

## 〈主 題〉

## 웨이블릿 변환을 이용한 영상압축

민병석 · 김승종 · 정제창

(한양대학교 대학원 전자통신학과)

## □차 례□

- |                              |                                    |
|------------------------------|------------------------------------|
| I. 서 론                       | IV. 시각 특성을 이용한 실험적 웨이블릿 기반 영상압축 기법 |
| II. 웨이블릿 이론, 다해상도 분해 및 필터 뱅크 | V. 실험 결과 및 고찰                      |
| III. 최근의 웨이블릿 기반 영상압축 기법     | VI. 결 론                            |

## I. 서 론

최근 영상, 비디오, 오디오 신호 처리 기술의 발달로 멀티미디어 통신의 서비스가 국내외에서 상용화되고 있으며 멀티미디어 관련 기술의 상업적 이용이 커다란 관심을 끌고 있다. 통신 채널(channel)의 대역폭(bandwidth)과 멀티미디어 응용분야에 따라 다양한 질(quality)의 정보를 제공하고자 하는 스케일러빌리티(scalability)의 개념에도 불구하고, 멀티미디어 정보 중 영상정보의 대용량은 시스템의 효율을 저해하는 요인으로 작용하고 있다. 대용량의 영상정보를 고능률로 압축하면서 응용 분야에 따른 요구조건을 만족하기 위한 방법으로 JPEG, MPEG-1, MPEG2, H.261과 H.263 등 영상부호화의 표준이 제정되었다. 또한, 차세대 영상부호화 방식으로 정지영상은 JPEG-2000을 중심으로 동영상은 MPEG-4를 중심으로 국제 표준의 제정이 진행되고 있으며 각국의 표준 및 기술 선점을 위한 연구가 활발히 진행되고 있다.

기존의 표준이외에 다른 방식의 영상부호화들도 연구되고 있으나 대부분 자연 영상의 계층적 분류에 따른 몇몇 모델에 기반한 것으로 볼 수 있고, 그 모델에 의해 특성화된 상호간의 의존성을 이용하는 것으로도 생각할 수 있다. 전형적인 변환과 부대역(subband)부호화 알고리즘은 영상을 통계적으로 서로 다른 협대역의 신호원으로서 모델링했다. 비록, VQ(vector quantization)[3]가 블록(block)간의 연관성

에 대한 모델은 제공하지 못했지만 기본적인 VQ 알고리즘은 영상 데이터의 블록간의 통계적 상관관계에 기반하고 있다.

다양하고 복잡한 영상부호화 알고리즘들이 블록간의 연관성 모델에 기반하여 제시되어 왔다. 즉, FSVQ[4], LOT(Lapped Orthogonal Transform), classified VQ[6-7], 영상 분할 모델 등이 그것이다. 이러한 접근 방법들의 성능은 양자화 전략(즉 밴드간 최적의 비트 할당[8], VQ[9-10], 최적의 비획일 스퀈라 양자화[11], trellis-coded 양자화[12] 등)을 주의깊게 고려해서 최적화되었다.

1980년대 후반, 신호 해석의 방법으로써 웨이블릿(wavelet)이 소개된 이후 영상부호화에 있어서 웨이블릿의 응용에 대한 연구가 활발히 진행되고 있다 [1][2][8][16-18]. 두 채널(dyadic) 웨이블릿 변환에 의해서 얻어진 계층적 신호들은 영상의 통계적 특성을 이용하였고 또한, HVS(human visual system)에 잘 맞는 양자화기를 설계하는 데 편리한 구조를 제공하였다.

웨이블릿 변환은 신호를 다해상도(multiresolution)로 분해하며, 입력 신호는 서로 다른 해상도로 표현하기 위한 부 신호들과 나머지 신호들을 표현하기 위한 세부 신호들로 나누어진다. 영상이나 음성 등의 신호들은 에너지 분포가 주파수 영역에서 관찰할 때 고르지 못한 분포를 가지고 있다. 이로 인해 신호를 분해하는 것이 필요하게 되었으며, 기본적으로 도입

된 개념이 스펙트럼을 부밴드로 분할하여 각각의 부대역을 독립적으로 목적에 맞도록 처리하는 M 대역 분해이다. 신호를 다해상도로 분해하는 것은 두 가지 방법이 있다. 하나는 M개의 동일한 대역(equal bandwidth)으로 분할하는 방법이고 다른 하나는 dyadic 트리 구조로의 분해인데 웨이블릿 변환이 이에 속한다. 웨이블릿으로 변환된 영상신호는 서로 다른 주파수 특징을 갖고 이들간에는 어느 정도의 상관관계가 존재한다. 웨이블릿 변환을 이용하여 영상을 다해상도로 분해한 후 영상을 압축하는 것은 기존의 DCT를 이용한 방법보다 압축률을 높일 수 있고 블록킹현상이 일어나지 않아 인간의 시각에 보다 적절한 압축방법이다[8][27].

인간의 시각은 세밀한 부분보다 불연속성이나 갑작스런 밝기의 변화와 같이 해상도가 낮은 부분의 영상에 대해 더 민감한 반응을 보인다. 웨이블릿 변환으로 분해된 부대역 영상들은 이러한 인간의 시각적 특성을 잘 반영하고 있어 영상을 효율적으로 압축할 수 있다. 즉, 영상신호를 웨이블릿 변환을 이용하여 다해상도로 분해하고 인간시각 특성을 고려하여 추출된 에지정보를 이용하여 분해된 부밴드 영상을 양자화한 후 양자화된 계수들의 통계적 특성과 상관관계를 이용하면 효율적으로 압축할 수 있다.

이와 같은 바탕에서 최근 웨이블릿을 이용한 효과적인 영상압축 기법이 많은 관심을 끌고 있다. 본 논문에서는 최근 MPEG-4에서도 표준의 일부로 채택될 전망이다 웨이블릿 기반의 압축기법에 대해서 개관하고자 한다.

본 논문의 구성은 다음과 같다. 2장에서는 푸리에 변환과 웨이블릿 변환과의 연관성 및 웨이블릿 변환과 영상의 다해상도 분해에 대해 살펴본다. 또한, 필터뱅크와 이산 웨이블릿 변환과의 관계 및 필터뱅크의 설계에 대해 다룬다. 3장에서는 최근에 제안된 웨이블릿 기반 영상압축 알고리즘에 대해 다루고 4장에서는 시각특성을 이용한 실험적 웨이블릿 기반 영상압축 기법에 대해 설명한다. 5장에서는 4장에서 설명한 압축기법의 성능평가 및 고찰에 대해 다루고 6장에서는 결론을 맺는다.

## 1. 웨이블릿 이론, 다해상도 분해 및 필터뱅크

### 1. 이산 푸리에 변환, Short-time Fourier 변환과 웨이블릿 변환

영상 부호화를 위한 신호 처리는 응용분야에 따른 영상의 화질과 정보를 유지하면서 최소의 정보를 전송하는데 있다. 일반적으로 정지영상의 부호화는 공간적 용장성(redundancy)을 최소화시키기 위한 처리와 공간상에서 구조적으로 분해되거나 변환을 적용해서 대역별로 분해된 영상정보의 중요성에 따라 비트할당 등으로 구성된다[5][16].

영상신호의 분해를 위해 이산 직교 변환(discrete orthogonal transform)이 변환 및 역변환의 연산 효율성 때문에 널리 사용되고 있다. 신호처리에 사용되는 직교변환은 입력신호와 유한 차원(finite dimension)의 주어진 기저(basis)와의 내적(inner product) 연산으로 이루어진다. 두 개의 이산 신호(discrete time signal)들을 수열  $x(n), y(n) \ n \in \mathbb{Z}$ 으로 표현할 때, 수식의 간결한 표현을 위해 행렬(matrix)과 벡터(vector)로 표시하면, 신호는 열벡터(column vector)로 표현된다.  $x(n), y(n) \ n \in \mathbb{Z}$ 일 때, 내적은 식(1)과 같고 H는 Hermitian을 나타낸다.

$$\begin{aligned} \langle x(n), y(n) \rangle &= \sum_{n=-\infty}^{\infty} x(n)y^H(n) \\ &= \mathbf{y}^H \mathbf{x} \end{aligned} \tag{1}$$

이산 신호 처리에서 신호내의 주파수 특성을 파악하기 위한 식(2)의 이산 푸리에 변환은, 복소 지수 함수  $e^{j\omega n}$ 를 원시(prototype)함수로 해서 주파수 스케일링(scaling)을 적용한 기저를 사용한다.

$$\begin{aligned} X(\omega) &= \langle x(n), e^{-j\omega n} \rangle = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} \\ x(n) &= 1/2\pi \int_0^{2\pi} X(\omega)e^{j\omega n} d\omega \end{aligned} \tag{2}$$

이산 푸리에 변환은 시간-주파수간의 불확정성 원리(uncertainty principle)를 따르므로, 시간영역에서 넓은 버팀(support)을 갖는  $x(n)$ 이 주파수 영역에서 좁은 버팀을 갖는  $X(\omega)$ 로 표현되어, 신호의 국부적(local) 성질을 분석하기 곤란하다. 예를 들면, 단일 주파수 성분을 갖는  $x(n) = e^{j\omega_0 n}$ 신호는 시간 영역에서 무한(infinite) 버팀을 갖지만, 주파수 영역에서는  $X(\omega) = 2\pi\delta(\omega - \omega_0)$ ,  $0 \leq \omega \leq 2\pi$ 으로 극단적인 국부적 표현을 갖는다. 이러한 성질은 특히 시변(time

varying) 주파수 성질을 갖는 비정상적(nonstationary) 신호를 해석하는데 제약을 준다. 따라서, 시변 주파수 특성을 갖는 신호를 해석하기 위해서는 시간과 주파수를 동시에 표현하는 방법이 필요하다.

