

## 시각 하중 이산여현변환 영상부호화

Image Coding of Visually Weighted  
Discrete Cosine Transform

李 門 浩\*  
Moon Ho, Lee

朴 周 用\*\*  
Ju Yong, Park

### Abstract

Utilizing a cosine transform in image compression has several recognized performance benefits, resulting in the ability to attain large compression ratio with small quality loss. Also, various models incorporating Human Visual System (HVS) to Discrete Cosine Transform (DCT) scheme are considered. Using the exact frequency components of DCT basis function, the optimum modulation transfer function (MTF) is obtained analytically. The errors at a block boundary which is important factor in transform coder are criteria for error measurement. The HVS weight coding results in perceptually higher quality images compared with the unweighted scheme.

### 1. 서 론

1970년대 이후 반도체, 마이크로 프로세서 및 디지털 신호처리 기술 등이 발전함에 따라 디지털 통신 기술이 급속도로 발전하였다. 이러한 통신기술의 발전으로 음성신호 뿐 아니라 화상 신호 등도 디지털화 되어가고 있는 추세이다. 그러나 화상 신호의 디지털화가 이룩되는 데는 많은 어려움이 있으며, 그 중의 하나가 엄청난 데이터량을 어떻게 처리할 것인가가 문제이다. 이런 점에서 데이터 압축 문제가 거론되고 있고 여러가지 방법들이 제안되어 왔으며, 이들의 기본 원리는 화소간의 공간적 상관관계가 크다는 점에 착안하여 그 상관관계를 줄여줌으로써 데이

타를 압축하고, 화면내와 화면간에 변화된 정보량만 고려하여 부호화한다.

최근에는 인간의 시각시스템(HVS)을 이용한 부호화 방법이 제안되고 있으며, 이 방법은 인간의 눈이 특정한 공간 주파수에서 더욱 민감하고, white 영역보다는 black 부분의 변화에 더욱 민감하다는 사실을 이용한다[2]. 본 논문에서는 이러한 인간의 시각시스템을 변환 부호화의 일종인 이산여현변환(DCT)에 이용하여 영상데이터를 압축하며, 이때 인간의 눈이 보다 민감하게 느끼는 부분은 다른 부분 보다 더욱 세밀하게 부호화 한다. 하중치에 따라서는 문턱치 보다 높은 계수값만 유효하고 이들만이 수신단에 전송된다.

\* 通信(電氣通信)技術士, 工博, 全北大 情報通信工學科 副教授

\*\* 全北大 情報通信工學科

이 논문은 1988년도 문교부 지원 한국 학술진흥재단의 자유공모과제 학술 연구 조성비에 의하여 연구되었음.

## 2. HVS 모델을 이용한 영상 압축

### 1. 인간의 시각 특성

HVS는 시신경계통을 나타낸 것으로써 인간의 신경계통은 가장 강력한 컴퓨터 즉, 뇌에 의해 운용되는 복잡한 통신망을 이루고 있으며 neuron이라 불리는 신경세포에 의해 시각적 자극이 전달된다. 은하계내의 별의 수 만큼이나 많은 대략  $10^{11}$ 개의 neuron이 있고, neuron의 크기는  $5\sim 100\ \mu\text{m}$  정도이며, 몸체에는 축색돌기(axon)와 수지상 돌기(dendrite)가 가지처럼 뻗어 있다[1]. 정보는 neuron과 neuron 사이에서 전기 화학적으로 전달되는데 두 개의 neuron이 만나는 부분을 연결부(synapse)라 하고 송신부와 수신부는 각각 presynaptic neuron과 postsynaptic neuron이다. Presynaptic neuron에서 발생된 정보는 마치케이블에서 전기 신호가 흐르듯이 축색돌기를 통과하여 화학적 형태로 postsynaptic neuron에 저장된다. 이 화학 성분은 새로운 전기 신호를 만들게 되며, 하나의 neuron 사이에는 수천의 presynaptic neuron과 postsynaptic neuron이 결합되어 마치 neuron은 거대한 집적회로를 연상케 하는데 잘 설계된 집적회로일지라도 fan-out수는 불과 수십으로 제한되는 것과 비교할 수 있으며, 전달 신호의 형태는 100mv의 진폭과 1ms의 주기를 갖는 일련의 펄스열로서 자극의 세기에 따라 주파수 변조를 행하여 직렬, 병렬 그리고 제환등으로 연결된 neuron 사이에 통신이 행해진다. Neuron의 작용은 흥분과 금지상태의 두 가지로서 전자는 Postsynaptic neuron에 펄스를 발생시키며, 후자는 그것들을 금지억압하여 그림 1과 같은 mach band 효과, 시각주파수 반응의 대역통과 특성, 눈의 경계 및 선검출의 원리가 된다. 인접한 신경 상호간에 강한 흥분을 억제하는 것을 측방억제(lateral inhibition)라 하는데 Furman은 그림 2와 같이 모델을 설정하고  $g_1=e_1-b_{12}g_2$ ,  $g_2=e_2-b_{21}g_1$ 로써 설명하고 있다[12]. 배경의 밝기를  $L_b$ , 목표물 주변의 밝기를  $L$ 이라 할 때 목표물의 명암 차이를 느끼는 정도(contrast sensitivity)에 대한 Weber's Law에 따르면 차이가 어느

