

시계열 데이터의 혼돈도 분석 알고리즘에 관한 연구

0이 병 채, 정기삼, 이명 호
연세대학교 공과대학 전기공학과

A Study on Complexity Measure Algorithm of Time Series Data

Byungchae Lee, Keesam Jeong and MyoungHo Lee
Dept. of Electrical Engineering
Yonsei University

ABSTRACT

This paper describes a complexity measure algorithm based on nonlinear dynamics(chaos theory). In order to quantify complexity or regularity of biomedical signal, this paper proposed fractal dimension-1 and fractal dimension-2 algorithm with digital filter. Approximate entropy algorithm which measure a system regularity are also compared. In this paper investigate what we quantify of biomedical signal. These quantified complexity measure may be a useful information about human physiology.

1. 서론

최근에 심박 시계열 데이터의 본질적인 성질인 비선형성을 매우 중요하게 여기고, 비선형 동역학 분야에서 비롯된 프랙탈과 카오스 개념을 이용하여 건강한 심장 활동을 이해 하려고 노력하고 있다. 카오스 이론에 의한 생체 신호의 분석은 숨겨진 동역학 특성을 분석할 수 있다는 점에서 대단히 가능성이 많은 방법이다. 그러나 카오스 이론에서 제공하는 다양한 분석방법을 실제 실험 데이터에 적용하고 정량화하는 데는 많은 어려움을 수반한다. 카오스 이론에서 제공하는 시스템의 정량 값은 엔트로피와 디멘전인데, 필요조건 및 충분조건을 만족시키는 데에는 데이터의 갯수에서 문제가 발생한다. 본 논문에서는 Goldberger등이 주창한 병리현상과 혼돈도의 감소라는 가설을 근간으로 하여, 단순히 신호의 혼돈도만을 정량화 함으로써, 시계열 데이터에 나타난 임상적 의미를 해석하고자 한다.

본 논문에서 기술하고자 하는 혼돈도 분석 알고리즘은 시스템의 동역학 특성을 분석할 수 없으며, 단순히 시간 축에서 시계열 데이터의 혼돈도만을 정량화하는 알고리즘이다. 본 논문에서는 혼돈도를 측정하기 위해, 다음과 같은 3가지 형태의 분석법을 도입하고 평가한다.

첫째, 신호의 규칙성을 무시하고 신호의 진폭과 주기 변화와 관계하는 프랙탈 차원-1(FD-1) 둘째, 신호의 진폭, 주기 변화와는 관계 없이 규칙성만 관계하는 Approximate Entropy (ApEn) 셋째, 신호의 진폭 변화와 규칙성 모두

관계하는 프랙탈 차원2(FD2)

2. 혼돈도 분석 알고리즘의 설계

Mandelbrot[1]는 프랙탈 개념을 매우 다양한 분야에 적용하고 연구하였다. 그는 불규칙적이지만 자기유사성이 존재하는 시스템을 프랙탈이라고 정의하였다. 시간이 변화하여도 시간 간격에 영향을 받지 않고 유사한 특성을 간직하는 프로세스를 프랙탈 프로세스라고 한다. 이러한 특성을 '자기유사성' 또는 '척도 무의존성' 이라고 한다. 이러한 '자기유사성' 성질은 시계열 데이터에 적용할 수 있다. 본 논문에서 설계 하고자하는 혼돈도 분석 알고리즘은 이러한 프랙탈 측정법에서 유추된 것이다.

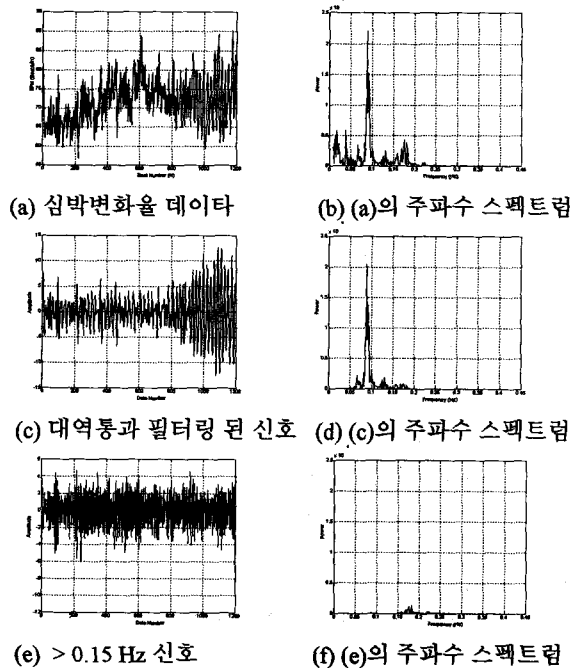


그림 1 제안된 디지털 필터를 이용한 UF, MF, HF 신호 및 주파수 스펙트럼
Fig.1 Time series data and its frequency spectrum through the proposed digital highpass and bandpass filter

