

LDA와 SVM 기반의 심실세동 검출에  
관한 연구

연세대학교 대학원  
의료공학협동과정  
의 공 학 전 공  
송 미 혜

# LDA와 SVM 기반의 심실세동 검출에 관한 연구

지도 이 경 중 교수

이 논문을 석사 학위논문으로 제출함

2004년 12월 일

연세대학교 대학원

의료공학협동과정

의 공 학 전공

송 미 혜

# 송미혜의 석사 학위논문을 인준함

심사위원 \_\_\_\_\_ 인

심사위원 \_\_\_\_\_ 인

심사위원 \_\_\_\_\_ 인

연세대학교 대학원

2004년 12월 일

## 감사의 글

본 논문이 있기까지 지도와 격려를 아끼지 않으신 이경중 교수님께 진심으로 머리 숙여 감사드립니다. 논문의 부족한 점을 지적해주시고 지도해주신 윤영로 교수님과 김경환 교수님께 감사드립니다. 저의 대학과 대학원 생활동안 많은 가르침을 주신 윤형로 교수님, 이윤선 교수님, 김동윤 교수님, 김영호 교수님, 신태민 교수님, 김법민 교수님, 김한성 교수님께 진심으로 감사드립니다. 또한 학교 생활에서 불편함이 없도록 배려해주신 안종수 선생님께 감사드립니다.

대학원 생활동안 옆에서 많은 격려와 도움을 아끼지 않으신 이전 선배님과 처음 대학원에 들어와 아무것도 몰랐던 저에게 기초부터 하나하나 직접 가르쳐주고, 많은 충고와 도움을 아끼지 않으신 박호동 선배님께 진심으로 감사드립니다. 대학원에 함께 입학하여 2년이란 시간동안 함께 고생하며 공부한 지종철, 안영호, 임예택 선배님께 감사드립니다. 대학원 생활동안 저에게 많은 도움을 주신 최호선 선생님과 박광리, 지정호, 이승형, 이재영, 서현순 선배님과 짧은 시간이나마 함께 공부한 장봉렬, 이효기, 김병균에게 감사합니다. 그 외 모든 의공학과 대학원 선배님, 후배들과 항상 마음속으로 힘이 되어 준 98학번 동기들에게 감사드립니다. 친구들과보다 더 가까이에서 정신적으로 많은 위로와 힘이 되어준 조한희, 사혜정 언니와 김정주에게 감사드리며 대학원 생활동안 항상 웃음을 잃지 않게 도와준 박지은, 박지현에게 감사의 말을 전합니다. 언제나 웃음으로 저에게 기쁨을 주고, 저의 부족한 부분을 항상 넓은 마음으로 이해해주며 대학원 생활동안 가장 큰 정신적 위로와 도움을 아끼지 않으신 조성필 선배님께 진심으로 감사드립니다.

마지막으로, 제가 하고자 하는 일을 끝까지 믿음으로서 지켜봐주고, 하나뿐인 딸과 동생을 위해 끝없는 사랑과 열성으로 도움을 아끼지 않으시는 세상에서 가장 사랑하는 아버지, 어머니, 하나뿐인 자랑스러운 오빠에게 온 마음을 다해 진심으로 감사드립니다. 저를 아껴주신 모든 분들께 이 책을 바칩니다.

2004.12

송 미 혜 올림

# 차 례

그림 차례 .....	i
표 차례 .....	iv
국문 요약 .....	v
제 1 장 서 론 .....	1
제 2 장 본 론 .....	3
2.1 심실세동 .....	4
2.2 특징 추출 .....	6
2.2.1 웨이브렛을 이용한 잡음제거 .....	7
2.2.2 QRS 콤플렉스 검출 및 구간설정 .....	8
2.2.3 특징점 추출 .....	13
2.2.3.1 RR 감격을 이용한 특징점 추출 .....	13
2.2.3.2 웨이브렛 변환을 이용한 특징점 추출 .....	14
2.3 선형판별분석기법(linear discriminant analysis) .....	15
2.4 Support Vector Machine(SVM) .....	18
제 3 장 실험 결과 및 고찰 .....	23
3.1 실험 방법 .....	23
3.2 QRS 콤플렉스 검출율 .....	24
3.3 파라미터에 따른 SVM의 성능비교 .....	25
3.4 심실세동 검출 알고리즘 성능 평가 .....	26
3.4.1 신경망(neural network) .....	27
3.4.2 퍼지 추론 시스템(fuzzy inference system) .....	28

3.4.3 SVM 다원 분류기 .....	29
3.4.4 성능 평가 .....	30
3.5 제안한 알고리즘을 이용한 심실세동 검출 .....	32
제 4 장 결 론 .....	34
참고 문헌 .....	35
ABSTRACT .....	39

## 그림 차례

그림 2.1 심실세동 검출 알고리즘에 대한 전체 블록도 .....	3
그림 2.2 정상심전도와 부정맥 심전도의 비교 .....	4
(a) 정상심전도(NSR)	
(b) 심실빈맥	
(c) 심실세동	
그림 2.3 심실세동 파형 .....	5
(a) 거친 심실세동 파형	
(b) 약해진 심실세동 파형	
그림 2.4 특징 추출 과정에 대한 블록도 .....	6
그림 2.5 Dyadic 트리 구조 .....	6
그림 2.6 웨이브렛을 이용한 잡음 제거 .....	7
(a) 잡음이 포함된 ECG 신호	
(b) 잡음이 제거된 ECG 신호	
(c) 필터링 전 스펙트럼	
(d) 필터링 후 스펙트럼	
그림 2.7 심전도 신호와 웨이브렛 변환을 이용한 합산신호 .....	8
(a) 정상과 심실세동 심전도 신호	
(b) 합산신호	
그림 2.8 QRS 콤플렉스가 존재할 예상구간 설정 .....	9
(a) 합산신호와 문턱치	
(b) 차분신호와 예상구간 설정	
그림 2.9 웨이브렛을 이용한 QRS 콤플렉스 검출 .....	10
(a) 정상과 심실세동 심전도 신호	

(b) QRS 콤플렉스	
그림 2.10 심전도신호(레코드106)의 웨이블릿을 이용한 QRS 콤플렉스 검출 .....	11
(a) 심전도 신호	
(b) 합산신호	
(c) QRS 콤플렉스 검출	
그림 2.11 분석구간 .....	12
그림 2.12 RR 간격을 이용한 특징점 .....	13
(a) 평균 RR 간격과 이전 RR 간격의 비율	
(b) 평균 RR 간격과 이후 RR 간격의 비율	
그림 2.13 웨이블릿 계수를 이용한 특징점 .....	14
(a) D4 신호	
(b) D5 신호	
(c) D6 신호	
(d) D7 신호	
그림 2.14 LDA를 이용한 특징 추출 및 차원감소(x: 정상심전도, o: 비 정상심전도) .....	15
(a) LDA 방향	
(b) 추출된 특징과 분포	
그림 2.15 LDA를 적용한 새로운 특징점 .....	17
(a) 특징점1	
(b) 특징점2	
(c) 특징점3	
(4) 특징점4	
그림 2.16 SVM 다원 분류기 블록도 .....	18



그림 2.17 SVM의 최적 분리면과 support vector( $x$ : 정상심전도, $o$ : 비 정상심전도) .....	19
그림 3.1 성능 평가에 대한 전체 블록도 .....	26
그림 3.2 심실세동 검출 블록도 .....	32

## 표 차례

표 2.1 웨이브렛 변환 레벨에 의한 주파수 분할 (단위: Hz) .....	6
표 2.2 커널 함수와 SVM 분류기의 종류 .....	22
표 3.1 QRS 콤플렉스 검출 결과 .....	24
표 3.2 C값과 $\alpha$ 값에 따른 분류 성능 비교 .....	25
표 3.3 신경망을 이용한 정상심전도 검출 결과 .....	27
표 3.4 퍼지 추론 시스템을 이용한 정상심전도 검출 결과 .....	28
표 3.5 제안한 알고리즘을 이용한 정상심전도 검출 결과 .....	29
표 3.6 세 가지 알고리즘 검출 결과 비교 .....	30
표 3.7 신경망과 제안한 알고리즘의 학습 데이터 수 비교 .....	31
표 3.8 제안한 알고리즘을 이용한 심실세동 및 다양한 부정맥의 검출 결과 .....	33

## 국문 요약

### LDA와 SVM 기반의 심실세동 검출에 관한 연구

본 논문은 RR 간격과 웨이브렛 변환을 통해 추출한 특징들을 선형판별분석법(linear discriminant analysis: LDA)과 Support Vector Machine(SVM) 다원 분류기에 적용하여 향상된 검출 성능을 갖는 심실세동 검출 알고리즘을 제안하였다. 제안한 알고리즘은 웨이브렛을 이용한 잡음제거, QRS 콤플렉스 검출 및 구간설정, 특징점 추출, LDA를 이용한 차원감소 그리고 SVM 다원 분류기로 구성되어 있다. RR 간격과 웨이브렛 변환을 통해 획득한 4번째, 5번째, 6번째, 7번째 상세계수(D4, D5, D6, D7)의 특징 포인트에서의 값을 특징점으로 추출하였다. 추출된 특징점을 LDA에 적용하여 최적의 특징벡터를 추출하고 이를 SVM 다원 분류기에 적용하여 심실 세동을 검출하였다. 제안한 알고리즘의 성능을 평가하기 위해 MIT-BIH 부정맥 데이터베이스, Creighton University Ventricular Tachyarrhythmia 데이터베이스, MIT-BIH Malignant Ventricular Arrhythmia 데이터베이스의 심전도 데이터를 사용하였고, 차원 감소의 대표적인 방법인 주성분분석(principal component analysis: PCA)과 심실세동 및 부정맥 검출 알고리즘에서 널리 사용되어지고 있는 신경망과 퍼지 추론시스템을 이용하여 검출 성능을 비교, 평가하였다. 본 논문에서 제안한 알고리즘의 성능 평가를 통해 알고리즘의 검출 성능의 우수성과 효율성을 확인하였다. LDA를 적용하여 추출한 최적의 입력은 기존의 많은 양의 입력을 사용한 경우보다 우수한 검출 성능을 보였으며, 적은 수의 입력은 SVM 다원 분류기를 학습 시키는 과정에서도 학습속도를 단축시키는 효과가 있었다. 또한 SVM 다원 분류기는 기존의 다른 분류 분류기에 비해 우수한 분류 성능을 보였으며, 적은 양의 학습 데이터만으로도 높은 검출 성능을 보이는 것을 확인하였다. 제안된 알고리즘의 이용한 심실세동 검출율은 Sensitivity(%)는 99.75( $p < 0.05$ ), Specificity(%)는 99.99( $p < 0.05$ ), Accuracy(%)는 99.88( $p < 0.05$ )로

우수한 성능을 보였다.

---

핵심이 되는 말 : 심실세동, 웨이브렛 변환, 선형판별분석법(LDA), Support Vector Machine(SVM)

# 제 1 장 서 론

본 논문은 RR 간격과 웨이브렛 변환을 통해 추출한 특징들을 선형판별분석법(linear discriminant analysis: LDA)과 Support Vector Machine(SVM) 다원 분류기에 적용하여 향상된 검출 성능을 갖는 심실세동 검출 알고리즘을 제안하였다.

심전도 신호는 일반적으로 동방결절(Sinus Node)에서 발생하는 전기적인 신호에 의해 심장이 규칙적으로 수축과 이완 할 때 체표면에서 기록한 신호이다. 심전도 신호 중 정상적인 범위내의 빈도를 벗어나거나 정상빈도를 갖더라도 패턴상의 이상이 있을 때를 부정맥(Arrhythmia)이라고 한다[1]. 부정맥은 여러 가지 원인에 의하여 심장 내 전기적 신호의 전달 경로나 그 주위 심장 부위에 이상이 생겨 발생한다. 부정맥은 크게 심박동수가 심하게 느려지는 서맥성 부정맥(Bradyarrhythmia)과 비정상적으로 빨라지면서 발생하는 빈맥성 부정맥(Tachyarrhythmia), 그리고 심박동이 예정보다 한 박자 빨리 나오는 조기 수축(Premature Contraction)으로 분류되고, 같은 부정맥이라도 다시 여러 형태의 부정맥으로 구분된다[2].

심실세동(ventricular fibrillation: VF)은 심장의 무질서한 전기적 활동으로 인해 심근 수축이 동시에 이뤄지지 않게 되어 발생한다. 이로 인해 심장의 혈액 공급이 중단되어, 신체 기관 및 뇌에 산소공급이 중단되게 된다. 뇌에 산소공급이 중단되면 뇌는 손상을 입게 되고, 결국은 몇 분 안에 뇌의 기능이 정지하여 급성심장사(sudden cardiac death: SCD)에 이르게 된다[3]. 즉, 심실세동은 환자의 생명을 위협하는 가장 치명적인 부정맥으로 즉각적인 치료를 하지 않을 경우 환자는 바로 사망하게 된다. 따라서 심장 질환의 증가와 더불어 급성 심장사에 이를 수 있는 심실성 부정맥 중의 하나인 심실세동 검출은 매우 중요하다.

