

퍼지 클러스터링을 이용한 심전도 신호의 구분 알고리즘에 관한 연구

A Study on Labeling Algorithm of ECG Signal using Fuzzy Clustering

공 인 옥, 권 혁 제, 이 정 환, 이 명 호

(In-wook Kong, Hyuk-je Kweon, Jeong-whan Lee, and Myoung-ho Lee)

Abstract : This paper describes an ECG signal labeling algorithm based on fuzzy clustering, which is very useful to the automated ECG diagnosis. The existing labeling method compares the crosscorrelations of each wave forms using IF-THEN binary logic, which tends to recognize the same wave forms as different things when the wave forms have a little morphological variation. To prevent this error, we have proposed an ECG signal labeling algorithm using fuzzy clustering. The center and the membership function of a cluster is calculated by a cluster validity function. The dominant cluster type is determined by RR interval, and the representative beat of each cluster is determined by MF (Membership Function). The problem of IF-THEN binary logic is solved by FCM(Fuzzy C-Means). The MF and the result of FCM can be effectively used in the automated fuzzy inference-ECG diagnosis.

Keywords : ECG, labeling, fuzzy clustering, fuzzy C-means, backsearch

I. 서론

심전도(Electro Cardio Gram: ECG)신호는 심방과 심실의 탈분극(depolarization)과 재분극(repolarization)에 의해서 발생하는 전기적 활동, 즉 전류의 크기와 방향의 변화를 체표면에 부착된 전극에 의하여 측정후 증폭하여 오실로스코프나 ECG 기록지에 각 파형성분(waves)과 파형군(complexes)으로 기록한 것이다. 심방의 탈분극으로 발생한 전류는 P 파로, 심실의 탈분극으로 발생한 전류는 Q, R, S 파, 즉 QRS군으로 각각 기록되며, 심실의 재분극은 T파로 기록된다. 정상적 심장사이클에서 심전도파형은 P파, QRS군 및 T파의 순으로 발생하여 한 주기를 이루며, 한주기의 크기는 정상인의 경우 심박수가 약 70회/분 정도이다. 각 성분파형과 QRS군사이의 ECG 부분을 세그먼트(segment)와 인터벌(interval)이라 부른다. 즉 PR세그먼트, ST세그먼트, TP세그먼트, PR인터벌, QT인터벌 및 RR인터벌로 구성된다. 인터벌에는 각 성분파형과 QRS군이 포함되나 세그먼트에는 포함되지 않는다. 심장내과 전문의들은 이와같은 심전도 신호의 진단변수들을 임상적 진단기준에 따라 판독하여 심장의 이상 유무를 판정한다.

그러나 최근 심장질환의 급증으로 전문의가 분석해야 할 환자데이터의 양이 방대해짐에 따라 전문의의 노력과 시간의 소요도 급격히 증가되었다. 이와같은 전문의의 분석 노력과 소요시간을 절감시키기 위하여 컴퓨터를 사용한 심전도 자동 진단시스템에 관한 연구가 그 어느 때보다도 활발해지고 있다. 심전도 자동 진단시스템은 패턴인식과정과 진단과정의 2단계를 통하여 환자의 심장상태의 이상 유무를 결정한다[1]. 패턴인식과정은 심장의

전기적 활동을 반영하는 각 파형(P파, QRS군, T파)의 패턴을 분류하고 특성점을 인식하는 과정으로서, 심전도 신호에 포함되어있는 여러잡음을 제거하기 위한 필터링 과정, QRS군 검출과정, 검출된 파형을 군단위로 구분(labeling)하여 대표비트를 선택하는 과정 등으로 구성된다. 한편 진단과정에서는 구분과정에 의해 구한 대표비트의 배열순서와 대표비트로부터 추출한 진단 파라미터에 정해진 진단기준을 적용하여 환자의 심장상태의 이상 유무를 판정한다. 이들 과정중 군단위로 구분하는 과정은 기록한 심전도로부터 검출된 QRS군을 여러 군으로 분류하는 과정으로 대표비트의 선택, 신뢰성 있는 부정맥 진단, 24시간 홀터모니터링시스템에서의 비정상비트 검출 등에 매우 중요하다.

기존의 구분 방법은 심전도신호의 진폭이나 간격등의 변수들을 이용하여 파형간의 상호관계를 IF-THEN에 의하여 분류하는 2진논리를 근거로 한 결정론적 방법[1]이었기 때문에, 동일한 형태의 파형으로 분류되어야 함에도 불구하고 형태의 변화가 약간만 있어도 일부 심전도 비트가 다른 군으로 분류되는 문제점을 내포하고 있다.

본 논문에서는 이러한 2진 구조의 문제점을 해결하기 위하여, 모호한 비트들에 대한 구조적 처리가 가능한 퍼지클러스터링(fuzzy clustering)[4]을 사용하며, 이 근거한 심전도 신호구분방법을 적용하기 위하여, 시간지연 좌표계에 기반한 새로운 특성벡터를 제안하고자 한다. 퍼지 클러스터링을 위한 각 심전도비트를 대표하는 2차원의 특성벡터를 구하기 위해 위상공간에 심전도신호를 재구성하였으며, 원심전도 신호에 정수필터를 적용하고, 2차원 시간지연 좌표계상에 재구성하여 기저선 동요 등의 저주파 잡음의 영향을 제거하였다[5]. 각 심전도 비트들은 특성벡터에 의해 군을 형성하는데 FCM(Fuzzy C-Means)알고리즘을 적용하여 클러스터링하고, 클러스

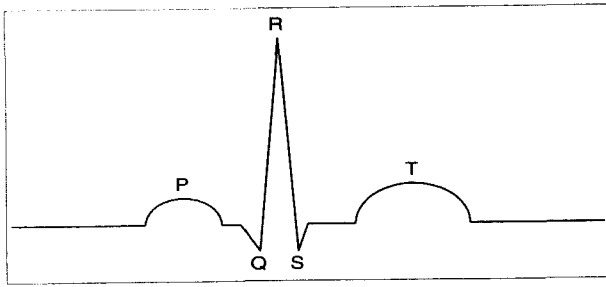


그림 1. 심전도의 기본 파형성분(P파, QRS 군, T파).
Fig. 1. The basic wave components of electrocardiogram.

터 수는 적합도 평가함수에 의해 사용하였다. FCM에 의해 각 특성벡터가 소속정도를 나타내는 MF를 최종적으로 얻을 수 있는데, 역퍼지화(defuzzification)를 통해 각각의 심전도비트의 형태를 구분할 수 있는 일련번호로써 구분하고, 지배형태의 결정에도 RR간격과 함께 MF를 이용하였다.

II. 심전도 비트의 특성벡터

심전도신호가 위상공간 좌표계에 투시되는 궤적은 그 파형에 따라 독특한 모양을 나타내며, 특히 R파는 가장 큰 궤적을 그린다. R파의 검출은 극대점중 가장 큰 궤적 변위량을 형성하는 것을 탐색하여 결정된다.

기저선 동요와 같은 저주파 잡음은 궤적의 모양을 변형시키지 않지만, 근전도 등의 고주파 잡음은 위상공간상의 궤적에 심각한 영향을 미치므로 이를 제거하기 위해 정수필터를 설계하여 저역통과 필터링을 수행하였다.

특성벡터로는 R파 좌측의 궤적변위량과 우측의 궤적 변위량 간의 비, R파를 중심으로 한 구간의 궤적변위량의 합을 이용하였다.

