

論文2002-39SP-6-1

비트율-왜곡 최적화된 DCT 계수 분할을 이용한 다중 표현 동영상 압축 방법

(Multiple Description Video Coding Using Rate-Distortion Optimized DCT Coefficient Splitting)

金 壹 求 * , 趙 南 翊 *

(Il Koo Kim and Nam Ik Cho)

요 약

본 논문에서는 다중 표현(Multiple Description)과 비트율-왜곡 연산자(rate-distortion operator)를 이용하여 에러가 발생하기 쉬운 환경에서 사용될 수 있는 강인한 동영상 압축 알고리즘을 제안한다. 하나의 비디오 스트림을 두 개의 연관된 표현 (description)으로 나누기 위해 과잉 비트율-왜곡(redundancy rate-distortion : RRD) 기준을 사용하였다. 제안하는 방법은 새로 추가한 비트율-왜곡 연산자와 새로 제안된 최적화 방법을 통해 기존의 RRD 기반 방법들보다 최적화 성능을 향상시킬 수 있다. 또한 실질적인 에러 환경을 고려한 새로운 왜곡 측정법(distortion measure)을 정의하였다. 제안하는 다중 표현 동영상 압축기는 H.263 표준을 기반으로 하기 때문에 각각의 표현은 H.263 표준 디코더를 사용하여 독립적으로 디코딩될 수 있다. 또한, 기존의 방법과는 달리 손실된 표현에 대한 복잡한 예측과정이 없어도 간단한 병합 과정을 통해 여러 개의 표현을 하나의 스트림으로 디코딩 할 수 있다. 모의 실험 결과, 제안하는 다중 표현 동영상 압축 방법은 기존의 다중 표현 분할(multiple description split) 방법보다 모든 비트율에서 더 좋은 성능을 보임을 알 수 있다.

Abstract

We propose an algorithm for the robust transmission of video in error prone environment based on multiple description (MD) scheme and simple rate-distortion operators. The redundancy rate-distortion (RRD) criteria is used to split a one-layer compressed video stream into two correlated descriptions. The proposed algorithm can find more candidate points on the RRD curve than the conventional RRD based methods. A new distortion measure is also defined in this paper, which considers more realistic error environments. Since the proposed MD video coder is based on the standard H.263 coder, each description can be decoded independently by the standard H.263 decoder. Also, several descriptions can be decoded into a single stream by additional merge stage and the H.263 decoder. Simulation results show that the proposed MD video coder yields better performance than the conventional MD splitting algorithm at all bitrates both in single and two description cases.

Keywords : error resilient coding, multiple description coding, redundancy-rate-distortion, H.263

* 正會員, 서울大學校 電氣工學部

(School of Electrical Eng. Seoul National Univ.)

接受日字:2002年7月15日, 수정완료일:2002年10月25日

I. 서 론

하나 이상의 독립적인 채널이 존재할 경우, 다중 표

현 압축(Multiple Description Coding : MDC)은 채널 에러에 대한 강인성을 크게 높일 수 있다. 소스 신호에 대한 다중 표현 문제는 정보 이론 분야와 신호 처리 분야에서 널리 연구되어 왔다.^[1~4] 신호처리 분야에서 MDC 개념은 음성 압축에 제일 먼저 적용되었으며,^[5] 그 후에 영상과 동영상 압축에 차례로 적용되었다.^[6~8] MDC의 가장 큰 장점은 수신단에서 모든 표현을 받지 못한 경우에도 소스 신호를 복원 할 수 있다는 점이다. 엄밀히 말하면, MDC는 하나의 스트림을 서로 연관성을 갖도록 설계된 여러 개의 표현으로 나눈다. 결과적으로 모든 표현을 성공적으로 전송하지 못한 경우라도 성공적으로 수신된 표현을 이용하여 손실된 표현들을 예측, 복원 할 수 있다. MDC와 관련된 연구분야에는 각 표현 사이에 효과적이고 구조화된 연관(correlation)을 만들어 주는 것과 이러한 연관성의 양을 조절하는 것이 중요한 연구내용이다. <그림 1>은 일반적인 다중 표현 압축에 대한 개요를 보여주고 있다. 다중 표현 동영상 압축에서 가장 중요한 문제는 인코더와 디코더 사이의 불일치(mismatch)를 처리하는 것이다. 인코더와 디코더 사이의 불일치는 기본적으로 움직임 보상 예측(motion compensated prediction)에서 기인한다. 전송시 표현이 손실되면 인코더에서의 움직임 보상 프레임과 디코더에서의 복원 프레임이 다르게 된다. 이러한 불일치는 움직임 보상에 의해 시간 축으로 전파되고 결과적으로 비디오 화질을 저하시키게 된다.

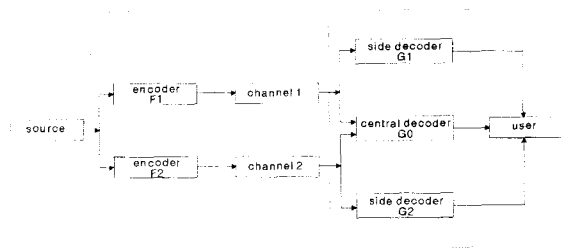


그림 1. 일반적인 다중 표현 압축 개요

Fig. 1. General multiple description coding overview.

