

論文98-35S-6-9

웨이브릿 패킷 영역에서의 방향성 벡터양자화

(Directional Vector Quantization on the Wavelet Packet Domain)

姜 東 郁 *

(Dong Wook Kang)

요 약

웨이브릿 패킷을 이용한 방향성 벡터양자화 기법을 제안하였다. 영상에 대한 웨이브릿 패킷 계수들을 대응되는 방향에 따라 9개의 에지로 분할한 후, 국지적으로 지배적인 에지들을 부호화하여 전송한다. 에지의 방향은 조건부 부호책 교환이 있는 가변길이부호화를 통하여 부호화하며, 에지의 내용은 벡터양자화와 그 인덱스의 가변길이부호화를 이용하여 부호화하였다. 제안된 부호화기법은 기존의 여러 영상압축기법에 비하여 PSNR 성능이 우수하며, 특히 0.1~0.3 bpp의 낮은 부호율의 경우 상대적으로 더욱 우수하다. 제안된 알고리즘은 인간시각 특성상 가장 중요한 에지 정보들을 보전함으로써, 낮은 부호율에서도 우수한 주관적 화질을 나타낸다.

Abstract

A method is proposed for directional vector quantization using the wavelet packets. After partitioning the wavelet packet coefficients into 9 edges according to the corresponding directions, it encodes and transmits locally dominant edges. The directions of the edges are encoded with a variable length coding and conditional switching of codebooks, while the contents of them with the vector quantization followed by the variable length index coding. The proposed algorithm is superior to various conventional image coding algorithms in the sense of PSNR, which is relatively more significant at very low bit rate such as 0.1~0.3 bpp. As the proposed algorithm preserves the edges which is the most important for the human visual system, it also provides the reconstructed images of good subjective quality even at very low bit rate.

I. 서 론

영상신호를 전송하거나 데이터베이스에 저장하기 위하여 영상압축은 필수적이다. 영상을 효과적으로 압축하기 위한 방법에는 예측부호화와 변환부호화가 두 축을 이룬다. 지금까지 변환부호화라 하면 주로 이산여현변환(DCT) 영역에서 그 변환계수를 부호화하는 것으로 인식되어 왔다. 이유는 영상의 경우 DCT의 변

환특성이 매우 우수할 뿐만 아니라 고속 알고리즘이 구현됨으로써 매우 효율적이고도 빠르게 영상을 변환할 수 있기 때문이다. 즉, DCT는 JPEG이나 MPEG-1, MPEG-2, H.261 등에서 영상의 표준압축기법으로 이용되고 있는 변환기법이다.

기존의 표준영상압축기법들에서는 DCT 계수를 양자화하기 위하여 스칼라양자화기법을 사용하고 있는데, 그 이유는 신호변환에 의한 탈상관화성질(decorrelation property)에 따라 변환계수들 사이의 상관성이 크지 않게 되어 스칼라양자화하여도 벡터양자화하는 것에 비하여 그 양자화 성능이 크게 뒤지지 않을 뿐만 아니라, 벡터양자화에 비하여 부호화 과정이 복잡하지

* 正會員, 國民大學校 電子工學部
(School of Elec. Eng., Kookmin University)
接受日字: 1997年9月24日, 수정완료일: 1998年5月25日

않고 부호화하는데 걸리는 시간이 현저히 짧아서, 매우 실용적이기 때문이다.

부호화 성능을 향상시키거나 복호기를 단순화하기 위한 방안으로서 DCT 영역에서의 벡터양자화는 유망한 방안으로서 인식되어 왔다^[1,2]. 그러나 벡터양자화 과정에서 계산상의 부담(복잡도)을 줄이기 위해서는 2차원 DCT 변환계수영역을 적당한 차원의 벡터들로 분할하여 product-VQ를 행할 필요가 있다. 벡터양자화기법에서 부호화 성능은 양자화하고자 하는 벡터의 구성방법에 좌우되기 때문에 최적의 벡터 구성 방법을 찾고자 하는 시도들이 계속되고 있으며, 이는 앞으로 계속 연구되어야 할 분야이다. 지금까지의 연구 결과들을 살펴보면, 영상의 시각적, 통계적 특성을 벡터 분할에 이용함으로써, 스칼라양자화에 비하여 벡터양자화를 통하여 우수한 성능을 얻을 수 있음을 알 수 있다.

DCT와 같은 구획변환을 이용하는 경우 발생하는 문제점으로는 복원영상에서 발생하는 구획화현상(blocking artifact)이 있다. 특히 낮은 비트율로 부호화하는 경우 구획화현상이 두드러지게 나타나는데, 이는 시각적으로 매우 불쾌감을 불러일으키기 때문에 주관적 화질을 급격히 떨어뜨린다. 또한 이러한 구획화현상은 변환계수를 스칼라양자화하느냐 벡터양자화하느냐에 상관없이 나타난다. 따라서 초저전송율을 목표로 하는 영상의 부호화 기법으로서 DCT가 부적합하다.

최근에 웨이브릿 변환(WT)을 이용하여 영상을 변환함으로써 구획변환에 의한 구획화현상을 방지하여 매우 우수한 주관적 화질을 나타내는 연구결과들이 속속 발표되고 있다^[3-6]. 즉, WT는 DCT와는 달리 변환을 위하여 영상을 구획화하는 과정이 필요하지 않으므로 복원영상에서의 구획화현상은 근본적으로 발생하지 않는다.