Gabor에 의해서 착안된 STFT(short time Fourier Transform)는, 기저가 주파수의 스케일링만을 갖는 푸리에 변환과는 달리 시간 천이와 주파수의 스케일링을 갖는 기저를 사용함으로써, 시간과 주파수 특성을 동시에 표현할 수 있다. 식(3)은 이산신호에 대한 STFT를 표현하고 있다[17][19].

$$X_{STFT}(\omega, m) = \sum_{n=-\infty}^{\infty} w(n-m)x(n)e^{-j\omega n}$$

$$x(n) = 1/2\pi \int_0^{2\pi} \sum_{m=-\infty}^{\infty} X_{STFT}(\omega, m)w^*(n-m)e^{j\omega n} d\omega \quad (3)$$

$w(n)$  : 창함수(window function) ,  $\sum_m |w(m)|^2 = 1$

$x(n)$ 에 시간축으로 일정하게 이동하는 시간 윈도우(time window)  $w(n)$ 을 적용하고, 푸리에 변환을 수행함으로써, 신호에 대한 스펙트로그램(spectrogram)이라고 하는 시간-주파수 표현을 얻을 수 있다[18]. 시간축으로 이동하는 윈도우에 의해 시간 국부성을 확보하고, 이것을 푸리에 변환해서 주파수 특성을 파악함으로써 시변 주파수 특성을 해석할 수 있다. Fourier 변환에서와 같이 STFT는 윈도우에 의해 결정되는 시간-주파수간의 국부성이 상충관계(trade-off)를 갖는다. 시간 영역에서 버팀을 작게 해서 해상도(resolution)를 높이면 주파수 영역에서의 해상도는 낮아지고, 주파수 영역에서의 해상도를 높이면 시간영역에서의 해상도가 낮아진다. 따라서, 분석하고자 하는 신호의 성질에 따라 시간-주파수간의 해상도를 결정하는 윈도우를 설계해야 한다. 그림 1의 (b)는 STFT의 시간-주파수 평면에서의 격자를 나타내었으며, 그림 1의 (a)에서 격자에 해당하는 STFT 커

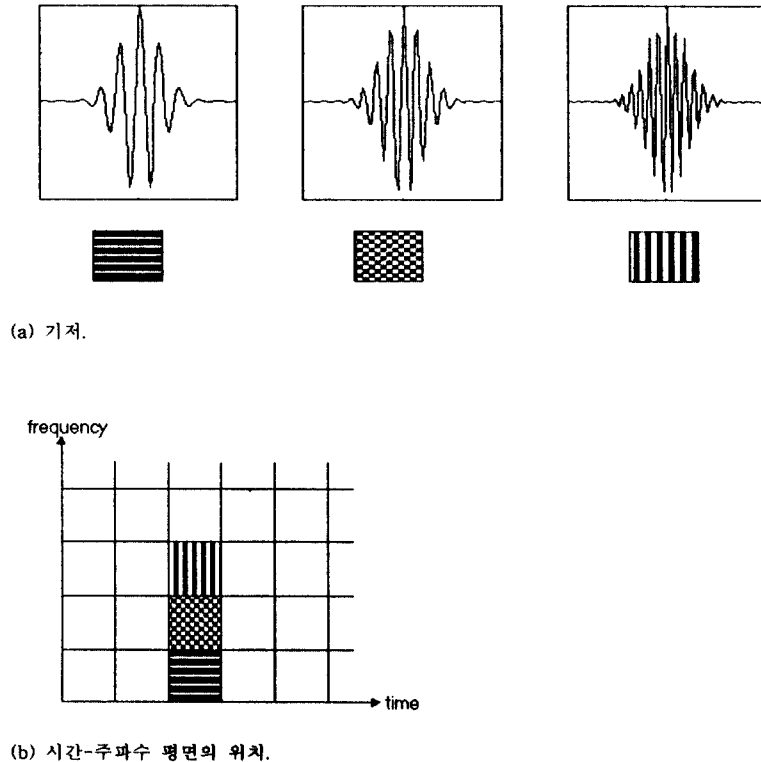


그림 1. STFT의 기저와 시간-주파수 해상도.  
 (a) 기저. (b) 시간-주파수 평면의 위치.  
 Fig. 1. Basis functions and time-frequency resolution of STFT.  
 (a) Typical basis. (b) Coverage of the time-frequency plane.

널을 시간영역에서 나타내었다.

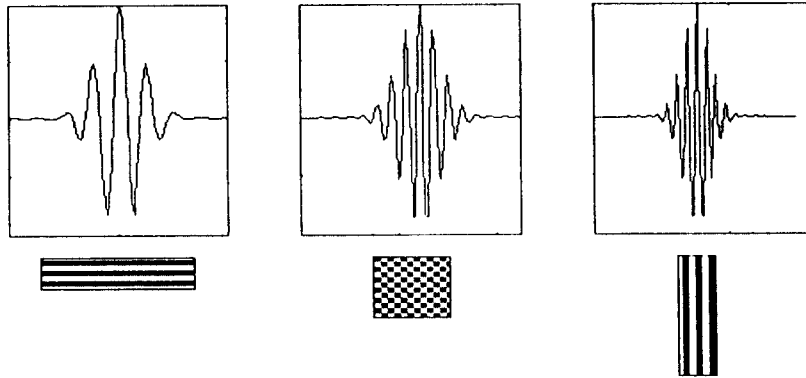
직교(orthogonal) 웨이블릿 변환은 모 웨이블릿(mother wavelet)  $\psi(t)$ 를 스케일링(scaling)과 천이(translation)시켜 생성한 웨이블릿-기저  $\psi_{ab}(t)$ 을 이용한다.

$$\psi_{ab}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \quad a \in \mathbb{R}^+, \quad b \in \mathbb{R} \quad (4)$$

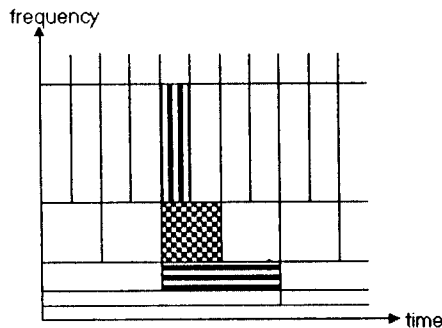
$$X(a, b) = \int_{-\infty}^{\infty} \psi_{ab}(t)x(t) dt$$

신호처리에 주로 응용되는 웨이블릿은, 주파수 영

역에서 주파수 국부성을 갖는 대역통과 필터(band pass filter)의 특성을 나타내고, 시간영역에서는 옹골버팀(compactly support)를 가지므로서 시간 윈도우와 같은 시간 국부성을 내포한다. 식(4)에서 변수  $b$ 는 시간 윈도우의 시간축에서의 이동을 의미하고 변수  $a$ 는 웨이블릿의 스케일링을 나타내는데,  $a$ 가 감소하면 웨이블릿의 시간 국부성은 낮아지는 반면 주파수 국부성은 높아진다.  $a, b$ 를 축으로 2차원 평면에 웨이블릿 변환 값을 계조도나 색깔로 표현한 것을 스케일로그램(scalogram)이라고 하며, STFT의 스펙트로그램과 유사한 해석을 하므로써 신호의 시간-주파수 특징을 나타낼 수 있다.



(a) 기저.



(b) 시간-주파수 평면의 위치.

그림 2. 웨이블릿 변환과 시간-주파수 해상도.

(a) 기저. (b) 시간-주파수 평면의 위치.

Fig. 2. Basis functions and time-frequency resolution of wavelet transform.

(a) Typical basis. (b) Coverage of the time-frequency plane.

일반적으로 영상에서의 에지 등과 같은 높은 주파수를 갖는 신호들은 짧은 공간(시간)영역에서 존재하고, 계조도가 평탄한 영역으로 이루어진 낮은 주파수를 갖는 신호들은 넓은 공간에 분포한다. 이러한 특성을 나타내는 신호를 분석하기 위해 고정된 윈도우를 갖는 STFT을 사용하는 것보다 높은 주파수일 때는 시간영역에서 작은 버팀을 갖고 낮은 주파수일 때는 넓은 버팀을 갖는 가변 윈도우를 적용하는 것이 시간-주파수 해상도의 관점에서는 유리하다.

2. 이산 웨이블릿 변환과 신호의 다해상도 분해

신호의 다해상도 표현은 신호  $x(t)$ 를 일련의 내포 부공간(embedded subspace)  $V_m \subset L^2(\mathbb{R})$ ,  $m \in \mathbb{Z}$  상으로 투영(projection)시킴으로써 이루어 진다. 신호  $x(t)$ 의 다해상도 표현  $X_m(t)$ 과 내포 부공간들의 관계는 식(5)와 같다.

$$V_m \subset V_{m-1}, \quad m \in \mathbb{Z}$$

$$\bigcap_m V_m = \{0\}$$

$$\bigcup_m V_m = L^2(\mathbb{R}) \tag{5}$$

$$x(t) \in V_m \iff x(2t) \in V_{m-1}$$

$$X_m(t) = \sum_{n=0}^{\infty} \langle x(t), \phi_{m,n}(t) \rangle \phi_{m,n}(t)$$

또한, 부공간  $V_0$ 의 정규 직교 기저  $\phi(t-n)$ ,  $n \in \mathbb{Z}$

를 생성하는 스케일링 함수  $\phi(t)$ 가 존재한다. 이러한 각 내포 부공간의 기저들은 식(6)에서와 같이 스케일링 함수를 시간천이 및 스케일링을 통해 생성된다.

$$\phi_{m,n}(t) = 2^{-m/2} \phi(2^{-m}t - n), \quad m, n \in \mathbb{Z} \tag{6}$$

신호를 높은 해상도 공간  $V_{m-1}$ 에서 낮은 해상도 공간  $V_m$ 으로 근사화할 때, 정보의 손실이 발생하는 데, 이러한 신호를 잔여신호(residual signal)라고 하며,  $V_m$ 의 직교 여 부공간(orthogonal complement subspace)  $W_m$ 상에 존재한다. 상하위 해상도 공간과 직교하는 공간 사이의 관계는 식(7)과 같고, 그림 3에서 이상적인 저대역 부공간  $V_m$ 과 밴드대역 부공간  $W_m$ 들의 스펙트럼을 나타내었다.

$$V_{m-1} = V_m \oplus W_m, \quad V_m \perp W_m \tag{7}$$

부공간  $V_m$ 의 스케일 함수  $\phi(t)$ 와 유사하게 부공간  $W_m$ 을 구성하는 정규직교 기저가 존재하며 기저를 생성하는 모 웨이블릿(mother wavelet)  $\psi(t)$ 이 존재한다. 식(8)에서와 같이 각 부공간의 기저는 웨이블릿  $\psi(t)$ 을 스케일링과 시간천이를 수행함으로써 생성된다.

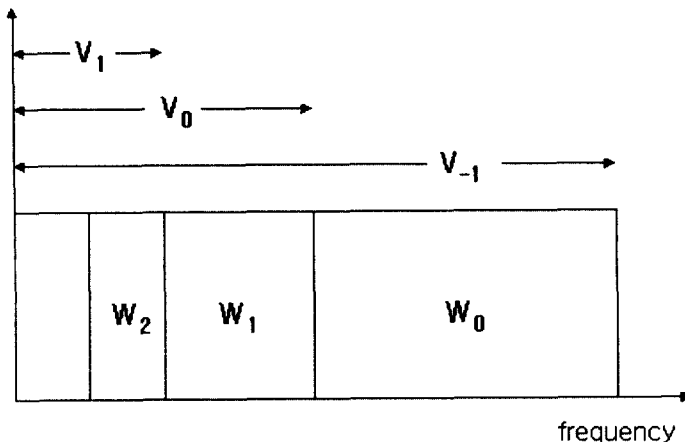


그림 3. 이상적인 저대역 부공간  $V_m$ 과 밴드대역 부공간  $W_m$ 의 스펙트럼  
Fig. 3. The spectrum of ideal lowpass spaces  $V_m$  and ideal bandpass spaces  $W_m$

$$\begin{aligned} \psi_{m,n}(t) &= 2^{-m/2}\psi(2^{-m}t-n) \quad , \quad m,n \in Z \\ x_{m-1}(t) &= x_m(t) + r_m(t) \quad (8) \\ r_m(t) &= \sum_n \langle x(t), \psi_{mn}(t) \rangle \psi_{mn}(t) \end{aligned}$$

