

특집논문-04-09-1-03

다중 표현을 이용한 에러에 강인한 동영상 부호화 방법

김 일 구*, 조 남 익*

Error Resilient Video Coding Techniques Using Multiple Description Scheme

Il Koo Kim* and Nam Ik Cho*

요 약

본 논문에서는 다중 표현(multiple description) 개념을 이용하여 에러에 강인한 동영상 부호화 방법을 제안한다. 제안하는 방법은 DCT 계수의 최적 분할 방법과 채널 환경에 따른 단일표현/다중표현 전환 방법으로 구성되어 있다. DCT 계수 최적 분할 방법에서는 입력 신호를 주어진 중복량(redundancy)에서 최적의 과잉 비트율-왜곡(redundancy rate-distortion, RRD) 성능을 갖는 두 개의 표현으로 분할한다. 최적화 방법으로는 라그랑제 최적화 방법(Lagrange optimization method)을 사용하였고 재귀적 구조를 사용한 다이내믹 프로그래밍 기법을 사용하여 분할의 복잡도를 줄인다. 단일표현/다중표현 전환 방법에서는 재귀적 최적 화소단위 예측(recursive optimal per-pixel estimate, ROPE)를 이용하여복원 에러를 예측한 후, 낮은 패킷 손실율에서는 압축 효율을 위하여 단일 표현을 사용하고 패킷 손실율이 큰 환경에서는 에러에 대한 강인성을 위해 다중표현을 사용한다. 모의 실험 결과, 제안하는 다중표현 동영상 부호화 방법은 이상적인 다중표현 채널에서 뿐만 아니라 다양한 패킷 손실율을 갖는 채널 환경에서도 기존의 단일표현 및 다중표현 에러 내성 부호화 방법보다 더 좋은 성능을보임을 알 수 있다.

Abstract

This paper proposes an algorithm for the robust transmission of video in error prone environment using multiple description coding by optimal split of DCT coefficients and rate-distortion optimization framework. In MDC, a source signal is split into several coded streams, which is called descriptions, and each description is transmitted to the decoder through different channel. Between descriptions, structured correlations are introduced at the encoder, and the decoder exploits this correlation to reconstruct the original signal even if some descriptions are missing. It has been shown that the MDC is more resilient than the single description coding(SDC) against severe packet loss rate condition. But the excessive redundancy in MDC, i.e., the correlation between the descriptions, degrades the RD performance under low PLR condition. To overcome this problem of MDC, we propose a hybrid MDC method that controls the SDC/MDC switching according to channel condition. For example, the SDC is used for coding efficiency at low PLR condition and the MDC is used for the error resilience at high PLR condition. To control the SDC/MDC switching in the optimal way, RD optimization framework are used. Lagrange optimization technique minimizes the RD-based cost function, $D + \lambda R$, where R is the actually coded bit rate and D is the estimated distortion. The recursive optimal per-pixel estimate technique is adopted to estimate accurate the decoder distortion. Experimental results show that the proposed optimal split of DCT coefficients and SD/MD switching algorithm is more effective than the conventional MDC algorithms in low PLR conditions as well as in high PLR condition.

Keywords : Multiple description coding, error resilient coding

* 서울대학교 전기컴퓨터공학부 뉴미디어통신공동연구소
 School of Electrical and Computer Engineering, Seoul National University, Institute of New Media & Communications
 ※ 본 연구는 장은공의재단의 지원으로 이루어 졌습니다.

I. 서론

음성, 오디오, 정지 영상, 동영상을 포함하는 대부분의 멀티미디어 정보들은 패킷이라 불리는 작은 크기의 단위로 나뉘어 네트워크를 통해 전송된다. 그러나 패킷 스위칭 네트워크는 종단간의 서비스 품질(quality of service, QoS)을 보장해 주지 않는다. 널리 사용되고 있는 인터넷의 경우 데이터 전송량이 증가하면 데이터 체증이 발생한다. 이로 인해 네트워크 상의 중간 노드들에서 버퍼 오버플로우와 긴 큐잉 지연이 발생하고, 중간 노드들은 이에 대응하기 위해 임의로 패킷을 제거하게 된다. 더욱이, 무선 네트워크에서는 다중경로 페이딩과 가리어짐(shadowing) 때문에 패킷 손실율이 매우 크다. 이러한 문제를 해결하기 위하여 많은 네트워크에는 자동재전송요구(automatic repeat request, ARQ)와 같은 패킷 재전송 방법이 구현되어 있다. 수신기에서 패킷 손실을 발견하면, 수신기는 승인 프로토콜(acknowledgment protocol)을 통해서 패킷 손실을 송신기에 알려준다. 그러면 송신기는 손실된 패킷을 파악하여 재전송해 준다. 따라서 이러한 재전송 방법은 최소 한번의 라운드 트립 만큼의 불가피한 지연을 발생시키게 된다. 불가피한 지연은 대부분의 응용에서 무시될 수 있으나, 실시간 통신과 같은 경우에는 용인될 수 없다. 이러한 문제를 해결하기 위해 여러 내성 부호화 방법, 에러 은닉 방법, 부호기-복호기 상호작용을 이용한 방법들이 제안되었다.

실시간 통신에서 에러에 대한 강인성을 높이기 위하여 여러 내성 부호화방법의 일종으로 다중 표현 압축 방법(multiple description coding, MDC)이 연구되어 왔다. 다중 표현 압축에서는 입력 신호는 표현(description)이라고 불리는 여러 개의 비트열로 나뉘어 물리적 혹은가상적으로 독립적인 채널들을 통해서 복호기로 전송된다. 다중 표현 압축의 강점은 복호기에서 모든 표현을 받지 못했을 때에도 신호를 복원할 수 있다는 데 있다. 자세히 설명하면, 다중 표현 압축은 입력 신호를 서로 연관된 여러 개의 표현으로 나누어 준다. 결과적으로 정상적으로 받지 못한 표현이 있더라도 성공적으로 복호된 신호를 이용하여 손실된 표현을 예측할 수 있는 것이다. 이러한 다중 표현 압축의 중요한 이슈는 각 표현 사이의 연관을 어떻게 효과적으로 만들 것인가 하는 점과 연관의 양을 효과적으로 조절하는 것이다.

다중표현 문제는 1970년대 부터 정보이론 분야와 신호처리 분야에서 널리 연구되어 왔다^[1]. 신호처리 분야에서는 제일 먼저 음성 압축^[2]에 이용되었고 정지 영상 압축^[3]으로 확장되었으며, 현재 동영상의 다중 표현에 대한 연구가 활발

히 수행되고 있다^[4]. 다중 표현 동영상압축에서 가장 중요한 문제는 부호기와 복호기 사이의 불일치(mismatch)를 최소화 시키는 것이다. 부호기-복호기 불일치는 움직임 보상 예측에서 기인한다. 자세히 설명하면, 표현들이 전송 과정에서 손실되면 복호기에서는 정확히 받아진 정보와 손실된 표현의 예측 정보를 이용하여 저화질의 동영상을 얻는다. 그러나 이렇게 복호된 저화질의 동영상을 이용한 움직임 보상은 부호기에서 사용한 에러가 없는 영상을 이용한 움직임 보상과 차이가 있다. 이러한 불일치가 에러를 시간축으로 전파시키고 전체 동영상의 화질을 저하시키게 되므로 다중표현 동영상 압축 방법은 이와 같은 불일치를 줄이는 것이 목표이다.

이와 같은 기존의 다중 표현 동영상 압축방법들을 소개하면, 대부분의 다중 표현 동영상 압축 방법은 움직임 벡터와 부가 정보는 모든 표현에 중복적으로 포함시키고 단지 DCT 계수만을 여러 개의 표현으로 나누는데 집중하고 있다^{[5][4][6][7]}. 더욱이 단순한 구조를 위하여 하나의 입력 신호를 두 개의 표현으로 분할하는 경우만이 고려되어 왔다. 보다 자세히, 가장 먼저 제시된 대칭 인터 프레임 다중 표현 동영상 압축방법^[5]을 소개하면, 이 방법에서는 두 개의 예측 루프와 다중 표현 스칼라 양자화기(multiple description scalar quantizer, MDSQ)를 채택하였다. 예측을 수행한 후에 나머지 신호는 MDSQ를 이용하여 두 개의 대칭적인 표현으로 양자화 한다. 이러한 접근 방법은 예측 루프에 채택된 등성등성 양자화기 때문에 압축 성능이 크게 저하된다. 그 후에 세계의 예측 루프(중앙 예측, 두개의 측면 예측)를 사용하는 다중 표현 동영상 압축방법^[4]이 제안되었다. 이 방법에서는 중앙 예측 루프에서 압축 성능과 에러에 대한 강인성을 고려하여 입력 신호를 두 개의 표현으로 분할한다. 두 개의 측면 예측 루프는 손실이 발생하여 불완전한 표현이 전달되었을 때 발생하는 부호기-복호기 불일치를 보상하기 위한 나머지 신호를 생성한다. 모든 표현이 손실 없이 완벽하게 전달된다면 측면 예측 루프에서 생성한 나머지 신호는 전혀 사용되지 않는 중복적인 정보가 된다. 이러한 문제의 해결책으로 에러가 발생하지 않을 경우에는 측면 예측 루프에서 생성된 나머지 신호가 동영상의 화질을 향상시키는데 사용되도록 하는 방법^[8]도 제안되었다.

