

# SD/MD 전환을 이용한 하이브리드 다중 표현 동영상 압축 방법

김일구, 조남익

서울대학교 전기컴퓨터공학부

lit2eng@ispl.snu.ac.kr, nicho@snu.ac.kr

## Hybrid Multiple Description Video Coding Using SD/MD Switching

Il Koo Kim, Nam Ik Cho  
Seoul National Univ. EECS

### Abstract

본 논문에서는 하이브리드 다중 표현(hybrid multiple description) 개념을 이용하여 에러가 발생하기 쉬운 환경에서 사용될 수 있는 강인한 동영상 압축 알고리즘을 제안한다. 다중 표현 압축은 단일 표현 압축(single description coding, SDC) 보다 패킷 손실율(packet loss rate, PLR)이 높은 환경에서 에러에 더 강인성을 갖는다. 그러나 PLR 이 낮은 환경에서는 과도한 중복량(redundancy) 때문에 성능이 저하되는 문제가 발생한다. 이러한 MDC 의 문제를 해결하기 위해서 채널 환경에 따라 SD/MD 를 전환할 수 있는 방법을 제안한다. 좀더 자세히 설명하면, 낮은 PLR 에서는 압축 효율을 위하여 SDC 를 사용하고 높은 PLR 환경에서는 에러에 대한 강인성을 위해 MDC 를 사용한다. SD/MD 전환을 최적화 시키기 위해 비트율-왜곡 최적화 프레임워크(rate-distortion optimization framework)을 사용한다. 부호화시에 복호기에서의 왜곡(distortion)을 정확히 예측하기 위해서 ROPE(recursive optimal per-pixel estimate) 방법을 사용한다. 모의 실험 결과 제안된 SD/MD 전환 방법이 기존의 SDC-ROPE 와 MDC-ROPE 보다 모든 에러 환경에서 더 효과적임을 알 수 있다.

### 1. 서론

음성, 오디오, 이미지, 비디오를 포함하는 대부분의 멀티미디어 정보들은 패킷(packet)이라 불리는 작은 크기의 단위로 나뉘어져 네트워크를 통해 전송된다. 그러나 패킷 스위칭 네트워크(packet switching network)은 중단간의 QoS(quality of service)를 보장해 주지 않는다. 널리 사용되는 인터넷(internet)의 경우 데이터 전송량이 증가하면 트래픽 썸(traffic jam)이 발생한다. 이로 인해 발생한 버퍼 오버플로우(buffer overflow)와 긴 큐잉 지연(queueing delay)에 의해 네트워크 상의 중간 노드들이 임의로 패킷을 제거하게 된다. 이러한 환경에서 에러에 대한 강인성 높이기 위해서 다중 표현 압축(multiple description coding, MDC)이 연구되어 왔다. 다중 표현 압축에서는 입력 신호는 디스크립션(description)이라고 불리는 여러 개의 비트열로 나뉘어져 물리적으로 독립적이거나 가상적으로 만들어진 독립적인 채널들을 통해서 복호기로 전송된다. 다중 표현 압축의 장점은 복호기에서 모든 디스크립션을 받지 못했을 때에도 입력 신호를 복원할 수 있다는 데 있다. 더 자세히 설명하면, 다중 표현 압축은 입력 신호를 서로 연관된 여러 개의 디스크립션으로 나누게 된다. 결과적으로 정상적으로 받지 못한 디스크립션이더라도 성공적으로 복호된 신호를 이용하여 예측(estimation) 할 수 있는 것이다. 다중 표현 압축의 중요한 이슈는 각 디스크립션 사이의 연관을 어떻게 효과적

로 만들 것인가 하는 점과 연관의 양을 효과적으로 조절하는 것이다.