심실세동 검출을 위한 기존의 방법으로는 비선형 분석법[4]-[6], 시주파수 분석법[7], 신경망[8], 퍼지 추론[9]을 이용한 방법 등이 있으며, 최근에는 검출의 정확성을 높이기 위한 알고리즘이 많이 연구되어 지고 있다. 비선형 분석법[4]-[6]과 시주파수 분석법[7]은 시간 영역 알고리즘에 비해 연산이 복잡하지만, 심실세동과 같

은 불규칙한 리듬 검출에 매우 안정적이고 효과적이다. 신경망[8]은 패턴인식, 식별, 분류와 같은 다양한 분야에서 복잡한 함수를 실행 할 수 있도록 훈련되어 실용화되고 있는 시스템으로서 검출 성능은 좋으나 많은 양의 심전도를 학습시켜야 하고 회로망의 구성이 복잡하다는 단점이 있다. 퍼지 추론[9]은 비교적 연산이 간단하고 신경망 같은 초기학습을 요구하지 않은 장점이 있으나, 멤버십 함수를 선정할 때 전문가의 의견과 반복적인 실험을 통해 범위를 정해야 하기 때문에 주관성을 배제할 수 없다는 단점이 있다[10].

이에 본 논문에서는 추출된 특징간의 최적의 구별이 가능케 하는 벡터를 추출하는 방법인 선형판별분석법[11]과 적은 양의 학습 데이터만으로 높은 분류 성능을 보이는 Support Vector Machine(SVM)[12]을 이용한 심실세동 검출 알고리즘을 제안하였다. 제안한 알고리즘은 적은 수의 최적화된 입력특징을 사용하므로 기존의 모든 특징을 사용한 경우에 비해 심실세동 검출 성능이 우수하며, 학습과정에서 소비되는 시간도 크게 단축하는 장점이 있다. 또한 최근 우수한 분류 성능으로 각광받기 시작한 SVM은 신경망의 단점인 많은 양의 학습 데이터를 요구하지 않고, 적은 수의 학습 데이터만으로도 우수한 검출 성능을 보인다. 본 논문에서 제안한 알고리즘의 성능 평가를 위해 사용한 심전도 데이터는 MIT-BIH 부정맥 데이터베이스, Creighton University Ventricular Tachyarrhythmia 데이터베이스, MIT-BIH Malignant Ventricular Arrhythmia 데이터베이스를 사용하였다. 또한 제안한 알고리즘의 성능 평가를 위해 선형판별분석법과 함께 차원 축소를 위해 널리 사용되어지고 있는 주성분분석(principle component analysis: PCA)[11]과 심실세동 검출에서 주로 이용되고 있는 분류기로서 신경망과 퍼지 추론을 사용하였다.

## 제 2 장 본 론

심실세동 검출 알고리즘은 크게 특징 추출, 선형관별분석법(LDA)을 이용한 차원 감소, Support Vector Machine(SVM) 다원 분류기로 나눌 수가 있다. 그림 2.1은 심실세동 검출 알고리즘에 대한 전체 블록도이다.

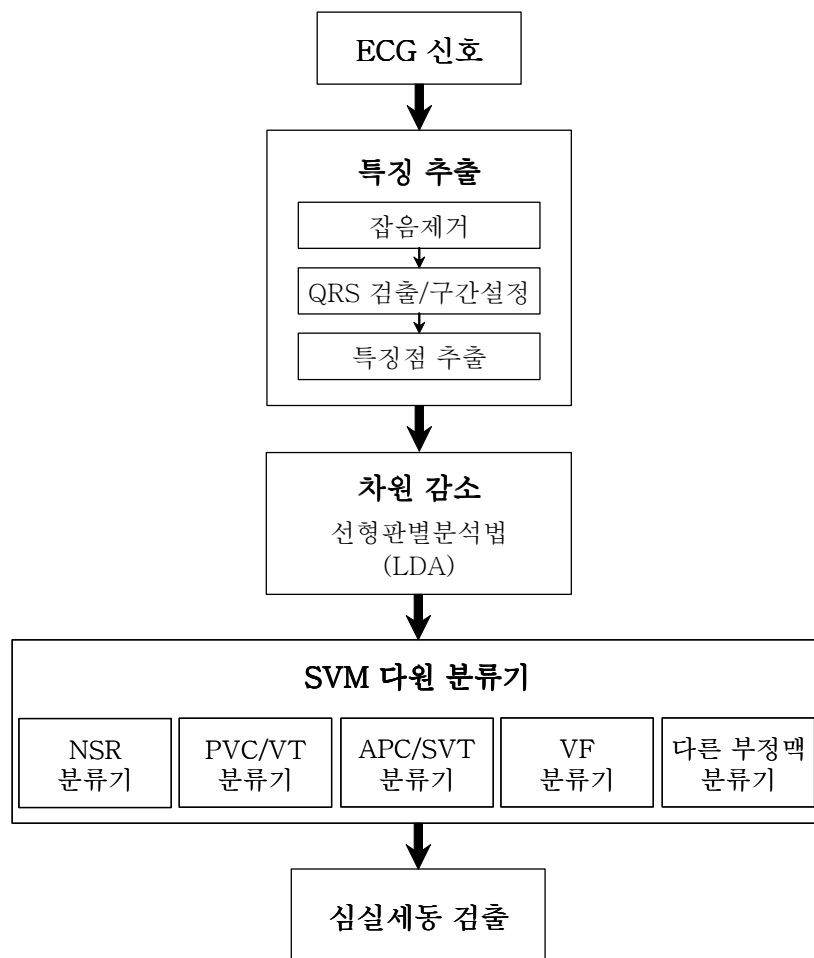


그림 2.1 심실세동 검출 알고리즘에 대한 전체 블록도

## 2.1 심실세동

심실세동은 흥분파가 심실의 한 부분을 자극한 후 다른 부분을 자극하고, 또 다른 부분을 자극하면서 지속적으로 이어져 결국은 처음 부분을 다시 자극하는 반복적인 양상을 지속적으로 나타내는 상황을 말한다. 이때 심실 내에서 많은 작은 부분들이 동시에 수축하는 반면, 다른 부분들은 이완되는 상황이 나타난다, 따라서 심장의 원활한 수축력을 위해 필요한 심실의 동시 수축이 일어날 수 없게 된다. 그러므로 많은 흥분성 신호들이 심실을 흐르고 있긴 하지만 심실은 수축하지도 이완되지도 못한 채로 적당히 부분적인 수축 상태를 유지한다. 따라서 혈액을 거의 배출하지 못하거나 전혀 배출하지 못하기도 한다. 그러므로 심실세동이 시작되고 4 내지 5초가 지나면 의식을 잃게 되고 몇 분이 지나면 조직의 불가역적 손상으로 인해 사망하게 된다[13]. 그림 2.2은 정상심전도(normal sinus rhythms: NSR), 심실 빈맥(ventricular tachycardia: VT), 심실세동시 심전도를 나타낸 것이다. 그림 2.2(c)에서 보는 바와 같이 심실세동의 심전도는 어떠한 규칙성도 없는 기이한 양상으로 그려지고, QRS군이나 T파를 감별할 수 없이 불규칙한 진동파로 나타난다.

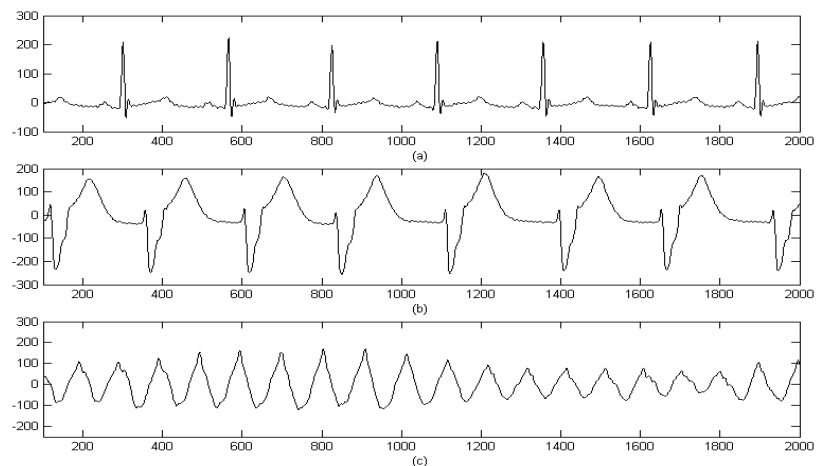


그림 2.2 정상심전도와 부정맥 심전도의 비교

(a) 정상심전도(NSR)

(b) 심실빈맥

(c) 심실세동



심실세동을 유발시킬 수 있는 요인으로는 심장에 갑자기 전기 충격이 가해질 경우, 흥분전도계와 같은 심근에 허혈이 있을 경우에 발생된다. 세동파의 크기에 따라서 진폭이 0.1mV 이상인 경우는 거친(coarse) 심실세동, 0.01mV 이하인 경우는 약해진(fine) 심실세동으로 구분된다[14]. 그림 2.3은 거친 심실세동 파형과 약해진 심실세동 파형을 나타낸 것이다.

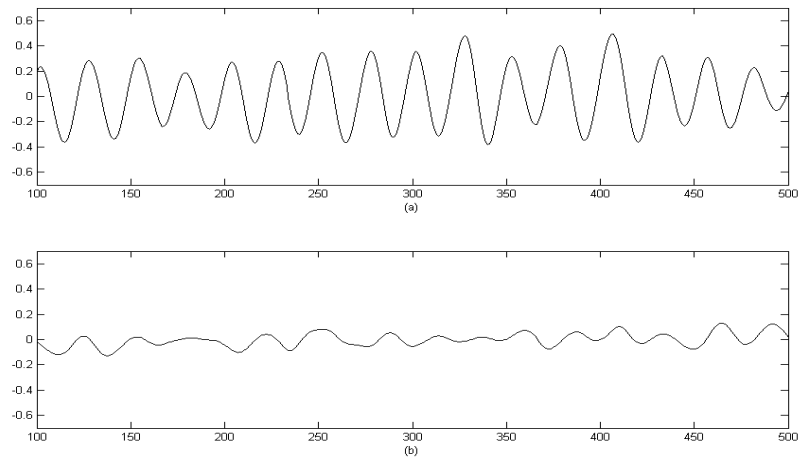


그림 2.3 심실세동 파형  
 (a) 거친 심실세동 파형  
 (b) 약해진 심실세동 파형

심실세동 초기에는 비교적 많은 심근이 동시에 수축하므로 불규칙하고 거친 형태의 심전도가 그려진다. 몇 초 후에 거친 심실수축이 없어지고 낮은 전위의 아주 불규칙한 약해진 심실세동 파형들이 그려진다. 따라서 어떤 반복적인 심전도 패턴이 나타나지 못한다. 다만 전류가 한 방향으로 흐르다 다음에는 다른 방향으로 흐르면서 어떤 특정한 주기를 갖지 못하기 때문에 전위는 일정하면서도 주기적으로 변한다[13]. 이미 언급한 것처럼 심실세동이 일어나면 혈액의 배출이 일어날 수 없기 때문에 즉각적인 전기 충격과 같은 방법으로 처치하여 멈추지 않으면 사망하게 된다.

## 2.2 특징 추출

그림 2.4는 심실세동 검출 알고리즘의 특징 추출 과정에 대한 블록도이며, 잡음 제거, QRS 콤플렉스 검출 및 구간 설정, 특징점 추출로 구성되어 있다.

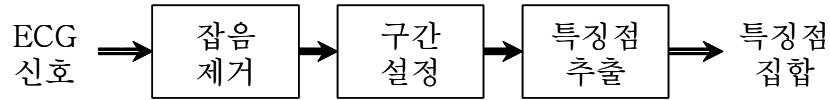


그림 2.4 특징 추출 과정에 대한 블록도

특징 추출을 위한 모든 과정은 웨이블릿 변환을 이용하였으며 *coif5* 웨이블릿을 사용하였다. 그림 2.5은 웨이블릿을 3번 취했을 경우( $j=-3$ )의 dyadic 트리 구조를 나타내고 있다.  $A_n$ 은  $n$ 번째 웨이블릿 변환을 취하여 얻은 저주파 성분을 나타내고,  $D_n$ 은 웨이블릿 변환을 취하여 얻은 고주파 성분을 나타낸다.

