

---

# 다중 웨이브렛을 이용한 음성신호 데이터 압축에 관한 연구

김태형\* · 박재우\* · 윤동한\* · 노석호\*\* · 조익현\*

\*금오공과대학교, \*\*안동대학교

## A Study on the Data Compression of the Voice Signal using Multi Wavelet

Tae-Hyung Kim\* · Jae-Woo Park\* · Dong-Han Yoon\* · Seok-Ho Noh\*\* · Ig-Hyun Cho\*

\*Kumoh National Institute of Technology, \*\*Andong National University

E-mail : thkim@kumoh.ac.kr

---

본 연구는 2004년도 PoP-iT 인력양성사업에 의하여 지원되었음

---

### 요 약

급변하는 정보 및 통신기술의 발달에 따라 멀티미디어 데이터의 효율적인 압축 기술에 관한 심층적 연구 필요성이 더욱 커지고 있다. 본 논문에서는 멀티미디어 신호 중 음성 및 음향 신호 데이터 압축과 관련하여 웨이브렛 기반의 압축 알고리즘 구조를 설계하였다. 2-band 구조와 웨이브렛 packet 구조에서의 압축의 효율성에 대하여 조사하고, 각 구조에서 웨이브렛 기저함수로 Daubechies 웨이브렛 계수와 Coiflet 계수를 사용하여 구조에서의 효율성 및 재생오차를 조사하였다. 최종 압축은 Huffman code를 사용하여 압축율(CR) 및 재생오차(PRD)를 기존의 DCT와 비교분석하였다.

### ABSTRACT

According to the rapid development of the information and communication technology, the demand on the efficient compression technology for the multimedia data is increased magnificently. In this Paper, we designed new compression algorithm structure using wavelet base for the compression of ECG signal and audible signal data. We examined the efficiency of the compression between 2-band structure and wavelet packet structure, and investigated the efficiency and reconstruction error by wavelet base function using Daubechies wavelet coefficient and Coiflet coefficient for each structure. Finally, data were compressed further more using Huffman code, and resultant Compression Rate(CR) and Percent Root Mean Square difference(PRD) were compared with those of existent DCT.

### 키워드

Wavelet Packet, DWT, Discrete Cosine Transform, 2-band structure

### I. 서 론

웨이브렛을 이용한 신호의 분해(analysis)와 재구성(reconstruction)은 스케일 함수(scale function)  $\phi(t)$ 와 웨이브렛 함수(wavelet function)  $\psi(t)$ 를 이용한다. 스케일 함수를 해상도의 개념을 이용하여 정의하고 이 함수로부터 웨이브렛 함수를 도출해 낸다[1]. 이 두 함수는 다음과 같이 정의된다.

$$\phi(t) = \sum_n h(n) \sqrt{2} \phi(2t-n) \quad (1)$$

$$\psi(t) = \sum_n h_1(n) \sqrt{2} \phi(2t-n) \quad (2)$$

여기서  $h(n)$ 과  $h_1(n)$ 을 각각 스케일링 필터계수와 웨이브렛 필터계수라고 부르며, 두 함수는 다음과 같이 서로 직교성(orthogonality)의 관계

를 가진다[2].

$$h_1(n) = (-1)^n h(1-n) \tag{3}$$

따라서 웨이블릿 함수는 스케일 함수에서 부호가 변하고 역의 형태로 스케일 함수와의 컨볼루션(convolution)에 의해 얻어진다. 즉 식(1)과 (2)의 스케일 함수와 웨이블릿 함수는 서로 직교한다. 또한 이들 필터 계수는 다음과 같은 조건을 만족하여야 한다[1,3].

$$\sum_n h(n) = \sqrt{2} \tag{4}$$

$$\sum_n h_1(n) = H_1(0) = 0 \tag{5}$$

이들 조건은 계수가 웨이블릿이 되기 위한 조건으로 알려져 있다. 또한 이렇게 만들어진 식(2)의 웨이블릿 함수는 다음식과 같이 확장함수를 위한 모 웨이블릿(mother wavelet)을 제공한다.

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k) \tag{6}$$

여기서  $2^j$ 는  $t$ 의 스케일링( $j$ 는  $\log_2$ 의 스케일),  $2^{j/2}k$ 는  $t$ 에서의 천이(translation)이고,  $2^{j/2}$ 는 단위 노름(norm)을 유지하기 위한 요소이다.

