

論文 2002-39SP-6-3

## 비디오의 에러내성 전송을 위한 DCT 계수의 새로운 분할 기법

## (A New Data Partitioning of DCT Coefficients for Error-resilient Transmission of Video)

盧奎燦\*, 金在均\*\*

(Kyu-Chan Roh and Jae-Kyoon Kim)

## 요약

이동 보상 예측과 DCT 기법을 이용하는 비디오 부호화에서 전형적인 데이터 분할 (data partitioning) 기법은 움직임 정보와 매크로블록 헤더를 texture 정보와 분리시키는 것이다. 이 방식은 에러가 있는 환경에서 비디오의 전송에 효율적이다. 그러나, 화면내 부호화 프레임 (Intra frame)의 경우에는 따로 분리시킬 움직임 정보가 없기 때문에 DCT 계수의 손실은 치명적이다. 그리고 화면간 부호화 프레임 (Inter frame)의 경우에도 DCT 계수 부분에 에러가 발생하면, 패킷 내의 모든 DCT 계수를 버리게 되므로 이로 인한 화질 저하가 발생한다. 본 논문에서는 비디오의 에러내성(error-resilient) 전송을 위한 DCT 계수의 효율적인 분할 기법과 부호화 기법을 제안한다. 양자화된 DCT 계수는 짝수-근사(even-value approximation) 부분과 그 나머지 부분으로 분리된다. 모의실험을 통해 제안 방식이 기존의 기법에 비해 우선순위 (선순위) 분할 데이터 부분에서 더 좋은 화질을 제공하고, 에러환경에서도 에러에 더 강한 성능을 나타냄을 보였다.

## Abstract

In the typical data partitioning for error-resilient video coding, motion and macroblock header information is separated from the texture information. It can be an effective tool for the transmission of video over the error prone environment. For Intra-coded frames, however, the loss of DCT (discrete cosine transform) coefficients is fatal because there is no other information to reconstruct the corrupted macroblocks by errors. For Inter-coded frames, when error occurs in DCT coefficients, the picture quality is degraded because all DCT coefficients are discarded in those packets. In this paper, we propose an efficient data partitioning and coding method for DCT-based error-resilient video. The quantized DCT coefficients are partitioned into the even-value approximation and the remainder parts. It is shown that the proposed algorithm provides a better quality of the high priority part than the conventional methods.

**Keywords** : data partitioning, error-resilience, two-layered coding, MPEG-4

## I. 서론

\* 學生會員, 韓國科學技術院 電子電算學科

(Department of EECS, KAIST)

\*\* 正會員, 韓國科學技術院 電子電算學科

(Department of EECS, KAIST)

接受日: 2002年7월15일, 수정완료일: 2002年10월22일

데이터 분할은 비디오 데이터를 둘 또는 그 이상의 부분으로 나누는 기법이다. MPEG-4 또는 H.263++ 등에서 사용되는 전형적인 데이터 분할 기법은 움직임 정보와 매크로블록 헤더를 DCT 계수와 분리시키는 것이다.<sup>[1,2]</sup> 비디오 전송시에 DCT 계수의 일부가 손실된

