
다중 클래스 SVM을 이용한 EMD 기반의 부정맥 신호 분류

이금분* · 조범준**

EMD based Cardiac Arrhythmia Classification using Multi-class SVM

Geum-boon Lee* · Beom-joon Cho**

요 약

심전도 신호 분석 및 부정맥 분류는 환자를 진단하고 치료하는데 중요한 역할을 한다. 부정맥은 맥박이 불규칙한 상태로 심실빈맥(VT)이나 심실세동(VF) 환자에게 심각한 위협이 될 수 있다. 심방조기수축(APC)과 상심실성빈맥(SVT), 심실조기수축(PVC)은 심실빈맥(VT)만큼 치명적이지는 않지만 심장질환을 진단하는데 중요한 부정맥이다. 본 논문은 2~3개의 부정맥 분류만을 고려한 기존의 방법을 극복하고 다양한 부정맥을 분류하기 위한 새로운 방법을 제시한다. 심전도 신호의 특징 추출을 위해서 EMD 방법으로 신호를 분해하여 IMFs를 얻는다. 입력 데이터의 양은 분류기 성능에 영향을 미치므로 신호 데이터의 차원을 감소시키기 위해 Burg 알고리즘을 IMFs에 적용하여 AR 계수를 구하고 여러 개의 이진 분류기를 결합한 다중 클래스 SVM의 입력으로 사용한다. 최적의 SVM 성능 파라미터를 선택하고 부정맥 분류에 적용한 결과 검출의 정확성은 96.8%~99.5%였다. 실험 결과는 제안한 EMD 방법에 의한 전처리 및 특징 추출과 다중 클래스 SVM에 의한 부정맥 분류의 유용성을 보여준다.

ABSTRACT

Electrocardiogram(ECG) analysis and arrhythmia recognition are critical for diagnosis and treatment of ill patients. Cardiac arrhythmia is a condition in which heart beat may be irregular and presents a serious threat to the patient recovering from ventricular tachycardia (VT) and ventricular fibrillation (VF). Other arrhythmias like atrial premature contraction (APC), premature ventricular contraction (PVC) and supraventricular tachycardia (SVT) are important in diagnosing the heart diseases. This paper presented new method to classify various arrhythmias contrary to other techniques which are limited to only two or three arrhythmias. ECG is decomposed into Intrinsic Mode Functions (IMFs) by Empirical Mode Decomposition (EMD). Burg algorithm was performed on IMFs to obtain AR coefficients which can reduce the dimension of feature vector and utilized as Multi-class SVM inputs which is basically extended from binary SVM. We chose optimal parameters for SVM classifier, applied to arrhythmias classification and achieved the accuracies of detecting NSR, APC, PVC, SVT, VT and VF were 96.8% to 99.5%. The results showed that EMD was useful for the preprocessing and feature extraction and multi-class SVM for classification of cardiac arrhythmias, with high usefulness.

키워드

심전도, 부정맥, EMD, AR 계수, 다중 클래스 SVM

Key word

ECG, Cardiac Arrhythmias, AR coefficient, Multi-class SVM

* 조선대학교 컴퓨터공학과 박사과정
** 조선대학교 컴퓨터공학과 교수(교신저자)

접수일자 : 2009. 12. 07
심사완료일자 : 2009. 12. 10

I. 서 론

다양한 조건하에서 자동으로 심전도 신호를 검출하고 부정맥을 분류하는 것은 심장병 환자를 진단하고 치료하는데 중요한 역할을 한다. 부정맥은 맥박이 불규칙한 상태로 심전도 신호 분석을 통하여 정상신호와 이상 패턴 신호를 검출함으로써 분류할 수 있다. 부정맥 중 심실빈맥(ventricular tachycardia: VT)과 심실세동(ventricular fibrillation: VF)은 환자에게 심각한 위협이 될 수 있으며, 심방조기수축(atrial premature contraction: APC)과 상심실성 빈맥(supraventricular tachycardia: SVT), 심실조기수축(premature ventricular contraction: PVC)은 치명적이지는 않지만 심장 질환을 진단하는데 중요하다.

부정맥을 분류하기 위한 방법으로 신경망[1][2], 퍼지 이론[3][4], 시주파수 분석[5][6], 비선형 분석법[7][8] 등이 연구되어 왔으나 2~3개의 부정맥을 분류하는데 그쳤을 뿐이며 다양한 부정맥을 분류하는데 어려움을 보여주고 있다.