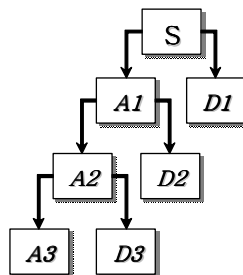


그림 2.5 Dyadic 트리 구조

표 2.1은 본 논문에서 사용한 360Hz 샘플링 주파수를 웨이블릿 레벨에 의해 분할한 결과를 나타내고 있다.

표 2.1 웨이블릿 변환 레벨에 의한 주파수 분할 (단위: Hz)

	-1	-2	-3	-4	-5	-6	-7	-8
저주파	0-90	0-45	0-22.8	0-11.4	0-5.7	0-2.9	0-1.4	0-0.7
고주파	90-180	45-90	22.8-45	11.4-22.8	5.7-11.4	2.9-5.7	1.4-2.9	0.7-1.4

## 2.2.1 웨이브렛을 이용한 잡음제거

본 연구에서는 ECG 신호의 기저선 및 고주파 성분을 제거하기 위해 웨이브렛 계수(A2, A8)을 사용하였다. 정상 및 비정상 파형은 대부분 20Hz 이내에서 큰 특성을 보이므로, 전처리 과정에서는 0.7Hz~45Hz 대역통과필터를 이용하여 잡음을 제거하였다[15][16]. 그림 2.6는 웨이브렛을 이용하여 기저선 및 고주파 성분의 잡음을 제거한 결과와 필터링 전·후의 스펙트럼 분포를 보여주고 있다. (a)는 정상 심전도에서 심실세동으로 전이되는 과정을 보여주고 있으며, 심실세동의 경우 정상 심전도에서 P-QRS-T의 모양이 모두 소실된 잡음에 가까운 매우 무질서한 모양을 보이고 있다. (b)는 기저선 및 고주파 성분을 제거한 결과를 보여주고 있고, 필터링을 통해 (c)에 나타난 잡음이 (d)에는 제거된 것을 확인할 수 있다.

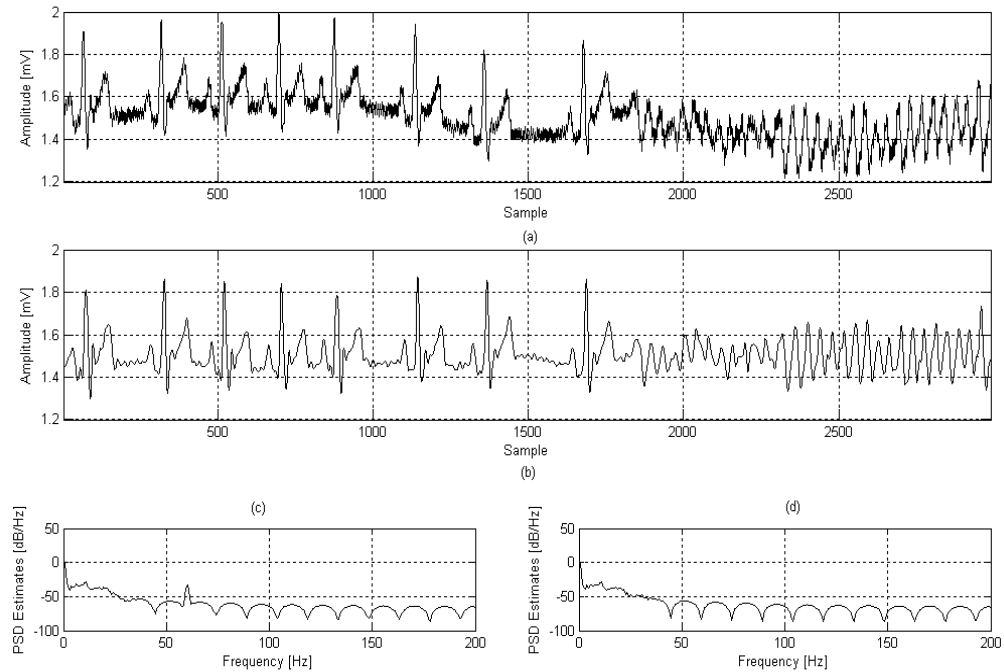


그림 2.6 웨이브렛을 이용한 잡음 제거

- (a) 잡음이 포함된 ECG 신호
- (b) 잡음이 제거된 ECG 신호
- (c) 필터링 전 스펙트럼
- (d) 필터링 후 스펙트럼

## 2.2.2 QRS 콤플렉스 검출 및 구간설정

본 연구에서는 QRS 콤플렉스 검출을 위해 웨이블릿 변환에 의해 얻은 계수(D3, D4)를 사용하였다. 전처리 과정을 통해 기저선 및 고주파 성분의 잡음을 제거한 ECG 신호를 제공하여 QRS complex를 부각시킨 후, QRS 콤플렉스가 포함되어 있는 웨이블릿의 3번째 레벨(D3)과 4번째 레벨(D4)의 신호를 더하여 합산신호( $D_{add}$ )를 식(2.1)와 같이 구성하였다.  $D_{add}$ 의 주파수 대역은 약 11.4Hz~45Hz으로 QRS 콤플렉스의 주파수 대역을 포함하고 있다.

$$D_{add} = D3 + D4 \quad (2.1)$$

그림 2.7는 정상과 심실세동이 합쳐진 심전도 신호와 이를 웨이블릿 변환하여 구성한 합산신호를 나타낸 것이다.

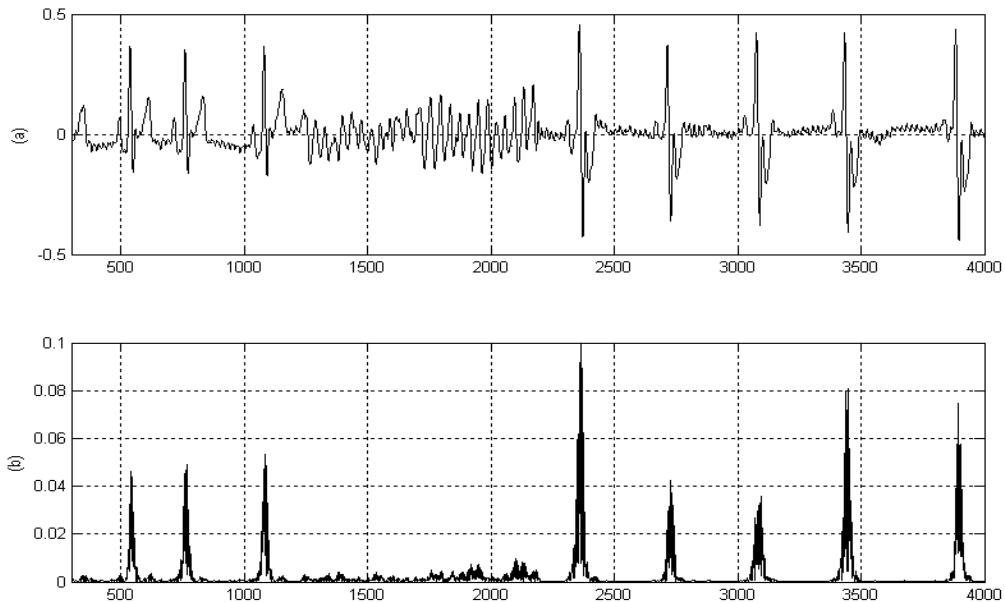


그림 2.7 심전도 신호와 웨이블릿 변환을 이용한 합산신호

(a) 정상과 심실세동 심전도 신호

(b) 합산신호

$Dadd$ 의 평균을 초기 문턱치로 설정하고,  $Dadd$ 가 문턱치보다 크면  $Dadd$ 에 가중치 1, 작으면 0을 적용하였으며, 가중치가 적용된  $Dadd$ 의 전·후 값의 차로 차분신호( $Ssub$ )를 식(2.2)와 같이 구성하였다.

$$Ssub = Dadd(i + 1) - Dadd(i) \quad \text{for } i = 1, \dots, N \quad (2.2)$$

$Ssub$ 가 급격히 감소하는 구간에서부터 0이 연속적으로 200ms 이상 지속되는 구간까지를 QRS 콤플렉스가 존재할 예상 구간으로 설정하고 이 구간 내에서의 최대값 및 최소값을 R점으로 검출하였다. 그림 2.8은 문턱치와 가중치를 적용한  $Dadd$ 와  $Dadd$ 의 전·후 값의 차로 구성된  $Ssub$ 를 나타낸 것이다.

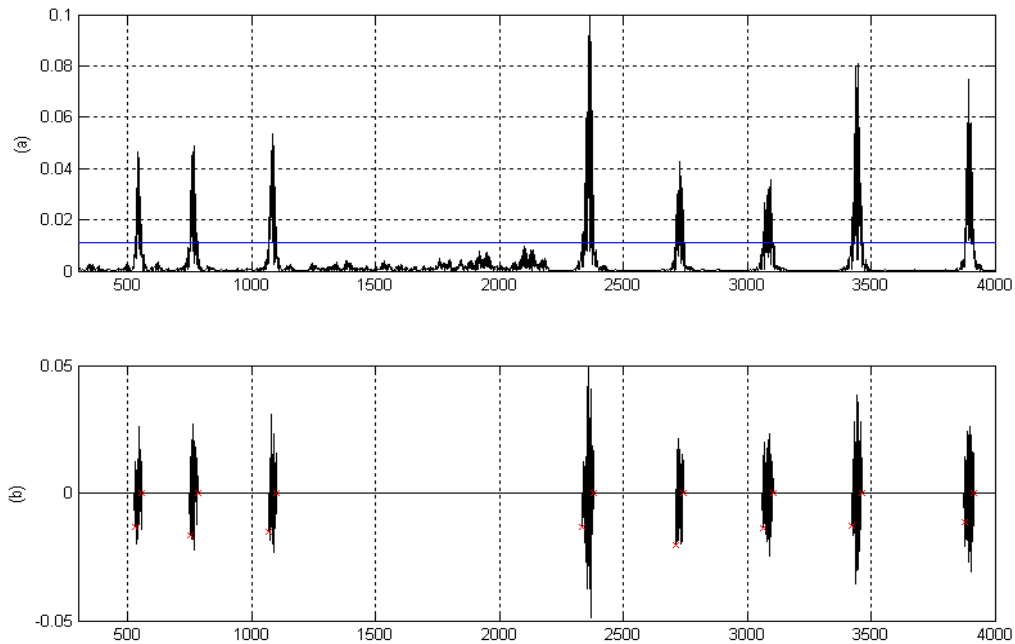


그림 2.8 QRS 콤플렉스가 존재할 예상구간 설정

(a) 합산신호와 문턱치

(b) 차분신호와 예상구간 설정

검출 오차를 줄이기 위하여 R점의 규칙성을 확인하였으며, RR 간격이 평균 RR의 0.5배 미만인 경우나 1.5배 이상인 경우에는 문턱치를 1%씩 감소하여 조정한 후 다시 QRS 콤플렉스를 검출하였다. 문턱치를 90%까지 조정하여 에러를 보정한 결과 RR 간격이 평균 RR 간격의 2~3배 이상인 경우에는 그 구간에 QRS 콤플렉스가 존재하지 않는 구간으로 간주하고, 웨이브렛의 5번째 레벨(A5)를 이용하여 신호를 신호의 peak을 검출하도록 하였다. 일반적으로 VF의 경우에는 R점이 존재하지 않기 때문에 본 연구에서는 구간 설정을 위해 신호의 peak점을 검출하였다 [17]. 그림 2.9은 정상과 심실세동이 포함된 심전도 신호로서 웨이브렛을 이용하여 QRS 콤플렉스를 검출 결과를 나타낸 것이다. x축은 샘플수, y축은 전압값(V)을 나타낸다.

