

論文2001-38SP-1-5

JPEG과 호환 가능한 Wavelet CODEC

(A Wavelet CODEC that is compatible with JPEG)

金容奎*, 金度圭**, 趙錫八***

(Yongkyu Kim, Dokyu Kim, and Sok Pal Cho)

요 약

JPEG의 부호화 기법에서 DCT의 단점인 블록효과를 제거하기 위해 웨이브릿 변환을 사용하였다. 기존의 JPEG과 호환성 유지하기 위하여, JPEG 부호화 기법과 동일한 기법을 사용할 수 있는 부호화기를 제안한다. 이를 위하여 웨이브릿 변환영역에서 각 대역의 신호를 8x8로 재구성하였고, 각 대역의 에러분석과 인간 시각특성을 고려하여 양자화 가중치를 구하여 양자화 하였다. 실험을 통하여 웨이브릿을 사용한 제안한 기법이 기존의 JPEG 보다 (Peak Signal to Noise Ratio), WMSE(Weighted Mean Square Error) 척도에서 다소 우수함을 보였다.

Abstract

WT(Wavelet Transform) is used to avoid blocking effect that is the disadvantage of JPEG using DCT(Discrete Cosine Transform). Because the proposed coding scheme is the same as JPEG, the proposed algorithm is compatible with that of JPEG. To achieve the goal, WT'ed image is reconstructed into 8x8 coding block. Each coding block is quantized with the proposed weighting matrix that is derived from human visual characteristic and error analysis in WT'ed domain. By experiments, the proposed algorithm is superior to JPEG, in terms of PSNR(Peak Signal to Noise Ratio) and WMSE(Weighted Mean Square Error).

I. 서 론

디지털 영상압축 기술은 DCT(Discrete Cosine Transform)와 허프만 부호화를 이용한, MPEG(Moving Pictures Expert Group), JPEG(Joint Photographic Experts Group) 등의 표준화된 기법으로 상용화에 많은 진보를 보였다.

그러나 변환기법의 측면에서 볼 때 기존의 표준화된 영상압축 기법은 DCT를 기반으로 구성되어 있으며 이는 영상을 여러 개의 균일한 블록으로 분할하여 각 블록에 여현함수(cosine) 커널(kernel)을 적용하여 변환을 할 때 얻어지는 에너지 집중(energy compaction) 효과와, 영상의 중복성 제거효과를 이용한 것이다. DCT는 거의 대부분의 표준부호화 기법에서 기본적으로 채택하여 사용하는 기법으로 고속 변환기법이 존재하고, 이미 상용화된 IC가 공급되고 있다는 장점이 있다. 그러나 이 방법은 변환기법의 특성상, 고압축시 블록효과(block effect)가 나타나므로 고압축 알고리즘에는 무리가 따른다. 특히 JPEG의 경우 거의 1bpp의 압축률에서 최고의 성능을 발휘하도록 조정되어 있기에 고압축을 요구하는 응용분야에서는 적용하는데 많은 애로사항이 있었다.

* 正會員, 韓國 GMX
(Global Media Exchange)

** 正會員, 尙州大學校 電子電氣工學部
(Sangju National University)

*** 正會員, 聖潔大學校 컴퓨터學部
(Sungkyul University)

接受日字:2000年10月24日, 수정완료일:2000年12月22日

기존 영상압축 기법들의 단점과 한계를 극복하기 위해 최근 새로운 변환기법인 웨이브릿(Wavelet Transform)이 도입되어 의미 있는 연구결과가 보고되고 있다. 웨이브릿은 시간과 주파수에 대한 국부성(locality)을 갖고 영상신호를 표현하므로 비정적(non-stationary) 신호의 분석에 유리하며 변환의 방식이 인간의 시각 인식 체계와 유사하다는 장점이 있다. 웨이브릿은 영상을 피라미드 구조로 표현하는 기존의 방식인 다해상도 분석(Multiresolution Analysis)과 대역분할 부호화(Subband Coding)를 하나로 통합한 이론이다. 변환된 각 부영상들은 저주파 대역부터 고주파 대역으로의 계층적 정보를 갖고 있으므로 이를 이용하여 적절한 부호화를 행할 수 있다

JPEG(Joint Photographic Experts Group) 2000은 기존 이미지 포맷인 JPEG의 2000년 버전을 말한다. JPEG 2000은 자연영상, 과학영상, 의료영상, 원격탐사영상, 문서영상, 그래픽영상 등과 같이 서로 다른 특성을 갖는 다양한 형태의 정지영상에 하나의 통합된 부호를 부여하는 것이다. 기존의 JPEG은 DCT를 사용하여 고압축시 이미지가 손상되는 단점이 있기에 JPEG 2000 위원회가 표준화를 추진중인 JPEG 2000은 압축률을 더욱 높이면서도 이미지 품질을 보존할 수 있는 웨이브릿변환을 이용한 압축기법을 채택하고 있다. 이 밖에 JPEG 2000은 JPEG을 통신과 인터넷 환경에 적합하도록 여러 가지의 요구사항을 제시하고 있다. JPEG 2000이 JPEG 기법을 계승한다는 관점에서 기존의 JPEG을 사용했던 많은 자원들과 호환성을 유지해야한다는 문제점에 기인하여 기존 JPEG과의 호환성 또한 요구되는 중요한 사항 중에 하나이다.

본 논문에서는 기존 JPEG의 DCT 변환을 JPEG 2000의 웨이브릿 변환으로 대체하고 그 밖의 부호화기는 기존의 JPEG 영상부호화기를 사용할 수 부호화기법을 제안한다. 제안된 부호화기법은 JPEG 부호화기법과 호환성을 유지하기 위해 설계되었다. 이를 위하여 웨이브릿 변환영역에서 각 부영역을 부호화 블록으로 재구성하는 과정이 필요하며, 각 대역에 대한 에러분석과 인간의 시각특성을 고려한 양자화 마스크를 유도하였다. 부가적으로 이와 같은 과정이 DCT 변환과 웨이브릿 변환의 변환 성능을 객관적으로 비교할 수 있는 방법이 되기도 한다.

II. Wavelet 변환 개요

하나의 신호를 여러 척도(scale)에서 관찰하고 그것을 여러 해상도(resolution)로 분해하지는 생각은 응용수학, 물리학, 공학의 각기 다른 많은 분야에서 독자적으로 출현하였던 웨이브릿은 최근에 이르러 단일화된 체계로 정비되었다. 영상인식(Computer Vision)분야의 다해상도 신호처리, 음성과 영상신호 압축에 이용되는 대역분할 압축기법(Subband Coding)^[1], 그리고 응용수학에서 이용되는 웨이브릿 수열 전개(Wavelet Series Expansion)등이 모두 서로 다른 관점에서 발전되어온 웨이브릿 이론의 한 분야로 볼 수 있다^[2].

