

# 통계적 모델링 기법을 이용한 연속심음신호의 자동분류에 관한 연구

## Automatic Classification of Continuous Heart Sound Signals Using the Statistical Modeling Approach

김 희 근\*, 정 용 주\*\*

(Hee-Keun Kim\*, Joo-Young Chung\*\*)

\*(주)글로벌테크, \*\*계명대학교 전자공학과

(접수일자: 2007년 3월 19일, 수정일자: 2007년 5월 15일, 채택일자: 2007년 5월 29일)

기존의 심음분류를 위한 연구들은 인공신경망을 이용하여 주로 이루어졌다. 그러나 심음신호의 통계적 특성을 분석한 결과 HMM의 의한 신호모델링이 적합한 것으로 나타났다. 본 연구에서는 다양한 질병을 나타내는 심음신호를 HMM을 이용하여 모델링 하고 인식능력이 심음신호의 클러스터링에 따라서 많이 좌우되는 것을 알 수 있었다. 또한 실제 환경에서의 심음신호는 그 시작과 끝나는 시점이 정해지지 않은 연속신호이다. 따라서 HMM을 이용한 심음분류를 위해서는 연속적인 심음신호로부터 한 사이클의 분할된 심음을 추출할 필요성이 있다. 일반적으로 수동분할은 분할오류를 발생시키며 실시간 심음인식에 적합하지 않으므로 분할과정이 필요치 않은 ergodic형 HMM을 변형하여 사용할 것을 제안하였다. 그리고 제안된 HMM은 연속심음을 이용한 분류실험에서 매우 높은 성능을 보임을 알 수 있었다.

핵심용어: HMM, 심음인식, 신호처리, 자동진단

투고분야: 음향 신호 처리 분야 (1.7)

Conventional research works on the classification of the heart sound signal have been done mainly with the artificial neural networks. But the analysis results on the statistical characteristic of the heart sound signal have shown that the HMM is suitable for modeling the heart sound signal. In this paper, we model the various heart sound signals representing different heart diseases with the HMM and find that the classification rate is much affected by the clustering of the heart sound signal. Also, the heart sound signal acquired in real environments is a continuous signal without any specified starting and ending points of time. Hence, for the classification based on the HMM, the continuous cyclic heart sound signal needs to be manually segmented to obtain isolated cycles of the signal. As the manual segmentation will incur the errors in the segmentation and will not be adequate for real time processing, we propose a variant of the ergodic HMM which does not need segmentation procedures. Simulation results show that the proposed method successfully classifies continuous heart sounds with high accuracy.

Key words: Hidden Markov Model, Heart sound classification, Signal processing, Automatic diagnosis.

ASK subject classification: Acoustic Signal processing (1.7)

### I. 서론

청진기를 통해서 심장의 소리 (심음)를 듣고 심장병을 진단하는 방법은 매우 유용하다. 최근에 들어서는, 초음

파나 MRI를 이용한 첨단진단도 많이 이루어지고 있지만, 청진기를 이용한 기법은 저렴한 비용으로 언제 어디서나 비교적 손쉽게 접근이 가능하다는 장점으로 인하여 여전히 중요한 진단방법이다. 그러나 심음을 듣고서 질병을 진단하는 것은 많은 경험을 필요로 하는 비교적 고도의 숙련도가 요구된다. 따라서 심음에 익숙하지 않은 의사들을 도와주는 자동심음 진단 시스템은 매우 유용하다.

디지털신호처리 기술의 발달과 더불어 artificial neural network (ANN) 기반의 자동전단 기법이 심음신호의 분류에 사용되어 왔으나 [1-5], 최근에 들어서는 hidden Markov model (HMM)을 이용한 새로운 방법들이 우수한 효과를 보이고 있다 [6][7]. 동적이며 비정상적인 (non-stationary) 심음신호의 특성은 HMM이 심음을 모델링 하는데 있어서 매우 적합하게 하며, HMM에 의한 심음신호분류 성능은 ANN에 비해서 매우 우수한 것으로 나타났다 [7]. 이러한 결과는 비정상적인 시계열 신호의 분류에 있어서 HMM이 매우 우수한 성능을 보인 기존의 연구결과들과 일맥상통한다.

심음신호는 그 파형이나 주파수의 특성이 매우 다양한 형태로 나타난다. 예를 들면, 동일한 질병의 원인에 의해서 발생한 심음일지라도 심음신호를 획득하는 신체의 부위의 위치나 환자에 따라서 서로의 특성이 상당히 차이가 나는 경우가 많다. 반대로 서로 다른 질병에 의해서 발생하는 두 가지의 심음이 실제로는 매우 유사하게 들리는 경우도 있다. 이러한 현상은 우리가 심음신호의 분류를 통해서 심장병의 종류를 판단하고자 할 때 많이 혼동하게 되는 이유가 된다. 이것은 또한 자동심음분류기의 성능을 저하 시키는 주요한 원인이 될 것으로 생각된다. 본 연구에서는 질병의 종류와 획득위치에 따른 다양한 종류의 심음신호에 대한 분류성능을 높이기 위해서 심음신호의 클러스터를 변화시키면서 그 분류성능의 변화를 검토하였다.

한편 기존의 HMM을 이용한 심음분류 방법에서는 여러 사이클 (cycle)로 이루어진 연속적인 심음신호로부터 분리된 1 사이클에 해당하는 고립된 심음들을 이용한다 [7]. 연속심음으로부터 분할된 이러한 고립된 심음들은 HMM에 기반한 인식기의 훈련과 인식시에 입력패턴으로 사용된다. 그러나 실제 환경에서 자동심음분류기가 온라인상에서 작동할 경우에는 연속심음을 고립단위의 심음으로 실시간 수동 분할하는 것은 거의 불가능하다. 자동 분할을 위한 몇 가지 방식들이 기존에 제안되었지만, 이러한 방법들에서는 분할 오류가 발생하므로 결국 인식성능의 저하가 발생하는 문제가 있다. 이와 같은 심음의 분할에 대한 문제점을 해결하기 위해서 본 연구에서는 심음의 분류 (인식)를 위해서 ergodic HMM의 변형을 제안한다. Ergodic HMM은 기본적으로 모든 상태들 간의 천이가 가능하므로 분할정보가 없어도 연속심음신호를 인식하기에 매우 효과적이기 때문이다.

