

흑백 정지 영상의 시각 하중 DCT 부호화

(Visual-Weighted DCT Coding for Monochrome Still Images)

黃 在 政*, 梁 根 鎬**, 李 門 浩***

(Jae Jeong Hwang, Guen Ho Yang, and Moon Ho Lee)

要 約

정지 화상의 시각 하중 변환에서 인간의 눈에 의해 감지되는 화질을 결정하기 위해서는 최저 공간 주파수와 명암 대조비를 중요한 요소로 삼아야 하는데 지금까지는 DCT의 주파수 영역 처리에 따라 시각의 주파수 특성만이 주요 시각 특성으로서 적용되어 왔다. 제안된 하중 DCT 기법은 8 x 8 블록내에서 4 [cyc/deg]의 최저 주파수를 갖고 DC 계수에 명암 특성을 적용하였다. 성능 평가에 적합한 오차 기준을 적용하였으며 복원된 화질이 비하중 기법이나 다른 하중 기법에 비해 우수함을 증명하였다.

Abstract

HVS based transform coding techniques for still images take the lowest spatial frequency and luminance contrast sensitivity as the most crucial factors in determining the image quality perceived by human eye. Only the frequency weighting has been main issue. The proposed Weighted Discrete Cosine Transform (WDCT) has the lowest frequency of about 4 [cyc/deg] in 8x8 subblock with the contrast-weighted DC coefficient. The error criteria have been rederived to adapt the contrast weighting for performance evaluation. The proposed scheme has produced higher quality images compared to the unweighted and other schemes.

I. 서 론

인간의 시각 시스템과 영상 신호와의 관계를 살펴 보는 것은 궁극적으로 인간이 영상을 관측하며 신호 자체의 성질이 시각 시스템에 의하여 변화된 모습으로 받아들여

지기 때문이다. 우리가 영상 신호와 관련하여 시각 시스템에서 고려해야 할 사항은 첫째, 명암의 세기에 따른 감도의 차이 즉 콘트라스트 감도이다. Weber-Fechner의 법칙에 의하면 배경의 밝기에 대해 약 2% 이상의 명암 변화량이 있을 때 변화된 화면이라고 인식한다.^[1] 또한 아주 낮은 레벨이나 높은 레벨에서는 선형성을 잃어 대체로 대수적인 명암 특성을 가지며, 불균일 양자화와 유사한 성질을 가진다. Mach 밴드 효과에 의하면 세기가 일정하게 증가하는 계단 파형에서 인간의 눈이 미분 작용을 일으켜 급격히 변화하는 지점에서 더욱 큰 감도를 가진다.^[2] 이런 것들이 명암에 관련된 것이고, 둘째로, 주파수 반응을 일컫는데 화면내 공간 주파수에 따라 인식의 차이가 있다.^{[3][4]} 공간 주파수는 화면의 높이에 대해 6배의 거리에서 관찰하였을 때 단위 각도당 진동하는 주파수를

*正會員, 群山大學校 通信學科

(Dept. of Telecommun. Kunsan Nat'l Univ.)

**正會員, 全北大學校 病院 醫工科

(Dept. of Biomed. Eng., Chonbuk Univ. Hospital)

***正會員, 全北大學校 情報通信工學科

(Dept. of Telecom. & Inform. Eng., Chonbuk Univ.)

接受日字 : 1992年 3月 2日

의미하며, 정현과 격자를 이용하여 실험해 보면 매우 높거나 낮은 주파수에서는 감도가 떨어지며 중간 주파수에서 감도가 높다. 셋째로, 시간적 (temporal) 인 주파수 즉 화면간 (interframe) 주파수에 따라 공간 (spatial) 주파수 반응이 달라지는 Robson^[5], Kelly^[6] 등의 모델을 이용할 수 있다. 넷째로, 색 신호의 인식에 대해서도 비직선성이 존재한다.^[7] 만일 본 연구와 같이 흑백 정지 화상을 취급하면 첫째, 둘째의 현상들을 영상부호화에 이용하면 충분하다.

1974년 Mannos와 Sakrison은 흑백 정지 화상의 비선형 명암 특성 및 주파수 반응에 대해 광범위한 실험을 통하여 변조 전달 함수를 구하였다.^[8] 이것이 시각 특성을 영상 부호화에 응용하는 계기가 되었으며, 단순한 자승 오차 기준보다는 시각적 오차 기준을 설정함으로써 인간의 눈에 적합한 화질을 재생하도록 하였다.^[9,11] 80년대 이후 몇몇 연구자들은 블록 효과 (block effect) 와 모기 효과 (mosquito effect) 등이 시각적 화질 저하 요인으로 등장하는 변환 부호화에 시각 하중을 적용하고자 하였다. 이들은 주로 압축 효율이 높은 이산 역현 변환 (DCT)에 대해 연구하였는데, DCT는 신호의 높은 상관도를 낮게 만드는 비상관률이 98.05%로 높고, 고속 알고리즘의 개발로 실시간 처리가 가능하므로 국제 화상 처리 표준 알고리즘으로 채택되었지만, 단순한 DCT 자체로는 1세대 부호화의 한계를 벗어날 수 없으므로, 지금까지의 연구는 양자화, 비트 할당등에 적응성을 부여하거나 시각의 주파수 특성을 적용하여 중복된 정보를 제거하고자 하였다. 실제로 JPEG (Joint Photographic Expert Group) 은 몇가지 제안 중에 ADCT (Adaptive Discrete Cosine Transform) 를 가장 우수한 것으로서 권장하고 있다.^[12] 시각적으로 민감한 양자화 오차를 줄이기 위해 시각 시스템 특성에 의한 양자화 행렬을 적응적으로 선택 사용하므로 시각 시스템 모델에 대한 연구가 필요한 것이다. Ericsson,^[13] Nill,^[14] Ngan,^[15] Rao,^[16] Hwang^[17] 등이 주파수 모델을 변환부호화에 적용하였다. 이들은 감도가 최대로 되는 최대치 주파수를 3 - 8 [cyc/deg]로 변화시키면서 자승 오차 또는 하중 오차가 최소로 되는 주파수를 근거로 각자의 모델을 설정했기 때문에, 실제 인간이 느끼는 공간 주파수보다는 사용한 부호화 방법과 오차 기준에 따라 달라지는 결과를 가져왔다. 기존의 시각 하중은 두가지 문제를 안고 있는데, 첫째로, 공간 주파수 개념이 모호하고 이론적 배경이 생략된 채 나름대로의 주파수 모델만을 적용하였다. 둘째로, 명암 특성 또한 중요하여 감도가 낮은 주파수에서도 명암의 세기가 크면 인식할 수 있다는 것이 고려되어야 한다. 즉 흑백 정지 화상에서 주파수 및 명암 레벨 특성 모두 부호화에 적용되어야 한다. 이러한 문제의 해결 방법으로서

