 부호화기에 현재 부호화하는 프레임의 보조정보가 유효하지 않기 때문에 현재 프레임과 보조정보간의 차이를 일종의 잡음으로 간주한다. 이 잡음은 전송채널에서 발생하는 것이 아니라 복호화기에서 현재 프레임에 유사한 보조정보를 만들 때 그 정교성이 떨어져 발생하는 것이므로 이를 ‘가상의 채널 잡음(Virtual Channel Noise)’이라 한다. Wyner-Ziv 부호화 기술에서는 이 가상의 채널 잡음을 제거하기 위해 부호화기가 터보코드나 LDPC코드와 같은 오류정정 능력이 뛰어난 채널코딩 기법을 이용하여 소스를 부호화하며 생성된 패리티 정보를 복호화기에 전송한다. 복호화기에서는 부호화기로부터 전송받은 패리티를 이용하여 보조정보에 있는 잡음을 제거하여 영상을 복원 한다.

따라서 이 기술에서는 복호화기에서 보조정보를 정교하게 만들수록, 즉 제거되어야 할 잡음이 적게 생성될수록 전체적인 시스템의 부호화 효율이 좋아진다. 따라서 보조정보를 얼마나 잘 만드는가는 전체 시스템의 성능과 직결된 요소라 할 수 있다. 하지만 실제로 복호화기에 부호화하는 현재 프레임에 대한 정보가 없는 상태에서 보조정보를 정밀히 만드는 것은 상당히 어려운 일이며 특히 프레임 간 움직임이 복잡하거나 occlusion이 많은 경우 더욱 그러하다. 따라서 보조정보를 정교하게 만들기 위해서는 부호화하는 현재 프레임에 대한 정보가 복호화기에 필요하다. 기존에는 이러한 목적을 위해 부호화기에서 복호화기로 부호화하는 프레임에 대한 일부 정보를 담고 있는 hash 비트를 별도로 전송하는 방식의 접근이 있었다^[4,5]. 하지만 기본적으로 Wyner-Ziv 부호화 기술에서는 부호화기에서 복호화기로 부호화하는 프레임에 대한 패리티 비트를 전송하며 이러한 패리티 비트는 부호화하는 프레임에 대한 정보를

포함하고 있으므로 추가적인 정보를 전송하지 않더라도 복호화기에서 전송받은 패리티 비트를 이용하여 보조정보를 더욱 정밀히 할 수 있다. 그리고 Wyner-Ziv 부호화 기술에서는 부호화기에는 보조정보가 그리고 복호화기에는 부호화하는 현재 프레임에 대한 정보가 없어 채널잡음에 대한 정보가 부/복호화기 모두에 유효하지 않다. 따라서 잡음을 제거하기 위해 필요한 패리티 비트의 양 또한 알 수 없어 역방향 채널을 두고 복호화기가 부호화기로부터 잡음이 제거될 때까지 패리티 비트를 점진적으로 전송받는다. 따라서 이러한 구조에서는 점진적으로 복호되는 현재 프레임의 정보를 이용하여 점진적으로 보조정보를 향상시키는 접근이 가능하다. 하지만 실제로 패리티비트를 이용하여 보조정보를 향상시키기 위해서는 복호화기에 보조정보 향상을 위해 주어지는 정보의 신뢰도가 중요하다. 따라서 기존의 방법^[6]은 각 비트프레임을 복호화한 이후 보조정보를 향상시켰었다. 이는 만약 각 패리티비트 전송 시마다 보조정보 향상을 시도할 경우 발생하는 채널복호화 오류로 인해 주어진 정보의 신뢰도가 떨어져 오히려 보조정보가 저하되는 문제가 발생하기 때문이다. 따라서 이전 방법^[7]에서는 이러한 채널 코드의 복호화 오류를 해결하기 위한 방안으로 공간적 유사도를 기반으로 하는 선택적 미디언 필터^[8,9]를 사용하는 점진적 보조정보 향상방법을 사용하였었다.

그러나 복호결과를 이용하여 보조정보를 향상시키기 위해서는 복호결과 정보의 에러율에 못지않게 복호결과로부터 유효한 정보의 양 또한 중요하다. 그리고 현재 Wyner-Ziv 부호화 기술에는 화소영역에서 작동하는 화소영역 Wyner-Ziv 부호화 기술(PDWZ: Pixel-domain Wyner-Ziv coding)^[3]과 변환영역에서 작동하는 변환영역 Wyner-Ziv 부호화 기술(TDWZ: Transform-domain Wyner-Ziv coding)^[10,11]이 있으며 변환영역의 Wyner-Ziv 부호화 기술이 부호화에 영상의 공간적인 유사도도 이용하므로 유효성능이 화소영역의 Wyner-Ziv 코딩 기술보다 좋다.

따라서 화소영역의 Wyner-Ziv 부호화 기술에 기반하였던 이전 제안 방법을 변환영역의 Wyner-Ziv 부호화 기술로 확장할 필요가 있으며 이에 본 논문은 점진적으로 복호되는 현재 프레임에 대한 정보의 에러율 및 유효한 정보 양을 고려한 신뢰도에 기반하여 점진적으로 보조정보를 향상시

키는 개선된 변환영역의 Wyner-Ziv 부호화 방법을 제안한다.

본 논문의 구성은 다음과 같다. 우선 II장에서는 제안하는 변환영역 Wyner-Ziv(TDWZ) 부호화 기술의 전반적인 구조 및 복호화 과정에 관하여 설명한다. III장에서는 제안하는 점진적 보조정보 향상방법을 상세히 설명한다. 그리고 IV장에서는 실험 및 그 결과에 대해서 논의한다. 마지막으로 V장에서는 결론과 향후 연구과제에 대해서 논의한다.

