

# 인간시각특성을 이용한 블록기반 DCT 영상 부호화기의 정량적 화질 평가

## Quantitative Image Quality Assessment for Block-based DCT Image Coder using Human Visual Characteristics

정태윤  
Tae-Yun Chung

강릉대학교 정보전자공학부

### 요 약

본 논문은 블록기반 DCT 부호화기의 성능 검증에 필수적인 부호화된 영상의 열화 정도를 측정하는 새로운 정량적 화질평가 모델은 제안한다. 제안된 모델은 HVS 기반의 다채널 시각 모델을 이용하여 영상신호에 있어서의 주파수 민감도와 채널의 마스킹 효과를 고려한 공간 영역상의 전반적인 왜곡의 가시도 뿐만 아니라 블록기반 DCT 부호화에서 발생할 수 있는 구획화 현상, 번짐, ringing 등의 국부적인 왜곡 특성을 측정하여 이를 화질 열화 특성에 반영하였다.

### ABSTRACT

This paper proposes a new quantitative image assessment model which is essential to verify the performance of block-based DCT coding. The proposed model considers not only global distortions such as frequency sensitivity and channel masking using HVS based visual model, but also distortions including several local distortions caused by block-based coding.

**Key Words** : 화질 평가, 인간시각 시스템, 영상 부호화

### 1. 서 론

디지털 정보화 사회라는 사회, 경제, 기술적 변화의 한 축에는 시각적인 정보를 제공하는 영상정보 서비스가 자리잡고 있다. 영상정보의 디지털화는 화질, 신뢰도, 보안성 그리고 정보 처리의 편이성에 있어 아날로그 방식에 비해 월등히 우수한 성질을 갖고 있으나 그 정보량의 방대한 때문에 전송 및 저장매체의 효율적인 사용을 위해 정보 압축 부호화가 필수적이다[1]. 영상정보 압축 부호화 기술은 응용에 따라 JPEG, MPEG 등 많은 국제 표준안들이 마련되었으며 이러한 표준 규격을 채용하는 시스템들의 사용이 더욱 확산될 전망이다.

영상정보의 손실 압축 부호화 기술은 사용되는 매체의 용량과 화질이라는 trade-off 관계를 고려하여 설계되며 정보량 감축에 따른 화질 열화가 불가피하다. 따라서 부호화기 설계에 있어 가장 중요한 문제는 원영상(original image)의 어떤 정보를 어느 정도로 왜곡시키는 것이 복호화된 영상의 화질 열화를 최소화 할 수 있는가 하는 것이며 이를 위한 정확한 화질측정 방법은 부호화기 성능 향상에 있어 가장 중요한 문제로 인식되어왔다[2].

영상정보의 화질평가 방법으로는 크게 인간이 직접 왜곡 영상을 평가하는 주관적 방법과 정량적 측정도구로 왜곡영상

의 화질을 평가하는 객관적 방법으로 나눌 수 있다[3]. 주관적 화질평가 방법은 화질에 대한 최종적인 평가는 결국 인간에 의해 이루어진다는 점에서 그 신뢰도가 높은 반면 평가환경 구축시의 비용 상승 및 장시간의 평가 소모 시간 등에 문제점이 있다. 또한 현재 객관적 화질평가 방법으로 널리 사용되는 MSE(mean square error) 및 PSNR (peak signal to noise ration) 측정법은 아날로그 영상의 화질 측정에는 비교적 적합하나 영상의 왜곡특성이 아날로그와는 다른 디지털 압축 부호화된 영상의 화질평가에는 적절하지 않다[4]. 따라서 최근에 주관적 화질평가 방법과 대등한 성능을 보일 수 있는 객관적, 정량적 화질평가 방법에 대한 연구가 활발히 진행되고 있다[5].

본 논문에서는 대부분의 영상정보 압축 부호화 표준안에서 채용하고 있는 블록기반 DCT 기법을 근간으로 하는 영상 부호화기의 성능 향상 및 검증에 필수적인 정량적이며 객관적인 화질평가 모델을 제안한다. 제안된 모델은 HVS (human visual system)의 다채널(multi-channel) 모델을 신호처리 기법으로 해석하여 영상 부호화시 발생하는 전반적 왜곡영상 특성을 분석하고 이를 바탕으로 블록기반 DCT 부호화기에서 발생하는 구획화 현상, 번짐, ringing 등의 국부적 화질열화 특성을 정량적 화질평가에 반영한다. 제안된 알고리즘의 성능은 JPEG으로 부호화된 정지영상을 대상으로 한 실험을 통해 주관적 화질평가와 비교함으로써 타당성을 입증하였다.

접수일자 : 2002년 2월 25일

완료일자 : 2002년 7월 4일

## 2. 인간시각 특성

영상정보 압축 부호화에 의해 왜곡된 화질을 객관적으로 평가하기 위해서는 눈의 생리학적 구조 및 그에 따른 시각 현상을 이용한 HVS를 모델링하는 것이 필수적이다. HVS를 모델링하기 위해 주로 사용되는 방법은 정신 물리학(psychophysics)적 실험을 통해 서로 다른 배경 영상위에 어떤 시각적 자극(stimulus)이 가해졌을 때 눈이 이를 감지하기 시작한 자극의 크기를 측정하여 이것을 정량화하는 것이다[6]. 이때 자극이 시각적으로 관측 가능한 상태가 되는 크기를 가시도 문턱값(visibility threshold)라 하며 여기서 얻어진 결과들은 공간 영역 및 주파수 영역에서 해석된다.

### 2.1 다채널 구조

HVS를 이용한 객관적 화질평가 모델의 목적은 인간의 눈으로부터 시각경(visual cortex)까지의 시각 시스템의 처리 과정을 신호처리 기법으로 표현하여 이를 원영상에 대한 왜곡영상의 상대적 화질열화 정도의 측정에 활용하는 것이다.

