

구조적 왜곡특성 측정을 이용한 블록기반 DCT 영상 부호화기의 객관적 화질평가

論文

52D-7-8

Objective Image Quality Metric for Block-Based DCT Image Coder Using Structural Distortion Measurement

鄭泰允*

(Tae-Yun Chung)

Abstract - This paper proposes a new quantitative and objective image quality metric which is essential to verify the performance of block-based DCT image coding. The proposed metric considers not only global distortion of coded image such as spatial frequency sensitivity and channel masking using HVS based multi-channel model, but also structural distortions caused block-based coding. The experimental results show a strong correlation between proposed metric and subjective metric.

Key Words : Image quality metric, vision model, human visual system, blocking artifacts, ringing artifacts

1. 서론

최근 멀티미디어와 고속의 유·무선 정보 통신망의 발전으로 인해 디지털 영상 처리 서비스의 응용 분야가 급속히 확대되고 있다. 영상 정보의 디지털화는 화질, 신뢰성, 보안성 그리고 정보 조작의 편의성에 있어 기존의 아날로그 방식에 비해 월등히 우수한 성능을 갖고 있지만 그 정보량의 방대함 때문에 사용되는 전송 및 저장 매체의 효율성을 고려하여 정보 압축 부호화 기술이 필요하다. 영상 정보의 압축 부호화 기술은 응용 시스템의 목적에 따라 JPEG, MPEG1/2/4, H.261/263 등의 많은 국제 표준안들이 마련되었으며 이러한 표준 규격을 채택하는 시스템들의 사용이 더욱 확산될 전망이다[1].

영상 정보의 손실 압축 부호화 기술은 사용되는 매체의 용량과 복원 영상의 화질이라는 trade-off 관계를 고려하여 설계되며 이 경우, 정보 손실에 따른 화질 열화가 발생한다. 따라서 영상 부호화기 설계에 있어 가장 중요한 문제는 원영상의 어떤 정보를 어느 정도 왜곡시키는 것이 복원 영상의 화질 열화를 최소로 할 수 있는가 하는 것이며 이를 위해 정확한 화질 평가 방법은 부호화기의 성능 평가 및 향상에 있어 매우 중요한 문제로 부각되고 있다[2].

일반적으로 영상 정보의 화질 평가는 크게 주관적 평가 방법과 객관적 평가 방법으로 나눌 수 있다. 주관적 화질 평가 방법은 ITU(International Telecommunication Union)-R 권고안 BT. 500을 통해 평가 조건과 기준, 그리고 평가 절차와 분석 방법 등의 표준화가 마련되어있다[3]. 이러한 주관적 화질 평가 방법은 영상의 화질은 결국 인간에 의해 최

종적으로 평가된다는 점에서 그 신뢰도가 높은 반면에 장시간의 평가 소요 시간과 고비용의 평가 환경 구축 등의 문제점이 있다[2].

객관적 화질 평가는 정량적인 측정 도구를 사용하여 주관적 화질 평가의 문제점을 해결하고 또한 성능면에서는 주관적 화질 평가와 대등한 신뢰성을 확보하는데 그 목적이 있다. 전통적인 객관적 화질 평가 방법으로 사용되는 MSE (Mean Square Error) 또는 PSNR(Peak Signal to Noise Ratio)은 원영상과 왜곡영상에 대해 공간적 화소 단위로 화질 열화 정도를 측정하는 방식으로서 아날로그 영상의 화질 측정에는 비교적 적합하지만 영상의 왜곡 특성이 아날로그 영상과는 다른 디지털 압축 부호화된 영상의 화질 평가에는 적절하지 않다[4]. 따라서 최근에 ITU 후원하에 VQEG (Video Quality Experts Group)가 결성되어 인간의 시각 특성을 고려한 디지털 압축 부호화 영상의 객관적 화질 평가 표준화 방안을 논의 중이다[5].

본 논문에서는 대부분의 영상 정보 압축 부호화 표준화에서 채용하고 있는 블록 기반 DCT 부호화기의 성능 향상 및 검증에 활용할 수 있는 객관적 화질 평가 모델을 제안한다. 제안된 모델은 HVS(Human Visual System)을 다채널 모델의 신호 처리 기법으로 해석하여 영상 부호화시 발생하는 전반적 왜곡 특성을 분석하고 이를 바탕으로 블록기반 DCT 부호화기에서 발생하는 블록화 현상(blocking artifacts), ringing 등의 구조적 화질 열화 특성을 정량적으로 측정하여 최종 화질 평가에 반영한다. 제안된 모델의 성능은 JPEG 부호화된 정지영상을 대상으로 한 주관적 화질 평가와 비교함으로써 타당성을 검증하였다.