## II. 통계적 모델링 방식에 의한 심음분류

패턴인식을 위한 방법은 전통적으로 인공신경망을 많이 이용하였다. 특히, 인공신경망은 문자인식이나 지문인식과 같은 정적인 패턴의 인식에 매우 효과적으로 사용되었다. 그러나 인공신경망은 음성신호와 같이 시간에 따라서 그 형태나 주파수 특성이 변하는 비정상적 (non-stationary)의 패턴의 분류에는 다소 문제가 있다는 것이 알려져 있다. 심음신호는 그 특성이 음성신호와 같이 동적이므로 인공신경망보다는 HMM을 이용하여 모델링하기에 적합하다고 생각된다. HMM은 이종의 통계 모델인데, 상태간의 이동을 나타내는 숨겨진 확률 과정과 각 상태별로 관측데이터를 발생시키는 확률 과정으로 이루어져 있다. HMM을 구성하는 주요소는 상태천이 확률과 출력 밀도함수이다. 상태천이 확률은 HMM의 어느 상태에서 임의의 다른 상태로 천이할 확률을 나타내며 출력 확률밀도함수는 HMM의 각 상태에서 관측데이터를 발생할 확률을 나타낸다. 임의의 상태  $j$ 에서 관측데이터 (벡터)  $O_t$ 에 대한 확률밀도함수  $b_j(O_t)$ 는 가우시안 분포의 혼합 형태로 규정되어지며, 다음식과 같다 [8].

$$b_j(O_t) = \sum_{m=1}^M c_{jm} N(O_t, \mu_{jm}, \Sigma_{jm}) \quad j=1, \dots, N \quad (1)$$

여기서  $M$ 은 혼합성분의 개수,  $N$ 은 상태의 개수를 나타내며,  $c_{jm}$ 은 각 혼합성분에 대한 가중치이고,  $N(O_t, \mu_{jm}, \Sigma_{jm})$ 은 평균값이  $\mu_{jm}$  이고 공분산값이  $\Sigma_{jm}$ 인 가우시안 확률밀도함수 (Gaussian probability density function)이다. 상태간의 천이는 천이확률 값으로 규정되며 상태  $i$ 에서 상태  $j$ 로의 천이확률값은  $a_{ij}$ 로 표시된다. 공분산은 추정의 용이함을 위하여 diagonal 행렬을 사용하였다.

HMM을 이용한 패턴인식 과정의 첫 번째 임무는 관측데이터  $O = \{O_1, O_2, \dots, O_T\}$ 와 HMM 파라미터값  $A = \{c_{jm}, \mu_{jm}, \Sigma_{jm}, a_{ij}, 1 \leq i, j \leq N, 1 \leq m \leq M\}$ 이 주어진 경우에 관측데이터의 발생 확률값  $P(O|A)$ 를 구하는 것이다. 두 번째 임무는 관측데이터  $O = \{O_1, O_2, \dots, O_T\}$ 를 발생시킨 최적의 상태열  $I = \{i_1, i_2, \dots, i_T\}$ 을 찾는 것이다. 이 첫 번째와 두 번째 문제는 HMM을 이용한 분류 (인식)과정에서 발생하는 문제이며 HMM 네트워크상에 Viterbi 디코딩 기법을 적용하여 해결한다. 마지막 임무는 주어진 관측데이터에 대하여  $P(O|A)$ 이 최대가 되도록

록  $\lambda$  값을 추정하는 것이다. 이 문제는 HMM모델의 훈련 과정에서 나타나며 Baum-Welch 알고리즘을 이용하여 해결이 가능하다.

HMM은 그 구조에 따라서 여러 가지 형태가 존재하는데, 음성신호와 같이 신호의 특성이 시간에 따라서 변하는 경우에는 자기천이 외에 상태간의 천이가 왼쪽에서 오른쪽으로만 가능한 left-to-right HMM이 가장 많이 사용된다. 같은 맥락에서 심음신호도 left-to-right HMM이 가장 적합한 것으로 보인다. 4 state (상태)의 left-to-right HMM이 그림 1에 나타나 있다. 여기서 HMM은 1 사이클 (cycle)의 심음을 모델링 하고 있으며 HMM의 각 상태는 심음의 각 요소들 (S1, S2, systole 그리고 diastole)을 나타낸다 [3][7].

HMM의 상태의 개수는 보통 모델링되는 신호의 특성에 의해서 결정되어 진다. 그림 1에서 각각의 상태는 하나의 심음신호 성분에 해당하는데, 이것은 심음신호의 각 성분은 그 주파수 특성이 동질성을 지녔다는 가정에서 비롯된다.

기존의 연구에서는 심음신호의 모델링을 위해서는 상태의 개수가 4개면 충분한 것으로 나타났지만, 심음신호의 다양성에 비추어 보면 상태의 개수에 대해서도 보다 많은 고찰이 필요한 것으로 보인다. 그림2에서는 일반적인 형태로부터 다소 변이가 발생한 심음신호들을 보여주고 있는데, 심음신호의 diastole 이나 systole 부분에 주변과 특성이 다른 murmur 나 click 성분이 존재하거나 S1 부분이 split 되는 현상들이 나타나는 것을 알 수 있다. 이와 같은 경우에는 diastole 이나 systole 그리고 split 된 S1 부분을 기존의 연구처럼 하나의 상태로 두는 것 보다는 2~3 정도의 상태로 구분하는 것이 더 바람직할 수도 있다. 본 연구에서는 이와 같이 다양한 특성과

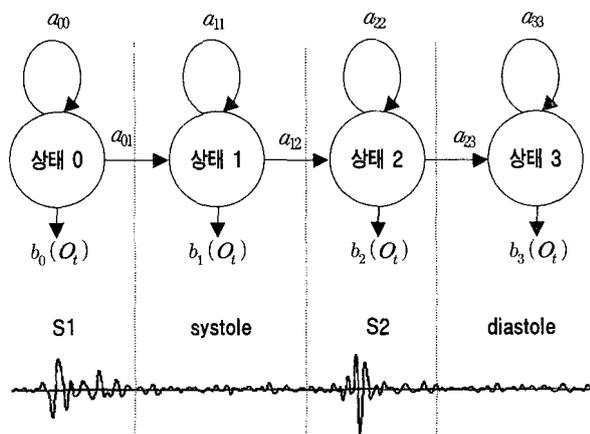


그림 1. 한 주기의 심음신호에 대한 left-to-right형 HMM  
Fig. 1. A left-right type HMM for a cycle of the heart sound signal.

