

論文97-34C-1-5

# 학습 횟수 조절 신경 회로망을 이용한 영상 신호의 벡터 양자화

## (Vector Quantization of Image Signal using Learning Count Control Neural Networks)

柳大鉉\*, 南基坤\*\*, 尹台焄\*\*, 金在昌\*\*

(D. H. Ryu, K. G. Nam, T. H. Yoon, and J. C. Kim)

### 요 약

벡터 양자화는 영상처리, 음성처리 등에 관련하여 데이터를 압축하는데 유용한 방법이다. 벡터 양자화는 경쟁 학습과 같은 비교사 학습 알고리즘을 갖는 신경 회로망으로 수행될 수 있다. 본 논문은 영상 압축에 적용하기 위하여 빈도 감응 학습 알고리즘에 기반한 학습 횟수 조절 방안을 제안한다. 이 방식은 입력 영상의 특성에 따라 학습 횟수를 달리하여 코드북을 훈련한다. 그 결과 인간 시각에 민감한 영역에 더 많은 코드가 할당되어 복원 영상의 주관적 화질을 향상시킨다. 본 논문에서는 대역 통과 특성을 갖는 변조 전달함수와 밝기 변화에 대한 비선형 특성으로 인간 시각을 모델화하여 사용하였다.

### Abstract

Vector quantization has been shown to be useful for compressing data related with a wide range of applications such as image processing, speech processing, and weather satellite. Neural networks with unsupervised learning algorithm, e.g. competitive learning, can perform the vector quantization of images. This paper proposes a efficient neural network learning algorithm, called learning count control algorithm based on the frequency sensitive learning algorithm. This algorithm can train a codebook in different learning count according to the characteristics of the input image contents. In result more codewords can be assigned to the sensitive region of the human visual system and the quality of the reconstructed image can be improved. We use a human visual system model that is a cascade of a nonlinear intensity mapping function and a modulation transfer function with a bandpass characteristic.

### I. 서 론

벡터 양자화는 프로토타입 패턴으로 구성된 코드북을 사용하여 입력 블록을 인코딩(encoding)하는 방법으로서 일종의 패턴 매칭 기법이라고도 할 수 있다. Shannon의 rate-distortion 이론에 의하면, 스칼라 양

자화 방식보다 벡터 양자화 방식이 항상 나은 성능을 가진다고 알려져 있으며, 종래의 준최적(sub-optimal) 기법들과 rate distortion 이론이 제시하는 소스 코딩의 이론적 한계 사이의 공백을 메워준다<sup>[1,2]</sup>. 벡터 양자화는 이러한 관점에서 연구가 시작된 비교적 새로운 압축 방식으로서 많은 연구가 이루어져 왔다.

벡터 양자화의 문제점은 벡터의 크기가 커짐에 따라 코드북의 생성과 최적화 및 인코딩시의 최적 코드어 탐색에 필요한 계산량이 지수 함수적으로 증가한다는 것이다. 이러한 이유 때문에 실제의 구현에서는 작은 블록 크기의 벡터 양자화가 이용되어 왔다. 블록의 크기가 작을 경우 영상에 존재하는 공간적 상관성을 충

\* 正會員, 韓國 電子通信 研究所 (ETRI)

\*\* 正會員, 釜山大學校 電子工學科 (Dept. of Electronics Engineering, Pusan National University)

接受日字:1996年11月12日, 수정완료일:1997年1月17日

분히 이용하지 못하게 되어 코딩의 효율성이 떨어지고, 저 전송율에서 블럭의 경계에 눈에 매우 거슬리는 블럭 효과를 발생시킨다. 한편, 블럭의 크기를 크게하면 화질은 향상될 수 있지만 계산 비용이 크게 증가된다<sup>13)</sup>.

이러한 문제점을 해결하기 위하여 신경 회로망을 이용한 벡터 양자화 방법들이 많이 연구되어 왔다. 벡터 양자화를 비교사 경쟁 학습(unsupervised competitive learning; CL) 알고리즘을 갖는 신경 회로망으로 수행하면 신경 회로망의 병렬 처리 구조로 인해 실시간 처리가 가능하고 신경 회로망의 다양하고 유연한 학습 알고리즘에 의해 벡터 양자화의 성능을 개선할 수 있는 장점이 있다<sup>14)</sup>.

본 논문에서는 벡터 양자화를 위한 코드북 생성시 인간의 시각 특성을 반영할 수 있는 학습 알고리즘을 제안한다. 일반적으로 자연 영상의 경우 인간의 시각이 민감한 에지 등의 영역이 전체 영상에서 차지하는 비율이 상대적으로 적다. 따라서 이러한 자연 영상을 이용하여 코드북을 학습하는 벡터 양자화에서 에지 등 인간의 시각이 민감한 영역의 학습이 잘 이루어지지 않는다.

이러한 점에 착안하여 학습 과정에서 입력 패턴의 시각 민감도에 따라 학습 횟수를 조절함으로써 신경 회로망의 코드북 학습 과정에서 인간 시각 특성을 반영하도록 한다<sup>15)</sup>. 이를 위하여 배경의 밝기 변화에 대한 비선형성과 공간 주파수 상에서의 대역 통과 특성으로 인간 시각 시스템을 모델화하여 사용하였다.

