

論文96-33B-11-5

Karhunen-Loève변환과 분류 벡터 양자화에 의한 효율적인 영상 부호화

(Karhunen-Loève Transform-Classified Vector Quantization for Efficient Image Coding)

金泰容*, 崔興文**

(Tae-Yong Kim and Heung-Moon Choi)

요 약

본 논문에서는 PCNN(principal component neural network)을 사용하여 KLT (Karhunen-Loève transform)를 고속화하고 질감 에너지 척도(texture energy measure)에 의한 영상 블럭의 통계적 특성 분류에 따라 분류 벡터 양자기를 사용함으로써 주어진 압축률에서 복구영상의 화질을 개선 할 수 있는 KLT-CVQ(classified vector quantization)를 제안하였다. 제안된 방법에서는 PCNN이 가지는 완벽한 무상관성(decorrelation), 에너지 집중 현상(energy compaction)과 VQ가 가지는 확률 밀도 함수의 형상(probability density function shape)특성과 공간 영역 채우기(space-filling) 잇점을 최대한으로 이용함으로써 높은 압축률에서 양호한 복구화질을 얻을 수 있는 영상 부호화가 되도록 하였다. 특히 질감 에너지 척도를 사용하여 입력 영상블럭을 통계적인 성질에 따라 유형 분류하고 각 유형별로 독립적인 KLT를 수행하는 PCNN를 사용하여 복구 영상의 에지 열화를 줄이도록 하였다. 컴퓨터 시뮬레이션 결과 제안된 KLT-CVQ에서는 기존의 KLT-CSQ (classified scalar quantization)나 DCT(discrete cosine transform)-CVQ에 비해 주어진 비트율에서 복구 영상의 화질이 개선됨을 확인하였다.

Abstract

This paper proposes a KLT-CVQ scheme using PCNN to improve the quality of the reconstructed images at a given bit rate. By using the PCNN and classified vector quantization, we exploit the high energy compaction and complete decorrelation capabilities of the KLT, and the pdf(probability density function) shape and space-filling advantages of the VQ to improve the performance of the proposed hybrid coding technique. In order to preserve the perceptual features such as the edge components in the reconstructed images, we classified the input image blocks according to the texture energy measures of the local statistics and vector-coded them adaptively, and thereby reduces the possible edge degradation in the reconstructed images. The results of the computer simulations show that the performance of the proposed KLT-CVQ is higher than that of the KLT-CSQ or the DCT-CVQ in the quality of the reconstructed images at a given bit rate.

* 正會員, 龜尾專門大學 電子科

(Department of Electronics, Kumi College)

** 正會員, 慶北大學校 電子電氣工學部

(The School of Electronics and Electrical Engineering, Kyungpook National Univ.)

接受日字:1996年9月11日, 수정완료일:1996年10月15日

I. 서 론

최근 복구 영상의 적정 화질을 유지하면서 압축율을 높여 영상의 저장과 전송 효율을 높이기 위한 영상 부호화 기법에 대하여 많은 연구가 진행되고 있다^[1-4]. 이들 연구 중에는 특히 DCT-VQ나 KLT-SQ와 같이 변환 부호화와 양자화 기법 각각의 장점을 적절히 활

용하고자 하는 혼합 부호화 방법에 대한 연구가 활발하다. 변환 부호화에 사용되는 변환함수에는 여러 종류가 있으나, 변환계수들의 에너지 집중성과 무상관성 등 그 성능에는 각각 차이가 있다. KLT는 변환계수의 에너지 집중성과 무상관성 측면에서 모든 선형 변환함수 중에서 최고의 성능을 가지지만, 종류가 다른 입력 데이터에 따라 변환을 위한 기저 행렬(basis matrix)이 변화하는 신호 의존적(signal-dependent) 특성을 나타낸다^[5]. 또한 변환행렬은 해석적으로 잘 알려져 있지 않으며, 알려진 경우라도 변환을 위해 수행해야 할 연산량이 방대하기 때문에 KLT는 주로 속도보다는 정확한 복구 영상이 요구되는 의료 영상 부호화에 사용된다. 반면에 DCT는 주로 영상통신에서와 같이 적정 화질 범위내에서 속도가 요구되는 경우에 사용된다. DCT는 비록 그 변환 계수들의 에너지 집중성, 무상관성 면에서는 KLT보다 열세인 준최적방법이지만 고속의 변환기법이 알려져 있고 인접 화소간의 상관도가 1에 가까운 경우에는 KLT 성능에 근접할 수 있으므로 JPEG(joint photographic experts group)이나 MPEG(moving pictures experts group)의 표준방식으로 권고되고 있다.

한편 벡터 양자기는 같은 크기의 부호책(codebook)에 대해 N차공간에 부호벡터(codevector)를 적절하게 위치시킴으로써 스칼라 양자기에 비해 공간 채우기, 확률 밀도 함수의 형상 잇점을 가지고 있다. 이는 주어진 왜곡 측도(distortion measure)가 최소가 되도록 부호 벡터를 위치시킬 수 있는 자유도가 주어짐을 의미하며, 확률 밀도가 높은 영역에서는 부호 벡터를 밀집시키고, 확률밀도가 낮은 영역에서는 부호 벡터를 드물게 위치시킴으로써 측정 왜곡의 최소화가 가능하다. 또한 벡터 양자기는 소수 비트 할당(fractional bit assignment)이 가능하여 정수를 사용하는 스칼라 양자기에 비해 효율적인 설계가 가능하다.