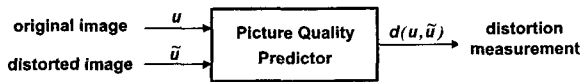


그림 1. 화질평가 모델  
Fig. 1. Image quality assessment model

그림 1은 화질평가 모델을 간단히 나타낸 것이다. 만일  $u$ ,  $\tilde{u}$ 가 각각 원영상과 왜곡영상이면 모델에 의해 평가되는 화질 측정치  $d(u, \tilde{u})$ 를 통해 인간에 의한 주관적 화질평가와 동일한 성능을 나타내고자 하는 것이다.

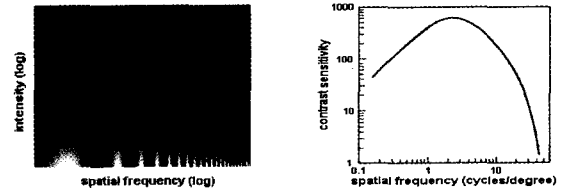
지금까지 연구되어온 시각 모델은 크게 단일채널 모델과 다채널 모델로 나눌 수 있다. Mannos 등은 MSE에 기반을 둔 단순한 왜곡 추정으로는 인간의 눈에 의해 인지되는 화질을 평가할 수 없다는 인식하에 처음으로 화질 평가 모델로 단일채널 모델을 제시하였다[7]. 이러한 단일채널 모델은 단일패턴에 의한 실험을 통해 HVS의 특성을 간단히 평가할 수는 있으나 많은 경우에 있어 인간의 복잡한 시각특성을 평가할 수는 없었다. 또한 Sakrison 등은 정신 물리학적 실험을 통해 인간의 시각 인지특성은 다해상도로 구성되어 있다는 것을 규명함으로써 다채널 모델을 제시하였다[8]. 따라서 대부분의 화질평가를 위한 시각모델은 좀 더 복잡한 자연영상의 패턴의 가시도 평가를 위해 채널의 상호 작용 및 마스크(masking)효과를 고려한 다채널 모델을 사용한다.

### 2.2 대비 민감도

인간의 눈은 밝기의 세기 그 자체보다는 밝기의 차이에 더욱 민감하다. 즉, 인간의 눈은 밝기의 변화를 인지하는데 있어 비선형 특성을 보인다. 이러한 영상의 대비(contrast) 특성은 그 응용에 따라 수 많은 방법으로 정의되고 있으나 정현파 격자(sinusoidal grating)와 같은 주기패턴으로 대비 특성을 측정하는 Michelson 대비와 균일한 밝기를 갖는 배경에 대한 균일자극의 변화량을 측정하는 Weber 대비등의 2가지 방법이 주로 사용된다.

이러한 대비를 공간 주파수에 대한 시각 민감도로 표현하는 것이 MTF(modulation transfer function)이다. Campbell 등은 공간 주파수 변화에 따른 정현파 격자의 가시도 문턱값을 관찰함으로써 MTF를 측정하였다[9]. 이때 대비 민감도

(CSF, contrast sensitivity function)는 가시도 문턱값의 역으로 정의된다. 그림 2는 공간 주파수 함수로서의 CSF를 나타낸 것이다. CSF는 대역통과필터의 특성을 보이며 따라서 동일한 양의 잡음성분에 대해서도 잡음의 주파수 분포에 따라 실제 시각특성은 다르게 나타난다.



a) 정현파 격자영상

b) CSF

그림 2. 정현파 격자 영상에 대한 CSF

Fig. 2. CSF for sinusoidal grating image

### 2.3 마스크 효과

대부분의 자연영상은 배경의 밝기가 불균일하고 복잡하다. 따라서 공간적으로 밝기의 큰 변화가 있는 영역에서의 자극에 대한 가시도를 측정하는 것이 필요하다. Fiorentini 등은 밝기의 변화에 대한 가시도 문턱값을 측정하여 밝기의 변화가 큰 인접부분은 자극에 대한 가시도의 민감도가 낮아져서 자극을 인식하기 어렵다는 것을 밝혀냈다[10]. 이러한 현상을 공간 마스크이라 한다. 즉, 같은 양의 잡음이라도 주변 및 배경의 밝기의 변화 및 주파수 분포에 따라 인간시각이 인지하는 왜곡량은 다르게 나타날 수 있다.

## 3. 블록기반 DCT 부호화기의 화질열화

부호화된 영상의 화질열화 특성의 고찰은 객관적 화질평가 뿐만 아니라 부호화기 설계에 있어서도 매우 중요하다. 즉, 화질열화 특성에 대한 정확한 분석은 부호화기내의 양자화기 설계에 반영되어 부호화기 효율 및 화질을 향상시킬 수 있다.

현재 가장 많은 응용 분야를 갖고 있는 영상 부호화 표준안은 JPEG과 MPEG이다. JPEG과 MPEG은 시스템 구현시의 복잡도와 화질을 고려하여 8x8 블록(block) 단위의 DCT를 근간으로, 공간적, 시간적인 영상의 잉여정보를 제거한다. 따라서 블록기반의 DCT 부호화기를 채용하는 JPEG, MPEG의 경우, 시각적으로 여러 형태로 나타나는 개별적인 화질열화 특성의 분석을 통한 통합적인 열화특성의 제시는 객관적, 정량적 화질평가에 필수적이다[11]. 그림 3은 블록기반 DCT 부호화에 의해 발생하는 화질열화 특성을 나타낸 것이다.

### 3.1 구획화(blocking) 현상

구획화 현상은 인접 블록과의 경계에서 나타나는 불연속으로 정의되며 블록기반 DCT 부호화에 의한 가장 시각적으로 두드러진 화질열화 특성이다. 즉, 이러한 현상은 인접 블록과의 상관성은 고려하지 않고 해당 블록을 독립적으로 부호화함으로써 발생한다.

DCT 계수 중 DC와 AC 계수의 양자화 방법이 다르며 따라서 영상의 화소값이 수평 또는 수직방향으로 단조 증가 또는 감소하고 또한 양자화 스텝 사이즈가 크다면 DC 계수를 제외한 나머지 영역의 양자화 오차가 커지게 되어 복원된 영

상은 8화소 단위로 불연속하게 나타난다.