II. 제안하는 변환영역 Wyner-Ziv 부호화 방법

1. 부호화기의 구조

그림 1은 본 논문에서 사용하는 Wyner-Ziv 이론 기반의 경량화 비디오 부호화 구조를 나타낸다. Wyner-Ziv 부호화 기술에서는 프레임을 크게 키프레임과 Wyner-Ziv 프레임으로 나눈다. 이 중 키프레임은 기존의 인트라 코덱으로 부호화되어 복호화기로 전송된다. 본 논문에서는 매 홀수 번째 프레임이 이에 해당하며 부호화에는 H.264/AVC intra

codec이 사용되었다. 키프레임 부호화후에는 Wyner-Ziv 프레임에 대한 부호화가 이뤄지며 이는 DCT 변환, DCT 변환계수들에 대한 주파수 밴드 형성, 양자화, 비트플레인 생성, LDPC 부호화를 거쳐 최종적으로 Wyner-Ziv 프레임에 대한 패리티 비트가 생성되는 과정으로 요약 할 수 있다.

1.1 DCT 변환 (Discrete Cosine Transform)

키프레임이 인트라 부호화되어 복호화기로 전송되고 나면 부호화기는 Wyner-Ziv 프레임을 식 (1)과 같이 4x4 정수 DCT 변환한다. 이때 변환과정의 스케일링은 역변환 과정에 포함된다.

$$F = ICT \times f \times ICT^T$$

$$f' = ICT^T \times (F * S * S') \times ICT \tag{1}$$

식 (1)에서 F는 변환영역의 4x4블록, ICT는 4x4 integer cosine transform matrix, f는 화소영역의 4x4블록, ICT^T는 4x4 transposed integer cosine transform matrix, f'는 역 변환된 화소영역의 4x4블록, *는 두 행렬의 동일 위치 요소

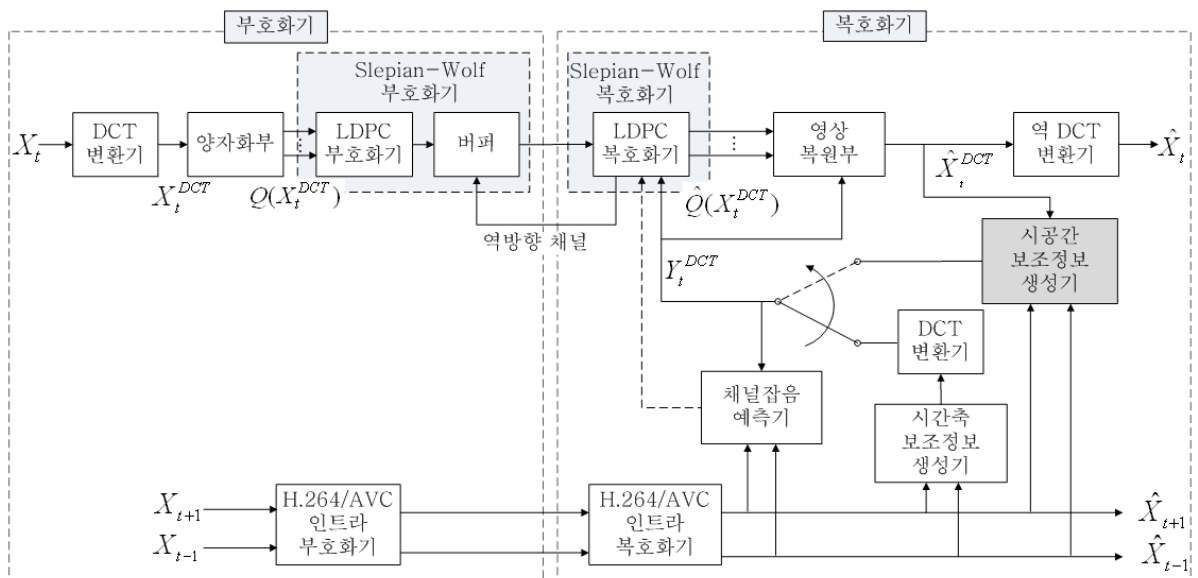


그림 1. 제안하는 변환영역 Wyner-Ziv 부호화 구조
Fig. 1. Proposed Transform-domain Wyner-Ziv coding structure

들 간 곱 연산을 뜻하며 S 는 4×4 스케일링 행렬, S' 은 4×4 역 스케일링 행렬을 의미한다.

1.2 양자화기 (M_k -bits Quantizer)

부호화기는 변환 후 프레임 내 동일 변환계수 성분들을 모아 주파수 밴드를 형성한다. 각 주파수 밴드는 그림 2의 저전송률(M1)에서 고전송률(M8)에 상응하는 양자화 표를 이용하여 양자화 되며 표 안의 숫자는 각 밴드에 해당하는 양자화 레벨의 수를 나타낸다^[11]. 그리고 각 밴드의 경우 DC 밴드에는 uniform scalar 양자화기를 사용하며 dynamic range는 $[0, 2^{12}]$ 로 고정하는 반면 AC 밴드에 대해서는 deadzone 양자화기를 사용하며 DC 밴드와는 달리 각 밴드의 dynamic range를 고정하지 않고 $2 \cdot \max val_k$ 로 사용한다. 여기서 $\max val_k$ 는 k 번째 밴드의 계수 중 그 크기가 가장 큰 값으로 부호화되는 모든 AC 밴드에 대해 복호화기로 전송된다. 한편 양자화 시 k 번째 AC 밴드의 양자화 스텝크기와 양자화 심볼값 q 가 가리키는 변환계수 값의 범위는 아래의 식 (2)와 같다^[12].

$$W_k = \left\lfloor \frac{2 \cdot \max val_k}{2^{M_k}} \right\rfloor$$

$$I_k^q = \begin{cases} [(q-1)W_k, qW_k) & q < 0 \\ [-W_k, W_k) & q = 0 \\ [qW_k, (q+1)W_k) & q > 0 \end{cases} \quad (2)$$

식 (2)에서 W_k 는 k 번째 주파수 밴드의 양자화 스텝크기, 2^{M_k} 는 k 번째 밴드에 할당되는 양자화 레벨의 수, I_k^q 는 k 번째 밴드의 양자화 심볼값 q 가 가리키는 변환계수 값의 범위를 의미한다.