본 논문의 구성은 다음과 같다. II장에서는 벡터 양자화 및 신경 회로망 벡터 양자화의 개념에 대하여 간단히 설명한다. III장에서는 먼저 잘 알려진 인간 시각 특성인 밝기 변화에 대한 비선형 특성과 공간 주파수 특성을 반영한 인간 시각 시스템을 모델화한다.

그리고 인간 시각 특성을 반영한 코드북을 학습하기 위해 학습 횟수 조절 방식을 제안한다. IV장에서는 실제 영상에 대해 모의 실험을 수행하고 그 결과를 보인다. V장에서는 본 논문의 내용 및 중요한 결과를 요약한다.

## II. 신경 회로망 벡터 양자화

본 장에서는 벡터 양자화 및 신경 회로망을 이용한 벡터 양자화의 기본 개념과 특성에 대하여 설명한다.

### 1. 벡터 양자화

벡터 양자화는 입력 데이터에 내재하는 통계적 특성을 이용하는 양자화 기법이라 할 수 있다. 먼저 입력 벡터 공간을 여러 개의 영역으로 나누고 각 영역에 대해 대표 벡터를 계산한다. 다음, 양자화될 데이터 벡터가 주어지면 그 벡터가 속한 영역을 찾고 그 영역에 대한 대표 벡터의 인덱스에 의해 그 입력 데이터 벡터를 표현한다. 다시 말하면 벡터 양자화란 적절한 거리 척도(distance measure)에 의해 이미 저장된 기준 벡터(reference vector) 중에서 가장 잘 일치하는 벡터로 매핑시켜주는 코딩 방식이다<sup>11) 12)</sup>. 따라서 이 벡터 양자화 매핑은

$$VQ: x = (x_1, x_2, \dots, x_k) \rightarrow m \quad (1)$$

과 같이  $k$ 차원 벡터 공간으로부터  $M$ 이라는 인덱스의 집합으로의 매핑으로 정의될 수 있다. 여기서  $m \in M$  이고,  $M$ 의 크기는  $M$ 이다. 또한 인덱스  $m$ 으로 부터 재생 벡터  $y_m$ 을 찾을 수 있다.

벡터 양자화에서는 먼저 시험 데이터에서 추출한 훈련 벡터들을 이용하여 입력 벡터의 대표 벡터들의 집합인 코드북을 만든다. 코드북을 구성하기 위해서는 모든 훈련 데이터와, 그와 가장 유사한 코드어 사이의 평균 왜곡이 최소화되도록 하는 LBG(Linde, Buzo and Gray) 알고리즘이 많이 이용된다. 코딩 시에는 입력 벡터와 가장 유사한 코드어를 코드북에서 찾는다. 또한 전송시 이 코드어를 전송하는 것이 아니라 코드어의 인덱스(index)를 전송함으로써 데이터량을 줄인다.

주어진 성능 기준에 대해 최적인 벡터 양자화 시스템의 설계는 주어진 기준에 대해 최적인 코드북을 결정하는 것과 같다. 이것은 일반적으로 입력 데이터에 대한 통계적 특성을 알아야만 가능하지만 입력 데이터에 대한 통계적 특성은 보통 알려져 있지 않으므로 코드북은 일반적으로 훈련(training)이라고 부르는 과정에 의해 구성된다. 따라서 훈련 과정에서는 실제의 데이터와 유사한 통계적 특성을 갖는 데이터를 사용하여야만 최적의 코드북을 생성할 수 있다.

훈련 과정에서는 왜곡 척도  $d(x, y)$ 를 사용하여 입력 데이터가 어느 영역에 속하는지를 결정한다. 왜곡 척도는  $x$ 를  $y$ 로 표현하는 비용(cost)으로 볼 수 있다. 훈련 데이터 벡터가 같은 영역에 속하는지를 결정함으로써  $k$ 차원의 입력 데이터 벡터를 셀들로 분할한다.

## 2. FSCL 을 이용한 벡터 양자화

앞 절에서 언급된 바와 같이 코드북을 생성하기 위한 훈련과 인코딩 과정은 매우 복잡하고 계산량도 많다. 또한 LBG 알고리즘과 같이 현재 사용되는 대부분의 알고리즘은 배치 모드 알고리즘(batch mode algorithm)이며, 훈련 과정에서는 훈련 데이터를 모두 액세스 해야 할 필요가 있다. 또한 입력 데이터의 통계적 특성이 상황에 따라 변할 수 있는 경우에는, 입력이 들어올 때 마다 코드북 벡터를 변경하는 적응 벡터 양자화 방식을 사용하는 것이 유용하다. 신경 회로망에는 많은 우수한 훈련 알고리즘들이 있어 이러한 적응 벡터 양자화를 적용하는데 효과적이다<sup>[6,7]</sup>.

비교사(unsupervised) 경쟁 학습 알고리즘을 갖는 신경 회로망으로 벡터 양자화를 수행하면 신경 회로망의 병렬 처리 구조로 인해 실시간 처리가 가능하고 신경 회로망의 다양한 학습 방법에 의해 벡터 양자화 알고리즘이 개선될 수 있는 장점이 있다. 신경 회로망을 이용한 벡터 양자화는 다음과 같이 인코딩 과정과 훈련 과정으로 이루어진다.

