

다중 비디오 프로그램 전송을 위한 비트율-왜곡 추정 기반의 결합 비디오 부호화

正會員 홍 성 훈*, 김 성 대*

Joint Video Coding for Multiple Video Program Transmission Based on Rate-Distortion Estimation

Sung-Hoon Hong*, Seong-Dae Kim* *Regular Members*

요 약

최근 MPEG-2와 같은 비디오 압축기술과 디지털 전송기술의 발전으로 기존의 일정 비트율 (constant bit rate: CBR) 채널을 통하여 디지털 압축된 다수의 비디오 프로그램들이 전송 가능하다. 본 논문은 이러한 환경에서 전송되는 모든 비디오 프로그램들의 전송률의 합은 일정하지만 각 비디오 프로그램에 대해서는 가변 비트율 (variable bit rate: VBR)을 허용함으로써 VBR 압축의 장점을 얻을 수 있는 결합 비디오 부호화 방식을 제시한다. 본 논문은 다음 두 가지 관점에서 기여한다. 첫째, MPEG-2 비디오에 대한 비트율-왜곡 추정 방식을 제안하는데, 이로 인하여 부호화에 사용되는 양자화 스텝 크기에 따른 한 영상에서 발생하는 비트량과 왜곡을 예측하거나 역으로 발생하는 비트량이나 왜곡에 대한 양자화 스텝 크기를 예측할 수 있다. 제안된 추정 방식의 장점은 실제의 비디오 부호화 응용에 적용할 수 있을 만큼 계산량이 적고 정확하다는 것이다. 둘째, 제안된 비트율-왜곡 추정 결과를 이용한 효율적이고도 정확한 결합 부호화 방식을 제시하고, 실험을 통하여 성능을 검증한다. 실험 결과 제안된 부호화 방식은 소수의 비디오 프로그램이 결합 부호화되어도 한 프로그램뿐만 아니라 프로그램들간에도 일관된 화질이 유지되고, 동일한 채널대역에 추가적인 프로그램의 전송과 같은 두드러진 이득을 얻을 수 있음을 확인하였다.

ABSTRACT

A conventional CBR channel is now capable of delivering several digitally compressed video programs due to recent advances in video compression, such as MPEG-2, and digital transmission technology. This paper presents a joint video coding scheme that is to maintain a constant sum of bit rates for all the programs but to allow the variable bit rate for individual program in the transmission environment mentioned above. Thus advantages of VBR video compression can be obtained. This paper contributes in two aspects. First, a rate-distortion estimation method

*한국과학기술원 전기 및 전자공학과
論文番號: 97330-0912
接受日字: 1997年 9月 12日

for MPEG-2 video is proposed, which enables us to predict the amount of bits and the distortion generated from an encoded picture at a given quantization step size and vice versa. The most attractive features of the proposed rate-distortion estimation method are its accuracy and a computational complexity low enough to be applied to real-time video coding applications. Second, this paper presents an efficient and accurate joint rate control scheme using the rate-distortion estimation results and verifies its performance with experiments. The experimental results show that our coding scheme gives a significant gain even though a small number of video programs are coded jointly. For example, a stable picture quality is maintained among the video programs as well as within a program, and additional extra programs can be transmitted over the same channel bandwidth if the proposed joint video coding scheme is used.

I. 서 론

부호화된 영상이 일정한 화질을 유지하기 위해서는 부호화에 사용되는 비트량이 영상의 복잡도에 따라 변해야 한다. 가변 비트율 (variable bit rate: VBR) 압축은 부호화기의 출력 전송률이 영상이 요구하는 비트량에 따라 변화할 수 있으므로 부호화된 영상의 화질이 일정 비트율 (constant bit rate: CBR) 압축보다 일정하게 유지될 수 있다. 또한 일반적으로 복잡한 영상이 단순한 영상보다 비트량 증가에 의한 왜곡 감소 효과가 크므로 영상의 복잡도에 따라 가변적인 전송률을 수용할 수 있는 VBR 압축이 CBR 압축보다 향상된 화질을 얻을 수 있다. 이러한 측면에서 VBR 압축이 CBR 압축에 비하여 효율적인 압축 부호화 방식으로 알려져 있다 [4]. 이러한 사실에도 불구하고 일정 비트율만을 수용하는 기존의 CBR 채널을 통하여 하나의 영상정보를 전송하는 응용에서는 CBR 압축만을 사용해야 했다. 그러나 현재 MPEG-2 [1] 등과 같은 새로운 영상 압축기술과 디지털 전송기술의 발달로 인하여 기존 전송 채널 대역에 다수의 비디오 프로그램들을 동시에 전송하는 다중 비디오 프로그램 전송 환경이 가능하게 되었다. 이와 같은 환경의 예로는 디지털 위성방송, 디지털CATV, 디지털 지상 방송 등이 있다.

다수의 비디오 프로그램들을 하나의 CBR 채널을 통하여 전송하는 가장 전형적인 방법은 모든 프로그램들을 독립적으로 CBR 압축하여 전송하는 독립 부호화 (independent coding) 방식인데, 이 방식에는 다음 두 가지 중요한 문제가 있다. 첫째, 프로그램 내용이 한 프로그램 내에서 뿐만 아니라 프로그램들간에도 큰 차이를 갖을 수 있다. 따라서 한 프로그램 내의

영상들이 일관되지 못한 화질을 갖을 수 있고, 프로그램들 간에도 큰 화질 차이가 발생할 수 있다. 이는 비디오 프로그램을 수신하는 모든 시청자들에게 만족할만한 품질의 영상을 공정하게 제공하지 못함을 의미한다. 둘째, 모든 프로그램들에 대하여 만족할만한 화질을 보장할 수 있는 충분한 대역을 할당할 경우 단순한 영상을 갖는 프로그램에서는 대역의 낭비를 초래하므로 채널 사용에 있어서 비효율적이다. 이는 한 채널을 통하여 전송 가능한 프로그램의 수가 감소함을 의미한다.

다중 비디오 프로그램 전송 환경에서는 다수의 비디오 프로그램들을 VBR로 압축하여 전송함으로써 VBR 압축의 장점을 얻을 수 있는데, 수행 방법으로는 통계적 다중화 방식 [3]-[5]과 본 논문에서 제시하는 결합 부호화 방식을 고려할 수 있다. 통계적 다중화는 주로 ATM 등에서 사용하는 방식으로 다수의 VBR 트래픽들을 망에서 수용할 경우 전체 비트율의 합이 거의 일정해 지는 사실에 근거한다. 그러나 다중화되는 프로그램 수가 적으면 다중화 효과를 얻기 어렵다. 또한 각 프로그램들이 동시에 높은 비트율의 비트열을 출력하면 부호기 버퍼의 넘침이 발생할 수 있고 이로 인하여 많은 정보 손실이 발생할 수 있다. 이를 방지하기 위해서는 주어진 채널 용량을 모두 사용할 수 없고, 이로 인하여 채널 사용 효율이 감소한다 [6]. 한편 결합 부호화는 그림 1에 나타난 바와 같이 VBR로 압축된 비디오 프로그램을 발생시키는 다수의 부호화기들을 중앙 제어부에서 동시에 제어함으로써 각 비디오 프로그램의 전송률은 가변적이지만 다중화되어 CBR 채널을 통하여 전송되는 모든 프로그램들의 전송률 합은 일정하도록 하는 방식이다. 따라서 이 방식은 통계적 다중화와는 달리 소수의 프로그램

들이 다중화되는 경우에도 채널 대역을 모두 사용할 수 있으므로 다중 비디오 프로그램 전송 환경에 보다 효율적인 부호화 방식이다.

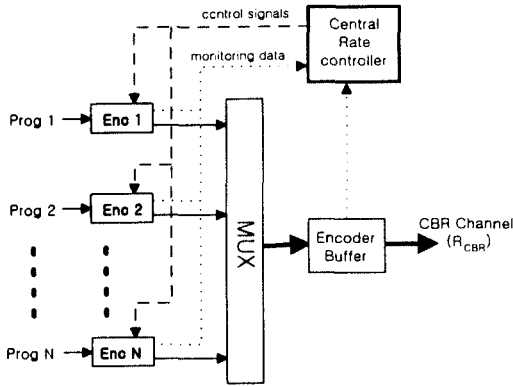


그림 1. 결합 부호화기의 개념적 구성도
Fig. 1 Conceptual diagram of joint video coding scheme.

다중 비디오 프로그램 전송 환경에서 결합 부호화를 수행할 때 고려해야 할 사항은 다음과 같다.

- 1) 모든 비디오 프로그램들의 전송률의 합이 CBR 채널 대역보다 적어야 한다. 만일 전송률의 합이 CBR 채널 대역보다 크면 정보 손실이 발생하므로 재생된 영상의 화질이 크게 훼손된다.
- 2) 각 프로그램들 간에 화질 차이가 적어야 한다. 만일 이 조건을 만족하지 못하면 모든 시청자에게 공정한 서비스를 제공하지 못한다.

본 논문에서는 위의 고려 사항들을 정확하게 만족시키는 MPEG-2 비디오에 대한 결합 비디오 부호화 방식을 제안하는데, 이 부호화 방식은 III장에서 제시하는 비트율-왜곡 추정 결과를 이용한다.

제안된 결합 부호화의 성능을 평가하기 위하여 서로 다른 특성을 갖는 여러 비디오 프로그램들을 MPEG-2로 부호화하여 하나의 CBR 채널을 통하여 전송하는 환경에 제안된 방식과 독립 부호화 방식을 적용하여 실험하였다. 실험 결과 제안된 방식이 각 프로그램을 CBR 압축하는 독립 부호화에 비하여 재생된 프로그램들 사이의 화질이 매우 유사하고 각 프로그램 내에서

도 비교적 일정한 화질을 유지함을 확인할 수 있었다.

본 논문의 구성은 다음과 같다. II장에서 기존 결합 부호화 방식에 대한 간단한 설명과 제안된 방식의 기본 개념을 설명하고, III장에서는 제안된 결합 부호화에 적용되는 비트율-왜곡 추정 방식을 제시한다. 제안된 결합 부호화의 상세한 설명을 IV장에서 다루고, V장에서는 실험을 통하여 제안된 방식의 성능을 평가한 후 VI장에서 결론을 맺는다.

