

웨이블릿 패킷 분해를 이용한 EEG 신호압축

조현숙*, 이 형*, 황선태*

EEG Data Compression Using the Feature of Wavelet Packet Coefficients

HyunSook Cho, Hyoung Lee, Suntae Hwang

Abstract

This paper is concerned with the compression of EEG signals using wavelet-packet based techniques. EEG data compression is desirable for a number of reasons. Primarily it decreases for transmission time, archival storage space, and in portable systems, it decreases memory requirements or increases channels and bandwidth. Upon wavelet decomposition, inherent redundancies in the signal can be removed through thresholding to achieve data compression. We proposed the energy cumulative function for deciding of the threshold value and it works very innovative of EEG data.

Keywords : EEG 신호 압축(EEG data compression),
웨이블릿 패킷 변환(wavelet Packet Transform)

1. 서 론

EEG는 뇌의 매커니즘을 이해하고 가능한 병적 증상을 이해하기 위한 중요 신호이다. 그러나 이러한 목적을 위하여 데이터의 양이 너무나 방대해서 샘플링 율을 최소로 낮춘다 하더라도 압축은 불가피하다. 따라서 EEG 데이터의 전송 시간이나 저장 공간을 감소시키고 필요한 채널이나 대역폭을 감소시키는 등의 압축에 관한 연구가 진행되어 왔다[1]. EEG 신호의 압축에 있어서 중요한 것은 의사가 신호의 흐름을 보고 환자의 병증을 판단하거나 뇌의 상태를 결정짓는데 필요한 정보를 손실해서는 안 된다는 것이다.

웨이블릿은 EEG와 같은 비정상 신호의 해석에 적합한 것으로[2] 주어진 신호를 웨이블릿 계수들로 표현하고 이를 해석하는 과정에서 생기는 신호에 영향을 미치지 않는 계수들을 제거하는 방식으로 수행된다. 임계값(threshold) 아래로 떨어지는 값을 잘라내어 신호를 표현하기 위한 최소의 데이터만 얻어 내게 된다. 이러한 기법이 웨이블릿을 데이터 압축 분야에서 훌륭한 도구로 사용하도록 만들었다[3].

본 논문에서는 웨이블릿 패킷 변환을 이용하여 신호의 중복을 제거하기 위한 임계값을 설정하는 방법을 제안하였다. 웨이블릿의 계수들이 분해되는 단계에서 그 값이 작아져 거의 "0"에 가까워지는 특성을 이용하고, 에너지 축적 함수를 이용하여 임계값을 설정하였다. 기존의 기본 웨이블릿 변환과 비교하여 패킷 변환을 사용함으로써 신호의 압축률을 향상시켰다.

본 논문의 내용은 제 2장에서 뇌파와 EEG 신호에 관한 내용을 설명하였고, 제 3

장에서 기본 웨이블릿 변환과 이를 일반화시킨 웨이블릿 패킷 변환을 비교 설명하였다 [4,5]. 제 4 장에서는 제안하는 EEG 데이터의 압축 방법을 설명하였고, 웨이블릿 변환과 웨이블릿 패킷 변환을 이용하여 EEG 신호를 압축한 결과를 비교하고 앞으로의 연구 과제에 관하여 설명하였다.

2. 뇌파의 웨이블릿 분석

2.1 뇌파

뇌파는 뇌피질 표면이나 두피상에 전극을 놓고 검출되는 자발성, 율동성의 전기 현상으로 일반적으로 1Hz에서 60Hz 까지의 주파수 범위에서의 파동을 말한다. 뇌파의 전압은 $100\sim 300\mu V$ 이고 시상수는 0.1초 또는 0.3초, 고역 차단 여파기에서 15Hz 또는 25Hz 이상인 성분을 제거하게 된다. 현재 뇌 신경세포의 활동상황을 검출할 수 있는 방법으로 사용되고 있는 장치는 Electric Potential을 검출하는 EEG, 신경전류에 의한 자기장을 검출하는 MEG(Magneto EncephaloGraphy), PET(Positron Emission Tomography), 신경 세포 활동 시 증가하는 산소소모량의 증가를 영상화하는 functional-MRI, 신경세포 활동 시 산소소모량을 적외선을 이용하여 검출하는 NIRS(Near Infrared Spectroscopy) 등이 있다.

