

원저

고차통계와 Hermite 모델을 이용한 계층적 심전도 비트 분류

박관수¹, 조백환³, 이도훈¹, 송수화¹, 이종실¹, 지영준², 김인영¹, 김선일¹

한양대학교 의용생체공학과¹, 울산대학교 공과대학 의공학과²,
Dept. of Biomedical Engineering, Carnegie Mellon Univ.³

Hierarchical Classification of ECG Beat Using Higher Order Statistics and Hermite Model

Kwan Soo Park¹, Baek Hwan Cho³, Do Hoon Lee¹, Su Hwa Song¹,
Jong Shill Lee¹, Young Joon Chee², In Young Kim¹, Sun I. Kim¹

Dept. of Biomedical Engineering, Hanyang Univ.¹,
Dept. of Biomedical Engineering, College of Engineering, University of Ulsan²,
Dept. of Biomedical Engineering, Carnegie Mellon Univ.³

Abstract

Objective: The heartbeat classification of the electrocardiogram is important in cardiac disease diagnosis. For detecting QRS complex, conventional detection algorithm have been designed to detect P, QRS, T wave, first. However, the detection of the P and T wave is difficult because their amplitudes are relatively low, and occasionally they are included in noise. Furthermore the conventional multiclass classification method may have skewed results to the majority class, because of unbalanced data distribution. **Methods:** The Hermite model of the higher order statistics is good characterization methods for recognizing morphological QRS complex. We applied three morphological feature extraction methods for detecting QRS complex: higher-order statistics, Hermite basis functions and Hermite model of the higher order statistics. Hierarchical scheme tackle the unbalanced data distribution problem. We also employed a hierarchical classification method using support vector machines. **Results:** We compared classification methods with feature extraction methods. As a result, our mean values of sensitivity for hierarchical classification method (75.47%, 76.16% and 81.21%) give better performance than the conventional multiclass classification method (46.16%). In addition, the Hermite model of the higher order statistics gave the best results compared to the higher order statistics and the Hermite basis functions in the hierarchical classification method. **Conclusion:** This research suggests that the Hermite model of the higher order statistics is feasible for heartbeat feature extraction. The hierarchical classification is also feasible for heartbeat classification tasks that have the unbalanced data distribution. (*Journal of Korean Society of Medical Informatics 15-1, 117-131, 2009*)

Key words: Electrocardiogram, Higher Order Statistics, Hermite Basis Function, Support Vector Machine, Hierarchical classification

논문투고일: 2008년 11월 10일, **심사완료일:** 2009년 2월 9일

교신저자: 김인영, 서울시 성동구 행당1동 17번지 한양대학교 의용생체공학과 (133-791)

Tel: 02-2291-1713, **Fax:** 02-2296-5943, **E-mail:** iykim@hanyang.ac.kr

* 본 논문은 지식경제부에서 지원하는 실버의료기기 핵심기술개발과제 지원에 의한 결과입니다.

I. 서론

심전도(electrocardiogram, ECG)신호는 심장활동 시에 나타나는 전기적 신호로써 심장 상태와 질환을 알아볼 수 있는 중요한 자료로 쓰인다¹⁾. 부정맥(arrhythmia)은 심전도 형태를 띤 연속된 비정상적 형태로 드물게 발생할 수 있으며, 이를 찾기 위해 24시간 이상 환자의 심전도를 측정할 수 있는 홀터(holter)를 이용하여 장시간의 심전도데이터를 저장하게 된다. 따라서, 이러한 대량의 데이터를 사람이 직접 분석하는 데에는 많은 시간과 비용이 들기 때문에, 환자의 부정맥 진단을 위한 자동화 기술들이 개발되고 있다. 부정맥을 식별하는데 있어서 가장 중요한 단계는 각 심장박동이 어떤 클래스(class)에 속하는지를 정확히 분류(classification)하는 것이며, 그 결과 중 비정상적인 심장박동이 연속적으로 나타나는 특정 패턴에 따라 다양한 부정맥 리듬으로 진단하게 된다²⁾.

심전도 파형은 일반적으로 P, QRS, T파로 구분되고, 이들 파형 중 QRS 콤플렉스(complex)는 심장박동이 건강한 리듬 비트인지 아니면 부정맥으로 인한 비정상 리듬 비트인지를 결정하고 해석하는데 가장 중요하게 쓰이는 부분이다. 심전도의 심장박동클래스 분류를 위한 심전도 파형 특징점 추출방법에는 심전도 형태학적 특징추출, 심장박동 간격의 특징추출, 그리고 주파수 관련 특징추출 등 여러 가지가 있다. 여기서, 전형적인 심전도 특징점 추출방법은 기본적으로 P, QRS, T파의 진폭과 위치 정보 그리고 심장박동의 간격 등의 특징들을 추출하여 특징점으로 사용한다. 이 방법은 특징점을 추출함에 있어 P, QRS, T파가 정상적으로 검출되었을 때에는 가능하지만, 종종 P파나 T파의 검출을 정확히 못한 경우에는 특징점 추출에 문제가 있을 수 있고, 결국 정확한 분류에 악영향을 줄 수 있다³⁾. 이를 해결하고자 다른 연구에서는 P, QRS, T파의 검출절차 없이 심장박동의 형태를 이용하여 특징점을 찾을 수 있는 고차통계(higher order statistics, HOS)와 Hermite basis function(HBF)이 사용되어 왔다³⁻⁵⁾.

또한, 심전도 심장박동 분류를 위한 분류기(classifier)에는 linear discriminants, back propagation neural networks, self-organizing maps 등이 널리 사용되어 왔다²⁾. 이들 방법들은 모든 클래스를 한꺼번에 학습하여 분류기를 생성하며, 모든 클래스를 동시에 찾아야 하는 문제점

이 있다. 다중클래스(multiclass) 분류방법으로써, 일대다(one-against-all)분류법은 해당 클래스와 그 외의 나머지 클래스를 비교하여 가장 높은 출력값을 나타내는 클래스로 분류하는 방법이고, 일대일(one-against-one, pairwise)은 클래스들을 서로 쌍을 이루어 비교하여 가장 많은 투표값을 갖는 클래스로 분류하는 방법이다. 일반적으로 일대일 방법이 더 좋은 성능을 나타낸다고 알려져 있지만, 한 클래스에 속하는 데이터의 수가 적을 때는 데이터 분포의 불균형으로 인한 치우친 결과를 나타낼 수도 있다^{6,7)}. 또 다른 연구에서는 다중클래스를 분류하는데 있어, 모든 클래스를 동시에 찾아내는 분류보다 클래스를 여러 단계를 거쳐 각각의 클래스를 하나씩 찾아내는 연구가 진행되어 왔다. 다중클래스를 여러 개의 이진(binary)분류를 사용하여 분류하는 방법인 계층적(hierarchical)분류방법이 사용되어 다중클래스 분류에서 좋은 성능을 발휘하고 있다⁸⁾.

심전도를 심장박동 별로 비교해 보면, 정상심장박동이 비정상심장박동보다 훨씬 많다. 이러한 불균형한 데이터 분포의 경우, 분류기가 이들 데이터를 학습하여 대다수(majority) 클래스로 치우치는 결과(skewed result)가 나올 가능성이 높다. 이를 해결할 수 있는 가장 간단한 방법으로는 majority 클래스 데이터 개수를 소수(minority) 클래스 데이터 개수와 같게 샘플링을 하거나, minority 클래스 데이터를 majority 클래스 데이터 개수만큼 복제(replication)하여 분류하는 방법이다. 이러한 방법으로 생성된 분류기 규칙이나 예측 모델은 본래의 집단을 잘 나타낸다고 할 수 없고, 테스트 성능 또한 보장하기 어렵다는 점이 있다⁹⁾. 이는 receiver operating characteristics(ROC) 곡선분석방법을 사용하여 문턱치(threshold)를 조절함으로써 데이터의 민감도(sensitivity)와 특이도(specificity) 변화를 비교하여 문제를 해결할 수 있다. 이를 통해 한쪽으로 치우친 데이터 분포결과에서 majority 클래스와 minority 클래스를 잘 분류할 수 있는 분류기 모델을 만든다¹⁰⁾.

따라서, 본 실험에서는 심전도 심장박동의 특징점 추출 방법으로 HOS, HBF, 그리고 고차통계의 Hermite model(Hermite model of higher order statistics, HMH)을 사용하였다. 또한, 계층적 분류는 심전도의 특정영역지식을 이용한 가정을 통해 계층적 분류단계를 구성하였고, 심장박동 클래스를 분류하기 위한 분류기로는 support vector machine(SVM)을 사용하였다. 전

형적인 다중클래스 분류방법과 계층적 분류방법의 결과를 비교하였으며, 계층적 분류방법에서는 ROC 곡선분석을 통해 데이터 불균형 문제를 해결하여 분류기의 성능을 향상하고자 했다.

