

論文99-36S-8-11

멀티미디어 정보검색에 적합한 영상정보의 벡터 양자화 코드북 설계 및 특징추출

(VQ Codebook Design and Feature Extraction of Image Information for Multimedia Information Searching)

徐錫培*, 金大鎭**, 姜大星*

(Seok-Bae Seo, Daijin Kim, and Dae-Seong Kang)

요 약

본 논문에서는 멀티미디어 정보검색을 위하여 영상정보의 특징추출에 적합한 벡터 양자화 코드북 설계 방법을 제안한다. 기존의 벡터 양자화의 경우 영상에 대한 특징을 추출할 경우 보통 영상을 복원한 다음 수행하므로 많은 시간과 메모리가 소요되며, DCT(discrete cosine transform)를 이용한 방법처럼 블록화 현상을 동반한다. 이를 개선하기 위하여 본 논문에서는 웨이블릿 변환과 주성분 해석을 이용한 벡터 양자화 코드북 설계 방법을 제안한다. 웨이블릿 변환은 높은 압축률에서도 블록화 없는 영상을 복원하기 위해서 도입되었으며, 주성분해석은 데이터를 여러 그룹으로 분할하기 위해 도입되었다. 신경회로망인 SOM(self-organizing map)을 이용한 벡터 양자화와 비교실험에서 비슷한 성능을 보이면서도 처리 시간을 대폭 단축시킬 수 있음을 볼 수 있었다.

Abstract

In this paper, the codebook design method of VQ (vector quantization) is proposed as method to extract feature data of image for multimedia information searching. Conventional VQ codebook design methods are unsuitable to extract the feature data of images because they have too much computation time, memory for vector decoding and blocking effects like DCT (discrete cosine transform). The proposed design method consists of the feature extraction by WT (wavelet transform) and the data group divide method by PCA (principal component analysis). WT is introduced to remove the blocking effect of an image with high compressing ratio. Computer simulations show that the proposed method has the better performance in processing speed than the VQ design method using SOM (self-organizing map).

* 正會員, 東亞大學校 電氣電子컴퓨터工學部

(School of Electrical, Electronic, and Computer Engineering, Dong-A University)

** 正會員, 浦港工科大学校 컴퓨터工學科

(Dept. of Computer Engineering, Pohang Univ. of Science and Technology)

※ 이 논문은 1997년 한국학술진흥재단 학술연구조성비에 의해 지원되었음.

接受日字:1999年2月3日, 수정완료일:1999年7月5日

I. 서 론

최근 컴퓨터와 통신의 발달로 멀티미디어 데이터가 증가함에 따라서 이를 효율적으로 관리하는 기술에 관한 연구가 활발히 진행 중이다. 멀티미디어 데이터는 문서, 음성, 영상, 비디오 등으로 구성되는데, 그 중 영상이나 비디오 등 영상에 관련된 데이터가 많은 용량을 차지하며 중요한 정보를 가지고 있다. 따라서 이를 효율적으로 저장하고 전송하기 위한 영상 압축 기술의

개발이 요구되며, 이는 영상 관련 데이터 뿐 아니라 전체 멀티미디어 데이터의 관리 및 활용에 직접 관련된다^[1-5].

본 논문에서 제안하는 압축 방법은 여러 영상 압축 기술 중에서 특성이 우수한 벡터 양자화(vector quantization)의 코드북 설계(codebook design)에 주성분 해석(Principal Component Analysis; PCA) 알고리즘을 도입함으로써, 기존의 신경회로망을 이용한 코드북 설계 방법과 달리 학습과정 없이 입력 데이터의 통계적 특성에 기반하여 벡터 양자화 코드북을 자율적으로 설계한다. 그리고 웨이브렛 변환을 이용한 특징추출에 의하여 복원 영상의 블록화 현상을 제거하고 압축 영역에서 원 영상의 정보를 추출할 수 있는 방법을 제시한다.

여러 가지의 영상 압축 기술 중에서 벡터 양자화는 압축 능력이 우수하여 영상 압축, 음성 압축을 비롯한 여러 분야에 널리 사용되고 있다^[6-8]. 벡터 양자화를 위해서는 입력 벡터와의 평균 제곱 오차(mean square error)를 최소화하는 코드북의 결정이 요구되며 이를 위해서 자기조직 능력을 갖는 SOM(Self-Organizing Map)이 흔히 이용된다. 하지만 SOM을 이용하는 경우 다른 신경회로망과 같이 학습과정이 요구되며, 따라서 여러 가지 학습 파라미터 설정이 쉽지 않고 원하는 해를 얻기까지 수렴하는데 많은 학습 시간이 요구되는 단점이 있다. SOM의 구조를 변화시킨 DSNN(Doubly Self-organizing Neural Network), SOM의 학습에 퍼지 알고리즘을 접목한 FSOM(Fuzzy Self-Organizing Map) 등 여러 가지 방법이 제안되었으나 위의 문제점을 근본적으로 해결하지 못하였다^[9-11]. 본 논문에서는 학습에 의존하지 않고 입력 벡터의 전체적인 분포 특성을 고려하여 코드북을 설계하는 고속의 알고리즘을 제안한다. 이 방법은 기존의 신경회로망과 다르게 학습과정 없이 입력 데이터의 통계적 특성에 기반하여 벡터 양자화 코드북을 자율적으로 설계함으로써, 파라미터 설정과 처리시간 장대화 문제를 개선하였다.

제안하는 알고리즘을 적용하기 전에 특징추출과정으로 웨이브렛 변환(wavelet transform)을 적용한다^[12]. 본 논문에서는 입력 영상을 4-band 웨이브렛 변환한 다음 가장 저주파 성분을 갖는 부대역(subband)만을 입력 데이터로 추출하였다. 나머지 밴드에는 영상의 수직, 수평, 대각에 대한 정보가 저장되는데, 보통 전

체적인 압축률을 높이기 위해 이 정보를 버리거나 높은 비율로 손실압축을 한다. 본 논문에서는 가장 저주파 성분과 부대역의 부대역만 압축하여 영상의 정보로 저장하고 나머지 부대역은 제거하였다. 이렇게 압축된 영상을 복원하면 영상이 좀 흐려질 뿐 영상의 대부분의 정보는 그대로 유지되며, 압축 비율이 높아지면 영상의 흐려짐의 정도가 점점 심해지게 된다. 따라서 본 실험에 사용한 특징추출 방법은 압축비율이 높아져도 DCT를 이용한 영상 압축처럼 영상의 블록화 현상이 나타나지 않으므로 인간의 시각적 특성에 유사한 압축기법이라 할 수 있다.

