

Wavelet Denosing을 이용한 전력신호 압축

정영식
인천대학교 전기공학과

Power Signal Compression by using Wavelet Denosing

Young-Sik Chung
Dept. of Electrical Engineering, University of Incheon

Abstract - According to the characteristics of power signal, the compression algorithm based on wavelet packet transform and denoising is presented. Its performance is assessed in terms of percentage of zero coefficients and energy retained in reconstructed signals.

1. 서 론

전력시스템의 자동화가 급속히 개발됨에 따라 전력품질 신호 모니터로부터 획득되는 데이터양이 증가되어 분석, 특성 파악 및 분류 등에 대한 연구가 매우 활발하게 진행되고 있다. 따라서 방대한 양의 데이터를 저장하고 전송하는 문제가 대두됨에 따라 전력신호 데이터 압축에 대한 관심 또한 증가되어 높은 정확도와 효율적인 압축 및 저장에 대한 새로운 방법들을 전력관련업체들에서 요구하게 되었다. 압축은 음성신호 및 영상신호 등과 같은 디지털 신호처리 분야에서 적용되어 다양한 알고리즘들이 개발되었으며, 매우 활발하게 연구되고 있는 분야로 상당한 압축률과 정확도를 보여주고 있다.

데이터 압축은 같은 양의 정보를 더 작은 크기의 데이터로 표현하는 기법으로 무손실 압축과 손실 압축으로 크게 두 부류로 구분할 수 있다. 무손실 압축은 복원된 데이터가 원 데이터와 동일한 정보를 갖도록 하는 기법으로 압축률은 상대적으로 많이 떨어지나 정확한 정보를 복원해 낼 수 있다. Huffman 부호화, Lempel-Ziv 부호화 등이 있다. 손실압축은 복원된 데이터가 원 신호로 완전히 복원되지 않고 일정한 정도의 데이터 손실이 있지만 매우 높은 압축율을 얻을 수 있다. 예측 부호화, 변환 기법 등이 있다. 기본적으로 데이터 압축은 원 데이터가 갖고 있는 중복성을 제거하여 데이터 양을 줄이는 것으로 통계적 중복성, 공간적 중복성, 시간적 중복성 등이 있다.

전력 시스템에서 과도현상들의 데이터를 저장 전송하기 위해 사용된 압축은 푸리에 변환을 이용한 알고리즘, Huffman 부호화를 사용한 알고리즘, 웨이블릿 패키지를 사용한 변환 방법들이 사용되었다.

전력 시스템의 과도현상들에 대한 분석 등에 주파수 영역에서 신호들을 다루며, 다양한 주파수 대역들에 대한 정보를 얻기 위해 신호들을 분해한다. 실제 응용들에서 신호를 처리하는데 주파수 대역들의 모든 정보가 요구되는 건 아니기 때문에 전력품질 왜란이 발생된 부분들이 더 중요하고 다른 부분들은 왜란 해석에 덜 중요하다. 전력 데이터 압축은 가능한 한 중요한 특성들을 보유한 상태로 데이터 용량을 줄여서 전송 속도를 빠르게 하고 저장 용량을 줄이면서 복원시 왜곡 정도를 줄여서 데이터를 더 정확하게 빨리 분석할 수 있도록 하는 것이다.

특정 압축기법이 특정 신호에는 잘 적용될 수 있으나, 다른 신호에도 똑 같이 효율적으로 적용되는 것은 아니다.

웨이블릿 변환은 음성 및 영상 신호 압축에 적용되어 좋은 결과들을 보여주었으며, 최근에 전력품질 압축에 적용되어 왔다. 웨이블릿 변환 알고리즘 중의 하나인 웨이블릿 패킷 변환(wavelet package transform)은 신호를 다양한 대역으로 분해하여 하부 대역별 정보를 갖고 있는 계수들을 얻을 수 있으며, 이 계수들에 대해 압축을 시도한다.

전력품질 신호는 다른 신호와 다른 특성을 갖고 있다. 기본적으로 정현파로 구성된 신호에 짧은 기간 동안 비정상외 왜곡된 신호들이 중첩되어 있는 신호이다. 이러한 왜곡을 과도현상이라 하며, 이 현상의 주파수 스펙트럼은 전 주파수 영역에 걸쳐 퍼져있다. 전력품질 신호의 주요한 특성들은 이 과도현상에 포함되어 있어 이 부분에 대한 압축은 가능한 손실이 없게 압축을 시도해야 한다.