$L^2(R)$  영역을 모두 스패น(span) 할 수 있는 함수  $g(t) \in L^2(R)$ 를 웨이블릿과 스케일 함수를 사용하여 표현하면 다음과 같이 표현된다.

$$g(t) = \sum_{k=-\infty}^{\infty} c(k)\phi(t) + \sum_{j=0}^{\infty} \sum_{k=-\infty}^{\infty} d(j,k)\psi_{j,k}(t) \tag{7}$$

식(7)에서 첫 번째 항의 함수  $g(t)$ 의 저해상도 혹은 근사적인 표현이며, 두 번째 항은  $j$ 가 매번 증가에 따라 고해상도가 추가되는 항이다. 이는 푸리에 급수 전개와 다소 유사하다. 이 함수에서 만약 정규직교 기저(Orthonormal basis)를 사용한다면 위의 급수 전개에서 계수들은 내적의 합으로 구할 수 있다.

$$c_{jk} = \langle g(t), \phi_{j,k}(t) \rangle = 2^{j/2} \int_{-\infty}^{+\infty} g(t)\phi(2^j t - k) dt \tag{8}$$

$$d_{jk} = \langle g(t), \psi_{j,k}(t) \rangle = 2^{j/2} \int_{-\infty}^{+\infty} g(t)\psi(2^j t - k) dt \tag{9}$$

위 식에서  $d_{j,k}$ 는 웨이블릿 계수로 상세(detail) 계수라고 부른다. 이는 신호  $g(t)$ 와 웨이블릿 패밀리  $\psi_{j,k}(t)$ 와의 내적을 취함으로써 계산된 것으로 연속시간에서의 이산 웨이블릿 변환(Discrete Wavelet Transform: DWT)라고 부른다. 또한 유사한 방법으로 입력 신호와 확장된 패밀리  $j$ 와 천이  $k$ 의 스케일 함수  $\phi(t)$ 와의 내적  $c_{j,k}$ 를 근사 계수(approximation coefficient)라고 부른다. 위의 두식을 웨이블릿을 이용한 완전한 신호의 분해(analysis)를 위한 절차로 정의된다. 효율적인 계산을 위하여 적분대신 식(1)와 (2)에서의 스케일 필터계수와 웨이블릿 필터계수를 이용하여 다음과 같이 계산하는 것이 계산상 많은 이점을 제공한다.

$$c_j(k) = \sum_m h(m-2k)c_{j+1}(m) \tag{10}$$

$$d_j(k) = \sum_m h_1(m-2k)c_{j+1}(m) \tag{11}$$

이와 같은 재귀적 분해 알고리즘은 cascade 알고리즘이나 피라미드(pyramid)알고리즘으로 언급되며, 이 두 식을 고속 웨이블릿 변환 알고리즘이라고 부른다. 낮은 레벨의 근사계수는 높은 레벨의 근사계수로부터 구해지며, 낮은 레벨의 상세계수도 역시 높은 레벨의 근사계수로부터 구해진다. 결과적으로 이들에 의해 분해(analysis)된 신호는 각각 저역통과필터와 고역통과필터에 대응된다.

이와 같은 2-밴드 구조의 경우 각 단계에서 구해진 계수에서 저주파 성분을 추출한 스케일링 계수만을 다음 단계로 분해하는 구조로 Mallat구조라고도 부르며 다음 그림 2의 상위 부분에 3단의 2밴드 구조를 표현해 놓았다.

만약 구조가 2-밴드가 아니라 M-밴드로 확장된다면 식(4)의 조건은 다음과 같은 조건을 만족하여야 한다.

$$\sum_n h(n) = \sqrt{M} \tag{12}$$

또한 웨이블릿 함수에 대한 식(6)에서의  $2^j$ 의 항목에서 2는 모두 M으로 바뀌어야 한다.

전형적으로 사용되는 웨이블릿은 M=2의 값으로 로그스케일의 주파수 해상도를 가진다. 따라서 낮은 주파수는 좁은 대역폭을 가지고, 높은 주파수는 넓은 대역폭을 가진다. 따라서 2-밴드 구조의 경우 낮은 주파수 해석에 좀더 적합할 것이다. 반대로 높은 주파수만을 해석할 경우는 이와 반대의 경우를 취하는 경우도 있다.

결국 다양한 주파수 성분이 포함된 신호의 경우, 2-밴드(two-band)구조의 시스템으로는 정밀한 해석을 하는데에는 한계가 있다. 따라서 전-밴드(full-band)구조의 웨이브렛 패킷구조가 제안되었다[1,3,4].

이 구조는 저역통과와 고역통과 밴드가 모든 단계에서 같이 분리된다. 따라서 이 구조는 필터뱅크 구조에서 완전한 구조에 대응되는 것으로 그림2의 아래에 나타내었다.