### 3.2 번짐(blurring) 현상

번짐 현상은 texture 영역 또는 물체의 에지(edge) 주변과 같은 공간적으로 높은 활성도(activity)를 갖는 영역에서 발생하며 에지의 예리함(sharpness)의 감소 및 공간적인 손실로 나타난다.

즉, 번짐은 성긴(coarse) 양자화에 의한 고주파 영역 AC 계수의 억압과 직접적으로 관련되며 따라서 블록내의 영상은 저주파 영역의 계수에 의해 표현된다. 결국 이 현상은 저역 통과 필터로 나타나는 화질열화의 특성을 보인다.

### 3.3 Ringing 현상

Ringing 현상은 영상의 비교적 균일한 영역에서 에지와 같은 영상의 경계면 주위에 어른거리거나 물결(ripple) 형태와 같은 주기적인 의사(pseudo) 에지로 나타난다. 이 현상은 DCT 계수중 고주파 성분의 잘림(truncation)에 기인하며 에지의 대비값이 커질수록 물결형태의 첨두(peak)값이 커지고 골(trough)이 깊어지는 Gibbs 현상과 관련이 있다.

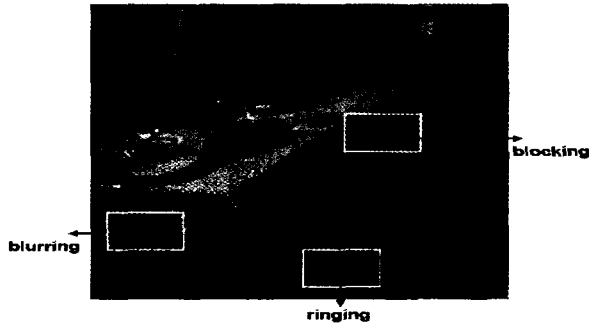


그림 3. 블록기반 DCT 부호화의 화질열화 특성  
Fig. 3. Coding artifacts of block-based DCT

## 4. 제안된 정량적 화질평가 모델

기존의 객관적 화질평가 모델의 대부분은 시공간적인 주파수 민감도 및 왜곡의 가시도에 대한 시각적 마스킹만을 고려하여 이를 원영상과 왜곡영상의 차영상에 대해 가중치로 반영함으로써 시각적 왜곡에 대한 정량적인 화질평가를 시도하였다. 그러나 이는 앞 절에서 언급한 구획화 현상, 번짐 및

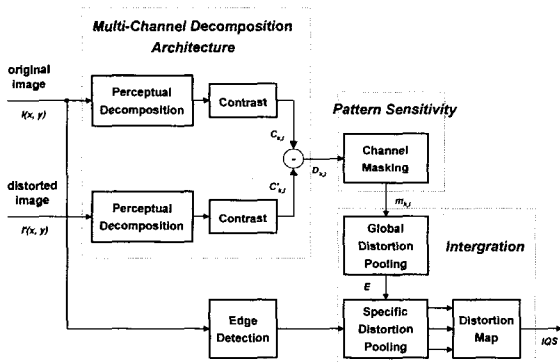


그림 4. 제안된 정량적 화질평가 모델

Fig. 4. Proposed quantitative quality assessment model

ringing 등 다양한 왜곡특성을 보이는 블록기반 DCT 부호화를 채용하는 영상 부호화기의 정량적이며 객관적인 화질평가 방법으로는 불충분하다.

따라서 본 논문에서는 HVS 기반의 주파수 민감도와 다채널 구조에 따른 시각적 마스킹에 의한 부호화된 영상의 전반적 왜곡의 가시도를 측정하고 이를 기반으로 영상 부호화시 발생하는 국부적인 왜곡을 고려함으로써 블록기반 DCT 부호화기에 적합한 정량적 화질평가 모델을 제안한다. 그림 4는 제안된 정량적 화질평가 모델의 블록도를 나타낸 것이다.

### 4.1 인지채널 분해

인간의 눈의 원시 시각영(V1 영역)에서의 세포구조는 전자 생리학적인 실험을 통해 Gabor 형태의 수용계 구조를 갖는 것으로 밝혀졌다[12]. 즉, 시각영 수용계는 Gabor 형태의 비원형 구조로서 방위(orientation)와 주파수에 대해 선택적 민감성을 갖는 흥분성과 억제성 영역으로 구성된다. 따라서 기존의 대부분의 다채널 모델에서는 원영상과 부호화에 의한 왜곡영상에 대해 시각 시스템의 방위와 주파수 성분에 대한 다채널 구성을 위한 인지채널 분해(perceptual channel decomposition)를 위해 Gabor 필터를 사용하였다. 그러나 Gabor 필터는 스펙트럼 영역상에서 매우 선택적(selective)이며 또한 필터링이 주파수 영역에서 수행되어야 하는 단점을 안고 있다. 이것은 Gabor 프로파일들이 FFT 연산을 필요로 하며 또한 과중한 메모리 사용을 요구한다. 그러나 정량적 화질평가에 있어 정확한 필터의 형태는 필요하지 않으며 시스템 구현시의 복잡도와 평가 정밀도의 trade-off 관계를 고려해야 한다[13]. 따라서 본 논문에서는 steerable 피라미드 변환에 의한 인지채널 분해를 시행한다.

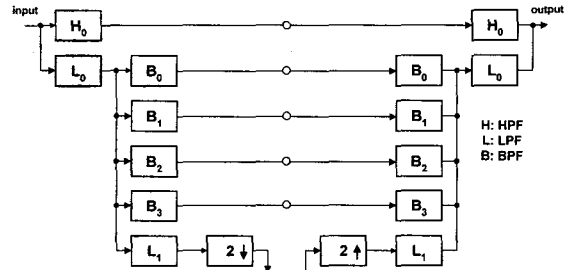


그림 5. Steerable 피라미드의 블록도  
Fig. 5. Block diagram of steerable pyramid

그림 5는 steerable 피라미드 변환의 블록도를 나타낸 것이다. steerable 피라미드 변환은 decimation과 컨볼루션을 이용하여 영상을 octave 대역의 부대역(subband)으로 구성된 공간주파수 영역으로 분해하고 각 부대역은 여러 개의 방위 대역으로 나눈다. Steerable 피라미드 변환의 장점은 자기 역변환(self-inverting)이 가능하며 동일 부대역 레벨에서 나누어진 방위 대역의 선형조합을 통해 어떠한 방위 성분의 주파수 영역도 표현이 가능한 회전 불변(rotation-invariant) 구조라는 것이다[14]. 따라서 이러한 steerable 피라미드 변환 필터를 설계하기 위해서는 다음의 조건을 만족시켜야 한다.