Kim 등^[1]은 에지열화(edge degradation)를 줄이기 위해 영상 블럭을 분류하여 분류된 블럭을 평균 벡터의 기하학적인 분산값에 따라 부블럭으로 나누어 비트할당 알고리즘을 적용한 DCT-CVQ를 제안하여 주어진 비트율에서 만족할 만한 성능을 얻었다. 한편 Kung 등^[2]은 KLT의 변환계수인 주성분(principal components)을 동시 추출 할 수 있는 병렬처리 알고리즘으로 APEX(adaptive principal component extraction)를 제안하고 영상 부호화를 수행한 결과 변

환계수를 DCT계수나 KLT계수의 어느 한가지만 사용하여 영상 부호화를 하는 것보다 이들을 조합하여 사용하는 것이 효과적임을 보였다. 또한 Bannour 등^[3]은 KLT를 PCNN으로 구현하기 위해 RLS(recursive least square)학습법을 이용하여 수렴이 보장되고 속도와 변환 계수의 정확도 측면에서 우수한 알고리즘을 개발하여 영상 압축에 적용한 결과, 대부분의 경우 전체 변환 계수들이 가지는 에너지의 99%가 처음 1/4의 변환 계수에 집중분포됨을 확인하였고, 이들 변환 계수만을 이용하여 영상 부호화를 수행하여도 우수한 복구 성능을 가짐을 보였다. 또한 Dony 등^[4]은 KLT는 전역적으로는 최적이지만 국부적으로는 준최적이고, 에지부근에서의 통계적인 특성은 영상전체의 통계적인 특성과 상이하기 때문에 KLT를 이용한 영상 부호화의 경우 에지부근에서 성능이 저하한다는 점에 착안하여 영상의 국부적인 변화에 적응할 수 있도록 입력 블럭을 유형분류하여 영상 부호화하는 OIAL(optimally integrated adaptive learning)를 제안하였다. 이는 전체적인 구조로 보아 KLT-CSQ에 해당된다고 볼 수 있다. 이 방법에서는 시스템구조에 비해 동일한 비트율에서 블럭의 국부특성을 고려하지 않고 부호화한 경우 보다는 약 1dB 개선되었으나, 비교적 낮은 복구성능을 얻었다. 따라서 KLT의 대량의 복잡한 계산으로 인한 느린 속도와 입력블럭의 통계적 특성에서 생길 수 있는 에지열화 등 두가지 문제에 대한 대책을 마련하면 KLT와 VQ와의 혼합부호화는 낮은 비트율에서 양호한 복구 영상 화질을 얻기 위해 시도할만 하다고 보겠다.

본 논문에서는 PCNN을 사용하여 KLT를 고속화하고 질감 에너지 척도에 의한 영상블럭의 통계적 특성 분류에 따른 분류 벡터 양자기를 사용함으로써 주어진 압축률에서 복구 영상의 화질을 개선할 수 있는 효과적인 영상 부호화 방법을 제안하였다. 제안된 방법에서는 KLT가 가지는 완벽한 무상관성, 에너지 집중 현상과 VQ가 가지는 확률 밀도 함수의 형상특성과 공간영역 채우기 잇점을 최대한으로 이용함으로써 높은 압축률에서 양호한 복구화질을 얻을 수 있는 영상 부호화가 되도록 하였다. 특히 KLT의 신호의존적인 특성을 역이용하여 통계적 특성이 유사한 블럭별로 영상블럭을 질감 에너지 척도에 따라 유형 분류하고, 각 유형별로 유형의 성질을 반영하는 독립적인 KLT를 수행하는 PCNN를 사용함으로써 복구 영상의 에지열화를 줄이

도록 하였다. 변환 계수 중 99%의 에너지가 분포되어 있는 주요 성분계수들에 대해서 KLT-CVQ와 KLT-CSQ를 수행하고 각각의 성능을 같은 비트율에서 DCT-CVQ 및 기존의 KLT-CSQ에 해당하는 OIAL과 성능을 비교하였다.

II. 혼합 영상 부호화에 의한 KLT-CVQ의 제안

1. PCNN을 이용한 KLT구현과 VQ

영상 부호화 방법중 변환 부호화법에서는 영상을 구성하는 화소 상호간에 존재하는 중복성 정보를 제거하기 위해 고차의 입력영상 공간으로부터 저차의 표상공간으로의 사상(mapping)을 통해 영상 화소간의 중요한 특징 값만을 추출하여 영상 부호화한다. KLT를 이용한 변환 부호화법에서는 먼저 주어진 입력영상을 $n \times n$ 크기의 블록으로 서로 중첩되지 않게 나눈다. 각 블록은 $N(=n \times n)$ 차의 벡터 \mathbf{x} 로 나타낼 수 있는데 \mathbf{x} 의 모든 블록에 대한 평균벡터를 \mathbf{m}_x 이라 하면 $M \times N$ ($M \leq N$)차의 행렬인 기저벡터(basis vectors) \mathbf{W} 를 이용하여 선형변환이 이루어진다. 따라서 각 영상블록에 대한 M 차의 계수벡터 \mathbf{y} 는