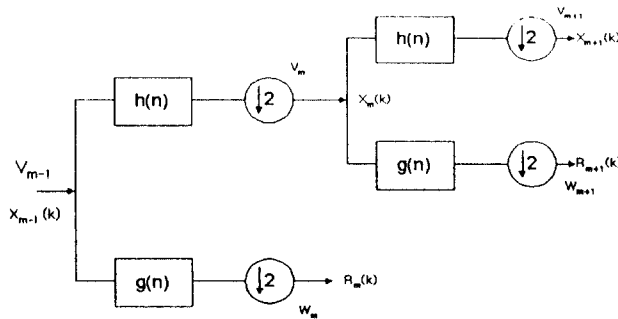
하나의 해상도 부공간은 더 낮은 해상도 부공간과 적고 여 부공간으로 표현가능하고, 반복적용을 하면 여 부공간만으로 표현가능하다. 따라서, 식(9)와 같이 신호  $x(t)$ 는  $W_m$ 의 기저들로 표현할 수 있다.

$$\begin{aligned} V_m(t) &= W_m \oplus W_{m-1} \oplus W_{m-2} \dots \\ L^2(R) &= \bigcup_m (\oplus W_m) \quad , \quad m \in Z \quad (9) \\ x(t) &= \sum_{m,n} \langle x(t), \psi_{m,n}(t) \rangle \psi_{m,n}(t) \end{aligned}$$

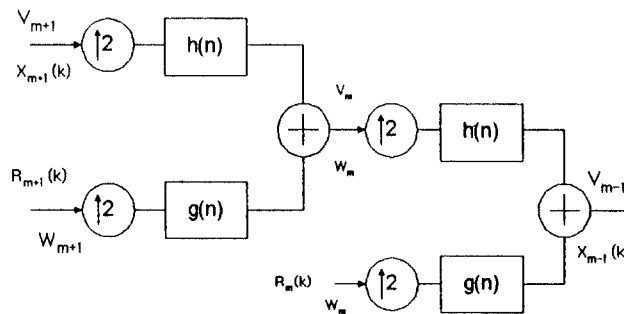
한편  $\psi(t) \in V_m$ 이면  $\psi(2t) \in V_{m-1}$ 이고, 식(10)과 같은 관계를 만족한다.

$$\begin{aligned} \phi(t) &= \sqrt{2} \sum_n h(n)\phi(2t-n) \quad , \quad n \in Z \\ \psi(t) &= \sqrt{2} \sum_n g(n)\psi(2t-n) \quad , \quad n \in Z \\ \sum_n h(n) &= \sqrt{2} \\ g(n) &= (-1)^n h(1-n) \\ \sum_n g(n) &= 0 \end{aligned} \quad (10)$$

신호  $x(t)$ 에 대해  $C_{mn} = \langle x(t), \psi_{m,n}(t) \rangle$ ,  $d_{mn} = \langle x(t), \phi_{m,n}(t) \rangle$ 일 때, 부공간  $V_m$ ,  $W_m$ 과  $h(n)$ ,  $g(n)$  간의 관계는 그림 4와 같이 2진 트리(binary tree) 구조를 형성한다.



(a) 웨이블릿 분해.



(b) 웨이블릿 합성.

그림 4. 2진 트리구조의 웨이블릿 분해와 합성

(a) 웨이블릿 분해. (b) 웨이블릿 합성.

Fig. 4. Binary tree structure of analysis and synthesis of wavelet.

(a) wavelet analysis, (b) wavelet synthesis.

3. 필터 뱅크와 이산 웨이블릿 변환

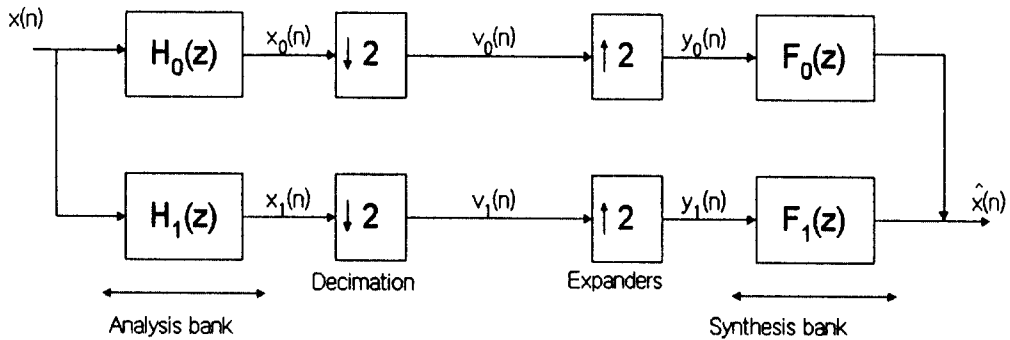
입력  $V_0$ 를 이산 신호  $x(n)$ 이라고 하면,  $C_{m-1,n}$ 과  $C_{m,n}$ ,  $d_{m,n}$ 의 관계는 완전 복원 필터 QMF(perfect reconstruction quadrature mirror filter)에서의 분해 및 합성과 동일하다. 그림 5에 2채널(two channel) QMF의 기본 구조와 분해 필터의 특성을 나타내었다[17].

그림 5의 2채널 필터뱅크의 분해 및 합성에서 입력 신호를  $X(z)$ 하고 할때, 출력신호  $\hat{X}(z)$ 는 식(11)로 표현된다.

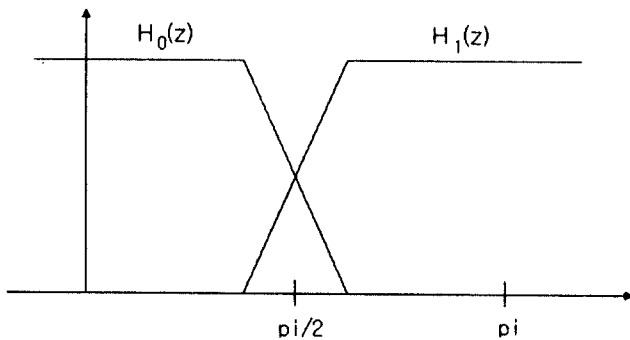
$$\hat{X}(z) = \frac{1}{2} [H_0(z)F_0(z) + H_1(z)F_1(z)]X(z) + \frac{1}{2} [H_0(-z)F_0(z) + H_1(-z)F_1(z)]X(-z) \quad (11)$$

$$\hat{X}(z) = T(z)X(z) + S(z)X(-z)$$

$S(z)$ 는 aliasing을 발생시키는 전달함수이다. 따라서  $S(z)=0$ 으로 놓아 aliasing을 제거시키고,  $T(z)=Cz^m$ 의 형태가 되도록 설계하면 진폭왜곡(amplitude



(a) 분해 및 합성 블록도.



b) 분해 및 합성 필터의 주파수 응답.

그림 5. QMF의 기본구조.

(a) 분해 및 합성 블록도.

(b) 분해 및 합성 필터의 주파수 응답.

Fig. 5. Structure of QMF.

(a) Block diagram of analysis and synthesis.

(b) Frequency response of analysis and synthesis filters.

distortion)과 위상왜곡(phase distortion)을 제거할 수 있다. 이 조건을 만족하는 출력  $\hat{X}(z)$  는 식(12)와 같다.

$$\begin{aligned} \hat{X}(z) &= Cz^{-n_0}X(z), \hat{x}(n) \\ &= Cx(n-n_0), C \neq 0 \text{인 상수} \end{aligned} \quad (12)$$

식(12)는 합성된 신호  $\hat{X}(z)$ 는 입력신호  $X(z)$ 가 증폭 또는 감쇄되고 시간천이된 형태를 의미하며, 이것을 완전복원이라고 한다. 완전복원의 해를 구하기 위해  $S(z) = 0$  이라고 하면 식(13)을 얻을 수 있다.

$$\frac{H_0(-z)}{H_1(-z)} = -\frac{F_1(z)}{F_0(z)} \quad (13)$$

완전복원 조건에 조건  $T(z) = Cz^{-n_0}$ 을 추가하면 각 필터의 관계 식(14)를 얻을 수 있으며 완전복원을 위한 충분조건이 된다.

$$\begin{aligned} H_1(z) &= -z^{-N}\hat{H}_0(-z), \\ F_0(z) &= z^{-N}\hat{H}_0(z), \quad F_1(z) = z^{-N}\hat{H}_1(z) \end{aligned} \quad (14)$$

여기서,  $\hat{A}(z) = A^*(z^{*-1})$ 이다. 식(14)를 식(11)

에 대입하면 식(15)와 같다.

$$\begin{aligned} \hat{X}(z) &= \frac{1}{2}z^{-N}[H_0(z)\hat{H}_0(z) \\ &+ H_1(z)\hat{H}_1(z)]X(z) \end{aligned} \quad (15)$$

식(15)에서  $H(z) = H_0(z)\hat{H}_0(z), F(z) = H_1(z)\hat{H}_1(z)$  이라고 하면,  $H(z)$ 와  $F(z)$ 는 각각 zero phase를 갖는 저역 통과 필터와 대역 통과 필터가 된다.

#### 4. 필터뱅크의 설계

영상부호화를 목적으로 FIR 필터 설계시 고려사항 으로서는 완전 복원, 필터의 직교성, 필터의 선형위상 응답, 평활도, 필터의 길이, 에너지 집중도, 시간-주파수 지역성 등이 있으나, 본 논문에서는 웨이블릿과 관련되어 완전복원, 직교성, 선형위상응답에 주목한다. 이산 웨이블릿 변환을 이용한 영상의 다해상도 분해에 사용되는 필터들은 완전복원을 만족해야 하지만 필터 뱅크 구현시 양자화 에러가 발생한다. 따라서, 응용분야에 따라 선형성이 요구되는 경우 완전복원 조건을 제외해서 필터설계의 자유도를 높일 수 있다[16-17]. 영상부호화에 널리 이용되는 웨이블릿 관련 필터 설계 방법을 기술하고자 한다. 대표적인 필터 설계 방법에 대해 필터의 탭수, 필터간의 직교성, 선형위상, 완전 복원 조건 등의 특징을 표1에 나타내었다.

표 1. 각 필터들의 특징.

Table 1. Characteristics of filters

필터 \ 특성	직교성	선형위상	완전복원
Equiripple filter	O	X	O
Daubechies' filter	O	X	O
Coiflets filter	O	△	O
Biorthogonal filter	X	O	O
Johnstone's filter	X	O	X

X: Poor, △: Almost good, O: Good



A. Equiripple 필터 설계(Smith and Barnwell filter)

Equiripple 필터는 완전복원을 위해서 다음 세 가지 조건하에서 설계된다.

- 1)  $G(z) = \sum_{n=-N}^N g(n)z^{-n}$  이고 차수가  $2N$ 인 power symmetric 필터를 설계한다. 이때 이 필터의 특성에 맞는 설계 알고리즘으로 Parks-McClellan 알고리즘을 이용한다.  $N = \text{홀수}$
- 2)  $H(z) = G(z) + \delta$ ,  $\delta = \text{peak subband ripple}$ ,
- 3)  $H(z)$ 에서 Minimum phase를 갖도록 단위 원 안의 Pole과 Zero를 선택하여 필터의  $H_0(z)$ 를 설계한다.