뇌파의 데이터 처리는 목적이나 규모에 따라 다양하다. 주된 방법들로는 첫째, 뇌파 분석기라는 대역 필터를 사용한 주파수 분석 장치로 기초 파동을 정량화, 객관화하는 방

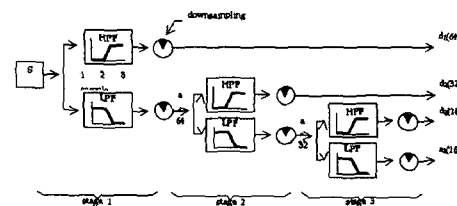
법, 둘째 뇌파 유발 전위를 평균 가산 전용 컴퓨터로 검출하는 방법[6], 셋째 상관함수, 파워 스펙트럼, 고차 모멘트, 코히어런스 함수 등을 계산해서 뇌파의 통계적 특징을 추출하는 방법[7,8,9] 및 패턴인식과 의사의 진단 이론에 기초한 뇌파의 자동해독 방법 등이다.

본 논문은 이와 같은 데이터 처리에 있어서 방대한 양의 뇌파기록을 저장하고 전송하는 일에 초점을 맞추어 데이터를 압축하고 복원한 후 스펙트럼을 분석하여 보았다. EEG 신호의 경우 환자의 병증을 EEG 신호 자체를 가지고 평가하므로 복원 후의 신호가 원 신호와 가시적인 차이를 보이지 않도록 압축하는 방법이 중요하다고 하겠다.

2.2 표준 웨이블릿 분해

웨이블릿은 광범위한 스펙트럼을 가진 신호를 분석하는데 적당한 함수들의 모임이라 할 수 있다. 이러한 특성으로 신호처리 분야, 특히 의용 공학(biomedical) 분야에 적용되고 있다. Gabor(1946)는 푸리에 변환의 시간 영역의 신호를 주파수 영역으로 분석하는 과정에서 생기는 시간에 대한 정보를 잃게 되는 단점을 극복하고자 고정된 크기의 윈도우 사이즈를 통하여 신호를 분석하는 방법(STFT : Short Time Fourier Transform)을 제안하였다. 웨이블릿 분석은 이를 발전시켜 다양한 크기의 윈도우를 사용하여 신호의 시간-주파수 지역화를 가능하게 하여 EEG와 같은 비대칭적이고 불규칙적이며 빠른 변화를 보이는 신호를 해석하는데 유용하게 사용되고 있다. 웨이블릿 분석은 어떤 신호를 모 함수 (original wavelet)의 이동(shift)과 팽창(scale)으로 얻어지는 신호의 가중 합으

로 정보를 손실 없이 표현한다. 효과적인 분석을 위하여 이동과 팽창은 크게 두 부분으로 양분된다. 즉 필터뱅크 알고리즘과 쿼드미러 필터를 통하여 신호가 계층적으로 나뉘어지는데 각 단계에서 다운 샘플링된 신호는 각각 고주파 필터와 저주파 필터를 거친 형태로 나타난다. 근사 정보(approximation)는 신호가 저주파 필터를 통과한 결과를 나타내며 상세 정보(detail)는 신호가 고주파 필터를 통과한 결과이다. 다음 [그림 1]은 신호를 3단계 DWT(Discrete Wavelet Transform)를 수행하는 과정이다.



[그림 1] 웨이블릿 분석
(Wavelet Decomposition)

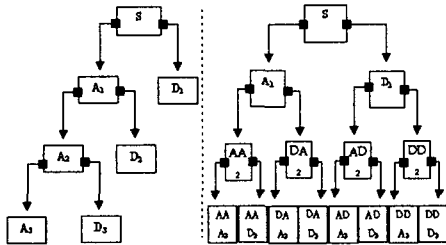
2.3 웨이블릿 패킷 분해

표준 웨이블릿은 저주파 필터를 통과한 신호의 다중 다운샘플링을 통하여 원 신호를 표현하기 때문에 고주파 성분에 중요한 정보를 담고 있을 경우에는 유용하지 못하다. 기본 웨이블릿 변환을 일반화한 패킷 변환은 시간-주파수 지역성을 조절하여 특성에 맞는 분할을 가능하게 한다. 3-단계로 신호를 웨이블릿 분석하면 입력신호는

$$S = A3 + D3 + D2 + D1$$

과 같이 나타낼 수 있다. 이 과정을 통하여 신호는 자연스럽게 벡터 $a3, d3, d2, d1$ 을 생성하여 웨이블릿 계수에 대응하게 된다. 이

계수들은 아래의 [그림 2]과 같이 주어진 스케일에 따라 서로 다른 주파수로 신호의 정보를 나타낸다. 그리고 이들 대역의 크기는 서로 다르게 나타난다.