WT를 이용하는 영상부호화의 경우, 일반적으로 입력영상을 몇 단계의 WT를 통하여 대역통과영상들로 분해한다. 이때 각단계의 분해는 입력영상에 대한 2차원 여파와 표본감축의 결합으로 구성된다. 또한 분해는 단지 저역성분영상에만 순환적으로 적용되는 것이 일반적이다^[3]. 그 이유는 고역성분영상을 분해함으로써 오히려 부호화 효율을 떨어뜨리는 경우가 종종 발생하기 때문이다. 그러나 최근의 몇몇 연구 결과에 따르면, 특히 다수의 강한 에지를 포함하고 있는 영상의

경우, 고역성분영상을 분해하는 것이 훨씬 효율적임이 알려졌다^[5,8]. 이처럼 저역성분영상 뿐만 아니라 고역성분영상에 대한 분해를 포함하는 웨이브릿 변환을, 일반화된 웨이브릿 분해로서, 웨이브릿 패킷(wavelet packets)이라 한다.

웨이브릿 계수의 부호화 방법으로는, 초기의 DCT 변환부호화에서와 마찬가지로, 스칼라양자화에 이어지는 가변길이부호화의 구조를 이용하는 것이 일반적인 경향이다. 이에 따라 효율적인 가변길이부호화가 가능하도록 하기 위한 양자화계급의 재배열 문제가 주로 연구되어 왔는데, 이에 대한 자세한 설명과 여러 가지 알고리즘들은 참고문헌^[3]에 자세히 설명되어져 있다. 최근에 웨이브릿 변환계수에 대한 벡터양자화의 가능성과 효율성에 대한 연구결과가 발표되었다^[6,7]. 정보이론(Information theory)으로부터, 벡터양자화는 스칼라양자화에 비하여 우수한 부호화 성능을 나타낼 것으로 예측할 수 있으나, 사실은 이와 달리 벡터양자화가 스칼라양자화에 비하여 더 나은 부호화 왜곡 특성을 보여주지 못하고 있는 실정이다. 특히 스칼라양자화된 웨이브릿 계수의 zerotree를 가변길이로 부호화하는 Shapiro의 알고리즘(WT-EZW)^[4]에 비하여, 지금까지 알려진 대부분의 벡터양자화 기법들은 오히려 열등한 부호화 성능을 보이고 있다.

본 논문에서는 웨이브릿 패킷으로 표현되는 대역분할영상으로부터 방향성영상을 구성하고 이를 벡터양자화함으로써, 방향성영상부호화(directional image coding)의 장점을 충분히 살려서 주관적, 객관적 부호화 성능이 우수하고, 부호시의 환경이나 조건에 따라서 양자화 파라미터를 가변시킴으로써 비교적 간단하게 가변적인 부호율을 달성할 수 있는 웨이브릿 패킷 영역에서의 벡터양자화 알고리즘을 소개하고자 한다.

II. 부호화 알고리즘

1. 웨이브릿 패킷의 구성

웨이브릿 패킷은 STFT(short time Fourier transform)와 웨이브릿 변환의 중간적인 특성을 나타낸다. 그림 1은 1차원 신호의 시간-주파수 tiling 구조를 통하여 웨이브릿 패킷과 STFT 및 웨이브릿 변환의 특성을 비교하여 보여주고 있다. 그림에서 알 수 있는 바와 같이 STFT는 전체 주파수 영역을 동일한 대역폭을 갖는 몇 개의 대역성분으로 채운다. 이때 각

각의 대역성분들은 시간영역에서 동일한 시간간격의 정의구역을 갖는 필터를 사용하여 얻게 된다. 이에 반하여 웨이브릿 변환은 저역성분은 보다 세밀하게 고역성분은 보다 큼직한 대역폭을 갖도록 전체 주파수 영역을 분해한다. 그리고 고역성분의 추출을 위한 시간영역에서의 정의구역은 저역성분 추출을 위한 정의구역보다 짧도록 하여 신호의 비정상성에 적절히 대응하도록 하고 있다. 한편 웨이브릿 패킷은 고역성분이라 하더라도 필요에 따라서 저역성분 만큼 세밀한 조각으로 주파수 영역을 채운다. 그러나 각 대역성분으로 분해하기 위해 사용하는 필터의 정의구역은 웨이브릿 변환과 동일하다.

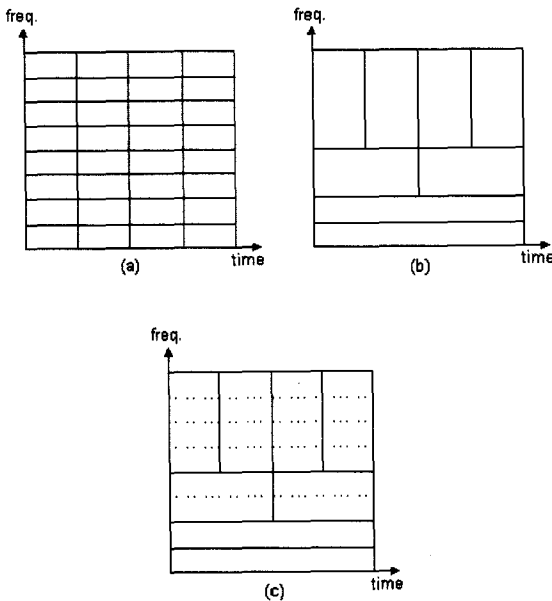


그림 1. 1차원 신호의 시간-주파수 영역 타일링 (tiling): (a) STFT, (b) 웨이브릿 변환, (c) 웨이브릿 패킷

Fig. 1. Tiling of the time-frequencies of the one-dimensional signal: (a) STFT, (b) wavelet transform, and (c) wavelet packets.