형태를 갖는 심음신호의 변이를 모델링하기 위해서 기존의 연구에서 주로 사용하던 4 state의 HMM 외에도 상태의 수를 가변적으로 설정한 HMM 모델을 사용하여 인식 성능의 향상을 꾀하였다.

또한 그림1과 같은 left-to-right HMM은 고립심음을 모델링하기에는 적합하지만 연속적인 심음의 경우에는 문제가 발생한다. Left-to-right HMM의 경우에는 모델링하고자 하는 심음의 첫 번째 성분이 S1 이라고 가정하므로, 우리는 left-to-right HMM을 사용하기 위해서는 연속심음을 분할하여 그 첫 번째 성분이 반드시 S1이 되도록 해 주어야 한다. 그러나 이러한 방식은 off-line 상에서의 HMM의 훈련을 위해서는 사용 가능하지만, 온라인 (on-line)의 인식과정에서는 수동분할이 어렵기 때문에 실제로 사용이 불가능하다. 심음의 자동분할을 위한 방식들이 존재하지만, 이러한 자동분할과정은 반드시 분할오류를 낳게 되고 이것은 심음인식에 있어서 치명적

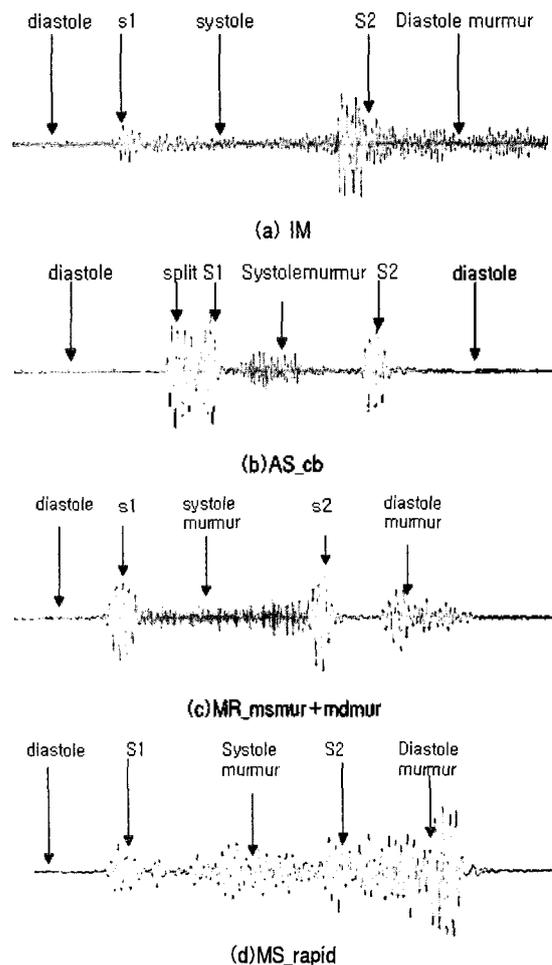


그림 2. 다양한 변이를 가진 심음신호의 예들  
Fig. 2. Examples of heart sound signals with varying signal characteristics.

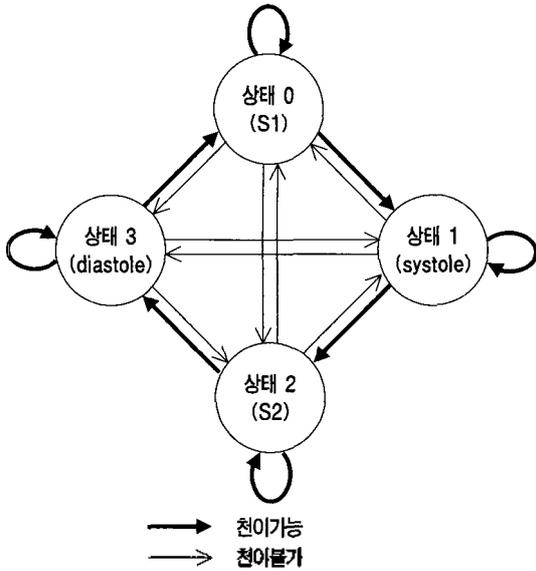


그림 3. 연속심음을 분류를 위해서 제안된 변형된 ergodic HMM  
 Fig. 3. The structure of the proposed modified ergodic HMM for classifying continuous heart sound signals.

인 인식오류를 만들어서 결국은 심음분류 성능을 현저히 떨어뜨리게 된다.

본 논문에서는 그림 3에 나타난 것과 같은 ergodic HMM의 변형을 사용하여 분할정보를 사용하지 않고도 심음신호를 모델링 할 수 있었다. Ergodic HMM은 기본적으로 모든 상태간의 전이가 가능하지만, 본 연구에서는 심음신호의 성분 간의 시간정보를 이용하여 그림 3의 굵은 화살표에 대한 전이만을 허용하였다. 따라서 제안된 형태의 HMM은 순수한 ergodic 형은 아니고 S1→diastole→S2→systole→S1의 순환형 구조를 가진다. 기존의 left-to-right HMM과는 달리 ergodic 형 HMM의 처음과 마지막 상태는 4개중 어떤 것이 되어도 상관없다. 따라서 ergodic HMM를 이용하면 심음신호의 지속시간이나 처음과 마지막 성분의 종류와 상관없이 연속심음을 모델링 할 수 있게 된다.