최근들어 비트율-왜곡 최적화(rate-distortion optimization)에 기반을 둔 다중 표현 비디오 압축 방법이 제안되었다[1]. 이 방법에서는 독립적인 세개의 예측 루프를 사용하지만, 측면 예측 루프에서 만들어진 나머지 신호는 복호기에 전송되지 않는다. 측면 예측 루프는 패킷 손실이 발생했을 때 부호기와 복호기 사이의 불일치(mismatch)를 줄이기 위한 비트율-왜곡 최적화에만 사용된다. 측면 예측 루프에서 만들어진 나머지 신호는 완전히 부가적인 정보이기 때문에 이러한 신호를 보내지 않음으로 해서 높은 압축 효율을 얻을 수 있다. 비슷한 시기에 이와 유사한 방법이 제안되었다[2]. [2]에서는 입력 신호를 높은 레졸루션(resolution)의 디스크립션과 낮은 레졸루션(resolution)의 디스크립션으로 불균등하게 나눈다. 부호기와 복호기 사이의 불일치를 막기 위해서 부호화할 때 비트율-왜곡 관점에서 복호기에서의 정확한 왜곡을 예측하는 것이 필요하다. [3]에서 화소 단위로 복호기의 왜곡을 예측하는 ROPE(recursive optimal per pixel estimate) 방법을 제안하였다(SDC-ROPE). ROPE는 SDC에서 인트라 블록의 삽입 위치를 최적화하기 위해서 제안되었다. [3]에서는 SDC-ROPE가 확장되어 MD 부호기에서 최적의 인트라 블록의 삽입과 MDC 중복량을 동시에 결정하기 위해서 사용되었다. MDC-ROPE는 2% 이상의

PLR에서 SDC-ROPE 보다 성능이 향상된다. 그러나 PLR이 2% 이하가 되면 피할 수 없는 중복량 때문에 SDC-ROPE이 보다 성능이 저하된다.

본 논문에서는 기존의 MDC에서 과도한 중복량 문제를 해결하기 위해서 비트율-왜곡에 기반을 두고 SD/MID 전환을 수행하는 하이브리드 MDC 알고리즘을 제안한다. 예를 들면, SDC는 압축 효율을 위해서 낮은 PLR에서 사용되고 MDC는 에러에 대한 강인성을 높이기 위해서 PLR이 높은 경우에 사용된다. SD/MID 전환을 최적화하기 위해서 비트율-왜곡 최적화 프레임워크를 사용한다. 라그랑제 승수법을 사용하여 목적 함수  $J = D + \lambda R$ 를 최소화하는 모드를 선택하게 된다.

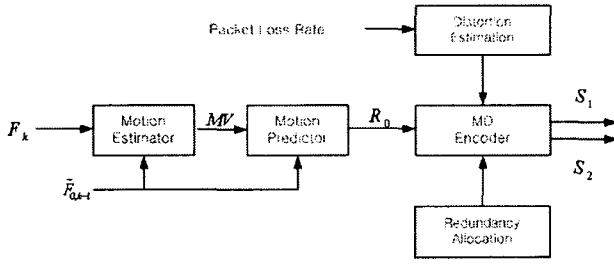


그림 1. 하이브리드 MDC 전체 프레임 워크

## 2. 하이브리드 다중 표현 비디오 압축 방법

### 2.1. 전체 프레임 워크

본 논문에서는 전통적인 부호기의 출력 DCT 계수들을 두 개의 디스크립션으로 나누는 단순한 방법을 사용한다. 완전한 불일치 조절을 위해서는 측면 예측 에러 신호를 전송해 주어야 하지만, [1]에서 계수들의 단순한 번갈음(alternation)과 중복(duplication)만으로도 효과적으로 계수들을 분리할 수 있다는 것을 보였기 때문에 측면 예측 에러 신호를 보내는 대신, 비트율-왜곡 최적화 방법을 사용한다.

그림 1.은 두개의 채널을 가정했을 때, 제안하는 다중 표현 비디오 압축기의 구조를 보여주고 있다. 2개 이상의 채널에 대한 구조 및 알고리즘들도 이와 같은 2개 채널의 경우로부터 쉽게 유도할 수 있다. k 번째 프레임  $F_k$ 와 k-1 번째로 복호된 프레임  $\tilde{F}_{0,k-1}$ 가 주어져 있을 때 부호기는 블록 기반 움직임 벡터, MV를 구한다. MV는 움직임 예측기를 통해서 예측 에러 신호  $R_0$ 를 생성한다. 예측 에러 신호  $R_0$ 는 MD 부호기에 의해서 두개의 표현으로 나눠지고 두개의 독립적인 채널을 통해서 복호기로 전송된다. MD 부호기는 주어진 PLR 환경과 중복량 조건을 고려하여 DCT 계수 최적 분할과 SD/MID 전환을 수행하게 된다. DCT 계수 최적 분할, SD/MID 전환, 복호기 왜곡 예측 방법은 아래에서 더 자세히 설명한다.

