# HEART SOUND DETECTION IN RESPIRATORY SOUND USING HIDDEN MARKOV MODEL

*Hamed Shamsi and I. Yucel Ozbek*

Ataturk University Department of Electrical and Electronics Engineering, Erzurum, Turkey
*e-mail: hamed.shamsi@atauni.edu.tr, iyucelozbek@gmail.com*

## ABSTRACT

In this work, we have investigated the heart sound (HS) detection performance of Hidden Markov Model (HMM) in respiratory sound. Respiratory sound is composed of heart sound and lung sound, and the main frequency components of these two sounds overlap with each other. To detect the locations of heart sound segments in such adverse condition accurately, the proposed method employs following steps. First, the Shannon entropy feature is extracted for robust representation of respiratory signal for different flow rates. Second, the probabilistic models are constructed by training HMM. Finally, the location of heart sound segments are efficiently estimated by the Viterbi decoding algorithm. The experimental results showed that the proposed heart sound detection method outperforms the three well-known heart sound detection methods in the literature. The average false negative rate (FNR) values for the proposed method are $5.4 \pm 2.4$ and $6.3 \pm 1.3$ for both low and medium respiratory flow rate, respectively, which are significantly lower than that of the compared methods in the literature.

## 1. INTRODUCTION

Respiratory sound carries valuable information about the state of the respiratory system including lung and trachea. Hence auscultation of respiratory sound is essential in detection various pulmonary diseases with non-invasive manner. The main difficult in auscultation process is to extract relevant information from respiratory sound. Most of the auscultation case, it is necessary to focus on the lung sound in respiratory sound and this requires removing heart sound signal from respiratory signal. The difficulty of cancellation heart sound from lung sound comes from overlapping regions of the both sound in time and frequency domains. Hence, over the recent years, various techniques of heart sound cancellation methods have been proposed [1, 2, 3] for this purpose. In most of the cancellation methods, the first step is to determine the location of main components of heart sound segment in respiratory sound, and then a heart sound cancellation algorithm is mainly applied to those segments to remove the effect of heart sound. Since the performance of these types of cancellation methods depend upon how accurately the locations of heart sound signals are estimated, the determination of the location of heart sound in respiratory sound has been studied by using different methods such as multi-resolution [1], [4] and [5] variance fractal dimension [6], the adaptive filtering [3] and [7], the non-linear prediction [2] and recently temporal fuzzy clustering technique [8]. Various methods described above use an adaptive threshold to make decision about locations of HS and non-HS segments. The main drawback of threshold based methods is that they are easily affected by rapid signal variations, and due to this reason, the misclassifications occur in such regions frequently. Respiratory sound is made up both heart and lung sound components which contains inhale and exhale sounds of breathing mechanism. To detect heart sound locations in respiratory signal accurately, it is necessary to use more sophisticated method, robust against to signal variation, than basic thresholding methods. Therefore, in this work, we propose heart sound detection method based on Hidden Markov Models (HMM) to improve the performance of the detection process. HMM is powerful probabilistic model,which is broadly used in speech recognition and segmentation applications. The main contribution of this work is to establish a sophisticated (relative to threshold based method) and trainable probabilistic model, which is capable of struggling with the statistical variations of respiratory signal, for the purpose of detection the location of heart sound segments in respiratory signal accurately. The performance of the proposed heart sound detection method is compared with the three well-known HS detection algorithms namely EFT [4], SSA [5] and TFCM [8], and the experimental results show that the proposed method outperforms other three method in terms of false negative rate (FNR), false positive rate (FPR). The structure of this paper is as follows: Section 2 presents the proposed algorithm. The experimental framework and experimental results are given in Section 3. We summarize our main conclusions and future work plan in Section 4.

## 2. PROPOSED METHOD

This work proposes a new framework for heart sound detection in respiratory sound based on the Hidden Markov

Model (HMM) which is a powerful statistical tool for modeling generative sequences. The proposed algorithm has three main stages: feature extraction, description and training of the model parameter, and detection. In the feature extraction stage, the Shannon entropy feature is extracted for each frame of respiratory sound. In the modeling stage, heart and lung sounds are assumed to be generated by HMM and the aim of this stage is to estimate the models parameters using available training data. In the third stage, the detection of heart sound in respiratory sound is done by the Viterbi algorithm based on the estimated models.

