

# 율-왜곡 기반 선택적 분산 비디오 부호화 기법

이병탁<sup>01)</sup>, 김진수<sup>2)</sup>, 김재곤<sup>1)</sup>

한국항공대학교 전자 및 항공전자공학부<sup>1)</sup>, 한밭대학교 멀티미디어공학과<sup>2)</sup>

[venezia\\_jgkim@kau.ac.kr](mailto:venezia_jgkim@kau.ac.kr), [jskim67@hanbat.ac.kr](mailto:jskim67@hanbat.ac.kr)

## Rate-Distortion Based Selective Encoding in Distributed Video Coding

Byung-Tak Lee<sup>01)</sup>, Jin-soo Kim<sup>2)</sup>, and Jae-Gon Kim<sup>1)</sup>

Korea Aerospace University<sup>1)</sup>, Hanbat National University<sup>2)</sup>

### 요 약

분산 비디오 부호화(Distributed Video Coding: DVC)는 복호화보다 훨씬 적은 계산량으로 부호화를 가능하게 함으로써 계산 성능 및 전력이 제한된 환경을 위한 비디오 부호화 기법으로 주목 받고 있다. 그러나 DVC의 부호화 성능은 아직 기존의 비디오 압축 표준에 많이 미치지 못하고 있으며 이를 극복하기 위한 많은 연구가 진행되고 있다. 본 논문에서는 DVC의 부호화 성능 개선을 위한 율-왜곡 기반의 영역별 선택적 부호화 기법을 제안한다. 제안 기법은 복호기 측에서 보조정보를 얻는 과정에서 얻어지는 움직임 벡터가 부호기 측으로 피드백된다는 가정하에 이를 이용하여 움직임 보상 예측 신호를 구하고 율-왜곡(Rate-Distortion) 기반의 블록 단위로 선택적 부호화를 수행한다. 모의실험을 통하여 본 논문에서 제안한 움직임 정보 피드백을 이용한 율-왜곡 기반의 선택적 블록 부호화 기법이 기존의 DVC 부호화 기법보다 시퀀스에 따라 최대 약 2.25 dB 개선된 성능을 가짐을 보인다.

### 1. 서 론

움직임 추정 예측 부호화를 수행하는 MPEG-1/2, H.264/AVC 등의 부호화는 복호화에 비해 5 ~ 10배 이상의 많은 연산을 수행한다. 이와 같은 '고복잡도 부호화와 저복잡도 복호화'의 비대칭구조의 코덱은 방송 또는 VoD 등과 같은 하향링크(downlink) 응용에 적합하다. 하지만, 계산 성능 및 전력이 제한되는 환경의 무선 저전력 감시 카메라, 멀티미디어 센서 네트워크 등의 상향링크(uplink) 응용에서는 복호기에서는 상대적으로 높은 복잡도가 수용되는 반면 부호기의 복잡도는 최소화하여야 하기 때문에 기존의 코덱은 부적합하다. 또한 양방향의 모바일 통신에서도 저복잡도의 비디오 부호화가 요구된다.

이러한 기존의 고복잡도 부호화기의 문제를 해결하는 방안으로서 분산 비디오 부호화(DVC)가 최근 활발히 연구되고 있다. DVC는 Slepian-Wolf[1]와 Wyner-Ziv[2]의 기본적인 정보이론에 기반을 두고 있다. Slepian-Wolf 이론[1]은 기존의 부호화 기법과 같이 부호기측과 복호기측에서 예측참조 영상(또는 보조정보(Side Information: SI)을 이용하여 부호화하거나, 또는 부호화기에서 독립적으로 부호화하고 복호기측에만 예측참조영상을 이용하여 복호화를 하더라도 원 영상 정보에 대한 최소의 정보량은 변하지 않음을 보인 것이다. Slepian-

Wolf 이론은 무손실 압축을 가정한 반면 Wyner-Ziv 이론[2]은 손실 압축을 할 경우에도 기존과 동일한 비트율-왜곡 특성을 얻을 수 있음을 증명 하였다. 결국 이러한 두 이론은 기존의 부호기에서 구현된 움직임 예측/보상 과정을 복호기측에서 처리하도록 하여 낮은 복잡도의 부호기를 구현할 수 있는 이론적 배경을 제시하고 있다.