보다 최근에는 이러한 중복 비트율까지 고려한 중복비트율-왜곡 최적화(redundancy rate-distortion optimization, RRD)에 기반한 다중 표현 동영상 압축 방법이 개발되었다^[6]. 이 방법에서는 세 개의 예측 루프가 사용되지만 이전의 방법과는 달리 측면 예측 루프에서 생성하는 나머지 신호가

명시적으로 복호기에 전송되지 않는다. 두 개의 측면 예측 루프는 에러가 발생했을 때 복원 에러와 부호기-복호기 불일치를 최소화하는 중복비트율-왜곡 최적화 과정에서만 이용된다. 부가적인 불일치 제어를 수행하지 않더라도 완전히 중복적인 정보를 보내지 않기 때문에 에러에 대한 강인성 뿐만 아니라 높은 압축 효율도 얻을 수 있다. 이와 유사한 방법으로 입력 신호를 높은 해상도 표현과 낮은 해상도 표현으로 비대칭적으로 나누는 비대칭적 다중 표현 동영상 압축 방법도 제안되었다^[7]. 부호기-복호기 불일치를 최소화하기 위하여는 부호화할 때 중복비트율-왜곡 관점에서 정확한 복호기 왜곡을 예측하는 것이 필요하다. [9]에서는 이러한 복호기의 왜곡을 예측하기 위하여 재귀적 최적 화소단위 예측(recursive optimal per-pixel estimate, ROPE) 방법이 사용되었는데, 이는 단일표현 압축에서 최적의 인트라 블록의 위치를 선택하기 위하여 소개되었다^[10]. [9]에서 이러한 단일표현 ROPE는 다중표현 압축에서 각 블록에 대하여 최적의 중복량을 할당하는 것 뿐만 아니라 인트라 블록의 위치를 최적화하는 것으로 확장되었다. 다중표현 ROPE는 2% 이상의 패킷손실율에서 성능 향상을 보인다^[9]. 그러나 2% 미만의 패킷손실율에서는 움직임 벡터와 코딩 모드 등과 같은 불필요한 과도한 중복량 때문에 단일 표현 ROPE보다 성능이 떨어진다.

본 논문에서는 일반적인 다중표현 압축 방법에서는 피할 수 없는 과도한 중복량을 제거하기 위하여 중복비트율-왜곡 최적화에 기반한 DCT 계수 최적 분할 방법과 단일표현/다중표현 전환 방법을 제안한다. 즉, 다중표현 압축 방법에서는 각 표현들 사이의 상관관계를 만들어 주기 위한 피할 수 없는 중복량이 존재하며 낮은 패킷손실율에서는 이러한 중복량은 압축 효율을 크게 저하시킨다. 이러한 다중표현 압축의 문제점을 해

결하기 위하여 채널 환경에 따라 단일표현/다중표현을 전환하는 하이브리드 다중표현 압축 방법을 제안한다. 예를 들면, 낮은 패킷손실 환경에서는 압축 효율을 높이기 위하여 단일표현 압축 방법을 사용하고 높은 패킷손실 환경에서는 에러에 대한 강인성을 높이기 위해 다중표현 압축을 사용한다. 단일표현/다중표현 전환을 최적화 하기 위하여 중복비트율-왜곡 최적화 프레임워크를 설정하며 라그랑제 승수법을 사용하여 목적 함수 $D + \lambda R$ 를 최소화하는 모드를 선택하게 된다. R 은 실제 압축된 비트율이고 D 는 예측된 왜곡이다. 모의 실험 결과 제안하는 다중 표현 동영상 부호화 방법은 이상적인 다중표현 채널에서 뿐만 아니라 다양한 패킷손실율을 갖는 채널 환경에서도 기존의 단일표현 및 다중표현 에러 내성 부호화 방법보다 더 좋은 성능을 보임을 알 수 있다.

본 논문의 구성은 다음과 같다. 제 2절에서는 제안하는 다중 표현 동영상 부호기의 구조 및 DCT 계수 최적 분할 알고리즘에 대해 구체적으로 설명하고 제 3절에서는 단일표현/다중표현 전환 방법을 제시한다. 제 4절에서는 비트율 할당 방법 및 부호기와 복호기의 구현 방법에 대하여 설명하고, 제 5절에서는 실제로 구현한 부호기를 에러 환경에서 모의 실험한 결과를 보여준다. 마지막으로 제 6절에서는 본 논문을 정리하고 결론을 내린다.

II. 제안하는 다중 표현 동영상 부호기

1. 일반적인 다중 표현 동영상 부호기

그림 1은 두 개의 독립적인 채널이 존재할 경우의 다중

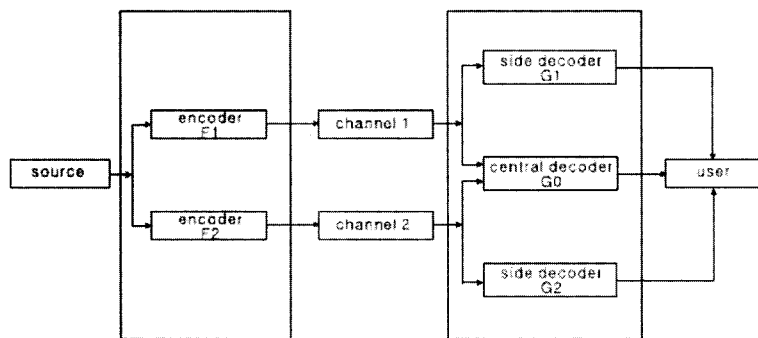


그림 1. 두 개의 독립적인 채널이 존재할 경우의 다중표현 부호화 개념도
Fig.1. Multiple description coding scheme with two independent channels

표현 부호화에 대한 개요를 보여주고 있다. 입력(source)은 분할 알고리즘에 의해 두 개의 표현으로 분할되고 분할된 각각의 표현들은 독립된 부호기들(F1, F2)에 의해 부호화 된다. 이렇게 분할 부호화된 표현들은 각각 독립적인 두개의 채널(channel 1, channel 2)를 통해 복호기에 전달된다. 복호화 과정은 두 가지로 나뉘어 진다. 첫 번째 경우는 두개의 표현을 모두 받은 경우이다. 이 경우에는 중앙 복호기(central decoder, G0)를 이용하여 복호화된다. 두 번째 경우는 하나의 표현만을 받은 경우이다. 이 경우에는 측면 복호기(side decoder, G1, G2)를 이용하여 복호화된다. 중앙 복호기를 이용한 복호화 과정에서는 부호화된 모든 정보를 수신하였기 때문에 부호화된 비트열을 모두 얻을 수 있다. 그러나, 측면 복호기를 이용한 복호에서는 오직 하나의 표현만을 이용하여 복호가 수행되기 때문에 복호된 비트열은 부호화된 비트열과 다르게 된다. 다중 표현 동영상 부호화에서 가장 중요한 문제는 이러한 부호기와 복호기 사이의 불일치(mismatch)를 최소화하는 것이다. 전송시 표현이 손실되면 부호기에서의 움직임 보상 프레임과 복호기에서의 복원 프레임이 다르게 된다. 이러한 부호기와 복호기 사이의 불일치는 움직임 보상 예측(motion compensated prediction)으로 인해 확대 전파되고 결과적으로 동영상 화질을 저하시키게 된다.

2. 제안하는 부호기 전체 구조

본 논문에서는 전통적인 단일계층, 단일표현 부호기에 의

해 생성된 DCT 계수를 두 개의 표현으로 분할하는 방법을 사용한다. 완전한 불일치 제어를 위하여는 측면 예측 루프에서 나머지 신호가 전송되어야한다. 그러나 [6][11]에서는 이러한 나머지 신호를 전송해 주지 않고, 중앙 예측 루프에서 생성된 DCT 계수를 단순히 번갈음과 중복시키는 방법만으로 분할하여도 성능이 크게 저하되지 않음을 보였다. 즉, 나머지 신호를 전송하지 않는 대신에 중복비트율-왜곡 최적화 방법을 이용하여 주어진 에러 환경에서 중복비트율-왜곡 성능을 개선한다.

그림 2는 두개의 독립적인 채널을 가정했을 경우 제안하는 다중표현 동영상 부호기의 전체 구조를 보여준다. 2개 이상의 채널에 대한구조 역시 두 개의 채널일 경우로 부터 쉽게 유도될 수 있다. 그림 2의 구조를 자세히 설명하면, 부호기는 k 번째 프레임 F_k 와 $(k-1)$ 번째 복호된 프레임들($\tilde{F}_{0,k-1}$, $\tilde{F}_{1,k-1}$, $\tilde{F}_{2,k-1}$)을 모두 이용하여 블록 기반으로 움직임 벡터를 예측한다. 움직임 벡터 MV 는 각각의 움직임 예측기에 적용되어 나머지 예측 오차 신호들 R_0 , R_1 , R_2 을 발생시킨다. 중앙 예측 에러 신호 R_0 는 두 개의 표현으로 분할되어 두 개의 독립적인 채널을 통해서 복호기로 전달된다. 측면 예측 에러 신호 R_1 와 R_2 는 측면 예측 오차(D_1 , D_2)를 계산하는데 사용된다. 만약 채널이 끊어진다면, 부호기에서의 참조 프레임 $\tilde{F}_{0,k-1}$ 과 복호기에서의 복호된 프레임 \hat{F}_{k-1} 과 달라지게 된다. 따라서 부호기-복호기 불일치가 발생하게 되고 성능이 저하된다. 제안하는 방법은 다이내믹 프로그래밍 기법을 사용하여 불일치를 최소화하고 에러에 대한 강인성을 높인다.