### 2.1. Feature Extraction

The various studies in the literature show that Shannon entropy is a powerful features in heart sound detection task in terms of both accuracy and robustness [4], [5] and [8]. Hence, following these studies, the Shannon entropy is selected as feature value in this work to represent respiratory sound efficiently. The feature extraction procedure is described in [4] in detail and summarized as follows. A respiratory sound is first divided into overlapping frames and then, for each frame the probability density function (pdf) is estimated by the non-parametric normal kernel estimation method [9]. Let $P(k,i)$ be discretized version of the estimated pdf, where value $i$ denotes the index of the bin of probability density function and $k$ denotes the frame number of respiratory sound. Then, the Shannon entropy, $z_k$ for $k$th frame is calculated as follows

$$z_k = -\sum_{i=1}^{M} P(k,i)\log_2(P(k,i)) \tag{1}$$

where $M$ is the total number of bin in the pdf.

### 2.2. Topology of the Hidden Markov Model (HMM)

In the proposed HS detection algorithm, the generation of respiratory sound is modeled by ergodic two states HMM. The first state represents the respiratory sound contains heart sound with lung sound, and it is denoted as HS. The second sate is characterized the respiratory sound contains only lung sound, and it is called as non-HS. The topology of the proposed HMM is shown in Fig.1. The a formal characterization of the proposed HMM is explained as follows

- *Initial state probabilities:* The probability that the $i$th state is the initial state, and defined as,

$$\pi_i \triangleq P(s_1 = i) \quad i \in \{1,\ldots,r\} \tag{2}$$

where, $\sum_{i=1}^{r} \pi_i = 1$.

- *State transition matrix:* $s_k \in \mathcal{S} = \{1,\ldots,r\}$ is the finite state variable. In this work, each state represents a respiratory segments: HS and non-HS. The state's Markov chain dynamics are governed by a transition probability matrix $\Pi = [\pi_{ij}]$ defined as

$$\pi_{ij} \triangleq P(s_{k+1} = j | s_k = i), \quad i,j \in \{1,\ldots,r\} \tag{3}$$

where, $\sum_{j=1}^{r} \pi_{ij} = 1, \quad j \in \{1,\ldots,r\}$

- *Observation likelihood:* The observation probability density function defines the likelihood that state $i$ generates the output $z_k$.

$$p(z_k|s_k = i) \triangleq \mathcal{N}(z_k;\mu_i,\Sigma_i), \quad i \in \{1,\ldots,r\} \tag{4}$$

where the notation $\mathcal{N}(z;\mu,\Sigma)$ stands for a Gaussian probability density function for dummy variable $z$ which has a mean $\mu$ and covariance $\Sigma$.

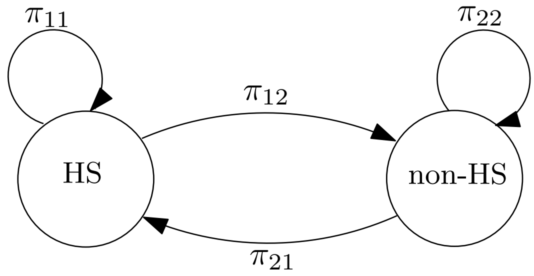


**Fig. 1**. The topology of the proposed ergodic HMM

### 2.3. Learning of the Model Parameters

As explained in previous section, the generation of respiratory signal is modeled by HMM. The parameter set for the $j$th state can be written as $\{\pi_j,\mu_j,\Sigma_j\}$. Considering two states' parameters and the state transition matrix $\Pi = [\pi_{ij}]$, the total parameter set, $\Theta$ can be defined as

$$\Theta \triangleq \left\{\{\pi_{ij},\pi_i,\mu_i,\Sigma_i\}_{i,j=1}^{r}\right\} \tag{5}$$

The total parameter set of the proposed method $\Theta$ is estimated by the maximum likelihood (ML) algorithm using available training data, as described in [10].

### 2.4. Detection of the heart sound locations

Detection of heart sound segments in respiratory signal is performed by estimating the most likely sate sequence using observation sequence and HMM's parameter set $\Theta$. In HMM theory, the estimation of the state sequence is known as decoding problem [10]. The decoding process estimates the underlying unobservable state sequence in optimal manner such that it best explains observation sequence, and it is explained briefly as follows. Let $Z \triangleq [z_1,\ldots,z_N]$ be observation features, $S \triangleq [s_1,\ldots,s_N]$ be related state sequence and $\Theta$ be the corresponding HMM. The state sequence is estimated by solving following optimization problem

$$\hat{S} = \arg\max_{S} p(S|Z, \Theta) \tag{6}$$

where $\hat{S}$ is the single best state sequence out of all possible state sequences. The solving equation (6) is done well-known Viterbi algorithm based on dynamic programming algorithm [10].