설계된 필터를 그림 6에 나타내었다. 이 방법으로 설계된 필터는 regularity가 낮으므로 영상신호의 손실 부호화에 적용하기에는 부적절하다[13]. 이를 해결하기 위해 Daubechies는 영상신호에 적합하도록 regularity가 높고 또한 완전복원 조건을 만족하는 필터설계 방법을 제안하였다[19].

B. Daubechies의 필터

Daubechies는 영상신호를 처리하기에 적절한 regularity가 높은 필터를 설계하였는데 설계조건은 1), 2), 3)과 같다[14].

- 1)  $z = -1$  즉,  $w = \pi$ 에서  $N$ 개의 zero를 갖도록  $P_N(z)$ 를 설계한다.
- 2) 필터  $P_N(z)$ 에서 1)로 설계된 부분 이외의 부분 즉,  $W(z)$ 에 진폭왜곡이 제거되도록 조건을 주어 삼각함수의 멱급수로 전개한다.
- 3) 설계된 필터  $P(z)$  계수 중 최소 위상 응답을 갖도록 계수를 선택한다.

이상의 과정은 식(16)-(25)와 같다.

$N \geq 2$  일때의 해를 구하면,

$$C(z) = \cos(\pi w) = (z^{-\frac{1}{2}} + z^{\frac{1}{2}})/2$$

$$S(z) = \sin(\pi w) = (z^{-\frac{1}{2}} - z^{\frac{1}{2}})/2i$$
(16)

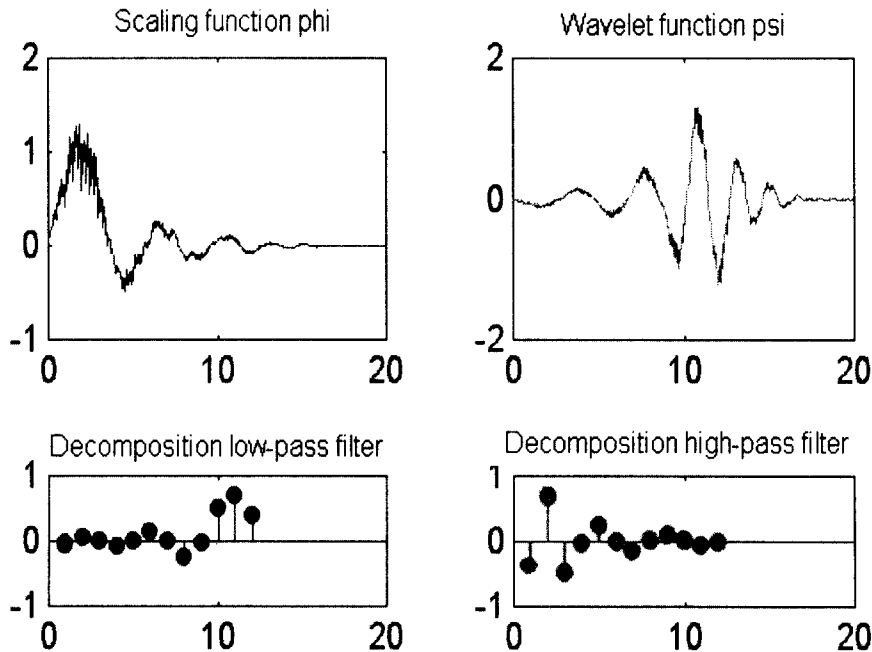


그림 6. Equiripple 필터  
Fig. 6. Equiripple filter

$$1 = (C^2 + S^2)^{2N-1} = \sum_{k=0}^{2N-1} \binom{2N-1}{k} C^{2N-2-2k} (1-C^2)^{2k} \geq 0 \quad (17)$$

여기서  $|P(z)|^2$  를  $P_N(C)$  의 형태로 표현하면,

$$P_M(C) = C^{2N} W_M(S) = \left(\frac{1+z}{2}\right)^{2N} W_M(S) \quad (18)$$

$$W_M(S) = \sum_{k=0}^{N-1} \binom{2N-1}{k} (1-S^2)^{N-1-k} S^{2k} \geq 0$$

$$W(z) = \sum_{n=0}^{N-1} w_n z^n \quad |W(z)|^2 = W_N(S) \quad (19)$$

$$P(z) = \left(\frac{1+z}{2}\right)^N W(z) = \left(\frac{1+z}{2}\right)^N \sum_{n=0}^{N-1} w_n z^n \quad (20)$$

$$W(1) = 1 \equiv P(1) = 1$$

$$P(z) = \frac{(1+z)}{2} = e^{-\pi w} \cos(\pi w) \quad (21)$$

$$P(-z) = \frac{(1-z)}{2} = ie^{-\pi w} \sin(\pi w)$$

$$W_2(S) = 1 + 2S^2 = 2 - \frac{z+z^{-1}}{2}$$

$$W(z) = a + bz, \quad W(-z) = W(z)^* = a + bz^{-1} \quad (22)$$

$$|W(z)|^2 = (a + bz)(a + bz^{-1}) = a^2 + b^2 + ab(z + z^{-1})$$

여기서  $N = 1$  이면, Haar system에 해당된다.

따라서 a,b는 식(23)과 같이 결정된다.

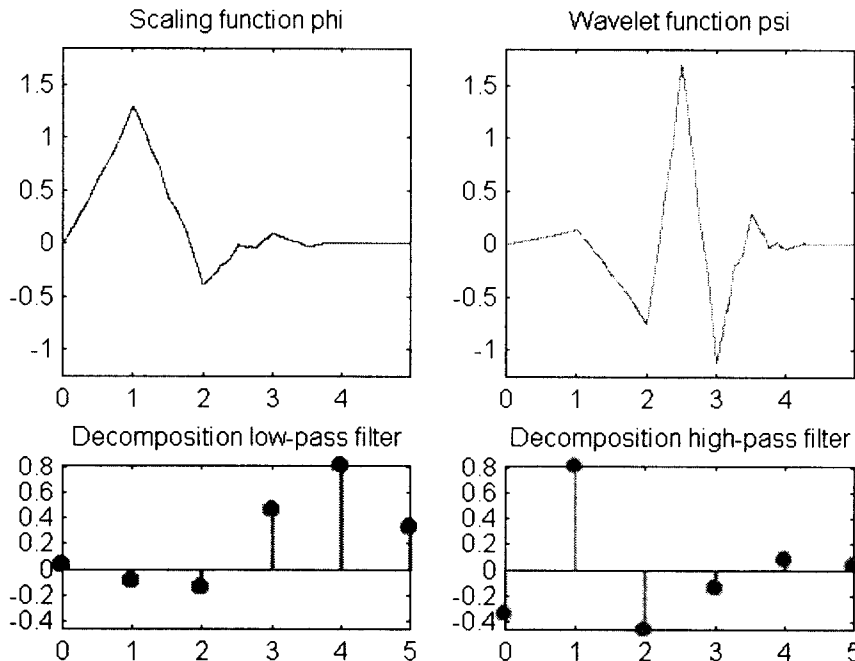


그림 7. Daubechies 필터.

Fig. 7. Daubechies' filter

$$a^2 + b^2 = 2, \quad ab = -\frac{1}{2}$$

$$a = \frac{1}{2}(1 \pm \sqrt{2}), \quad b = \frac{1}{2}(1 \mp \sqrt{3}) \quad (23)$$

$$W_{\pm z} = \frac{1}{2}[(1 \pm \sqrt{3}) + (1 \mp \sqrt{3})z]$$

$$P_{\pm z} = \frac{1}{2}\left(\frac{1 \pm z}{2}\right)^2[(1 \pm \sqrt{2}) + (1 \mp \sqrt{3})z] \quad (24)$$

- maximum phase , + minimum phase

N = 2 인 경우, 필터의 계수값은 (25)와 같다.

$$\frac{1}{2} \begin{bmatrix} 1 & 2 & 1 \\ 1 \pm \sqrt{3} & 1 \mp \sqrt{3} \end{bmatrix} =$$

$$\frac{1}{2} \begin{bmatrix} 1 \pm \sqrt{3} & 3 \pm \sqrt{3} & 3 \mp \sqrt{3} & 1 \mp \sqrt{3} \end{bmatrix} \quad (25)$$

그림 7에 설계된 6 tap 필터를 나타내고 Equiripple 필터 보다 regularity가 높아진 것을 알 수 있다 [15][19].

### C. Coiflets 필터

Coiflets filter는 완전복원을 만족하면서 가능한 선형위상에 가깝도록 설계한 것이다. 필터의 선형위상 성질을 증가시키기 위해 식(26)과 같은 조건을 부가하였다[16].

$$\int \phi(t) dt = 1$$

$$\int t^\nu \phi(t) dt = 0 \text{ for } \nu=1,2,\dots,L-1 \quad (26)$$

$$\int t^\nu \psi(t) dt = 0 \text{ for } \nu=1,2,\dots,L-1$$

여기서  $\psi(t)$ 는 스케일링 함수,  $\phi(t)$ 는 웨이블릿 함수를 나타낸다. 식(26)의 조건들을 주파수 영역에서 보면 식(27)과 같이 주파수 응답의 평활한 정도 (flatness)를 의미한다.

$$\frac{d^\nu}{d\Omega^\nu} \phi(t) = 1$$

$$\frac{d^\nu}{d\Omega^\nu} \phi(t) = 0 \text{ for } \nu=1,2,\dots,L-1 \quad (27)$$

$$\frac{d^\nu}{d\Omega^\nu} \psi(t) = 0 \text{ for } \nu=1,2,\dots,L-1 \quad (1049)$$

식(27)을 필터  $H_0(e^{j\omega})$ 로 치환하면 식(28)과 같다.

$$\left. \frac{d^\nu}{d\omega^\nu} H_0(e^{j\omega}) \right|_{\omega=0} = 0 \text{ for } \nu=1,2,\dots,L-1$$

$$\left. \frac{d^\nu}{d\omega^\nu} H_0(e^{j\omega}) \right|_{\omega=\pi} = 0 \text{ for } \nu=1,2,\dots,L-1 \quad (28)$$

식(28)의 조건들을 만족하려면  $H_0(e^{j\omega})$ 는 식(29)와 같은 형태를 가져야 한다.

$$H_0(e^{j\omega}) = 1 + (1 - e^{j\omega})^L V(e^{j\omega})$$

$$H_0(e^{j\omega}) = \left[ \frac{1}{2}(1 + e^{j\omega}) \right]^L P(e^{j\omega}) \quad (29)$$

여기서,  $V(e^{j\omega})$ 와  $P(e^{j\omega})$ 는 삼각다항식이다.  $H_0(e^{j\omega})$ 에 완전복원 조건을 추가하면  $H_0(e^{j\omega})$ 는 식(30)과 같이 얻을 수 있다.

$$H_0(e^{j\omega}) = 1 + (\sin^2 \frac{\omega}{2})^K \left[ - \sum_{k=0}^{K-1} \binom{K-1+k}{k} \right]$$

$$\left[ (\cos^2 \frac{\omega}{2})^k + (\cos^2 \frac{\omega}{2})^{K-k} \right] f(e^{j\omega}) \quad (30)$$

여기서  $L = 2K$ 이다.  $f(e^{j\omega})$ 는 식(29)의 두 가지 조건을 만족하도록 선택되어야 한다. 그림 8에 설계된 6 tap 필터를 나타내었다.