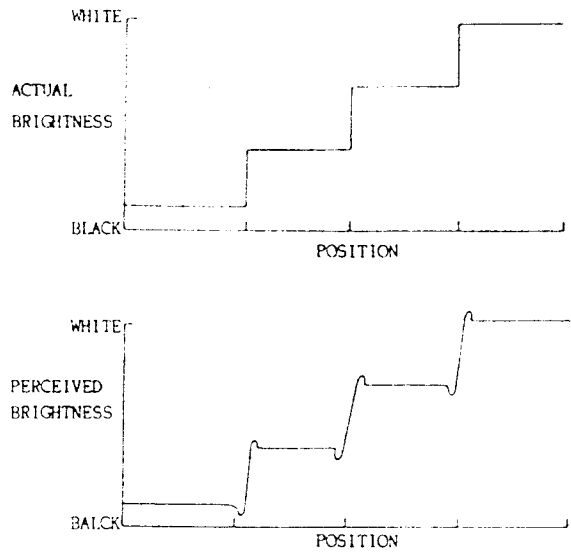


그림 1. Mach Band 효과

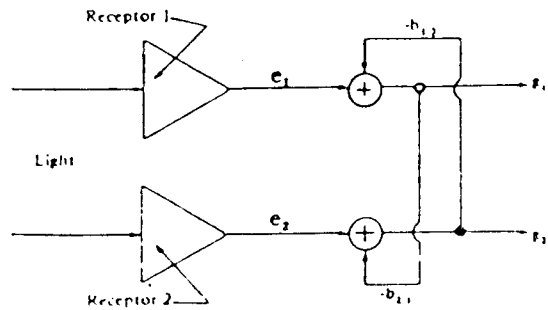


그림 2. 측방 억제 모델

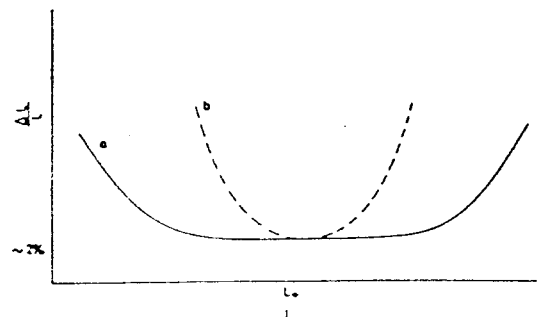


그림 3. 밝기에 따라 감지 가능한 Contrast 비율  
(a)  $L_b=L$  (b)  $L_b=L_o$

중간 주파수 대역에서 약 2% 이상이면 감지할 수 있음을 그림 3과 같이 보여준다. 그림에서 점선 부분은  $L_b=L_o$  즉, 배경과 목표물 주변이

다른 것을 의미하고 배경과 주변이 일치할 때 감도는 보다 넓은 범위에 걸쳐 일정하게 된다.

## 2. 시각 시스템의 모델링

비선형적 특성을 보이는 시각 시스템에서 고려되어야 할 사항들은 첫째, 명암의 세기에 따른 감도차이(contrast sensitivity) 둘째, 공간 주파수에 따른 감도의 차이 셋째, 시간적(temporal)으로 변화되는 성분, 마지막으로 칼라성분의 비선형성을 들 수 있는데 흑백정지화상의 경우에 명암과 공간주파수 성분을 주로 다루게 된다. 명암의 세기에 관하여는 선형적 분포를 갖는 패턴 영상에 대해 지수함수적 감도를 갖는다는 것이 알려져 있고 공간주파수 감도에 대하여는 Sakrison은 HVS를 isotropic 모델로 가정하여 다음과 같은 형식의 MTF(modulation transfer function)를 제시하였다[3].

