

SYNTACTIC 패턴인식에 의한
생체신호처리

○ 김용만, 김정훈, 정희교, 이명호
연세대학교 전기공학과

Biological signal processing using syntactic pattern recognition

Yongman Kim, Junghun Kim, Heekyo Jeong, MyoungHo Lee
Dept. of Electrical Engineering, Yonsei University

= Abstract =

A method of quantitative electrocardiogram analysis, based on concepts drawn from syntactic pattern recognition theories, is described. The algorithm can be used for removing the interference noises and base line drift as a filter function, and for reducing the number of points representing the digitized ECG waveform. The parsing is performed with simple finite state automata inferred by experiments and suitable to be updated during experiment execution. Two parameters are utilized for defining the noise and these make the algorithm flexible. The examples for testing the algorithm is real ECG waveforms with noise. Some experimental results are presented.

1. 서론

생체신호는 그 신호와 계속의 특성상 원래의 신호에 잡음이 많이 혼입되고 그 잡음의 크기가 크고 주파수가 신호대내에 있어 원 신호를 많이 손상시키게 된다. 이러한 잡음을 제거하기 위하여 최근에는 디지털 신호처리에 의한 각종 필터(LPF, BPF, HPF, BSF)의 설계 및 적응제어(Adaptive control) 개념의 신호처리방식들이 주로 사용되어지고 있다. 그러나 생체신호의 주파수 대역폭이 좁고 주파수가 아주 낮기 때문에 계산량 및 정확도, 경제성등의 문제로 인하여 실시간처리에 아직 많은 연구를 필요로 한다. 그러나 최근 인공지능(Artificial Intelligence)의 발달에 힘입어 생체신호처리에 이를 이용한 패턴인식 알고리즘이 많이 연구되고 있다. 특히 Syntactic 패턴인식은 종래에 널리 쓰이고 있는 decision-theoretic 패턴인식방법에 비하여 임의의 패턴을 정의 하고 분류하는데 구조적 정보(structural information)를 사용하여 영상해석에 일반적으로 많이 응용되어 왔다(10). 신호해석에도 응용되어 왔으며(1-4) 생체신호 해석에도 응용되었다. (6-9)

Syntactic 패턴인식은 영상분야에서와 마찬가지로 복잡

한 생체신호를 해석하는데 있어서 사람의 판단방식과 매우 유사한 것이 종래의 decision-theoretic 방법과의 커다란 차이라고 할 수 있다. 패턴이 복잡하고 class의 수가 많은 경우에도 간단한 패턴 primitive로 분석이 가능하고 computer의 처리시간도 단축되는 장점이 있다. 본 연구에서는 이러한 syntactic 패턴인식방법으로 생체신호의 기본과정인 잡음제거 및 baseline drift 제거등과 data의 압축을 위한 syntactic grammar를 제시하였으며 이를 실제의 신호를 통해 실험하였다. (그림 1 참조)

2. Syntactic Pattern Recognition

Syntactic 패턴인식의 기본은 decision-theoretic 방법과의 비슷하다. 임의의 신호를 알고있는 몇개의 class로 분류하기 위해서는 각 class의 특징을 알고 있어야 이를 바탕으로 structural feature인 primitive와 각 class와 주어진 신호와의 복합관계를 표시하는 grammar에 의해 판단하여 인식하게 된다. Horowitz는 이를 이용하여 디지털화된 심전도 신호에서 peak를 검출하기 위한 syntactic 알고리즘을 제안하였다.

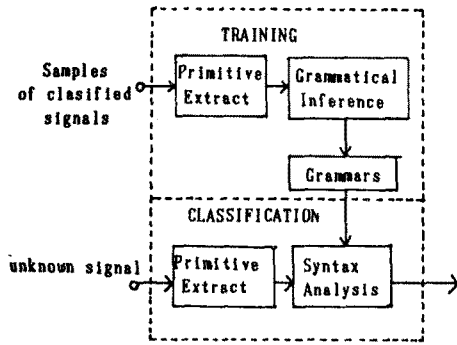


Fig. 1 A general syntactic pattern recognition system

이 알고리즘은 piecewise-linear approximation 기법을 사용하여 파형을 여과하였으며 이는 앞서 제시한 'split-and-merge' 알고리즘과 함께 원신호에 혼합된 잡음을 제거하는 필터의 기능을 수행하는 것이었다. (12, 13)

본 연구에서의 syntactic 알고리즘은 attribute grammar의 notation을 이용하여 표시할 수 있는데 이 알고리즘은 artifact(교류잡음, 근잡음등)로 인한 잡음과 전극의 움직임으로 인한 baseline drift를 제거하기 위한 것으로 생체신호처리의 본격적 자동인식을 위한 전처리과정으로 볼 수 있다. 이 notation은 syntax와 semantic을 표시하는데 적합하며 problem definition language로써 제안되어져 왔다.