다면, 이 정보는 완전히 버려지고 움직임 정보만으로 재생 화면을 구성하게 된다. 그러나 화면내 부호화 프레임(Intra-coded frame)의 경우에는 움직임 정보가 없기 때문에 DCT 계수의 손실은 재생된 화면의 화질에 치명적인 영향을 끼친다. 지금까지 에러내성을 위한 DCT 계수 분할 방식은 주로 스펙트럼 분리 (spectral separation) 기법이란 이름으로 제안되었다.<sup>[3-6]</sup> 스펙트럼 분리 기법은 DCT 계수 중에서 시각적으로 중요한 저주파 성분을 우선순위 (선순위) 부분 (high priority part)으로, 나머지 고주파 부분을 낮은순위 (후순위) 부분 (low priority part)으로 구별하여 나눈다.<sup>[3]</sup> 그러나, 스펙트럼 분리 기법이 DCT 계수에 대한 에러내성을 제공하지만, 선순위 부분의 화질이 같은 비트율의 단일 스트림에 비해 현저히 떨어진다는 단점이 있다.<sup>[7]</sup> 따라서 후순위 부분에 에러가 발생해서 높은순위 부분만으로 화면을 구성한다고 해도, 그리 큰 이득을 볼 수 없게 된다. 이러한 이유로 MPEG-4, 또는 H263++ 등에서 사용되는 데이터 분할 기법에서는 매크로블록 헤더와 함께 단지 DC 계수만을 선순위 부분에 두는 방식을 사용하고 있다.<sup>[1,2]</sup> 그러므로 화면내 부호화 프레임에서 DCT 계수에 에러가 발생하여 DC 계수만으로 복호화가 이루어질 때, 재생된 화면의 화질은 크게 떨어지게 되고, 이 에러는 이후의 화면간 부호화 프레임 (Inter frame)들에게 계속 전파된다.

본 논문에서는 이와 같은 문제를 극복하기 위한 새로운 data partitioning 기법을 제안한다. 먼저, II장에서 DCT 계수를 짝수-근사(even-value approximation) 부분과 나머지(odd remainder) 부분으로 나누는 방법을 제안하고, 아울러 나머지 부분의 효율적인 부호화 방법도 제안한다. III장에서는 에러가 없는 환경과 랜덤 에러 환경에서의 모의 실험을 통해, 제안 방식과 기존의 데이터 분할 방식의 성능이 비교되며, IV장에서 결론을 맺는다.

## II. 제안된 데이터 분할 방법

<그림 1>은 데이터 분할을 포함한 제안된 부호기(encoder)의 구조도를 나타낸다. <그림 1>에서 알 수 있듯이, 기존의 스펙트럼 분리 방식과 가장 큰 차이는 제안방식의 구조가 재양자화의 방식을 따르고 있다는 것이다. 본 장에서는 먼저 재양자화 방식과 스펙트럼 분리 방식과의 차이점을 분석해 보고, 짝수-근사화 기법에 대

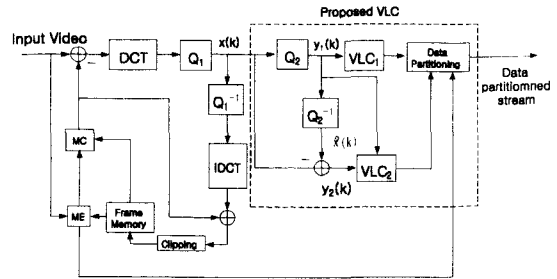


그림 1. 제안된 부호기의 구조도

Fig. 1. Block diagram of the proposed encoder.

해 설명한다.

### 1. 스펙트럼 분리 방식과 재양자화 방식의 비교

양자화된 DCT 계수,  $x(k)$ 에 대해 스펙트럼 분리 방식에 의해 나누어진 두 데이터를 각각,  $Z_1, Z_2$ 라 하면, 이들은 아래 식들과 같이 표현된다.

$$Z_1 = \{x(k) | k \in S\} \quad (1)$$

$$Z_2 = \{x(k) | k \in S^c\} \quad (2)$$

여기서  $S$ 는 전체 DCT 계수 집합에서 저주파 성분의 계수 집합을 의미하는 주파수 성분의 인덱스 (index) 집합이다.

제안된 재양자화 방식을 따르는 경우, DCT 계수는 <그림 1>에서와 같이 2단계의 양자화기, 즉 재양자화기  $Q_2$ 에 의해 만들어진 데이터인  $y_1(k)$ 와, 이에 의한 양자화 오차 신호인  $y_2(k)$ 로 나누어 진다. 여기서 전자를 선순위 부분 (primary part), 후자를 후순위 부분 (secondary part)이라 정의한다. 재양자화기의 양자화 간격 (step size)을  $N$ 이라 하면,  $y_1(k)$ 는 다음과 같이 표현된다.