화면과 관측 거리의 관계로부터 DCT의 최저, 최대 주파수를 결정하여야 하며, 명암 레벨의 양자화 과정에서 일반적으로 이용되는 비선형 양자기를 시각적으로 명암 대조비에 의하여 적용하면 된다.

본 논문에서는 최대치 주파수 및 모델을 Mannos와 Sakrison의 것으로 제한하였는데, 왜냐하면 이들의 결과는 부호화와 관계 없이 실험적으로 인간의 시각 특성을 관측한 것이기 때문에, 부호화에 따라 결정된 모델과는 달리 객관적으로 인정되어야 한다. 다만 화면과 인간의 눈 사이에 존재하는 공간 주파수는 블록으로 나누어 처리할 때 블록 크기에 따라 달라지는 점을 이용하여, DCT 계수의 최저 공간 주파수를 블록 크기에 따른 최저 주파수로서 적용하였다. 명암 특성을 보상하기 위해 역시 DCT 영역에서 블록 평균을 의미하는 첫번째 계수에 대수적 하중을 하여 효율이 향상되면서 블록 효과가 줄어들음을 보였다.

II. 주파수 하중 DCT의 역사

그간 변환 부호화에 시각 시스템을 적용하고자 하는 연구가 몇몇 연구자에 의해 행해졌다. 1974년 Mannos와 Sakrison의 연구가 이분야에서 처음으로 이루어졌는데, HVS 모델의 변수를 설정하기 위해 광범한 실험을 하였으며, 영상이나 HVS를 등방성으로 가정하여 가우시안 정보원의 최적 부호화를 시도하기위해 다음과 같은 변조 전달 함수 (MTF)의 일반형을 설정하였다.^[8]

$$H(f) = a(b+cf) \exp(-cf)^d \quad (1)$$

단 f 는 공간 주파수 [cycles/degree]를 의미하며 a , b , c , d 는 상수이다. 이 상수들을 변화해가면서 MTF의 모양과 최대 주파수를 가변하였는데, 512×512 개의 화소를 갖는 영상을 시각 1° 당 65 화소가 보이는 지점에서 관측하여 이산 푸리에 변환에 의해 주파수 영역으로 변환하였다. 푸리에 계수들은 MTF에 의해 하중되며 임계치보다 큰 계수가 선정되어 송신되었다. 수신기에서는 역과정 이 수행되며 이 결과는 주관적으로 평가되었다. 최종적인 MTF는 다음과 같다.

$$H(f) = 2.6(0.0192 + 0.114f) \exp(-(0.114f)^{1.1}) \quad (2)$$

이것은 7.9 [cyc/deg]에서 최대 감도를 가진다.

한편 최근 연구자들은 이산 역현 변환에 대한 MTF를 연구하였다. 1984년 Ericsson은 상기 모델에 기초하고 DCT 기본 함수의 주파수 영역 오차에 의하여 표 1과 같은 하중 행렬을 제시하였다.^[13] Ericsson의 하중 행렬은

표 1. Ericsson에 의한 주파수 하중 계수
Table 1. Frequency weighting coefficients by Ericsson.

| | | | | | | | |
|------|------|------|------|------|------|------|------|
| 1.00 | 1.00 | 1.00 | 1.00 | 0.95 | 0.86 | 0.74 | 0.50 |
| 1.00 | 1.00 | 1.00 | 0.98 | 0.93 | 0.83 | 0.72 | 0.49 |
| 1.00 | 1.00 | 0.97 | 0.93 | 0.86 | 0.76 | 0.65 | 0.45 |
| 0.91 | 0.90 | 0.87 | 0.81 | 0.74 | 0.65 | 0.56 | 0.38 |
| 0.74 | 0.73 | 0.70 | 0.65 | 0.59 | 0.52 | 0.45 | 0.32 |
| 0.56 | 0.55 | 0.53 | 0.49 | 0.45 | 0.40 | 0.35 | 0.25 |
| 0.40 | 0.40 | 0.38 | 0.36 | 0.33 | 0.29 | 0.26 | 0.18 |
| 0.19 | 0.18 | 0.18 | 0.17 | 0.16 | 0.14 | 0.13 | 0.09 |

화면의 4:3 비율을 고려하여 수평 및 수직 샘플링 주파수를 각각 60, 40 [pel/deg]로 하고 설정한 등방성 모델에서 저주파 부분 9계수의 값이 균일하게 1로 정해지는데, 계수상의 분산값이 저주파 부분에 몰린다는 일반적인 부호화 개념에서 비롯되어 시각 시스템의 주파수 특성에 적절하지 못하다.

