

# EREC 기반 비트열을 위한 Source-Channel 결합 최적 양자화기 설계 및 이를 통한 안정적 영상 전송에 관한 연구

정회원 김 용 구\*, 송 진 규\*\*, 최 윤 식\*

## A Study on the Reliable Video Transmission Through Source/ Channel Combined Optimal Quantizer for EREC Based Bitstream

Yong-Goo Kim\*, Jin-Gyu Song\*\*, Yoonsik Choe\* *Regular Members*

### 요 약

오류를 수반하는 통신망을 통한 멀티미디어 데이터의 응용은 최근 그 수요가 급증하고 있다. 하지만 그 구현은 많은 문제점들을 야기하는데, 전송된 비디오 데이터에 발생한 오류를 처리하는 문제가 그 중 하나이다. 이는 압축된 비트열에 발생한 오류가 영상의 시-공간 방향으로 심각한 전파 현상을 수반하기 때문이다. 이러한 심각한 오류 전파를 완화하기 위해 본 논문에서는 EREC라 알려진 오류 제한 기법을 적용하고, 적용된 EREC의 오류 전파 특성을 분석하였다. 이를 통해, 압축 부호화된 하나의 기본 블록(매크로 블록)이 복호시 오류가 생길 확률을 추정하였으며, 추정된 확률의 근사를 통해 양 끝단(전송단과 수신단)에서의 비디오 화질 열화를 예측하였다. 추정 확률의 근사는 매 기본 블록에서 발생된 비트수에 대한 그 기본 블록이 복호시 오류가 생길 확률을 간단한 1차식을 통한 선형 회귀법으로 모델링 되었으며, 따라서 간단한 방법을 통해 양 끝단의 화질 열화를 효과적으로 예측할 수 있었다. 부호화된 비트열이 전송 오류에 보다 강인하게 되도록 하기 위해, 본 논문에서 개발된 화질 열화 모델을 양자화기 선택에 적용함으로써, 새로운 최적 양자화 기법을 제시하였다. 본 논문에서 제안된 최적 양자화 기법은, 기존의 양자기 최적화 기법들과는 달리, 복호단에서의 복원 영상 화질이 주어진 비트율에서 최적이 되도록 양자화를 수행한다. H.263 비디오 압축 규격에 적용한 제안 양자화 기법의 실험 결과를 통해, 제안 기법이 매우 적은 계산상의 부하를 비용으로 객관적 화질은 물론 주관적 화질까지 크게 개선할 수 있음을 확인 할 수 있었다.

### ABSTRACT

Current outburst needs for multimedia communication over error prone network such as mobile channel pose many challenging problems. One of the most important issues is the error handling technique for transmitted bitstream including video because the errors severely propagate through spatio-temporal direction in video and thus may cause serious degradation in reconstructed video quality. In order to prevent such devastating error effects, error resilient entropy coding (EREC) technique was applied to the proposed scheme in this paper. We analyze the error propagation in EREC based on a specific searching pattern, and then develop an estimation technique for the probability of a basic coding unit being in error at its decoding. Then we approximate the estimated probability to model the end-to-end degradation by the function of generated number of bits for each basic coding unit. Based on the approximation, we propose a new optimal quantization scheme, in order to make the transmitted video more resilient to the channel errors. Our method is optimal in the sense that the distortion

\* 연세대학교 공과대학 기계전자공학부  
논문번호: 99518-1230, 접수일자: 1999년 12월 30일

\*\* LG전자 정보통신 CDMA 단말연구소

of reconstructed video at decoder is minimum at the given rate. Through simulations with H.263 video codec, we demonstrate that the proposed method generates error-robust bitstream which improves objective as well as subjective video quality at the cost of negligible additional computations.

## I. 서론

이동 통신이 급속도로 발전하는 현 상황에서 휴대용 단말기에서의 멀티미디어 서비스가 많이 개발되고 있다. 그 중에서도 동영상의 전송은 필수적인 기술이 되었다. 하지만, 채널의 상태가 안 좋은 경우 기존의 부호화 기법으로는 만족할 만한 성능을 기대할 수 없는 상황이어서, 무선 통신 채널과 같이 오류가 많은 환경에서 안정적으로 동영상을 전송하는 것은 매우 어려운 일이다. 따라서 동영상의 안정적인 전송을 위한 새로운 코덱의 개발이 필요한 때이다.

부호화된 영상을 전송하는 경우, 채널의 상태가 좋지 않으면 높은 정확도의 소스 코딩이 더 부정확한 결과를 낳는다는 연구 결과는 알려지지 오래이다. 위 결과는 모두 엔트로피 부호화 부분을 고려하지 못한 것이기는 하지만, 엔트로피 부호화를 추가하더라도 그 결과는 비슷할 것으로 예측된다<sup>[1][3]</sup>. 엔트로피 부호화를 고려하는 경우에는, 가변장 부호의 오류 전파에 대해서 그 동안 지나치게 과대 평가된 부분이 있었기 때문에 그 부분을 해결해야 한다. 그러면서도 효과적으로 발생 오류를 처리할 수 있는 '오류 회복 (Error Resilience)' 기법이 첨가된다면 오류에 대한 특성이 더욱 좋아질 것이다.

EREC(Error Resilient Entropy Coding) 알고리즘<sup>[4]</sup>은 발생한 비트 오류를 효과적으로 코덱시킴으로써 오류가 있는 재동기 단위 내의 오류 발생 이후 정보들을 모두 버리지 않을 수 있다. 또한 최소한의 부가 정보만으로, 부호화된 매 기본 블록마다 재동기가 가능한 우수한 특성이 있다. 본 논문에서는, 비트 오류가 복원 영상에 미치는 영향을 효과적으로 분석하기 위해 계층적 EREC 구조<sup>[8]</sup>를 부호화된 비트열에 적용하고, 그 전파 특성을 분석하였다. EREC의 오류 전파 특성은 [4]에서도 분석된 바 있지만, 이는 영상 전체에서 (또는 영상 전체를 표현하는 비트열에서) 발생한 비트가 어떤 비율로 전파되는지만을 알 수 있는 평균적인 개념이다. 본 논문에서는 복원 영상에 발생 오류가 어떻게 나타나는지를 효과적으로 분석하기 위해서, 특정한 탐색 패턴에 따른 각 기본 부호화 블록 (매크로블록) 의 오류 전파를 분석하였다. 이를 통해, 부호화된 하나

의 기본 블록이 복호시 오류가 발생할 확률을 추정하였으며, 추정된 확률의 근사를 통해 전송단과 수신단의 양 끝단에서의 비디오 화질 열화를 예측하였다. 추정된 확률의 근사는 간단한 1차식을 통해 선형 회귀법으로 모델링 되었기 때문에 간단한 계산만으로 복원 영상의 화질 열화를 예측할 수 있다. 또한 전송 오류에 강인한 영상 부호화를 위해서 채널의 상태를 고려한 최적 양자화기를 설계한다. 본 논문에서 제안하는 최적 양자화 기법은, 예측된 복원 영상의 화질 열화를 최소로 되도록 함으로써 복호단에서 나타나는 영상 화질 입장에서 최적이다. 기존의 최적 율-왜곡 (rate-distortion) 양자화가 에러를 최소화함으로써 전송 오류 없이 복원된 영상의 화질을 최적화한 것에 비해, 본 논문에서 제안하는 최적 양자화 기법은 보다 일반화된 기법이라 생각할 수 있다.

본 논문의 구성은 다음과 같다. II장에서는 기존의 논문에 발표된 채널 오류에 대한 복원 영상의 화질 열화 특성과 EREC 구조에 대해서 설명한다. III장에서는 [9]에서 제시된 최적 율-왜곡 양자화 기법에 대한 이론을 설명하고, 본 논문에서 개발한 EREC 오류 전파 특성을 제시한다. 또한, 그 전파 특성의 모델링 및 채널 오류를 고려한 새로운 최적 율-왜곡 양자화기의 설명도 III장에서 다룬다. TMN (Test Model Nearterm) 8 이후에 채택된 H.263+ 양자화기와 제안한 양자화기에 대한 성능 실험 및 그 결과를 IV장에서 보인다. 마지막으로 본 논문의 결론과 향후 과제들에 대한 언급이 V장에 실렸다.

## II. 이론적 배경

일반적으로 영상 데이터는 그 양이 방대하기 때문에, 저장이나 전송을 위해서는 압축 부호화 된다. 영상 압축의 기본적인 원리는 변환을 통해 에너지 집중도를 향상시키고, 변환 계수에 양자화를 수행하여 그 표현을 용이하게 하며, 더욱 높은 압축율을 얻기 위해 엔트로피 부호화를 수행하는 것이다. 이처럼 압축 부호화된 영상 정보의 비트열에 오류가 발생하면 일반적으로 접두 조건 (prefix condition) 을 만족하는 엔트로피 가변장 부호의 복호에 혼란

을 초래하여, 그 오류가 비트열을 따라 (영상의 입장에서 보면, 공간적으로) 전파된다. 또한, 비디오 데이터의 경우에는 압축 부호화에서 일반적으로 움직임 보상 기법을 적용하여, 기준 영상을 사용하므로, 한번 발생한 오류는 일정 시간 동안 다음에 오는 영상들로 그 훼손 영역이 확대 전파된다. 따라서, 이러한 시간-공간적 방향으로의 오류 전파 현상은 비디오 데이터의 오류를 수반하는 채널을 통한 전송에서 반드시 고려되어야 할 중요한 문제이다. 본 절에서는 이와 같은 영상 정보에 있어서의 오류 영향에 대한 기존의 연구 결과들을 소개하고, 이들로부터 복원 영상의 훼손을 줄이기 위해 어떠한 접근 방법이 필요한지를 설명함으로써 본 논문에서 해결하고자 하는 문제를 정의한다.