$$\mathbf{y} = \mathbf{W}(\mathbf{x} - \mathbf{m}_x) \quad (1)$$

로 표시된다. 변환행렬 \mathbf{W} 는 샘플 자기공분산(auto-covariance) 행렬 $\Sigma = E[\mathbf{x} \mathbf{x}^T]$ 에서 M 개의 가장 큰 고유치에 대응하는 M 행의 고유벡터(eigenvector)로 구성된다. 여기서 \mathbf{x}^T 는 \mathbf{x} 의 전치행렬(transpose matrix)을 나타낸다. 원래의 블록으로 복구된 블록을 $\hat{\mathbf{x}}$ 로 나타내면

$$\hat{\mathbf{x}} = \mathbf{W}^T \mathbf{y} + \mathbf{m}_x \quad (2)$$

로 주어지는데, 여기서 \mathbf{W}^T 는 \mathbf{W} 의 전치행렬을 나타낸다. 선형변환의 성능은 평균제곱오차 $\|\mathbf{x} - \hat{\mathbf{x}}\|^2$ 에 의해 변환의 효율성이 좌우되는데 KL변환이 최적으로 알려져 있다^[5]. 또한 벡터 \mathbf{y} 의 계수들은 계수의 분산 값이 큰 주성분 순서대로 배열되는데, 대부분의 경우 변환 계수 에너지의 99%가 처음 25%의 주성분에 집중분포된다. 따라서 이들 25%의 주성분만을 이용하여 영상 부호화를 하여도 복구 영상의 화질은 원래의 영

상과 거의 구분이 되지않을 정도의 화질을 얻을 수 있다^[3].

KLT를 수행하기 위해서는 공분산 추정을 위한 계산과 고유치 및 고유벡터의 계산 등 많은량의 계산이 요구되는데, 입력 데이터 세트로부터 곧바로 정상 벡터 확률 과정(stationary vector stochastic process)의 주성분을 추출할 수 있는 PCNN을 이용하면 이 문제는 쉽게 해결된다. 영상블록의 KLT를 위해 RLS 학습 규칙을 사용하여 첫번째 주성분을 추출해내는 PCNN의 구조를 그림 1에 나타내었다. 그림에서 신경망의 입력 $\mathbf{x}(n)$ 은 영상블록의 각 화소값을 나타내며 입력층의 가중치 $\mathbf{W}_1(n-1)$ 가 변환행렬의 첫번째 행에 대응하는 고유벡터가 되며, 출력층의 가중치 $\mathbf{W}_1^T(n-1)$ 는 변환행렬의 전치행렬인 역변환(복구)행렬의 고유벡터를 나타낸다. 따라서 입력 공분산 행렬의 p 개의 가장 중요한 고유치에 대응하는 처음 $p < N$ 개의 고유벡터에 순차적으로 수렴하는 p 개의 가중치 벡터를 구하면 그 벡터가 바로 KLT를 위한 변환행렬이 된다.

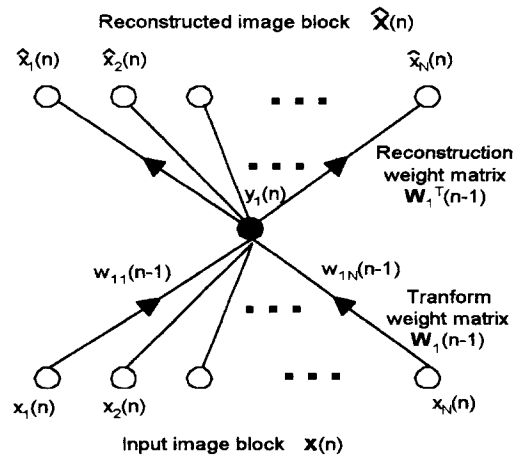


그림 1. 첫번째 주성분 추출을 위한 PCNN
Fig. 1. PCNN for the extraction of the first principal component.

PCNN에 의해 추출된 주성분을 벡터 양자화하면 혼합영상 부호화의 성능을 극대화할 수가 있다. 즉, 벡터 양자기는 스칼라 양자기에 비해 벡터 성분간의 확률 밀도 함수의 형상에 따라 생기는 잊짐과 벡터 차수에 따른 벡터 성분의 최적 공간배치에 의한 공간채우기 잊짐이 존재하는데, polytopes이 공간을 채울수 있는 효율을 의미하며 입력원의 확률분포와는 무관하다. 이

공간채우기 잇점은 벡터 양자기에 항상 존재하는 장점으로써 스칼라 양자기보다 우수한 성능을 내는 요인이 된다^[6].

