

Why should you model time when you use Markov Models for heart sound analysis*

Jorge Oliveira¹, Theofrastos Mantadelis¹, Miguel Coimbra¹

Abstract—Auscultation is a widely used technique in clinical activity to diagnose heart diseases. However, heart sounds are difficult to interpret because a) of events with very short temporal onset between them (tens of milliseconds) and b) dominant frequencies that are out of the human audible spectrum. In this paper, we propose a model to segment heart sounds using a semi-hidden Markov model instead of a hidden Markov model. Our model in difference from the state-of-the-art hidden Markov models takes in account the temporal constraints that exist in heart cycles. We experimentally confirm that semi-hidden Markov models are able to recreate the “true” continuous state sequence more accurately than hidden Markov models. We achieved a mean error rate per sample of 0.23.

I. INTRODUCTION

The phonocardiogram (PCG) signal is obtained during an auscultation using a traditional or an electronic stethoscope. The PCG contains important information concerning the mechanical activity of the heart valves [1]. The signal processing of a PCG has two main goals: the first one, is to split the PCG into heart cycles. Each heart cycle is composed by the first heart sound (*S1*), the systolic period (*siSys*), the second heart sound (*S2*), and the diastolic period (*siDia*). The second goal is the detection of other sounds such as the third and fourth heart sounds (*S3* and *S4* respectively) as well as heart murmurs that may be associated to cardiac pathologies.

For heart sound segmentation, the methods can be divided depending on which domain they are applied: the time domain (Shannon energy [2]), the frequency domain (homomorphic filter [3]) and the information domain (entropy gradient [4]). For heart sound classification different classifiers have been proposed like Artificial Neural Networks (ANN) [5], k-Nearest Neighbors (k-NN) [6], Support Vector Machines (SVM) [7] and Hidden Markov Models (HMM). HMMs seems to be the ideal statistical model for the highly dynamic and non-stationary nature of the cardiac system, since it is assumed that a sequence of events happens sequentially, and they are highly correlated both temporally and spatially.

Recently, HMMs has shown to be very effective in modeling the heart sound signals: in Gill [8], the signal

is pre-processed and a subset of candidates (peaks) are extracted from homomorphic envelopogram, these candidates are classified using a discrete-time HMM, where the state-distribution is modeled using the time-duration from the preceding candidate to the current one. Chung [9], detected and classified heart sounds using first a left-right HMM model (the first state is assumed to be known) and later an ergodic HMM (the initial state is assigned a probabilistic distribution), the variability in each state is modeled by using multiple mixtures of a Gaussian multivariate distribution.

Schmidt [10], implemented a duration-dependent HMM using the homomorphic filtering envelopogram as observation to the system. This has the advantage (compared to the traditional HMM) that every state duration is explicitly modeled in the state transition matrix. The state duration distribution function is modeled by a Gaussian distribution, where the systolic (*siSys*) and diastolic (*siDia*) duration parameters are estimated through autocorrelation analysis of the homomorphic filtering envelopogram.

Our contributions are: (1) we present an alternative approach for modeling the sojourn time (waiting time) by a semi-hidden Markov Model (HSMM) [11]; (2) we approximate the sojourn time distribution by using a Poisson distribution; (3) we conduct experiments over 32 different models (4 suitable observation features \times 4 dataset fractions \times 2 model types) over a real life dataset of 13 individuals.

The paper is organized as follows: in the second section we present a survey of HMMs and HSMMs. In the third section, we explain the used materials, the extracted features, and the experimental setup. In the fourth section, the results are presented. Finally, we present conclusions and future work.

II. HIDDEN MARKOV MODELS

HMMs are stochastic finite state automata. Where the observation sequence $X = x_1, x_2, \dots, x_n$ depends on the underlying hidden state sequence $S = s_1, s_2, \dots, s_n$ and the unobserved Markov process. The ergodic hidden Markov model assumes that the state transition probability matrix Γ is homogeneous:

$$\gamma_{ij} = Pr(s_t = j | s_{t-1} = i), \quad (1)$$

the γ_{ij} is the probability of being in state j knowing that the previous state was i , is independent of current evaluation time t [9].