본 논문에서는 그림 1(c)에서 보인 차원 웨이브릿 패킷을 수직 및 수평 방향으로 각각 적용하여 총 64 개의 2차원 웨이브릿 패킷으로 영상을 분해하였다. 웨이브릿 패킷의 구성 예로서 Bank 영상을 분해한 결과를 그림 2에 보였다. 영상부호화에서 일반적으로 이용되어온 웨이브릿 변환(wavelet transform) 대신 웨이브릿 패킷(wavelet packet)으로 영상을 분해한

목적은 각 대역성분 즉 각 웨이브릿 계수들에 대하여 동일한 격자구조를 이용한 표본줄임을 행하기 위함이다. 이렇게 함으로써, 공간 좌표상에서 동일한 격자상의 웨이브릿 계수들은 영상내의 동일한 위치의 국지적인 영상정보를 표현하게 된다. 따라서 영상의 내용에 따라서 매우 큰 상관성을 보이는 계수들이 존재하게 되며, 이 계수들을 묶어서 벡터양자화함으로써 스칼라 양자화보다 우수한 부호화 특성을 기대할 수 있게 된다.

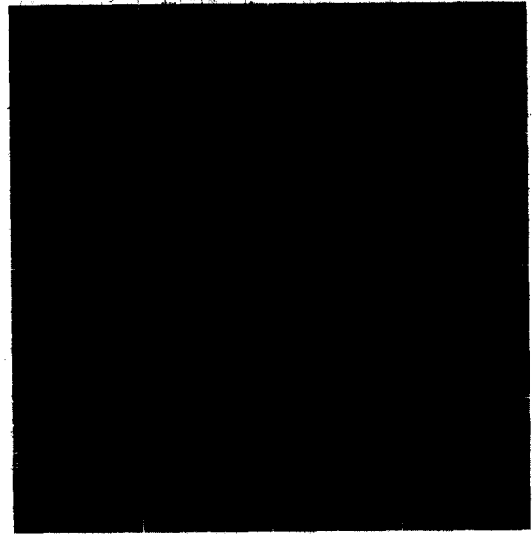


그림 2. 웨이브릿 패킷의 예. Bank영상을 64개의 대역영상으로 분해하였다. 좌측하단 영상이 최저역영상이고, 우측으로 갈수록 수평 방향의 고역성분영상을, 위로 갈수록 수직 방향의 고역성분영상을 각각 나타내고 있다.

Fig. 2. Example of wavelet packets. Bank is decomposed into 64 subband images. The one at the left-bottom is the lowest-band image, and horizontally higher-band images are displayed from left to right while vertically higher-band images from bottom to top.

2. 방향성영상들의 구성

영상 에지(edge) 정보들은 그 에지의 방향에 따라 결정되는 한정된 몇 개의 웨이브릿 패킷들에 의해 표현된다. 그리고 이 계수들 사이에는 통계적 상관성이 매우 크다. 벡터양자화는 벡터 원소들 사이의 선형 상관성(linear correlation) 뿐만 아니라 비선형 또는 통계적 상관성이 큰 경우에도 스칼라양자화에 비하여 우수한 부호화특성을 나타낸다는 것은 자명하다. 그림 3은 통계적 상관성이 큰 계수들을 그룹화함으로써 웨이

브릿 패킷으로부터 9개의 방향성영상을 구성하기 위한 2차원 그룹화 패턴을 보여주고 있다.

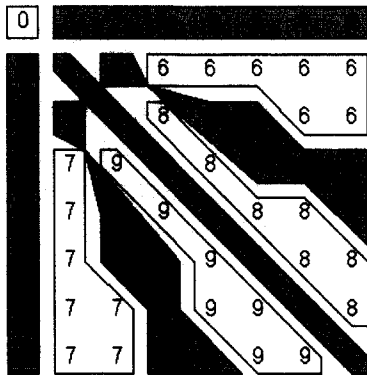


그림 3. 2차원 영상의 웨이브릿 패킷의 방향성 그룹화
Fig. 3. Directional grouping of the wavelet packets of the two-dimensional image.

통계적 상관성이 매우 큰 패킷 계수들을 그룹화 (grouping)하면, 각 계수 그룹은 하나의 방향성영상을 표현한다^[7]. 구체적으로 먼저 3단계의 웨이브릿 분해를 통하여 64개의 2차원 웨이브릿 패킷들을 구성한 후, 이 패킷들을 9개의 그룹으로 그룹화함으로써 방향성 영상을 구성한다. 그룹화하는 과정에서, 각 웨이브릿 패킷에서 동일한 공간 좌표상의 계수들이 원영상의 동일한 국지적 정보를 표현하고 있다는 점을 이용하여 웨이브릿 패킷을 그룹화하게 되며, 각각의 그룹은 각각 한 방향의 에지를 표현하게 된다. 이 원리는 DCT 계수를 이용한 방향성 영상의 구성에서도 이용된 바 있다^[2].