II. 결합 부호화의 기본 개념과 관련 연구

다중 비디오 프로그램 전송 환경에서는 서로 다른 프로그램들로부터 발생된 비트열들이 다중화되어 하나의 CBR 채널을 통하여 전송된다. 지금까지 여러 비디오 프로그램들을 하나의 채널을 통하여 전송하는 다중 비디오 프로그램 전송에 대한 다수의 연구들이 수행되었다. [7]-[9]들은 버퍼의 상태에 따라 양자화 파라미터와 같은 부호화 변수를 조절하여 버퍼의 넘침이나 고갈을 방지하는데 관심을 뒀다. 그러나 이 방식들은 화질에 대한 고려를 하지 않으므로 부호화된 프로그램들 간에 큰 화질 차이가 생길 수 있고, 화질이 버퍼 상태에 따라 큰 변화가 발생할 수 있다. 따라서 이 방식들은 모든 시청자들에게 만족할 만한 화질을 갖는 공정한 비디오 서비스를 보장하지 못한다. 이와 다른 접근 방식들 [10]-[13]은 본 논문의 목적과 같이 여러 비디오 프로그램들을 통합적으로 제어하여 부호화된 비디오 프로그램들의 화질을 향상시키고자 하였다. [10]에서는 간단한 비트율-왜곡 모델을 이용하여 다중화되는 프로그램들의 전체 왜곡을 최소화할 수 있도록 프레임 주기마다 부호화 되는 각 프로그램에 대역을 할당하는 방식을 제시했다. 그러나 간단한 비트율-왜곡 모델은 적용되는 영상에 따라 큰 오차가 발생할 수 있다. 만일 오차를 줄이기 위하여 고차의 모델을 적용할 경우 이 모델의 변수를 구하기 위해서 반복적인 부호화 처리와 같은 추가되는 연산이 증가하므로 구현에 어려움이 있다. 또한 이 방식은 전체 왜곡을 최소화 하는 것이 목적이므로 다중화되는 프로그램들 사이에 화질 차이가 발생할 수 있고, 이로 인하여 모든 시청자들에게 공정한 비디오 서비스 제공을 보장하지 못한다. [11][12]에서는 다중화되는 모든 프로그램들의 프레임과 GOP들로 구성

되는 super-프레임과 super-GOP 개념을 도입하여 마치 하나의 프로그램을 부호화 하는 것과 같은 처리를 수행했다. 그리고 super-프레임에 대하여 할당된 비트량에 가장 근접한 비트를 발생시키는 하나의 양자화 파라미터를 부호화될 모든 영상들에 동일하게 적용함으로써 모든 프로그램들이 균일한 화질을 유지하고자 하였다. 그러나 이 영상들을 동일한 양자화 파라미터를 사용하여 부호화하더라도 영상의 내용에 따라 서로 다른 화질을 가질 수 있다. 또한 부호화에 적용되는 양자화 파라미터를 얻기 위해서는 적용 가능한 양자화 파라미터에 대한 비트 발생량을 미리 구해야 하므로 구현상의 어려움이 있다. [13]에서는 매프레임 주기마다 각 프로그램에 할당하는 비트량을 MPEG-2의 TM5에서 제시한 영상의 복잡도에 비례하게 할당하여 프로그램들간의 화질 차이를 최소화하고자 하였다. 그러나 비트 발생량과 평균 양자화 파라미터의 곱으로 구해지는 TM5의 복잡도는 실험적으로 구해진 근사화 결과이므로 동일한 영상에 대해서도 적용되는 양자화 파라미터에 따라 서로 다른 결과를 나타낸다. 따라서 모든 프로그램들에 대하여 동일한 화질을 보장하기 어렵다. 이러한 문제점들을 해결하기 위하여 본 논문에서는 적용되는 양자화 파라미터에 대한 비트 발생량과 왜곡의 추정에 근거한 보다 직접적인 결합 부호화 방식을 제안한다.

본 논문에서 제안하는 결합 부호화의 목적은 하나의 CBR 채널을 통하여 전송되는 프로그램들을 시청하는 모든 시청자들에게 만족할만한 화질의 영상을 공정하게 제공하는 것인데, 이는 부호화될 각 프로그램의 영상에 대한 비트 할당에 의해서 달성될 수 있다. 즉, 모든 시청자들에게 만족할만한 화질의 영상을 제공하기 위해서는 결합 부호화되는 모든 프로그램들 중에서 최대 왜곡을 발생시키는 프로그램의 왜곡이 최소화하도록 각 프로그램에 비트를 할당해야 한다. 한편 모든 시청자들에게 공정한 서비스를 제공한다는 측면에서는 부호화된 모든 프로그램의 화질이 동일해야 한다. 그런데 영상에 할당되는 비트량과 이에 따른 왜곡은 반비례하므로 부호화된 모든 프로그램의 영상들이 동일한 왜곡 (또는 화질)을 갖을 때 최대 왜곡은 최소화된다. 따라서 본 논문에서는 각 프로그램에 대한 비트 할당을 모든 프로그램들이 동일한 화질을 갖도록 하였다.

다중화되어 CBR 채널을 통하여 전송되는 비디오 프로그램의 수가 N 이고 현재 프레임 기간 동안 부호화될 모든 영상들에 할당되는 전체 비트량을 B_T 라 하면, 모든 프로그램들이 동일한 화질을 유지하기 위해서는 식 (1)을 만족하도록 B_T 를 각 프로그램에 재 할당하면 된다.

$$B_1(D) + B_2(D) + \dots + B_k(D) + \dots + B_M(D) = B_T \quad (1)$$

where $B_T \leq B_{max}$

여기서 $B_k(D)$ 는 부호화된 k 번째 프로그램의 영상으로부터 발생하는 왜곡이 D 가 되기 위해서 필요한 비트량이고, B_{max} 는 부호기 버퍼가 현재 프레임 기간 동안 버퍼의 넘침이 발생하지 않으면서 수용 가능한 최대 비트량을 의미한다. 그러나 실제적으로는 모든 프로그램들의 화질이 절대적으로 동일할 수는 없으므로 각 프로그램에 대한 비트 할당은 전체 할당 비트량이 B_T 보다 크지 않으면서 임의의 두 프로그램 사이의 화질 차이가 최소화되도록 해야 한다.

제안된 결합 부호화에서 전체 부호화 비트량 B_T 를 각 프로그램에 재 할당하는 과정은 정보이론에서 알려진 사실에 근거한다. 만일 하나의 복합 신호 (composite signal)가 두개 이상의 독립적인 신호 성분으로 분해 가능하면 복합 신호에 대한 비트율-왜곡 함수는 독립적인 신호의 비트율-왜곡 함수로부터 직접 구해진다 [14]. 이 사실에 근거한 비트 재할당 과정의 기본 개념은 그림 2에 나타난 바와 같이 현재 부호화될 각 프로그램의 영상들에 대한 비트율-왜곡 함수 $R_k(D)$ 들을 비트율축 (수직축)에 대하여 더함으로써 전체 비트율-왜곡 함수 $R_T(D)$ 를 구하고, $R_T(D)$ 로부터 B_T 에 대응되는 왜곡 D_T 를 구한 후, D_T 와 $R_k(D)$ 로부터 현재 부호화될 k 번째 프로그램의 영상에 할당할 비트량 B_k 을 구하는 것이다. 결국 현재 부호화될 k 번째 프로그램의 영상에 할당되는 비트량은 B_k 이고 왜곡은 모든 프로그램들에 대하여 D_T 가 된다.

그러나 이 비트 재 할당 과정에서 사용되는 비트율-왜곡 함수를 구하기 위해서는 MC-DCT 방식의 영상 부호화기에서 사용되는 양자화기의 종류가 31개인 경우 31번의 양자화, 역양자화, IDCT, VLC, 그리고 왜곡 측정을 위한 처리과정이 추가로 필요하므로 실제 부호화 과정에 이를 적용하기는 어렵다. 따라서 본

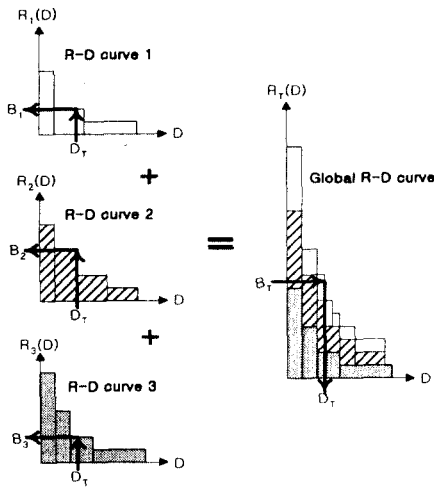


그림 2 부호화 비트 재 할당의 기본 개념
Fig. 2 Basic concept of bits re-allocation

논문에서는 적용되는 양자화기에 대응되는 비트율-왜곡을 추정하는 방식을 제안하여 적용함으로써 추가되는 연산 처리과정을 대폭 감소시켰다.

Ⅲ. 비트율-왜곡 추정

MPEG-2 [1]와 같이 MC-DCT 방식의 영상 부호화를 수행하는 표준 비디오 부호화 방식들 [17]-[19]은 일반적으로 양자화 과정에 적용되는 양자화 스텝 크기 (quantization step size: QS)를 조절하여 발생 비트량 또는 화질을 제어한다. 따라서 영상 부호화에 적용되는 QS에 대한 발생 비트량과 왜곡의 정확한 추정은 효율적이고도 정확한 비디오 부호화 제어에 도움을 준다. 여기서 양자화 스텝 크기 QS는 양자화 파라미터 QP의 함수로써 n 번째 양자화 파라미터를 QP_n (보통 $QP_n = n$)이라 하고 이에 대응하는 QS를 QS_n 이라 하면 QS와 QP의 관계는 보통 $QS_n = 2 \times QP_n$ 이고 $QP_n < QP_{n+1}$ 이다.