제안하는 주성분 해석 알고리즘을 이용한 코드북 설계 방법은 웨이브렛 변환으로 얻어진 특징 데이터의 통계적 특성을 이용하여 벡터 양자화 코드북을 생성한다. 먼저 모든 입력 특징 데이터를 한 노드에 할당한다. 다음 초기 노드에 대한 주축을 구한다. 주축은 해당 노드의 데이터의 변화가 가장 많은 축으로, 해당 노드 내의 데이터를 이 축으로 근사하면 데이터의 차원을 낮추면서도 효과적으로 데이터의 정보를 표현할 수 있다. 본 논문에서 입력 벡터는 2차원인데 주성분 해석에 의하여 1차원의 데이터로 표현이 가능하다. 다음으로 차원이 줄어든 데이터에 대하여 오차가 가장 적은 분할점을 계산하여 그 점을 기준으로 데이터를 양분한다. 양분된 데이터에 대해서도 위의 방법이 똑같이 적용되어 원하는 노드 수에 도달할 때까지 반복하게 된다. 입력 벡터의 차원을 1차로 줄여서 계산량이 감소하긴 하였지만 최적의 분할점을 찾는 계산의 양은 여전히 문제점으로 남아있다. 이를 해결하기 위해서 수치해석에서 흔히 쓰이는 평균 구하는 방법을 도입하였다. 이 방법은 현재의 평균과 데이터 개수를 알고 있으면 데이터가 하나 추가될 때 모두의 평균을 계산하는 것이 아니고 한 번의 계산으로 추가된 데이터를 포함한 평균치를 계산한다. 이상의 방법에 의하여 코드북을 설계하면 코드북의 생성시간이 짧을 뿐만 아니라 파라미터 설정의 어려움 없이 좋은 특성의 코드북을 설계할 수 있다.

II. 주성분 해석을 이용한 코드북 설계와 영상정보의 특징추출

1. 주성분 해석(Principal Component Analysis)

주성분 해석 기법은 입력 공간내의 변동량이 많은

방향으로 주축을 발견하여 데이터의 중복성을 제거하는 기법 중 널리 사용되는 방법으로 주성분 분석기법이라고도 한다. 이 기법은 입력 데이터를 보다 적은 차원으로 근사시켜 표현할 수 있으므로 데이터 압축에 유용하게 적용할 수 있다. n 차원의 정상적 확률 신호(stationary random signal)에서 입력 신호 X 는 식 (1)과 같이 벡터로 표시되며 그의 자기상관행렬 (autocorrelation matrix) R_X 는 식 (2)와 같다.

$$X = [x_1, x_2, \dots, x_n]^T \quad (1)$$

$$R_X = E[XX^T] \quad (2)$$

여기서, 자기상관행렬 R_X 의 정규화된 고유치(normalized eigenvectors) $\{e_i\}$ ($\|e_i\|=1, i=1, 2, \dots, m$)에 의하여 확률신호 X 는 1차원 부밴드(sub-space)로 투영(projection)될 수 있다. 따라서 X 의 주성분(principal component)은 식 (3)과 같이 R_X 의 정규화된 고유치 중 가장 큰 고유치(e_1)의 고유 요소(eigen-component)로 표현할 수 있으며, m 개의 first principal component를 사용하면 n 차원의 확률신호를 m 차원으로 축소할 수 있다.

$$a = e_1^T X \quad (3)$$

영상 압축에 주성분 해석을 이용하려면 먼저 n 차원 영상의 특징 데이터에 대하여 자기상관행렬 R_X 와 그의 정규화된 고유치를 구한다. 그 다음 추출한 R_X 의 정규화된 고유치 $\{e_i\}$ 를 크기 역순으로 배열하고 그 중 $m(m \leq n)$ 개를 이용하여 특징 데이터를 나타내면 n 차원 특징 데이터를 $n:m$ 의 비율로 압축할 수 있다. 즉, 영상의 통계적 특성을 이용하여 영상 특징 데이터의 차원을 $n:m$ 의 비율로 압축한 것이다.

2. 주성분 해석을 이용한 코드북 설계

벡터 양자화에서 코드북의 설계는 전체의 압축 성능을 좌우하므로 매우 중요하다. 코드북의 설계방법은 여러 가지가 제안되어 있으나 우수한 성능의 SOM을 이용한 방법이 널리 쓰인다. 하지만 SOM은 실제로 적용할 때 학습 시간의 장대화, 파라미터 설정 문제의 단점이 있어서 이를 개선하기 위한 다양한 연구가 있었다. 그 대표적인 예가 인간의 판단능력을 모방한 퍼지 알고리즘(Fuzzy Algorithm), 생물의 진화 과정을 모델링한 진화 알고리즘(Genetic Algorithm)과 SOM

의 결합이다. 그리고 최근에 제안된 gas는 SOM의 웨이트 크기의 정보를 계층(rank)을 두어 학습함으로써 SOM의 단점을 개선하였다^[10]. 열거한 알고리즘들에 의하여 성능이 많이 개선되었지만 SOM을 이용할 때 가장 문제가 되는 점인 학습의 장대화 문제와 파라미터 설정 문제는 여전히 존재한다. 일반적인 경우 실제 사용하는 값은 학습이 끝난 웨이트 값이므로 학습시간이 길더라도 큰 문제가 되지 않지만, 만약 학습할 대상이 자주 변화한다면 SOM을 사용하는 것은 비효율적이다. 이를 개선하기 위하여 주성분 해석을 이용한 새로운 알고리즘이 제안되었다. 제안된 방법은 SOM과 유사한 특성을 가지며 빠른 코드북 생성시간을 가지고 있으나 입력 파라미터가 오차이므로 원하는 압축률을 얻기 위해서는 오차값을 바꾸어가며 여러 번 실험 해야하는 단점이 있다. 그리고 이를 개선한 논문 [13]에서는 파라미터를 실제 압축률로 설정하여 파라미터 설정 문제와 학습의 장대화 문제를 개선하였지만 압축비율이 높을 때 성능이 저조하였다. 본 논문에서는 위 방법을 더욱 개선하여 오차가 가장 큰 노드를 반복하여 나누어 갈 때 최적의 경계를 설정하는 방법을 보충함으로써 파라미터 설정문제, 학습의 장대화 문제, 압축 성능 개선문제를 동시에 해결하였다. 전체 알고리즘과 세부 알고리즘을 다음에 설명한다.

(가) 전체 알고리즘

Step 1. 입력 영상에서 특징을 추출한다.

Step 2. 최종노드 개수를 설정한다. 최종노드 알고리즘의 압축률을 결정하는 파라미터로 압축률은 식 (4)와 같이 계산할 수 있다.

$$\text{압축률} = \frac{\text{입력벡터의개수} \times 4}{\text{최종노드개수}} \quad (4)$$

Step 3. 1개의 초기노드를 생성한 다음 최소오차 경계결정 알고리즘을 적용한다.

Step 4. 최소오차경계를 기준으로 초기노드를 2개로 분할한다.

Step 5. 새롭게 생성된 2개의 노드에 대해서 최소 오차경계결정 알고리즘을 적용한다.

Step 6. 각 노드들을 오차값에 의해 내림차순 정렬한다.

Step 7. 마지막에 위치한 노드(오차값이 가장 큰 노드)를 최소오차경계를 기준으로 2개로

분할한다.

Step 8. 전체의 노드의 개수가 초기에 설정한 최종노드의 개수와 같아질 때까지 Step 5, Step 6, Step 7을 반복한다.