양자화 후 각 밴드의 양자화 심볼값은 각 밴드에 대해서 동일 위치의 비트를 모아 만든 MSB에서 LSB까지의 비트 플레인으로 구성되며 이후 순차적으로 LDPC 부호화기로 입력된다.

1.3 LDPC 부호화기 (LDPC Encoder)

그림 1의 LDPC 부호화기^[13]는 각 밴드의 양자화 심볼스트림의 각 비트 플레인을 입력으로 받아 syndrome 비트를 생성하며 생성된 syndrome 비트는 버퍼에 저장한다. 이때 복호화기에서 보조정보는 원본 프레임에 가상의 채널 잡음이 인가된 것으로 간주되므로 생성되는 systematic 정보는 복호화기로 전송되지 않는다. 그리고 버퍼에 저장되었던 syndrome 비트는 복호화기의 요청에 따라 그 중 일부를 점진적으로 전송한다.

2. 복호화기의 구조

2.1 시간축 보조정보 생성기 (Temporal Side Information Generator)

Wyner-Ziv 부호화 기술에서는 보조정보를 만드는데 일

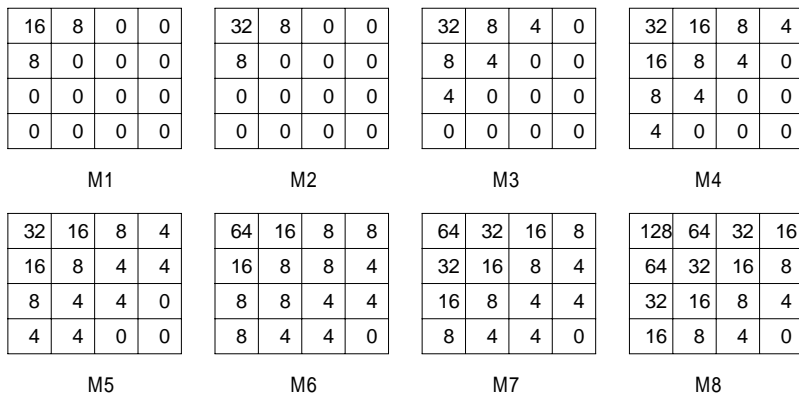


그림 2.율제어를 위한 8개의 양자화 표(저전송률(M1) ~ 고전송률(M8))^[11]
 Fig. 2. Eight quantization matrices for rate control(low bit-rate(M1) ~ high bit-rate(M8))^[11]



(a) 14th 프레임

(b) 78th 프레임

그림 3. 생성된 보조정보(QP = 25, 포만 시퀀스)
Fig. 3. Generated side information (QP = 25, Foreman sequence)

반적으로 프레임 간 선형적인 움직임을 가정한 프레임 보 간법을 사용한다^[14]. 그림 3은 이에 대한 예로써 생성된 Foreman 영상의 보조정보를 보여준다. 보조정보를 만들 때 프레임 간에 움직임이 많지 않은 경우에는 그림 3(a)와 같이 정상적으로 보조정보가 생성된다. 하지만 이와는 달리 프레임 간 움직임이 심한 경우 또는 occlusion이 존재할 경우에는 그림 3(b)와 같이 잘못된 블록을 보조정보로 생성하게 되어 보조정보 내에 많은 오류가 발생하게 되며 이는 복원된 영상의 주관적 그리고 객관적인 화질을 심히 저하시킨다. 본 논문에서는 이를 해결하기 위하여, 대응하는 블록 주변의 유사성에 기반한 보간오류 판별 및 보정방법을 사용한다^[15].

2.2 채널잡음 예측기 (Channel Noise Estimator)

일반적으로 Wyner-Ziv 부호화 기술에서 부호화되는 현재 프레임과 보조정보 간 채널잡음의 통계적인 특성은 laplacian 분포로 모델되며 이는 laplacian 계수값만으로 간단히 정의가 가능하다. 하지만 보조정보 상에 발생하는 채널잡음은 주로 프레임 간 움직임에 의해 발생하므로 한 프레임 내에서 non-stationary한 특성을 띤다. 따라서 채널잡음을 모델링하기 위해 모든 변환 계수에 각각 laplacian coefficient 값을 구하여 적용한다^[16].

2.3 LDPC 복호화기 (LDPC Decoder)

본 논문의 LDPC 복호화기^[13]는 보조정보와 잡음모델로부터 intrinsic information을 계산하며 여기에 부호화기로

부터 전송되는 syndrome 비트를 더해 복호 과정을 수행한다. 따라서 더 좋은 율-왜곡 성능을 얻기 위해서는 intrinsic information이 주는 정보의 양이 많아야 하며 또한 그 신뢰도가 높아야 한다. joint bit-plane 복호화 방법은 식 (3)과 같이 엔트로피의 증가 없이 심볼레벨의 계산을 비트레벨의 계산으로 치환한다^[17].

$$\begin{aligned}
 H(q_{k,i} | Y_{k,i}^{DCT}) &= H(q_{k,i}^{MSB}, \dots, q_{k,i}^{LSB} | Y_{k,i}^{DCT}) \\
 &= H(q_{k,i}^{MSB} | Y_{k,i}^{DCT}) + H(q_{k,i}^{MSB+1} | q_{k,i}^{MSB}, Y_{k,i}^{DCT}) \\
 &\quad + \dots + H(q_{k,i}^{LSB} | q_{k,i}^{MSB}, \dots, q_{k,i}^{LSB-1}, Y_{k,i}^{DCT})
 \end{aligned}
 \tag{3}$$

식 (3)에서 $q_{k,i}$ 는 k 번째 주파수 밴드의 i 번째 양자화 심볼, $Y_{k,i}^{DCT}$ 는 보조정보의 k 번째 밴드의 i 번째 변환계수, $q_{k,i}^m$ 은 k 번째 밴드의 i 번째 양자화 심볼의 m 번째 비트를 뜻한다.