2. 유형분류기

제안된 KLT-CVQ에서 KLT의 복잡하고 방대한 연산을 PCNN으로 고속병렬화하고 변환 계수의 에너지집중 현상을 이용하여 주요 변환 계수만을 벡터 양자화하면 공간채우기 잇점과 확률밀도함수의 형상에 의한 잇점을 최대한으로 활용하게 되어 낮은 비트율에서 양호한 복구화질을 얻을 수 있는 효율적인 혼합 영상 부호화가 가능하다. 따라서 전체성능을 좌우하는 것은 분류양자화에서 최적의 분류를 하여 유형별로 최적의 기저행렬을 사용하는 것이 관건이 된다. 주어진 블록의 통계적 특성, 즉 블록내의 화소간의 방향성과 레벨의 크기와 관련된 특징 추출에 있어서 KLT를 사용할 경우 식(1)로 주어지는 변환 계수벡터 \mathbf{y} 는 최소의 평균 제곱오차를 출력하므로 표현위주 특징(most expressive features: MEFs)이라고 알려져 있다^[11]. 이는 선형변환의 의미에 있어 블록 집단을 가장 잘 표현하기 때문이다. 하지만 MEFs는 블록 표현을 하는데는 적합하지만 블록이 어떤 유형을 나타내는 지를 판단하기 위한 특징으로는 부적합하다. 왜냐하면 MEFs 즉, 변환 계수들 중 첫번째값은 블록내에 변화가 제일 심한 정도 예를들면 블록의 평균밝기를 나타내는 특징이며 나머지 계수들은 블록내의 변화가 심한 정도를 순서대로 나타내는 값이므로 이들로부터는 어떻게 유형을 구분할 것인지에 대한 판단이 어렵기 때문이다. 따라서 MEFs에 적절한 투영법을 사용하여 출력된 특징을 판별 위주 특징(most discriminating features: MDFs)이라 부르는데 이는 학습세트에 나타난 유형을 최적으로 구별(discrimination)할 수 있게 하는 특징이기 때문이다.

\mathbf{V} 를 MDF영역으로 입력 블록 벡터를 투영하는 투영행렬이라고 하면, 각 유형의 평균벡터가 \mathbf{m}_i , $i=1, 2, \dots, c$ 인 c 개의 유형을 가진 샘플로 부터의 새로운 특징 벡터는

$$\mathbf{z} = \mathbf{V}^T \mathbf{y} \quad (3)$$

로 표시할 수 있다. 그러면 유형내의 분산행렬(within-class scatter matrix)은 유형 i 의 샘플수가 n_i 일때

$$\mathbf{S}_{wi} = \sum_{j=1}^{n_i} (\mathbf{y}_j - \mathbf{m}_i)(\mathbf{y}_j - \mathbf{m}_i), \quad \mathbf{S}_w = \sum_{i=1}^c \mathbf{S}_{wi}, \quad (4)$$

로 정의된다. 모든 유형으로 부터의 샘플들에 대한 전체 평균벡터를 \mathbf{m} 이라 하면, 유형간의 분산행렬(between-class scatter matrix)은

$$\mathbf{S}_b = \sum_{i=1}^c (\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m}) \quad (5)$$

로 정의된다. 판별 분석법(discriminant analysis)에 의하면 $\frac{\det[\mathbf{S}_b]}{\det[\mathbf{S}_w]}$ 의 비를 최대화하는 투영행렬 \mathbf{V} 을 결정하면 동일 유형내의 분산은 최소화하는 반면 유형간의 분산은 최대화함으로써 효과적인 유형분류가 가능하다. [12]에 의하면 이 비는 투영행렬 \mathbf{V} 의 열(column) 벡터가 가장 큰 고유치와 관계된 $\mathbf{S}_w^{-1} \mathbf{S}_b$ 의 고유벡터일 때 최대가 된다. 또한 \mathbf{z} 의 스칼라 성분들이 주어진 샘플의 특징값이 되고 \mathbf{V} 의 열(column) 벡터가 MDF 특징벡터가 된다. 하지만 MDF를 유형 분류에 적용할 경우 사전에 어떤 블록이 어느 유형에 속하는 지를 알고 있어야 \mathbf{S}_w , \mathbf{S}_b 를 구할 수 있으나, 본 논문의 경우에는 입력블록에 대한 사전정보가 없으므로 이 방법을 적용할 수가 없다. 한편 블록크기가 4x4인 경우는 DCT계수의 수직 및 수평방향의 크기와 부호를 고려하면 블록의 통계적 특성으로 인해 유형분류를 위한 효과적인 특징 추출이 가능하여 DCT가 주로 사용되었으나 영상블록의 크기가 큰 8x8 혹은 16x16의 경우에 대해서는 고려해야 할 방향성과 크기가 다양하여 이에 맞는 효과적인 다른 특징 추출방법이 요구된다.

Laws^[7]는 블록내의 통계적 특성 추출이 가능한 제안한 각각의 마스크를 사용하여 입력블록에 상승연산(convolution)을 수행하여 얻은 절대값을 더하여 특징을 추출하는 질감 에너지 척도를 사용하여 질감영상을 분류한 결과, 각 마스크는 우수한 분류성능을 나타냄을 보였다. 즉, 중심에 가중치(center-weighted)를 두어 국부적인 평균을 취하는(local averaging) L3 ≡ (1, 2, 1), 에지검출을 위해 대칭적 1차 미분을 출력하는(symmetric first differencing) E3 ≡ (-1, 0, 1) 및 블록내의 점을 찾기 위해 2차 미분(second differencing)인 S3 ≡ (-1, 2, -1) 등의 기본적인 벡터로부터 이들 벡터를 자기자신 혹은 다른것과 상승연산을 하면 국부 평균값(local average)을 출력하는 L5 = (1, 4,