II. 재료 및 방법

1. 데이터 수집

MIT-BIH 부정맥 데이터베이스는 48명 환자의 두 채널 심전도 데이터로써, 이들 기록(recording)은 환자 별로 30분 정도의 데이터를 가지고 있다. 이 기록들 중 23개 기록은 입원환자(약 60%)와 외래환자(약 40%) 4,000명으로부터 임의로 선택된 것이며, 나머지 25개 기록은 위 환자 중 임상적으로 중요한 부정맥이 있는

환자기록이 선택된 것이다¹¹⁾. MIT-BIH 부정맥 데이터베이스는 0.1~100Hz band-pass 필터링한 신호를 11 bit로 360Hz 샘플링한 것이다. 심장박동은 15개 타입으로 나뉘어 있고, 각 기록 안에는 심장박동에 따른 타입과 기준점(fiducial point)의 시간에 대한 정보가 주석 안에 포함되어있다¹¹⁾. Association for the Advancement of Medical Instrumentation(AAMI) 권고에 따라 48개 기록들 중 심장박동 조절기에 의한 데이터가 포함된 4개 기록은 삭제하여 44개 기록을 사용하였다¹²⁾. Figure 1은 심전도로부터 심장박동 클래스 분류를 하기 위한 전 과정을 나타내었다. 심전도에서 클래스 분류까지 3단계로 구성되어 있는데, 첫 번째 단계는 심전도에 포함된 artifact를 제거하기 위한 전처리 단계이고, 두 번째 단계는 QRS 콤플렉스와 R피크 기준점을 사용하여 추출 방법에 따른 특징점 추출단계이다.

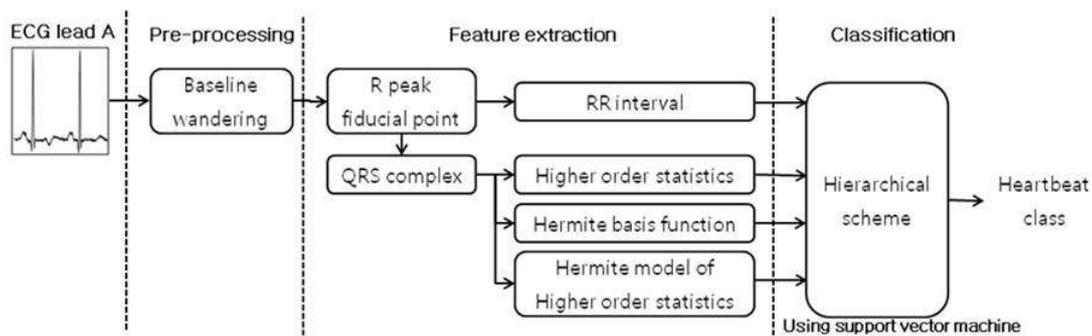


Figure 1. The procedure for classification of heartbeat class from electrocardiogram

Table 1. Merging the MIT-BIH heartbeat types to the AAMI heartbeat classes

AAMI heartbeat class	Description	MIT-BIT heartbeat types
N	Any heartbeat not in the S, V, F or Q classes	normal beat (NOR)
		left bundle branch block beat (LBBB)
		right bundle branch block beat (RBBB)
		atrial escape beat (AE)
		nodal (junctional) escape beat (NE)
S	Supraventricular ectopic beat	atrial premature beat (AP)
		aberrated atrial premature beat (aAP)
		nodal (junctional) premature beat (NP)
V	Ventricular ectopic beat	supraventricular premature beat (SP)
		premature ventricular contraction (PVC)
F	Fusion beat	ventricular escape beat (VE)
		fusion of ventricular and normal beat (fVN)
Q	Unknown Beat	paced beat (P)
		fusion of paced and normal beat (fPN)
		unclassifiable beat (U)

마지막 단계는 SVM을 이용한 계층적 방법을 통해 심장박동 클래스를 분류하는 단계이다.

AAMI 권고에 따라 본 실험에서는 15개 MIT-BIH 심장박동 타입을 5개 심장박동 클래스로 합쳐 Table 1과 같이 구성하여 연구를 하였다¹²⁾. Q 클래스에는 분류할 수 없는 비트(unclassifiable beat)와 페이스메이커 비트(paced beat)가 포함되어 있고, 더욱이 Q 클래스로 속한 심장박동비율이 너무 적어 본 실험에서는 Q 클래스를 제외한 나머지 클래스데이터를 이용하여 실험을 진행하였다.

또한, MIT-BIH 부정맥 데이터베이스 44개 기록을 각 클래스의 분포에 비례하게 22개씩 기록을 나누어 두 개의 데이터세트로 구성하였다. 각 데이터세트에는 약 50,000개의 심장박동들이 있으며, Table 2는 각 데이터세트의 심장박동 클래스 별 구성과 비율을 보여준다.

Table 2. Organized Dataset 1(DS1) and Dataset 2(DS2) from MIT-BIH dataset

	N	S	V	F	Total
DS1	45,868	943	4,259	415	51,013
(Ratio, %)	(89.08)	(1.83)	(8.24)	(0.81)	(100)
DS2	44,259	1,837	3,221	388	49,705
(Ratio, %)	(89.03)	(3.71)	(6.48)	(0.78)	(100)

두 개의 데이터세트 중 Dataset1(DS1, validation set)은 분류기에서 가장 좋은 성능을 보이는 매개변수를 찾기 위한 validation 데이터로 사용하였고, dataset2(DS2, testing set)는 DS1에서 결정된 분류기의 최종 성능평가를 위한 test 데이터로 사용하였다²⁾.

2. 전처리

본 연구에서 사용된 MIT-BIH 부정맥 데이터들은 심전도 신호의 baseline wandering 제거를 위해 두 번의 메디안 필터(median filter)를 사용하였다. 처음에는 QRS 콤플렉스와 P-파 제거를 위해 200ms 메디안 필터를 사용하였고, T-파 제거를 위해 600ms 메디안 필터를 사용하였다. 두 번의 메디안 필터를 거쳐 나온 신호가 심전도 신호의 baseline이므로, 원 신호(original signal)에서 baseline을 제거한 심전도 신호를 본 실험에 사용하였다²⁾.

3. 특징점 추출(Feature extraction) 방법

(1) 고차통계(higher order statistics, HOS)

HOS 방법 중 하나인 2차, 3차, 4차 cumulants를 사용하였는데, cumulants는 cumulants 생성 함수인 테일러 전개(Taylor expansion)의 계수를 말하며, cumulants의 선형 또는 비선형 조합들을 통계적 모멘트(statistical moments)로 표현하기도 한다³⁾. 평균이 0인 통계적 프로세스일 때, $x(t)$ 의 2차, 3차 cumulants는 이와 대응되는 2차, 3차 통계적 모멘트로 나타낼 수 있다³⁾.

$$\begin{aligned} c_{2x}(\tau_1) &= m_{2x}(\tau_1) \\ c_{3x}(\tau_1, \tau_2) &= m_{3x}(\tau_1, \tau_2) \end{aligned} \tag{1}$$

그리고 4차 cumulants는 4차와 2차 모멘트의 정보를 사용하여 만들어 진다.

$$\begin{aligned} c_{4x}(\tau_1, \tau_2, \tau_3) &= m_{4x}(\tau_1, \tau_2, \tau_3) \\ &\quad - m_{2x}(\tau_1) m_{2x}(\tau_3 - \tau_2) \\ &\quad - m_{2x}(\tau_2) m_{2x}(\tau_3 - \tau_1) \\ &\quad - m_{2x}(\tau_3) m_{2x}(\tau_2 - \tau_1) \end{aligned} \tag{2}$$

c_{nx} 와 m_{nx} 는 $x(t)$ 의 n 차 cumulants와 n 차 통계적 모멘트를 나타내고, 이때 τ_1, τ_2 , and τ_3 는 시간간격(time lags)이다. 2차 cumulants는 신호의 자기상관(autocorrelation)을 나타내고, 3차와 4차 cumulants는 왜도(skewness)와 첨도(kurtosis)를 나타낸다. 같은 사람의 심장박동이라고 할지라도 같은 심장박동 클래스에서 QRS 콤플렉스의 진폭은 서로 다르게 나타난다. 이들 심장박동에 cumulants를 적용함으로써 같은 심장박동 클래스에서 진폭간의 퍼짐이 줄어들게 된다.

본 실험에서는, 심전도 비트의 R피크를 기준으로 전후 250ms(90개 샘플 포인트)를 취한 $x(t)$ 를 구성하여, 즉 0.5초의 데이터를 이용하여 2차, 3차, 4차 cumulants를 구한다. R피크는 MIT-BIH 부정맥 데이터베이스가 제공하는 annotation에 있는 R 피크 기준점을 사용하여 수행하였다. 2차, 3차, 4차 cumulants를 수행하여 나온 결과를 Figure 2에 나타내었다. 각 클래스에 따른 2차, 3차, 4차 cumulants를 비교해 보면 각 클래스간의 서로 다른 점을 찾을 수 있으며, 차수가 높아짐에 따라 클래스간의 차이가 더욱 더 잘 구별됨을 알 수 있다(Fig. 3), (Fig. 4).