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

단,  $f$ 는 공간주파수(cycles/degree)이며  $a, b, c, d$ 는 상수이다. 시각적 감도가 최대가 되는 peak 주파수는  $d$ 를 1로 하였을 때 다음과 같이 표현된다.

$$fp = 1/c - a/b \text{ (cycles/degree)} \quad (2)$$

상수들을 변경하면 MTF의 peak 주파수가 변하며  $512 \times 512$  영상을 DFT 영역으로 변환하여 주파수 감도를 측정하였다. 푸리에 계수들은 MTF에 따라 하중되며 미리 정해진 임계치 보다 큰 계수들을 전송하여 수신단에서 역과정을 거쳐 주관적 평가를 하였다. 이 때  $f=8$  cycles/degree에서 peak 주파수를 갖는 MTF는 다음과 같다.

$$H(f) = 2.6(0.0192 + 0.114f) \exp(-0.114f)^{1.1} \quad (3)$$

최근에 Nill은 실제영상 신호가 isotropic 하지 않고 푸리에 변환을 통한 주파수 해석은 관측시간 이후의 것이어야 한다는 점을 고려할 때 고주파계수 성분의 의미를 강조하는 보정항으로 MTF는 다음과 같다[6].

$$H(f) = (0.2 + 0.45f) \exp(-0.18f) \quad (4)$$

이 함수는 5 cycles/degree에서 peak 주파수를

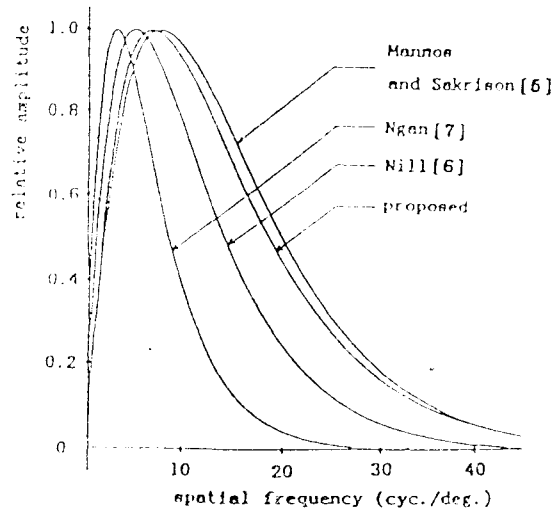


그림 4. HVS 모델의 비교

가지며 실제적으로 영상부호화에 적용하지는 않았다. Ngan은 이것을 실제 부호화에 적용하였으며 기존의 MSE 측정을 통하여 식 (5)와 같이 3 cycles/degree의 peak 주파수를 제안하였다.

$$H(f) = (0.31 + 0.69f) \exp(0.29f) \quad (5)$$

이와 같이 몇몇 HVS를 모델링한 MTF를 살펴보면 peak 주파수는 서로 다르다. 사용하는 코딩, 양자화 및 비트할당 알고리즘에 따라 주관적 평가에 의존하거나 MSE 측정에 의한 결과는 서로 다른 값을 갖게 되지만 그림 4에 도시한 것처럼 대략 3-8cycles/degree에서 인간의 눈이 최대 감도를 갖는 것으로 나타나고 있으므로 이점에 유의하여 부호화했을 경우보다 우수한 화질을 얻을 수 있다.

## 3. DCT의 주파수 스펙트럼

1차원 DCT는 다음과 같은 기본함수를 갖는  $N$ 포인트 시퀀스로 표시할 수 있다.

$$U_1(n) = \sqrt{\frac{1}{N}} V(n) \\ U_i(n) = \sqrt{\frac{2}{N}} \cos[\pi(n-1/2)(i-1)/N] V(n) \quad (6)$$

여기서 참함수(window function)는 다음과 같다.

$$V(n) = \begin{cases} 1, & n=1, 2, \dots, N \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

DCT 기본함수의 푸리에 변환은 다음과 같다.

$$U_1(q) = \sqrt{\frac{1}{N}} V(q)$$

$$U_i(q) = \sqrt{\frac{2}{N}} \exp\left(-j2\pi q \frac{N+1}{2}\right) \exp\left(j\pi \frac{i-1}{2}\right) \left\{ D\left(q - \frac{i-1}{2N}\right) + (-1)^{i-1} D\left(q + \frac{i-1}{2N}\right) \right\},$$

$$i=2, 3, \dots, N$$

$$\text{where } D\left(q - \frac{i-1}{2N}\right) = \frac{\sin\{(q - (i-1)/2N)N\pi\}}{\sin\{(q - (i-1)/2N)\pi\}} \quad (8)$$

DCT 기본함수는 식 (8)로부터 구해지며, peak 주파수는 index 에 따라 증가한다.