표준 비디오 스트림은 움직임 벡터, 코딩 모드와 같은 부가 정보 그리고 DCT 계수로 구성되어 있다. 이 중에 움직임 벡터에 대한 정교한 다중 표현 동영상 압축 방법이 발표된 바 있으나,^[9] 대부분의 다중 표현 동영상 압축 방법은 움직임 벡터와 부가 정보는 모든 표현에 중복적으로 포함시키고 단지 DCT 계수만 여러 개의 표현으로 나누는데 집중하고 있다. 이러한 다중 표

현 동영상 압축 방법은 크게 세 가지로 나눌 수 있다. 첫 번째 방법은 미리 설계된 다중 표현 스칼라 양자화기(MDSQ)를 이용한 방법이다. 두 번째 방법은 켈레별 상관 변환(pairwise correlating transform : PCT)을 이용한 방법이고, 마지막 방법은 비트율-왜곡 최적화를 이용한 방법이다.^[10~12] MDSQ를 이용한 동영상 압축은 각각의 표현에 연관성을 분포시키는 스칼라 양자화기를 설계해서 사용한다. 이러한 양자화기는 한번 설계되면 입력 신호의 성질에 따라서 적응적으로 변화시킬 수 없다는 단점이 있다. 반대로 PCT를 이용한 방법은 입력 신호의 통계적 성질에 따라서 연관의 정도를 조절 할 수 있으나 PCT에 사용되는 매개변수를 정확히 구하기 위해서는 입력 신호의 전체적인 통계적 성질이 필요하고 어떤 방식으로 켈레를 만드는가 하는 전략에 대한 부가 정보가 디코더에 전송되어야 한다. 이 방법은 알고리즘을 복잡하게 만들고 비트율을 증가시킨다. 앞의 두 방법보다 최근에 제시된 비트율-왜곡 최적화에 기반한 방법은 각 표현 사이의 연관성의 양을 조절 할 수 있고 부가적인 정보의 전송도 필요 없다는 장점을 갖는다. 구체적으로, [12]의 다중 표현 분할 방법에서는 DCT 계수를 각 표현에 분할하기 위해서 우선 프레임들을 구성하는 8x8 크기의 각각의 블록에 대하여 하나의 문턱치(threshold)를 고려한다. 문턱치 보다 큰 계수는 모든 표현에 중복적으로 저장되고 문턱치보다 작은 계수는 각 표현 중 오직 하나에만 저장된다. 최적의 문턱치 계산을 위해서는 각 문턱치에 대하여 과잉 비트율(redundancy rate, ρ)과 측면 복원 에러(side reconstruction distortion, D_1)가 계산된다. 0에서부터 모든 계수가 번갈음 될 때 까지 문턱치를 증가시키며 목적 함수를 계산하며 RRD 곡선 위에서 서로 다른 동작점들을 찾는다. 이러한 동작점 중에서 목적 함수($D_1 + \lambda\rho$)를 최소화시키는 동작점이 최적이며, 이 동작점의 문턱치가 최종적으로 다중 표현 동영상 압축에 사용된다. 디코더에서는 이러한 문턱치가 없어도 표현들을 디코딩 할 수 있다. 디코더에서는 모든 표현을 받았을 경우에는 각 표현의 계수들을 병합한다. 병합된 스트림은 H.263 표준 디코더를 이용하여 디코딩된다. 일부의 표현을 받았을 경우에는 전송된 표현만을 이용하여 H.263 표준 디코더와 동일한 방법으로 스트림을 디코딩 할 수 있다. 이 경우에는 열화된 비디오를 얻게 된다.

다중 표현 분할 방법은 표준에 대응하는 방법을 통해서 디코딩이 가능하고 구현이 단순하며 코비트율에서는 PCT를 사용한 방법보다도 성능이 더 좋다. 그러나 [12]에서는 문턱치를 사용하여 DCT 계수의 크기 정보만 계수를 분할하는데 이용되었고, 계수들이 각 표현에 중복적으로 저장되거나 하나의 표현에만 번갈아 가면서 저장되는 두 가지 방법을 사용하기 때문에 최적화 성능이 나쁘다. 따라서 본 논문에서는 DCT 계수의 크기 뿐만 아니라 주파수 특성을 이용하고, 중복시키거나 번갈아 저장하는 방법 뿐만 아니라 비대칭 분할(asymmetric split)을 고려한 새로운 최적화 알고리즘을 제안한다. 이러한 방법은 기존의 RRD 기반 방법들보다 최적화 성능을 향상시킬 수 있다. 또한 측면 복원 에러(side reconstruction distortion) 뿐만 아니라 중앙 복원 에러(central reconstruction distortion)도 고려한 새로운 왜곡 측정법(distortion measure)을 정의해서 사용한다.

II. 제안하는 다중 표현 동영상 압축 방법

A. 전체 프레임워크

표준 비디오 스트림은 움직임 벡터, 코딩 모드와 같은 부가 정보 그리고 DCT 계수로 구성되어 있다. 제안하는 알고리즘은 기존의 알고리즘과 동일하게 DCT 계수의 분할에 중점을 두고 있다.

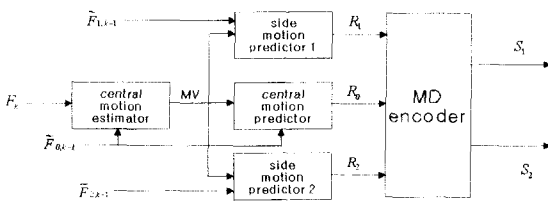


그림 2. 제안하는 MD 비디오 시스템의 P 모드에서의 전체 프레임워크
 Fig. 2. General framework of proposed MD video systems in the P mode.

<그림 2>는 제안된 다중 표현 동영상 압축기의 전체적인 구조이다. 입력 스트림을 두 개의 표현으로 나누는 것을 가정하고 있다. k 번째 프레임 F_k 와 (k-1) 번째 복원 프레임들 ($F_{0,k-1}$, $F_{1,k-1}$, $F_{2,k-1}$)이 주어지면 인코더는 움직임 벡터(MV)를 예측할 수 있다.