(2) 벡터 양자화를 위한 특징추출

본 논문에서 벡터 양자화를 위한 특징추출은 영상을 웨이브렛 변환한 다음 그 결과를 정규화(normalization)해서 사용하였다. 기존의 DCT를 이용한 압축에서는 원 영상을 일정한 매크로 블록단위인 8×8로 분할하기 때문에 복원한 영상에서 블록화 현상을 볼 수 있다. 이 현상은 압축률이 높아지면 높아질수록 심해지는 단점이 있어서 이를 개선하기 위한 연구결과가 많이 발표되었다. 블록들을 겹치도록 하여 블록간 상관성을 가질 수 있도록 한 방법^[14]이나, 블록 경계면들의 각 화소들에 저대역통과 필터링 한 방법^[15] 등이 좋은 예라 할 수 있다. 본 논문에서는 원 영상을 4-band 웨이브렛 변환한 다음 중요한 정보를 가지고 있는 가장 저역밴드를 영상의 정보로 이용한다. 영상의 웨이브렛 변환에서 가장 저역밴드의 정보는 저주파 성분으로 영상의 대부분의 정보가 포함되어 있으며 나머지 밴드에는 각각 가로성분, 세로성분, 대각선 성분이 포함되어 있다. 본 논문에서는 압축비율을 높이기 위하여 가로성분, 세로성분, 대각선 성분의 밴드는 제거하였다. 가장 저역밴드의 정보만으로 영상을 복원하면 영상 전체가 희미해지는 경향은 있으나 블록화 현상이 생기지 않으므로 DCT를 이용한 방법에 비해서 시각적 특성이 뛰어난을 볼 수 있다.



그림 1. 특징추출 알고리즘의 구성
Fig. 1. Block diagram of the feature extraction algorithm.

웨이브렛 변환의 가장 저역밴드의 정보를 영상의 특징으로 나타내기 위해서 먼저 저역밴드의 정보를 0과 100사이로 정규화한다. 그 다음 데이터를 2개씩 묶어 2차원 벡터로 나타내어 벡터 양자화에 적합한 형태로 변환한다. 데이터를 2차원 벡터로 표현하는 방법은 데이터를 순서대로 두 개씩 묶어서 저장하면 된다. 원 영상이 256×256 영상이므로 2차원 벡터의 개수는

$((256 \times 256)/4)/2$ 개, 즉 8192개이다. 그림 1은 특징추출 알고리즘의 구성을 나타낸다.

(3) 주성분 해석에 의한 데이터 차원감소

본 논문에서 제안하는 최적경계결정 알고리즘은 오차가 가장 작은 경계를 찾아야 하므로 많은 계산량이 요구되며 속도가 매우 느린 단점이 있다. 이를 개선하기 위해서 최소오차경계결정 알고리즘을 적용하기 전에 주성분해석을 도입하여 데이터의 차원을 감소시킴으로서 계산속도를 증가시켰다. 주성분 해석에 의하여 계산된 주축은 입력 데이터의 최대 고유치의 고유 벡터의 방향과 일치하며 입력 데이터의 변동량이 가장 큰 방향이다. 이 축은 2차원 데이터를 1차원으로 감소시키는 기준으로, 차원 감소시 최소의 오차를 가지는 방향이다. 2차원의 입력 벡터를 1차원으로 감소하는 방법을 6단계로 설명한다. 여기서, 입력 벡터의 축을 X 와 Y 로 표시하며 입력 벡터를 $V = [V_x, V_y]^T$ 로 나타낸다.

주성분 해석에 의하여 계산된 주축은 입력 데이터의 최대 고유치의 고유 벡터의 방향과 일치하며 입력 데이터의 변동량이 가장 큰 방향이다. 이 축은 2차원 데이터를 1차원의 데이터로 감소시키는 기준으로, 차원 감소시 최소의 오차를 가지는 방향이다

Step 1. 입력 벡터에 대한 기하학적 평균 (geometrical mean) M 을 구한다.

$$M = [m_x, m_y]^T \tag{5}$$

$$m_x = \frac{1}{n} \sum_{i=0}^{n-1} v_{x_i}, \quad m_y = \frac{1}{n} \sum_{i=0}^{n-1} v_{y_i}$$

여기서, v_{x_i}, v_{y_i} 는 V_x, V_y 의 i 번째 요소를 나타내며 n 은 입력 벡터의 개수로 본 실험에서는 8192이다.

Step 2. 입력 벡터에 대한 공분산 행렬 (covariance matrix) C 를 구한다

$$C = \begin{bmatrix} c_1 & c_2 \\ c_3 & c_4 \end{bmatrix} \tag{6}$$

여기서,

$$c_1 = \frac{1}{n} \sum_{i=0}^{n-1} v_{x_i}^2 - m_x^2$$

$$c_2 = c_3 = \frac{1}{n} \sum_{i=0}^{n-1} v_{x_i} v_{y_i} - m_x m_y$$

$$c_4 = \frac{1}{n} \sum_{i=0}^{n-1} v_{y_i}^2 - m_y^2$$

Step 3. 가장 큰 고유치(the largest eigenvalue)를 구한다

$$\lambda_{\max} = \frac{A + \sqrt{A^2 - 4B}}{2} \quad (7)$$

여기서, $A = c_1 + c_4$, $B = c_1 c_4 - c_2 c_3$.

Step 4. 주축의 각도(θ)를 계산한다.

$$\theta = \text{atan}\left[\frac{\lambda_{\max} - c_1}{2}\right] \quad (8)$$

Step 5. 주축의 좌표를 P로 표시하면 주축상의 일차원 데이터 p_i 는 아래의 식과 같다.

$$p_i = [(v_x - m_x) \cos \theta - (v_y - m_y) \sin \theta] + m_x \quad (9)$$

여기서, $i = 0, 1, \dots, n-1$.

그림 2에서 2차원 데이터가 1차원 데이터로 변환되는 과정을 설명한다. 2차원 데이터는 주축인 P에 수직으로 투영되어 X-Y좌표에서는 각도가 θ 인 직선의 형태가 됨을 볼 수 있다.

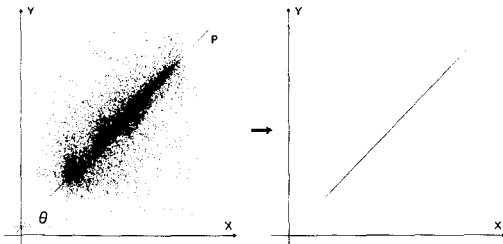


그림 2. 주성분 해석에 의한 데이터 차원감소
Fig. 2. Data dimension reduction by PCA.

(4) 최소오차경계결정 알고리즘

최소오차경계결정 알고리즘은 일차원 데이터를 두 부분으로 나눌 때 최적의 경계를 찾는 알고리즘이다. 기존의 경계결정은 2차원 데이터에 대하여 주축을 구하고 주축과 수직하며 데이터의 중점을 지나는 직선에 의하여 데이터를 두 부분으로 분할하였다^[13]. 이 방법은 계산의 시간이 빠른 반면 데이터의 분포가 균일하지 않을 때는 특성이 좋지 않다. 제안하는 방법은 차원이 1차원으로 감소된 데이터에 대하여 최적인 경계를 계산하여 분할한 다음 원래의 차원인 2차원으로 데이터를 다시 복원함으로써 위 알고리즘의 특성을 개선하였다. 본 장에서는 1차원 데이터에 대해서 최적의 경계를 결정하는 방법에 대하여 설명한다.

그림 3에서 1차원 데이터 p_i 를 두 그룹 P_a, P_b 로 나누는 경우를 예로 들면, 전체 데이터 p_i 가 최적으로 나누어지는 조건은 두 그룹의 평균오차의 합이 가장 작을 때이다. 두 그룹이 임의의 경계 $d(0 \leq d < n-1)$ 를 기준으로 왼쪽을 P_a 그룹, 오른쪽을 P_b 그룹이라 하면 분할된 각 그룹의 평균 avg_{P_a}, avg_{P_b} 는 식 (10)과 식 (11)과 같이 나타낼 수 있다.



