이산 웨이브렛 변환을 이용한 Audio 신호의 기호화 및 압축

백 한 욱, 정진현
광운대학교 제어계측공학과

The study of discrete wavelet transform for the coding and the compression of the audio data

Han Wook Baek, Chin Hyun Chung
Department of Control and Instrumentation Eng. Kwangwoon University

Abstract - This paper propose a new method for the discrete signal : Discrete Wavelet Transform(DWT). This paper is a brief introduction to the DWT and applies the DWT coding for the audio data as an example. We can have a number of hint about the compression algorithm of multimedia resources and the high performance of transmission and storage.

1. 서 론

본 논문은 최근 많은 응용 분야로 개척되고 있는 Wavelet 변환을 통해 신호의 해석 및 그 의미와 방법을 고찰하고 그 응용의 예로서 Audio 데이터의 기호화를 통해 그 타당성을 증명하는데 목적이 있다.
1800년 Joseph Fourier는 sin과 cos 함수를 통해 다른 함수(Signal)를 해석하는데 반하여 Wavelet 분석의 경우는 그 신호가 갖고 있는 Scale에 주안점을 갖고 각 데이터의 Scale이 하는 역할을 분석하는 것이다. Wavelet 알고리즘은 데이터를 서로 다른 Scales와 Resolutions에서 처리하게 된다. DWT의 직접적인 계산을 위하여 사용하게될 Pyramidal Coding Algorithm은 다해상도(Multi-resolution)의 성격으로 DWT의 공학적인 응용 전에 제시되었던 내용이다. 본 논문에서 설명되어질 내용들은 수학적인 내용 보다 그 활용의 직접적인 내용에 주안점을 두기로 한다.

2. 본 론

2.1 DWT(Discrete Wavelet Transform)

본 논문에서 사용하게될 DWT(Discrete Wavelet Transform)은 최근에 신호 해석 분야에서 많은 응용이 시도되고 있다. 본 논문에서는 Wavelet에 관한 수학적 내용은 수학자 Gilbert Strang[1]의 논문을 참고하였다. Fourier 변환에서 그 기초 함수(Basis Function)로 정형파(Sinusoid)를 사용하고 있음에 반하여 wavelet은 다음과 같은 식으로 기초 함수와 표현된다.

$$\Phi(x)=\sum_{k=0}^{M-1} c_k \cdot \Phi(2x-k) \qquad (1)$$

위의 합계의 범위는 특정된 0이 아닌 계수(係數)의 개수 M에 의해 결정된다. 이 계수(係數)의 개수 M은 Wavelet의 차수로도 표현된다. 다음은 계수를 결정하는 조건들이다.

● **Orthogonality**

Translations에서의 직교됨을 다음 식은 나타내고 있다.

$$\int \Phi(x)\cdot\Phi(x-k)dx=0$$

설계된 계수는 그 dilation과 scale에 대하여 다음의 식은 직교적임을 나타내고 있다.

$$\int \psi(x)\cdot\psi(2x-k)dx=0$$

$$\psi(x)=\sum_k(-1)^k\, c_{1-k}\cdot\Phi(2x-k)$$

위의 식들에서 알 수 있듯이 $\psi(x)$는 $\Phi(x)$를 푸는데 달려 있다.

● **Normalization**

$$\sum_k(-1)^k\cdot\, c_{1-k}\cdot\, c_{k-2m}=0$$

위의 식은 0이 아닌 계수들에 의한 합이 0이 되어야 함을 의미한다.

위와 같은 조건들에 의해 계수(係數)는 정해지고 그 계수에 의한 변환 작업이 이루어지게 된다. 본 논문에서는 계수(係數)를 유도하는 과정은 다루지 않기로 한다.
DWT를 활용하기 위한 실험에서는 표[1]에서와 같은 주어진 값들을 사용하기로 한다. 다음은 주어진 표의 내용을 직접 프로그램하기 위하여 서두에 언급한 Pyramid Algorithm에 관하여 설명해 보기로 하겠다.

2.1.1 Pyramid Algorithm

Pyramid Algorithm은 2의 승수로 이루어진 데이터 개수 N을 처리한다. 이들 입력 값들은 두 개의 Convolution 함수에 의해 처리하게 되는데 이들 함수는 일종의 Filter의 개념의 것들이다. 이 Filter 함수들은 다음과 같이 두 종류이다.

