

DCT 맵과 이진 트리 구조 벡터 양자화기를 이용한 영상 부호화

조 성 환[†] 김 응 성^{**}

요 약

벡터 양자화기를 이용하여 영상의 부호화기를 설계하는데 있어서 2차원 이산여현변환(2D-Discrete Cosine Transform)에 근거한 DCT 맵(map)과 새로운 부호책(codebook) 설계 알고리즘을 제안한다. 영상을 작은 부분 블록으로 나누고 2차원 이산여현변환으로 대부분의 정보를 포함하는 부분, 즉 부호화하기 어려운 부분과 부호화하기 쉬운 적은 정보를 포함하는 영역으로 나누어 맵을 만들고 이 맵에 따라 영상의 중요한 특징들을 2차원 이산여현변환으로 추출한다. 부호책은 트리 구조에 근거한 2진 트리로 두 영역을 따로 학습세트로 나눔으로서 만들어 진다. 2진 트리의 중간 노드에서 각 학습 벡터는 그 노드에서의 문턱 값과 비교하여 두개의 아래 노드중 하나에 속하게 된다. 국제 표준 화상인 Lenna와 Boat 영상에 대하여 본 알고리즘으로 영상을 부호화했을 때 PNN과 CVQ 알고리즘에 비하여 수행 시간을 줄이고 PNN 알고리즘보다는 각각 약 0.45 dB과 0.33 dB만큼, CVQ 알고리즘보다 각각 약 0.05 dB과 0.1 dB만큼 더 좋은 영상의 화질을 얻을 수 있었다.

Image Coding Using DCT Map and Binary Tree-structured Vector Quantizer

Seong Hwan CHO[†] and Eung Sung KIM^{**}

ABSTRACT

A DCT map and new codebook design algorithm based on a two-dimension discrete cosine transform (2D-DCT) is presented for coder of image using vector quantizer. We divide the image into smaller subblocks, then, using 2D DCT, separate it into blocks which are hard to code but it bears most of the visual information and easy to code but little visual information, and DCT map is made. According to this map, the significant features of training image are extracted by using the 2D DCT. A codebook is generated by partitioning the training set into a binary tree based on tree-structure. Each training vector at a nonterminal node of the binary tree is directed to one of the two descendants by comparing a single feature associated with that node to a threshold. Compared with the pairwise nearest neighbor (PNN) and classified VQ (CVQ) algorithm, about 'Lenna' and 'Boat' image, the new algorithm results in a reduction in computation time and shows better picture quality with 0.45 dB and 0.33 dB differences as to PNN, 0.05 dB and 0.1 dB differences as to CVQ respectively.

1. 서 론

디지털 영상 데이터를 전송하거나 저장하기 위해서는 전송채널의 대역폭이 커야 되며 큰 용량의 메

모리가 필요하다. 그러므로 저장하거나 전송할 영상 데이터의 크기를 줄여 주기 위해 영상 압축(image compression)이라고 하는 처리 과정을 거치게 된다. 지금까지 연구된 정지 영상 신호의 데이터 압축 기법에는 예측 부호화(predictive coding), 변환 부호화(transform coding), 복합 부호화(hybrid coding), BTC(block truncation cod-

[†]정 회 원: 대우공업전문대학 전자계산기와 조교수

^{**}정 회 원: 성균관대학교 전자공학과 박사과정

논문접수: 1994년 3월 7일, 심사완료: 1994년 4월 18일

ing), 벡터 양자화(vector quantization) 등 1세대 부호화 방법과 계층적 부호화(hierarchical coding), anisotropic nonstationary 예측 부호화, contour-texture 부호화, directional decomposition 기반 부호화 등의 2세대 부호화 방법이 있다. 예측 부호화는 화소들간의 상관관계가 높은 점을 고려하여 현재 부호화될 화소 값과 이미 부호화된 인접 화소 값들로부터 구한 예측 값과의 차를 전송하는 방법으로 예측 오차의 분산이 원래 영상 신호의 분산에 비해 적어진다는 점을 이용한 것이고 [1]. 변환 부호화는 화소들간의 상관 관계를 낮추기 위해 원 영상을 적당한 변환 방법에 의해 변환하여 에너지의 집중현상을 이용한 것이다[2,3,4,5]. 복합 부호화는 예측 부호화와 변환 부호화를 합한 방법으로서 하드웨어의 구성을 용이하게 함과 동시에 변환 부호화의 장점 또한 가질 수 있는 방법이고[3], BTC는 영상을 중복되지 않는 블록으로 나누어 각 블록마다 통계적 특성에 따라 비트 맵(bit map)을 만들어 전송하는 방법이다[7]. 계층적 부호화는 한 영상을 계층적으로 부호화 하므로써 각 단계별로 다른 해상도의 영상을 복원할 수 있는 방법이며[8,9,10], anisotropic nonstationary 예측 부호화는 영상 복원에서의 추정문제를 부호화에 적용하여 복원 필터를 사용하여 부호화하는 방법이고[11], contour-texture 부호화는 영상을 중복성이 많은 texture 부분과 경계에 해당하는 contour 부분으로 나누어 각각을 따로 부호화하며 [12], directional decomposition 기반 부호화는 경계 요소만을 포함하는 directional 영상을 정의하여 고주파수 성분의 directional 영상 및 저역 통과 영상을 사용하여 부호화하는 방법이다[13]. 이외에도 variable velocity scanning, pseudo-random scanning, cluster and feature 부호화, edge 부호화, singular value decomposition 부호화 및 autoregressive 부호화 등이 있다 [14]. 본 논문에서는 이들 방법중에서 변환 부호화와 벡터 양자화를 결합한 방식을 이용하여 영상 부호화를 수행한다.