### 2.5. Multi-dimensional features

As explained in previous section, one dimensional Shannon entropy feature is extracted from respiratory signal. A one dimensional feature set can be converted into multi-dimensional feature set by a feature augmentation procedure. In this work, the feature augmentation is employed as follows. Let $Z \triangleq [z_1, \ldots, z_N]$ be $N$ single dimensional features and $m$ be the dimension of augmented feature to be generated. The $m$ dimensional augmented feature at $n$th frame, $\mathbf{z}_n^m$ is defined as

$$\mathbf{z}_n^m = [z_{n-l}, \ldots, z_n, \ldots, z_{n+r}]^T. \tag{7}$$

where, $m \triangleq l + r + 1$. The overall multi-dimensional feature set $\mathcal{Z}$ can be defined as follows

$$\mathcal{Z} \triangleq [\mathbf{z}_l^m, \ldots, \mathbf{z}_{N-r}^m]. \tag{8}$$

In this work, the multi-dimensional observation feature set defined in (8) is also employed in training and testing of the proposed HS detection method based HMMs.

## 3. EXPERIMENTAL STUDIES

### 3.1. Database and Database labeling

The database used in this study (originally used in [4] and [11]) is a real respiratory sound which contain five healthy subjects with three females. It was recorded using Siemens EMT25C from third intercostal space anteriorly on the right upper lung lobe. Recording of database is performed for three target flow rates: low, medium and high flow rate. Following the studies [4] and [5], the proposed method and other compared methods are tested under two respiratory flow rates: low and medium. Each part of the respiratory signal has about length of 20s. The database is labeled as HS and non-HS by three trained persons under the control of a cardiologist to measure the performance of the methods. The labeling is performed by using the WaveSurfer toolkit [12], and decisions for boundaries of heart sound are made by examining time waveform, spectrogram, and listening of the sound of data.

### 3.2. Performance Measures

To measure the performance of the methods, we calculated four quantitative results: 1) true-positive (TP) when a HS sound is correctly detected by the algorithms; 2) true-negative (TN) when a non-HS sound is correctly detected by the algorithms; 3) false-negative (FN) when a HS sound is missed; and 4) false-positive (FP) when a non-HS is detected as a HS sound. To evaluate the performance of the proposed localization algorithm, the FN rate (FNR) and the FP rate (FPR) are calculated. The calculation rules of the metrics can be seen as follows

$$FNR \triangleq \frac{FN \times 100}{TP + FN}, \quad FPR \triangleq \frac{FP \times 100}{TN + FP} \tag{9}$$

Moreover, we use the detection error trade-off (DET) curve as a performance measure which visually depicts the performance and performance trade-off of a classification model [13]. The DET curve is created by plotting the FNR values against the FPR values at all threshold values.

### 3.3. Experimental Setup

All experiments for heart sound localization are conducted and all the methods are evaluated on the same database described above. Mean and standard deviation values in all figures and tables are obtained using a set of five subjects. The Shannon entropy features are calculated using a 20 ms window with 10 ms shift. The HMMs used in this work are subject-dependent. For each person two separate HMMs are trained for low and medium respiratory flow rates. They are tested using 5-fold cross-validation. For each fold, four-fifth the data (15s) are used for training and one-fifth (5s) for testing. Cross-validation performance measures are computed as the average of all 5 folds.

### 3.4. Experimental Result and Discussion

This section presents our experimental results. For a better presentation and assessment, we divide this subsection into two subsections as measurement dimension selection and method comparison. The performance of the various dimension of observation in HMM is examined in the first subsection. In the second subsection, the heart sound detection performance of the proposed method is compared in detail with three well-known methods available in the literature.

#### *3.4.1. Experimental Results for Feature Dimension Selection*

In this subsection, various experiments with different dimensional feature set defined in (7) are carried out, and their HS detection performance are compared with each other. Fig.2 shows the Detection Error Rate (DER) performance of various $m$ values for both low and medium respiratory flow rates, respectively. From this figure one can note that the best feature dimension, $m$, is 3 ( $l = 1, r = 1$) for HS detection task among all others and the corresponding DER value is 4.85 and 4.9 for low and medium respiratory flow rates, respectively. Considering the experimental results given in this figure, we