2. 인간 시각 특성

영상 정보의 손실 압축 부호화에 의한 화질 열화 정도를 정량적으로 평가하기 위해서는 인간 시각 특성을 모델링하

* 正會員 :江陵大 情報電子工學部 助教授·工博
接受日字 : 2003年 4月 24日
最終完了 : 2003年 5月 14日

는 것이 필요하다[6]. 그러나 시각 정보의 취득에서부터 취득한 정보를 뇌에서 인지하는 과정을 다루는 HVS는 그 구조 및 현상이 매우 복잡하여 수 많은 연구에도 불구하고 오늘날까지도 시스템적인 성질이 정확히 규명되고 있지 않은 상태이다[7]. 다만 눈의 광학계에서 원시 시각영(primary visual cortex)에 이르는 HVS의 전단은 전자생리학 및 정신물리학적 실험에 의해 어느 정도 그 메카니즘이 근사적으로 밝혀졌다[4]. HVS를 모델링하기 위해 주로 사용되는 방법은 서로 다른 배경 영상위에 임의의 시각적 자극이 가해졌을 때 눈이 이를 감지하기 시작한 자극의 크기를 측정하여 정량화하는 것이다[8]. 이때 시각적으로 관측 가능한 상태가 되는 자극의 크기를 가시도 문턱값(visibility threshold)이라 하며 이것의 결과들을 공간 영역 및 주파수 영역에서 해석한다.

2.1 시각 모델 구조

그림 1은 현재 VQEG의 Phase 1에서 논의 중인 기준 영상(full reference image)에 대한 왜곡 영상의 화질 평가 모델을 간단히 나타낸 것이다.

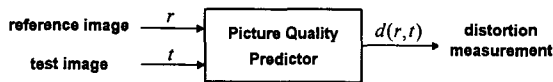


그림 1 화질 평가 모델
Fig. 1 Image quality metric model

객관적 화질 평가의 목적은 만일 r 과 t 가 각각 원영상과 왜곡영상이면 주관적 화질 평가와 동일한 성능을 보이는 화질 측정값 $d(r, t)$ 를 구하고자 하는 것이다.

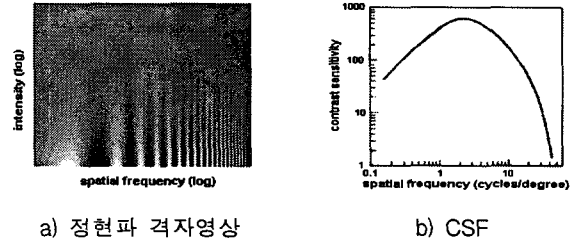
이러한 객관적 화질 평가를 위한 시각 모델은 크게 단일 채널 모델과 다채널 모델로 나눌 수 있다. 단일 채널 모델은 단일 자극 패턴에 의한 실험을 통해 HVS의 특성을 비교적 간단히 평가할 수는 있으나 많은 경우에 있어 인간의 복잡한 시각특성을 충분히 반영할 수 없다. 따라서 Sakrison 등은 정신물리학적 실험을 통해 인간의 시각 인지 특성은 다해상도로 구성되어 있다는 것을 규명함으로써 다채널 모델을 제시하였다[9]. 따라서 대부분의 객관적 화질 평가를 위한 시각 모델은 자연 영상 패턴의 가시도 평가를 위해 채널간, 채널내의 상호 작용 및 시각적 마스크(masking) 효과를 고려한 다채널 모델로 구성한다.

2.2 시각 민감도(Visual Sensitivity) 측정

인간 시각은 밝기의 세기보다는 밝기의 차이 즉, 대비(contrast)에 더욱 민감하다. 이러한 영상의 대비 특성을 그림 2와 같이 공간 주파수에 대한 시각 민감도로 표현하는 것이 CSF(Contrast Sensitivity Function)이다. 그림 2에서 알 수 있듯이 CSF는 대역통과필터의 특성을 보이며 따라서 동일한 양의 잡음성분에 대해서도 잡음의 주파수 분포에 따라 실제 시각이 인지하는 왜곡 특성은 다르게 나타난다.