## II. 본 론

본 논문에서는 원신호(original signal)와 복원신호의 재생오차를 최대한 줄이기 위하여 다중 웨이브렛 기저함수를 이용하여 심전도(ECG) 및 음향신호 데이터 압축에 대하여 연구하였다. 신호는 비교적 저주파 성분이 많이 포함되어있는 ECG 신호와 음향신호를 선택하여 그림 1에 나타낸 압축구조로 신호를 압축하였다.

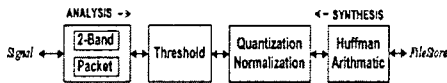


그림 1. 심전도 및 음향신호의 압축구조

분석은 먼저 2-밴드 구조와 전-밴드구조에 동시에 신호가 가해져 각각의 웨이브렛 변환을 행하였다. 이 구조는 그림 2에 나타내었다. 여기서 사용한 웨이브렛은 Daubechies 웨이브렛, Coiflet 웨이브렛 등 다수에 웨이브렛을 사용하였고, 각각에 대하여 1단에서 5단까지의 변환을 행하였다. 다음으로 이들 신호를 각 분해 단계에서 설정된 기준값(threshold)를 조정하여 변환영역에서 손실 압축을 하였다. 또한 중요하지 않은 특정 고주파 대역의 주파수들에 계수는 영으로 만들고, 다시 이 데이터를 합성하여 신호를 재구성하였다. 변환만 사용하였을 경우의 압축율(CR)과 퍼센트오차(PRD)를 구하기 위하여 다음에 주어진 식(13)과 (14)를 이용하여 각각 압축율(CR)과 PRD를 구하였다.

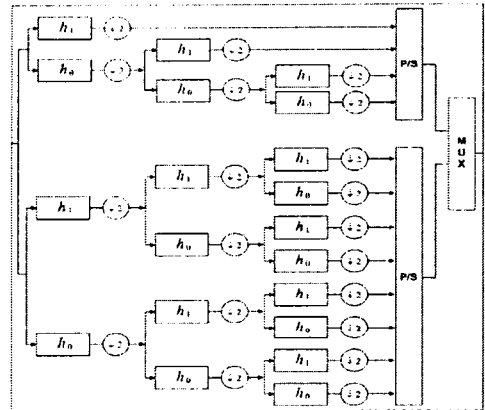


그림 2. 압축구조에서 2-밴드와 Packet 구조

또한 신호의 분해이후 구해진 데이터를 정규화(normalize)시킨 후, 각 데이터의 범위를 0~256으로 양자화 시켰다. 또한 보다 효율적인 압축을 위하여 무손실 압축인 허프만 코드(Huffman code)와 산술연산(Arithmetic)코드를 이용하여 최종 압축을 하였다. 이들 압축된 데이터는 파일로 저장하였고, 이들 압축된 파일과 원 파일의 크기를 비교하여 압축율(compression ratio)을 구하였다.

$$CR = \frac{\text{Original Signal Bit}}{\text{Compressed Data Bit}} \quad (13)$$

$$PRD = \sqrt{\frac{\sum_{i=1}^n [x(i) - r(i)]^2}{\sum_{i=1}^n x(i)^2}} \times 100 \quad (14)$$

또한 압축율의 효율성을 비교하기 위하여 같은 구조에서 웨이브렛 변환 대신 기존의 이산 코사인 변환(DCT)을 사용하여 같은 방법으로 압축율 및 퍼센트 오차를 구하였다.

## III. 실험결과

본론에서 정의한 방법에 따라 심전도(ECG)와 음향 신호에 대한 압축율 및 퍼센트 오차(PRD)에 대하여 조사하였다. 사용된 ECG 신호는 MIT-BIH 데이터베이스의 ECG106를 사용하였고, 음향 신호는 윈도우프로그램에서 제공하는 음향신호를 사용하였다[5].

먼저 ECG신호의 경우 원신호를 그림 3의 위쪽에 나타내었고, 웨이브렛 패킷을 사용하여 압축

후 허프만 코드를 이용하여 저장한 데이터를 다시 재구성한 신호는 그림 3에 아래쪽에 나타내었다. 이 경우 압축율은 2.78에 PRD는 2.48%의 결과이다. 압축율과 오차에 대한 결과는 그림 4에 그래프로 표현하였다.

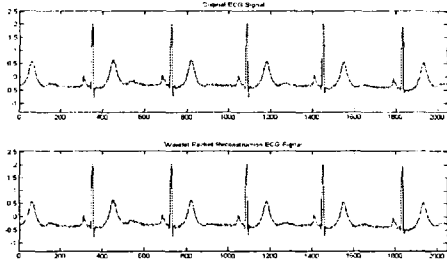


그림 3. 원 ECG신호와 재생 ECG 신호(웨이브렛 packet사용)

그림 4에서 보는바와 같이 DCT보다는 2-밴드의 구조에서 상당히 좋은 압축율(CR) 및 퍼센트 오차(PRD)를 보였다. 같은 PRD에서 평균적으로 약 2.5배정도의 압축율을 보였다. 또한 웨이브렛 패킷의 경우 낮은 PRD에서는 다소 좋지 못한 CR을 보였으나 전반적으로 DCT와 2-밴드보다는 좋은 결과를 보였다.