본 논문은 심전도 신호를 분석하고 부정맥을 분류하기 위하여 EMD 방법과 다중 클래스 SVM을 적용한다. 심전도 신호의 전처리로 노이즈 제거 및 QRS complex 검출, 부정맥 신호의 특징 검출을 위해 EMD(empirical mode decomposition) 방법을 제시하며, 정상신호(NSR), 심방조기수축(APC), 상심실성 빈맥(SVT), 심실조기수축(PVC) 및 심실빈맥(VT)과 심실세동(VF) 등의 부정맥을 분류하기 위하여 패턴 인식 및 분류에 있어 일반화 성능이 뛰어난 SVM을 다중 클래스 문제에 적용한다.

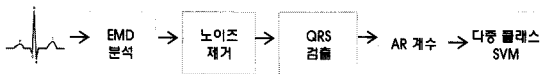


그림 1. 전체 시스템 블록도
Fig. 1. System block diagram

II. EMD(Empirical Mode Decomposition)

2.1 EMD 방법을 이용한 특징 추출

비선형과 비정상 신호들을 특징을 분석하기 위한 방법인 EMD는 생체 신호 처리에서 폭넓은 유용성을 보여주고 있다.

EMD 방법은 원신호를 ‘체과정 (shifting processing)’이라 불리는 반복적인 신호 분해과정을 통하여 IMFs (Intrinsic Mode Functions)라 불리는 함수의 집합으로 분해하는 것이다. 각각의 IMF는 원신호의 고유 속성을 보존하며 신호의 기저로 사용된다[9].

분해 단계가 증가할수록 IMF의 복잡도는 감소하며, 신호의 스케일 또한 감소한다. 심전도 신호를 EMD 방법으로 분해된 IMFs는 다양한 스케일 상의 신호 상태를 보여주고 있으며, 원신호가 다르던 분해된 IMFs도 다르다는 것을 보여준다. 그러므로 각 단계의 IMF는 신호의 기저가 되며, 심전도 신호를 구분할 수 있는 특징으로서 사용될 수 있다.

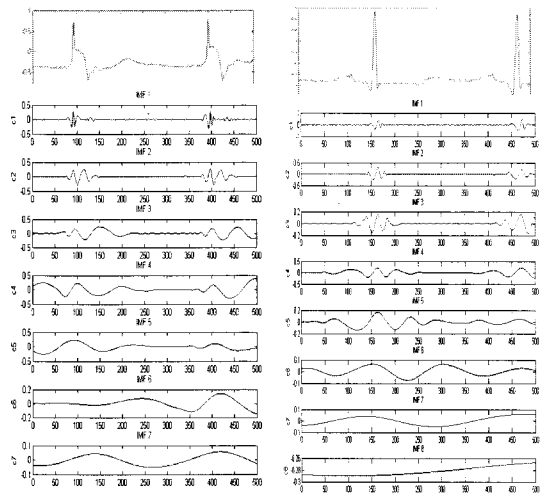


그림 2. 특징 추출을 위한 심전도 분석
Fig. 2. ECG analysis for feature extraction

$$S(t) = \sum_{i=1}^n C_i(t) + R_n(t) \quad (1)$$

원신호 $S(t)$ 는 N 개의 $IMFs$ 인 $C_1 \sim C_n$ 와 한 개의 잔여 신호 $R_n(t)$ 로 분해된다. 저차 $IMFs$ 는 신호의 고주파 정보를 나타내며, 고차 $IMFs$ 는 신호의 저주파 정보를 보여준다[10].