1985년 Nill은 다음과 같은 MTF를 제안하였다.^[14]

$$H(f) = (0.2 + 0.45f) \exp(-0.18f) \quad (3)$$

이 함수는 약 5.1 [cyc/deg]에서 최대 감도를 가진다. HVS모델을 DCT에 적용하기 위해서는 DCT의 특징인 원 영상의 대칭적 확장성을 고려하여야 한다. 관측자는 이러한 변화된 영상을 보는 것이 아니므로 DCT영상은 물리적 의미를 상실한다는 것이 그의 주장으로서, 이를 극복하기 위해서는 MTF를 보정하는 함수 |A(f)|를 다음과 같이 제안하였다.

$$|A(f)| = \left[\frac{1}{4} + \frac{1}{\pi^2} \left[\log_e \left(\frac{2\pi f}{\alpha} + \left(\frac{4\pi^2 f^2}{\alpha^2} + 1 \right)^{1/2} \right) \right]^2 \right]^{1/2} \quad (4)$$

단, α 는 11.636 deg⁻¹로 주어진다. 이리하여 수정된 MTF는 다음과 같다.

$$H'(f) = |A(f)| H(f) \quad (5)$$

이 식의 최대주파수는 6.9 [cyc/deg]로 높아진다. 1986년 Ngan^[15]은 Nill의 보정함수를 이용하여 전달함수의 최대 주파수가 3 [cyc/deg]임을 보였는데, 그 전달함수는 다음과 같다.

$$H(f) = (0.31 + 0.69f) \exp(-0.29f) \quad (6)$$

H(f)에 보정함수 |A(f)|를 곱하면 결과는 최대 주파수가 4.1 [cyc/deg]로 된다. DCT계수를 Zig-zag 주사하여 저전송률에서 만족스런 결과 영상을 얻었다.

1988년 Rao^[16]는 DCT의 계층적 부호화에서 식 (7)의 MTF를 제시하였고 이에 따른 하중 행렬은 표 2와 같다.

$$H(f) = 2.46(0.1 + 0.25f) \exp(-0.25f) \quad (7)$$

이 함수는 3.75 [cyc/deg]에서 최고치 주파수가 되는데, 각각의 상기 모델들은 직류 교차점과 최고치 주파수 그리고 고주파에서 약간 다른 감쇠 특성을 갖는다. 각자 사용한 양자기 및 부호기에 따라 변하는 오차를 기준으로 MTF를 설정함으로써, 인간의 공간 주파수 감도가 부호기마다 달라야만 이론적으로 성립하는데, 최종 영상을 보는 인간의 눈은 달라질 수 없는 일정한 것이다.

표 2. Rao에 의한 주파수 하중 계수
Table 2. Frequency weighting coefficients by Rao.

| | | | | | | | |
|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.4942 | 1.0000 | 0.7023 | 0.3814 | 0.1856 | 0.0849 | 0.0374 | 0.0160 |
| 1.0000 | 0.4549 | 0.3085 | 0.1706 | 0.0845 | 0.0392 | 0.0174 | 0.0075 |
| 0.7023 | 0.3085 | 0.2139 | 0.1244 | 0.0645 | 0.0311 | 0.0142 | 0.0063 |
| 0.3814 | 0.1706 | 0.1244 | 0.0771 | 0.0425 | 0.0215 | 0.0103 | 0.0047 |
| 0.1856 | 0.0845 | 0.0645 | 0.0425 | 0.0246 | 0.0133 | 0.0067 | 0.0032 |
| 0.0849 | 0.0392 | 0.0311 | 0.0215 | 0.0133 | 0.0075 | 0.0040 | 0.0020 |
| 0.0374 | 0.0142 | 0.0103 | 0.0067 | 0.0040 | 0.0022 | 0.0022 | 0.0011 |
| 0.0160 | 0.0175 | 0.0063 | 0.0047 | 0.0032 | 0.0020 | 0.0011 | 0.0006 |

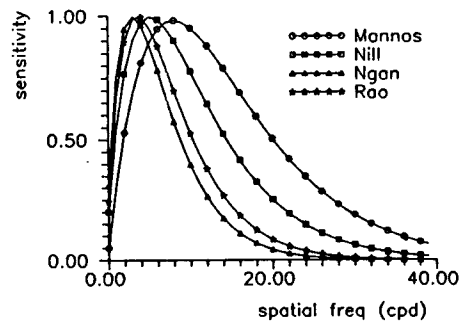


그림 1. 다양한 시각 모델의 비교
Fig. 1. Comparison of various MTFs.

Ⅲ. 주파수 및 명암 특성의 모델링

1. 공간 주파수 모델링

화면내에서 변화하는 신호 성분은 화면을 보는 인간의 눈으로부터 떨어진 거리에 따라 공간적으로 주파수가 다시 정의된다. 이것을 공간 주파수 [cyc/deg]라 하며 그림 2의 θ_x 로 주어지는 시각 (visual angle) 에서 단위 각도당 진동하는 주파수를 말한다.