$$\hat{X}(w) = \{ |H_0(w)|^2 + |L_0(w)|^2 ( |L_1(w)|^2 + \sum_{k=0}^3 |B_k(w)|^2 ) \} X(w) + a.t. \quad (1)$$

여기서  $X(w)$ 와  $\hat{X}(w)$ 와 각각 입력 영상과 복원된 영상

을 나타내며  $H_0(w)$ 는 고역통과 필터를,  $L_0(w)$ 와  $L_1(w)$ 는 각 레벨에서의 저역통과 필터를, 그리고  $B_k(w)$ 는 대역통과 방위필터를 나타내며  $n$ 은 각 부대역에서의 방위성분의 수를 표현한 것이다. 식 (1)에서  $a.t.$ 는 aliasing 항을 표현한 것으로 대역통과 필터의 aliasing을 최소화 하기 위해서는 식 (2)와 같이  $L_1$  필터는  $\pi/2$  이상의 주파수에서는 0의 응답특성을 가져야 한다.

$$L_1(w) = 0, \text{ for } |w| > \pi/2 \quad (2)$$

또한 진폭왜곡을 방지하기 위해 시스템 전달함수는 식 (3)과 같이 단위 응답 진폭(unity response amplitude)의 특성을 보여야 하며 각 부대역의 저역통과 필터는 시스템의 순환적 할당에 영향을 받지 않도록 식 (4)의 조건을 만족시켜야 한다.

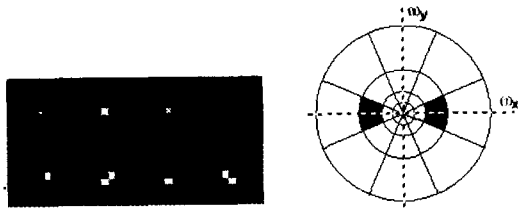
$$|L_0(w)|^2 \left[ |L_1(w)|^2 + \sum_{k=0}^n |B_k(w)|^2 \right] + |H_0(w)|^2 = 1 \quad (3)$$

$$|L_1(w/2)|^2 \left[ |L_1(w)|^2 + \sum_{k=0}^n |B_k(w)|^2 \right] = |L_1(w/2)|^2 \quad (4)$$

또한 대역통과 필터  $B_k(w)$ 는 식 (5)의 회전불변 특성을 만족해야 한다.

$$B_k(w) = B(w) [-j \cos(\theta - \theta_k)]^n \quad (5)$$

여기서  $\theta = \arg(w)$ ,  $\theta_k = \pi k / (n+1)$ 이다.



a) 필터 커널                      b) 주파수 응답특성  
(상단 : L1, L0, H0 커널    하단 : Bk의 방위 커널)

그림 6. Steerable 피라미드의 필터특성

Fig. 6. Filter characteristics of steerable pyramid

본 논문에서는 식 (1)에서 식 (5)의 조건을 만족하는 그림 6의 필터 커널을 이용하여 4개의 부대역과 각 부대역 당 4개의 방위성분으로 구성된 steerable 피라미드 변환을 통해 인지채널 분해를 시행하였다.

#### 4.2 대비 측정

인간의 눈은 밝기의 세기보다는 밝기의 변환에 민감하다. 따라서 steerable 피라미드 변환에 의해 다채널의 시각적 인지채널로 분해된 영상신호의 화소값은 HVS의 전자 생리학적 실험결과를 반영하기 위해 대비값으로 변환되어야 한다. 앞 장에서 언급했듯이 Mechelson 대비법 및 Weber 대비법은 정현파 격자와 같은 단순한 주기패턴을 바탕으로 한 방법이며 따라서 통계적 성질이 다양하고 복잡한 자연영상의 대비측정에는 적합하지 않다. Peli는 이러한 문제를 해결하기 위해 자연영상의 대비측정으로 식 (6)의 LBC(Local band-limited contrast)법을 제안하였다[12].

$$c_k(x, y) = \frac{bp_k(x, y)}{lp_{k+1}(x, y)} \quad (6)$$

여기서  $c_k$ 는 인지채널  $k$ 에서의 국부 대비값이며 이때  $x$ ,  $y$ 는 각각 영상의 수평, 수직의 화소 위치를, 그리고  $bp_k$ 와  $lp_{k+1}$ 는 각각 인지채널  $k$ 의 화소값 및 채널  $k$ 이하 대역의 저역통과 필터링된 화소값을 나타낸다. 그러나 이 방법 또한  $lp_{k+1}$ 가 0이 될 경우가 발생할 수 있기 때문에 자연영상의 대비 측정에는 결함이 있다. 따라서 본 논문에서는 LBC 법을 수정하여 공간주파수 부대역 채널  $k$ 와 방위의 채널대역  $l$ 에서의 대비를 식 (7)을 통해 측정한다.

$$c_{k,l}(x, y) = \left| \frac{bp_{k,l}(x, y)}{lp_{k+1}(x, y) + b_{mean} + 1} \right| \quad (7)$$

여기서  $b_{mean}$ 은 입력영상의 평균값이고 분모에 1을 더함으로써 0으로 나누어지는 것을 방지한다. 따라서 식 (7)을 통해 각 채널의 대비값은 0~1의 값을 갖는다.