6, 4, 1), 점과 에지를 나타내는 $S5 = (-1, 0, 2, 0, -1)$, $E5 = (-1, -2, 0, 2, 1)$ 블럭내의 리플 및 웨이브를 검지 할 수 있는 $R5 = (1, -4, 6, -4, 1)$, $W5 = (-1, 2, 0, 2, 1)$ 로 주어지는 5개의 벡터를 구할 수 있다. 또한 길이 3 혹은 길이 5의 열(column) 벡터를 동일한 길이의 행(row) 벡터로 곱하면 3×3 혹은 5×5 크기의 마스크를 구할 수 있다. 본 논문에서는 입력블럭의 유형분류를 위해 블럭내의 통계적 특징인 점, 선, 리플 및 웨이브등의 특징 추출이 가능한 Laws의 3×3 및 5×5 마스크를 이용한 질감 에너지 척도에 의해 입력블럭의 특징을 추출하였다.

3. KLT CVQ의 제안

본 논문에서 제안한 KLT-CVQ의 블록도는 그림 2와 같으며, 전체적인 시스템동작은 다음과 같다. 먼저 $n \times n$ 크기의 입력 블럭들을 에지 정보를 포함한 국부적인 통계특성에 따라 유형 분류하기 위해 질감 에너지 척도에 의해 블럭의 특징을 추출하였다. 제안된 시스템은 이 특징을 SOFM(self-organization feature maps)에 입력시켜 입력 블럭을 유형 분류하였다. 또한 입력 블럭은 그 국부적인 특성에 따른 주성분을 구하기 위해서 PCNN으로 구성되는 독립적인 KLT에 의해 변환되도록 하였으며, 이때 각각의 변환행렬은 입력블럭의 통계적인 특성을 반영하는 기저벡터로 구성된다. KLT의 에너지 집중현상을 고려하여 전체 주성분 중에서 처음 1/4에 해당하는 주성분들만을 벡터 양자화하였다. 또한 이들 주성분들 중에서 첫번째 주성분은 인접 블럭의 상관도를 고려하여 DPCM(differential pulse code modulation)으로 부호화하고 나머지 주성분은 벡터 양자화하여 입력 블록 전체에 대한 부호화를 하였다. 영상의 복구를 위해서 부호기는 각 블럭의 유형에 대한 색인, 첫 번째 주성분 색인 및 나머지 주성분에 대한 벡터 양자화 색인을 복호기에 전송하여 이들 값으로부터 원래의 입력 영상을 복구하도록 하였다.

제안된 방법에서는 먼저 해당 입력영상을 8×8 화소를 가지는 중첩되지 않은 블럭들로 나눈다. 복구 영상의 에지들을 잘 보존하기 위해서는 이들 국부적인 특성을 잘 나타내는 블럭의 통계적 특성이 유사한 블럭별로 분류하여 부호화를 수행하는 것이 효율적이므로, 유형분류 벡터 양자화를 하였다. 영상분류를 위해서는 Laws가 질감 영상 분류에 사용한 마스크중에서 분류 성능이 우수한 3개의 3×3 마스크와 2개의 5×5 마스크

를 이용하여 특징을 추출하였다. 또한 블럭내의 에지 방향을 찾기위해 각 마스크별로 그를 $0^\circ, 90^\circ, 180^\circ, 270^\circ$ 회전시킨 마스크를 포함하여 4개씩의 마스크를 사용하였다.

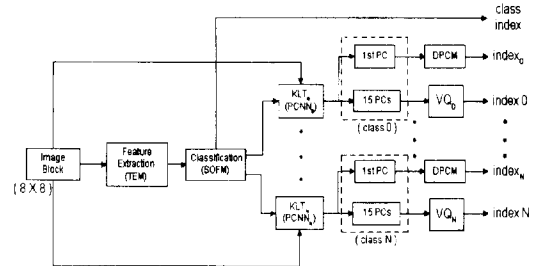


그림 2. 제안된 KLT-CVQ의 블록도

Fig. 2. Block diagram of the proposed KLT-CVQ.

질감 에너지 척도에 의해 추출된 특징을 SOFM에 입력시켜 각 블럭의 유형을 분류하였다. KLT는 입력 신호에 의존하는 변환행렬을 출력하므로 국부적인 통계 특성의 변화에 잘 적응하도록 각 블럭의 통계적인 특성이 유사한 블럭에 해당된 변환행렬을 사용함으로써 복구 영상의 화질을 향상시키도록 하였다.

한편, VQ를 위한 부호책의 설계는 일종의 군집(clustering) 문제로서 이제까지 LBG(Linde Buzo Gray) 알고리즘을 주로 이용되어 왔으나 최근에는 SOFM도 많이 이용하고 있다. 이는 SOFM이 초기 가중치에 무관하게 학습이 진행됨에 따라 입력벡터의 분포를 적절히 반영하는 가중치를 가지게 되어 LBG를 사용하는 경우에 비해 복구 영상의 화질이 우수할 뿐만 아니라 승리 뉴런의 가중치와 이웃 뉴런의 가중치를 동시에 변화시키므로 가중치간에는 상관성이 강하게 되어 코드벡터의 예측에 의한 빠른 부호책 탐색이 가능하기 때문이다. 제안된 KLT CVQ방법에서는 SOFM을 적용하여 영상블럭의 통계적 특성에 따라 분류된 각각의 유형에 대해 독립적인 부호책을 설계하였다. 또한 8×8 크기의 블럭에 대해 PCNN에 의해 변환된 64개 계수의 처음 1/4에 해당하는 16개의 주성분 중 변화값의 변화가 제일 심한 첫번째 주성분은 인접 블럭과의 상관도를 고려하여 DPCM (differential pulse code modulation)을 사용하여 변환 부호화에서 발생할 수 있는 복구 영상의 구획(contouring) 효과를 줄이도록 하였고, 나머지 15개 주성분들은 벡터 양자화

를 수행하여 주어진 비트율에 해당하는 부호책을 생성시켰다.