이미 비트율-왜곡 추정에 관한 여러 연구들 [14]-[16]이 발표되었으나 현행 표준 비디오 부호화 방식들에 대한 정확한 모델을 제시하지 못하고 여전히 많은 계산량이 요구된다. 따라서 본 장에서는 MPEG-2를 기본 부호화 방식으로 사용하는 결합 부호화에 적용하기 위한 비트율-왜곡 추정방식을 제시하고 실험 결과

를 보인다. 실험에 사용되는 부호화기는 MPEG-2의 TM5 [2]를 적용하였다. 제안된 비트율-왜곡 추정의 주된 연산은 입력 신호에 대하여 히스토그램 또는 가중 히스토그램을 구하는 것이므로 계산량이 적고, 최종 결과식이 간단하다. 또한 제안된 추정 과정은 H.261 [17], MPEG-1 [18], H.263 [19]등과 같은 다른 표준 비디오 부호화에도 동일한 방식으로 적용 가능하다.

1. 적용되는 양자화 파라미터(QP)에 대응하는 비트 발생량 추정

MPEG-2 비디오 부호기의 출력 정보는 움직임 벡터, 각종 헤더 등과 같은 부가 정보와 DCT 계수로부터 발생하는 신호 정보로 구성된다. 여기서 양자화 파라미터 QP의 변화에 영향을 받는 신호 성분은 intra-DC를 제외한 DCT 계수들이고, 높은 비트율로 부호화할 경우 부호화기의 출력 비트량은 대부분 DCT 계수로부터 발생된다. 양자화 처리된 DCT 계수들은 run-length 부호화에 의하여 (zero-run, level)인 codeword로 변환된 후 가변 길이 부호화(VLC)하여 최종적인 비트열로 만들어진다. 여기서 (zero-run, level)로 표현되는 codeword의 수는 양자화 처리된 DCT 계수 중 영이 아닌 성분 (non-zero coefficient: NZC)의 수와 일치한다. 따라서 직관적으로 양자화된 DCT 계수로부터 발생하는 비트량은 NZC의 수에 비례할 것이라고 생각할 수 있다. 만일 한 영상에서 발생하는 codeword들에 대한 평균 코드 길이가 부호화에 적용되는 QP에 무관하게 일정하다면 DCT 계수로부터 발생하는 비트량과 NZC의 수는 선형적으로 비례하게 된다. 한편 부가 정보로부터 발생하는 비트량은 동일한 부호화 방식으로 부호화된 이전 영상에서 발생하는 비트량과 비슷한 양을 갖는다. 이와 같은 가정으로부터 적용되는 QP에 대응하는 비트 발생량 추정이 가능하다.

그림 3은 부호화 방식이 다른 I, P 그리고 B 프레임을 부호화하였을 때 MB (MacroBlock)당 양자화된 DCT 계수로부터 발생하는 비트량과 NZC 개수의 관계를 나타낸 것으로, 발생 비트량과 NZC의 개수는 비교적 선형적으로 비례함을 보여준다. 한편 그림 4는 I, P 그리고 B 프레임에 대하여 적용되는 QP 값을 변화 시킬 때 DCT 계수로부터 발생하는 비트량과 NZC 수의 관계를 나타낸 것이다. 여기서 한 프레임에 포함된 모든 MB들에는 동일한 QP 값이 적용되고, P 또

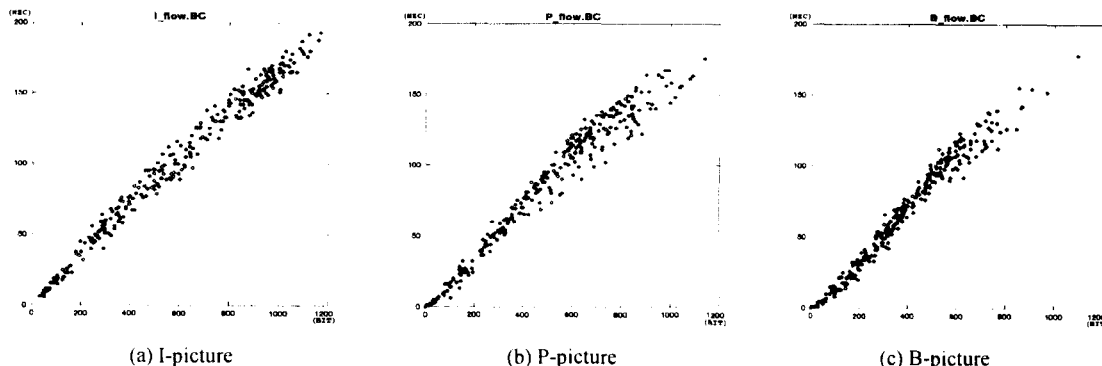


그림 3. MB당 발생하는 codeword 수와 발생 비트량의 관계:
Fig. 3 Relationship between generated codewords and bits per MB

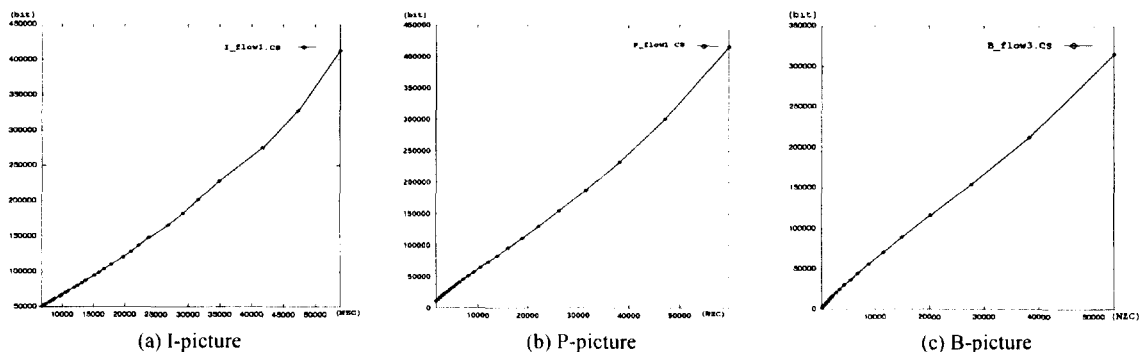


그림 4. 양자화 파라미터에 따른 프레임당 발생 비트량과 발생 codeword와의 관계.
Fig. 4 Relationship between generated codewords and bit amount per frame according to the quantization parameter.

는 B프레임을 부호화할 때 기준 프레임 (I 또는 P 프레임)은 QP=6으로 부호화한 영상을 사용하였다. 이 결과는 적용되는 양자화 파라미터 QP가 1인 경우를 제외하면 영상의 내용이나 부호화 방식에 무관하게 적용되는 QP에 따라 발생하는 비트량과 NZC 개수의 관계가 비교적 선형적으로 비례함을 보여준다. 그런데 보통 QP가 1인 경우는 부호화기 버퍼의 고갈을 방지하기 위해서만 사용되고 정상적인 부호화 과정에서는 사용되지 않는다. 따라서 본 논문에서는 적용되는 양자화 파라미터에 따라 DCT 계수로부터 발생하는 비트량 $Sbit$ 을 NZC의 개수를 추정함으로써 식 (2)와

같이 추정한다.

$$Sbit(QP_n) = \alpha \times NZC(QP_n) \tag{2}$$

여기서 비례계수 α 는 현재 부호화할 영상과 같은 부호화 방식 (picture coding type)을 갖는 이전 영상의 발생 비트량과 NZC의 수로부터 구한 값을 사용하는데, 보통 5.5에서 7.0 사이의 값을 갖는다.

MPEG-2에서 사용되는 양자화 과정은 부호화 방식에 따라 intra와 non-intra 양자화 과정으로 분류되는데, MPEG-2의 TM5에서 제시한 양자화기는 다음과

같다 [2].

- Intra quantizer

$$\overline{ac}_I(i, j) = (16 \times ac(i, j)) / W_I(i, j) \quad (3)$$

$$QAC_I(i, j) = \frac{[\overline{ac}_I(i, j) + \text{sign}(\overline{ac}_I(i, j)) \times ((3 \times QP) / 4)]}{2 \times QP} \quad (4)$$

- Non-intra quantizer

$$\overline{ac}_N(i, j) = (16 \times ac(i, j)) / W_N(i, j) \quad (5)$$

$$QAC_N(i, j) = \frac{[\overline{ac}_N(i, j)]}{2 \times QP} \quad (6)$$

Intra-DC를 제외한 DCT 계수 $ac(i, j)$ 에 대한 양자화 과정은 우선 식 (3) 또는 (5)와 같이 양자화 매트릭스 (quantization matrix)에 의한 개별적인 양자화를 수행한다. 여기서 $W_I(i, j)$ 와 $W_N(i, j)$ 는 각각 intra 그리고 non-intra 양자화 매트릭스의 (i, j) 성분이다. 양자화 매트릭스 처리된 DCT 계수 (즉, $\overline{ac}_I(i, j)$ 또는 $\overline{ac}_N(i, j)$)는 식 (4) 또는 (6)과 같이 양자화 파라미터 QP에 의해서 양자화 처리된다. 여기서 QP는 보통 비트율 제어 알고리즘에 의해서 MB 또는 이 이상의 계층 단위로 조절될 수 있다. 한편 MPEG-2에서 사용되는 역 양자화는 각 MB의 부호화 방식에 따라 식 (7) 또는 (8)로 표현되는 다음 역 양자화기를 적용하여 수행된다 [1].