$$y_1(k) = \text{sgn}(x(k)) \left\lfloor \frac{|x(k)|}{N} \right\rfloor, \quad 1 \leq k \leq 64 \quad (3)$$

$$\text{sgn}(x) = \begin{cases} 1 & , x \geq 0 \\ -1 & , x < 0 \end{cases} \quad (4)$$

여기서  $\lfloor x \rfloor$ 는  $x$ 를 넘지 않는 최대 정수를 의미하고, 재양자화로부터 복원된 근사신호,  $\hat{x}(k)$ 와 후순위 신호는 다음과 같다. 그리고  $x$ 는 DCT 계수의 주파수 인덱스이다.

$$\hat{x}(k) = N y_1(k), \quad (5)$$

$$y_2(k) = x(k) - \hat{x}(k) \\ = x(k) - Ny_1(k), \quad 1 \leq k \leq 64. \quad (6)$$

스펙트럼 분리 방식의 경우, DCT 계수의 주파수 인덱스에 의해 데이터를 둘로 나누는 것이다, 반면에 재양자화 방식에서는 고주파 성분까지 다 포함하여 재양자화에 의한 하나의 근사적 비트열을 만든 다음, 이 신호와 원래 신호와의 차이를 가지고 또다른 하나의 비트열을 구성하는 것이다.

이와 같은 두가지 방식의 데이터 분할에서 선순위 부분의 왜곡(distortion)은 다음과 같이 정의할 수 있다.

$$D_{spec} = \sum_{k \in U} (x(k) - z_1(k))^2 = \sum_{k \in S^c} (x(k))^2 \quad (7)$$

$$D_{prop} = \sum_{k \in U} (x(k) - \hat{x}(k))^2 \quad (8)$$

여기서  $U$ 는 DCT 영역에서의 전체 화소 집합이다.

식 (7)과 (8)에서 알 수 있듯이 선순위 부분의 화질은, 스펙트럼 분리의 경우에는 선순위 부분에 포함될 저주파 성분의 개수에 따라 결정되고, 제안 방식의 경우에는 재양자화기의 양자화 간격에 의해 결정된다. 즉, 스펙트럼 분리의 경우, 왜곡의 크기는 후순위 영역의 크기와  $x(k)$ 의 크기 제곱에 비례하고, 재양자화의 경우에는  $x(k)$ 와 재양자화에 의한 오차와의 크기 제곱에 비례한다.

<그림 2>는 300 프레임의 "Foreman" 영상을 프레임울 10Hz에서 양자화 계수가 15이고, MPEG-4 Simple Profile 부호기를 사용하여 부호화 할 때, 각각의 AC계수들을 부호화 하는데 필요한 화면당 평균 비트량 증가분을 각 주파수 인덱스별로 나타낸 그래프이다. 즉,  $k$ 번째 DCT 계수에 대한 비트 증가량은 처음부터  $k$ 번째 DCT 계수까지 부호화 할 때 필요한 비트량과  $(k-1)$ 번째 계수까지 부호화 하는데 필요한 비트량의 차이로 구해진다. 이때 GOP 크기는 15이고, B-화면은 없다. 그리고, <그림 3>은 위와 같은 방법으로 재생된 화면의 왜곡에 대한 변화량(증가량)을 각 DCT 계수별로 PSNR로 나타내었다. 각 계수별 왜곡을 구하는 방식은 앞에서 각 계수별 비트량을 구하는 것과 같은 방식으로 구하였다. 각 주파수별 PSNR을 모두 더하면 화면 전체의 PSNR이 구해진다.

<그림 2>에서 알 수 있듯이 저주파 부분을 부호화

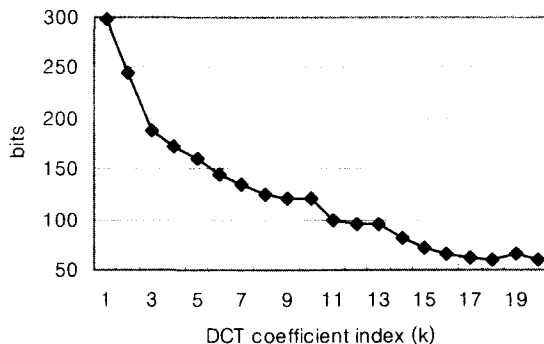


그림 2. DCT 계수의 index에 따른 화면당 평균 비트 증가량

Fig. 2. Average amount of coded bits increment per frame for each DCT coefficient index.

