

변환영역에서의 지능형 분류벡터양자화를 이용한 영상압축

Image Compression using an Intelligent Classified Vector Quantization Method in Transform Domain

이현수 · 공성곤

Hyun-Soo Lee and Seong-Gon Kong

숭실대학교 전기공학과

요 약

이 논문에서는 영상데이터를 여러개의 영상블록들로 나누고 이산 코사인변환 영역에서 물체의 에지에 해당하는 영상블록을 에지방향을 고려하여 적절히 분류함으로써 영상데이터를 효과적으로 압축하였다. 벡터양자화에 의한 영상데이터의 압축은 높은 압축률을 실현할 수 있지만 영상내 물체의 에지부근이 손상되어 시각적인 화질이 저하되는 단점이 있다. 높은 압축률을 유지하면서도 시각적인 화질의 열화를 피하기 위하여 영상블록의 이산 코사인변환계수의 에너지 분포에 따라 에지블록을 8개의 부류로 분류하였다. 또한 이 분류과정을 통하여 얻어진 데이터를 가지고 신경회로망을 학습하여 구현한 에지블록의 분류과정과 성능을 비교하였다. 에너지분포에 의한 에지분류방법과 신경망으로 학습한 분류과정은 에지특징벡터에 의한 분류벡터양자화에 비해 더 높은 PSNR과 시각적으로 좋은 화질을 보여주었다.

ABSTRACT

This paper presents image data compression using a classified vector quantization (CVQ) which categories edge blocks according to the energy distribution of subimages in the discrete cosine transform domain. Classifying the edge blocks enhances visual quality of the compressed images while maintaining a high compression ratio. The proposed classification method categories subimages into eight types of edge features according to an energy distribution. A neural network, trained with the data generated from the proposed classification method, can successfully classify subimages to eight edge categories. Experimental results are given to show how the CVQ method incorporated with a neural network can produce faithful compressed image quality for high compression ratios.

1. 서 론

벡터양자화를 이용한 영상압축은 주어진 영상데이터를 일정한 크기의 벡터로 분할하고, 입력벡터를 미리 결정된 한정된 개수의 코드벡터로 근사화하여 표현함으로써 높은 압축률을 실현하기 위한 방법이다 [1, 2]. 그러나 벡터양자화에 의한 영상압축은 영상신호에 포함되어 있는 물체의 에지부근이 손상되어 시각적인 화질이 저하되는 단점이 있으므로 입력벡터를 에지의 종류에 따라 분류하여 시각적인 화질의 저하를 줄이기 위한 분류벡터양자화(CVQ)가 제안되었다. 분류벡터양자화는 에지의 특징에 따라 입력벡터를 몇 개의 부류로 분류하여 부류별로 별도의 코드북을 생성하여 압축함으로써 압축된 영상의 시각적인

화질을 향상시킬 수 있다. 또한 분류벡터양자화는 각 부류마다 별도로 구성된 작은 규모의 코드벡터들 중에서 대표벡터를 탐색하므로 탐색시간을 줄일 수 있다는 장점이 있다.

이산 코사인변환(DCT) 영역에서 분류벡터양자화에 의해 영상데이터를 압축하는 것은 공간영역에서 보다 데이터의 상관성이 크게 줄어들어 영상을 효율적으로 압축할 수 있는 잇점이 있고 간단한 분류방식으로 입력벡터를 분류할 수 있으나, 실제영상에 존재하고 있는 복잡한 부분을 정의하기 어려우므로 에지부근이나 복잡한 부근이 손상되는 문제가 발생한다. 변환영역에서의 에지블록을 분류하는 방법중에 에지특징벡터를 이용하는 방법[3]은 영상블록을 2차원 이산코사인변환을 하고 AC성분에 해당하는 계수들중

*이 연구는 1995년도 한국과학재단 연구비지원에 의한 결과임(과제번호 951-0906-025-2).

에서 값이 가장 큰 두 계수에 의해 에지블록을 분류하는 방법으로, 간단하면서도 영상블록을 에지의 특징에 따라 효과적으로 분류하고 있으나 복잡한 부분을 잘 정의하지 못하는 단점이 있다.

본 논문에서는 영상블록의 이산 코사인변환계수의 에너지분포에 따라 지능적으로 에지블록을 8개의 부류로 분류함으로써 분류벡터양자화의 성능을 향상시켰고 높은 압축률을 유지하면서도 압축된 영상의 시각적 화질의 저하를 줄였다. 또한 변환계수의 에너지 분포에 의한 에지블록의 분류과정에서 얻어진 입출력 데이터를 오차역전과 알고리즘에 의해 신경회로망을 학습하여 구현하였다. 신경회로망에 의한 영상블록의 분류기는 변환계수의 에너지분포에 있어서의 불확실성에 대한 대처능력을 향상시킬 수 있다.

2. 벡터양자화

벡터양자화는 학습데이터로부터 코드벡터를 생성하는 과정과 주어진 입력벡터를 가장 잘 표현하는 코드벡터를 탐색하는 패턴매칭 과정으로 이루어진다. 코드벡터 생성과정은 입력데이터를 벡터로 나타내고 유사한 성질을 가지고 있는 벡터들을 하나의 그룹으로 클러스터링하고 그 클러스터를 가장 잘 표현하는 코드벡터를 찾아내는 과정이다. 입력벡터들을 N개의 그룹 D_1, D_2, \dots, D_N 으로 클러스터링하였을 때, 각 클러스터를 대표하는 N개의 코드벡터 $\{y_1, y_2, \dots, y_N\}$ 가 존재한다. 데이터의 확률분포가 $p(x)$ 일 때 코드벡터 중에서 벡터양자화의 평균제곱오차를 최소로 하는 코드벡터는 식 (1)과 같은 클러스터의 무게중심(centroid)이다[2].

$$y_j = \frac{\int_{D_j} xp(x) dx}{\int_{D_j} p(x) dx} \quad (1)$$

패턴매칭과정은 클러스터링을 통하여 코드벡터 $\{y_1, y_2, \dots, y_N\}$ 가 생성되었을 때, 주어진 입력벡터 x 를 가장 잘 표현하는 코드벡터 y_m 을 대응시키는 과정이다. 잘 정의된 벡터공간에서 입력벡터들 간의 유사성은 거리척도에 의해 표현되므로, 입력벡터를 가장 잘 표현하는 코드벡터는 식 (2)와 같이 입력벡터와의 거리척도가 가장 작은 코드벡터로 정의될 수 있다.

$$d(x, y_m) = \min_{i=1, \dots, N} d(x, y_i) \quad (2)$$

여기서 $d(x, y_i)$ 는 입력벡터 x 와 코드벡터 y_i 사이의 거리함수값을 나타내며 일반적으로 Euclidean norm을 사용한다.