1. 2차원 시간지연 좌표계에서의 재구성

R파의 검출과 특성벡터의 생성을 위해서는 정수필터 [8]를 이용하여 근전도 잡음 등의 고주파 잡음을 제거한 심전도신호를 위상공간상에 재구성하였다. 본 논문에서는 F. Takens[11]의 임베딩 이론을 심전도신호에 적용하여 상태변수를 지연시간으로 하는 시간지연 좌표계를 이용하여 위상공간상에 재구성하였다. 임베딩 차원을 d_E , 지연시간을 τ 라 하고 시계열 데이터 $\xi(t_k)$ 를 Takens의 임베딩 이론에 따라 위상공간에 재구성하면 (1)과 같이 표현된다.

$$x(t_s) = \begin{bmatrix} \xi(t_s) \\ \xi(t_s + \tau) \\ \vdots \\ \xi(t_s + \tau \cdot (d_E - 1)) \end{bmatrix} \quad (1)$$

여기서, $t_s = sT_a$, $s \in S_0$

$$S_0 = \{s \in N_0; s < N_{dat} - (\tau/T_a) \cdot (d_E - 1)\}$$

(1)의 $x(t_s)$ 는 d_E 차원의 위상공간에 임의의 궤적을 형성하게 된다. d_E 와 τ 의 최적치는 카오스의 동역학 이론을 이용하여 결정할 수 있는데, 궤적을 어트랙터로 간주하고, 어트랙터가 d_E 차원의 공간을 폭넓게 사용하도록

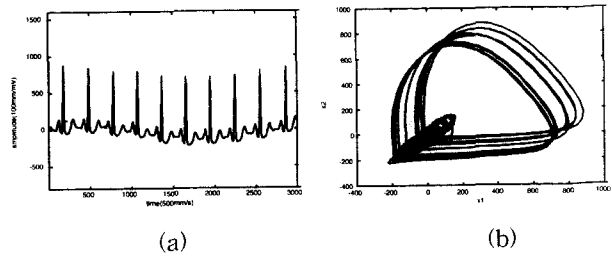


그림 2. 심전도 신호와 위상공간상의 궤적. (a) 심전도 신호, (b) 위상공간상의 궤적.
Fig. 2. ECG signal and the portrait on the phase space.

τ 와 d_E 를 결정하면 된다. 본 연구에서는 어트랙터의 공간이용도를 반영하는 fill-factor를 계산하여 결정하였다.

심전도신호에 가장 적합한 차원과 지연시간을 결정하기 위해 MIT-BIH 데이터 베이스에 대해 fill-factor를 산출하면, Fill-factor 차원 d_E 보다는 지연시간 τ 에 크게 영향을 받는다. 본 연구에서는 연산부담을 줄이기 위해 차원 d_E 의 값을 2로 설정하였다. 심전도신호의 특성에 가장 적합한 지연시간 τ 는 20ms가 44%, 16ms인 경우가 31%, 24ms인 경우가 21%,기타의 경우가 2%의 분포(표준편차 3.8ms)를 이루었으므로, 지연시간 τ 는 20ms로 결정하였다.

결정된 τ 와 d_E 를 이용하여, 500Hz로 샘플링된 심전도 데이터 $x(n)$ 을 x_1, x_2 의 2차원 시간지연 좌표계에 재구성한 궤적은 (2)와 같다.

$$\begin{bmatrix} x_1(n) \\ x_2(n) \end{bmatrix} = \begin{bmatrix} x(n) \\ x(n+10) \end{bmatrix} \quad (2)$$

그림 2(a)는 CSE 데이터베이스 SET3 #101의 심전도 데이터를 저역통과 필터에 의해 필터링한 심전도 데이터를 나타내고, 그림 2(b)는 (2)에 따라 시간지연좌표계에 심전도 데이터를 재구성한 것이다. 그림 2(b)에서 심전도 신호상의 임의의 피크는 재구성된 시간지연 좌표계에서 하나의 루프에 대응되며, 이 루프중 크기가 가장 큰 (800,0)부근의 루프는 R파에 의해 형성된다. 또한 T파와 P파는 R파에 비해 상대적으로 작은 궤적을 형성한다.

2. QRS 군의 검출

본 논문에서는 시간지연 좌표계에서 심전도신호를 재구성하여 R파를 검출하는 방법을 이용하였는데, 이 방법은 실시간임에도 불구하고 높은 검출율(99.58%)과 기저선 동요 등의 저주파 잡음에 궤적의 형태가 영향을 받지 않는다. 시간지연 좌표계를 이용한 기존의 QRS군 검출 방법으로는 궤적의 방향을 4방향으로 나눈뒤 QRS군일 경우의 방향 형태를 조사하여 QRS군으로 검출하는 방법이 있다. 이 방법은 정확한 R파의 피크 위치를 검출할 수 없는 단점을 내포하고 있고, 발생 가능한 궤적 방향 경로에 대한 사전 학습 데이터를 가지고 있어야 한다.

본 논문에서는 정확한 R파 위치 검출과 연산시간의 감소를 위해서 궤적변위량을 이용하여 그 값이 문턱치 이상인 구간에 대해서 지역 극대점을 검색하는 방법을 제안하였다. 이는 R파에 대응되는 궤적은 P파, T파에 비

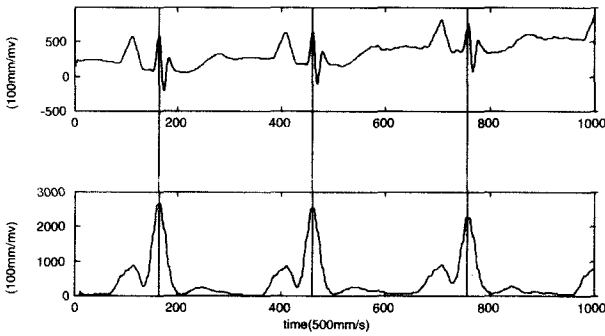


그림 3. 심전도 신호와 위상공간의 궤적 변위량.
Fig. 3. ECG signal and displacement of portrait on phase space.

해서 매우 큰 루프를 형성하므로 그 궤적의 변화량도 P 파와 T파에 비해서 크게 된다는 특성을 이용한 것이다. 즉, 궤적변위량이 문턱치 이상인 곳에서 지역 극대점을 검색하여 그 극대점 중 최고의 궤적변위량을 가지는 것을 R파로 검출한다.

시간지연 좌표계상의 궤적변위량 $dvec(n)$ 은 (3)에 의해서 결정된다. 궤적변위량은 R파의 극점보다는 R파의 피크점 전후가 크므로 검출알고리즘의 편의를 위해서 (3)과 같이 20개의 변위량의 이동합창(moving sum window)으로하여 R파의 피크에서 최대의 변위량을 가질 수 있게 한다. 또한 연산부담을 줄이기 위해 (4)와 같이 간략화된 식을 이용하여 계산한다.

$$dvec(n) = \sum_{i=n-10}^{n+9} |(x_1(i), x_2(i)) - (x_1(i+1), x_2(i+1))| \quad (3)$$

$$dvec(n) \cong \sum_{i=n-10}^{n+9} |x_1(i) - x_1(i+1)| + |x_2(i) - x_2(i+1)| \quad (4)$$

(4)의 Σ 를 전개하면, (5)와 같이 변화분만을 반영하는 식으로써 연산량을 줄일 수 있다.

$$\begin{aligned} dvec(n+1) &= dvec(n) \\ &+ |x_1(n+10) - x_1(n+11)| \\ &+ |x_2(n+10) - x_2(n+11)| \\ &- |x_1(n-10) - x_1(n-9)| \\ &- |x_2(n-10) - x_2(n-9)| \end{aligned} \quad (5)$$

그림 3은 (5)에 의해 생성된 시간지연 좌표계상의 궤적변위량과 실제 심전도 시계열데이터를 비교하여 표시한 것인데, 궤적의 변위량이 R파에서 가장 큰 값을 가지고, 비슷한 크기를 갖는 P파에 대해서는 상대적으로 작은 값을 갖게 되어, R파와 P파, T파를 구분할 수 있음을 확인할 수 있다. 구해진 변위량이 문턱치 이상이면 그 시점에서부터 탐색구간에서 R파를 탐색한다. 문턱치는 고정된 값을 사용하지 않고 예측되는 RR간격내에 R파가 없을 경우 문턱치를 낮추어 다시 검색하고, 검출된

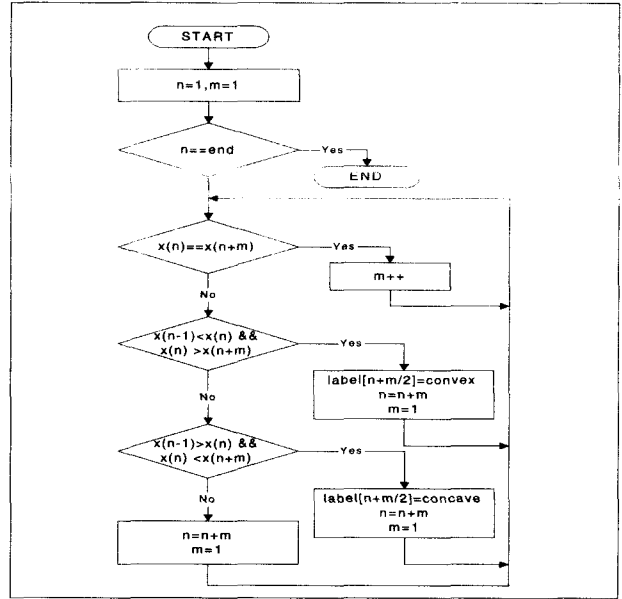


그림 4. 극대점,극소점의 결정에 관한 흐름도.
Fig. 4. Flow chart of the decision of local extrema.

