# Segmentation of Heart Sound Recordings from an Electronic Stethoscope by a Duration Dependent Hidden-Markov Model

SE Schmidt[1], E Toft[1], C Holst-Hansen[2], C Graff[1], JJ Struijk[1]

[1]Department of Health Science and Technology, Aalborg University, Denmark
[2]Department of Cardiology, Aalborg Hospital, Århus University Hospitals, Aalborg, Denmark

## Abstract

*Digital stethoscopes offer new opportunities for computerized analysis of heart sounds. Segmentation of hearts sounds is a fundamental step in the analyzing process. However segmentation of heart sounds recorded with handheld stethoscopes in clinical environments is often complicated by recording and background noise. A Duration-dependent Hidden Markov Model (DHMM) is proposed for robust segmentation of heart sounds. The DHMM model was developed and tested with heart sounds recorded at bedside with a commercially available handheld stethoscope. In a population of 60 patients, the DHMM identified 739 S1 and S2 sounds out of 744 which corresponded to a 99.3% sensitivity. There were seven incorrectly classified sounds which corresponded to a 99.1% positive predictive value. Our results suggest that DHMM could be a suitable method for segmentation of clinically recorded heart sounds.*

## 1. Introduction

Computer-Aided auscultation has the potential to give an accurate and objective interpretation of the heart sounds, which may improve the early diagnostic process of cardiovascular diseases [1]. Several algorithms for automatic analysis of heart sounds have been proposed, including numerous methods for evaluation of murmurs related to heart valve diseases [1, 2, 3], but also methods for detection of weak murmurs related to coronary artery disease [4,5]. The appearance of electronic stethoscopes with digital recording capabilities provides a platform for Computer-Aided auscultation in the clinical setting.

Identification of states in the heart cycle, such as the diastolic and systolic periods is fundamental in almost all heart sound algorithms. The first heart sound (S1) and the second heart sound (S2) are the dominating audible reflections, and indicates the beginning of the systole and the diastole, respectively.

Detection of S1 and S2 is complicated by background noise, variations in heart rhythm, anatomical variations, different recording sites, recording artifacts and pathological heart sounds. Segmentation of heart sounds recorded with handheld stethoscopes in clinical settings are especially challenging due to background noise and friction noise between the stethoscope and the skin. Direct application of earlier algorithms developed for signals recorded with special equipment using fixed microphones in a low noise environment [6-10], is therefore not reasonable.

Hidden Markov Models (HMM) have been used for segmentation of heart sounds [6, 9, 11-13]. The HMM is well suited since it assumes a double stochastic process, consisting of an underlying hidden Markov process, in our case a heart cycle, which generates an observable stochastic output, the heart sound. However the standard the Markov model does not model the duration of the states explicitly, which is a clear limitation since the probability of state transition are highly related to the duration of the states. In the present study we evaluate whether a Duration-dependent Hidden Markov Model (DHMM) [15] can be used for the segmentation of heart sounds recorded with a commercially available handheld stethoscope in a clinical environment

## 2. Methods

### 2.1. Data collection and pre-processing

Heart sounds were recorded at bedside in a multiple patient room, using an electronic stethoscope (3M Littmann E4000, State, USA), from patients referred for coronary arterial angiography at the Department of Cardiology, Aalborg Hospital. The recording site was lateral to the sternum in the 4th intercostal space on the left side. Each recording was 8 seconds long, recorded in 16 bits resolution and sampled with 4000 samples per second. The data analysis and processing were conducted in Matlab. Recordings from 100 subjects were included and divided into a training set (N=40) and a test set

(N=60). Patients with arrhythmia, dominating heart valve disease or extremely noisy recordings were excluded before the test. To allow training and test of the DHMM, the recordings where manually segmented into 4 states: S1, silent systole (siSys) , S2 and silent diastole (siDia), see fig 1. The recordings were band-pass filtered with a 4th order butterworth filter using corner frequencies at 25 and 400 Hz. To emphasize the S1 and S2 sounds a Homomorphic envelogram was created [11]. Inter-patient variation was reduced by normalizing the envelope with the 97% percentile value of the envelope.

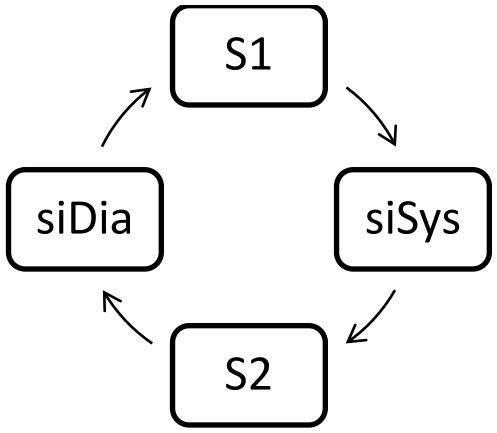


Figure.1. The 4 defined states of the heart cycle.

