

심전도 신호의 선택적 패턴인식

남 승우, 이 병채, 신 건수, 이 재준, 이 명호
연세대학교 전기공학과

(Syntatic Pattern recognition of the ECG)

Seung-woo Nam, byung-cha Lee, Kun-su Sin, Jae-jun Lee, Myung-hoo Lee
Yonsei Univ. Dept of Electrical Engineering

ABSTRACT

This paper describes the ECG pattern recognition using the syntatic pattern recognition algorithm. The algorithm uses the BNF rule with the semantic evaluation which has the structural information of the ECG. This algorithm is constructed with (1) removing the baseline drift by the Cubic spline function and extract the significant point by the line-approximation algorithm, (2) syntatic peak recognition algorithm with the extracted significant point, (3) produce the token which is used pattern recognition, (4) pattern recognition of the ECG by the syntatic pattern recognition algorithm, (5) extract the parameter with the pattern recognized ECG signal

1. 서론

본 논문에서는 심전도 신호의 구조적 정보와 심전도 파형에 대한 지식을 근거로 하여 심전도 신호의 패턴을 인식하고 진단 파라미터를 추출하는 선택적 패턴 인식 알고리즘을 제안하려고 한다. 선택적 패턴 인식 알고리즘은 심전도 신호가 가질 수 있는 패턴을 BNF(Backus Normal Form)문법으로 표현하고 어의론적(semantic)평가를 더해 심전도 신호를 인식하는 것이다.

선택적 패턴 인식 알고리즘은 크게 (1) 기저선 동요를 제거하고 심전도 신호의 중요점(significant point)를 추출하는 전처리 과정, (2) 추출한 중요점으로 피이크를 검출하는 선택적 피이크 인식 과정, (3) 인식된 피이크로 패턴 인식에 필요한 토큰(token) 생성 과정, (4) 선택적 패턴 인식 알고리즘에 의한 심전도 신호의 패턴 분류 과정, (5) 진단 파라미터 추출 과정등으로 이루어져 있다.

2. 전처리 (Preprocessing)

2.1 기저선 동요 제거 알고리즘

심전도 파형을 자동 인식하기 위해서는 먼저 파형에 존재하는 저주파 성분의 기저선 동요(baseline drift)를 제거하여야 한다. 이와 같은 잡음을 제거하기 위한 방법으로 고역통과 필터(high pass filter), 직선 보정에 의한 방법등이 널리 사용되어 왔다. 고역 통과 필터를 사용할 경우 기저선 동요를 만족스럽게 제거하기 위해서는 차단 주파수가 0.5Hz 이상이어야 한다. 그러나 이것은 ST 세그먼트와 같은 저주파 성분에 영향

을 미치게 된다. 또한 직선에 의한 보정은 저주파 성분의 기저선 동요는 제거할 수 있으나 고주파 성분의 기저선 동요를 제거하는 데는 부적합하다. 따라서 본 연구에서는 ST 세그먼트에 영향을 미치지 않고 동시에 고주파 성분의 기저선 동요도 제거할 수 있는 3차 스피라인(Cubic spline function) 알고리즘을 사용하였다.

2.2 선형 근사화 알고리즘

본 연구에서는 피이크를 인식하기 위하여 선택적 피이크 검출 알고리즘의 전처리 단계로서 선형근사화 방법을 사용하였다. 본 연구에서는 선형근사화를 수행하기 위해 초기 세그먼트의 수를 임의로 할 수 있고, 속도도 빠른 "분할 - 융합(split-and-merge)" 알고리즘을 사용하였다.

분할 - 융합 알고리즘은 다음과 같다.

[Step 1] 끝점이 일치하고 길이가 같은 임의개의 세그먼트로 파형을 나눈다.

[Step 2] 다음 식과 같이 각 세그먼트에 대해 에러범(error norm)을 계산한다.

$$E_i = \sum_{i=1}^n [y_i - (A_r \times x_i + b)]$$

여기서 y_i : 각 세그먼트에 속하는 점 x_i 에서의 값
 A_r : 선형 근사함수의 기울기
 b : 선형근사함수의 절편
 n : 각 세그먼트의 길이 (세그먼트에 속하는 포인트 수)

E_i 를 최소로 하는 선형함수의 기울기와 절편은 다음과 같다.

$$A_r = \frac{(n \times \sum x_i y_i - \sum x_i \sum y_i)}{(n \times \sum x_i - (\sum x_i)^2)}$$

$$b = \frac{1}{n} (\sum y_i - A_r \sum x_i)$$

그 다음, 위에서 계산한 에러를 모두 합한다.

$$E = \sum_{i=1}^k E_i$$

여기서 k : 세그먼트의 수

- [Step 3] E > ETOL인 경우, 가장 큰 E_i를 갖는 세그먼트를 찾아내어 두개의 세그먼트로 분할한다. E < ETOL이 될때까지 반복수행한다.
- [Step 4] 모든 세그먼트를 분할한 후, E가 최소가 되도록 끝점을 수정한다.
- [Step 5] E가 최소로 되도록 끝점을 수정한 후 E < ETOL로 되게하는 인접 세그먼트를 결합한다.