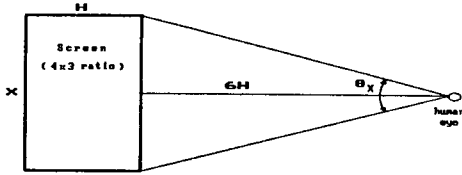


그림 2. 화면과 인간의 공간 접속
Fig. 2. Spatial relation of TV screen to human eye.

공간 주파수 f_r 는 하나의 화소가 갖는 정규화 주파수 f_n 과 단위 각도당 샘플링 주파수 f_s 로 식 (8)과 같이 분리된다.^[16]

$$f_r[\text{cyc/deg}] = f_n[\text{cyc/pel}] f_s[\text{pel/deg}] \quad (8)$$

여기서 f_n 은 DCT 합수 특성상 우수 확장으로부터 2N 포인트인 점을 고려하여 정의한다. 즉,

$$f_n[\text{cyc/pel}] = \frac{K}{2N}, k=0, 1, \dots, N-1 \quad (9)$$

그림 2는 화면 비율 4 : 3 이며 720×576 개의 화소를 갖는 CCIR 601 화면을 화면 높이의 6배 거리에서 관측하도록 하였다. 중심부에서 멀어질수록 공간 주파수는 밀해지는 것이 사실이나 화소의 등간격 그리고 등방성 (isotropic) 을 가정하여 다음과 같이 계산된다. 수평 방향 샘플링 주파수 f_{sx} 는

$$f_{sx} = 720/\theta_x [\text{pel/deg}] \quad (10)$$

로 주어지며 수직 방향 샘플링 주파수 f_{sy} 는

$$f_{sy} = 576/\theta_y [\text{pel/deg}] \quad (11)$$

이다. 이것들이 등방성으로 분포된다면 방사상 주파수

(radial sampling frequency) f_{sr} 는

$$f_{sr} = (f_{sx}^2 + f_{sy}^2)^{1/2} [\text{pel/deg}] \quad (12)$$

로 결정되며 이것이 식 (8)의 f_s 에 해당된다. f_n 는 하나의 화소가 갖는 정규화 주파수로서 공간 영역에서 0.5 [cyc/pel] 로 가정한다. 따라서 화면내 최대 공간 주파수는 다음과 같이 결정된다.

$$f_r = f_n f_{sr} = 41.5 [\text{cyc/deg}] \quad (13)$$

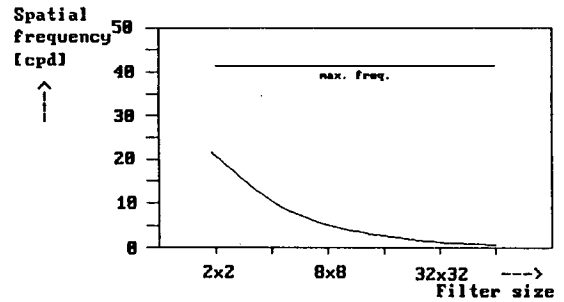


그림 3. 블록내에 존재하는 공간 주파수
Fig. 3. Possible spatial frequency in subblocks.

식 (13)는 화면내에 존재하는 최대 주파수를 의미하며 직교 변환 또는 영상 필터링등 일정한 크기와 블록으로 나누어 처리할 때 블록내 공간 주파수는 그림 3과 같이 존재한다. 즉 최대 주파수는 일정한 것에 비해 최저 주파수는 블록 크기에 따라 반비례하므로 영상의 블록 처리시에 중요한 변수로 고려되어야 한다.

한편 디지털 영상의 512×512 화소를 가정하고 시각 1°당 64 화소가 보이는 위치에서 관측하면,^[16] 최대 주파수는 32 [cyc/deg]이며 블록 크기 8×8 에서 처리 가능한 최저 주파수는 약 4 [cyc/deg]로서 제로 주파수는 이 블록내에 존재하지 않는다. 이것이 주파수 모델에서 고려되어야 할 오프셋이며, DCT 영역에서 직류 계수는 공간 주파수상으로 제로가 아니라 4 [cyc/deg]로서 모델에 적용된다.

하나의 블록내에서 처리 가능한 공간 주파수 분포를 앞에서 보였고, Mannos와 Sakrison의 실험 결과를 만족하면서 얻은 하중계수를 표 3에 제시하였다.

2. Weber-Fechner 법칙의 수식화

Weber-Fechner의 법칙에 의하면 화소 주위의 배경에

표 3. 제안된 주파수 하중 계수

Table 3. Proposed frequency weighting coefficients.

| | | | | | | | |
|--------|--------|--------|--------|--------|--------|--------|--------|
| 0.9314 | 0.9724 | 0.8513 | 0.6666 | 0.4856 | 0.3365 | 0.2248 | 0.1459 |
| 0.9724 | 0.9069 | 0.7684 | 0.5990 | 0.4383 | 0.3058 | 0.2056 | 0.1343 |
| 0.8513 | 0.7684 | 0.6432 | 0.5027 | 0.3710 | 0.2616 | 0.1778 | 0.1173 |
| 0.6666 | 0.5990 | 0.5027 | 0.3963 | 0.2963 | 0.2118 | 0.1459 | 0.0975 |
| 0.4856 | 0.4383 | 0.3710 | 0.2963 | 0.2249 | 0.1632 | 0.1142 | 0.0774 |
| 0.3365 | 0.3058 | 0.2616 | 0.2118 | 0.1632 | 0.1205 | 0.0857 | 0.0590 |
| 0.2248 | 0.2056 | 0.1778 | 0.1459 | 0.1142 | 0.0857 | 0.0620 | 0.0434 |
| 0.1459 | 0.1343 | 0.1173 | 0.0975 | 0.0774 | 0.0590 | 0.0434 | 0.0309 |