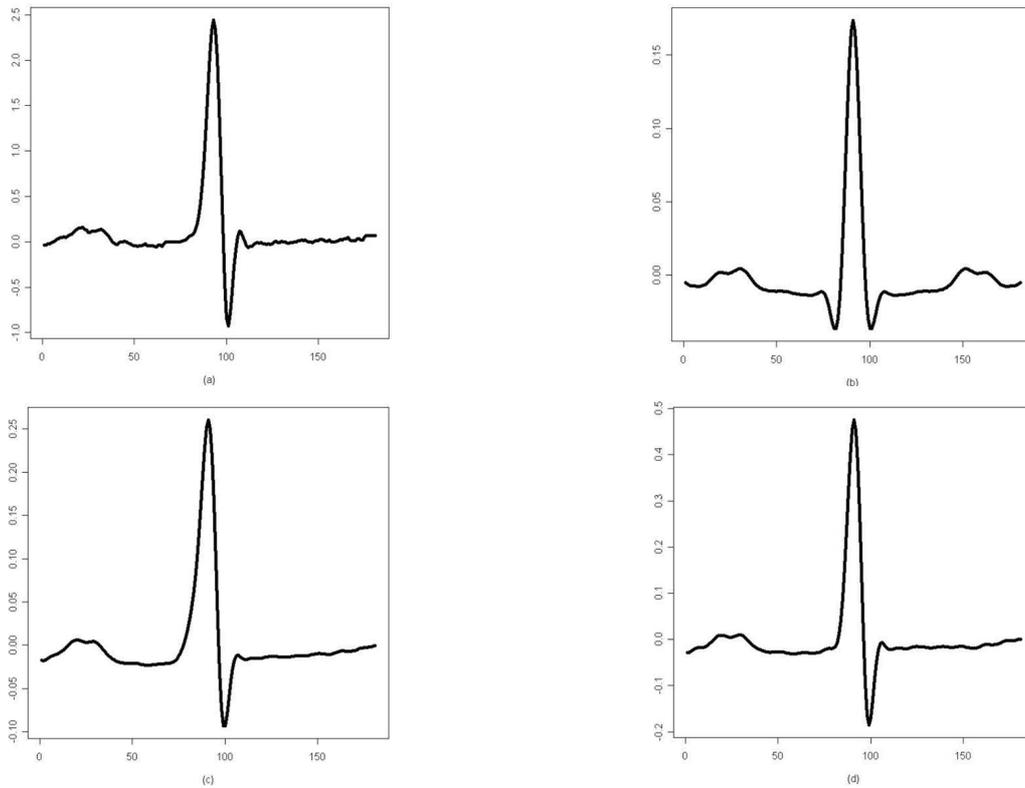


Figure 2. (a) The N class waveform of ECG beat and its cumulant representations of the (b) second, (c) third, and (d) fourth order

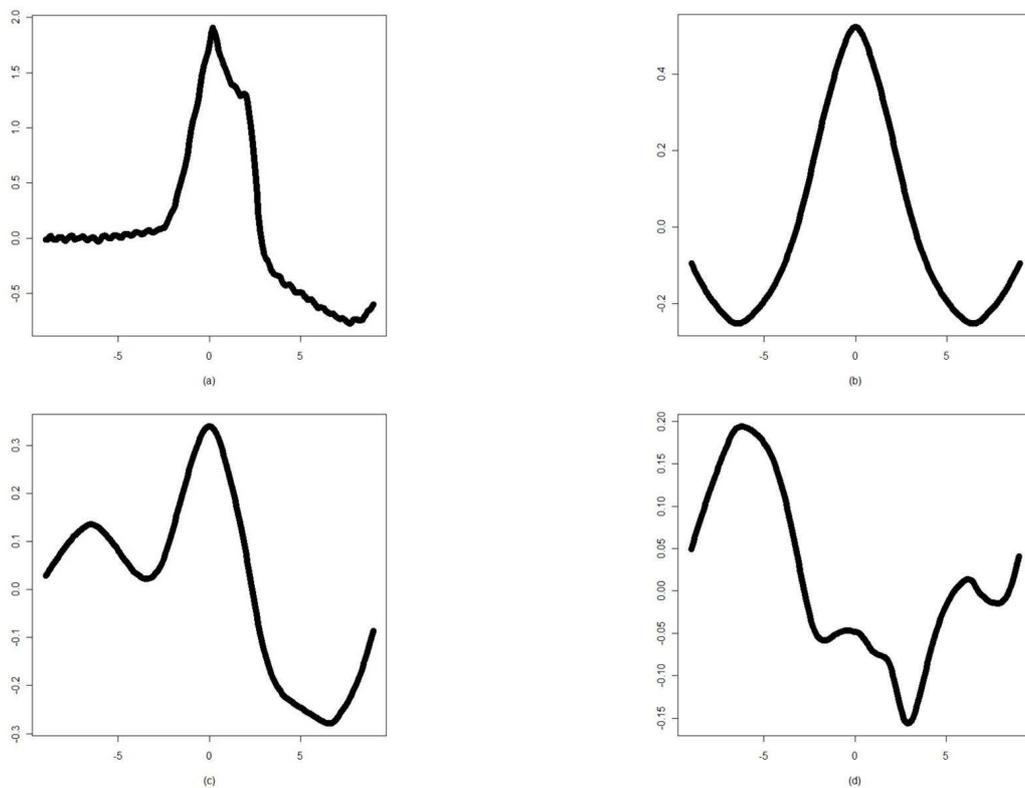


Figure 3. (a) The V class waveform of ECG beat and its cumulant representations of the (b) second, (c) third, and (d) fourth order

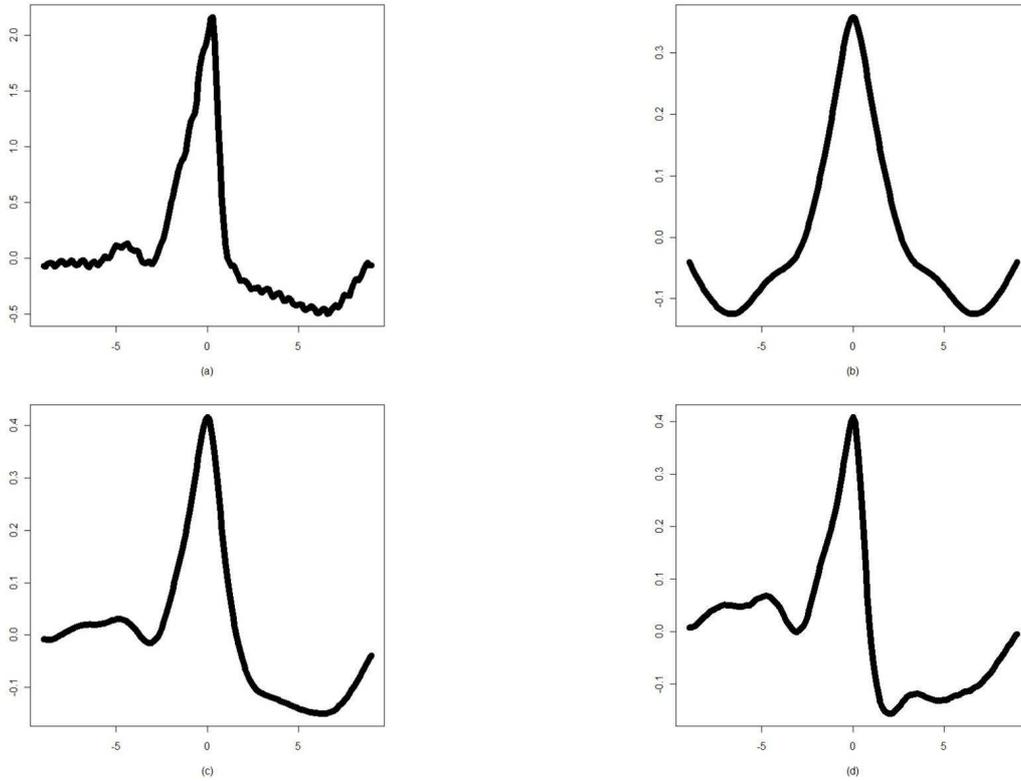


Figure 4. (a) The F class waveform of ECG beat and its cumulant representations of the (b) second, (c) third, and (d) fourth order

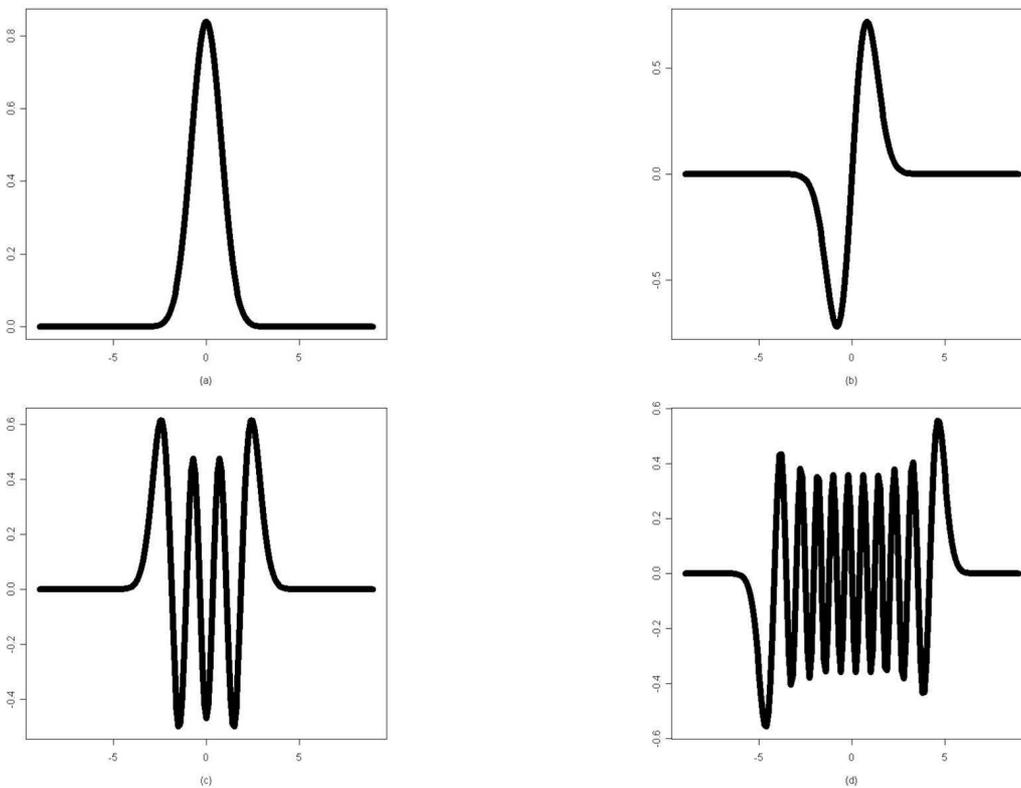


Figure 5. For $\sigma=0.8$, (a) $n=1$, (b) $n=2$, (c) $n=7$, and (d) $n=20$ Hermite basis function