디지털 필터의 설계

본 논문에서는 주파수 특성 성분 시계열 데이터를 얻기 위해 실시간 처리가 가능하며, 안정적인 위상보정 필터를 설계 하였다. 위상보정필터는 Potta[2] 등이 제안한 Bilinear Null Phase Filter(BL/NP) 가 전방향, 후방향 처리를 하여, 실시간 환경에 부적합한 면을 개선 하였다. 그림 1에는 본 논문에서 구현한 디지털 필터를 이용하여 UF, MF 그리고 HF 신호를 추출하고 각각의 신호에 대한 주파수 스펙트럼을 나타내었다.

전방향만을 처리한 후 지연된 시간 만큼을 원 데이터에서 빼줌으로써, 위상보정 효과를 갖음과 동시에 실시간 환경에 적합 하도록 하였다. 원신호(UF), 저주파 신호(0.05 Hz - 0.15 Hz) 신호 그리고 고주파신호 (> 0.15 Hz) 이상의 세 종류의 신호로 분리한 후 각각의 프랙탈 차원을 계산하고 sympathovagal balance 를 측정하기 위한 교감 신경 활동 지시자로서 MF /HF 비를 계산 한다.

프랙탈 차원1

일반 곡선의 프랙탈 차원은 $FD = \log(L) / \log(d)$ 로 나타낼 수 있다. 여기서 L 은 전체 곡선의 길이 이며, d 는 곡선의 직경이다. 곡선의 직경에 따라 편차가 심해 질 수 있다. Katz[3,4] 등은 다음과 같이 위의 일반식을 변형 하였다.

$$D = \log(L/a) / \log[(K/a)A] \quad (1)$$

여기서 D 는 프랙탈 차원이며, L 은 전체 곡선의 길이,

a 는 평균 거리, $K = 2 / \pi$, A 는 잠재적으로 곡선이 채울 수 있는 원의 면적으로 정의 하였다.

식 (3) 에서 상수 K 를 제거하면,

$$D = \log(na/a) / \log[((2/\pi)/a)\pi d^2] \\ D = \log(n) / \log(nd/L) \\ = \log(n) / [\log(n) + \log(d/L)] \quad (2)$$

여기에서 L 은 각 샘플간의 거리를 모두 더한 전체 곡선의 길이이며, n 은 전체 샘플 간격의 수이다.

일반적으로 d 는 곡면의 지름을 이용할 수 있다. 시계열 데이터의 미묘한 변화를 감지할 수 있는 프랙탈 차원을 얻기 위해서 본 논문에서는 1차 회귀선을 구하여, 이 회귀선의 거리를 척도로 한 회귀선 척도 프랙탈 차원을 설계하였다. 회귀선위에서 변동하는 신호의 전체 길이를 측정하여 회귀선의 거리로 나누어 준다면, 1차원을 갖는 직선에 비해 아주 작은 거리의 증감의 정도를 정량화 할 수 있다. 따라서, 프랙탈 차원 1은 진폭의 변동, 주기의 변동을 민감하게 정량화 할 수 있다.

프랙탈 차원2

회귀선 척도 프랙탈 차원과 함께 본 논문에서는 시계열 데이터를 2차원으로 임베딩한 2차원 어트랙터의 공간에서의 프랙탈 차원을 식(2)에서 공간에서 가장 긴 샘플의 길이를 d로 하여 프랙탈 차원 2를 계산 하였다. 따라서 프랙탈 차원1 보다 편차를 크게 하였으며 데이터 좌표점 사이의 상관 관계를 고려하였다.

Approximate Entropy

Kaplan[5], Pincus[6]은 KS-entropy에서 유추한 ApEn (approximate entropy) 을 이용하여 심박시계열 데이터의 규칙성을 정량화 하였다. 이 측정치는 유사한 패턴들이 비교를 위해 시간이나, 단위를 증가시켜도 계속 로그적으로 유사한 특성을 평가하는 것이다. 그러므로 엔트로피가 클수록 랜덤 시계열 데이터와 비슷해 지는 것이다.

Approximate entropy는 확률 통계적으로 규칙성이나 복잡도를 정량화하는 알고리즘으로 각광 받고 있다.

동역학적인 정보나 시스템의 특성 분석 보다는 단순히 시스템의 복잡도 또는 규칙성 정도를 분석하는 데 사용되며, 표준편차와 매우 밀접한 관계가 있으므로 표준편차의 차이가 매우 큰 데이터들에 대한 비교에는 불리하다. 제안된 알고리즘의 특성을 보이기 위해 표 1에는 주기신호인 $\sin(10\pi t)$ 의 크기를 증가 시키며 비교하였다.