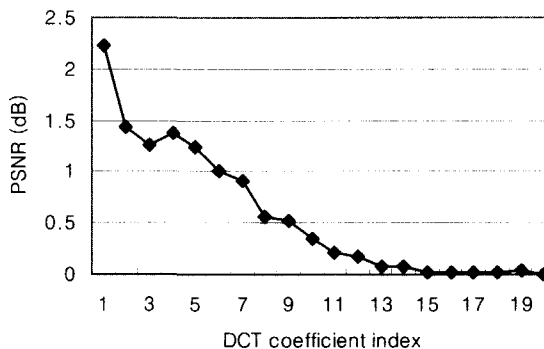


그림 3. DCT 계수의 인덱스에 따른 PSNR 증가량 (dB)

Fig. 3. PSNR increment (dB) for each DCT coefficient index.

하는데 고주파 부분보다 더 많은 비트를 필요로 한다. 그러므로 하나의 비트열을 스펙트럼 분리 방식에 의해 비슷한 비트량을 갖는 두 부분으로 나눌 경우, 선순위 부분에서 저주파 부분이 상당량 손실되고, <그림 3>에서 알 수 있듯이 이로 인한 심각한 화질 저하를 예측할 수 있다.

### 2. 짝수-근사화에 의한 데이터 분할

식 (3)-(6)에서 재양자화기의 양자화 간격인  $N$ 의 크기를 2로 둘 경우, 선순위 부분의 재생값,  $\hat{x}(k)$ 는 원래의 양자화 값에 대한 짝수-근사 계수(even-approximated coefficients)값이 된다. 이 경우, 후순위 부분,  $y_2(k)$ 는 크기가 0과 1로만 표현되는 나머지 값이 되므로, 2진수에 대한 새로운 1차원 연속길이 부호(run-length code)로 효율적으로 부호화 할 수 있다. 더구나  $y_2(k)$ 의 부호는  $y_1(k)$ 와 같으므로, 단지  $y_1(k)=0$

이고  $y_2(k) \neq 0$ 인 경우에만 1bit의 추가적인 부호정보가 삽입된다. 즉 후순위 부분에서 부호정보의 삽입 여부는 선순위 부분의 계수의 부호에 따라 결정된다. 복호기에서는 후순위 부분의 계수에 부호정보가 없을 때, 같은 위치에 있는 선순위 부분의 계수로부터 부호를 알 수가 있다.

<그림 4>는 제안된 기법을 이용해서 만든 비트열의 구조도이다. 두개의 데이터 영역으로 분할된 MPEG-4 데이터 분할 기법과 비교해 볼 때, DCT 계수들이 'Texture marker'에 의해 새로운 두개의 영역으로 분할됨으로써, 전체적으로 세개의 영역으로 구성된다. 화면내 부호화 프레임의 경우, partition 0에 매크로블록 헤더정보와 함께 원래의 DC 계수들이 아닌 짝수 근사화된 DC 계수들이 위치하게 되어 부호어 (codeword)의 길이가 줄어들게 된다. 따라서, 전송에러가 있는 환경에서 정보의 손실시에 가장 치명적인 partition 0의 손실확률을 더 줄일 수가 있다.

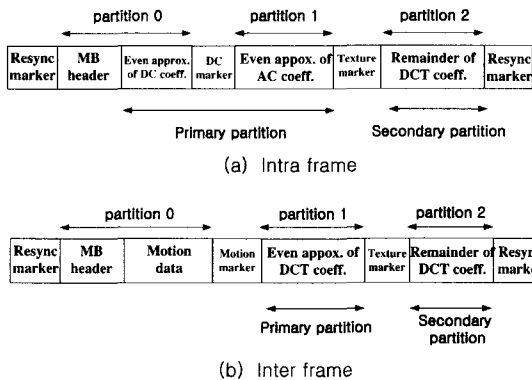


그림 4. 제안된 데이터 분할 방식에 의한 비트스트림 구조

Fig. 4. Bitstream organization of the proposed data partitioning.