● **"Low-Pass" Filter 함수**

$$a_i = \frac{1}{2} \sum_{j=1}^{N} c_{2i-j+1} \cdot f_j, \quad i=1,\ldots,\frac{N}{2} \qquad (2)$$

● **"High-Pass" Filter 함수**

$$b_i = \frac{1}{2} \sum_{j=1}^{N} (-1)^{j+1} \cdot c_{j+2-2} \cdot f_j, \quad i=,\ldots,\frac{N}{2} \qquad ...(3)$$

High-Pass 출력 값들은 원래 입력 값과 Low-Pass 출력 값에 의해 재구성(Reconstructed Input)된 값들과 차를 의미한다. 일반적으로 높은 차수의 DWT(계수의 개수가 보다 많을 경우)는 High-Pass 출력 값 보다 Low-Pass 출력에 많은 양의 정보를 갖게 되고, 만약 High-Pass 출력 값이 0에 가까운 값을 갖게 하는 계수를 구할 수 있다면 DWT에 의한 압축률은 보다 높아 질 수 있을 것이다.

2.2 DWT의 구현

이해가 용이한 Haar Wavelet 변환을 통해 DWT를 계산하는 방법을 기술하여 보겠다.

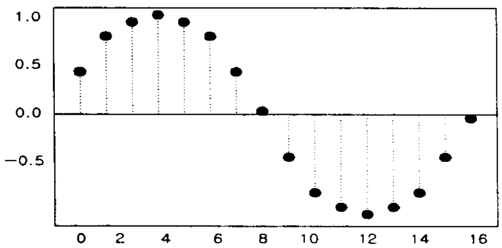


그림 [1]. Sin 함수의 Sample된 값들이다.

HWT(Haar Wavelet Transform)의 경우 Low-Pass의 출력 값들은 두 개의 입력 값들의 평균값이 되고 High-Pass 출력 값은 그림 [2]와 같이 작은 양의 정보를 갖게 된다.

$$f_j^L = \sum_{i=1}^{N/2} c_{2i-j} \cdot a_i, \quad j=1,\ldots,N \qquad (4)$$

위의 식은 Low-Pass 출력 값(그림[2])에 보여주는 것 같이 각각의 왼쪽(데이터 순차적인 측면)으로 복사하는 UP-Sampling을 통해 복원되게 된다. High-Pass 출력 값 또한 다음과 같은 식을 통해 복원된다.

$$f_j^H = \sum_{j=1}^{N/2} (-1)^{j+1} c_{j+1-2i} \cdot b_i, \quad j=1,\ldots,N \qquad (5)$$

N	입력 데이터의 개수
c	계수(係數)
f	입력 값
a	Low-Pass 출력 값
b	High-Pass 출력 값

완벽한 원 입력 값의 복원은 $f = f^L + f^H$이 된다.
만약 High-Pass 출력 값이 각 단계에서 0에 가까운 값으로 나온다면 High-Pass 출력값을 삭제하는 방법을 통해 데이터의 압축을 이룰 수 있게 되는 것이다. 위에서
언급한 Pyramid Algorithm에 의한 프로그램을 통해 최소의 단위로 압축을 한 후 그 일련의 데이터 값들을 데이터 허용 범위 내에서 삭제함으로써 최고의 압축을 하게 되는 것이다.