그림 3. 일차원 데이터의 분할
Fig. 3. The divisions of one dimensional data.

$$avg_{P_a} = \frac{1}{d} \sum_{i=0}^{d-1} p_i \quad (10)$$

$$avg_{P_b} = \frac{1}{n-d} \sum_{i=d}^{n-1} p_i \quad (11)$$

다음으로 두 그룹의 오차를 각각 $D(P_a), D(P_b)$ 라고 하고 최적의 경계를 결정하기 위한 설정한 파라미터 R을 식 (14)와 같이 정의한다.

$$D(P_a) = \sum_{i=0}^{d-1} (p_i - avg_{P_a})^2 \quad (12)$$

$$D(P_b) = \sum_{i=d}^{n-1} (p_i - avg_{P_b})^2 \quad (13)$$

$$R = [D(P_a) + D(P_b)] \quad (14)$$

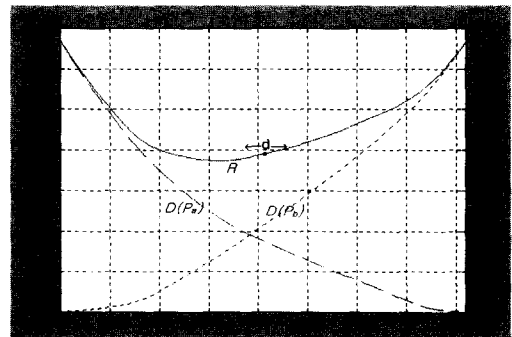


그림 4. Lenna 영상에 대한 에러그래프
Fig. 4. Error graphs for Lenna image.

그림 4는 Lenna영상에 대해서 식 (12), 식 (13), 식 (14)를 계산한 그래프이다. 그래프에서 $D(P_a)$ 에 대한 계산은 P_b 그룹 없이 P_a 그룹만 데이터를 증가시키며 오차를 계산하였으며, 반대로 $D(P_b)$ 의 계산

시에는 P_b 그룹에 대한 데이터만 증가시키며 오차를 계산하였다. 그림 4에서 두 그룹의 오차를 최소로 분할되기 위해서는 R 이 최소가 되는 점을 기준으로 분할하여야 한다. 하지만 R 이 최소가 되는 점을 계산하려고 하면 많은 $D(P_a)$ 와 $D(P_b)$ 를 계산한 다음 R 까지 계산해야 하므로 계산량이 무척 많아진다. 본 논문에서는 이 계산량을 줄이기 위해서 먼저 데이터 전체 개수의 중심점을 계산한 다음 오차가 줄어드는 방향으로 데이터의 분할점을 이동하는 방법을 적용하였다. 그림 5의 초기 d 점은 전체 데이터 개수의 중심을 나타낸다. 이 점에서 좌우로 하나씩 이동하며 R 을 계산하고, 두 방향 중 R 이 줄어드는 방향으로 데이터의 분할점을 이동한다. 여기서 R 값의 분포는 극값이 하나인 오목(concave)한 형태이므로 국부적인 최적값에 수렴할 염려는 없다. 그림 4의 R 의 경우는 분할점이 왼쪽으로 이동하는데, 분할점이 왼쪽으로 이동하면서 그때마다 R 을 계산하고 R 이 감소하는 구간 중 가장 왼쪽의 데이터의 개수를 최소오차인 분할점으로 결정한다. 이렇게 함으로서 전체적인 오차를 모두 계산하지 않고도 최적의 경계를 찾을 수 있다. 자세한 알고리즘은 다음의 4단계로 설명한다.

- Step 1. 초기 분할점의 위치 d 를 전체 데이터 개수의 중심으로 설정한다.
- Step 2. $d+1$ 에서와 $d-1$ 에서의 $D(P_a)$, $D(P_b)$, R 을 계산한 다음 R 이 감소하는 방향을 결정한다.
- Step 3. R 이 증가할 때까지 Step 2에서 결정된 방향으로 d 를 이동시키며 R 을 계산하고, R 이 증가하면 정지한다.
- Step 4. R 이 증가하는 점을 두 그룹을 분할하는 최적의 경계(d^*)로 설정한다.

위의 알고리즘에서 데이터가 하나 추가되면 그 때의 평균값과 오차를 다시 계산하여야 한다. 여기서 데이터가 추가되기 전 평균값을 알고있을 때 새로운 데이터가 추가될 때 평균값을 간단하게 계산하는 방법 공식 (15)에서 소개한다^[16].

$$avg^+ = avg^- + \frac{(data^+ - avg^-)}{cnt^- + 1} \quad (15)$$

여기서 $data^+$ 는 추가되는 데이터를, avg^- 는 데이터가 추가되기 전 평균값을, cnt^- 는 데이터가 추가되

기 전 그 그룹의 데이터 개수를 나타내며, avg^+ 는 $data^+$ 가 추가된 다음의 그룹의 평균값을 나타낸다.

그림 5는 제안한 알고리즘에 의하여 데이터가 분할되는 과정을 설명한 그림이다. (a)는 Lenna 영상에 2차원 벡터의 분포와 그의 주축을 나타낸 그림이고, (b)는 모든 데이터를 주축으로 근사한 그림이다. (c)에서는 최소오차경계 d^* 를 나타내었고, (d)는 d^* 를 기준으로 입력 데이터가 두 그룹으로 나누어지는 그림이다. 그림 6은 그림 5의 Group 1이 다시 두 그룹으로 나누어지는 과정으로 분할의 과정은 그림 5와 동일하다.

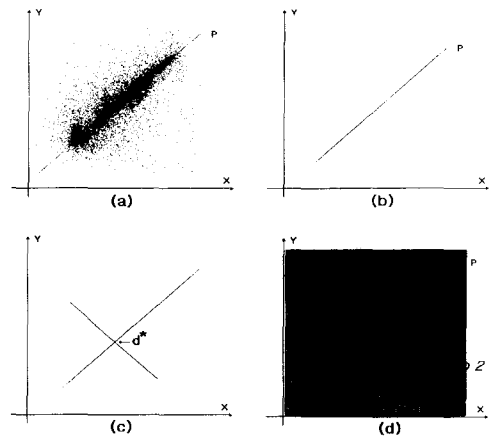


그림 5. 제안한 알고리즘에 의한 데이터 분할(1)
Fig. 5. Data divisions by the proposed algorithm(1).

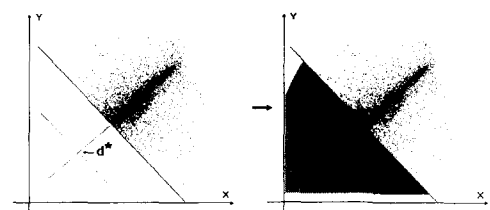


그림 6. 제안한 알고리즘에 의한 데이터 분할(2)
Fig. 6. Data divisions by the proposed algorithm(2).

(5) 코드북 생성

전체 알고리즘을 수행하고 나면 입력 벡터가 여러 개의 그룹으로 분할된다. 각 그룹에 속하는 데이터의 평균값을 그 그룹의 대표값으로 하여 코드워드로 저장하고 그룹 이름을 코드북의 인덱스로 저장하면 코드북이 생성된다. 그림 7은 최종 노드의 개수를 7개로 설정하였을 때 입력 벡터(a)와 각 데이터 그룹(b)-(h)를 나타내었고, 그림 8은 입력 벡터(a)와 생성된 7개의 코드워드 값(b)을 나타내었다. 즉, 모든 입력 벡터를

그림 8(b)의 코드워드 중 가장 근사한 코드워드의 인덱스로 나타냄으로서 데이터를 압축하게 된다.