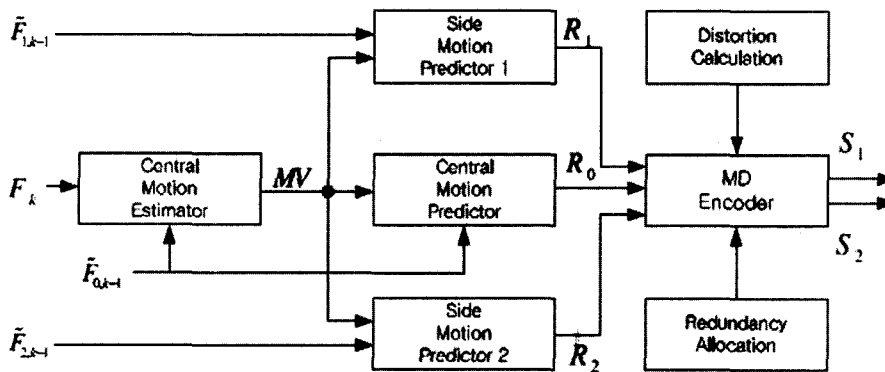


그림 2. 제안하는 다중 표현 동영상 부호기의 전체 구조
Fig. 2. Overall structure of proposed multiple description video coder

3. 다이나믹 프로그래밍을 이용한 최적화

DCT 변환된 입력 블록은 양자화 계수(Quantization Parameter)에 의해 양자화 되고 이 블록에 다이나믹 프로그래밍을 이용한 최적화 방법이 적용된다. 최적화의 목표는 원본 블록의 계수(집합 \tilde{C} 의 원소들)들을 각 표현에 모두 똑같이 할당하거나 배타적으로 번갈아 할당함으로써 만들어지는 각 표현들의 계수들의 집합(\tilde{C}_1, \tilde{C}_2)을 찾는 것이다. 즉, 경우에 따라 \tilde{C}_1, \tilde{C}_2 가 같은 집합이 될 수도 있고 \tilde{C} 의 원소를 \tilde{C}_1 과 \tilde{C}_2 에 나눌 수도 있다. 최적의 집합을 찾기 위하여 계수분할을 위한 세개의 연산(operation)을 정의한다. "duplication(dup)"은 모든 표현에 계수를 중복시키는 것을 의미하고, "alternation into 1(alt 1)"은 첫번째 표현 (S_1)에 계수를 할당하는 것을 의미하고, "alternation into 2(alt 2)"은 두번째 표현 S_2 에 계수를 할당하는 것을 의미한다. 이러한 연산들을 이용하여 찾아진 계수들의 집합 \tilde{C}_1 는 첫번째 표현(S_1)에 할당되고 \tilde{C}_2 는 두번째 표현(S_2)에 각각 할당된다. 이 문제는 다음과 같은 최소화문제로 공식화 할 수 있다.

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

위의 식에서 D 와 ρ 는 각각 복원 오차와 과잉 비트율이고 ρ_{budget} 은 미리 정해진 과잉 비트율의 최대값이다. 자세히 설명하면, C_1 과 C_2 를 각각의 측면 예측 에러 R_1 과 R_2 의 DCT 변환된 계수라고 할때 복원 오차 D 는 측면 복원 오차 $D_1(C_1, \tilde{C}_1)$ 와 $D_2(C_2, \tilde{C}_2)$ 의 합이다. $D_1(C_1, \tilde{C}_1)$ 은 다음과 같은 식으로 구해진다. ($D_2(C_2, \tilde{C}_2)$ 도 유사한 방법으로 구할 수 있다.)

$$D_1(C_1, \tilde{C}_1) = \sum_{i=0}^{63} d_{n,m}^i = \sum_{i=0}^{63} (f_{n,m,1}^i - \tilde{f}_{n,m,1}^i)^2 \quad (2)$$

위의 식에서 $d_{n,m}^i$ 는 n 번째 프레임의 m 번째 입력 블록의

i 번째 계수의 복원 오차를 나타내고, $f_{n,m,1}^i$ 와 $\tilde{f}_{n,m,1}^i$ 는 각각 첫번째 표현의 원본 계수와 복원된 계수를 나타낸다. 과잉 비트율 ρ 는 $\rho = \{Rate(\tilde{C}_1) + Rate(\tilde{C}_2)\} - Rate(\tilde{C})$ 가 되고, 식(1)은 라그랑제 최적화 방법에 의해 해결할 수 있다^{[12][13][14]}. 예를 들면 이 문제는 다음 식의 최소화문제로 변형하여 풀 수 있다.

$$J(\lambda) = D(\tilde{C}_1, \tilde{C}_2) + \lambda \rho(\tilde{C}_1, \tilde{C}_2). \quad (3)$$

위 식에서 라그랑제 승수 λ 는 복원 오차와 과잉 비트율을 고려해서 결정된다. 식(3)을 풀기 위하여^[15]에서 사용된 다이나믹 프로그래밍 방법을 수정하여 적용한다. [15]에서는 비트율이 주어졌을 경우에 양자화 테이블과 DCT 블록의 계수들의 최적 제거를 이용하여 비트율-왜곡 측면에서 최적적인 부호화를 수행한다. 제안하는 방법은 비트율과 양자화 계수가 주어졌을 때 입력 비트열을 RRD 관점에서 두개의 최적 표현으로 나누어 준다. 구체적인 알고리즘의 설명을 위하여 j 번째 계수를 최적화하는 경우를 생각해 보자. [15]에서는 블록의 $j+1$ 번째 계수부터 마지막 계수까지 모두 제거해 버린다. 그런데, [15]에서와는 달리 다중 표현에서는 제거된 계수들은 복호기에서 원래 값으로 복원될 수 없기 때문에 제안하는 알고리즘에서는 계수들을 제거해 버리는 대신에 미리 정의된 기본연산(default operation, 예를 들면, dup, alt 1, alt 2 중 하나 선택)을 사용하여 처리한다. 즉, j 번째 계수까지의 최적화 해를 알고 있다면 다음과 같은 식을 이용하여 k 번째 계수($k > j$)까지의 해를 구할 수 있다.

$$J_k = J_j^* + \Delta J_{j,k}^{op} (k > j). \quad (4)$$

위의 식에서 $\Delta J_{j,k}^{op}$ 는 특정한 연산(위의 dup, alt1, 또는 alt2)이 주어졌을 경우의 라그랑제 코스트의 증가분이다. 이러한 재귀적인 구조가 다이나믹 프로그래밍을 가능케 한다. 실제 구현에서는 0 번째 계수가 가장 먼저 최적화된다. 인트라 블록(intra block)의 경우는 0 번째 계수는 중요한 DC 성분을 가지기 때문에 모든 표현에 항상 중복적으로 할당되어야 한다. 그러나 인터 블록(inter block)의 경우에는 다

음과 같은 라그랑제 코스트를 최소화 하는 연산(op_o^*)이 선택된다.

$$J_0^* = \min_{op} J_0^{op}, op = \{ \text{dup}, \text{alt 1}, \text{alt 2} \}. \quad (5)$$

다음으로 첫번째 계수($k=1$)에 대한 최적화를 고려한다. 이 경우에는 재귀적인 성질을 이용하여 다음과 같은 식에서 J_0^* 로부터 J_1^* 를 구한다.

$$J_1^* = \min_{op} \{ J_0^* + \Delta J_{0,1}^{op} \} \quad (6)$$

그러면 첫번째 계수의 최적의 선행자(optimal predecessor)는 0 번째 계수가 되고선택된 연산(op)은 최적의 연산(op_1^*)이 된다. 두번째 계수는 최적의 선행자로 두개의 후보를 가질 수 있다. 첫번째 선행자는 첫번째 계수이다. 첫번째 선행자가 최적의 선행자로 선택되면 다음의 식이 성립한다.

$$J_{0 \rightarrow 2} = \min_{op} \{ J_0^* + \Delta J_{0,2}^{op} \} \quad (7)$$

두번째 선행자는 첫번째 계수이다. 두번째 선행자가 선택되면 다음의 식이성립한다.

$$J_{1 \rightarrow 2} = \min_{op} \{ J_1^* + \Delta J_{1,2}^{op} \} \quad (8)$$

[15]에서는 $J_{0 \rightarrow 2}$ 를 계산하기위하여 첫번째 계수를 제거한다. 그러나 이러한 방법은 복원 규칙을 깨기때문에 계수를 제거하는 방법 대신에복원 규칙을 만족시키는 기본 연산을 사용한다. 최적화 과정은 마지막 계수까지 동일한 방법으로 진행된다. 일반적인 k 에 대해서 최소의 코스트는 다음과 같다.

$$J_k^* = \min_{j,op} \{ J_j^* + \Delta J_{j,k}^{op} \}, \text{ for } j = 0, \dots, k-1. \quad (9)$$

본 논문에서와 같은 응용에서는 이 방법을 이용하여 다음과 같은모든 최적의 라그랑제 코스트를 계산할 수 있다.

$$J_k^* \text{ for } k = 0, \dots, 63 \quad (10)$$

라그랑제 코스트 뿐만 아니라 최적의 선행자 j 와 k 번째 계수의최적의 연산 op_k^* 도 따로 저장된다. J_k^* 를 최소화하는 k 를 k^* 로 표현하면 $J_{k^*}^*$ 는 전체 블록에 대한최소의 라그랑제 코스트가 된다. 전체 블록에 대한 최소의 라그랑제 코스트가 구해지면, 최적 선행자와 최적 연산과 같은 미리 구해진정보를 사용하여 최적 선행자를 따라 거슬러 올라가 0 번째 계수까지 최적화된 계수를 구한다.