또한 대부분의 자연 영상은 배경의 밝기가 불균일하고 복잡하다. 이러한 배경 영상의 밝기의 변화에 의해 목적 영상

의 가시도 민감도가 변화하는 현상을 공간 마스크(masking)라 한다. 즉, 같은 양의 잡음이라도 주변 및 배경의 밝기 변화에 의해 인간 시각의 민감도는 다르게 나타나며 결국 마스크 현상은 시각 모델의 다채널 특성을 유발하는 주요 원인이 된다[10].



a) 정현파 격자영상 b) CSF

그림 2 정현파 격자 영상에 대한 CSF
Fig. 2 CSF for sinusoidal grating image

3. 블록기반 DCT 부호화의 화질 열화 특성

손실 압축 부호화 영상의 화질 열화 특성 분석은 효율적인 부호화기 설계 뿐만 아니라 객관적 화질 평가 모델 개발에 있어서도 매우 중요하다. 현재 가장 많은 응용 분야에 적용되고 있는 영상 부호화 표준안은 JPEG과 MPEG으로서 이 방식들은 시스템 구현의 복잡도와 화질을 고려하여 8×8 블록단위의 DCT를 근간으로 양자화와 엔트로피 부호화 과정을 거쳐 공간적, 시간적인 영상의 잉여정보를 제거한다.

이러한 영상 압축 부호화는 결과적으로 시각적으로 두드러진 구조적 화질 열화를 유발하며 따라서 개별적인 구조적 화질 열화 특성의 분석을 통해 통합적인 정량적, 객관적 화질 평가 모델을 제시하는 것이 필요하다[11].

3.1 블록화 현상

블록화 현상은 블록 기반 DCT 부호화에 의한 시각적으로 가장 두드러진 화질 열화 특성으로서 인접 8×8 블록과의 경계에서 나타나는 불연속으로 정의된다. 이 현상은 인접 블록과의 상관성은 고려하지 않고 해당 블록만을 독립적으로 부호화함으로써 발생한다.

DCT 계수 중 DC와 AC 계수의 양자화 방법이 다르며 따라서 영상의 화소값이 수평 또는 수직 방향으로 단조 증가 또는 감소하고 또한 양자화 스텝 사이즈가 크다면 양자화 오차가 커지게 되어 복원된 영상은 블록 단위로 불연속하게 나타난다.

3.2 Ringing 현상

Ringing 현상은 영상의 비교적 균일한 영역에서의 에지(edge)와 같이 큰 대비값을 갖는 영상의 경계면 주위에 어긋나거나 물결(ripple) 형태와 같은 주기적인 의사(pseudo) 에지의 형태로 보이는 시각적 열화 특성이다. 이 현상은 DCT 계수 중 고주파 성분의 잘림에 기인하며 에지의 대비값이 커질수록 물결 형태의 침투값이 커지고 골(trough)이 깊어지는 Gibbs 현상과 관련이 있다.

3.3 번짐(blurring) 현상

번짐 현상은 텍스처 영역 또는 에지 주변과 같이 공간적으로 높은 활성도를 갖는 영역에서 발생하며 공간적인 정보 손실로 나타나는 화질 열화 특성이다. 즉, 성긴(coarse) 양자화에 의한 고주파 영역의 AC 계수의 억압과 직접적으로 관련되며 따라서 블록내의 영상은 저주파 영역의 계수에 의해서만 표현되는 저역 통과 필터의 특성을 보인다.

4. 제안된 객관적 화질 평가 모델

그림 3은 본 연구에서 제안한 객관적 화질 평가 모델의 블록도를 나타낸 것이다. 기존의 객관적 화질 평가 모델 연구의 대부분은 영상의 왜곡 부분에 대한 시공간적 주파수 민감도와 시각적 마스킹 현상만을 고려하여 왜곡 정도에 따른 단일 척도를 산출함으로써 왜곡 영상에 대한 객관적 화질 평가를 시도하였다.

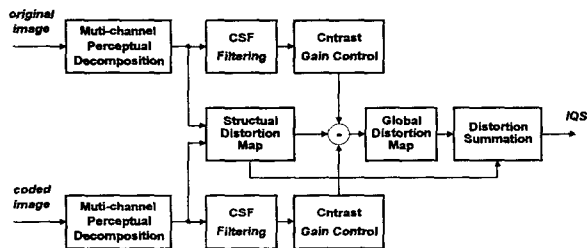


그림 3 제안된 객관적 화질 평가 모델
Fig. 3 Proposed objective image quality metric model

그러나 이러한 모델은 블러킹 현상, ringing 및 번짐 등의 가시적으로 왜곡 현상이 두드러진 구조적 화질 열화 특성을 보이는 블록 기반 DCT를 채용하는 영상 압축 부호화기의 객관적 화질 평가 방법으로는 불충분하다.