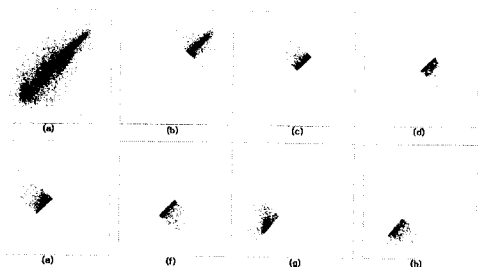


그림 7. 입력 벡터의 분할
Fig. 7. The divisions of input vectors.

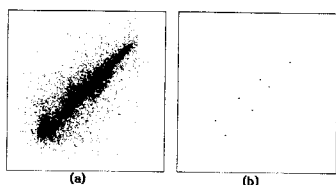


그림 8. 입력 벡터(a)와 생성된 코드워드(b)
Fig. 8. Input vectors(a) and generated code-words(b).

3. 멀티미디어 검색을 위한 특징추출

영상의 검색 중 특징에 의해서 검색하는 방법을 내용 기반 검색이라고 하며, 종류로는 색상 정보를 이용한 검색, 질감정보를 이용한 검색, 형태 정보를 이용한 검색 등이 있다. 제안하는 알고리즘은 내용기반 영상 검색을 수행하기 위한 전단계로서, 압축 영역에서 특징을 직접 추출할 수 있는 알고리즘이다^[17].

웨이브렛 변환의 가장 저역 밴드에는 영상의 대부분의 정보가 포함되어있다. 이 정보는 웨이브렛 계수를 0에서 255 사이로 정규화(normalization)하면 시각적으로 볼 수 있는데, 본 논문에서는 이 성질을 멀티미디어 검색을 위한 특징추출에 이용한다. 웨이브렛 변환된 데이터의 가장 저역 밴드의 정보를 0번부터 순서대로 번호를 부여한 다음, 그 번호가 4로 나누어 나머지가 0인 데이터만 추출한 다음 정규화 한다. 이 데이터는 크기는 작지만 원 영상과 거의 유사한 영상임을 볼 수 있다. A 를 ' $\sqrt{\text{저역 밴드의 데이터의 수}}/4$ '라 하면 특징으로 추출한 영상의 크기는 ' $A \times A$ '이다. 그림 9는 영상을 압축한 다음 압축 데이터에서 멀티미디어 정보검색을 위한 영상의 특징을 추출하는 과정을 설명한 그림이다. 입력 영상이 먼저 4-band 웨이브렛 변환된 다음, 주성분 해석과 최소오차경계결정 알고리즘

을 이용한 벡터 양자화에 의하여 압축된다. 멀티미디어 검색을 위한 영상 정보 추출은 압축된 데이터를 복원하지 않고 압축 데이터의 일부분만 추출하여 정규화하면 된다. 입력 영상이 256×256 이면 영상 정보검색을 위한 데이터의 크기는 64×64 이다. 전체 영상의 $1/8$ 크기인 이 특징데이터는 영상의 색인과 검색에 원 영상 대신 이용된다.

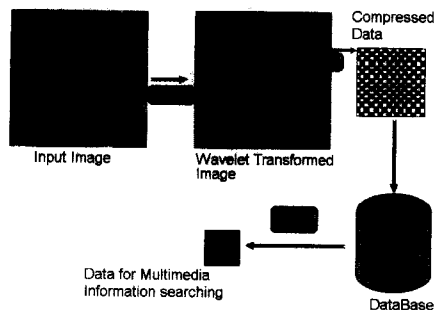


그림 9. 멀티미디어 정보검색을 위한 특징추출
Fig. 9. Feature extraction for multimedia information searching.

III. 실험결과 및 고찰

1. 웨이브렛 변환에 의한 특징추출

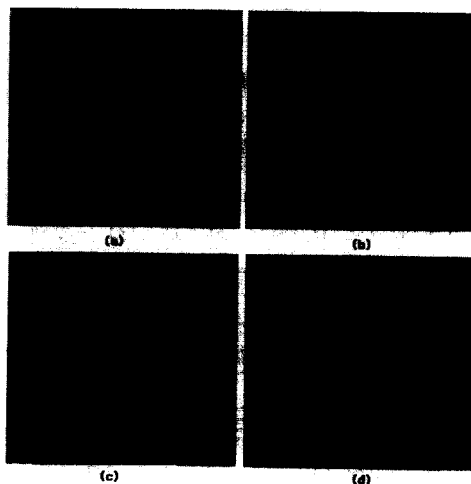


그림 10. 원영상(a,c)과 웨이브렛 변환된 영상(b,d)
Fig. 10. Input Images(a,c) and wavelet transformed images(b,d).

그림 10은 그레이 스케일(gray scale)의 256×256 Baboon 영상과 Lenna 영상, 그리고 그의 웨이브렛 변환 결과이다. (b)와 (d)는 시각적으로 보이기 위해 웨이브렛 변환된 결과를 0과 255사이로 정규화하여

그레이 스케일의 영상으로 표현하였다. 본 논문에서는 특징을 추출할 때 데이터 압축률을 높이기 위해서 가장 저역 밴드를 제외한 나머지 부밴드들의 값들은 제거하였다. 다음의 주성분 해석 단계를 위해서 웨이브렛 변환된 결과를 0과 100 사이로 정규화한다. 이상의 결과에 의하여 입력 벡터는 개수가 25%로 줄어들고 크기는 0과 100 사이로 정규화된 데이터가 된다.

2. 데이터의 차원감소와 입력 벡터의 분할

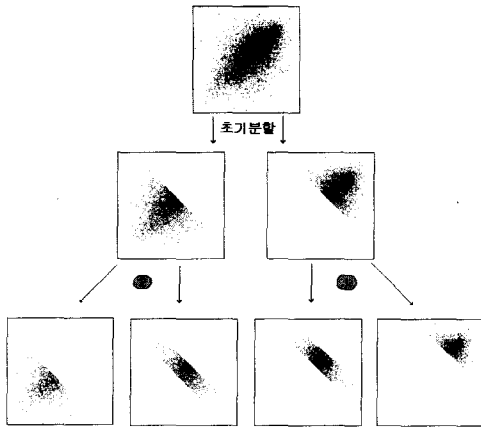


그림 11. Baboon 영상의 입력 특징 벡터의 분할
Fig. 11. The division of input vectors for Baboon image.