그림 3은 5개의 계수를 가질 때 제한한 알고리즘의 방향성이 있는 비순환 그래프(directed acyclic graph, DAG)^[16]로 표현한 것이다. 마지막 최적 계수가 다섯 번째 계수일때, 굵은 선으로 표시된 것이 최적의 경로 이고 회색으로 칠해진 노드가 그 경로의 원소이다. 이 경우 다음과 같은 최적 연산들의 집합을 얻을 수 있다. { dup, dup, alt 1(default op), alt 1, alt 2 }

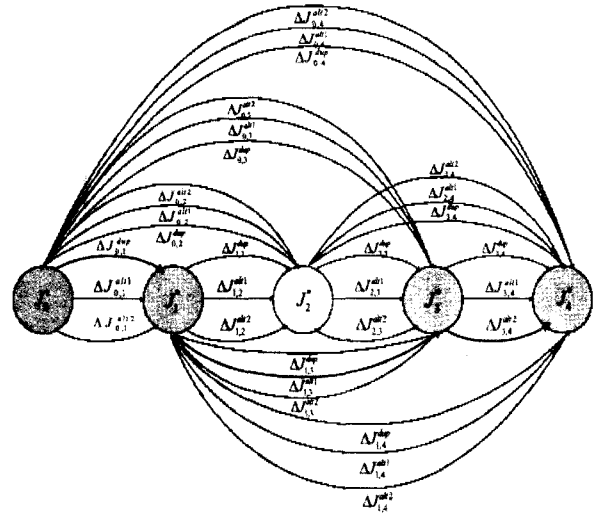


그림 3. 계수가 5개일 경우의 DAG 표현. 굵은 선이 최적의경로를 나타내고 이 경우 회색으로 칠해진 노드가 그 경로의 원소이다.

Fig. 3. DAG representation of 5 coefficients. The thick line is the optimal path. Then, the gray nodes are the elements of the optimal path and the white nodes are not.

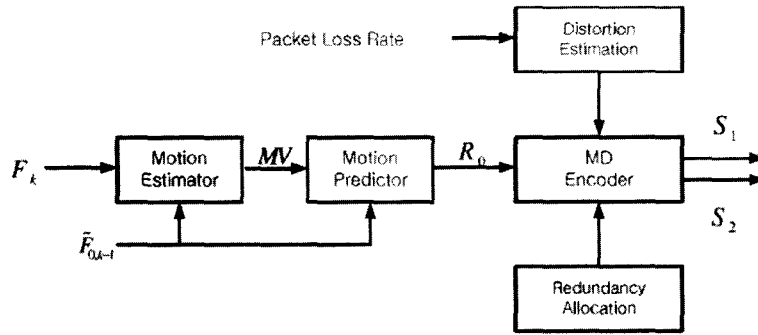


그림 4. 패킷 손실 환경으로 확장된 부호기 전체 구조
Fig.4 Overall structure of proposed coder extended to packet loss environment

III. 패킷 손실 환경을 고려한 알고리즘

1. 부호화시에 복호기 왜곡예측

지금까지는 하나 이상의 채널이 완전히 끊어지는 이상적인 다중표현 전송 환경을 가정하였다. 그러나 이러한 가정은 패킷 스위칭 네트워크에서는 성립하지 않는다. 패킷은 아무런 공지없이 중간 노드에서 버려질 수 있다. 따라서 손실된 정확한 패킷을 알 수 없고 일정 시간 동안의 패킷손실을만 측정할 수 있을 뿐이다. 이러한 패킷손실을 이용하여 부호화시에 복호기의 왜곡을 예측하는 것이 필요하다. 패킷손실을 고려하면 식(2)는 다음과 같이 수정되어야 한다.

$$D(\tilde{C}_1, \tilde{C}_2) = \sum_{i=0}^{63} d_{n,m}^i \quad (1)$$

위 식에서 $d_{n,m}^i$ 은 n 번째 프레임의 m 번째 블록의 i 번째 화소의복원 오차를 나타내며 다음과 같이정의된다. (이제 부터 표현의 간결성을 위해 블록 인덱스 m 을 제거하고 표현한다.)

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

f_n^i 은 n 번째 프레임의 원본 화소 값이고 \tilde{f}_n^i 는 복호기에서 복원된 화소값이다. 부호기에서는 \tilde{f}_n^i 을 알 수 없기 때문에 부호화시에는 확률변수가 된다. 이러한 개념은 제일 먼저

[10]에서 단일표현 동영상 압축에 ROPE(recursive optimal per-pixel estimate)라고 제안되었고 [9]에서 다중표현 압축 방법에 확장되었다. 본 논문에서는 편의상 [10]의 방법을 SDC-ROPE라고 하고 [9]의 방법을 MDC-ROPE라고 지칭한다. 제안된 방법에서는 부호화시에 복호기의 왜곡을 예측하기 위하여 MDC-ROPE를사용한다. 식2를 이용하여 복호기왜곡을 예측하기 위하여는 $E\{\tilde{f}_n^i\}$ 와 $E\{(\tilde{f}_n^i)^2\}$ 를 계산해야 한다. 인트라 프레임에 대하여 이 값들은 아래와 같이계산될 수 있다.

$$E\{\tilde{f}_n^i\} = (1-p)^2 \hat{f}_n^i + p(1-p)((\hat{f}_n^1)_1 + (\hat{f}_n^i)_2) + p^2 E\{\tilde{f}_{n-1}^i\} \quad (3)$$

$$E\{(\tilde{f}_n^i)^2\} = (1-p)^2 (\hat{f}_n^i)^2 + p(1-p)((\hat{f}_n^1)_1^2 + (\hat{f}_n^i)_2^2) + p^2 E\{(\tilde{f}_{n-1}^i)^2\} \quad (4)$$

여기서 p 는 패킷손실이 발생할 확률이다. 식 (3)의 첫번째 항은 패킷 손실이 발생하지 않은 경우를 나타내고, 두번째와 세번째 항은 각각 하나의 패킷 손실이 발생한 경우와 두 패킷이 모두 손실된경우를 나타낸다. $(\hat{f}_n^i)_1$ 과 $(\hat{f}_n^i)_2$ 는 각각 하나의 패킷이 손실되었을 때의 복원값이고, \tilde{f}_{n-1}^i 은 두개의 패킷이 모두 손실되었을 경우에 디코더에서 에러 은닉 방법을이용하여 복원한 값이다. 식 (4)의 각 항들도 동일한 경우를 나타낸다. 인터 프레임에 대한 값들도 다음과 같이 유사한 계산을 이용하여 구할 수 있다.

$$\begin{aligned}
& E\tilde{f}_n^i(P) \\
&= (1-p)^2(\hat{e}_n^i + E\tilde{f}_{n-1}^j) + \\
& P(1-p)((\hat{e}_n^1)_1 + E\tilde{f}_{n-1}^h + (\hat{e}_n^i)_2 + E\tilde{f}_{n-1}^{j_2}) + \\
& p^2 E\tilde{f}_{n-1}^i
\end{aligned} \quad (5)$$

$$\begin{aligned}
& E(\tilde{f}_n^i)^2(P) \\
&= (1-p)^2((\hat{e}_n^i)^2 + 2\hat{e}_n^i E\tilde{f}_{n-1}^j + E(\tilde{f}_{n-1}^j)^2) + \\
& P(1-p)((\hat{e}_n^1)_1^2 + 2(\hat{e}_n^i)_1 E\tilde{f}_{n-1}^h + \\
& E(\tilde{f}_{n-1}^{j_1})^2) + (\hat{e}_n^1)_2^2 + 2(\hat{e}_n^i)_2 E\{\tilde{f}_{n-1}^{j_2}\} + \\
& E\{(\tilde{f}_{n-1}^{j_2})^2\} + p^2 E\{(\tilde{f}_{n-1}^i)^2\}
\end{aligned} \quad (6)$$

앞의 식에서 \hat{e}_n^i 은 i 번째 화소의 움직임 보상된 나머지 화소값을 나타내고, $\tilde{f}_{n-1}^{j_1}$ 과 $\tilde{f}_{n-1}^{j_2}$ 은 하나의 패킷손실이 발생했을 경우 움직임 보상에러 은닉을 사용하여 이전 프레임의 j_1 번째와 j_2 번째의 화소 값을 이용하여 복원된 화소 값이다. \tilde{f}_{n-1}^j 의 경우 인트라 프레임의 경우와 동일하게 두개의 패킷이 모두 손실되었을 경우에 디코더에서 에러 은닉 방법을 이용하여 복원한 값이다.

2. 단일표현/다중표현 전환을 이용한 하이브리드 다중 표현 압축

높은 패킷손실율을 가지는 환경에서는 다중표현압축 방법이 단일표현압축 방법보다 에러에 더 강인하다. 그러나 낮은 패킷손실율에서는 표현들 사이의 연관을 구성하는 과도한 중복량 때문에 중복비트율-왜곡 성능이 저하된다. 이러한 문제점을 해결하기 위하여 채널환경에 따라서 단일표현/다중표현을 전환하는 하이브리드 다중표현 압축 방법을 제안한다. 예를 들면, 낮은 패킷손실율에서는 압축 효율을 위하여 단일표현압축을 사용하고 높은 패킷손실 환경에서는 에러에 대한 강인성을 위해 다중표현압축을 사용한다. 단일표현/다중표현 전환을 최적화 시키기 위해 중복비트율-왜곡 최적화 프레임웍(redundancy rate-distortion optimization framework)을 사용한다.