### 3. 분할된 데이터의 가변길이 부호화

앞에서 설명한 바와 같이  $\{y_2(k)\}$ 의 분포특성이  $\{x(k)\}$ 나  $\{y_1(k)\}$ 와는 차이가 크기 때문에, 후순위 부분은 부호정보의 감소와 함께, 새로운 적절한 부호테이블로 부호화 할 수 있다. 후순위 부분에 대한 부호테이블을 구하기 위해, 후순위 부분에서 계수들의 0의 연속 길이를 의미하는 확률변수를  $R$ 이라 하고, 연속길이에 대한 확률분포함수를  $P_R(r)$ 이라 하면, 이는 식 (9)와 같이 근사적으로 기하분포의 형태로 모델링 할 수 있다.

$$P_R(r) = p_o^r (1 - p_o) \quad (9)$$

여기서  $p_o$ 는  $x(k)$ 가 짝수일 확률, 즉  $y_2(k) = 0$ 일 확률이다.

식 (9)에서  $p_o$ 의 값을 알 수 있다면, 연속길이에 대한 확률분포함수를 얻을 수 있다. 그러나,  $p_o$ 는 영상에 따라서, 그리고 양자화 계수의 크기에 따라서 그 값이 달라지기 때문에, 실제 영상에 대한 여러가지 모의실험을 통해 그 평균치를 얻게 된다. 모의실험 결과 얻어진  $p_o$ 의 값을 이용하여 구해진 확률분포 함수를 통하여 0의 연속길이에 대한 각각의 확률값을 구할 수 있다. 이 확률값을 이용하여 후순위 부분의 부호화를 위한 1차원 허프만 부호테이블을 얻게 된다.

선순위 부분의 경우, 양자화 계수를 2배 한 것과 거의 비슷한 효과를 가지므로 기존의 부호테이블을 그대로 이용할 수 있다.

## III. 모의실험 결과

본 장에서는 스펙트럼 분리 기법과 MPEG-4 데이터 분리 기법이 제안방식과 비교, 검증된다. 성능의 비교를 위해, 300 frame의 News, Foreman, 그리고 Stefan 영상이 사용되었고, 공간 해상도는  $176 \times 144$  화소이다. 식 (9)에서 정의된  $p_o$ 의 값을 구하기 위하여 News, Foreman, Coastguard, 그리고 Stefan 실험영상들에 대해서 비트율이 32 kbps부터 320 kbps까지 여러 번의 모의실험을 하였다. 그 결과로 얻어진  $p_o$ 의 값은 0.65로 설정되었고, 그 외 모의실험 환경은 <표 1>에 나타나 있다.

표 1. 모의실험을 위한 실험환경

Table 1. Test conditions for simulations.

Frame rate	10 Hz
Period of I	15
Period of P	1
Quantization	H.263
MV range	16
Rate control	TM5
Approximate size of resynchronization packet	500 bits

스펙트럼 분리방식은 <그림 5>와 같이 DCT 계수를 저주파 부분과 고주파 부분으로 분리한 것이다. 여기서 선순위 부분, 즉 저주파 부분에 포함될 계수의 개수는 같은 양자화 계수를 사용했을 때, 제안방식에 의한 선순위 부분의 비트량과 동일하게 되도록 선택된다.