2차, 3차, 4차 cumulant를 수행하여 나오는 모든 결과값을 특징점으로 사용하게 되면 학습(learning)과정에서 오버피팅(over-fitting)되어 테스트과정에서 나쁜 성능을 나타내는 경우가 있을 수 있다. 이를 방지하고자 2차, 3차, 4차 cumulants 결과값에서 QRS 콤플렉스에서 임의로 정한 10개 시간간격 점에 해당하는 cumulants 값을 추출하여 총 30개의 특징점을 사용하였다. 여기서 시간간격 점으로 15, 30, 45, 60, 75, 90, 105, 120, 135 그리고 150을 선택하였다.

(2) Hermite basis function (HBF)

HBF은 푸리에 시리즈(Fourier series)와 유사하게 여러 개의 waveform의 합으로 새로운 신호를 생성할 수 있고, 따라서 이를 이용하여 심전도의 QRS 콤플렉스를 근사화하는데 사용할 수 있다⁴⁾. HBF 전개는 계수들은 심전도 비트의 QRS 콤플렉스 커브를 대표하는 특징점들이 된다. x(t)를 심전도 비트의 QRS 콤플렉스 파형이라 했을 때, 다음은 Hermite 시리즈 전개이다.

$$x(t) = \sum_{n=0}^{N-1} c_n \phi_n(t, \sigma) \quad (3)$$

여기서 c_n 은 전개 계수이고 σ 는 width 매개변수이며 $\phi_n(t, \sigma)$ 는 n차의 HBF이다. $\phi_n(t, \sigma)$ 는 다음과 같이 정의된다⁴⁾¹³⁾.

$$\phi_n(t, \sigma) = \frac{1}{\sqrt{\sigma 2^2 n! \sqrt{\pi}}} e^{-\frac{t^2}{2\sigma^2}} H_n\left(\frac{t}{\sigma}\right) \quad (4)$$

$H_n(t/\sigma)$ 는 n차 Hermite 다항식이다. Hermite 다항식들은 재귀관계를 이용하여 구할 수 있다.

$$H_n(x) = 2xH_{n-1}(x) - 2(n-1)H_{n-2}(x) \text{ with } H_0(x) = 0, \text{ and } H_1(x) = 2x \text{ for } n=2,3,\dots \quad (5)$$

Figure 5는 n차 HBF을 도시한 것이며, Hermite 다항식이 고차로 갈수록 타임 도메인(time domain)내에서 신호의 변화가 빈번하게 일어나게 되어 심전도 신호의 빠른 변화에 빠르게 적응될 수 있게 한다.

전개계수 c_n 은 오차 제곱의 합(Sum of Squared Errors)의 최소화를 통해서 계산할 수 있으며, 오차는 다음과 같이 정의된다.

$$E = \sum_i \left\| x(t_i) - \sum_{n=0}^{N-1} c_n \phi_n(t, \sigma) \right\|^2 \quad (6)$$

오차 함수는 계수 c_n 을 구하기 위한 선형 방정식이고 위 방정식을 풀기 위해 특이값 분해(singular value decomposition, SVD)와 pseudo-inverse 기술을 이용하였다¹⁴⁾.

HOS 방법과 같이 심전도 비트는 R피크를 기준으로 앞뒤 90개 샘플(250ms)을 취하여 총 181개 샘플이 QRS 콤플렉스를 나타내는 x(t)가 되어 HBF 전개를 수행하였다. 경험적으로 20차 Hermite 다항식을 선택하였고, 이 때의 20개의 Hermite 계수를 특징점으로 구성하여 실험하였다. Figure 6은 심전도 비트 클래스 N에 적용한 것이고, 이에 해당하는 20차 Hermite 다항식으로 나타낸 그래프가 QRS 콤플렉스를 잘 표현하고 있음을 알 수 있다.

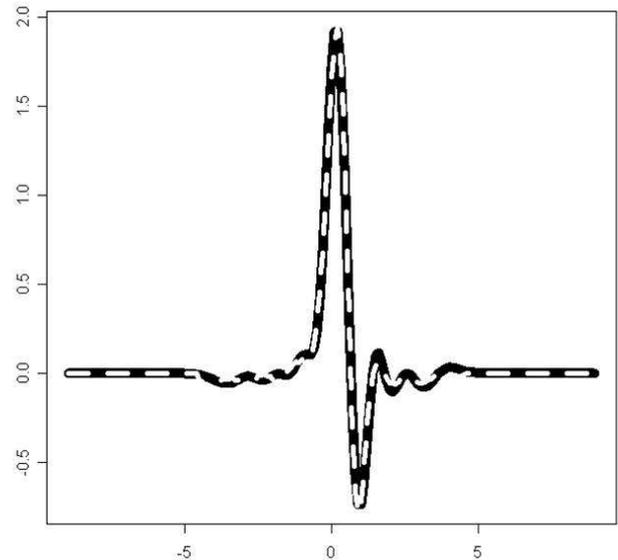


Figure 6. The normal QRS complex of an ECG beat (black solid line) and its estimation using a Hermite polynomial of the 20th order(white dashed line)

(3) 고차통계의 Hermite 모델(Hermite model of higher order statistics, HMH)

HOS인 cumulants는 통계적인 방법을 이용하여 심전도 심장박동을 재해석함으로써 각각의 심장 클래스 간의 차이점을 잘 구별하여 나타낼 수 있고, HBF은 원상태의 심전도 QRS 콤플렉스를 Hermite 다항식의

계수들의 조합으로 형상화하여 잘 나타내고 있다²⁴⁾. 고차통계인 경우 2차, 3차, 4차 cumulants에서 나온 결과를 샘플링하여 사용하면 샘플링된 특징점들이 QRS 콤플렉스의 특징을 잘 나타내고 있는지 확인이 불가능하며, 2차, 3차, 4차 cumulants결과 모두를 특징점으로 사용하면 학습과정에서 오버피팅되어 테스트과정에서 나쁜 성능을 나타낼 수 있다. 또한 HBF의 경우, 잡음이 포함된 QRS콤플렉스에 HBF를 적용하면 과형에 남아있는 잡음이 특징점 추출에 영향을 주어 분류의 성능을 떨어질 수 있다. 이런 문제들을 해결하고자 HOS와 HBF의 조합방법을 사용하게 되는데, 이는 HOS를 통해 재해석된 심전도 심장박동의 2차, 3차, 4차 cumulants에 HBF를 적용함으로써, HOS에서 나온 특징점의 수를 줄이는 동시에 잡음을 제거한다. 먼저, QRS콤플렉스 $x(t)$ 에 HOS를 적용하여 2차, 3차, 4차 cumulants를 구하고, 각각의 cumulants과형에 20차 Hermite 다항식을 적용하여 나온 계수를 특징점으로 사용하여 실험을 하였다.

(4) RR interval

앞서 제시한 특징점 추출 방법들뿐만 아니라, R피크를 이용한 세가지 RR간격 특징점을 추가적으로 추출하여 실험을 수행하였다. 첫 번째 RR간격 특징점은 현재 R피크와 이전 R피크 사이의 RR간격을 특징점으로 사용하였고 두 번째 특징점은 현재 비트로부터 그전 10개 RR간격의 평균을 특징점으로 사용하였다¹⁴⁾. 마지막 간격 특징점은 RR간격을 그전 10개 RR간격평균으로 나눈 값을 특징점으로 사용하였다. 같은 심전도 심장박동 클래스라도 사람에 따라 심전도의 RR간격이 달라지고, 그전 10개 RR평균 간격 역시 사람에 따라 달라진다. 이 때문에 첫 번째와 두 번째 특징점으로는 심전도 심장박동 클래스 분류에 어려움이 있어 RR간격과 그전 10개 RR간격평균의 비를 추가적으로 적용한다.