### III. 분류방법

그림 4에서는 연속심음신호에 대한 훈련과 인식 과정이 나타나 있다. 인식사와 달리 훈련과정에서는 수동으로 분할된 고립 심음이 left-to-right HMM을 모델링하기 위한 데이터로 사용된다. Left-to-right HMM의 파라미터값은 EM (Expectation-Maximization) 방식에 기반을 둔 Baum-Welch 알고리즘을 이용하여 반복적으로 추정

된다 [10]. 훈련된 HMM 파라미터는 인식시에 사용될 ergodic 형 HMM를 구성하는데 이용된다.

그림 5는 Baum-Welch 알고리즘에 대한 흐름도를 보다 상세히 보여주고 있다. 특징벡터를 이용하여 HMM 파라미터를 추정하기 위해서는 파라미터 값에 대한 초기치를 먼저 추정해야 된다. 효과적인 초기값을 추정하기 위해서는 전체 훈련데이터를 각 상태별로 균등 분할하여 각 상태별 확률분포에 대한 평균값과 분산 값을 정한다. 이러한 초기값이 설정되면, Maximum likelihood (ML) 기준에 따라서 파라미터값을 재 추정하게 된다. 일반적으로 재추정과정은 새로운 파라미터 추정값인  $\hat{\lambda}$ 에 대한  $P(O|\hat{\lambda})$  값이 그 이전의  $P(O|\lambda)$  보다 어느 정도 이상 증가하는 동안 계속되거나 반복횟수가 어느 정도 회수가 될 때 까지 계속된다.

인식과정은 주어진 입력 특징 벡터열  $O = \{O_1, O_2, \dots, O_N\}$ 에 대해서 최대의  $P(O|\lambda)$  값을 나타내는 HMM이 어느 클래스에 해당하는지를 찾아내는 것이다.  $P(O|\lambda)$ 를 계산하기 위해서는 forward-backward 알고리즘을 이용해야 되지만, 이 방식은 연산량이 많기 때문에  $P(O|\lambda)$ 의 로그 값에 대한 근사치를 간편하게 얻을 수 있는 Viterbi 알고리즘을 사용한다 [8]. 특징 추출을 위해서는 fast Fourier transform (FFT)을 이용한 필터뱅크 (filter bank) 출력을 사용하였다. 이는 기존의 연구에서 필터뱅크 출력이 MFCC (Mel-frequency cepstral coefficient)보다 우수한 것으로 나타났기 때문이다 [7]. 특징을 얻는 구체적인 과정은 다음과 같다.

먼저, 심음신호  $s(n)$ 을 길이가 N인 Hamming 윈도우 함수  $w(n)$ 에 곱하여 신호  $x(n)$ 을 얻는다.

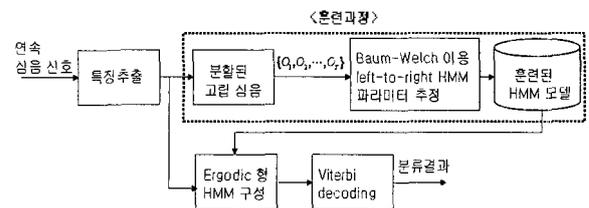


그림 4. 연속심음신호의 분류절차  
 Fig. 4. Classification procedure of a continuous heart sound signal.

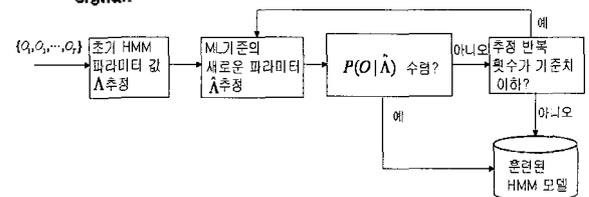


그림 5. Baum-Welch 알고리즘을 이용한 훈련의 흐름도  
 Fig. 5. The flowchart of training based on the Baum-Welch algorithm.

표 1. 분류실험에 이용한 심음신호의 종류

Table 1. The various heart sound signals used in the present classification experiment.

set A		set B					
심음의 종류	데이터 개수	심음의 종류	데이터 개수	심음의 종류	데이터 개수	심음의 종류	데이터 개수
normal sound	15	AR_mur	16	AS_msmur	25	MR_hsmur_apex	14
innocent murmur(IM)	14	AR_ejection	16	AS_msmur_apex	19	MR_msmur+mdmur	14
Aortic Regurgitation(AR)	14	AR_sysmur	15	AS_prebeat	10	MR_hsmur_absent	18
Aortic Stenosis(AS)	18	AR+AS	20	AS_subval	10	MS_presys	19
Coarctation of the Aorta(CA)	20	CA_apex	20	AS_supval	18	MS+PH_lb	12
Mitral Regurgitation(MR)	21	TR+PH	42	AS_cb	12	MS_rapid	12
Mitral Stenosis(MS)	14	MVP_msc	15	AS_cb+AR	15	AF_MS	19
Mitral Valve Prolapse(MVP)	13	MVP_ismur	26	IM_2lis	14	TR_hsmur	17
Tricuspid Regurgitation(TR)	20	MVP_multi	17	IM_musical_4lis	14	VSD	13
Ventricular Septal Defect(VSD)	10	MVP_mslick	14	MR_hsmur	13	normal sound	15
TOTAL	159						504

$$x(n) = s(n) w(n) \tag{2}$$

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right), 0 \leq n \leq N-1$$

$$= 0, \text{ otherwise} \tag{3}$$

그리고  $x(n)$  에 대해서 N-포인트 FFT (Fast Fourier Transform)을 취하여  $X_k$  값을 얻는다.

$$X_k = \sum_{n=0}^{N-1} x(n) e^{-j\frac{2\pi}{N}nk}, 0 \leq k \leq N-1 \tag{4}$$

이때, 필터뱅크출력  $b_j, j=1, \dots, B$ 를 얻기 위해서 우리는 전체 주파수 범위 ( $k=1, \dots, N$ )을 mel-scale [11]에 맞추어 B개로 비균등 분할한 후, j 번째 영역에 할당된  $|X(k)|$  값들을 더해 주게 된다. 이와 같이 얻은  $b_j, j=1, \dots, B$  값들은 필터뱅크 출력으로 사용하게 되며, B는 필터뱅크 특징의 차원을 나타내면 본 논문에서 18을 사용하였다.