#### 4.3 공간 마스킹에 의한 패턴 민감도 측정

부호화 영상의 왜곡의 가시도는 결국 원영상과 부호화 영상의 차영상으로 표현되는 오차의 가시도이며 이는 HVS에 따른 배경 영상의 밝기 변화 및 공간적 구조, 그리고 오차신호의 주파수 분포등에 영향을 받는다. 따라서 식 (8)과 같이 원영상의 대비값  $c_{k,l}(x, y)$ 와 부호화 영상의 대비값  $c'_{k,l}(x, y)$ 의 차영상  $D_{k,l}(x, y)$ 에 대한 왜곡의 가시도를 정량화하는 것이 필요하며 이를 위해 마스킹에 의한 패턴 민감도를 측정한다.

$$D_{k,l}(x, y) = |c_{k,l}(x, y) - c'_{k,l}(x, y)| \quad (8)$$

본 논문에서는 원영상과 식 (8)의 차영상을 각각 마스킹 패턴(배경영상) 및 왜곡의 자극 패턴으로 간주하여 공간 주파수상의 CSF를 고려한 문턱값 상승을 표현하는 비선형 변환 함수로부터 원영상에 의한 차영상의 마스킹 효과를 반영함으로써 패턴 민감도를 측정한다. 식 (9)는 차영상에 대한 마스킹 문턱값을 나타낸 것이다.

$$C_{k,l}^T(x, y) = \begin{cases} C_{T0} & \text{if } c_{k,l}(x, y) < C_{T0} \\ C_{T0} \left( \frac{c_{k,l}(x, y)}{C_{T0}} \right)^\epsilon & \text{otherwise} \end{cases} \quad (9)$$

여기서  $C_{T0} = c_0 / k_{CSF}$ 이며 본 실험에서는 동일한 대역에서 패턴 민감도를 측정하므로  $k_{CSF} = 1$ 이며  $c_0$ 와  $\epsilon$ 는 Legge[15]의 실험값을 바탕으로 각각 0.02 및 0.8을 사용하였다. 그림 7은 공간 마스킹을 위한 식 (9)의 비선형 함수를 도시한 것이다.

따라서 원영상의 마스킹에 의한 부호화 영상의 왜곡 패턴 민감도는 식 (10)을 통해 구해진다.

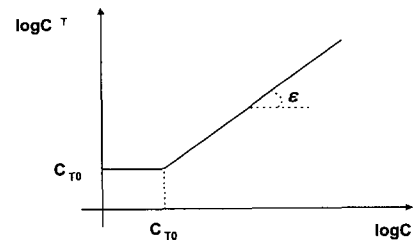


그림 7. 마스킹 문턱값 함수

Fig. 7. Masking threshold function

$$m_{k,i}(x, y) = \frac{D_{k,i}(x, y)}{C_{k,i}^T(x, y)} \quad (10)$$

#### 4.4 왜곡 가시도의 규합(pooling)

공간주파수와 방위성분의 각 인지채널로부터 CSF 및 마스킹 효과가 고려된 왜곡의 가시도는 공간영역상의 왜곡량에 대한 최종적인 인지를 위해 규합된다. 공간영역상의 전반적(global) 왜곡의 가시도  $E(x, y)$ 는 식 (11)과 같이 왜곡량에 대한 가중치를 부여하기 위해 Minkowski 합을 통해 측정된다.

$$E(x, y) = \left\{ \sum_{k=1}^{N_s} \sum_{i=1}^{N_o} [m_{k,i}(x, y)]^\beta \right\}^{1/\beta} \quad (11)$$

여기서  $N_s$ 와  $N_o$ 는 각각 공간 부대역 및 방위 성분의 채널 수이고  $\beta$ 는 Minkowski 합을 위한 지수상수로서 본 실험에서는 2를 사용하였다.

#### 4.5 블록기반 DCT 부호화기의 화질열화 측정

블록기반 DCT를 채용하는 영상 압축 부호화 과정에서 발생하는 화질열화의 특성은 매우 다양하며 따라서 공간영역상의 전반적인 왜곡의 가시도 측정만으로는 정량적이며 객관적인 화질평가에 불충분하다. 따라서 본 논문에서는 인간의 눈에 민감한 왜곡특성인 구획화 현상, 번짐, ringing등의 국부적인 화질열화 요인을 측정하여 이를 정량적 화질평가 측정에 반영한다.

##### (1) 구획화 현상의 열화 측정

구획화 현상은 인접 블록과의 상관성을 고려하지 않고 해당 블록만을 독립적으로 부호화하는 DCT등의 블록기반 부호화 과정에서 발생하는 화질열화의 요인으로 블록경계면을 따라 인접 블록과의 불연속으로 나타난다. 따라서 공간영역상의 왜곡의 가시도  $E(x, y)$ 로부터 식 (12)을 이용하여 각각 수평, 수직 블록의 경계면에서 발생하는 불연속의 양  $B_h$ 와  $B_v$ 를 측정 후 식(13)을 통해 구획화 현상의 화질열화 성분  $L_{block}$ 을 정량적으로 평가한다.

$$B_k(x, y) = \sum_{x,y} \{E(x, y)_{\in m,n} - E(x, y+1)_{\in m,n}\}^2 \quad (12)$$

여기서  $m, n$ 은 각각 수평, 수직의 블록경계면의 화소위치치를 나타내며 이때  $B_v$ 도 동일한 방법으로 구한다.

$$L_{block}(x, y) = \sqrt{B_h(x, y)^2 + B_v(x, y)^2} \quad (13)$$

##### (2) 번짐의 열화 측정

번짐 현상은 물체의 에지 주변 또는 texture 영역과 같이 공간적으로 높은 활성도를 갖는 영역에서 생긴 양자화에 의한 고주파 성분의 억압에 의해 발생하는 화질열화 특성이다. 따라서 번짐 현상은 가시적으로는 저역통과 필터의 현상으로 나타나며 결국 화소간의 상관관계의 감소를 초래한다. 따라서  $E(x, y)$ 로부터 식 (14) 및 (15)를 이용하여 화소간의 국부적인 상관도  $L_{blur}(x, y)$ 를 측정한다.

$$B_r(i, j) = \frac{1}{N_w - 1} \left\{ \sum_{i,j,m,n \in w} E(i, j)E(i+m, j+n) - \frac{1}{N_w} \sum_{i,j,m,n \in w} E(i, j) \sum_{i,j,m,n \in w} E(i+m, j+n) \right\} \quad (14)$$

여기서  $w$ 는 수평, 수직의 화소위치  $x, y$ 에서의  $5 \times 5$  창(window)크기이며  $N_w$ 는 창 내부의 화소의 수를 의미한다.

$$L_{blur}(x, y) = \sum_{i,j \in w} |B_r(i, j)|^\gamma \quad (15)$$

여기서  $\gamma$ 는 왜곡의 상대적인 크기를 감쇄하기 위해 지수상수이며 본 실험에서는 0.5를 사용하였다.