표2에는 주기를 증가시키며 비교 하였다. 주기신호인 경우, 프랙탈 차원2와 ApEn은 진폭이 아무리 크게 변하여도 항상 일정한 값을 갖는 것을 볼 수 있다. 프랙탈 차원 1은 주기 변동과 진폭 변동에 각각 민감하게 반응하여 구분하며, 프랙탈 차원2와 ApEn은 패턴이 일정한 경우, 진폭 변동에는 변화하지 않으며, 주기가 증가함에 따라 정량값도 함께 증가하는 것을 알 수 있다. 프랙탈 차원 2는 다른 정량값들 보다 훨씬 큰 편차로 변화된다. 표 3에는 랜덤 신호 발생기를 이용하여, 진폭을 증가 시키며, 일정한 한계 크기 내의 랜덤한 시계열을 발생 시킨 후 이에

대한 평가 결과를 보였다. 프랙탈 차원 1은 표준편차에도 민감하게 반응하며, 프랙탈 차원 2와 ApEn는 거의 유사한 값의 증감을 보인다.

표 1 진폭증가와 프랙탈 차원 및 ApEn 비교
Table 1 Comparison FD1,FD2 and ApEn with amplitude

데이터	표준편차	FD1	FD2	ApEn
x 1	0.7	1.000036	1.499651	0.089191
x 10	7	1.003473	1.499651	0.089191
x 20	14	1.012936	1.499651	0.089191
x 50	35	1.058295	1.499651	0.089191
x 100	70	1.137380	1.499651	0.089191

표 2 주기 증가와 프랙탈 차원 및 ApEn 비교
Table 2 Comparison FD1,FD2 and ApEn with frequency

데이터	표준편차	FD1	FD2	ApEn
1 주기	3.2293	1.000001	1.006488	0.003388
10 주기	3.2293	1.000143	1.067299	0.035586
20 주기	3.2293	1.000567	1.120811	0.072828
50 주기	3.2293	1.003446	1.240570	0.198861
100 주기	3.2293	1.012571	1.383577	0.230640

표 3 랜덤 신호에 대한 프랙탈 차원 및 ApEn 비교
Table 3 Comparison FD1, FD2 and ApEn of random signal

데이터	표준편차	FD1	FD2	ApEn
랜덤1	1.0186	1.079473	5.449728	1.244703
랜덤2	9.8051	1.558465	5.347710	1.258851
랜덤3	19.8549	1.814713	4.665635	1.202946
랜덤4	50.4008	2.431092	5.113823	1.221004
랜덤5	101.7031	3.147249	5.212919	1.241478

그림 2에는 여러 유형의 신호에 대한 프랙탈 차원 및 ApEn 값을 나타내었다.

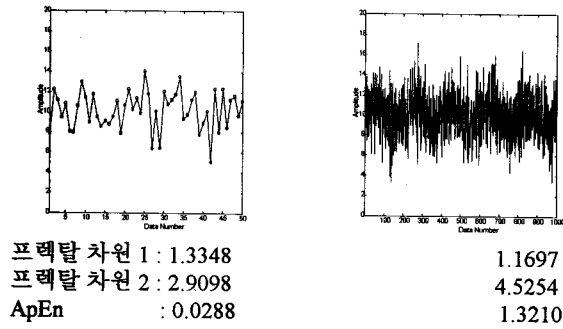
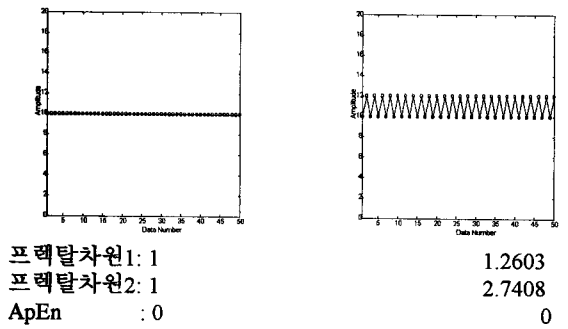


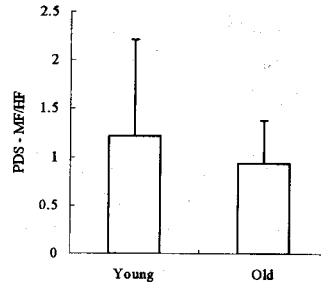
그림 2 여러 유형 신호에 대한 프랙탈 차원과 ApEn
Fig. 2 FD1, FD2 and ApEn of various signal

본 논문에서 제안된 알고리즘과 ApEn은 각기 다른 특성을 가지며, 시계열 데이터의 변화 및 혼돈도를 정량화할 수 있음을 보였다. 3장에서는 이를 이용한 간단한 시물레이션을 통해 이용 가능성을 살펴보기로 한다.