그림 2에서 64개의 대역성분영상들이 등간격으로 나뉘어진 대역들을 표현하기 때문에 일견 STFT와 동일한 특성을 나타낸다고 생각할 수 있으나, STFT와는 달리 변환 필터의 정의구역이 대역에 따라 다르기 때문에 STFT로부터 얻게 되는 부영상과는 본질적으로 다른 영상들이다. 즉, 저역성분 웨이브릿 패킷을 얻기 위한 필터의 정의구역은 넓고, 고역성분은 좁다. 한편 영상복원을 위한 필터가 분해를 위하여 사용한 필터와 동일하므로, 각각의 패킷에서 발생한 잡음이 복원 영상에서 미치는 범위 또한 저역성분과 고역성분에 따라서 차이가 남을 의미하기도 한다.

3. 방향성영상의 부호화

영상의 직류(DC) 성분은 최저역 패킷 계수에 대한

DPCM을 이용하여 부호화한다. DPCM을 위하여 수평방향의 1차원 예측을 행하였으며, 예측오차를 균일양자화(uniform quantization)하였다. 균일양자화기의 양자화 계단폭(quantization step size)은 복원영상의 목표화질(target quality: 출력 PSNR(peak signal-to-noise ratio))에 따라서 다음과 같은 식을 가변적으로 적용하여 결정하였다.

$$\Delta = \alpha \cdot \sqrt{\frac{255^2}{10 \cdot (S/M) \cdot 7/10}} \tag{1}$$

여기서 (S/M),는 복원영상의 목표화질로서 부호화할 때 외부 입력을 통하여 결정하는 상수이고, α 는 비례 상수이다. 그리고 Δ 는 그에 따른 양자화 계단폭이다. 만일 α 가 $2\sqrt{3}$ 이면, 균일양자화잡음과 계단폭 사이의 잘 알려진 관계식, 즉 $\sigma^2 = \Delta^2/12$ 에 의한 예측 PSNR을 목표화질로 삼는 것을 의미한다. 그러나 본 논문의 경우에서는 목표화질에 따라 엄밀하게 계단폭을 결정하는 것이 중요한 것이 아니라 PSNR에 대한 계단폭 결정의 비례관계를 설정하는 것이 목적이므로 계산상 가장 간단한 상수인 $\alpha=1$ 을 이용하였다.

영상의 교류(AC) 정보의 전달은 국지적으로 지배적인 에지(locally dominant edge) 또는 중요하다고 판단되는 에지(significant edge)들을 전송함으로써 달성된다. 이처럼 방향성 에지들을 전송하는 것은 인간의 시각인식과정을 효과적으로 이용하는 한 방법으로서 잘 알려져 있다^{[7] [9]}. 어떤 에지의 중요성(significance)을 판단하기 위하여, 부호화 과정에서 외부입력으로 주어질 목표화질에 따라 가변적으로 결정되는 문턱값과 비교하여 그 에지의 에너지가 문턱값보다 크면 중요한 에지로 판단한다. 즉, 에지의 에너지 $E(e)$ 에 따라서 부호화할 가치가 있음을 나타내는 지시함수(indicator function)를 다음과 같이 결정한다:

$$I(e) = \begin{cases} 1 & \text{if } E(e) > T \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

여기서, $E(e) = \sum_{i=0}^{D-1} e_i^2$ 이고, $e = [e_0, e_1, \dots, e_{D-1}]$ 는 에지를 나타내는 분할된 패킷 계수로 구성된 D-차원 벡터이다. 또 문턱값 T 는 다음과 같이 결정된다:

$$T = \beta \cdot D \cdot \frac{255^2}{10 \cdot (S/M) \cdot 7/10} \tag{3}$$

여기서 β 는 비례상수이고, D 는 에지 벡터의 차원 (dimension)이다. 식 (3)에서 우변의 세 번째 항은 식 (1)에서 나타난 양자화 계단폭의 제곱에 해당된다. 이는 DC 계수와 AC 계수를 목표화질에 따라 같은 비율로 섬세함(fineness)을 제어하면서 양자화하기 위해서 정한 결과이다. 각 벡터마다 그 차원에 비례하는 문턱값을 사용하는 이유는 양자화 과정에서 벡터 내의 각각의 패킷 계수에 대한 문턱값의 영향이 서로 같도록 하기 위함이다. 또한 비례상수 β 에 따라서 DC와 AC 계수 사이의 중요성을 인위적으로 정해줄 수 있는데, 본 논문에서는 $\beta=1$ 을 이용하였으나, 시각적으로 최적의 β 를 찾는 것이 앞으로 더욱 연구되어야 할 내용 중의 하나이다.