#### IV. 실험결과 및 토의

인식실험을 위해서 사용된 심음신호는 의사들의 청진 훈련을 위해서 제작된 상업용 CD로부터 추출되었다 [9]. 여기에 포함된 심음신호는 임상과정에서 획득되었으며 다수의 환자로부터 발생한 다양한 질병을 나타낸다. 원래의 CD상의 심음신호는 16KHz, 16bit로 변환되었다. 본 연구의 인식실험에 사용된 심음데이터는 2가지 set으로 이루어져 있다. 첫 번째 set에 포함된 심음데이터 (set A)는 비교적 그 형태가 단순하며 기본적인 심음신호의 구조인 S1->systole->S2->diastole 순서를 정확히 따르는 신호들이다. Set A에 해당하는 심음신호는 질병별로 레이블링 되었으며 정상적인 심음신호 (normal sound)를

비롯한 10개의 클래스 (class)에 해당하는 159개의 심음 신호가 있다. 두 번째 set에 해당하는 심음데이터 (set B)는 set A에 해당하는 심음신호에 비해서 그 형태가 다소 복잡하고 변화가 심한 것들이다. 예를 들어 동일한 질병이 원인으로 발생하는 심음신호의 경우에도 한 가지 형태의 심음신호가 존재하는 것이 아니라 환자의 상태나 질병이 심한정도 그리고 신호의 획득위치에 따라서 그 형태가 상이하게 나타나는데 set B에서는 이러한 다양한 심음신호를 포함함으로써 좀 더 실제적인 연구결과를 얻을 수 있도록 하였다. 그림6에는 MVP 질환에 해당하는 여러 가지 형태의 심음신호가 나타나 있는데 신호의 특성이 서로 상당히 다를 수 있다. set B에는 이러한 다양성을 포함하는 전체 504개의 심음신호가 포함되어 있으며 전체 30개의 클래스가 존재한다. Set A 와 set B에 포함된 심음신호들에 대한 보다 자세한 클래스의 종류와 데이터 개수는 표1에 나타나 있다. 표1에 나타난 클래스의 이름은 질병의 이름을 근간으로 하되 변이의 원인을 참조하여 임의로 작성되었다.

각각의 클래스에 해당되는 심음신호들을 이용하여 훈련과정을 통하여 클래스 당 한 개씩의 left-to-right HMM과 이를 이용한 ergodic 형 HMM이 구성되었다. 또한, 인식실험 시에 부족한 데이터 문제를 해결하기 위해서 Jack-knifing 방식을 사용하였다. 즉, HMM을 훈련할 경우에는 인식실험을 위해서 사용될 한 개의 심음신호를 제외하고 모든 심음데이터가 사용된다. 이 과정을 반복하여 모든 심음신호가 인식실험에 사용 될 수 있도록 하는 것이다. 최종적인 인식율은 반복된 실험치의 결과를 종합하여 나타나게 된다.

심음신호는 HMM 모델링을 위해서 프레임단위로 처리된다. 심음신호의 1 사이클의 길이는 일반적으로 500 ms에서 1000 ms 사이가 되며 우리는 프레임의 간격과 길이

표 2. Left-right HMM을 이용한 연속 심음의 고립된 사이클에 대한 분류 실험 결과

Table 2. The classification results on the isolated cycle of the continuous heart sound signal using the left-to-right HMM.

	0-900(Hz)		0-210(Hz)		200-900(Hz)		200-300(Hz)	
	Accuracy(%)	Correct/Total	Accuracy(%)	Correct/Total	Accuracy(%)	Correct/Total	Accuracy(%)	Correct/Total
normal	100	15/15	93.33	14/15	100	15/15	93.33	14/15
IM	100	14/14	100	13/14	100	14/14	92.86	13/14
AR	100	14/14	100	14/14	100	14/14	100	14/14
AS	100	18/18	100	18/18	100	18/18	100	18/18
CA	95	19/20	95	19/20	95	19/20	95	19/20
MR	100	21/21	100	21/21	100	21/21	100	21/21
MS	100	14/14	100	14/14	100	14/14	100	14/14
MVP	100	13/13	92.31	12/13	100	13/13	92.31	12/13
TR	100	20/20	100	20/20	100	20/20	100	20/20
VSD	100	10/10	100	10/10	100	10/10	100	10/10
Average	99.37	158/159	98.74	157/159	99.37	158/159	97.48	155/159

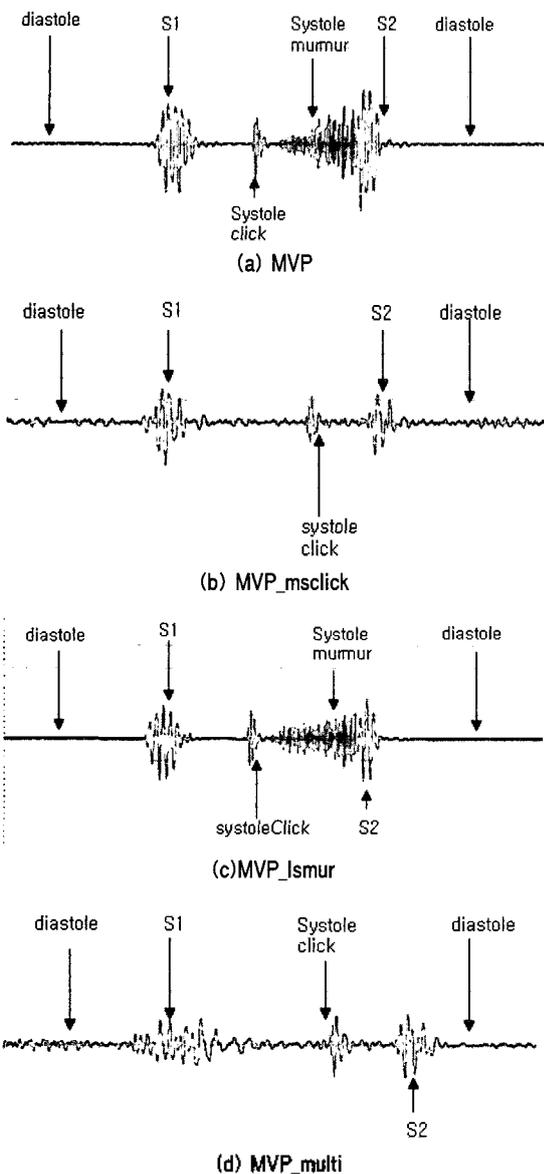


그림 6. MVP와 관련된 다양한 심음신호의 예  
Fig. 6. Examples of the heart sound signals related with the MVP.