전력품질 신호에 웨이블릿 변환을 적용하여 얻게 되는 계수들은 왜란이 발생하는 곳의 계수들 값이 발생하지 않은 정상 신호보다 크게 나타나며, 계수들이 sparse한 성질을 갖고 나타난다. 이러한 성질을 이용하여 효율적인 전력신호 압축을 이룰 수 있다.

본 논문에서는 전력신호의 과도현상을 효율적으로 압축 전송할 수 있는 알고리즘을 제안한다. 웨이블릿 패킷 변환을 적용하여 왜란신호들을 다

수의 하부대역으로 분리한다. 이때 전력품질 신호를 획득하는 과정에서 발생하는 잡음들을 제거하기 위해 denoising 방법을 적용하여 잡음을 제거하는 동시에 중요도가 떨어지는 계수들을 제거하여 압축 효율을 높인다.

2. 웨이블릿 패킷 변환

웨이블릿 변환은 다양한 신호들의 데이터 압축 알고리즘에 성공적으로 적용되어 왔다. 이산 웨이블릿 변환은 직교변환(orthogonal transform)이기 때문에 신호를 직교 웨이블릿 계수들로 표현할 수 있어, 신호에 포함된 중복성이 제거가 용이하다. 또한 웨이블릿 변환은 웨이블릿의 양단이 소멸되기 때문에 에너지가 한 곳에 집중되는 성질을 갖고 있다. 다해상도 알고리즘(multi-resolution algorithm, MRA)을 적용하게 되면 각 해상도에 대해 최적의 압축을 얻을 수 있게 문턱값(threshold value)을 적용할 수 있다. 이산 웨이블릿 변환의 분해(decomposition) 트리구조는 입력신호 S 를 저역통과 필터를 적용하여 근사 성분에 해당되는 근사 계수 $A1$ 을 얻고 고역통과 필터를 적용하여 상세성분에 해당되는 상세 계수 $D1$ 을 얻으며, 다음 단계들에서도 상세 계수들에 대해서만 하부 단계를 내려가면서 근사 계수에 대해서만 저역통과 및 고역통과 필터를 적용하는 트리구조를 갖고 있다.

웨이블릿 패킷 변환(WPT)은 이산 웨이블릿 변환을 직접 확장한 이진 트리구조로 그림 1에 분해 트리구조를 보였다. WPT는 각 해상도에서 근사계수와 상세계수 모두에 대해 저역통과 필터와 고역통과 필터를 적용한다.

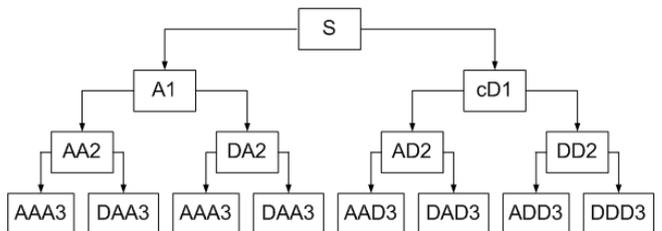


그림 1. WPT의 분해 트리구조

WPT의 장점은 응용에 적합한 웨이블릿 기저함수를 통해 신호를 더 정확하게 표현할 수 있다는 것이다. 참고문헌 [1]에서 제안한 가장 적합한 기저함수는 근사계수와 상세계수들로부터 엔트로피를 측정하여 얻는다. 신호 $x = \{x_0, x_1, \dots, x_{N-1}\}$ 에서 엔트로피는 다음과 같이 정의된다.

$$H(x) = - \sum_{n=0}^{N-1} |x_n|^2 \log |x_n|^2 \quad (1)$$

식 (1)은 Shannon 엔트로피로 알려져 있으며, 가장 최소값을 나타내는 기저함수가 가장 적합한 기저함수로 계수들로부터 최대의 정보를 얻을 수 있다.

잡음이 중첩되어 있는 신호에서 잡음을 제거하고 최적의 압축을 이룰 수 있는 가장 작은 수의 계수를 얻기 위한 기준으로 최소표현길이 (minimum description length, MDL)를 적용하였다[1]. 이산 신호

$$f = x + n \quad (2)$$

여기서 f 는 잡음이 중첩되어 있는 신호이며, x 는 추정할 정보를 갖고 있는 신호 벡터이며, n 은 잡음 벡터이다. Saito[1,2]에 의해 제안된 MDL은 인덱스 (k, n) 에서

$$MDL(k, n) = \min \left\{ \frac{3}{2} k \log N + \frac{N}{2} \log \|\tilde{\alpha}_n - \tilde{\alpha}_n^{(k)}\|^2 \right\} \quad (3)$$