따라서 본 연구에서는 HVS 기반의 다채널 구조에 따른 주파수 민감도와 시각적 마스킹 현상을 고려하여 부호화 영상의 전반적 왜곡의 가시도 맵을 산출하고 또한 블록 기반 DCT를 채용하는 부호화기에서 발생하는 가시적으로 민감한 구조적 왜곡 특성을 측정하여 이를 단일 평가 척도의 객관적 화질 평가에 반영한다.

4.1 다채널 인지 분해

전자 생리학 실험을 통해 인간의 눈의 시각령 세포구조는 Gabor 형태의 수용계 구조를 갖는 것으로 밝혀졌다[12]. 즉, 시각령 수용계는 Gabor 형태의 비원형 구조로서 주파수와 방위(orientation)에 대해 선택적으로 반응하는 흥분성과 억제성 영역으로 구성된다. 따라서 원영상과 부호화 영상에 대한 HVS의 방위와 주파수 성분을 표현하는 다채널 인지 분해(perceptual decomposition)는 Gabor 필터로 구현될 수 있다. 그러나 Gabor 필터는 스펙트럼 영역상에서 매우 선택적이며 또한 필터링이 주파수 영역에서 수행됨으로서 과중한 메모리와 연산 시간을 요구한다. 또한 객관적 화질 평가에 있어 사실상 정확한 필터의 형태는 필요하지 않으며 시

스템 구현시의 복잡도와 평가 정밀도의 trade-off 관계를 고려해야 한다[7]. 따라서 본 연구에서는 steerable 피라미드 변환을 통한 다채널 인지 분해를 시행한다.

그림 4는 steerable 피라미드 변환의 블록도를 나타낸 것이다. Steerable 피라미드 변환은 decimation과 컨볼루션을 이용하여 영상을 옥타브 대역의 부대역(subband)으로 구성된 공간주파수 영역으로 분해하고 각각의 부대역은 여러개의 방위대역으로 나눈다.

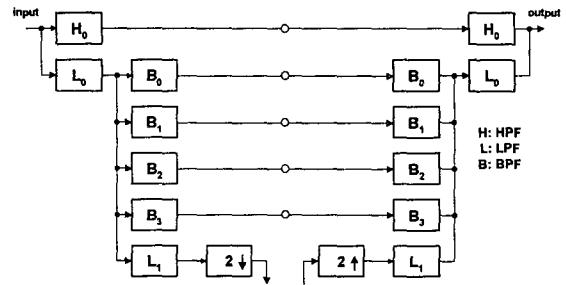


그림 4 Steerable 피라미드의 블록도
Fig. 4 Block diagram of steerable pyramid

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[13]. 따라서 이러한 steerable 피라미드 변환 필터를 설계하기 위해서는 다음의 조건을 만족시켜야 한다.

$$\begin{aligned} \hat{X}(w) = & \{|H_0(w)|^2 \\ & + |L_0(w)|^2 \left(|L_1(w)|^2 + \sum_{k=0}^n |B_k(w)|^2 \right) \} X(w) + a.t. \end{aligned} \quad (1)$$

여기서 $X(w)$ 와 $\hat{X}(w)$ 와 각각 입력 영상과 복원 영상의 스펙트럼을 나타내며 $H_0(w)$ 는 고역통과 필터, $L_0(w)$ 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 최소로 하기 위해서는 식 (2)와 같이 L_1 필터는 $\pi/2$ 이상의 주파수에서는 0의 응답특성을 가져야 한다.

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

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.

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

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

- 2 단계 : step 1에서 검출된 수평 및 수직 방향의 블록의 에지점을 계수하여 각 방향에서의 8×8 블록화 경계면을 결정한다.

$$B_v(8m, y+j) = \begin{cases} 1, & j \in [0, 7], \text{ if } \sum_{j=0}^7 EP_v(j) > Th_2 \\ 0, & j \in [0, 7], \text{ otherwise} \end{cases} \quad (9)$$

식 (9)로 부터 수직 8 line의 블록화 경계면을 결정하며 여기서 $EP_v(j)$ 는 j 화소가 step 1에서 얻어진 수직 방향의 블록 에지점이면 1이고 그 밖의 경우는 0을 의미하며 Th_2 는 블록화 영역을 결정하는 문턱값이다. 본 실험에서는 Th_1 과 Th_2 를 각각 7.0과 4.0으로 정의한다.

