

# 시계열 데이터의 추정을 위한 웨이블릿 칼만 필터 기법

## The wavelet based Kalman filter method for the estimation of time-series data

홍 찬 영\*, 윤 태 성\*\*, 박 진 배\*

(Chan Young Hong, Tae Sung Yoon and Jin Bae Park)

\* 연세대학교 전기전자공학과(전화:(02)2123-2773, 팩스:(02)362-0413, E-mail : cyhong@control.yonsei.ac.kr)

\*\* 창원대학교 전기공학과(전화:(055)279-7513, 팩스:(055)263-9956, E-mail : tsoon@sarim.changwon.ac.kr)

**Abstract** : The estimation of time-series data is fundamental process in many data analysis cases. However, the unwanted measurement error is usually added to true data, so that the exact estimation depends on efficient method to eliminate the error components. The wavelet transform method nowadays is expected to improve the accuracy of estimation, because it is able to decompose and analyze the data in various resolutions. Therefore, the wavelet based Kalman filter method for the estimation of time-series data is proposed in this paper. The wavelet transform separates the data in accordance with frequency bandwidth, and the detail wavelet coefficient reflects the stochastic process of error components. This property makes it possible to obtain the covariance of measurement error. We attempt the estimation of true data through recursive Kalman filtering algorithm with the obtained covariance value. The procedure is verified with the fundamental example of Brownian walk process.

**Keywords** : estimation, Kalman filter, error covariance, wavelet coefficient

### 1. 서론

대부분의 경우에 있어서 분석에 필요한 데이터들은 시간에 따라서 얻는 시계열 데이터(time-series data)이다. 그러나 데이터를 얻는 측정 과정에서 측정 오차(measurement error)가 발생하기 때문에, 측정 데이터를 바탕으로 오차 요소를 제거하여 참값을 추정(estimation)하는 과정이 필요하다.

근래에 들어 활발히 연구되고 있는 웨이블릿 변환(wavelet transform) 이론은 Daubechies 등의 공헌으로 발전하였으며, 신호 데이터를 다양한 차원에서 분해하여 해석한다. 웨이블릿 변환은 신호들을 어떤 시간 구간에 대해 국소화(localization)하고, 확장되거나 축소될 수 있는 구획들로 분해시킴으로써 다른 스케일(scale)의 해상도에서 신호를 분석하도록 한다. 이러한 특성 때문에 웨이블릿은 영상 신호의 압축이나 디지털 신호의 잡음 제거, 또는 여러 알고리즘 수행의 속도 증진 등에 사용되고 있다[1]-[2].

한편, 칼만 필터 기법은 신호의 통계적 특성을 바탕으로 참값을 추정하는 전통적인 방법이다. 칼만 필터 기법에서는 프로세스(process)와 측정(measurement) 과정을 상태 방정식으로 표현하고 각각의 과정에서 발생하는 오차의 공분산(covariance)을 통해 반복적으로 참값을 추정해나간다. 그러나 칼만 필터를 적용하기 위해서는 오차의 공분산 값에 대한 선행 지식이 있어야 하므로, 이 값을 추정하기 위한 다양한 방법론이 제안되었다[3]-[5].

본 논문에서는 칼만 필터의 오차 공분산 값의 결정에 웨이블릿 변환을 이용하는 간단한 방법을 제시한다. 측정된 신호를 웨이블릿 분해(decomposition)하여 얻을 수 있는 웨이블릿 계수는 주파수 대역별로 나뉘어진 근사(approximation) 신호와 세부(detail) 신호의 정보를 담고 있다. 세부 신호의 웨이블릿 계수는 오차에 대한 통계적 추이를 반영하고 있으므로, 이 계수들의 분산값을 구하여 칼만 필터 알고리즘에 필요한 파라미터로 이용한다.

2장에서는 칼만 필터와 웨이블릿 변환에 대한 기본 이론을 살펴보고 제안된 방법을 설명한다. 3장에서는 대표적인 임의 데이터 집합인 브라운 과정(Brownian process)을 대상으로 모의실험을 하며, 4장에서 결론을 제시한다.