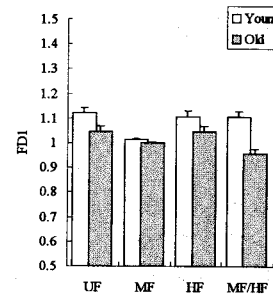
3. 응용 시물레이션

본 논문에서는 시계열 데이터의 시간축에서 혼돈도 분석 알고리즘에 관해 논하였다. 각 알고리즘의 특성과 응용의 예를 보이기 위해 20-30대(27±2.3 세)와 60-70대(69±5.7세)의 실험 대상자 각각 10 명에 대한 심박변화율을 시간영역에서 비교하였다. 인제대학교 서울 백병원 신경외과 중환자실에서 회복기에 있는

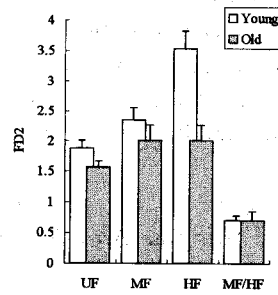
환자들로서 모니터링 시스템의 리이드II 심전도신호를 1시간동안 TEAK 사의 4 채널 레코더로 기록하였다. 저장된 심전도 신호는 R 파검출후 2 Hz의 연속적인 심박변화율 시계열 데이터로 변환하였다.



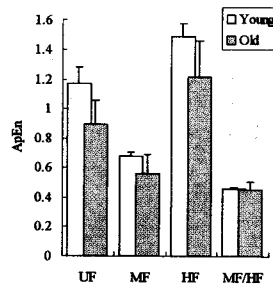
(a) 주파수 스펙트럼의 MF/HF 비



(b) 프랙탈차원-1 (FD1)



(c) 프랙탈 차원-2 (FD2)



(d) Approximate entropy

그림 3. 심박변화율 시계열 데이터의 엔트로피 측정
Fig. 3 Entropy measure of heart rate variability

그림 3에 분석결과를 보였다. 현재 널리 이용되고 있는 주파수 스펙트럼 분석에 의한 MF/HF 비는 그림 3의 (a)에 보인바와 같이 20-30대의 경우 1.2169 ± 0.9945 (평균 \pm 확신도(95%))이며 60-70대의 경우 0.9392 ± 0.4452 로 나타났는데, 이는 신뢰할 수 없는 수준의 오차 분포였다. FD1, FD2 그리고 ApEn에 의한 MF/HF 비율도 주파수 스펙트럼과 비슷한 패턴을 보이며 신뢰성이 훨씬 높음을 알 수 있다. MF/HF에 의한 sympathovagal balance 평가의 문제는 FFT가 갖는 선형성 및 stationary 조건등, 이용의 한계성에 비추어 볼때 이용 가능성이 매우 높음을 알 수 있다. 제안된 알고리즘과 ApEn의 결과에서 20-30대 심박변화율이 60-70대 실험대상자들보다 높은 혼돈도를 보임을 알 수 있다.

이러한 새로운 정보들은 기존의 주파수 분석법 그리고 새롭게 대두되고 연구되고 있는 비선형성을 이용한 카오스 분석법등과 함께 사용되어 임상적으로 매우 유용한 정보를 제공해 줄것으로 생각된다.

참고문헌

- [1] B.B. Mandelbrot, *The Fractal Geometry of Nature*, W.H.Freeman & Company, New York, 1982
- [2] E.W.Pottala et.al, "Suppression of Baseline Wander in the ECG Using Bilinearly Transformed, Null-phase Filter", *Journal of Electrocardiology*, Vol.22, pp.243-247, 1990
- [3] M.J. Katz "How Straight do Axons Grow?", *Journal of Neuroscience*, Vol.5, pp.589-595, 1985
- [4] M.J. Katz et.al., "Fractals and the Analysis of waveforms", *Comput. Bio.Med.*, Vol.18, pp.145-156, 1988
- [5] D.T. Kaplan et.al., "Aging and the Complexity of Cardiovascular Dynamics", *Biophysics Journal*. Vol.19, pp.945-949, 1991
- [6] S.M. Pincus, "Approximate Entropy as a Measure of System Complexity", *Proc. Natl. Acad. Sci.USA*, Vol.88, pp.2297-2301, 1991
- [7] Myoungho Lee and Byungchae Lee. "A study on ECG pattern recognition using chaos theory", *Proceedings of the 33rd SICE Annual conference*, pp. 981-984, Tokyo, Japan. 1994.7
- [8] Byungchae Lee, Jaeho. Jang, Jeongwhan Lee and Myoungho Lee. "Relationship Between Autonomic Nervous System and Chaotic Dynamics on Heart Rate Variability and Peripheral Blood Pressure" *Proceedings of 16th International Conference on EMBS, IEEE*, pp. 1256-1257, Baltimore, U.S.A, 1994