어떤 에지가 의미있는 에지로 판단되면, 반복적 근사법(successive approximation)에 따라 벡터양자화한다. 반복적 근사란, 높은 목표화질의 경우 어떤 에지가 실제로 양자화되어 전송된다 하여도 제한된 크기의 불완전한 부호책을 사용함으로써 인하여 잔류오차신호의 전력이 너무 커서 목표화질을 만족하지 못하게 되는데, 이 경우 같은 에지에 대하여 벡터양자화를 다시 수행하여 양자화잡음을 더욱 낮추어 가는 방법을 말한다^[2]. 이때 매 반복마다 잔류오차신호의 특성에 최적화된 서로 다른 부호책을 사용하여 벡터양자화할 수도 있으나, 그러면 결과적으로 참조해야 하는 부호책의 크기를 증대시켜야 된다는 단점이 있다. 연구 결과에 따르면, 반복 단계에 상관없이 동일한 부호책을 사용하여 양자화하여도 양자화 이득이 약간 떨어지지만 전체적으로는 보다 유사한 벡터로의 근사가 가능하다^{[2][7]}. 따라서 본 연구에서는 하나의 부호책을 사용하여 각 반복 단계에서의 양자화를 수행하는 경우를 다루었다. 하나의 에지를 반복적으로 근사하다가 양자화이득이 일정 수준에 미달하게 되면 반복적 근사를 멈추게 된다. 양자화이득이란 에지를 양자화하여 전송함에 따라 줄어든 잔류오차신호의 전력을 말한다. 양자화이득을 계산하여 양자화 인덱스를 실제로 전송할지를 다시 한번 판단하는 이유는 벡터양자화를 위한 부호책의 크기가 작아서 벡터양자화에 의한 이득이 거의 없음에도 불구하고 양자화함으로써 발생하는 부호율의 급격한 증가를 억제하기 위한 방안이다. 이때 문턱값은 (3)식에서 사용한 문턱값을 같이 이용하였다.

각 방향성 에지의 계수 분포특성과 시각적 중요성이 서로 다르므로, 이를 고려하기 위하여, 각각의 에지 방

향마다 서로 다른 크기를 갖는 부호책을 설계하여 벡터양자화에 사용한다. 일반적으로 저역 성분이 고역 성분보다 시각적으로 더 중요하다. 즉, 저역 성분에서의 양자화잡음이 고역 성분에서의 그것에 비하여 눈에 더 거슬린다. 따라서 저역 성분을 더 섬세하게 양자화하는 것이 주관적 화질의 면에서 유리하다. 본 논문에서는 계수의 분포특성은 각 에지마다 분산만이 다를 뿐 분포의 형태는 동일한 것으로 고려하였다. 이때 계수들의 분산은 혼련열을 이용하여 실험적으로 추정하였다. 결과적으로 시각적 중요성에 대한 고려는 부호책 크기 결정을 위한 비트 할당 과정에서 서로 다른 가중치 할당으로 실현하였다. 한편, 각 에지 벡터에서 가장 큰 에너지를 나타내는 원소는 벡터내의 첫 번째 최저역 패킷 성분이므로, 이 성분을 기준으로하여 제 6,7,8,9 에지 벡터들에는 가중치 1을, 제 4,5,6 에지 벡터들에는 가중치 4를, 제 1,2,3 에지 벡터들에는 가중치 16을 각각 할당하였다. 각 방향성 에지에 대한 계수의 추정된 분산과 비트 할당을 위한 가중치 및 이에 따라 결정된 부호어의 길이를 표 1에 요약하였다.

표 1. 각 방향성 에지에 대한 계수의 분산, 가중치 및 결과적인 부호책의 크기.

Table 1. The variances of the coefficients, weighting factors and the resulting sizes of the codebooks for the directional edges.

에지 방향	1	2	3	4	5	6	7	8	9
계수의 분산	235.	237.	73.2	14.1	16.6	19.0	20.4	0.43	0.62
가중치	1	5	0	7	1	8	7	12	27
부호어 길이 (비트)	16	16	16	4	4	1	1	1	1
부호어 길이 (비트)	10	10	9	7	7	7	7	4	4

부호어길이 b_i 에 대하여 부호책의 크기 N_i 는 $N_i = 2^{b_i}$ 로서 결정된다. 계수의 분산과 가중치에 따라 부호어의 길이를 결정하는 방법은 다음과 같다^[10]:

$$b_i = \bar{b} + \frac{1}{2} \log_2 \frac{\sigma_i^2}{\rho^2} + \frac{1}{2} \log_2 \frac{w_i}{W}, \quad (4)$$

여기서 \bar{b} 는 평균부호어길이이고, σ_i^2 는 각 에지벡터들의 성분 계수의 분산, w_i 는 가중치를 각각 나타낸다. 그리고, ρ^2 와 W 는 각각 σ_i^2 와 w_i , $i=1, \dots, 9$ 의 기하평균들이다.

$$\rho^2 = \left(\prod_{i=1}^9 \sigma_i^2 \right)^{1/9} \quad (5)$$

$$W = \left(\prod_{i=1}^9 w_i \right)^{1/9} \quad (6)$$

위의 과정을 통하여 양자화된 에지에 대하여 그 에지의 방향(direction) 인덱스와 내용(content)에 대한 양자화 인덱스를 가변길이부호화하여 전송함으로써 부호화 과정이 완성된다. 그림 4에 지금까지 설명한 방향성 에지의 부호화 과정을 개념적으로 도식화하여 나타내었다. 벡터양자화 인덱스를 가변길이부호화하는 이유는, 국지적으로 지배적인 에지는 그 에너지의 변동이 완만하여 에지의 방향을 따라서 군집하여(burst) 나타날 뿐만 아니라 그 분포가 특정 값에 편중하여 나타나기 때문에, 인덱스를 가변길이부호화함으로써 비트율을 상당히 떨어뜨릴 수 있기 때문이다.