[그림 2] 웨이블릿 분해와 패킷 분해의 비교

패킷 분해의 경우 주어진 신호는 웨이블릿 분해보다 더욱 다양한 방법으로 신호를 표현할 수 있다. 예를 들어 같은 방법으로 3-단계 분해를 수행하면 입력신호

$$S = A1 + AD2 + DDA3 + DDD3 \text{ 또는 } S = AA2 + DA2 + ADA3 + ADD3 + DDA3 + DDD3 \text{ 등으로 표현할 수 있다.}$$

웨이블릿 패킷 분해는 기본 분해 방법과 달리 각 단계에서 동일한 주파수 대역으로 신호를 분할하고, 서로 다른 웨이블릿 패킷의 기저로 원 신호를 표현할 수 있다. 이 때 모든 표현 방법 중에서 가장 효율적으로 신호를 나타낼 수 있는 기저를 선택하는 것이 가장 중요하다고 할 수 있다. 여기서 “효율적”이란 원 신호를 적은 수의 웨이브 형태, 또는 웨이블릿 패킷으로 표현한다는 의미이다 즉, 신호의 분해를 위한 기저가 계수들의 웨이트에 따라 웨이블릿 패킷의 최소 개수로 집중되고, 계수 중 가장 큰 수가 “0”에 근접하도록 선택되어야 한다는 것이다. 주어진 신호에서 최적의 기저를 선택하기 위한 일반적인 방법으로 최소 엔트로피 기준법 (Coifman and Wickerhauser (1992)[10])이

사용된다. 최소 엔트로피 기준법은 각 단계의 웨이블릿 계수를 $\{V_i\}$ 라 할 때, 엔트로피는 다음과 같이 정의된다.

$$E^2(V_i) = - \sum_i \frac{v_i^2}{\|v\|^2} \log_2 \frac{v_i^2}{\|v\|^2}$$

(단, $\|v\|^2 = \sum_i v_i^2$) (1)

최적의 기저는 최소의 엔트로피를 만들어야 하는 것이라 할 때, 엔트로피는 몇 개의 유효한 요소가 특정 기저 안에서 신호 표현에 필요한가를 나타낸다고 할 수 있다. 따라서 엔트로피가 작으면 작을수록 신호를 표현하는데 필요한 특정 계수는 작아진다.

3. 제안하는 압축 알고리즘

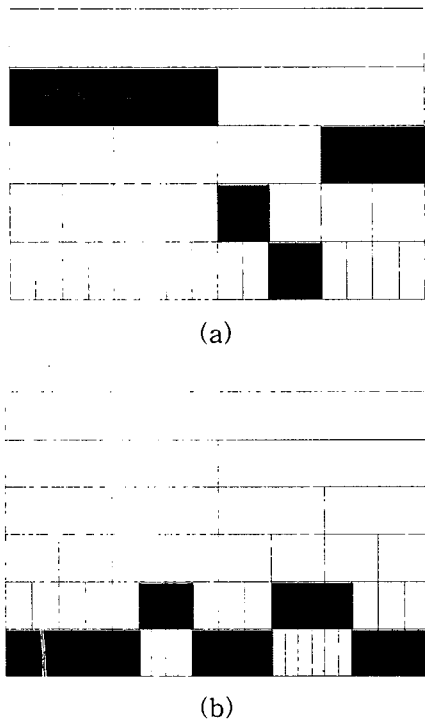
기본적으로 웨이블릿 분해 후 계수를 살펴보면, 많은 공간에 작은 에너지를 가지고 있는 값들이 존재한다. 이 값들은 원 데이터에는 거의 영향을 미치지 않으므로 임계값을 부여하여 제거하는 방식으로 신호를 압축하게 된다. 본 논문은 이 임계값을 결정하기 위하여 계수의 에너지 축적 함수를 만들어 총 에너지가 99% 이내에 포함되는 계수들을 압축하였다. 제안하는 압축 알고리즘을 단계별로 살펴보면 다음과 같다.