단일표현/다중표현 전환을 위하여는, 비디오 시퀀스 안에

프레임 당 M 개의 블록을 가지는 N 장의 프레임이 있다는 가정하에, 다음의 식을 최소화하는 최적의 모드(단일표현/다중표현)를 찾는 것이다.

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

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

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

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

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

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

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

3. SD/MD 전환 정보 전송 및 비트율 할당 방법

제안된 다중표현 압축 부호기-복호기는 ITU-T H.263을 기준으로 구현되었다. 하이브리드 다중표현 압축 방법에서는 기존의 단일표현 압축, 다중표현 압축 방법과는 달리 복호기에 단일표현/다중표현 전환 값들은 전달해 주어야만 한다. H.263의 문법 구조를 수정하지 않기 위하여픽처 헤더에 포함된 PEI(extra insertion information)와 PSUPP (supplemental enhancement information)을 사용한다^[17]. PSUPP 데이터는

4비트의 함수 형 표시자 FTYPE와 데이터 크기를 나타내는 4비트의 DSIZE로 구성된다. 지원되지 않는 함수형 표시자를 복호기가 받게 되면 함수의 인자 정보는 버려지게 된다. 단일 표현/다중표현 전환 정보를 전달하기 위하여 확장 함수 형을 사용하기 때문에 이러한 함수형을 지원하지 않는 복호기는 단일표현/다중표현 정보를 잃게 된다.

비트율을 효과적으로 조절하기 위하여 TMN8 비트율 조절 알고리즘과 라그랑제 승수 선택 방법을사용한다. 라그랑제 승수 선택 알고리즘은 [18]에서 제안되었다. 오프라인 실험을 통해서 가장 좋은 성능을 갖는 라그랑제 승수는 다음과 같이 결정된다.

$$\lambda = c \cdot (QP)^2 \tag{11}$$

이 식에 대한 이론적인 증명도 동일한 논문에 주어져 있으며 일반적인 H.263 코덱에 대해서는 0.85로 알려져 있다. 따라서, 본 논문에서는 구현된 SD-ROPE에서도 $c = 0.85$ 를 사용하여 실험하였다. 그러나 다중표현압축 동영상 방법은 일반적인 단일표현 동영상 압축방법과 특성이 다르기 때문에 새로운 라그랑제 승수 선택방법이 필요하다. 다중표현압축에서는 라그랑제 승수는 압축 효율을 결정할 뿐만아니라 에러에 대한 강인성도 조절한다. 일반적으로 높은 패킷손실율에서는 패킷 손실에 대항하기 위하여 더 많은 중복량이 필요하다. 따라서 본 논문에서는 패킷손실율이 라그랑제 승수 선택 알고리즘에 포함될수 있도록 다음과 같이 새로운 라그랑제 승수를 정의하였다.

$$\lambda = c \cdot (QP)^2 \cdot (-\log(PLR)). \tag{12}$$

MD-ROPE와 Hybrid MDC에서는 수정된 식에 $c=0.15$ 를 대입하여 사용하였다.

IV. 절 모의 실험 결과

1. 단일 채널 복호 성능

움직임 보상되고 양자화된 DCT 블록은 다이내믹 프로그

래밍을 이용하여 개선된 최적화 알고리즘을 이용하여 두개의 표현으로 나뉘어지고, 각각의 표현들은 H.263 표준으로 부호화 된다. 하나의 채널이 끊어졌을 경우에 대하여 제안한 다중 표현 동영상부호기와 기존의 MD split6] 부호기와 UMDC 부호기^[7]를 비교한다. 모의 실험에 사용된 동영상 시퀀스는 Foreman, Hall monitor, Carphone, News으로서 각각 그림 5(a)에 나타나 있다. Foreman은 빠른 카메라 움직임이 있고 끝부분에 카메라가 좌우로 크게 이동한다. Carphone은 Foreman과 유사한 빠른 카메라 움직임을 가지고 있으나 차창밖으로 풍경이 빠르게 지나가는 차이점이 있다. Hall monitor와 News는 카메라가 고정되어 있고 화

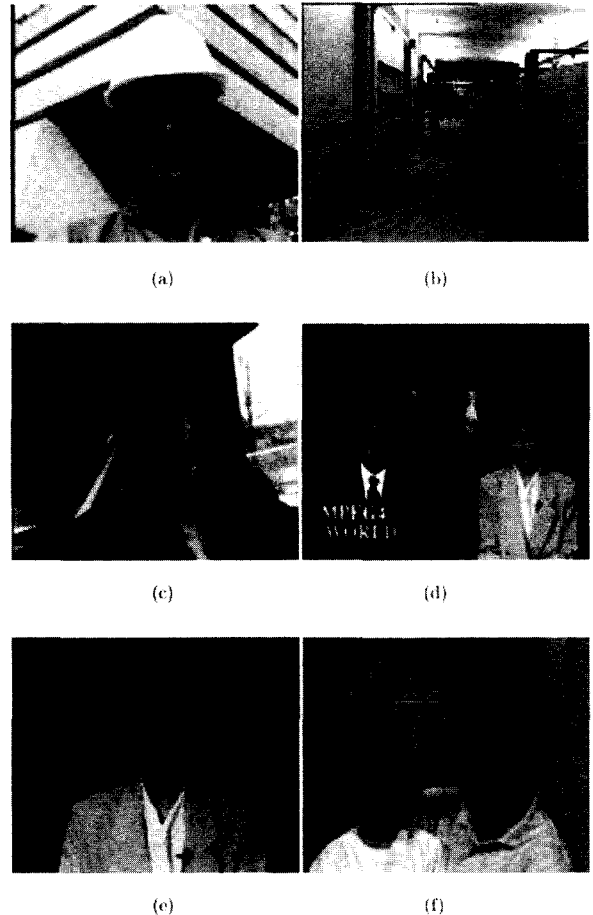


그림 5. 모의 실험에 사용된 시퀀스, (a) Foreman (b) Hall monitor (c) Carphone (d) News (e) Akiyo (f) Mother & Daughter
Fig.5 Sequences, (a) Foreman (b) Hall monitor (c) Carphone (d) News (e) Akiyo (f) Mother & Daughter

면 안 물체의 움직임이 상대적으로 작다. 이러한 각각의 시퀀스에 대하여 고정된 프레임율(frame rate)과 일정한 양자화 스텝크기(quantization stepsize)를 사용하여 부호화하였다. 또한, H.263의 부록 D, F, J가 사용되었으며 GOB 헤더가 포함되었고 인트라 프레임에 매 10 프레임마다 한번씩 갱신하였다. 모의 실험의 결과는 시간에 대한

평균 휘도의 PSNR 대 기준 비트율에 대한 과잉 비트율로 표현된다.

그림 6은 $QP = 12$ 일 때 주어진 시퀀스들의 PSNR 곡선을 나타내고 있다. 이 그림들에서 과잉 비트율 100%는 원본 비트율의 모든 정보가 두 개의 표현에 모두 중복적으로 저장 되었음을 의미한다. 이 경우에는 각각의 표현들은