R파를 기반으로하여 매번 새로운 문턱치를 적용하는 후방탐색기법(back search)기법[3][21]을 사용하였다. 후방탐색기법의 파라미터를 설정할 경우 F.N.(QRS군이 존재하는 데도 검출하지 못함)는 발생할 수 있지만, F.P(QRS군이 존재하지 않는데도 검출함)는 절대로 발생하지 않도록 문턱치의 초기값,최소치,변화량,예상 RR간격, 최소 RR간격을 설정해야 한다. 탐색구간은 인접한 다음 R파를 검출할 가능성을 배제하기 위해 최소RR간격(300ms) 이하의 값인 [-200ms,+200ms]으로 설정한다. 만약 검색구간내에 R파가 존재하지 않는다면 QS 형태로 간주하고 극소점을 검출한다. 그림 4는 극대점, 극소점을 결정하는 과정을 나타내는데,저역통과 필터를 거친 심전도 신호에 대해서 $|x(n-1)-x(n)| < 4$ 인 경우에는 $x(n) = x(n-1)$ 로 간주하는 전처리 과정을 적용하여, 미세한 변동으로 인해 극대,극소점을 잘못 검출하는 오류를 피하도록 하였다.

3. 특성벡터의 생성

심전도비트를 특성벡터로 나타내려면, 그 심전도비트로부터 특성벡터의 차원 d_f 개의 파라미터를 추출하여야 한다. 이 경우 신호의 1차 미분값의 최대/최소값과 같이 심전도비트의 기준점 위치에 무관한 파라미터만을 사용하는 경우에는 파라미터를 추출하기 위해 정확한 기준점을 설정하는 정렬과정이 필요하지 않지만, QRS군의 진폭 시계열 데이터는 정렬과정을 거쳐서 각각의 심전도비트에 대하여 기준점을 설정한 뒤 그 기준점을 이용하여 파라미터를 추출한다. 본 논문에서는 R파의 극대점을 검출하는 것에 의해 정렬하였고, 산출된 기준점과 궤적변위량으로부터 파라미터를 추출하여 각 심전도비트를 (6)에서처럼 2차원의 특성벡터에 대응시킨다.

$$f: E \rightarrow SF_1 \times SF_2 \quad (6)$$

여기서, $E = \{Beats\ of\ ECG\}$

$$SF_1 = \{x | x = sf_1(e) ; e \in E\}$$

$$SF_2 = \{x | x = sf_2(e) ; e \in E\}$$

파라미터 산출 구간은 최소 RR간격(MINIMUM_RR1)보다 적은 값으로 설정하여야 하는데 이 구간이 최소 RR간격에 가까우면, RR간격이 평균보다 작은 부정맥의 경우 인접한 QRS의 성분을 파라미터에 포함하여 오류가 발생할 수 있다. 하지만, 그 구간이 지나치게 좁다면 QRS군의 특성을 잘 나타낼 수가 없어서, 그 형태가 다름에도 불구하고 비슷한 파라미터를 산출하는 오류가 역시 발생할 수 있다

위상공간에 분포한 특성벡터를 비교하기 위해서는 특성벡터의 각 파라미터 SF_1, SF_2 를 [0,1]의 범위로 정규화하였다. 정규화 과정은 전체 파라미터중에서 최대값으로써 정규화시키는 방법과 그 파라미터 자체에 대한 정보만으로 정규화시키는 두가지의 방법으로 나눌 수 있다. 최대값에 대한 비율로 산출하는 방법은 심전도 비트의 절대적인 크기가 특성벡터의 분포를 변화시켜야 하는 경우에 사용하게 된다. 반면 자기 자신에 관한 정보만으로 정규화하는 방법은 각각의 심전도비트의 크기에 관한 정보가 특성벡터에 미치는 영향을 제거할 수 있다. 본 논문에서는 SF_2 를 산출하는 과정에는 전자의 정규화 방법을, SF_1 을 산출하는 과정에는 후자의 정규화 방법을 사용하였다.

SF_1 은 정해진 구간($\pm 80ms$)내에서 검출된 기준점으로부터 좌측의 궤적변위량과 우측의 궤적변이량을 사용하여 구한다. 파라미터의 범위를 [0,1]로 정규화하도록 하기 위해서 (좌측의 궤적 변위량)/(좌측의 궤적변위량 + 우측의 궤적변위량)으로 SF_1 을 설정한다. 궤적변위량은 위상공간상에서의 절대적 위치에는 거의 영향을 받지 않고, 궤적 자체의 모양에 의해 그 값이 많은 영향을 받는 특성을 가진다. 이러한 특성으로 인해서, 궤적의 절대적 위치를 이동시키는 기저선 변동 등의 저주파 잡음은 궤적변위량으로 부터 산출되는 파라미터에는 영향을 주지 않는다. SF_1 을 산출하는 상세한 과정은 (7)과 같다.

SF_2 는 R파를 기준으로 하여 정해진 구간($\pm 80ms$)내에서의 궤적 변위량의 합으로 구한다. SF_2 는 SF_1 을 보완하는 성격을 가지게 되는데, SF_1 이 심전도신호의 궤적 변위량으로부터 구한 심전도비트의 형태에 관한 것이라면, SF_2 는 심전도 비트의 크기에 관한 것이라고 볼 수 있다. 실제 심전도 신호의 특성상 비정상비트는 정상비트에 비해 그 크기에 있어서 많은 차이를 보이고, 두 파형의 형태가 비슷하더라도 그 크기가 많은 차이를 보인다. 이와 같은 필요성으로 인해서 형태에 관한 특성을 나타내는 SF_1 외에 심전도 비트의 크기에 관한 특성을 나타내는 SF_2 가 필요하다. SF_2 를 [0,1]의 범위로 정규화하기 위해서 SF_2 의 최대값으로써 모든 SF_2 를 나누는 방법을 사용하였다. SF_2 를 산출하는 상세한 과정은 (8)과 같다.

$$sf_1(q) = \frac{\sum_{i=r(q)-40}^{r(q)} dvec(i)}{\sum_{i=r(q)-40}^{r(q)+40} dvec(i)} \tag{7}$$

q=QRS 군의 번호

r(q)=q번째 QRS군에서 R파의 절대위치

$$sf_2(q) = \frac{\sum_{i=r(q)-40}^{r(q)+40} dvec(i)}{\max(SF_2)} \tag{8}$$

q=QRS군의 번호

r(q)=q번째 QRS군에서 R파의 절대위치

III. ECG 신호의 구분

ECG신호의 구분은 QRS군중 같은 형태적 특성을 가진 것끼리 군으로 분류하는 과정으로서 각 QRS군을 다차원의 특성벡터로 나타낼 수 있다고 한다면, 구분의 문제는 특성벡터 공간에서 분할 구간을 찾아 각 그룹으로 그 특성벡터를 분류하여 클러스터링하는 문제로 생각할 수 있다.