### 1) 인코딩

양자화될 벡터를  $k$  차원 벡터 공간으로 두고, 왜곡 척도  $d(x, y)$ 가 이 공간에서 정의된다고 하자. 코드북의 크기를  $M$ 으로 두고 코드어는  $w_i = 1, \dots, M$ 으로 둔다.  $M$ 개의 신경 유닛을 갖는 신경 회로망에서  $i$ 번째 코드어  $w_i$ 를 신경 유닛  $i$ 에 대한 가중치 벡터로 둔다. 인코딩될 어떤 벡터  $x$ 가 주어질 때,  $x$ 는  $M$ 개의 신경 유닛에 병렬로 입력된다. 각 신경 유닛에서 입력 벡터와 그 가중치 벡터와의 왜곡은

$$d_i = d(x, w_i), i = 1, \dots, M \quad (2)$$

에 의해 계산된다. 그리고, 입력 벡터는

$$d_i = \min_j d_j \quad (3)$$

과 같이 최소 왜곡을 갖는 신경 유닛의 인덱스  $i^*$ 로 인코딩된다. 최소 왜곡을 갖는 신경 유닛을 결정하는 것과 그 인덱스를 부여하는 것을 제외한 이러한 일련의 계산 과정은 모두 병렬로 수행될 수 있다. 또한 최소 왜곡을 갖는 신경 유닛을 결정하는 데는 여러 가지의 신경 회로망 학습 알고리즘이 사용될 수 있다.

### 2) 훈련(training)

벡터 양자화를 신경 회로망으로 구성하는 가장 큰

장점은 벡터 양자화를 훈련하는데 CL, KSF(M(Kohonen self-organizing feature map) 등 다양한 신경 회로망 학습 알고리즘을 적용할 수 있다는 것이다. CL 신경 회로망에서는 먼저 가중치 벡터  $w_i$ 를 랜덤한 값으로 초기화시킨다. 그리고 각 입력 벡터  $x$ 가 주어질 때 마다

$$z_i = \begin{cases} 1 & \text{if } d(x, w_i) \leq d(x, w_j), j = 1, 2, \dots, M \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

에서 왜곡  $d(x, y)$ 를 최소로 하는 승리 출력 신경 유닛  $z_i$ 와 승리 가중치 벡터(winning weight vector)  $w_i$ 를 찾는다. 새로운 승리 가중치 벡터  $w_i(n+1)$ 는

$$w_i(n+1) = (x - w_i(n+1)) z_i \quad (5)$$

으로 계산된다. 위 식에서 파라미터  $\epsilon$ 은 학습율(learning rate)이며 이는 학습이 진행됨에 따라 서서히 감소되어 영이 된다. 이 알고리즘은 학습 데이터를 반복 사용하여 훈련하며 이러한 학습 알고리즘의 문제점은 사용되지 않는 가중치 벡터가 발생하여 전체 평균 왜곡을 증가시키는 것이다. 이 문제를 해결하기 위한 방법으로 KSF(M은 학습 과정에서 승리 가중치 벡터들과 이웃하는 가중치 벡터들과 상호 연관성을 두어 승리 가중치의 값 뿐 아니라 이웃한 가중치 벡터들도 같이 그 값을 조정하도록 하였다. 그러나 이 방법은 계산량이 많고 구조가 복잡하다는 단점이 있다.

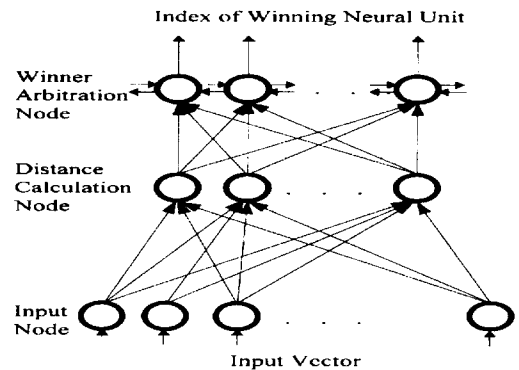


그림 1. FSCL 의 구조

Fig. 1. Structure of FSCL.

FSCL에서는 학습 과정에서 사용되는 왜곡 척도를

$$d(x, w_i) = d(x, w_i(n))F(u_i) \quad (6)$$

과 같이 수정하여 이러한 문제를 해결하였다. 여기서  $F(u_i)$ 는 신경 유닛  $i$ 가 학습 과정에서 승리한 횟수인  $u_i$ 의 함수이다. 이 경우에서도 승리 신경 회로망 유닛은 학습 과정에서  $d$ 가 최소인 신경 유닛이 된다. 그림 1에 이러한 FSCL의 구조가 나타나 있다<sup>[6]</sup>.

### III. 학습 횟수 조절 신경 회로망 벡터 양자화를 이용한 영상 압축

#### 1. 인간 시각 특성의 모델화

일반적으로 영상에서 주파수가 높은 영역의 오차는 인간이 잘 인식하지 못하며 배경 부분의 밝기에 따라 포함되어야 한다. 예를 들면 주파수에 따른 효과를 이용하기 위해서는 영상을 DCT 또는 웨이브렛 변환하여 변환 영역에서 처리하는 것이 효과적일 수 있다. 본 논문에서는 그림 2에 나타난 것과 같은 밝기에 따른 시각의 감도 차이와 공간 주파수에 따른 감도의 차만을 고려한 모델을 사용하였다<sup>[8]</sup>.