그리고 여기에 이전에 복호된 비트플레인의 결과를 다음 비트플레인의 복호화 과정에 이용하여 점진적으로 노이즈 모델을 정교하게 함으로 intrinsic information의 엔트로피를 더욱 낮출 수 있다.

각 비트플레인 복호화 시 LDPC 복호화기에 입력되는 intrinsic information은 다음의 log-likelihood ratio 값으로 주어진다.

$$LLR(k,i,m) = \log \frac{P(q_{k,i}^m = 0 | Y_{k,i})}{P(q_{k,i}^m = 1 | Y_{k,i})}
 \tag{4}$$

식 (4)에서 $LLR(k,i,m)$ 는 k 번째 밴드의 i 번째 양자화 계수의 m 번째 비트에 대한 intrinsic information 값이다.

LDPC 복호화기는 intrinsic information과 syndrome 비트로부터 복호화과정을 수행하고 그 결과로 연판정-출력(Soft-output)을 낸다. 그리고 그 결과로부터 에러율을 산출하나 이 값이 신뢰도가 낮기 때문에 에러율이 임계 에러율(예: 10^{-4}) 이하로 내려갔는지의 여부를 정확하게 판단하기 어렵다. 따라서 실제적으로는 부호화기로부터 CRC비트를 전송받아 이를 판별하며 CRC 체크 후 에러율이 임계값 이하로 내려간 경우 복호를 끝마치고 그렇지 못할 경우에는

피드백 채널을 통하여 부호화기에 추가적인 패리티 요청을 한다^[18].

2.4 영상 복원기 (Reconstruction)

채널 복호화가 끝나면 복호화기는 원본 변환계수의 양자화 심볼스트림을 복원한다. 영상 복원기는 이 심볼스트림에 양자화 과정에서 버려진 나머지 하위 비트들을 식 (5)와 같이 보상한다. 이는 보조정보와 복원된 양자화 심볼 그리고 원본정보와 보조정보간의 통계적 상관관계(즉, 원본정보와 보조정보간의 예측된 잡음모델)에 기반하여 원본정보의 기대값을 구하는 것이다.

$$\hat{X}_{k,i}^{DCT} = E[X|Y_{k,i}^{DCT}, \hat{q}_{k,i}] \quad (5)$$

식 (5)에서 $\hat{X}_{k,i}^{DCT}$ 는 복원된 k 번째 밴드의 i 번째 변환계수, $\hat{q}_{k,i}$ 는 복원된 k 번째 밴드의 i 번째 양자화 심볼을 의미한다.

2.5 시공간 보조정보 생성기 (Spatio-temporal Side Information Generator)

영상이 복원되고 나면 우선 시공간 보조정보 생성기는 선택적 미디언 필터를 통하여 복원영상에 존재하는 채널 복호화 오류를 정정한다^[8,9]. 공간적 유사도를 이용하는 상기의 과정이 끝나면 이어서 영상 간에 존재하는 시간적 유사도를 이용하여 복원영상에 잔여 하는 에러를 정정하며 이 과정은 복원된 영상과 양 키프레임 사이의 양방향 움직임 예측/보상과정을 통하여 수행된다. 그리고 보정된 복원 영상을 새로운 보조정보로 갱신함으로써 다음의 채널 복호화 과정에 사용한다.

III. 제안하는 점진적 보조정보 향상 방법

1. 점진적 보조정보 향상 방법 및 과정

채널 복호화기에서 각 비트프레임을 임계 에러율 이하로 복호함으로써 부호화하는 현재 프레임에 대한 정보가 복호화

기에 유효하게 된다. 복호화기는 이 정보를 이용하여 양 키프레임 상에서 양방향 움직임 예측/보상 과정을 수행함으로써 보조정보를 향상시킨다. 이때 보조정보의 향상정도는 복원된 정보의 양과 신뢰도, 즉 에러율, 그리고 영산간의 시간적 유사도에 의존한다. 따라서 기존방법^[6]에서는 각 비트프레임을 완전히 복원한 이후에만 양방향 움직임 예측/보상을 통해 보조정보를 향상시켰다. 하지만 비트프레임을 완전히 복원하지 않아 에러가 발생하더라도 복원된 정보의 신뢰도가 충분하다면 양방향 움직임 예측/보상을 통해 보조정보를 향상시킬 수 있다. 이는 아래의 식 (6)에서 복호화기에 유효한 현재 프레임에 대한 정보의 양이 프레임 보간으로 생성된 기존 보조정보 향상에 필요한 최소한의 정보량보다 커지는 시점이 된다.

$$E[|X_{ij} - E[\hat{X}_{ij}|p_e^{ij}]|] \leq \epsilon \quad (6)$$

식 (6)에서 \hat{X}_{ij} 은 (i,j) 위치의 복원된 변환계수, p_e^{ij} 는 (i,j) 위치의 복원된 변환계수의 에러율, X_{ij} 는 (i,j) 위치의 현재 프레임의 변환계수, ϵ 은 보조정보 향상에 필요한 현재 프레임에 대한 정보의 최소치가 복호화기에 유효할 때 복원정보와 원본정보 간 에러의 평균값이다. 하지만 ϵ 은 영상의 움직임 특성에 따라 달라지며 또한 p_e^{ij} 는 (i,j) 위치의 복원된 비트프레임의 수(즉, 복원된 정보의 양)와 각 위치의 비트 에러율에 의존한다. 하지만 복호화기에서는 이의 정확한 값을 알 수 없으므로 실제로는 복호화기에서 예측한 복원정보의 에러율에 의존하여 영상에 무관하게 보조정보 향상이 가능한 시점을 실험적으로 추정하였다.