클러스터링 방법으로는 각 특성벡터마다 그룹에의 소속도(MF)를 부여하는 FCM을 이용하여 2진 논리에 기반한 구분의 단점을 해결하는 방법을 채택하였다. FCM은 각 클러스터의 중심과 각 특성벡터가 클러스터에 속하는 정도를 MF를 이용하여 나타낼 수 있다. 클러스터의 수(c)는 적합도 판정함수를 평가하여 최적의 적합도를 가지는 값으로 결정하였다. 또한, 심전도신호의 클러스터링은 특성벡터가 1차원 심전도신호로부터 산출된 것이므로, 특성벡터의 시간적인 선후관계에 관한 정보를 가지는데, 이를 클러스터링의 적합도 판별에 이용하였다. QRS군이 1종류 이상 존재할 때에는 어느 QRS 형태가 지배형태인지 결정하여야 하는데, 보통 지배형태는 가장 많은 형태를 가지는 QRS군 형태로 결정한다. 그러나, 심실성 이단현상의 경우 첫번째와 마지막 QRS군이 기외수축[14]을 나타내는 경우 비정상 비트가 정상 비트보다 많아질 수도 있으므로, 지배형태를 결정할 때는 가장 많이 나타나는 형태만으로 결정해서는 않된다. 본 논문에서는 지배형태의 결정시에 RR간격에 관한 정보를 이용하여 심실성 이단현상을 구분하였다.

1. Fuzzy C-Means 알고리즘

FCM은 각각의 데이터가 어느 정도로 특정 클러스터에 속하는지를 나타내는 MF를 클러스터링 과정에 이용함으로써 다른 클러스터에 속한 데이터가 특정 클러스터의 중심결정에 미치는 영향을 점진적으로 줄여서 보다 좋은 클러스터링 결과를 얻을 수 있고, 부수적으로 각각의 데이터가 각 클러스터에 어느 정도로 속하는지를 MF를 이용해서 나타낼 수 있다.

분석될 데이터는 d차원의 벡터형태로 표시된다. 다시 말하면, 특정한 개체(심전도 비트)는 다른 개체들과 비교하기 위한 척도로써 d개의 파라미터를 가지게 되는 것이다. FCM 알고리즘의 출력으로는 크게 다음의 세가지의

종류로 나눌 수 있다.

- 1) d차원의 벡터들로 표시된 데이터들을 c개의 클러스터로 나눈다.
- 2) 각각의 클러스터를 대표하는 벡터 v_i 를 구한다.
- 3) 각각의 데이터들이 각 클러스터들에 어느 정도로 속하는지를 나타내는 MF인 u_{ik} 를 구한다.

$X = \{x_1, \dots, x_n\}$ 으로 n개의 R^d 차원 벡터들의 집합이다. 이때 X를 c개의 클러스터로 나누게 되면 S_1, \dots, S_c 의 부분 집합으로 볼 수 있다. X를 c개의 클러스터로 나누었으므로, $S_1 \cup \dots \cup S_c = X$ 이다. u_{ik} 는 벡터 x_k 가 i번째 클러스터에 어느 정도의 소속도를 가지는지를 나타내는 MF이므로 모든 클러스터에 대해 (9)를 만족하여야 한다.

$$\sum_k u_{ik} = 1 \quad (i=1, \dots, c, \quad k=1, \dots, n) \quad (9)$$

데이터의 퍼지클러스터링 결과를 일반적인 개념의 클러스터링 결과로 변환할 경우에는 (10)과 같은 역퍼지화 과정이 필요하다. S_i 클러스터에 속하는 벡터 x를 결정할 경우, $u_1(x), \dots, u_c(x)$ 중 $u_i(x)$ 가 가장 큰 MF 값을 가진다면 벡터 x는 S_i 클러스터에 속하는 것으로 결정된다.

$$S_j = \{x \in X : u_j(x) \geq u_i(x) \text{ 단, } j=1, \dots, c\} \quad (10)$$

단, $i=1, \dots, c$

FCM 알고리즘은 평가함수를 최소화 시키도록 $u_i(x)$ 와 각 클러스터의 중심 v_i 를 변화시킨다. 평가함수 U는 (11)과 같이 설정한다.

$$U = \sum_i \sum_k (u_{ik})^m |x_k - v_i|^2 \quad (11)$$

여기서, $u_i : X \rightarrow [0,1]$

$$\sum_i u_i = 1$$

$$v_i \in R^d \quad i=1, \dots, c$$

u_{ik} 는 k번째 데이터 x_k 에 대한 i번째 클러스터의 MF
 $m > 1$; m은 실수

(11)의 평가함수는 각 데이터(x_k)의 유사성을 벡터사이의 거리로써 측정한다. 심전도 신호의 경우에도, 동일한 형태의 심전도 비트를 나타내는 특성벡터를 위상공간상에 표시해보면 각 파라미터의 변이가 일정 수준 이내이므로 비슷한 형태를 가진 비트는 비슷한 위치에서 군을 형성하게 된다.

평가함수 U의 m은 잡음을 제거시키는 역할을 하게 된다. m 값이 클수록 클러스터 중심에서 멀리 떨어진 노이즈성의 데이터를 무시하게 된다. 심전도 신호의 경우 비정상 비트는 그 수가 적으므로, m값을 낮은 값으로 설정하여 비정상 비트에 관한 데이터가 무시되지 않도록 한다. 일반적인 데이터의 클러스터링 경우에는 $m=2.0$ 을 사용하지만, 심전도 신호의 경우에는 $m < 2.0$ 을 사용할 것이 요구된다. 또한, $m < 1$ 인 경우에는 M.F.가 큰 특성벡터가 M.F.가 작은 특성벡터에 비해 상대적으로 적게 평가되므로 $m \geq 1$ 을 사용하는 것이 유리하다. 이와같은 이유로 인해서, 비정상 심전도 신호가 적은 수의 클러스터를 형성할 수 있다는 특징을 고려하여, m값을 2.0과 1.0의 중간값인 $m=1.5$ 로 설정하였다.

u_{ik} 와 v_i 는 (11)의 평가함수 U를 최소화시키도록 결정되어야 하는데, U를 최소화시키는 u_{ik} 와 v_i 의 분석적 해는 존재하지 않는다[4][5]. 또한 (11)의 U는 (9)의 제한 조건을 가지고 있으므로, 제한조건이 있는 최적화 문제에 이용되는 라그랑주 승수법을 이용하여 u_{ik} 와 v_i 를 산출한 뒤, 평가함수 U가 일정오차 ϵ 보다 적어지도록 반복법에 의해서 최종적인 해를 구하는 방법을 이용하게 된다.

먼저 라그랑주 승수 λ 를 이용하고 제한조건 (11)을 포함시킨 평가함수 J는 (12)와 같이 구성할 수 있다.

$$J = \sum_i \sum_k (u_{ik})^m |x_k - v_i|^2 + \lambda \cdot \sum_k (\sum_i u_{ik} - 1) \quad (12)$$

(12)의 평가함수 J를 최적화하는 u_{ik} 와 v_i 를 구하면 (13), (14)와 같다.

$$u_{ik} = (1 / |x_k - v_i|^2)^{1/(m-1)} / \sum_j (1 / |x_k - v_j|^2)^{1/(m-1)} \quad (13)$$

$$v_i = \sum_k (u_{ik})^m x_k / \sum_k (u_{ik})^m \quad (14)$$

(14)를 유심히 살펴보면 이 식이 각각의 벡터에 대해 가중치가 곱해져서 구한 무게중심이라는 것을 직관적으로 추측할 수 있다. 가중치로 인해서 v_i 는 다른 클러스터에 속하는 데이터에는 거의 영향을 받지 않고, i번째 클러스터에 속하는 데이터에 의해서 구해진 클러스터의 중심을 가르키게 될 것이다. (13)에서 분모항은 $\sum_j u_{ik} = 1$ 을 만족 시키기 위한 것임을 알 수 있다. 또한 각 클러스터의 중심 v_i 에서 거리가 멀수록 클러스터에 대한 소속도는 낮아지게 설정되어 있음을 확인할 수 있다.

(13), (14)를 클러스터링에 적용하기 위해서 v_i 와 u_{ik} 의 초기치를 설정하여야 하는데, 그 방법이 적합하지 않으면 원하는 결과를 얻을 수 없다. 본 연구에서는 v_i 는 모두 (0.5,0.5)로 초기화하였고 모든 u_{ik} 는 $1/c$ 로 초기화하였다. 본 연구에서는 u_{ik}, v_i 를 갱신시키는 방법으로 그림 5에 표시된 FCM 알고리즘의 흐름도와 같은 방법으로 (13), (14)를 적용시켜 u_{ik}, v_i 를 산출하였다.