분산 비디오 부호화 기법에서 보조정보는 송신측에서 제공하는 WZ 프레임과 매우 유사한 신호 성분으로 구성되지만 동일하지 않으며, 이러한 WZ 프레임과 보조정보의 차이 성분은 전송 중에 발생한 잡음으로 가정하고 가상 채널(virtual channel) 잡음으로 모델링 한다. 그러나, WZ 복호기에서 원 영상에 대한 정보를 갖지 않으므로, 가상 채널 잡음에 대한 정확한 모델링은 매우 어려우며, 가상 채널에서 삽입된 잡음의 양과 잡음이 발생된 위치를 정확하게 파악하는 것은 현실적으로 불가능하다. 그래서 수신측에서 가상채널 잡음에 의한 왜곡을 복구하도록 하기 위해, 송신측에서는 WZ 프레임을 채널 부호화하고 그 패리티 정보를 전송하고, 수신측에서는 수신된 패리티 정보를 이용하여 가상 채널 잡음으로 왜곡된 신호를 효과적으로 복구하는 방안에 대해 많은 연구가 진행되었다[3]-[5].

그러나, 분산 비디오 부호화 기법에서는 부호기의 연산 능력 등의 자원이 제한으로 인해 그 부호화 성능이 기존의 부호기의 부호화 성능에 비해 떨어지는 한계가 있다.

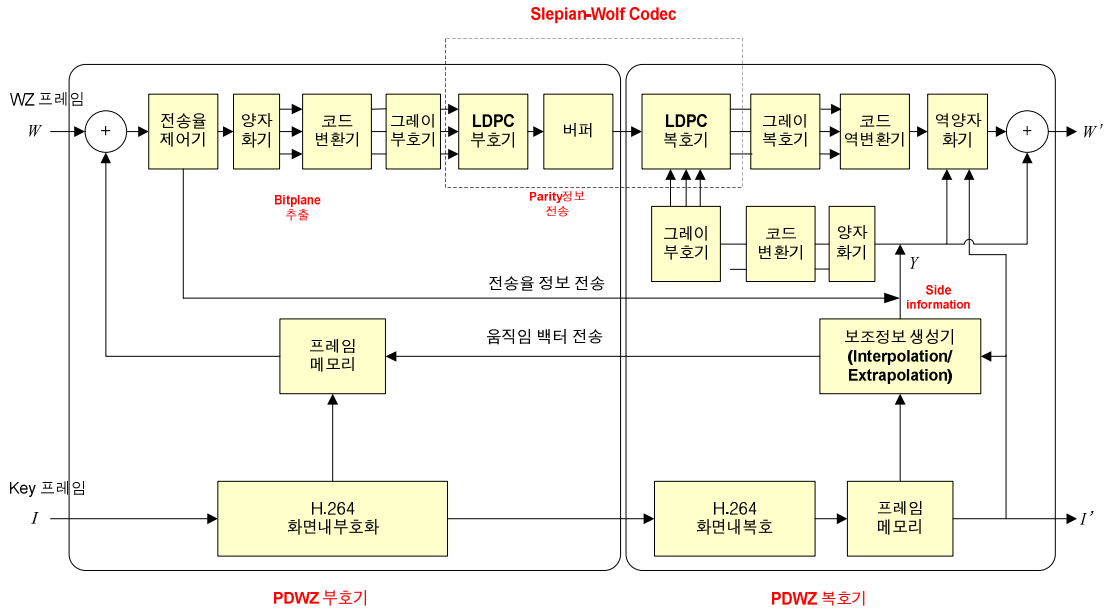


그림 1. 제안하는 PDWZ 코덱의 구조

이와 같은 분산 비디오 부호화 기법의 성능 한계를 극복하기 위한 많은 기법이 제안되었다. J. Kim[6]은 부호기에서 코드변환과 그레이코드 변환과 같이 간단한 연산을 통해 부호화 성능이 개선될 수 있음을 보였다. S. Park[7]은 생성된 보조정보를 구성하는 블록들을 움직임 정합 척도와 주위 블록간의 상관성에 기초하여 비용을 정의하고, 높은 비용을 갖는 블록에 대해서만 패리티 정보를 요구하도록 함으로써 전송 비트량을 절감시킬 수 있는 방안을 제시하였다. 또한 J. Kim[8]은 높은 비용이 소요되는 블록에 대해 움직임 정보를 피드백시키고, 송신 측에서는 움직임 보상 예측오차 신호에 대해서 [6]에서 사용된 코드변환을 이용해 패리티 정보를 전송함으로써 부호화 효율을 향상시킬 수 있는 구조를 제시하였다.