[1단계] 최적의 기저를 찾아 웨이블릿 패킷 변환(4-단계)을 수행한다.

$$DWPD(Y) = W_y \tag{2}$$

(단, $DWPT$: Discrete Wavelet Packet Decomposition, Y : 원신호, W_y : 웨이블릿 변환계수)

최적의 기저를 찾는 데는 2절에서 소개한 최소 엔트로피 기준법을 사용하였다. 다음 그림은 동일한 신호에 대하여 4 단계 웨이블릿 패킷 분해를 수행하여 최적의 기저를 찾은 결과를 나타내고 있다.

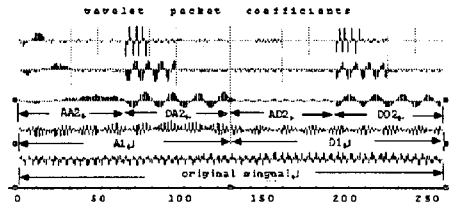


[그림 3] 필터에 따른 최적 기저

위의 [그림 3]에서 (a)는 Coeflit4 필터를 사용한 결과이고 (b)는 LeastAsymmetric 필터를 사용한 결과이다. 필터에 따라서 최적의 기저는 달라지게 되는데, 본 논문에서는 Coeflit4 필터를 사용하였다. 최적의 기저를 찾아 계수들을 분해하는 이유는 웨이블릿 분해보다 더 적은 수의 계수에 에너지를 집중시키는 결과를 가져오기 위해서이다.

다음 그림은 Grass-Telefactor사의 Aurora EEG 장비를 사용하여 표본 추출한 정상인 남자의 Fp1(그림 6참조) 신호에 대하여 웨이

블릿 패킷 분해한 결과이며, [그림 2]에서 설명한 패킷 변환의 단계를 2단계까지만 기입하였다.



[그림 4] EEG 신호의 웨이블릿 패킷 분해

[2단계] 에너지 축적 함수를 이용하여 압축을 위한 임계값 “ τ ”를 결정한다.

압축 방법은 다음 두 가지로 분류할 수 있다.

첫째, 압축률을 M/N 으로 정해 놓는 방법으로 전체 웨이블릿 계수의 개수 N 개중 M 개를 선택하고, 나머지를 제거하는 방법이 있고,

둘째, 임의의 수 “ τ ”를 정하여 “ $|c| < \tau$ ”를 만족하는 웨이블릿 계수 “ c ”를 제거하는 방법이 있다.

본 논문에서는 EEG 신호의 스펙트럼이 손상될 것을 우려하여 후자를 선택하였다. 임계값 “ τ ”를 선택하기 위해서 에너지 축적 함수를 정의하였다.

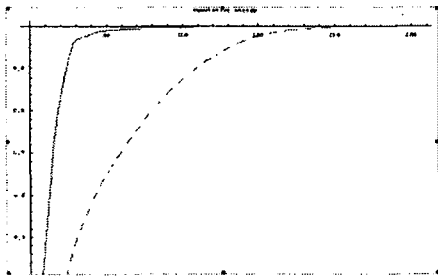
에너지 축적 함수 $E(N)$ 은 다음 식과 같다.

$$E(n) = \sum_{i=0}^n |c_i|^2 \quad (3)$$

(여기서, $|c_i| \geq |c_{i+1}|$)

[3단계] “ τ ”를 구하기 위해 웨이블릿 계수의 축적 함수를 정규화 시킨 후, 총 에너지 축적 함수의 99%에 들어오는 계수의 개수를 구한다[11].

다음 [그림 5]는 256개의 웨이블릿 계수에 대한 에너지 축적 함수의 그래프이다. 실제로 처음 패킷 분해의 경우 70여 개의 웨이블릿 계수에 총 에너지가 집중되는 반면 웨이블릿 분해는 200개 정도까지로 거의 대부분의 계수에 에너지가 퍼져 있는 것을 알 수 있다.



[그림 5] 정규화 에너지 분포 그래프

위에서 실선은 웨이블릿 패킷 분해, 점선은 웨이블릿 분해를 나타낸다.

[4단계] [3단계]에서 결정된 “ τ ”의 개수만큼 압축한 후 역 변환한다.

$$IDWPD(\hat{x}) = W^{-1}(X) \tag{4}$$

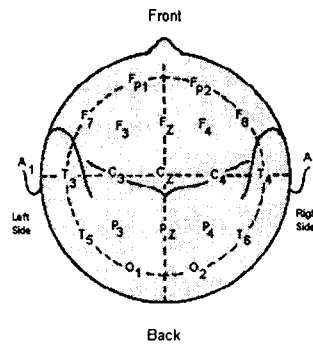
여기서 IDWPD : Inverse Discrete Wavelet Packet Decomposition, (\hat{x}) 임계값 내의 웨이블릿 계수들을 나타낸다.