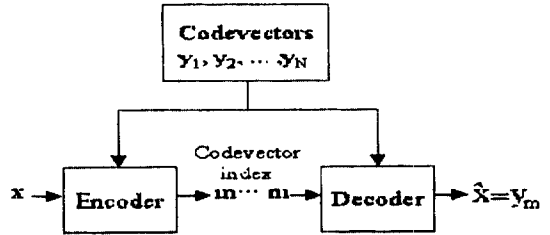


그림 1. 벡터양자화에 의한 데이터압축 과정.
Fig. 1. Data compression procedure by the vector quantization.

그림 1은 벡터양자화에 의한 데이터압축 과정을 나타낸다. 입력벡터가 주어지면 부호기(encoder)는 미리 결정해 둔 코드벡터들 중에서 입력벡터 x 를 가장 잘 표현하는 코드벡터 y_m 을 선택하여 그 지표 m 을 전송한다. 수신측에서 복호기(decoder)는 수신된 지표 m 에 해당하는 코드벡터 y_m 을 입력벡터의 양자화된 벡터 \hat{x} 로 대응시켜 원래 데이터를 복원한다.

이 논문에서 코드벡터들은 LBG 알고리즘[4]을 이용하여 생성하였다. 이 알고리즘은 학습벡터 x_i 와 코드벡터 y_i 와의 평균왜곡값(average distortion) D 가 주어진 임계치에 도달할 때까지 반복적으로 학습벡터들을 분할하고 분할된 학습벡터의 표본평균으로 코드벡터를 갱신하는 것이다.

$$D = \frac{1}{T} \sum_{i=0}^{T-1} \min_{i=1, \dots, N} \|x_i - y_i\|^2 \quad (3)$$

여기서 T는 학습벡터의 수이고, N은 코드벡터의 수를 나타낸다.

LBG 알고리즘에 의해 코드벡터를 생성하는 과정을 요약하면 다음과 같다.

0) 수렴판정을 위한 임계치 ϵ 와 초기 평균왜곡치 $D_{i=\infty}$ 를 정하고 코드벡터의 집합을 초기화한다: $A_0 = \{y_i, i=1, \dots, N\}$

1) 학습벡터의 분할과 평균왜곡치를 계산한다.

m 번째 코드벡터의 집합 $A_m = \{y_1, y_2, \dots, y_N\}$ 에 대하여 $\|x_i - y_i\|^2 \leq \|x_i - y_j\|^2$ 이면, 1번째 학습벡터 x_i 를 S_i 로 분할하고, 이때의 평균왜곡 D_m 을 계산한다.

$$D_m = D(\{A_m, P(A_m)\}) = \frac{1}{T} \sum_{i=1}^T \min_{y_i \in A_m} \|x_i - y_i\|^2 \quad (4)$$

여기서, S_i 는 학습벡터를 분할하는 i 번째 분할집합이고, $P(A_m) = \{S_i, i=1, \dots, N\}$ 이다.

2) 평균왜곡치의 수렴을 판정한다.

$|D_{m-1}-D_m|/D_m \leq \epsilon$ 이면 중단하고, 아니면 과정 3)을 실행한다.

3) 학습벡터의 표본평균으로 새로운 코드벡터를 구한다.

$$y_i = \frac{1}{\|S_i\|} \sum_{j: x_j \in S_i} x_j \quad (5)$$

$\|S_i\|$ 는 S_i 에 있는 학습벡터의 개수를 나타낸다. 이때, $\|S_i\|=0$ 이면, 이전의 i 번째 코드벡터로 대신한다. 새로운 코드벡터의 집합 $A_{m+1}=\{y_i, i=1, \dots, N\}$ 을 형성하고, $m \leftarrow m+1$ 로 하고, 과정 1)로 간다.

3. 이산 코사인변환영역에서의 분류벡터양자화

변환영역에서의 영상신호 압축은 영상신호의 변환계수들간에 상관성이 적고 낮은 공간주파수에 해당하는 변환계수에 에너지가 집중되며 인간의 시각계통이 높은 공간주파수 성분의 영상신호에 대해서는 별로 민감하지 않다는 성질을 이용하고 있다. 공간영역에서의 영상신호는 에너지 분포가 골고루 분산되어 있지만, 직교변환에 의해 표현되는 영상신호는 중복성이 작고 적은 수의 변환계수들에 에너지가 집중되어 있다. 따라서 에너지가 큰 변환계수들만 선택하여 보존하고 낮은 에너지를 갖는 높은 공간주파수의 변환계수들은 무시해 버림으로써 영상신호를 압축할 수 있다. 여러 종류의 직교변환중에서 이산 코사인 변환은 에너지의 집중능력이 우수하고 고속 알고리즘이 존재하므로 영상압축에 널리 사용되고 있다. $M \times M$ 크기의 영상신호 $x(m, n)$ 의 2차원 이산코사인 변환계수 $X(u, v)$ 는 다음과 같이 계산된다.

$$X(u, v) = \frac{2}{M} c(u)c(v) \sum_{m=0}^{M-1} \sum_{n=0}^{M-1} x(m, n) \cdot$$

$$\begin{aligned} & \cos \frac{\pi u(2m+1)}{2M} \cos \frac{\pi v(2n+1)}{2M} \\ x(m, n) = & \frac{2}{M} \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} c(u)c(v) X(u, v) \cdot \\ & \cos \frac{\pi u(2m+1)}{2M} \cos \frac{\pi v(2n+1)}{2M} \quad (6) \end{aligned}$$

여기서 계수는 $c(0)=1/\sqrt{2}$, $c(k)=1, k=1, \dots, M-1$ 이다. 벡터양자화는 주어진 영상블록을 몇 개의 벡터로 표현하고 각 벡터들을 가장 잘 대표하는 코드벡터를 찾아 전송함으로써 높은 압축률로 영상데이터를 압축하는 방법이다. 벡터양자화에서는 코드벡터를 생성하는 과정과 입력벡터를 가장 잘 표현하는 코드벡터를 탐색하는 과정으로 이루어진다. 벡터양자화에서는 식 (2)에 의해 입력벡터와 가장 유사한 코드벡터를 탐색하므로 에지의 특성을 고려하지 않아 에지부근이 손상되어 시각적인 면에서 화질이 저하되는 문제가 있다. 또한 벡터양자화에서 필요한 계산량은 코드벡터의 수에 비례하여 증가하므로 벡터의 차수 또는 전송률의 증가에 따라 복잡성은 기하급수적으로 증가하게 된다. 이러한 계산의 복잡성 및 에지부근의 손상을 해결하기 위하여 분류벡터양자화[5, 6]가 제안되었다.