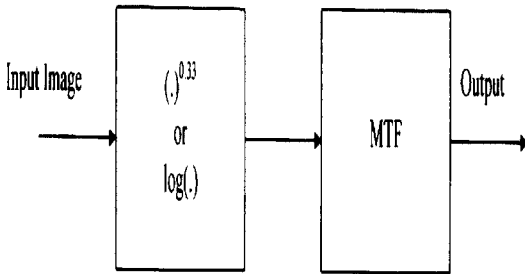


그림 2. 시각 시스템 모델  
Fig. 2. Human visual system model.

#### 1) 배경의 밝기 변화에 따른 비선형 특성

인간의 눈은 밝은 영역 보다 어두운 영역에서의 변화에 더욱 민감하다. 영상의 어떤 국부적 영역에 대한 인간 시각의 반응은 그 영역의 배경의 밝기가 밝을수록 민감도가 떨어지며, 배경이 어두울 경우의 잡음에 대해서는 매우 민감하다. 이는 배경의 밝기에 대한 비선형적인 특성으로서 Weber의 법칙으로 표현된다<sup>[9]</sup>. 따라서 영상 압축시 배경의 밝기를 고려함으로써 시각 심리학적인 중복성(psychovisual redundancy)을 제거하여 압축 효율을 높일 수 있을 뿐아니라, 재생 영상

의 주관적인 화질도 향상시킬 수 있다.

이와 같은 특성은 그림 3과 같이 로그 함수 혹은 지수 함수에 의해 근사화 할 수 있다. 영상 압축 시스템에서는 양자화 단계 이전에 영상의 밝기 값에 대하여 이러한 비선형 변환을 수행한 다음 양자화하고 수신측에서는 복원 영상에 대해 역변환을 수행함으로써 이를 보상한다<sup>[9,10,11]</sup>.

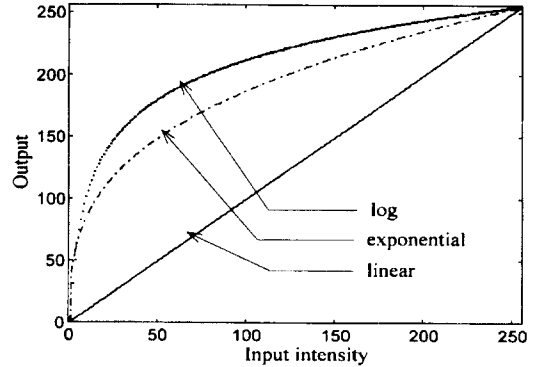


그림 3. 밝기에 대한 비선형 특성 모델화  
Fig. 3. Nonlinear characteristic modeling of brightness.

#### 2) 공간 주파수 특성

영상 압축 시스템에서 양자화 잡음에 대한 시각의 민감도는 공간 주파수 특성에 따라 다르다. 즉 인간의 눈은 저주파나 고주파 대역에 비해 중간 주파수 대역에서 보다 높은 감도를 갖는다. 이와 같은 대역 통과 필터링 특성은 광학적인 측면에서 눈의 불완전성에 의한 저역 통과 필터링 특성과 망막(retina)에 존재하는 광수용기(photo-receptor)들 상호간의 측방 억제(lateral inhibition)에 의한 고역 통과 필터링 특성이 결합되어 나타나는 것으로서 공간 주파수 영역에서 변조 전달 함수(modulation transfer function: MTF)로 모델화 된다. 변조 전달 함수는 인간 시각의 공간 주파수 특성을 수식으로 모델화한 것으로서 일반적으로 격자 모양의 정현파 패턴을 이용한 심리 시각적인 실험을 통하여 얻어진다. 이러한 공간 주파수와 관련된 변조 전달 함수의 특성을 이용하여 영상의 각 주파수 대역에 대하여 각각 다른 방법으로 양자화하면 주관적으로 보다 우수한 화질을 얻을 수 있다.

영상에서의 공간 주파수는 화면과 인간의 눈과의 거리에 따라 재정의 되는데 그 단위는 보통 cpd(cycle per degree)를 사용한다. 연구에 의하면 인간의 시각

은 공간 주파수 4.7 ~ 8 cpd에서 최대의 민감도를 나타내며 0.1과 30 ~ 40 cpd에서 최대치의 3%의 민감도를 나타낸다고 알려져 있다. 결국 공간 주파수가 중간 정도인 부분에서는 인간 시각이 민감하므로 양자화 시 이를 고려하여야 한다. Chitprasert와 Rao가 제시한 변조 전달 함수는

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

로 주어지며 그림 4에 이 변조 전달 함수의 1차원 특성을 나타내었다<sup>12)</sup>.

이러한 특성을 영상 압축 시스템에 적용하기 위해서는 일반적으로는 영상을 DCT나 웨이브렛 변환 등을 통하여 공간 주파수 영역으로 변환하여 변환 영역에서 처리한다. 한편 이러한 시각의 공간 주파수 특성에 대한 신경 생리학적인 연구 결과에 의하면 변조 전달 함수는 가우스 함수를 2차 미분한 형태인 LOG(Laplacian of a Gaussian) 함수로 나타나며 이는 분산비가 서로 다른 두 가우스 함수의 차인 DOG(Difference of two Gaussians)과 유사함이 알려져 있다<sup>13)</sup>.