를 각각 2.5 ms 와 7.5 ms 로 설정하였다. 프레임의 길이와 간격은 기존의 음성인식에서 사용하는 값을 참고로 하되 심음신호가 음성신호에 비해서 지속시간이 짧다는 것을 고려하였다. 실제로 몇 가지의 프레임 간격과 길이 값을 이용하여 인식실험을 한 결과 제안된 값이 우수한 성능을 보임을 알 수 있었다.

먼저, 기반 인식실험을 위해서 기존의 left-to-right HMM을 이용한 고립 심음인식 실험을 수행하였다.

기존의 연구에서 분석된 심음신호의 주파수 분포를 바탕으로 필터뱅크 출력을 위한 주파수 대역은 0-90Hz, 0-210Hz, 200-900Hz, 그리고 200-300Hz 등으로 정하였고 각각의 주파수 대역에 따른 심음신호 인식의 결과는 표2에 나타나 있다. 여기에서는 set A 심음신호만을 이용하였다.

표 2의 인식결과를 보면 0-900Hz 대역과 200-900Hz 대역이 가장 우수한 성능을 보임을 알 수 있으나 주파수 대역에 따른 인식성능 차이는 그리 크지 않음을 알 수 있다. 전반적으로 주파수 대역의 폭이 넓을수록 인식성능이 향상됨을 알 수 있어서 일반적인 예상과 그리 차이가 나지 않은 것으로 보인다.

표3은 표2에서 사용된 left-to-right HMM을 연속심음신호에 직접 적용했을 때의 결과를 보여준다.

좀 더 자세한 실험결과를 위해서 본 연구에서는 0-420Hz 대역을 추가로 사용하였다. 실험을 위해서 우리는 표2에서 사용된 고립심음신호로부터 인공적으로 연속심음신호를 만들었다. 그 첫 번째 방법에서는 원래의 1 사이클의 고립심음신호를 반복 재생하여 3 사이클의 심음신호를 하나 생성하였다. 이 3 사이클의 심음신호로부터 시작점을 서로 달리하는 8 개의 연속심음신호가 만들어졌다. 이들의 시작점들은 각각 100ms 의 간격을 두도

표 3. Left-right HMM을 이용한 연속 심음에 대한 분류 실험 결과  
Table 3. Classification results on continuous heart sound signals using the left-to-right HMM.

Frequency Range(Hz)	일정간격연속심음		무작위연속심음	
	Accuracy (%)	Correct /Total	Accuracy (%)	Correct /Total
0-900	75.47	960/1272	73.75	1759/2385
0-210	64.15	816/1272	61.25	1461/2385
200-900	73.11	930/1272	72.99	1741/2385
200-300	49.05	624/1272	43.27	1032/2385
0-420	64.77	824/1272	60.04	1432/2385

표 4. Ergodic 형 HMM을 이용한 연속 심음에 대한 분류 실험 결과  
Table 4. Classification results on continuous heart sound signals using the ergodic HMM.

Frequency Range(Hz)	일정간격연속심음		무작위연속심음	
	Accuracy (%)	Correct /Total	Accuracy (%)	Correct /Total
0-900	99.37	1264/1272	99.28	2368/2385
0-210	95.59	1216/1272	94.67	2258/2385
200-900	98.74	1256/1272	98.57	2351/2385
200-300	93.23	1186/1272	92.87	2215/2385
0-420	100	1272/1272	99.07	2363/2385

록 하였다. 이들 8개의 연속심음신호의 길이를 같게 만들기 위해서 시작점 이전에 있었던 3 사이클 심음신호의 샘플들은 새로이 생성되는 연속심음신호의 끝부분에 첨가하였다.

두 번째 방법은 첫 번째 방법과 유사하나 시작점을 랜덤하게 선택 하였으며 원래의 3 사이클 심음신호로부터 15개씩의 연속심음신호가 생성되었다. 표기의 편리함을 위해서 첫 번째 방법으로 생성된 연속심음데이터를 일정 간격연속심음 이라하고 두 번째 방법으로 생성된 연속심음데이터는 무작위연속심음으로 구분했다.

표3에서 보여주는 바와 같이 left-to-right HMM을 연속심음신호를 인식하기 위해서 사용할 경우 인식성능이 매우 저조한 것을 알 수 있었다. 이것은 left-to-right HMM이 연속심음신호의 첫 번째 성분과 마지막 성분이 고정된 것으로 가정하기 때문인 것으로 생각된다.

표4는 표3에서 사용된 연속심음신호를 ergodic 형 HMM을 이용하여 인식한 결과를 보여준다. 고립심음에 대한 인식결과를 나타내는 표2와 비교하면 필터뱅크 출력을 구하기 위한 주파수 대역이 달라짐에 따라서 인식성능이 비교적 예민하게 반응하는 것을 볼 수 있다. 특히, 주파수 대역이 0-210Hz 와 200-300Hz 인 경우에는 표

2의 결과와는 달리 인식성능이 다소 저하되는 것을 볼 수 있다. 이것은 left-to-right HMM에 비해서 ergodic 형 HMM의 경우에는 상태간의 천이가 비교적 자유롭기 때문에, HMM의 상태와 심음성분간의 불일치가 다소 자주 발생하기 때문인 것으로 보인다. 따라서 연속심음신호의 인식을 위해서는 좀 더 넓은 주파수 대역을 사용함으로써 이러한 불일치에 의한 영향을 감소시킬 수 있다고 생각된다. 예상대로 특징 추출시 고려된 주파수 대역이 넓은 경우에는 표2의 고립심음의 분류결과와 비교하여 거의 차이가 없는 것을 알 수 있었다. 연속심음데이터의 생성 방법에 따라서는 큰 인식성능의 차이가 없는 것으로 보인다.