분류벡터양자화는 영상블록을 에지의 종류에 따라 분류하여 각 에지부류마다 독립적으로 벡터양자화하는 것으로 에지종류를 고려하여 분류하기 때문에 에지방향을 보전할 수 있어 압축된 영상의 시각적 화질을 개선할 수 있다. 또한 각 에지부류마다 별도로 구성된 적은 수의 코드벡터들 중에서 대표벡터를 탐색하므로 부호화시 계산의 복잡성을 줄일 수 있다. 그러나 영상블록을 에지의 특성에 따라 효과적으로 분류할 수 있는 알고리즘이 필요하다.

그림 2는 이산 코사인변환영역에서의 분류벡터양자화에서 영상신호를 압축하는 과정을 나타낸 것이

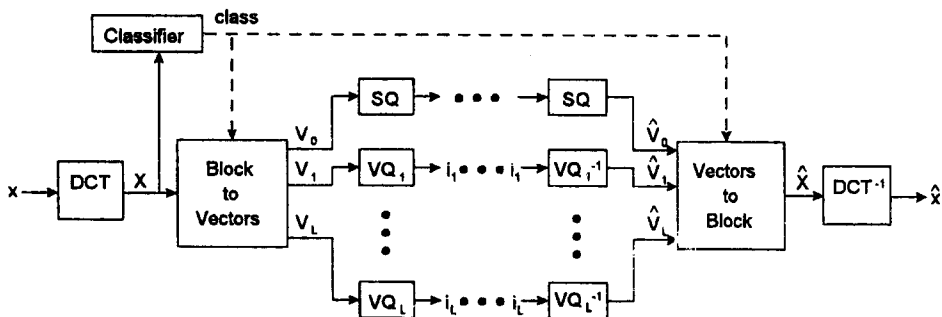


그림 2. 변환영역에서의 분류벡터양자화 과정.
Fig. 2. Classified vector quantization in the discrete transform domain.

다. 이산 코사인변환영역에서의 분류벡터양자화는 영상데이터를 중복성이 작은 변환계수로 표현하는 변환부호화[7, 8]와 에지의 특성에 따라 영상블록을 분류하여 코드벡터를 별도로 생성하여 영상신호를 압축하는 분류벡터양자화의 장점을 결합한 것이다. 이산 코사인변환영역에서의 분류벡터양자화는 영상블록을 2차원 이산코사인변환에 의해 변환계수로 표현하고, 변환계수의 에너지의 분포에 따라 에지방향을 고려하여 몇 개의 에지부류로 분류하여 압축한다.

벡터양자화에서 코드벡터의 수 N , 비트율 $R(\text{bits/pixel})$, 그리고 벡터의 차수 k 사이의 관계는 다음 식과 같다.

$$N = 2^{kR} \quad (7)$$

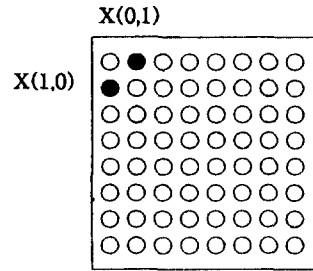
4. 에지특성에 따른 영상블록의 분류

4.1 에지특징벡터에 의한 에지블록의 분류

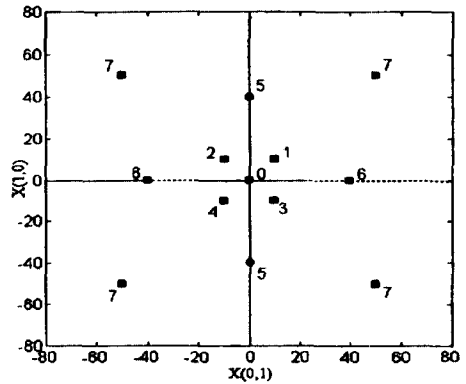
영상신호를 4×4 크기의 블록으로 나누고 이산 코사인 변환을 하였을 때 얻어지는 변환계수들 중에서 첫 번째 열과 첫 번째 행을 이루고 있는 변환계수들을 에지특징벡터로 이용하면 영상블록을 에지특성에 따라 효과적으로 분류할 수 있다.[3] 블록을 주어진 영상을 8×8 크기의 블록으로 나누고 2차원 코사인변환을 하였을 때 얻어지는 변환블록 $X(u, v)$, $u, v=0, 1, \dots, 7$ 에서 두 개의 DCT 변환계수 $X(0, 1)$ 와 $X(1, 0)$ 은 다른 변환계수들에 비하여 상대적으로 에지블록의 특징을 결정하는데 중요한 역할을 한다. 변환블록의 DCT계수의 에너지는 낮은 공간주파수영역에 거의 집중되어 있고, 직육성분을 제외한 다른 모든 계수들보다 상대적으로 큰 값을 가지고 있기 때문에 DCT계수 $X(0, 1)$ 와 $X(1, 0)$ 을 에지방향을 나타내는 특징벡터로 선택할 수 있다. 이 두 변환계수를 에지특징벡터(edge feature vector)로 정의하고 이 변환계수의 값에 따라 성공적으로 에지블록을 분류하였다.

그림 3(a)는 특징벡터를 구성하는 영상블록의 두 변환계수들을 나타낸다. 이 변환계수들은 대부분의 영상신호에 대해서 모든 AC 계수들 중에서 가장 큰 값에 해당한다. 그림 3(b)는 위에서 정의된 8가지의 부류별로 특징벡터의 분포를 나타낸 것이다. 특징벡터에 의해 에지블록을 분류하는 분류벡터양자화에서는 특징벡터를 각 부류를 대표하는 코드벡터로 하고, 입력 영상블록의 특징벡터와의 거리가 최소인 부류로 분류한 다음 벡터양자화에 의해 영상데이터를 압축한다.