- Intra dequantizer

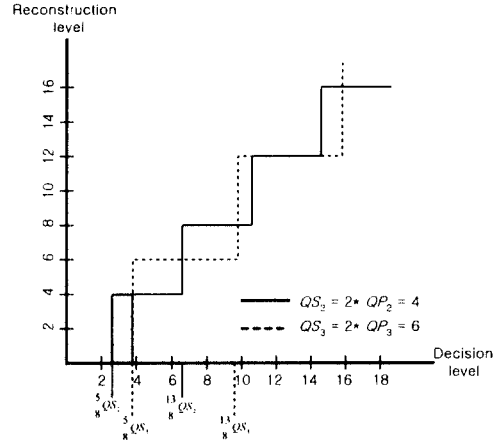
$$REC_I(i, j) = (QAC_I(i, j) \times W_I(i, j) \times QS) / 16 \quad (7)$$

- Non-intra dequantizer

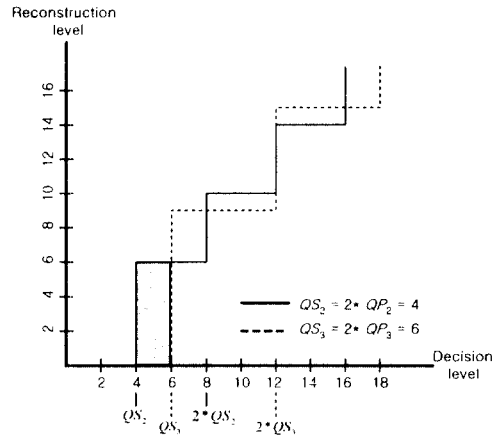
$$REC_N(i, j) = \{ [2 \times QAC_N(i, j) + \text{sign}(QAC_N(i, j))] \times W_N(i, j) \times QP \} / 32 \quad (8)$$

그림 5는 양자화 파라미터 QP가 2, 3일 때 식 (4)와 (7)로부터 얻어지는 intra 양자화와 식 (6)과 (8)로부터 얻어지는 non-intra 양자화의 입출력 특성을 나타낸 것이다. 빗금 친 영역에 존재하는 입력 신호 (즉 $\overline{ac}_I(i, j)$ 또는 $\overline{ac}_N(i, j)$)들에 대한 양자화 처리 결과는 양자화에 적용되는 양자화 파라미터 QP가 2일 때는 영이 아니고 3일 때는 영이 된다.

MPEG-2에서 제시한 양자화기를 사용하여 부호화할 때 적용되는 양자화 파라미터에 따라 발생하는



(a) Intra 양자화



(b) Non-intra 양자화

그림 5. MPEG-2 양자화 특성

Fig. 5 MPEG-2 quantization characteristics

NZC의 수는 다음 절차에 의해서 쉽게 구할 수 있고, 이 결과를 식 (2)에 대입함으로써 DCT 계수로부터 발생하는 비트량을 추정할 수 있다.

S1) $AC(i, j) = \text{clipping}[\beta \times |\overline{ac}(i, j)|]$ 의 히스토그램을 구한다.

$$NZC_pdf[AC(i, j)] + + \quad (9)$$

여기서 $\text{clipping}[\cdot]$ 은 입력을 0에서 31로 제한된 정수

로 변환하는 연산자이고, β 는 식 (4)또는 (6)으로부터 유도되는 상수로서 intra 양자화기의 경우는 0.8이고 non-intra 양자화기의 경우는 0.5이다. 따라서 $AC(i, j)$ 는 양자화 처리된 DCT 계수의 절대값 1이 되기 위한 QP 값을 의미한다. 예를 들어, non-intra 양자화기의 경우 식 (6)으로부터 $QAC_N(i, j)$ 의 절대값을 1로 만드는 QP 값은 $QP = 0.5 \times |\overline{ac}(i, j)|$ 이다. 결국 $NZC_pdf[QP_n]$ 에는 양자화 과정에 적용되는 QP가 OP_n 에서 OP_{n+1} 로 증가할 때 이에 대응하는 NZC의 감소량이 저장된다. 즉 OP_n 을 적용하여 양자화했을 때는 영이 아니고 한단계 증가된 OP_{n+1} 를 적용하여 양자화하면 영으로 바뀌는 DCT 계수의 개수가 $NZC_pdf[QP_n]$ 에 저장되는 것이다. 예를 들어 만일 양자화에 적용되는 QP가 2라면 $NZC_pdf[2]$ 에 저장되는 값은 그림 5에 나타난 양자화의 입출력 특성에서 빗금 친 영역에 존재하는 입력 신호(즉 $\overline{ac}_I(i, j)$ 또는 $\overline{ac}_N(i, j)$)의 개수이다.

S2) $NZC_pdf[QP]$ 의 누적 값을 구한다.

$$NZC_cdf[QP_n] = \sum_{QP=n}^{31} NZC_pdf[QP], 1 \leq n \leq 31 \quad (10)$$

결국 $NZC_cdf[QP_n]$ 에는 적용되는 양자화 파라미터가 QP_n 일 때 발생하는 NZC의 개수가 저장된다. 그림 6은 식 (9)와 (10)으로부터 얻어진 $NZC_pdf[QP]$ 와 $NZC_cdf[QP]$ 의 예를 보여준다.

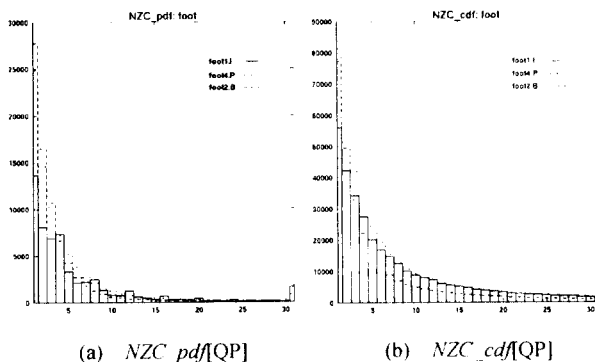


그림 6. I, P, B 프레임에 대한 $NZC_pdf[QP]$ 와 $NZC_cdf[QP]$ 의 예

Fig. 6 An example of $NZC_pdf[QP]$ and $NZC_cdf[QP]$ for I, P, and B frame

이와 같이 DCT 계수를 양자화 매트릭스 처리한 결과인 $\overline{ac}_I(i, j)$ 또는 $\overline{ac}_N(i, j)$ 의 절대값에 0.8 또는 0.5을 곱한 값들의 히스토그램과 히스토그램의 누적 값을 구함으로써 적용되는 양자화 파라미터에 따라 발생하는 NZC의 수를 구할 수 있고, 이 값에 식 (2)의 관계를 적용하면 양자화 파라미터에 따라 DCT 계수들로부터 발생하는 비트량을 추정할 수 있다. 따라서 한 영상으로부터 발생하는 전체 비트량은 식 (11)로 추정할 수 있다.

$$Ebit(QP_n) = Sbit(QP_n) + Obit \quad (11)$$

여기서 $Obit$ 는 같은 부호화 방식을 갖는 이전 영상에서 DCT 계수를 제외한 부가 정보로부터 발생한 비트량이다.

2. 적용되는 양자화 파라미터에 대응하는 왜곡 추정

제안된 왜곡 추정 과정은 다음 두 가지 사실에 근거한다. 첫째, 평균 자승 오차 (mean squared error: MSE) 관점에서의 왜곡은 양자화에 적용되는 양자화 스텝 크기 QS의 제곱에 비례한다. 실제적으로 표준 비디오 부호화 방식들에 사용되는 양자화기는 균일 양자화기 (uniform quantizer)인데, 만일 양자화 과정이 이상적인 균일 양자화이고 양자화기의 입력 신호가 균일 분포를 갖는다면 적용되는 QS에 대한 왜곡(MSE)은 $QS^2/12$ 이다. 그런데 MPEG-2에서의 양자화는 모든 DCT 계수들에 대하여 8×8 block단위로 양자화 매트릭스 처리를 수행한다. 따라서 양자화 스텝 크기가 QS_{n-1} 에서 QS_n 으로 증가할 때 양자화 매트릭스의 영향을 고려한 왜곡의 증가량은 $(QS_n^2 - QS_{n-1}^2) \times W(i, j)$ 에 비례한다. 둘째, 양자화 스텝 크기가 QS_n 에서 QS_{n+1} 로 한단계 증가할 때 왜곡의 증가에 영향을 미치는 입력 신호 성분들은 QS_n 으로 양자화 처리했을 때 영이 되지 않은 성분들이다. 즉 QS_{n-1} 을 적용했을 때 이미 영인 신호 성분들은 왜곡의 증가에 영향을 미치지 못하고 다만 QS_{n-1} 을 적용했을 때 존재하는 영이 아닌 신호 성분들만이 왜곡의 증가에 영향을 미친다. 본 절에서는 이와 같은 사실을 기반으로 하여 MPEG-2 부호화에 있어서 양자화 과정에 적용되는 QS에 대한 MSE관점에서의 왜곡을 추정하는데, 제안된 왜곡 추정은 다음 절차에 따라 구해지는 가중 히스토그램과

이 가중 히스토그램의 누적 값을 사용한다.

S1) 양자화 매트릭스를 고려한 $AC(i, j)$ 의 가중 히스토그램을 구한다.

$$QM_pdf[AC(i, j)] + W^2(i, j) \quad (12)$$

여기서 $AC(i, j)$ 는 식 (9)에서 정의된 값이고, $W^2(i, j)$ 는 양자화 매트릭스의 (i, j) 번째 성분을 제공한 값으로 미리 계산된 값을 사용하면 되므로 실제적인 연산은 덧셈 처리이다.

S2) 가중 히스토그램 $QM_pdf[QP]$ 의 누적 값을 구한다.

$$QM_cdf[QP_n] = \sum_{QP=n}^{31} QM_pdf[QP], \quad 1 \leq n \leq 31 \quad (13)$$

식 (13)으로부터 얻어지는 $QM_cdf[QP_n]$ 에는 적용되는 양자화 파라미터가 QP_n 일 때 양자화된 DCT 계수들 중 NZC에 해당되는 양자화 매트릭스의 제공 값들의 합이 저장된다. 그림 7은 MPEG-2의 기본 양자화 매트릭스를 사용했을 때 식 (12)로부터 얻어진 가중 히스토그램과 식 (13)으로부터 얻어진 $QM_cdf[QP]$ 의 예를 보여준다.