4. Support Vector Machine (SVM)

SVM은 지도학습(supervised learning)을 하는 분류 기이며, 좋은 일반화 성능(generalization performance)을 가지고 있어 데이터 분류에 응용되고 있다. SVM은 입력공간에 주어진 두 클래스의 데이터를 분류시키기 위해 초월평면(hyper-plane)을 만들게 된다. 이때

최대 마진(margin)을 갖는 초월평면을 찾게 되는데, 마진은 초월평면과 가장 가까이에 있는 두 클래스 샘플간의 거리를 말한다. 이는 트레이닝에서 나온 최대 마진 초월평면을 갖고 테스트 데이터를 분류할 때, 보다 정확한 예측을 할 수 있다는 가정이 있기 때문이다. 여기서 초월평면과 가장 가까이 있는 데이터 샘플들을 support vector(SV)라고 한다¹⁵⁾¹⁶⁾.

일반적으로 샘플들을 두 클래스로 구분할 수 있는 초월평면이 존재하는 경우와 존재하지 않을 경우로 나뉜다. 초월평면이 존재하지 않을 경우, 트레이닝 시에 나오는 에러 페널티(error penalty)를 어느 정도 허용하면서 초월평면에서 가장 가까이 있는 샘플간의 거리가 최대가 되는 소프트마진 방법을 사용하고 에러 페널티가 선형일 때 다음과 같이 구할 수 있다.

$$\begin{aligned} &\text{minimize} && \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^n \varepsilon_i \\ &\text{subject to} && y_i(W \cdot X_i + b) \geq 1 - \varepsilon_i, \quad i = 1, \dots, n \\ &&& \varepsilon_i \geq 0, \quad i = 1, \dots, n \end{aligned} \tag{7}$$

여기서, W 는 가중치 벡터(weight vector)를 나타내고, X_i 는 평면상의 입력 벡터이고, b 는 바이어스이며, ε_i 는 오분류 척도로 정의한다. 소프트 마진 방법은 하드마진 방법과 비교하여 볼 때 C 값을 포함한 에러 페널티 부분이 추가되어 있고, 위 식을 최적화하기 위해서는 최대 마진을 가지면서 최소 에러 페널티를 가져야 한다. 에러 페널티가 선형일 때에는 dual problem을 이용하여 계산상의 효율 향상을 기대할 수 있다. 다음은 dual problem을 나타낸 것이다¹⁶⁾.

$$\begin{aligned} &\text{minimize} && \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j \alpha_i \alpha_j K(x_i, x_j) \\ &\text{subject to} && 0 \leq \alpha_i \leq C, \quad i = 1, \dots, n \\ &&& \sum_{i=1}^n \alpha_i y_i = 0 \end{aligned} \tag{8}$$

여기서 α 는 Lagrange multiplier로써, 모든 트레이닝 데이터 중 SV 데이터는 양수 값을 갖고 SV외에 다른 데이터는 0값을 갖게 된다. 또한 K 는 커널(kernel)이며, 커널은 선형커널과 radial basis function(RBF)과 같은 비선형커널로 나뉘는데 이는 연구 목적과 데이터 분포에 따라 적절히 선택하여 사용하게 된다. 최종적으로 다음과 같은 결정 함수(decision function)식을 구

할 수 있고, 테스트 데이터를 이 함수에 적용했을 때, 결과값으로 양수값이 나오면 양성, 음수값이 나오면 음성으로 판별한다.

$$F(x) = \text{sgn} \left(\sum_{i=1}^n \alpha_i y_i K(x, x_i) + b \right) \quad (9)$$

멀티클래스 분류 문제에 있어서는 주로 pair-wise 방식을 사용하였다. Pair-wise 방식은 k개의 클래스를 분류하기 위해 (k-1)*2개의 이진분류를 구성하여 SVM 수행 후 나온 결정값을 통해 가장 많은 투표값을 가진 클래스를 판별하여 채택하는 방식이다. SVM을 이용한 분류방법은 확률 정보를 이용하여 클래스 라벨(label)을 예측하는 것이 아니라 결정 함수에서 나온 결정값을 이용한다. 결정값이 아닌 확률값을 추정하기 위해서는 SVM에서 나온 결정값을 이용하여 다음과 같이 구한다. 어느 x에 관한 주어진 k클래스들의 확률은 $p_i = p(y = i|x), i = 1, \dots, k$ 이고, 멀티클래스 분류를 위해 pair-wise를 이용한 클래스 확률은 $r_{ij} \approx p(y = i|j) = i \text{ or } j, x$ 이다. 이때, pair-wise 클래스 확률을 다음과 같이 확장하여 정의 할 수 있다.

$$r_{ij} \approx \frac{1}{1 + e^{A\hat{f} + B}} \quad (10)$$

여기서 A와 B는 트레이닝 데이터에서 negative logarithmic likelihood를 최소로 하는 값이며, \hat{f} 는 결정값이다¹⁷⁾.

5. Hierarchical classification

본 실험에서는 클래스 N과 S의 QRS 콤플렉스는 클래스 V와 F에 비하여 형태학상으로 유사성이 매우 높은 반면, RR간격이 두 그룹을 구분하는데 유용할 것이라는 가정을 하였다. 이는 클래스 S의 경우 P와 Q의 부재로 인해 클래스 N과 RR간격에서 큰 차이를 보이게 되나 QRS 콤플렉스는 클래스 N과 S간에 큰 차이를 보이지 않는다¹⁸⁾. 따라서 클래스 N과 S를 하나의 그룹으로 먼저 묶고, 나머지 클래스를 다른 한 그룹으로 묶어서 먼저 분류하도록 계층적 분류 방식을 구성하였다.

Figure 7은 본 연구에서 사용한 계층적 분류에 사용된 특징점과 순서를 나타내었다. 여기서 멀티 클래스를 전형적인 분류방법으로 분류하는 것이 아니라 형태학상의 유사성을 이용하여 클래스를 묶어줌에 따라 각 단계(step) 분류 때마다 이진분류를 하게 되었고 첫 번째 단계에서는 클래스 NS와 VF로 분류를 하고, 두 번째 단계에서는 세부 클래스분류로 나뉘게 되는데, 클래스 N과 S로 분류하는 부분과 클래스 V와 F로 분류하는 부분으로 구성되어 계층적 분류는 총 3번의 분류 과정을 거쳐서 결과를 나타내게 된다.

첫 번째 단계에서는 QRS 콤플렉스 파형의 형태학적 정보를 나타내는 특징점들(HOS, HBF, HMH)중 하나와 RR간격 특징점을 이용하여 분류하도록 하였다. 두 번째 단계 중에서 클래스 N과 S를 분류시킬 때에는 QRS 콤플렉스 특징점을 제외한 RR간격 특징점을 활용하였고, 클래스 V와 F를 분류할 때에는 첫번째 단계에서 사용한 특징점과 RR간격 특징점을 이용하여 분류하였다.

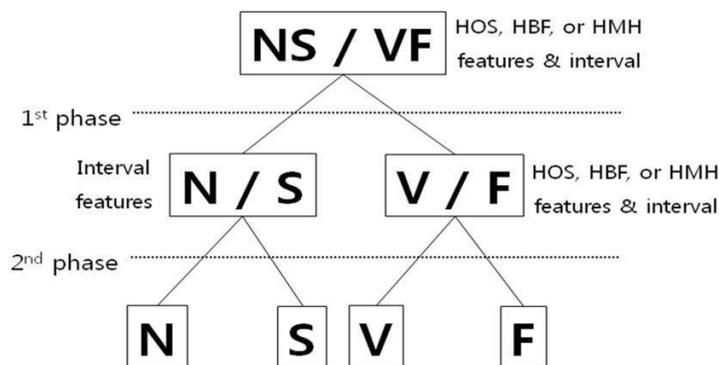


Figure 7. Hierarchical classification is that the ECG beat classes were combined into two classes based on the morphological similarity at first phase. The second phase classifies individual classes

6. 실험절차 및 평가방법

전처리 과정에서 언급했듯이, 클래스의 분포에 비례하여 MIT-BIH 부정맥 데이터베이스를 각각 22개의 기록을 포함하는 두 개의 데이터셋으로 나누게 된다²⁾. DS1은 매개 변수의 변화에 따른 분류기의 성능 평가를 통해 좋은 성능을 가진 분류기 모델을 구축하는데 사용하였다. 또한, DS2는 구축된 분류 모델을 이용하여 심전도 비트클래스 분류성능을 최종적으로 테스트 하였다. DS1을 이용하여 분류기의 성능 평가를 할 때에는 11 fold cross validation을 하여 성능 평가를 하였다. 본 연구에서 사용한 cross validation에서는 22개 기록에 포함된 50,000개 심전도 비트들을 임의로 선택하여 수행하는 beat-based cross validation을 이용하지 않고, 22개 기록을 2개 기록씩 묶어 11개 데이터군으로 분리하는 recording-based cross validation을 진행하였다. 한꺼번에 beat-based cross validation을 하게 되면 학습시에 사용되었던 데이터는 테스트시에 사용될 데이터와 동일한 기록, 즉 동일한 환자의 데이터가 될 확률이 높다. 그런 경우 예측 모델의 성능이 과대평가(overestimate)될 가능성이 높다. Figure 8은 본 연구의 실험절차를 도시하였다.

본 실험에서는 DS1의 총 데이터 개수가 50,000개가 넘어 학습을 하고 테스트를 하는데 많은 시간이 걸리기 때문에 학습시간의 감소시키기 위해 DS1에서 데이터의 20%를 각 클래스의 비율을 유지하면서 무작

위로 선택하는 stratified sampling을 하여 본 실험에 사용하였다. Table 3은 DS1과 sampled DS1의 클래스 양과 비율을 나타내었다.