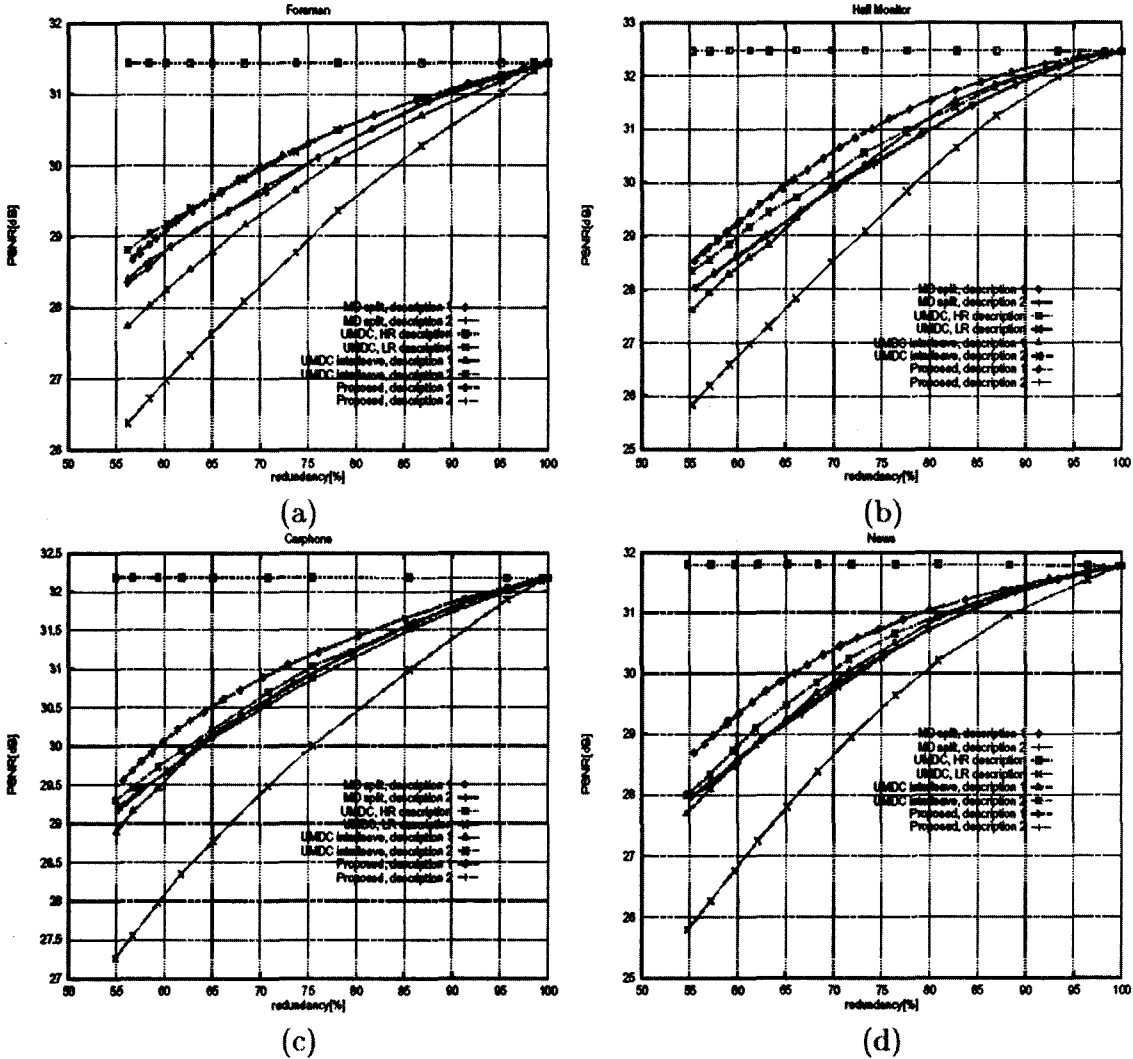


그림 6. 하나의 채널만 유효할 경우의 비트율-왜곡 성능 (a) Foreman(두개의 채널이 유효할 경우의 왜곡 32.4dB, 기준 비트율 51.7kbps) (b) Hall monitor(두개의 채널이 유효할 경우의 왜곡 33.2dB, 기준 비트율 38.2kbps) (c) Carphone(두개의 채널이 유효할 경우의 왜곡 32.2dB, 기준 비트율 52.9 kbps) (d) News(두개의 채널이 유효할 경우의 왜곡 31.8dB, 기준비트율 48.9 kbps)

Fig. 6. One channel rate-distortion performance (a) Foreman(two channel performance : 32.4dB, bit rate : 51.7kbps) (b) Hall monitor(two channel performance : 33.2dB, bit rate : 38.2kbps) (c) Carphone(two channel performance : 32.2dB, bit rate : 52.9kbps) (d) News(two channel performance : 31.8dB, bit rate : 48.9kbps)

원본 비트열과 동일하다. 그리고 MD split은 [6]의 다중 표현 분할 알고리즘을 구현한 것이고 UMDC는 [7]의 비대칭 다중 표현 알고리즘을 구현한 것이다. UMDC 알고리즘은 본래 비대칭이지만, 인터리브 패킷화(interleaved packetization)를 통해서 대칭적인 표현을 생성할 수 있다. 고화질 프레임의 짝수번째 GOB와 저화질 프레임의 홀수번째 GOB를 하나의 패킷에 저장하고 고화질 프레임의 홀수번째 GOB와 저화질 프레임의 짝수번째 GOB를 다른 패킷에 저장한다. UMDC interleave는 이러한 인터리브 패킷화 개념을 이용한 알고리즘을 나타낸다. 실험 결과 그래프에서 "description 1"은 두 번째 표현을 모두 수신하지 못했을 경우의 결과이고 "description 2"의 경우는 첫 번째 표현을 모두 수신하지 못했을 경우를 나타낸다. 모의 실험의 결과 그림을 살펴보면, 인터리브 패킷화를 사용하지 않은 UMDC 방법의 경우, 저화질의 채널만 손실됐을 경우에는 손실이 없는 경우와 동일한 좋은 성능을 보이지만, 고화질의 채널이 손실됐을 경우에는 성능이 급격히 악화된다. 인터리브 패킷화를 사용한 경우에는 사용하지 않은 경우보다 차이가 줄어들지만 완전한 균형을 이루지는 못한다. 이에 반해 MD split 방법과 제안하는 방법은 각각의 경우에 엄밀히 균형잡힌 표현을 만들어 내는것을 알 수 있다. 또한 제안하는 방법이 동일한 과잉 비트율에서 MD split 방법과 UMDC 방법보다 더 좋은 성능을보임을 알 수 있다. 결론적으로 제안된 다중 표현

동영상 부호기는 엄밀히 대칭적이고 비트율-왜곡 분할 방법^[6] 과 비대칭 다중표현 방법^[7]보다 모든 과잉 비트율에서 부호화 성능이 좋음을 알 수 있다.

2. 패킷 손실 환경에서의 성능

다음으로 패킷 손실이 있는 환경에서 인터/인트라 전환 방법과 제안하는 단일표현/다중표현 전환 방법을 비교한다. 그림 7은 실험에 사용된 3개의 알고리즘을 개념적으로 설명하고 있다. 그림에서 첫번째 행은 패킷손실율이 낮은 경우를 나타내고 두번째 행은 패킷손실율이 높은 경우를 나타낸다. 검은색으로 칠해진 블록은 인트라로 코딩된 블록을 나타낸다. SDC-ROPE와 MDC-ROPE는 인터/인트라 모드 전환을 사용하여 예러의 전과를 막는데 반하여, Hybrid MDC는 채널 환경에 따라단일표현/다중표현 전환을 수행한다. 먼저 비트율 제어의 효과를 제거한 상태에서 제안한 알고리즘의 성능을 분석하기 위하여 양자화 계수를 일정하게 고정하고 실험하였다. 표 1과 표 2는 Akiyo(그림 5(e)), Carphone(그림 5(c)), Foreman(그림 5(a)), Mother & Daughter(그림 5(f)) 시퀀스에 대한 결과를 요약하여 보여 주고 있다. 동일한 비트율에서의 비교는 아니지만, 제안된 방법의 성능을 고찰해 볼 수 있다. SDC-ROPE가 제일 적은 비트를 생성해 내고 MDC-ROPE가 제일 많은 비트를받

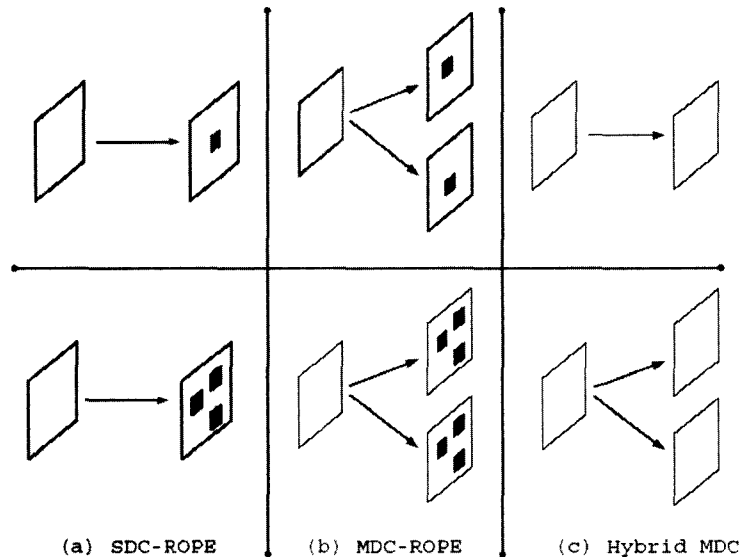


그림 7. 모의 실험에사용된 세 방법에 대한 개념도
Fig. 7. Three algorithms used in experiments

표 1. 패킷 손실 환경에서의 성능, QP=12
Table. 1 Packet loss performance, QP=12

		PLR	0.001	0.01	0.1	0.2
Akiyo	SDC-ROPE	Rate	26639	26911	23462	29566
		PSNR	34.07	34.02	33.31	32.54
	MDC-ROPE	Rate	43983	44059	46493	51768
		PSNR	34.05	33.81	32.96	32.70
	Hybrid MDC	Rate	29990	29990	31894	33933
		PSNR	34.06	34.01	33.36	32.89
Carphone	SDC-ROPE	Rate	64624	64761	67879	73218
		PSNR	31.95	31.76	29.23	27.72
	MDC-ROPE	Rate	96146	97088	104292	120304
		PSNR	32.01	31.70	30.68	30.15
	Hybrid MDC	Rate	68093	68093	93190	111015
		PSNR	31.94	31.70	30.09	29.42
Foreman	SDC-ROPE	Rate	82910	83149	93804	111399
		PSNR	31.18	30.84	27.71	25.72
	MDC-ROPE	Rate	126675	129596	141974	180085
		PSNR	31.22	30.92	29.88	29.17
	Hybrid MDC	Rate	86574	86574	133548	151390
		PSNR	31.09	30.70	29.54	28.25
M & D	SDC-ROPE	Rate	35503	35653	36652	38620
		PSNR	32.22	32.11	30.86	29.81
	MDC-ROPE	Rate	56872	57254	59890	66931
		PSNR	32.21	32.02	31.22	30.92
	Hybrid MDC	Rate	39044	39044	44399	51848
		PSNR	32.21	32.08	31.00	30.31