그러나, [7]과 [8]에서 사용한 비용은 보조정보 생성에 있어서, 시공간적으로 얼마나 정합이 잘 되어 있는지를 나타내는 척도이며, 복호기에서는 WZ 프레임을 알 수 없기 때문에 정의된 비용이 실제 WZ 프레임과 보조정보 사이의 오차의 양을 근사화하기 어렵다. 또한 [7]과 [8]에서는 비용에 대한 임계값을 두지 않고, 비용이 높은 50%만 전송하기 때문에 선택적 부호화 방식이 영상 특성에 적응적으로 효과적이지 못하다는 단점이 있다.

따라서 본 논문은 [7]과 [8]에서 사용된 기법을 개선하여 부호화 효율을 개선시키기 위한 방안을 제안한다. 제안한 부호화 기법은 [8]에서 사용된 움직임 정보 피드백 방식을 적용하여, 부호화기에서 기 저장된 키 프레임과 피드백된 움직임 정보를 이용하여 보조정보를 복원하고, 복원된 보조정보와 WZ 프레임 간의 비트 플레인 별 비트 오류율(Bit Error Rate: BER)을 구하고 이를 이용하여 블록 별로 전송되어질 패리티 정보의 양을 예측한다. 패리티 전송 시에 WZ 프레임과 보조정보 간의 비트 오류가 완벽히 복구된다고 가정하였을 때의 WZ 프레임과 복호기 측의 복호화된 프레임 사이의 왜곡을 이용하여 블록 단위의 윌-왜곡 비용을 정의하고, 블록 별로 패리티 정보를 보낼 경우와 보내지 않을 경우 비용을 비교

하여 패리티 정보를 보낼 영역을 적응적으로 선택하여 움직임 보상 오차 신호를 [6]에서 사용한 코드변환과 그레이코드로 변환하여 부호화 효율을 향상시킬 수 있는 구조를 제시한다.

본 논문의 구성은 다음과 같다. 우선, 제 II 장에서는 화소영역(Pixel Domain) Wyner-Ziv (PDWZ) 코덱의 구조를 제시하고, 부호화 성능을 개선할 수 있는 윌-왜곡 기반 선택적 부호화 기법을 제안한다. 제 III 장에서는 모의실험을 통하여, 본 논문에서 제안한 윌-왜곡 기반 선택적 부호화 기법의 부호화 효율을 기존의 기법과 비교 분석한다. 끝으로 제 IV 장에서는 결론 및 향후 연구 과제에 대해 논의한다.

## 2. 선택적 부호화 기법

### 2.1 PDWZ 코덱 구조

PDWZ 코덱은 홀수 프레임은 기존의 화면 내 부호화 기법으로 부호화하여 전송하고, 짝수 프레임에 대해서는 양자화를 수행하고 그 결과를 채널 부호화하여 얻은 패리티 정보를 전송한다. PDWZ 코덱은 부호기에서 단순히 화소값을 부호화하거나 참조 영상과의 화소 차이 값을 부호화함으로써, 부호기의 구조를 간단화하고 계산량을 절감시킬 수 있는 장점이 있다.

홀수 프레임은 키 프레임으로서 기존의 H.264/AVC 화면 내 부호화 방식을 통하여 부호화되어 전송된다. 짝수 프레임은 WZ 프레임으로서 스칼라 양자화를 적용하여 정보량을 줄이고, 각 비트 플레인 단위로 채널 부호화 후에 패리티 정보만을 복호기 측으로 전송한다. 복호기 측에서는 홀수 프레임으로 수신되는 키 프레임을 복호하고, 복호된 영상은 프레임 메모리에 저장된다. 프레임 메모리에 저장된 키 프레임의 정보를 이용하여 WZ 프레임의 예측신호인 보조정보를 생성한다. 생성된 보조정보는 부호기 측의 원 WZ 프레임과 매우 유사하도록 생성되며, 이 정보는 부호기측과 동일하게 양자화되고 각 비

트 플레인 단위로 수신되는 패리티 정보에 대한 유료부하로 사용되어, LDPC 복호를 수행하고, 복호된 신호는 역양자화되어 WZ 프레임에 대한 복원된 영상을 얻게 된다.