본 연구에서는 'split-and-merge' 방법을 사용하였는데 이를 사용하면 앞서 언급한 piecewise-linear approximation이 필요치 않게되어 처리속도를 훨씬 빠르게 향상시킬 수 있었다. 또 syntactic 알고리즘은 파형의 다른 부분은 손상시키지 않고 제거가 필요한 부분만 선택적으로 제거할 수 있으므로 다른 방법에 비하여 커다란 장점을 가지게 된다. 더 나아가 계산이 간단하고 각각의 처리과정이 단순하여 빠르게 진행되므로 실시간 온-라인 시스템에 최적이라 할 수 있다.

3. Syntactic 알고리즘

attribute grammar는 각각의 nonterminal이 attribute의 집합으로 표시된 context-free grammar이다. 이 attribute는 nonterminal의 특성을 나타내며 값은 주어진 syntactic rule에 의한 semantic rule에 의해 결정되는데 이 semantic rule

은 동일한 syntactic rule내에서 어떤 다른 attribute의 값으로부터 다른 attribute의 값을 결정하는 함수이다. 이 경우 parse tree의 decendent node로부터 값이 결정되는 synthesized attribute와 ancestor 혹은 sibling node로부터 값이 결정되는 inherited attribute로 구분할 수 있다.

파형은 유한한 쌍의 집합 W로 나타낼 수 있다.

$$W = C_1, C_2, \dots, C_i, \dots, C_m$$

$$C_i = i\text{-번째 샘플 포인트의 좌표 } (x_i, y_i) \quad 1 \leq i \leq m$$

2개의 attribute 즉 attribute의 값이 시(starting)좌표인 inherited attribute 's'와 종(ending)좌표인 synthesized attribute 'e'로 syntactic structure를 이루고 여기에 샘플 포인트 값의 평균좌표 'a' attribute와 샘플 포인트의 수 'n' attribute등 2개의 attribute가 더 사용되었다.

a_i 를 $\Delta x = \epsilon_x$ 의 간격 I_i 에서의 좌표의 평균으로 하자. 만약 모든 샘플 포인트 C_j 가 간격 I_i 에서 j 가 $2 \leq j \leq n$ 일 때 $|x_{2j} - y_{2j}| < \epsilon_i$ 이면 이는 급격한 경사가 없는 경우이므로 a_i 를 제거시킬 포인트로 간주한다. 만약 위조건이 만족되지 않으면 다음의 알고리즘을 수행한다.

- Step 1 : 좌표 C_1, C_2, \dots, C_{j-1} 의 평균을 filtering된 포인트로 간주하고
 - Step 2 : 계속되는 좌표의 배열 C_j, C_{j+1}, \dots 의 upgoing과 downgoing portion을 구한다
 - Step 3 : upgoing 혹은 downgoing의 끝 좌표를 filtering된 포인트로 간주하고 step 1 - step 3를 계속해서 샘플 포인트의 끝 까지 반복한다.
- 본 알고리즘을 위한 grammar G은 다음과 같다.

$$G = (V_N, V_T, P, S)$$

$$V_N = \{ECC, WAVE1, WAVE2, WAVES, WAVEHD1, WAVEHD2, NOISE1, NOISE2, NOISE3, NOISEHD1, NOISEHD2, UPGO1, UPGO2, UPGO3, DOWNGO1, DOWNGO2, DOWNGO3, POSLP1, POSLP2, POSLP3, NEGSLP1, NEGSLP2\}$$

$$V_T = \{E, A\}$$

$$S = \{ECC\}$$

또 여기서 P는

$$\begin{aligned} ECC &\rightarrow WAVE2 & WAVE1 &\rightarrow WAVEHD2 \mid WAVES \\ WAVE1 &\rightarrow WAVEHD2 & WAVEHD1 &\rightarrow NOISE2 \\ WAVEHD1 &\rightarrow UPGO2 & UPGO2 &\rightarrow E \end{aligned}$$

WAVEHD1 → DOWNGO2 DOWNGO2 → E
 UPGO1 → POSLP2 | UPGO3 UPGO1 → POSLP2
 POSLP1 → A DOWNGO1 → NEGSLP2 | DOWNGO3
 DOWNGO1 → NEGSLP2 NEGSLP1 → A
 NOISE1 → NOISEHD2 NOISEHD1 → A
 NOISE1 → NOISEHD2 | NOISE3

$\epsilon_1=150 \quad \epsilon_2=0.02$ 인 경우

로 표시된다.