Ⅲ. 시뮬레이션 결과 및 고찰

제안된 영상 부호화기법인 KLT-CVQ와 KLT-CSQ에 대한 성능을 평가하기 위해 컴퓨터 시뮬레이션을 수행하였다. 학습에는 256 명암도의 512x512 크기의 영상 12개와 256x256 크기의 영상 8개를 사용하였다. 일반적으로 8x8크기보다 작은 크기의 블록을 KL 변환할 경우 충분한 공간상의 상관도를 포착하지 못해 주성분의 정확한 추정치에 곤란하므로 실험에서는 입력 영상을 8x8크기의 블록으로 나누어 변환 부호화하였다. 또한 학습후 시험 영상으로는 학습에 사용되지 않았던 256 명암도의 512x512 크기의 "Lena" 와 "Bridge" 영상을 사용하였다. 먼저 학습에 사용된 전체 블록의 각각에 대해 질감 에너지 척도를 이용하여 특징을 추출한 후, 이를 SOFM에 입력시켜 16개 및 32개의 유형으로 분류하였다.

준비된 각 유형의 학습 데이터로부터 입력 뉴런이 64개이고 은닉 뉴런이 16개인 PCNN을 학습시키기 위해 초기 가중치를 무작위로 초기화하였다. 또한 RLS 학습 알고리즘을 이용하여 개개의 뉴런에 대해 순차적인 학습을 통해 가중치를 구하였다. 구해진 가중치 즉 변환행렬을 이용하여 전체 변환 계수의 1/4인 16개의 주성분만을 추출하였다.

먼저 제안된 KLT-CVQ의 성능을 알아보기 위해 추출된 주성분중에서 첫번째 주성분은 DPCM을 사용하여 양자화하였으며 Laplacian 확률 밀도 함수에 최적화된 Max 양자화기를 사용하였고, 나머지 15개 주성분은 벡터 양자화하여 부호책을 만들었다. 또한 KLT-CSQ의 성능과도 비교하기 위해 상기 PCNN에서 출력된 16개의 주성분에 대해 Laplacian 분포를 가지는 각 주성분에 대해 최적 비트할당 알고리즘과 Max 양자화기를 이용하여 스칼라 양자화하였다^[9].

복구화질의 척도로는 PSNR(peak signal-to-noise ratio) 즉,

$$PSNR = 10 \log_{10} \left[\frac{G^2}{\frac{1}{P} \sum_{i=0}^{P-1} (x_i - \hat{x}_i)^2} \right] \text{ dB} \quad (6)$$

를 사용하였다. 여기서 G는 영상의 최대 명암도 값이고 x_i 는 원래의 화소값을 나타내며 \hat{x}_i 는 복구된 화

소값을 나타낸다. 또한 P는 입력영상의 전체 화소수를 나타낸다.

유형분류 벡터 양자기에서는 에지 등을 고려한 유형 분류에 사용되는 분류기의 특성과 분류된 유형의 수가 주어진 비트율에서 복구화질의 성능에 영향을 주는데, 사용된 분류기의 특성에 따라 최적의 유형수가 결정된다. 질감 에너지 척도를 유형분류에 사용하는 경우 블록내의 통계적 특성에 의해 유형이 분류되기 때문에 분류된 유형은 시각적으로는 유형 상호간의 구분이 어려우며, 입력블럭의 유형은 사용된 마스크의 일반적인 형태 즉, 점, 선, 에지, 리플 등에 영향을 받아 유형이 결정된다. 입력블럭을 16개 및 32개의 유형으로 구분한 경우 제안된 KLT-CVQ와 KLT-CSQ에 대한 각각의 PSNR을 표 1에 나타내었다. II절에서 언급한 바와같이 KLT-CVQ는 KLT의 우수한 에너지 집중 현상, 완벽한 무상관성을 가짐과 동시에 VQ의 공간채우기 및 확률 밀도 함수의 형상 잇점을 지니고 있지만, KLT-CSQ는 스칼라양자화를 사용하기 때문에 공간채우기 및 확률 밀도 함수의 형상 잇점을 가지지 못하므로 0.375 bpp에서 16개 및 32개의 유형으로 분류할 경우, KLT-CVQ가 KLT-CSQ에 비해 "Lena" 영상인 경우 각각 약 0.4dB 및 1.2dB 성능향상을 보였으며, "Bridge" 영상에 대해서는 각각 약 0.6dB 및 약 1dB 개선되었다. 따라서 유형수가 증가함에 따라 영상의 통계적 특성이 보다 세밀하게 고려되므로 복구성능이 개선됨을 알 수 있으며, 이로써 영상블럭의 통계적 특성을 고려하여 유형분류를 한 다음 부호화를 수행하는 것이 효율적임을 알 수 있다. 질감 에너지 척도를 유형분류에 사용한 경우 주어진 비트율에서 최상의 복구화질을 얻을 수 있는 최적의 유형수에 관한 연구는 더 진행되어야 할 것이다.