그림 7에는 정상 심음에 대해 ergodic 형 HMM을 사용한 경우에 Viterbi 디코딩의 결과로 얻어지는 상태 시퀀스 (state sequence) 정보와 연속심음사이의 정합정도를 나타내고 있다. 연속심음신호의 동일 성분 내에서도 상태들의 천이가 발생하는 것을 볼 수 있는데, 이것은 ergodic 형 HMM의 자유로운 천이 가능성에 기인한다. S1과 S2의 경우에는 비교적 정합이 정확히 이루어지는 것을 볼 수 있는 반면 diastole 과 systole의 경우에는 서로 간에 혼동이 자주 발생하는 것을 볼 수 있는데 이것은 기본적으로 정상심음의 경우에 있어서 diastole 영역과 systole 영역의 주파수 특성이 매우 유사하기 때문으로 생각된다.

Set B에 해당하는 심음신호는 표1에 따르면 전체 30개의 클래스가 있다. 여기에는 질병종류뿐만 아니라 획득위치 그리고 질병이 심한 정도에 따라서도 각각의 클래스가 정의되어 있다. 예를 들어 AR\_mur, AR\_ejection 그리고 AR\_sysmur들은 모두 AR 질병을 나타내지만 질병의 정도에 따라서 나타나는 신호의 패턴이 서로 상이하어 다른 클래스로 분류되었다. 그러나 심음을 분류하고자 하는 원래의 목적에 비추어 보면 심음신호를 표1에 있는 것처럼

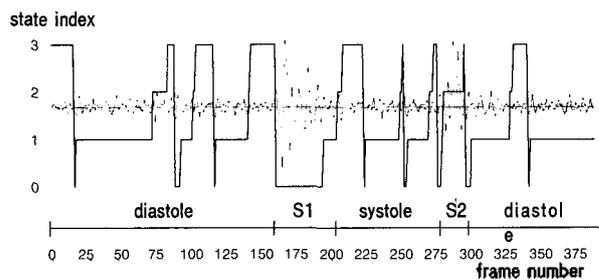


그림 7. 정상 심음에 대한 연속심음신호와 HMM 상태 열 간의 정합  
Fig. 7. The matching between the state sequence and the corresponding continuous heart sound signal in the normal condition.

30개의 클래스로 분류하는 것 보다는 질병의 종류에 따라서 단순하게 분류하는 것이 더 나을 것이라 생각되었으며 또한 클래스의 수를 줄이면 분류성능이 좋아지리라 생각되었다. 따라서 본 연구에서는 표1에 있는 30개의 클래스를 단순히 질병의 종류에 따라서 16개의 클래스로 나누어 보았다. 하지만 그림6에서 보았듯이 같은 종류의 질병에 해당하는 심음신호들 간에도 그 특성이 너무나 상이한 경우가 많으므로 16개의 클래스에 기반한 심음분류는 예상보다 저조한 성능을 보였다. 따라서 우리는 같은 질병을 나타내는 심음신호일지라도 그 음색과 심음신호의 형태에 따라서 서로 다른 클래스로 나누어서 전체 21개의 클래스를 구성하였다. 또한 set B에 속한 심음신호들은 다양한 변이를 일으키는데 그림2에 나타나 있는 것처럼 다소 불규칙한 형태를 띠기도 한다. 이와 같은 심음신호의 변이에 대응하기 위해서 set B에 대한 분류실험에서는 상태의 개수를 클래스별로 다르게 한 경우에 대해서도 분류실험을 수행하였다.

그림8에는 set B에 대한 고립심음 분류결과가 나타났다. 클래스의 개수를 30개, 16개 그리고 21개로 다르게 했을 경우의 분류성능을 보이고 있으며, 또한 HMM 모델당 상태의 개수를 4개로 한 경우, 6개로 한 경우와 함께 HMM 모델별로 상태의 개수를 다르게 한 경우 (variable)에 대해서도 분류성능을 나타내 보았다. 그림8의 결과를 보면 16개의 클래스를 적용한 경우에 가장 분류성능이 저조하고 21개의 클래스를 이용한 경우 가장 높은 분류성능을 보임을 알 수 있다. 이것은 동일한 질병이라 할지라도 그 음색과 형태에 따라서 서로 다른 클래스로 취급하는 것이 옳다는 것을 보여준다. 또한 너무 세세하게 클래스

를 선정하는 것 (30개 클래스) 보다는 신호의 획득위치 등의 차이와 같이 신호의 특성에 그리 크지 않은 영향을 주는 요인들은 클래스 선정시 무시하는 것이 더 나은 성능을 나타냄을 의미한다. 또한 각 HMM모델별의 상태의 개수에 따라서 분류성능의 차이가 발생하는데 전반적으로 각 클래스별 HMM 모델 당 6개의 상태를 사용하는 것이 제일 나은 성능을 나타내고 있다. 이것은 set B에 포함되어 있는 심음신호들이 diastole 이나 systole 에 다양한 변이를 가지고 있으므로 이들을 효과적으로 모델링하기 위해서는 기존에 사용되던 4개의 상태로는 다소 부족함을 의미한다 하겠다. HMM 모델별로 일정한 수의 상태 개수를 할당하지 않고 각 클래스의 특성에 따라서 4~6개 사이의 임의의 상태 수를 할당하는 방식 (variable)은 생각보다는 성능이 좋지 못함을 알 수 있다. 이것은 어느 한 HMM 모델의 상태의 개수가 늘어나면 상대적으로 다른 HMM 모델의 상태의 수가 줄어드는 효과를 발휘하여 상태의 수가 늘어난 클래스에 대해서는 분류성능이 좋아지지만 다른 클래스에 대해서는 성능이 나빠지는 현상이 발생하기 때문인 것으로 판단된다.