## 2.2 율-왜곡 기반 선택적 부호화 기법

그림 1은 본 논문에서 제안하는 PDWZ 코덱의 구조를 나타낸다. 본 논문에서는 식 (1)과 같이 율-왜곡 비용을 [9]을 정의하고 이를 바탕으로 블록 별로 선택적으로 WZ 프레임을 부호화한다.

$$J_i = D_i + \lambda \cdot R_i \quad (1)$$

$D_i$ 와  $R_i$ 는  $i$  번째 블록에 대한 왜곡과 전송률 예측값을 의미하며,  $\lambda$ 는 가중치이다. 상기 비용함수를 이용하여 프레임 차 신호의 각 블록에 대한 패리티 정보를 전송할 경우와 전송하지 않을 경우의 비용을 비교하여 패리티 정보를 전송할 경우의 비용이 적은 경우에는 해당 블록의 패리티 정보를 전송하도록 결정하고, 패리티 정보를 전송하지 않을 경우의 비용이 적은 경우에는 해당 블록의 패리티 정보를 전송하지 않도록 결정한다.

수식 (1)에서  $D_i$ 는 식 (2)와 같이 구해지는데, 이는 패리티 정보가 전송되는 비트 플레인의 비트 에러는 모두 완벽하게 정정이 된다는 가정을 한다.

$$\begin{aligned} D_i &= \sum_b D_i^b \\ &= \sum_b 2^{(8-b)} \cdot Ber_i^b \\ 0 \leq i &< height/8, \quad 0 \leq j < width/8 \end{aligned} \quad (2)$$

이때,  $b$ 는 전송될 패리티의 비트 플레인을 인덱스이고,  $Ber_i^b$ 은 WZ 프레임의  $i$  번째 블록과 보조정보의 동일 위치의 블록 간의  $b$  번째 비트 플레인에서의 비트 오류 수이다.

한편, 수식 (1)의 비용함수에서  $R_i$ 를 구하기 위해 하나의 LDPCA 프레임 크기( $K$ )에 대해서 그림 2와 같이 실험적으로 모델을 찾을 수 있다. 이는  $K$ 에 발생한 비트 오류 수에 대해서 모든 에러를 정정하기 위해 전송이 요구되는 비트 수를 실험 데이터로부터 모델링한 것으로 식 (3)으로 근사화 된다.

$$\begin{aligned} R(K) &= 100\sqrt{Ber(K)} \quad , \quad Ber(K) < 300 \\ R(K) &= 3 \cdot Ber(K) + 800 \quad , \quad Ber(K) \geq 300 \end{aligned} \quad (3)$$

이 때,  $R(K)$ 는  $K$  크기의 LDPCA 블록에 대한 예측 소요 비트량이며,  $Ber(K)$ 는  $K$  크기의 LDPCA 블록에 발생한 비트 오류의 수이다. 따라서  $i$  번째 블록에 대한 예측 소요 비트량은 식 (3)에서 구한 하나의 LDPCA 프레임 크기에 대한 예측 소요 비트량을 이용하여 하나의 블록 크기에 대하여 식 (4)와 같이 근사화 한다.

$$R_i = R(K)/n \quad (4)$$

이때,  $n$ 은  $K$  크기의 LDPCA 블록을 구성하는 블록의 수이며, 모든 블록의 예측되는 소요 비트량은 동일하도록 예측된다.

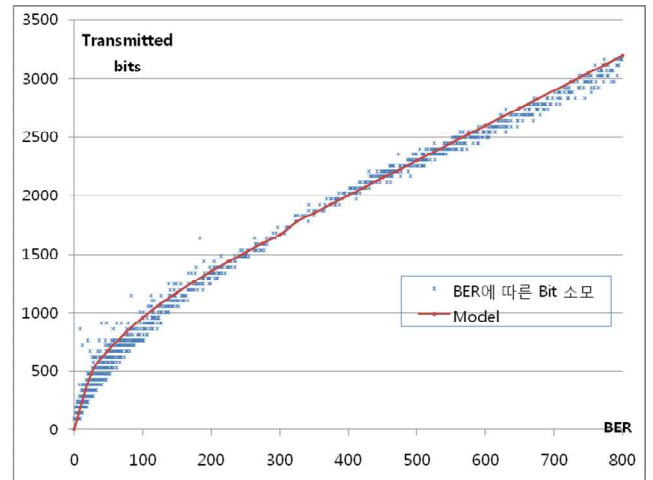


그림 2. 소요 비트량의 실험 데이터 및 모델

$\lambda$ 는 식 (5)와 같이 구해진다[9].