여기서 부류 0은 평탄, 부류 1, 2, 3, 4는 약한 대각, 부류 5는 강한 수평, 부류 6은 강한 수직, 부류 7은 강



(a)



(b)

그림 3. (a) 에지특징벡터 및 (b) 부류별 대표 특징벡터의 분포.

Fig. 3. (a) Edge feature vector (b) Distribution of representative feature vectors at each class.

한 대각으로 정의된다. 이 방법은 에지특징벡터만을 이용하여 영상블록들을 분류하기 때문에 간단하게 분류할 수 있다는 장점이 있으나 두개의 계수만으로 에지방향을 표현하기에는 부족하므로 에지블록이 잘못 분류될 수 있으며 또한 실제 영상에 존재하는 복잡한 영역이 정의되어 있지 않은 단점이 있다.

표 1. 에지부류의 정의

Table 1. Definition of the edge classes

부류	에지의 종류
0	평탄
1	약한 수평
2	약한 수직
3	약한 대각
4	강한 수평
5	강한 수직
6	강한 대각
7	복잡

4.2 변환계수의 에너지 분포에 의한 에지블록의 분류

영상 데이터내에 포함되어 있는 물체의 에지성분은 시각적으로 인식하는 화질에 민감한 영향을 주므로 영상압축과정에서 시각적으로 화질을 향상시키기 위해서는 에지성분을 보전하는 것이 필수적이다. 영상블록을 에지 및 실제 영상에 존재하는 복잡한 블록을 효율적으로 분류하고 에지부류의 분류정보를 이진수로 표현하여 전송할 때 전송에 필요한 비트수를 효율적으로 이용하기 위하여 8가지의 부류로 분류하였다. 이 논문에서는 영상블록을 에지특성에 따라 다음과 같은 8가지의 부류를 고려하였다.

대부분의 영상블록의 이산 코사인변환계수의 에너지는 낮은 공간주파수 영역에 집중되어 있고, 에지의 특성에 따라 에너지의 분포가 다르다. 이 논문에서는 영상블록의 전체 AC 에너지와 변환계수의 에너지분포에 따라 영상블록내에 포함되어 있는 에지의 종류를 분류하였다. 영상블록이 가지고 있는 전체 AC 에너지 E_T 는 DC성분 $X(0, 0)$ 를 제외한 전체 에너지의 합으로 정의된다.

$$E_T = \sum_{u=0}^{M-1} \sum_{v=0}^{M-1} |X(u, v)| - |X(0, 0)| \quad (8)$$

그림 4는 에지의 특성에 따른 영상블록내 변환계수의 분포를 나타낸 것이다. 검정색으로 나타낸 변환계수의 절대값의 합을 각각 H, V, 그리고 D로 정의하고, 이 값 중에서 가장 큰 값을 E_M 으로 정의하였다. 특히 대각방향의 에지패턴 D의 경우 저주파수 영역에서의 값이 크고 고주파수 영역으로 갈수록 작아지므로 비대칭적으로 설정하였다.

그림 5는 변환계수의 에너지분포에 의해 영상블록을 8가지 에지부류로 분류하는 과정을 나타낸 순서도이다. 밝기의 변화가 거의 없는 평탄한 영상블록의 에너지는 주로 낮은 공간주파수 영역에 집중되어 있으므로, E_M 값이 임계치 θ_1 보다 작을 경우에는 평탄부류로 분류하였다. 밝기의 변화가 심한 복잡한 성분을 많이 포함하고 있는 영상블록의 경우에는 변환계수의

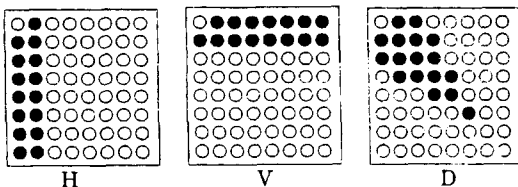


그림 4. 변환계수의 에너지분포에 따른 에지블록의 구분.
Fig. 4. Classification of edge blocks according to energy distribution in the transform domain.

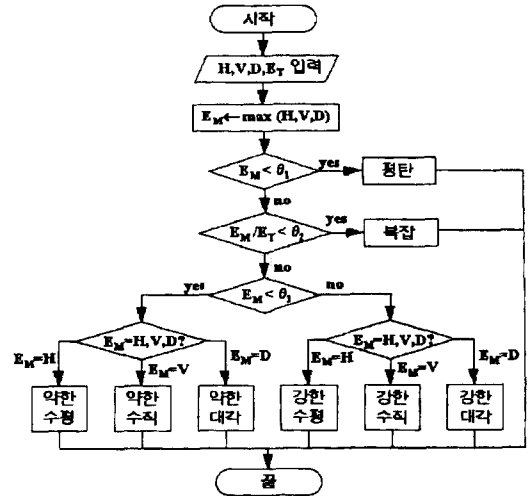


그림 5. 변환계수의 에너지분포에 의한 에지부류의 분류.
Fig. 5. Classification of edge classes according to the energy distribution of transform coefficients.

에너지가 거의 모든 주파수대역에 걸쳐 분포하므로 전체에너지 E_T 와 E_M 값의 비율에 의해 복잡부류와 에지부류를 구분하였다. E_M 값이 임계치 θ_1 보다 크고 E_M/E_T 값이 임계치 θ_2 보다 작을 경우에는 복잡부류로 분류하였다. 평탄부류나 복잡부류에 해당하지 않는 영상블록에 대해서는 변환계수의 에너지분포를 나타내는 H, V, 그리고 D값 중 가장 큰 값 E_M 을 이용하여 에지부류를 결정하였다. 이 경우 E_M 값이 임계치 θ_3 보다 작은 경우에는 약한 에지성분으로, 그리고 임계치보다 큰 경우에는 강한 에지성분으로 세분하여 분류하였다. 세부 그룹내에서도 에너지분포에 따라 각각 수평, 수직, 그리고 대각의 세가지 부류로 분류하였다. 임계치는 실험에 의하여 $\theta_1=7.0$, $\theta_2=0.65$, 그리고 $\theta_3=120$ 으로 정하였다.

4.3 신경회로망에 의한 에지블록의 분류

신경회로망은 기본 처리단위인 뉴론과 뉴론들을 연결하고 있는 연결강도(synaptic weight)에 의하여 이루어지는 네트워크이다. 신경회로망은 기능별로 입력을 받아들이는 입력층, 내부적인 표현을 담당하는 중간층, 그리고 출력을 발생하는 출력층으로 구분되어 있다[9, 10]. 중간층도 여러개의 층으로 구성될 수 있으나 하나의 중간층을 가진 신경회로망도 임의의 함수를 원하는 정도까지 근사화할 수 있다는 것이 알려져 있다. 그림 6은 에지블록의 분류를 위한 신경회로망의 구성을 나타낸다.