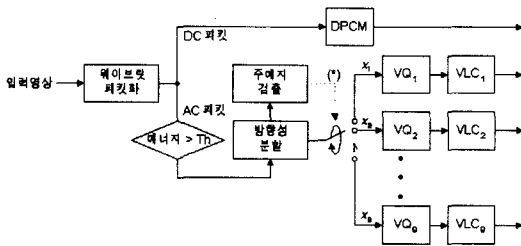


그림 4. 웨이브릿 패킷을 이용한 방향성 벡터양자화의 구획도. (*)는 부호기 출력버퍼의 상태에 따라서 양자화를 중단하기 위한 제어신호.

Fig. 4. Block diagram of the directional vector quantizer using the wavelet packets. (*) represents the control signal to stop the encoding procedure according to the status of the output buffer of the encoder.

본 논문에서 제안하고 있는 부호화 알고리즘을 제2세대 부호화의 하나인 방향성 영상부호화 기법과 비교하여 보면, 제안 알고리즘은 수평, 수직 방향으로 각각 8:1로 표본줄임된 방향성 영상들에 대한 방향성 벡터 양자화에 해당된다고 하겠다. 예를 들어 512×512 크기의 입력영상을 부호화하는 경우, 웨이브릿 패킷으로 분해된 영상들은 64개의 64×64 대역성분영상이 된다. 이러한 패킷 영상들을 그룹화하여 64×64 크기의 격자에서 각 격자점마다 9개의 에지 벡터들을 구성함으로써 총 9×64×64개의 에지 벡터를 구성하게 된다.

이 벡터들을 전송하는 방법은 다음과 같다. 64×64

크기의 격자에서 격자점을 raster scan 형식으로 주사해 가면서 각 격자점마다 9개의 에지 벡터 중에서 미리 정한 문턱값보다 큰 에너지를 갖는 에지 벡터들을 전송한다. 이 경우 주사순서상의 격자점에서 전송할 가치가 있는 중요한 에지가 더 이상 존재하지 않게 되면, 에지의 방향 인덱스에 대한 부호로서 그 위치에서의 부호화가 끝났음을 알리는 특수부호를 전송함으로써 정해진 다음 위치에서의 양자화 과정으로 이동하게 된다. 참고로 에지의 방향을 표현하기 위한 부호어에는 9개의 서로 다른 에지 방향을 나타내기 위한 9개의 부호어와 위에서 설명한 양자화의 끝을 알리기 위한 1개의 특수부호어를 포함하여 총 10개이다.

III. 실험결과

본 논문에서 제안한 웨이브릿 패킷 영역에서의 벡터 양자화 알고리즘을 적용하여 크기가 512×512이고 256 계급을 갖는 휘도영상 Lena, Bridge, F16에 대하여 부호화하였다(그림 5). 에지 내용의 벡터양자화를 위한 부호책을 설계하기 위해서 같은 크기와 같은 계급을 갖는 12장의 휘도영상들을 이용하였는데, 부호화하고자 하는 실험영상은 부호책 설계에서 제외한 outside training을 실시하였다.

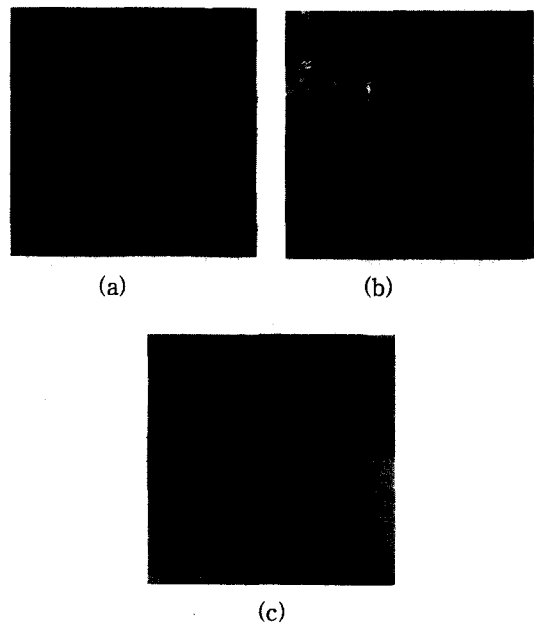


그림 5. 실험 영상: (a) Lena, (b) Bridge, (c) F16.
Fig. 5. Simulated Images: (a) Lena, (b) Bridge, (c) F16.

Lena 영상에 대해서는 웨이브릿 계수를 스칼라양자화한 Shapiro의 embedded zerotree(WT-EZW)^[4]와의 PSNR 성능을 비교하여 그림 6에 나타내었다. 또한 DCT 계수를 벡터양자화하는 Kim 등의 DCT-CVQ^[1], 강동욱 등의 DCT-PEVQ^[2], 그리고 웨이브릿 계수의 벡터양자화를 사용한 김철우의 WT-VQ^[6]의 성능도 같은 그림에서 비교하였다. 실험결과로부터 본 논문에서 제안한 알고리즘이 전반적으로 가장 우수한 성능을 나타내고 있음을 알 수 있다. 그리고 특히 0.3 bpp 이하의 비트율에서 다른 부호화 알고리즘과의 성능 차이가 두드러진다. 예를 들어, 0.29 bpp에서 부호화한 경우, DCT-CVQ(0.316 bpp)와 DCT-PEVQ(0.316 bpp)에 비하여는 각각 2.6 dB와 1.6 dB, WT-VQ(0.25 bpp)와 WT-EZW(0.25 bpp)에 비하여는 각각 5.5 dB와 0.8 dB 정도 우수하다.