Shannon은 특정 통신율(rate)에 대한 평균 왜

곡(distortion)을 최소화한다는 의미에서 데이터 압축성능에 대해 원 샘플을 각각 하나보다는 블록이나 벡터로 부호화하므로서 이론적인 한계에 가깝게 접근할 수 있다는 것을 보였다[15]. 그러한 벡터 부호나 벡터 양자화는 LBG 알고리즘[16]과 같은 실제적인 벡터 양자화기 설계 알고리즘이 개발된 후 널리 사용되게 되어 특정 응용분야 뿐만 아니라 일반적인 벡터 양자화에 관한 몇몇 논문들이 발표되었다[17,18,19]. 벡터 양자화는 n 차원 벡터 X 의 원 집합(source set)으로부터 유한한 인덱스 집합(index set) I 로의 맵핑(부호화), 그리고 I 로부터 n' 차원 벡터 y 의 집합으로 재구성하는 맵핑(복호화)으로 정의할 수 있다. 벡터 양자화는 I 에서의 요소수가 m 일 때 입력 샘플당 최대 $\log_2 m/n$ 비트가 요구되는 인덱스만으로 원 샘플의 재구성이 가능하기 때문에 데이터 압축에 사용될 수 있는데 이는 일반적으로 X 내의 요소를 직접 나타내는데 필요한 비트 수보다 훨씬 적다.

벡터 양자화의 인덱스 집합 I 를 부호책(codebook)이라 부르는데, 일반적인 벡터 양자화기에서는 입력을 부호책의 일정수인 벡터집합에 속하게 하여 복호화기로는 그 집합의 인덱스만을 보내 복호화기에서는 부호화기와 같은 부호책을 이용하여 간단한 테이블 참조(table look-up)방법을 사용하여 원 벡터를 재생한다. 벡터 양자화의 관건은 벡터를 나타내는 좋은 부호책을 설계하는 것인데, 잘 알려진 부호책 설계는 Lloyd 방법[20]과 K-means clustering 방법[21]을 이용한 Linde, Buzo, Gray의 방법[16]이다. 이 알고리즘에서는 모든 학습 벡터가 부호벡터와의 최소 왜곡 법칙에 따라 클러스터를 이루게 된다. 이들 클러스터의 중심값(centroid)이 다음 반복에서 새로운 부호 벡터가 된다. 이와 같은 단계를 부호벡터와 클러스터 구성원사이의 전체 왜곡이 정해진 기준 이하가 될 때까지 계속한다. 그러나 LBG 알고리즘은 부호책 설계를 시작하기 위해서는 초기 부호책이 제공되어야 하고 그 단계가 반복적이며, 각 부호 벡터와 클러스터 구성원 사이에 왜곡(distortion)계산으로 인한 많은 수렴시간을 요구한다.

또한 Equitz에 의해 PNN(pairwise nearest neighbor)이라는 효과적인 알고리즘이 제안되었는데[22,23], 이 알고리즘은 초기 부호책이 필요 없으며 약간의 화질 저하를 가져오긴 하지만 LBG 알고리즘보다 수행시간이 훨씬 더 빠르다. 이 PNN 알고리즘의 기본 개념은 두 클러스터의 중심 값에 의해 두개의 가장 가까운 클러스터를 다른 클러스터로 연속적으로 합병해 나가는 것이다. 즉 전체 학습 집합을 가지고 시작하여 벡터 집합을 원하는 크기로 줄이거나 부호 왜곡에 도달할 때까지 단계를 반복한다. PNN 알고리즘의 계산상의 이득은 학습 벡터를 k차원 트리로 분할하는 전처리 과정에 의해 이뤄지는데 이때 일차원 벡터와 그의 평균 값이 분할시에 사용된다. 이는 각 반복 과정에서 검색과 클러스터의 합병 속도를 빠르게 한다. PNN 알고리즘과 LBG 알고리즘의 실행 시간을 비교하면 PNN 알고리즘의 실행시간은 LBG 알고리즘 실행 시간의 약 5% 정도이다[22].

벡터 양자화의 한 예로써 영상의 시각적인 특징(perceptual feature)을 유지하기 위해서 classified VQ(CVQ)가 Ramamurthi와 Gersho에 의해 제안되었다[24]. 이 CVQ에서는 윤곽선(edge)의 방향, 위치등에 의해서나 배경의 밝기 영역에 따라 각 블럭을 분류하고, 준 부호책(sub-codebook)이라고 하는 분리된 부호책이 각각의 분류에 대해 속하는 학습 벡터로부터 만들어 지는데, 이 각 부호책의 부호벡터는 LBG 알고리즘을 사용하여 할당한다. 처음 영상의 입력벡터가 부호기로 들어오면 분류기는 그 블럭을 각 특징 클래스중의 하나로 분류하여 그 클래스의 부호단어를 전송하고, 복호기는 전송된 부호단어로부터 전송된 부호책을 이용하여 블럭에 해당되는 벡터를 만들어 준다. 이 CVQ는 각 준 부호책에 반복적이고 초기 값이 요구되는 LBG 알고리즘을 사용하므로 알고리즘 수행 시간이 상당히 길다.

본 논문에서는 분류화 문제에 있어서 계산량을 줄이는, 즉 데이터의 차원(dimension)을 줄이는 특징의 추출방법과 분류기의 설계를 제안하는데 2차원 이산여현변환(2D-DCT)으로 영상을 부호화

하기 어려운 대부분의 정보를 포함하는 블럭과 부호화하기 쉬운 적은 정보를 포함하는 블럭으로 나누어 맵을 만들고 이 맵에 따라 영상의 중요한 특징들을 2차원 이산여현변환으로 추출한다. 두가지의 블럭에 대해 따로 부호책을 만드는데 부호책의 부호단어(codeword)를 만들기 위해 2진 트리 분류기를 사용한다. 2진 트리의 각 비단말 노드에서 학습 벡터 집합을 분할하기 위해 하나의 특징이 사용된다. 제안된 알고리즘도 PNN 알고리즘과 마찬가지로 초기 부호책이 필요 없다. 시뮬레이션 결과 새로 제안된 방법으로 영상 부호화를 하면 PNN이나 CVQ 알고리즘과 비교하여 계산 시간이 줄고 더 좋은 화질의 영상을 얻을 수 있었다.