웨이브릿 변환은 비정적신호(non-stationary signal)의 분석에 유리하며 STFT (Short Time Fourier Transform)과 Garbor Transform의 대응으로 이용된다^[3]. 기존의 STFT는 하나의 분석 윈도우를 이용하지만 웨이브릿 변환은 고주파에서는 짧은 윈도우, 저주파 신호에는 긴 윈도우를 이용하므로 융통성 있게 신호를 분석할 수 있다. 웨이브릿 이론의 이러한 특성을 줌인(zoom-in), 줌아웃(zoom-out) 특성이라 부르며 불연속(discontinuity)신호의 분석에 유리하다^[3].

80년대 중반, Morlet, Grossmann, Meyer등 여러 학자들에 의해 이론적인 체계를 갖추기 시작한 웨이브릿은 임의의 기저함수(basis function, mother wavelet)를 천이(translation), 확장(dilation)/수축(contraction)으로 파생되는 함수를 의미한다.

$L^2(R)$ 공간상에서 웨이브릿 함수 $\psi_{j,k}(x)$ 는 다음과 같다. 여기서 $2^{j/2}$ 는 정규화 요소이다.

$$\psi_{j,k}(x) = 2^{j/2} \Psi(2^j x - k), \quad j, k \in Z \quad (1)$$

식 (1)에서와 같이 정의된 함수의 집합 $\{\psi_{j,k}\}$ 가 $L^2(R)$ 의 직교기저인 경우를 직교 웨이브릿이라 한다.

$$\langle \psi_{j,k}, \psi_{l,m} \rangle = \delta_{j,l} \cdot \delta_{k,m} \quad j, k, l, m \in Z \quad (2)$$

식 (2)을 만족하는 웨이브릿 기저에 대해 모든 함수 $f \in L^2(R)$ 는 식 (3)으로 표시할 수 있다.

$$f(x) = \sum_{j,k=-\infty}^{\infty} c_{j,k} \psi_{j,k}(x) \quad (3)$$

식 (3)에서 웨이브릿의 계수는 식 (4)로 주어진다.

$$c_{j,k} = \langle f, \Psi_{j,k} \rangle \quad (4)$$

영상에 웨이브릿 변환을 적용하는 것은 2차원 신호에 웨이브릿을 도입하는 것이다. 일차원 웨이브릿으로부터 이차원의 웨이브릿으로의 확장은 각 스케일 함수 (scaling function)와 웨이브릿의 곱으로 구해지는 분리 가능한 웨이브릿(separable wavelet)을 이용하여 다해상도 필터뱅크(multidimensional filter bank)를 구성하는 것이다^{3, 4}. 2차원의 분리 가능한 웨이브릿은 다음 식으로 얻어진다.

$$g_s(n_x, n_y) = g(n_x) \cdot g(n_y) \quad (5)$$

$$h_s^{(1)}(n_x, n_y) = g(n_x) \cdot h(n_y) \quad (6)$$

$$h_s^{(2)}(n_x, n_y) = h(n_x) \cdot g(n_y) \quad (7)$$

$$h_s^{(3)}(n_x, n_y) = h(n_x) \cdot h(n_y) \quad (8)$$

III. 웨이브릿 변환영역에서의 에러분석

웨이브릿 변환을 이용한 영상 부호화 시스템에서 정확한 복원 에러를 분석하기 위하여, 우선 간단한 웨이브릿 변환인 1차원 모델^[1]에 대하여 그림 1과 같은 에러 분석을 실시하고, 이 결과를 영상신호의 2차원 형태로 확장한다. 그림 2의 입력신호인 $x_1(n)$, $x_2(n)$ 는 원 신호 $x(n)$ 을 웨이브릿 변환하여 얻은 결과로서 길이는 N 이며 각각 저역통과필터와 고역 통과필터를 통과한 신호이다. 만약 양자화 에러 또는 부호화 에러가 없다면 원 신호 $x(n)$ 은 완벽하게 복원될 수 있다. 그러나 실제적으로 신호를 압축하기 위해서는 최소한 양자화

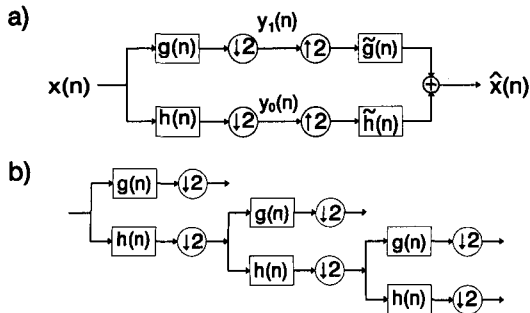


그림 1. 웨이브릿 변환의 개념도
Fig. 1. wavelet analysis system.

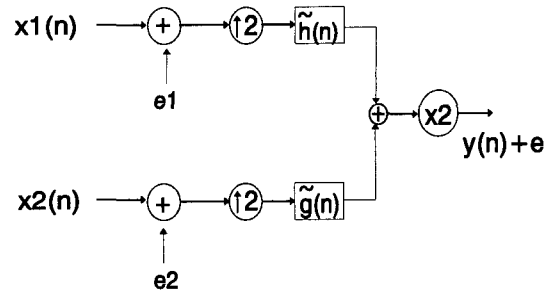


그림 2. 복원 에러를 분석하기 위한 1-D 모델
Fig. 2. 1-D Model for reconstruction error.

를 실시해야 하고 이때 발생한 에러는 복원신호 $y(n)$ 에서 노이즈 성분으로 해석할 수 있다.