### 2.1 채널 오류 환경에서의 양자화

채널 오류에 의한 복원 영상 훼손의 정도에 대한 양자화 기법의 영향은 매우 큰 것으로 알려져 있으며, 따라서, 이에 대한 연구가 활발히 이어져오고 있다<sup>[1],[3]</sup>. 전송 오류를 고려하지 않는 경우에 양자기 선택의 문제는 어떻게 하면 주어진 비트율을 넘지 않도록 영상을 부호화 하면서 가장 좋은 복원 영상의 화질을 얻을 것인가에 대한 것이었다. 일반적으로 양자화 간격이 좁을수록 양자화에 의한 화질 열화<sup>1)</sup>는 줄어드는 반면 발생하는 비트가 많아지므로 최적의 타협점을 모색하려는 것이었다. 이러한 문제들은 양자기의 선택에 따르는 발생 비트와 화질 열화를 모델링하여, 제한 조건이 포함된 최적화 기법을 통해 해결하였다.

오류 환경에서의 양자기 선택 문제도 같은 맥락에서 연구되고 있지만, 전송 오류에 의한 가변장 부호의 혼란 영향을 정확히 파악하기는 거의 불가능하기 때문에, 일반적으로 오류가 발생한 다음 비트에서부터 영상의 압축 기법에서 제공되는 재 동기 단위까지의 모든 정보를 유실한 것으로 가정한다. 이러한 가정은 비트 오류를 지나치게 확대 해석하는 것으로, 양자기 선택에 따른 복원 영상의 화질 열화 모델을 실제와는 크게 차이가 나는 최악의 경우에 대한 모델이 되도록 한다. 따라서, 기존의 연구에서는 양자기 선택에 따르는 정확한 화질 열화

모델을 구하기보다는 양자화 기법이 복원 영상의 화질에 어떠한 영향을 미치며, 오류에 보다 강한 부호화를 위해서는 어떤 식의 접근이 바람직한지에 대한 결론을 내리고 있다. [1],[2]에서 Modestino 등은 변환 방법으로 DPCM (Differential Pulse Coded Modulation)과 블록 DCT (Discrete Cosine Transform)를 사용하고, 양자화를 수행한 다음 고정길이 부호를 사용하여 영상을 부호화하는 경우에 있어 양자화 기법이 복원 영상의 화질에 미치는 영향을 고찰하였다. 고정 길이 부호를 사용하는 경우에는 발생한 오류가 이어지는 비트열의 심볼들을 통해 전파되지 않기 때문에, 쉽게 양자화의 영향을 해석할 수 있는데, 여기서 Modestino 등은 채널의 오류 정도가 높을수록 간격이 넓은 (정확도가 떨어지는) 양자기를 사용할 때 최적의 복원 화질을 얻을 수 있다는 결론을 얻었다. 이 결과는 다음과 같은 예를 통해 쉽게 이해할 수 있다. 우선, (0, 1)사이의 실수 값을 가지는 부호화 할 입력 신호를  $x_i$ 라 하고, 양자화 간격이 각각 0.25, 0.5인 균일 양자화기  $Q_{.25}$ 와  $Q_{.5}$ 를 생각해 보자. 입력 신호의 분포가 균일하고 역 양자화에서 양자 구간의 중간 값을 복원 값으로 한다고 가정하면, 전송 오류가 없는 경우 양자화 오차는 다음과 같이 나타낼 수 있다. (단,  $r_i$ 는  $i$ 번째 심볼에 대한 복원 값을 나타낸다.)

$$E = \int_{(n-1)d}^{n \cdot d} (x_i - r_i)^2 \times f(x) dx = \frac{d^2}{12}$$

즉,  $E(Q_{.25}) = 0.25^2/12 = 0.005$  이고

$E(Q_{.5}) = 0.5^2/12 = 0.020$ 으로, 양자화 간격이 큰 양자기가 더 큰 복원 오류를 나타낸다. 이제 각 양자기로 양자화 된 심볼들을 각각 2bit, 1bit의 고정 길이 비트열로 나타내기로 하고, 전이 확률  $\epsilon$ 을 가지는 이진 대칭 채널 (Binary Symmetric Channel)을 통해 전송한다고 가정해 보자. 우선  $Q_{.5}$ 를 사용하는 경우를 생각해 보면, 양자화 된 심볼은 0.5보다 작은 경우와 0.5보다 큰 2가지의 경우가 각각 0, 1의 비트로 부호화 되었을 것이다. 각 비트에 오류가 생길 확률은  $\epsilon$ 이고, 오류가 생길 경우 복원되는 복원치 사이에는 0.5의 차이가 있으므로, 채널 오류에 의해 추가되는 복원 열화의 평균값은  $\epsilon \times 0.5$ 가 되고 따라서 전체 열화는  $0.020 + \epsilon \times 0.5$ 로 표현할 수 있다. 이제 양자기  $Q_{.25}$ 를 사용하는 경우를 고려해 보자. 이 양자기로 입력 신호를 양자화 하는 경우 (0, 0.25], (0.25, 0.5], (0.5, 0.75], (0.75, 1)의

1) 앞서 설명한 일반적인 영상 손실 압축 기법에서, 변환이나 엔트로피 부호화는 모두 무손실 기법이다. 손실 압축에서 유일하게 복원 영상에 훼손을 초래하는 부분이 양자화 부분이므로, 따라서 양자기가 복원 영상의 화질에 미치는 영향은 지대하다.

구간에 속하는 4개의 심볼이 발생하고 각각을 00, 01, 10, 11의 비트열로 부호화 하였다고 가정하자. 우선 한 심볼을 표현하는 두 비트 중에서 앞의 비트에만 오류가 생기는 경우를 생각해 보면, 오류가 생길 확률은  $\epsilon(1-\epsilon)$ 이고 이때의 열화 값은 0.5 (두 단계의 복원치 사이 거리)이므로 평균 열화는  $\epsilon(1-\epsilon) \times 0.5$ 가 된다. 마찬가지로, 두 번째 비트에만 오류가 발생하는 경우 평균 열화 값은  $\epsilon(1-\epsilon) \times 0.25$ 가 된다. 마지막으로 두 비트 모두에 오류가 생기는 경우를 생각해 보자. 이 경우에는 비트열 00 혹은 11에 오류가 생기는 경우에는 열화 값이 0.75가 되지만, 01 혹은 10에 오류가 생기는 경우에는 0.25가 된다. 하지만 입력 신호가 균일 분포를 가진다고 가정하였으므로, 평균 열화 값은  $\epsilon^2 \times 0.5$ 가 된다. 따라서, 전체 열화는  $0.005 + \epsilon(1-\epsilon)(0.5 + 0.25) + \epsilon^2 \times 0.5$ 로 표현할 수 있다. 종합해 보면, 오류가 없을 때  $E(Q_{.5}|\epsilon=0) = 0.005 < E(Q_{.5}|\epsilon=0) = 0.020$ 으로 양자화 간격이 큰 양자기에서 더 큰 양자화 오차가 생겼던 것과는 대조적으로, 채널 오류의 확률이 0.1이 되면 ( $\epsilon=0.1$ ),  $E(Q_{.5}|\epsilon=0.1) = 0.079 > E(Q_{.5}|\epsilon=0.1) = 0.025$ 으로 양자화 간격이 더 작은 양자기에서 더 큰 평균 복원 열화치가 발생한다. 즉, 오류가 없는 상황에서 가장 좋은 양자화 방법일지라도, 전송에 오류가 발생하는 것을 고려하면 양자화 간격이 더 큰 양자화기를 사용하여 적은 양의 비트로 입력 신호를 부호화 하는 것이 보다 우수한 복원 결과를 가져올 수 있다는 것이다.

[3]에서 Farvardin과 Vaishampayan은 양자기의 선택 뿐 아니라 양자화된 각 심볼에 어떤 고정 길이 부호를 할당할 것인가도 채널 오류의 복원 신호 훼손에 영향을 줄 수 있다는 문제를 제시했다. 이를 해결하기 위해, 그들은 양자화 단계와 엔트로피 부호 단계를 따로 구분하지 않고, 대신 하나의 부호화 가능 블록으로 간주하였다. 따라서, 입력된 신호를 직접 고정길이 부호에 매핑하도록 했는데, 다만 매핑될 고정 길이 부호의 수를 달리하여 그 결과를 보였다. 그들의 결론 또한 앞서 살펴본 [1],[2]에서와 마찬가지로, 채널 상황이 열악하면 열악할수록 더욱 적은 개수의 고정길이를 사용하여 부호화 발생하는 전체 비트수를 줄이는 것이 보다 우수한 복원 신호를 얻을 수 있다는 것이었다.

이들 연구들은 모두 가변장 부호를 고려하지 못한 다소간은 일반적이거나 합리적이지 못한 실험

방법을 통해 얻은 결과이지만, 가변장 부호를 고려 하더라도 비슷한 결론을 얻을 수 있을 것으로 기대 된다. 따라서, 채널의 오류가 심할수록 양자화 단계의 수를 줄이는 것이 정보의 왜곡을 더욱 줄일 수 있는 방법이라는 결론을 주목할 필요가 있다.

## 2.2 EREC 알고리즘

EREC 알고리즘은 가변길이의 데이터를 비트 오류에 강한 구조로 재배치한다. 각 기본 부호화 블록 (예를 들어, 매크로블록)의 시작위치를 복호화기와 부호화기가 모두 알 수 있게 배치하는 방식을 사용 하는데, 이로 인해 매 기본 부호 블록마다 자동적으로 동기화가 되는 효과를 얻을 수 있어, 가변장 부호화의 가장 큰 문제점인 오류 전파 현상을 막을 수 있다. 또한 이와 같이 빠른 재 동기를 얻을 수 있으면서도, 추가적으로 전송해야 할 오버헤드가 거의 없다는 장점을 가진다. 오류가 있는 채널에서 EREC는 동기를 놓치지 않을뿐더러 정보의 손실이 크지 않기 때문에 BER (Bit Error Rate)이 높아져도 이 부호 영상의 화질은 그에 비해서 조금만 저 하된다. 따라서, 기존의 오류 특성이 화질 열화에 미치는 영향을 연구하는 경우, 오류 발생 비트부터 다음 번 재 동기까지의 모든 비트열을 의미 없는 것으로 가정했던 것과는 대조적으로, 효과적인 오류 고립화를 피할 수 있기 때문에 비트 오류의 지나친 확대 해석을 피하고 효과적으로 오류 영향을 분석 할 수 있게 된다. 다음은 EREC를 수행하는 구체적인 방법에 대한 설명이다.