2. DCT 맵 설계

영상을 작은 부분블럭으로 나누어 벡터 양자화와 같은 손실 있는 압축방법(lossy compression)을 적용할 때 다른 블럭들에 비해 부호화하기 어려운 블럭들이 있게 된다. 즉 압축비나 평균 제곱 오차(mean square error:MSE)의 관점에서 모든 블럭들에 대해 같은 크기의 부호책을 설계하는 것 보다는 부호화하기 쉬운 부분, 즉 윤곽선을 적게 포함하고 있는 배경 블럭은 적은 수의 부호책을, 윤곽선 즉 고주파 성분이 많은 블럭은 더 큰 수의 부호책을 설계하면 같은 평균 제곱 오차를 갖는 복원 영상에 대해 압축율을 높일 수 있다. 따라서 본 논문은 영상의 각 블럭에 대해 고저주파의 포함 여부를 결정하기 위하여 2차원 이산여현변환을 취함으로써 DCT 맵을 설계한다.

2.1 맵 설계

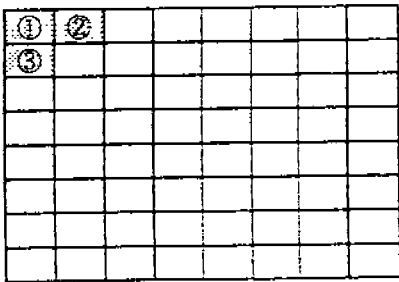
DCT 맵은 다음 세 단계로 만들어 진다.

[단계 1] 2차원 원 영상을 좌측상단으로부터 우측으로 8 픽셀, 하단으로 8 픽셀씩 8×8의 부분블럭(subblock)으로 나누어 2차원 이산여현변환을 계산한다.

계산 결과에 대해 (그림 1)과 같이 좌측 상단의 그 블럭의 DC 평균값인 ①의 값과 가장 낮은 주파수의 AC 값인 ②, ③의 값을 포함

한 3개 항의 값만 가지고 나머지 값은 0으로 만들어 2차원 이산변환행렬의 역변환을 취한다.

그 결과 값들과 원 영상의 블록값들과의 평균 제곱 오차를 계산하여 그 값이 문턱값(threshold) δ 보다 크면 이 블록에 대한 DCT 맵은 '1'로, 적으면 '0'으로 할당 값을 결정한다.



(그림 1) 8x8 블록에서의 DCT값 선택

(Fig. 1) Selection of DCT values in 8x8 subblock

[단계 2] DCT 맵에서 '1'로 된 각각의 8x8 블록에 대해 4개의 4x4 블록으로 나누어 단계 1과 마찬가지로 방법으로 4개의 4x4 블록에 대해 DCT 맵의 값을 결정한다. '0'으로 된 8x8 블록에 대해서는 4개의 4x4 블록에 대한 DCT 맵 값을 '0'으로 한다.

[단계 3] DCT 맵이 '1'로 정해진 4x4 블록과 '0'인 4x4 블록을 따로 모아서 두개의 다른 부호책 크기를 갖는 2진 트리 구조 양자화기의 입력으로 들어간다.

3. 2진 트리 분류기를 이용한 부호책 설계

3.1 2진 트리 분류기

결정 트리 분류기는 분류하려는 특징에 대한 확률 밀도 함수가 비교적 알려져 있지 않은 경우에 좋은 분류기술로 알려져 있고[25], 문자인식[26]과 질량 스펙트럼 분석[25], 벡터 양자화[27, 28, 29], 그리고 외용공학[25, 30] 분야 등에서 널리 사용되고 있다. 2진 트리는 가장 간단하고 널리 사용되는 결정 트리 구조로서 하나의 뿌리(root) 노드와 일련의 단말(terminal), 비단말(non-

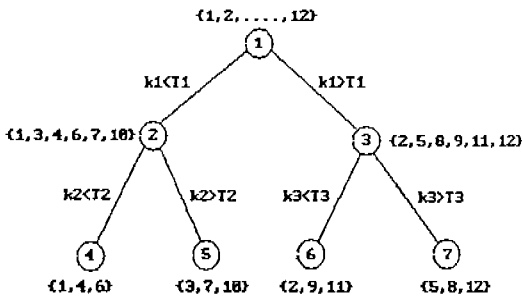
terminal) 노드로 구성되며, 비단말 노드는 2개의 하위 노드를 가진다. 패턴 벡터를 뿌리 벡터로부터 각 단말까지 트리를 통하여 이동함으로써 분류한다. 각 비단말 노드는 그 하위노드 중 하나에 입력 패턴을 할당하기 위해 특징 집단을 사용하고 단말노드는 분류된 패턴들에 할당되는 패턴 클래스, 즉 노드의 중심값을 가진다. 전체 성능을 최적화하기 위해 여러 가지 트리구조[31, 32, 33]가 적용되나 이들 최적화에는 많은 계산 시간과 저장 공간이 필요하다. 최적 트리 구조를 설계하는 것은 어려운 일이기 때문에 많은 초기 논문들은 균형 트리 구조(balanced tree structure)에 근거한 특징의 선택이나 결정 법칙(decision rule)을 연구했고[34, 35, 36] 이는 계산의 복잡도와 저장 공간 측면에서 비교적 좋은 성능을 나타낸다.