The state depended distribution matrix P is assumed to be a continuous Gaussian function:

$$p(x_t | \mu_s, \sigma_s) = \frac{1}{\sigma_s \sqrt{2\pi}} e^{-(x_t - \mu_s)^2 / (2\sigma_s^2)}, \quad (2)$$

*This article is a result of the project NanoSTIMA, NORTE-01-0145-FEDER-000016, supported by Norte Portugal Regional Operational Programme (NORTE 2020), through Portugal 2020 and the European Regional Development Fund.

¹J. Oliveira and M. Coimbra are with the Instituto de Telecomunicações, Faculdade de Ciências da Universidade do Porto; T. Mantadelis is with CRACS & INESC TEC, Faculdade de Ciências da Universidade do Porto. Rua do Campo Alegre 1021/1055, 4169-007 Porto, Portugal (oliveira_jorge@dcc.fc.up.pt, theo.mantadelis@dcc.fc.up.pt, mcoimbra@dcc.fc.up.pt).

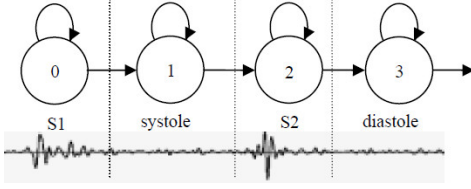


Fig. 1. Four state HMM for a cycle of a normal heart sound signal (adapted from [9]).

with σ_s being the standard deviation and μ_s the mean value in state s .

It is assumed that each state of a HMMs corresponds to an element of the heart sound signal because the signal characteristics in each element are thought to be homogeneous, as it is depicted in Figure 1. For simplicity, this model ignores S3, S4 and murmur sounds.

The likelihood of a state sequence S of a HMM with observation sequence X is:

$$P(X, S, \Theta) = \pi_1 \left\{ \prod_{t=2}^n \gamma_{s_{t-1}s_t} \right\} \prod_{t=1}^n P_{s_t}(x_t), \quad (3)$$

where $\Theta = \{\pi_1, \Gamma, P\}$ denotes the model parameters such as the initial state distribution π_1 , state transition probability matrix Γ and state depended distribution matrix P .

A. Semi-hidden Markov Models

In standard HMMs, the sojourn time (waiting time) is geometrically distributed over all states. This is an unrealistic assumption in heart sound signals, since the state transition probabilities are constantly changing over time. The solution we propose is to model explicitly the sojourn time by using a HSMM.

We first define, D as the sojourn time distribution matrix. In our case we use the Poisson distribution function:

$$d_{s_k}(u_k | \lambda_{s_k}) = \frac{e^{-\lambda_{s_k}} \lambda_{s_k}^{u_k}}{u_k!}, \quad (4)$$

$d_{s_k}(u_k)$ is the probability of spending u_k units of time in the state s_k . Furthermore we define, $d_{s_k}^*(u_k)$ as the survivor function for the sojourn time:

$$d_{s_k}^*(u_k) = \sum_{v \geq u_k} d_{s_k}(v). \quad (5)$$

We also define r as the total number of state transitions that occurred until time n ; and finally, we also define as $N(t)$ as the current state at time t .

The likelihood of a state sequence S of a HSMM is:

$$P(X, S, \Theta) = \pi_1 \left\{ \prod_{k=2}^{r-1} \gamma_{s_{k-1}s_k} d_{s_k}(u_k) \right\} \gamma_{s_{r-1}s_r} d_{s_r}^*(u_r) \cdot \prod_{l=1}^n P_{s_{N(l)}}(x_l), \quad (6)$$

where s_k is the k^{th} visited state and u_k is the sojourn time of the k^{th} state. Therefore, a HSMM is specified by the quadruple $\Theta = \{\pi_1, \Gamma, P, D\}$ [11].