EREC의 구조는 길이가  $i$ 번째 슬롯의 길이가  $s_i$ 인 전체  $N$ 개의 슬롯으로 구성된다. 이 경우에 전송 되는 전체 비트 수는  $T = \sum_{i=1}^N s_i$ 가 된다. 중요한 점은 부호단과 복호단에서 모두  $T, N, s_i$ 의 값을 알고 있어야 부복호화가 가능하다는 것이다. 하지만, 실제로는  $T$  값만을 전송하게 되는데, 이는 일반적인 동영상 부호화에서 영상의 크기가 이미 정해져 있기 때문에 기본 부호화 블록의 개수  $N$ 을 따로 전송 할 필요가 없을 뿐만 아니라,  $T, N$ 의 값을 이용하면  $s_i$ 를 계산할 수 있기 때문에  $s_i$  또한 전송할 필요가 없기 때문이다. 따라서,  $T$  값이 오류에 의해서 손상되면 전 프레임의 복호가 불가능하기 때문에 반드시 강력한 오류 정정 부호를 사용하여 전송해야 한다.

EREC 기법은 각 EREC 기본 프레임 ( $N$ 개의 슬롯으로 구성됨)에  $N$ 개의 가변 길이 블록을 재배치

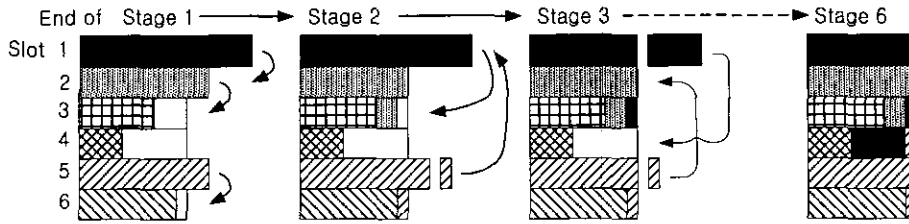


그림 1. EREC 비트 재 정렬 알고리즘

하고 이를 전송하는 것인데, 이렇게 되면  $i$ 번째 부호화 기본 블록에서 발생한 비트 수가  $b_i$ 인 경우, 그 전체 비트 수의 합이  $T$ 보다는 작아야 한다.

$$T = \sum_{i=1}^N s_i > \sum_{i=1}^N b_i$$

이제 EREC에서 사용하는 비트열 재 정렬 방법을 알아보도록 하자.

그림 1에 알고리즘의 간단한 예가 나타나 있다. 첫 번째 단계에서는 매 블록의 자료를 해당하는 슬롯에 배치한다. 배치 할 때에는 최대한 많은 비트를 배치하게 된다. 그러므로 첫 번째 단계가 끝나고 나면 1, 2, 5번 슬롯은 다 차게 된다. 다음 단계부터는 더 재배치 할 비트가 남아 있는 블록들은  $i + \phi_n \pmod{N}$ 으로 정해지는 순서대로 남은 공간이 있는 슬롯을 찾게 된다.  $i$ 는 현재 재배치 할 블록의 번호이고,  $n$ 은 진행중인 단계 (stage)의 수이다.  $\phi_n$ 은 미리 정해진 탐색 패턴 (searching pattern)으로서 빈 슬롯을 찾아가는 순서이다. 그림 1에서의  $\phi_n$ 은 {0, 1, 2, 3, 4, 5}이다. 오류가 없는 채널 환경이라면, 복호화 단계에서는 이와 같은 알고리즘을 그대로 사용하여서 오류 없이 복호화 할 수 있다. H.263의 경우라면 모든 허프만 부호들이 접두 조건 (prefix condition)을 만족하고 있어 그 부호의 끝을 알 수 있고, 모든 블록은 last 부호가 있기 때문에 그 블록의 끝 위치를 알 수 있다. 따라서, 복호화 단계에서는 각 슬롯위치의 기본 블록에서 복호를 반복하면서, 복호가 완료되지 않은 블록들 (last 부호는 복호가 완료됨을 뜻한다)은 여분의 데이터가 존재하는 슬롯을 탐색함으로써 점진적으로 복호를 마칠 수 있게 된다.

### 2.3 계층적 EREC

앞에서 설명한 EREC 기법은 기본적으로 비트 오류의 고립화를 통해서 엔트로피 부호의 오류 전파 현상을 막는 데 있다. 그 외에 영상 신호를 어떻

게 부호화 할 것인지에 대해서는 매우 개방적이어서 블록 단위로 부호화를 수행하는 MPEG-1, 2, 4 혹은 JPEG이나 H.261 등의 어떠한 부호화 기법에도 적용시킬 수 있다. 하지만 본 논문에서는 H.263의 문법 구조를 기준으로 [8]에 근거한 EREC의 구체적인 적용 방법을 설명하도록 하겠다.

우선 H.263의 비트열의 구조를 간단히 알아보면, 가장 큰 단위로서 picture layer가 있다. 여러 가지 picture header와 GOB(group of blocks) layer의 정보가 있다. GOB layer에는 마찬가지로 헤더 정보와 Macroblock(MB) 정보가 있다. 영상의 복호에 실제 사용되는 가장 많은 정보를 포함하는 MB layer는 움직임 벡터와 같은 헤더 정보와 DCT 계수들로 이루어진 block layer로 구성된다. EREC의 적용에서 가장 중요한 문제 중 하나는 EREC 기법의 중요한 구성 요소인 블록과 슬롯을 어떤 단위로 선택할 것인가이다. 직관적으로 생각했을 때 매크로 블록 단위의 EREC가 가장 합리적이다. 왜냐하면 매크로 블록은 영상 신호를 부호화할 때에 중요한 정보를 많이 담고 있고, 영상을 구성하는 기본적인 단위이기 때문이다. 그리고 H.263에서는 GOB의 헤더 정보를 부호화 할 것인가를 선택사항으로 두고 있기 때문에 GOB 헤더가 없다면 picture layer 다음은 곧바로 MB layer가 된다. 그러나, 매크로 블록 단위로 EREC를 할 경우에는 매크로 내의 한 블록이 채널 오류에 의해서 손상됐을 경우 해당하는 매크로 블록의 끝까지 오류가 전파된다. 따라서 기본 단위를 매크로 블록으로만 하는 단순한 EREC 보다는 그 내부의 블록에까지 계층적 구조를 가지도록 하여 발생한 오류에 적절히 대처할 수 있도록 하는 구조가 바람직하다. 이를 위해 본 논문에서는 [8]에서 정의된 계층적 구조의 EREC를 매크로 블록과 블록 단위에 적용하는데, 그 상세 알고리즘은 다음과 같다. 우선, 명확한 설명을 위해서 다음과 같은 몇 가지 기호를 정의하자.

- $T$  : 한 프레임에서 picture 헤더를 제외하고 발생한 비트 수
- $N$  : 프레임에 포함된 매크로 블록의 수
- $s_M$  : 매크로 슬롯의 길이
- $s_B^i$  :  $i$ 번째 매크로 슬롯에서 매크로 블록 헤더의 길이를 빼고 난 후의 슬롯의 길이
- CBP: H.263의 MCBPC (Macroblock Type & Coded Block Pattern for Chrominance) 및 CBPY (Coded Block Pattern for Luminance) 정보로부터 계산된 한 Macroblock 내에 존재하는 Block의 코딩 패턴 (Coded Block Pattern)으로 이를 통해 # of blocks 를 알 수 있다.

여기서 매크로 슬롯이란 매크로 블록에 대응하는 슬롯을 말한다. 계층적 EREC 기법을 구현하기 위한 알고리즘은 다음과 같다.

1.  $s_M = \lceil \frac{T}{N} \rceil$  를 계산한다.  $\lceil \cdot \rceil$  은 ceiling function이다.
  2. CBP에 의해  $s_B^i = \frac{s_M - \text{MB header}}{\text{\# of blocks}}$  를 계산한다.
  3. 블록단위의 EREC를 수행한다.
  4. 매크로 슬롯에 매크로 블록 헤더를 앞에서부터 차례로 넣은 후 단계 3에서 EREC가 끝난 블록 데이터를 슬롯에 넣는다.
  5. 매크로 슬롯에 다 들어가지 못한 데이터를 순서대로 정렬한 뒤, 비어있는 매크로 슬롯에 EREC에 의해 넣는다. 이때 빈 매크로 슬롯 내에서 비어있는 슬롯의 뒤에 블록 순서대로 넣는다.
- 위의 알고리즘은 그림 2를 참조하면 쉽게 이해할 수 있다. 여기에서 주의할 점은 매크로 블록 헤

더의 정보는 중요하기 때문에 오류의 전파 현상을 방지하기 위해서 블록 단위의 EREC에서 제외한다. 따라서 헤더는 다른 매크로 슬롯에 나누어서 저장되지 않기 때문에 오류가 생길 확률이 매우 낮아진다. 예를 들어  $i$ 번째 매크로 블록의 헤더 정보에 해당하는 비트열의 길이를  $s_B^i$ 라 하자. 부호화 된 비트열이 전이 확률  $\epsilon$ 을 가지는 이진 대칭 채널 (Binary Symmetric Channel)을 통해 전송된다고 가정하면, 그 매크로 블록의 헤더에 오류가 생길 확률은  $(1 - (1 - \epsilon)^{s_B^i})$ 로  $\epsilon = 10^{-3}$ 으로 표현되며, 전송한 매크로 블록의 헤더가 10비트의 길이를 갖는다면 그 값은 0.00996이 된다. 만약  $176 \times 144$ 의 영상이 라면 한 프레임 99개의 매크로 블록 중에서 매크로 블록 헤더에 오류가 생기는 개수는 한 개 이하로 볼 수 있다.