$e_1(n)$, $e_2(n)$ 을 입력신호 $x_1(n)$ 과 $x_2(n)$ 에 각각 부가된 에러라고 하고, 편리성을 위하여 $E[e_1(n)] = E[e_2(n)] = 0$ 이라고 하자. 따라서 입력신호 $x_1(n)$ 과 $x_2(n)$ 의 자승평균오차(MSE) σ_1^2 , σ_2^2 는 다음 식 (9)로 주어진다.

$$\begin{aligned} \sigma_1^2 &= \frac{1}{N} \sum_n e_1^2(n) = E[e_1^2] \\ \sigma_2^2 &= \frac{1}{N} \sum_n e_2^2(n) = E[e_2^2] \end{aligned} \quad (9)$$

또 용이한 분석을 위하여 입력 신호는 주기적 확장을 사용하는 것으로 가정한다. 따라서 에러가 첨가된 입력신호를 주기적 확장에 의해 복원된 신호 $y(n)$ 은 다음 식 (10)과 같이 나타낼 수 있다.

$$y(n) + e(n) = 2 \sum_k \{ \tilde{h}(n-2k) [x_1(k) + e_1(k)] + \tilde{g}(n-2k) [x_2(k) + e_2(k)] \} \quad (10)$$

그러므로 복원 에러는 다음 식 (11)로 구해진다.

$$e(n) = 2 \sum_k \{ \tilde{h}(n-2k)e_1(k) + \tilde{g}(n-2k)e_2(k) \} \quad (11)$$

위 식에서 $e(n)$ 의 길이는 각 신호 $x_1(n)$ 과 $x_2(n)$ 이 2배로 확장(up_sampling) 되므로 $2N$ 된다. 이 결과로부터 복원신호의 MSE σ^2 은 다음과 같이 계산된다.

$$\begin{aligned} \sigma^2 &= E[e^2(n)] = \frac{1}{2N} \sum_n e^2(n) \\ &= \frac{2}{N} \sum_n \left\{ \sum_k \sum_l \tilde{h}(n-2k) \tilde{h}(n-2l) e_1(k) e_1(l) \right. \\ &\quad + \sum_k \sum_l \tilde{g}(n-2k) \tilde{g}(n-2l) e_2(k) e_2(l) \\ &\quad + \sum_k \sum_l \tilde{h}(n-2k) \tilde{g}(n-2l) e_1(k) e_2(l) \\ &\quad \left. + \sum_k \sum_l \tilde{h}(n-2l) \tilde{g}(n-2k) e_1(l) e_2(k) \right\} \end{aligned} \quad (12)$$

$$\begin{aligned}
 &= \frac{2}{N} \sum_k \sum_l \left\{ \sum_n \tilde{h}(n-2k) \tilde{h}(n-2l) e_1(k) e_1(l) \right. \\
 &\quad + \sum_n \tilde{g}(n-2k) \tilde{g}(n-2l) e_2(k) e_2(l) \\
 &\quad + \sum_n \tilde{h}(n-2k) \tilde{g}(n-2l) e_1(k) e_2(l) \\
 &\quad \left. + \sum_n \tilde{h}(n-2l) \tilde{g}(n-2k) e_1(l) e_2(k) \right\}
 \end{aligned}$$

사용된 필터가 정규직교 조건을 만족한다면 다음 식 (13)이 성립한다.

$$\begin{aligned}
 \sum_n \tilde{h}(n-2l) \tilde{h}(n-2k) &= \frac{1}{2} \sigma_{lk} \\
 \sum_n \tilde{g}(n-2l) \tilde{g}(n-2k) &= \frac{1}{2} \sigma_{lk} \\
 \sum_n \tilde{h}(n-2l) \tilde{g}(n-2k) &= 0
 \end{aligned} \tag{13}$$

식 (13)을 식 (12)에 대입하면 다음 식 (14)과 같은 복원신호의 MSE σ^2 가 구해진다.

$$\begin{aligned}
 \sigma^2 &= \frac{1}{N} \sum_k e_1^2(k) + \frac{1}{N} \sum_k e_2^2(k) = \sigma_1^2 + \sigma_2^2 \\
 &= \frac{1}{2N} \sum_k \{\sqrt{2}e_1(k)\}^2 + \frac{1}{2N} \sum_k \{\sqrt{2}e_2(k)\}^2 \\
 &= E \{[\sqrt{2}e_1]^2\} + E \{[\sqrt{2}e_2]^2\}
 \end{aligned} \tag{14}$$

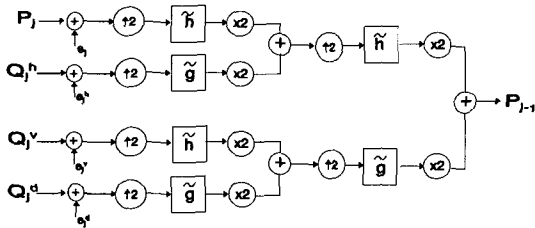


그림 3. 복원 에러를 분석하기 위한 2-D 모델
Fig. 3. 2-D Model for reconstruction error.

위의 결과 식으로부터 복원신호의 MSE는 두 입력신호의 MSE의 합과 같고, 각 입력신호의 에러 $e_1(k)$, $e_2(k)$ 는 복원신호에서 각각 $2^{1/2}$ 의 가중치를 갖는다. 식 (14)의 1차원 웨이브릿 변환에 대한 결과는, 비슷한 방법으로 2차원 웨이브릿 변환의 복원 신호로 확장될 수 있다. 그림 3과 같이 영상에 대하여 2차원 웨이브릿 변환을 실시한 경우 P_{j-1} 의 영상을 구하기 위해서는 P_{j1} , Q_{jh} , Q_{jv} , Q_{jd} 의 부영상이 필요하고 e_{j-1} , e_{j1} , e_{jh} , e_{jv} , e_{jd} 를 각각 부영상에서 발생한 양자화 에러라고 할 때 복원영상 P_{j-1} 의 MSE는 다음과 같이 계산된다.

$$\begin{aligned}
 \sigma_{j-1}^2 &= \sigma_j^2 + \sigma_{jh}^2 + \sigma_{jv}^2 + \sigma_{jd}^2 \\
 &= E \{[(2e_j)^2]\} + E \{[(2e_{jh}^k)^2]\} + E \{[(2e_{jv}^k)^2]\} \\
 &\quad + E \{[(2e_{jd}^k)^2]\}
 \end{aligned} \tag{15}$$

좀더 일반화시켜 전체 영상을 복원하였을 때의 MSE σ_0^2 를 구하면 다음과 같다.

$$\begin{aligned}
 \sigma_0^2 &= \sigma_j^2 + \sum_{j=1}^J (\sigma_{jh}^2 + \sigma_{jv}^2 + \sigma_{jd}^2) \\
 &= E \{[(2^j e_j)^2]\} \\
 &\quad + \sum_{j=1}^J \{E \{[(2^j e_{jh}^k)^2]\} + E \{[(2^j e_{jv}^k)^2]\} + E \{[(2^j e_{jd}^k)^2]\} \}
 \end{aligned} \tag{16}$$

위 식 (16)을 통하여 웨이브릿 변환 영상을 부호화할 때 각 부영상의 양자화 에러가 전체 영상을 복원했을 때 미치는 영향이 균일하지 못하고, 각 부영상의 위치에 따라 각각 다른 가중치를 갖는다. 즉 각 부영상의 양자화 에러는 해상도에 따라 2의 지수승으로 전체 복원영상에 영향을 미친다.