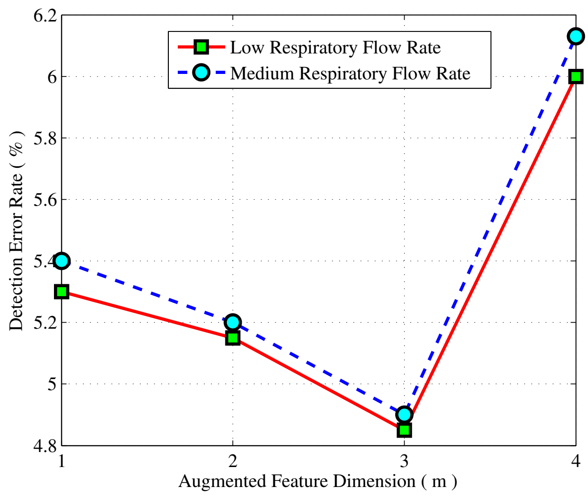


**Fig. 2**. DER results (%)of the proposed method for the various $m$ values at low and medium respiratory flow rates

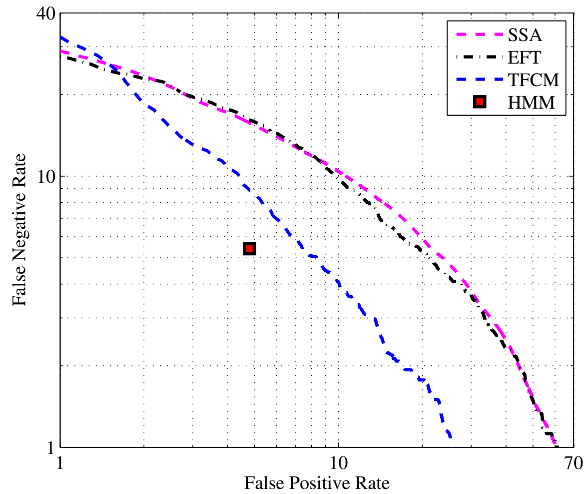


**Fig. 3**. Comparative results based on DET graphs for the TFCM, EFT and SSA methods in low respiratory flow rate. The result of HMM method is given in single point

draw following conclusion. Since it adds temporal information to the detection system, the multi-dimensional observation set improves the performance of the detection task.

### *3.4.2. Experimental Results for Method Comparison*

In this subsection, our aim is to evaluate the heart sound detection performances of the proposed HMM and compare it with three well-known methods: (EFT) [4] ,(SSA) [5] and TFCM [8] in the literature for the same experimental condition. DET curves of the other methods and the performance of the proposed method as a single point are shown in Fig.3 for low respiratory flow rate as well as in Fig.4 for medium respiratory flow rate. From these figures we note that the proposed method gives considerably lower error rates than that of the other methods with the SSA and EFT methods having the highest DET curve values. For a quantitative comparison of the performance of the all methods, we computed the FNR values of each method at same FPR values of HMM, and the corresponding results are given in Table 1 and Table 2. Based on the results of these tables, we note that the proposed HMM method yields the lowest FNR values for low and medium respiratory flow rates. For instance, the HMM method for low respiratory flow rate has a mean FNR value (5.4) lower than that of the TFCM (8.9) , indicating approximately a 39.3% lower FNR value. Considering medium respiratory flow rate (Table 2), we notice that, as expected, errors at the medium flow rate become higher than the ones at the low flow rate. Moreover the HMM has better performances than the other methods at the medium flow rate. It has a mean FNR value (6.3) lower than that (10.4) by the TFCM method, indicating approximately 39.4% lower FNR value.

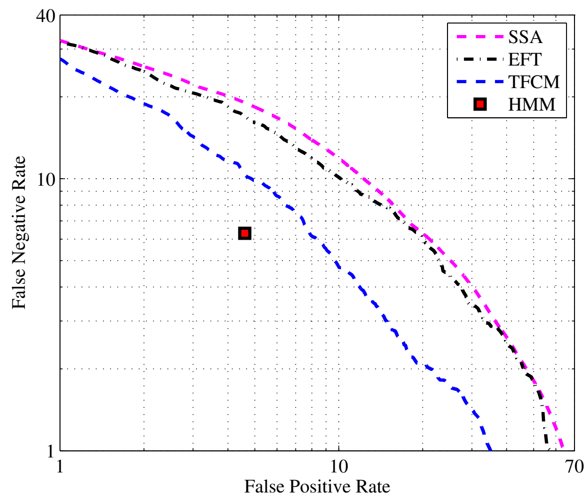


**Fig. 4**. Comparative results based on DET graphs for the TFCM, EFT and SSA methods in medium respiratory flow rate. The result of HMM method is given in single point.