### 2.2. DCT 계수 최적 분할

한 개의 레이어를 두개의 표현으로 분할 하기 위해서 [5]에서 제안된 DCT 계수 최적 분할을 사용한다. 다이나믹 프로그래밍 접근법을 사용하여 계수의 집합  $\tilde{C}_1$ 과  $\tilde{C}_2$ 를 구한다.  $\tilde{C}_1, \tilde{C}_2$ 의 각각의 계수들은 원본 블록  $\tilde{C}$ 의 계수들은 번갈음하거나 중복시켜서 구한다. 분할 된 계수들의 집합,  $\tilde{C}_1$ 은 첫번째 표현  $S_1$ 에 할당되고  $\tilde{C}_2$ 는 두 번째 표

현  $S_2$ 에 할당된다. 이러한 문제는 다음과 같은 제한 문제로 표현될 수 있다.

$$\min_{\tilde{C}_1, \tilde{C}_2} [D(\tilde{C}_1, \tilde{C}_2)] \text{ subject to } \rho(\tilde{C}_1, \tilde{C}_2) \leq \rho_{\text{budget}} \quad (1)$$

여기에서  $D$ 와  $\rho$ 는 각각 왜곡과 중복량이고  $\rho_{\text{budget}}$ 은 최대 사용할 수 있는 중복량이다. 식 (1)은 라그랑제 승수법을 이용하여 다음과 같은 목적 함수를 최소화하면 해결할 수 있다.

$$J(\lambda) = D(\tilde{C}_1, \tilde{C}_2) + \lambda \rho(\tilde{C}_1, \tilde{C}_2) \quad (2)$$

라그랑제 승수  $\lambda$ 는 왜곡과 중복량의 최대값을 고려해서 결정한다. 그런데, 만약 j 번째 계수 까지의 최적의 해를 가지고 있다면 다음과 같은 식을 사용하여 다음 계수인 k 번째 계수의 문제를 풀 수 있다.

$$J_k = J_j^* + \Delta J_{j,k}^{op} \quad (k > j) \quad (3)$$

위 식에서  $\Delta J_{j,k}^{op}$ 는 MD 연산자 op를 사용했을 때의 라그랑제 코스트의 증가분이다. 이러한 재귀적인 구조는 다이나믹 프로그래밍 접근법을 가능하게 한다. 일반적인 k에 대하여 최소 코스트는 다음과 같은 식으로 구해 질 수 있다.

$$J_k^* = \min_{j, op} \{J_j^* + \Delta J_{j,k}^{op}\}, \text{ for } j=0, \dots, k-1 \quad (4)$$

이러한 재귀적인 구조를 이용하면 첫번째 계수에서부터 63번째 계수까지 최적의 라그랑제 코스트  $J_k^*$ 를 구할 수 있다. 라그랑제 코스트 뿐만 아니라 최적의 상속자 j와 최적의 연산자  $op_k^*$ 도 저장된다. 그 후,  $J_k^*$ 를 갖는 최적의 계수 k는  $k$ 로 표현될 수 있다. 명확히하면  $J_k^*$ 는 전체 블록에 대한 최소의 라그랑제 코스트가 된다. 이 값과 계산 과정에서 저장된 상속자와 최적 연산자를 이용하여 최적의 계수를 모두 찾아 나간다.