##### (3) Ringing의 열화 측정

Ringing 현상은 비교적 균일한 배경영상에서 에지와 같이 주변부와의 대비가 큰 영역의 고주파 성분의 잘림에 의해 나타나는 화질열화 특성으로 가시적으로 물체의 경계면 주변등에 어린거리는 주기적인 의사 경계면으로 보이게 된다. 따라서 원영상으로부터 물체의 경계면을 표현하는 에지를 검출하여 공간영역상의 왜곡의 가시도  $E(x, y)$ 로부터 에지 주변부의 국부적인 활성도를 측정함으로써 ringing의 화질열화 특성을 정량적으로 평가한다. 먼저 원영상에서의 에지 검출은 식 (16)과 같이 Sobel 연산자를 사용한다.

$$S(x, y) = \begin{cases} Sobel[I(x, y)] & \text{if } sobel[I(x, y)] \geq Th \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

여기서  $I(x, y)$ 는 원영상이며  $Th$ 는 경계면 검출을 위한 문턱값이다. 이와 같이 얻어진 물체의 경계면 영상에 대해 각각 수평, 수직방향의 국부 활성도  $A_h(x, y)$ ,  $A_v(x, y)$ 를 식 (17)로부터 구한 후

$$A_k(x, y) = \frac{|I(x, y-1) - I(x, y+1)|}{2} \quad (17)$$

식 (18)에서 구한 각각 수평, 수직의 마스킹 함수  $M_h(x, y)$ 와  $M_v(x, y)$ 를 통해 식 (19)의 각 화소당 ringing 현상에 의한 마스킹 오차  $L_{ringing}(x, y)$ 를 구한다.

$$M_k(x, y) = e^{-a \cdot A_k(x, y)} \quad (18)$$

여기서  $a$ 는 마스킹함수의 지수감쇄 상수로서 본 실험에서는 0.2를 사용하였으며  $A_v(x, y)$  및  $M_v(x, y)$ 도 동일한 방법으로 구한다.

$$L_{ringing}(x, y) = E(x, y) \{M_h(x, y) + M_v(x, y)\} \quad (18)$$

#### 4.6 최종적인 화질평가

주관적인 화질평가 방법은 ITU-R 권고안 500을 통해 평가 조건, 테스트 영상 선정 및 평가 절차와 분석 방법등의 기준이 마련되었으며 일반적으로 DSCQS(double stimulus continuous scale)법, DSIS(double stimulus impairment scale)법 등이 사용되고 있다[16]. 이러한 평가법들은 표 1에 나타낸 5가지 등급 척도를 통해 왜곡 영상의 화질을 평가한다.

표 1. 주관적 화질평가 등급 척도

Table 1. Rating scale for subjective quality assessment

등급	열화	화질
5	Imperceptible	Excellent
4	Perceptible, not annoying	Good
3	Slightly annoying	Fair
2	Annoying	Poor
1	Very annoying	Bad

따라서 본 논문의 정량적 화질평가 결과는 주관적 화질평가와 동일한 단일 등급 척도로 표현하기 위해 식 (20)의 전반적 화질열화 특성 요인과 국부적 화질열화 특성을 고려한 최종 화질열화 특성 맵으로부터 식 (21)을 통해 구해진다.

$$E_{map}(x, y) = E(x, y) \cdot \left( \sum_{i=1}^k L_i(x, y) \right)^\lambda \quad (20)$$

여기서  $L_i$ 는 각각 구획화 현상, 번짐, ringing의 화질 열화 측정 값이며  $\lambda$ 는 스케일링 지수로서 본 실험에서는 0.2를 사용하였다.

$$IQS = \frac{5}{1 + Er} \quad (21)$$

여기서  $Er = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} \{E_{map}(x, y)\}^2$ 이며  $N_x$ 와  $N_y$ 는 각각 영상의 수평, 수직 화소 수이다. 또한 PSNR과 직접적인 비교 평가를 위해 식 (21)으로부터 식 (22)의 IQPSNR를 산출한다.

$$IQPSNR(dB) = 10 \log_{10} \frac{255^2}{Er \cdot R} \quad (22)$$

여기서  $R$ 는  $IQPSNR$ 의 최대 동적범위(dynamic range)을 약 48dB에 맞추기 위한 상수로서 본 실험에서는 50을 사용하였다.

## 5. 실험 및 결과 고찰

### 5.1 실험 내용

실험에 사용된 정지영상은 표준 테스트 영상인 512x 512 크기의 Lena, Einstein, Fruits, Goldhill 흑백 영상이며 부호화 방식은 baseline JPEG을 사용하였다. 또한 제안된 모델 평가의 벤치마킹을 위해 주관적 화질평가는 DSCQS 법으로 6명의 영상 전문가 집단을 통해 각 2회씩 시행하여 MOS(mean opinion score)로 표현한다.

### 5.2 결과 고찰

그림 8은 0.3 bits/pixel 로 부호화된 Lena 영상에 대해 본 논문에서 제안한 왜곡 측정치를 나타낸 것이다.

제시된 결과는 디스플레이를 위해 16 ~ 235의 계조도로 스케일링 한 것이다. 그림에서 알 수 있듯이 원영상과 부호화 영상의 차영상 c)는 복잡한 배경을 갖는 영역에서 마스킹 효과에 의해 d) 영상과 같이 왜곡 정도가 상당히 완화되며 또한 블록기반 DCT 부호화가 갖는 독특한 화질열화 특성도 제안된 모델에 의해 e), f), g) 영상을 통해 비교적 정확히 측정되어 h) 영상의 화질열화 특성 맵의 결과로 나타난다.