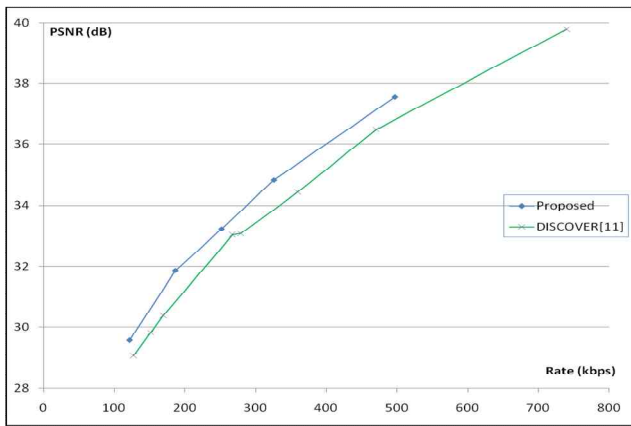
$$\lambda = - \frac{dD}{dR(D)} \quad (5)$$

식 (5)는 비용함수의 값이 최소가 될 때의  $\lambda$  값이며, 이와 같은 방법으로 율, 왜곡을 추정하고 식 (1)에 정의한 비용함수를 이용하여 패리티 비트를 전송할 경우와 전송하지 않을 경우에 대한 비용을 비교하여, 패리티 정보를 전송할 경우의 비용이 적은 경우에는 해당 블록의 패리티 정보를 전송하도록 결정하고, 패리티 정보를 전송하지 않을 경우의 비용이 적은 경우에는 해당 블록의 패리티 정보를 전송하지 않도록 결정한다. 또한 블록 별 패리티 전송 여부에 대한 정보를 추가적으로 전송한다.

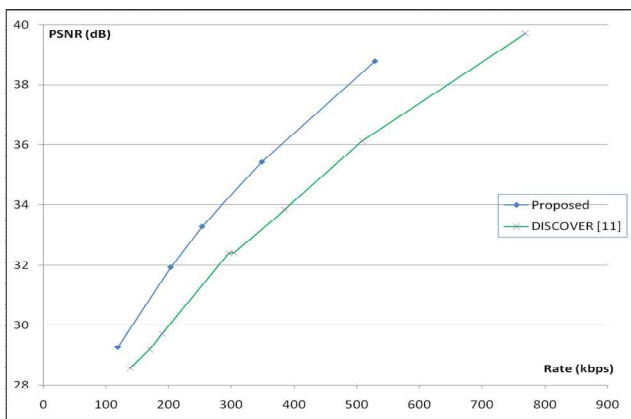
## 3. 실험결과 및 분석

모의실험을 위해 QCIF, 30Hz의 Foreman (291프레임), Salesman (291)프레임을 사용하였다. 먼저, 홀수 프레임은 키 프레임으로 설정하고 H.264/AVC의 화면내 부호화 방식을 이용하여 양자화 파라미터(QP)를 고정시킨 후에 가변 비트율로 부호화하였다. 그리고, 키 프레임 사이에 위치하는 WZ 프레임은 그림 1에 나타낸 PDWZ 구조로 부호화 하였다. 제안된 PDWZ 구조에 사용된 움직임 정보의 피드백 전송에 의한 선택적 부호화 기법의 효과 및 성능을 모의실험을 통하여 분석하기 위해 LDPCA를 각 비트 플레인에 동일하게 적용하였다[10].

그림 3은 Foreman과 Salesman 시퀀스에 대한 모의실험 결과를 나타내고 있다. Foreman 시퀀스는 공간적인 복잡도보다 시간적인 복잡도가 상대적으로 높은 시퀀스이며, DISCOVER 코덱[11]의 경우에는 영역별 보조정보의 화질, 소요 비트량 등을 고려하지 않고 모든 영역의 패리티 정보를 전송한다. 그러나 제안하는 방식은 프레임에 따라 패리티 정보를 전송하는 블록의 수를 적응적으로 결정할 수 있기 때문에 제안한 방식이 DISCOVER 코덱[11]에 비해 최대 약 1 dB 정도의 우수한 성능을 보인다.