본 절에서 제안하는 왜곡 추정은 위에서 언급한 MPEG-2의 양자화 특성에서 결정 준위(decision level)

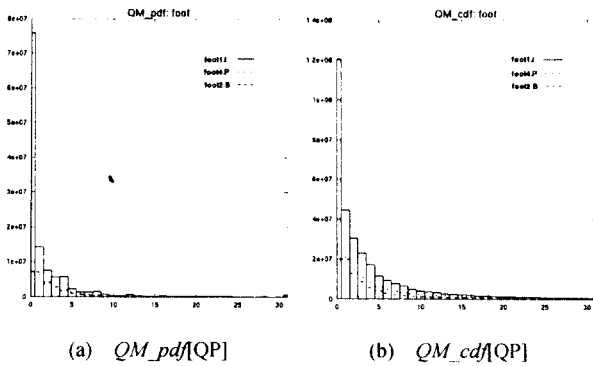


그림 7. I, P, B 프레임에 대한 $QM_pdf[QP]$ 와 $QM_cdf[QP]$ 의 예

Fig. 7 An example of $QM_pdf[QP]$ and $QM_cdf[QP]$ for I, P, and B frame

사이에 존재하는 NZC들의 분포가 균일하다는 가정 하에서 유도된다. 또한 MPEG-2의 양자화 처리는 부호화 방식에 따라 intra와 non-intra 양자화 처리로 분류되고, 각 양자화 처리의 특성이 서로 다르므로 제안된 왜곡 추정방식은 적용되는 양자화 과정에 따라 다른 결과식을 갖는다.

A. Intra 양자화에 대한 왜곡 추정

MPEG-2의 intra 양자화 특성에서 결정 준위 사이에 존재하는 입력 신호가 균일 분포를 갖는다면, 적용되는 양자화 스텝 크기가 $QS_n (n > 1)$ 일 때 발생하는 왜곡(MSE)은

$$\frac{1}{QS_n} \int_{\frac{5}{8}QS_n}^{\frac{13}{8}QS_n} \frac{1}{QS_n} (x - QS_n)^2 dx = \frac{19}{192} QS_n^2$$

이다(그림 5의 (a) 참조). 따라서 양자화 스텝이 QS_{n-1} 에서 QS_n 로 한단계 증가할 때 $\frac{5}{8} QS_n$ 보다 큰 값을 갖는 입력 신호 성분들에 의해 증가된 왜곡의 합은 (14)와 같다.

$$\Delta RE(QP_n) = \frac{19 \times (QS_n^2 - QS_{n-1}^2)}{192} \times \frac{QM_cdf[QP_n]}{16^2}, \quad (n > 1) \quad (14)$$

한편 $\frac{5}{8} QS_{n-1}$ 에서 $\frac{5}{8} QS_n$ 사이에 존재하는 균일 분포를 갖는 입력 신호 성분들에 의한 왜곡의 증가량은

$$\frac{1}{\frac{5}{8}(QS_n - QS_{n-1})} \int_{\frac{5}{8}QS_{n-1}}^{\frac{5}{8}QS_n} x^2 dx - \frac{19}{192} QS_{n-1}^2 = (6QS_{n-1}^2 + 25QS_{n-1}QS_n + 25QS_n^2)/192.$$

이 된다. 따라서 $\frac{5}{8} QS_{n-1}$ 에서 $\frac{5}{8} QS_n$ 사이에 존재하는 NZC들에 의해 증가되는 왜곡의 합은 식(15)로 표현할 수 있다.

$$\Delta RE(QP_n) = \frac{(6QS_{n-1}^2 + 25QS_{n-1}QS_n + 25QS_n^2)}{192} \times \frac{QM_pdf[QP_{n-1}]}{16^2},$$

$$(n > 1) \tag{15}$$

결국 양자화 스텝 크기가 QS_{n-1} 에서 QS_n 으로 한 단계 증가할 때 intra 양자화에 의한 왜곡의 증가량은 식 (16)으로 추정되고 왜곡은 식 (17)로 추정할 수 있다.

$$\Delta D_I(QP_n) = \Delta E_I(QP_n) + \Delta RE_I(QP_n), \quad (n > 1) \tag{16}$$

$$D_I(QP_n) = D_I(QP_{n-1}) + \Delta D_I(QP_n) / I_pixel_num \tag{17}$$

여기서 I_pixel_num은 intra 방식으로 부호화되는 부호화 단위 내에 존재하는 화소의 개수이다. 예를 들어 부호화 단위가 block이라면 I_pixel_num은 64이고, 부호화 단위가 한 프레임이라면 한 프레임 내에 존재하는 화소의 개수이다.

한편 제안된 왜곡 추정 과정에서 intra 양자화에 대한 초기 왜곡은 0에서 $\frac{5}{8} QS_1$ 사이에 존재하는 균일 분포를 갖는 입력 신호 성분에 대한 왜곡이

$$\frac{1}{QS_1} \int_0^{\frac{5}{8} QS_1} x^2 dx = \frac{25}{192} QS_1^2$$

이므로 NZC에 의한 초기 왜곡은 식 (18)로 추정할 수 있다.

$$D_I(QP_1) = \frac{DC_mse + (25QS_1^2 \times QM_pdf[0] + 19QS_1^2 \times QM_cdf[1]) / (192 \times 16^2)}{I_pixel_num} \tag{18}$$

여기서 DC_mse는 DCT계수 중 intra-DC 성분으로부터 발생하는 왜곡의 합이다.

B. Non-intra 양자화에 대한 왜곡 추정

Non-intra 양자화에 대한 왜곡 추정과 동일한 유도 과정을 따른다. Non-intra 양자화 특성에서 결정 준위 사이에 존재하는 입력 신호가 균등 분포를 갖는다면, 적용되는 양자화 스텝 크기가 QS_n 일때 왜곡(MSE)은

$$\frac{1}{QS_n} \int_{QS_n}^{2QS_n} \left(x - \frac{3}{2} QS_n\right)^2 dx = \frac{1}{12} QS_n^2$$

이 된다. 따라서 양자화 스텝 크기가 QS_{n-1} 에서 QS_n 로

한단계 증가할 때 QS_n 보다 큰 값을 갖는 입력 신호 성분에 대한 증가된 왜곡의 합은 식 (19)와 같다.

$$\Delta E_N(QP_n) = \frac{(QS_n^2 - QS_{n-1}^2)}{12} \times \frac{QM_cdf[QP_n]}{16^2}, \quad (n > 1) \tag{19}$$

한편 QS_{n-1} 에서 QS_n 사이에 존재하는 균등분포를 갖는 입력 신호 성분에 의한 왜곡의 증가는

$$\begin{aligned} & \frac{1}{(QS_n - QS_{n-1})} \int_{QS_{n-1}}^{QS_n} x^2 dx - \frac{1}{12} QS_{n-1}^2 \\ &= \frac{1}{3} \left(QS_n^2 + QS_n QS_{n-1} + \frac{3}{4} QS_{n-1}^2 \right) \end{aligned}$$

이다. 따라서 QS_{n-1} 에서 QS_n 사이에 존재 NZC들에 의해 증가되는 왜곡의 합은 식 (20)과 같다.

$$\begin{aligned} \Delta RE_N(QP_n) &= \frac{\left(QS_n^2 + QS_n QS_{n-1} + \frac{3}{4} QS_{n-1}^2 \right)}{3} \\ &\times \frac{QM_pdf[QP_{n-1}]}{16^2}, \quad (n > 1) \tag{20} \end{aligned}$$

결국 양자화 스텝 크기가 QS_{n-1} 에서 QS_n 으로 한 단계 증가할 때 non-intra 양자화에 의해 증가되는 왜곡의 합은 식 (21)로 추정되고, 왜곡은 식 (22)로 추정할 수 있다.

$$\Delta D_N(QP_n) = \Delta E_N(QP_n) + \Delta RE_N(QP_n), \quad (n > 1) \tag{21}$$

$$D_N(QP_n) = D_N(QP_{n-1}) + \Delta D_N(QP_n) / N_pixel_num \tag{22}$$

여기서 N_pixel_num은 non-intra 방식으로 부호화되는 부호화 단위 내에 존재하는 화소의 개수이다.

한편 제안된 왜곡 추정 과정에서 non-intra 양자화에 대한 초기 왜곡은 0에서 QS_1 사이에 존재하는 균등 분포를 갖는 입력 신호 성분에 대한 왜곡이

$$\frac{1}{QS_1} \int_0^{QS_1} x^2 dx = \frac{1}{3} QS_1^2$$

이므로 NZC에 의한 초기 왜곡은 식(23)과 같다.

$$D_N(QP_1) = \frac{QS_1^2 \times (4 * QM_pdf[0] + QM_cdf[1])}{12 \times 16^2 \times N_pixel_num} \tag{23}$$

만일 양자화 처리를 프레임 단위로 처리한다면 I 프레임에는 intra 양자화 처리만이 사용되지만 P 또는 B 프레임의 경우는 각 MB의 부호화 방식에 따라 intra와 non-intra 양자화 처리가 사용된다. 따라서 만일 왜곡 추정을 한 프레임 단위로 수행할 경우는 식 (24)를 사용하여 왜곡을 추정해야 한다.

$$D(QP_n) = D(QP_{n-1}) + \{\Delta D_I(QP_n) + \Delta D_N(QP_n)\} / \text{pixel_num} \quad (24)$$

여기서 pixel_num은 한 프레임 내에 존재하는 화소의 개수이고, 왜곡 증가량의 합인 $\Delta D_I(QP_n)$ 과 $\Delta D_N(QP_n)$ 의 산출에 사용되는 가중 히스토그램과 가중 히스토그램의 누적 값은 부호화 방식에 따라 각각 구해진 값들이다. 즉 $\Delta D_I(QP_n)$ 의 계산에는 intra 양자화 처리되는 DCT 계수들로부터 얻어진 가중 히스토그램과 누적 값이 사용되고, $\Delta D_N(QP_n)$ 의 계산에는 non-intra 양자화 처리되는 DCT 계수들로부터 얻어진 가중 히스토그램과 누적 값이 사용된다.