또한 질감 에너지 척도에 따라 32개로 유형분류한 제안된 KLT-CVQ 및 KLT-CSQ의 성능과 DCT계수에 따라 13개로 유형분류한 DCT-CVQ^[11] 및 투영 벡터의 노름(norm)에 따라 512개로 유형분류한 OIAL (optimally integrated adaptive learning)^[14] 등 기존 연구들의 성능을 표 2에 비교하였다. DCT-CVQ^[11]에서 사용된 유형분류기는 에지열화를 줄이기 위해 영상블럭에 DCT를 수행하여 그 변환 계수 중 에지의 크기와 방향에 관계된 수평 및 수직방향에 대한 9개의 DCT계수를 사용하여 입력 영상을 시각적으로 구분이 가능한 형태인 평탄(shade)블럭, 수평블럭, 수직블럭,

대각선블럭 등을 포함하여 이 분류기에 의해 만족할 만한 성능을 얻을 수 있도록 13개의 유형으로 분류한 후 벡터 양자화하였다. 또한 OIAL^[4]에서는 입력블럭의 국부적인 변화에 잘 적응하도록 유형분류하고, 변환된 계수의 일부를 스칼라양자화한 KLT-CSQ의 일종으로 볼 수 있는데, 유형분류기는 선형 변환을 가능케하는 유형수 만큼의 독립적인 모듈들과 이들 모듈에서의 출력 중 한 개를 선택하는 분류기로 구성된다. 입력블럭의 유형을 결정하기 위해서 각 모듈은 선형변환이 가능한 기저영상(basis image)들과 입력블럭을 내적시킨 각각의 변환 계수벡터와 이를 역변환시킨 투영벡터의 노름을 출력하는데, 제일 큰 노름을 출력하는 모듈의 색인으로 해당유형을 결정한다. 여기서도 변환된 전체계수 중 적절한 수의 변환 계수만으로 스칼라양자화한 경우 유형의 수가 증가할수록 주어진 비트율에서 복구화질이 개선됨을 보였다. 표에서 보는 바와같이 "Lena" 영상에 대해 0.37 bpp에서 본 논문에서 제안한 KLT-CVQ는 KLT-CSQ보다 약 1.2dB 성능이 우수할 뿐만 아니라 DCT-CVQ^[1]에 비해 약 2.8dB, OIAL^[4]에 비해 약 5.4dB 개선되었다. 또한 본 논문에서 질감 에너지 척도에 따라 32개로 유형분류한 KLT-CSQ는 기존의 KLT-CSQ로 볼 수 있는 OIAL^[4]에 비해 약 4.1dB 개선을 보이는데, 이는 본 논문에서 유형분류에 사용한 질감 에너지 척도가 OIAL에서의 분류척도인 투영벡터의 노름보다 분류 성능이 우수한데 기인하는 것으로 보인다. 표 2의 오른쪽 하단의 "*"는 참고문헌상에는 해당 실험데이터가 없어서 인용비교할 수 없음을 나타낸다. 그림 3에는 시험에 사용된 "Lena" 영상에 대해 질감 에너지 척도에 의한 유형분류기를 사용하여 32개의 유형으로 분류 벡터 양자화했을 때 원래영상, 복구된 영상 및 차이영상을 나타내었다. 그림에서 차이영상은 시각적으로 확인될 수 있도록 3배 강조하여 나타낸 것이다.

IV. 결론

본 논문에서는 PCNN를 이용하여 KLT를 고속화하고 영상을 분류 양자화함으로써 주어진 압축률에서 복구화질을 개선할 수 있는 KLT-CVQ를 제안하였다. 제안된 방법에서는 복구 영상의 에지를 잘 보존하기 위해서 질감 에너지 척도를 사용하여 블럭의 통계적인 특성에 따라 분류 벡터 양자화하였다.

표 1. 0.375bpp에서 유형수에 따른 제안된 KLT-CVQ와 KLT-CSQ의 PSNR [dB] 비교

Table 1. PSNR [dB] comparison of the proposed KLT-CVQ and KLT-CSQ for different number of classes at 0.375bpp.

Method Test image	16 - Class		32 - Class	
	KLT-CVQ	KLT-CSQ	KLT-CVQ	KLT-CSQ
Lena	33.02	32.61	35.06	33.83
Bridge	26.86	26.21	27.66	26.61

표 2. 512x512, 8비트 시험영상에 대한 PSNR [dB] 비교

Table 2. PSNR [dB] comparison for the 512x512, 8-bit test images.