### III. Rate Control 기법과 오류 전파 모델

#### 3.1 TMN8 (Test Model Nearterm 8)에서의 Rate Control 기법

최근 블록 DCT를 사용하는 비디오 코덱을 위한 율-왜곡 (Rate-Distortion) 최적 rate control에 관한 논문이 발표되었고, 그 내용이 MPEG-4와 TMN8의 rate control 참고 자료로 채택되었다<sup>[9]</sup>. 본 논문에서 제안하는 양자화 기법은 [9]의 rate control 기법 중 매크로 블록 레벨에 적용한 율-왜곡 최적 기술을 기반으로 그 비용함수 (Cost Function)를 설계하였다. 따라서, 본 절에서는 [9]에서 제안한 내용을 간략히 살펴봄으로써 본 논문에서 제시하는 양자화기 선택 기법에 대한 이해를 돕도록 한다.

우선 영상의 한 프레임이  $N$ 개의 매크로 블록으로 구성되어지고, 원하는 프레임 당 비트 수가  $B$ 라고 가정하자. 그렇게 되면 각 매크로 블록의 최적 양자화기  $Q^* = \{Q_1, Q_2, \dots, Q_N\}$ 는 다음의 식을 만족한다.

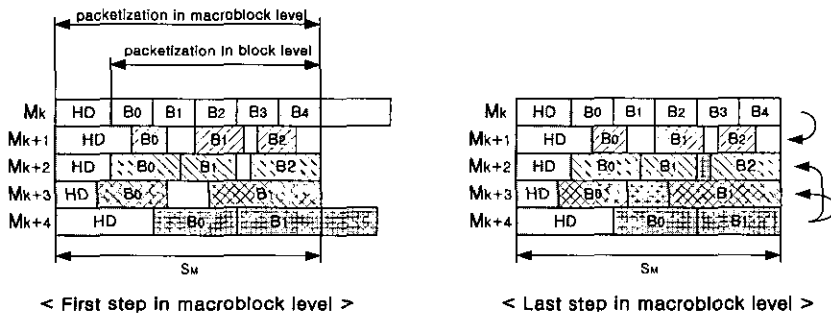


그림 2. 계층적 구조에서의 비트 재배열

2) 계층적 EREC의 자세한 구조는 참고문헌 [8]을 참고한다.

$$Q^* = \underset{Q}{\operatorname{argmin}} D(Q) \text{ subject to } R(Q) \leq B \quad (1)$$

여기에서  $D(Q)$ 와  $R(Q)$ 는 양자화기  $Q$ 의 선택에 따른 복원 영상의 왜곡과 비트율을 나타낸다. 이 식에서 구해진 최적의 양자화기  $Q^*$ 는 압축 부호화된 영상의 화질을 최대화하면서 정해진 비트율  $B$ 를 유지하게 된다. 이제 양자기의 선택에 따른 복호 영상의 왜곡  $D(Q)$ 와 발생한 비트율  $R(Q)$ 이 어떻게 정의되고, 또 최적의 양자화기는 어떻게 선택되는지 자세히 살펴보도록 하자.

움직임 보상된 차이 영상의 DCT 계수들은 분산  $\sigma^2$ 를 가지는 라플라시안 (Laplacian) 분포가 된다고 볼 수 있기 때문에, 양자화 단계  $Q$ 로 양자화 된 계수들의 엔트로피는 다음의 식으로 근사 될 수 있다<sup>[10][11]</sup>.

$$H(Q) = \begin{cases} \frac{1}{2} \log_2 \left( 2e^2 \frac{\sigma^2}{Q^2} \right) & \frac{\sigma^2}{Q^2} > \frac{1}{2e} \\ \frac{e}{\ln 2} \frac{\sigma^2}{Q^2} & \frac{\sigma^2}{Q^2} \leq \frac{1}{2e} \end{cases} \quad (2)$$

따라서  $R(Q) \approx H(Q)$ 이기 때문에 양자화기 선택에 따라 발생하는 프레임 당 비트 수는 다음의 식으로 표현할 수 있다.

$$R(Q) = \sum_{i=1}^N B_i = \sum_{i=1}^N A \left( K \frac{\sigma_i^2}{Q_i^2} + C \right) \quad (3)$$

이 식에서  $A$ 는 매크로 블록 안의 픽셀의 개수이고,  $K$ 는 식(2)에서처럼  $e/\ln 2$ 로 설정할 수 있으나, 식(2)도 근사식이기 때문에 부호화 과정에서 계속 값을 갱신해 주도록 한다. 상수  $C$ 는 움직임 벡터와 같은 헤더의 길이를 나타내는 값이다.

또한 [9]에서는 양자화기 선택에 따른 왜곡 모델로서 다음과 같은 가중 양자화 왜곡에 대한 식을 사용한다.

$$D(Q) = \frac{1}{N} \sum_{i=1}^N D(Q_i) = \frac{1}{N} \sum_{i=1}^N a_i^2 \frac{Q_i^2}{12} \quad (4)$$

여기서  $a_i$ 가 의미하는 것은 매크로 블록마다 왜곡의 가중치이다.

식 (1)에서 정의된 조건이 있는 최적화 문제를 풀기 위해서 Lagrange multiplier  $\lambda$ 를 사용할 수 있다. 식 (1)에  $\lambda$ 를 이용하면, 식 (5)처럼 하나의 식으로 조건이 없는 비용함수로 표현이 가능하다.

$$Q_1^*, \dots, Q_N^*, \lambda^* = \underset{Q_1, \dots, Q_N, \lambda}{\operatorname{argmin}} \frac{1}{N} \sum_{i=1}^N a_i^2 \frac{Q_i^2}{12} + \lambda \left[ \sum_{i=1}^N A \left( K \frac{\sigma_i^2}{Q_i^2} + C \right) - B \right] \quad (5)$$

이 왜곡 함수는 인자  $Q$ 에 대한 2차 정방식 형태이고 미분 가능하기 때문에 Lagrange 이론에 의하면 유일해를 찾을 수 있다. 이 식을 편미분하여 꼭지점을 찾으면 식 (6)으로 표현되는 최적 양자화 단계를 찾을 수 있다.<sup>3)</sup>

$$Q_i = \sqrt{\frac{AK}{B - ANC} \cdot \frac{\sigma_i}{a_i} \sum_{k=1}^N a_k \sigma_k} \quad (6)$$

### 3.2 계층적 EREC에서의 오류 전파 특성

채널 오류를 고려한 최적 양자화기를 설계하기 위해서는 오류가 생겼을 때에 어느 정도 전파되는 특성이 있는지 알아야 한다. 계층적 EREC의 구조가 워낙 복잡하기 때문에 그 오류 전파 특성 또한 쉽게 분석하고 수식화 하기 힘들다. 따라서 우선 몇 개의 기호에 대해서 정의가 필요하다.

- $s_i$  :  $i$ 번째 매크로 슬롯의 길이
- $b_i$  :  $i$ 번째 매크로 블록에서 발생한 비트열의 길이
- $DMB_i$ :  $i$ 번째 MB의 복호화
- $e^k(\cdot)$ : EREC 알고리즘에서  $k$ 번째 단계 이내에 오류가 발생하는 사건
- $\bar{e}^k(\cdot)$ : 오류가 발생하지 않는 사건.  $e^k(\cdot)$ 의 여사건
- $p(k)$ :  $k$ 번째 단계의 탐색 패턴
- $b_{i+p(k)}$ :  $(i+p(k) \bmod N)$ 번째 슬롯에 저장된  $i$ 번째 매크로 블록의 비트열
- $I_{i+p(k)}$ :  $(i+p(k) \bmod N)$ 번째 슬롯의  $i$ 번째 MB 데이터 이전에 저장된 비트열들의 인덱스
- $h_{i+p(k)}$ :  $(i+p(k) \bmod N)$ 번째 슬롯으로부터  $i$ 번째 MB 복호를 위해 전달된 정보

환경은 전이 확률  $\epsilon$ 의 BPSK(Binary Phase Shift Keying)로 전송하는 AWGN(Additive White Gaussian Noise)채널으로 가정한다. 본 절에서는 복호화기가 채널을 통해 전송된 비트열로 복호화를 하는 중 각 매크로 블록의 복호화에 오류가 나타날 확률을 분석하려는 것이 목표이므로, 이는 다음과 같이 표현되어 질 수 있다.

3) 자세한 전개 과정은 부록을 참조한다.

$$\Pr\{e^k(DMB_i)\} \text{ (단, 탐색 패턴 } p(k) \text{에 의존함)} \quad (7)$$

식 (7)은  $i$ 번째 매크로 블록이  $k$ 번째 단계에서 오류가 있을 확률을 의미한다. 이 때에는  $(i+p(k) \bmod N)$ 번째 슬롯의 빈 공간을 탐색하게 될 것이다. 위의 정의에서  $b_{i+p(k)}$ 는 이 때에 저장된  $i$ 번째 매크로 블록의 비트열을 의미한다. 그리고 그 비트열이 저장되어 있는  $(i+p(k) \bmod N)$ 번째 슬롯에  $b_{i+p(k)}$  이전에 저장되어 있는 모든 비트열들의 매크로 블록 인덱스의 집합을  $I_{i+p(k)}$ 라 표현한다.