그림 8은 특성이 다른 여러 영상에 대하여 식 (24)에 의해 추정된 왜곡과 실제로 양자화 파라미터를 변화 시키면서 측정한 왜곡의 관계를 보여준다. 이 결과들은 추정된 왜곡이 영상의 특성에 무관하게 측정한 왜곡과 선형적으로 비례함을 보여준다. 또한 그림 9의 (a)는 제안된 왜곡 추정의 결과가 부호화 방식에 무관

하게 측정한 왜곡과 선형적으로 비례함을 보여주고, (b)는 적용되는 양자화 파라미터에 따른 추정 왜곡과 측정 왜곡의 관계를 나타낸 것으로, 제안된 왜곡 추정이 실제 부호화 과정에 적용할 수 있을 만큼 정확함을 보여준다.

IV. 제안된 결합 부호화

본 장에서는 제안된 결합 부호화 처리과정에 대한 세부적인 내용을 기술한다. 제안된 결합 부호화는 매 프레임 주기마다 다음의 처리과정에 의하여 수행된다. 첫번째 단계에서는 현재 부호화할 각 영상에 대하여 III장에서 제시한 추정방식에 의한 비트율-왜곡 함수를 구한다. 이 과정에 의하여 적용되는 양자화 파라미터에 따라 각 영상으로부터 발생하는 비트량과 왜곡을 예측할 수 있다. 두 번째 단계에서는 현재 부호화할 모든 영상들에 할당할 전체 목표 비트량을 구한다. 이때 이 전체 목표 비트량은 부호기 버퍼의 넘침이 발생하지 않는 범위에 있도록 조정되어지므로 버퍼 넘침에 의한 정보손실을 방지한다. 세 번째 단계는 전체 목표 비트량으로 결합 부호화될 모든 프로그램의 영상들에 대한 목표 왜곡을 구하는 과정이다. 네 번째 단계는 각 영상에 대한 왜곡 추정 결과를 이용하여 부호화한 영상의 왜곡이 목표 왜곡에 근접하도록 하는 양자화 파라미터를 구하고, 이 양자화 파라미터를 이

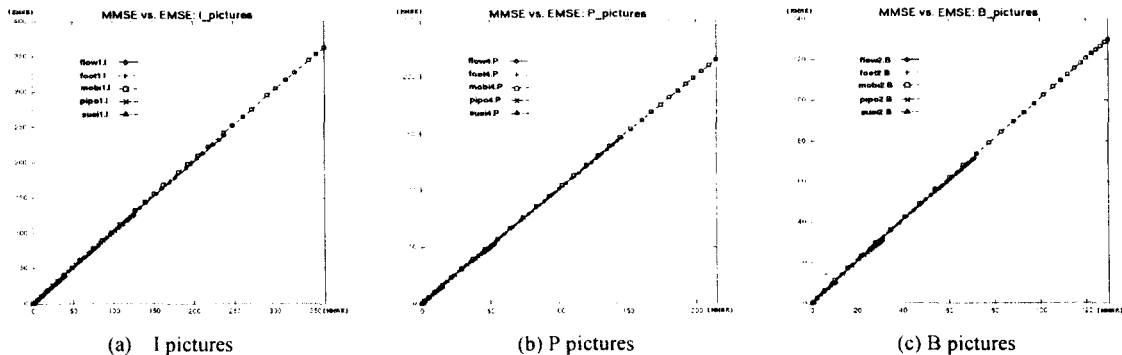


그림 8. 여러 실험 영상에 대한 측정 왜곡(MMSE)과 추정 왜곡(EMSE)의 관계
 Fig. 8 Relationship between measured distortion (MMSE) and estimated distortion (EMSE) for several test images

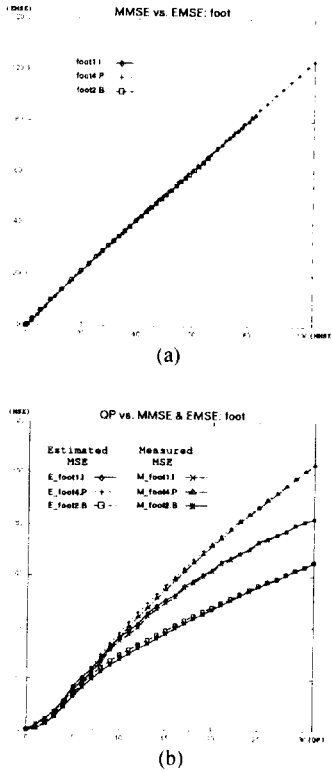


그림 9. 다른 영상 부호화 방식에 대한 왜곡 추정 결과
 (a) 측정왜곡(MMSE)과 추정왜곡(EMSE)의 관계
 (b)QP변화에 따른 측정왜곡과 추정왜곡

Fig. 9 Distortion estimation results for the different picture coding type
 (a) relationship between MMSE and EMSE
 (b)MMSE and EMSE according to the applied QP

용하여 부호화하는 과정이다.

1. 전체 목표 비트량 결정

전체 목표 비트량은 다음 절차에 의해서 구해진다. 현재 부호화할 각 프로그램 영상에 대한 개별적인 목표 비트량을 독립적으로 구하고, 구해진 각 프로그램 영상의 목표 비트량들을 더하여 전체 목표 비트량을 구한다. 그리고 이 전체 목표 비트량에 따라 결합 부호화한 결과가 부호화기 버퍼의 넘침이 발생하지 않는 범위에 있도록 전체 목표 비트량을 재조정한다.

각 프로그램 영상에 대한 개별적인 목표 비트량은 MPEG-2 TM5에서 제시하는 목표 비트 할당(target bit

allocation) 방식을 확장하여 구한다. 현재 부호화할 k 번째 프로그램 영상의 부호화 방식이 $pct \in \{I, P, B\}$ 일 때 이 영상에 할당되는 비트량 $T_{pct}(k)$ 은 다음과 같다.

$$T_{pct}(k) = \frac{X_{pct}(k)/K_{pct}}{\sum_{pct \in \{I, P, B\}} \frac{X_{pct}(k)}{K_{pct}}} R(k)$$

여기서 $X_{pct}(k)$ 는 k 번째 프로그램의 영상에 대한 복잡도로써 현재 영상과 동일한 부호화 방식을 갖는 이전 영상에 적용된 평균 양자화 파라미터와 발생 비트량의 곱이다; K_{pct} 는 양자화 매트릭스에 관련된 상수로 $K_I = K_P = 1.0$ 이고 $K_B = 1.4$ 이다; $N_{pct}(k)$ 는 k 번째 프로그램의 현재 GOP에서 부호화 방식이 pct 인 아직 부호화되지 않은 영상의 수이다; $R(k)$ 는 k 번째 프로그램의 현재 GOP에 할당된 비트량 중 잔여 비트량으로 다음과 같이 갱신된다:

매 프레임 주기마다 모든 영상이 부호화된 후 $R(k)$ 는 다음과 같이 갱신된다.

$$R(k) = R(k) - E_T/N$$

여기서 E_T 는 현재 프레임 기간동안 결합 부호화되는 모든 영상들로부터 발생한 비트량이고, N 은 결합 부호화되는 프로그램의 수이다; 한편 각 프로그램의 GOP 주기마다 $R(k)$ 는 다음과 같이 갱신된다.

$$R(k) = G(k) + R(k)$$

$$G(k) = (R_T \times N_{GOP}(k)) / (N \times \text{frame_rate})$$

여기서 R_T 는 전체 프로그램에 대한 채널 전송률이고, $N_{GOP}(k)$ 는 k 번째 프로그램의 한 GOP에 포함된 영상의 수이다.

이와 같이 구해진 각 프로그램에 대한 목표 비트량을 더하여 현재 부호화할 모든 프로그램 영상에 할당할 전체 목표 비트량 B_T 를 구한다.

$$B_T = \sum_{k=1}^N T_{pct}(k). \tag{25}$$

그러나 비트를 추정의 부정확성으로 인하여 추정된 전체 목표 비트량과 실제로 부호화하여 발생하는 비트량이 일치하지 않을 수 있으므로 전체 목표 비트

량 B_T 가 지나치게 높으면 부호기 버퍼의 넘침이 발생할 수 있다. 따라서 추정된 전체 목표 비트량은 부호기 버퍼 상태를 고려하여 다음과 같이 재 조정할 필요가 있다.

모든 프로그램들의 비트열들이 버퍼 크기가 B_{max}^e 인 부호기 버퍼에 저장된 후 채널 전송률이 B_T 인 CBR 채널을 통하여 전송될 경우 부호기 버퍼의 넘침이 발생하지 않을 조건은 다음과 같다.

$$B_i^e \leq B_{max}^e, \quad \forall i \quad (26)$$

$$B_i^e = B_{i-1}^e + E_i - R_T T \quad (27)$$

여기서 T 는 프레임 주기이고 B_i^e 는 iT 인 순간 부호기 버퍼의 총만도이며 E_i 는 $(i-1)T$ 에서 iT 기간 동안 부호기 버퍼에 입력되는 비트량이다. 식 (26)과 (27)로부터 부호기 버퍼의 넘침이 발생하지 않을 E_i 의 범위는

$$E_i \leq B_{max}^e + R_T T - B_{i-1}^e (= E_i^{max})$$

이다. 따라서 식 (25)에서 구한 전체 목표 비트량 B_T 는 식 (28)과 같이 재 조정된다.

$$B_T = \text{Min}(B_T, (1-\gamma) \cdot E_i^{max}) \quad (28)$$

여기서 γ 는 부호기 버퍼 상태가 B_{max}^e 에 지나치게 근접하는 것을 막기 위한 상수로 본 논문에서는 0.2를 적용하였다.

2. 목표 왜곡 결정

이 과정은 현재 부호화할 모든 영상들을 전체 목표 비트량 B_T 로 결합 부호화할 때 모든 영상들이 가장 유사한 화질을 유지하기 위한 왜곡을 추정하는 과정으로, III장에서 제안된 비트율-왜곡 추정 결과를 이용한다.