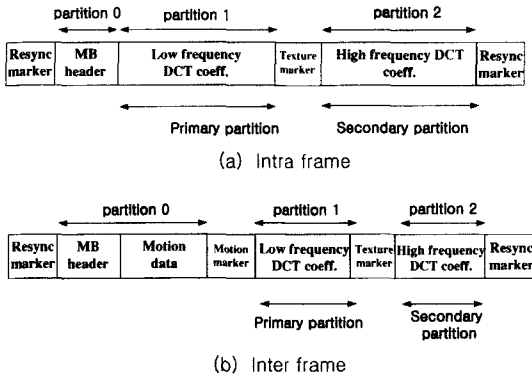
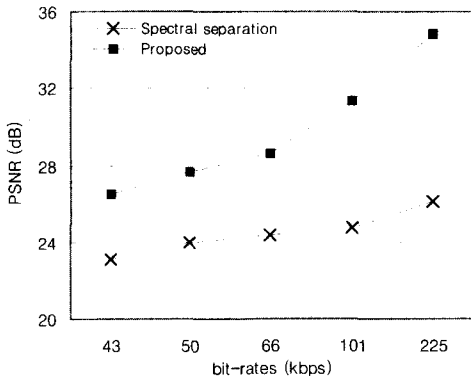
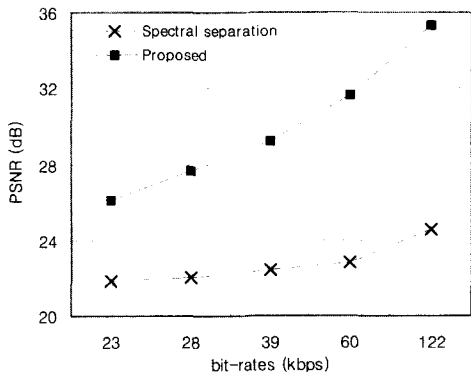


그림 5. 스펙트럼 분할 방식에 의한 비트열 구조  
Fig. 5. Bitstream organization of the spectral separation.



(a) Foreman



(b) News

그림 6. 선순위 부분에서의 PSNR 비교  
Fig. 6. PSNR comparison for the primary partition.

<그림 6>은 전통적 스펙트럼 분리 기법과 제안방식의 선순위 부분, 즉 <그림 4>와 <그림 5>에서 데이터 분할 0, 1 (partition 0, 1) 부분에서의 각 비트율에 대한 평균 PSNR을 나타낸다. 그림에서 가로축은 각각 고정된 양자화 계수를 사용하여 부호화 하였을 때 얻어진 비트율을 표시한 것이다. 데이터 분할 기법 이외의 부호화 방식은 MPEG-4 Simple Profile 부호기를 사용하였다. <그림 6>에서 보여진 바와 같이, 선순위 부분에서의 화질은 제안방식이 동일한 비트율에서 스펙트럼 분리방식에 비해 월등히 우수함을 알 수 있다. 비트율이 증가할수록 제안방식의 성능은 더욱더 좋아진다. 이는 높은 비트율일수록 DCT 계수의 고주파 성분 많이 존재하는데, 이들 중 많은 부분이 스펙트럼 분리에 의해 버려지기 때문이다. 선순위 부분에서 제안방식의 PSNR 이득은 3.3 dB~10.8 dB 정도이다.

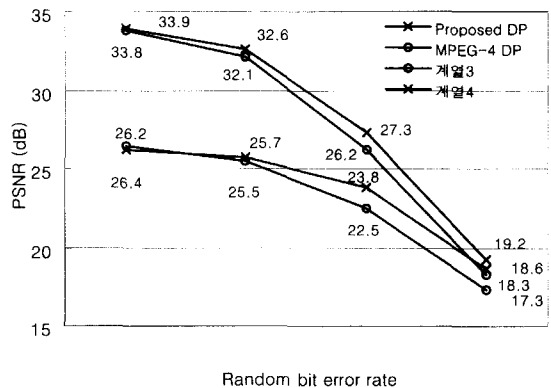


그림 7. 랜덤비트에러환경에서 재생영상의 PSNR 비교  
Fig. 7. Comparison of average PSNR in random-bit-error prone environment.