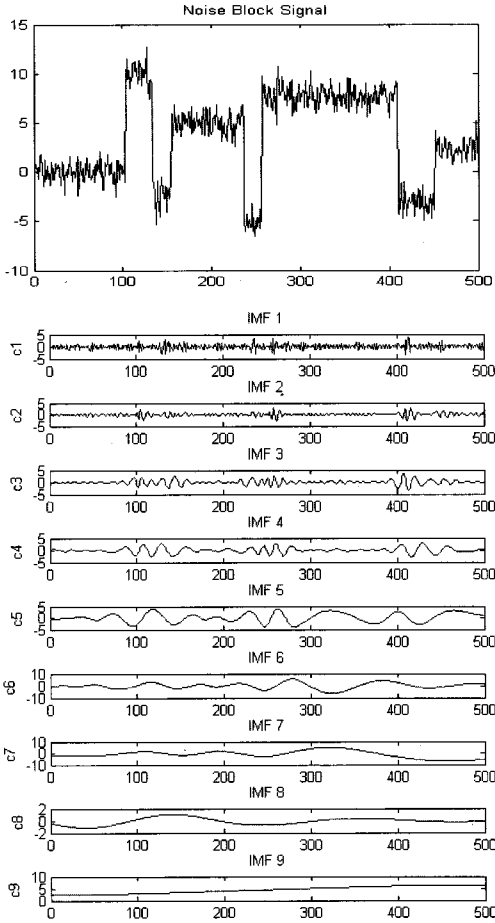


그림 3. EMD 방법에 의한 신호 분해
Fig. 3. Signal decomposition by EMD method

잔여 신호 $R_n(t)$ 를 포함하면서 고차 $IMFs$ 로부터 시작하여 저차 $IMFs$ 의 구성 요소인 C_n, \dots, C_2, C_1 의 결합으로 신호를 재구성할 수 있다. 그림 3은 IMF 구성요소들의 결합으로 원신호를 복원할 수 있음을 보여주고 있다.

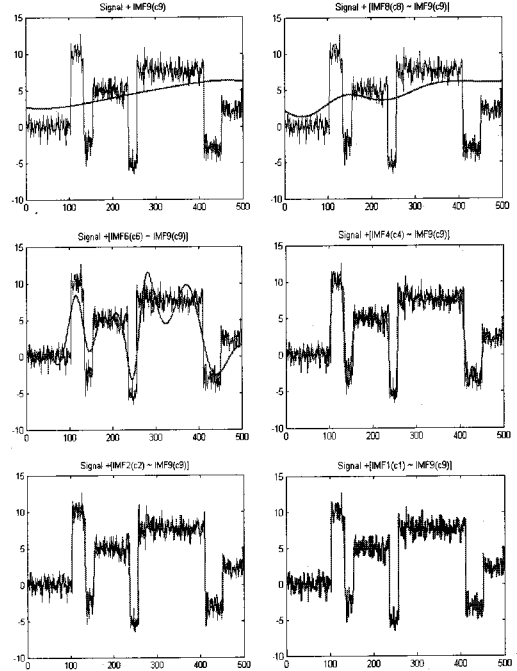


그림 4. IMF 구성요소에 의한 신호 복원
Fig. 4. Signal reconstruction from the IMF components

분해된 $IMFs$ 를 재구성하여 심전도 필터를 구현한 후 노이즈 제거 및 QRS complex 검출에 관한 자세한 내용은 [11]에서 제시하였다.

2.2 AR 모델링을 이용한 특징 추출

심전도 신호의 전처리로 노이즈 제거 및 부정맥 신호의 특징을 검출하기 위해 **EMD** 방법이 적용되었으며, **Burg** 알고리즘을 적용하여 분해된 신호 $IMFs$ 에 대해 **AR** 계수를 계산하고 다중 클래스 **SVM**의 입력으로 적용한다. 미리 선택된 모델 차수 P 는 **AR** 계수를 추정하기 위해 사용된다.

$$e_p^f(k) \equiv \sum_{l=1}^p a_l^p x(k-l) + x(k) \tag{2}$$

$$e_p^b(k) \equiv \sum_{l=1}^p a_l^p x(k+l-p) + x(k-p)$$

AR 계수는 재귀적인 방법으로 계산된다.

$$a_p^p = \frac{-E(e_{p-1}^f(k)e_{p-1}^{b*}(k-1))}{\frac{1}{2}E(|e_{p-1}^f(k)|^2 + |e_{p-1}^b(k-1)|^2)} \quad (3)$$

Levinson의 재귀 관계식을 사용하여 오류 신호들 사이의 재귀 관계를 얻고 다음 차수로 AR 계수를 추정하기 위한 계산에 사용된다.