#### 4. DCT 하중치의 도출

여러가지 HVS 모델들이 위에서 제안되었는데, 정규화된 모델은 식 (9)와 같이 표현될 수 있다.

$$H(f) = a(b+cf)\exp(-cf) \quad (9)$$

정규화된 곡선에서 상부  $a, b, c$  는 peak 주파수  $f$  에 따라서 다음과 같은 관계를 가지고 있다.

$$f_{\max} = \frac{1-b}{c}$$

$$a = \exp(1-b) \quad (10)$$

낮은 주파수 부분에서는 두 가지 상충되는 현상이 있다. 첫째는 intercept 값이 올라감에 따라 낮은 주파수 부분이 강조되어 영상이 더욱 흐려지고 block 현상이 심해진다. 그러나 특정지점보다 낮은 intercept 값은 영상의 평균값을 낮추게 한다. Mannos 와 Sakrison 은 peak 의 높이를 1로 정규화시키면 intercept 값이 0.05 일때가 가장 우수하다는 것을 발견하였다[3].

$$H(0) = 0.05 \quad (11)$$

그래서 식 (9)를 식(11)에 이용하면 주파수에 따라 해석적으로  $a, b, c$  값이 구해진다. 영상을

simulation 하는데는 MSE 와 블럭 경계부분에서 error 의 변화를 나타내는 BMSE 가 사용되었다.

$$BMSE = [E\{\Delta e(m, jN)\}^2 + E\{\Delta e(kN, n)\}^2]^{1/2}$$

$$\Delta e(m, jN) = (m, iN) - e(m, jN+1)$$

$$\Delta e(kN, n) = e(kN, n) - e(kN+1, n) \quad (12)$$

여기서  $(m, n)$  은 pixel 의 주소이고,  $(j, k)$  는 블럭을 나타내며,  $N$  은 블럭크기이다. HVS 는 7 cycles/degree 부근에서 peak 값을 가질때 가장 좋은 것으로 알려졌다. 그래서 다음과 같은 모델이 얻어 지고

$$H(f) = 2.667(.0187 + .140f)e^{-0.14f} \quad (13)$$

또한 제안된 가중치 matrix 는 표 1 과 같다. 가변주파수  $f$  는 다음과 같이  $f_r$  과  $f_s$  로 표시될 수 있다.

$$j(\text{cycles/deg.}) = f_r(\text{cycles/pel})$$

$$f_s(\text{pels/deg.}) \quad (14)$$

여기서  $f_r$  는 DCT domain 의 주파수이고,  $f_s$  는 conversion factor 이다.  $256 \times 256$  영상은 pel 수가  $512 \times 512$  의 절반이므로 해당 radial 주파수에서 symmetric 하고 isotropic 한 성질을 갖도록 전개된다. 그래서 최대 radial 주파수는  $8 \times 8$  sub-block 에서 23 cycles/degree 를 갖고,  $f_r = (f_x^2 + f_y^2)^{1/2}$  으로 구해진다.

#### 5. 부호화 방법

시각 시스템의 모델링 방법에 따라 위와 같이 서로 다른 견해를 보이고 있지만 기본적으로 대역통과 특성을 가짐은 명백하다. 영산신호는 기존의 방법에서 화소, 블럭내의 화소 집합, 푸리에 또는 다른 변환에 의한 계수, 예측치 또는 미분치등으로 표시되는데 적절한 방법에 의해 대역통과 특성이 만족되도록 부호화 하여야 한다. DCT 는 각기 정해진 주파수 성분으로 입력 신호를 처리한다. 주파수 성분이라 함은 푸리에 영역에서 보다 더 엄밀한 표현이 가능하지만 real