그림 11은 입력 벡터인 Baboon 영상의 특징 데이터를 주성분 해석을 이용하여 입력 벡터를 분할하는 과정을 설명한다. 전체 입력 벡터는 일단 2개의 그룹으로 나누어지고, 두 그룹 중 오차가 큰 그룹이 다시 반으로 나누어진다. 그림에서 숫자는 노드가 나누어지는 순서를 표시한 것이다. 각각 나누어진 그룹에 대해서도 같은 방법으로 노드를 나누어 가다가 초기에 설정한 노드개수에 이르면 노드 분할을 마친다. 표 1은 Baboon 영상과 Lenna 영상의 노드 분할 실험결과 중 covariance matrix, λ_{max} , 주축과 좌표의 X축과의 각도(θ), 최소오차경계 등 초기 노드의 분할을 위한 정보를 나타내었다. 초기 데이터에 대한 주축과 좌표의 X축과의 각도(θ)는 웨이브렛 변환의 특성에 의하여 Baboon 영상이나 Lenna 영상 모두 45° 정도로 비슷함을 볼 수 있었다. 그리고 최소오차경계의 경우도 비슷한 함을 볼 수 있는데, 여기서 D_{3838} 의 의미는 2차원 데이터가 1차원 데이터로 차원감소 한 후 일차원 데이터를 크기대로 나열한 순서가 3838번째라는 것을 나타낸다.

표 1. 초기 노드 분할을 위한 정보

Table 1. The informations for division of initial nodes.

		Baboon	Lenna
Covariance matrix	c1	245.63	280.74
	c2	140.77	216.16
	c3	140.77	216.16
	c4	248.56	282.94
λ_{max}		387.87	498.00
θ [degree]		45.30	45.15
d(최소오차경계)		D_{3838}	D_{3836}

그림 12는 Lenna 영상과 Baboon 영상의 입력 벡터를 나타내고 그림 13은 주축으로 근사시킨 결과를 나타낸다. 데이터의 주축으로 근사한 결과 두 데이터가 비슷한 형태가 됨을 볼 수 있다. 그림 14부터 그림 17까지는 Lenna와 Baboon 영상에 대하여 각각의 코드워드 개수에 의한 입력 벡터 분할 형태와 그에 대한 차원감소 결과를 나타내었다. 그림에서 회색으로 나타낸 부분은 입력벡터가 분포하는 영역을 나타낸다.

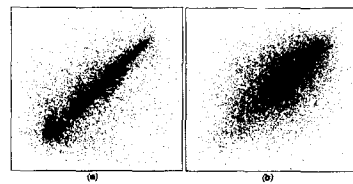


그림 12. 입력 벡터 (a) Lenna (b) Baboon
Fig. 12. Input vectors (a) Lenna (b) Baboon.

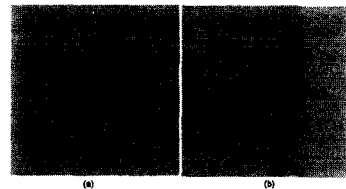


그림 13. 입력 벡터의 차원감소 결과 (a) Lenna (b) Baboon
Fig. 13. The results of dimension reduction for input vectors (a) Lenna (b) Baboon.

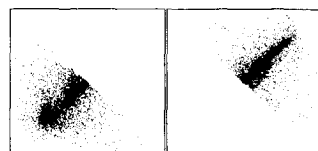


그림 14. Lenna 영상의 입력 벡터 분할 결과 (2 코드워드)
Fig. 14. The results of input vector divisions for Lenna image (2 codewords).

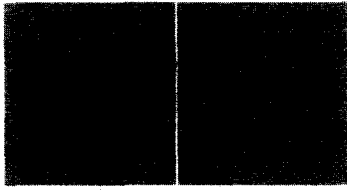


그림 15. 그림 14 데이터에 대한 차원감소 결과
Fig. 15. The results of dimensions reduction for Fig. 14.

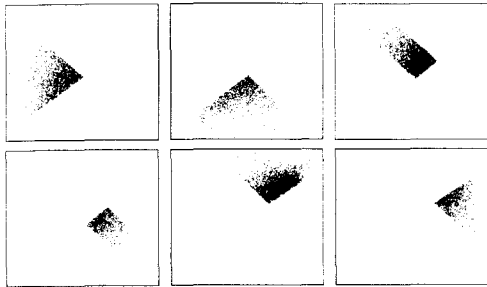


그림 16. Baboon 영상의 입력 벡터 분할 결과 (6 코드워드)
Fig. 16. The results of input vector division for Baboon image (6 codewords).

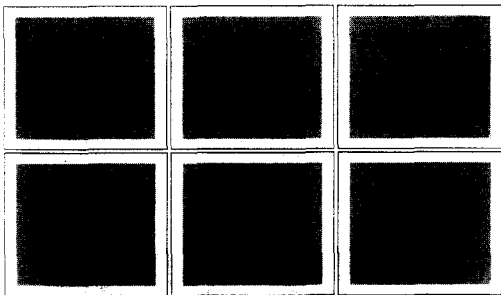


그림 17. 그림 16에 대한 차원감소 결과
Fig. 17. The results of dimension reduction for Fig. 16.

3. 영상압축 결과

그림 18은 Baboon 영상과 Lenna 영상에 대하여 코드워드 수를 달리하여 생성한 코드북으로 압축한 다음 복원한 결과를 나타낸다. (a)와 (d)는 코드워드의 수가 4일 때, (b)와 (e)는 36일 때, (c)와 (f)는 100일 때이다. 그림 19는 그림 18에 대한 각각의 코드워드 값을 0과 100사이로 정규화 해서 2차원 좌표에 나타낸 것이다.

Lenna 영상과 Baboon 영상에 코드워드 수를 설정하고 실험하였는데, Lenna 영상이 더 높은 PSNR을 보였다. 여기서 제안한 알고리즘은 Baboon 영상처럼 평범한 영상보다는 Lenna 영상처럼 데이터가 주축방

향으로 밀집된 영상에 더 효과적임을 볼 수 있었다. 표 2에서는 코드워드의 수에 따른 압축 결과를 비교하였다.

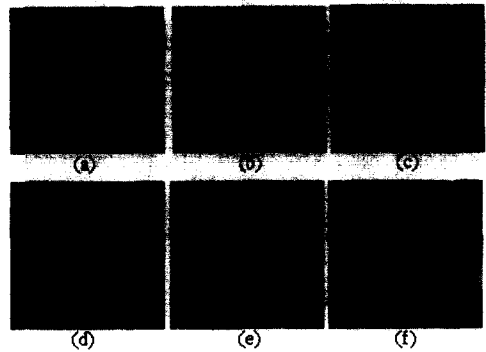


그림 18. 코드워드 개수에 따른 압축 결과
Fig. 18. The results of comparison for codeword numbers.

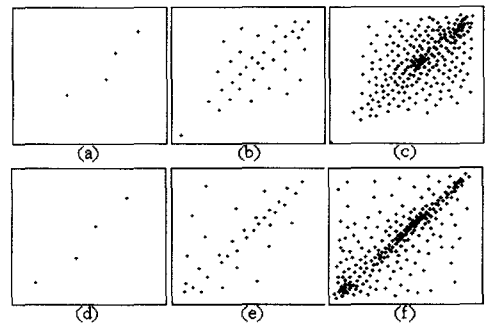


그림 19. 그림 18에 대한 코드워드
Fig. 19. Codewords for Fig. 18.

표 2. 코드워드 수에 따른 압축 결과 비교
Table 2. Comparison of compression results for codeword number.

특성	Baboon			Lenna		
	(a)	(b)	(c)	(e)	(f)	(g)
코드워드 수	4	36	225	4	36	225
PSNR[dB]	21.07	22.71	22.86	23.38	28.10	29.47

4. 기존의 알고리즘과 압축 성능 비교

SOM을 이용한 코드북 설계 실험과 비교하기 위해서 여러 가지 SOM의 크기에서 반복회수는 50000회로 설정하였고, 영상은 Lenna를 사용하여 학습시켰다. 초기 학습률은 0.4로 설정하여 반복회수에 비례하여 8×10^{-6} 까지 감소하였다. 학습반경은 반복회수가 증가할수록 줄어들며, 7000회 이상부터는 0으로 하여 승리노드의 웨이트값만 갱신하였다.