그림 2에 나타난 바와 같이 현재 부호화할 각 프로그램 영상에 대한 추정된 비트율-왜곡 곡선들을 비트축 (수직 축)에 대하여 더함으로써 전체 비트율-왜곡 곡선 $R_T(D)$ 를 구한다. 여기서 각 프로그램 영상에 대한 비트율-왜곡 곡선은 계단 형태를 갖는다. 특히 본 논문에서와 같이 적용 가능한 양자화 파라미터가 31 종류이고 한 영상에 대하여 동일한 양자화 파라미터

를 적용하여 부호화할 경우는 각 프로그램 영상에 대한 비트율-왜곡 곡선이 31 단계의 왜곡 값을 갖을 수 있다. 또한 N 개의 프로그램이 결합 부호화되는 경우 전체 비트율-왜곡 곡선은 최대 $N \times 31$ 단계의 왜곡 값을 갖을 수 있다. 이와 같이 구성된 전체 비트율-왜곡 곡선 $R_T(D)$ 로부터 구해지는 목표 왜곡 D_T 는

$$D_T = \text{Min}\{D \mid R_T(D) \leq B_T\} \quad (29)$$

이다. 즉 목표 왜곡 D_T 는 그림 2에서 전체 비트율-왜곡 곡선 $R_T(D)$ 와 전체 목표 비트량 B_T 가 교차하는 점에 해당된다.

3. 양자화 파라미터 선택 및 부호화

이 과정은 각 프로그램의 영상에 대하여 목표 왜곡 D_T 에 근접한 왜곡을 발생시키는 양자화 파라미터를 선택하여 부호화하는 것으로 각 프로그램의 영상에 대한 왜곡 추정 결과로부터 다음 식을 만족하는 양자화 파라미터 QP 를 구하여 부호화 한다.

$$QP = \text{Min}_{1 \leq QP \leq 31} \{QP \mid D_T \leq D_k(QP)\} \quad (30)$$

여기서 $D_k(QP)$ 는 현재 부호화될 k 번째 프로그램 영상에 대하여 양자화 파라미터 QP 를 적용했을 때 왜곡의 추정 값이다. 따라서 선택되는 QP 는 목표 왜곡보다 큰 추정된 왜곡 값을 발생시키는 QP 중 가장 작은 값이다.

V. 실험

본 장에서는 제안된 결합 부호화의 성능을 평가하기 위하여 6종류의 비디오 프로그램들을 적용하여 실험을 수행하고, MPEG-2 TM5 방식에 의한 독립 부호화 결과와 비교하였다. 적용된 비디오 프로그램들은 SIF 형식(352×240 spatial resolution, 30 frames/sec, 4:2:0 format)의 표준 MPEG 비디오 시퀀스로써 결합 부호화되는 비디오 프로그램의 수를 2개에서 6개까지 변화 시키면서 성능을 평가했다. 결합되는 프로그램의 순서는 *mobile*, *flower garden*, *football*, *table tennis*, *popple* 그리고 *suzie* 시퀀스로 영상이 복잡한 순서와 같다. 한편 GOP 구조는 IBBPBBPBBPBB형식의

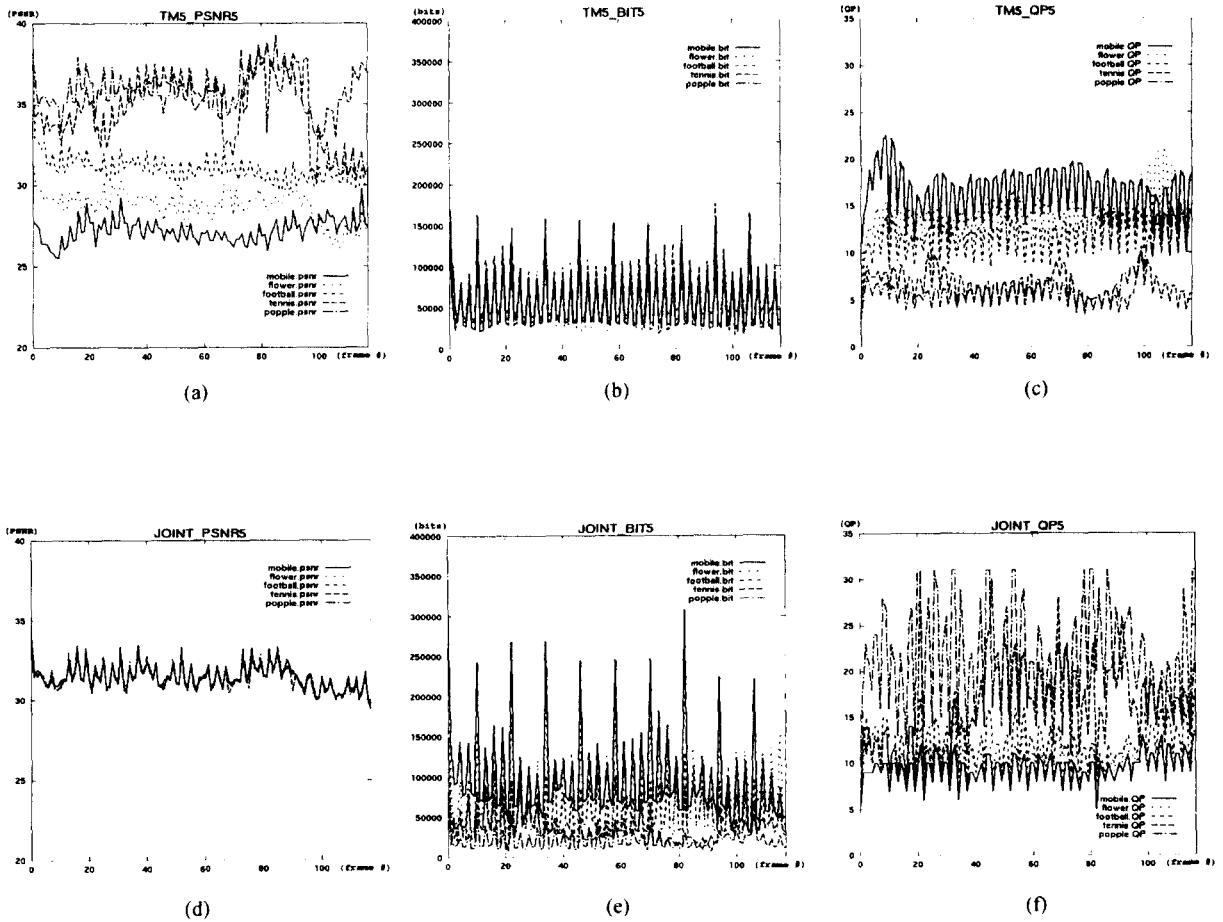


그림 10. 독립 부호화와 제안된 결합 부호화의 성능 비교 ($5 \times 2.25\text{Mbps}$)

독립 부호화: (a)PSNR, (b)발생 비트량, (c)평균 QP

제안 방식: (d)PSNR, (e)발생 비트량, (f)평균 QP

Fig. 10 Performance comparison between the independent coding and the proposed joint coding ($5 \times 2.25\text{Mbps}$):

Independent coding: (a) PSNR, (b) generated bits, (c) average QP

Proposed joint coding: (d) PSNR, (e) generated bits, (f) average QP

고정된 구조를 사용하였고, 최악의 경우를 고려하기 위하여 I 프레임들이 서로 다른 프로그램으로부터 동시에 발생하는 것으로 가정 했다.

결합 부호화에 적용한 채널 전송률 R_{CBR} 은 $N \times 2.25\text{Mbps}$ 와 $N \times 1.675\text{Mbps}$ 인데, 여기서 N 은 결합되는 프로그램의 수이다. 그리고 방송 서비스의 경우 전체

지연을 보통 300ms 설정하므로 버퍼의 크기는 $B_{max}^e = R_{CBR} \times 300\text{ms}$ 로 정하였다. 한편 독립 부호화에 적용한 채널 전송률은 각 프로그램에 대하여 2.25Mbps 와 1.675Mbps 를 적용하였다.

그림 10은 5개의 프로그램들을 $5 \times 2.25\text{Mbps}$ 의 채널 전송률에 대하여 독립 부호화와 결합 부호화한 결

과를 보여준다. (a)와 (d)는 PSNR관점에서의 성능을 나타낸 것으로 제안된 결합 부호화는 부호화된 프로그램들간의 화질을 매우 일정하게 유지함을 알 수 있다. 이는 (e)에 나타난 바와 같이 제안된 방식은 상대적으로 단순한 영상에 할당된 비트량의 일부를 복잡한 영상에 할당함으로써 모든 프로그램들간에 균일한 화질을 유지하게 되는 것이다. (c)와 (f)는 각 프로그램의 영상에 적용된 평균 양자화 파라미터를 나타낸 것으로 독립 부호화는 단순한 영상이 복잡한 영상에 비하여 작은 양자화 파라미터 값을 사용하여 부호화됨으로써 프로그램들간에 큰 화질 차이가 발생하지만, 제안된 결합 부호화 방식에서는 상대적으로 단순한 영상에 큰 양자화 파라미터 값을 적용함으로써 프로그램들간에 화질을 균일하게 유지한다.

표 1과 표 2는 채널 전송률이 $N \times 2.25\text{Mbps}$ 와 $N \times 1.675\text{Mbps}$ 일 때 결합되는 프로그램의 수에 따른 독립

부호화와 결합 부호화의 성능을 비교한 것이다. 제안된 결합 부호화 방식으로 부호화된 비디오 프로그램들에 대한 최고 화질(MAX PSNR)과 최소 화질(MIN PSNR)은 결합되는 프로그램의 수에 무관하게 매우 유사한 값을 갖는다. 따라서 적은 수의 비디오 프로그램들을 결합 부호화하더라도 모든 시청자들에게 공정한 서비스를 제공할 수 있다. 특히 제안된 결합 부호화 방식은 독립 부호화 방식과 비교하여 최소 화질을 크게 향상시킨다. 따라서 제안된 결합 부호화 방식을 적용하면 독립 부호화 방식에 비하여 동일한 채널 대역에 보다 많은 비디오 프로그램을 전송할 수 있다. 또한 평균 화질(avg. PSNR 또는 avg. MSE) 관점에서도 제안된 결합 부호화 방식이 독립 부호화 방식에 비하여 보다 향상된 화질을 제공할 수 있다. 또한 왜곡의 표준편차 (avg. σ_{MSE})들이 적은 값을 갖으므로 일관된 화질이 유지됨을 알 수 있다.