$F_{0,k-1}$ 은 두 개의 표현 모두를 이용한 복원 프레임, $F_{1,k-1}$, $F_{2,k-1}$ 은 각각 하나씩의 표현만을 이용한 복원 프레임이다. MV는 각각의 움직임 예측기에 적용되어 예측 신호($P_{0,k}$, $P_{1,k}$, $P_{2,k}$)와 예측 에러 신호($R_{0,k}$, $R_{1,k}$, $R_{2,k}$)를 만들어 낸다. 예측 신호 $P_{0,k}$ 는 중앙 예측(central prediction)이라고 하고, $P_{1,k}$ 와 $P_{2,k}$ 를 측면 예측(side prediction)이라고 한다. 중앙 예측 에러(central prediction error) $R_{0,k}$ 는 두 개의 표현 $S_{1,k}$ 와 $S_{2,k}$ 로 나누어져 두 개의 독립적인 채널을 통해 디코더에 전송된다. 측면 예측 에러(side prediction error) $R_{1,k}$ 와 $R_{2,k}$ 는 측면 복원 에러(side loop distortion, $D_i = R_{i,k} - S_{i,k}$)을 계산하는데 사용된다. 이 측면 복원 에러(side reconstruction distortion)는 다중 표현 동영상 압축기에서 인코더와 디코더에서의 예측 루프(prediction loop) 사이에서의 불일치의 중요한 원인이 된다. 만약 채널이 끊어지거나 에러가 많이 발생하면, 인코더에서의 참조 프레임(reference frame) $F_{0,k}$ 와 디코더에서의 복원 프레임(reconstruction frame) \hat{F}_k 는 D_i 만큼의 차이가 생기게 된다. 따라서 인코더와 디코더 사이에 불일치가 발생하고 성능이 저하되게 된다. 인코더와 디코더 사이에 불일치를 미리 제거하기 위해서는 과잉 비트율 제한 조건 아래에서 왜곡 D_i 를 최소화하여야 한다. <그림 2>의 MD encoder 블록에서 제안하는 알고리즘에 따라서 불일치를 최소화한다.

B. 비트율-왜곡 연산자를 이용한 최적화 방법

기존의 다중 표현 분할 알고리즘에서는 문턱치보다 큰 계수는 두 개의 표현에 중복적으로 저장되고 문턱치보다 작은 계수들은 두 개의 표현에 번갈아 가면서 저장되었다. 최적 문턱치 선택은 문턱치를 변화시켜가며 목적 함수를 최소화시키는 것으로 결정하였다. 이 방법에서는 하나의 블록에 대해 하나의 매개변수, 즉 문턱치만이 사용되었기 때문에 복잡도가 작다는 장점이 있으나 계수의 크기만을 사용하고 있기 때문에 효율이

- 1) 일반적으로 움직임 벡터 MV는 $F_{0,k-1}$ 을 참조 프레임(reference frame)으로 하여 구하지만 움직임을 위해 복원 프레임 전부를 이용할 수 있다. 본 논문에서는 $F_{0,k-1}$ 만을 이용하여 움직임을 예측을 수행하였다.

그다지 높지 못하다. 즉, DCT 계수들 중 일반적으로 저주파 쪽의 계수들이 고주파 계수들보다 화질에 미치는 영향이 더 큰데 기존의 방법에서는 모든 계수를 똑 같이 취급하였다. 따라서 본 논문에서는 계수의 크기 뿐만 아니라 주파수 특성도 사용하도록 한다.

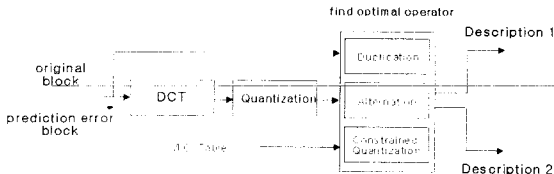


그림 3. 비트율-왜곡 조작을 이용한 제안 알고리즘
Fig. 3. Proposed algorithm using rate distortion optimization operator.

또한, 기존의 방법에서는 문턱치보다 큰 계수들과 작은 계수들이 하나의 따로 단위로 연산되었으나 제안하는 방법에서는 한번에 하나의 계수만이 연산된다. 기존에는 각 계수를 문턱치에 따라 중복(duplication) 시키거나 번갈음(alternation) 시키는데 반하여 본 논문에서는 주파수 위치와 주어진 목적 함수에 따라 중복, 번갈음 그리고 비대칭 분할(asymmetric split)의 세가지 방법을 적용한다. 중복은 계수를 두 개의 표현에 동일하게 저장한다. 번갈음은 두 개의 표현 중 하나에만 계수를 저장한다. 비대칭 분할은 양자화된 계수의 합이 원본 계수와 같고 두 값이 동일하지 않은 두 개의 새로운 계수로 양자화한다. 나누어진 계수들은 각각의 표현에 저장된다. 중복 연산과 번갈음 연산은 오직 하나의 계수에만 적용된다는 점만 제외하고는 다중 표현 분할 방법과 유사하다.

이론적으로 64개의 계수를 갖는 입력 블록을 세 개의 연산자를 가지고 최적화하기 위해서는 3^{64} 번의 연산이 필요하다. 그러나 이러한 방법은 구현이 불가능하기 때문에 그리디(greedy) 접근 방법을 사용한다. 즉, 작은 인덱스를 갖는 계수부터 큰 인덱스를 갖는 계수까지 순서대로 최적화하는 것이다. 이러한 경우 최대의 연산 횟수는 64×3 번이 된다. 이러한 접근 방법을 이용하면 계산량의 감소 뿐만 아니라 주파수 특성까지 효과적으로 이용할 수 있다. 비트율-왜곡 분할을 이용한 방법에서는 단순히 문턱치를 두어 계수의 크기만으로 중복 및 번갈음을 결정된다. 이런 경우는 고주파 영역에 크기가 큰 계수가 존재한다면 그 계수는 항상 중복

되어 비트율-왜곡 관점에서 성능을 저하시킨다. 따라서 계수의 크기를 기준으로 중복, 번갈음을 결정하는 것이 아니라, 저주파 영역에서 고주파 영역으로 순차적으로 최적화해 나가면 결과적으로 고주파 영역에 비교적 큰 계수가 존재한다고 하더라도 저주파 영역에 비트가 더 많이 할당되는 결과를 얻게 된다. 하나의 블록에 대한 전체적인 최적화 과정은 다음과 같다.