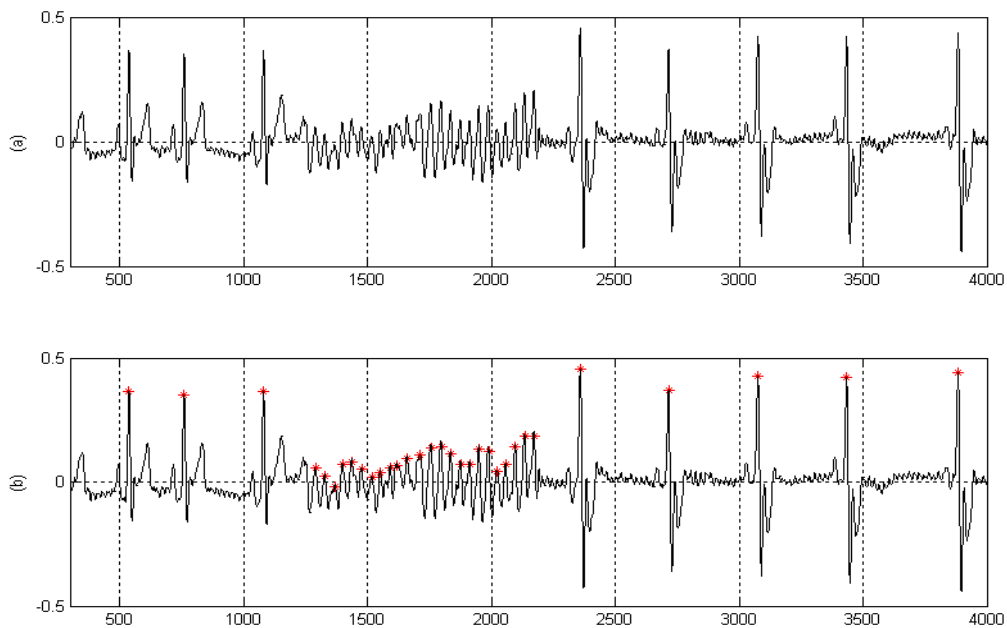


그림 2.9 웨이브렛을 이용한 QRS 콤플렉스 검출

(a) 정상과 심실세동 심전도 신호

(b) QRS 콤플렉스

그림 2.10은 MIT-BIH 데이터베이스의 106번 데이터와 웨이블릿 변환에 의해서 구성된 합산신호를 나타내고 있다. (a)에서 보는 바와 같이 106번 데이터는 다른 신호들에 비해 다양한 부정맥을 포함하고 있다. 웨이블릿 계수를 이용한 합산신호는 QRS 콤플렉스를 확연하게 구별할 수 있는 정보를 제공하여 정확한 QRS 콤플렉스를 검출할 수 있게 한다. QRS 콤플렉스 예상 구간을 설정하면, 0을 기준으로 positive 피크와 negative 피크의 크기를 고려하여 최종의 QRS 콤플렉스를 검출하게 된다.

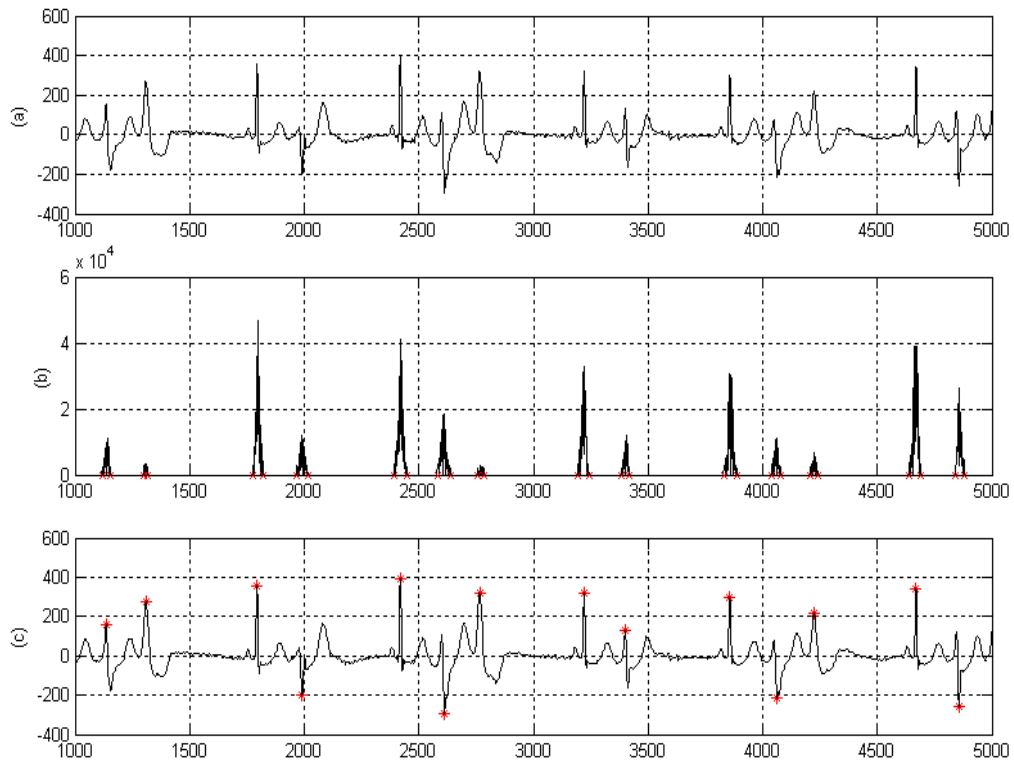


그림 2.10 심전도 신호(레코드 106)의 웨이블릿을 이용한 QRS 콤플렉스 검출  
 (a) 심전도 신호  
 (b) 합산신호  
 (c) QRS 콤플렉스 검출

본 논문에서는 신호분석을 위해 R점을 기준으로 전·후의 200ms를 분석구간으로 설정하였다. 그림 2.11은 R점을 기준으로 전·후 200ms로 나눈 것으로 심실세동을 비롯한 다양한 부정맥은 R점을 기준으로 200ms 전·후의 구간에서 확연한 차이를 보이므로, 본 논문에서는 구간을 설정하여 심전도의 형태 및 주파수 성분을 분석하였다.

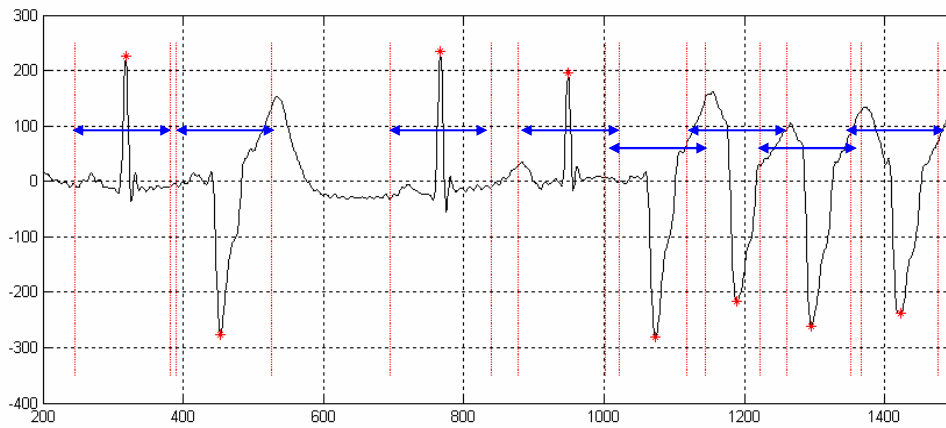


그림 2.11 분석 구간



## 2.2.3 특징점 추출

본 연구에서는 심실세동을 검출하기 위한 파라미터로 RR 간격, 웨이브렛 변환에 의한 상세계수를 특징점으로 추출하였다. 특징점 추출을 위해 MIT-BIH 부정맥 데이터베이스 중 VF 등 다양한 부정맥이 포함되어 있는 10개의 레코드를 사용하였다.

### 2.2.3.1 RR 간격을 이용한 특징점 추출

본 논문에서는 RR 간격의 비율을 이용하여 특징점을 추출하였다. 정상 파형의 경우는 RR 간격은 대부분 R점 전·후 RR간격이 비슷하거나, 이전 RR 간격이 이후 RR 간격보다 더 긴 특성을 보이는 반면에 대부분의 비정상 파형의 경우에는 이전 RR 간격이 이후 RR 간격보다 짧거나, 특히 심실세동의 경우 R점 전·후의 RR 간격은 정상 파형의 평균 RR 간격에 비해 매우 짧은 특성이 있다[18]. 따라서 RR 간격을 식(2.3), (2.4)에 적용하여 특징 벡터로 추출하였다. 그림 2.12은 4 가지 심전도에서 추출된 RR 간격을 이용한 특징점의 평균과 표준편차를 나타낸 것으로서 (a)는 평균 RR 간격과 R점을 기준으로 이전의 RR 간격의 비, (b)는 평균 RR 간격과 이후의 RR 간격의 비를 나타낸 것이다.

$$Feature\ 1 = \frac{\text{평균 } RR \text{ 간격}}{\text{이전 } RR \text{ 간격}} \quad (2.3)$$

$$Feature\ 2 = \frac{\text{평균 } RR \text{ 간격}}{\text{이후 } RR \text{ 간격}} \quad (2.4)$$

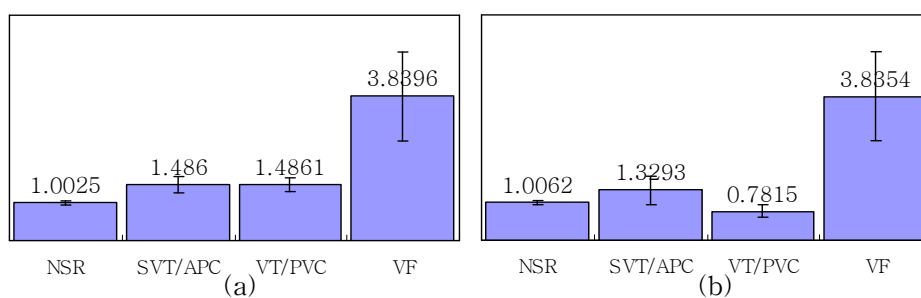


그림 2.12 RR 간격을 이용한 특징점

(a) 평균 RR 간격과 이전 RR 간격의 비율

(b) 평균 RR 간격과 이후 RR 간격의 비율

### 2.2.3.2 웨이브렛 변환을 이용한 특징점 추출

본 논문에서는 웨이브렛 변환에 의한 상세계수(detail coefficient)를 이용하여 특징점을 추출하였다. 웨이브렛 변환으로 추출된 계수들은 각각의 레벨마다 포함하고 있는 주파수 성분만을 시간축 상에 나타내게 된다. 따라서 본 논문에서는 정상 및 다양한 부정맥의 주요 주파수 성분을 확연하게 구분할 수 있는 4번째 레벨(D4), 5번째 레벨(D5), 6번째 레벨(D6) 그리고 7번째 레벨(D7)을 선택하여, 15개의 특정 포인트에서 나타나는 값을 특징 벡터로 추출하였다. 그림 2.13은 다양한 심전도에 대한 웨이브렛 계수(D4, D5, D6, D7)의 나타낸 것으로 15개의 특정 포인트에서의 값의 범위가 크게 차이가 나는 것을 확인할 수 있다.

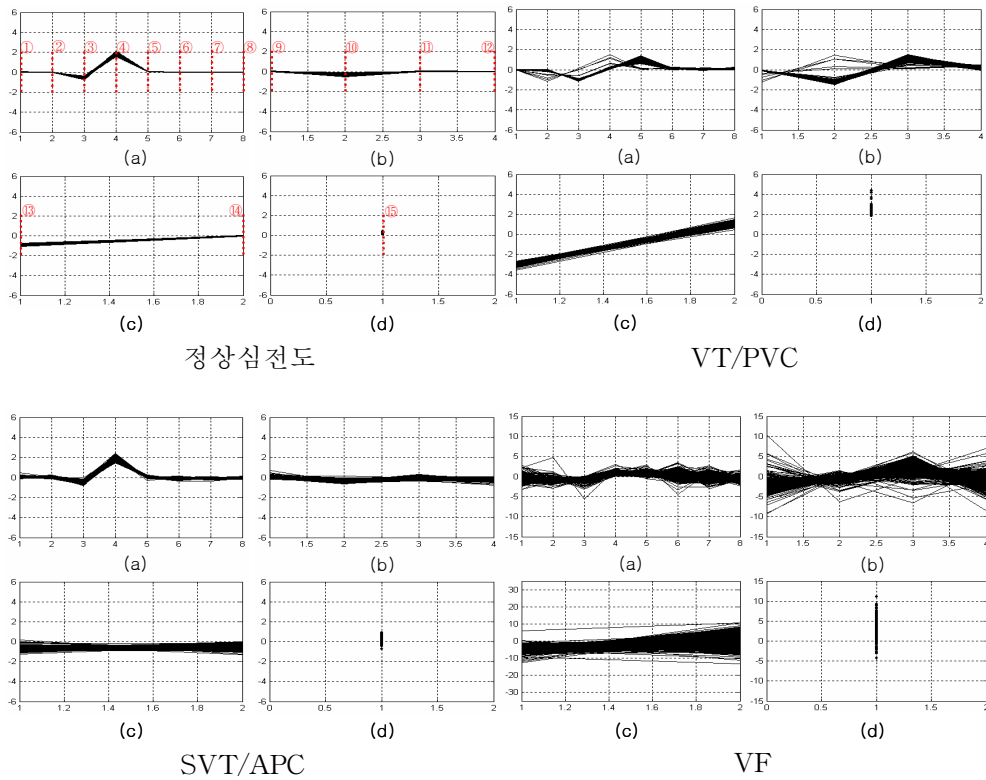


그림 2.13 웨이브렛 계수를 이용한 특징점

- (a) D4 신호
- (b) D5 신호
- (c) D6 신호
- (d) D7 신호

## 2.3 선형판별분석기법(linear discriminant analysis)

본 논문에서는 추출된 특징간의 최적의 구별을 가능케 하는 벡터를 추출하는 방법으로 선형판별분석법(linear discriminant analysis: LDA)을 사용하였다. 상기 특징 추출 과정에서 RR 간격과 웨이브렛 상세계수로부터 추출한 17개의 특징점들에 LDA를 적용하여 선형변환에 의해 새롭게 구성된 4개의 특징 벡터를 추출하였다.