3.2 부호책 설계

본 논문에서는 벡터 양자화기 부호책을 만들기 위해 2진 트리 분류기를 사용하는데 학습 벡터로부터 추출된 특징들을 문턱값(threshold)으로 간단히 분류한다. 이는 부호화기의 설계를 간단히 하고 계산을 쉽게 해 준다. 즉, 학습벡터에서 '키(key)'로서 단일 특징이 추출되고 이를 노드에서 분류시에 사용하는데 이때의 결정 법칙은 단순히 문턱값 계산만 수행하면 되는 것이다. 비단말 노드에서 학습 벡터를 그 벡터의 키값이 문턱값보다 작으면 왼쪽 하위 노드로, 크면 오른쪽 하위 노드로 분류한다. 이 분류 과정을 원하는 부호책 크기에 도달할 때까지 계속하고 각 단말 노드에서는 학습 벡터의 중심값(centroid)을 부호책의 부호 벡터로 선택한다. 학습 벡터의 분류에 사용되는 특징들은 2차원 이산변환을 사용하여 추출하고, 추출된 특징중의 하나가 문턱값과 비교하는 키로서 사용되고, 문턱값은 특징들의 평균값으로 한다.

이와 같은 2진 트리 분류기를 사용한 부호책 설계의 예가 (그림 2)에 나타나 있는데 (그림 2)에서 학습 벡터 $\{x_1, x_2, \dots, x_{12}\}$ 로부터 4개 크기의 부호책을 만든다고 하면, 뿌리(①) 노드에서 이 12개의 학습벡터를 두개의 하위(②, ③) 노드로

분류하는데 각 벡터의 키 값(k_1)이 문턱값(T_1)보다 작으면 ②번 노드로, 크면 ③번 노드로 각 벡터를 분류하게 된다. 그 다음 ②번 노드에서도 학습벡터 $\{x_1, x_3, x_4, x_6, x_7, x_{10}\}$ 을 키 값(k_2)과 문턱값(T_2)을 비교하여 작으면 ④번 노드, 크면 ⑤번 노드로 분류하고 ③번 노드도 마찬가지로 수행한다. 이렇게 해서 만들어진 단말 노드(④, ⑤, ⑥, ⑦)의 분류된 벡터들의 중심값을 계산하여 그 값을 부호책의 부호 벡터로 사용한다. 예를 들어, ⑤번 노드의 부호벡터 C_5 은 $\left(\frac{1}{3}\right) \{x_3 + x_7 + x_{10}\}$ 과 같이 계산한다.



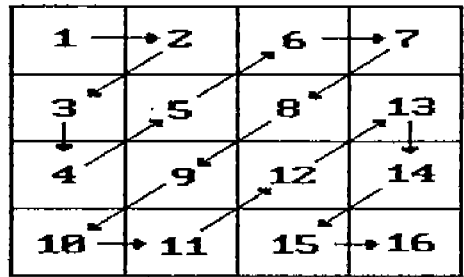
(그림 2) 2진 트리 분류기를 이용한 부호책 설계의 예
(Fig. 2) Example of codebook design using binary tree classifier

4. 2D-DCT를 사용한 특징 추출

변환(transform)의 목적은 통계적으로 연관된 화상의 화소를 비연관 계수로 바꾸는 것인데 변환된 데이터는 원 화상보다 분류목적에 있어서 훨씬 효과적인 특징 집합이 된다. 여러 변환 기법 중에서 본 논문은 DCT가 실(real)함수 계산이고, 계산의 복잡도가 비교적 낮으며, 거의 최적인 비연관(decorrelation) 효과가 있고, 고속 알고리즘[2, 3, 37]이 존재하기 때문에 DCT를 변환에 사용한다.

본 논문에서는 주파수 성분의 포함 여부에 따라 두개의 다른 부호책 크기를 갖는 부호화기를 설계한다. 2절에서의 DCT 맵을 이용하여 맵값이 '0'으로 정해진 블록들에 대해서는 '1'인 블록보다

적은 수의 부호책 크기를 갖게 하고, 그 블록의 특징 벡터로서 (그림 1)에서와 같이 3개의 DCT 계수 값을 사용하여 그 계수 값 특징 벡터를 2진 분류 트리 분류기로 입력한다. DCT 맵값이 '1'인 4x4 블록에 대해서는 2D-DCT를 취하여[38] 계산된 2D-DCT 계수들은 (그림 3)과 같이 지그재그 스캔으로 1차원으로 재 배열하여 이 재 배열된 계수값들을 그 블록의 특징 벡터로 한다. 이와 같은 1차원 배열의 특징 벡터를 2진 분류 트리의 분류기로 입력한다.



(그림 3) DCT 계수의 지그재그 스캔
(Fig. 3) The zigzag scan of the DCT coefficients

이때 각 노드에서 분류를 위한 키의 선택은 각 특징중에서 가장 큰 분산값을 갖는 특징을 키로 한다. 이는 데이터가 특정된 특징에 대해 넓게 퍼져 분포되어 있다면 그 특징에서의 차이가 가장 중요한 점이 되기 때문이다. 그리고 선택된 키에 해당하는 특징에 대한 평균값을 그 노드의 문턱값으로 사용하는데 이는 평균값이 특징의 대부분의 통계적 정보를 포함할 뿐만 아니라 계산이 간단하기 때문이다.

예를 들어 DCT 맵 값이 '1'인 4x4 블록에 대해, (그림 2)에서 설명한 트리 분류기에서 ②번 노드의 학습벡터를 1차원으로 나타낸 것을 z라 할 때 이 값들이 (그림 4)에 나타나 있다. (그림 4)에서와 같이 각 벡터에서 각 특징에 대한 평균값 $[M(1), \dots, M(16)]$ 과 분산값 $[V(1), \dots, V(16)]$ 을 계산한 후 가장 큰 분산값을 갖는 특징이 키로서 선택되고 그 키에 해당되는 평균값을 문턱값으로 한다. 예로서 $V(1)$ 값이 가장 크다면 이 ②

번 노드의 문턱값은 $M(1)$ 이 되는 것이다. 그래서 각 벡터의 첫번째 특징 벡터값, 즉 $z_1(1), z_2(1), \dots, z_n(1)$ 과 $M(1)$ 과 비교하여 작으면 그 벡터는 ④번 노드로, 크면 ⑤번 노드로 분류한다. DCT 맵 값이 '0'인 블록에 대해서는 16개의 특징 벡터 중 앞의 3개 벡터를 사용하는 것만이 다를 뿐이다.