입력 블록은 DCT를 이용하여 변환되고 표준 비디오 코딩과 같이 양자화 계수(QP)를 이용하여 양자화된다. 비트율-왜곡 연산자를 이용한 최적화는 양자화된 블록(B)에 대하여 수행된다. 전체 과정은 <그림 3>에 기술되어 있다. 그림에서 점선으로 표시된 블록이 중복 연산(T_{DUP}), 번갈음 연산(T_{ALTER}), 비대칭 분할(T_{AS}) 중에서 최적의 연산자를 찾는 비트율-왜곡 연산자를 나타낸다. 블록의 첫 번째 계수(b_0)는 DC 계수로서 화질에 매우 중요하므로 항상 중복시킨다. 영이 아닌 계수들은 인덱스가 작은 계수부터 연산자들이 적용되고 목적 함수를 최소화시키는 연산자가 선택된다. 즉 $C = Tb_k$ 로써, b_k 는 입력 블록의 k 번째 계수이고 $C = [c_1, c_2]^T$ 는 분할된 계수이며 T 는 수행되는 연산자이다. 그리고 분할된 계수 c_1 과 c_2 는 각각의 표현 S_1 과 S_2 에 저장된다. 즉 $s_{1,k} = c_1, s_{2,k} = c_2$ 로써, $s_{1,k}$ 는 S_1 의 k 번째 계수이고 $s_{2,k}$ 는 S_2 의 k 번째 계수이다. 목적 함수를 최소화하기 위해 라그랑주 승수법을 사용한다.^[13,14]

$$J = D([S_1]_{k-1}, c_1, [S_2]_{k-1}, c_2) + \lambda \cdot \rho([S_1]_{k-1}, c_1, [S_2]_{k-1}, c_2) \quad (1)$$

위 식에서 D 는 복원 오차이고 ρ 는 과잉 비트율이며 $[S_1]_{k-1}$ 은 $k-1$ 번째 계수까지 최적화된 S_1 이다. $[S_2]_{k-1}$ 은 S_2 의 경우이다. 자세히 설명하면, D 는 S_1 의 복원 오차 D_1 과 S_2 의 복원 오차 D_2 의 합이다. ρ 는 하나의 원본 블록을 대신해서 두 개의 표현을 코딩했을 경우에 실제로 발생하는 부가적인 비트량이다. 라그랑주 계수 λ 는 과잉 비트율을 고려해서 결정된다. λ 가 작아지면 과잉 비트율이 증가하고 λ 가 커지면 과잉 비트율이 감소한다. 하나의 계수에 대해 최적화가 완료되면 선택된 비트율-왜곡 연산이 수행되고 나누어진 두 개의 계수는 각각의 표현에 저장된다. 두 개의 표현 S_1 과 S_2 의 k 번째 계수 $s_{1,k}$ 와 $s_{2,k}$ 는 최적의 연

산자를 이용해 구해진 c_1^* 과 c_2^* 로 각각 갱신되고 다음 계수의 왜곡과 과잉 비트율을 구하는데 사용된다. 이러한 과정은 양자화된 블록의 모든 계수가 전부 최적화 될 때까지 계속된다. 최적화가 종료되면 H.263 표준과 동일한 방법으로 압축되어 디코더에 전송된다.

C. 새로운 왜곡 측정법

[12]에서는 최적화 과정에서 오직 측면 복원 에러만 고려하였다. 측면 복원 에러는 측면 참조 프레임 오직 하나의 표현만을 이용하여 복원된 프레임과의 차이이다. 그러나 하나의 표현이 완전히 손실되는 것은 매우 드문 경우이므로 중앙 복원 에러도 역시 고려해야 한다. 중앙 복원 에러는 중앙 참조 프레임과 오직 하나의 표현만을 이용하여 복원된 프레임간의 차이이다. 제안된 알고리즘에서는 다음과 같이 새로운 왜곡 측정법을 정의하여 (1)식의 복원 오차 D 대신 사용한다.

$$D = (1 - \mu)D_{side} + \mu D_{central} \tag{2}$$

여기서 D_{side} 는 측면 복원 에러이고 $D_{central}$ 은 중앙 복원 에러이다. μ 는 균형 인자(balancing parameter)이다. 손실이 많이 발생하는 환경에서는 하나의 표현만을 받을 확률이 크기 때문에, 중앙 복원 에러의 효과를 감소시키기 위해 μ 를 작게 만들어 준다. 그러나 채널이 끊어질 확률이 작은 경우에는 중앙 복원 에러가 중요하기 때문에 μ 를 크게 만들어 주어야 한다.

III. 모의 실험 결과

제안하는 알고리즘의 구현은 H.263 표준을 기반으로 하고 있다. 인트라/인터 방식(intra/inter mode) 교환도 역시 표준에 따라 수행된다. 움직임 보상 후 양자화된 DCT 블록은 비트율-왜곡 연산자를 이용하여 두 개의 표현으로 나누어지고, 각각의 표현들은 H.263 표준과 동일한 방법으로 압축된다. 중복 연산은 입력 블록의 k 번째 계수가 두 개의 S_1, S_2 의 k번째 계수로 치환된다. 번갈음 연산은 목적 함수의 값에 따라 k번째 계수가 하나의 표현에만 저장된다. 목적 함수값이 동일한 경우에는 약간의 번갈음 전략이 필요하다. 계수가 홀수번째 블록의 홀수번째 계수이거나 짝수번째 블록의 짝수번째 계수일 경우 S_1 에 저장하고, 그 외의 경우에는 S_2 에 저장한다. 이러한 전략은 번갈음 연산이 비대칭적인

연산이기 때문에 각 표현의 비트율을 맞추기 위해 필요하다. 비대칭 분할은 계수들의 합이 원본 계수와 동일하도록 분할하여 각각의 표현에 저장한다. 만약 계수가 짝수인 경우 각각의 계수가 동일하게 나누어지는 것은 피해야 한다. 왜냐하면, 그럴 경우에는 디코더에서 중복 연산과 제한 양자화 연산을 구별 할 수 없기 때문이다.