이 가중치를 적용하는 방법은, 웨이브릿 변환된 영상에 대하여 각 대역이 같은 조건이 되도록 정규화시키는 방법과, 웨이브릿 변환을 실시할 때 이 가중치를 각 필터에 미리 고려하여 변환 모델을 설정하는 방법이 있다. 두 가지 모두 같은 결과를 제공하고, 후자의 경우 웨이브릿 변환은 그림 4와 같이 구현된다.

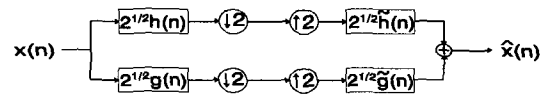


그림 4. 에러의 가중치를 고려한 웨이브릿 변환
Fig. 4. Wavelet transform for error weighting.

IV. 변환 영역의 변경된 Block 구조

JPEG에서 사용하는 양자화기 및 부호화기는 DCT를 이용한 8×8 단위 블록변환 방식에 적합하게 구성되어 있으므로 웨이브릿계수에 적용하기에는 적합하지 않다. 따라서 JPEG에서 사용하는 양자화기 및 부호화기의 성능 및 요구 조건에 잘 부합되도록 웨이브릿계수를 재구성해야만 한다. 이는 웨이브릿계수를 8×8 블록단위로 나누어야 하며, 이 블록의 계수들은 DCT 계수들이 보이는 통계적 특징(Energy compaction)과 유사한 성질을 가져야 한다. 웨이브릿은 블록 변환기법이 아니라 대역분할기법의 일종이지만 일반적인 블록변환 영상에서 나타나는 에너지 집중 특성과 유사하게 저주파 대역의 계수는 큰 값을 나타내며 고주파대역의 값은 작은 값을 나타낸다. 이는 그림 5에 나타난 것처럼, 블록변환 기법과 대역분할기법의 관계에 기인한다. 따라

서, 8×8 matrix를 구성할 때 저주파대역과 고주파대역 간의 자기유사성을 고려하여 단위 블록을 구성하였다. 웨이브릿 변환은 영상을 분할할 때 $1/2$ 만큼 간축되므로 가장 저주파대역을 제외한 부대역은 각 방향 성분의 부대역들에 대해 각각 공간적으로 $1:4$ 의 공간 크기를 갖는다. 따라서, 저주파대역의 계수와 고주파대역의 계수의 샘플링은 $1:4$ 의 비율로 이루어지게 되므로 8×8 matrix를 구성하기 위해서는 웨이브릿의 분할 횟수를 3으로 제한하였다. 웨이브릿변환에서 변환영역의 대역 분할된 형태와 부호화블록과 관계를 그림 5에 나타내었다.

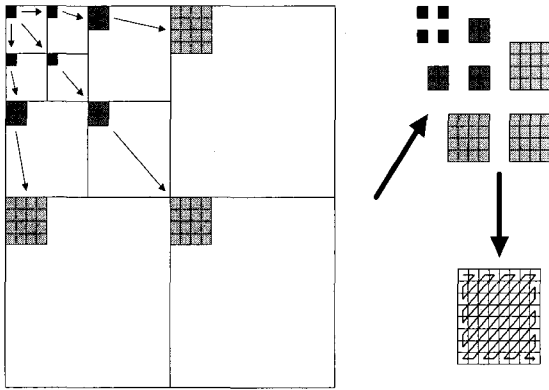


그림 5. 변환영역과 부호화블록의 관계
Fig. 5. The relationship between transformed image and coding block.

V. HVS

재생된 영상의 화질은 인간의 시각에 의해 평가되며 인간의 시각이 느끼지 못하는 영상정보를 위해 자원을 할당하는 것은 낭비를 초래하게 된다. 그러므로 인간 시각 체계(HVS: human visual system)를 이용한 영상시스템은 최적 시스템의 전제조건이 되어야 한다. 물론 영상처리 시스템에 HVS를 도입하는 것은 정량적인 화질 평가에서 어느 정도의 열화를 발생시키지만 궁극적인 화질평가는 인간의 시각에 의해 이루어지므로 HVS를 도입하는 것이 타당하다.

1. 공간주파수특성

공간 주파수 특성은 인간 시각장치가 모든 주파수 대역에서 균등한 민감도를 갖지 않고 각 주파수 대역마다 서로 다른 민감도를 나타내고 있다는 것이다. 따라서 가장 이상적인 영상 압축 기법은 인간의 시각 특

성에 근거하여, 인간이 민감하게 반응하는 주파수 대역에서는 양자화 에러를 최소화하고, 인간이 비교적 둔감한 주파수 대역에서 대부분의 양자화 에러가 발생하게 하는 것이다.

Mannos와 Sakrison^[8]은 푸리에 변환 영역에서 HVS를 이용한 영상 부호화 기법을 최초로 제시하였고, Nil^[9]은 DCT(discrete cosine transform) 영역에서 인간 시각 특성을 이용하기 위한 HVS 모델을 제안하였다. 이 두 가지 변환 기법은, 필터 बैं크(filter bank) 이론 측면에서 볼 때 모두 협대역(narrow band) 필터에 의하여 구현된다고 볼 수 있다. 따라서 단일 주파수에 대한 인간 시각의 민감도를 MTF(Modulation Transfer Function)로 구할 수 있다. 그러나 웨이브릿 변환은 고주파 영역에서는 광대역(wide band) 필터, 저주파수 영역에서는 협대역 필터로 구현된다고 볼 수 있다. 따라서 고주파 영역처럼 광대역 필터에 의해 변환되는 경우 단일 주파수에 대한 인간 시각의 민감도는 많은 오차를 내포하게 된다.

그러므로 웨이브릿 변환에 적합한 HVS 모델은 MTF처럼 단일 주파수로 인간의 시각 특성을 정의한 것이 아니고, 주파수 대역(frequency band)에 따른 인간의 시각 특성이 정의된 모델이 필요하다. 일부 웨이브릿 변환을 이용한 영상 부호화기에서 인간의 시각 특성을 반영하기 위하여, 각 주파수 대역의 민감도를 중심 주파수에 대응하는 민감도로 대체하여 사용하기도 했다^{[7][14]}. 그러나 고주파영역에서는 대역폭이 넓기 때문에, 중심주파수의 민감도를 이용하는 것보다는 기존의 MTF를 사용하되, 각 대역에 대응하는 MTF의 면적비를 이용하는 것이 타당하다.