### D. Biorthogonal 필터

Biorthogonal 필터는 orthogonality보다는 선형위상과 완전복원을 만족하도록 설계한 필터이다[13]. 완전복원의 조건으로부터  $S(z)=0$ ,  $T(z)=Cz^L$ , 식(13), 식(14), 식(31), 식(32)를 만족하도록 설계하면 Biorthogonal 필터를 얻을 수 있다.

$$\langle \tilde{f}_0(n-2k), h_1(n-2l) \rangle = 0$$

$$\langle \tilde{f}_1(n-2k), h_0(n-2l) \rangle = 0 \quad (31)$$

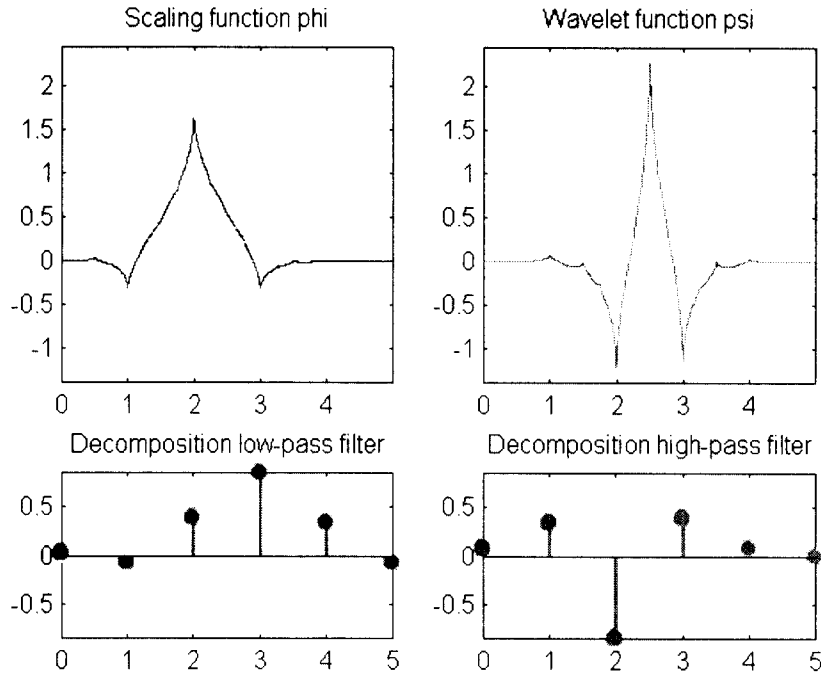


그림 8. Coiflets 필터.  
Fig. 8. Coiflets filter.

$$\begin{aligned}
 \tilde{f}_i(n) &= f_i(-n) \\
 \langle \tilde{f}_0(n-2k), h_0(n) \rangle &= \delta_k \\
 \langle \tilde{f}_1(n-2k), h_1(n) \rangle &= \delta_k
 \end{aligned} \tag{32}$$

9 tap, 7 tap Biorthogonal 필터를 그림 9에 나타내었다.

E. Johnstone 필터

Johnstone 필터는 완전복원을 만족하지 않지만, 선형위상을 만족하면서 차단(cut-off) 주파수 대역의 에너지와 ripple 에너지를 필터 설계의 조건으로 하고 있다.

Ripple 에너지

$$E_r = 2 \sum_{\omega=0}^{\frac{\pi}{2}} [ |H(\omega)|^2 + |H(\omega - \pi)|^2 - 1 ]^2$$

차단 주파수대역 에너지

$$E_s = \sum_{\omega=f_{co}}^{\pi} |H(\omega)|^2$$

최소화하는 비용함수(cost function)는 식(33)과 같고  $\alpha$ 는 설계 자유도를 나타낸다.

$$E = E_r + \alpha E_s(f_{SB}) \tag{33}$$

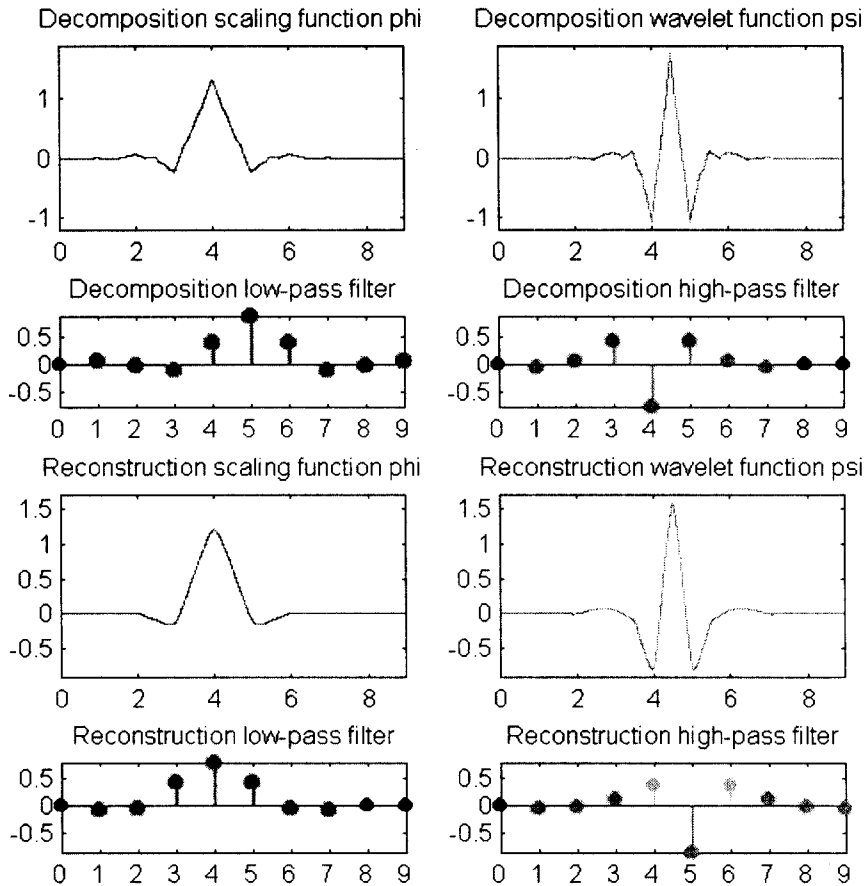


그림 9. Biorthogonal 필터

Fig. 9. Biorthogonal filter.

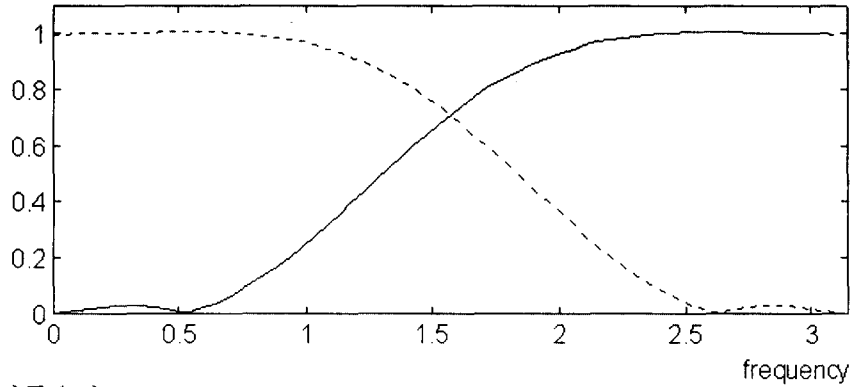
Johnstone 필터의 주파수 특성은 그림 10에서 나타나는 것과 같이 완전복원필터의 진폭응답과 유사하며 선형 위상 특성을 갖고 있다.

### III. 최근의 웨이블릿 기반 영상압축 기법

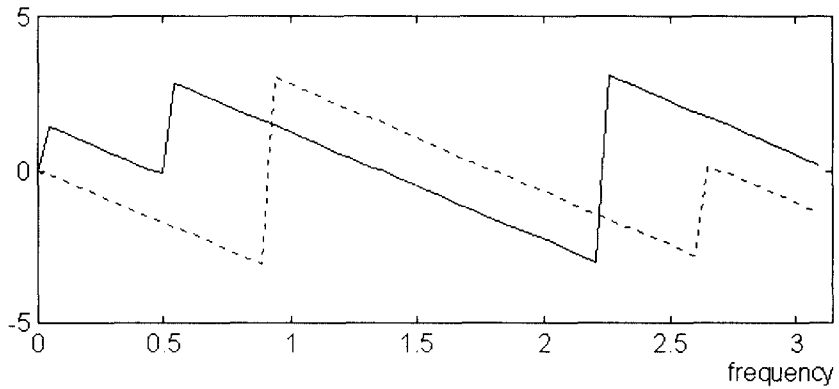
영상압축 기술은 영상내에 존재하는 통계적인 용장성을 제거함으로써 가능하며, 이산여현변환(DCT), DPCM, 부밴드 영상들의 엔트로피 부호화는 이러한 방식을 사용한다. 용장성을 제거하는 것은 단지 영상 압축에 있어서 제한적인 양만을 이용할 뿐이다. 따라서, 높은 압축률을 얻기 위해서는 비용장성(non-redundancy)정보의 일부분도 제거되어야만 한다. 그러나, 이러한 성질을 영상압축에 적용하면 시각

적으로 중요한 부분에 에러를 야기함으로써 복원 영상의 화질 열화를 일으킨다. 한편, low level에서의 영상은 에지에 의해서 경계 지워진 텍스처나 칼라의 표면들로 구성된 구조로서 생각될 수 있다. 또한 영상의 색과 텍스처는 방향(orientation)과 밝기(illumination)에 의해서 크게 영향을 받으므로 좋은 영상압축 방법은 에지 변형(distortion)을 최소화 하도록 해야한다.

인간의 눈은 영상을 많은 수의 부밴드들로 나누어 영상을 필터링하고, 각각은 주파수에서 하나의 옥타브(octave)범위를 가지고 있다. 공간적인 측면에서 영상은 많은 수의 다른 스케일로 구성된 정보의 조합으로 생각될 수 있다[20]. Marr는 영상에서 밝기의 변화를 일으키는 물리적 현상이 공간적으로 지역화(locality)되어 있다는 것을 제안하였다.



(a) 진폭응답.



(b) 위상응답.

그림 10. Johnstone 8탭 필터의 주파수 특성

(a) 진폭응답. (b) 위상응답.

Lewis[1] 등은 인간시각을 고려한 Marr의 공간적 지역성 (spatial locality)에 기반한 웨이블릿 영상압축 알고리즘을 제안하였다. 즉, 가장 낮은 밴드의 웨이블릿 계수들은 그대로 전송하고, 나머지 부밴드는 부밴드내에 0이 아닌 웨이블릿 계수들 중에 어느 것이 잡음에 상응하며 혹은 어느정도 시각적으로 중요한지를 Marr의 공간적 지역화의 제약에 기반한 알고리즘을 적용하였다. 따라서, 중요한 상세(detail) 정보가 영상에서 발생했다면 그 위치에 상응하는 웨이블릿 계수들은 한 옥타브 이상에서 임계치를 넘을 것으로 판단할 수 있다.