표 2. 패킷 손실 환경에서의 성능, QP=5
Table. 2 Packet loss performance, QP=5

		PLR	0.001	0.01	0.1	0.2
Akiyo	SDC-ROPE	Rate	64403	66047	70998	77290
		PSNR	39.33	39.20	37.63	36.47
	MDC-ROPE	Rate	85682	88961	98825	120417
		PSNR	39.30	38.85	37.59	37.60
	Hybrid MDC	Rate	67801	71555	78169	92680
		PSNR	39.32	39.17	37.87	37.12
Carphone	SDC-ROPE	Rate	186604	188029	210536	236414
		PSNR	37.30	36.86	33.27	31.26
	MDC-ROPE	Rate	227645	230538	267413	354952
		PSNR	37.35	36.71	35.15	34.79
	Hybrid MDC	Rate	189937	198077	301128	328533
		PSNR	37.23	36.75	35.19	33.35
Foreman	SDC-ROPE	Rate	233446	237732	308740	375712
		PSNR	36.22	35.69	31.90	29.85
	MDC-ROPE	Rate	298331	304363	376815	554416
		PSNR	36.31	35.68	34.36	34.09
	Hybrid MDC	Rate	238002	251322	391060	416014
		PSNR	36.10	35.53	34.03	31.34
M & D	SDC-ROPE	Rate	114793	115474	123969	136266
		PSNR	37.10	36.83	34.61	33.08
	MDC-ROPE	Rate	145212	151089	164466	203077
		PSNR	37.10	36.67	35.44	35.16
	Hybrid MDC	Rate	118410	118410	153282	183351
		PSNR	37.08	36.77	35.14	34.37

표 3. 인터/인트라, 단일표현/다중표현 전환 횟수, QP=12
Table. 3 Number of switching, Intra/Inter, SD/MD, QP=12

		PLR	0.001	0.01	0.1	0.2
Akiyo	SDC-ROPE	0	50	218	342	
	MDC-ROPE	0	2	54	120	
	Hybrid MDC	0	0	1	18	
Carphone	SDC-ROPE	270	304	684	1260	
	MDC-ROPE	62	104	238	414	
	Hybrid MDC	0	0	66	91	
Foreman	SDC-ROPE	446	496	1645	3422	
	MDC-ROPE	88	214	587	1223	
	Hybrid MDC	0	0	93	98	
M & D	SDC-ROPE	32	52	176	400	
	MDC-ROPE	8	22	82	180	
	Hybrid MDC	0	0	15	47	

표 4. 인터/인트라, 단일표현/다중표현 전환 횟수, QP=5
Table. 4 Number of switching, Intra/Inter, SD/MD, QP=5

		PLR	0.001	0.01	0.1	0.2
Akiyo	SDC-ROPE	9	118	374	712	
	MDC-ROPE	23	69	120	237	
	Hybrid MDC	0	1	26	72	
Carphone	SDC-ROPE	556	656	2022	3520	
	MDC-ROPE	155	262	489	1056	
	Hybrid MDC	0	7	94	99	
Foreman	SDC-ROPE	980	1312	4996	8298	
	MDC-ROPE	299	603	1266	2049	
	Hybrid MDC	9	19	99	100	
M & D	SDC-ROPE	86	130	628	1328	
	MDC-ROPE	78	208	292	574	
	Hybrid MDC	0	0	46	87	

생시킨다. Carphone과 Foreman의 경우 배경과인물의 움직임이 크기 때문에 PLR이 증가하면 SDC-ROPE의 경우 PSNR이 급격히 하락하지만, MDC-ROPE와 Hybrid MDC의 경우에는 PSNR의 하락이 상대적으로 작다. 움직임이 작은 Akiyo와 Mother & Daughter의 경우에도 Carphone과 Foreman과 비슷한 성질을 보이지만, SDC-ROPE의 경우도 PSNR의 저하가 적다. 이는 움직임이 작은 경우에는 SDC-

ROPE가 에러에 대한 효과적으로 대처할 수 있는 것을 보여준다. 이에 반해, MDC-ROPE의 경우에는 PSNR의 저하가 작지만, 에러 강인성을 위해 발생된 비트량이 크다. 제안하는 Hybrid MDC의 경우에는 발생된 비트는 SDC-ROPE 보다 많고 MDC-ROPE보다는 작고, PLR이 낮은 환경에서는 SDC-ROPE와 유사한 성능을 보이고 PLR이 높은 경우에는 MDC-ROPE와 유사한 성능을 보여준다. 이

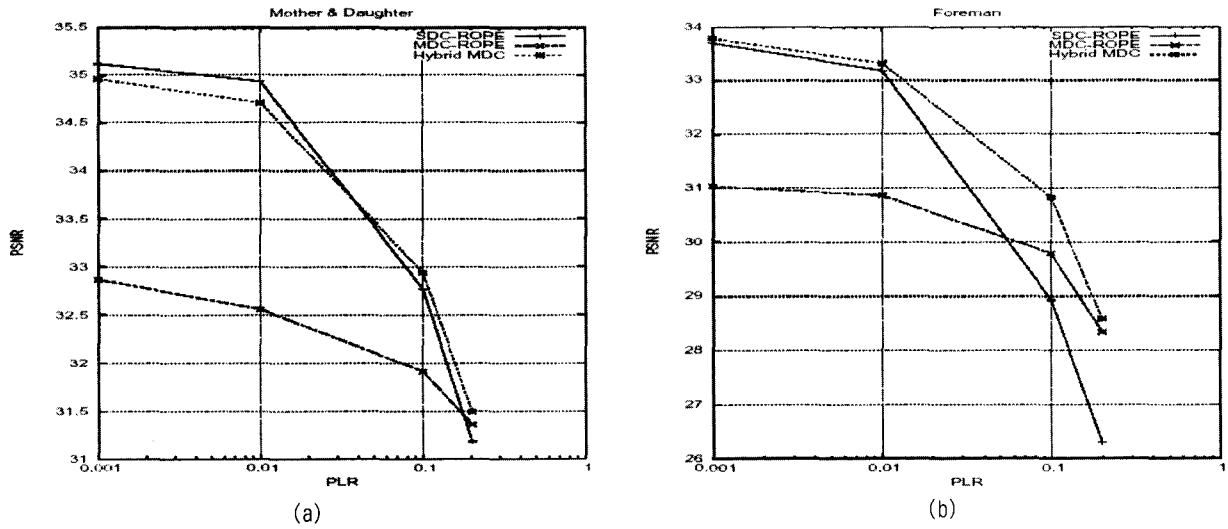


그림 8. 패킷 손실 환경에서의 성능 (a) Mother & Daughter, 64kbps (b) Foreman, 128kbps
 Fig. 8. Packet loss performance (a) Mother & Daughter sequence at 64kbps (b) Foreman sequence at 128kbps

는 PLR이 낮은 경우에는 SDC-ROPE와 유사한 작은 비트량으로 높은 PSNR을 얻고, PLR이 높은 경우에는 MDC-ROPE와 같이 비트량을 증가시켜 에러에 대한 강인성을 얻는 것을 보여준다. 표 3와 표 4는 각각의 경우 인터/인트라 전환과 단일표현/다중표현 전환이 발생한 횟수를 나타내고 있다. 움직임이 큰 시퀀스인 경우에 그렇지 못한 경우보다 인터/인트라 전환과 단일표현/다중표현 전환이 더 많이 발생한다. Hybrid MDC의 경우, PLR=0.001, PLR=0.01에서는 주로 SD를 이용하여 압축 효율을 높이고, PLR=0.1, PLR=0.2인 경우에는 MD 전환을 수행하여 에러에 대한 강인성을 높이고 있음을 알 수 있다.

다음으로 TMN8 비트율 조절 알고리즘과 수정된 라그랑주 승수 선택 방법을 사용하여 SDC-ROPE, MDC-ROPE, Hybrid MDC의 성능을 비교한다. 각 비디오 시퀀스에 대하여 30개의 서로 다른 랜덤 패킷 손실 패턴에 대한 실험을 수행한 후 평균값을 구하였다. 그림 8는 QCIF 크기의 Mother & Daughter 시퀀스와 Foreman 시퀀스에 대한 패킷 손실 환경에서의 성능을 나타내고 있다. Mother & Daughter 시퀀스의 경우 64kbps 비트율에서 실험을 수행하였고 Foreman 시퀀스의 경우 128kbps 비트율에서 실험하였다. Mother & Daughter의 경우 움직임 작기 때문에 에러가 발생하더라도 단순한 에러 은닉 방법을 사용하여 에러의 영향을 줄일 수 있다. 따라서 SDC-

ROPE가 거의 모든 패킷손실율에서 MDC-ROPE보다 더 좋은 성능을 보인다. 제안하는 Hybrid MDC는 패킷손실율이 낮은 경우에는 SDC-ROPE와 유사한 성능을 보이고, 패킷손실율이 높아지면 MDC-ROPE의 성능을 보인다. 그러나 패킷손실율이 증가할 수록 각 방법에 대한 차이가 점차 줄어든다. 이에 반해 Foreman 시퀀스의 경우는 배경과 인물이 동시에 움직이기 때문에 단순한 에러 은닉만으로는 에러의 효과를 감소시킬 수 없다. 따라서 패킷손실율이 증가하면 MDC-ROPE가 SDC-ROPE 보다 성능이 좋아진다. Foreman 시퀀스의 경우에도 제안한 방법은 두 가지 방법보다 더 좋은 성능을 보인다. 이러한 결과로부터 제안하는 Hybrid MDC 알고리즘이 높은 패킷손실율에서 뿐만 아니라 낮은 패킷손실율에서도 기존의 SDC-ROPE와 MDC-ROPE 보다 더 좋은 성능을 보임을 알 수 있다. 그림 9와 그림 10은 각각 Akiyo와 Foreman의 복호된 프레임임을 나타내고 있다. Akiyo의 경우 64kbps로 압축하고 PLR이 10%인 환경에서 실험을 하였고, Foreman의 경우에는 128kbps로 압축하고 PLR이 10%인 환경에서 실험을 수행하였다. 그림에서 알 수 있듯이 SDC-ROPE의 경우에는 에러가 발생하면 에러 은닉을 수행하더라도 눈에 거슬리는 블록 효과가 발생한다. 이에 반해 MDC-ROPE와 Hybrid MDC의 경우에는 이러한 효과가 거의 제거된다. 그러나 MDC-ROPE는 상대적으로 비트를 많이 발생시키