에러환경에서 에러에 대한 강인성을 비교하기 위해, 제안방식과 기존의 MPEG-4 데이터 분할 방식의 에러 내성 성능을 비교한다. <그림 7>은 128 kbps의 비트율로 부호화된 비트열을 여러 랜덤비트에러 환경을 통해 전송한 뒤 복호화 했을 때의 평균 PSNR을 나타낸다. 에러로 인해 손실된 매크로 블록은 이전의 화면에서 대체하는 방식을 사용하였다. <그림 7>에서와 같이 에러가 없는 환경에서는 제안기법의 추가적인 오버헤드로 인해서 같은 비트율에서 0.1 dB~0.2 dB 정도의 화질저하가 있다. 그러나, 랜덤에러 환경에서의 성능은 에러율에 따라 다르지만 MPEG-4에 비해 화질이 0.2 dB~1.3 dB 정도 향상되며, 에러율이 높을수록 더 좋은

성능을 나타낼 수 있다. 이는 텍스처 정보를 새로운 두 부분들로 나눔으로써, 각 부분에서 에러가 발생할 확률이 줄어들었기 때문이다.

#### IV. 결 론

DCT 계수의 데이터 분할에서, 스펙트럼 분리에 의한 선순위 부분의 현저한 화질저하 문제가 제안된 방식에 의해 개선되었다. 선순위 부분에서 기존의 방식과 비교한 PSNR 이득은 3.3 dB~10.8 dB이다. MPEG-4 데이터 분할 방식과 비교해 볼 때, 에러가 없는 환경에서는 새로운 부분을 구별하기 위한 오버헤드에 의한 비트량 증가가 있지만, 에러 환경에 대한 모의실험에서는 제안 방식이 MPEG-4 데이터 분할 기법에 비해 더 우수한 화질을 제공하였다. 제안방식의 PSNR 이득은 에러율에 따라 다르지만 0.2 dB~1.2 dB 정도이다. 이와 같이, 제안된 데이터 분할 기법은 에러 환경에서 비디오의 에러내성 전송을 위한 효율적인 부호화 기법으로 적용될 수 있다.

#### 참 고 문 헌

[1] ISO/IEC 14496-2 MPEG-4 International Standard, "Information technology - generic coding of audio-visual objects", October 1998.

- [2] ITU-T, "Recommendation H.263, Video coding for low bitrate communication", 2000.
- [3] S.L.Blake, S.A.Rajala, F.Gong and T.L.Mitchell, "Efficient techniques for two-layer coding of video sequences", in *Proc. of Int. Conf. on Image Processing*, pp. 253~257, Austin, TX, October 1994.
- [4] W. Luo and M. El Zarki, "Adaptive data partitioning for MPEG-2 video transmission over ATM based networks", in *Proc. of Int. Conf. on Image Processing*, pp. 17~20, Washington DC, USA, October 1995.
- [5] R.Talluri, I.Moccagatta, Y. Nag and G. Cheung, "Error concealment by data partitioning", *Signal processing: Image Communication*, Vol. 14, No. 6-8, pp. 505~518, May 1999.
- [6] R. Mathew and J.F. Arnold, "Efficient layered coding using data partitioning", *Signal processing: Image Communication*, Vol. 14, No. 9, pp. 761~782, July 1999.
- [7] L.R.Mendes, T.del Pietro and M.H.M.Costa, "Two-scale transmission of DCT-coded video over losst packet networks", *Proc. of ITS*, pp. 510~515, 1998.

#### 저 자 소 개

##### 盧奎燦(學生會員)

1995년 한국과학기술원 전기및전자공학과 학사. 1997년 한국과학기술원 전기및전자공학과 석사. 1997년 3월~현재, 한국과학기술원 전자전산학과 영상통신연구실에서 박사과정 재학중.

##### 金在均(正會員)

1962년 국립항공대학교 응용전자공학과 학사. 1967년 서울대학교 전자공학과 석사. 1971년 University of Southern California 전기공학과 (전자통신) 박사. 1972년 4월~1973년 3월 미국 NASA GSFC 연구원. 1990년~1993년 ISO/IEC JTC1-SC29(Multimedia Coding) 한국위원회 위원장. 1993년 1월~1994년 12월 한국통신학회 회장. 1993년 8월~1996년 7월 The Founder Chair, APCC International Steering Committee. 1973년 4월~현재 한국과학기술원 전자전산학과 교수 재직중.