4. 성능평가 및 결론

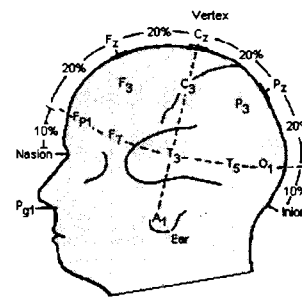
4.1 EEG 측정과 성능평가

EEG 데이터 측정에서 피험자는 정상인

남자를 대상으로 하고, Aurora EEG 장비를 사용한다. 샘플링 주파수는 400Hz, 16개의 채널(F7, T3, FP1, F3, C3, O1, F8, T4, T5, T6, FP2, F4, C4, P4, O2)을 다 채널 동시 기록 방법으로 측정하며, 전극은 지름 8mm의 산플라 접시 전극을 사용한다. 다음 [그림 6]은 국제 뇌과학회가 나타낸 유도 전극의 위치로 성능평가에 사용한 EEG 채널은 Fp1 채널이다.



(a)

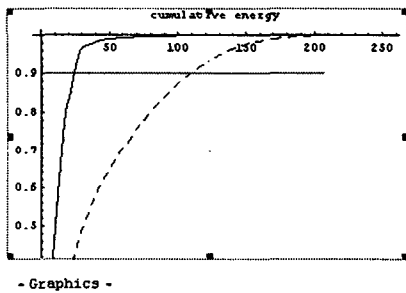


(b)

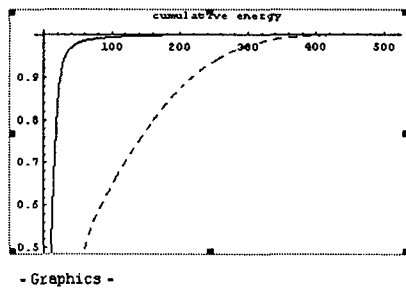
[그림 6] 뇌파 유도부위 상단모습(a)과 측면모습(b)

성능평가를 위하여 EEG 데이터의 개수를 늘려가면서 축적 에너지를 계산한 후 정규화시켰다. 다음 그림은 축적 에너지의 정

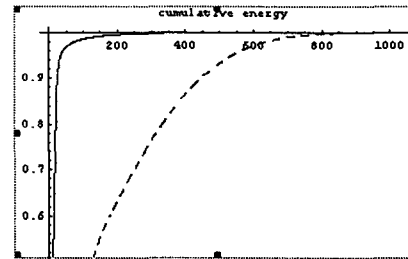
규화 그래프로 (a)는 1초간의 웨이블릿 전체 계수 256 개 (b)는 2초간의 웨이블릿 총계수 512개 (c)는 4초간의 웨이블릿 총 계수 1024 개 (d)는 40초간의 웨이블릿 10240개를 나타낸다. 각 그래프에서 실선은 웨이블릿 패킷 방법, 점선은 웨이블릿 방법을 나타낸다.



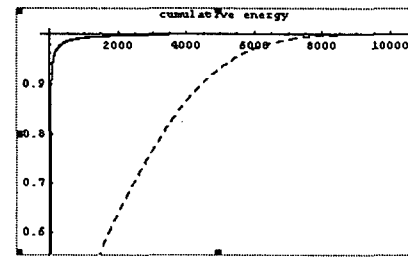
(a)



(b)



(c)



(d)

[그림 7] 축적 에너지 그래프

위의 [그림 7]중 (a)에서 전체 웨이블릿 계수 256개 중 90%의 에너지를 축적하고 있는 계수의 개수를 비교해 보면, 패킷 분해의 경우 약 20여 개 기본 웨이블릿 분해의 경우 약 110개로 나타난다. 즉, 전체 신호 256을 표현하기 위해 패킷 분해는 20개의 계수만으로 표현이 가능하고 기본 웨이블릿의 경우 110개가 필요해서 5배 이상의 차이를 보인다는 것이다.