V_N : 비 종단 기호로 이루어진 집합
 V_N = { S, <pos>, <neg>, <zero> }
 P : 생성 규칙으로 이루어진 집합

3. 신택틱 피이크 인식 알고리즘

3.1 피이크 인식을 위한 프리미티브 추출

심전도 신호의 패턴을 분류하기 위해서는 우선 심전도 신호를 형성하고 있는 피이크들을 인식하여야 한다. 피이크를 형성하고 인식할 수 있는 요소는 피이크가 가지고 있는 값의 크기와 피이크를 구성하는 선분이다. 그래서 피이크의 인식을 위한 프리미티브를 두가지 형태로 추출하였다.

1) 선분의 기울기를 나타내는 프리미티브

파형의 두 점 [(x₁, y₁), (x₁₊₁, y₁₊₁)]을 연결하는 세그먼트에 대한 기울기의 특성을 a_i라 하면

$$a_i = \begin{cases} A_p > stol & \longrightarrow "/ \\ A_p < -stol & \longrightarrow "\ \\ |A_p| < stol & \longrightarrow "o" \end{cases}$$

여기서 A_p : 각 세그먼트에 대한 선형 근사식의 기울기
 stol : 기울기에 대한 문턱치

2) 피이크의 크기를 나타내는 프리미티브

심전도 파형에는 "+" 피이크와 "-" 피이크가 존재하며 이들을 구별하여야 한다. 이를 위해 각 세그먼트의 시작점(또는 끝점)이 기저선에 대해 어떤 위치에 놓이는 가에 대한 기준을 프리미티브 심볼로 선택하였다.

$$b = \begin{cases} Y_p > btol & \longrightarrow "+" \\ Y_p < -btol & \longrightarrow "-" \\ |Y_p| < btol & \longrightarrow "*" \end{cases}$$

여기서 Y_p : 각 세그먼트의 시작점(끝점)에서의 값
 btol : 기울기에 대한 문턱치

3.2 문맥 자유 문법(Context free grammar)의 구성

프리미티브 심볼을 선택한 후, 인식하고자 하는 피이크의 구조를 묘사하기 위한 언어를 생성하는 즉, 주어진 부류에 속하는 모든 피이크를 묘사할 수 있는 문법(grammar)을 설계하여야 한다. 프리미티브 심볼(primitive symbol) { /, \, o }을 이용하여 그림 2와 같은 "+" 피이크와 "-" 피이크를 묘사하기 위한 정규 표현(regular expression)은 다음과 같다.

$$"+" \text{ 피이크의 좌측부 } L = /:(/(:o)*/ \quad (3)$$

$$"+" \text{ 피이크의 우측부 } R = \backslash:(\backslash(:o)\ \quad (4)$$

여기서 ":" : 합집합의 연산자

"*" : reflexive transitive closure of the set

식 (3)과 (4)를 결합함으로써

$$"+" \text{ 피이크} = LO^*R$$

$$"- \text{ 피이크} = RO^*L$$

이와 같은 "+" 피이크와 "-" 피이크의 성질로 부터 다음과 같은 결정적 자유 문법을 설계하였다.

$$G = \{ V_s, V_t, P, |S \}$$

여기서 |S| : 시작 기호의 집합으로서 S ∈ V_N
 V_T : 종단 기호로 이루어진 집합
 V_T = { /, \, o }

4. 심전도 신호의 패턴 인식 알고리즘

4.1 패턴 인식을 위한 토큰(token) 생성

심전도 패턴 해석을 위해 본 논문에서 제시한 파서(parser)의 입력은 심전도 신호를 문자열로 표현한 것이다. 피이크 인식 알고리즘에서 검출된 피이크로서 심전도를 기울기가 큰 피이크, 기울기가 작은 피이크, 세그먼트의 형태를 가진 토큰으로 다음과 같이 표현할 수 있다.