$0 \leq k \leq N, 1 \leq n \leq M$

여기서 $\tilde{\alpha}_n = W_n f$ 는 웨이블릿 필터 n 을 f 에 적용하여 얻은 웨이블릿 계수들의 벡터를 의미하며, $\tilde{\alpha}_n^{(k)}$ 는 문턱 값을 적용하여 얻은 0이 아닌 웨이블릿 계수들의 벡터이다. MDL이 최소가 되는 k 의 값이 웨이블릿 계수들로 표현할 수 있는 최적의 길이가 된다. 참고문헌 [1]에서 다수의 웨이블릿 기저함수, 즉, Daubechies(2,4,6,8,10, 12,14,16,18,20), Coiflets(2,4,6,8,10) 및 Symlets(4,6,7,8)을 적용하여 엔트로피와 최소표현길이를 기준으로 최적 기저함수의 후보로 Symlet 7과 8을 제안하였다.

3. 웨이블릿 denoising

잡음은 모든 주파수에 영향을 미치며, 신호는 대부분의 경우 주파수영역에서 작은 부분에 존재하고 있다. 따라서 특정 주파수 성분 위에 있는 높은 주파수 성분들은 잡음에 의한 것으로 고려할 수 있다. 신호 f 에 웨이블릿 변환을 적용하여 구한 계수를 d 라 하면,

$$d = Wf = \theta + \eta \quad (4)$$

을 얻을 수 있으며, 계수 d 는 잡음이 없는 상태의 신호에서 구한 계수 θ 와 잡음 η 합으로 표현될 수 있다. 여기서 W 는 웨이블릿 변환이며, η 는 i.i.d이며, 가우시안 분포를 갖고 있다. 식 (4)로부터 잡음이 존재하는 신호에서 얻은 웨이블릿 계수에는 잡음 성분에 의한 양이 포함되어 있음을 볼 수 있다.

선형 다해상도 해석의 근사화는 데이터를 평탄화하는 경향이 있으며, 계수들이 잡음에 영향을 받으므로, 이러한 단점을 극복하기 위해 Donoho와 Johnston[3]은 데이터의 작은 도약 또는 내려앉음을 정확하게 표현하기 위한 비선형 근사화 방법인 shrinkage를 바탕으로 한 웨이블릿 변환을 제안하였다. 크기가 큰 웨이블릿은 중요한 성분으로 선택하고 잡음을 나타내는 비교적 작은 웨이블릿 계수들은 0으로 설정하는 비선형 방법은 더 작은 수의 계수들로 신호를 나타낼 수 있는 효과적인 방법이다. Denoising 방법은 우선 웨이블릿 계수 벡터에서 잡음에 기여되는 부분이라고 생각되는 곳을 문턱 값을 적용하여 제거한다.

$$\tau = Th(d) \quad (5)$$

여기서 $Th(\cdot)$ 는 웨이블릿 계수 벡터에서 문턱 값을 구하는 연산자이다. 문턱 값 τ 를 적용하여 denoising 연산자 $D(\cdot, \cdot)$ 를 적용한다.

$$z = D(d, \tau) \quad (6)$$

웨이블릿 역 변환을 적용하여 z 로부터 재구성된 신호를 얻을 수 있다.

$$\tilde{f} = W^{-1}z \quad (7)$$

웨이블릿 denoising은 그림 2에 보여진 hard thresholding(HTH)와 soft thresholding(STH) 두 방법을 사용한다.

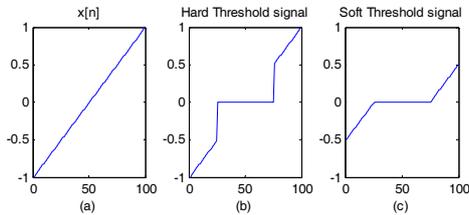


그림 2. Hard thresholding and Soft Thresholding.

HTH는 특정 문턱 값보다 작은 웨이블릿 계수들을 0으로 설정하는 것으로 비교적 간단한 방법이다. 즉,

$$D(d, \tau) = \begin{cases} d & |d| > \tau \\ 0 & \text{그 이외} \end{cases} \quad (8)$$

SHT는 계수들의 절대 값이 특정 문턱 값보다 크면 일정한 양만큼 줄이고, 아래에 존재하면 0으로 설정하는 방법이다. 즉,

$$D(d, \tau) = sgn(d) \max(0, |d| - \tau) \quad (9)$$

여기서 $sgn(\cdot)$ 는 웨이블릿 계수의 부호를 나타내는 연산자이다.