LDA는 클래스 내의 데이터 분산(within class variance)을 최소화 하고, 클래스 간의 분산(between class variance)를 최대로 하는 결정 초평면(decision hyperplane)을 분석하는 알고리즘이다. LDA는 클래스 분리를 최대화 시켜 주므로, 클래스 간의 특징 벡터들을 비교적 정확하게 분류하는 장점이 있다[11][19]. 그림 2.14은 정상심전도와 비정상심전도의 특징점 2개를 x, y축으로 나타낸 것으로, LDA를 적용한 차원 감소를 간단히 설명하기 위해 상기에서 추출된 15개의 특징점 중 2개의 특징점을 이용하여 2차원 그래프로 나타낸 것이다. (a)는 두 클래스 간 분리를 최대로 하는 방향(LDA 방향)을 나타내고 있고 (b)는 LDA 방향으로 투영된 새로운 데이터와 그것의 가우시안 분포를 보여주고 있다.

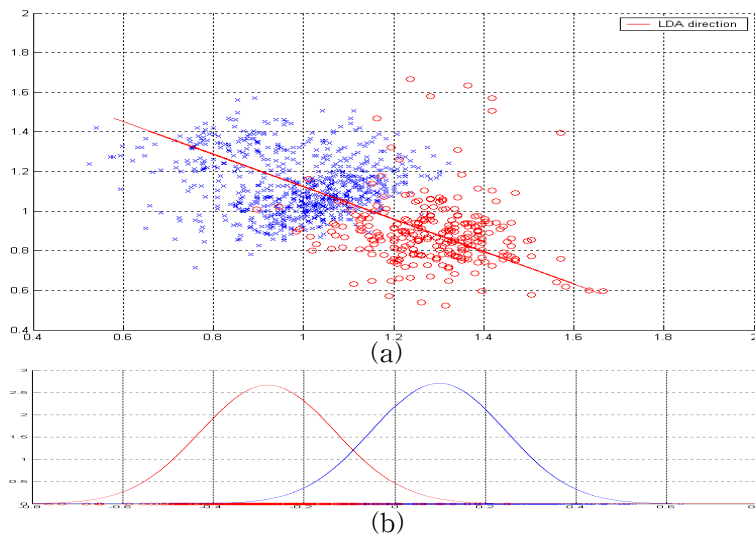


그림 2.14 LDA를 이용한 특징 추출 및 차원 감소

(a) LDA 방향

(b) 추출된 특징과 분포

(x: 정상심전도, o: 비정상심전도)

n 차원 벡터 공간의 열벡터  $\overrightarrow{X_{ij}}$ 를  $m(m < n)$ 차원 특징 공간으로 매핑하는 선형 변환에 의해 새로운 특징 벡터  $\overrightarrow{Y_{ij}}$ 는 식(2.5)와 같이 정의된다.

$$\overrightarrow{Y_{ij}} = W^T \overrightarrow{X_{ij}} \quad (2.5)$$

여기서,  $\overrightarrow{X_{ij}}$ 와  $\overrightarrow{Y_{ij}}$ 는  $W$ 에 의한 선형 변환 이전 및 이후의  $i$ 번째 클래스에 속한  $j$ 번째 열벡터를 나타낸다. 클래스 간의 분산행렬은 식(2.6), 클래스 내의 분산행렬을 식(2.7)과 같이 정의한다.

$$S_b = \sum_{i=1}^C (\overrightarrow{M_i} - \overrightarrow{M})(\overrightarrow{M_i} - \overrightarrow{M})^T \quad (2.6)$$

$$S_w = \sum_{i=1}^C \sum_{j=1}^{N_i} (\overrightarrow{X_{ij}} - \overrightarrow{M_i})(\overrightarrow{X_{ij}} - \overrightarrow{M_i})^T \quad (2.7)$$

여기서,  $\overrightarrow{M_i}$ 는  $i$ 번째 클래스의 평균,  $\overrightarrow{M}$ 은 전체 평균,  $c$ 는 클래스의 수이며  $N_j$ 는  $j$ 번째 클래스의 데이터 수이다. 클래스 간 분산 행렬의 결정과 클래스 내의 분산 행렬의 비를 최대화 하는 최적의 선형변환행렬  $W_{opt}$ 는 식(2.8)를 통해 계산하게 된다.

$$F(W) = \frac{|W^T S_b W|}{|W^T S_w W|} \quad (2.8)$$

LDA를 통해 감소할 수 있는 최대의 차원 수는 클래스가  $c$ 개 일 때,  $c-1$ 까지 차원을 감소할 수 있다. 그림 2.15은 LDA를 적용하여 새롭게 구성된 특징 벡터를 나타낸 것으로 4가지의 심전도에 대한 평균과 표준편차를 보여주고 있다. 그림 2.15에서 볼 수 있듯이 LDA에 의해 추출된 특징벡터는 다양한 심전도끼리 확연히

구분 할 수 있도록 기존의 특징점들을 재구성했다.

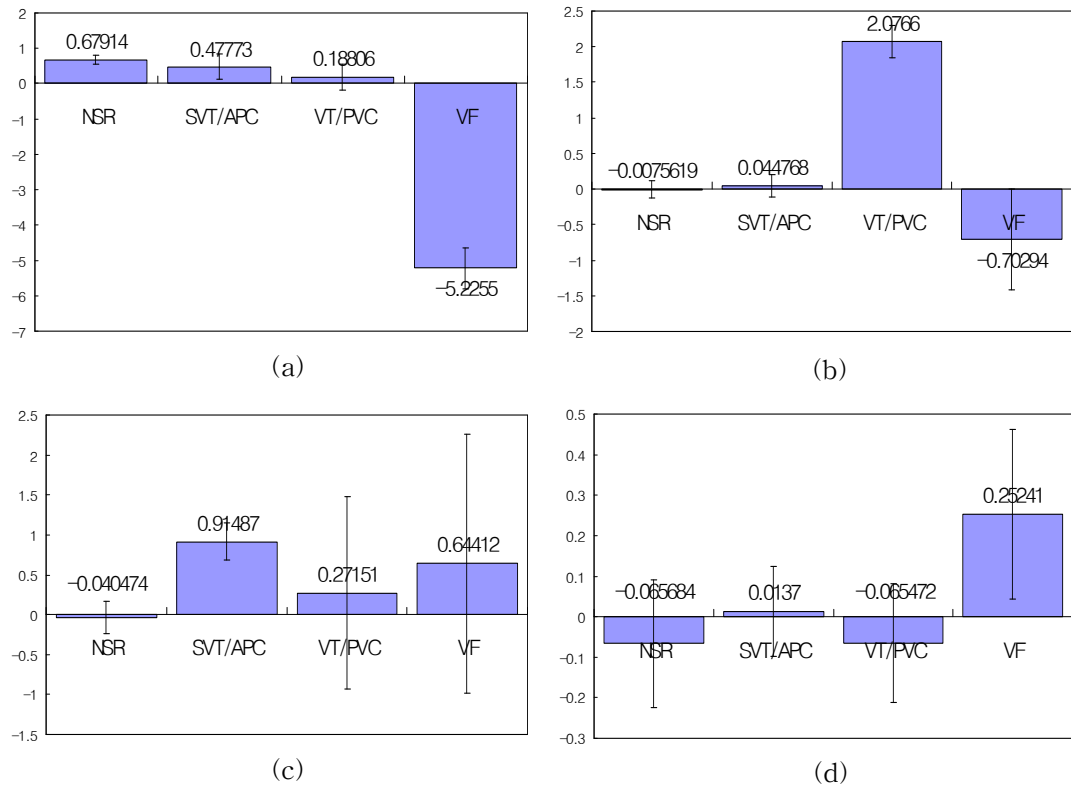


그림 2.15 LDA를 적용한 새로운 특징점

- (a) 특징점1
- (b) 특징점2
- (c) 특징점3
- (d) 특징점4

## 2.4 Support Vector Machine(SVM)

본 논문에서는 심실세동 검출을 위해 SVM 다원 분류기를 사용하였다. 그림 2.16는 SVM 다원 분류기의 블록도를 나타낸 것으로서 LDA를 적용하여 추출한 특징을 5개의 SVM 분류기의 입력으로 사용하였다. 본 논문에서는 5가지의 다양한 SVM 분류기를 이용하여 심실세동을 검출하였다. SVM 다원 분류기는 NSR 분류기, PVC/VT 분류기, APC/SVT 분류기, VF 분류기 그리고 다른 부정맥 분류기로 구성되어 있다. VF 분류기는 입력된 특징을 이용하여 VF인지 아닌지를 분류하게 된다. 만약 VF면 positive value를 출력하게 되고, VF가 아니면 negative value를 출력하게 된다. 다른 4가지 분류기들도 상기와 같은 방법으로 결과값을 출력하게 되고, 최종 분류는 5가지 분류기들의 결과값 중 최대값을 출력하는 분류기로 분류된다.

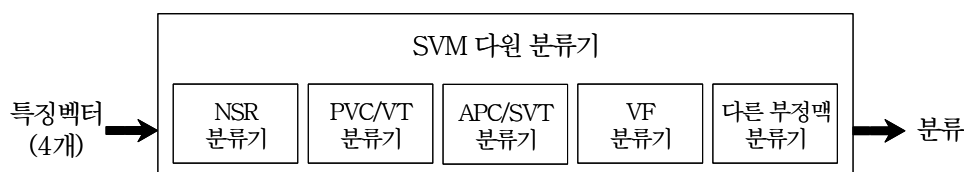


그림 2.16 SVM 다원 분류기 블록도

SVM은 기본적으로 두 범주를 갖는 객체들을 분류하는 초평면(hyperplane)을 분석하는 것으로 객체들의 분리를 최대화 할 수 있는 최적 분리 경계면을 제공한다. 기존의 분류기들 대부분이 경험적 위험 최소화(empirical risk minimization)에 기초하는 반면에 SVM은 일반화 에러의 상한(upper bound)을 최소화하는, 구조적 위험 최소화(structural risk minimization) 기법에 그 기초를 두고 있다[12][20]–[25]. 훈련 데이터가 식(2.9)와 같이 주어졌을 때, 입력데이터( $x_i$ )는 두 클래스 중의 하나에 속하고,  $y_i$ 는 해당 클래스를 표시하는 라벨의 역할을 한다.

$$[(x_1, y_1), \dots, (x_n, y_n)] \in \mathbb{R}^d \times \{\pm 1\} \quad i = 1, \dots, N \quad (2.9)$$

SVM은 각 클래스를 구분하는 최적의 분리경계면(separating hyperplane)을 구하기 위해 분리 경계면과 분리 경계면에 가장 인접한 점(support vector)과의 거리를 최대화 한다. 만약, 두 클래스가 선형적으로 분리가 가능하다면, 두 클래스로 구분할 수 있는 최적의 선형 분리 경계면은 식(2.10)으로 정의할 수 있다.

$$f(x) = w^T x + b \quad (2.10)$$

식(2.10)에 의해, support vector와  $f(x)$ 의 거리를  $\frac{1}{\|w\|}$ 로 나타낼 수 있고, SVM은  $\|w\|$ 를 최소화하여 분리 간격이 최대가 되도록 하는 최적 분리면을 구하게 된다. 이러한 기본적인 형태에서 SVM은 학습데이터를 분류하고 가장 작은 가중치 벡터를 가지는 분리면을 찾는다. 이 분리면은 positive 학습 예제들과 negative 학습 예제들을 최대 마진(margin)으로 분류한다. 그림 2.17는 정상심전도와 비정상심전도의 특징점 2개를 x, y축으로 나타낸 것으로 최적 분리면이 정상심전도와 비정상심전도의 특징들을 최대 마진(margin)을 갖도록 분류한 것을 볼 수 있다. 이때 최적 분리면에 가장 가까운 특징들을 support vector라고 하며, 그림 2.17에서는 원으로 표시했다.