그림9에는 set B의 심음신호들을 연속심음으로 확장하여 분류실험을 할 결과를 나타내고 있다. 전반적으로 그림8의 고립심음의 경우에 비해서는 저조한 성능을 보이고 있음을 알 수 있다. 이것은 그림7에서 보였듯이 ergodic HMM의 구조상 상태간의 천이가 보다 자유롭게 일어남으로서 인하여 분류성능의 강인성이 다소 떨어짐으로 인해 발생하는 것이라 생각된다. 향후 연속심음에 대한 성능을 높일 수 있는 방안에 대한 노력이 필요할 것으로 보인다.

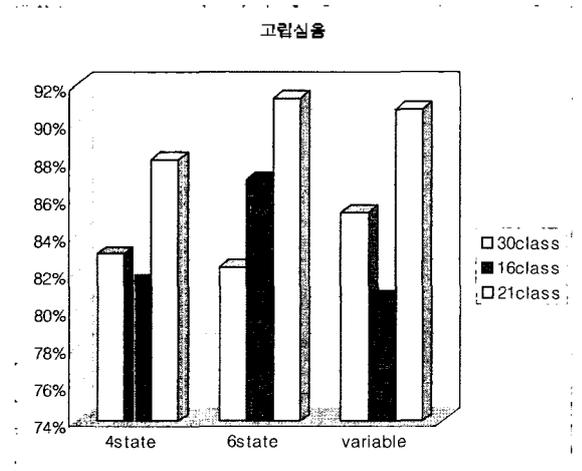


그림 8. set B의 심음데이터에 대한 고립심음 분류결과  
Fig. 8. Classification rates for the isolated heart sound signals in the set B.

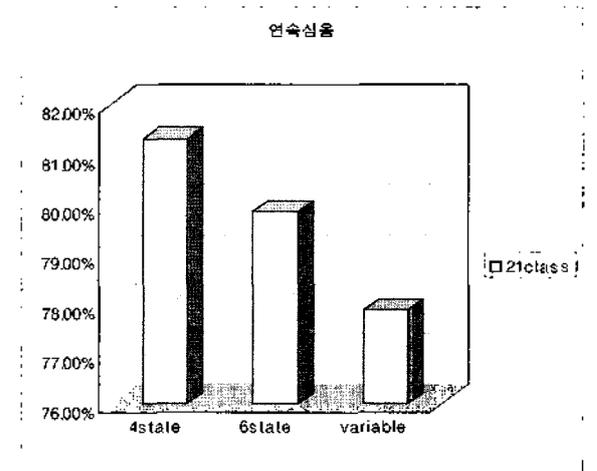


그림 9. set B의 심음데이터에 대한 연속심음 분류결과  
Fig. 9. Classification rates for the continuous heart sound signals in the set B.

## V. 결론

본 연구에서는 통계적 신호 모델링 기법 중에서 효과적인 것으로 알려져 있는 HMM을 이용하여 심음분류를 수행하였고, 특히 ergodic 형 HMM을 이용한 연속심음신호 분류 방법을 제안하였다. 기존의 left-to-right HMM은 고립심음신호의 인식에는 효과적이었으나 분할정보가 주어지지 않은 연속심음신호의 인식에서는 매우 저조한 성능을 보여 주었다. Ergodic 형 HMM은 그 구조적 특성으로 인하여 이러한 분할문제를 해결할 수 있었으며 연속심음신호에 대해서도 우수한 인식성능을 나타내었다. 그러나 심음신호의 특성이 다양하게 변하는 경우에 있어서는 전반적으로 연속심음신호 분류 성능이 저조하게 나타났는데, 이에 대해서는 보다 강인한 인식방법을 통해서 분류성능의 향상을 꾀할 필요가 있다고 생각된다.

## 감사의 글

본 연구는 산업자원부 지방기술혁신사업 (RT104-01-01)지원으로 수행되었음.

## 참고 문헌

1. T.S.Leung, P.R.White, W.B.Collis, E.Brown, A.P.Salmon, "Acoustic diagnosis of heart diseases", Proceedings of the 3rd international conference on acoustical and vibratory surveillance methods and diagnostic techniques, Senlis, France, 389-398, 1998.
2. Cather I, "Neural network assisted cardiac auscultation", Artificial Intelligence in medicine, 7, 53-66, 1995.
3. Bhatikar S, DeGroff C, Mahajan R, "A classifier based on the artificial neural network approach for cardiology auscultation in pediatrics", Artificial Intelligence in medicine, 34 (1) 65-76, May 2005.
4. DeGroff C, Bhatikar S, Hertzberg J, Shandas R, Valdes-Cruz L, Mahajan R, "Artificial neural network-based method of screening heart murmur in children", Circulation, 103, 2711-6, 2001.
5. Gill D, Intrator N, Gavriely N, "A Probabilistic Model for Phonocardiograms Segmentation Based on Homomorphic Filtering", 18-th Biennial International EURASIP Conference Biosignal 2006, Brno, Czech Republic, 87-89, 2006.
6. Ricke A, Provinelli R, Johnson M, "Automatic segmentation of heart sound signals using hidden Markov models", Computer in Cardiology, 9, 953-956, 2005.
7. 김희근, 정용주, "Hidden Markov Model을 이용한 심음분류에 관한 연구", 한국음향학회지, 25 (3), 166-170, 2006.

8. L. R. Rabiner, "A tutorial on hidden Markov model and selected applications in speech recognition", IEEE Proc., 77, 257-286, Feb. 1989.
9. Daniel Mason, *Listening to the heart: A Comprehensive Collection of Heart Sounds and Murmurs*, (F. A. Davis Company, Philadelphia, 2000)
10. L. E. Baum, G. S. T. Petrie and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains", Ann. Math.Statist., 41, 164-171, Jan. 1970.
11. Davis S. B. and Mermelstein P. "Comparison of parametric representation for monosyllabic word recognition in continuously spoken sentences", IEEE Trans. Acoust., Speech, Signal Processing, 28, 357-366, 1980.

---

## 저자 약력

---

• 김희근 (Hee-Keun Kim)

2007년 2월: 계명대학교 전자공학과 (공학석사)  
2007년 3월~현재: (주)글로벌테크



• 정용주 (Yong-Joo Chung)

한국음향학회지 25권 3호 참조