따라 감지 가능한 콘트라스트 비율 $\Delta L/L$ 은 가해주는 휘도의 세기에 따라 변화하나 Graham은 휘도 세기 0.1 - 1000 [mL]에서 2% 정도임을 실험하였다. 즉,

$$\Delta L/L = K \Delta B \approx 0.02 \quad (14)$$

K는 일정한 상수이며 ΔB 는 가해진 휘도 세기이다. 이것은 오차를 수반하는 영상 처리에서 감지 가능한 최소 오차를 규정하는 기준이 된다. 즉 낮은 레벨에서 허용되는 오차는 높은 레벨의 그것보다 더 작게 되므로 낮은 레벨에서 밀하게 양자화하는 것이 콘트라스트 감도와 일치하게 된다. 일반적으로 가우시안 등의 확률 분포에 따라 비직선 양자화 하는 것과 유사하나, 시각적 Weber-Fechner의 법칙에 준하여 다음과 같이 유도한다. 하중된 명암 x에 대한 복원치 y를 다음과 같은 지수함수적 관계식으로 설정한다.

$$y(x) = a (\exp(bx)-1) \quad (15)$$

여기서 y(x)는 명암 특성을 데이터 영역이 아닌 DCT 영역에서 이미 구해진 DC 성분으로서 두 영역의 에너지 보존성에 의해 논리에 합당하며, 블록 효과의 주원인으로 나타나는 DC 성분의 오차를 줄이는 효과를 가져온다. 최대 레벨에서 불변의 조건을 설정하면 식 (16)과 같다.

$$a (\exp(bx_m) - 1) = x_m \quad (16)$$

수신측에서 선형 특성을 얻기 위해 식 (15)의 역을 취하면 식 (17)과 같다.

$$x(y) = 1/b \ln(y/a+1) \quad (17)$$

변환 계수를 256 레벨로 양자화하면 최저치는 0이며 최대치는 255이므로 최저치와 최대치에서 입력값을 변동

시키지 않는 조건하에 상수를 결정한다. 상수 a, b의 결정은 식(15)를 미분하여 얻은 $y'(x) = ab \exp(bx)$ 로부터 최솟치 0와 최댓치 x_m 에서 다음과 같이 정의된다.

$$\begin{aligned} y'(0) &= ab \\ y'(x_m) &= ab \exp(bx_m) \end{aligned} \quad (18)$$

여기서 미분치의 증가비 V를 정의하면

$$V = y'(x_m)/y'(0) = \exp(bx_m) \quad (19)$$

식 (19)를 식 (16)에 대입하면

$$a = \frac{x_m}{V-1}, \quad b = \frac{\ln V}{x_m} \quad (20)$$

식 (20)의 a, b 값을 식 (17)에 대입하면

$$x(y) = \frac{x_m}{\ln V} \ln \left(\frac{y}{x_m} (V-1) + 1 \right) \quad (21)$$

이것을 y에 대해 미분하면

$$\begin{aligned} \frac{dx(y)}{dy} &= \frac{x_m}{\ln V} \frac{1}{y+x_m/(V-1)} \\ dx &= \frac{x_m}{\ln V} \frac{dy}{y+x_m/(V-1)} = \frac{x_m}{\ln V} \frac{dy}{y'} \end{aligned} \quad (22)$$

결국 식 (22)는 Weber-Fechner의 법칙과 유사하여 다음과 같이 놓을 수 있다.

$$\frac{dy}{y'} = \frac{\ln V}{x_m} dx \approx 0.02 \quad (23)$$

여기서 $\ln V/x_m$ 와 dx 는 각각 식 (14)에서 상수 K와 ΔB 를 의미한다. 최댓치 x_m 은 255로 제한되고 시각은 $x_m=255$ 에서 0.02의 감도를 조건으로 하면 $255 \times 0.02=5.1$ 의 Weber 비율 (fraction)이다. 식 (23)으로부터 $dx=5.1$ 로 하면 $V \approx 3$ 에서 등식에 만족된다. 따라서 $V=3$ 은 Weber-Fechner의 법칙을 고려한 명암 하중 상수이며 이때 $a=127.5$, $b=4.3 \times 10^{-3}$ 을 갖는다. DC 계수에 대한 명암 하중은 식 (21)에 의해 이루어지며 그림 4와 같은 비직선 특성을 갖는다. 역하중은 식 (15)에 의해 그림 4의 'reverse' 부분에 도시되어 있다.

DCT 계수중 직류 계수는 블록 전체의 평균 밝기를 의미하며 이것을 오프셋으로 하여 교류 신호가 변하게 된다. 명암 특성을 적용하는 것은 Weber-Fechner의 법칙

과 같이 비선형적 대조 (contrast) 특성을 만족시키는 것이다. 변환 부호화에서 명암의 변화 즉 부호화로 인해 변화된 출력 신호는 주로 블록 평균이 변하는 것이므로, 직류 계수는 주위 배경에 대하여 인간이 느끼는 명암 대조 계수이다. 따라서 비선형 대수 함수에 의한 명암 하중을 부호화 전에 행하고 복호화 전에 다시 지수 함수적 보상 특성으로 재하중한다. 그림 4에 식 (20)의 $V=3$ 인 경우에 대한 하중 및 역하중 특성을 도시한다.