부영상 분류를 위한 신경회로망의 구성은 중간층

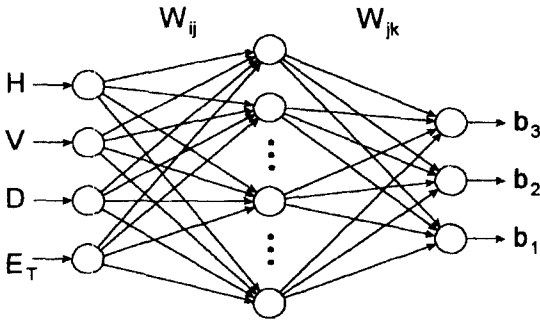


그림 6. 에지블록의 분류를 위한 신경회로망의 구성.
Fig. 6. Structure of the neural network for classification of edge blocks.

을 포함하는 다층구조로 하고, 오차역전과 학습 알고리즘을 이용하여 학습 하였다. 신경회로망의 입력은 수평성분 H, 수직성분 V, 대각성분 D, 그리고 전체 에너지 E_T로 하였고, 출력층은 8개의 부류를 이진수 000부터 111까지로 나타내기 위하여 b₁, b₂, b₃의 3개로 하였다. 중간층의 뉴론수는 입력데이터의 수와 학습에 소요되는 계산량을 고려하여 16으로 하였다. 입력 데이터는 부류별로 균일하게 학습되도록 하기 위하여 에너지 분포에 의한 분류로부터 얻어진 데이터를 부류별로 20개씩 무작위로 추출하여 총 160개의 입출력 데이터를 사용하여 오차 역전과 알고리즘으로 신경회로망을 학습하였다.

이 신경회로망에서 중간층과 출력층의 활성화함수는 0과 1사이의 범위내 있는 비선형함수인 로그 시그모이드함수를 사용하였다. 모든 뉴론은 0과 1사이의 실수값을 나타내도록 정의되었으므로 신경회로망의 학습시 입력데이터로 사용하기 전에 모두 0과 1사이로 정규화하여야 한다. 그러나 출력함수가 아주 큰 입력값에 대해서도 0이나 1에 가까워지기는 하지만 만나지 않으므로 출력뉴론이 0이나 1이 나오도록 학습시키는 데는 많은 시간이 소요된다. 따라서 이러한 문제를 해결하기 위하여 모든 입력과 출력은 식 (9)에 의하여 학습패턴을 a에서 b로 정규화하였다. 이 논문에서는 a=0.1, b=0.9의 값으로 정규화하였다.

$$u' = (b - a) \left[\frac{u - u_{\min}}{u_{\max} - u_{\min}} \right] \quad (9)$$

여기서 u는 원래의 학습 데이터이고, u'는 a에서 b사이의 값으로 정규화된 학습 데이터를 나타낸다. u는 입력데이터 H, V, D, E_T중 하나이고, u_{max}와 u_{min}은 각각 학습 데이터에서의 최대값과 최소값을 나타낸

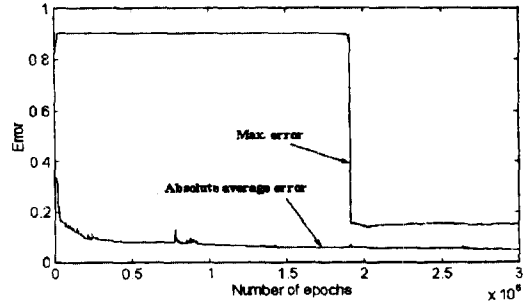


그림 7. 반복실행 횟수에 따른 최대오차와 절대치 평균오차의 변화.
Fig. 7. Max error and average absolute error according to the number of epochs.

다. 입력과 원하는 출력이 0.1에서 0.9로 정규화되었기 때문에 실제 출력과 원하는 출력사이에 나올 수 있는 최대 오차값은 0.9이다.

다음은 전체 30만번의 반복실행에서 매 실행마다 최대오차 E_{max}와 평균 절대값오차 E_{avg}값의 변화를 나타내고 있다.

$$E_{\max} = \max_{p=1, \dots, P} [\max_{k=1, \dots, K} |d_{pk} - o_{pk}|] \quad (10)$$

$$E_{\text{avg}} = \frac{1}{PK} \sum_{p=1}^P \sum_{k=1}^K |d_{pk} - o_{pk}| \quad (11)$$

여기서 P는 전체 학습패턴의 개수, K는 입력패턴의 차수이며, d_{pk}, o_{pk}는 각각 p번째 패턴에 대한 출력층의 k번째 뉴론의 원하는 출력값과 실제 출력값을 나타낸다.

그림 7에서 모든 입력데이터에 대한 최대치 오차값이 18만번의 반복실행에서 0.2이하로 줄어드는 것으로 미루어 부영상의 부류를 나타내는 각 비트의 오차값이 0.2이하인 것을 알 수 있으며 신경회로망의 출력값을 사사오입하여 모든 학습에 쓰인 데이터를 분류할 경우 잘 분류할 수 있음을 나타내고 있다.

5. 에지블록 분류방법의 성능비교

5.1 변환블록의 부벡터의 분할 및 비트할당

에지의 특성에 따라 분류된 영상블록의 DCT계수들은 분산이 크게 변화하기 때문에 모든 계수에 동일한 양자화방법을 적용하는 것은 비효율적이다. 압축된 영상의 질을 크게 저하시키지 않으면서 높은 압축률을 실현하기 위해서는 계수들의 분산특성을 고려하여 차등적으로 양자화하는 것이 유리하다. 영상블록의 변환계수들을 에너지의 상관관계가 거의 동일

한 즉, 분산값이 유사한 계수들로 구성된 적은 차수의 부벡터로 분할한다. 그리고 분산값에 비례하여 코드북의 크기를 할당하고 차등적으로 각각의 부벡터를 벡터양자화한다. 분산이 큰 부벡터에는 많은 비트를 할당하고, 분산값이 작은 부벡터에는 적은 수의 비트를 할당하여 분할된 벡터들을 독립적으로 벡터양자화한다. 이때 DC 성분에 해당하는 DCT계수는 그 부영상의 평균값을 나타내고 시각적인 화질에 미치는 영향이 크므로 항상 최대값인 8비트로 스칼라 양자화한다.