Σ = { ksp, ksn, klp, ksn, e }
 ksp : positive peak with small slope
 ksn : negative peak with small slope
 klp : positive peak with large slope
 kln : negative peak with large slope
 e : straight line segment

피이크를 나타내는 토큰에는 8가지의 속성(attribute)이 부여되어있다.

(x_{1k}, y_{1k}) : left boundary of the peak P_k
 (x_{mk}, y_{mk}) : peak extream of the peak P_k
 (x_{rk}, y_{rk}) : right boundary of the peak P_k
 E_k : energy of the peak P_k
 angle_k : angle of the peak P_k

피이크가 가지고 있는 에너지의 계산 방법은 다음식과 같다.

$$E_k = \sum_p^q (y_i - y_{i-1})^2, \quad p = x_{1k} + 1, \quad q = x_{rk}$$

세그먼트를 나타내는 토큰에는 다음과 같은 4가지의 속성을 부여하였다.

(x_{1s}, y_{1s}) : start point of the segment S
 (x_{rs}, y_{rs}) : end point of the segment S

4.2 패턴 인식을 위한 문법의 구성

심전도의 경우 질병으로 인하여 수많은 형태나 패턴을 가지고 있고 더구나 잡음으로 인하여 패턴분류에 어려움이 있으므로 문장론적(syntax) 분석과 의미론적(semantic) 평가를 같이 행할 수 있는 강력한 문법이 요구된다. 따라서 본 논문에서는 BNF(Backus Normal Form) 문법에 의미론적 평가가 더해진 패턴 그램머로 심전도 신호를 표현하였다. 심전도 신호가 가질 수 있는 파형별 형태중 패턴 인식할 수 있는 것들은 다음과 같다.

P wave -> positive P
 negative P
 P missing
 QRS complex -> [(q)<r>(s)<r'>(s')<r'>[(s')]
 -> [(q)<r>(s)<r'>[(s')]
 -> [(q)<r>[(s')]
 -> <qs>
 T wave -> positive T
 negative T
 positive-negative T

negative-positive T
T missing

단 [x]는 option을 의미한다.

선택적 패턴 인식을 위한 입력 토큰에는 심전도 신호를 형성하는 토큰외에 노이즈 성분을 피이크로 인식한 토큰이 많이 포함되어있다. 앞에서 설계한 파서는 단순히 문장론적인 파서이므로 심전도 신호와 노이즈 성분을 인식할 수 있는 기능을 가지고 있지 못하고 있다. 이런 문제를 해결하기 위해 파싱중에 심전도 신호의 구조적 정보와 속성을 계산하여 노이즈 성분을 구분하고 심전도 신호의 패턴을 분류하는 의미론적 분류 과정이 필수적이다.

기존의 선택적 패턴 인식 알고리즘에서는 피이크의 패턴을 노이즈 성분으로부터 분류하기 위해 단순히 에너지와 피이크가 이루는 각(angle)에대한 문턱치를 두어 패턴 분류를 행하였다. 그러나 입력된 신호에 노이즈 성분이 가미되어 있으면 에너지가 작고 노이즈 성분이 섞여있으면 에너지가 크게되어 정량적인 문턱치로서는 정확한 패턴 분류에 어려움이 있었다. 본 논문에서는 심전도의 구조적인 정보를 고려하여 각 파의 패턴분류를 행하였기 때문에 좀 더 정확한 패턴 분류를 할 수 있었다.

본 논문에서 제기한 각 피이크들의 패턴분류 방법은 다음과 같다.

1) P 파 인식

일반적으로 P 피이크는 R-R interval 의 1/2 의 내부에 존재하므로 그 영역안에 존재하는 peak 들의 에너지와 각(angle)등 형태(morphology)를 계산하여 P파 인식을 행하였다.

2) QRS 인식

QRS complex의 폭은 적어도 0.03sec이상이므로 QRS라고 인식되는 토큰의 폭이 이보다 작으면 고주파의 노이즈라고 인식하였다.

3) T 파 인식

P파 인식과 같은 방법으로 R-R 간격의 2/3되는 지점 안에 T 피이크가 존재한다고 생각하였다. 그 구간내에 존재하는 피이크들중 에너지와각(angle) 등 형태를 계산하여 T 피이크를 인식하였다.

4.3 선택적 패턴 인식을 위한 파서(Parser)의 설계

선택적 패턴 인식을 위한 파서는 크게 top-down 파서와 bottom-up 파서가 있다. 본 논문에서 사용한 파서는 bottom-up 파서로서 대표적으로 사용되는 LALR 파서를 사용하였다. LR 파서를 위한 파싱표(parsing table)는 사용한 문법이 복잡할 경우 사람이 작성하기에 매우 어려우므로 LALR(1) parser generator (YACC)를 사용하였다. YACC는 작성한 BNF 문법을 컴파일 하여 자동으로 파싱표를 만들어준다.