## 2.2. Hidden Markov models

The heart cycle is defined as a Markov Model with 4 states. It is non-ergodic because the states in the heart cycle only occur in a fixed order, see Fig 1. Central in the standard Markov model are the transition probabilities $a_{ij}$ which defines the probability of state $j$ at next time instances given state $i$ at the current time instance [14].

$$a_{ij} = P(q_{t+1} = S_j | q_t = S_i) \qquad (Eq.1)$$

where $q_t$ is the time-state vector defining the states at time t, S denotes the individual states as $S = \{S_1, S_2 \ldots, S_N\}$, which in the current implantation corresponds to $S=\{siSys, siDia, S1, S2\}$. However, equation 1 is a limited model of the cardiac cycle sine the probability of transition is not independent of the time spend in a given state. For example the probability of transition from diastole to S1 is more likely in the end of the diastolic period than in the beginning. In the duration dependent Markov model the transition probability is supplemented with a duration probability distribution for each state $p_j(d)$. The conceptual difference between HMM and DHMM is illustrated in Fig. 2.

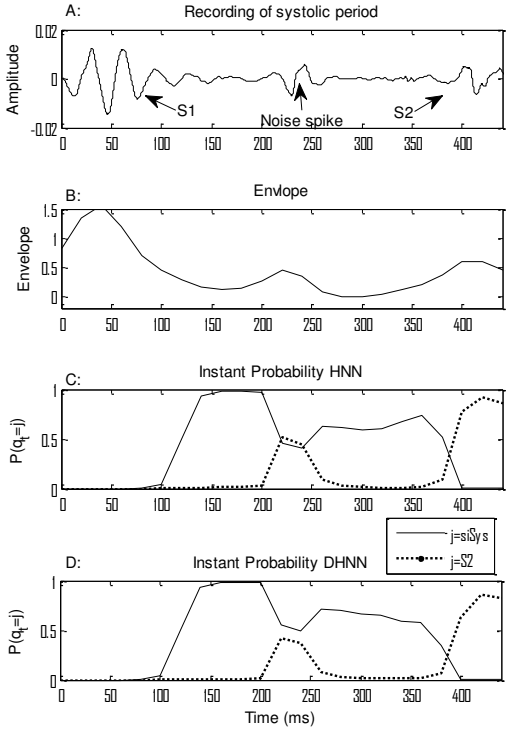


Figure 2. The difference between the HMM and DHMM illustrated over a systolic period. A: a systolic period bounded by S1 and S2 and contaminated by a noise spike. B: Signal envelope. C: the probability that $q_t$ is in state siSys or state S2 calculated with a forward HMM algorithm. The noise spike at 220 ms erroneously makes the probability of S2 higher than siSys. D: Simlar to C but with the probability calculated using DHMM which reduces the effect of the noise spike.

In the HMM and DHMM the true states: $Q=\{q_1,q_2,\ldots.q_t\}$ are not known, the Markov model is hidden, but an observation sequence $O$ is known. The observation is related to the states with the observation probability distribution: $B=\{b_j(O_t)\}$, which defines the probability that state *"j"* generates the output $O_t$. The task in the current application is to estimate the state sequence which is most likely to produce the given observations [14]

$$Q^* = \text{argmax}_Q \, P(Q|O,\lambda) \qquad (Eq.\ 2)$$

where $Q^*$ is the state sequence of all possible state sequences which is most likely to produce O. Lambda ($\lambda$) denotes the model parameters such as the transition matrix, observation probability distribution and in the case of the DHMM, the duration distribution. Solution of equation 2 requires calculation of all combinations of Q, which can be shown to correspond to $T*N^T$ multiplications were T is the number of samples and N

the number of states [14]. Therefore the DHMM is implemented as a forward algorithm calculating an estimated instantaneous probability $\delta_t(j)$ of $q_t$ changing to a new state at the next time instant [14].