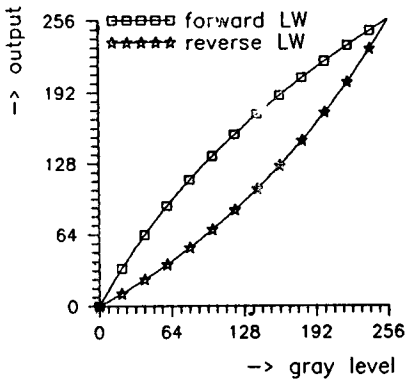


그림 4. 명암 하중 특성
Fig. 4. Luminance weighting characteristics.

IV. 실험 및 결과

실험에 사용한 부호기는 시각 부호화의 성능을 쉽게 비교하기 위해 적응 기법을 적용하지 않은 전형적 기법을 채택하였으며, 256×256 화소, 256 레벨의 데이터를 처리한 후 256 레벨로 균일 양자화하였다. 명암 레벨에 대한 하중은 앞절에서 제시한 비직선 함수를 DCT 계수의 DC 성분에 하중함으로써 주위 배경에 대한 콘트라스트 감도를 향상시키는 효과를 얻었다. 전체 부호기의 계통도가 그림 5에 도시된다. 부호기에서 채널 잡음 및 부호화는 고려하지 않았으며, 양자화 및 비트 할당 과정에서 특이성을 배제하기 위해 가장 일반적인 직선 양자화와 분산 에너지에 따른 고정 비트 할당을 하였다.

하중에 의한 부호기에서는 오차 측정 기준이 중요하다. 실험에 사용한 세가지 오차 기준은 다음과 같이 정의된다.

- ① 전형적인 자승 오차 기준 (PMSE)은 $10 \log(255^2 / \text{MSE})$ 이다.
- ② CMSE는 Weber-Fechner의 법칙에 의한 콘트라스트 감도를 의미하며, PMSE 기준에서 감지 가능

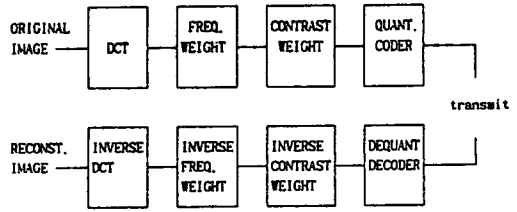


그림 5. 하중 DCT의 계통도
Fig. 5. Blockdiagram of the weighted DCT

한 2% 이상의 오차만을 측정하였다.

- ③ WMSE는 주파수 영역에서 하중에 의한 오차 기준이다. 일관성을 위해 제안한 주파수 모델로써 일률적으로 하중하였다.

표 4 부터 표 7 까지 네개의 영상에 대한 실험 결과를 보였다. 주파수 하중(FW)과 명암 하중(LW)을 부가한 결과는 표 4 부터 7 가지에서 보듯이 FW-LW 로 표시하였으며, FW는 주파수 특성만을 하중한 경우이다. 명암을 하중하여도 PMSE 와 WMSE 는 별 차이가 없는데, 양자화 오차가 적고 직류 계수의 주파수 하중이 크지 않기 때문이다. 그러나 감지 가능한 큰 세기의 오차만을 고려하는 CMSE는 하중한 경우가 하지 않은 것보다 우수한 성능을 나타냈다.

표 4. GIRL. DAT의 복원 오차 (1.0 bpp)
Table 4. Reconstruction error of GIRL. DAT (1.0 bpp).

| | | no weight | Ericsson | Rao | proposed |
|------|-------|-----------|----------|--------|----------|
| PMSE | FW | 29.374 | 29.509 | 29.045 | 31.115 |
| | FW-LW | 29.376 | 29.509 | 29.053 | 31.117 |
| CMSE | FW | 48.578 | 48.392 | 36.212 | 45.750 |
| | FW-LW | 51.056 | 50.906 | 40.365 | 51.734 |
| WMSE | FW | 34.729 | 34.831 | 36.011 | 36.222 |
| | FW-LW | 34.733 | 34.835 | 36.045 | 36.228 |

표 5. LENNA. DAT의 복원 오차(1.0 bpp)
Table 5. Reconstruction error of LENNA. DAT(1.0 bpp).

| | | no weight | Ericsson | Rao | proposed |
|------|-------|-----------|----------|--------|----------|
| PMSE | FW | 25.660 | 25.755 | 25.778 | 27.327 |
| | FW-LW | 25.660 | 25.755 | 25.781 | 27.327 |
| CMSE | FW | 41.666 | 41.510 | 33.162 | 40.016 |
| | FW-LW | 51.853 | 51.268 | 37.345 | 52.196 |
| WMSE | FW | 31.680 | 31.788 | 33.899 | 33.558 |
| | FW-LW | 31.680 | 31.789 | 33.560 | |

표 6. CRONK.DAT의 복원 오차(1.0 bpp)
Table 6. Reconstruction error of CRONK.DAT(1.0 bpp).

| | | no weight | Ericsson | Rao | proposed |
|------|-------|-----------|----------|--------|----------|
| PMSE | FW | 27.944 | 28.230 | 29.243 | 30.898 |
| | FW-LW | 27.928 | 28.213 | 29.250 | 30.899 |
| CMSE | FW | 45.573 | 45.711 | 38.016 | 44.419 |
| | FW-LW | 41.979 | 42.124 | 40.595 | 49.728 |
| WMSE | FW | 32.381 | 32.485 | 33.777 | 33.857 |
| | FW-LW | 32.343 | 32.446 | 33.794 | 33.859 |