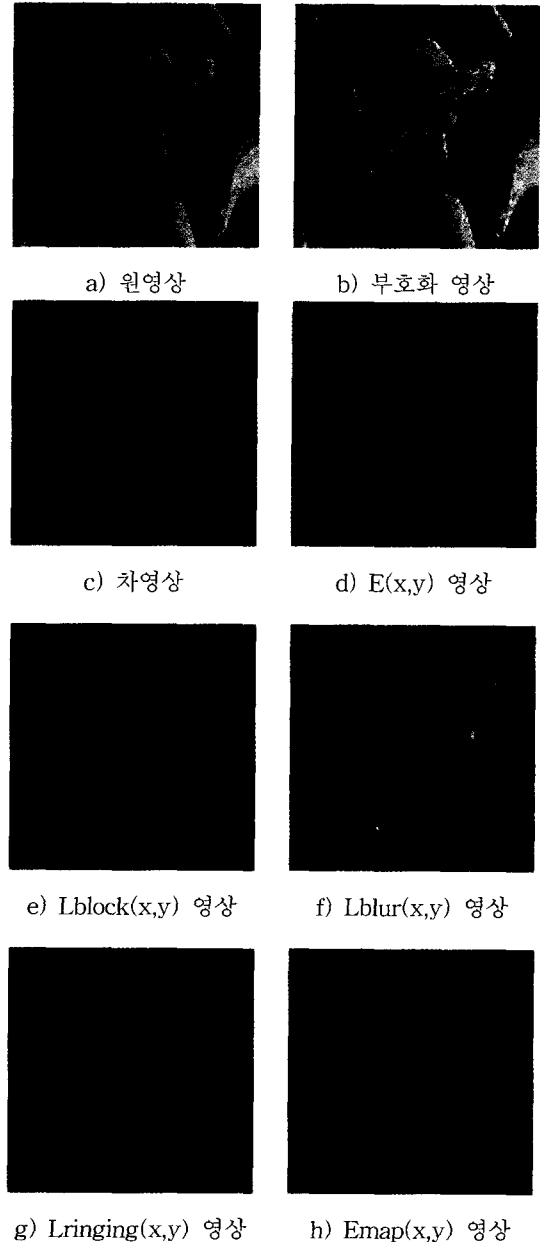


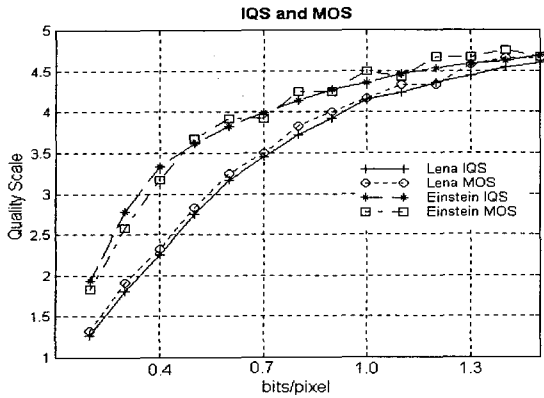
그림 8. 왜곡 측정의 결과 영상

Fig. 8. Result images of distortion measurement

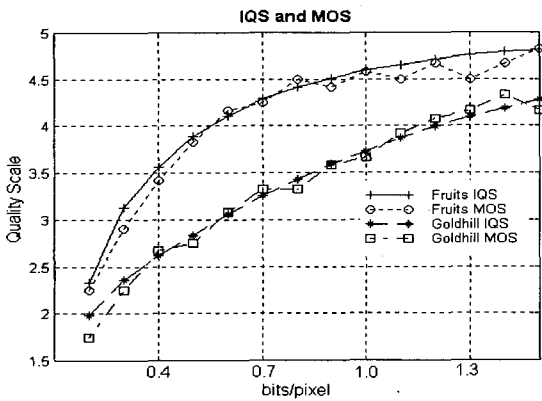
그림 9는 0.2~1.5bits/pixel의 비트율로 부호화된 각 테스트 영상에 대해 본 논문에서 제안한 IQS와 주관적 MOS에 의해 평가된 결과를 나타낸 것이다.

그림에서 알 수 있듯이 주관적 화질 평가인 MOS와 제안한 IQS의 RMSE는 0.085 ~ 0.12로서 거의 대등한 결과를 나타내고 있다. 또한 MOS는 부호화 비트율이 약 1.0 bits/pixel 이상인 부분에서 비트율이 증가해도 시각적인 착시 현상에 의해 화질 등급의 감소 부분이 간혹 발생하였으나 제안한 IQS는 이 부분에서도 완만한 화질 상승을 보여주고 있다. 이러한 현상은 부호화기 설계시 주관적 화질 평가 뿐 아니라 정량적 화질평가가 반드시 필요하다는 사실을 의미한다.

그림 10은 각 비트율에 대해 제안한 IQPSNR과 기존의 PSNR의 결과를 도시한 것이다.



a) Lena와 Einstein 영상



b) Fruits와 Goldhill 영상

그림 9. IQS와 MOS 결과  
Fig. 9. Results of IQS and MOS

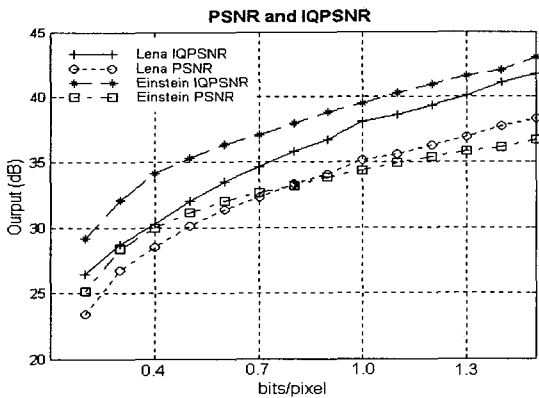


그림 10. IQPSNR과 PSNR 결과  
Fig. 10. Results of IQPSNR and PSNR

결과에 그림에서 알 수 있듯이 0.8bits/pixel 이하의 비트율에서는 Einstein 영상의 PSNR이 Lena 영상 보다 높으나 그 이상의 비트율에서는 Lena 영상의 PSNR이 Einstein 영상보다 높다. 그러나 제안한 IQPSNR은 모든 비트율에 걸쳐 Einstein 영상이 Lena 영상 보다 높다. 이것은 Einstein 영상이 Lena 영상에 비해 시각적으로 덜 민감한 특성을 갖고 있

으며 따라서 동일한 비트율에서 정량적 화질 특성이 좋으며 또한 PSNR이 부호화 왜곡에 대한 실제 인간의 화질 인지와는 다르다는 것을 보여주는 것이다.