SOM의 경우 Pentium II-233Hz 사용시 144 코드워드의 코드북 설계를 위한 학습시간이 3시간 정도 소요되었으며, 제안한 알고리즘의 경우 37초의 짧은 시간에 코드북을 생성하였다. 그리고 압축 성능인 PSNR의 경우 SOM은 29.38 [dB] 이었고 제안한 방법은 29.30 [dB] 정도로 나타났다. 즉, SOM을 이용한 방법과 비교해서 제안한 알고리즘은 코드북 생성 시간은 3시간 정도 단축시키며 비슷한 PSNR 성능을 나타내었다. SOM의 경우 학습률과 반복, 반복회수 등을 조정하면서 실험을 계속한다면 보다 더 좋은 특성이 나올 수 있지만 결과를 얻을 때까지 무척 많은 시간이 요구된다. 즉, 세 개의 파라미터를 조정하며 적당한 특성의 코드북을 설계하는 과정을 오랜 시간 반복해야 하며, 한 번 설정한 파라미터에 대한 결과를 확인할 때까지도 시간이 무척 걸린다. 그리고 일반적인 클러스터링 알고리즘인 K-mean의 경우 수행 속도가 빠르고 코드워드의 수가 적을 때 좋은 특성을 보이지만 초기노드를 랜덤으로 설정하는 과정에서 전체적인 특성이 변하며, 코드워드 수가 늘어날수록 다른 방법에 비하여 그 특성이 저조함을 볼 수 있었다. 반면 제안하는 알고리즘이나 논문 [13]은 하나의 파라미터인 코드워드의 개수만 설정하면 원하는 코드북을 단 한번에 설계할 수 있으므로 파라미터 설정에 대한 어려움이 없다. 그리고 제안한 알고리즘의 경우 코드워드의 수가 작을 때는 압축성능이 가장 뛰어나므로 멀티미디어 정보 중 영상에 대한 압축알고리즘으로 적당하리라 사료된다.

표 3에서는 256×256 Lenna 영상에 대하여 SOM, K-mean, 논문 [13], 제안한 알고리즘을 이용하여 압축한 다음 복원한 영상의 PSNR 성능과 처리시간을 나타내었다. 코드워드의 개수가 적을 때는 제안한 방법과 K-mean 알고리즘이 가장 우수하였고, 코드워드 수가 늘어나면 세 방법은 비슷한 결과를 보이지만 K-mean 알고리즘은 특성이 저조하였다. 처리시간의 경우 SOM을 이용한 방법은 코드워드의 수가 늘어날수록 기하급수적으로 증가하지만 나머지 세 알고리즘은 많이 증가하지 않음을 볼 수 있다.

실험결과, 코드워드 수가 적을 때는 제안한 방법을, 코드워드 수가 늘어나면 논문 [13]의 방법이나 제안한 방법을 적용하는 것이 바람직 할 것이다. 그림 20에서는 3개의 알고리즘에 대하여 압축성능을 그래프로 나타내었다.

표 3. 기존 알고리즘과의 비교

Table 3. Comparison of the PSNR performances and processing time.

크기	성능(PSNR) [dB]			
	SOM	K-mean	논문 [13]	제안한 방법
4	23.35	23.37	22.11	23.38
36	28.32	28.11	27.80	28.10
100	29.20	29.00	29.09	29.12
225	29.51	29.25	29.47	29.47

크기	처리시간 [sec]			
	SOM	K-mean	논문 [13]	제안한 방법
4	350	4	6	26
36	3000	4	6	30
100	7110	4	6	33
225	14760	4	7	42

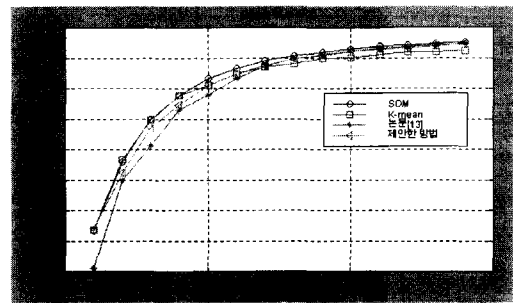


그림 20. 기존의 알고리즘과 성능 비교 그래프
Fig. 20. Comparison graphs of PSNR performance.

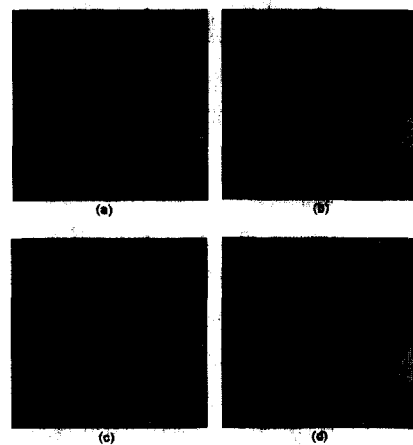


그림 21. 코드워드 개수가 9일 때 복원영상 비교 (a) SOM 이용한 방법 (b) K-mean 이용한 방법 (c) 논문 [13] (d) 제안한 방법
Fig. 21. Comparison of reconstruction image in 9 codeword (a) SOM (b) K-mean (c) Ref. [13] (d) Proposed algorithm.

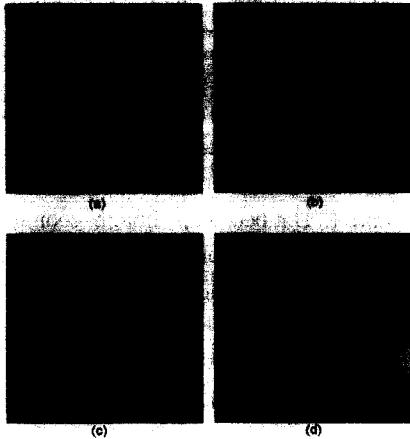


그림 22. 코드워드 개수가 225일 때 복원영상 비교 (a) SOM 이용한 방법 (b) K-mean 이용한 방법 (c) 논문 [13] (d) 제안한 방법
 Fig. 22. Comparison of reconstruction image in 225 codeword (a) SOM (b) K-mean (c) Ref. [13] (d) Proposed algorithm.

5. 멀티미디어 검색을 위한 특징추출

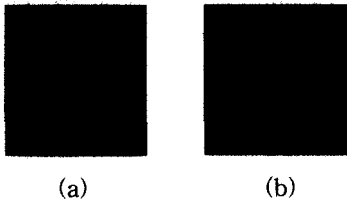


그림 23. 압축 데이터에서 추출한 영상 정보검색 특징 데이터 (a) Lenna (b) Baboon
 Fig. 23. Feature data which are extracted from compressed data for image information searching (a) Lenna (b) Baboon

그림 23은 Lenna 영상과 Baboon 영상에 대하여 압축 영상에서 특징 데이터를 추출한 것을 나타낸다. 원영상은 256×256이고 이를 검색하기 위해 추출한 특징의 영상 크기는 64×64이다. 크기는 작지만 원영상의 대부분의 정보를 가지고 있으며, 압축 영역에서 데이터 복원 없이 바로 추출할 수 있다. 지금까지는 영상 압축의 주된 관심의 압축률의 정도였다. 하지만 생성되는 정보의 양이 기하급수적으로 증가하고, 그 중에서 영상이 차지하는 비율이 점점 높아지는 현재에는 영상정보의 검색은 정보 수집하고 이용하는데 없어서는 안 될 중요한 수단이다. 이를 위해서는 영상이나 여러 가지 정보를 큰 비율로 압축하는 방법을 연구하는 것도 중요하지만, 압축된 정보를 효율적으로 관리하고 이용 가능한 압축 방법에 대한 연구가 더 중

요하다고 할 수 있다.