먼저 0번째 단계에서의 오류 발생 확률에 대해서 알아보면, 이 단계에서는 매크로 블록의 비트열이 자신의 슬롯에 저장되기 때문에  $b_i$ 와  $s_i$ 의 크기에 따라서 약간의 확률이 달라지게 된다. 이는 다음 식과 같이 적어도 한 개의 오류가 발생할 확률이기 때문에 전체 확률에서 오류가 없는 경우를 뺀 값과 같다.

$$\Pr\{e^0(DMB_i)\} = \begin{cases} \sum_{k=0}^{b_i} \binom{b_i}{k} \varepsilon^k (1-\varepsilon)^{b_i-k} = 1 - (1-\varepsilon)^{b_i}, & \text{if } b_i < s_i \\ \sum_{k=0}^{s_i} \binom{s_i}{k} \varepsilon^k (1-\varepsilon)^{s_i-k} = 1 - (1-\varepsilon)^{s_i}, & \text{if } b_i \geq s_i \end{cases} \quad (8)$$

만약에  $b_i < s_i$ 인 경우에는  $i$ 번째 매크로 블록의 부호화가 끝난 것이기 때문에 다른 슬롯의 오류 여부에 전혀 관계가 없게 된다. 따라서,  $\Pr\{e^k(DMB_i)\} = \Pr\{e^0(DMB_i)\}$ 와 같이 단계에 무관한 오류 확률을 가진다.

그러나,  $b_i \geq s_i$ 인 경우에는 단계가 높아질수록 확률은 복잡해진다.  $k=1$ 인 경우를 생각해 보면, 부호화의 경우  $(i+p(k) \bmod N)$ 번째 슬롯의 정보가 필요한데, 이 정보는  $(i+p(k) \bmod N)$ 번째 슬롯에 저장된  $b_{i+p(k)}$ 에 국한되지 않는다. 이는 다음의 두 가지 상황으로 분석할 수 있다.

- ①  $b_{i+p(k)} \neq \emptyset$  ( $(i+p(k) \bmod N)$ 번째 슬롯에 정보가 저장된 경우
- ②  $b_{i+p(k)} = \emptyset$  ( $(i+p(k) \bmod N)$ 번째 슬롯이 모두 차 (full) 있는 경우  
(공간이 없어서 정보를 저장하지 못한 경우)

우선 ①번의 경우에는  $b_{i+p(k)}$ 에는 반드시 오류가 없어야 한다. 그 뿐만 아니라,  $(i+p(k) \bmod N)$ 번째

슬롯에 오류가 있다면  $b_{i+p(k)}$ 에도 전파될 것이기 때문에 슬롯 자체에 오류가 없어야 한다. 그리고 ②번의 경우에도 마찬가지로 현 슬롯에는 가져올 정보가 저장되어 있지는 않지만, 슬롯에 오류가 발생한다면 잘못된 비트열을  $i$ 번째 매크로 블록에 전달할 수 있다. 따라서 현재 탐색하고 있는  $(i+p(k) \bmod N)$ 번째 슬롯에 기록된 다른 모든 매크로 블록이 오류가 없어야 정상적인 비트열  $b_{i+p(k)}$ 을 전달할 수 있다. 그러므로, 단계가 2번째 이상인 경우에는 더욱 복잡해진다.

그럼 3을 참조하면,  $i$ 번째 매크로 블록의 비트열을 저장하기 이전에 이미  $j$ 번째 매크로 블록의 비트열이 저장되어 있는 것을 볼 수 있다. 앞에서 설명했듯이 이 경우에  $i$ 번째 매크로 블록 복호에 오류가 없기 위해서는  $(i+p(k) \bmod N)$ 번째 슬롯과  $j$ 번째 매크로 블록에 모두 오류가 없어야 한다. 여기에서 문제가 되는 것은  $j$ 번째 매크로 블록인데, 그 이유는 다음과 같다.  $j$ 번째 매크로 블록에 오류가 없기 위해서는  $j$ 번째 매크로 블록이  $(k-1)$  단계 이내에 검색을 했던 모든 슬롯들이 오류가 없어야 한다. 그리고 그 슬롯들에는 또 다른 매크로 블록들이 정보를 저장했기 때문에 그 매크로 블록들도 오류가 없어야 한다. 그러므로 한가지 블록의 수렴 단계가 커질수록 오류가 발생할 확률은 더욱 높아지게 된다. 이와 같은 내용을 식으로 표현하면 다음과 같다.

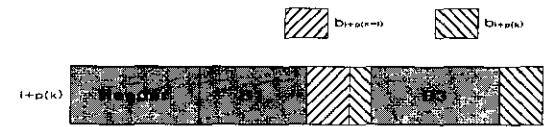


그림 3.  $k (\geq 2)$ 번째 탐색 단계에서의  $i$ 번째 매크로 블록의 복호화

$$\Pr\{e^k(DMB_i)\} = \Pr\{e^{k-1}(DMB_i)\} + \Pr\{\tilde{e}^{k-1}(DMB_i), e^k(h_{i+p(k)})\} \quad (9)$$

여기에서  $h_{i+p(k)}$ 는 위의 정의에 따라서  $(i+p(k) \bmod N)$ 번째 슬롯으로부터 전달된 총 정보로서 다음의 식으로 설명된다.

$$h_{i+p(k)} = \left[ \bigcup_{j \in I_{i+p(k)}} e^{k-1}(DMB_j) \right] \cup e^k(b_{i+p(k)}) \quad (10)$$

따라서 계층적 EREC에서의 오류 전파 특성은 위의 두 식 (9), (10)에 의해서 점화식으로 규정할 수 있다.



그러나 이 식은 계산하는 것이 거의 불가능하기 때문에 계산상의 간편함을 위해서 확률 식을 다음과 같이 정의했다.

$$\Pr\{e^k(DMB, i)\} = N \times [1 - (1 - \epsilon)^{s_i}] - \binom{N}{2} \times [1 - (1 - \epsilon)^{s_i}]^2 + \binom{N}{3} \times [1 - (1 - \epsilon)^{s_i}]^3 - \dots \approx N \times [1 - (1 - \epsilon)^{s_i}] \quad (11)$$

식 (11)에서  $N$ 은  $i$ 번째 매크로 블록이 복호화를 하는 과정에서  $k$ 번째 단계까지에서 복호화에 영향을 미칠 수 있는 슬롯들의 개수를 나타낸다. 영향을 미칠 수 있는 슬롯이란, 현재 복호화 중인 블록이 탐색했던 슬롯을 포함해서 그 슬롯을 탐색했던 다른 블록들이 거처온 슬롯들을 지칭한다. 하나의 슬롯에 오류가 생길 확률은 각각  $[1 - (1 - \epsilon)^{s_i}]$ 로 같다. 따라서 한 개의 슬롯에 오류가 생길 확률과 영향력이 있는 슬롯의 개수를 곱하면 전체 확률을 구할 수 있다.

### 3.3 제한한 채널 오류를 고려한 최적 양자화기

앞 절에서 유도한 것과 같이  $N$ 의 경향만 밝혀낸다면 정해진 BER에서 특정 매크로 블록에 오류가 발생할 확률은 쉽게 구할 수 있을 것이다. 그렇다면  $N$ 을 어떤 함수로 모델링 하느냐는 것이 관건이다.  $N$ 은  $i$ 번째 매크로 블록이 탐색하는 횟수에 관계하고, 매크로 블록이 탐색하는 횟수는 보통 발생한 비트 수에 관계하기 때문에  $N$ 은  $(R(Q) - s_i)$ 의 함수라고 볼 수 있다. 이 두 인자 사이의 관계를 실험적으로 얻은 결과가 그림 4에 표현되어있다. 그림 4에서 가로축은 각 매크로 블록에서 발생한 비트 수에서 슬롯의 길이를 뺀 값을 나타내고, 세로축은 양자화기  $Q$ 에 대한 함수  $N(Q)$ 를 나타낸다. 그리고 실선은 실제 영상을 부호화하면서 구한 실험치이고, 점선은 결과를 이용해서 일차 함수로 근사해 본 것이다. 그래프는 가로축의 오른쪽으로 갈수록 거칠어지는데, 이는 샘플의 부족으로 통계적 특성을 구하기 힘들었기 때문이다. 그래프에서 볼 수 있는 바와 같이  $N(Q)$ 는  $(R(Q) - s_i)$ 에 대한 일차 함수로 모델링 하는 것이 적절하다. 이전 프레임까지 발생한  $N$ 값을 기준으로 하여서 linear regression을 수행하면 직선의 기울기  $m$ 을 쉽게 구할 수 있는데 이는 다음과 같다.

$$N'(Q) = \begin{cases} m(R(Q) - s_i) + 1 & R(Q) \geq s_i \\ 1 & R(Q) < s_i \end{cases} \quad (12)$$

$$m = \frac{S_{xy} - S_x S_y}{S_{xx}} \quad (\text{where, } S_{xy} = \sum_{i=1}^N (b_i - s_i) N_i, \quad S_x = \sum_{i=1}^N (b_i - s_i), \quad S_{xx} = \sum_{i=1}^N (b_i - s_i)(b_i - s_i)) \quad (12-1)$$

$$N(Q) = \min(N'(Q), 99) \quad (12-2)$$

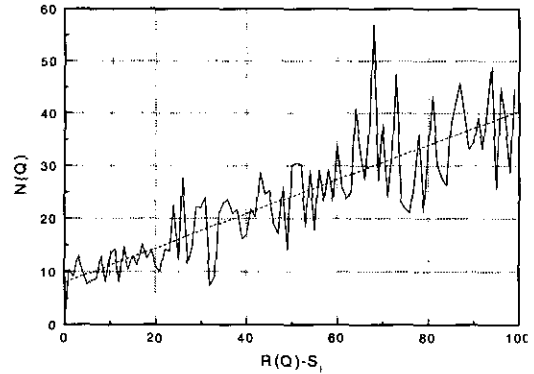


그림 4. 발생한 비트에 따른  $N$ 의 그래프

이 결과를 이용하면 채널 오류에 따른 복원 영상의 화질 열화에 관한 왜곡 함수를 정의할 수 있다. 우선, 위의 근사를 식 (11)에 적용하면 다음과 같이 표현할 수 있다.