### 2.3. 복호기 왜곡 예측

[5]에서는 두개의 채널 중 하나의 채널이 모두 실패하는 이상적인 MD 채널 환경을 가정하였다. 그러나 이런 가정은 패킷 교환 네트워크에서는 더 이상 타당하지 않다. 패킷은 어떠한 공지없이 중간 노드에서 누락될 수 있다. 따라서 어떤 패킷이 사라졌는지 정확히 알 수 없고 일정 시간 동안의 패킷 손실율을 예측할 수 밖에 없다. 패킷 손실율을 사용하여 부호화 시에 복호기에서의 왜곡을 예측하는 것이 필요하다. PLR을 고려하면 왜곡 D는 다음과 같은 식으로 표현된다.

$$D(\tilde{C}_1, \tilde{C}_2) = \sum_{n=0}^{63} d_n^i \quad (5)$$

위 식에서  $d_n^i$ 는 다음과 같이 정의 된다.

$$d_n^i = E\{(f_n^i - \tilde{f}_n^i)^2\} = (f_n^i)^2 - 2f_n^i E\{\tilde{f}_n^i\} + E\{(\tilde{f}_n^i)^2\} \quad (6)$$

$f_n^i$ 는 n 번째 프레임의 i 번째 화소의 원본 값이고  $\tilde{f}_n^i$ 는 에러 은닉후 복호기에서 복원된 화소값이다. 부호기 측에서 보면,  $f_n^i$ 는 고정된 값이 아니라 PLR에 따른 확률 변수이다. 이러한 개념은 [4]에서 ROPE(recursive optimal per-pixel estimate)라 명명 되고 SD 비디오에 처음 적용되었고, [3]에서 확장되어 MD 비디오에 적용되었다. 본 논문에서는 부호기 측에서 복호기의 왜곡을 예측하기 위해서 절반의 픽셀 정확도를 갖는 ROPE를 사용한다.

## 2.4. SD/MD 전환

SD/MD 전환을 위해 필요한 것은, 비디오 시퀀스 안에 프레임 당  $M$  개의 블록을 가지는  $N$  장의 프레임이 있다고 가정하면 SD와 MD 중에서 다음의 식을 최소화하는 최적의 모드를 찾는 것이다.

$$\text{mode}^* = \min_{\text{mode}} E_{i,\text{mode}} \quad (7)$$

위 식에서  $E_{i,\text{mode}}$  는 다음과 같이 정의된 목적 함수이다.

$$E_{i,\text{mode}} = D_{i,\text{mode}} + \mu R_{i,\text{mode}} \quad (8)$$

위 식에서  $\mu$  는 상수이고,  $D_{i,\text{mode}}$  와  $R_{i,\text{mode}}$  는 각각  $i$  번째 프레임이 복원되었을 때의 왜곡과 비트율이다.  $\mu$  는 DCT 계수의 최적 분할에서 사용된 라그랑제 승수의 평균을 사용한다.  $D_{i,\text{mode}}$  와  $R_{i,\text{mode}}$  를 더 자세히 풀어 써 보면 다음과 같다.

$$D_{i,\text{mode}} = \sum_{j=0}^{M-1} d_{i,j,\text{mode}} \quad (9)$$

$$R_{i,\text{mode}} = \sum_{j=0}^{M-1} r_{i,j,\text{mode}} \quad (10)$$

위 식에서  $d_{i,j,\text{mode}}$  와  $r_{i,j,\text{mode}}$  는 각각  $i$  번째 프레임의  $j$  번째 블록의 복원 왜곡과 비트율이다.  $d_{i,j,\text{mode}}$  와  $r_{i,j,\text{mode}}$  는 위의 섹션에서 설명한 DCT 계수의 최적 분할을 통해서 얻어지는 값들이다.

## 3. 코덱 구현 및 비트 할당

제안된 MDC 코덱은 ITU-T H.263을 기준으로 구현되었다. 하이브리드 MDC에서는 기존의 SDC, MDC와는 달리 복호기에 SD/MD 전환 값들은 전달해 주어야만 한다. H.263의 문법 구조를 수정하지 않기 위해서 픽처 헤더에 포함된 PEI(extra insertion information)와 PSUPP(supplemental enhancement information)을 사용한다[6]. PSUPP 데이터는 4비트의 함수형 표시자 FTYPPE와 데이터 크기는 나타내는 4비트의 DSIZE로 구성된다. 지원되지 않는 함수형 표시자를 복호기가 받게 되면 함수의 인자 정보는 버려지게 된다. SD/MD 전환 정보를 전달하기 위해서 확장 함수형을 사용하기 때문에 이러한 함수형을 지원하지 않는 복호기는 SD/MD 정보를 잃게 된다.