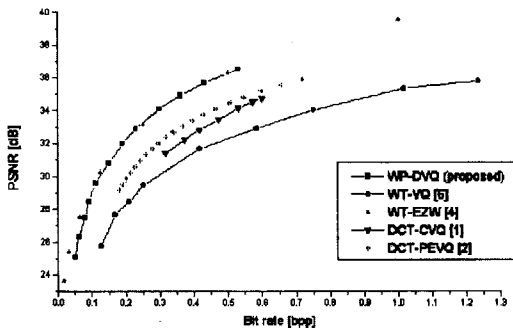


그림 6. Lena 영상에 대한 부호율-PSNR 특성 비교
Fig. 6. Comparison of the rate-PSNR characteristics for encoding the Lena.

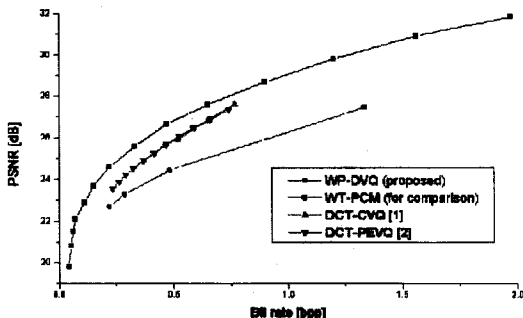


그림 7. Bridge 영상에 대한 부호율-PSNR 특성 비교
Fig. 7. Comparison of the rate-PSNR characteristics for encoding the Bridge.

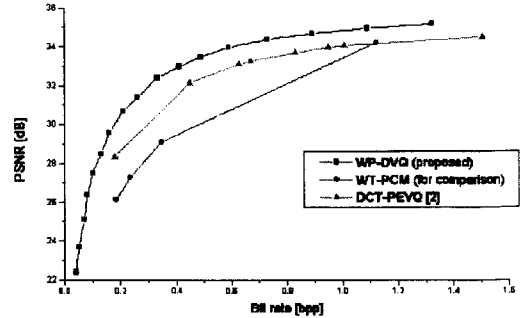


그림 8. F16 영상에 대한 부호율-PSNR 특성 비교
Fig. 8. Comparison of the rate-PSNR characteristics for encoding the F16.

Bridge 영상과 F16 영상을 이용한 실험 결과는 각각 그림 7과 8에 보였다. WT-EZW와의 PSNR 성능비교는 보고된 자료가 없어서 불가능하였지만, 그밖의 발표된 결과가 있는 부호화 알고리즘과 성능을 비교하였다. 이와 아울러 웨이브릿 변환 계수를 PCM으로 스칼라양자화하는 경우의 부호율-PSNR 곡선을 같은 그림에서 나타내어 참고하도록 하였다. 이 영상의 경우에도 본 논문에서 제안한 알고리즘이 다른 알고리즘들에 비하여 상당히 우수한 성능을 나타냄을 확인할 수 있다.

복원영상의 주관적 화질에 대한 평가를 위하여 복원 영상을 보였다. 웨이브릿 변환부호화의 경우, Lena나 F16 영상과 같이 비교적 단순한 영상에서는 약 34 dB, Bridge 영상과 같이 매우 복잡한 영상에서는 약 28 dB 이상의 PSNR을 나타내는 경우, 벡터양자화나 스칼라양자화에 상관없이 어느 알고리즘이나 매우 우수한 주관적 화질을 나타낸다. 그러나 PSNR이 점차로 낮아지는 경우, 스칼라양자화를 이용하는 부호화 알고리즘이 벡터양자화를 이용한 기법에 비하여 같은 PSNR에 대해서 주관적 화질은 약간 우수하게 인식된다. 그러나 본 논문에서 제안된 알고리즘의 경우, 매우 낮은 비트율로 부호화한 경우에도 상당히 우수한 주관적 화질을 보여주고 있는데, 이는 복원영상에 인간의 시각특성에서 가장 중요한 에지 또는 컨투어(contour)가 일정 이상 잘 보존되고 있기 때문이다. 그림 9와 10은 0.18 bpp로 부호화된 Lena와 0.32 bpp로 부호화된 Bridge의 복원영상과 그 확대화면을 보여준다. 강한 에지 근처에서 미약하지만 ringing이 발생하여 눈에 약간 거슬리나, JPEG 등에서 보이는 구획화현상

이 나타나지 않으며, 낮은 부호율의 경우 발생하는 전 반적인 blurring을 인식하지 못할 수준이어서, 제안하고 있는 알고리즘이 매우 우수한 주관적 화질을 제공하고 있음을 확인할 수 있다.