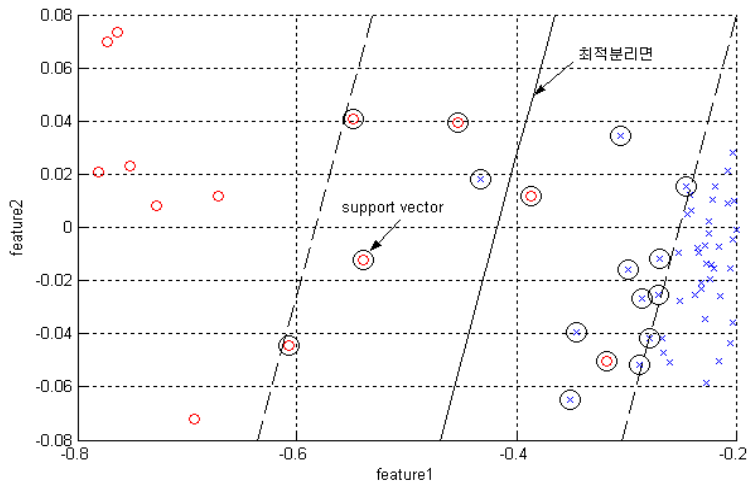


그림 2.17 SVM의 최적 분리면과 support vector  
(x: 정상심전도, o: 비정상심전도)

이 분리면을 찾는 것은 다음의 최적화 문제로 해석될 수 있다. SVM은 식(2.12)의 조건하에서 식(2.11)을 최소로 하는 가중치 벡터  $w$ 를 구하여 분리 간격이 최대화되는 최적 분리면을 구하게 된다.

$$\min \frac{1}{2} w^T w \quad (2.11)$$

$$y_i (w^T x_i + b) \geq 1 \quad i = 1, \dots, N \quad (2.12)$$

식(2.11)을 라그랑주 승수(Lagrange Multiplier)로 다시 쓰면 식(2.14)의 조건하에서 식(2.13)을 최대로 하는 최적화 문제를 푸는 것과 같게 된다.

$$\max Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (2.13)$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \quad \text{and} \quad \alpha_i > 0 \quad i = 1, \dots, N \quad (2.14)$$

$Q(\alpha)$ 를 최대로 하는  $\alpha$ 를  $\alpha_0$ 라 하면, 최적 가중치 벡터  $w_0$ 는 식(2.15)와 같고, 최적 바이어스  $b_0$ 는 식(2.16)을 이용하여 구할 수 있다. 여기서  $x^{(s)}$ 는 support vector이다.

$$w_0 = \sum_{i=1}^N \alpha_0 y_i x_i \quad (2.15)$$

$$b_0 = 1 - w_0^T x^{(s)} \quad y^{(s)} = 1 \quad (2.16)$$

입력 공간에서의 데이터들이 선형적으로 분리 가능 하지 않는 경우에는, 입력공간에서의 데이터를 비선형변환을 통해 보다 높은 차원의 공간에서 데이터를 선형적으로 분리가능하게 만들 수 있다. 입력공간에서 특징공간으로의 비선형 변환을



$\varphi(x)$ 라 두면 특징공간에서의 최적 분리면은 식(2.17)으로 구할 수 있다.

$$\sum_{i=1}^N \alpha_i y_i \varphi^T(x) \varphi(z) = 0 \quad (2.17)$$

그러나 특징공간은 매우 높은 차원의 공간이므로 내적(inner-kernel) 커널을 이용하여 특징공간에서 직접 연산을 수행하지 않고 입력공간에서 데이터를 처리할 수 있다. 커널 함수를 식(2.18)이라 한다면, 식(2.13)과 (2.14)는 식(2.19)와 (2.20)과 같이 나타낼 수 있다.

$$k(x, z) = \varphi^T(x) \varphi(z) \quad (2.18)$$

$$\max Q(\alpha) = \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j k(x_i, x_j) \quad (2.19)$$

$$\sum_{i=1}^N \alpha_i y_i = 0 \quad \text{and} \quad 0 \leq \alpha_i \leq C \quad i = 1, \dots, N \quad (2.20)$$

여기서 C는 페널티 파라미터이다. 라그랑주 승수  $\alpha_i$ 를 구하면 특징공간에서 최적 초평면을 구할 수 있고, 식(2.21)은 최적 초평면 함수를 정의한다.

$$f(x, \alpha^*) = \sum_{i=1}^N \alpha_i^* y_i k(x, x_i) \quad (2.21)$$

여기서  $\alpha_i^*$ 는 라그랑주 승수,  $k(x, x_i)$ 는 커널함수,  $y_i$ 는 학습데이터의 라벨,  $x$ 는 입력데이터,  $x_i$ 는 SV(support vector)이다. 일반적으로 커널 함수로는 표 2.2에 나타난 세 가지가 주로 사용되어지고 있고, 커널 함수의 타입에 따라 SVM은

polynomial 분류기, radial basis 함수(RBF) 분류기, 2-layer sigmoid 분류기로 나눌 수 있다[20]-[25].

표 2.2 커널 함수와 SVM 분류기의 종류

커널 함수	분류기 형태
$k(x, x_i) = (\tau + x_i^T x)^p$	p차 다항식
$k(x, x_i) = \exp(-\frac{1}{2\sigma^2} \ x - x_i\ ^2)$	Radial-Basis 함수
$k(x, x_i) = \tanh(\kappa_1 x_i^T x + \kappa_2)$	다층퍼셉트론

\* $\tau, \sigma, \kappa_1, \kappa_2$ : 커널 함수 파라미터

본 논문에서는 SVM 분류기에 사용되는 커널 함수로 RBF(Radial-Basis Function) 커널을 사용하였고, 커널의 폭인  $\sigma$ 와 마진폭과 분류 오류 사이의 타협점을 찾아주는 역할을 담당하며, 분리할 수 없는 데이터에 대한 오류 패널티 변수인 C값은 실험적인 방법을 통해 각각 1, 10으로 결정하였다. SVM 다원 분류기의 학습을 위해 정상심전도, 심실세동 및 다양한 부정맥을 포함한 4135 비트의 심전도데이터를 사용하였다.

## 제 3 장 실험 결과 및 고찰

본 논문에서는 제안한 알고리즘의 성능을 평가하기 위하여 차원 감소의 대표적인 방법인 주성분 분석(principal component analysis: PCA)[11][19]과 심실세동 및 부정맥 검출 알고리즘에서 널리 사용되어지고 있는 신경망[8]과 퍼지 추론시스템[9]을 이용하여 검출 성능을 비교, 평가하였다.

### 3.1 실험 방법

본 논문에서는 QRS 콤플렉스 검출 성능 평가, 파라미터에 따른 SVM 성능비교와 심실세동 검출 알고리즘 성능 평가를 위해 360Hz로 샘플링 된 MIT-BIH 부정맥 데이터베이스의 ECG 신호를 사용하였다. 제안한 알고리즘을 이용한 심실세동 검출에서는 MIT-BIH 부정맥 데이터베이스와 250Hz로 샘플링 된 Creighton University Ventricular Tachyarrhythmia 데이터베이스, MIT-BIH Malignant Ventricular Arrhythmia 데이터베이스 중 VF가 포함된 ECG 신호를 360Hz로 up-sampling하여 사용하였다.

검출 성능 평가를 위한 방법으로는 식(3.1), 식(3.2)와 식(3.3)에 정의된 민감도(sensitivity: ST)와 특이도(specificity: SF), 정확도(accuracy: AC)를 사용하였다 [29,30]. 여기서 TP은 true positive, FN은 false negative, TN은 true negative, FP는 false positive의 약자이다.

$$Sensitivity(ST) = \frac{TP}{TP + FN} \times 100 \quad (3.1)$$

$$Specificity(SF) = \frac{TN}{TN + FP} \times 100 \quad (3.2)$$

$$Accuracy(AC) = \frac{(TP + TN)}{(TP + FN + TN + FP)} \times 100 \quad (3.3)$$

### 3.2 QRS 콤플렉스 검출율

본 논문에서는 웨이브렛 변환을 이용한 QRS 콤플렉스 검출 알고리즘의 성능을 평가하기 위해 MIT-BIH 부정맥 데이터베이스 중 24개 레코드를 사용하였다. 표 3.1은 검출 결과를 나타낸 것으로서 평균 99.48%의 우수한 검출 성능을 보여주고 있다. 다른 신호에 비해 잡음의 영향을 많이 받은 105번 데이터와 다양한 형태의 부정맥을 포함하고 있는 데이터들에서 많은 검출 오류를 나타냈다.

표 3.1 QRS 콤플렉스 검출 결과

데이터 종류	총갯수 (비트)	FP (비트)	FN (비트)	검출오류 (비트)	에러율 (%)
105	2576	23	10	33	1.28
106	2019	45	1	46	2.28
112	2528	0	5	5	0.20
113	1878	1	3	4	0.21
114	1871	1	5	6	0.32
115	1944	0	0	0	0
116	1402	1	22	23	0.96
117	1528	0	1	1	0.07
119	1981	0	0	0	0
121	1849	0	0	0	0
122	2401	0	0	0	0
123	1511	0	4	4	0.27
124	1611	0	1	1	0.06
200	2591	21	3	24	0.93
205	2592	2	0	2	0.08
208	2943	30	1	31	1.05
209	2993	3	0	3	0.10
213	328	0	3	3	0.09
215	3352	0	2	2	0.06
220	2039	0	0	0	0
223	2417	0	8	8	0.33
228	2047	51	5	56	2.74
233	2029	28	0	28	1.38
234	2742	0	0	0	0
평균	55085	206	74	280	0.52

### 3.3 파라미터에 따른 SVM의 성능 비교

본 논문에서는 학습 데이터가 정확히 두 범주로 나뉘지지 않는 경우 일 때, 학습 데이터에 어느 정도의 오류를 허용하기 위한 임계치 C값과 RBF 커널의 파라미터인  $\sigma$ 값에 따른 SVM의 성능을 평가하였다. MIT-BIH 부정맥 데이터베이스 중 10개 레코드를 사용하여 정상심전도(NSR)을 분류하였다. 표 3.2는 C값과  $\sigma$ 값에 의한 SVM 분류 성능을 비교해 놓은 결과로서, C값이 0.1, 1, 10일 때  $\sigma$ 값을 1부터 4까지 변화시켜가며 성능을 비교하였다. 표 3.2에 나타난 결과를 보면  $\sigma$ 값의 변화에는 성능의 차이가 거의 없었으며, C값이 0.1일 때 평균적으로 가장 낮은 분류 성능을 보였다. C값이 10이고  $\sigma$ 값이 1일 때 가장 좋은 성능을 보였다.

표 3.2 C값과  $\sigma$ 값에 따른 분류 성능 비교

$\sigma$	C=0.1			
	sensitivity (%)	specificity (%)	accuracy (%)	평균 (%)
1	99.969	92.862	97.851	96.894
2	100	91.114	97.351	96.155
3	100	88.711	96.635	95.115
4	100	85.652	95.723	93.791
$\sigma$	C=1			
	sensitivity (%)	specificity (%)	accuracy (%)	평균 (%)
1	99.907	94.829	98.393	97.909
2	99.969	92.935	97.872	96.925
3	100	92.061	97.634	96.565
4	100	91.697	97.525	96.407
$\sigma$	C=10			
	sensitivity (%)	specificity (%)	accuracy (%)	평균 (%)
1	99.938	95.776	98.697	<b>98.137</b>
2	99.907	94.610	98.328	97.615
3	99.938	93.518	98.024	97.160
4	100	92.935	97.894	96.943

### 3.4 심실세동 검출 알고리즘 성능 평가

본 논문에서는 제안한 알고리즘의 성능을 비교, 평가하기 위해서 두 가지 실험을 하였다. 첫째, LDA의 성능을 평가하기 위해 LDA를 적용하지 않은 기존의 특징점 (original features: OFS) 집합, 주성분 분석(PCA)을 적용하여 특징점과 본 논문에서 제안한 LDA를 적용한 특징점을 분류기의 입력으로 사용하여 검출 성능을 비교 하였다. 둘째, SVM의 분류 성능을 평가하기 위해 심실세동 및 부정맥 검출 알고리즘에서 주로 사용되어지고 있는 신경망[8], 퍼지 추론 시스템[9]과 본 논문에서 제안한 SVM 다원 분류기를 사용하여 성능을 비교하였다. 그림 3.1은 심실세동 검출 알고리즘 성능 평가를 위한 전체 블록도이다.

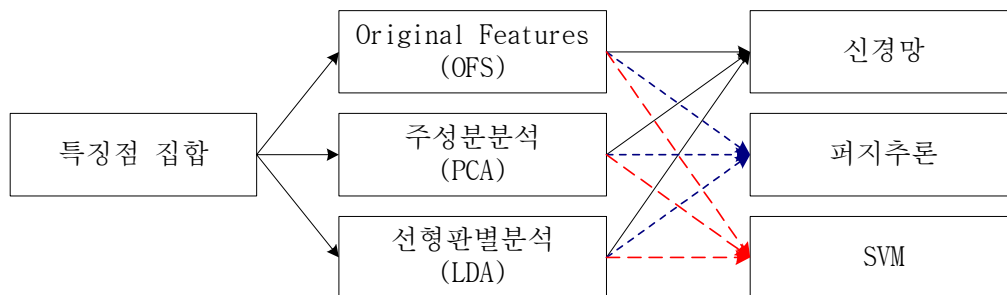


그림 3.1 성능 평가에 대한 전체 블록도

주성분 분석(principal component analysis: PCA)은 LDA와 함께 차원 축소를 위해 널리 알려진 기법 중의 하나로서 고차원의 입력 데이터를 저차원의 데이터로 변환하여 몇 개의 주성분 값으로 나타내는 것으로, 서로 다른 클래스를 차원을 줄여서 간단하게 표현할 수 있는 체계적이고 실용적인 방법이다[19].