*(Eq.3)*

$$\delta_t(j) = P(O_1, O_2, \ldots . O_t, q_t = S_j, q_{t+1} \neq S_j | \lambda)$$

Given the duration $d$ and that the previous state was $i$ at time $t$-$d_j$, the forward calculated probability that $q_t$ is the last time instance in state $j$ is calculated as

$$\delta_t(j,i,d) = \delta_{t-d}(i)\, a_{ij}\; p_j(d) \prod_{s=0}^{d-1} b_j(O_{t-s}) \quad \textit{(Eq.4)}$$

where $p_j(d)$ is the density distribution of the state duration, $\delta_{t-d}(i)$ is the probability that the previous state $i$ ended at $t$-$d$, and $a_{ij}$ is the transition probability when a transition to a new state occurs. The expression to the right of the product sign in (4) is the probability that the state $j$ generated the output observed in the period $t$-$d$ to $t$. From (4) $\delta_t(j)$ is estimated by maximizing according to both the duration $d$ and the previous state $i$ [15].

*(Eq 5)*

$$\delta_t(j) \approx \max_d \max_{i \neq j} \left[ \delta_{t-d}(i)\, a_{ij}\; p_j(d) \prod_{s=0}^{d-1} b_j(O_{t-s}) \right]$$

The $d$ and $i$ which maximized (Eq. 4) is stored and later used in a backtracking algorithm. The q* is found by backtracking though the stored values of $d$ and $i$. See [14, 15] for additional information about the DHMM.

## 2.3. Estimation of parameters

Analysis of the data showed that the inter-subject variation in mean and variance of the envelopes was small since the envelopes were normalized. The same was observed for the distribution of the S1 and S2 durations. Fixed values of theses parameters were therefore estimated from the 40 training recordings. In contrast, the durations of the diastole and systole vary significantly from subject to subject. As a consequence, the distribution parameters for the durations of diastole and systole were estimated individually from each subject, using the autocorrelation of the signal envelope, where the distance from lag null to the first distinct peaks reflects the durations of the systole and diastole.

## 2.4. Test

The DHMM performance was compared to a standard HMM, with both models using the same envelope as input. The ability of both models to locate S1 and S2 correctly was measured by sensitivity and positive predictive value for the 60 test recordings. A sound was correctly located if the center of the detected sound was closer than 60 ms to the center of a similarly predefined sound. All other detected sounds were defined as false positive. Since the models need a short time period in the beginning and end of the recordings to overcome end effects, the first, the second and the last 1.5 seconds of the recordings were excluded from the sensitivity and positive predictive value test. A Wilson interval was used to define the 95 % confidence intervals (CI) of sensitivity and positive predictive values [16].

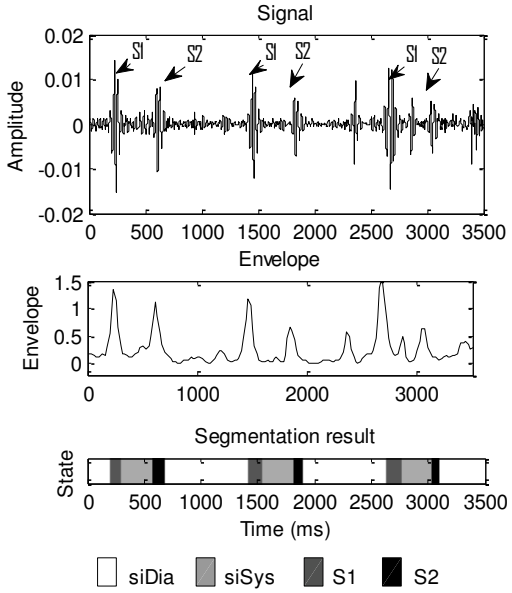


Figure 3. Example of segmentation of heart sounds.

## 3. Results

In the 60 recordings that were included in the test, the DHMM model identified 739 S1 and S2 sounds out of 744 which corresponded to a sensitivity of 99.3% (CI: 98.4-99.7%) see table 1. In 57 recordings out of 60 recordings no sounds were missed.

*Table 1.*

| *Models* | *Sensitivity* | *Positive predictive value* |
|---|---|---|
| DHMM | 99.3% (CI: 98.4-99.7%) | 99.1% (98.1-99.5%) |
| HMM | 59.5% (CI: 56-63%) | 55% (CI: 51.6-58.4%) |

The DHMM misplaced 7 sounds more than 60 ms, which corresponds to 99.1% (98.1-99.5%) positive predictive value. The standard HMM had a sensitivity of

59.5% (CI: 56-63%) and positive predictive value of 55% (CI: 51.6-58.4%), which was considerably lower than the DHMM. A typical error for the HMM was confusion between noise spikes and S1 and S2 sounds.