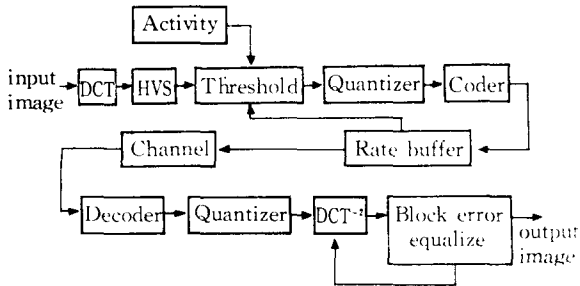


그림 5. 전체 처리 구성도

데이터를 취급하는 경우에는 real 변환 또한 적절한 주파수 표시방법이 되기 때문에 변환계수에도 적용되며, 변환후에 변환계수상에서 처리하는 계통도는 그림 5와 같다. 그림 6은 256×256 GIRL 영상을 DCT 처리한 것과 HVS 변환 처리한 것을 나타내고 있다. 계수 분포는 DC 성분에 Gaussian 분포를, AC 성분에 Laplacian 분포를 가정하여 양자화하였을 때 우수한 화질이 얻어짐을 보이고 있으며, 여기에 공간주파수 8 cycles/degree 에서 가장 민감한 Mannos와 Sakrison의 모델을 계수상에서 하중시킴으로써 그림 6과 같이 전송을 1bpp에서 아주 우수한 화질을 얻게 된다. HVS가 기존의 코딩에 주는 영



(a) original



(b) DCT 처리된 영상, 1bpp



(c) DCT와 HVS 처리된 영상, 1bpp

그림 6. 처리된 256×256 GIRL 영상

향중에 단적인 것은 peak 주파수 성분을 밀하게 하고, 그렇지 않은 곳에서 거칠게 양자화 하는 것으로 MSE는 크게 줄어들지 않지만 블럭에러를 감소시켜 보다 부드러운 영상을 재현하고 있다.

## 6. 에러 측정

기존의 MSE는 시각 시스템의 제반 특성으로 인하여 더 이상 절대적인 의미를 갖지 못하는 것이 분명하다. 영상의 품질은 잡음에 의해서 뿐만 아니라 변환부호화후 발생하는 blocking 현상에 의해서 더욱 손상되므로 기존의 MSE와 다른 평가방법이 요구된다. 몇가지 에러 측정의 기준을 살펴 보면 다음과 같다.

### 1. WMSE(weighted mean square error)

$$e = \frac{\sum \sum e'(m, n)^2}{\sum \sum x(m, n)^2}$$

$e'$  = error weighted by  $W(f)$  (15)

에리에 하중되는 함수  $W(f)$ 는 contrast 감도와 공간주파수 반응을 고려하여 결정한다.

### 2. 임계치에 의한 WMSE

$$e = \frac{\sum \sum \{e'(m, n)^2 \times T\}}{\sum \sum x(m, n)^2}$$

$T=1, |e'| > \text{임계치}$  (16)  
 $0, |e'| < \text{임계치}$

윗 식은 시각적으로 감도가 낮은 에리에 대해서는 계산에서 제외하고, 임계치 이상의 것만 측정하는 WMSE의 진보된 형태이다.

### 3. 에러의 자기 상관

$$e = \left[ \sum \sum R \{e(m, dn)\}^2 + \sum \sum R \{e(dm, n)\}^2 \right]^{1/2}$$

$$R \{e(dm, n)\} = E \{e(m+dm, n) e(m, n)\} / e^2_{\max} \quad (17)$$

$$R \{e(m, dn)\} = E \{e(m, n+dn) e(m, n)\} / e^2_{\max}$$

에러의 상관도가 높으면 보다 부드러운 영상이 되며 불규칙적 에러가 발생될 때 시각의 거부 반응이 증가 된다.

### 4. 블럭 경계에서 에러의 변화

$$e = \left[ E \{\Delta e(m, kN)\}^2 + E \{\Delta e(LN, n)\}^2 \right]^{1/2}$$

$$\Delta e(m, kN) = \{e(m, kN) - e(m, kN+1)\}^2$$

$$\Delta e(LN, n) = \{e(LN, n) - e(LN+1, n)\}^2 \quad (18)$$

변화 부호화의 문제점은 블럭 내에서의 분산 효과와 경계 부분에서의 분산효과 및 경계부분에서의 에러 증가이다.

### 5. PMSE(perceptual mean square error)

$$e = \frac{\sum \sum \{Z(m, n) - \hat{Z}(m, n)\}^2}{\sum \sum \{Z(m, n)\}^2} \quad (19)$$

$Z(m, n)$  : 변환 계수  $\hat{Z}(m, n)$  : 부호화된 계수

윗 식은 변환 영역에서 계수 값과 부호화된 계수값과의 에러를 비교하는 것으로 주파수 영역에서 에러 계산을 한다.