비트율을 효과적으로 조절하기 위해서 TMN8 비트율 조절 알고리즘과 라그랑제 승수 선택 방법을 사용한다. 라그랑제 승수 선택 알고리즘은 [7]에서 제안되었다. 오프라인 실험을 통해서 가장 좋은 성능을 갖는 라그랑제 승수  $\lambda$  는 다음과 같이 결정된다.

$$\lambda = c \cdot (QP)^2 \quad (11)$$

이 식에 대한 이론적인 증명도 동일한 논문에 주어져 있다. 일반적인 H.263 코덱에 대해서는  $c = 0.85$  로 알려져 있다. 구현된 SD-ROPE에서  $c = 0.85$  를 사용하여 실험하였다. 그러나 MD 비디오는 일반적인 SD 비디오와 특성이 다르기 때문에 새로운 라그랑제 승수 선택방법이 필요하다. 따라서 새로운 방법을 제안한다.

$$\lambda = c \cdot (QP)^2 \cdot (-\log(PLR)) \quad (12)$$

MDC에서는 라그랑제 승수는 압축 효율을 결정할 뿐만 아니라 에러에 대한 강인성도 조절한다. 높은 PLR에서는 패킷 손실에 대항하기 위해서 더 많은 중복량이 필요하다. 따라서 PLR이 라그랑제 승수 선택 알고리즘에 포함되어야 한다. MD-ROPE와 Hybrid MDC에서는 수정된 식에  $c = 0.15$  를 대입하여 사용하였다. 이 상수는 다양한 모의 실험을 통해서 구해 졌다.

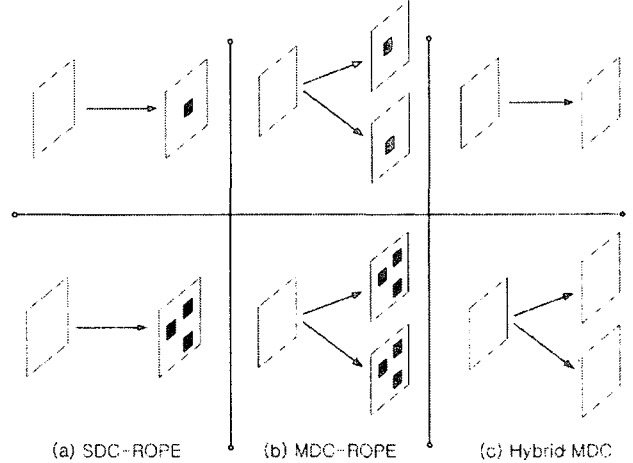


그림 2. 실험에 사용한 알고리즘들

## 4. 실험 결과

TMN8 비트율 조절 알고리즘과 수정된 라그랑제 승수 선택 방법을 사용하여 SDC-ROPE, MDC-ROPE, Hybrid MDC의 성능을 비교한다. 그림 2는 실험에 사용된 3개의 알고리즘을 개념적으로 설명하고 있다. 그림에서 첫번째 행은 PLR이 낮은 경우를 나타내고 두번째 행은 PLR이 높은 경우를 나타낸다. 검은색으로 칠해진 블록은 인트라 코딩된 블록을 나타낸다. SDC-ROPE와 MDC-ROPE는 인트라/인터 모드 전환을 사용하는데 반하여, Hybrid MDC는 채널 환경에 따라 SD/MD 전환을 수행한다. 그림 3과 4는 QCIF크기의 Mother & Daughter 비디오와 Foreman 비디오에 대한 패킷 손실 환경에서의 성능을 나타내고 있다. 각 비디오 시퀀스에 대하여 30개의 서로 다른 패킷 손실 패턴에 대한 실험을 수행한 후 평균값을 구하였다. 이러한 결과로부터 제안하는 Hybrid MDC 알고리즘이 높은 PLR 뿐만 아니라 낮은 PLR에서도 기존의 SD-ROPE와 MD-ROPE 보다 더 좋은 성능을 보임을 알 수 있다.