본 논문에서 사용한 문법을 YACC로 컴파일하여 보면 파싱표에 7개의 충돌(Conflict)현상이 발생한다. 이와같은 현상을 제거하기 위하여 look-ahead합수를 사용해 충돌 현상을 방지하였다.

5. 실험 및 결과고찰

본 논문에서 심전도 신호의 패턴 분류를 수행하기 위하여 입력한 데이터는 12 bit 500 samples/sec로 A/D 변환된 CSE(Common Standard for quantitative Electrocardiography) 데이터 베이스를 사용하였다. 그림 1은 본 논문에서 제안한 알고리즘의 흐름도이다.

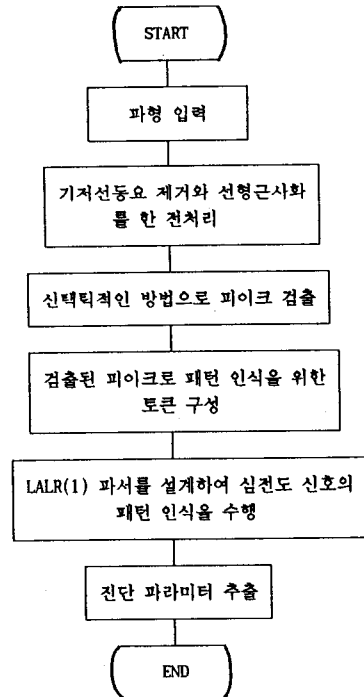


그림 1. 선택적 패턴 인식을 위한 알고리즘

그림 (2)는 입력된 데이터를 보여주고 있다. 그림 (3)은 기저선 동요의 제거와 선형 근사화 알고리즘을 수행한것이고 그림 (4)는 피이크 인식 알고리즘을 수행하여 나타난 결과이다. 그림 (5)는 패턴 인식된 심전도 신호를 보여주고 있다. 그림 (6)은 심전도 신호의 진단 파라미터를 나타낸다.

6. 결론

본 논문에서는 심전도 신호의 패턴을 분류하기위해 선택적 패턴인식 알고리즘을 제안하였다. 본 연구에서 제안한 알고리즘을 적용하여 다음과 같은 결론을 얻어내었다.

(1) 본 알고리즘은 패턴 그램머에서 정의된 문장론적 패턴인식과 의미론적 패턴 인식을 병행하기 때문에 근접음, 전원잡음등과 같은 잡음성분에 영향을 받지 않고 심전도 신호의 패턴분류를 수행할 수 있었다.

(2) 심전도 신호를 표현하는데 있어 패턴 그램머를 사용하기 때문에 질병으로 인한 심전도 신호의 여러가지 형태를 정의하여 패턴 분류를 수행할 수 있었다.

(3) 본 알고리즘은 정상파형 뿐만 아니라 부정맥등의 비 정상파형에서도 정확한 패턴 분류를 수행할 수 있었다.

(4) 본 논문에서 사용한 패턴인식 알고리즘은 심전도의 문장론적 해석과 더불어 의미론적 평가를 같이 수행하기 때문에 심전도 신호의 정확한 패턴 분류를 할 수 있었다. 문장론적 해석에서는 구현하기 쉽고 파싱속도도 빠른 LALR(1) 파서를 사용하였고 의미론적 평가에서는 심전도의 구조적 정보와 형태(morphology)로서 심전도의 패턴을 분류하기 때문에 정확하고 빠른 패턴 인식을 할 수 있었다.

7. 참고 문헌

1. E. Skodalkis, "Syntactic ECG processing: A review," in Pattern Recognition, vol. 19, no. 4, pp. 305-313, 1986
2. 신건수, 이병재, 이명호, "신터틱 패턴 인식에 의한 심전도 신호의 피크 검출에 관한 연구", 전자공학회는 문집, 제 28권, 1991
3. G. Papakonstantinou, E. Skordalakis, F. Gritzal, "An attribute grammar for QRS detection", Pattern Recogn., vol. 19, pp. 297-303, 1986
4. G. Papakonstantinou, "An interpreter of Attribute Grammar and its Application to Waveform Analysis", IEEE transaction on software eng., vol. se.7, no.3, may 1981
5. P. Trahanias, E. Skordalkis, "Syntactic Pattern Recognition of the ECG", IEEE transaction on pattern analysis and machine intelligence, vol. 12, no 7, July 1991
6. Alfred V.Aho, Ravi Sethi, Jeffery D.Ullman, Compilers Principles, Techniques, and Tools, Bell Telephone Laboratories, Incorporated, 1986
7. Allen I.Holub, Compiler Design in C, Prentice-hall international Editions, 1990