본 논문에서는 DOG 함수를 이용하여 영상을 필터링하여 시각적으로 민감한 부분을 강조한 다음 신경 회로망 벡터 양자화 시스템에서 이러한 영역을 효과적으로 벡터 양자화할 수 있도록 하여 HVS의 밝기에 대한 비선형 특성과 공간 주파수 특성이 반영되도록 하였다.

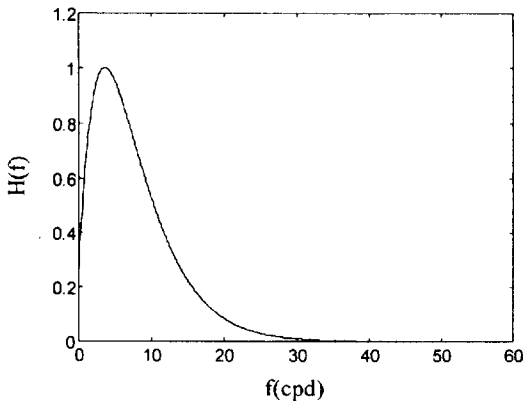


그림 4. MTF 특성

Fig. 4. MTF characteristic.

2. 신경 회로망 벡터 양자화를 이용한 영상 압축  
신경 회로망 벡터 양자화를 이용한 영상 압축 시스템에서도 맨 먼저 코드북을 생성하여야 한다. 신경 회

로망을 이용한 코드북 생성 과정이 그림 5에 나타나 있다. 코드북 생성을 위해서는 시험 영상을 입력으로 하여 훈련 또는 학습이라는 과정을 거치게 되는데, 시험 영상은 먼저  $nm$ 의 부블럭으로 나누어져 신경 회로망에 입력되어 신경 회로망을 학습한다. 신경 회로망의 입력층의 뉴런의 수는  $k(=nm)$ 개가 되고 이는 결국 입력 벡터의 차수가 된다.

각 입력 뉴런은 모두 두번째 계층인 왜곡 오차 계산 뉴런으로 연결되어 왜곡 오차를 계산한다. 왜곡 오차 계산 뉴런에서 계산된 왜곡 오차는 출력층에 전달되고, 출력층에서 가장 작은 오차를 갖는 승리 뉴런을 결정한다. 승리 뉴런에 연결된 가중치 벡터를 경쟁 학습의 경우에 식 (5)에 의해 갱신한다. 갱신 과정에서 가중치 벡터는 입력 데이터 벡터 쪽으로 정해진 학습율 만큼 이동한다.

이러한 학습 과정은 기준 출력을 필요로 하지 않으므로 비교사 학습 과정이며, 입력 데이터들을 비슷한 통계적 특성을 가지는 군들로 분류하는 기능을 수행한다. 학습은 왜곡 오차가 주어진 임계값 이하로 떨어질 때까지 계속되며 이들 뉴런에 관련된 가중치 벡터가 결국 코드북이 된다.

일단 코드북이 얻어지면 송신측과 수신측이 나누어 가진다. 송신측에서는 입력 데이터 벡터에 대하여 코드북의 코드어들 중에서 가장 유사한 코드어를 찾고 그 인덱스를 전송한다. 이때 유사성의 척도로서 코드북 생성 과정에서 사용한 왜곡 척도를 사용한다. 코드북의 코드어들 중에서 입력 데이터 벡터와 가장 유사한 코드어를 찾는 과정은 입력 데이터 벡터의 차수와 코드북의 크기가 큰 경우 계산량이 매우 많다.

일반적으로 코드북 생성 과정은 오프라인으로 이루어 질 수도 있으나 송신측의 인코딩 과정은 실시간으로 처리되어야 하므로 계산량의 증가는 하드웨어 구현에 있어서 장애가 된다. 그러나 신경 회로망 벡터 양자화는 코드북 생성 과정과 인코딩 과정이 모두 병렬로 처리되므로 계산시간을 크게 줄일 수 있다. 또한 인코딩 과정의 신경 회로망은 가중치 갱신 과정이 없을 뿐 코드북 생성을 위한 신경 회로망 구조와 유사하므로 동일한 하드웨어로 구현하여 사용할 수 있다<sup>14)</sup>.

수신측에서는 수신된 인덱스를 주소로 하여 테이블 참조에 의해 수신 데이터 벡터를 복원한다. 고정된 하나의 코드북을 사용하는 시스템에서는 송신측에서만 코드북을 생성하므로 수신측의 구조는 매우 단순하다.

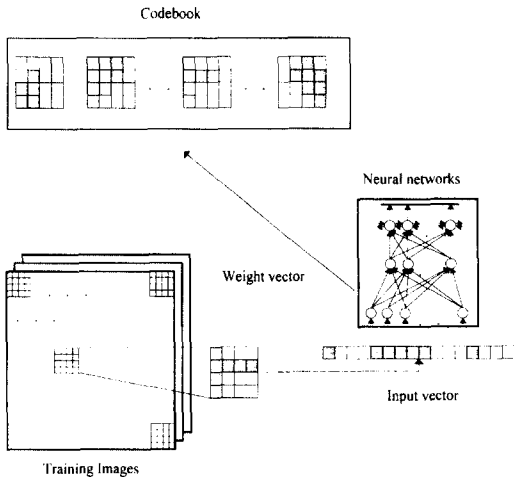


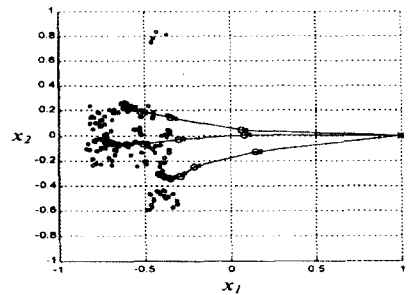
그림 5. 신경 회로망 벡터 양자화에 의한 코드북 생성 과정  
 Fig. 5. Codebook generation procedure by neural networks vector quantization.