그림 4는 제안하는 점진적 보조정보 향상 방법의 흐름도이다. 우선 복호화기는 시간축 보조정보 생성단계에서 만들어진 보조정보를 이용하여 복호화 과정을 수행한다. 이때 채널 복호화기는 그림 4에서와 같이 목표 임계 에러율(Pe^{target})보다 더 높은 임계 에러율(Pe^{1st})에서 복호를 마친다. 발생하는 채널 복호화 오류는 다음의 시공간 보조정보 생성단계에서 선택적 미디언 필터를 적용하여 정정한 후 양방향 움직임 예측/보상을 통해 복원영상을 향상시키고 이를 새로운 보조정보로 갱신하여 다음의 복호 및 복원

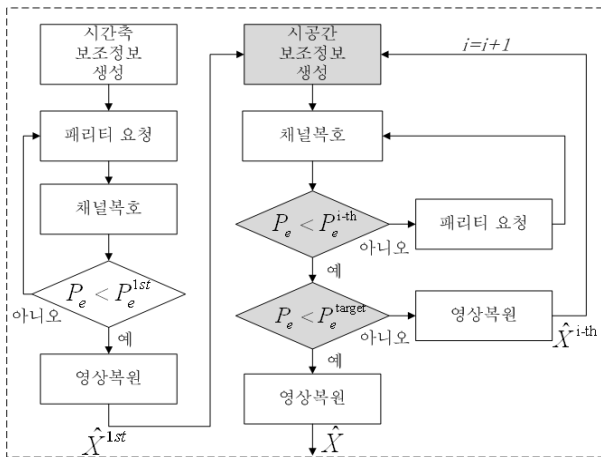


그림 4. 제안하는 점진적 보조정보 향상방법
 Fig. 4. Proposed method to successively improve side information

과정에 사용한다. 이때 보조정보 개선을 시도하는 임계 에러율의 단계를 너무 많이 하면 증가하는 복호화기 복잡도에 비해 얻을 수 있는 이득이 미비하므로 적정한 단계의 임계 에러율을 두고 보조정보 개선을 시도한다.

2. 시공간 보조정보 생성방법

보조정보 향상을 위해서는 복원결과의 신뢰도뿐만이 아니라 그 정보의 양도 중요하다. 따라서 본 논문에서는 실제적으로는 상위의 몇 개의 비트플레인만 보조정보 개선 없이 완전히 복원하고 그 이하의 비트플레인에 대해서만 에러율 레벨에서 보조정보 개선을 수행한다.

그리고 보조정보 개선 가능한 채널 복호화기의 임계 에러율을 높이기 위해 선택적 미디언 필터를 사용하여 발생하는 채널 복호화 오류를 정정한다. 하지만 화소영역에서 적용되던 이전 방법⁷⁾과는 달리 본 논문의 선택적 미디언 필터는 변환영역에서 적용된다. 따라서 선택적 미디언 필터의 임계값을 크게 하여 주변계수들에 비해 상당히 다른 계수들에 대해서만 채널 복호화 오류가 발생한 것으로 판별하고 필터를 적용한다.

양방향 움직임 예측/보상과정은 영상간의 시간적 유사도를 이용하여 복원영상내의 여분의 잡음을 정정한다. 하지만 통상적인 움직임 예측/보상 과정과는 다른 처리가 필

요한데 이는 복원영상 내에 발생하는 에러에 의해 신뢰도가 높은 정보와 그렇지 않은 정보가 혼재하기 때문이다. 이 중 신뢰도가 높은 정보는 채널복호화기로부터 정상적으로 복호된 변환계수의 양자화 심볼값으로 이는 변환계수의 MSB'에 해당하는 정보이다. 반면 신뢰도가 낮은 정보는 MSB'에 채널복호화 오류가 발생한 변환계수와 복원 후 여전히 LSB'에 많은 양자화 에러를 갖고 있는 변환계수, 그리고 아직 복호가 수행되지 않아 보간오류를 그대로 갖고 있는 상위 주파수 밴드의 변환계수들이다.

복호화기는 복원영상이 갖고 있는 정보의 신뢰도를 화소영역이 아닌 변환영역에서 상기의 설명과 같이 구분할 수 있으므로 움직임 예측/보상 과정은 모두 변환영역에서 이루어져야만 한다.

이 중 움직임 예측과정은 우선 현재 복호를 진행하는 주파수 밴드가 특정 임계 주파수 이하일 때는 움직임 예측시 각 블록의 이미 복원된 저주파 성분만을 이용하여 움직임 예측을 수행한다. 이 후 복호화를 상기 임계 주파수까지 진행하여 대부분의 잡음이 제거 되면 복호가 진행되지 않은 상위의 고주파 밴드를 포함한 모든 주파수를 이용하여 움직임 예측을 수행한다. 그리고 각 블록의 정합도 계산 시에는 복원영상의 보정대상 블록 내 변환계수 중 미디언 필터가 적용된 계수가 있을 경우, 이는 비록 공간적 유사도에 기반하여 보정되지만 여전히 보정된 값에 에러가 남아 있을 수 있기 때문에 정확도를 위해 계산에서 제외한다. 그 외의 변환계수들에 대해서는 키프레임의 탐색영역에서도 출되는 후보블록의 대응하는 변환계수가 동일한 양자화 심볼값을 갖는 경우 정합도 계산에 가중치를 둔다⁶⁾.

움직임 보상 시에는 복원영상의 보정대상 블록의 변환계수 중 아직 복호가 수행되지 않은 상위의 고주파 변환계수나 이미 복호되었거나 또는 복호가 진행 중인 저주파 변환계수 중 선택적 미디언 필터가 적용된 변환계수는 움직임 벡터가 가리키는 후보 블록의 대응 변환계수로 보상한다. 그리고 이외의 변환계수에 대해서는 움직임 벡터가 가리키는 후보블록의 변환계수가 동일한 양자화 심볼값을 갖는 경우 하위 LSB'를 보상하고 그렇지 않은 경우에는 후보블록의 대응 변환계수에 잡음이 있는 것이므로 보상하지 않는다.