III. 다중 클래스 SVM을 이용한 부정맥 분류

SVM은 이진 분류를 위해 고안된 분류기이다. 다중 클래스 문제에 SVM을 적용하는 문제는 두 가지 방법으로 구분되는데 몇 개의 이진 분류기를 결합하여 다중 클래스 분류기를 구성하거나 모든 클래스들을 동시에 고려하는 일반화된 분류기를 만드는 것이다. 다중 클래스 문제를 푸는 것은 계산적으로 복잡하기 때문에 심전도의 부정맥 분류와 같은 다중 클래스 문제를 해결하기 위해서는 입력 데이터의 차원을 줄일 수 있는 방법을 고려해야 한다.

본 논문에서는 다양한 부정맥 신호를 분류하기 위해 다중 클래스 SVM 방법을 제시한다.

$$\begin{aligned} \min_{w^j, b^j, \xi^j} & \frac{1}{2}(w^j)^T w^j + C \sum_i \xi_i^j \\ (w^j)^T \phi(x_i) + b^j & \geq 1 - \xi_i^j, \quad \text{if } y_i = i, \\ (w^j)^T \phi(x_i) + b^j & \leq -1 + \xi_i^j, \quad \text{if } y_i = j, \\ \xi_i^j & \geq 0 \end{aligned} \quad (4)$$

다중 클래스 문제를 해결하기 위해서 $k(k-1)/2$ 개의 분류기를 구성하며 i 와 j 번째 클래스들로 훈련 데이터에 대해 이진 분류 문제를 해결한다. 만일 식 (5)의 x 가 i 번째 클래스에 있다고 하면 i 번째 클래스에 대해 1을 증가시키는 'Max Wins' 전략을 사용한다.

$$\text{sign}((w^j)^T \phi(x) + b^j) \quad (5)$$

식 (4)의 훈련 데이터 x_i 는 함수 ϕ 에 의해 고차원 공간에 투영되며, 함수 ϕ 는 커널 함수로 RBF (radial basis function)를 사용하며, C 는 페널티 파라미터이다.

$\frac{1}{2}(w^j)^T w^j$ 를 최소화 하는 것은 데이터의 두 부류 사이의 마진인 $\frac{2}{\|w^j\|}$ 를 최대화 하는 것을 의미하며, 데이터가 선형 분리 불가능할 때 훈련 에러의 수를 줄일 수 있는 페널티 항 $C \sum_{j=1}^l \xi_j^i$ 이 있다. SVM은 마진 $\frac{1}{2}(w^j)^T w^j$ 과 훈련 오류 사이에 균형을 찾는 것이며 RBF 커널의 폭인 γ 과 C 의 최적값은 실험적인 방법을 통해 결정된다.

IV. 실험 결과

4.1 실험환경 및 데이터

심전도 신호를 분석하여 부정맥을 분류하기 위한 실험은 Windows XP에서 Pentium 2.4GHz, 1G RAM으로 Visual C++ 6.0과 Matlab 7.0을 사용하였다.

360Hz로 샘플링된 MIT-BIH 부정맥 데이터베이스와 250Hz로 샘플링된 MIT-BIH 심실 부정맥 데이터베이스, 그리고 128Hz로 샘플링된 MIT-BIH SVA 데이터베이스로부터 심전도 데이터 셋을 360Hz로 다시 샘플링하여 사용하였다. 데이터 셋은 심전도 정상신호(NSR), 심방 조기수축(APC), 상심실성빈맥(SVT), 심실조기수축(PVC) 및 심실빈맥(VT)과 심실세동(VF) 신호 각각 120 쌍의 세그먼트를 포함한다.

4.2 특징 추출

EMD를 통하여 어떤 신호이든지 몇 개의 IMF로 분해될 수 있으며, 분해가 데이터의 시간 스케일의 지역적인 특징에 기반을 두므로 심전도와 같은 비정상(non-stationery) 신호처리에 적용할 수 있음을 알 수 있다. 잡음이 있는 심전도 신호는 식 (6)과 같이 EMD 방법으로 분해한 후, 저차 IMF를 제거함으로써 저주파 필터

를 구현하여 적용하였다.

$$S_i(t) = \sum_{i=p}^n C_i(t) + R_n(t) \quad (6)$$

심전도 신호의 속성을 보전하면서 잡음이 제거된 신호 $S_i(t)$ 는 SVM의 훈련 데이터로 적용되기에 앞서 Burg 알고리즘에 의해 AR 계수값으로 전환되었다. SVM은 입력 데이터의 크기에 따라 성능 저하가 심하기 때문에 AR 계수값으로 특징을 추출하여 사용하였다.