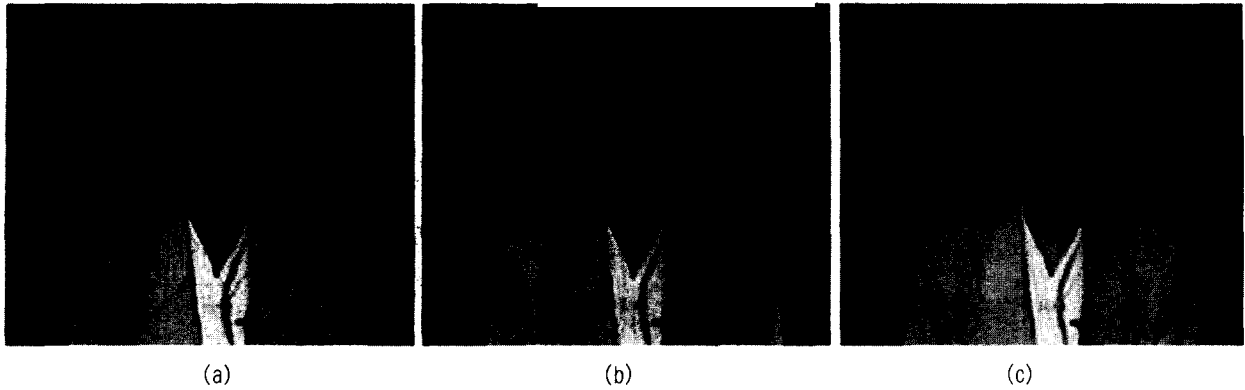


그림 9. Akiyo에 대한 주관적 화질 비교, 64kbps, PLR 10% (a) SDC-ROPE (b) MDC-ROPE (c) Hybrid MDC
 Fig. 9. Comparison of decoded picture quality, Akiyo, 64kbps (a) SDC-ROPE (b) MDC-ROPE (c) Hybrid MDC

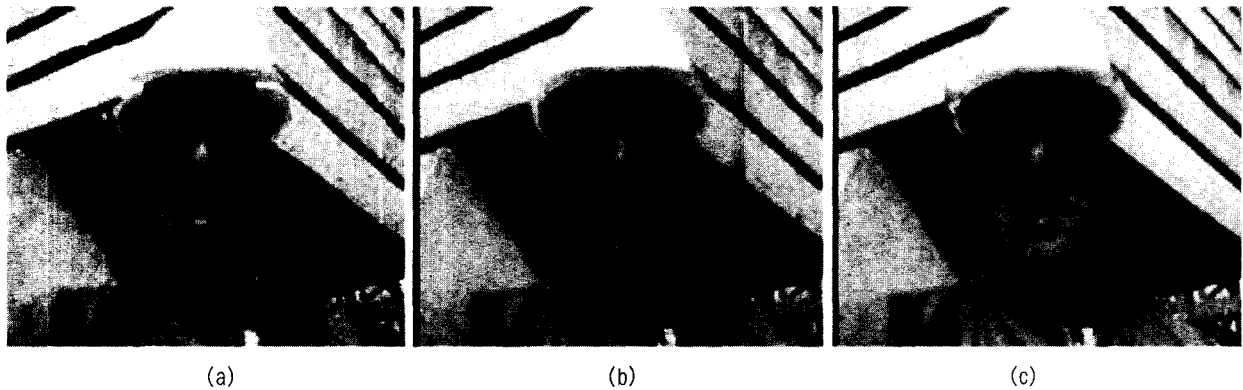


그림 10. Foreman에 대한 주관적 화질 비교, 128kbps, PLR 10% (a) SDC-ROPE (b) MDC-ROPE (c) Hybrid MDC
 Fig. 10. Comparison of decoded picture quality, Foreman, 128kbps (a) SDC-ROPE (b) MDC-ROPE (c) Hybrid MDC

기 때문에, 고정된 비트율로 압축된 경우에는 전체적으로 화질의 선명도가 저하된다. 이에 반해 Hybrid MDC의 경우에는 각 PLR에 적응적으로 비트를 생성하기 때문에 눈에 거슬리는 블록 효과를 제거하는 동시에 화질의 저하를 막고 있다.

V. 결론

본 논문에서는 다중 표현 개념을 이용하여, 에러에 강인한 동영상 부호화 방법을 제안하였다. 제안하는 방법은 DCT 계수의 최적 분할과 채널 환경에 따른 단일표현/다중표현 전환 방법으로 구성되어 있다. DCT 계수 최적 분할

방법에서는 입력 신호를 라그랑제 최적화 방법을 사용하여 주어진 중복량에서 최적의 과잉 비트율-왜곡 성능을 갖는 두 개의 표현으로 분할한다. 또한, 재귀적 구조를 사용한 다이내믹 프로그래밍 기법을 사용하여 분할의 복잡도를 줄인다. 단일표현/다중표현 전환 방법에서는 ROPE를 이용하여 복원 에러를 예측한 후, 낮은 패킷 손실율에서는 압축 효율을 위하여 단일표현을 사용하고 패킷 손실율이 큰 환경에서는 에러에 대한 강인성을 위해 다중표현을 사용한다. 모의 실험 결과, 제안하는 다중표현 동영상 부호화 방법은 이상적인 다중표현 채널에서 뿐만 아니라 다양한 패킷 손실율을 갖는 채널 환경에서도 기존의 단일표현 및 다중표현 에러 내성 부호화 방법보다 더 좋은 성능을 보임을 알 수 있었다.

참 고 문 헌

[1] V. Goyal, "Multiple description coding : Compression meets the network," IEEE Signal Processing Mag., vol. 18, pp. 74-93, Sept. 2001.

[2] A. Ingle and V. Vaishampayan, "DPCM system design for diversity systems with applications to packetized speech," IEEE Trans. Speech and Audio Processing, vol. 3, pp. 48-57, Jan. 1995.

[3] Y. Wang, M. Orchard, and A. Reibman, "Multiple description image coding for noisy channels by pairing transform coefficients," in Proc. First Workshop on Multimedia Signal Processing, June 1997, pp. 419-424.

[4] A. Reibman, H. Jafarkhani, Y. Wang, M. Orchard, and R. Puri, "Multiple description coding for video using motion compensated prediction," in Proc. Int. Conf. Image Processing, Oct. 1999, pp. 837-841.

[5] V. Vaishampayan and S. John, "Balanced interframe multiple description video compression," in Proc. Int. Conf. Image Processing, Oct. 1999, pp. 812-816.

[6] A. Reibman, H. Jafarkhani, Y. Wang, and M. Orchard, "Multiple description video using rate-distortion splitting," in Proc. Int. Conf. Image Processing, Oct. 2001, pp. 978-981.

[7] 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.

[8] Y.-C. Lee, Y. Altunbasak, and R. Mersereau, "A two-stage multiple description video coder with drift-preventing motion compensated prediction," in Proc. Int. Conf. Image Processing, Sept. 2002, pp. 557-560.

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

[10] 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.

[11] 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. 2003.

[12] H. Everett, "Generalized lagrange multiplier method for solving problems of optimum allocation of resources," Oper. Res., vol. 11, pp. 399-417, 1963.

[13] A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," IEEE Signal Processing Mag., vol. 15, pp. 23-50, Nov. 1998.

[14] G. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," IEEE Signal Processing Mag., vol. 15, pp. 74-90, Nov. 1998.

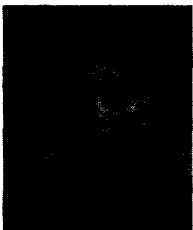
[15] K. Ramchandran and M. Vetterli, "Rate-distortion optimal fast thresholding with complete JPEG/MPEG decoder compatibility," IEEE Trans. Image Processing, vol. 3, pp. 700-704, 1994.

[16] L. Kondi and A. Katsaggelos, "An operational rate-distortion optimal single-pass SNR scalable video coder," IEEE Trans. Image Processing, vol. 10, pp. 1613-1620, 2001.

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

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

저 자 소 개



김 일 구

- 1995년 : 수원고등학교 졸업
- 1999년 : 서울대학교 전기공학부 학사
- 2001년 : 서울대학교 전기공학부 석사
- 2001년~현재 : 서울대학교 전기공학부 박사과정



조 남 익

- 1982년~1986년 : 서울대학교 제어계측공학과, 학사
- 1986년~1988년 : 서울대학교 제어계측공학과, 석사
- 1988년~1992년 : 서울대학교 제어계측공학과, 박사
- 1991년~1994년 : 서울대학교 제어계측신기술연구소, 연구원
- 1994년~1998년 : 서울시립대학교 전자전기공학부, 조교수
- 1999년~현재 : 서울대학교 전기공학부, 부교수