## 4. Discussion and conclusions

The present study shows that the DHMM could be a well suited method for segmenting heart sounds recorded at bedside with a commercially available electronic stethoscope. Consistent and reliable segmentation was achieved without the use of other signals as the ECG. A high precision was obtained even though recordings were contaminated with background noise and noise from the recording process such as friction noise. The DHMM outperformed the HMM, indicating that the DHMM used considerable information about the duration distribution, which was included in the DHMM. Previous studies applying the HMM have obtained considerably better results than 59.5% sensitivity, but results obtained with fixed microphones in a low noise environment are not directly comparable with the current results.

The present study was based on a selected population of patients referred for coronary arterial angiography, making the method suited for acoustical detection of coronary artery disease. The current study did not include patients with arrhythmias or heart valve disease. Arrhythmias may challenge the DHMM algorithm since the variation of the diastolic and systolic durations will increase. Heart valve disease may cause an amplitude increase in the diastolic or systolic periods. Therefore estimates of the diastolic and systolic amplitude distributions could be erroneous and a more inconsistent separation of S1 and S2 from periods with murmurs would be expected. Further adaption and tests in patients with arrhythmias or heart valve disease, is needed.

The current method is an important step in the development of methods for computer-aided auscultation with electronic stethoscopes in a clinical setting.

## Acknowledgements

The authors thank the personnel at the Department of Cardiology at Aalborg Hospital for their cooperativeness in the data collection process.

## References

[1] Watrous RL. Computer-aided auscultation of the heart: From anatomy and physiology to diagnostic decision support. EEE/EMBS 2006;1:140-3.

[2] Tavel ME, Katz H. Usefulness of a new sound spectral averaging technique to distinguish an innocent systolic murmur from that of aortic stenosis. American Journal of Cardiology 2005;95:902-4.

[3] Kim D, Tavel ME. Assessment of Severity of Aortic Stenosis Through Time-Frequency Analysis of Murmur. Chest 2003;124:1638-44.

[4] Akay M, Semmlow JL, Welkowitz W, Bauer MD, Kostis JB. "Noninvasive detection of coronary stenoses before and after angioplasty using eigenvector methods. IEEE Trans.Biomed.Eng. 1990;37:1095-104.

[5] Schmidt SE, Holst-Hansen C, Graff C, Toft E, Struijk JJ. Detection of Coronary Artery Disease with an Electronic Stethoscope. Computers in Cardiology 2007;34:757−760.

[6] Gamero LG, Watrous R. Detection of the First and Second Heart Sound Using Probabilistic Models. Acta Paediatr 2007;96:1036-42.

[7] Liang H, Lukkarinen S, Hartimo I. Heart sound segmentation algorithm using wavelet decomposition and reconstruction. EEE/EMBS 1997;4:1630-3.

[8] Oskiper T, Watrous R. Detection of the first heart sound using a time-delay neural network. Computers in Cardiology 2002;29:537-40.

[9] Ricke AD, Povinelli RJ, Johnson MT. Automatic segmentation of heart sound signals using hidden Markov models. Computers in Cardiology 2005;32:953-6,.

[10] Wang P, Lim CS, Chauhan S, Foo JY, Anantharaman V. Phonocardiographic signal analysis method using a modified hidden Markov model. Ann.Biomed.Eng 2007; 35:367-74.

[11] Gill D, Gavrieli N, Intrator N. Detection and identification of heart sounds using homomorphic envelogram and self-organizing probabilistic model. Computers in Cardiology 2005;32:957-960.

[12] Lima CS, Cardoso MJ. Phonocardiogram segmentation by using hidden Markov models. BioMed 2007;5:415-8.

[13] Y. Chung. Classification of Continuous Heart Sound Signals Using the Ergodic Hidden Markov Model. Pattern Recognition and Image Analysis 2007;4477:563-70.

[14] Rabiner LR. Tutorial on hidden Markov models and selected applications in speech recognition. Proceedings of the IEEE 1989;77:257-86.

[15] Moore MD, Savic MI. Speech reconstruction using a generalized HSMM (GHSMM). Digital Signal Processing: A Review Journal 2004;14:37-53.

[16] Brown LD, Cai TT, DasGupta A. Interval Estimation for a Binomial Proportion. Statistical Science 2001;16:101-3.

Address for correspondence

Samuel Emil Schmidt
Department of Health Science and Technology,
Aalborg University
Fredrik Bajers Vej 7 E1-207 9220 Aalborg Ø, Denmark
E-mail address: sschmidt@hst.aau.dk