### II. 웨이블릿 변환과 칼만 필터 기법

웨이블릿 변환은 다루는 데이터에 따라 연속시간 웨이블릿 변환과 이산시간 웨이블릿 변환으로 나눌 수 있다. 본 논문에서는 이산시간 웨이블릿 변환으로 데이터를 분석한다.

웨이블릿 변환은 웨이블릿 함수라고 불리는  $\psi(t)$ 를 빌딩블록(building block) 함수로 선정하여 주어진 신호  $f(t)$ 를 재구성하는 방법이다. 이 때 기본 함수  $\psi(t)$ 는 식 (1)과 같이 팽창(dilate) 또는 수축(contract)으로 척도변형(scaling)되고 평행이동(shift)된 여러 개의 빌딩블록으로 변경된다.

$$\psi_{k,j} = 2^{k/2} \psi(2^k t - j), \quad k, j \in Z \quad (1)$$

where  $k$ : scaling factor,  $j$ : shifting factor

웨이블릿 함수  $\psi$ 와 마찬가지로 스케일링 함수  $\phi$ 도  $\phi_{k,j}$ 와 같은 형태로 표현한다. 이 때 웨이블릿 변환을 통해 함수  $f(t)$ 는 식 (2)와 같은 형태로 표현된다.

$$f(t) = \sum_j c_{k0,j} \phi_{k0,j} + \sum_{k=k_0}^{N-1} \sum_j d_{k,j} \psi_{k,j} \quad (2)$$

근사 계수  $c_{k,j}$ 는  $f(t)$ 의 근사화 정보를 가지고 있으며, 세부 계수  $d_{k,j}$ 는  $f(t)$ 의 세부 정보를 가지고 있다. 각 계수는 식 (3)과 같이 구할 수 있다.

$$\begin{cases} c_{k,j} = \langle f, \phi_{k,j} \rangle = \sum_{l \in Z} h_{l-2j} c_{k+1,l} \\ d_{k,j} = \langle f, \psi_{k,j} \rangle = \sum_{l \in Z} g_{l-2j} c_{k+1,l} \end{cases} \quad (3)$$

$\langle a, b \rangle$ 는 두 함수  $a$ 와  $b$ 의 내적을 나타내며,  $g$ 와  $h$ 는 각각  $\phi(t)$ 에 관련된 HPF 계수와  $\psi(t)$ 에 관련된 LPF 계수를 의미한다.

웨이블릿 변환에 의해 함수는 여러 단계로 분해될 수 있으며 이 과정을 도식적으로 나타내면 다음의 그림 1과 같다.

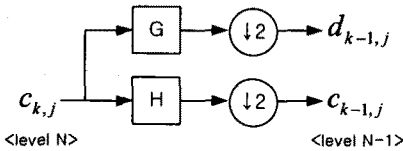


그림 1. 웨이블릿 분해  
Fig. 1. wavelet decomposition

일반적인 이산 시간 칼만 필터는 식 (4)와 같은 상태 방정식으로 표현된다.

$$\begin{cases} \mathbf{x}_{t+1} = \mathbf{F}_t \mathbf{x}_t + \mathbf{w}_t, & E\{\mathbf{w}_t^2\} = Q \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{v}_t, & E\{\mathbf{v}_t^2\} = R \end{cases} \quad (4)$$

- where  $\mathbf{x}_t$ : 상태변수 벡터  
 $\mathbf{z}_t$ : 측정치 벡터  
 $\mathbf{w}_t$ : 시스템 잡음 (white sequence)  
 $\mathbf{v}_t$ : 측정 잡음 (white sequence)  
 $\mathbf{F}$ : 상태변수  $\mathbf{x}_{t+1}$ 과  $\mathbf{x}_t$ 의 관계 행렬  
 $\mathbf{H}$ : 상태변수와 측정치와의 관계 행렬

칼만 필터(KF)에서는 보통 측정시 생기는 측정 잡음  $\mathbf{v}_t$ 의 공분산값을 안다고 가정하고 시작하지만, 실제로는 측정 데이터에서 공분산값을 미리 알 수 없는 경우가 있다. 이 때 측정 잡음이 백색성(whiteness) 성질을 가지고 있다면, 웨이블릿 변환(WT)을 통해 세부 계수를 구함으로써 그 값을 얻어낼 수 있다. 그림 2는 이러한 방법으로 원래 데이터의 추정치  $\hat{x}$ 를 구하는 과정을 블록 다이어그램으로 나타낸 것이다.