부벡터로의 분할과정에서 각 부류별로 영상블록의 변환계수들의 분산은 다음과 같이 계산한다.

$$\sigma^2(u, v) = \frac{1}{T} \sum_{i=1}^T (X_i(u, v) - m(u, v))^2$$

$$m(u, v) = \frac{1}{T} \sum_{i=1}^T X_i(u, v) \quad (12)$$

여기서 T는 각 부류별로 변환계수들의 분산을 계산하는 데 이용된 학습벡터의 수를 나타낸다. 변환계수들의 분산값의 차이가 급격하므로 다음과 같이 로 그를 취하고, 크기 순서대로 정렬한다음 비슷한 값들끼리 묶어 부벡터를 구성하여 변환블록을 분할한다.

$$\sigma^2(u, v)^* = \log_2(\sigma^2(u, v)) \quad (13)$$

이렇게 분할된 변환블록의 부벡터들은 벡터양자화하기 위하여 코드북의 크기를 나타내는 비트를 할당하여야 한다. 이 때 큰 분산값을 갖는 부벡터들은 재구성할 영상에 보다 중요하게 기여하고 작은 분산값을 갖는 부벡터들은 재구성할 영상에 덜 기여하므로 코드북의 크기는 부벡터의 분산에 의존한다. 그러므로 코드북의 크기를 나타내는 비트수는 각 부류마다 부벡터의 로그분산값의 평균을 식 (14)와 같이 구하고 이것에 적절한 값을 더하거나 빼주어 원하는 비트율로 조정한다.

$$\hat{\sigma}_i^2 = \frac{1}{\|F_i\|} \sum_{u, v \in F_i} \sigma^2(u, v)^* \quad (14)$$

여기서 $\sigma^2(u, v)^*$ 은 u, v번째 분산에 로그값을 나타내며 i는 i번째 부벡터를 나타내는 변환계수의 위치집합, $\|F_i\|$ 는 F_i의 집합의 크기 즉, i번째 부벡터의 차수, $\hat{\sigma}_i^2$ 는 i번째 부벡터의 로그분산의 평균값을 의미한다.

직류성분을 제외한 나머지 벡터에 대한 비트 할당 과정은 다음과 같다.

1) 원하는 평균비트율 M_b 와 원하는 비트율과 실제 평균비트율의 오차 임계치 ϵ 의 초기값을 설정하고,

바이어스 Δ 를 입력한다.

2) 부류 c의 비트맵중에서 i번째 부벡터의 비트수 b_{ci} 를 식 (15)와 같이 그 부벡터의 분산 $\hat{\sigma}_i^2$ 에 비례하도록 계산하고, 만일 이 값이 0보다 작으면 0으로 한다.

$$b_{ci} = \hat{\sigma}_i^2 - \Delta \quad (15)$$

3) 평균비트율 m_b 를 다음과 같이 계산한다.

$$m_b = \sum_{c=1}^C p_c \sum_{i=1}^{I_c} b_{ci} \quad (16)$$

C: 부류의 수

p_c : 주어진 영상블록이 부류 c에 속하는 확률

I_c : 부류 c에서 b_{ci} 가 0이 아닌 부벡터의 개수

4) $|M_b - m_b| \epsilon$ 일 경우 비트할당 과정을 종료한다. 그렇지 않으면 바이어스 Δ 를 다시 설정하고, 상기 과정을 반복수행한다. 여기서는 $\epsilon=1.0$ 으로 하였다.

에지특성의 분류결과를 포함한 부영상의 평균비트율 $R[bpp]$ 은 식 (17)과 같이 계산된다.

$$R = (B_c + m_b) / k [bpp] \quad (17)$$

여기서 B_c 는 부영상의 분류정보를 전송할 때의 비트수를 나타내며, 예를들어 8개의 부류로 분류한다면 $B_c = \log_2 8$ 가 된다. DC요소는 균일분포가 아닌 다른 확률분포를 가지고 있다는 가정하에 평균제곱오차를 최소로 할 수 있는 비균일 양자기를 사용하여 8비트로 스칼라 양자화하였다.

5.2 실험조건

분류벡터양자화에서 영상블록을 분류하는 방법에 따른 영상데이터 압축성능을 비교해보기 위하여 실제 영상데이터에 적용하여 보았다. 이를 위하여 256×256 크기, 256레벨의 픽셀값을 갖는 흑백영상을 압축하는데 에지특징벡터에 의해 에지블록을 분류하는 방법, 변환계수의 에너지의 분포에 의한 분류방법, 그리고 신경회로망에 의하여 변환계수의 에너지분포에 따른 분류방법을 적용하여 부호화하여 세가지 분류방식의 성능을 비교하였다. 시험영상으로 사용한 Lenna와 Cablecar 영상은 모두 다양한 에지방향을 포함하고 있는 자연영상으로, 하나는 사람의 얼굴이고 또 하나는 도심의 풍경으로 세가지 분류방식을 평가하기에 적당한 대표적인 영상이라고 할 수 있다. 부영상의 크기는 효율적인 벡터 분할법과 압축율을 얻기 위하여 8×8 로 하였으며 벡터양자화에 필요한 코드벡터는 8개의 256×256 크기의 영상, 그리고 3개의 512×512 크기의 영상을 학습데이터로 사용하

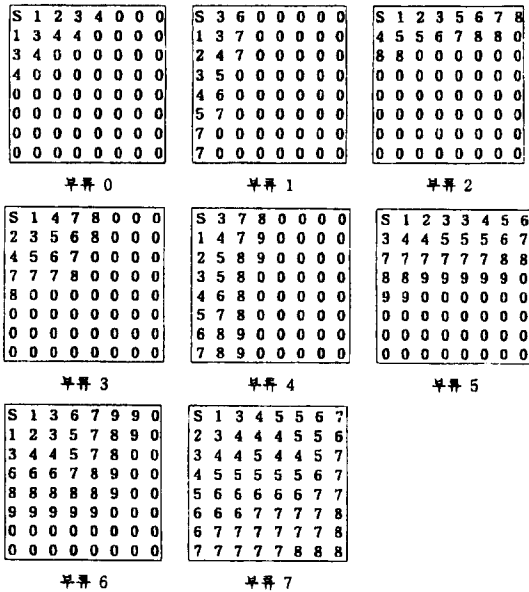


그림 9. 에너지 분포에 의한 분류방식에 의한 각 부류별 부벡터맵.
 Fig. 9. Subvectors at each class in the classification by energy distribution.