IV. 실험 및 결과

1. 실험 방법

보조정보의 화질과 영상의 움직임 복잡도 간의 관계를 살펴보기 위해 Hall monitor, Coastguard, Foreman, Stefan 영상에 대하여 다양한 키프레임 양자화 파라미터에 따른 보조정보의 화질을 측정하여 비교하였다. 그리고 본 논문에서 제안하는 변환영역의 Wyner-Ziv 부호화 방법의 성능을 평가하기 위해 상기의 실험 영상에 대하여 다음 다섯 가지 부호화 방법의 율-왜곡 성능을 비교하였다.

- SKKU-TDWZ(with refinement): 점진적으로 보조정보를 향상시키는 제안 TDWZ
- SKKU-TDWZ: 점진적으로 보조정보를 향상시키지 않는 기본구조의 TDWZ
- H.264/AVC inter(IPIP): JM 12.2
- H.264/AVC intra only: JM 12.2
- DISCOVER: DISCOVER 그룹의 TDWZ^[12,19]

구체적 실험 조건은 다음과 같다.

- 영상: 15Hz 150프레임의 QCIF 사이즈 Hall Monitor,

표 1. 양자화 파라미터
Table 1. Quantization setting

WZ frame (Quantization matrix number)		1	2	3	4	5	6	7	8
Key frame (QP for H.264/AVC Intra)	Hall monitor	37	36	36	33	33	31	29	24
	Coastguard	38	37	36	34	33	30	30	26
	Foreman	40	39	38	34	34	32	29	25
	Stefan	40	39	38	34	34	32	29	25

Coastguard, Foreman, Stefan 영상

- 양자화: M1 ~ M8의 8개의 양자화 표에 대하여 상응하는 키프레임의 QP는 표 1과 같다.
- 시간 축 보조정보 생성기: 양 키프레임 간 움직임을 위한 블록 크기는 8x8이며 움직임을 위한 탐색영역은 ±32.
- 시공간 보조정보 생성기: 양방향 움직임을 위한 탐색영역의 크기는 ±8.

전송률과 PSNR은 Wyner-Ziv 프레임과 키프레임을 포함하는 모든 프레임에 대한 값이며 PSNR은 원본 영상과 복원된 영상 간에 측정되었다.

표 2. 생성된 보조정보의 화질 [dB]
Table 2. Image quality of generated side information [dB]

Sequence	QM(WZ frame)	1	2	3	4	5	6	7	8
	QP(Key frame)	40	39	38	34	34	32	29	25
Hall monitor	SI	29.01	29.70	30.24	32.43	32.43	33.43	34.50	35.62
	Key	29.05	29.84	30.49	33.51	33.51	35.04	37.26	40.26
Coastguard	SI	27.20	27.67	28.06	29.74	29.74	30.37	31.16	31.89
	Key	27.51	28.09	28.62	31.23	31.23	32.53	34.69	38.12
Foreman	SI	26.30	26.66	26.94	28.05	28.05	28.46	28.92	29.44
	Key	28.68	29.33	29.90	32.61	32.61	33.93	36.05	39.25
Stefan	SI	21.71	21.97	22.15	22.70	22.70	22.81	22.93	23.03
	Key	25.79	26.55	27.27	30.65	30.65	32.35	35.02	38.89

* QM은 WZ(Wyner-Ziv)프레임에 대한 양자화 표 값이며, QP는 키프레임에 대한 양자화 파라미터 값이다. 그리고 SI는 보조정보의 화질을 나타내며 Key는 키프레임의 화질을 나타낸다.