A. 하나의 채널이 완전히 손실됐을 경우의 복원 성능

표 1. QCIF 시퀀스(10초)
Table 1. 10 SECOND QCIF SEQUENCES.

	Frame rate	Quantization Parameter	PSNR without loss	Reference bit-rate
Foreman	10 fps	12	31.1 dB	55.2 kbps
Foreman	10 fps	8	33.1 dB	89.3 kbps
Hall monitor	10 fps	12	31.9 dB	20.9 kbps
Hall monitor	10 fps	8	34.1 dB	32.8 kbps

먼저 하나의 채널이 완전히 손실되었을 경우를 가정하고 제안하는 다중 표현 동영상 압축 방법의 성능과 기존의 다중 표현 분할 알고리즘[12]의 성능을 비교한다. 실험에 사용한 동영상들은 Foreman과 Hall monitor 이다. 이 동영상들의 성질은 <표 1>에 나타나 있다. 여기서는 고정된 프레임율(frame rate)과 일정한 양자화 크기(quantization stepsize)를 사용하였고 H.263의 기본

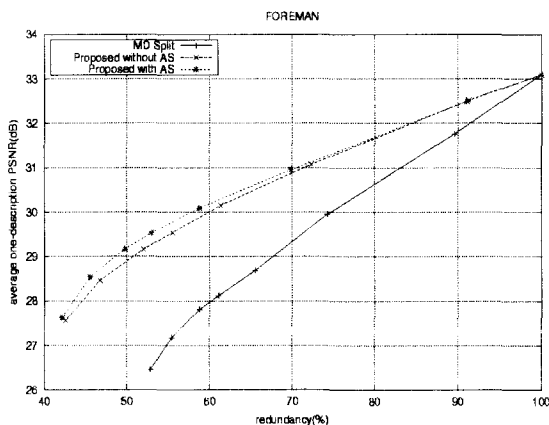


그림 4. Foreman : 하나의 채널만 유효할 경우의 비트율-왜곡 성능(두개의 채널이 유효할 경우의 왜곡 33.1dB, 기준 비트율 89.4kbps)

Fig. 4. Foreman : Single channel redundancy rate-distortion performance (two channel distortion 33.1 dB, reference bit rate 89.4 kbps).

모드를 사용하였으며 GOB 헤더를 포함시켰다. 이 경우에는 하나의 표현이 모두 손실 된 경우이므로 식 (2)에서의 균형 인자 μ 는 0으로 정한다. 실험의 결과는 시간에 따른 평균 밝기의 PSNR(Peak Signal to Noise Ratio)과 과잉 비트율의 백분율로 나타내었다.

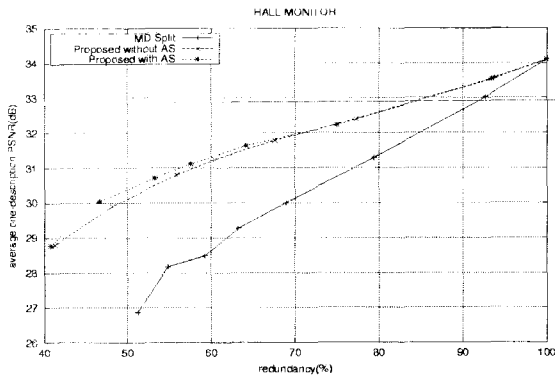


그림 5. Hall monitor : 하나의 채널만 유효할 경우의 비트율-왜곡 성능(두개의 채널이 유효할 경우의 왜곡 34.1dB, 기준 비트율 32.8kbps)

Fig. 5. Hall monitor : Single channel redundancy rate-distortion performance(two channel distortion 34.1 dB, reference bit rate 32.8 kbps).

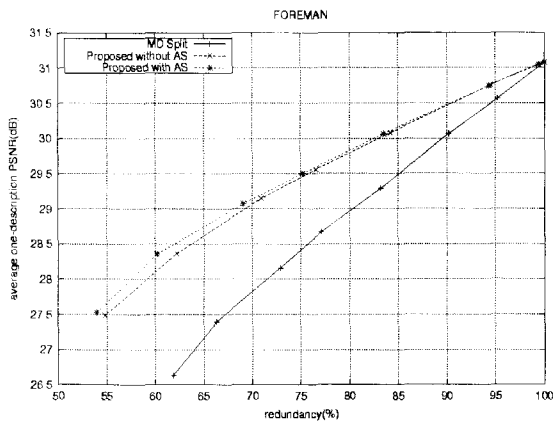


그림 6. Foreman : 하나의 채널만 유효할 경우의 비트율-왜곡 성능(두개의 채널이 유효할 경우의 왜곡 31.1dB, 기준 비트율 55.2kbps)

Fig. 6. Foreman : Single channel redundancy rate-distortion performance(two channel distortion 31.1 dB, reference bit rate 55.2 kbps).

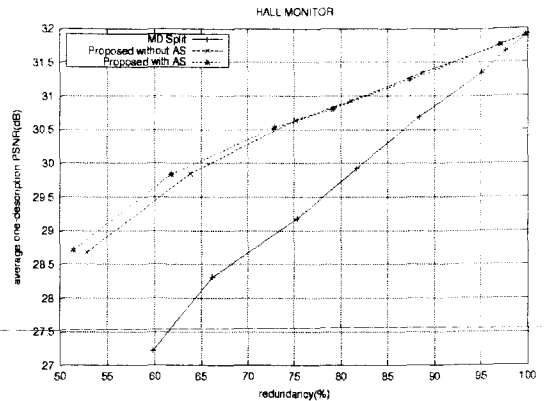


그림 7. Hall monitor : 하나의 채널만 유효할 경우의 비트율-왜곡 성능(두개의 채널이 유효할 경우의 왜곡 34.1dB, 기준 비트율 32.8kbps)

Fig. 7. Hall monitor : Single channel redundancy rate-distortion performance(two channel distortion 34.1 dB, reference bit rate 32.8 kbps).