2. 주파수와 웨이브릿 영역의 관계

웨이브릿 변환은 $L^2(R)$ 공간을 부공간인 W_j 로 분해하는 과정으로 이는 다음 식으로 표현할 수 있다.

$$L^2(R) = \bigoplus_{j \in Z} W_j := \dots \oplus W_1 \oplus W_0 \oplus W_{-1} \oplus \dots \quad (17)$$

식 (17)에서 웨이브릿 기저함수인 ψ 가 직교함수인 경우 모든 부공간 역시 서로 직교하게 되며 각각의 공간은 반복 필터 बैं크(iterative filter bank)의 j 번째 단에서의 대역폭과 부합한다. 제2장에서 언급된 바와 같이 웨이브릿 함수 ψ 는 고역통과 특성을 갖는 FIR(finite impulse response) 필터 $g(n)$ 에 의하여 구현되고, 척도 함수 ϕ 는 저역통과 특성을 갖는 FIR

필터 $h(n)$ 에 의하여 구현된다. 이때 저역통과 필터 $h(n)$ 와 고역통과 필터 $g(n)$ 는 다음 식 (18)과 같은 관계가 성립한다.

$$g(n) = (-1)^n h(n) \tag{18}$$

변환영역에서는 다음과 같다.

$$G(e^{j\omega}) = H(e^{j(\omega-\pi)}) \tag{19}$$

여기서 $H(e^{j\omega})$ 와 $G(e^{j\omega})$ 는 $h(n)$ 과 $g(n)$ 의 주파수 표현이다. 이때 $h(n)$ 과 $g(n)$ 이 직교 웨이브릿^[4, 10]이 되기 위한 조건 중의 하나는 아래의 식 (20)이다.

$$\sum_n h(n)h(n+2k) = \delta(k) \tag{20}$$

여기서 $\delta(k)$ 는 $k=0$ 일 때 1이고 $k \neq 0$ 이면 0이다. 이 조건은 반대역(half band)필터의 조건과 일치한다^{[4][10]}. 이는 모든 주파수에 대한 $H(e^{j\omega})$ 와 $H(e^{j(\omega-\pi)})$ 의 합은 1이 됨을 의미한다. 즉 $H(e^{j\omega})$ 는 반대역(half band) 주파수 $\pi/2$ 에 대하여 대칭을 이룬다. 결과적으로, 웨이브릿 변환은 반대역 필터 $h(n)$ 과 $g(n)$ 으로 구현되므로, 영상 주파수 영역에서 전체 대역폭을 1/2로 분해하여 나가는 과정으로 볼 수 있다. 이상적인 필터를 가정한 것은 일반적으로 웨이브릿 변환을 위하여 사용된 모든 FIR 필터는 이상적인 필터를 근사시킨 것이다. 실제 사용된 FIR 필터로부터 분할된 주파수 영역을 계산하는 것도 가능하지만, 여러 가지 웨이브릿 필터에 모두 공통적으로 적용할 수 있는 HVW(Human Visual Weight)를 구하기 위하여 이상적인 필터로 가정하여 각 주파수 대역의 민감도를 계산한다.

3. 주파수와 MTF의 관계

공간 주파수는 시야각 1°에 대한 정현파 패턴의 변화율로 정의되고 단위는 cpd(cycles per degree)로 나타내며, 이 공간 주파수에 대한 시각의 반응 특성을 표현하기 위하여 MTF가 사용된다. 지금까지 MTF모델은 Mannos, Sakrison^[8], Nill^[9]등 많은 사람들에 의하여 제안되었다. 인간 시각의 대역통과 특성을 나타내는 MTF는 3~10 cpd에서 최대값을 나타내며 고주파 영역에서는 지수 함수적으로 감소하는 것으로 알려져 있다. 일반적으로 MTF는 다음과 같이 나타내어진다.

$$H(f_s) = p(q + rf_s) \exp(-sf_s)^t \tag{21}$$

여기서 f_s 는 공간 주파수이며 p, q, r, s, t 는 상수이다. Sakrison의 경우 $p=2.6, q=0.192, r=s=0.114, t=1.1$ 로 된다. 인간의 시각 특성을 고려한 가중치를 반영하기 위해서 공간주파수(cpd) 영역에서 정의된 MTF를 영상주파수(cpp: cycle per pixel) 영역으로 변환하여야 한다. MTF의 공간주파수는 식 (22)과 같이 나타낼 수 있다.

$$f_s(\text{cycles/degree}) = f_i(\text{cycles/pixel})f_n(\text{pixels/degree}) \tag{22}$$

$$= \frac{f_i}{\sin^{-1} \left[\frac{1}{\sqrt{1+d^2(h^2+v^2)}} \right]}$$

여기서 f_i 는 영상을 푸리에 변환하여 얻어진 영상주파수를 의미하며 정규화 요소(factor) f_n 은 화면의 대각선 크기의 d 배의 거리에서 영상을 보았을 때 시각 1° 안에 포함되는 화소의 수이다. 단일 화면 대각선 크기의 4배 정도의 거리에서 영상을 보았다면 512×512 영상에서 f_n 은 50.55이다. Mannos와 Sakrison이 제안한 MTF를 주파수 영역으로 변환시킨 결과가 그림 6에 나타나 있다. 그림 6은 1-D MTF와 함께 동일한 주파수 영역에서의 웨이브릿 변환의 주파수 분할을 보여주고 있다.

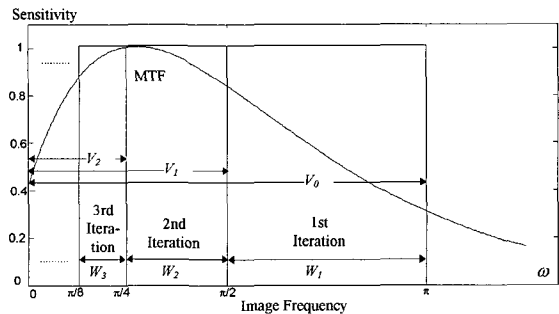


그림 6. 1-D MTF와 웨이브릿 변환의 주파수 분할 관계

Fig. 6. The relationship between 1-D MTF and frequency division of wavlet transform.