임계치값들은 인간 시각 특성을 주파수, 배경의 밝기 및 질감 마스크의 형태로 모델링하여 적응적으로 결정되었다. 즉, 식(34) ~ 식(37)과 같이 저주파 성분

으로부터 배경의 밝기를 구하고 저주파 및 고주파 계수를 이용하여 질감 크기를 구하여 이를 웨이블릿 계수의 양자화에 이용하였다.

$$qstep(r, s, x, y) = q_0 * frequency(r, s) * luminance(r, x, y) * texture(r, x, y)^{0.034} \quad (34)$$

$$frequency(r, s) = \begin{cases} \sqrt{2}, & \text{if } s = GG \\ 1, & \text{otherwise} \end{cases} * \begin{cases} 1.00, & \text{if } r = 0 \\ 0.32, & \text{if } r = 1 \\ 0.16, & \text{if } r = 2 \\ 0.10, & \text{if } r = 3 \end{cases} \quad (35)$$

$$luminance(r, x, y) = 3 + \frac{1}{256} \sum_{i=0}^{2^k-1} \sum_{j=0}^{2^k-1} I^{3,LL}(i+1 + \frac{x}{2^k}, j+1 + \frac{y}{2^k}) \quad (36)$$

$$texture(r, x, y) = \sum_{k=1}^3 16^{-k} \sum_{i=0}^{2^k-1} \sum_{j=0}^{2^k-1} (I^{k+r}(i + \frac{x}{2^k}, j + \frac{y}{2^k}))^2 + 16^{3-r} var(I^{3,LL}((1,2) + \frac{x}{2^k}, (1,2) + \frac{y}{2^k})) \quad (37)$$

[1]에서 제안된 방법의 결과는 HVS를 기반으로한 필터링과 시각적으로 화질 열화를 덜 일으키는 양자화기는 블록 변환 방법에 비해 압축률과 화질에 있어서 월등한 향상을 가져왔지만, 예측에러의 측면에서 보면 종종 큰 에러를 야기시키기 때문에 고압축률에서 불연속이 크게 나타나는 부근에 보풀과 같은 artifacts가 발생하여 화질 열화가 발생하였다.

Queiroz[21] 등은 기존의 JPEG 코더와 호환성이 있으며, 높은 압축율에서 블록화(blocking artifacts) 및 링잉(ringing artifacts)현상 등으로 복원화질이 저

하되는 단점을 보완하여 JPEG구조와 유사한 JPEG+Like코더를 제안하였다. 즉, 그림 11과 같이 하나의 웨이블릿 블록은 서로 다른 부밴드(subband)에서 같은 공간적인 위치에 존재하는 웨이블릿 계수들을 그룹화 하여 구성된다. 이렇게 구성된 웨이블릿 블록의 계수들은 적절한 양자화 테이블을 이용하여 양자화 한 후, 부밴드의 밴드별 방향 특성을 이용한 스캔방법으로 스캔하여 엔트로피 부호화 된다.

그림 12는 JPEG 베이스라인 인코더의 구조와 DWT-JPEG 인코더 구조를 나타내고 있다. 그림 12에서 (a) 와 (b)의 차이점은 변환방식이 서로 다르다. 즉, JPEG에서는 DCT를 쓰는 반면 DWT-JPEG에서는 DWT을 채택하고 있으며, 스캔방식에 있어서도 JPEG은 지그재그 스캔을 수행하는 반면 DWT-JPEG에서는 그림 11과 같이 변형된 스캔방식을 사용한다.

DWT-JPEG 코더는 기존의 JPEG과 비교하여 모든 비트율에서 약 1dB 이상의 화질 향상을 보였으며, 4레벨 DWT-JPEG의 성능은 3레벨 DWT-JPEG과 비교하여 낮은 비트율에서는 보다 좋은 성능을 보이지만, 비트율이 높아지면 오히려 3레벨 DWT-JPEG의 성능이 좋을 수 있다.

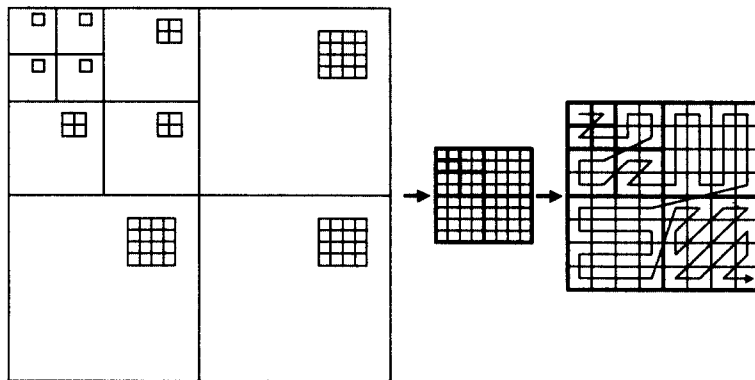


그림 11. 3단계 DWT의 경우 웨이블릿 블록 구성 과정

Fig. 11. Illustration of the block construction procedure for a three-level DWT.

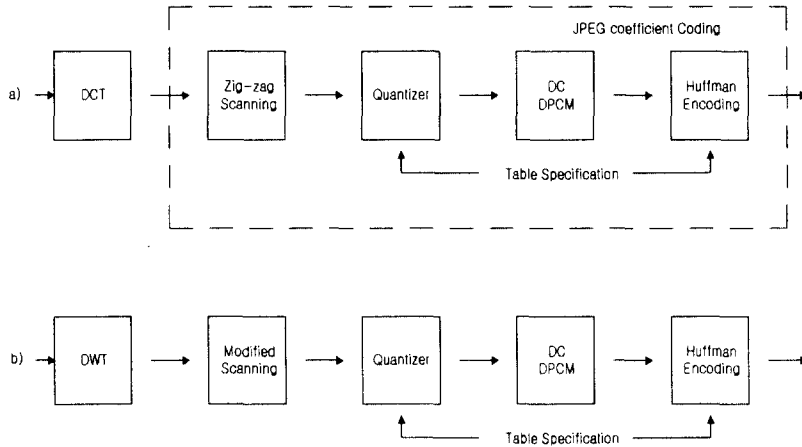


그림 12. (a) 베이스라인 JPEG 인코더. (b) JPEG 구조 기반의 DWT-JPEG 인코더.

Fig. 12. (a) Baseline JPEG basic encoding flow diagram. (b) DWT-JPEG encoder based on a JPEG structure.

[22]에서는 JPEG의 DCT 서브루틴[23]과 Shapiro가 제안한 embedded 제로 트리 양자화[2]를 결합하여 embedded 비트 스트림을 발생시키는 DCT 기반의 embedded 영상압축 코더를 제안하였다. 제안된 코더는 입력 영상을 8×8 블록으로 나누고 DCT를 수행한다. 다음에는 변환된 64개의 DCT 계수들에 대하여 그림 13과 같이 레이블링(labeling)한 후, embedded 제로 트리 양자화를 수행한다. 한편, 제로 트리 양자화를 수행하는 동안 모든 DCT 블록에서 같은 인덱스를 갖는 계수들은 함께 모아져서 DC 계수에서부터 스캔된다. 즉 AC계수들에 적용하기 이전에 먼저 모든 블록의 DC계수를 양자화 한다.

DCT 기반의 embedded 영상압축 코더는 Shapiro의 EZW(embedded zerotree wavelet)코더에 의해 얻어진 좋은 결과가 부분적으로 웨이브릿 변환의 공헌 때문인 반면, DCT 또한 embedded 제로 트리 양자화기와 결합될 수 있음을 제시하였고 복원된 영상은 JPEG과 Shapiro의 EZW 코더보다 향상된 PSNR을 가져왔다.

0	1	4	5	16	17	20	21
2	3	6	7	18	19	22	23
8	9	12	13	24	25	28	29
10	11	14	15	26	27	30	31
32	33	36	37	48	49	52	53
34	35	38	39	50	51	54	55
40	41	44	45	56	57	60	61
42	43	46	47	58	59	62	63

그림 13. 8×8 DCT 블록 계수의 depth-3 트리 구조 표현.

Fig. 13. Depth-3 tree structure of 8×8 DCT block coefficients.



[24]에서는 부분집합 분할(subset partitioning) 알고리즘을 기반으로 웨이블릿 계수들의 크기에 따른 부분 오더링(ordering), 재정렬된 비트들의 비트 플레인 전송, 웨이블릿 변환 영상의 서로 다른 스케일의 자기 유사성(self-similarity)등을 이용한 EZW 코더를 제안하였다.

부분 오더링은 변환 계수의 크기를 집합 분할 정렬(set partitioning sorting) 알고리즘을 이용하여 레벨별로 감소하는 임계치들의 집합과 비교함으로써 수행되며, 주어진 임계치에 대해서 변환 계수의 크기가 임계치를 넘느냐, 아니냐에 따라 significant/insignificant를 결정한다.

전송 순서는 그림 14와 같이 크기 순서대로 정렬된 웨이블릿 계수에 대하여 MSB를 먼저 전송하고 LSB는 나중에 전송하는 점진적 전송이 가능한 방식을 채택함으로써, 즉 전송 priority를 바꿈으로써 하나의 embedded 파일을 가지고 모든 비트율에 대해 가장 좋은 성능을 보였다.

[25]에서는 최근의 웨이블릿 코딩 알고리즘[1-2]을 기반으로 자연 영상이 주파수와 공간 모두에서 에너

지 집중도의 선형결합으로서 잘 특징지워질 수 있다는 사실(영상내의 에너지 대부분이 저주파 영역에 집중되어 있으며 남아있는 고주파성분의 대부분 에너지가 에지 주위에 몰려있음)을 이용한 매우 효율적인 영상압축 방법을 제안하였다. 즉, 매우 간단한 양자화를 통해 웨이블릿 변환의 주파수와 공간상 모두에서의 에너지 집중도 특성을 이용하고 웨이블릿의 공간상의 집중도 특성을 이용하기 위해 고주파 성분 계수의 공간적 영역은 0의 값을 갖도록 제로 트리 양자화하였다.

제로 트리 양자화는 고주파 성분 계수들이 모여있는 영역을 가리키는 하나의 메카니즘으로 볼 수 있으며, 영상 모델에 의해서 예측된 고주파 성분 계수들의 공간적 군집화(spatial clustering)을 이용할 수 있다. 따라서, 제로 트리 양자화에 의해서 0으로 되지 않은 계수들에 대하여 계수들의 주파수 대역과는 무관한 균일 스칼라 양자화(uniform scalar quantizer)를 적용하였다. 다시말하면, 공간적 양자화 모드(제로 트리 양자화)와 각 계수의 주파수 밴드의 스칼라 양자화를 동시에 최적화 할 수 있는 영상압축 알고리즘을 제안하였다.