<그림 4, 5>는 QP = 8일 경우에 Foreman과 Hall monitor의 PSNR 곡선을 나타내고 있고, <그림 6, 7>은 QP = 12일 경우의 PSNR 곡선을 나타낸다. "MD Split"은 다중 표현 분할 방법을, 나머지는 제안한 최적화 방법에서 비대칭 분할 연산자의 유무에 따른 방법의 차이를 나타낸다. 100%의 과잉 비트율은 소스 스트림의 모든 정보가 두 개의 표현에 모두 중복되어 저장되었다는 것을 나타낸다. 이 경우에 각 표현은 소스 스트림을 표준 H.263 코더를 이용하여 압축한 것과 동일하다. 모든 경우에 대하여 제안된 다중 표현 동영상 압축 방법의 성능이 기존의 다중 표현 분할 방법보다 좋다는 것을 알 수 있다. 또한 그림들은 비대칭 분할 연산이 알고리즘의 복잡도를 증가시키지만, 작은 과잉 비트율에서 효과적이라는 것을 나타내고 있다. 이러한 사실은 <표 2>에서도 확인해 볼 수 있다. <표 2>는 기준 비트율이 55.2kbps인 Foreman 시퀀스에 대하여 과잉 비트율에 따른 각 연산자의 발생 빈도를 나타내고 있다. 표를 통해서 과잉 비트율이 작아질수록 비대칭 분할 연산자가 최적의 연산자로 선택될 확률이 증가함을 알 수 있다.

표 2. 과잉 비트율에 따른 각 연산자의 발생 빈도

Table 2. relative frequency of each operators.

redundancy	duplication	alternation	asymmetric split
100	44606(100%)	0(0%)	0(0%)
99.4	44406(99.5%)	85(0.2%)	115(0.3%)
94.4	42324(94.9%)	1079(2.4%)	1203(2.7%)
83.5	34571(77.5%)	4951(11.1%)	5084(11.4%)
75.2	25733(57.7%)	9001(20.2%)	9872(22.2%)
69.1	18284(50.0%)	12754(28.6%)	13588(30.4%)
60.2	7451(16.7%)	16908(37.9%)	20247(45.4%)
54.0	1756(3.9%)	19398(43.5%)	23452(52.6%)

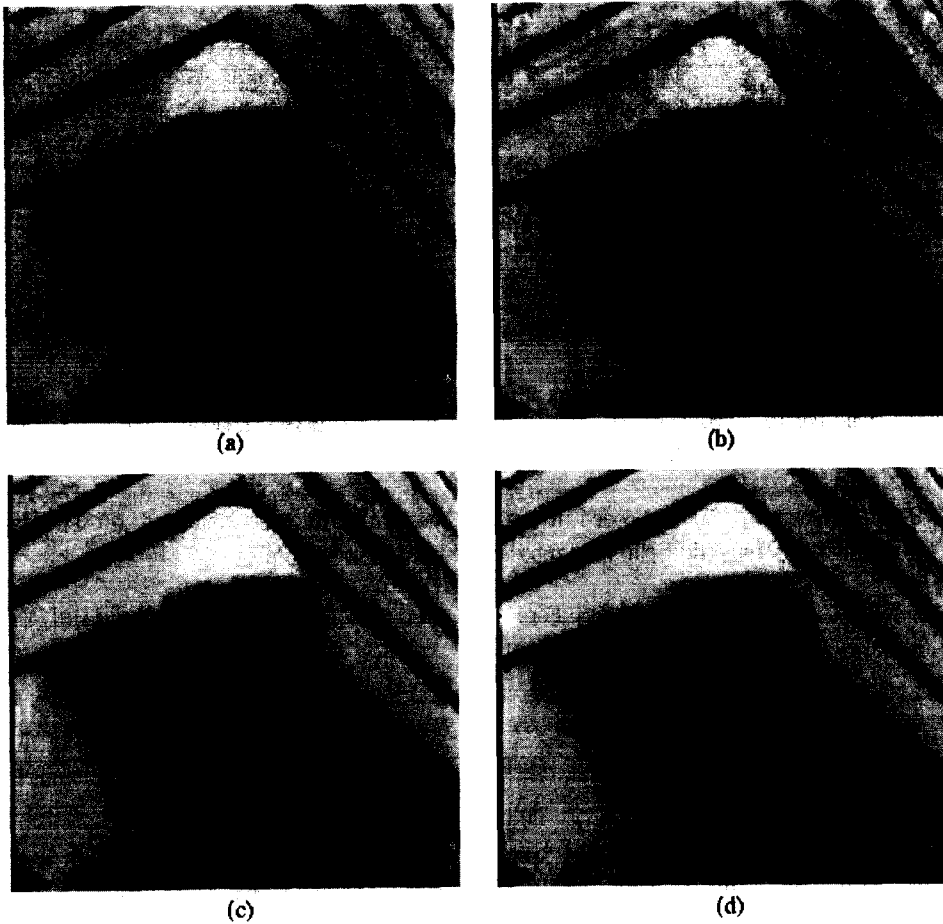


그림 8. Foreman : 53번째 복원 프레임 (a) 에러가 없는 경우 (b) [12]을 사용한 경우 (c) 제안하는 방법을 사용한 경우(AS 제외) (d) 제안하는 방법을 사용한 경우(AS 포함)

Fig. 8. Foreman : 53th decoded frame. (a) no error (b) MD Split [12] (c) Proposed without Asymmetric Split (d) Proposed with Asymmetric Split