4. 주파수에 대한 양자화 가중치(HVW : human visual weight)

대역폭 $(0, \pi)$ 의 이산 입력신호에 고역 통과 필터 (high-pass filter) $g(n)$ 을 적용시키면 대역폭 $(\pi/2, \pi)$ 의 신호가 얻어진다. 반복 필터뱅크에서 $g(n)$ 의 계속적인 적용은 전체 대역폭을 반으로 분해하는 과정과 동일하며 j 번째 필터적용의 결과로 얻어지는 대역폭은

다음과 같다.

$$WB_j = \left(\frac{\pi}{2^j}, \frac{\pi}{2^{j-1}} \right) \quad (23)$$

주파수 대역의 민감도는 각 대역에서 주파수에 대한 MTF의 변적비(적분비)로써 정의된다. $g(n)$ 이 이상적인 필터라면 j 번째 필터뱅크의 HVW는 식 (24)에서 보너바와 같이 계산된다.

$$Weight_j = \frac{\int_{WB_j} dw}{\int_{WB_j} MTF(w) dw} \quad (24)$$

인간 시각 특성의 이방성(anisotropy)을 보상하기 위하여 MTF를 2차원(dimension)으로 확장시키는 이는 수직, 수평방향에서 대각선 방향으로 갈수록 민감도가 감소한다는 사실을 반영함으로써 구할 수 있다. 이 감소현상은 식 (25)에서와 같이 공간 주파수를 스케일링함으로써 모델링할 수 있다^[12].

$$f_s = \frac{(f_x^2 + f_y^2)^{1/2}}{s(\theta)} \quad (25)$$

where $s(\theta) = \cos(4\theta)(1-k)/2 + (1+k)/2$
 $\theta = \tan^{-1}(f_y/f_x)$

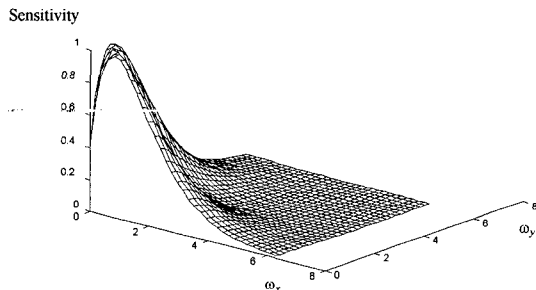


그림 7. 영상 주파수 영역에서의 2-D MTF
 Fig. 7. The 2-D MTF in image frequency domain.

각 성분(angular component)에 대한 정규화 요소 k 는 45° 각도에서 0.7 정도의 값이 되는데 이는 70%의 유효대역폭 감소효과를 의미한다. 그림 7은 정규화된 주파수 영역에서의 2-D MTF를 보여주고 있다. 모든 필터는 이상적인 웨이브릿 필터뱅크로 간주되었다. 본 연구에서는 512×512 영상에 대하여 3회 반복된 웨이브릿 변환과, 관찰자와 영상과의 거리는 화면 대각선 크기의 4배를 가정하였다. 이상적인 필터가 아닌 실제의 웨이브릿 필터에 사용되는 HVW는 반복되는 필터 बैं크

의 실제 주파수 응답으로 쉽게 구해진다. Sakrison^[8]의 MTF에 대한 각 주파수 대역의 가중치가 표 1에 제시되어 있다.

표 1. 웨이브릿 변환을 위한 각 주파수밴드의 가중치

Table 1. The weight of each frequency band for wavelet transform.

| Iteration | Band | Weight |
|-----------|------------------------------------|---------|
| 1 | $[(\pi/2, \pi), (0, \pi/2)]$ | 2.75945 |
| | $[(\pi/2, \pi), (\pi/2, \pi)]$ | 12.1359 |
| | $[(0, \pi/2), (\pi/2, \pi)]$ | 2.75945 |
| 2 | $[(\pi/4, \pi/2), (0, \pi/4)]$ | 1.18572 |
| | $[(\pi/4, \pi/2), (\pi/4, \pi/2)]$ | 1.93598 |
| | $[(0, \pi/8), (\pi/8, \pi/4)]$ | 1.18572 |
| 3 | $[(\pi/8, \pi/4), (0, \pi/8)]$ | 1.01871 |
| | $[(\pi/8, \pi/4), (\pi/8, \pi/4)]$ | 1.06558 |
| | $[(0, \pi/8), (\pi/8, \pi/4)]$ | 1.01871 |
| 3 | $[(0, \pi/8), (0, \pi/8)]$ | 1.19563 |

VI. 양자화

단위 블록으로 구성된 64의 웨이브릿계수를 그대로 부호화하여 복호화기에 전송하면 8×8화소의 블록 화상을 얻을 수 있으나, 높은 부호화 효율은 얻을 수 없으므로 양자화과정에서 복원영상의 화질과 부호화된 정보량을 조절해야 한다. JPEG에서는 양자화 테이블의 값을 변화시켜 DCT계수에 나누어줌으로써 화질과 부호량을 자유롭게 조절할 수 있는데, 이 양자화 테이블은 색 성분 단위로 설정이 가능하며 응용과 영상 특성, 출력장치에 따라 최적화하여 부호화 정보를 보내기 전에 수신측에 전송한다. JPEG에서는 인간의 시각특성을 이용해서 고주파 성분에는 큰 폭으로 양자화 함으로써 부호화 효율을 높이게 되어 있다. 본 연구에서도 인간 시각 특성을 이용해서 양자화를 하는데 웨이브릿의 특성상 각각의 계수에 양자화 가중치를 가하는 것이 아니라 계수가 위치한 대역별로 가중치를 나누어주어야 하므로 표 1에 나타난 값을 사용했다.

VII. 실험 및 결과

웨이브릿 변환은 Antonini^[6]가 제안한 9-7 탭 필터를

사용하였고, 주기 확장(periodic signal extension) 방법으로 웨이브릿 변환을 구현하였다. 이 필터는 비교적 적은 탭(tap)으로 PR(perfect reconstruction)조건과 선형 위상(linear phase)조건을 만족하는 쌍직교 필터이다.

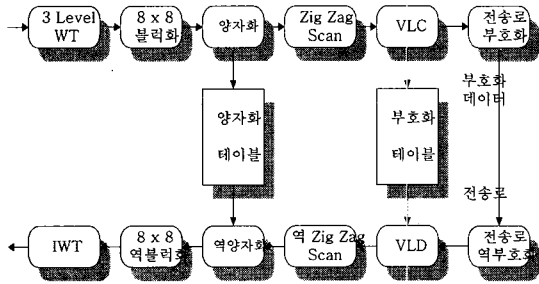


그림 8. 제안한 알고리즘의 블록도
Fig. 8. The block diagram of the proposed algorithm.

제안된 알고리즘의 블록도가 그림 8에 나타내었다. JPEG과 호환 가능한 부호화기를 위하여 인간시각특성을 고려하여야한다. 본 실험에서는 512×512 영상에 대하여, 관찰자가 영상으로부터 화면 대각선 길이의 4배 떨어진 거리에서 관찰한 경우를 가정하였다. 실제로 이 정도의 거리에서 영상을 관찰할 때 TV의 주사선을 느끼지 못한다^[8, 13]. 본 실험에서는 Sakrison^[8]이 제안한 MTF를 이용하여 각 주파수 대역에 대한 가중치를 구하였다. 시각 특성을 반영하기 위하여 표 1의 가중치로 각 해당 주파수 대역의 변환계수를 나누고, 복호화기에서는 이 가중치를 곱하여 인간의 시각 특성을 고려한 영상 부호화기를 구현하였다.