므로 공간영역의 에지 방향을 충분히 반영하고 있으며 또한 실제 영상에 존재하는 복잡한 영역까지도 정의하고 있다. 부류 0의 경우 평탄을 분류하므로 부벡터들이 저주파수 대역에만 존재하고 있음을 알 수 있고, 부류 1에서 부류 6까지 각각 부벡터의 분할맵이 각각 좌측 수직 수평 대각방향으로 편중되어 있는 것을 알 수 있다. 그리고 복잡한 영역을 나타내는 부류 7의 경우 전대역에 걸쳐 부벡터 맵이 분포하므로 제안한 분류방식이 효율적으로 에지방향을 잘 분류하고 있음을 알 수 있다.

특징벡터를 이용한 분류에서는 각 부류마다 에지

표 2. 에너지 분포에 의한 분류방식에서의 각 부류별 부벡터의 비트할당

Table 2. Bit assignment for each subvector in the classification by energy distribution

부류	부 벡 터									
	V ₁	V ₂	V ₃	V ₄	V ₅	V ₆	V ₇	V ₈	V ₉	V ₁₀
0	4	3	2	1	-	-	-	-	-	-
1	8	7	5	4	3	2	1	-	-	-
2	8	7	6	5	4	3	2	1	-	-
3	8	7	6	5	4	3	2	1	-	-
4	8	8	7	6	5	4	3	2	1	-
5	8	8	7	6	5	4	3	2	1	-
6	8	8	7	6	5	4	3	2	1	-
7	8	7	6	5	4	3	2	1	-	-

방향과는 관계없이 거의 모든 주파수대역에 부벡터들이 존재한다. 이는 학습에 사용된 부영상이 잘못 분류되어 부벡터로 분할하기 위한 분산행렬에 영향을 미쳐 원하는 에지방향의 에너지분포를 나타내지 못하고 있음을 나타낸다. 그림 9에서 보는 바와 같이 부벡터맵이 모든 부류에서 에지의 방향에 따라 방향성을 가지고 있으므로 원하는 에지방향의 에너지분포를 반영하고 있다.

표 2는 특징벡터를 이용한 분류방법에서 식 (17)에서 구한 평균비트율 R이 0.52[bpp]인 경우에 각 부류별 부벡터에 해당하는 코드벡터의 수를 나타내는 비트수이다. 예를들어 부류 3에서 부벡터 5에 대응하는 비트수 4는 부류 3의 5번째 부벡터 V₅를 벡터양자화 하는데 필요한 코드벡터의 수를 16(=2⁴)으로 한다는

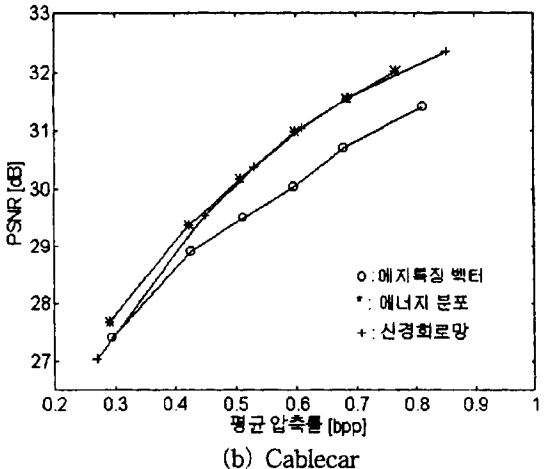
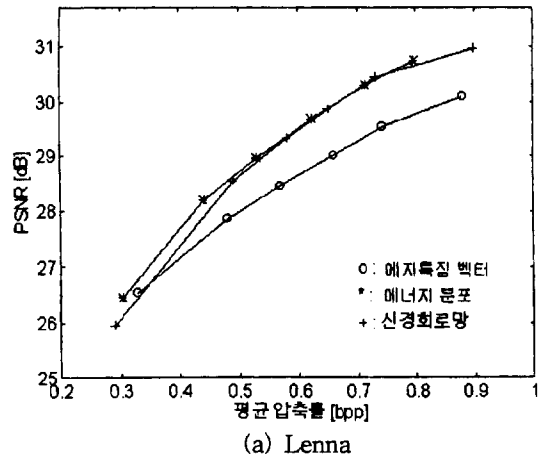


그림 10. 각 분류방식에 의한 (a) Lenna 와 (b) Cablecar 영상의 PSNR.

Fig. 10. PSNRs of the images (a) Lenna and (b) Cablecar by the classification methods.

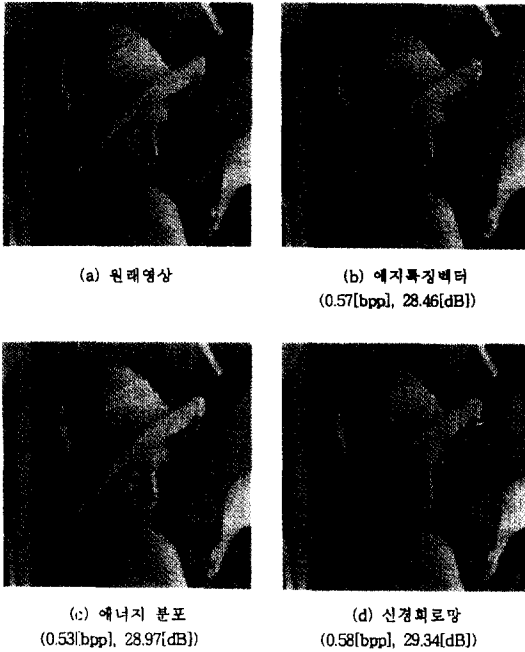


그림 11. 각 분류방식에 의한 Lenna영상의 압축결과.
Fig. 11. Compression result of the Lenna by the classification methods.