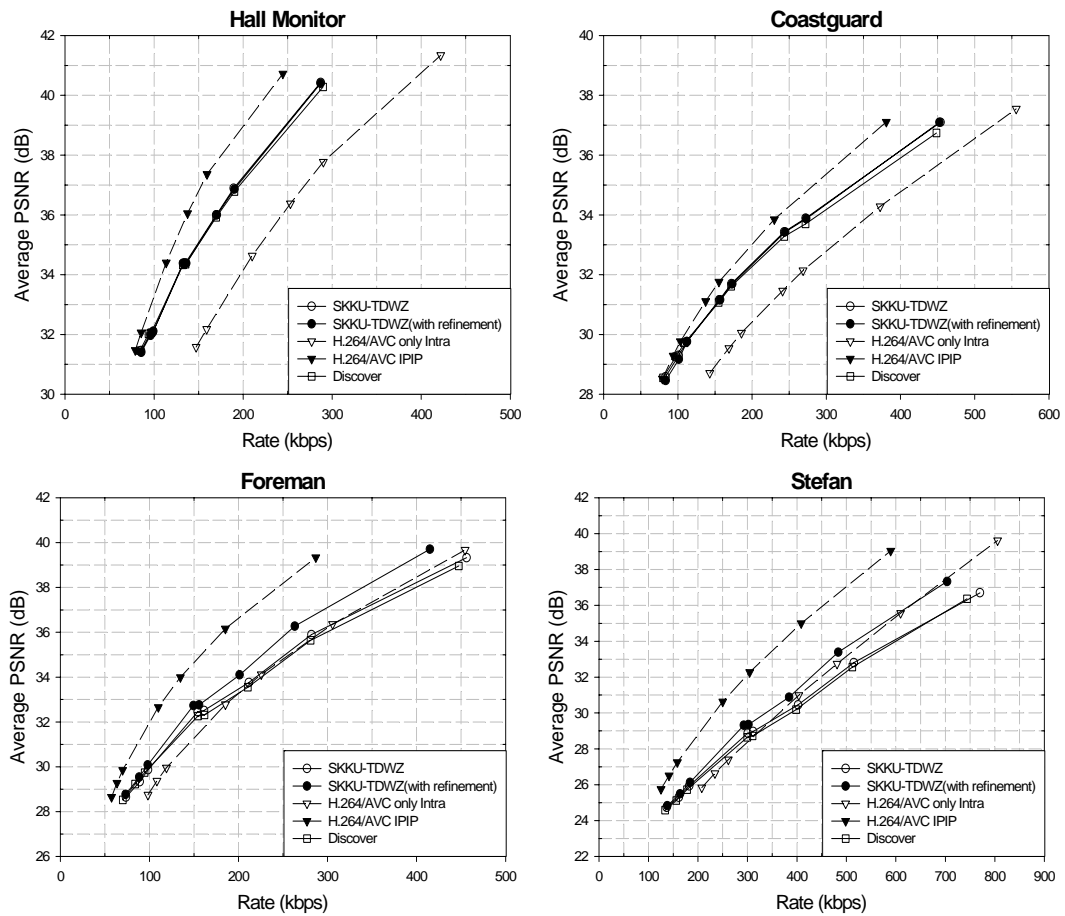


그림 5. R-D 성능 비교
Fig. 5. R-D performance comparison

2. 실험 결과

표 2는 여러 영상에 대하여 생성된 보조정보의 화질을 나타낸 것으로 보조정보의 화질은 생성된 보조정보와 원본 프레임간에 측정된 PSNR 값이다. 표로부터 Hall Monitor 나 Coastguard와 같이 움직임이 단순한 영상은 비록 참조되는 키프레임의 화질이 더 낮음에도 불구하고 Foreman이나 Stefan과 같은 움직임이 복잡한 영상보다 보조정보의 화질이 더 좋음을 알 수 있다. 이로부터 보조정보의 화질이 영상 간 움직임에 상당히 의존함을 확인 할 수 있는데, 이는 복호 화기에서 보조정보를 생성할 때 현재 프레임에 대한 정보

가 없어 프레임간 움직임을 선형으로 가정하는데 상당부분 기인한다. 따라서 특히 Foreman이나 Stefan과 같이 움직임이 복잡한 영상의 경우, 복호화기에서 점진적으로 복원되는 현재 프레임에 대한 정보를 이용하여 양방향 움직임 예측/보상을 함으로써 영상 간 비선형적인 움직임을 잘 예측하여 보조정보를 상당히 향상시키므로 개선된 율-왜곡 성능을 얻을 수 있다.

그림 5는 실험에 사용한 영상들에 대한 율-왜곡(R-D: rate-distortion) 성능을 나타낸다. 우선 DISCOVER 방법은 현재 DVC의 선두 연구 그룹 중 하나인 DISCOVER가 최근 공개한 TDWZ 코덱을 비교 방법들과 동일 조건 하에서 실

힘한 것이다. 이와 비교하였을 때 제안 TDWZ에서 점진적 보조정보 향상방법을 사용하지 않는 방법(SKKU-TDWZ)은 고전송률에서 다소의 성능이득을 보이는데 이는 프레임 보간으로 보조정보 생성 후 부정확하게 보간 되는 블록들을 side matching^[15]으로 정정하여 나타나는 결과이다.

제안하는 점진적 보조정보 향상방법의 성능을 평가하기 위하여 제안방법을 적용한 SKKU-TDWZ(with refinement) 방법과 제안 방법을 적용하지 않은 SKKU-TDWZ 방법의 실험결과를 비교하면 우선 Hall Monitor나 Coastguard와 같이 움직임이 적은 영상에서는 제안방법을 통한 성능 향상이 거의 없음을 알 수 있다. 이는 영상 내 움직임이 적거나 단순한 경우 시간 축 보조정보 생성 시 이미 상당한 화질의 보조정보를 생성할 수 있어 제안 방법을 통한 보조정보 화질 개선의 여지가 적기 때문이다. 하지만 이와는 달리 Foreman이나 Stefan 영상과 같이 움직임이 복잡한 경우에는 제안 방법을 통하여 보조정보를 상당히 개선하므로 실험결과와 같이 상당한 윌-왜곡 성능 향상을 얻을 수 있다. Foreman 영상의 경우에는 제안하는 점진적 보조정보 향상방법을 적용할 경우, 점진적 보조정보 향상방법을 사용하지 않은 SKKU-TDWZ 방법과 비교하였을 때 최고 약 1.2dB에 해당하는 PSNR 이득 또는 약 60 kbps에 해당하는 전송률 감소효과를 보이며 Stefan 영상의 경우에는 최고 약 1.7 dB에 해당하는 PSNR 이득 또는 약 110 kbps에 해당하는 전송률 감소효과가 있음을 보인다. Stefan 영상이 Foreman 영상보다 움직임이 더 많아 보조정보를 만들기 더 어렵기 때문에 제안하는 방법을 통하여 더 많은 윌-왜곡 성능 향상을 보임을 확인할 수 있다. 그리고 결과로부터 두 영상 모두 저전송률일 때 보다 고전송률일 때 성능 향상이 더 많이 발생하는 것을 알 수 있는데 이는 우선 높은 전송률로 갈수록 양자화 심볼에 발생하는 잡음이 많아 보조정보를 개선하였을 때 얻을 수 있는 이득이 더 많기 때문이다. 그리고 여기에 움직임 예측/보상 시 복호화기에 유효한 정보가 저전송률일 때 보다 고전송률일 때가 더 많고 또한 두 키프레임 상에 존재하는 양자화 에러도 더 적어 더 정확한 움직임 예측/보상을 할 수 있기 때문이다.