제안한 기법들의 신뢰성을 위하여 모든 성능평가는 실제 압축된 결과로부터 복원된 영상으로 평가하였다. 제안한 알고리즘의 성능이 입력영상에 무관함을 입증하기 위해서 서로 다른 두 종류의 표준영상을 사용했다. 하나는 비교적 저주파 성분이 많고 단순한 Lena영상이고 다른 하나는 Peppers 이다. 모든 실험은 화소당 8비트의 계조도를 갖고, 해상도는 512×512이다. 부호화 성능을 정량적으로 평가하기 위한 척도로 PSNR(peak signal to noise ratio)과 가중자승평균오차(WMSE: weighted mean square error)를 사용하여 영상에 대한 객관적 평가를 실시하였다^[9]. 사용된 식은 각각 다음과 같다.

$$PSNR = 10 \log_{10} \frac{255^2}{E \{ [f(i, j) - f'(i, j)]^2 \}}$$

$$WMSE = \frac{1}{N^2} \sum_i \sum_j \left[\frac{F(w_i, w_j) - F'(w_i, w_j)}{MTF(w)} \right]^2$$

여기서 $F(w_i, w_j)$ 와 $F'(w_i, w_j)$ 는 $f(i, j)$ 와 $f'(i, j)$ 의 푸리에 변환이며 $MTF(w)$ 는 주파수에 대한 가중치를 의미한다.

비록 PSNR은 실제 인간에 의해 인식되는 영상에 대한 정확한 평가를 의미하지는 못하지만 기존의 평가방법의 대부분이 PSNR을 사용하였기 때문에, 본 논문에서도 다양한 부호화 기법과 비교할 때에는 이 척도를 사용했다. 또한 WMSE가 인간 시각 특성들을 모두 다 반영하지는 못하지만, 최소한 MTF에 의해 모델링되는 공간 주파수에 대한 특성만은 반영할 수 있다.

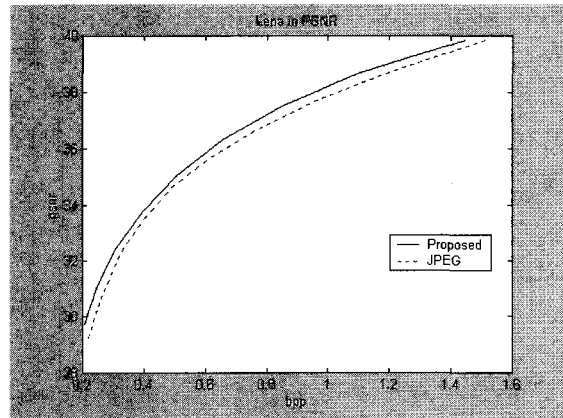


그림 9. Lena 영상에 대한 PSNR 비교
Fig. 9. PSNR for Lena image.

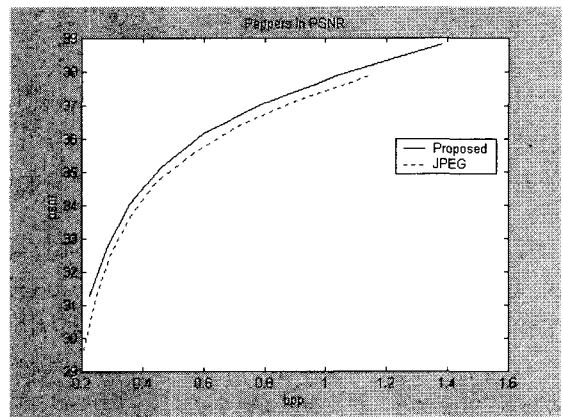


그림 10. Peppers 영상에 대한 PSNR 비교
Fig. 10. PSNR for Peppers image.

JPEG은 DCT를 이용한 부호화 기법으로 영상을 8×

8 블록(block) 단위로 변환하여 인간 시각 특성을 고려하여 설계된 8×8 양자화 행렬을 이용하여 양자화 한다. 그러나 입력 영상을 인접 블록과 무관하게 변환하였기 때문에 고압축시 인접 블록과 불연속된 영상이 나타난다. 이러한 현상을 블록 현상(block effect)이라고 부른다. 그러나 본 논문에서 제안한 부호화 기법은 웨이브릿 변환을 사용하기 때문에 변환 특성상 블록 현상은 발생되지 않는다. 일반적으로 인간 시각 특성을 반영한 기법은 PSNR과 같은 정량적인 화질 평가에서 어느 정도의 수치적 감소를 동반하지만 궁극적으로 인간의 시각에는 보다 좋은 화질을 제공한다. 따라서 인간 시각 특성을 이용한 기법과 이용하지 않은 기법을 PSNR 측면에서 비교하는 것은 의미가 없지만, 인간 시각 특성을 고려한 기법들끼리 PSNR 측면에서 비교하는 것은 타당한 비교가 된다. 그림 9와 그림 10에서 Lena 영상과 Peppers 영상에 대하여 JPEG의 결과와 인간 시각 특성을 고려한 제안한 부호화 기법을 PSNR 측면에서 비교하였다. 두 가지 경우 모두 인간의 시각 특성을 고려한 기법이다. 도표에 나타난 것과 같이 제안한 부호화 기법이 JPEG 보다 약간 우수함을 보이고 있고, 또한 제안한 부호화 기법은 블록 현상을 제거할 수 있다는 장점이 있다.

그림 11과 그림 12에서 Lena 영상과 Peppers 영상에 대하여 JPEG의 결과와 인간 시각 특성을 고려한 제안한 부호화 기법을 WMSE 측면에서 비교하였다. 도표에 나타난 것과 같이 Lena 영상과 Peppers 모두 낮은 비트율에서는 제안한 부호화 기법이 JPEG 보다 약간 우수함을 보이고 있다. 즉 PSNR 측면과, WMSE 측면

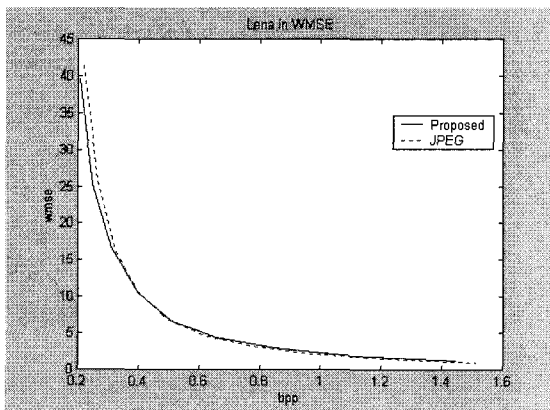


그림 11. Lena 영상에 대한 WMSE 비교
Fig. 11. WMSE for Lena image.