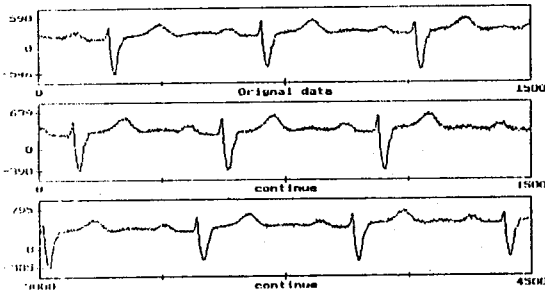


그림 2

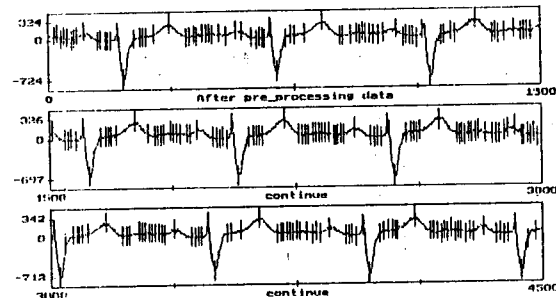


그림 3

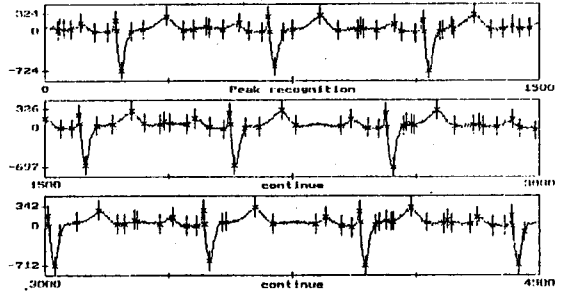


그림 4

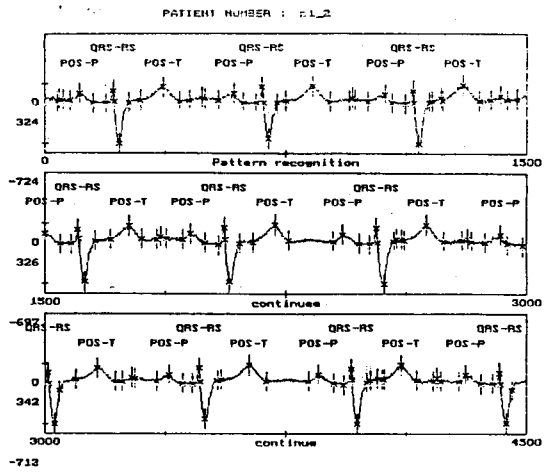


그림 5

```

*****
** ECG signal parameter of P wave **
*****
** type onset peak offset peak_asp **
*****
POS-P 83 110 150 162
POS-P 540 588 618 173
POS-P 1006 1060 1082 148
POS-P 1470 1502 1850 184
POS-P 1929 1956 2060 177
POS-P 2387 2426 2480 154
POS-P 2873 2919 2942 160
POS-P 3343 3387 3430 156
POS-P 3822 3853 3900 167
POS-P 4280 4312 4360 174

```

```

*****
** ECG signal parameter of QRS complex **
*****
** type onset peak offset peak_asp **
*****
QRS-RS 191 213 263 223
QRS-RS 639 677 723 250
QRS-RS 1124 1149 1198 221
QRS-RS 1583 1604 1656 251
QRS-RS 2040 2058 2106 279
QRS-RS 2511 2534 2587 278
QRS-RS 2990 3011 3047 197
QRS-RS 3461 3481 3537 291
QRS-RS 3932 3954 4002 275
QRS-RS 4397 4417 4458 207

```

```

*****
** ECG signal parameter of T wave **
*****
** type onset peak offset peak_asp **
*****
POS-T 288 308 422 287
POS-T 760 835 890 306
POS-T 1238 1274 1351 324
POS-T 1704 1752 1801 314
POS-T 2150 2219 2260 326
POS-T 2610 2686 2742 326
POS-T 3055 3102 3219 288
POS-T 3536 3607 3692 342
POS-T 4058 4111 4156 336

```

그림 6