만약 이러한 식으로 갱신하지 않고 모든 클러스터에 대해 v_i 를 갱신한뒤, 다시 모든 데이터에 대해 u_{ik} 를 갱신하게 되면 모든 v_i 가 항상 같은 u_{ik} 를 이용하여 갱신되므로 항상 같은 값을 가지고 위상공간상에서 이동하게 되어 v_i 가 각각의 클러스터의 중심을 대표하는 벡터가 될 수 없다.

모든 클러스터에 대해서 한번의 갱신이 끝나게 되면 평가함수의 결과 UNEW가 이전 평가함수의 결과 U에 비해 일정값($\epsilon, 10^{-8}$) 이하로 작다면 이때의 v_i 와 u_{ik} 를 클러스터의 중심과 각 데이터에 대한 M.F.로 설정하고 알고리즘의 수행을 끝낸다.

2. 클러스터링의 적합도

FCM 알고리즘을 적용하기 위해서는 클러스터의 수(c)를 결정할 필요가 있다. 클러스터 수(c)를 결정하기 위한 방법으로 적합도 판정함수를 사용하였다.

적합도 평가함수를 결정하기 위해서는 먼저 두가지 조건을 고려하여야 한다. 첫째, v_i 와 x_k 는 동일한 클러스터에 소속될 경우 그 거리가 가까울 수록 좋은 적합도를

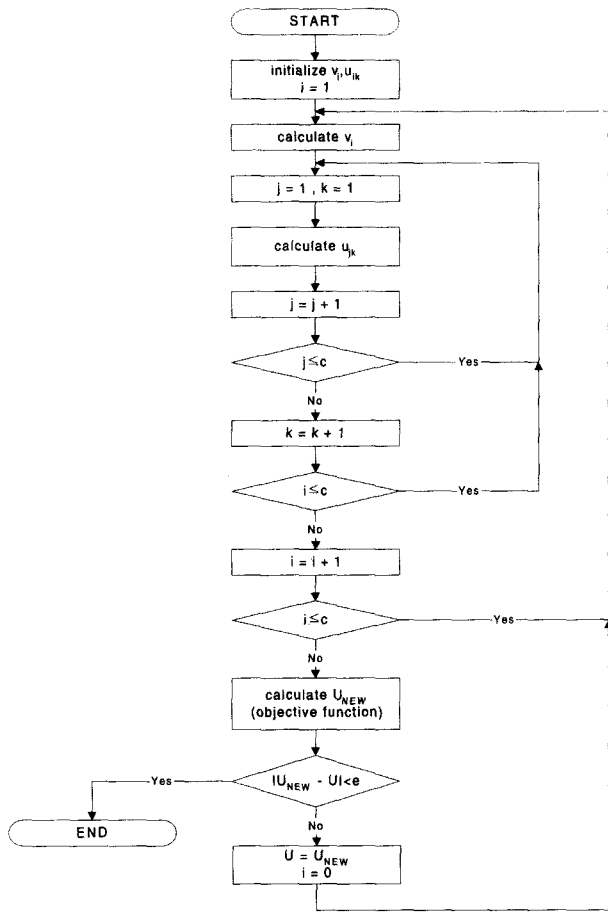


그림 5. Fuzzy C-Means 알고리즘 순서도.
Fig. 5. Flow chart of Fuzzy C-Means algorithm.

가지도록 평가되어야 한다. 둘째, 클러스터의 갯수(c)가 필요이상으로 많아져서 v_i 와 x_k 가 가까워져 적합도가 좋은 것으로 평가되는 것을 방지할 수 있어야 한다. 이러한 평가방법으로 (15)의 분할계수법, (16)의 최대 엔트로피법, (17)의 Xei-Beni법 등이 제안되었다. (15)은 그 값이 클 수록 적합도가 큰 것이며, (16), (17)은 그 값이 적을수록 그 클러스터링의 적합도가 큰 것으로 평가된다.

$$v_{PC} = \frac{\sum_{k=1}^c \sum_{i=1}^n u_{ik}^2}{n} \quad (15)$$

$$v_{PE} = -\frac{1}{n} \left\{ \sum_{k=1}^c \sum_{i=1}^n [u_{ik} \log_a(u_{ik})] \right\} \quad (16)$$

$$v_{XB} = \frac{\sum_{k=1}^c \sum_{i=1}^n u_{ik}^2 |x_k - v_i|^2}{n(\min_{i \neq j} |v_i - v_j|^2)} \quad (17)$$

분할계수를 이용하는 방법과 최대 엔트로피를 구하는 방법은 평가식 내부에 특성벡터와 가상의 클러스터 중심과의 거리에 관한 항이 포함되어 있지 않으므로, 심전도 신호의 경우와 같이 정상 비트의 수에 비해 비정상 비트의 수가 상대적으로 적어서 하나의 클러스터에 소속된 데이터가 다른 클러스터보다 그 수가 매우 적을 경우

다른 적합도를 평가하지 못한다. 이로 인해서, 다수의 개체수를 가진 클러스터에 의해 소수의 개체수를 가진 클러스터는 무시되고 적합도는 많은 개체를 가지는 클러스터만을 고려하게 된다. 이러한 문제는 각 클러스터에 포함된 데이터의 수가 비슷한 경우에는 각 클러스터의 데이터가 평가식에서 차지하는 비중이 비슷하므로 발생하지 않는 문제이다.

이러한 문제점을 해결하기 위해서 적은 수의 데이터로 구성된 클러스터도 독립된 클러스터로써 비슷한 비중으로 평가되도록 (17)의 식을 사용할 것을 제안한다. (17)의 평가식 내부에는 특성벡터와 가상의 클러스터 중심과의 거리에 관한 항이 포함되어 있어, 비록 독립된 클러스터에 포함된 데이터의 수가 적더라도 그 데이터가 별도의 클러스터로 인정되는 경우와 인정되지 않는 경우에 따라 전체적인 평가식이 많은 차이를 보이게 된다.

만약에 다수의 개체수를 가진 클러스터와 일정 거리만큼 떨어진 소수의 클러스터가 무시될 경우 거리에 관한 항으로 인해서 전체적인 평가식의 값이 증가하게 되어 소수의 클러스터도 독립된 클러스터로 인정하는 것이 오히려 평가식의 값을 적게 하는 바람직한 현상을 보이게 된다. (17)의 분자는 FCM에서의 평가함수 U와 일치하므로, 최적의 상태로 클러스터링이 되었다면 (17)의 분자항은 가장 적은 값을 유지할 것이다. 하지만, 이 값은 모든 데이터가 하나의 클러스터를 가지는 것으로 평가될 경우 가장 최소의 값을 나타내게 되므로, 이러한 단점을 보완하기 위해서 (17)의 분모항이 필요하게 된다.

(18)을 이용해 군의 개수가 상당한 차이를 보이는 인위적인 데이터 군을 생성한 후 FCM과 적합도 평가함수 VXB를 이용한 결과는 그림 6과 표 1에서 볼 수 있다. 예상했던대로, (17)이 올바른 적합도를 구한다.

$$(x_1, x_2) = (x_{1c} + 0.1 \times rand(1), x_{2c} + 0.1 \times rand(1)) \quad (17)$$

여기서, $rand(1)$ 은 $-1 < rand(1) < 1$ 범위의 임의의 수
 $0 < x_1 < 1, 0 < x_2 < 1$

심전도신호의 특성벡터들은 각 심전도 비트로부터 생성된 것이므로 특성벡터들 간에는 시간적 관계가 있다. 만약 동일한 형태를 가진다면 동일한 클러스터내에서만 특성벡터가 이동할 것이므로 그 이동량이 적을 것이다. 하지만 클러스터가 여러개 존재한다면 클러스터 사이를 이동하게 되므로 이동량이 크게 될 것이다. 이러한 성질을 이용하면, 하나의 특성벡터에서 다음 특성벡터로의 이동하는 거리를 구할 수 있게 되고, 이 값이 일정치 이상이면 다른 클러스터로의 이동이 발생한 것으로 간주하고 $c \geq 2$ 인 경우로 간주하여 적합도를 산출하도록 한다. 그렇지 않고, 그 값이 일정치 이하이면 하나의 클러스터만이 존재하는 것으로 간주하고 적합도를 산출하지 않는다. 이를 위해 본 논문에서는 실제 심전도 데이터에 적용하여, 하나의 클러스터를 가지는 특성벡터들의 집합에서의 지역 최대 이동량들중 최대값을 구하고, 두 개 이상의 클러스터를 가지는 특성벡터들의 집합에서의 산출한 최대 이동량들중 최소값을 구하여 그 중간값으로 $c=1$