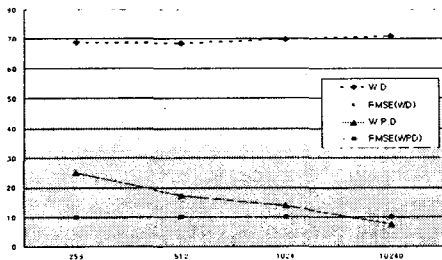
신호의 애러울을 비교하기 위하여 식(5)에 의하여 RMSE (Root Mean Square Error)를 계산한다.

$$RMSE = \sqrt{\frac{\sum_{k=1}^N |c'_k - c_k|^2}{\sum_{k=1}^N c_k^2}} \quad (5)$$

여기서, c_k : 원 데이터 c'_k : 복원 데이터이다.

아래의 표 1은 EEG 데이터에 따른 신호의 에러율 RMSE와 99%의 에너지를 가지고 있는 신호를 표현하는데 필요한 웨이블릿 계수(M) 그리고 압축률(M/N)을 나타낸다. (WD : Wavelet Decomposition, WPD : Wavelet packet Decomposition)

	Decomposition	M	M/N (%)	RMSE (%)
256 (i)	W.D	176	68.75	9.84
	W.P.D	64	25	9.88
512 (ii)	W.D	351	68.55	9.98
	W.P.D	88	17.18	9.96
1024 (iii)	W.D	715	69.82	9.96
	W.P.D	142	13.86	9.97
10240 (iv)	W.D	7267	70.96	9.99
	W.P.D	779	7.60	9.99



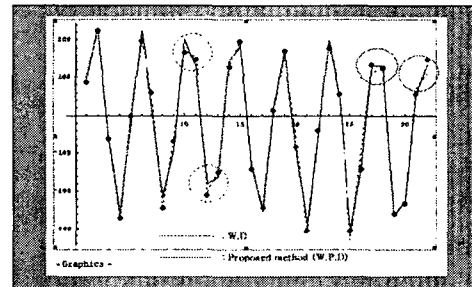
[표-1] 에러율과 압축률 비교표

위의 [표-1]에서 (i)의 경우, 에러율(RMSE)은 W.D의 방법과 W.P.D의 방법이 9.84와 9.88로 거의 유사하다. 압축률에서는 W.D는

68.75%, W.P.D는 25%로 약 2.75배로 W.P.D가 우수하다. (ii)의 경우, W.D는 68.55%의 압축률, W.P.D는 17.18%로 약 3.98배 (iii)의 경우, W.D는 69.82%의 압축률, W.P.D는 13.86%로 약 5.03배, (iv)의 경우, 40초간의 Fp1 채널의 웨이블릿 계수 10240는 W.D는 70.96%, W.P.D는 7.60%으로 약 9.09배가 우수하다.

위의 성능평가는 단일 채널(Fp1)에 대한 신호의 압축결과이다. 만일 16 채널 동시기록에 대하여 40초간 EEG 신호를 측정할 경우 위의 실험결과를 적용한다면 압축률이 약 145.44배 이상의 감소 효과를 얻을 수 있다고 본다.

본 논문에서 제안하는 방법이 EEG 신호에 적용되는 만큼 압축에서의 높은 성능을 유지하면서 복원 후 스펙트럼을 살펴본다.



[그림 8] 복원 후 결과 그래프

위의 [그림 8]은 표본추출된 32개의 디지털 뇌파 신호를 복원시킨 결과이다. 점(dot)으로 표시된 것은 원 신호, 점선으로 표시된 것은 웨이블릿, 실선은 웨이블릿 패킷을 이용한 방법으로 복원된 데이터 값이다. 등근 점선은 웨이블릿이 패킷분해보다 원 신호에 유사한 부분을 나타낸 것이고, 등근 실선은 패킷분해가 기본 웨이블릿보다 원 신호와 유

사한 부분을 나타낸 결과이다. 스펙트럼의 형태가 변형되거나 해석을 달리할 만큼 커다란 변화를 보이지는 않음을 확인할 수 있다.