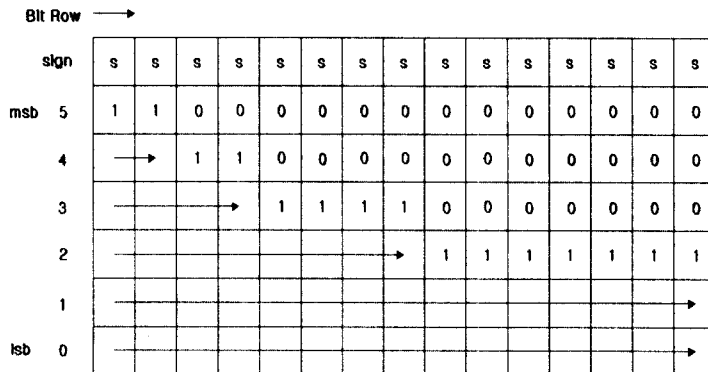


그림 14. 크기 순서화된 계수의 2진 표현.  
Fig. 14. Binary representation of the magnitude-ordered coefficients.

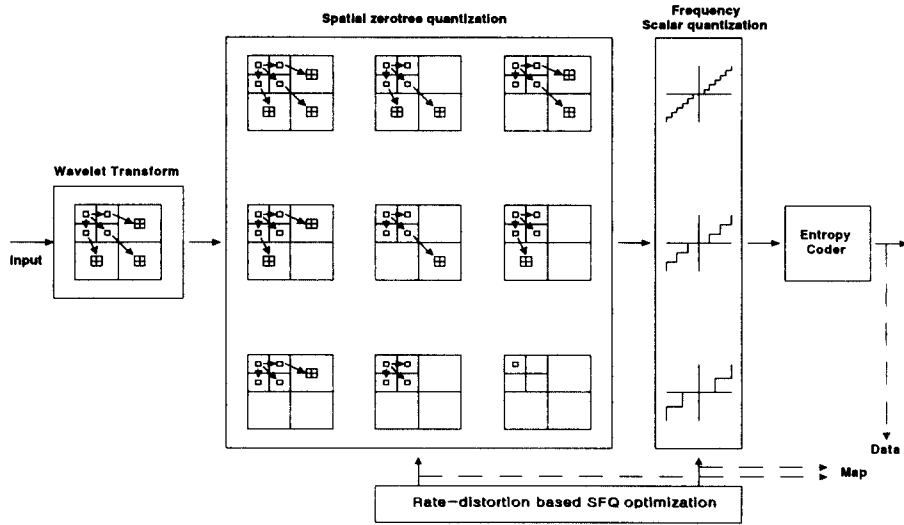


그림 15. SFQ 코더의 블록도.

Fig. 15. Block diagram of the SFQ(Space-Frequency Quantization) coder.

Lewis와 Knowles[1]는 같은 공간적 영역에 해당하는 저주파 밴드에서 보다 낮은 값을 갖는 계수에 기반한 고주파 계수들의 제로 트리를 예측하는 임계치 방법(thresholding scheme)을 적용하였다. 그러나, 이 방법은 밴드간 의존성을 이용하고 있는 반면, 예측에러의 측면에서 보면 종종 큰 에러를 야기한다. Shapiro[2]는 해당하는 트리의 모든 계수값이 0일 때, 제로 트리 심볼을 적용하였다. 그러나, 이러한 접근 방법은 부호화 방법의 왜곡(distortion)을 최소화 할 수 있는 반면에 주어진 비트율(bit-rate)을 만족하는 모든 전략상의 왜곡을 가장 최소로 하지는 못한다. 따라서, [25]에서 제안된 알고리즘은 주어진 비트율에 대해 왜곡을 최소로 하기 위한 제로 트리와 스칼라 양자화의 적용을 최적화하는 문제에 초점을 두었다.

그림 15는 SFQ(space-frequency quantization) 과정을 보이고 있으며, 입력 영상에 대하여 웨이블릿 변환을 수행하고 R-D(rate-distortion) 기반의 제로트리 양자화와 스칼라 양자화 과정을 거쳐 엔트로피 부호

화 된다.

공간적 제로트리 양자화와 각 계수의 주파수 밴드의 스칼라 양자화는 고정된 quality factor와 양자화 스텝 사이즈에서 왜곡(distortion)을 가장 최소로 하는 tree-map을 구한 후, 고정된 quality factor에서 왜곡을 가장 최소로 하는 스칼라 양자화 스텝 사이즈를 구하고 원하는 비트율을 만족하는 가장 최적인 quality factor를 구하는 과정으로 볼 수 있다.

#### IV. 시각 특성을 이용한 실험적 웨이블릿 기반 영상압축 기법

이상의 이론 및 최근의 연구 동향을 바탕으로 웨이블릿 변환을 이용한 간단한 실험적 영상압축 방식을 제안하여 성능을 고찰해 보기로 한다. 인간의 시각은 주위의 여러 가지 환경에 대하여 다르게 반응한다. 뿐만 아니라, 어떤 특정한 성분에 민감한 반응을 보

이기 때문에 영상압축시 시각 특성을 고려하여야 한다. 기본적으로 사람이 어떤 물체를 눈으로 판별할 수 있는 것은 빛이 눈의 시신경을 자극하기 때문이다. 인간이 인식하는 빛을 빛의 3원색(빨강, 파랑, 녹색)으로 나타내며, 많은 영상압축 표준에서 휘도(luminance)와 나머지 두 성분의 신호로 전환하여 사용하는데 이것은 휘도 성분이 다른 성분보다 더 중요하게 사람의 눈에 인식되기 때문이다.

인간의 시각특성에 주로 영향을 미치는 요소는 대비도(contrast), 인식한계, 주파수별 감도(sensitivity) 등이 있다. 대비도는 두 밝기 간의 차이를 의미하며 이 차이가 클수록 인간의 시각은 더욱 선명하게 보게 된다. 인간의 시각은 같은 차이로 변하는 휘도신호라도 인식하는 한계를 가진다. 즉 일정한 범위를 벗어난 휘도값들에 대해서는 변화를 잘 인식하지 못한다. 인간의 시각은 주파수에 따라 다른 감도를 보이며, 중간 주파수 영역이 다른 주파수 영역보다 민감하게 반응한다. 즉, 중간 주파수 부분을 보다 선명하게 느끼는 것은 인간의 눈이 일종의 대역통과(Band pass) 필터 역할을 하기 때문이다.

### 1. 분해 및 복원

웨이블릿 변환을 이용한 영상의 다해상도 분해에 사용되는 필터들은 완전복원을 만족해야 하지만 양자화를 하는 경우 어느 정도의 에러가 발생한다. 따라서, 이보다 작은 범위 내에서 완전복원을 만족하지 않아도 되므로 필터의 선택에 있어서 폭이 넓어진다. 사용되는 필터들의 특징으로는 필터의 길이, 필터간의 직교성(orthogonality), 선형위상(linear phase), 완전복원(perfect reconstruction)조건 등을 들 수 있다.

대부분의 일차원 신호와 영상신호는 유한 길이를 가지고 있다. 따라서 가장자리(boundary)값을 필터링하기 위해서는, 확장(extension)을 해야 한다. 특히 필터의 길이가 긴 경우 확장해야 할 가장 자리는 그만큼 증가하게 된다. 가장 간편한 방법은 zero-padding을 하는 것이지만 이렇게 할 경우 신호의 주기성이 사라져 전송 또는 저장하는 데이터량이 증가하게 된다. 가장자리 확장은 입력 신호와 사용하는 필터의 특성에 따라 다르게 하여야 한다. 보다 효율적인 padding 방법을 Li가 제안하였는데 이 경우 각 단계에서 1개 이하의 신호만 Padding하면 된다[26]. 일반적인 가장자리 확장 방법이 Bamberger등에 의해 정리되었다[27][28].

## 2. 양자화와 부호화

웨이블릿 변환으로 분해된 부밴드 영상들은 서로 다른 통계적 특성을 갖는다. 또한, 다른 단계의 부밴드 영상 계수들의 절대치 간에도 일정한 상관성이 존재하는데, 약 0.3정도의 값을 갖는 것으로 알려져있다[2]. 이 상하 단계의 상관성을 이용하여 분해된 부밴드 영상들을 양자화하고 압축하는 방법(embedded zerotree wavelet:EZW)이 Shapiro에 의해 제시되었다.

### 1) LL밴드 부호화

LL밴드 영상은 다른 밴드들과는 다른 통계적 특성을 가지고 있고 다해상도 압축에서 상대적으로 중요한 정보를 담고 있기 때문에 압축할 때 많은 비트를 할당하여야 한다. 이것은 JPEG과 같은 블록단위의 압축에서 DC 계수를 상대적으로 적은 Q값으로 나누어줌으로써 이 계수들의 오차를 적게 한다는 점에서 인간시각 특성을 고려한 것이기도 하다. Q값을 결정하는 것은 압축률을 고려하여 결정하여야 한다. 즉 압축률이 높을 경우 Q값도 큰 값으로 해야하며 낮을 경우 Q값도 작게 하여야 최상의 화질을 얻을 수 있다. 다해상도로 분해한 LL밴드 영상은 적절한 Q값으로 나누어 선형 양자화기를 거친 후 JPEG의 무손실 압축모드를 이용하여 DPCM하였다[23].

### 2) 다른 부밴드 부호화

웨이블릿 변환으로 영상을 다해상도로 분해하면 분해된 각 부밴드 영상들은 서로 다른 방향의 성분을 나타낸다. 따라서, 분해된 부밴드 영상 별로 스캔 방향을 다르게 하면 더 효과적으로 압축할 수 있다. 즉, LH밴드는 수직방향, HL밴드는 수평방향 그리고 HH밴드는 지그재그(zigzag) 방향으로 스캔한다. 이렇게 스캔 방향을 달리 할 경우 산술부호화시 효율이 증가됨을 알 수 있다. 스캔 후 LL밴드 영상을 제외한 다른 부밴드 영상들은 Shapiro가 제안한 EZW를 이용하여 양자화 및 압축한다.

## V. 실험 결과 및 고찰

실험에 사용한 영상은 512×512의 크기를 갖는 Lena이다. 영상을 다해상도로 분해한 후 EZW를 이

용하여 연속근사양자화하기 때문에 완전복원 특성보다는 선형위상 특성이 더 중요하다. 따라서, 필터는 선형위상을 가지면서 완전복원에 가까운 특성을 지닌 필터를 사용하였다[19]. 우선 목표비트율(Target bit rate)이 정해지면 LL밴드 영상을 적절한 Q값을 주어

DPCM 및 압축하고 남은 bit에 나머지 부밴드 영상들을 EZW를 사용하여 압축하였다. 그림 16(a)에 사용한 Lena 원영상을 나타내었고, (b)에 0.125bps로 압축한 영상을 나타내었다.