학습 벡터	계수 벡터	특 징 벡 터
X_1	$Z_1 = \{ z_1(1) z_1(2) \dots z_1(16) \}^T$	
X_3	$Z_3 = \{ z_3(1) z_3(2) \dots z_3(16) \}^T$	
X_4	$Z_4 = \{ z_4(1) z_4(2) \dots z_4(16) \}^T$	
X_6	$Z_6 = \{ z_6(1) z_6(2) \dots z_6(16) \}^T$	
X_7	$Z_7 = \{ z_7(1) z_7(2) \dots z_7(16) \}^T$	
\vdots	\vdots	\vdots
X_{10}	$Z_{10} = \{ z_{10}(1) z_{10}(2) \dots z_{10}(16) \}^T$	
		$\downarrow \quad \quad \quad \downarrow$ 평균 $M(1) \quad M(2) \quad \dots \quad M(16)$ 분산 $V(1) \quad V(2) \quad \dots \quad V(16)$

(그림 4) 노드의 키 선택
(Fig. 4) The choice of key at node

5. 부호책 설계 알고리즘

2진 트리 구조의 분류기와 2D-DCT에 근거한 부호책 설계 알고리즘은 다음과 같다.

[단계 1] 입력 영상을 겹치지 않는 $M \times M$ 크기의 블록으로 나눈다.

[단계 2] 2D-DCT를 이용하여 각 블록의 특징 벡터를 추출한다.

특징 벡터를 Z 라 하면 $Z = \{ Z(j), j=1, 2, \dots, k \}$ 이다.

여기서 $k = M \times M$: DCT 짝 = "1"

$k = 3 \times 1$: DCT 짝 = "0"

[단계 3] 1) 각 특징(계수)에 대한 평균과 분산을 계산한다.

노드 n 에 대한 j 번째 특징의 평균을 $M_n(j)$, 분산을 $V_n(j)$ 라 하면 각각은 다음과 같이 정의한다.

$$M_n(j) = N_n^{-1} \sum_{i=1}^{N_n} Z_i^n(j) \quad (1)$$

$$V_n(j) = N_n^{-1} \sum_{i=1}^{N_n} \{ Z_i^n(j) - M_n(j) \}^2 \quad (2)$$

여기서 $j = 1, 2, \dots, k$

N_n = 노드 n 에서의 학습 벡터의 수

$Z_i^n(j)$ = 노드 n 에서 학습 벡터에 해당하는 j 번째 특징

2) 가장 큰 분산을 가진 특징을 키로서 채택한다. 즉, $V_n(p) = \max\{ V_n(j), j=1, 2, \dots, k \}$ 이면 p 번째 특징이 n 노드의 키로서 선택되고 $M_n(p)$ 를 문턱값으로 선택한다. 특징 벡터의 p 번째 특징의 값이 $M_n(p)$ 보다 작으면 왼쪽 하위 노드로, 크면 오른쪽 하위 노드로 분류한다.

3) 요구되는 2진 트리의 단말 노드의 수, 즉 부호책의 부호 단어의 수가 될 때까지 1), 2)를 반복한다.

[단계 4] 각 단말 노드에서의 벡터들의 중심값을 계산하여 부호책의 부호 벡터로 사용한다. 즉, 단말 노드 q 의 중심값은 식 (3)과 같다.

$$Cent(q) = N_q^{-1} \sum_{i=1}^{N_q} X_i \quad (3)$$

여기서 X_i 는 노드 q 의 학습 벡터이고 N_q 는 이 노드의 학습 벡터 수를 나타낸다.

6. 실험결과

제안된 알고리즘의 평가는 국제 표준화상인 512 × 512 크기의 Lenna와 Boat 영상에 대해 수행하였고 DCT 맵 설계시의 문턱값(threshold) δ 값과 부호책(codebook)의 크기를 변경시키면서 실험하였다. <표 1>은 DCT 맵이 "0"인 블록의 부호책 크기는 128로, "1"인 블록의 부호책 크기는 256으로 고정시키고 문턱값 δ 를 10에서 70까지 변화시킬 때 Lenna와 Boat 복원 영상의 평균 제곱 오차(MSE)와 PSNR값을 보여주고 있다. 표에서 δ 값이 커질수록 PSNR이 커져서 복원 영상의 화질이 좋아짐을 알 수 있다. Lenna 영상의 경우