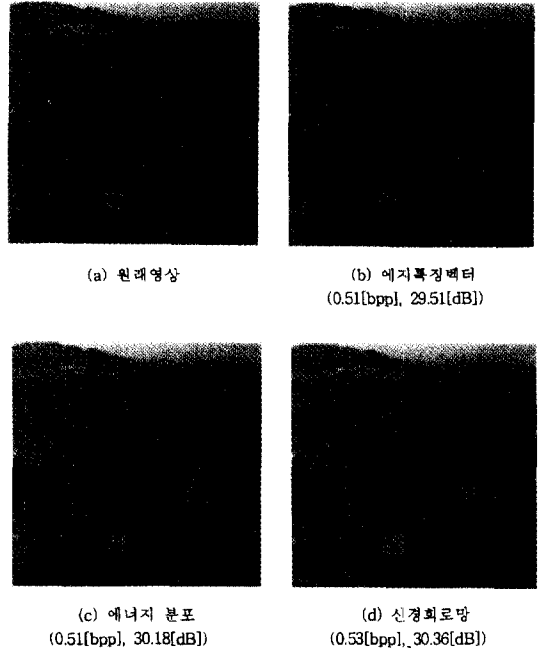


그림 12. 각 분류방식에 의한 Cablecar영상의 압축결과.
Fig. 12. Compression result of the Cablecar by the classification methods.

의미이다. 분할맵에서 부호화시 S는 가우시안 분포를 갖는 최적의 평균제곱양자기를 사용하여 8비트로 스칼라 양자화하였다. 양자화하기 위해서는 먼저 각 부영상의 변환계수의 직류요소를 평균이 0이고, 분산이 1인 신호로 변환하여 양자화하였다. 0인 부분은 무시한다.

그림 10은 Lenna영상과 Cablecar영상을 변환영역에서의 분류벡터양자화에 의하여 평균압축률을 0.3 [bpp]에서 0.8[bpp]로 변화시키면서 압축한 결과이다. 두 영상 모두 제안한 분류방식과 신경회로망에 의한 압축결과가 거의 비슷함을 나타내는데 이는 에지부근의 경우 제안한 분류기준이 잘 학습되었음을 알 수 있다. 특징벡터에 의한 기존의 분류방식은 0.3[bpp]의 평균압축률에서는 신경회로망에 의한 압축결과와 비슷하나 그 이상의 평균 압축률에서는 모든 경우에 작은 값으로 나타나고 있다.

그림 11과 그림 12는 변환영역에서의 분류벡터양자화에 의하여 약 16:1의 평균압축률로 압축한 결과를 나타내고 있다. 에지특징벡터에 의한 분류방식에서는 부분적으로 손상된 에지영역이 발견되고 있는 반면, 에너지분포에 의한 분류와 신경회로망에 의하여 학습된 분류방식은 거의 모든 에지가 잘 보존되어 있으며 에지부근의 손상이 거의 없다. 복잡한 부근에

서는 세가지 분류방식 모두가 시각적으로 차이가 별로 없었는데 이는 에지특징벡터에 의한 분류방식에서는 강한에지 부류가 복잡한 부근까지 포함하고 있다고 할 수 있다.

6. 결 론

벡터양자화에 의한 영상압축에서는 높은 압축률을 실현할 수 있지만 에지부근의 손상에 의해 시각적인 화질이 저하할 수 있으므로 에지특성에 따라 영상블록을 분류하여 별도로 압축하는 분류벡터양자화가 제안되었다. 변환영역에서의 분류벡터양자화는 공간영역에서보다 간단한 분류방식에 의해 에지블록을 분류할 수 있으나, 분류기준이 공간영역의 에지방향과 일치하지 않고 실제 영상에 존재하고 있는 복잡한 부분이 정의되지 않은 문제가 있다.

본 논문에서는 에지부근의 손상을 최소화하고 복잡한 부분을 잘 나타낼 수 있도록 블록내의 에너지의 분포에 의하여 영상블록을 분류하였고 분류된 학습 패턴들을 신경회로망에 의해 학습하여 에지 특성에 따른 영상블록의 분류에 사용하였다. 에너지분포에 의한 분류방식과 이를 학습한 신경회로망의 분류성을 확인하기 위하여 실제 영상의 압축에 적용하였

고 이 결과를 에지특징벡터에 의한 분류방식과 비교하였다. 비교결과 에지특징벡터에 의한 분류방식보다 에너지분포에 의한 분류 및 학습된 신경회로망에 의한 분류방법에 의한 영상압축 결과가 시각적으로 에지부근을 잘 나타내었으며 객관적인 성능지표인 PSNR도 향상되었다.

참고문헌

[1] R. M. Gray, "Vector Quantization," *IEEE ASSP Magazine*, Vol. 1, 4-29, 1984.

[2] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Publishers, 1991.

[3] D. S. Kim and S. U. Lee, "Image Vector Quantizer Based on a Classification in the DCT Domain," *IEEE Trans. on Communications*, Vol. 39, No. 4, 549-556, April 1991.

[4] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. on Comm.*, Vol. COM-28, No. 1, 84-95, Jan. 1980.

[5] B. Ramamurthi and A. Gersho, "Classified vector quantization of images," *IEEE Trans. on Comm.*, Vol. COM-34, 1105-1115, Nov. 1986.

[6] Yo-Sung Ho and Allen Gersho, "Classified Transform Coding of Images Using Vector Quantization," *Proc. of ICASSP*, 1890-1893, 1989.

[7] W. H. Chen and C. H. Smith, "Adaptive coding of monochrome and color images," *IEEE Trans. Comm.*, Vol. COM-25, 1285-1292, Nov. 1977.

[8] K. R. Rao and P. Yip, *Discrete Cosine Transform—Algorithm, Advantage, Applications*, Academic Press, 1990.

[9] J. M. Zurada, *Introduction to Artificial Neural Systems*, West information Access, 1992.

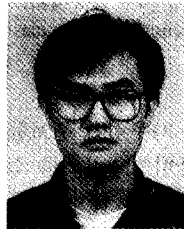
[10] B. Kosko, *Neural Networks and Fuzzy Systems*, Prentice-Hall, 1992.



공성곤 (Seong-Gon Kong) 종신회원

1982년: 서울대학교 전기공학과 (학사)
 1987년: 서울대학교 대학원 전기공학과 (석사)
 1991년: 미국 University of Southern California 전기공학과 (박사)
 1982년~1985년: 금성기전 연구원

1985년~1987년 : 한국 전자통신연구소 연구원
 1992년~현재: 숭실대학교 전기공학과 부교수
 주관심분야: 퍼지시스템, 신경회로망, 신호처리, 지능제어, 패턴분류



이현수 (Hyun-Soo Lee) 정회원

1993년: 숭실대 공대 전기공학과 졸업
 1996년: 동대학원 전기공학과 졸업 (석사)
 현 재: (주)건한 멀티미디어 기술연구소 연구원
 주관심분야: 영상 처리, 인공지능 및 네트워크 통신