- 3 단계 : 검출된 수평 및 수직 방향의 블록 경계면으로부터 각각 인접한 상하, 혹은 좌우에 해당하는 8×8 영역을 블록화 영역으로 결정한다.
- 4 단계 : 부호화 과정에서 발생하는 블록화 영역만을 검출하기 위해 원영상과 부호화 영상 모두에서 검출되는 블록화 영역은 제외한다.

이와 같은 과정으로 검출된 블록화 영역의 왜곡량 B_e 는 식 (7)에서 산출한 전반적 왜곡의 가시도 맵 $E(x, y)$ 로부터 식 (10)을 통해 결정한다.

$$B_e = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} E(x, y), \quad \forall (x, y) \in \text{블록화 영역} \quad (10)$$

(2) Ringing 왜곡 측정

ringing 현상은 부호화 영상에서 큰 대비값을 갖는 물체의 에지 주변부에서 발생하며 가시적으로는 에지의 어긋나리는 현상으로 나타난다. 따라서 이러한 에지 영역은 고역 통과된 원영상 $A(x, y)$ 에서 식 (11)과 같이 주변 5×5 영역으로부터 검출한다.

$$R(x, y) = \sum_{i=-2}^2 \sum_{j=-2}^2 (A(x+i, y+j) - A(x+i+1, y+j))^2 + \sum_{i=-2}^2 \sum_{j=-2}^2 (A(x+i, y+j) - A(x+i, y+j+1))^2 \quad (11)$$

식 (11)에서 $R(x, y) > Th_3$ 이면 해당 화소는 압축 부호화시 ringing 현상이 발생할 가능성이 높으며 따라서 해당 화소를 둘러싼 3×3 영역을 ringing 영역으로 결정한다. 단, 블록화 영역과의 중복을 피하기 위해 블록화 영역으로 결정된 화소는 ringing 영역에서 제외한다. 본 실험에서 Th_3 는 1500으로 정의한다. 이와 같은 과정으로 검출된 ringing 영역의 왜곡량 R_e 는 식 (12)로부터 구한다.

$$R_e = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} E(x, y), \quad \forall (x, y) \in \text{ringing 영역} \quad (12)$$

(3) 번짐 및 기타 왜곡 측정

번짐은 가시적으로 부호화에 의한 저역통과 필터의 특성을 보이는 구조적 왜곡 특성이며 전반적 왜곡의 가시도 맵 $E(x, y)$ 로부터 블록화 영역과 ringing 영역을 제외한 부

분에서 대부분 발생한다. 따라서 번짐 및 기타의 구조적 왜곡량 L_e 는 식 (13)으로부터 구한다.

$$L_e = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} E(x, y), \quad \forall (x, y) \in \text{기타 영역} \quad (13)$$

4.5 단일 평가 척도의 객관적 화질 평가

주관적인 화질 평가 방법은 ITU-R 권고안 500을 통해 평가 조건, 평가 절차와 분석방법 등의 기준이 마련되었으며 일반적으로 DSCQS(Double Stimulus Continuous Quality Scale)법이 많이 사용되고 있다. 본 논문에서는 부호화 영상을 DSCQS법에 의해 1에서 10점(10점이 가장 좋은 화질)의 등급으로 평가한 주관적 평가와의 성능 비교를 위해 식 (14)로부터 B_e , R_e 및 L_e 의 구조적 왜곡 특성을 선형 조합한 부호화 영상의 왜곡량을 산출한 후 식 (15)를 통해 객관적 화질 평가의 단일 평가 척도 IQS(Image Quality Scale)를 구한다.

$$E_t = \sum_{i=1}^3 s_i w_i \quad (14)$$

여기서 s_i 는 B_e , R_e 및 L_e 의 구조적 왜곡량을, 그리고 w_i 는 각 화질 열화 요인의 가중치를 나타낸 것으로서 본 실험에서는 주관적 평가의 MOS(Mean Opinion Score)와의 회귀모델로부터 구한 $w_1 = 1.73$, $w_2 = 1.19$, $w_3 = 0.82$ 를 사용한다.

$$IQS = \frac{10}{1 + E_t^{0.5}} \quad (15)$$

5. 실험 및 결과 고찰

5.1 실험 내용

제안된 객관적 화질 평가 모델의 타당성을 검증하기 위해 본 실험에서는 Texas 대학에서 제공하는 화질 평가용 흑백 정지화 데이터 베이스를 활용하였다. 실험 영상은 여러 가지 해상도를 갖는 총 23종의 원영상과 baseline JPEG으로 부호화된 120개 평가 영상으로 구성되어 있으며 또한 각각의 부호화 평가 영상에 대해 DSCQS법으로 총 33명의 전문가에 의해 시행된 1~10점의 등급을 갖는 주관적 화질 평가의 MOS를 제공한다.