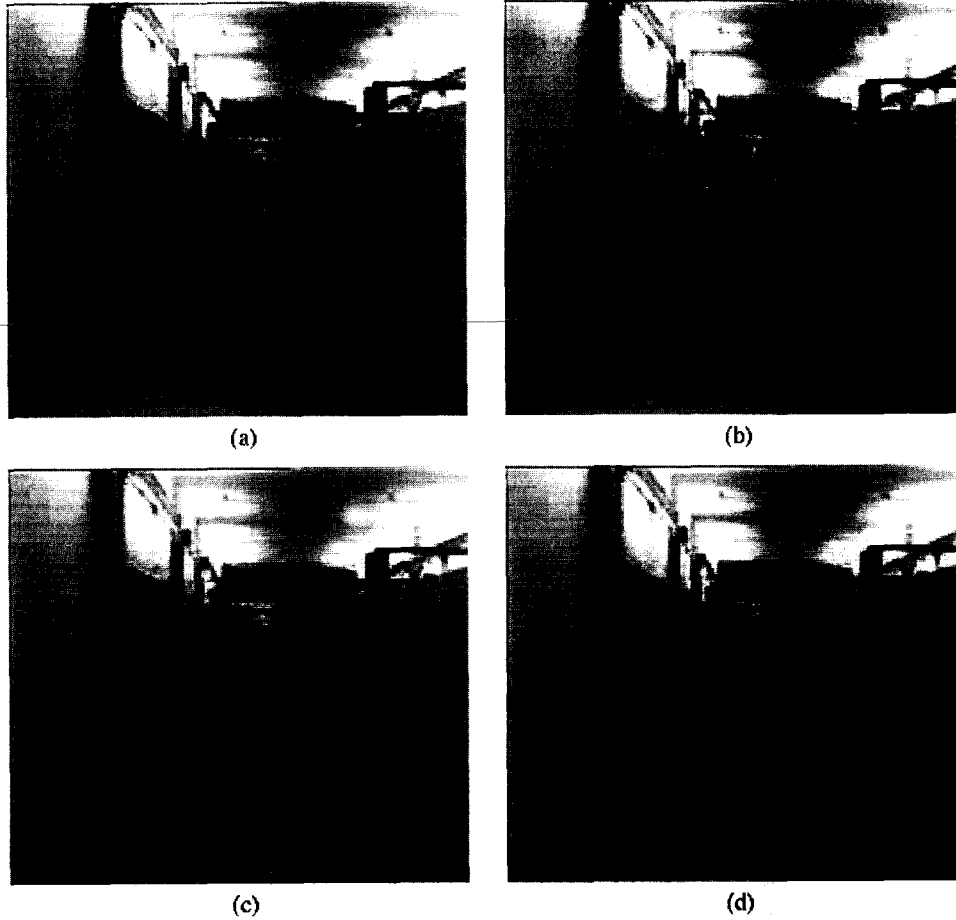


그림 9. Hall monitor : 60번째 복원 프레임 (a) 에러가 없는 경우 (b) [12]를 사용한 경우(MD Split) (c) 제안하는 방법을 사용한 경우(AS 제외) (d) 제안하는 방법을 사용한 경우(AS 포함)

Fig. 9. Hall monitor : 60th decoded frame. (a) no error (b) MD Split [12] (c) Proposed without Asymmetric Split (d) Proposed with Asymmetric Split

<그림 8>은 하나의 표현만을 받았을 경우에 Foreman 동영상에 대하여 비트율-왜곡 분할 방법에 의해서 복원된 영상과 제안하는 비트율-왜곡 연산자를 이용한 방법에 의해 복원된 영상을 나타내고 있다. 하나의 스트림으로 압축했을 경우의 비트율은 89.3kbps이고, 과잉 비트율(ρ)이 59%일 경우이다. 비교를 위하여 두 개의 표현을 모두 받았을 경우의 복원 영상도 포함되어 있다. <그림 8(b)>와 <그림 8(d)>를 비교해 보면 비슷한 과잉 비트율에서 제안하는 알고리즘이 비트율-왜곡 분할 방법보다 주관적 화질이 좋음을 알 수 있다. Hall monitor 동영상에 대해서도 <그림 9>와 같이 마찬가지로의 결과를 얻었다. 하나의 스트림으로 압축했을 경우의 비트율이 32.8kbps이고, 과잉 비트율이 53%일 경우이다.

B. 패킷 손실이 있는 채널 환경에서의 복원 성능

다음은 두 개의 채널이 사용가능하고 패킷 손실이 발생할 수 있는 환경을 가정하고 성능을 비교하였다. 하나의 프레임은 두 개의 패킷으로 나뉘어 지는 것을 가정하였다. 즉 하나의 표현 자체가 하나의 패킷이 된다. 실험을 위해서 두 개의 채널에 독립적으로 무작위적인 패킷 손실을 발생시킨다. 패킷 손실 비율이 각각 10%인 경우의 Foreman과 Hall monitor 시퀀스의 복원 성능은 <그림 10, 11>에 나타나 있다. 각각의 그래프들은 제안하는 방법에서 균형인자가 0인 경우, 균형인자가 0.8인 경우, 균형인자가 1.0인 경우, 그리고 다중 표현 분할 방법을 나타낸다. 그림들을 통해서 패킷 손실이 있는 환경에서도 제안한 알고리즘이 이전의 다중

표현 분할 방법보다 효과적이며, 균형 인자가 에러에 대한 강인성에 영향을 줄 수 있다. 10%의 프레임 손실 비율에서는 1에 근접하는 균형 인자가 작은 균형 인자보다 효과적임을 알 수 있다. 채널의 패킷 손실 비율에 따라 균형 인자를 효과적으로 적용하면 복원 성능을 향상시킬 수 있다.