Table 3. Class ratio of DS1 and sampled DS1

	N	S	V	F	Total
DS1	45,868	943	4,259	415	51,493
(Ratio, %)	(89.08)	(1.83)	(8.24)	(0.81)	(100)
Sampled DS1	9,172	194	854	86	10,306
(Ratio, %)	(89.00)	(1.88)	(8.29)	(0.83)	(100)

이진분류에서 학습을 통한 분류기의 매개변수 선정(parameter tuning)을 위해서, ROC곡선의 곡선아래면적(area under curve, AUC)을 비교하여 가장 높은 AUC를 나타내는 매개변수를 선정하였다. 또한 불균형적인 데이터의 분포로 인해 한쪽으로 치우친 결과가 예상되는 데이터에 있어서는, ROC 곡선 분석을 통해 문턱치를 찾아 적용함으로써 데이터의 불균형 분포로 인한 한쪽으로 치우친 결과를 해결할 수 있다.

계층적 분류에서는 각 분류 단계 및 과정에 따라 따로 학습하였다. DS1을 이용하여 전형적인 분류와 계층적 분류에 따른 성능평가 비교는 SVM 학습 시 선형 커널과 비선형 커널을 이용하여 나온 결과를 비교하였다. 분류기를 통해 나온 테스트데이터는 0에서 1사이의 확률값으로 나오게 되며, 전형적인 분류에서

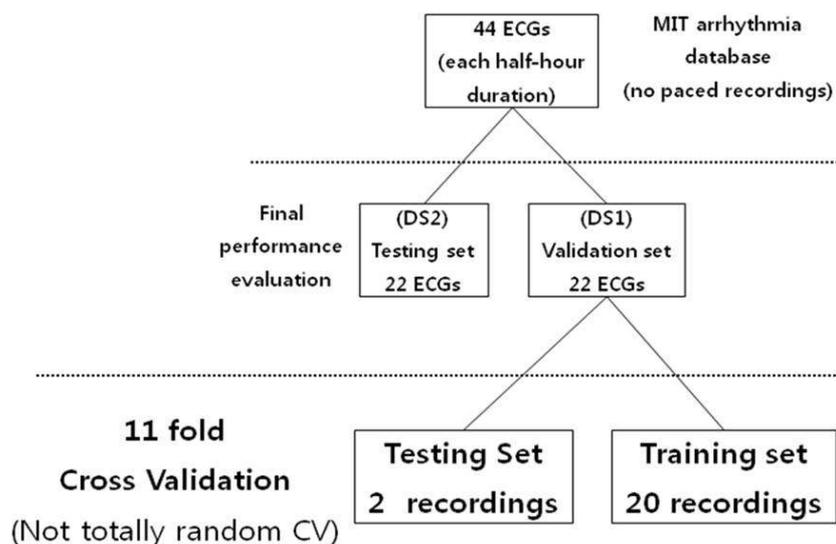


Figure 8. Division of validation set (DS1) into training and testing set for classifier evaluation using 11 fold cross validation. Final performance evaluation is performed to test testing set (DS2).

는 확률 값이 0.5이상이면 양성 데이터로 판단하고, 그 미만이면 음성 데이터로 판단하게 된다.

계층적 분류가 다중클래스 데이터에서 얼마나 성능을 향상시키는지 알아보기 위해서 전통적인 분류방법을 이용하여 성능을 비교하였다. 전통적인 분류는 모든 클래스를 학습하여 한꺼번에 클래스를 분류하는 방법이다. 전통적인 분류인 경우는 최종적으로 나오는 결정 매트릭스(decision matrix)에서 나오는 정확률(accuracy), 각 클래스의 민감도(sensitivity), 양성예측도(positive predictivity) 및 평균민감도(mean of sensitivity)를 구하여, 평균민감도가 가장 좋은 것을 선택하였다. 계층적 분류인 경우 분류방법에 따른 ROC 곡선의 AUC값을 구하고, 각 분류방법에서 AUC 값이 가장 좋은 것을 선택하였다.

전통적인 분류 방법의 경우 평균 민감도가 가장 좋은 결과의 특징점 추출 방법, 커널 그리고 매개변수 값을 이용하여 DS1의 모든 심전도 비트를 트레이닝 데이터로 이용하여 SVM 학습을 하게 되고, 여기서 나온 SVM 분류기를 DS2에 테스트함으로써 결과를 얻었다. 계층적 분류 방법인 경우, 분류 방법에서 나온 가장 좋은 AUC값에 해당하는 ROC 곡선분석을 이용하여 원하는 지점의 문턱치를 구하게 된다. 이들 문턱치 값은 분류기의 기존 문턱치 값(0.5)을 대신하여 각각 분류방법의 문턱치 값으로 사용하여 DS2를 테스트를 하게 된다. 또한, 전형적인 SVM 분류와 계층적 SVM 분류를 비교하기 위하여 결정 매트릭스로 비교하였으며, 정확률, 각 클래스의 민감도, 양성예측도 및 평균민감도도 같이 비교 하였다.

계산을 통해 나온 정확률, 각 클래스의 민감도 및 평균 민감도 중에서 평균 민감도를 주요 성능 평가 지표로 삼아 실험을 수행하였다. 정확률은 특정 클래스에 영향을 많이 받고 있으며, 각 클래스의 민감도 또한 측정할 수 있는 기준이 불확실하고, 양성예측도 특정 클래스에 많은 영향을 받고 있어 평가 기준으로 정하기 힘들다. 평균 민감도의 경우 각 클래스의 민감도의 평균을 이용하기 때문에 특정 클래스의 영향을 받지 않고 결과를 나타낼 수 있는 장점이 있어 본 연구에서는 각 클래스의 평균을 평가 기준으로 하여 분류기의 성능을 평가했다.

III. 결과

1. Conventional multiclass classification

먼저 계층적 분류를 적용하지 않고 HOS, HBF, 그리고 HMM으로 계산한 특징점들을 이용하여 전통적인 다중 클래스 분류방법으로 실험한 결과를 Table 4에 나타내었다. Table 4는 DS1을 이용하여 매개 변수의 변화에 따른 결과값을 선형 커널을 이용할 때와 비선형 커널을 이용할 때로 나누어 비교하였다.

실험 결과 특징점 추출에 따른 민감도 N은 선형 커널이 비선형 커널보다 조금 나은 성능을 보였고, 민감도 V는 비선형 커널이 선형 커널보다 조금 나은 성능을 보였다. 민감도 N만을 놓고 볼 때에는 모두 85% 이상의 성능을 보였고, 민감도 V는 대부분의 성능이 70%미만으로 결과는 만족스럽지 못하고, 민감도 S

Table 4. Result of DS1 using conventional multiclass classification by linear and RBF kernel

Kernel Parameter	Linear			Radial		
	HOS [*]	HBF [†]	HMM [‡]	HOS	HBF	HMM
N Sensitivity	99.11	91.39	98.58	98.82	87.70	98.07
S Sensitivity	0.00	1.70	0.00	0.00	0.85	0.00
V Sensitivity	58.07	27.70	59.23	65.25	70.02	64.22
F Sensitivity	0.24	0.24	0.24	1.44	0.96	1.45
Accuracy	93.43	84.26	93.04	93.71	84.08	92.96
Mean of Sensitivity	39.36	30.26	39.51	41.38	39.88	40.94
+P [§] of S	0.00	0.99	0.00	0.00	0.44	0.00
+P [§] of V	65.54	34.90	62.72	78.20	39.88	78.78

* HOS: 고차통계

† HBF: Hermite basis function

‡ HMM: 고차통계의 Hermite 모델

§ +P: 양성예측도

와 F의 경우 모든 실험에서 거의 정확하게 분류하지 못하고 있음을 보여준다. 정확률은 두 커널 모두 80% 이상의 높은 결과 값을 보여주고 있으나, 평균민감도의 경우 최대 값이 약 41%인 것으로 보아 높은 결과라 할 수 없다. 양성 예측도의 경우 클래스 S는 Table 4와 같이 클래스 S를 거의 정확하게 분류를 하지 못한 결과를 보여주고 있고, 클래스 V는 비선형 커널이 선형 커널보다 조금 나은 성능을 보였다.

평균 민감도 값이 가장 좋은 비선형 커널에서 HOS를 SVM학습하여 나온 분류기에 DS2를 테스트한 결과를 Table 5에 나타내었다.

Table 5. Result of testing DS2 using conventional multiclass classification

Parameter	Result
N Sensitivity	99.65
S Sensitivity	0.00
V Sensitivity	84.48
F Sensitivity	0.52
Accuracy	94.21
Mean of Sensitivity value	46.16
+P ^o fS	0.00
+P ^o fV	72.37

⁺P: 양성 예측도

Table 5를 보면 여전히 데이터의 불균형으로 인해 한쪽으로 치우친 결과로써, 클래스 N과 V의 성능은 80%이상으로 비교적 높은 성능을 나타내었지만, 클래스 S와 F의 경우 민감도가 1%미만으로 매우 낮은 성능을 나타내었다. 클래스 S의 양성 예측도는 정확한 분류를 하지 못하여 0%를 나타내고 있고, 클래스 V는 72.73%의 양성 예측도를 나타내고 있다. 정확률의 경우 90%의 결과를 보이고 있지만 평균 민감도의 경우 50%미만의 좋지 못한 결과를 나타내었다.