표 7. COUPLE.DAT의 복원 오차(1.0 bpp)
Table 7. Reconstruction error of COUPLE.DAT(1.0 bpp).

| | | no weight | Ericsson | Rao | proposed |
|------|-------|-----------|----------|--------|----------|
| PMSE | FW | 31.981 | 32.027 | 30.297 | 32.812 |
| | FW-LW | 31.984 | 32.030 | 30.308 | 32.816 |
| CMSE | FW | 43.001 | 42.875 | 34.226 | 42.103 |
| | FW-LW | 45.957 | 45.769 | 37.678 | 44.981 |
| WMSE | FW | 37.160 | 37.241 | 37.114 | 38.240 |
| | FW-LW | 37.169 | 37.250 | 37.160 | 38.254 |

실험 결과 영상으로서, 첫째 명암 성분에 하중한 것과 하중하지 않은 것을 그림 6에 제시하였으며, 이는 표 4에서 PMSE와 WMSE는 거의 같지만 CMSE가 하중에 의해 향상됨을 증명하는 결과이다. 둘째로 그림 7은 명암과 주파수 하중의 결과를 보이기 위해 비하중 기법, 다른 연구자의 주파수 하중법에 명암 하중을 더한 것, 그리고 본 논문에서 제안한 결과 영상을 제시하였다.

V. 결 론

본 논문에서는 흑백 정지 화상의 시각 특성에서 고려해야 할 두가지 요소를 모두 적용한 하중 DCT 기법이 제안되었다. 시각 시스템을 영상의 변환 부호화에 적용하기 위해 실제 공간 주파수 특성을 이용하여 DCT 계수상에서 하중 행렬을 제안하였고 기존의 다른 연구 결과와 비교하였다. 그들은 시각의 공간주파수 개념을 인정하면서도 변환계수의 DC성분에 의해 발생하는 오차를 중시하여 임의로 하중 행렬을 제안하였으나, 본논문에서는 가장 객관적으로 알려진 최대 감도의 공간주파수 8 [cyc/deg]를 유지한 채 화면과 인간의 공간적 관측에 의한 공간주파수 개념을 첨가하였고, 수치상 큰 차이가 없는 하중 계수를 가지고 영상에 따라 MSE는 1 - 3 dB 정도



(a)



(b)

그림 6. 명암 하중 효과를 설명하는 결과 영상 (GIRL.DAT, 1.0 bpp)
(a) 비명암하중 기법
(b) 명암하중 기법

Fig. 6. Result images describing the luminance weighting effect (GIRL.DAT, 1.0 bpp).
(a) non-weighting method,
(b) weighting method.

의 개선된 결과를 보였다. 주파수 특성뿐 아니라 명암 특성 또한 중요하게 적용되었는데, 직류 계수에 비직선 명암 하중을 행한 결과 최대 12 dB 이상의 향상을 가져왔다. 이것은 감지 가능한 오차의 세기가 명암 하중에 의해 줄어드는 것을 의미하며, MSE 또는 하중 오차 기준 WMSE에서 큰 차이가 없음에도 불구하고 시각적 화질에 중요한 요소임을 증명한다. 이것이 콘트라스트 오차 기준(CMSE)이며, 시각의 명암 대조 특성을 보이기 위해 변환 계수 영역에서 DC성분이 2% 이상의 오차를 가지는 즉 Weber-Fechner 법칙의 응용이다. 낮은 MSE의 영상이라 할지라도 블록 현상으로 인해 거슬리는 결과를 설명할 수 있는 측정 기준의 하나이다. 이러한 시각 특성을 DC성분에 대해 시각적 비선형 양자화함으로써 시각 부호화에 적절한 결과를 얻었다. 결과로 제시한 영