IV. 결 론

본 논문에서는 기존의 다중 표현 분할 알고리즘^[2]보다 개선된 비트율-왜곡 최적화 방법과 새로운 왜곡 측정법을 이용한 다중 표현 동영상 압축 방법을 제안하였다. 이 방법에서는 기존의 다중 표현 동영상 압축 방법에서 필요했던 미리 설계된 양자화 테이블이 필요하지 않고 비디오의 통계적 특성을 미리 알 필요도 없다. 또한 제안된 비트율-왜곡 연산자와 최적화 방법을 이용하면 다중 표현 분할 알고리즘^[2]보다 최적화 성능을 향상시킬 수 있다. 제안한 다중 표현 비디오 압축 방법은 표준 H.263을 기반으로 하고 있으며, 따라서 각각의 표현들은 표준 H.263 디코더를 사용하여 독립적으로 디코딩할 수 있으며 여러 표현들은 간단한 병합 과정을 거친 후 H.263 디코더를 이용하여 하나의 스트림으로 디코딩 할 수도 있다. 광범위한 모의 실험 결과 제안하는 방법이 기존의 다중 표현 분할 알고리즘^[2]보다 중복 비트율이 70%일 경우에 평균적으로 약 1dB 정도 더 좋은 성능을 보인다는 것을 확인하였다.

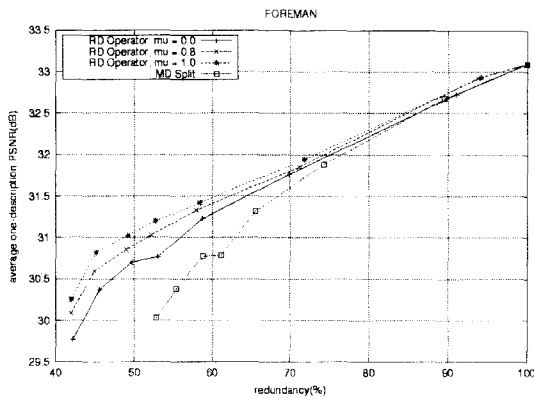


그림 10. Foreman : 두 개의 채널이 사용 가능하고 각 채널이 10%의 패킷 손실 비율(packet loss rate)을 가질 경우의 성능 그래프

Fig. 10. Foreman : Two channel redundancy rate-distortion performance for 10% packet loss rate.

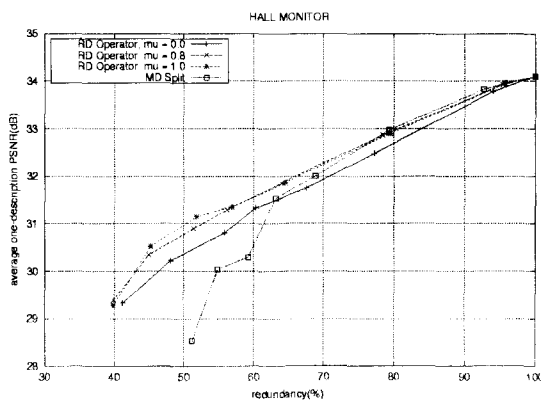


그림 11. Hall monitor : 두 개의 채널이 사용 가능하고 각 채널이 10%의 패킷 손실 비율(packet loss rate)을 가질 경우의 성능 그래프

Fig. 11. Hall monitor : Two channel redundancy rate distortion performance for 10% packet loss rate.

참 고 문 헌

- [1] L. Ozarow, "On a source coding problem with two channels and three receivers," Bell Syst. Tech. J., Vol. 59, pp. 1909~1921, Dec. 1980.
- [2] V. A. Vaishampayan, "Design of multiple description scalar quantizer," IEEE Trans. Inform. Theory, Vol. 39, pp. 821~834, May 1993.
- [3] V. K. Goyal and J. Kovacevic, "Optimal multiple description transform coding for gaussian vectors," Proc. Data Compress. Conf, Mar. 1998, pp. 388~397.
- [4] V. K. Goyal, "Multiple description coding : compression meets the network," IEEE Signal Processing Mag., Vol. 18, pp. 74~93, Sept. 2001.
- [5] A. Ingle and V. A. 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.
- [6] M. Orchard, Y. Wang, V. A. Vaishampayan, and A. Reibman, "Redundancy rate distortion analysis of multiple description image coding using pairwise correlating transforms," Proc. Int. Conf. Image Processing, Oct. 1997, pp. 608~611.
- [7] Y. Wang, M. Orchard, V. A. Vaishampayan, and A. Reibman, "Multiple description coding using pairwise correlating transforms," IEEE Trans. Image Processing, Vol. 10, pp. 351~366, Mar. 2001.
- [8] 김이한, 김성대, 영상 압축 기술의 동향 및 알고리즘, 전자공학회지. 1995년 4월 pp. 64~75
- [9] Chang-Su Kim and Sang-Uk Lee, "Multiple description coding of motion fields for robust video transmission," IEEE Trans. Circuits Syst. Video Techno., Vol. 11, pp. 999~1010, Sept. 2001.
- [10] V. A. Vaishampayan and S. John, "Balanced interframe multiple description video compression," Proc. Int. Conf. Image Processing, Oct. 1999, pp. 812~816.
- [11] A. Reibman, H. Jafarkhani, Y. Wang, M. Orchard, and R. Puri, "Multiple description coding for video using motion compensated prediction," Proc. Int. Conf. Image Processing, Oct. 1999, pp. 837~841.
- [12] A. Reibman, H. Jafarkhani, Y. Wang, and M. Orchard, "Multiple description video using rate-distortion splitting," Proc. Int. Conf. Image Processing, May 2000, pp. 978~981.
- [13] H. Everett, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," Oper. Res., Vol. 11, pp. 399~417, 1963.
- [14] A. Ortega and K. Ramchandran, "Rate-distortion methods for image and video compression," IEEE Signal Processing Mag., Vol. 15, pp. 23~50, Nov. 1998.

 저 자 소 개



金壹求(正會員)

1995 : 수원고등학교 졸업. 1995~1999 : 서울대학교 전기공학부 학사. 1999~2001 : 서울대학교 전기공학부 석사. 2001~ : 서울대학교 전기공학부 박사과정



趙南翊(正會員)

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