본 논문에서는 LDA를 적용하지 않은 기존의 특징점 집합과 PCA, LDA를 적용한 특징점을 신경망, 퍼지추론 시스템과 SVM 다원 분류기의 입력으로 각각 사용하여 정상심전도(NSR)의 검출 성능을 평가하였다. MIT-BIH 부정맥 데이터베이스 중 VF 등 다양한 부정맥이 포함되어 있는 23개 레코드를 사용하였다.

### 3.4.1 신경망(Neural Network)

본 연구에서는 다양한 신경망 중에서 오류 역전파(error back propagation) 학습 알고리즘을 사용하였다. 본 연구에서는 입력층, 은닉층, 출력층을 각 하나씩 갖으며 입력층은 4개의 노드, 출력층은 5개의 노드를 갖는 신경망을 구성하였다. 은닉층은 시그모이드 함수를 적용하였으며 최적의 성능을 보이는 10개의 노드수를 사용하였다. 신경망은 26512개의 ECG 신호를 사용하여 학습시켰다. 표 3.3는 신경망을 이용하여 정상심전도를 검출한 결과를 나타낸 것이다.

표 3.3 신경망을 이용한 정상심전도 검출 결과 (단위 : %)

데이터 종류	OFS			PCA			LDA		
	ST	SF.	AC	ST	SF	AC	ST	SF	AC
101	100	100	99.88	100	100	99.88	100	100	100
103	99.86	100	99.76	100	100	99.90	100	100	99.90
106	100	97.29	99.31	96.93	98.26	97.27	100	99.81	99.95
112	100	100	99.92	100	100	99.92	100	100	99.92
113	100	100	99.66	100	100	99.66	100	100	100
114	96.45	96.61	98.56	99.95	86.44	99.52	99.95	91.53	99.68
115	100	-	100	100	-	100	100	-	100
116	100	98.49	99.92	100	98.49	99.92	100	98.49	99.92
117	99.93	100	99.93	100	100	99.93	100	100	100
119	100	100	100	100	100	100	100	100	100
121	97.57	100	97.57	98.86	50	98.81	99.84	100	99.84
122	100	-	100	100	-	100	100	-	100
123	99.80	100	99.80	100	100	100	99.80	100	99.80
200	99.65	92.63	97.33	99.54	93.45	97.53	100	94.62	98.22
202	99.46	92.16	97.32	99.23	92.48	97.32	100	94.77	98.47
205	99.96	87.06	99.55	100	83.53	99.47	100	84.71	99.51
208	99.30	100	99.57	99.49	99.56	99.53	99.81	98.89	99.46
213	100	89.34	99.09	99.73	90.16	98.92	100	88.53	99.03
215	97.93	91.07	97.58	99.91	92.86	99.55	98.65	86.31	98.03
223	98.91	48.96	87.85	99.46	73.87	93.79	99.90	74.91	97.37
228	99.99	99.18	99.02	99.46	97.81	99.17	99.76	98.90	99.61
233	99.96	95.50	98.73	99.32	88.76	96.41	99.87	97.04	99.09
234	100	-	100	99.84	-	99.84	100	-	100
평균	99.49	94.68	98.70	99.66	92.65	98.95	99.89	95.64	99.32

\*ST=sensitivity, SF=specificity, AC=accuracy

\*OFS=original features, PCA=principal component analysis, LDA=Linear discriminant analysis

### 3.4.2 퍼지 추론 시스템(fuzzy inference system)

논문에서는 입력된 특징들을 if-then 규칙으로 구성된 지식기반(fuzzy rule base)을 바탕으로 추론값을 출력하는 퍼지 추론을 사용하였다. 퍼지 추론은 일반적으로 가장 널리 사용되는 min-max 방법을 사용한 Mamdani 추론 방법을 사용하였으며, 비퍼지화 방법으로 무게중심(centroid) 방법을 사용하였다. 퍼지 추론 시스템의 성능을 좌우하는 소속 함수는 가우시안 분포 곡선을 적용하여 반복적인 실험을 통해 획득하였다. 표 3.4는 퍼지 추론 시스템을 이용하여 정상심전도를 검출한 결과를 나타낸 것이다.

표 3.4 퍼지 추론 시스템을 이용한 정상심전도 검출 결과 (단위 : %)

데이터 종류	OFS			PCA			LDA		
	ST	SF	AC	ST	SF	AC	ST	SF	AC
101	100	100	100	100	100	99.88	100	100	99.88
103	99.03	100	98.94	100	100	99.90	100	100	99.90
106	99.80	98.07	99.36	99.53	95.16	98.41	100	93.42	98.31
112	98.26	100	98.18	83.76	50	73.73	99.96	100	99.88
113	95.33	83.33	95.29	100	100	99.66	100	100	99.66
114	97.89	84.75	97.48	99.95	84.75	99.46	99.83	88.14	99.46
115	98.19	-	98.19	100	-	100	100	-	100
116	100	95.46	99.76	100	98.48	99.92	100	98.49	99.92
117	99.40	100	99.41	100	100	99.93	100	100	99.93
119	99.67	99.32	99.59	100	100	100	100	100	100
121	95.67	100	95.68	91.75	100	31.82	97.24	100	97.25
122	99.92	-	99.92	99.64	-	99.63	99.96	-	99.96
123	94.62	100	94.63	100	100	100	99.980	100	99.80
200	96.88	95.01	96.29	99.59	93.45	97.57	99.94	91.35	97.10
202	96.62	96.41	96.56	99.32	92.16	97.22	100	90.20	97.13
205	99.92	81.17	99.32	99.96	84.71	99.48	99.96	84.71	99.47
208	95.57	68.19	85.04	98.61	100	99.14	99.11	100	99.45
213	99.66	66.39	96.83	99.73	90.57	98.95	99.96	89.34	99.06
215	93.94	79.18	93.19	94.50	75.59	93.55	96.57	66.07	95.05
223	99.80	67.94	92.75	99.60	72.13	93.52	99.85	39.55	86.45
228	89.41	97.81	90.91	84.64	98.90	87.19	93.63	98.90	94.57
233	99.91	94.44	98.40	99.37	84.85	95.37	99.64	73.01	92.30
234	100	-	100	100	-	100	100	-	100
평균	97.89	90.56	96.90	95.42	91.46	94.70	99.39	91.10	98.08

\*ST=sensitivity, SF=specificity, AC=accuracy

\*OFS=original features, PCA=principal component analysis, LDA=Linear discriminant analysis



### 3.4.3 SVM 다원 분류기

본 논문에서는  $C$ 가 10이고,  $\sigma$ 가 1인 RBF 커널함수를 이용한 SVM을 사용하였다.

표3.5는 SVM 다원 분류기를 이용한 정상심전도 검출 결과를 나타낸 것이다.

SVM은 상기에서 적용한 신경망과 퍼지추론 시스템에 비해 평균적으로 우수한 검출 성능을 보여주고 있으며, 그 중 LDA를 사용하여 최적의 특징을 추출한 경우가 높은 검출율을 보였다.

표 3.5 제안한 알고리즘을 이용한 정상심전도 검출 결과 (단위 : %)

데이터 종류	OFS			PCA			LDA		
	ST	SF.	AC	ST	SF	AC	ST	SF	AC
101	100	100	99.88	100	100	99.88	100	100	100
103	100	100	99.90	100	100	99.90	100	100	99.95
106	100	97.71	99.16	99.27	95.55	98.31	100	100	100
112	100	100	99.92	100	100	99.92	100	100	100
113	99.77	-	99.78	100	-	100	99.21	-	99.22
114	97.41	94.34	97.32	99.72	86.79	99.36	99.61	96.23	99.52
115	100	-	100	100	-	100	100	-	100
116	100	98.48	99.92	100	98.49	99.92	100	100	100
117	100	100	100	100	100	99.93	100	100	100
119	100	100	100	100	100	100	100	100	100
121	99.24	100	99.24	99.24	50	99.19	99.03	100	99.03
122	100	-	100	100	-	100	100	-	100
123	100	100	100	100	100	100	99.60	100	99.60
200	99.88	92.15	97.33	99.94	94.02	97.99	99.65	96.86	98.73
202	99.96	91.48	97.42	100	93.44	98.09	99.32	98.03	98.95
205	99.92	95.95	99.81	99.96	95.95	99.85	99.88	97.65	99.85
208	98.79	100	99.26	99.30	100	99.57	98.16	100	98.87
213	100	89.34	99.09	99.92	89.75	99.06	99.88	93.03	99.30
215	95.42	86.67	94.99	98.21	86.06	96.63	99.81	98.79	99.76
223	99.61	94.7	97.27	99.22	94.98	94.56	99.37	94.29	98.30
228	98.63	99.17	98.73	98.45	98.62	98.48	97.63	99.17	97.89
233	99.67	82.30	92.23	99.01	89.26	93.64	99.37	100	99.54
234	100	-	100	99.92	-	99.92	100	-	100
평균	99.51	96.06	98.74	99.67	93.65	98.86	99.61	98.75	99.52

\*ST=sensitivity, SF=specificity, AC=accuracy

\*OFS=original features, PCA=principal component analysis, LDA=Linear discriminant analysis

### 3.4.4 성능 평가

표 3.6 세 가지 알고리즘 검출 결과 비교 (단위 : %)

알고리즘	신경망	퍼지 추론	SVM
OFS	97.72	95.12	96.10
PCA	97.09	93.86	97.39
LDA	98.28	96.18	99.29

\*OFS=original features, PCA=principal component analysis, LDA=Linear discriminant analysis

표 3.6은 상기에서 제시한 표 3.3, 3.4, 3.5의 결과를 식 (3.4)에 적용하여 획득한 평균 검출율을 나타낸 것이다.

$$\text{평균검출율} (\%) = \frac{(\text{sensitivity} + \text{specificity} + \text{accuracy})}{3} \quad (3.4)$$

LDA를 적용하여 추출한 특징들을 입력벡터로 사용한 경우는 LDA를 적용하지 않은 것과 PCA를 적용한 것에 비해 평균적으로 0.97%, 1.8% 향상된 검출 결과를 보였다. LDA를 적용하지 않고 모든 특징점을 사용한 경우는 PCA를 적용하여 특징을 추출한 것에 비해 0.83% 높은 검출 성능을 보였다. PCA는 어떤 클래스를 잘 축약해서 표현하는 데는 유용하나 클래스 간을 잘 분리해서 나타내지는 못하는 단점과 클래스와의 관계를 고려하지 못하고 전체 데이터의 분산을 최대로 하는 방향으로만 선형 투사하기 때문에 클래스 간의 겹침이 발생할 수 있다. SVM 다원 분류기를 사용한 경우 검출 성능은 신경망, 퍼지 추론에 비해 0.48%, 2.99% 우수한 검출 성능을 보였으며, 신경망의 경우 퍼지 추론보다 높은 검출 성능을 나타냈다. 퍼지추론 시스템의 경우 세 알고리즘 중에서 평균적으로 가장 낮은 검출 성능을 보였는데, 이는 퍼지 추론 시스템의 멤버십 함수 설정시 발생할 수 있는 오류의 영향에 의한 것이다. SVM 다원 분류기는 많은 학습 데이터를 필요로 하

는 신경망에 비해 적은 학습 데이터만으로도 우수한 분류 성능을 보이는 장점이 있다. 표 3.7은 본 연구에서 사용한 신경망과 SVM 다원 분류기에서 사용한 학습 데이터의 양을 나타낸 것으로서 신경망의 경우 SVM의 약 6.4배 정도 많은 학습 데이터를 사용하였다.

표 3.7 신경망과 제안한 알고리즘의 학습 데이터 수 비교

	신경망	제안한 알고리즘
학습 데이터 수	26212 (비트)	4135 (비트)

본 논문에서 제안한 알고리즘의 성능 평가를 통해 알고리즘의 검출 성능의 우수성과 효율성을 확인하였다. LDA를 적용하여 추출한 최적의 입력벡터는 기존의 많은 양의 입력벡터를 사용하였을 경우보다 우수한 검출 성능을 보였으며, 적은 수의 입력벡터는 SVM 다원 분류기를 학습 시키는 과정에서 학습속도를 단축시키는 효과가 있었다. 또한 SVM 다원 분류기는 기존의 다른 분류 알고리즘에 비해 우수한 분류 성능을 보였으며, 적은 양의 학습 데이터만으로도 높은 검출 성능을 보이는 것을 확인하였다.