표 1. 심전도 신호의 평균 AR 계수값
Table 1. Average AR coefficients of ECG

부정맥	AR 계수값			
NSR	-2.22	1.74	-0.69	-0.07
PVC	-2.27	2.18	-1.32	0.47
APC	-2.38	1.86	-0.43	-0.09
SVT	-2.81	3.25	-1.87	0.46
VT	-1.42	0.07	0.32	-0.01
VF	-1.75	0.36	0.537	-0.15

미리 선택된 모델 차수 $P=4$ 로 AR 계수를 추정하기 위해 사용되었으며 Burg 알고리즘은 예측 오차의 추정을 보여준다.

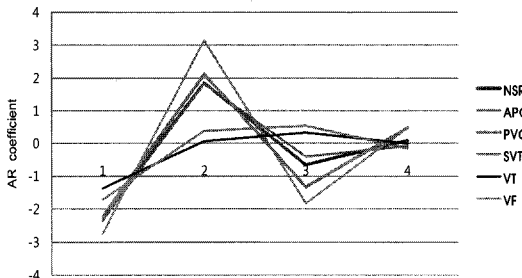


그림 5. 부정맥과 AR 계수값
Fig. 5. Cardiac arrhythmia and AR coefficients

4.3 다중클래스 SVM

본 실험에서는 부정맥 분류 문제를 해결하기 위하여

비선형 다중 클래스 SVM을 구성하고 마진과 오류 허용을 조절하기 위한 최적의 페널티 파라미터 C 값과 RBF 커널 파라미터 γ 값을 찾아 SVM 성능을 실험하였다. 비선형 SVM에서는 인식 시간이 서포트 벡터의 개수에 비례하므로 본 실험에서는 훈련 데이터로 네 개의 AR 계수값만을 사용하였으며, 네 개의 서포트 벡터가 되도록 하였다.

표 2. 파라미터에 따른 SVM 성능
Table 2. SVM performance parameters

$\gamma \backslash C$	10	50	100	200
0.5	75.5%	80.7%	83.5%	85.6%
1.0	95.3%	97.3%	97.6%	98.0%
2.0	95.0%	96.6%	97.5%	97.8%
3.0	85.7%	95.1%	97.0%	97.0%

표 2는 페널티 파라미터 C 값의 증가에 따라 성능이 향상되고 있음을 보여주고 있다. 마진을 희생하는 대신 분할 띠 안에 잘못 분류되는 훈련 데이터의 수를 줄이고 있음을 알 수 있다.

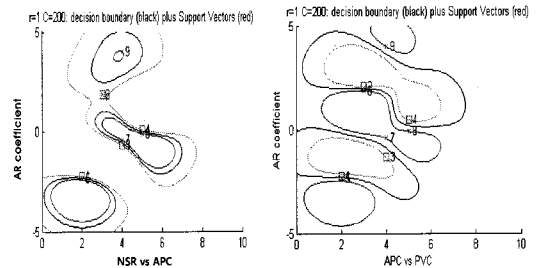


그림 6. $\gamma = 1, C = 200$ 에 따른 부정맥 분류
Fig. 6. Cardiac classification with $\gamma = 1, C = 200$

부정맥 분류를 위한 다중 클래스 SVM 성능을 평가하기 위하여 정확도를 사용하였으며 식(7)과 같다.