() : bpp

Method Test image	KLT-CVQ	KLT-CSQ	DCT-CVQ[1]	OIAL[4]
	32 classes	32 classes	13 classes	512 classes
Lena	35.06 (0.375)	33.83 (0.375)	32.21 (0.373)	29.67 (0.375)
Bridge	27.66 (0.375)	26.61 (0.375)	25.94 (0.518)	*



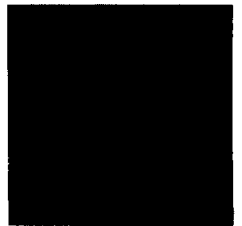
(a)



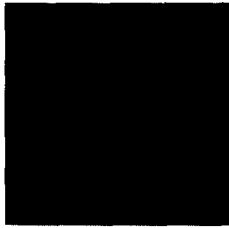
(b)



(c)



(d)



(e)

그림 3. 0.375bpp에서 32 유형으로 분류한 512x512, 8-bit "Lena"에 대한 제안된 KLT-CVQ와 KLT-CSQ의 성능비교 (a) 원래 영상, (b) KLT-CVQ에 의한 복구 영상, (c) KLT-CSQ에 의한 복구 영상, (d) KLT-CVQ때 3배 강조된 차이영상, (e) KLT-CSQ때 3배 강조된 차이영상

Fig. 3. Comparison of the results of the proposed KLT-CVQ and KLT-CSQ at 0.375bpp classified into 32 classes by texture energy measure for "Lena" 512x512, 8-bit. (a) original image, (b) image reconstructed by KLT-CVQ, (c) image reconstructed by KLT-CSQ, (d) intensified difference image (3-fold) for KLT-CVQ, and (e) intensified difference image(3-fold) for KLT-CSQ.

따라서 KLT와 CVQ가 가지는 장점을 최대한으로 이용함으로써 높은 압축률에서 양호한 복구화질을 얻을 수 있도록 하였다. 실험결과, 질감 에너지 척도에 대한 분류 유형의 수가 증가할수록 복구화질이 개선되어 블럭의 통계적 특성의 적응도가 유형 수와 직접적으로 관계됨을 확인하였다. 또한 전체 주성분의 처음 1/4만 벡터 양자화해도 양호한 복구 화질을 얻는데 충분함을 확인하였다. 컴퓨터 시뮬레이션 결과, "Lena" 영상에 대해 0.375bpp에서 질감 에너지 척도에 의해 32개로 유형분류한 제안된 KLT-CVQ는 KLT-CSQ보다 약 1.2dB 성능이 우수할 뿐만 아니라 DCT계수에 의해 13개로 유형분류한 DCT-CVQ^[1]에 비해 약 2.8dB, 투영벡터의 노움에 의해 512개로 유형분류한 OIAL^[4]에 비해 약 5.4dB 개선되었다. 향후 제안된 KLT-CVQ에서 주어진 질감 에너지 척도를 사용하여 분류벡터 양자화 할 경우 주어진 비트율에서 최대의 복구성능을 얻을 수 있는 최적의 유형수를 구하는 문제가 더 연구되어야 할 것으로 사료된다.

참 고 문 헌

[1] J. W. Kim and S. U. Lee, "A transform

domain classified vector quantizer for image coding," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 2, no. 1, pp. 3-14, Mar. 1992.

[2] S. Y. Kung, K. I. Diamantaras, and J. S. Taur, "Adaptive principal component extraction (APEX) and applications," *IEEE Trans. on Signal Processing*, vol. 42, no. 5, pp. 1202-1217, May 1994.

[3] S. Bannour and M. R. Azimi-Sadjadi, "Principal component extraction using recursive least squares learning," *IEEE Trans. on Neural Network*, vol. 6, no. 2, pp. 457-469, Mar. 1995.

[4] R. D. Dony and S. Haykin, "Optimally adaptive transform coding," *IEEE Trans. on Image Processing*, vol. 4, no. 10, pp. 1358-1370, Oct. 1995.

[5] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice Hall, 1984.

[6] T. D. Lookabaugh and R. M. Gray, "High resolution quantization theory and the vector quantizer advantage," *IEEE Trans. on Information Theory*, vol. 35, no. 5, pp. 1020-1033, Sep. 1989.

[7] K. I. Laws, "Textured image segmentation measures," Image Processing Inst., of Univ. of Southern California, Rept. 940, Jan. 1980.

[8] M. Pietikainen, A. Rosenfeld, and L. S. Davis, "Experiments with texture classification using averages of local pattern matches," *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 13, no. 3, pp. 421-426, May/June 1983.

[9] M. Unser and M. Eden, "Multiresolution feature extraction and selection for texture segmentation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 717-728, July 1989.

[10] J. Makhoul, S. Roucos, and H. Gish, "Vector quantization in speech coding," *Proceedings of IEEE*, vol. 73, no. 11, pp. 1551-1588, Nov. 1988.

[11] D. L. Swets and J. J. Weng, "Efficient

image retrieval using a network with complex neurons," *Proceedings of IEEE International Conference on Neural Networks*, vol. 4, pp 1600-1605, 1995.

[12] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, New York: Wiley, 1973.

저 자 소 개



金泰容(正會員)

1957년 5월생. 1980년 2월 경북대학교 전자공학과 졸업(공학사). 1982년 2월 경북대학교 전자공학과 석사과정 졸업(공학석사). 1987년 ~ 1992년 산업과학기술연구소(RIST). 1994년 2월 경북대학교 전자공학과 박사과정

수료. 1992년 ~ 1996년 11월 현재 구미전문대학 전자과 교수. 주관심분야는 영상압축, 신경회로망 및 패턴인식 등임

崔興文(正會員) 第33卷 B編 第3號 參照

현재 경북대학교 전자공학과 교수