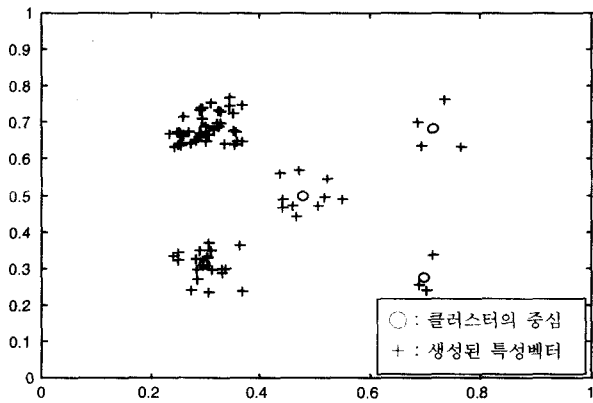


그림 6. 인위적 데이터에 대한 FCM의 적용.
Fig. 6. FCM algorithm on artificial data set.

표 1. 인위적인 데이터에 대한 FCM을 적용한 결과.

Table 1. Results applying FCM to the artificial data.

클러스터 #	개 수	x_{1c}	x_{2c}	FC-M에 의한 x_{1c}	FC-M에 의한 x_{2c}
1	40	0.3	0.7	0.2980	0.6847
2	20	0.3	0.3	0.2986	0.3100
3	10	0.5	0.5	0.4777	0.4975
4	5	0.7	0.7	0.7147	0.6830
5	3	0.7	0.3	0.6984	0.2760

인 경우를 판별하도록 설정하였다. CSE SET3 심전도 데이터 베이스를 이용하여 산출해보면 하나의 클러스터를 가지는 경우 특성벡터의 최대 이동량의 최대값은 0.15384767이고, 두 개 이상의 클러스터가 존재하는 경우 최대 이동량의 최소값은 0.34253642이었다. 이와 같은 결과를 이용하여, 본 논문에서는 특성벡터간의 위상공간상에서의 최대 이동량이 0.25보다 적은 경우 $c=1$ 인 경우로 판별하게 하였다.

3. 클러스터링 결과를 이용한 심전도 신호의 라벨링

FCM을 이용하여 심전도 비트로부터 산출한 특성벡터를 퍼지 클러스터링을 하게 되면 비슷한 형태를 가진 심전도 비트는 위상공간($SF_1 \times SF_2$)상에서 일정한 군을 형성하게 된다. 기존의 라벨링 방법이 IF-THEN에 의한 2진 논리를 이용한 결정론적 방법이었기 때문에 같은 파형이지만 약간의 형태변화가 있는 경우, 일부는 같은 군으로, 일부는 다른 군으로 분류하는 오류를 범할 수 있다. 이러한 2진논리의 문제점으로 인해서 발생하는 오류를 줄이기 위해 퍼지 클러스터링의 적용이 필요하다. 심전도 신호의 라벨링에 퍼지 클러스터링을 적용하게 되면 같은 파형을 가지는 심전도 비트의 특성벡터들은 위상공간상에서 군(클러스터)을 형성하게 될 것이다. 이 때, 약간의 형태변화가 있는 심전도 비트의 특성벡터는 클러스터의 중심에서 약간 벗어날 뿐 그 군에서 다른 군을 형성할 정도로 멀어지지 않기 때문에, 클러스터링을 행하

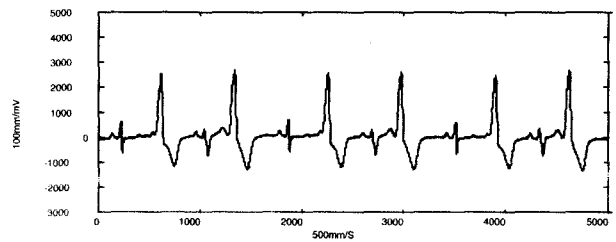


그림 7. 심실성 이단 현상(CSE set3 #117).
Fig. 7. Ventricular bigeminy (CSE set3 #117).

게 되면 동일한 클러스터로 분류되게 된다. 또한, 그 속도가 MF의 형태이므로 두 클러스터에 비슷한 정도로 속한 특성벡터의 경우에도 2진논리를 이용한 경우처럼 어느 한쪽 클러스터에 소속하는 것으로 판단하는 오류도 발생하지 않게 된다. 퍼지 클러스터링의 결과로써 각각의 심전도 비트가 특정 클러스터에 소속하는 정도를 MF의 형태로 얻을 수 있는데, 이를 기존의 2진논리에 의한 라벨링 결과와 비슷한 양식으로 표시하기 위해서는 각 특성벡터(x_k)가 어느 클러스터(S_i)에 속하는지를 결정하기 위한 역퍼지화 과정이 필요하다. 역퍼지화는 $u_1(x_k), \dots, u_c(x_k)$ 의 형태를 가지는 M.F.을 특성벡터 x_k 가 어느 클러스터에 속하는 지를 0 또는 1의 값으로 나타내는 과정인데, (14)를 이용하여 역퍼지화를 하게되면 일반적인 클러스터링을 했을 때의 결과와 유사한 양식으로 특성벡터 x_k 가 어느 클러스터에 속하는지 알 수 있다.

심전도비트가 2 종류이상 존재할 경우 어떤 심전도비트의 형태가 지배형태인지를 결정하여야 한다. 역퍼지화를 통해서 모든 심전도비트는 자신이 어떤 클러스터에 속하는지를 나타내는 하나의 라벨을 가지게 되었지만, 이는 단순히 형태를 구분할 뿐 어느 형태가 지배형태인지 구분되어 있지 않다. 지배형태는 대개의 경우 가장 많은 수를 가지는 형태로 결정되지만, 예외적인 경우로 그림 7과 같은 심실성 이단현상의 경우 첫번째와 마지막 QRS 군이 기외수축을 나타내는 경우 비정상비트가 정상비트보다 많아지는 경우가 발생할 수도 있다. 그러므로, 단순히 가장 많은 수를 가진다고 해서 지배형태를 결정해서는 않된다. 이러한 오류를 피하기 위해서 본 논문에서는 지배형태의 결정시 RR간격에 관한 정보를 이용하여 구분토록 하였다.

구체적으로, 가장 많이 발생하는 QRS군 형태가 두번째로 많이 발생한 QRS군 형태보다 3배이상 많이 발생하지 않는다면 심실성 이단현상으로 간주하고, 첫번째로 많이 발생한 QRS군 형태가 발생한 직후의 RR간격과 두번째로 많이 발생한 QRS군 형태가 발생한 직후의 RR간격을 비교를 하게 한다. 만약 1번째로 발생한 QRS군의 RR간격이 2번째로 많이 발생한 QRS군의 RR간격보다 150ms이상 크다면 2번째 많이 발생한 QRS 형태를 지배형태로 결정하고, 그렇지 않으면 첫번째로 많이 발생한 형태를 지배형태로 결정하도록 하였다. 각 QRS군이 클러스터링에 의해 결정되면, 각 그룹을 대표하는 대표비트를 설정하는 과정이 필요하다. 대표비트는 동일한 군

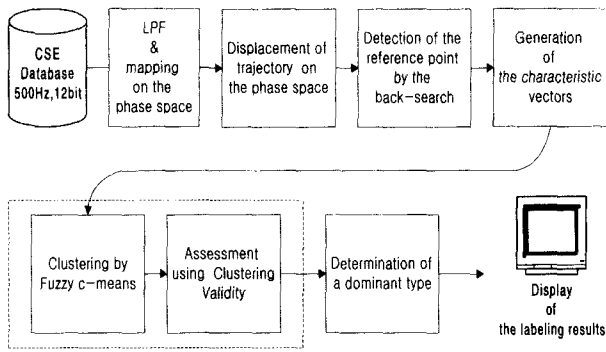


그림 8. 심전도 신호 구분 알고리즘의 블록선도.
Fig. 8. Block diagram of ECG signal labeling algorithm.