2. Hierarchical classification

(1) ROC curve analysis

특징점 추출 방법을 계층적 SVM분류로 실험하였을 때, 첫 번째 단계에서는 RBF 커널에서 HOS방법이 가장 좋은 성능을 나타내었고 Figure 9는 이때의 결과를 ROC 곡선으로 나타내었다. Figure 9의 ROC 곡선은 sampled DS1(10,306개)에서 나온 확률 결정값을 이용하여 0~1까지 문턱치를 변화시키면서 나온 특이도와 민감도의 좌표를 이용하여 ROC 곡선을 도시하였다. ACU값은 0.909이고, 각각의 화살표에 해당하는 지점들은 민감도 90% 지점(Sen_90), 특이도(specificity)

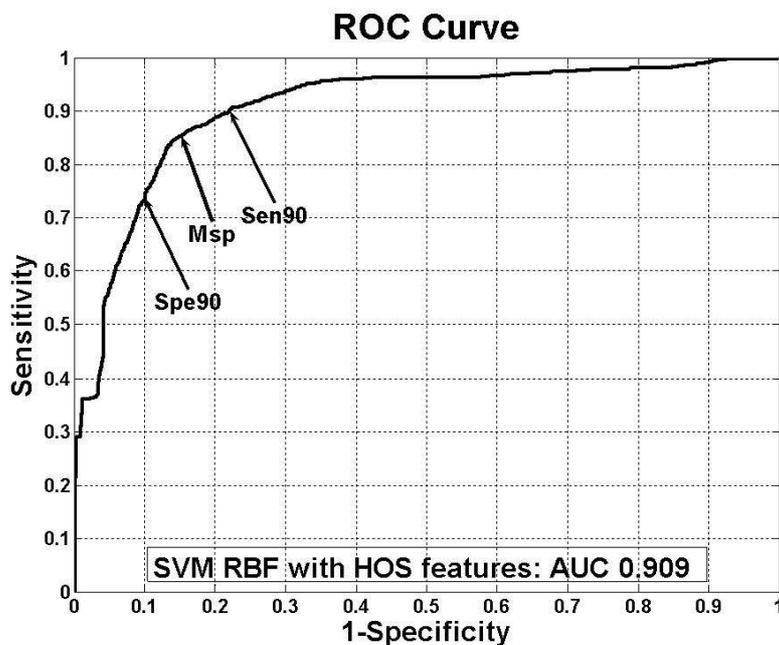


Figure 9. ROC curve of the best classifier. The arrows indicate sensitivity 90% point (Sen_90), specificity 90% point (Spe_90), and minimum distance point (Msp) between ROC curve and (1,0) point

90% 지점(Spe_90) 및 민감도 100%와 특이도 100% 지점에서 가장 가까운 ROC 곡선상의 지점(Msp)을 나타낸다. Table 6은 각 지점에 대한 민감도, 특이도, 정확률 및 문턱치를 나타내었다.

Table 6. Sensitivity, specificity, accuracy and threshold of each arrow points in NS vs. VF ROC curve

Parameter	Sen_90 [*]	Msp [†]	Spe_90 [‡]
Sensitivity	90	84	74
Specificity	78	86	90
Accuracy	79	86	88
Threshold	0.955	0.925	0.835

^{*} Sen_90: 민감도 90% 지점

[†] Msp: 민감도 100%와 특이도 100% 지점

[‡] Spe_90: 특이도 90% 지점

Table 6에서 민감도는 클래스 VF를 나타내고 특이도는 클래스 NS를 나타낸 것이고, 민감도와 특이도의 변화에 따라 클래스 NS와 VF가 변화가 있음을 것을 확인할 수 있다. Table 3과 6을 볼 때, 클래스 NS의 데이터가 클래스 VF의 데이터에 비해 많기 때문에, 정확률은 민감도 보다는 특이도에 많은 영향을 받는다. 따라서, 본 실험에서는 정확률이 가장 높은 Spe_90 지점의 문턱치를 사용하여 다음 실험을 진행 하였다. HOS, HBF 그리고 HMH에서 나온 특징점을 계층적 분류방법으로 분류하여 나온 AUC 값들 중 가장 높은 값을 갖는 매개변수를 선택하였다. 선택된 매개변수를 이용하여 ROC곡선 분석을 통해 Spe_90지점의 문턱치를 구하였다. Table 7은 각각의 추출방법에 따른 AUC값, 커널을 나타내었다.

Table 7. AUC value and kernel by feature extraction method

Phase	Class	Parameter	HOS [*]	HBF [†]	HMH [‡]
1 st step	NS vs. VF	AUC	0.909	0.893	0.899
		Kernel	Radial	Radial	Radial
2 nd step	N vs. S	AUC		0.897 [§]	
		Kernel	Linear	Linear	Linear
	V vs. F	AUC	0.770	0.656	0.815
		Kernel	Linear	Linear	Linear

^{*} HOS: 고차통계

[†] HBF: Hermite basis function

[‡] HMH: 고차통계의 Hermite 모델

[§] RR 간격 특징점만을 이용한 결과

Table 7을 보면 첫 번째 단계의 경우 HOS가 가장 높은 AUC 값을 가지고 있고, 두 번째 단계에서 클래스

V와 F 분류에서는 HMH가 가장 높은 AUC 값을 갖고 있었다.

(2) Feature extraction method

Table 7에서 특징점 추출 방법을 계층적 분류의 분류과정에 따른 각각의 문턱치 값들을 적용하여 DS2를 테스트한 결과를 Table 8에 나타내었으며, Table 5의 전통적인 다중클래스 분류 결과를 계층적 분류 결과와 비교하기 위해 좋은 성능을 나타낸 다중클래스 분류 결과를 포함 하였다.

Table 8. Classification performance of hierarchical classification and multiclass classification on DS2

Parameter	DS2_multi [*]	DS2_HOS [†]	DS2_HBF [‡]	DS2_HMH [§]
N Sensitivity	99.65	81.23	86.25	82.98
S Sensitivity	0.00	57.65	82.63	75.23
V Sensitivity	84.48	83.11	80.88	84.17
F Sensitivity	0.52	79.90	54.90	82.47
Accuracy	94.21	80.47	85.56	82.80
Mean of Sensitivity	46.16	75.47	76.16	81.21
+P of S	0.00	21.50	24.78	24.12
+P of V	72.37	46.11	89.09	84.69

^{*} multi: 전형적인 다중클래스

[†] HOS: 고차통계

[‡] HBF: Hermite basis function

[§] HMH: 고차통계의 Hermite 모델

^{||} +P: 양성예측도

Table 8을 보면 전형적인 다중분류를 통한 DS2_multi의 경우, 정확률 및 클래스 N과 V의 민감도의 경우 계층적 분류보다 높은 결과를 나타내고 있으나, 클래스 S와 F의 민감도 경우 1% 미만의 결과를 나타내었다. 정확률은 90%이상의 결과를 나타내고 있지만, 평균 민감도는 46%를 나타내고 있다. 클래스 S에 대한 양성 예측도는 0%의 결과를 나타내어 계층적 분류에 비해 낮은 결과를 나타내고 있으나 클래스 V에 대한 양성 예측도는 70%이상의 결과를 보이고 있다. DS2_multi와 HOS를 비교해 보면 정확률 및 클래스 N과 V의 민감도의 감소는 있었으나 클래스 S와 F의 경우 전통적인 다중클래스 분류 비해 많은 S와 F를 찾고 있는 것을 볼 수 있다. HOS의 경우 클래스 S민감도와 양성 예측도를 제외한 결과는 70%이상의 결과를 보이고 있어 DS2_multi에 비해 비교적 높은 결과를 나타낸다.

HBF에 있어서는, 클래스 N과 S민감도는 80%이상의 결과를 보이고 HOS에 비해 잘 찾아내고 있으나, 클레

스 V와 F의 경우 V는 소폭 감소하였지만 80%이상의 결과를 유지하고 있으나 F는 감소가 많이 생겨 60%미만의 결과를 나타내고 있다. 정확률은 DS2_multi에 비해 감소하였으나, 80%이상의 좋은 결과를 나타내고 있고 평균 민감도는 75%의 결과를 보이고 있다.

HMH를 HOS 및 HBF와 비교해 보면, 정확률 및 클래스 N과 S 민감도의 경우 HOS에 비해 나은 결과를 내었지만, HBF에 비해 나쁜 결과를 내었고, V와 F 민감도의 경우는 HOS과 HBF에 비해 나은 결과를 보였다. HMH의 클래스 S와 V의 양성 예측도 경우 DS2_multi와 HOS에 비해 나은 결과를 높은 결과를 나타내고 있지만 HBF보다는 낮은 결과를 나타내고 있다. 여기서, 중요 타겟으로 잡은 평균 민감도의 경우 HMH가 나머지 3개와 비교하여 가장 좋은 결과를 내었고, 결과 역시 80%를 넘게 나와 좋은 결과를 나타내었다.

Table 9은 평균 민감도가 가장 좋은 HMH로 DS2를 테스트한 결과를 결정 매트릭스로 나타내었다.