5.2 결과 고찰

그림 6은 0.32bpp로 부호화된 480×720의 Sailing 영상에 대해 본 논문에서 제안한 부호화 왜곡 측정치를 나타낸 것이며 c)~f)영상은 디스플레이를 위해 16~235의 계조도로 스케일링 한 것이다. 원영상과 부호화 영상의 차영상 c)와 본 논문의 전반적 왜곡 가시도 맵 d)영상을 비교해 보면 c) 영상에서 복잡한 배경을 갖는 영역은 d) 영상에서는 마스킹 효과에 의해 왜곡 정도가 상당히 완화되며 그 반대로 c) 영상에서의 왜곡량은 상대적으로 작지만 시각적으로 민감한 영역에서는 d) 영상은 왜곡 정도가 강조되는 결과를 보이고

있다. 또한 블록 기반 DCT 부호화기에서 나타나는 구조적 화질 열화 특성인 블록화 영역과 ringing 영역도 e)와 f) 영상에서 알 수 있듯이 비교적 정확히 검출되어 시각적으로 민감한 구조적 화질 열화 특성을 객관적 화질 평가에 반영할 수 있다.

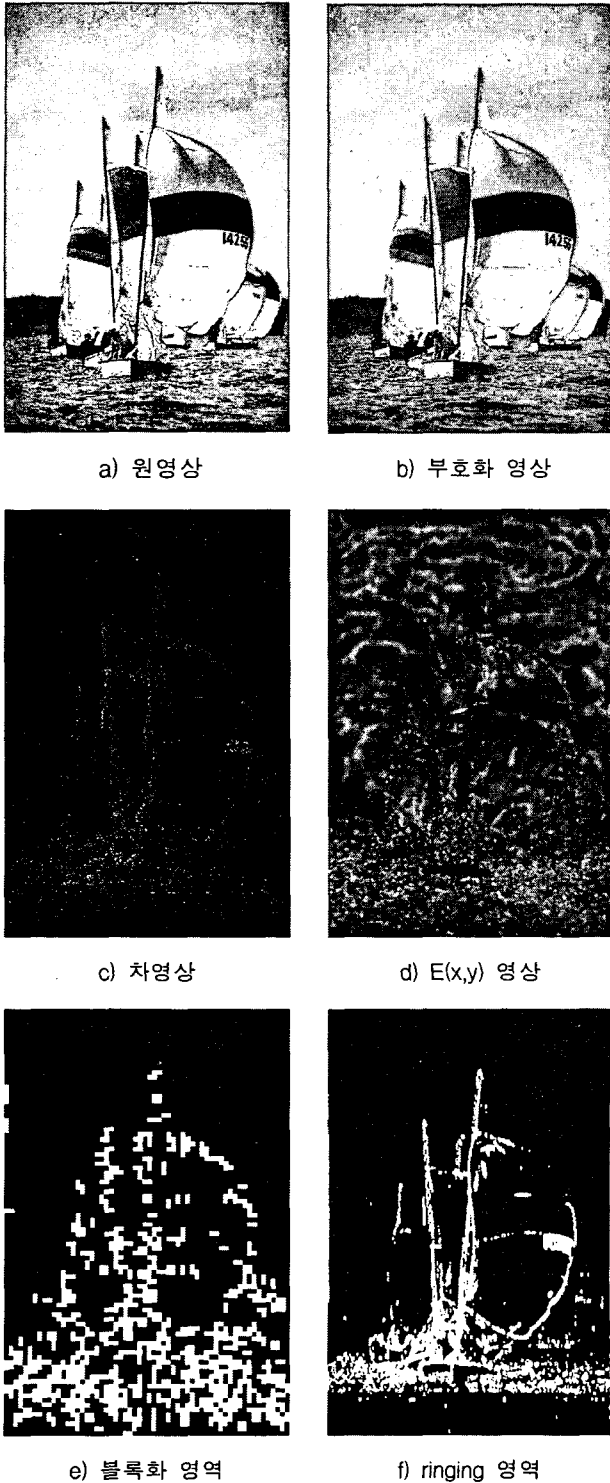
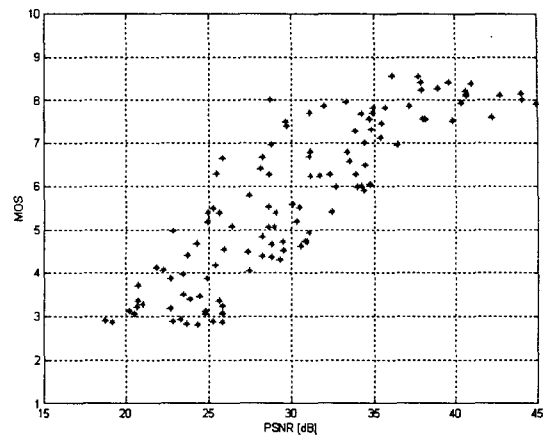