### 7. 결 론

HVS와 DCT를 이용하여 DCT 함수의 푸리에 스펙트럼을 계산하고, MSE와 BMSE에 근거하여 최적 MTF를 얻었다. HVS가중치를 이용하여 야자화 범위와 bit rate를 줄였을 때 variable length coding이 간단해졌다. 또한 HVS가중치는 유효한 항만을 시각에 이용하기 때문에 하나의 filter로 여길 수 있으므로 교류성분에 대해서 inverse가중치를 주지 않고서도 좋은 화질을 얻을 수 있었다. 예를 들면 'GIRL' 영상

에서 BMSE는 0.5 bit rate일 때 -31.3[dB]이다. 결국 HVS가중치 방법을 DCT에 적용한 결과 감각적으로 보다 좋은 영상을 얻을 수 있었고 bit rate를 1bpp로 낮출 수 있었다.

표 1. 제안된 가중치 matrix

.050	.753	.921	.996	.985	.910	.804	.640
.753	.889	.968	1.000	.973	.893	.788	.628
.921	.968	.995	.994	.953	.869	.765	.610
.996	1.000	.994	.966	.913	.827	.727	.580
.985	.973	.953	.913	.855	.771	.677	.542
.910	.893	.869	.827	.771	.694	.611	.491
.804	.788	.765	.727	.677	.611	.539	.436
.640	.628	.610	.580	.542	.491	.436	.357

### Reference

- [1] M. Kunt, A. Ikonomopoulos and M. Kocher, "Second-generation image coding techniques," Pros. IEEE, vol. 73, no. 4, Apr. 1985.
- [2] 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.
- [3] J.L. Mannos and D.J. Sakrison, "The effects of a fidelity criterion on the encoding of images," IEEE Trans. on Inf. Theory, vol. IUT-20, pp. 525~536, July 1974.
- [4] E.L. Hall, Computer Image Pricessing and Recognition, Academic Press, 1979.
- [5] F. X. J. Lukas and Z. L. Badrikis, "Picture quality prediction based on a visual model," IEEE Trans. on Comm., vol. COM-30, no. 7, Jurt 1982.
- [6] N.B. Nill, "A visual model weighted cosine transform for image compression and quality assessment," IEEE Trans. on Comm., vol. COM-33, pp. 551~557, June 1985.
- [7] D.K. Sharma and A.N. Netravali, "Design of quantizers for DPCM coding of picture signals," IEEE Trans. on comm., vol. COM-25, no. 11, Nov. 1977.
- [8] R.M. Gray, "Vector quantization," IEEE Magazine, Apr. 1984.
- [9] R.C. Reininger and J.D. Gibson, "Distribution

- of the two dimensional DCT coefficients for images," IEEE Trans. on Comm, vol.COM-31, no. 6, June 1983.
- [10] S. Yuan and K.B. Yu, "Zonal sampling and bit allocation of HT coefficients in image data compression," IEEE Trans. on Comm., vol. COM-34, no.12, DEc. 1986.
- [11] J.W. Modestino, N. Farvardin and M.A. Vgrinc, "Performance of block cosine image coding with adaptive quantization," IEEE Trans. on Comm., vol.COM-33, no. 3, Mar. 1985.
- [12] G.G. Furman, "Comparison of models for subtractive and shunting lateral-inhibition in receptor-neuron fields," Kybernetik, vol. 2, pp. 257~274, 1965.
- [13] F. Kretz, "Subjectively optimal quantization of pictures," IEEE Trans. on Comm., vol. COM-23, pp.1288~1292, Nov. 1975
- [14] K.V. Ngan, K.S. Leong and H. Shingh, "Cosine transform coding incorporated human visual system model," SPIE Fiber LASE 86, Cambridge MA, vol.707 pp.165~177, Sep. 1986.
- [15] M.H. Lee, "The Center Weighted Hadamard Transform," to be appear, IEEE Trans. on Circuit and Systems.
- [16] J.J. Hwang and M.H. Lee, "A Visual Model Weighted Hadamard Transform," PCS '88, 12~14th, Sep. 1988, Torino, Italy.
- [17] S.C. Ahn and M.H. Lee, "Variable Threshold Detection with Weighted BPSK/PCM Speech Signals Transmitted over Gaussian Channel," to be appear IEEE Trans. on Communications.