3. 학습 횟수 조절 방식

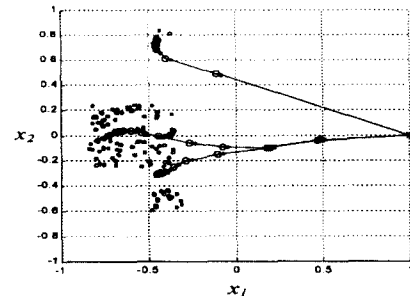
일반적으로 자연 영상은 밝기의 변화가 많지 않은 평탄한 영역이 많은 부분을 차지하고, 인간의 시각이 민감한 에지와 같은 부분은 차지하는 비중이 적다. 이러한 자연 영상에 대해 등화률 원리를 구현하는 LBG 알고리즘이나 FSCL 같은 일반적인 학습 알고리즘을 이용하여 코드북을 생성하는 경우 많은 비중을 차지하는 평탄한 영역에 많은 코드어가 할당된다. 이 경우 객관적인 화질의 척도인 PSNR 등의 측면에서는 좋은 결과를 얻을 수 있으나 전체 영상에서 적은 비중을 차지하는 인간의 시각이 민감한 영역은 충실히 재생하지 못하는 결과를 낳을 수 있다.

본 논문에서는 학습시 훈련 벡터의 특성에 따라 학습 횟수를 조절함으로써 특정 성질을 갖는 입력 패턴에 대해서 더 많은 수의 코드어를 할당하게 하는 방식을 제안한다. 이러한 방식은 영상 데이터를 벡터 양자화할 때 인간의 시각이 보다 민감한 영역에 더 많은 코드어를 할당하도록 코드북을 설계함으로써 주관적인 화질의 측면에서 성능을 개선할 수 있다. 제안된 학습 횟수 조절 방식은 입력 블록의 시각적 활동도에 따라 학습 횟수를 조절한다. 즉 시각적 활동도가 큰 블록에 대해서는 학습 횟수를 상대적으로 크게 함으로써 학습이 잘 이루어지도록 한다. 시각적 활동도는 여러가지 방법으로 정의 가능하나 본 논문에서는 인간 시각 모델을 통과한 영상의 블록 분산 값을 이용한다.

2차원 벡터들에 대한 학습 과정을 그림 6에 나타내었다. 그림 6(a)에서 좌측 상단의 갯수가 적은 군은 제대로 코드북에 반영되지 못함을 보여주고 있다. 그림 6(b)는 학습 횟수 조절 방식으로 학습하는 과정을 나타내었다. 이 그림에서 특정 입력 패턴군에 대하여 학습 횟수를 증가시킴으로써 뉴런의 가중치 벡터가 특정 군 쪽으로 이동해 감을 볼 수 있다. 따라서 영상 데이터를 벡터 양자화할 때 인간의 시각이 민감한 영역에 더 많은 코드어를 할당하여 잘 표현 할 수 있도록 코드북을 학습함으로써 주관적 화질의 향상을 얻을 수 있다.



(a)



(b)

그림 6. 코드북 학습(3개의 데이터 군) : (a) FSCL에 의한 학습, (b) 학습 횟수 조절에 의한 학습  
 Fig. 6. Codebook training (3 data cluster) : (a) training by FSCL, and (b) training by learning count control.

IV. 모의 실험 및 고찰

일반적으로 코드북 생성을 위한 훈련 영상은 여러 특성을 가지는 많은 영상을 이용하는 것이 바람직하다. 본 모의 실험에서는 그림 7(a)와 같이 크기가 256x256 pixel 이며 8 bit/pixel의 밝기 해상도를 갖는 4장의 영상을 훈련 영상으로 이용하고 그림 7(b)를 시험 영

상으로 이용하였다. 입력 데이터 벡터는 44의 부분력으로 구성하였으며 훈련 데이터 벡터의 수는 모두 16,384개가 된다. 벡터 양자화기의 코드어의 개수는 256개로 설계하였다. 따라서 입력 벡터의 차수  $k$ 가 16(=44)이며, 코드북의 크기  $N$ 이 256이므로 비트율은 0.5 bit/pel이 되며 압축비는 16이 된다.

본 실험에서는 객관적인 화질의 향상에 관심을 두고 있으나 객관적인 평가 기준으로는

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \quad (8a)$$

$$MSE = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} |F(i, j) - F'(i, j)|^2 \quad (8b)$$

으로 정의되는 PSNR도 함께 이용하였다.