$$\Pr\{e^\infty(DMB, i)\} = (1 - (1 - \epsilon)^{s_i}) \times (m(A(K \frac{\sigma_i^2}{Q_i^2} + C) - S_i) + 1) \quad (13)$$

따라서, 이 식을 이용하여 채널에 의한 왜곡 함수  $D_c(Q)$ 를 다음과 같이 정의할 수 있다.

$$D_c(Q) = \sigma_i^2 \times (1 - (1 - \epsilon)^{s_i}) \times (m(A(K \frac{\sigma_i^2}{Q_i^2} + C) - S_i) + 1) \quad (14)$$

식 (14)에서  $\sigma_i^2$ 는 실제로 전송된 차이 영상의 정보량을 뜻하는데, 복원 영상의 화질에 미치는 채널 오류만의 영향을 고려하기 위한 양자화를 거친 후의 영상의 분산 값을 나타낸다. 이제 채널 오류를 고려한 양자화기를 설계하기 위해 식 (5)의 비용 함수를 변형하여서 새롭게 정의하도록 하자.

$$Q_i^*, \dots, Q_N^*, \lambda^* = \underset{Q_1, \dots, Q_N, \lambda}{\operatorname{argmin}} [D_c(Q) + D_s(Q)] + \lambda [R(Q) - B] = \underset{Q_1, \dots, Q_N, \lambda}{\operatorname{argmin}} \left[ \frac{1}{N} \sum_{i=1}^N \left( \sigma_i^2 \frac{Q_i^2}{12} + \sigma_i^2 \times (1 - (1 - \epsilon)^{s_i}) \times (m(A(K \frac{\sigma_i^2}{Q_i^2} + C) - S_i) + 1) \right) + \lambda \left[ \sum_{i=1}^N A \left( K \frac{\sigma_i^2}{Q_i^2} + C \right) - B \right] \right] \quad (15)$$

식 (15)를 만족하는  $Q_i^*$ 는 비트율  $B$ 를 유지하면서 압축 부호화한 영상의 비트열을 BER  $\epsilon$ 의 이진 대칭 채널을 통해서 전송했을 때에 복원 영상의 화질 열화를 최소화 할 수 있는 양자화기의 단계를 나타낸다. 그러나 식 (15)의 최적화에서는 식 (6)과 같은 형태의 해를 직접 구할 수 없고, 다만 재귀적인 방법에 의해 각 양자화 단계의 설정에 따르는 비용함수를 계산할 수 있을 뿐이다. 따라서 각 매크로블록이 선택할 수 있는 모든 양자화 단계에 대해서 식 (15)의 비용함수를 계산하고, 그 중에서 최소의 비용함수를 가지는 양자화 단계를 최적 양자화기로 선택해야 한다. 본 논문에서는 이러한 최적 양자화기 선택의 문제를 H.263의 기본 모드에 구현하여 그 성능을 평가했는데, H.263에서는 각 매크로블록의 양자화 단계 변화를 표현하는 DQUANT라는 파라미터가 2비트의 고정장 부호로 +2~-2까지 0을 포함하여 5가지의 값만을 가질 수 있기 때문에 부가되는 계산상의 부하가 무시할만하다. 최근 완성된 MPEG-4 규격에서도 마찬가지로 2bit의 고정장 부호를 통해 5가지의 양자화 단계 변화만을 허용하므로 구현상의 문제는 없지만, MPEG-1이나 MPEG-2 규격의 경우에는 고려해야 할 사항이 있다. MPEG-1,2 규격에서 각 매크로블록에 허용하는 양자화 단계는 5bit의 quantizer\_scale 파라미터를 통해 표현되기 때문에, 모두 31가지의 선택이 가능하기 때문이다.<sup>4)</sup> 이렇게 각 매크로블록이 선택할 수 있는 양자화 단계의 경우가 많은 경우에는 그만큼 비용함수의 계산을 여러 번 수행해야 하기 때문에, 추가되는 계산상의 부하가 더 이상 무시할 수 있을 만큼 적지 않다. 따라서, 이러한 경우 계산상의 부담을 줄일 수 있는 빠른 알고리즘의 개발이 구현의 필수적인 문제가 될 수 있는데, 2.1절에서 설명했던 채널 오류 환경에서의 양자화에 대한 기존의 연구 결과를 통해 한 가지 간단한 알고리즘을 제시할 수 있다. 우선 채널 오류가 없을 경우의 최적 양자화 단계를 식 (6)을 통해 계산한다. 그리고, 채널의 오류가 심해질수록 양자화 단계의 개수를 줄이는 것이 바람직하다는 기존 연구 결과를 이용하여, 채널의 비트 오류율 (BER)의 정도에 따라 계산된 양자화 단계를 낮춘다. 이 과정을 수행하기 위해서는 식 (15)를 통해 구한 최적 양자화 단계와 식 (6)을 이용해 구한 최적 양자화 단계 사이의 차이를 채널의 BER에 따른 함수로 모델링

하는 과정이 반드시 필요하다. 그 후 낮추어진 양자화 단계를 시작점으로 작은 양자화 단계 변화만을 허용하여 H.263에서 구현했던 방법과 같이 작은 영역 탐색을 수행하는 것이다.

#### IV. 실험 및 결과

##### 4.1 실험 방법

본 논문에서 제안하는 소스-채널 결합 최적 양자화기가 기존의 울-왜곡 최적 양자화기의 성능에 비해 오류를 수반하는 채널에서 우수한 특성을 나타낼 수 있음을 보이기 위해 TMN8 코덱<sup>5)</sup>에 제안 시스템을 구현하였다. 부/복호에 사용된 테스트 영상은 비교적 움직임이 많은 Foreman 시퀀스를 사용하였고, 그 영상 포맷은 QCIF이다. 제안 시스템은 식 (15)로 대변되는 양자화기를 사용하여 테스트 영상을 부호화 한 후 발생된 비트열에 2.3절에서 설명한 계층적 EREC를 적용하였고, 이와 비교할 시스템에서는 TMN8에서 채택한 울-왜곡 최적 양자화기를 이용하여 부호화를 수행한 후 마찬가지로 EREC를 적용하였다. 각 비트열에 EREC를 수행하는 과정에서 사용되는 탐색 패턴으로는 아직까지 그 성능이 우수한 것으로 알려져 있는 FEREC(Fast EREC) 알고리즘<sup>5)</sup>을 사용하였다. 부호화에 사용된 option은 우선 테스트 시퀀스를 100번째 프레임까지 10프레임마다 I-picture로 갱신하면서 10 frame/second로 부호화 하였으며, 목표 비트율은 50kbps와 100kbps로 하였다. 부호화된 비트열은 BPSK(Binary Phase Shift Keying) 변조되어 채널 부호화 없이 전송된다고 가정하였으며, 일반적인 통신 시스템의 목표 BER인  $1 \times 10^{-2}$ 부터  $1 \times 10^{-4}$ 사이의 AWGN(Additive White Gaussian Noise) 채널을 통해 100번씩 반복하여 전송하였다. 합리적인 상황에서 형평성 있는 비교를 위해 매 프레임의 picture header 정보는 채널 잡음에 의해 손상되지 않는다고 가정한다. 이는 picture header가 손상되면 해당 프레임 전체의 복호화가 불가능해지기 때문에 두 시스템의 성능 비교가 무의미하기 때문이다. 또한 BSC(Binary Symmetric Channel)로 가정된 상황에서 그 채널의 전이 확률  $\epsilon$ 에 대한 정보는 부호화 전에 기저국 등을 통해서 전송 받을 수 있다고 가정하였다.

4) quantizer\_scale= 00000인 경우는 허용되지 않기 때문에 모두 31가지의 양자화 단계 선택이 가능하다.

5) 본 실험에서 사용한 부복호화기는 The University of British Columbia의 TMN coder version 3.0이다.

표 2. 목표 비트율과 BER에 따른 실제 부호화된 양자화 단계와 비트율

비트율 (kbps)	BER7		$1 \times 10^{-2}$	$5 \times 10^{-3}$	$1 \times 10^{-3}$	$1 \times 10^{-4}$
51.2	TMN8	Q의평균	11.04			
		실제 얻은 비트율	52.01			
	제한한 기법	Q의평균	12.03	11.49	11.02	11.03
		실제 얻은 비트율	50.27	51.35	52.03	52.02
102.4	TMN8	Q의평균	6.35			
		실제 얻은 비트율	104.88			
	제한한 기법	Q의평균	7.78	7.78	7.71	6.62
		실제 얻은 비트율	95.06	95.02	95.82	102.78