표 1. 독립 부호화와 결합 부호화의 성능 비교 ($N=2\sim6$ and $R_T = N \times 2.25\text{Mbps}$).

Table 1. Performance comparison between the independent coding and the proposed joint coding ($N=2\sim6$ and $R_T = N \times 2.25\text{Mbps}$).

	independent coding ($N \times 2.25$)							joint coding ($N \times 2.25$)						
	Avg. PSNR	Avg. MSE	Avg. σ_{MSE}	Avg. QP	MIN PSNR	MAX PSNR	MAX σ_{MSE}	Avg. PSNR	Avg. MSE	Avg. σ_{MSE}	Avg. QP	MIN PSNR	MAX PSNR	MAX σ_{MSE}
Join 2	29.43	74.09	16.56	11.61	28.77	30.21	16.58	30.68	55.55	12.67	11.81	30.68	30.69	12.75
Join 3	30.21	61.90	14.02	10.74	28.77	32.38	16.58	31.49	46.12	9.03	10.86	31.49	31.50	9.12
Join 4	31.14	50.02	12.60	9.30	28.77	36.56	16.58	32.30	38.29	7.60	10.46	32.24	32.35	7.87
Join 5	31.78	43.11	11.88	8.49	28.77	36.56	16.58	33.10	31.88	6.94	10.27	32.90	33.18	7.26
Join 6	32.51	36.50	10.85	7.52	28.77	42.79	16.58	34.36	23.85	5.21	11.54	34.16	35.26	5.70

표 2. 독립 부호화와 결합 부호화의 성능 비교 ($N=2\sim6$ and $R_T = N \times 1.675\text{Mbps}$).

Table 2. Performance comparison between the independent coding and the proposed joint coding ($N=2\sim6$ and $R_T = N \times 1.675\text{Mbps}$).

	independent coding ($N \times 1.675$)							joint coding ($N \times 1.675$)						
	Avg. PSNR	Avg. MSE	Avg. σ_{MSE}	Avg. QP	MIN PSNR	MAX PSNR	MAX σ_{MSE}	Avg. PSNR	Avg. MSE	Avg. σ_{MSE}	Avg. QP	MIN PSNR	MAX PSNR	MAX σ_{MSE}
Join 2	27.78	108.49	20.45	15.31	27.15	28.50	20.48	29.00	81.80	13.90	15.90	28.99	29.02	14.07
Join 3	28.57	90.45	17.31	14.19	27.15	30.78	20.48	29.79	68.17	11.20	14.46	29.77	29.82	11.38
Join 4	29.49	73.12	15.73	12.28	27.15	34.90	20.48	30.57	57.05	9.17	14.11	30.49	30.62	9.37
Join 5	30.14	62.95	15.03	11.19	27.15	34.90	20.48	31.32	47.98	8.72	14.07	31.23	31.37	9.13
Join 6	30.87	53.21	13.73	9.93	27.15	41.59	20.48	32.47	36.78	6.75	14.82	32.21	33.77	7.26

VI. 결 론

본 논문에서는 MPEG-2 비디오에 대한 비트율-왜곡을 추정방식을 제안하고 서로 다른 특성을 갖는 여러 영상에 제안된 추정방식을 적용한 결과를 보였다. 또한 이 추정 결과에 근거하여 다수의 비디오 프로그램들을 하나의 CBR 채널을 통하여 전송하는 다중 비디오 프로그램 전송 환경에 적용되는 결합 비디오 부호화 방식을 제안하고 실험을 통하여 성능을 평가했다.

제안된 비트율-왜곡 추정방식은 MPEG-2로 부호화된 한 영상에서 발생하는 비트량과 왜곡을 비교적 정확하게 추정하며, 특히 추정을 위한 주된 연산이 양자화에 입력되는 신호의 히스토그램 또는 가중 히스토그램을 구하는 것이므로 계산량이 매우 적다. 따라서 추정의 정확성과 추정을 위한 계산량 측면에서 제안된 방식은 실시간 MPEG-2 비디오 응용에 적용하기에 매우 적합한 방식이다. 또한 제안된 추정과정은 MPEG-2뿐만 아니라 H.261, H.263 또는 MPEG-1 등과 같은 다른 표준 부호화 방식에도 적용이 가능하다.

한편 제안된 결합 부호화 방식은 결합 부호화되는 모든 프로그램들의 전체 전송률은 항상 일정하게 유지하면서 각 프로그램에 대해서는 VBR 압축을 허용한다. 따라서 독립적으로 각 프로그램들을 CBR 압축하는 독립 부호화 방식에 비하여 평균 화질이 향상되고, 결합 부호화된 각 프로그램은 일관된 화질을 유지한다. 또한 각 비디오 프로그램의 한 영상에 대하여 동일한 양자화 파라미터를 적용하므로 영상 내에서도 균일한 화질이 유지된다. 특히 결합 부호화된 모든 프로그램들간에 화질이 매우 유사하므로 모든 시청자들에게 공정한 비디오 서비스를 제공할 수 있고, 결합 부호화된 프로그램들에 대한 최소 화질이 독립 부호화된 비디오 프로그램들에 비하여 크게 향상되므로 동일한 채널 대역폭에 보다 많은 수의 비디오 프로그램들을 전송할 수 있다.

실제적인 측면에서도 제안된 결합 부호화 방식은 비트율-왜곡 추정 과정을 포함하여 추가되는 연산량이 적으므로 위성방송, CATV, 디지털 지상방송과 같은 실시간 비디오 서비스에 적용이 용이하다.

참 고 문 헌

1. ISO-IEC/JTC1/SC29/WG11, "Generic coding of moving pictures and associated audio information: Video," ISO-IEC 13818-2, Nov. 1994.
2. Document ISO-IEC/JTC/SC29/WG11, "Test Model 5," Draft, Apr. 1993.
3. W. Veriest and L. Pinnoo, "A variable bit rate video codec for asynchronous transfer mode networks," *IEEE J. Select. Areas Commun.* Vol. 7, no. 5, June 1989.
4. R. Coellco and S. Tohme, "Video traffic smoothing and ATM multiplexer performance," *IEEE GLOBECOM'91*, pp. 239-243.
5. D. Reininger, D. Raychaudhuri, B. Melamed, B. Sengupta, and J. Hill, "Statistical multiplexing of VBR MPEG compressed video on ATM networks," *IEEE GLOBECOM'93*, pp. 919-926.
6. T. Tanaka, S. Okubo, H. Hashimoto and H. Yasuda, "A study on comparison between VBR and CBR video service in ATM environment," *IEEE ICC'92*, pp. 551-555.
7. T. Koga, Y. Iijima, Iinuma, and T. Ishiguro, "Statistical performance analysis of an interframe encoder for broadcast television signal," *IEEE Trans. Commun.*, vol. COM-29, no. 12, pp. 1868-1876, Dec. 1981.
8. B. G. Haskell and A. R. Reibman, "Multiplexing of variable rate encoded streams," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 4, no. 4, pp. 417-424, Aug. 1994.
9. A. Guha and D. J. Reininger, "Multichannel joint rate control of VBR MPEG encoded video for DBS applications," *IEEE Trans. Consumer Electron.*, vol. 40, no. 3, pp. 616-623, Aug. 1994.
10. S. Sakazawa, Y. Takishima, M. Wada, and Y. Hatori, "Coding control scheme for a multi-encoder system," *6th International Workshop on Packet Video*, Brisban, Australia, pp. 83-88, March 1996.
11. L. Wang, A. Vincent and P. Corriveau, "Multi-program video coding with joint rate control," *IEEE GLOBECOM'96*, pp. 1516-1520.

12. L. Wang and A. Vincent, "Bit allocation for joint coding of multiple video programs," *SPIE VCIP '97*, pp. 149-158.
13. G. Keesman and D. Elias, "Analysis of joint bit-rate control in multi-program image coding," *SPIE VCIP'94*, pp. 1906-1917.
14. T. Berger, *Rate Distortion Theory*. Englewood Cliffs, NJ: Prentice Hall, 1971.
15. H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 676-683, Sept. 1968.
16. N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice Hall, 1984.
17. CCITT, Working Party XV/1, "Draft of recommendation H.261: Video codec for audiovisual services at $P \times 64$ kbit/s," July 1990.
18. ISO-IEC/JTC1/SC29/WG11, "Coding of moving pictures and associated audio," CD 11172-2, Nov. 1991.
19. ITU-T, Study Group 15, "Draft of recommendation H.263: Video coding for low bitrate communication," 2 May 1996. T.



홍 성 훈(Sung-Hoon Hong) 정회원

1965년 11월 29일생

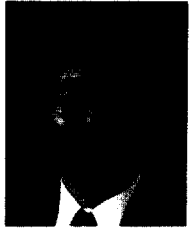
1988년: 영남대학교 전자공학과 졸업(공학사)

1991년: 한국과학기술원 전기 및 전자공학과 졸업(공학석사)

1995년 3월~현재: 한국과학기술원 전기 및 전자공학과 박사과정

1991년 2월~현재: LG 전자 멀티미디어 연구소
 ※주관심분야: 비디오 부호화, 영상통신시스템, 멀티미디어 통신 등

e-mail: hsh@sdvvisio.kaist.ac.kr



김 성 대(Seong-Dae Kim) 정회원

1953년 12월 26일생

1997년: 서울대학교 전자공학과 졸업(공학사)

1979년: 한국과학기술원 전기 및 전자공학과 졸업(공학석사)

1983년: 프랑스 ENSEEIHT 졸업(공학박사)

1984년~현재: 한국과학기술원 전기 및 전자공학과 교수
 ※주관심분야: 영상처리, 영상통신, 컴퓨터 비전, VLSI 구현 등