### 3.5 제안한 알고리즘을 이용한 심실세동 검출

본 논문에서는 제안한 알고리즘을 심전도 데이터에 적용하여 심실세동 및 다양한 부정맥을 검출하였다. 그림 3.2는 심실세동 검출 블록도로서, MIT-BIH 부정맥 데이터베이스 중 VF 등 다양한 부정맥이 포함되어 있는 22개의 레코드와 Creighton University Ventricular Tachyarrhythmia 데이터베이스, MIT-BIH Malignant Ventricular Arrhythmia 데이터베이스 중 VF가 포함된 각각 10개의 레코드에서 총 85630 비트의 심전도 신호를 사용하였다.

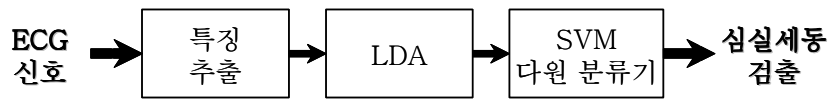


그림 3.2 심실세동 검출 블록도

표 3.8은 SVM 다원 분류기를 이용하여 심실세동을 검출한 결과로서 정상심전도와 심실세동은 평균적으로 98.39%, 99.88%의 우수한 검출 성능을 나타냈다. 다른 부정맥들은 정상심전도와 심실세동에 비해 평균적으로 낮은 검출율을 보였다. 이는 본 논문에서 고려한 부정맥들은 정상심전도와 심실세동에 비해 데이터 양이 크게 부족하였으며, 본 논문에서 고려하지 않은 RBBB(left bundle branch block), RBBB(tight bundle branch block) 또는 fusion of ventricular and normal beat 등과 같은 비정상과형에 의해 많은 검출 오류를 발생하였다.

표 3.8 제안한 알고리즘을 이용한 심실세동 및 다양한 부정맥의 검출 결과

Type of Rhythm	Sensitivity (%)	Specificity (%)	정확성 (%)	평균(%)
NSR	99.66	96.22	99.31	98.39
PVC	92.16	98.94	98.34	96.48
APC	88.29	99.95	99.27	95.84
VT	91.70	99.63	99.44	96.92
SVT	82.82	100	99.85	94.23
VF	99.75	99.99	99.88	99.88

\*PVC=prematuer ventricular contraction, APC=atrical premature contraction

VT=ventricular tachycardia, SVT=supraventricular tachycardia, VF=ventricular fibrillation

본 논문에서 제안한 알고리즘의 특징들은 QRS 콤플렉스를 기준으로 추출된 것이므로 심실세동 검출 및 다른 부정맥들의 검출 성능을 보다 향상시키기 위해서는 정확한 QRS 콤플렉스와 많은 학습 데이터, 다른 부정맥에 대한 연구가 필요하다.

## 제 4 장 결 론

본 논문에서는 선형판별분석법과 SVM 다원 분류기 기반의 심실세동 검출 알고리즘을 제안하였다. 웨이브렛 변환에 의해 추출된 계수를 이용하여 심전도 신호의 잡음을 제거하고, QRS 콤플렉스를 검출하였다. 검출된 QRS 콤플렉스를 기준으로 구간을 설정 한 후 웨이브렛 변환을 통해 획득한 계수와 RR 간격을 특징벡터로 추출하였다. 특징벡터를 LDA에 적용하여 최적의 입력벡터를 구성하고 이를 SVM 다원 분류기에 적용하여 심실세동을 검출하였다. 제안한 알고리즘의 성능을 평가하기 위해 심실세동 파형이 포함되어 있는 3 종류의 심전도 데이터베이스와 주성분 분석, 신경망, 퍼지추론 시스템을 사용하였으며, 성능 평가를 통해 제안한 알고리즘의 검출 성능의 우수성을 확인하였다. LDA를 적용하여 추출한 최적의 입력벡터는 검출 성능을 향상 시킬 뿐만 아니라 SVM 다원 분류기의 학습속도를 단축시키는 효과가 있었다. SVM 다원 분류기는 신경망과 퍼지추론 시스템에 비해 우수한 분류 성능을 보였으며, 적은 양의 학습 데이터로 높은 검출 성능을 보이는 것을 확인 할 수 있었다. 제안된 알고리즘의 이용한 심실세동 검출율은 Sensitivity(%)는 99.75( $p < 0.05$ ), Specificity(%)는 99.99( $p < 0.05$ ), Accuracy(%)는 99.88( $p < 0.05$ )로 우수한 성능을 보였다.

## 참고 문헌

- [1] 윤형로, "실시간 부정맥 검출을 위한 컴퓨터 시스템에 관한 연구", 연세대학교 박사학위논문, 1985.
- [2] 김준수, "부정맥이란", 삼성의료원 건강소식, 중년이후의 건강, 통권 5호, 1997.
- [3] 성홍모, "시-주파수 분석을 이용한 심실세동시 심전도 분석을 통한 제세동 예측에 관한 연구", 연세대학교 박사학위논문, 2004.
- [4] H. X. Zhang, Y. S. Zhu, Z. M. Wang, "Complexity measure and complexity rate information based detection of ventricular tachycardia and fibrillation", *Medical & Biological Engineering & Computing*, Vol. 38, pp. 553-557, 2000.
- [5] M. I. Ojws, A. H. Abou-Zied, A. M. Youssef, "Study of Features based on Nonlinear Dynamical Modeling in ECG Arrhythmia Detection and Classification", *IEEE Transactions on Biomedical Engineering*, Vol. 49, No. 7, pp. 733-736, 2002.
- [6] N. Srinivasan, M.T. Wong, S.M. Krishnan, "A new Phase Space Analysis Algorithm for Cardiac Arrhythmia Detection", *IEEE Engineering in Medicine and Biology Society*, pp. 82-85, 2003.
- [7] V. X. Afonso, W. J. Tompkins, "Detection Ventricular Fibrillation", *IEEE Engineering in Medicine and Biology Society*, pp. 152-159, 1995.
- [8] K. I. Minami, H. Nakajima, "Real-Time Discrimination of Ventricular Tachyarrhythmia with Fourier-Transform Neural Network", *IEEE Transactions on Biomedical Engineering*, Vol. 46, No. 2, pp. 179-185, 1999.
- [9] T. Sugiura, H. Hirata, Y. Harada, T. Kazui, "Automatic Discrimination of Arrhythmia Waveforms using Fuzzy Logic", *Proceeding of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Vol. 20, No. 1, pp. 108-111, 1998.
- [10] 박광리, "스트레스 심전도의 잡음 제거를 위한 WAF와 WIF의 설계", 연세대

- 학교 박사학위논문, 2000.
- [11] A. M. Martinez, A. C. Kah, "PCA versus LAD", *Patterns Analysis and Machine Intelligence*, IEEE Trans., Vol.23, No.2, pp.228-233, 2001.
- [12] S. R. Gunn, "Support Vector Machines for Classification and Regression", Faculty of Engineering and Applied Science, Technical Report, pp.1-54, 1998.
- [13] M. D. Guyton, J. E. Hall, "Medical Physiology", Saunders, pp.155-166, 2002.
- [14] 이승하, "시-주파수 공간 분석법을 이용한 심실세동의 자동 검출에 관한 연구," 연세대학교 석사학위논문, pp.24-27, 1999.
- [15] Minami, K., Ohkuma, Y., Nakajima, H., Toyoshima, T., "Real-time ventricular arrhythmia detection with Fourier analysis and neural network", *Computers in Cardiology*, pp. 545-548, 1996.
- [16] H. Gholam-Hosseini, H. Nazeran, K. J. Reynolds,, "ECG noise cancellation using digital filters", 2nd International Conference on Bioelectromagnetism, pp.151-152, 1998.
- [17] R. H. Clayton, A. Murray, R. W.F. Campbell, "Comparison of four techniques for recognition of ventricular fibrillation from the surface ECG", *Medical & Biological Engineering & Computing*, pp.111-117, 1993
- [18] 최진호, "초보자를 위한 심진도 가이드 : chapter 5 concept of arrhythmia", 2000.3.9. 제4판, 2000.3.23. 개정판
- [19] 홍은혜, 고병철, 변혜란, "PCA와 LDA를 이용한 실시간 얼굴 검출", 한국정보과학회 학술대회논문집, Vol.29, No.2, pp.538-540, 2002.
- [20] 어상준, "Support Vector Machines를 이용한 문서 정보 기반의 단백질 기능 분류", 서울대학교 석사학위논문, 2003.
- [21] W. Y. Choi, K. Lee, Y. Chung, "Support vector machines for robust speaker verification", *Proc. of the AICSST*, pp.262-267, 2002.
- [22] C. M. Rocco and J. A. Moreno, "Fast Monte Carlo Reliability Evaluation Using



- Support Vector Machine”, Reability Engineering system safety, Vol.76, pp.237-243, 2002.
- [23] 김만선, 이상용, “BP알고리즘과 SVM을 이용한 심전도 신호의 패턴 분류”, 퍼지 및 지능시스템학회 논문지, Vol.14, No. 1, pp.82-87, 2004.
- [24] 최우용, 황병희, 이경희, 반성범, 정용화, 정상화, “에스 브이 엠을 이용한 화자인증 알고리즘의 하드웨어 구현 연구”, 전자공학회논문지, 제 41 권, 제 3호, pp.365-372, 2004.
- [25] B. Scholkopf, C. J. C. Burges, A. J. Smola, “Advances in kernel methods”, The MIT Press, 1999.
- [26] Yang Wang, Yi-Sheng Zhu, Nitish V. Thakor, Yu-Hong Xu, “A short-Time Multifractal Approach for Arrhythmia Detection Based on Fussy Neural Network”, IEEE transactions on Biomedical Engineering, Vol. 48, No. 9, 2001.
- [27] J. H. T. Wang, B. W, “Using FNN for classification of cardiac arrhythmia”, IEEE Conference on Computers, Communications, Control and Power Engineering, Vol. 1, pp. 687-690, 2002.
- [28] L. Y. Shyu, Y. H. Wu, W. Hu, “Using Wavelet Transform and Fuzzy Neural Network for VPC Detection From the Holter ECG”, IEEE Transactions on Biomedical Engineering, Vol. 51, No. 7, pp. 1269-1273, 2004
- [29] Y. K. Jung, W. J. Tompkins, “Detecting and classifying Life-Threatening ECG Ventricular Arrhythmia using wavelet Decomposition”, 2003.
- [30] R. H. Clayton, A. Murray, “Comparison of techniques for time-frequency analysis of the ECG during human ventricular fibrillation”, 1998.
- [31] Xu-Sheng Zhang, Yi-Sheng Zhu, V. Thakor, Zhi-Zhong Wang, “Detecting Ventricular Tachycardia and Fibrillation by Complexity Measure”, 1999.
- [32] S. H. Lee, “Detection of Ventricular Fibrillation using Time-Frequency Analysis”, 1999.

- [33] N. Srinivasan, M.T. Wong, S.M. Krishnan, "A new Phase Space Analysis Algorithm for Cardiac Arrhythmia Detection", 2003.
- [34] "A Short-Time Multifractal Approach for Arrhythmia Detection based on Fuzzy Neural Network"
- [35] X. S. Ahang, Y. S. Zhu, N. V. Thakor, Z. Z. Wang, "Detecting Ventricular Tachycardia and Fibrillation by Complexity Measure", IEEE Transactions on Biomedical Engineering, Vol. 46, No. 5, pp.548-555, 1999.

# ABSTRACT

## A Study on the Detection of the Ventricular Fibrillation based on LDA and SVM

Song, Mi-Hye

Graduate Program in Medical Engineering

The Graduate School

Yonsei University

In this paper, we proposed a ventricular fibrillation detection algorithm based on linear discriminant analysis(LDA) and support vector machine(SVM). This algorithm consists of filtering using wavelet transform, detection of QRS complex and determination of analysis interval, feature extraction, dimensionality reduction using LDA and SVM multi-classifier. We selected RR intervals, the 4th, 5th, 6th and 7th wavelet detail coefficients(D4, D5, D6, D7) as features for classifying ventricular fibrillation. MIT-BIH Arrhythmia database, Creighton University Ventricular Tachyarrhythmia database and MIH-BIH Malignant Ventricular Arrhythmia database were used as test and learning data. To evaluate the performance of the proposed algorithm, we compared the result of the proposed algorithm with that of without LDA, principal component analysis(PCA), fuzzy inference and neural network. The proposed algorithm showed good performance that the classification rate of ventricular fibrillation was sensitivity(%) of 99.75( $p < 0.05$ ), specificity(%) of 99.99( $p < 0.05$ ) and accuracy(%) of 99.07( $p < 0.05$ ).

---

**Key words : Ventricular Fibrillation, Wavelet Transform, Linear Discriminant Analysis, Support Vector Machine**