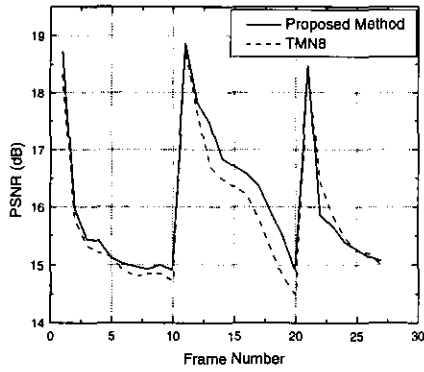
#### 4.2 제한한 rate control 기법의 성능 비교 평가

실험에 의해서 얻어진 목표 비트율에 대한 발생 비트율의 결과를 표 1에 나타냈다. 표에서 볼 수 있는 바와 같이 전반적으로 채널의 에러 확률이 커질 수록 더 낮은 발생 비트율을 나타낸다는 것을 알 수 있다. 이는 2.1절에서 설명했던 기존의 연구 결과와 일맥상통하는 것으로 기대했던 시스템 특성과 잘 맞아떨어진다. 하지만, 이 결과에서 한가지 주목해야 할 점은 채널의 오류율이 크게 변화함에도 불구하고, 발생된 비트율은 목표 비트율에서 크게 달라지지는 않는다는 점이다. 심지어 목표비트율 50Kbps의  $\epsilon = 10^{-3}$ 과  $\epsilon = 10^{-4}$  채널에서의 결과나 목표비트율 100Kbps의  $\epsilon = 10^{-2}$ 과  $\epsilon = 10^{-3}$  채널에서의 결과를 보면 기대했던 바와 반대로 더 높은 에러율을 가지는 채널에서 더 많은 비트가 발생되었다. 이러한 결과가 나타난 이유는 본 논문에서 사용한 근사 모델의 모델링 오차에 기인한 것으로 볼 수도 있지만, 그보다는 비트열에 적용한 EREC 기법의 특성에 의해 발생한 것으로 생각된다. 좀 더 자세히 설명하자면, 식 (15)로 표현된 비용함수에서 특정 매크로 블록에 오류가 발생할 확률은 EREC 복호 과정에서 나타나는 오류 전파 특성에 의해 크게 좌우되는데, 비록 비트 발생량에 따라 그 전파특성이 달라지는 것은 분명한 사실이지만, 사용된 탐색 패턴이나 기 부호화된 매크로 블록에서 실제 발생된 비트수의 형태에 따라 EREC 복호 에러 전파 특성이 다소 달라질 수 있다는 것이다. 따라서, EREC 복호시 나타나는 복호 에러의 추정을 간단한 모델로 근사하여 현재 부호화 할 매크로 블록에서 발생될 비트수만의 함수로 해석한 본 논문의 제안 방식이 적용된 EREC 기법의 특성에 의해 다소 왜

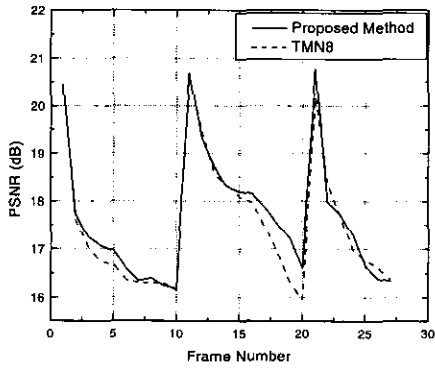
곡되었을 가능성이 있다. 하지만, 전반적인 비트 발생률의 특성과 그림 5, 6, 7을 통해 나타난 제안 시스템의 우수성을 고려하면, 본 논문에서 고려한 비트 발생량에 따른 복호 에러 추정이 복호 에러에 영향을 줄 수 있는 다른 인자들에 비해 복호 오류 전반을 지배하는 특성이라는 결론을 유추할 수 있다.

다음은 복호 영상의 화질 성능 비교를 위해 100 번씩 복호된 각 프레임의 평균 PSNR을 그림 5, 6에 도시하였다. 그림에서 보이는 바와 같이, 전반적으로 0.2-0.6dB의 평균 (전체 프레임에 대한 평균) PSNR 개선 효과를 관찰할 수 있으며, 그림 6-(d)의 경우에는 1dB 정도의 화질 개선을 보이는 프레임을 볼 수 있다. 하지만, 결과에서 제안 시스템의 성능이 기존 방법에 비해 월등히 향상되지 않았던 (예를 들어 그림 5-(d))이유 중 하나는 오류 은닉과 같은 후처리를 고려하지 않았기 때문이다. 즉, 본 논문에서의 실험은 복호중 오류가 발생하더라도 특별한 처리를 하지 않고 끝까지 복호를 수행하기 때문에, 예를 들어 고주파 DCT 계수에 오류가 발생하는 경우 심하게는 비독판 무늬의 블록이 발생하게 된다. 따라서, 제안 시스템의 복호 과정에서 이와 같은 치명적인 오류가 결과의 효율적인 비교를 다소 저해하는 것으로 생각된다. 하지만, 본 논문의 저자가 아는 한 TMN8의 울-왜곡 양자화 기법은 기존에 보고된 각종 양자화 방법에 비해 그 성능이 월등히 우수한 방법인데다가 여러 BER과 비트율 그리고 다양한 실험 영상을 통해 제안 시스템의 성능을 비교해 본 결과 제안 시스템의 성능이 기존의 다른 기법들에 비해 적지 않은 향상을 가져왔다는 사실을 확인할 수 있었다. 뿐만 아니라, 결과 영상으로 볼 때 PSNR 상의 개선도 있었지만, 실제 시각으

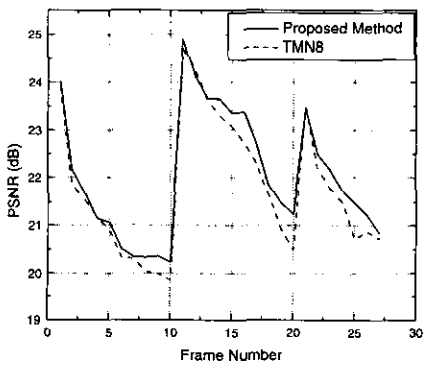
로 느끼는 성능의 개선도 확인함을 알 수 있었다. 특히 P 프레임의 경우는 이전 프레임의 오류가 시간적으로 전파되는 특성이 보이는데, 제안 기법은 이러한 단점을 억제하는 역할을 한다. 따라서 무선망을 통한 영상 통신과 같이 비트율의 제한으로 인해서 P 프레임을 많이 사용하게 되는 경우에 제안 시스템이 더욱 적합하다고 볼 수 있다



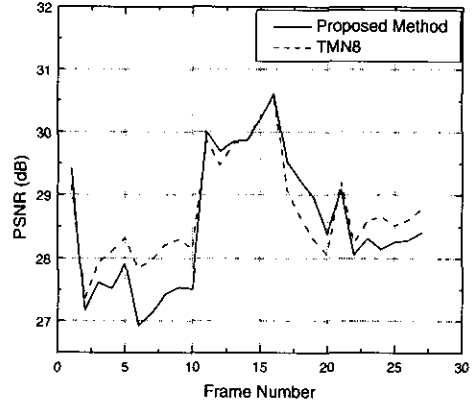
(a) BER  $1 \times 10^{-2}$



(b) BER  $5 \times 10^{-3}$

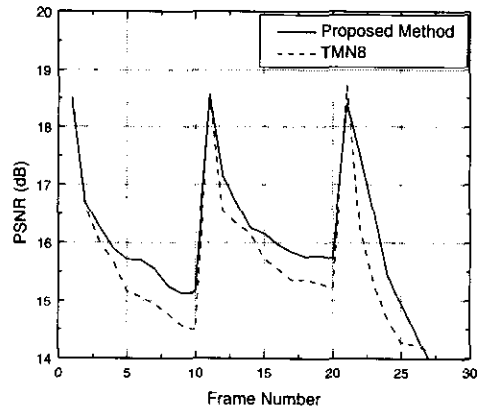


(c) BER  $1 \times 10^{-3}$

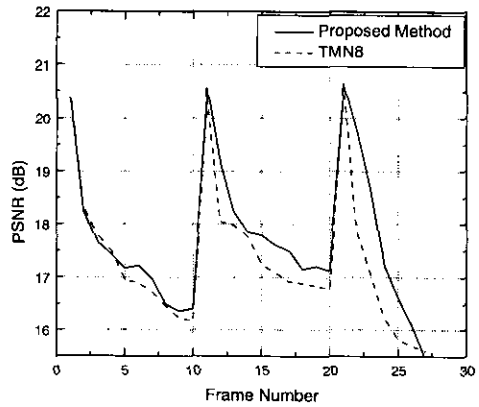


(d) BER  $1 \times 10^{-4}$

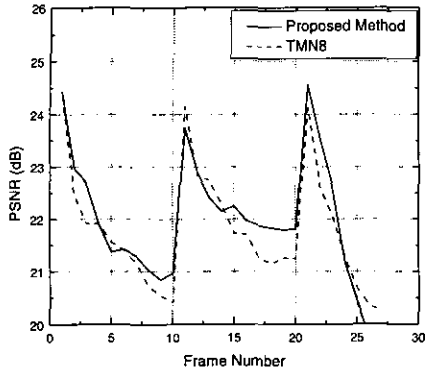
그림 5. 50kbps일 때의 BER에 따른 성능 비교



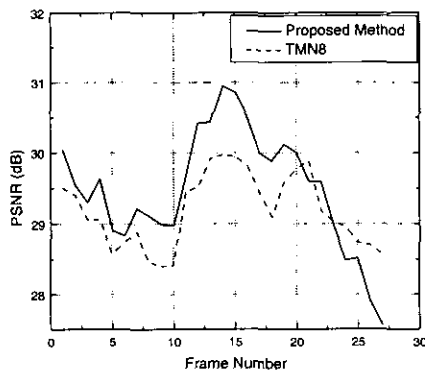
(a) BER  $1 \times 10^{-2}$



(b) BER  $5 \times 10^{-3}$



(c) BER  $1 \times 10^{-2}$



(d) BER  $5 \times 10^{-3}$

그림 6. 100kbps일 때의 BER에 따른 성능 비교



(a) Foreman 시퀀스 원 영상 #010



(b) TMN8



(c) 제안한 기법

그림 7. 비트율 100kbps, BER  $1 \times 10^{-2}$  에서의 결과 영상

### V. 결론

본 논문에서는 EREC 기반의 비트열을 위한 소스/채널 결합 최적 양자화를 설계하였고 이를 안정적 영상 전송 시스템에 적용하였다. 기존의 재동기 방식을 사용하는 영상 정보 통신 시스템에서는 발생한 비트 오류의 효과를 지나치게 과대 평가하여 그 화질 열화를 예측할 수 없었다. EREC는 발생한 채널 오류를 효과적으로 고립화함으로써 이러한 기존 방식의 한계를 극복하고, 발생 오류의 복원 영상에의 화질 열화를 효과적으로 분석할 수 있게 하였다. 따라서, 본 논문에서는 EREC를 기반으로 하는 시스템에서, 복호시 각 기본 부/복호화 블록(예를 들어 매크로 블록)들에 오류가 발생할 확률을 효과적으로 분석하였고, 이를 통해 복원 영상의 화질 열화를 예측하였다. 이는 기존의 오류를 수반하는 채널에서 양자화의 영향에 대한 연구를 보다 진보된 형태로 발전시킨 것으로, 예측된 화질 열화를 유효한 최적 양자화 문제에 전송 에러에 의한 왜곡 함수로 추가함으로써 이루어 졌다. 그 결과로 본 논문에서 제안하는 소스/채널 결합 양자화는 복호단에서의 복원 영상 화질이 주어진 비트율에서 최적일 수 있도록 한다.