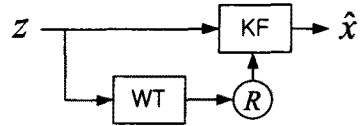


그림 2. 제안된 WT 기반 KF  
Fig. 2. Proposed WT based KF

### III. 모의 실험

제안된 방법의 검증을 위해 기본적인 임의의 신호인 Brownian walk 데이터를 대상으로 모의실험을 하였다. 1차원 Brownian walk는 다음의 식 (4)와 같이 표현되며 추정 과정은 식 (5)와 같이 모델링된다.

$$\begin{cases} \mathbf{x}(N, t+1) = \mathbf{x}(N, t) + \mathbf{w}(N, t) \\ E\{\mathbf{w}(N, t)\} = 0, \quad E\{\mathbf{w}^2(N, t)\} = Q \end{cases} \quad (4)$$

$$\begin{cases} \mathbf{z}(N, t+1) = \mathbf{x}(N, t) + \mathbf{v}(N, t) \\ E\{\mathbf{v}(N, t)\} = 0, \quad E\{\mathbf{v}^2(N, t)\} = R \end{cases} \quad (5)$$

where  $N$ : the highest resolution level  
 $t$ : time index

각각의 모델에서  $\sqrt{Q} = 0.1$ ,  $\sqrt{R} = 3$ 으로 설정한다. 위 모델에서 생성된 데이터의 참값  $x(N, k)$ 와 측정값  $z(N, k)$ 는 그림 3에 각각 나타내었다. 측정 데이터를 대상으로 Haar filter bank를 사용한 3단계 웨이블릿 변환을 수행한 후 웨이블릿 계수의 추이를 살펴보면 그림 4와 같다. 각 해상도 단계별 세부 웨이블릿 계수의 분산은 다음과 같이 비슷한 값으로 계산된다.

$$\sigma_{D1}^2 = 3.1682, \quad \sigma_{D2}^2 = 2.9412, \quad \sigma_{D3}^2 = 3.1239$$

여기서는 1단계 웨이블릿 변환으로 얻어진 세부 웨이블릿 계수의 분산값 3.1682를 측정 오차의 공분산값으로 설정하여 칼만 필터 알고리즘에 대입하였다. 칼만 필터를 통해 추정된 근사치는 그림 5 (a)에 제시되었

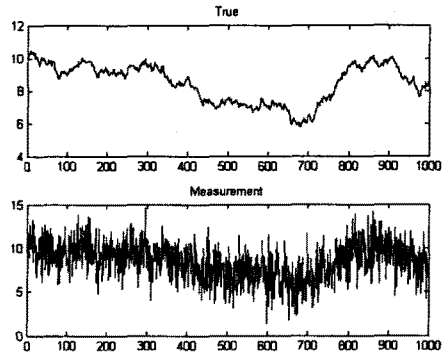


그림 3. Brownian walk의 참값과 측정값  
Fig. 3. True & measurement quantities of Brownian walk

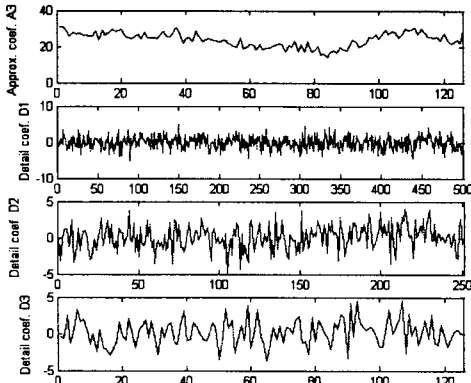


그림 4. 측정치를 3단계 웨이블릿 변환한 후의 웨이블릿 계수

Fig. 4. Wavelet coefficients after 3-level wavelet transform with measurement data