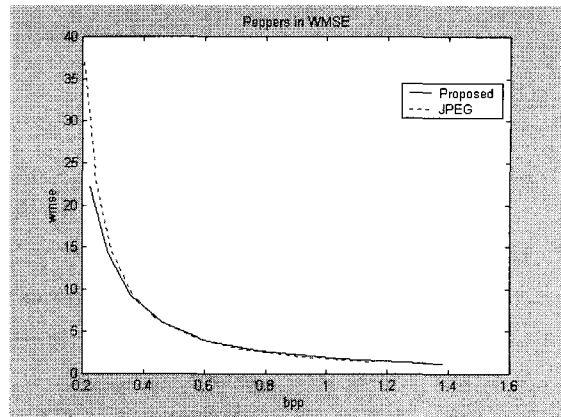


그림 12. Peppers 영상에 대한 WMSE 비교
Fig. 12. WMSE for Peppers image.

에서 모두 우수한 성능을 나타내고 있다. 높은 비트율에서 WMSE 측면에서는 거의 차이가 없음을 보였다.

일반적으로 웨이브릿 변환의 장점을 최대한 살리기 위해서는 5회 이상 반복변환을 실시하여야 최대의 변환성능을 나타냈다. 그러나 본 부호화기에서는 JPEG과 호환성을 위해 단 3회만의 반복변환을 실시하였다. 따라서 웨이브릿 변환의 최대성능이 아닌 상태에서 JPEG 부호화 결과와 비교가 되었다. 결과적으로 일반적인 영상에 대해서는, 웨이브릿 변환이 DCT 변환보다 우수하다는 것을 의미한다.

VIII. 결 론

본 연구에서는 웨이브릿 변환을 이용하여 JPEG 부호화기와 호환 가능한 형태의 부호화기를 제안하였다. JPEG 부호화기에서 채택된 DCT 변환기법은 낮은 비트율과 고압축에서 블록효과가 발생한다는 문제점이 있으나 웨이브릿 변환을 이용하여 이와 같은 단점을 극복할 수 있었다. JPEG 부호화기에서 변환 기법을 단순히 웨이브릿으로 대체하기 위하여 웨이브릿 변환영역에서 블록을 재구성하여 부호화 하였다. 이때 효과적인 양자화 기법을 위해 웨이브릿 변환영역에서 에러신호를 분석하여, 인간의 시각특성을 고려한 양자화 마스크를 제안하였다. 실험 결과 제안된 양자화 마스크가 타당함을 입증하였고, DCT와 웨이브릿 변환을 도입한 부호화 기법에 적용하여 단순히 변환기법의 부호화 성능을 간접적으로 비교할 수 있었다.

참 고 문 헌

- [1] A. Grossman and J. Morlet, "Decomposition of hardy function into square integrable wavelets of constant shape," SIAM J. Math., vol. 15, pp. 723-736, 1984.
- [2] Y. Meyer, "Orthogonal wavelets," in Wavelet, Time-Frequency Method and Phase Space, Springer-Verlag, pp.21-37, 1989.
- [3] P. P. Vaidyanathan, "Quadrature mirror filter banks, M-band extensions and perfect reconstruction techniques," IEEE ASSP Mag., vol.4, pp.4-20, 1987.
- [4] P. P. Vaidyanathan, Multirate System and Filter Banks, Prentice-Hall, New Jersey, 1993.
- [5] P. J. Burt and E. H. Adelson, "The laplacian pyramid as a compact image code," IEEE Trans. Commun., vol. 31, pp.523-540, 1983.
- [6] M. Antonini, M. Barlaud, P. Mathieu and I. Daubechies, "Image coding using wavelet transform," IEEE Trans. Image Proc., vol.1, pp.244-250, Apr. 1992.
- [7] A. S. Lewis and G. Knowles, "Image Compression Using the 2-D Wavelet Transform," IEEE Trans. Image Processing, vol.1, pp.244-250, Apr. 1992.
- [8] J. L. Mannos and D.J. Sakrison, "The effects of a visual fidelity criterion on the encoding of images," IEEE Trans. IT., vol.IT-20, pp.525-536, 1974.
- [9] N. B. Nill, "A visual model weighted cosine transform for image compression and quality assessment," IEEE Trans. Commun., vol.COM-33, pp.551-557, 1985.
- [10] Ali N. Akans and Richard A. Haddad, Multiresolution Signal Decomposition, Academic Press. INC. 1992.
- [11] Jie CHEN, Shuichi ITOH and Takeshi HASHIMOTO, "Scalar quantization noise analysis and optimal bit allocation for wavelet pyramid image coding," ICEIE Trans. Fundamentals, vol. E76-A, No.9, pp.1502-1514, Sep. 1993.
- [12] J. Sullivan, R. Miller and G. Pios, "Image halftoning using a visual model in error diffusion," J. Opt. Soc. Am. A, vol.10, pp.1714-1724, 1993.
- [13] B. Chitprasert and K.R. Rao, "Human visual weight progressive image transmission", IEEE Trans. COMM vol. 38, No. 7, pp. 1040-1044, 1990.
- [14] M. G. Perkins and T. Lookabaugh, "A psychophysically justified bit allocation algorithm for subband image coding system", Proc. ICASSP, pp. 1815-1818, 1989.

저 자 소 개



金容圭(正會員)

1985. 2 : 연세대학교 전자공학과 (학사). 1987. 2 : 연세대학교 전자공학과 (석사). 1995. 8 : 연세대학교 전자공학과 (박사). 1987 - 1995. 9 : 삼성전자. 1996 - 현재 : 성결대학교 컴퓨터학부 정보통신전공



金度圭(正會員)

1987. 2 : 경북대학교 전자공학과 (학사). 1989. 2 : 한국과학기술원 전기전자공학과 (석사). 1993. 2 : 한국과학기술원 전기전자공학과 (박사). 1993 - 1995. 2 : 정보통신부 정보통신정책실 통신사무관. 1995 - 현재 :

성결대학교 컴퓨터학부 정보통신전공



趙錫八(正會員)

1976. 2 : 광운대학교 전자통신학과 (학사). 1987. 8 : 한양대학교 전자통신학과 (석사). 1992. 8 : 경희대학교 전자공학과 (박사). 1975 : Control Data Corp. Computer SE. 1984 : 삼성전자 정보통신연구소 연구실장(수석연구원). 1995 - 현 : 성결대학교 컴퓨터학부 정보통신전공