본 논문에서 제안하는 양자화는 블록 기반 양자화를 수행하는 MPEG-1, 2, 4 혹은 H.261, 263 등의 다양한 압축 표준안에 적용이 가능하지만, H.263을 기준으로 제안 시스템의 성능을 비교 평가하였다. 비교의 대상이 되었던 양자화 기법은, 현재까지 알려진 rate control 기법들 중 그 성능이 가장 우수한 것으로 알려져 있는 TMN8의 양자화 기술이었으며, 제안 시스템과 동일하게 EREC를 적용한 후 복호 영상의 화질 및 발생 비트수를 비교하였다. 제안 시스템을 적용하여 부호화한 결과를 통해 기존 소스/채널 결합 양자화 연구에서 얻었던 결론, 즉, 채널 상황이 열악할수록 낮은 비트율을 발생시켜

부호화하는 것이 바람직하다는 사실을 다시 한번 확인할 수 있었다. 또한, AWGN 채널을 통해 전송한 후 복호된 영상의 품질 비교에서 제안 시스템의 기술이 주관적 및 객관적으로 보다 우수한 복원 영상을 얻을 수 있다는 것을 보였다.

### 부 록

식 (5)의 최적 해를 위한 필요, 충분 조건으로 식 (5)를  $Q_i$ 와  $\lambda$ 에 대해 각각 편미분 하여 그 결과를 0으로 놓음으로써 다음을 얻을 수 있다. (편의상 \*는 생략한다)

$$\frac{1}{N} a_i^2 \frac{Q_i^2}{12} = \lambda AK \frac{\sigma_i^2}{Q_i^2} \quad i=1,2,\dots,N \quad (A1)$$

$$\lambda \left[ \sum_{i=1}^N A \left( K \frac{\sigma_i^2}{Q_i^2} + C \right) - B \right] = 0 \quad (A2)$$

식 (A1)으로부터 다음을 얻는다.

$$(Q_i^2)^2 = \frac{\sigma_i^2}{a_i^2} \lambda AK \cdot 12N \quad \text{또는} \quad Q_i^2 = \frac{\sigma_i}{a_i} \sqrt{\lambda AK \cdot 12N}$$

따라서,

$$\begin{aligned} K \frac{\sigma_i^2}{Q_i^2} + C &= K \cdot \frac{a_i \times \sigma_i^2}{\sigma_i \sqrt{\lambda AK \cdot 12N}} + C \\ &= a_i \sigma_i \sqrt{\frac{K}{\lambda A 12N}} + C \end{aligned}$$

이고, 또한

$$\sum_{i=1}^N A \left( K \frac{\sigma_i^2}{Q_i^2} + C \right) = \sqrt{\frac{AK}{\lambda 12N}} \sum_{i=1}^N a_i \sigma_i + ANC$$

이다. 따라서, 식 (A2)는 다음과 같이 표현될 수 있고

$$\lambda \left[ \sum_{i=1}^N A \left( K \frac{\sigma_i^2}{Q_i^2} + C \right) - B \right] = \sqrt{\frac{\lambda AK}{12N}} \sum_{i=1}^N a_i \sigma_i + \lambda (ANC - B) = 0$$

이는 달리 다음과 같이 나타낼 수 있다.

$$\frac{\sqrt{AK}}{B-ANC} \sum_{i=1}^N a_i \sigma_i = \sqrt{\lambda \cdot 12N} \quad (A3)$$

반면, 식 (A1)에서  $i=k$ 인 경우를 생각해 보면

$$\lambda \cdot 12N = \frac{1}{AK} \cdot \frac{\alpha_k^2}{\sigma_k^2} Q_k^4 \quad \text{또는}$$

$$\sqrt{\lambda \cdot 12N} = \frac{1}{\sqrt{AK}} \cdot \frac{\alpha_k}{\sigma_k} Q_k^2$$

이고, 그러므로 식 (A3)으로부터

$$\frac{\sqrt{AK}}{B-ANC} \sum_{i=1}^N a_i \sigma_i = \frac{1}{\sqrt{AK}} \cdot \frac{\alpha_k}{\sigma_k} Q_k^2 \quad \text{또는}$$

$$Q_k = \sqrt{\frac{AK}{B-ANC} \cdot \frac{\sigma_k}{\alpha_k} \sum_{i=1}^N a_i \sigma_i}$$

### 참 고 문 헌

- [1] J. W. Modestino and D. G. Daut, "Combined Source-Channel Coding of Images," *IEEE Trans. Commun.*, vol. com-27, pp. 1644-1659, Nov. 1979.
- [2] J. W. Modestino, D. G. Daut and A. L. Vickers, "Combined Source-Channel Coding of Images Using the Block Cosine Transform," *IEEE Trans. Commun.*, vol. com-29, pp. 1261-1274, Sep. 1981.
- [3] N. Farvardin and V. Vaishampayan, "Optimal Quantizer Design for Noisy Channels : An Approach to Combined Source-Channel Coding," *IEEE Trans. Inform. Theory*, vol. 33, pp. 827-838, Nov. 1987.
- [4] D. W. Redmill and N. G. Kingsbury, "The EREC : An Error-Resilient Technique for Coding Variable-Length Blocks of Data," *IEEE Trans. Image Processing*, vol. 5, pp. 565-574, Apr. 1996.
- [5] R. Chandramouli, N. Ranganathan, and S. J. Ramadoss, "Adaptive Quantization and Fast Error-Resilient Entropy Coding for Image Transmission," *IEEE Trans. Circ. Syst. Video Technol.*, vol. 8, pp. 411-421, Aug. 1998.
- [6] N. T. Cheng and N. G. Kingsbury, "The ERPC : An efficient error-resilient technique for encoding positional information on sparse data," *IEEE Trans. Commun.*, vol. 40, pp. 140-148, Jan. 1992.
- [7] ITU-T DRAFT H.263, "Video Coding for Low Bit Rate Communication," Sep. 1997.
- [8] H. Jung, R. Kim and S. Lee, "On the Robust Transmission Technique for H.263 Video Data

Stream Over Wireless Networks," in *Proc. of the ICIP*, vol. 3, pp. 463-466, 1998.

[9] J. Ribas-Corbera and S. Lei, "Rate control in DCT video coding for low-delay communications", *IEEE Trans. CSVT*, vol.9, No. 1, pp.172-185, 1999.

[10] J. Ribas-Corbera and D. Neuhoff, "On the optimal motion vector accuracy for block-based motion-compensated video coders," in *Proc. IS&T/SPIE Dig. Video Comp.: Alg. & Tech.*, San Jose, CA, Jan.-Feb. 1996, pp. 302-314.

[11] J. Ribas-Corbera and D. Neuhoff, "Optimizing block size in motion-compensated video coding," *J. Electron. Imaging*, vol. 7, pp. 155-165, Jan. 1998.

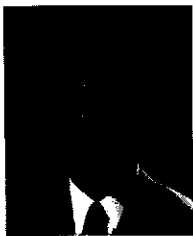
[12] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery, "*Numerical Recipes in C : The Art of Scientific Computing*," 2nd. Edition, Cambridge University Press, 1992.

[13] B. Sklar, "*Digital Communications, Fundamentals and Applications*," Prentice-Hall, 1988

[14] S. Whitehouse and N. Kungsbury, "Enhancements to the Error Resilient Entropy Code," *Multimedia Signal Processing*, 1998 IEEE Second Workshop, pp. 456-461, 1998.

[15] D. W. Redmill, D. R. Bull, J. Chung-How, N. G. Kingsbury, "Error-Resilient image and video coding for wireless communication systems," *Electronics & Communication Engineering Journal*, Vol. 104, pp. 181-190, Aug. 1998.

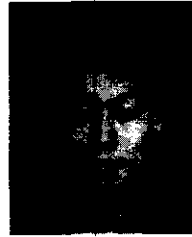
김 옹 구(Yong-Goo Kim)



1989~1993 : 연세대학교  
전기공학과 (학사)  
1993~1995 : 연세대학교  
전기공학과 (석사)  
1995~현재 : 연세대학교 전기  
및 컴퓨터 공학과  
(박사과정)

<주관심 분야> 멀티미디어 영상통신, Source/Channel 결합 부호화, Wavelet, Fractal, Error Resilient Video, Error Concealment, 통계신호처리

송 진 규(Jin-Gyu Song)



1994~1998 : 연세대학교  
전기공학과 (학사)  
1998~2000 : 연세대학교  
전기공학과 (학사)  
2000~현재 : LG 정보통신  
CDMA 단말연구소  
연구원

<주관심 분야> 멀티미디어 영상통신, Source/Channel 결합 부호화, 최적 양자화 시스템

최 윤 식(Yoonsik Choe)



1975~1979 : 연세대학교  
전기공학과 (학사)  
1981~1984 : Case Western  
Reserve Univ. (MSEE)  
1984~1987 : Penn. State Univ.  
(MS)

1987~1990 : Purdue Univ. (Ph.D)  
1990~1993 : 현대전자산업 산업전자연구소 책임연구원  
1993~현재 : 연세대학교 기계전자공학부 교수  
<주관심 분야> 영상 부호화, 차세대영상압축, 영상통신, 디지털 비디오 신호처리, 통계적 신호처리, 멀티미디어 시스템 등