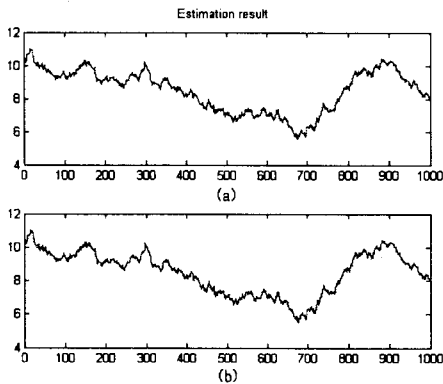


그림 5. 찾아낸 공분산으로 추정한 결과(a)와 알고 있는 공분산으로 직접 추정한 결과(b)  
Fig. 5. Estimation results from figured covariance (a) and known covariance (b)

다. 이와 비교하여, 원래 설정했던 측정오차의 공분산값 3을 알고 있다는 가정 하에 바로 칼만 필터 알고리즘을 수행한 결과는 그림 5 (b)에 제시하였다. 두 결과의 평균 제곱근 오차(RMS error)를 구하면 각각 0.3043과 0.3044로 유사하다. 위와 같은 시뮬레이션을 100번 시행하여 평균한 RMS error도 0.4176으로 차이가 없으므로, 제안된 방법을 통한 측정 오차의 공분산 추정은 정확하다고 할 수 있다.

#### IV. 결론

본 논문에서는 측정 오차가 섞인 시계열 데이터의 참값 추정 문제에 있어서 웨이블릿 변환을 통한 칼만

필터 방법을 적용하였다. 측정 데이터를 웨이블릿 변환하여 세부 웨이블릿 계수를 구하고 그 분산값을 얻어내어 칼만 필터의 오차 공분산 파라미터로 이용하였다. 이러한 방법은 측정 오차의 분산을 알지 못하는 데이터에 대해서도 칼만 필터의 알고리즘을 통해 효과적으로 참값을 추정해낼 수 있도록 한다.

#### 참고문헌

- [1] Graps, A, "An introduction to wavelet", Computational Science and Engineering, Volume: 2 Issue: 2, pp. 50-61, Summer 1995.
- [2] Sarkar, T.K.; Su, C.; Adve, R.; Salazar-Palma, M.; Garcia-Castillo, L.; Boix, R.R "A tutorial on wavelets from an electrical engineering perspective. I. Discrete wavelet techniques", Antennas and Propagation Magazine, IEEE, Volume: 40, Issue: 5, pp.49-68, Oct. 1998.
- [3] Mehra, R. "On the identification of variances and adaptive Kalman filtering" Automatic Control, IEEE Transactions on, Volume: 15 Issue: 2 pp. 175-184, Apr. 1970
- [4] Godbole, S. "Kalman filtering with no a priori information about noise -White noise case: Identification of covariances" Automatic Control, IEEE Transactions on, Volume: 19, Issue: 5, pp. 561-563, Oct. 1974
- [5] Lee, T. "A direct approach to identify the noise covariances of Kalman filtering" Automatic Control, IEEE Transactions on, Volume: 25, Issue: 4, pp. 841-842, Aug. 1980.
- [6] Hong, L.; Cheng, G.; Chui, C.K. "A filter-bank-based Kalman filtering technique for wavelet estimation and decomposition of random signals" Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on, Volume: 45, Issue: 2, pp. 237-241, Feb. 1998.
- [7] Hirchoren, G.A.; D'Attellis, C.E. "Estimation of fractional Brownian motion with multiresolution Kalman filter banks" Signal Processing, IEEE Transactions on, Volume: 47, Issue: 5, pp. 1431-1434, May. 1999.
- [8] Tongxin Zheng; Adly A. Girgis; Elham B. Makram "A hybrid wavelet-Kalman filter method for load forecasting" Electric Power Systems Research, Volume: 54, Issue: 1, pp. 11-17, Apr. 2000.