δ 값이 60일 때 가장 좋은 화질을 가지지만 70인 경우, 즉 δ 를 계속 늘린다고 영상의 화질은 개선되지 않는 것을 볼 수 있는데 이는 δ 의 큰 값으로 인해 DCT 맵 설계시 거의 모든 블록들이 한쪽으로 맵핑되므로 그 후에 수행되는 두 개의 구별된 트리 구조 양자화기의 한쪽에 많은 벡터가 몰려 균형된 분류기로 동작되지 못하여 발생한 결과이다. Boat 영상의 경우 Lenna 영상과 마찬가지로 δ 값이 커질수록 PSNR이 커지지만 δ 가 50과 60 일때 PSNR 값이 거의 비슷한 것을 볼 수 있다. Boat 영상이 Lenna 영상보다 윤곽선이 많이 존재하기 때문에 본 알고리즘의 수행결과 PSNR 값이 약간 적게 나타났다. <표 2>는 문턱값 δ 를 60으로 고정시키고 부호책의 크기를 각각 64에서 128, 64에서 256으로 변화시킬 때 두 복원 영상의 PSNR 값을 비교한 것이다. 부호책의 크기가 모두 64일 때 두 복원 영상의 화질이 가장 나쁘고, DCT 맵 값 "0"에 대한 분류기의 부호책의 크기가 128이고, "1"인 분류기의 부호책의 크기가 256인 경우가 가장 화질이 좋았다. 각 분류기의 부호책의 크기가 각각 64와 128인 두가지 경우를 비교해보면 DCT 맵이 "0", 즉 DCT 계수의 3개만 특징 벡터로 사용하는 경우 부호책의 크기가 64이고 16개의 모든 DCT 계수를 사용하는 맵이 "1"인 부호책의 크기가 128인 경우가 그 반대인 경우보다 Lenna 영상은 0.15dB, Boat 영상은 0.98 dB 정도 PSNR 값이 더 컸다. 이는 맵 "1"에 대한 분류기, 즉 고주파 성분에 대한 분류기의 부호책 크기가 복원 영상의 화질에 더 큰 영향을 미치기 때문이다. (그림 5)와 (그림 6)에 Lenna와 Boat 영상에 대한 PNN 알고리즘과 CVQ 알고리즘, 본 논문에서 제안한 알고리즘의 시뮬레이션 결과가 나타나 있는데 각각의 그림에서 (b)의 PNN 결과 영상과 (f)의 제안 알고리즘의 결과 영상을 비교해 보면 영상의 윤곽선(edge) 영역에서 PNN보다 본 알고리즘이 더 적게 왜곡이 일어났고, (d)의 CVQ 결과 영상과 (f)의 제안 알고리즘의 결과 영상은 PNN보다는 적게 왜곡이 발생한 것을 볼 수 있다. 더 자세히 비교하기 위해서 세개의 알고리즘과 원 영상

과의 오류(error)를 계산하여 화소값을 10배 확대한 그림인 (c), (e)와 (g)에서 본 논문에서 제안된 알고리즘의 오류 화상인 (g)의 오류가 (c)의 PNN 오류보다 적고, (e)의 CVQ 오류는 PNN의 오류보다는 적지만 제안 알고리즘보다는 약간 큰 정도의 오류가 발생한 것을 볼 수 있다. <표 3>에는 PNN과 CVQ, 제안된 알고리즘의 CPU 실행 시간과 MSE, PSNR값의 비교가 나타나 있는데 알고리즘 수행 시간은 Lenna 영상의 경우 본 알고리즘이 PNN보다 약 1.8배, CVQ보다는 4.3배 가량 빠르고, PNN의 경우 PSNR은 약 0.45 dB정도 증가하고 MSE도 4.5정도 감소되었으며, CVQ의 경우는 약 0.05 dB정도 PSNR이 증가, MSE는 0.42만큼 감소되어 실제 CVQ 결과 영상은 주관적인 판단으로 제안 알고리즘의 결과 영상과 거의 같았다. Boat 영상의 경우에도 제안 알고리즘이 PNN의 경우 수행시간은 약 1.8배, CVQ는 4.1배나 빠르고, PSNR은 각각 0.33 dB, 0.1dB정도 증가하고 MSE는 4.77, 1.54정도 감소되어 본 논문에서 제안된 알고리즘이 PNN이나 CVQ에 비하여 더 좋은 결과를 나타내는 것을 알 수 있다.

<표 1> 크기가 128 (DCT 맵-"0"), 256 (DCT 맵-"1")인 부호책을 갖는 분류기에서 DCT 맵 문턱값 δ 의 변화에 의한 복원 영상의 MSE와 PSNR 비교

<Table 1> Comparison of MSE and PSNR for reconstructed image based on DCT map threshold δ at TSVQ codebook size 128 and 256

영 상	문턱값 δ	10	20	30	40	50	60	70
		MSE	44.01	43.75	43.11	42.26	41.56	40.00
Lenna	PSNR (dB)	27.89	29.05	30.34	31.11	31.76	32.11	31.98
	MSE	46.26	45.98	44.06	43.01	41.10	41.01	42.26
Boat	PSNR (dB)	26.00	26.99	27.63	30.96	31.83	31.87	30.90

<표 2> DCT 맵 문턱값 δ 가 60일 때 부호책의 크기에 의한 복원 영상의 PSNR 비교

<Table 2> Comparison of PSNR for reconstructed image based on variable TSVQ codebook size at DCT map threshold δ , 60

DCT 맵 "0" 부호책 크기	64	64	64	128	128	128	
DCT 맵 "1" 부호책 크기	64	128	256	64	128	256	
PSNR (dB)	Lenna	28.66	29.79	30.89	29.64	30.86	32.11
	Boat	26.84	28.99	29.96	28.01	29.97	31.87

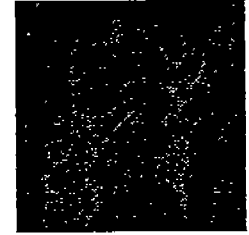
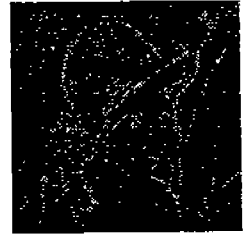
<표 3> PNN, C.VQ, 제안 알고리즘의 CPU 실행시간과 MSE, PSNR 비교

<Table 3> Comparison of CPU time and MSE, PSNR for PNN, C.VQ and proposed algorithm

영	상	PNN 알고리즘	C.VQ	제안 알고리즘
Lenna	CPU 실행시간	32	730	185
	MSE	44.41	41.42	40.00
	PSNR(dB)	11.51	12.66	32.11
Boat	CPU 실행시간	77	334	206
	MSE	43.73	40.55	41.01
	PSNR(dB)	11.50	11.77	31.87