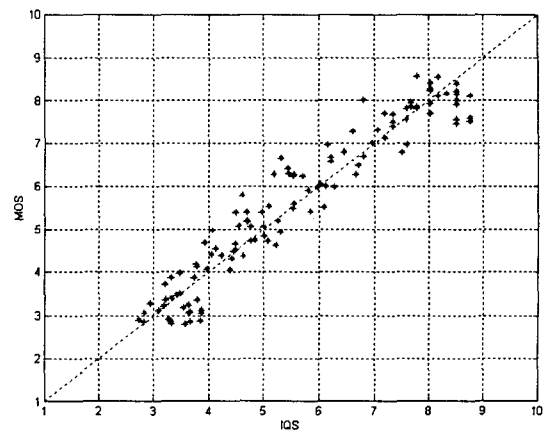
그림 6 왜곡 측정의 결과 영상
Fig. 6 Results images of distortion measurement

객관적 화질 평가 방법의 타당성을 논의할 때 고려해야 할 사항은 실험 영상에 대한 주관적 화질 평가와 객관적 화질 평가와의 오차 분포와 상관관계이다.

그림 7은 각각의 부호화 실험 영상에 대해 PSNR 및 식 (15)에서 제안된 IQS와 주관적 평가의 MOS 결과와의 관계를 도식화한 것이다. 그림의 결과에서 알 수 있듯이 동일한 MOS에 대하여 PSNR 측정값은 많게는 15dB 이상의 상당히 넓은 분포를 보이고 있어서 객관적 화질 평가의 측정 도구로 적합하지 않음을 확인할 수 있었으며 반면에 제안된 IQS 측정값은 MOS와의 RMSE(Root Mean Square Error)가 0.513으로서 MOS의 분산의 평균값($\overline{\sigma_{mos}^2}$)인 0.637보다 작아서 주관적 평가자에 의한 부호화 영상의 주관적 평가의 오차 범위에 있고 또한 MOS의 증가에 대해 선형적 증가 특성을 보이고 있다.



a) PSNR과 MOS 결과



b) IQS와 MOS 결과

그림 7 주관적 평가와 객관적 평가의 결과

Fig. 7 Subject assessment vs. object assessment

표 1은 주관적 평가의 MOS와 객관적 화질 평가 방법들의 상관관계를 정량적으로 분석한 결과이다. 본 실험에서 사용한 분석 방법은 식 (16)의 Pearson 상관계수와 식 (17)의 Spearman 순위(ranking order) 상관계수이며 상관계수가

1에 근접할수록 측정 변인들의 상관성은 높은 것이다.

$$P_r = \frac{\sum_i (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_i (X_i - \bar{X})^2 \sum_i (Y_i - \bar{Y})^2}} \quad (16)$$

여기서 X : 주관적 MOS값, \bar{X} : X 의 평균
 Y : 객관적 측정값, \bar{Y} : Y 의 평균

$$S_r = 1 - \frac{6 \sum_{i=1}^n D_i^2}{n(n^2 - 1)} \quad (17)$$

여기서 n : 측정 회수, D : MOS와 객관적 측정값의 순위차

표 1 PSNR과 IQS의 성능 비교

Table 1 Performance comparison of PSNR and IQS

평가방법	Pearson 상관계수 (P_r)	Spearman 순위상관계수 (S_r)	σ_{mos}^2
PSNR	0.828	0.839	0.637
IQS	0.960	0.957	

표의 결과에서 알 수 있듯이 제안된 객관적 화질 평가 방법은 PSNR에 비해 MOS와의 Pearson 및 Spearman 순위상관 계수가 0.13 정도 향상되어 거의 1에 접근하는 결과를 보이고 있다. 따라서 제안된 IQS 측정값은 주관적 평가 결과와의 오차 범위와 상관 관계를 고려할 때 주관적 화질 평가와 거의 대등한 성능을 보이는 객관적 화질 평가 방법으로 매우 적합함을 알 수 있다.