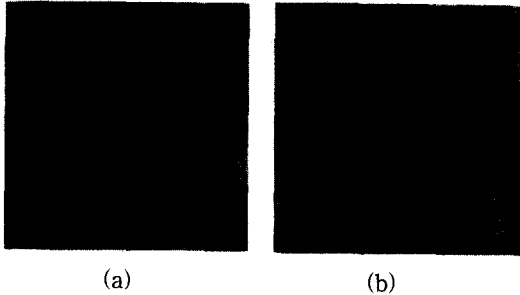


그림 9. Lena 복원영상과 그 부분확대: (a) 복원영상 (0.18 bpp, 32.0 dB); (b) 수직, 수평 각 4배로 부분확대한 영상

Fig. 9. Reconstructed Lena image and its magnified version. (a) Reconstructed image(0.18 bpp, 32.0 dB), (b) its part magnified by 4 both horizontally and vertically.

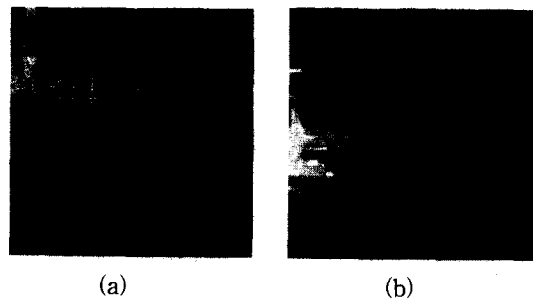


그림 10. Bridge 복원영상과 그 부분확대: (a) 복원영상(0.32 bpp, 25.5 dB); (b) 수직, 수평 각 4배로 부분확대한 영상

Fig. 10. Reconstructed Bridge image and its magnified version. (a) Reconstructed image(0.32 bpp, 25.5 dB), (b) its part magnified by 4 both horizontally and vertically.

IV. 결 론

이 논문에서는 웨이브릿 패킷 영역에서의 벡터양자화를 통하여 우수한 부호화 결과를 보여주고 있다. 특히, 제안된 영상압축 알고리즘은 에지에 기초한 부호화를 행함으로써 매우 낮은 부호율에서도 복원영상이 상당히 우수한 주관적 화질을 나타낸다. 이는 DCT 등을 이용하는 경우 발생하는 구획화현상이 없을 뿐만

아니라 인간시각특성상 중요한 에지 정보가 보존되기 때문에 가능하다. 또한 객관적 성능평가에 있어서도, DCT나 웨이브릿 변환 영역에서 계수를 벡터양자화하는 기존의 부호화 알고리즘들에 비하여 상대적으로 우수한 부호율-왜곡 특성을 나타낸다. 제안된 알고리즘은 벡터양자화를 행하기 위한 벡터의 차원(크기)이 크지 않고, 부호책의 크기도 비교적 작기 때문에 시스템의 복잡도 면에서도 우수하다. 앞으로 주어진 부호화 조건(사용가능한 비트수)에서 보다 높은 PSNR과 보다 우수한 주관적 화질을 제공할 수 있는 최적의 부호화 알고리즘을 구현하기 위해서, 엔트로피 제한 벡터양자화 기법 등 보다 고도의 벡터양자화 기법을 적용하는 등의 추가적인 연구가 필요하다.

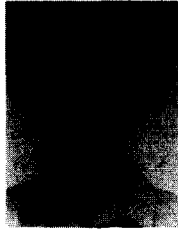
참 고 문 헌

- [1] J. W. Kim and S. U. Lee, "A transform domain classified vector quantizer for image coding," *IEEE Trans. Circuit and Syst.-Video Technology*, vol. CASVT-2, pp. 3-14, Mar. 1992.
- [2] 강동욱, 송준석, 이충웅, "영상부호화를 위한 DCT계수의 시간적 분해 및 순차적 주에지 벡터양자화," *대한전자공학회논문지 제32권 B편 제 1호*, pp. 64-72. 1995년 1월.
- [3] M. Vetterli and J. Kovacevic, *Wavelet and Subband Coding*, Prentice-Hall, 1995.
- [4] Jerome M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, No. 12, pp. 3445-3462, Dec. 1993.
- [5] T. Otake, K. Fukuda, and A. Kawanaka, "Adaptive wavelet packet image coding with zerotree structure," *Proc. IWISP'96*, pp. 281-284, 1996, Manchester, U.K.
- [6] 김철우, 웨이브릿 변환영역에서의 벡터양자화 연구, *서울대학교 대학원 공학석사학위논문*, 1992.
- [7] D. W. Kang, "Directional image coding on wavelet transform domain," *Proc. IWISP'96*, pp. 235-238, 1996, Manchester, U.K.
- [8] K. Ramchandran and M. Vetterli, "Best

wavelet packet bases in a rate-distortion sense," *IEEE Trans. Image Processing*, vol. 2, No. 2, pp. 160-175, April 1993.

- [9] M. Kunt, A. Ikononopoulos, and M. Kocher, "Second generation image coding techniques," *Proc. of IEEE*, vol. 73, pp. 549-574, Apr. 1985.
- [10] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, Kluwer Academic Pub., 1992.

저 자 소 개



姜 東 郁(正會員)

1963년 3월 7일생. 1986년 서울대학교 전자공학과 공학사. 1988년 서울대학교 대학원 전자공학과 공학석사. 1995년 서울대학교 대학원 전자공학과 공학박사. 1995년 ~ 현재 국민대학교 전자공학부 조교수. 주관심분야는 영상신

호처리, 멀티미디어 통신, 디지털텔레비전 등임.