7. 결 론

본 논문에서는 영상의 벡터 양자화를 위한 새로운 부호책 설계 알고리즘을 제안하였다. 이 알고리즘은 반복적이 아니며 초기 부호책을 필요로 하지 않고 2진 트리도 특징 벡터를 분류하였다. 영상을 작은 부분들로 나누고 2차원 이산여현변환으로 고주파 성분이 많이 포함된, 즉 대부분의 정보를 포함하는 부분과 부호화하기 쉬운 적은 정보를 포함하는 영역으로 나누어 DCT 값을 만들고 이 맵에 따라 영상의 특징 벡터를 2차원 이산여현변환으로 추출한다. DCT 값에 따라 두 개의 트리 구조 분류기를 사용하는 2진 트리의 각 노드에서 분류 기준은 간단한 평균과 분산 값의 계산만으로 수행



(그림 5) "Lenna" 영상에 대한 PNN, CVQ 알고리즘과 본 알고리즘의 비교

- (a) 원 영상
- (b) PNN의 결과 영상
- (c) PNN의 오류 영상
- (d) CVQ의 결과 영상
- (e) CVQ의 오류 영상
- (f) 제안 알고리즘의 결과 영상
- (g) 제안 알고리즘의 오류 영상

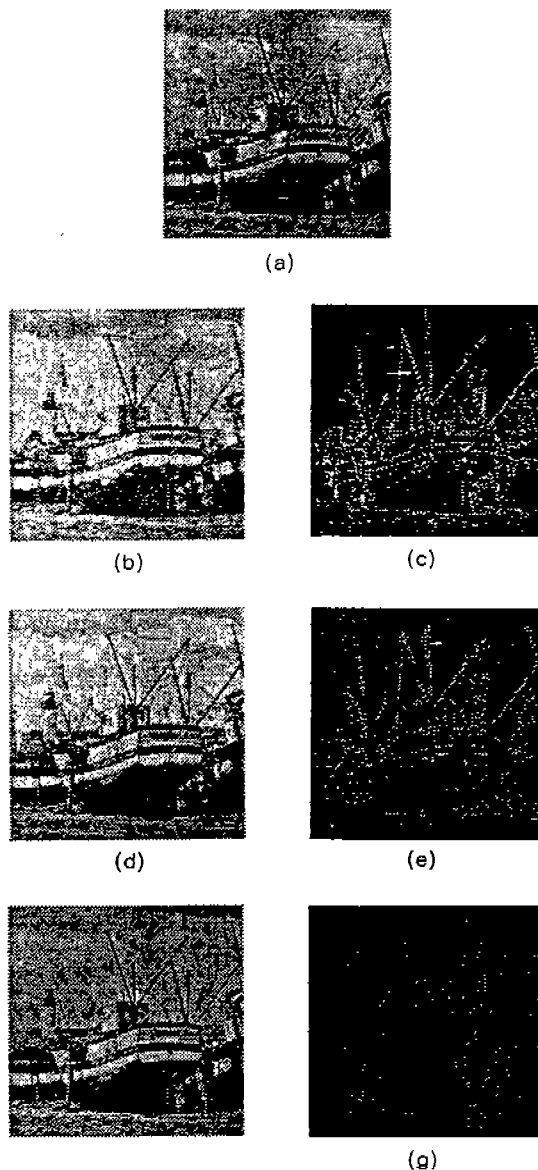
(Fig. 5) Comparison of PNN, CVQ and proposed algorithm of "Lenna"

- (a) Original image
- (b) Reconstructed image of PNN
- (c) Error image of PNN
- (d) Reconstructed image of CVQ
- (e) Error image of CVQ
- (f) Reconstructed image of proposed algorithm
- (g) Error image of proposed algorithm

하였다. 실험결과는 본 논문에서 제안한 알고리즘으로 수행시 PNN과 CVQ 알고리즘보다 실행 시간을 줄일 수 있었고 더 좋은 화질의 영상을 얻을 수 있었다.

참 고 문 헌

[1] R.C.Gonzalez and R.E.Woods, *Digital Image Processing*, Addison-Wesley, NY, 1992.



(그림 6) "Boat" 영상에 대한 PNN, CVQ 알고리즘과 본 알고리즘의 비교

- (a) 원 영상 (b) PNN의 결과 영상
- (c) PNN의 오류 영상
- (d) CVQ의 결과 영상
- (e) CVQ의 오류 영상
- (f) 제안 알고리즘의 결과 영상
- (g) 제안 알고리즘의 오류 영상

(Fig. 6) Comparison of PNN, CVQ and proposed algorithm of "Boat"

- (a) Original image
- (b) Reconstructed image of PNN

- (c) Error image of PNN
- (d) Reconstructed image of CVQ
- (e) Error image of CVQ
- (f) Reconstructed image of proposed algorithm
- (g) Error image of proposed algorithm

- [2] W.H.Chen, C.H.Smith, and S.C.Fralick, "A fast computational algorithm for the discrete cosine transform," IEEE Trans. Commun., Vol.COM-25, pp.1004-1008, Sept. 1977.
- [3] B.G.Lee, "A new algorithm to compute the discrete cosine transform," IEEE Trans. Acoust., Speech, Signal Processing, Vol.ASSP-32, pp.1243-1245, Dec. 1984.
- [4] N.I.Cho and S.U.Lee, "DCT algorithms for VLSI parallel implementation," IEEE Trans. Acoust., Speech, Signal Processing, Vol.ASSP-38, No.1, pp.121-127, Jan. 1990.
- [5] A.Habibi, "Survey of adaptive image coding techniques," IEEE Trans. Commun., Vol.COM-25, No.11, pp.1275-1284, Nov. 1977.
- [6] A.N.Netravali and J.O.Limb, "Picture coding : A review," Proc. IEEE, Vol.68, pp.366-406, Mar. 1980.
- [7] M.D.Lema and O.R.Mitchel, "Absolute moment block truncation coding and its application to color images," IEEE Trans. Commun., Vol.COM-32, No.10, pp. 1148-1157, Oct. 1984.
- [8] L.Wang and M.Goldberg, "Reduced-difference pyramid : A data structure for progressive image transmission," Opt. Eng., Vol.28, No.7, pp.708-716, July 1989.
- [9] H.Blume and A.Fand, "Reversible and irreversible image data compression using the S-transform and Lempel-Ziv coding," in Proc. SPIE Medical Imaging