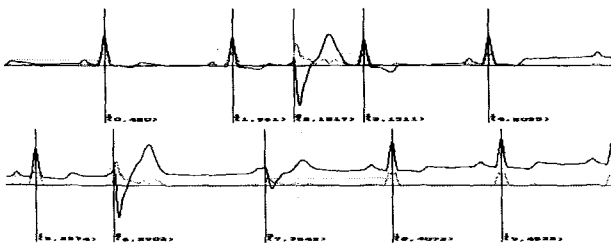


그림 9. SET3 #28을 이용하여 구분한 결과.
Fig. 9. Labeling results of SET3 #28.

에 속한 심전도 비트들을 평균화 과정을 통해 잡음을 제거시켜 구하게 되는데, 평균화 과정에서 동일한 군에 속해있지만 변이가 비교적 큰 것을 포함시키게 되면 올바른 평균화 과정에서 과정을 수행할 수 없다.

평균화 과정에서 좋은 결과를 얻기 위해서 그 그룹에의 소속도가 낮은 비트는 평균화과정에서 제외시키는 방법을 적용하였다. 이 경우 각 비트의 소속도는 퍼지 클러스터링의 결과로 생성된 MF를 이용하였고, MF가 일정이하(0.9)인 비트는 대표비트를 구하는 평균화 과정에서 제외시키도록 하였다.

IV. 실험 및 결과고찰

퍼지 클러스터링을 이용한 심전도 신호를 라벨링은 그림 8에 표시된 과정과 같이 처리된다. 구체적으로 원 심전도 신호를 저역통과필터를 거쳐 시간지연 좌표계상에 재구성하여 얻은 궤적변위량으로부터 기준점(R파)을 검출한다. 검출된 기준점으로부터 각 심전도 비트를 대표하는 특성벡터를 생성하여, 이 특성벡터를 FCM에 의해 퍼지 클러스터링한 뒤 그 결과를 이용해서 지배형태를 결정하고 각 심전도 비트를 그 종류에 따라 라벨링한다.

그림 9, 그림 11에서 심전도 파형주변에 표시된 파형은 위상공간상에서의 궤적변위량을 표시한 것이다. 궤적변위량과 연결된 수평선은 검출된 궤적변위량의 지역극대값을 의미하고, 이 값을 바탕으로 다음 R파 검출을 위한 문턱치를 산출하게 된다. 수직선은 검출된 기준점(R파)을 표시한다. 수직선의 우측하단에 표시된 수치는 구분된 결과를 나타낸다.

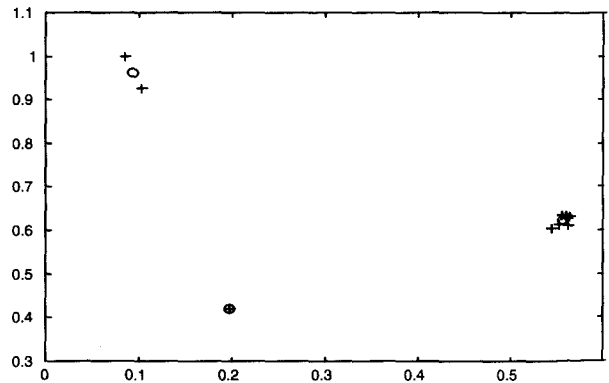


그림 10. 생성된 SET3 #28의 특성 벡터.
Fig. 10. The characteristic vectors of SET3 #28.

표 2. SET3 #28 심전도 신호에 대한 적합도 V_{XB} .
Table 2. Validity function V_{XB} of SET3 #28 ECG data.

클러스터의 수	Fc-M 평가함수 U	적합도 V_{XB}
2	0.13650875	0.08586104
3	0.00527815	0.00279100
4	0.00339136	0.09035113
5	0.00210896	0.44298357

표 3. SET3 #28 심전도 신호의 각 클러스터에 대한 MF.

Table 3. Membership functions of each clusters of set3 #28 ECG signal.

소속도 클러스터	0	1	2	3	4	5	6	7	8	9
1	0.9994	0.9944	0.0053	0.9804	0.9996	0.9965	0.0046	0.0000	0.9993	0.9994
2	0.0003	0.0035	0.9913	0.0104	0.0002	0.0021	0.9927	0.0000	0.0004	0.0004
3	0.0002	0.0022	0.0034	0.0092	0.0002	0.0014	0.0028	1.0000	0.0003	0.0003

실제 심전도 신호를 사용하여 산출한 특성벡터에 제안된 FCM과 V_{XB} 를 적용한 결과를 세 종류의 QRS군이 포함되어 있는 SET3 #28, 심실성 이단현상을 보이는 SET3 #117을 이용하여 살펴본다.

SET3 #28에 대해 구분한 결과는 그림 9와 같고, 그 특성벡터의 분포는 그림 10과 같다. 위상공간상에서 7번비트가 좌하단에, 2번,6번비트는 좌상단에 나머지는 우측에 군을 이루고 있음을 알 수 있다. 적합도 판정 결과는 그림 10의 특성벡터에서 볼 수 있듯이 3개의 클러스터에서 최소의 값을 가질것으로 예상할 수 있는데, 클러스터링의 적합도를 구한 결과인 표 2에서 보면, c=3인 경우에 최소값 0.00279100을 가져서 그 결과가 일치함을 확인할 수 있다. 표 3은 FCM에 의해 퍼지클러스터링이 완료되었을 때의 MF를 나타내는데, 이를 통해 각 심전도 비트가 각각의 클러스터에 어느정도로 소속되는지를 확인할 수 있다. 지배형태는 개체수가 가장 많은 0, 1, 3, 4, 5, 8, 9번 비트가 속한 클러스터로 결정되어 '1'로 구분되었음을 확인할 수 있다.

두 번째는 심실성이단현상을 나타내는 SET #117에

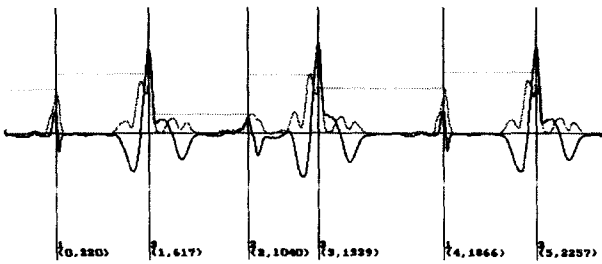


그림 11. SET3 #117을 이용하여 구분한 결과.
Fig. 11. Labeling results of SET3 #117.

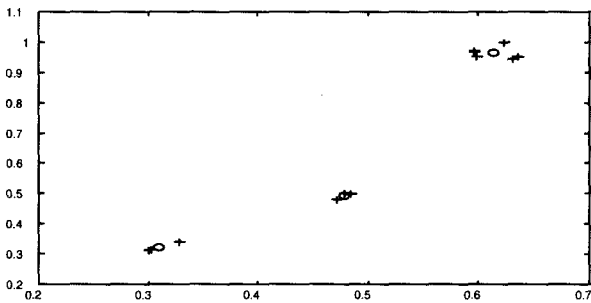


그림 12. 생성된 SET3 #117의 특성 벡터.
Fig. 12. The characteristic vectors of SET3 #117.

표 4. SET3 #117 심전도 신호에 대한 적합도 V_{XB} .
Table 4. Validity function V_{XB} of SET3 #117 ECG data.

클러스터의 수	Fc-M 평가함수 U	적합도 V_{XB}
2	0.08697301	0.02004879
3	0.00493356	0.00716284
4	0.00461218	0.85212559
5	0.00456972	8.44169752

대해서 퍼지클러스터링을 수행한다. 그림 11은 라벨링 결과를 나타내고, 그림 12는 각 심전도 비트의 특성벡터를 위상공간에 표시한 것이다. SET3 #117은 기외수축인 1,3,5번이 가장 그 수가 많지만 두번째로 많은 비트인 0,4번비트와 RR간격의 비교에 의해 기외수축을 나타내는 비록 그 수가 많음에도 불구하고 지배형태에서 제외되어 '3'으로 라벨링 되었음을 확인할 수 있다. 클러스터의 적합도 V_{XB} 는 3개의 클러스터에서 최소값을 가지는데 이는 기외수축이 아닌 비트가 그림 12에서와 같이 두 가지 형태가 있어 총 3가지 형태의 QRS군이 있기 때문이다. 비록 그 수가 많음에도 불구하고 지배형태에서 제외되어 '3'으로 라벨링 되었음을 확인할 수 있다.