$$Accuracy = \frac{(TP + TN)}{(TP + FN + TN + FP)} \times 100 \quad (7)$$

표 3. 부정맥 분류 성능 평가

Table 3. Cardiac arrhythmia performance evaluation

Arrhythmia	NSR	APC	PVC	SVT	VT	VF
Accuracy (%)	99.1	97.4	96.8	98.0	99.2	99.5

V. 결 론

본 논문에서는 EMD 방법을 사용하여 심전도 신호를 IMFs로 분해하고, IMF를 이용하여 노이즈 제거 필터 및 QRS complex를 추출한다. 그리고 심전도의 특징 벡터로 AR 계수를 구하여 다중 클래스 SVM의 입력 데이터로 사용하여 심전도 정상신호(NSR), 심방조기수축(APC), 상심실성빈맥(SVT), 심실조기수축(PVC) 및 심실빈맥(VT)와 심실세동(VF)의 부정맥을 분류한다. 다양한 부정맥 분류를 위해 이진 SVM 분류기를 조합하여 다중 클래스 분류기를 구성하였으며 커널 함수로 RBF를 사용하였다. 다중 클래스 SVM의 최적 파라미터 γ 과 페널티 파라미터 C 값을 찾아 부정맥 분류에 적용하여 우수한 성능을 얻었으며, 평가지수로 정확도를 측정하여 96.8%~99.5%의 결과를 보였다. 본 논문에서 제시한 방법은 신호 데이터의 다양한 특징 추출 연 구 및 다중 클래스 분류 문제를 해결하는데 유용성이 크다.

향후 생체 신호의 의료 정보에 대한 주석 (annotation) 처리에도 활용 가능할 것이다.

참고문헌

[1] S.L.Melo, L.P.Caloba, J.Nadal, "Arrhythmia analysis using artificial neural network and decimated electrocardiographic data," *Comp. in Cardiology*, vol. 27, pp. 355-358, 2000.
 [2] R.Silipo, C.Marchesi, "Artificial Neural Networks for Automatic ECG Analysis," *IEEE Trans. Signal Processing*, vol. 46, pp. 1417-1425, 1998.

[3] M.B.Anuradha, V.C.Veera Reddy, "Cardiac Arrhythmia Classification using Fuzzy Classifiers," *Journal of Theoretical and Applied Information Technology*, vol. 4, no. 4, pp. 353-359, 2005.
 [4] M.Nambakhsh, V.Tavakoli, N.Sahba, "FPGA -core defibrillator using wavelet-fuzzy ECG arrhythmia classification," in *Proc. of IEEE Conf. on Engineering in Medicine and Biology Society*, pp. 2673-2676, 2008.
 [5] I.Chouvarda, N.Maglaveras, C.Pappas, "Adaptive time-frequency ECG analysis," *Computers in Cardiology 2001*, pp.265-268, 2001.
 [6] M.G.Tsipouras, V.P.Oikonomou, Di Fotiadis, L.K.Michalis, D.Sideris, "Classification of atrial tachyarrhythmias in electrocardiograms using time frequency analysis," *Comp. in Cardiology*, pp. 245-248, 2004.
 [7] M.I.Oiws, A.H.Abou-Zied, A.M.Youssef, "Study of Features based on Nonlinear Dynamical Modeling in ECG Arrhythmia Detection and Classification," *IEEE Trans. Biomedical Eng.*, vol. 49, no. 7, pp. 733-736, 2002.
 [8] H.X.Zhang, Y.S.Zhu, Z.M.Wang, "Complexity measure and complexity rate information based detection of ventricular tachycardia and fibrillation," *Medical and Biological Eng. and Comp.*, vol. 38, no. 5, pp. 553-557, 2000.
 [9] N.E.Huang, Z.Shen, S.R.Long, M.C.Wu, E.H.Shih, Q.Zheng, Yen. N.-C., C.C.Tung and H.H.Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc., London. A*, vol. 454, pp. 903-995, 1998.
 [10] H.Zhang and Q.Gai, "Research on properties of empirical mode decomposition method," *Proc. of the 6th World Congress on Intelligent Control and Automation*, vol. 2, pp. 10001-10004, 2006.
 [11] 이금분, 조범준, "EMD 방법을 이용한 ECG 신호 필터링," *한국해양정보통신학회논문지*, vol. 13 no. 12, pp.2671-2676, 2009.

저자소개



이금분(Geum-Boon Lee)

2002년 대전대학교 컴퓨터공학과
공학 석사

2007년 조선대학교 컴퓨터공학과
박사수료

※ 관심분야: 영상처리, 신호처리, 패턴인식, 신경망



조범준(Beom-Joon Cho)

1980년 조선대학교(B.S., M.S.(82))

1988년 한양대학교 전기공학과
공학박사

2004년 KAIST 전자전산학과
공학박사

1980년~현재 조선대학교 컴퓨터공학부 교수

※ 관심분야: 인공지능, 패턴인식, 뉴로컴퓨터