IV. 결 론

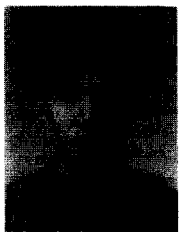
본 논문에서는 입력 벡터의 통계학적 특성을 이용하여 주축을 계산하고, 그것을 이용하여 최적의 코드북을 설계하는 알고리즘을 제안하였다. 기존의 코드북 설계 알고리즘 중 널리 사용되는 SOM을 이용한 벡터 양자화와 비교해 볼 때 파라미터 설정이 간단하고 계산시간이 무척 빠르다. SOM의 경우 학습률과 갱신반경, 반복회수를 조정하면 여러 번 실험해야만 원하는 결과를 얻을 수 있다. 하지만 제안한 방법의 파라미터는 코드워드의 개수 즉, 압축률이 입력 파라미터로 파라미터 설정이 매우 간단할 뿐만 아니라 코드워드의 수가 적은 부분에서는 압축성능이 더 우수함을 볼 수 있었다. 물론 SOM의 경우 여러 번의 실험을 하면 보다 더 우수한 성능을 얻을 수 있지만 압축비율과 입력 영상이 달라지면 매번 파라미터 설정을 다시 해야할 뿐만 아니라 한 번의 실험에도 많은 시간이 소요되는 단점이 있다. 그리고 제안한 알고리즘의 큰 장점은 웨이블릿 변환의 특성 중 가장 저역의 부밴드에 중요한 정보가 집중되는 성질을 이용하여 영상을 복원하지 않고 압축 영역에서 영상의 정보를 바로 추출할 수 있다는 것이다. 만약 제안한 방법으로 영상이 압축되어있다면 영상 정보를 검색하기 위해 압축된 영상을 복원하는 시간을 없애므로서 멀티미디어 검색 중 영상의 정보검색 시간을 단축시킬 수 있다. 앞으로의 연구 방향은 압축 영역에서 추출한 정보를 이용하여 효과적인 영상 정보를 검색하는 알고리즘을 개발하는 것이다. 본 연구와 유사한 분야의 연구가 활발해져서 우수한 특성의 알고리즘이 많이 발표된다면 영화 속의 한 장면을 단순한 스케치 등으로 검색이 가능한 날이 머지 않아 올 것이다.

참 고 문 헌

[1] Borko Furht, Joshua Greenberg, and Raymond Westwater, "Motion estimation algorithms for video compression," Kluwer Academic Publishers, pp.1-6, 1997.
 [2] Ahmed K. Elmagarmid, Haito Jiang, Abdelsam A. Helal, Anupam Joshi, and Magdy Ahmed, "Video database systems,

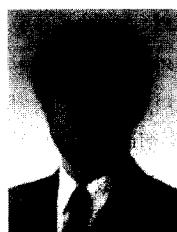
- issues, products, and applications," Kluwer Academic Publishers, pp.1-56, 1997.
- [3] Chee Sun Won, Dong Kwon Park, and Seong-Joon Yoo, "Extraction image feature for MPEG-2 compressed stream," Storage and retrieval for image and video database VI, vol.3312, February 1998.
- [4] 전승원, 최강선, 설상훈, 고성제, "특징 추출에 의한 고속 영상 검색 시스템," 신호처리합동학술대회, 제 11권 제 1호, pp.227-230, 1998
- [5] 김소연, 노용만, "정합추적 알고리즘에 의한 영상검색에 관한 연구," 신호처리합동학술대회, 제 11권 제 1호, pp.231-234, 1998
- [6] Allen Gersho and Robert M. Grey, "Vector quantization and signal compression," Kluwer Academic Publishers, 1991.
- [7] Bhaskar Ramamurthi and Allen Gersho, "Classified vector Quantization of image," IEEE Trans. on Communications, vol. Com-34 No.11, 1986.
- [8] G. Poggi, "Application of the Kohonen algorithm in vector quantization," European Trans. on Telecommunication and Related Techniques 6(2), pp.191-202, 1995.
- [9] T. Kohonen, "The self-organizing map," Proc. IEEE 78, pp.1464-1480, 1990.
- [10] T. M. Martinez, S. G. Bekovich, and K. J. Schulten, "'Neural-gas' network for vector quantization and its application to time-series prediction," IEEE Trans. Neural Networks 4, pp.558-568, 1993.
- [11] D. T. Pham and E. J. Bayro-Correchango, "Self-organizing neural network based pattern clustering method with fuzzy outputs." Pattern recognition, Vol.27, No.8, pp.1103-1110, 1994.
- [12] C. K. Cheong and K. Aizawa, T. Saito, and I. Daubechies, "Image coding using wavelet transform," IEEE Trans. on Image Processing, No.2, 1992.
- [13] 서석배, 강대성, 김대진, "개선된 주성분 해석을 이용한 벡터 양자화," 신호처리학회 합동학술대회, 제 11권 1호, pp.785-788, 1998.
- [14] H. S. Malvar and D. H. Staelin. "The LOT:transform coding without blocking artifacts," IEEE Trans. on Acoust Speech Signal Processing, Vol.37, No.4, pp.553-559, Apr. 1989.
- [15] T. Jarske, P. Haavisto and I. Defe'e, "Post Filtering Methods for Reducing Blocking artifacts from Coded Images." IEEE Int. Conf. Consumer Electronics, pp.218-219, 1994.
- [16] Richard O. Duda and Peter E. Hart, "Pattern classification and scene analysis," Wiley-Interscience Publishers, pp.221-228, 1973.
- [17] Y. Gong, C. H. Chuan, and G. Xiaoyi, "Image indexing and retrieval based on color histograms," Multimedia Tools and Applications, Vol.3, No.3, pp.133-156, 1996.

— 저 자 소 개 —



徐錫培(正會員)

1997년 2월 부경대학교 제어계측공학과 졸업(학사). 1999년 2월 동아대학교 전자공학과 졸업(석사). 1999년~현재 동아대학교 컴퓨터공학과 박사과정. 주관심분야는 신경망, 영상압축, 패턴인식 등



姜大星(正會員)

1984년 2월 경북대학교 전자공학과(학사). 1991년 8월 Texas A&M Univ., Electrical Eng.(석사). 1994년 5월 Texas A&M Univ., Electrical Eng.(박사). 1984년 3월~1989년 6월 국방과학연구소 연구원. 1994년 7월~1995년 8월 한국전자통신연구소 선임연구원. 1995년 9월~현재 동아대학교 전기전자컴퓨터공학부 조교수. 주관심분야는 영상처리, 패턴인식, 영상코딩, 통신시스템 등

金大鎮(正會員) 第34卷C編 第7號 參照

현재 포항공과대학교 컴퓨터공학과 부교수