## 4. CONCLUSIONS AND FUTURE WORK

In this paper, we examined the performance of Hidden Markov Model (HMM) in heart sound detection in respiratory signal. In the proposed method, Shannon entropy features are computed for each frame of the respiratory sound and then, using these features HMMs are trained for each segment of the respiratory sound (HS and non-HS) using available training data. The detection part of the proposed method is based on Viterbi algorithm. The experimental results confirm that the performance of the proposed method is better than the other implemented methods (EFT[4], SSA

**Table 1**. Mean and Standard Deviation Results (%) of the HMM and Other Compared Methods at Low Respiratory Flow Rates

| Region | Method | FNR% | FPR% |
|---|---|---|---|
| Low | SSA | $15.7 \pm 10.8$ | 4.8 |
| | EFT | $16.1 \pm 11.3$ | 4.8 |
| | TFCM | $8.9 \pm 4.8$ | 4.8 |
| | HMM | $5.4 \pm 2.4$ | 4.8 |

**Table 2**. Mean and Standard Deviation Results (%) of the HMM and Other Compared Methods at Medium Respiratory Flow Rates

| Region | Method | FNR% | FPR% |
|---|---|---|---|
| Medium | SSA | $19.6 \pm 6.9$ | 4.6 |
| | EFT | $19.0 \pm 6.3$ | 4.6 |
| | TFCM | $10.4 \pm 3.3$ | 4.6 |
| | HMM | $6.3 \pm 1.3$ | 4.6 |

[5] and TFCM [8]) given in the literature. The experimental results show that the proposed method have approximately 39.3% and 39.4% lower FNR than those obtained by TFCM method at low and medium respiratory flow rate, respectively. In this work, the HMMs are trained in a subject-dependent manner. Such models can be used for patients having chronic pulmonary disease such that gathering training data from those subjects is possible. On the other hand, it is also possible that there may be no available training data for a normal subject (not having chronic illness). In such case, it is necessary to use available training data collected from other subjects and generalize the training procedure in a way that it is performed in a subject independent manner with some adaptation procedures. This type of training and heart sound detection procedures constitutes our future work plan.


## Acknowledgment

The authors would like to thank Dr. Zahra Moussavi (from University of Manitoba, Canada) for providing the database used in this study.


### 5. REFERENCES


[1] M. T. Pourazad, Z. Moussavi, and G. Thomas, "Heart sound cancellation from lung sound recordings using time-frequency filtering," *Med. Biol. Eng. Comput.*, vol. 44, no. 3, pp. 216–225, Mar. 2006.

[2] C. Ahlstrom, O. Liljefeldt, P. Hult, and P. Ask, "Heart sound cancellation from lung sound recordings using recurrence time statistics and nonlinear prediction," *IEEE Signal Process. Lett.*, vol. 12, no. 12, pp. 812–815, Dec. 2005.

[3] V. K. Iyer, P. A. Ramamoorthy, H. Fan, and Y. Ploysongsang, "Reduction of heart sounds from lung sounds by adaptive filterng," *IEEE Trans. Biomed. Eng.*, vol. BME-33, no. 12, Dec. 1986.

[4] A. Yadollahi and Z. Moussavi, "A robust method for heart sounds localization using lung sounds entropy," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 3, pp. 497–502, Mar. 2006.

[5] F. Ghaderi, H. R. Mohseni, and S. Sanei, "Localizing heart sounds in respiratory signals using singular spectrum analysis," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 12, pp. 3360–3367, Dec. 2011.

[6] J. Gnitecki and Z. Moussavi, "Variance fractal dimension trajectory as a tool for hear sound localization in lung sounds recordings," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 3, Sep. 2003, pp. 2420–2423.

[7] M. Kompis and E. Russi, "Adaptive heart-noise reduction of lung sounds recorded by a single microphone," in *Proc. 14th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2, Oct. 1992, pp. 691–692.

[8] H. Shamsi and I. Y. Ozbek, "Heart sound localization in chest sound using temporal fuzzy c-means classification," in *Proc. 34th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Sep. 2012, pp. 5286–5289.

[9] B. W. Silverman, *Density Estimation for Statistics and Data Analysis*. Chapman and Hall/CRC, Apr. 1986.

[10] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," in *proceedings of the IEEE, vol 77, no. 2*, February 1989.

[11] J. Gnitecki, Z. Moussavi, and H. Pasterkamp, "Recursive least squares adaptive noise cancellation filtering for heart sound reduction in lung sounds recordings," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 3, Sep. 2003, pp. 2416–2419.

[12] K. Sjlander and J. Beskow, "Wavesurfer - an open source speech tool," 2000.

[13] A. Martin, G. Doddington, T. Kamm, M. Ordowski, and M. Przybocki, "The DET curve in assessment of detection task performance," 1997, pp. 1895–1898.