이 알고리즘은 실제의 신호의 peak가 아닌 잡음에 의한 peak가 원
 신호에서 제거되어야 하므로 어느정도의 파형정형이 좋은 것인지에
 대해서는 실제의 실험을 거치도록 하였다.



$\epsilon_1=70 \quad \epsilon_2=0.02$ 인 경우



4. 결과고찰

ECG신호는 IBM PC/XT로 12bit A/D변환 하여 얻었으며 샘플링
 rate는 200Hz로 하였다. 본 연구에서 사용한 알고리즘은
 2개의 parameter의 적당한 값을 찾기위해 미리 여러번의
 실험을 거쳤다. 이는 2개의 parameter의 값에따라 출력파형
 의 변화가 있으므로 임상에서의 실제용도에 따른 파형을 얻기
 위해 여러번 반복하여 실험을 거쳐야 하는 단점이 있다.

Fig. 3 Filtered ECG signal

그림2에서 잡음이 섞인 심전도를 볼 수 있으며 그림3에서는
 제안된 알고리즘에 의해 개선된 파형으로 2개의 parameter
 값에따라 나타난 차이를 볼 수 있다. 이는 앞으로 연구를 통해
 parameter값의 결정에 위한 개선된 알고리즘 및 grammar의
 개발이 이루어져야 하며 실제 임상에서의 적용가능성, 기존의
 아날로그 혹은 디지털 신호처리와의 성능비교가 이루어져야
 할 것이다.



Fig. 2 ECG signal with noise

본 연구는 앞으로 생체신호자체의 자동진단을 위한 신호
 성분의 패턴인식연구를 계속할 것이다.

5. References

(1). Stokman, G. C. and Kanal, L. N., Problem reduction repre-
 -sentation for the linguistic analysis of waveforms.
 IEEE Trans. Pattern Anal. Math. Intelligence, 5, 287, 1983

(2). Mottl, V. V. and Muchnik, I. B., Linguistic analysis of
 experimental curves, Proc. IEEE, 67, 714, 1979.

(3). Fu, K. S., Syntactic pattern recognition and its appli-
 -cation to signal processing in Digital Waveform
 Processing and Recognition. Chen. C. H. Ed. CRC Press.
 Boca Raton, Fla. 1982. chap 5

(4). Sankar, P. V. and Rosenfeld, A., Hierarchical represent-
 -ation of waveforms. IEEE Trans. Pattern Anal. March.
 Intelligence, 1, 73, 1979.

(5). Lin, H. H. and Fu, K. S., An application of syntactic pa-
 -tern recognition to seismic discrimination. IEEE.
 Trans. Geosci. Remote. Sens., 21, 125, 1983

(6). Albus, J. E., ECG interpretation using stochastic finite
 state model, in Syntactic Pattern Recognition Appli-
 -cations, Fu, K. S., Ed. Springer-Verlag. Berlin. 1976.

- (7). Smets, P., New quantified approach for diagnosis classification in Optimization of computer ECG Processing Wolf, H. K. and MacFarlane, P. W., 1980. 229
- (8). Belforte, G., De-Mori, R., and Ferraris., A contribution to the automatic processing of electrocardiograms using syntactic methods, IEEE Trans. Biomed. Eng., 26, 125, 1979
- (9). Papanikolaou, G. and Gritzali, F., Syntactic filtering of ECG waveforms Comput. Biomed. Res., 14, 158, 1981
- (10). Fu, K. S. Syntactic Pattern Recognition and Applications. Prentice-Hall, 1982.
- (11). Gonzalez, R. C. Syntactic Pattern Recognition. Addison-Wesley, 1974.
- (12). Horowitz, S. L. A syntactic algorithm for peak detection in waveforms with applications to cardiography. Comm. ACM 18, No. 5, 281, 1975.
- (13). Pavlidis, T., and Horowitz, S. L. Segmentation of plane curves. IEEE Trans. Comput. C-23, No8, 860, 1974.