(a) Foreman (30Hz)



(b) Salesman (30Hz)

그림 3. 선택적 부호화 방식의 부호화 성능

한편, Foreman 시퀀스와 달리, Salesman 시퀀스는 상대적으로 공간적인 복잡도는 높고 시간적인 복잡도는 매우 낮은 시퀀스이다. 따라서 모든 코덱에 대해 키 프레임의 화면내 부호화에 대한 성능은 낮다고 할 수 있으며, 본 논문에서 제안한 방식은 Salesman 시퀀스의 특성에 따라서 페리티 정보를 전송하지 않는 블록의 수를 적극적으로 더 높일 수 있기 때문에 시간적 복잡도를 효율적으로 이용할 수 있으므로 DISCOVER 코덱[11]보다 최대 약 2.2 dB 더 우수한 성능을 보인다.

#### 4. 결론

본 논문에서는 PDWZ 코덱의 부호화 효율을 개선하기 위해 복호기 측에서 보조정보 생성시에 발생하는 움직임 정보를 피드백하여 보조정보와 WZ 프레임 간의 유효-왜곡 최적화를 통해 선택적으로 부호화하는 분산 비디오 부호화 방식을 제안하였다. 제안한 방법은 부호기에서 영상 특성에 적응적으로 블록의 수를 정하여, 움직임 보상 오차 신호를 부호화함으로써 기존의 방식에 비해 시퀀스에 따라 최대 약 2.2 dB의 부호화 성능을 개선시킬 수 있었다.

본 논문에서 제안한 PDWZ 코덱 구조는 부호기와 복호기가 상호 정보 교환이 가능한 환경에 유용하게 활용될 수 있을 것으로 보인다. 앞으로 복호기 측으로부터 피드백된 움직임 정보를 이용하여, 부호기 측에서 블록 별

R-D 모델링에 대한 복잡도를 줄이기 위한 연구가 필요하며, 기존의 국제 영상 압축 표준안의 화면 간 부호화 방식의 부호화 효율에 더욱 근접하는 방식에 대한 연구가 필요하다.

#### 감사의 글

이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업[2010-0016880]과 경기도의 경기도지역협력연구센터(GRRC)사업[GRRC-항공2010]의 지원을 받아 수행된 것입니다.

#### 참고문헌

- [1] D. Slepian and J. Wolf, "Noiseless Coding of Correlated Information Sources", IEEE Trans. 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 Trans. Information Theory, vol. 22, pp. 1-10, Jan. 1976.
- [3] L. Qing, X. He, R. Lv, "Modeling Non-stationary Correlation Noise Statistics for Wyner-Ziv Video Coding," In Proc. ICWAPR '07, vol. 1, Nov. 2007, pp. 316 - 320.
- [4] Z. Li, L. Liu and E. J. Delp, "Rate-distortion Analysis of Motion Side Estimation in Wyner-Ziv Video Coding", IEEE Trans. Image Processing 16, pp. 98-113, Jan. 2007.
- [5] E. Peixoto, R. L. Queiroz and D. Mukherjee, "On Side Information Generation for Wyner-Ziv Video Coding", in Proc. of Simposio Brasileiro de Tel., Sep. 2008.
- [6] J.-s. Kim, J.-G. Kim, and K.-d. Seo, "A PDWZ Encoder Using Code Conversion and Bit Interleaver," J. Broadcasting Engineering, vol. 15, no. 1, pp. 52-62, Jan. 2010.
- [7] S.-U. Park, J.-W. Choi, C.-S. Kim, S.-U. Lee and J.-W. Kang, "Efficient Distributed Video Coding Using Symmetric Motion Estimation and Channel Division", in PACRIM09, Aug. 2009.
- [8] J.-s. Kim, J.-G. Kim, "A Selective Block Encoding Method with Feedback of Motion Information in Distributed Video Coding," submitted to J. Broadcasting, 2010.
- [9] G. J. Sullivan and T. Wiegand, "Rate-Distortion Optimization for Video Compression," IEEE Signal Processing Mag. Nov. 1998.
- [10] D. Varodayan, A. Aaron and B. Girod, "Rate-adaptive codes for distributed source coding", EURASIP Signal Processing, vol. 86, no.11, pp.3123-3130, Nov. 2006.
- [11] <http://www.discoverdvc.org>