V. 결론

본 논문에서는 심전도신호를 시간지연 좌표계에 재구성하여 특성벡터를 생성하고 FCM을 이용하여 심전도신호를 라벨링하는 기법을 제안하고, 이를 통해 다음과 같은 결론을 얻었다.

1) 심전도 신호를 2차원 시간지연 좌표계상에 재구성한

궤적의 변위량을 기반으로하여 R파 검출 및 특성벡터를 생성하여, 심전도신호와 중첩되는 주파수 대역으로 인해 제거하기 힘든 기저선동요 등과 같은 저주파 잡음이 특성벡터의 생성에 미치는 영향을 차단할 수 있었다.

2) 위상공간 상에 특성벡터의 형태로 분포하는 심전도비트에 FCM을 적용하여 클러스터링함으로써 IF-THEN의 2진논리를 이용해 구분할 경우에 문제가 되었던 약간의 형태변화가 있는 심전도비트일 경우에도 퍼지클러스터링의 특성상 일부를 다른 군으로 인식하는 오류를 구조적으로 발생시키지 않는다.

3) 심전도신호가 가진 특성을 클러스터링에 반영하기 위해서, $c=1$ 인 경우는 특성벡터간의 이동량으로 판별하였고, 비정상비트가 속한 작은 개체수의 클러스터가 독립된 클러스터로 인정되도록 하기 위해서 m 을 1.5로 설정하고, 클러스터의 중심과 데이터간의 거리에 관한 파라미터를 포함하고 있는 적합도 평가식 V_{XB} 를 사용하였다.

4) 지배형태를 결정할 경우 단순히 가장 많이 발생한 형태를 결정하지 않고, 각 클러스터에 속하는 비트의 평균 RR간격을 비교하여 심실성이단현상에 대해서도 효과적으로 지배형태를 결정할 수 있었다.

5) 각의 심전도 비트가 특정 클러스터에 속하는 정도를 MF로 제공하므로 퍼지 추론을 이용하여 진단을 행하는 경우 입력변수로서 본 연구의 결과를 활용할 수 있다.

6) 심전도 신호의 특성벡터 생성과정에서 고속화된 알고리즘을 제안하여 연산부담을 줄일 수 있었다.

현재 적용된 퍼지 클러스터링 알고리즘은 심전도상의 모든 특성벡터에 대해 연산을 하므로 심전도 데이터량이 많아지면 그에 비례하여 연산량이 증가하게 되어 실시간 진단을 어렵게 한다. 앞으로의 연구방향으로 실시간 진단을 위한 심전도 신호의 수집단계에서부터 점진적으로 퍼지 클러스터링을 적용하여 실시간으로 구분이 완료될 수 있도록 클러스터링 알고리즘을 개량하는 연구가 필요할 것이다.

참고문헌

- [1] 권혁제, "심전도 자동 진단 장치 구현 및 온라인 신호해석 알고리즘에 관한 연구," 연세대학교 박사 학위 논문, 1996.8
- [2] J. A. Kors and J. L. Talmon, "Multilead ECG analysis," *Comp. Biomed. Res.*, vol. 19, pp. 28-46, 1986.
- [3] 정석현, 이정환, 이병채, 이명호, "시간 지연 2차원 좌표계를 이용한 실시간 QRS 검출에 관한 연구," 대한의용생체공학회 춘계학술대회 논문집, vol. 17, no. 2, pp. 277-280, 1995.
- [4] J. C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithms*, New York: Plenum Press, 1981.
- [5] R. Cannon, J. Dave, J. Bezdek, "Efficient implementation of the fuzzy c-means clustering algorithms," *IEEE Trans. Pattern Analysis and*

Machine Intelligence, vol. 8, no. 2, pp. 248-255, 1986.

[6] N. R. Pal, J. C. Bezdek, "On cluster validity for the fuzzy c-means model," *IEEE Trans. Fuzzy Systems*, vol. 3, no. 3, pp. 370-379, 1995.

[7] X. L. Xie, G. A. Beni, "Validity measure for fuzzy clustering," *IEEE Trans. Pattern Anal. Machine Intell.* vol. 3, no. 8, pp. 841-846, 1991.

[8] J. L. Willems, *CSE Multilead Atlas: measurement results data set3*, Leuven, 1988.

[9] W. J. Tompkins, *Biomedical Digital Signal Processing*, New Jersey: Prentice Hall, 1993.

[10] P. A. Lynn, "Online digital filters for biological signals: some fast designs for a small computer," *Med. & Biol. Eng. & Comput.*, vol. 15, pp. 534-540, 1977.

[11] F. Takens, *Lectures Notes in Mathematics, Warwick*, vol. 898, pp. 230, 1980.

[12] 장재호, "Fill-factor 알고리즘을 이용한 생체카오스의 어트랙터 재구성에 관한 연구," 연세대학교 석사 학위 논문, 1995. 2.

[13] J. Lee, K. Jeong, J. Yoon, M. Lee, "A simple real-time QRS detect algorithm," *18th IEEE EMB Ann. Conference Proceeding CDROM*, 1996.

[14] D. Dubin, *心電圖 속성판독법, 고려의학*, 1989.

[15] M. P. Windham, "Cluster validity for fuzzy c-means clustering algorithm," *IEEE Trans. Pattern Analysis Machine Intelligence*, vol. PAMI4, no. 4, pp. 357-363, 1982.

[16] R.-P. Li, M. Mukaidono, "A maximumentropy approach to fuzzy clustering," *Fourth IEEE International Conference on Fuzzy Systems*, vol. 4, pp. 2227-2232, 1995.

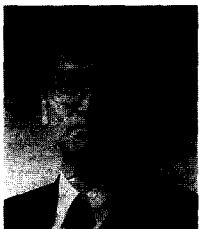
[17] H. J. Zimmerman, *Fuzzy Set Theory and Its Applications*, Boston: Kluwer-Nijhoff Publishing, pp. 191-212, 1984.

[18] 권혁제, 공인옥, 이명호, "클러스터링 알고리즘을 이용한 QRS 컴플렉스 특성 표시에 관한 연구," 대한 전자공학회 하계학술대회 논문집, vol. 17(1), pp. 680-694, 1995.

[19] S. Barro, R. Ruiz, and J. Mria, "Fuzzy beat labelling for intelligent arrhythmia monitoring," *Comp. Biomed. Res.*, vol. 23, pp. 240-258, 1990.

[20] H. J. Kweon, J. W. Suk, and M. H. Lee, "Intelligent QRS typification using Fuzzy clustering algorithm," *Proc. of 16th International Conference on EMBS IEEE*, pp. 760-761, Montreal, Canada, 1995.

[21] 공인옥, "퍼지 클러스터링을 이용한 심전도 신호의 라벨링에 관한 연구," 연세대학교 석사 학위논문, 1997. 2.



공 인 옥

1995년 연세대 전기공학과 졸업. 동 대학원 석사(1997). 1997년-현재 삼성전자반도체부품 연구원. 관심분야는 신호처리, 퍼지 및 신경회로망.



이 정 환

1994년 연세대 전기공학과 졸업. 동 대학원 석사(1996). 1996년-현재 동 대학원 박사과정 재학중. 관심분야는 의료기기설계 및 의료진단자동화. 뉴럴시스템.



권 혁 제

1990년 연세대 전기공학과 졸업. 동 대학원 석사(1992). 동대학원 박사(1996). 1996년-현재 LG산전연구소 선임연구원. 관심분야는 신호처리, 퍼지 제어 및 의용계측설계 및 ASIC.



이 명 호

1972년 연세대 전기공학과 졸업. 동 대학원 석사(1974) 및 박사(1978). 1980년-1996년 연세대학교 전기공학과 교수. 1996년-현재 동대학 기계전자공학부 교수. 1996년-현재 동대학 의료기기기술연구소 소장. 1983년-1984년 University of Wisconsin-Madison 객원교수. 관심분야는 의용계측제어기기, 생체신호처리, 의료진단자동화, 생체제어계측 및 뇌공학시스템.