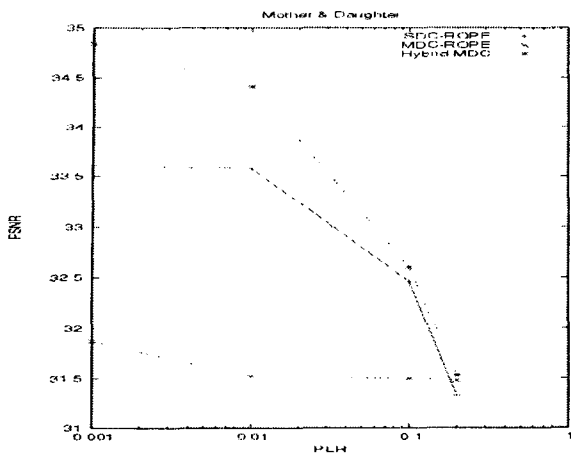


그림 3. Mother & Daughter 결과

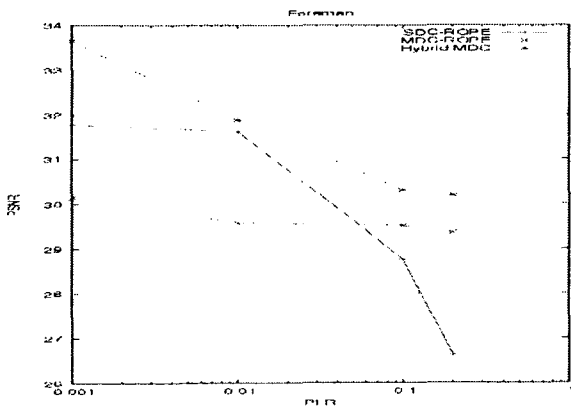


그림 4. Foreman 결과

### 5. 결론

SD/MD 전환을 이용한 하이브리드 MDC 방법을 사용하는 에러에 강인한 비디오 압축 알고리즘을 제안하였다. MDC의 과도한 중복량 문제를 해결하기 위해서 SD/MD 전환 개념을 사용한다. 낮은 PLR 환경에서는 압축 효율을 위해 SDC를 사용하고 높은 PLR 환경에서는 에러에 대한 강인성을 높이기 위해 MDC를 사용한다. SD/MD 전환을 최적화하기 위해서 비트율-왜곡 최적화 프레임워크를 사용한다. 라그랑주 최적화 방법을 이용하여 목적함수  $J = D + \lambda R$ 를 최소화한다. 부호기에서 복호기 왜곡을 정확히 예측하기 위해 ROPE 방법을 채용하였다. 모의 실험을 통해 제안하는 SD/MD 전환 방법이 기존의 SDC-ROPE, MDC-ROPE 방법보다 높은 PLR에서 뿐만 아니라 낮은 PLR에서도 더 좋은 성능을 보임을 확인하였다.

### 6. 참고 문헌

[1] A. Reibman, H. Jafarkhani, Y. Wang, and M. Orchard, "Multiple description video using rate-distortion splitting", in Proc. Int. Conf. Image Processing, May 2000, pp. 978-981

[2] D. Comas, R. Singh, and A. Ortega, "Rate-distortion

optimization in a robust video transmission based on unbalanced multiple description coding", In Proc. Of 2001 IEEE Fourth Workshop on Multimedia Signal Processing, Oct. 2001, pp. 581-586

[3] A. Reibman, "Optimizing multiple description video coders in a packet loss environment," In Proc. Int. Packet Video Workshop, Mar. 2002

[4] R. Zhang, S. Regunathan, And K. Rose, "Video coding with optimal Inter/Intra-mode switching for packet loss resilience," IEEE Journal on Selected Areas in Communications, Vol. 18, pp. 966-976. June 2000

[5] I. K. Kim and N. I. Cho, "Error resilient video coding using optimal multiple description of DCT coefficients," in Proc. Int. Conf. Image Processing, Sept.

[6] Z. Jia, K. Tang, and H.Cui,"A H.263 compatible error resilient video coder," In Proc. Int. Conf. WEE-ICCT 2000, Aug.2000, pp. 21-25

[7] T. Wiegand and B. Girod, "Lagrange multiplier selection in hybrid video coder control," in Proc. Int. Conf. Image Processing, Oct.2001, pp. 542-545