Table 9. Result of decision matrix on DS2 of Hermite model of higher order statistics

		Actual class			
		N	S	V	F
Predicted Class	n	36,728	221	42	16
	s	4,079	1,382	249	18
	v	319	171	2,711	34
	f	3,133	63	219	320
Sensitivity(%)		82.98	75.23	84.17	82.47
+P*			24.12	84.69	

*P: 양성예측도

IV. 토의 및 결론

본 연구에서는 QRS 콤플렉스에서 P, QRS, T 파의 검출 없이 형태학적인 특징점을 추출하여 심전도 비트클래스를 분류 해보고자 하였다. 특징점 추출방법들에 있어서, HOS은 통계적인 방법을 이용하여 QRS 콤플렉스를 재해석함으로써 각각의 클래스간을 잘 구분하여 분류할 수는 있으나, 많은 특징점 때문에 학습 과정에서 오버피팅되어 테스트과정에서 나쁜 성능을 나타낼 수 있다. HBF은 QRS 콤플렉스를 Hermite 다항식의 계수들의 조합으로 파형들을 잘 형상화할 수 있고 HOS보다 적은 특징점을 사용하여 클래스를 잘 구분할 수 있지만, 파형에 남아있는 잡음에 영향을 받

을 수 있어 분류성능에 영향을 미칠 수도 있다. HMH의 경우 HOS 특징점 추출을 통한 잡음에 민감하게 반응하지 않는 특징점을 추출할 수 있고, 이들 특징점은 HBF를 통해 HOS의 특징을 유지하면서 특징점의 개수를 줄일 수 있다. 이러한 HMH는 HOS와 HBF의 분류성능보다 나은 성능을 보이고 있다.

심전도 비트 데이터 다중분류를 목적으로 계층적 분류기법을 사용하였으며, 심전도 파형이 클래스 N과 S 그리고 클래스 V와 F가 형태학상으로 유사성이 있다고 가정한 영역지식을 이용하여 계층 구조를 구성하였다. 실험결과를 보면 계층적 분류를 최종적으로 수행한 결과가 전통적인 다중클래스 분류보다 나은 성능을 보이고 있다. 이를 통해 다중클래스가 각 단계에 따른 이진분류방법으로 나뉘면서 ROC곡선 분석이 가능케 하였다. 이를 통해 문턱치 조절에 따른 최적의 지점을 찾게 되어 불균형 데이터 분포 문제를 해결하였다.

측정 방법에 관해서는 모든 데이터를 사용하여 분류기의 성능을 평가하는 것보다 샘플링을 통해 분류기 성능을 평가한 것이 계산 속도 면에서 많은 이득을 보이고 있다. 동일한 조건하에 샘플링을 했기 때문에 어떠한 바이어스가 존재할 수가 없고 샘플링 된 데이터들은 특별한 다른 점들이 존재하지 않는다. 클래스 데이터의 대부분을 클래스 N이 차지하고 있고, 다음으로 클래스 V, S, F 순으로 차지하고 있다. 클래스 N과 V가 차지하고 있는 비율이 97%이상을 차지하고 있어 클래스 S와 F를 많이 찾더라도 두 개를 합친 비율이 전체에 1%미만이므로 정확률에는 아무런 영향을 끼치지 못하고 있다. 클래스 N과 V가 얼마나 잘 분리되느냐에 따라 정확률이 결정 되며, 특히 클래스 N에 정확률이 가장 많은 영향을 받고 있다. 양성 예측도의 경우 클래스 S와 V의 민감도와는 상관없이 예측 분류된 결과를 이용하여 클래스의 양성 예측도를 나타낸 것으로 이 또한 클래스 N이 얼마나 잘못 예측되었느냐가 양성 예측도 결과에 많은 영향을 주고 있다. 이에 반해, 평균 민감도는 특정 클래스에 의해 영향을 받지 않는다는 것을 이용하여 DS2를 테스트한 결과를 비교해 보니, HMH가 HOS와 HBF에 비해 평균 민감도에서 높은 결과를 나타내었다. DS2로 테스트한 결과가 DS1으로 학습한 결과보다 좋은 성능을 나타낸 이유는 DS1과 DS2는 무작위로 추출한 데이터 셋이어서 서로 바이어스가 존재하지 않을 것이며, DS1에서 단순히 트레이닝 성능만을 측정된 것

이 아니라, cross validation한 결과이기 때문이고, 또한 두 결과의 성능 차이가 확연하게 나타나지 않고 있다.

Osowski[3]의 연구 결과와 비교해서 살펴보면 HOS와 HBF기법에서 학습오차가 각각 5.19%(HOS)와 3.02%(HBF)로써, 본 연구의 Table 4에서 보이고 있는 93.43%(HOS)와 84.26%(HBF)의 정확률에 비해 우수한 보이고는 있다. 그러나 Osowski의 연구는 실험 방법으로 제시한 데이터의 경우 테스트데이터와 트레이닝데이터의 구분은 하고 있으나 기록에 따른 비트 분류가 아닌 비트 별로 데이터 나누어서 분류를 하였기 때문에 트레이닝데이터와 테스트데이터에 같은 기록의 데이터가 포함되어 있을 가능성이 높고 본 연구와 비교하기 힘들다. 이는 MIT-BIH 부정맥 데이터베이스[11]를 참고해보면 특정 클래스 비트는 특정 기록에 치우쳐져 있음을 확인할 수 있고, Osowski가 제시한 클래스 비트 개수도 또한 맞지 않음을 확인할 수 있다. 이 때문에 Osowski의 테스트데이터와 트레이닝데이터의 학습오차는 낮게 나올 수 밖에 없는 것이지만 본 연구는 beat-based cross validation을 이용하지 않고, recording-based cross validation을 진행하여 최적의 분류기를 찾아내고 찾아낸 분류기에 독립된 테스트데이터를 평가하였기 때문에 Osowski가 제시한 성능보다는 낮게 나온다. 본 연구와 Osowski의 연구를 직접적으로 성능을 비교하기에는 테스트데이터와 트레이닝데이터의 데이터 분포가 틀리기 때문에 곤란하다.

본 연구에서는 특징점 추출 방법에서 나온 모든 특징점을 이용하여 클래스를 분류하였으나 앞으로의 연구에서는 특징점 선택(feature selection)방법을 이용하여 불필요한 여분의 특징점을 제거하고 가장 최적화된 특징점만을 선택하여, 성능향상을 기대할 수 있고, 위험인자들의 시각화(visualization of risk factors)를 통해 다른 방향으로 성능 향상을 기대할 수 있다고 보기 때문에, 임상외에게 도움을 줄 수 있을 것으로 판단이 되므로, 향후에 이와 관련된 다양한 연구를 진행할 수 있을 것이다.

참고문헌

1. Hu YH, Palreddy S, Tompkins W. A patient adaptable ECG beat classifier using a mixture of experts approach. *IEEE Trans Biomed Eng* 2007;44:891-900.
2. Chazal P, O'Dwyer M, Reilly R. Automatic Classification of Heartbeats Using ECG Morphology and Heartbeat Interval Features. *IEEE Trans Biomed Eng* 2004;51:1196-1206.
3. Osowski S, Hoai LT, arkiewicz T. Support vector machine-based expert system for reliable heartbeat recognition. *IEEE Trans Biomed Eng* 2004;51:582-89.
4. Nikias C, Petropulu A. Higher Order Spectral Analysis. Prentice-Hall, 1993.
5. Lagerholm M, Peterson C, Braccini G, Edenbrandt L, and Sornmo L. Clustering ECG complexes using Hermite functions and self-organizing maps. *IEEE Trans Biomed Eng* 2000;47:838-47.
6. Vapnik V. *Statistical Learning Theory*. Wiley, 1998
7. Kreel U. *Advances in Kernel Methods: Support Vector Learnings*. MIT Press, 1999;255-68.
8. Casasent D, Wang YC. A hierarchical classifier using new support vector machines for automatic target recognition. *Neural networks* 2005;18:541-48.
9. Elkan C. The foundation of cost-sensitive learning. *Proceeding of the 17th International Joint Conference on Artificial Intelligence*; 2001 Aug 4-10; Seattle, USA.
10. Yang H, Carlin D. ROC surface: a generalization of ROC curve analysis. *Journal of biopharmaceutical statistics* 2000;10:183-96.
11. Mark R, Mody G. MIT-BIH Arrhythmia Database. 1997 Online available at: <http://www.physionet.org/>.
12. American National Standard. Testing and reporting performance results of cardiac. ANSI/AAMI EC57 : 1998/(R)2003.
13. Sansone G. *Orthogonal Functions*. Dover, 1991.
14. Golub G, Van Loan C. *Matrix Computations*. Johns Hopkins University, 1991.
15. Kim Y, Kang S, Park I, Noh G. Population Pharmacokinetic and Pharmacodynamic Models of Propofol in Healthy Volunteers using NONMEM and Machine Learning Methods. *Journal of Korean Society of Medical Informatics* 2008;14 Suppl 2:147-159.
16. Cho B, Lee J, Chee Y, Kim K, Kim I, Kim S. Prediction of Diabetic Nephropathy from Diabetes Database Using Feature Selection Methods and SVM Learning. *Korea Society of Medical & Biological Eng* 2007;28:255-62.
17. Lin HT, Lin CJ, Weng RC. A note on Platt's probabilistic outputs for support vector machines. *Taiwan Univ.*, 2003.