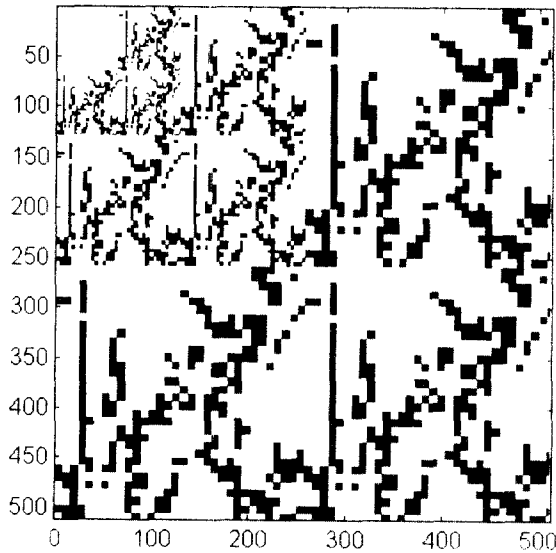


그림 16. (a) Lena 원영상 (b) 0.125bps (PSNR 30.12 dB)

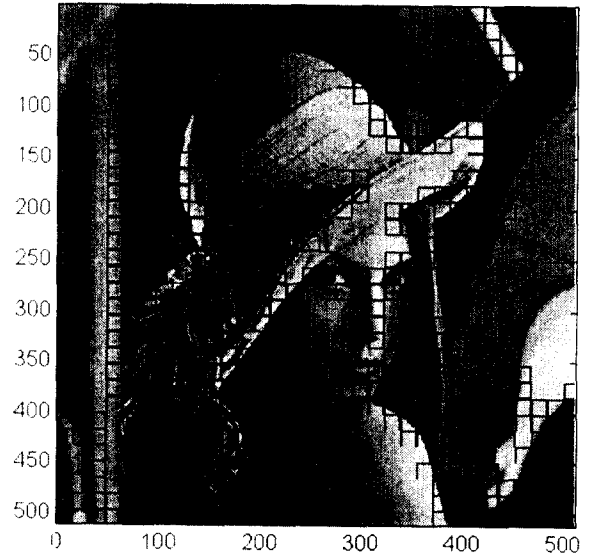
Fig. 16. (a) Original Lena image. (b) Result at 0.125bps (PSNR 30.12 dB)

Edge image

Edge image



(a)



(b)

그림 17. (a) 검출된 에지들의 2진 map. (b) 검출된 에지 map을 원화에 부영한 그림.

Fig. 17. (a) Binary map of edge. (b) Overlap edge to Lena.

LL밴드를 선형양자화한 LL 부영상을 이용하여 에지를 검출하고 각 밴드의 크기로 전환한 후 가중치를 주었다. 각 파라미터 값들은 실험으로 결정되었으며, 예지로 판별되는 문턱값으로 그 부밴드들의 분산값의 1.6배 정도에서 중요한 에지를 가장 잘 검출하는 것을 확인하였다. 그림 17(a)는 Lena 영상을 3단계로 분해하고 LL밴드 영상을 이용하여 에지를 검출한 후 각 단계의 크기로 전환한 것이고, 그림 17(b)는 검출된 에지를 원화에 투영한 것으로 비교적 중요한 에지를 잘 검출하고 있다. 실험 결과, 에지에 가중치를 주지 않고 압축한 영상과 거의 같은 PSNR을 나타내었고, Lena 영상의 눈이나 어깨선 그리고 턱선과 같은 부분에서 시각적 화질이 개선됨을 알 수 있다.

본 논문에서 구현한 방법과 JPEG의 성능을 그림 18에 비교하였다. 주관적인 화질이나 PSNR면에서 기존의 압축방법보다 우수한 성능을 보였다. 또한, Shapiro가 제안한 EZW와 비교할 때 거의 같은 성능을 나타냈으며 시각적 특성은 오히려 능가함을 알 수 있었다.

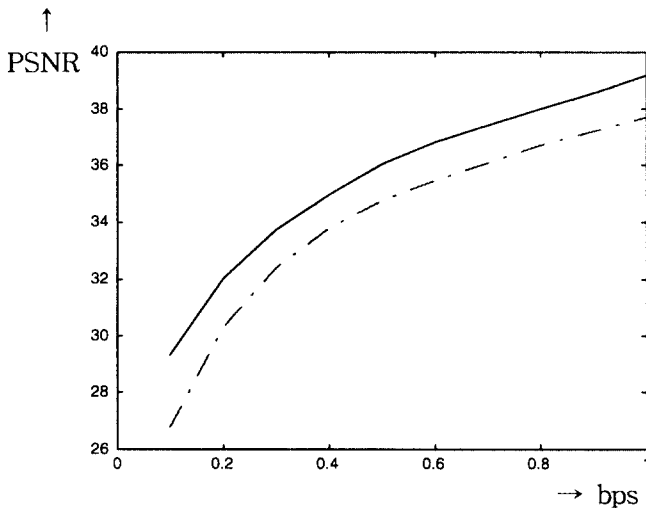


그림 18. JPEG과 본 논문과의 비교

Fig. 18. JPEG vs. proposed results.

## VI. 결 론

본 논문에서 푸리에 변환과 웨이블릿 변환과의 연관성 및 물리적인 차이점 등에 대한 이론적인 접근, 웨이블릿 변환과 영상의 다해상도 분해, 그리고 필터뱅크와 이산 웨이블릿 변환과의 관계에 대해 살펴보

았다. 또한, 기존의 웨이블릿 필터들에 대한 필터설계 방법과 각각의 특성에 대하여 살펴보았다. 한편, 최근에 제안된 웨이블릿 기반 영상압축 기법들이 제시하고 있는 주요 내용과 성능에 대하여 간략하게 설명하였다.

또한, 웨이블릿 변환을 이용하여 영상을 다해상도로 분해한 후 인간의 시각특성을 고려한 실험적 영상압축 기법을 소개하였다. 본 방식에서 LL밴드 영상은 DPCM과 허프만부호화 및 산술부호화를 수행하고 나머지 부밴드 영상들은 에지를 고려하여 적절한 가중치를 주어 EZW로 연속 근사 양자화한 후, 발생된 양자화 심벌들은 무손실 압축한다. 웨이블릿 변환은 인간의 시각특성을 잘 반영하고 있으므로 압축된 영상은 JPEG으로 압축할 때와 같은 블록간의 왜곡이 거의 발생하지 않으며 전체적으로 JPEG 및 다른 압축들과 비교할 때 우수한 성능을 나타낸다.

## 참고문헌

- [1] A. Lewis and G. Knowles, "Image compression using the 2-D wavelet transform," *IEEE Trans. Image Processing*, vol. 1, pp. 244-250, Apr. 1992.
- [2] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. Signal Processing*, vol. 41, pp. 3445-3463, Dec. 1993.
- [3] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston, MA: Kluwer, 1992.
- [4] G. D. Fomey, R. M. Gray, and M. Ostendorf Dunham, "Finite-state vector quantization for waveform coding," *IEEE Trans. Inform. Theory*, vol. 31, pp. 348-359, May. 1985.
- [5] H. S. Malvar, *Signal Processing with Lapped Transforms*. Norwood, MA: Artech House, 1992.
- [6] B. Ramamurthi and A. Gersho, "Classified vector quantization of images," *IEEE Trans. Commun.*, vol. 34, pp. 1105-1115, Nov. 1986.
- [7] E. A. Riskin, "Optimal bit allocation via generalized BFOS algorithm," *IEEE Trans. Inform. Theory*, vol. 37, pp. 400-402, Mar. 1991.

- [8] J. W. woods and S. O'Neil, "Subband coding of images," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 1278-1288, Oct. 1986.
- [9] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. 28, pp. 84-95, 1980.
- [10] P. H. Westerink, D. E. Boekee, J. Biemond, and J.W. Wood, "Subband coding of images using vector quantization," *IEEE Trans. Commun.*, vol. 36, pp. 713-719, June. 1988.
- [11] J. Max, "Quantization for minimum distortion," *IEEE Trans. Inform. Theory*, pp. 7-12, Mar. 1960.
- [12] M. W. Marcellin and T. R. Fischer, "Trellis coded quantization of memoryless and Gaussian-Markov sources," *IEEE Trans. Commun.*, vol. 38, no. 1, pp. 82-93, Jan. 1990.
- [13] A. Averbuch, D. Lazer and M. Israeli, "Image Compression Using Wavelet Transform and Multiresolution Decomposition," *IEEE Trans. Image Proc.*, vol. 5, no. 1, pp. 4-15, Jan. 1996.
- [14] S. G. Mallat, "Multifrequency channel decompositions of images and wavelet models," *IEEE Trans. ASSP*, vol. 37, no. 12, pp. 2091-2110, Dec. 1989.
- [15] O. Rioul, "Regular Wavelets: A Discrete-time approach," *IEEE Trans. Signal Proc.*, vol. 41, no. 12, Dec. 1993.
- [16] Ali N. Akansu and Richard A. Haddad, *Multiresolution Signal Decomposition*, Academic Press, Inc., 1992.
- [17] P. P. Vaidyanathan, *Multirate Systems and Filter Banks*, Prentice Hall, 1993
- [18] Jechang Jeong and William J. Williams, "Variable-Windowed Spectrograms: Connecting Cohen's Class and the Wavelet Transform," *IEEE 5th ASSP Workshop on Spectrum Estimation and Modeling*, pp. ? Oct. 10-12, 1990.
- [19] Ingrid Daubechies., *Ten Lectures on Wavelets*, SIAM, 1992
- [20] D. Marr, *Vision*, New York: Freeman, 1982.
- [21] R. Queiroz, C. K. Choi, Y. Huh, and K. R. Rao, "Wavelet Transforms in a JPEG-Like Image Coder," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 7, no. 3, pp. 419-424, 1997.
- [22] Z. Xiong, O. Guleryuz and M. T. Orchard, "A DCT-based Embedded Image Coder," *IEEE Singal Processing Letters*, vol. 3, 11, pp. 289-290, November, 1996.
- [23] W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression Standard*, New York: Van Norstand Reihold, 1992.
- [24] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 243-250, June 1996.
- [25] Z. Xiong, K. Ramchandran, and M. T. Orchard, "Space-Frequency Quantization for Wavelet Image Coding," *IEEE Trans. Image Processing*, vol. 6, no. 5, pp. 677-693, 1997.
- [26] J. Li, P. Y. Cheng and C. J. Kuo, "On the Improvements of Embedded [27] A. Aldroubi and M. Unser, *Wavelets in Medicine and Biology*, CRC Press, pp. 43-46, 1996.
- [27] A. Aldroubi and M. Unser, *Wavelets in Medicine and Biology*, CRC Press, pp. 43-46, 1996.
- [28] R. H. Bamberger, S. L. Eddins and Veyis Nuri, "Generalized Symmetric Extension for



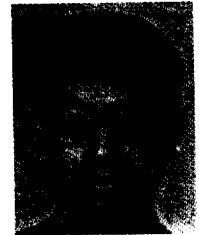
민 병 석

- 
- 1990년 3월 : 한양대학교 전자통신공학과 졸업(공학사)
  - 1992년 8월 : 한양대학교 대학원 전자통신공학과 졸업(공학석사)
  - 1993년 3월~현재 : 한양대학교 대학원 전자통신공학과 박사과정
  - 관심분야 : 신호처리, 영상통신, 멀티미디어통신



김 승 종

- 
- 현 재 : 한양대학교 대학원 전자통신공학과 박사과정
  - 참 조 : 한국통신학회논문지 제 22권 5호



정 제 창

- 
- 현 재 : 한양대학교 대학원 전자통신공학과 조교수
  - 참 조 : 한국통신학회논문지 제 22권 5호