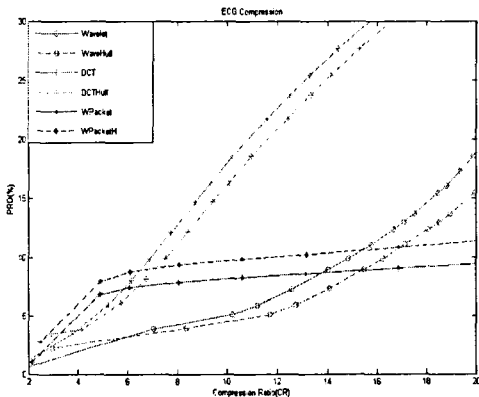


그림 4. ECG신호의 CR vs PRD

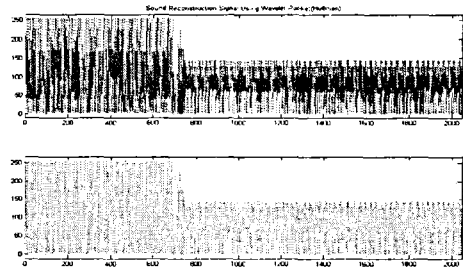


그림 5. 원 Sound 신호와 재생 Sound 신호 (웨이브렛 packet사용)

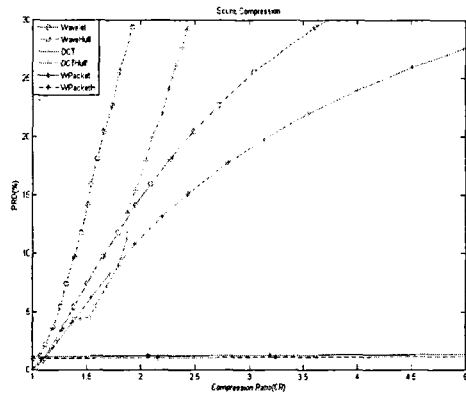


그림 6. Sound신호의 CR vs PRD

음향 신호의 경우 DCT의 경우가 2-밴드 구조보다는 좋은 결과를 보였다. 하지만 웨이브렛 패킷을 사용한 경우가 월등한 압축율 및 퍼센트 오차에 값을 보였다. 음향신호원과 재생신호원, 그리고 결과는 그림6에 나타내었다.

#### IV. 결 론

기존의 방법인 DCT의 경우 심전도신호나 음향 신호에 대하여 조사한 압축율(CR) 및 퍼센트오차 (PRD)는 전반적으로 좋은 결과를 얻을 수 있었다.

2-밴드 구조를 사용하였을 경우 심전도신호의 경우 비교적 DCT보다는 좋은 결과를 가져왔으나 음향 신호에서는 거의 유사한 결과를 가져왔다. 이는 2-밴드 구조가 가지는 주파수의 밴드에 기여한 것으로 판단된다. 또한 웨이브렛 패킷 구조를 사용하였을 경우 두 경우보다는 월등한 결과를 가져왔다. 이는 웨이브렛 패킷 구조가 가지는 높은 주파수 밴드와 선택 가능한 해상도를 가진다는 것을 증명해준다. 구조적으로 2-밴드 구조보다는 계산상의 복잡성과 계산시간이 다소 문제가 될 수 있지만 웨이브렛 패킷을 이용한 선택적으로 주파수 밴드를 설정할 수 있도록 구조를 만든

다면 나은 결과를 얻을 수 있을 것으로 판단된다.

#### 참고문헌

- [1] C.S.Burrus, "Introduction to Wavelet and wavelet Transform" Prentice Hall, 1998.
- [2] Daubechies Ingrid, "Orthonormal base of compactly supported wavelets", Comm. on Pure and Applied Math,41, pp.909-996, 1988.
- [3] Dwight F. Mix, Kraig J. Olejniczak, "Elements of Wavelets for Engineers and Scientists", Wiley Interscience, 2003.
- [4] Raghuv eer M. Rao and Ajit S. Bopardikar, "WAVELET TRANSFORMS Introduction to Theory and Applications", Addison-Wesley, 1998.
- [5] MIT-BIH Arrythmia Database, Havard-MIT Div. Health Sci. Tech., Biomed. Eng. Ctr., Mass. Inst. Tech, Rm. 20A-113, Cambridge, MA.