4.2 결론

본 논문에서는 웨이블릿 패킷 분해가 근사 정보는 물론 상세정보를 동일 대역으로 분할하고, 에너지가 소수의 계수에 집중되는 특성을 이용하여 EEG 신호를 압축하는 방법을 제안하였다. 최소 엔트로피 기준법을 이용하여 최적의 기저를 찾고 임계값 설정을 위해 에너지 축적 함수를 정의하여 1% 미만의 적은 에너지를 유지하면서도 EEG 데이터가 커질수록 높은 압축률을 유지하는 결과를 얻었다. 기본 웨이블릿 방법은 1초간의 신호인 경우, 256개중 176개에 해당하는 계수에 에너지가 집중되는 반면, 웨이블릿 패킷 분해를 이용한 경우, 64개에 에너지가 집중되어 2.75배 정도의 높은 압축률의 차이를 보임에도 불구하고 스펙트럼에는 거의 영향을 끼치지 않는 결과를 얻었다.

제안하는 방법이 PACS(디지털 의료 영상 저장 전송 시스템)에 적용될 때, 생체 신호의 실시간 전송을 위한 의료 기기 개발에서의 활용 방안이 기대된다.

참 고 문 헌

- [1] G.A.Peredana and P.Tonella, "EEG data compression techniques" IEEE Trans. Biomed.Eng., Vol.44,1997, pp.105-114
- [2] M.Unser and A. Aldroubi. "A review of wavelets in biomedical applications." In Processings of the IEEE, Vol 84, April,1996,pp.626-638
- [3] M.Key, "Wavelets in biomedical engineering", Annals of Biomed, Engg, Vol.23, 1995, pp.531-542
- [4] C.S. Burrus, H.Guo and R.A. Copinath. "Introduction to wavelets & wavelet transforms " Prentice Hall, Englewood Cliffs, NJ, 1997
- [5] G. Strang and T. Nguyen. "Wavelets and Filter Banks" Wellesley-Cambridge Press, Wellesley, MA, 1997
- [6] Pradhan, N. and D.N.Dutt. "A nonlinear perspective in understanding the neurodynamics of EEG" Comput. Biol. Med.23 ,1993, pp. 425-442
- [7] H.1 J. Ben and Wei-Kang Cheng, "Structural EEG engineer", IEEE Engineering in medicine and biology magazine, ,1998, pp. 41-45
- [8] Roeschke, J and E. Baar, "The EEG is not a simple noise", Strange attractors in intracranial structures In : Baar E.(ed.), Dynamics of Sensory and Cognitive Processing by the Brain, Springer, Brelin, Heidelberg, New york, 1988, pp.203-216
- [9] Skrandies, K. "Evoked potentials correlates of semantic meaning -A brain mapping study." Brain Res Cog Brain Res 6(3), 1998, pp.173-183
- [10] Ronald R. Coifman, Yves Meyer, and Victor Wickerhauser "Wavelet Analysis and Signal Processing" Yale University, New-Haven, Ct 06520,1992
- [11] Schult, and Wyld,(1992), "Using wavelets to solve the Burgers

equation: A comparative study,"
Physical Review A, 46, N0.12, 7953-7958

저자 약력



조현숙(Cho, Hyunsook)

2000년

1996년: 대전대학교 수학과 졸업

2001년: 대전대학교 정보통신공학과 석사

<관심분야> 신호처리, 영상처리



황신태 (Hwang, Suntae)

1979년: 서강대학교 수학과 학사

1987년: Case Western Reserve University(미국) 전자계산학과 석사

1993년: Case Western Reserve University(미국) 전자계산학과 박사

1978. 12월 ~ 1982. 8월: KIST 연구원

1993. 9월 ~ 1995. 2월: 현대 전자 연구소 책임연구원

1995. 3월 ~ 현재: 대전대학교 정보통신공학과 교수

<관심분야> Computer Architecture, Smart Card 및 컴퓨터 보안 분야



이 형(Lee, Hyoung)

1964년 2월: 서울대학교 수학교육과 졸업

1971년 2월: 성균관대학교 전자계산학과 석사

1992년 2월: 조선대학교 전기공학과 박사

1993. 3월 ~ 현재: 대전대학교 컴퓨터정보통신공학부 교수

1995. 3월 ~ 1997. 2월: 대전대학교 전자계산소 소장

1999. 3월 ~ 2001. 3월: 대전대학교 공과대학 학장

2000. 1월 ~ 현재: 한국정보기술응용학회 회장

2000. 3월 ~ 현재: 대전대학교 산학협력연구원장

2000. 9월 ~ 현재: 대전대학교 한의학응용공학연구소장

<관심분야> 계산이론, 컴퓨터그래픽스, 신호처리

◆ 이 논문은 2003년 6월 30일 접수하여 2차 수정을 거쳐 2003년 12월 3일 게재 확정 되었습니다.