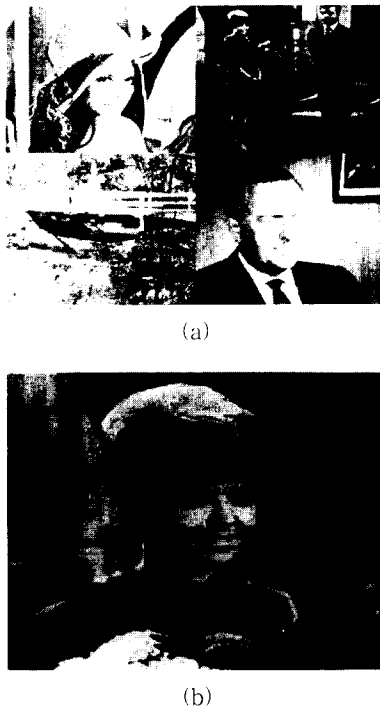


그림 7. 훈련 영상 및 시험 영상 : (a) 훈련 영상, (b) 시험 영상

Fig. 7. Training and test images : (a) training images, and (b) test image.

모의 실험에서는 먼저 종래의 벡터 양자화 방식중 가장 많이 사용되는 LBG 알고리즘을 이용하여 코드북을 생성하고 이 코드북을 이용하여 인코딩과 디코딩을 하였으며 그 결과를 그림 8(a)에 나타내었다. 또한 제안된 방식들과의 성능 비교를 위하여 FSCL의 초기 가운데 벡터를 영벡터로 하고 fairness 함수를 제공된 함

수로 하며 학습율을 0.1로 하여 모의 실험 한 결과를 기준 결과 영상으로 사용하였으며 그림 8(c)에 나타내었다. 본 논문에서 제안한 학습 횟수 조절 방식을 적용한 결과는 그림 8(e)에 나타내었다. 이 그림에서 (b), (d), (f)는 6464 크기의 부분 확대 영상이다. 실험에서는 훈련 영상을 변조 전달 함수에 의해 대역 필터링하고 비선형 변환한 다음 코드북을 생성한다. 이때 입력 블록의 시각적 활동도에 따라 학습 횟수를 조절한다. 학습 횟수 조절 방식에 의해 생성된 코드북으로 시험 영상을 코딩한 경우에 주관적 화질의 측면에서 뿐만 아니라 PSNR의 측면에서 볼 때도 FSCL을 적용한 방식보다 나은 성능을 보이고 있다.

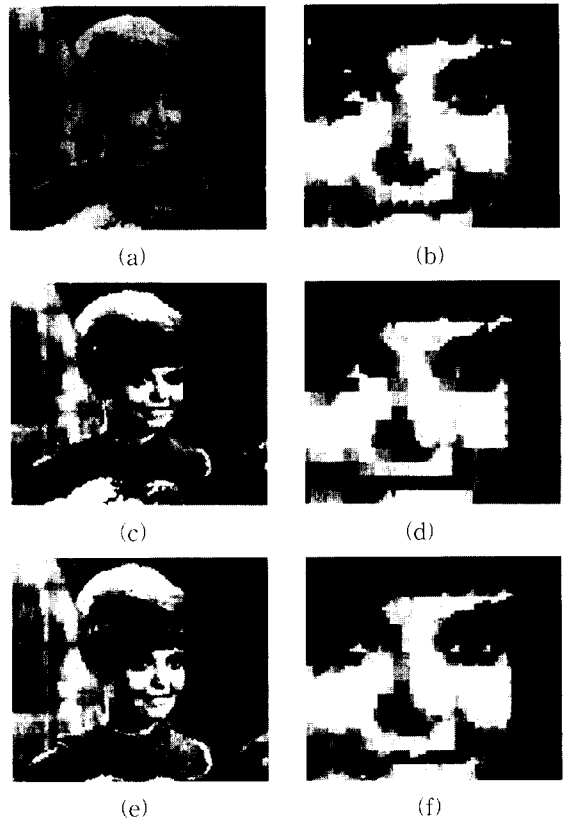


그림 8. 재생 영상 : (a) LBG(PSNR=25.8877), (c) FSCL(PSNR=27.7845), (e) 학습횟수 조절 방식(PSNR= 28.1566), (b) (a)의 부분 확대 영상, (d) (c)의 부분 확대 영상, (f) (e)의 부분 확대 영상

Fig. 8. Reconstructed images : (a) LBG(PSNR=25.8877), (c) FSCL (PSNR = 27.7845 ), (e) learning count control method (PSNR=28.1566), (b) enlarged part of (a), and (d) enlarged part of (c), and (f) enlarged part of (e).

제안된 방식들에 의하여 구해진 코드북의 코드어들이 시각적 엔트로피가 큰 영역에 많이 할당되었는지를 확인하기 위하여 그림 9에 결과 코드북의 코드어들의 분산값의 분포를, 그림 10에 그 히스토그램을 나타내었다. 이 그림에서 선택적 학습 또는 학습 횟수 조절 방식을 적용한 경우에 코드어들의 분산 값들이 FSCL만으로 학습한 경우보다 큰 것을 알 수 있다. 이는 제안된 방식을 적용한 경우에 코드어들이 인간의 시각이 보다 민감한 부근에 더 많이 할당 되었음을 의미한다.

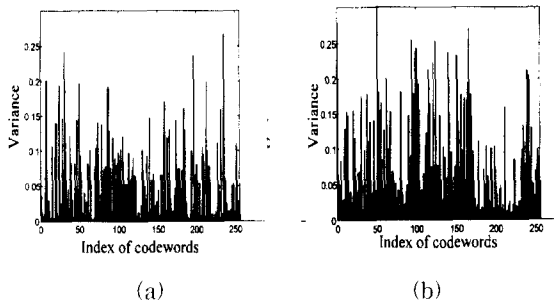


그림 9. 결과 코드북의 분산 값 : (a) FSCL, (b) 학습 횟수 조절  
 Fig. 9. Variance of the result codebook : (a) FSCL, and (b) learning count control.