IV. 결론 및 향후 연구과제

본 논문에서는 복호화기에서 점진적으로 복호되는 현재 프레임에 대한 정보의 에러율 및 유효한 정보의 양을 고려한 신뢰도에 기반하여 점진적으로 보조정보를 향상시키는 변환영역의 Wyner-Ziv 부호화 방법을 제안하였다. 움직임이 많아 보조정보를 잘 생성하기 어려운 실험 영상의 모의 실험에서는 제안방법을 통하여 복호화기에서 보조정보를 점진적으로 향상시키므로 채널 복호화기가 요구하는 패리티 비트양이 상당히 줄어들고 또한 양자화 에러도 상당량 보정되어 개선된 윌-왜곡 성능을 얻을 수 있음을 확인할 수 있었다. 이후의 연구에서는 선택적 미디언 필터와 움직임 예측/보상방법을 더욱 발전시켜 더 높은 임계 에러율에서부터 보조정보를 개선할 수 있는 방안에 대해서 연구할 예정이다.

참 고 문 헌

- [1] J. D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Theory*, vol. IT-19, pp. 471 - 480, July 1973.
- [2] A. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Theory*, vol. 22, pp. 1-10, July 1976.
- [3] A. Aaron, S. Rane, R. Zhang, and B. Girod, "Wyner-Ziv coding for video: Applications to compression and error resilience," *Proc. of IEEE Data Compression Conference, Snowbird, UT*, pp. 93-102, March 2003.
- [4] A. Aaron, S. Rane and B. Girod, "Wyner-Ziv video coding with hash-based motion compensation at the receiver," *Proc. IEEE International Conference on Image Processing, Singapore*, vol. 5, pp. 3097-3100, October 2004.
- [5] J. Ascenso and F. Pereira, "Adaptive hash based side information exploitation for efficient Wyner-Ziv video coding," *IEEE International Conference on Image Processing, San Antonio, USA*, vol. 3, pp. 29-32, September 2007.
- [6] J. Ascenso, C. Brites, and F. Pereira, "Motion compensated refinement for low complexity pixel based distributed video coding," *Proc. of IEEE International Conference on Advanced Video and Signal Based Surveillance, Como, Italy*, pp. 593-588, September 2005.
- [7] B. Ko, H. J. Shim, and B. Jeon, "Wyner-Ziv coding with spatio-temporal refinement based on successive turbo decoding," *IEEE International Conference on Multimedia and Expo*,

Hannover, Germany, pp. 785-788, June 2008.

[8] 고봉혁, 심혁재, 전병우, "센서네트워크상의 응용을 위한 오류정정 기법을 이용한 경량화 부호화 방법," 전자공학회논문지 제45권 제1호, pp. 11-21, 2008년 1월.

[9] 고봉혁, 심혁재, 전병우, "공간적 유사성과 심볼단위 오류정정 채널 코드를 이용한 경량화 비디오 부호화 방법," 방송공학회논문지 제13권 제2호, pp. 188-199, 2008년 3월.

[10] A. Aaron, S. Rane, E. Setton and B. Girod, "Transform-domain Wyner-Ziv codec for video," Proc. Visual Communications and Image Processing, San Jose, CA, pp. 520 - 528, January 2004.

[11] C. Brites, J. Ascenso and F. Pereira, "Improving Transform domain Wyner-Ziv video coding performance," IEEE International Conference on Acoustics, Speech and Signal Processing, Toulouse, France, vol. 2, pp. 525-528, May 2006.

[12] X. Artigas, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov and M. Oualet, "The DISCOVER codec: architecture, techniques and evaluation," Picture Coding Symposium, Lisbon, Portugal, November 2007.

[13] D. Varodayan, A. Aaron and B. Girod, "Rate-adaptive distributed source coding using Low-Density Parity-Check codes," Proc. Asilomar Conference on Signals and Systems, Pacific Grove, CA, pp. 1203-1207, November 2005.

[14] 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, Slovak Republic, July 2005.

[15] B. Ko, H. J. Shim, and B. Jeon, "Wyner-Ziv video coding with side matching for improved side information," IEEE Pacific Rim Symposium on Image Video and Technology, Santiago, Chile, pp. 816-825, December 2007.

[16] C. Brites, J. Ascenso and F. Pereira, "Studying temporal correlation noise modeling for pixel based Wyner-Ziv video coding," IEEE International Conference on Image Processing, Atlanta, USA, pp. 273-276, October 2006.

[17] R. Westerlaken, S. Borchert, R. K. Gunnewiek and I. Lagendijk, "Analyzing symbol and bit plane-based Ldpc in Distributed Video Coding," IEEE International Conference on Image Processing, San Antonio, USA, vol. 2, pp. 17-20, September 2007.

[18] D. Kubasov, K. Lajnef and C. Guillemot, "A hybrid encoder/decoder rate control for a Wyner-Ziv video codec with a feedback channel," IEEE MultiMedia Signal Processing Workshop, Chania, Crete, pp. 251-254, October 2007.

[19] <http://www.discoverdvc.org/>

— 저 자 소 개 —



고 봉 혁

- 2006년 : 아주대학교 전자공학부 졸업(학사)
- 2007년 ~ 현재 : 성균관대학교 정보통신공학부 석사과정
- 주관심분야 : 멀티미디어 영상압축, 신호처리



심 혁 재

- 2000년 : 성균관대학교 전자공학과 졸업(학사)
- 2002년 : 성균관대학교 정보통신공학부 졸업(석사)
- 2002년 ~ 현재 : 성균관대학교 정보통신공학부 박사과정
- 주관심분야 : 멀티미디어 영상압축, 신호처리

저 자 소 개



전 병 우

- 1985년 : 서울대학교 전자공학과 졸업(학사)
- 1987년 : 서울대학교 전자공학과 졸업(석사)
- 1992년 : Purdue Univ, School of Elec. 졸업(공학박사)
- 1993년~1997년 : 삼성전자 신호처리연구소 수석연구원
- 1997년~현재 : 성균관대학교 정보통신공학부 교수
- 주관심분야 : 멀티미디어 영상압축, 영상인식, 신호처리