- III: Image Capture and Display. Vol. 1031, pp.2-18, 1989.
- [10] P.J.Burt and E.H.Adelson, "The Laplacian pyramid as a compact image code," IEEE Trans. Commun., Vol.COM-31, No.4, pp.532-540, Apr. 1983.
- [11] R.Wilson, H.E.Knutsson, and G.H.Granlund, "Anisotropic nonstationary image estimation and its applications: Part II - Predictive image coding," IEEE Trans. Commun., Vol.COM-31, No.3, pp.398-403, Mar. 1983.
- [12] M.Kocher and M.Kunt, "A contour-texture approach to picture coding," ICASSP-82, Paris, France, pp.436-440, May 1982.
- [13] A.Ikonomopoulos and M.Kunt, "Image coding based on a directional decomposition," Proc. Int. Picture Coding Symp., pp.13-15, Davis, CA, Mar. 1983.
- [14] A.K.Jain, "Image data compression: A review," Proc. IEEE, Vol.69, pp.349-389, Mar. 1981.
- [15] C.E.Shannon, "Coding theorems for a discrete source with a fidelity criterion," in IRE Nat. Conv. Rec., Vol.4, pp.142-163, 1959.
- [16] 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.
- [17] A.Gersho and V.Cupermann, "Vector quantization: A pattern matching technique for speech coding," IEEE Commun. Mag., Vol.21, pp.15-21, Dec. 1983.
- [18] R.M.Gray, "Vector quantization," IEEE ASSP Mag., Vol.1, pp.4-29, April 1984.
- [19] N.M.Nasrabadi and F.B.King, "Image coding using vector quantization: a review," IEEE Trans. Commun., Vol.COM-36, pp.957-971, Aug. 1988.
- [20] S.P.Lloyd, "Least-squares quantization in PCM," IEEE Trans. Inform. Theory, Vol.IT-28, pp.129-137, March 1982.
- [21] J.A.Hartigan, Clustering Algorithm. New York: Wiley, 1975.
- [22] W.Equitz, "Fast algorithm for vector quantization picture coding," IEEE Int. Conf. on ASSP, pp.725-727, 1987.
- [23] W.Equitz, "A new vector quantization clustering algorithm," IEEE Trans. ASSP, Vol.37, pp.1568-1575, Oct. 1989.
- [24] B.Ramamurthi and A.Gersho, "Classified vector quantization of images," IEEE Trans. Commun., Vol.COM-34, No.11, pp.1105-1115, Nov. 1986.
- [25] L. Breiman et al. Classification and Regression Trees. The Wadsworth Statistics / Probability Series. Belmont, CA : Wadsworth, 1984.
- [26] R.G.Casey and G.Nagy, "Decision tree design using a probabilistic model," IEEE Trans. Inform. Theory, Vol.IT-30, pp.93-99, 1984.
- [27] A.Buzo et al., "Speech coding based upon vector quantization," IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-28, pp.562-574, Oct. 1980.
- [28] E.A.Riskin and R.M.Gray, "Lookahead in growing tree-structured vector quantizers," IEEE Int. Conf. on ASSP, pp.2989-2992, Mar. 1991.
- [29] P.A.Chou, T.Lookabaugh, and R.M.Gray, "Optimal pruning with applications to tree-structured source coding and modeling," IEEE Trans. Inform. Theory, Vol. 35, pp.299-316, Mar. 1989.
- [30] G.H.Landerweed et al., "Binary tree versus single level tree classification of white blood cells," Pattern Recognition, V.16, pp.571-577, 1983.
- [31] P.Swain and W.Meisel, "The decision tree classifier design and potential,"

IEEE Trans. Geosci.Electron., Vol.GE-15, pp.142-147, July 1977.

[32] A.V.Kulkarni and L.N.Kanal, "An optimization approach to hierachical classifier design," in Proc. Third Int. Joint Conf.on Pattern Recognition, 1976, pp. 459-466.

[33] H.J.Payne and W.S.Meisel, "An algorithm for constructing optimal binary decision trees," IEEE Trans. Comput., Vol. C-26, pp.905-916,Sept.1977.

[34] Y.K.Lin and K.S.Fu, "Automatic classification of cervical cells using a binary tree classifier," Pattern Recognition, Vol. 16, No.1, pp.69-80,1983.

[35] M.W.Kurzynski, "The optimal strategy of a tree classifier," Pattern Recognition, Vol.16, No.1, pp.81-87, 1983.

[36] A.V.Kulkarni, "On the mean accuracy of hierarchical classifiers," IEEE Trans. Comput., Vol.C-27, pp.771-776, Aug. 1978.

[37] H.S.Hou, "A fast recursive algorithm for computing the discrete cosine transform," IEEE Trans. Acoust., Speech, Signal Processing, Vol.ASSP-35, pp.1455-1461, Oct.1987.

[38] W.H.Chen and W.K.Pratt, "Scene adaptive coder," IEEE Trans. Commun., Vol. COM-32, pp.225-232, Mar.1984.



조 성 환

1980년 성균관대학교 전자공학과 졸업(학사)
 1982년 성균관대학교 대학원 전자공학과(공학석사)
 1991년 성균관대학교 대학원 전자공학과(공학박사)
 1982년~85년 해군사관학교 전기 및 전자공학과 전임강사
 1985년~현재 대우공업전문대학 전자계산기과 조교수
 관심분야: 신경회로망, 영상처리, 패턴인식



김 응 성

1989년 성균관대학교 전자공학과 졸업(학사)
 1992년 성균관대학교 대학원 전자공학과(공학석사)
 1992년~현재 성균관대학교 대학원 전자공학과 박사과정
 관심분야: 영상처리, 영상통신, HDTV, 신경회로망