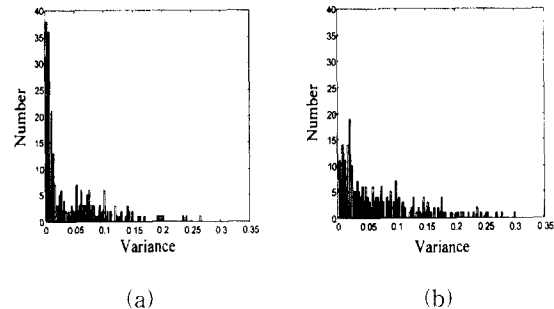


그림 10. 코드북의 분산 값의 히스토그램 : (a) FSCL, (b) 학습횟수 조절.  
 Fig. 10. Histogram of the codebook variance : (a) FSCL, and (c) learning count control.

V. 결론

신경 회로망을 이용한 벡터 양자화 방법은 기존의 벡터 양자화 방법들과 조금 나은 성능을 보이며 병렬 처리 특성으로 인하여 하드웨어로 구현될 경우 기존의 벡터 양자화의 계산량으로 인한 문제를 해결할 수 있을 뿐만 아니라 적응형 벡터 양자화에 적합한 특성을 가지고 있다.

본 연구에서는 벡터 양자화를 위한 코드북 생성시 인간 시각 특성이 반영될 수 있는 신경 회로망 학습 알고리즘을 제안하였다. 먼저 코드북 생성시 고려되어야 할 인간 시각 특성중 배경의 밝기 변화에 대한 비선형성과 공간 주파수 상에서의 대역 통과 특성을 반영할 수 있도록 시각 시스템을 모델화하였다. 이를 이용하여 코드북 생성을 위한 훈련 영상의 각 블록에 대한 시각 활동도를 구하였다. 다음, 신경 회로망이 학습 데이터를 이용하여 코드북을 학습하는 과정에서 각 블록의 시각 활동도에 따라 학습 횟수를 조절함으로써 인간 시각 특성을 반영하도록 하였다. 제안된 방식은 인간의 시각이 민감한 영역의 재생 화질을 향상시킬 수 있음을 모의 실험에 의해 확인하였다.

참 고 문 헌

- [1] C. E. Shannon, A mathematical theory of communication, *BSTJ*, vol. 27, pp. 379-423, 623-656, 1948.
- [2] R. M. Gray, Vector quantization, *IEEE ASSP Magazine*, vol. 1, pp. 4-29, Apr. 1984.
- [3] Y. Linde, A. Buzo, and R. M. Gray, An algorithm for vector quantizer design, *IEEE Trans. Commun.*, vol. COM-28, pp. 84-95, Jan. 1980.
- [4] S. C. Ahalt, A. K. Krishnamurthy, P. Chen and D. E. Melton, "Competitive learning algorithms for vector quantization," *Neural Networks*, vol. 3, no. 3, pp. 277-290, 1990.
- [5] 류대현, 남기곤, 윤태훈, 김재창, 선택적 학습을 이용한 영상 신호의 벡터 양자화, *제5회 한국화상학회 학술발표회 논문집*, pp. 48-60, 5월 1996
- [6] J. Naylor and K. P. Li, Analysis of a neural network algorithm for vector quantization of speech parameters, *Proceedings of the First Annual INNS Meeting*, New York, pp. 310-313, 1988.
- [7] N. M. Nasrabadi and Y. Feng, Vector quantization of images based upon the Kohonen self-organizing feature maps, *IEEE International Conference on Neural Networks*. San Diego, pp. 1101-1108, 1988.
- [8] B. Chitprasert and K. R. Rao, Human visual



- weighted progressive image transmission, *ICCS88, Intl. Conf. on Commun. System, Singapore*, pp 1101-1103, Nov. 1988.
- [9] J. L. Mannos and D. J. Sakrison, The effects of a visual fidelity criterion on the encoding of images, *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 525, July 1974.
- [10] D. J. Granrath, The role of human visual models in image processing, *Proc. IEEE*, vol. 69, pp. 552-560, May 1981.
- [11] 황재정, 이문호, 시각 시스템(Human Visual System)에 의한 영상 정보 압축 기술, *전자공학 회지*, 제 15권, 제2호, pp. 148-154, 4월 1988.
- [12] B. Chitprasert and K. R. Rao, Human visual weighted progressive image transmission, *IEEE Trans. Comm.*, vol. 38, pp. 1040-1044, 1990.
- [13] D. Marr and E. C. Hildreth, Theory of edge detection, *Proc. R. Soc. London*, vol. B207, pp. 187-217, 1980.

---

 저 자 소 개
 

---

柳大鉉(正會員) 第33卷 B編 1號 참조  
현재 한국전자통신연구소 재직 중

南基坤(正會員) 第33卷 B編 7號 참조  
현재 부산대학교 전자공학과 부교수

尹台燾(正會員) 第32卷 B編 10號 참조  
현재 부산대학교 전자공학과 부교수

金在昌(正會員) 第33卷 B編 7號 참조  
현재 부산대학교 전자공학과 부교수