6. 결 론

기존의 객관적인 화질 평가 방법으로 사용되는 PSNR은 실제 측정값과 인간이 시각적으로 인지하는 가시적인 화질과는 그 상관성이 낮아서 압축 부호화된 디지털 영상신호의 정량적이며 객관적인 화질 평가에는 적합하지 않았다.

따라서 본 논문에서는 대부분의 영상 정보 압축 부호화 표준에서 채용하고 있는 블록 기반 DCT 부호화기의 화질 검증 및 성능 향상에 필수적으로 사용되는 HVS 기반의 단일 척도의 객관적 화질 평가 모델을 제시하였다. 제안된 모델은 HVS 다채널 구조에 따른 주파수 민감도와 시각적 마스킹 현상을 고려하여 부호화 영상에 대한 공간 영역상의 전반적 왜곡의 가시도 맵을 산출하고 또한 블록 기반 DCT를 채용하는 부호화기에서 발생하는 블록화 현상, ringing 현상 등의 가시적으로 민감한 구조적 왜곡 특성을 화질 열화 특성에 반영함으로써 주관적 화질 평가와 거의 대등한 성능을 보임을 실험을 통해 검증하였다.

감사의 글

본 연구는 정보통신부의 정보통신기초기술연구지원 사업(정보통신연구진흥원)으로 수행한 연구결과임.

참 고 문 헌

- [1] K. R. Rao and J. J. Huang, Techniques and Standards for Image, Video and Audio Coding, Prentice Hall, 1996.
- [2] M. Miyahara, K. Kotani and V. R. Algazi, "Objective Picture Quality Scale(PQS) for Image Coding," IEEE Trans. Comm., Vol. 46, No. 9, pp. 1213-1226, 1998.
- [3] ITU-R Rec. BT. 500-9, "Methodology for the subjective assessment of the quality of television pictures," ITU, 1998.
- [4] C. J. van den Branden Lambrecht, "Perceptual Models and Architectures for Video Coding Applications," Ph.D. Thesis, Ecole Polytechnique Federale De Lausanne, 1996.
- [5] VQEG, "Final Report from the video quality experts group on the validation of objective models of video quality assessment," Mar. 2000.
- [6] Z. Wang and A. C. Bovik, "Why is mage quality assessment so difficult?," IEEE Int. Conf. Acoustcs, Speech, and Signal Processing, Vol. 4, pp. 3313-3316, May 2002.
- [7] Z. Yu, H. R. Wu, S. Winkler and T. Chen, "Vision Model based Impairment Metric to Evaluate Blocking Artifacts in Digital Video," Proceeding of IEEE, Vol. 90, No. 1, pp. 154-169, Jan. 2002.
- [8] J. Lubin and D. Fibush, "Sarnoff JND Vision Model," T1A1.5 Working Group Document #97-612, T1 Standard Committee, 1997.
- [9] 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.
- [10] S. Winkler, "A Perceptual Distortion Metric for Digital Color Video," Human Vision and Electronic Imaging, Proc. SPIE, Vol. 3644, pp. 175-184, 1999.
- [11] 정태윤, 홍민석, 박강서, 김현술, 박상희, "블록기반 동 영상 부호화기의 객관적 화질평가," 대한전기학회 논문지, 제 48권, 제 12호, pp. 1589-1598, 1999.
- [12] P. C. Teo and D. J. Heeger, "Perceptual Image Distortion," in Proc. SPIE, Vol. 1279, pp. 127-141, San Jose, CA, Feb. 1994.
- [13] 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.
- [14] A. B. Watson and J. Solomon, "Model of Visual Contrast Gain Control and Pattern Masking", Journal of the Optical Society of America, Vol. 14 pp. 2379-2391, 1997.
- [15] H. Ridder and M. C. Willemsen, "Percentage Scaling : a New Method for Evaluating Multiply Impaired Images," Proceeding of SPIE, Vol. 3959, 2000.

저 자 소 개



정 태 윤 (鄭 泰 允)

1964년 11월 21일생. 1987년 연세대학교 전기공학과 졸업. 1989년 동 대학원 전기공학과 졸업(석사). 2000년 동 대학원 전기·컴퓨터공학과 졸업(공학박). 1989년~1996년 삼성종합기술원 선임연구원. 1996~2001년 삼성전자 중앙연구소 책임연구원. 2001년~현재 강릉대학교 정보전자공학부 조교수.

Tel : 033-640-2425, Fax : 033-640-2244

E-mail : tychung@kangnung.ac.kr