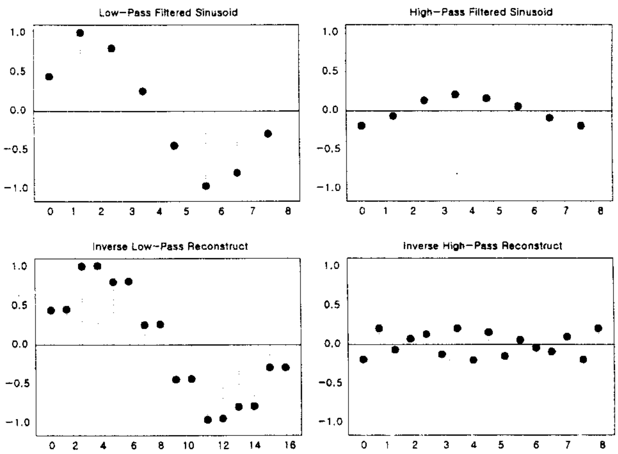
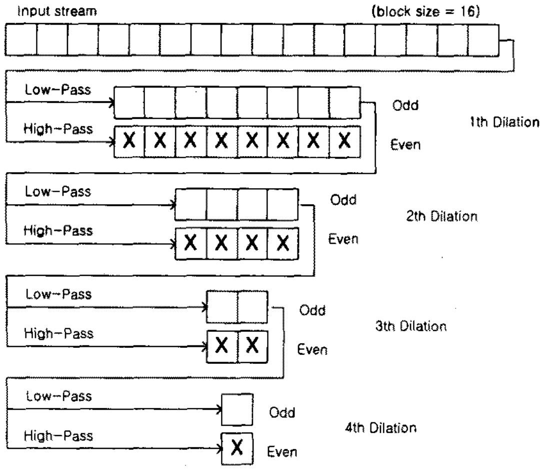


그림 [2]

2.2.1 Wavelet의 반복 수행에 의한 확장 (Dilation)

다음은 DWT을 반복 수행하여 얻어 질 수 있는 압축에 관해 기술하기로 하겠다. 여기서 일단 구현상의 이해와 편이를 위하여 Low-Pass 출력 값을 홀수(Odd)값으로 High-Pass 출력 값을 짝수(Even) 값으로 명칭하기로 하겠다. 다음의 그림은 Wavelet 변환의 확장 개념을 Block Diagram을 통해 기술한 것이다.

Wavelet Transfotrm Dilations



Wavelet	c_0	c_1
Haar	1.0	1.0

Wavelet	c_0	c_1	c_2	c_3
Daubechies-4	¼(1+√3	¼(3+√3	¼(3-√3	¼(1-√3

Wavelet	c_0	c_1	c_2	c_3	c_4	c_5
Daubechies - 6	0.332671	0.806891	0.459877	-0.13501	-0.08544	0.035226

표 (1)

3. 결론

본 논문에서 제시하였던 DWT를 통한 데이터의 기호화 실험은 Audio 데이터 중에서 일반 음악 부분과 음성 신호에 나누어 실험을 해 보았다. 연산 자체가 간편함으로 많은 양의 데이터를 빠른 시간에 처리할 수 있었다. 그리고 데이터 자체만을 직접 압축하는 방법을 통해 DWT가 어느 정도의 손실 속에서 압축이 가능한가를 실험하여 보았다. 다음은 음성 데이터를 Windows PCM Wave 파일을 통해 직접 압축한 것에 대한 PSNR 측정치이다.

압축률	x 2	x 4	x 8	x 16	x 32
PSNR	38.58	34.33	30.57	26.55	22.79

위에서 알 수 있듯 DWT에 의한 Audio 데이터의 변환 기호화의 경우 빠르고 간단한 편리함을 가지고 있으며 주파수 대역이 낮은 쪽에 몰여 있는 경우 그 압축 효율이 높아짐을 알 수 있었다. 세밀한 데이터의 압축을 위해서는 각각의 데이터(Frame or Segment) 마다 기초 함수와 압축 비율을 다르게 적용, 가장 효율적이고 적절한 압축을 얻을 수 있을 것이라 생각한다. 그리고 계속적인 계수의 개발에 있어 주파수 대역과의 관계를 통한 기초 함수의 개발이 가장 중요한 과제가 되리라 본다.

[참고 문헌]

[1] Strang, G. "Wavelets and Dilation Equations: A Brief Introduction" SIAM Review, vol 31, no.4, December 1989, pp 614-627

[2] Daubechies, I. "Orthonormal Bases of Compactly Supported Wavelets", Comm. Pure Applied Mathematics, vol 41, 1988, pp 909-996

[3] Johnson, J.D. "Transform Coding Of Audio Signals Using Perceptual Noise Criteria", IEEE Journal On Selected Areas In Communications, Vol.6, No.2, pp 314-323, Feb. 1988

[4] Martin Vetterli and Jelena Kovacevic "Wavelets And Subband Coding" Prentice Hall PTR

[5] Daubechies, I. "Ten Lectures on Wavelets", CBMS-NSF Series in Appl. Math.,SIAM, 1991