표 2는 각 비트율로 부호화된 테스트 영상의 화질 측정 결과의 평균을 정리한 것이다.

표 2. 측정 결과의 평균  
Table 2. Mean results of measurement

	IQS	MOS	IQPSNR(dB)	PSNR(dB)
Lena	3.48	3.55	35.51	32.86
Einstein	3.94	3.95	37.75	32.84
Fruits	4.18	4.11	39.64	34.67
Goldhill	3.38	3.36	34.63	29.16

## 6. 결 론

기존의 객관적인 화질평가의 척도로 사용되는 PSNR은 실제 측정치와 인간이 시각적으로 인지하는 가시적인 왜곡치와는 그 상관성이 낮아서 부호화된 디지털 영상신호의 정량적인 화질측정에 적합하지 않았다.

따라서 본 논문에서는 대부분의 영상정보 압축 부호화 표준안에서 채용하고 있는 블록기반 DCT 기법을 근간으로 하는 영상 부호화기의 성능 향상 및 검증에 필수적인 HVS 기반의 정량적인 화질평가 모델을 제시하였다. 즉, HVS의 다채널 모델을 통해 영상신호에 있어서의 주파수 민감도와 채널의 마스킹 효과를 고려한 공간 영역상의 전반적인 왜곡의 가시도 뿐만 아니라 블록기반 DCT 부호화에서 발생할 수 있는 구획화 현상, 번짐, ringing 등의 국부적인 왜곡 특성을 측정하여 이를 화질 열화 특성에 반영함으로써 주관적 화질 평가와 대등한 성능을 보임을 실험을 통해 검증하였다.

## 참 고 문 헌

- [1] J. S. Lim, "Two-Dimensional Signal and Image Processing," Prentice Hall, 1990.
- [2] S. Winkler, "Issues in Vision Modeling for Perceptual Video Quality Assessment," Signal Processing, Vol. 78, No. 2, pp. 231-252, 1999.
- [3] A. A. Webter, C. T. Coleen, "An Object Video Quality Assessment System based Human Perception," in SPIE Human Vision, Visual Processing and Digital Display IV, Vol. 1913, pp. 15-26, 1993.
- [4] M. Miyahara, K. Kotani, V. R. Algazi, "Objective Picture Quality Scale(PQS) for Image Coding," IEEE Trans. Comm., Vol. 46, No. 9, pp. 1213-1226, 1998.
- [5] C. J. van den Branden Lambrecht, "Perceptual Models and Architectures for Video Coding Applications," Ph.D. Thesis, Ecole Polytechnique Federale De Lausanne, 1996.
- [6] A. B. Watson, "The Cortex Transform : Rapid

- Computation of Simulated Neural Images," Computer Vision Graphics and Image Processing, Vol. 39, No. 3, pp. 311-327, 1987.
- [7] J. L. Mannos and D. J. Sakrison, "The Effect of a Visual Fidelity Criterion on a Encoding of Images," IEEE Trans. Inform. Theory, Vol. IT-20, pp. 525-536, July 1974.
- [8] D. J. Sakirison, "On the Role of the Observer and a Distortion Measure in Image Transmission," IEEE Trans. Comm., Vol. COM-25, pp. 1251-1267, Nov. 1977.
- [9] F. W. Campbell and J. G. Robson, "Application of Fourier Analysis to the Visibility of Grating," Journal of Physiology, Vol. 197, pp. 551-566, 1968.
- [10] A. Fiorentini and M. Jeanne, "Measurement of Differential Threshold in the Presence of Spatial Illumination Gradient," Atti Ford. Ronchi, Vol. 10, pp. 371-379
- [11] M. Yuen and H. R. Wu, "A Survey of Hybrid MC/DPCM/DCT Video Coding Distortions," Signal Processing, Vol. 70, No. 3, 1998.
- [12] E. Peli, "Contrast in Complex Images," Journal of the Optical Society of America, Vol. 7, No. 10, pp. 2032- 2040, 1990.
- [13] P. C. Teo and D. J. Heeger, "Perceptual Image Distortion," in Proc. SPIE, Vol. 1279, pp. 127-141, San Jose, CA, Feb. 1994.
- [14] E. P. Simoncelli, W. T. Freeman, D. J. Heeger, "Shiftable Multi-scale Transforms," IEEE Trans. Inform. Theory, Special Issue on Wavelets, Vol. 38, pp. 587-607, 1992.
- [15] G. E. Legge and J. M. Foley, "Contrast Masking in Human Vision," Journal of the Optical Society of America, Vol. 70, No. 12, pp. 1458-1471, 1980.
- [16] ITU-R Recommendation BT. 500-7, "Methodology for the Subjective Assessment of the Quality of Television Pictures," ITU, Geneva, Switzerland, 1995.

저 자 소 개



정태윤(Tae-Yun Chung)

1987년: 연세대 전기공학과 (학사)  
 1989년: 동대학원 전기공학과 (공학석사)  
 2000년: 동대학원 전기컴퓨터공학과 (공학박사)  
 1989년~1996년: 삼성종합기술원 (선임연구원)

1996년~2001년 2월: 삼성전자 중앙연구소 책임연구원  
 2001년 3월~현재: 강릉대학교 정보전자공학부 전임강사

관심분야 : 영상처리 및 부호화, 화질평가, 멀티미디어 정보보호, 멀티미디어 시스템 설계

Phone : +82-33-640-2425

Fax : +82-33-640-2244

E-mail : tychung@kangnung.ac.kr