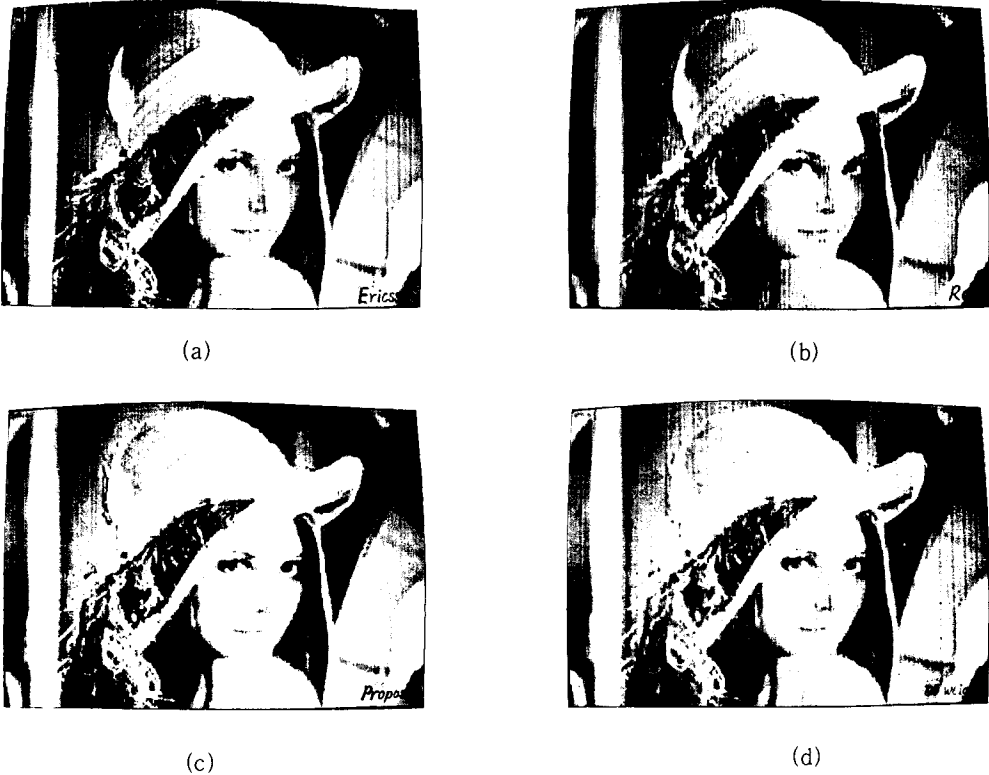


그림 7. 명암 및 주파수 하중에 의한 복원 영상 (LENNA, DAT, 1.0 bpp)

- (a) 비하중 (b) Ericsson
- (c) Rao (d) 제안된 기법

Fig. 7. Reconstruction images by luminance frequency weighting

- (LENNA, DAT, 1.0 bpp).
- (a) no weighted, (b) Ericsson,
- (c) Rao, (d) proposed.

상들은 시각 특성의 효과를 분명히 보이기 위해 가시적으로 열화된 상태에서 제시되었으며, 실험 과정에서는 보다 우수한 저전송률 결과 영상들이 출력되었다.

参 考 文 献

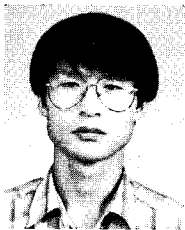
- [1] E. L. Hall, *Computer image processing and recognition*, Academic Press, 1979.
- [2] T. G. Stockham, "Image processing in the context of a visual model," *Proc. IEEE*, vol. 60, no. 7, pp. 828-842, July 1972.
- [3] J. J. DePalma and E. M. Lowry, "Sine-wave response of the visual system. II. Sine-wave and square-wave contrast," *J. Opt. Soc. Am.*, vol. 52, no. 3, Mar. 1962.
- [4] M. B. Sachs, "Spatial-frequency channels in human vision," *J. Opt. Soc. Am.*, vol. 61, no. 9, Sep. 1971.
- [5] J. G. Robson, "Spatial and temporal contrast-sensitivity functions of the visual system," *J. Opt. Soc. Am.*, vol. 56, pp. 1141-1142, 1966.
- [6] D. H. Kelly, "Visual contrast sensitivity," *Optica Acta*, vol. 24, no. 2, pp. 107-129, 1977.
- [7] W. Frei and B. Baxter, "Rate-distortion coding simulation for color images," *IEEE Trans. Commun.*, vol. COM-25, no. 11, pp. 1385-1392, Nov. 1977.
- [8] J. L. Mannos and D. J. Sakrison, "The effects of a fidelity criterion on the encoding of

- images," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 525-536, July 1974.
- [9] F. Kretz, "Subjectively optimal quantization of pictures," *IEEE Trans. Comm.*, vol. COM-23, pp. 1288-1292, Nov. 1975.
- [10] M. Miyahara and T. Fujimoto, "A measure of picture quality coded images-considerations of local errors-," *PCS '87*, pp. 16-17, Sweden, June 1987.
- [11] J. J. Hwang and M. H. Lee, "Image data compression techniques by human visual system," *J. KITE*, vol. 15, no. 2, pp. 148-154, Apr. 1988.
- [12] H. Yasuda, "A draft ISO standard for coded representation of still pictures," *Signal Processing of HDTV*, L. Chiariglione (ed.), Elsevier Science Pub., pp. 1-11, 1988.
- [13] S. Ericsson, "Frequency weighted interframe hybrid coding," *TRITA-TTT-8401*, Royal Inst. of Tech., Sweden, Jan. 1984.
- [14] N. B. Nill, "A visual model weighted cosine transform for image compression and quality assessment," *IEEE Trans. Comm.*, vol. COM-33, pp. 551-557, June 1985.
- [15] K. V. Ngan, K. S. Leong and H. Shingh, "Cosine transform coding incorporated human visual system model," *SPIE Fiver LASE 86*, Cambridge, MA, vol. 707, pp. 165-177, Sept. 1986.
- [16] B. Chitprasert and K. R. Rao, "Human visual weighted progressive image transmission," *ICCS '88, Int. Conf. on Commun. Sys.*, Singapore, no. 17. 1, Nov. 1988.
- [17] J. J. Hwang and M. H. Lee, "A simple human visual weighted Hadamard transform image coding," *J. KITE*, vol. 26, no. 4, pp. 98-105, Apr. 1989.

 著 者 紹 介

黄 在 政 (正會員) 第26卷 第4號 參照
 현재 군산대학교 통신학과
 교수

李 門 浩 (正會員) 第26卷 第4號 參照
 현재 전북대학교
 정보통신학과 교수



梁 根 鎬 (正會員)
 1963年 9月 6日生. 1989年 2月 전북
 대학교 전자공학과 (학사). 1989年
 3月~1991年 8月 전북대학교 대학원
 전자공학과 (석사). 1990年 8月~현
 재 전북대학교병원 의공학실 연구
 원. 주관심분야는 데이터통신 및 영

상신호처리 등임.