4. 모의실험 및 고찰

실험에 사용되는 그림 3(a)의 전력신호 파형은 전압상승, 강하 및 커패시터 스위칭 현상이 중첩되어 있는 신호이며, 그림 3(b)는 백색잡음이 중첩된 신호로 이 신호에 알고리즘을 적용한다. 웨이블릿 기저함수로는 제안한 Symlet 7을 사용하며, 다해상도 단계는 엔트로피가 최소가 되는 2단계까지 적용한다. Denoising은 STH를 사용하여 적용하였으며, 문턱 값은 다음 식을 통해 결정한다.

$$\tau = (1 - \mu) \max\{|d|\} \quad (10)$$

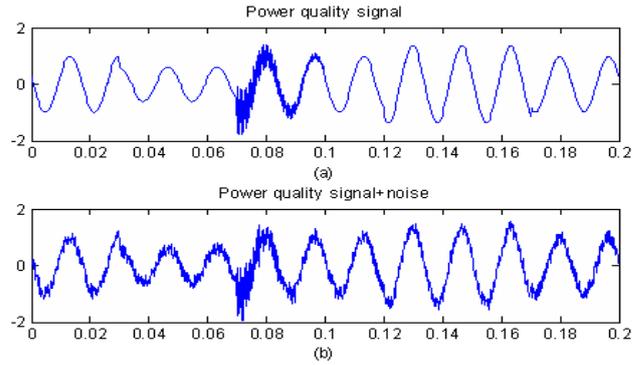


그림 3. (a) 전력신호 (b) 잡음이 중첩된 전력신호

식 (1)의 문턱 값에서 $\mu = 0.5$ 로 설정하여 알고리즘을 수행한 결과 0의 값을 갖는 비율이 73.47%였으며, 압축된 신호가 보유한 에너지는 97.17%로 그림 4(a) 신호는 압축된 결과로부터 복구한 신호로 그림 3(b)의 신호와 비교해보면 대부분의 잡음이 제거되었으며, 그림 3(a) 전력신호의 특성들을 그대로 보유하고 있음을 볼 수 있다. 따라서 denoising 효과를 볼 수 있으며, 그림 3(a)의 에너지를 그대로 보유하고 있다.

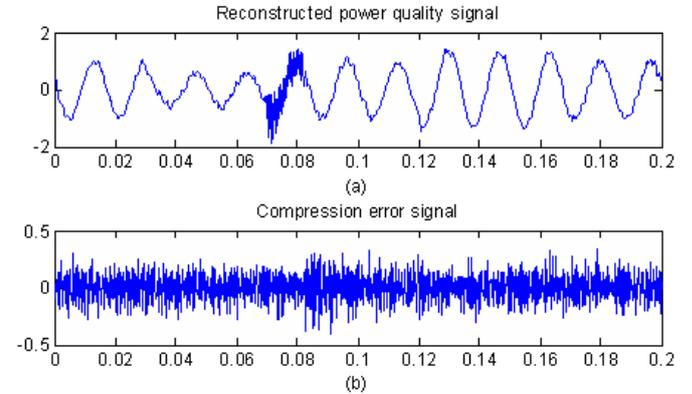


그림 4. (a) 재구성된 신호 (b) 압축오차 신호

5. 결론

웨이블릿 패킷 변환에 denoising을 결합한 전력신호 압축을 알고리즘을 제안하였다. 외부에서 중첩된 잡음과 계수에 반영된 잡음이 상당부분 제거되어 잡음이 제거되는 결과를 보여주었으며, 0의 값을 갖는 계수의 백분율이 매우 높으며, 에너지 보유율도 매우 높아 압축과 잡음제거 효과를 제안한 알고리즘을 통해 얻을 수 있었다. 이 계수들에 호프만 부화 또는 벡터 양자화 부호법 등을 통해 더욱 압축률을 높일 수 있을 것이다.

[참고 문헌]

- [1] E. Y. Hamid and Z. Kawasaki, "Wavelet-Based Data Compression of Power System Disturbances Using the Minimum Description Length Criterion", IEEE Tran. Power Delivery, Vol. 17, No. 2, pp. 460-466, 2002.
- [2] N. Saito, Simultaneous Noise Suppression and Signal Compression Using a Library of Orthonormal Bases and the Minimum Description Length Criterion: Wavelets in Geophysics, New York, Academic Press, 1994.
- [3] B. Vidakovic, Statistical Modeling by Wavelets, John Wiley & Sons, 1999.
- [4] D. L. Donoho and I. M. Johnstone, "Adapting to Unknown Smoothness in Wavelet Shrinkage", Journal of American Statistical Association, Vol. 90, 1995.
- [5] S. K. Meher, A. K. Pradhan and G. Panda, "An Integrated Data Compression Scheme for Power Quality Events using Spline Wavelet and Neural Network", Electric Power Systems Research Vol. 69, pp.213-220.