B. Optimizing the Parameters of HMM and HSMM

In order to solve Equations 3 and 6, it is necessary to choose a good parameter initialization. The parameters Θ are estimated using the expectation maximization (EM) algorithm, which assigns posterior probabilities to each component with respect to each observation. The method uses an iterative algorithm that converges to an optimal solution [9]. Unlike in HMMs, in HSMMs we also need to optimize D . For HSMM in the expectation step, we need to calculate the expected number of times η_{iu_k} that the model remains in state i for u_k time steps:

$$\eta_{iu_k} = \sum_{v=1}^{u_k} P(s_{u_k} \neq i, s_{u_k-v} = i | X, \Theta) + \sum_{v=0}^{u_k-1} \sum_{t=1}^n P(s_{t+u_k+1} \neq i, s_{t+u_k-v} = i, s_t \neq i | X, \Theta), \quad (7)$$

Again only for HSMM, in the maximization step, the Poisson parameter λ_{s_k} is calculated by:

$$\lambda_{s_k} = \sum_{v=1}^n \eta_{iv}(v - \beta), \quad (8)$$

where β is all possible shifts parameters, $\beta = 1, \dots, \min(u_k : \eta_{iu_k} > 0)$. The maximization step chooses the β which gives the maximum likelihood of Equation 6.

Finally, we use the Viterbi algorithm to determine the most likely state sequence.

C. Using the Poisson distribution

For our application, the duration probability distribution D is approximated by a Poisson distribution. We chose the Poisson distribution, as the more suitable to model the behavior of PCG signals because:

- 1) We need to count the number of state transitions in a large number of trials (n-sampling size).
- 2) The state transitions in PCG signals are rare events because we sample at a high frequency (our signals have a sampling rate $f = 4kH_z$). Heart sounds frequency f_h lies between 20Hz and 2 kHz and Nyquist rate imposes that the sampling rate f should be $f \geq 2 * f_h$. Attempting to sample in a lower frequency would create aliasing, resulting in a loss of information.
- 3) The successful events are also very unlikely in heart sound signals, because of some physiological time constraints that exist in the cardiac cycle, for example: the cardiac muscle (like any excitable tissue) exhibits a refractory period to re-stimulation. During this time interval normal cardiac impulse cannot re-excite an already excited area of cardiac muscle. The normal refractory period of the ventricle is 0.25 to 0.30 second.

To use a Poisson distribution, we have made an assumption that the outcome trials in different time instances are “weakly” dependent but not necessarily independent.

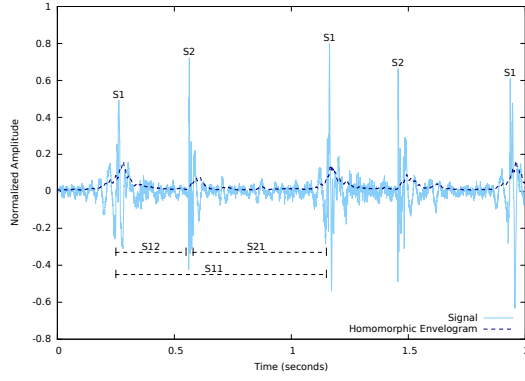


Fig. 2. Signal and the extracted Homomorphic envelopegram.

III. METHODOLOGY

A. Materials

The dataset we use is composed by samples from 13 healthy individuals from six months to 17 years old. The heart sounds have been collected in Real Hospital Português (Recife, Brasil) using a Littmann 3200 stethoscope embedded with the DigiScope technology. The DigiScope technology was developed within the homonymous project to collect, transmit and record heart sounds without interfering with clinical routine [12]. The heart sounds are recorded in the mitral spot, for about 15 seconds at 4000Hz sampling frequency. Two cardiapulmonologists manually annotated the $S1$ and $S2$ states beginning and ending using the Audacity software.

B. Features extraction

The system first filters the original signal using a zero-phase Butterworth bandpass filter of order 10. We use a lower cutoff frequency of 30Hz and a higher cutoff frequency of 460Hz. Then the signal is normalized (mean subtraction) and scaled. From the filtered signal, different envelopegrams are extracted: Shannon energy in the frequency domain; homomorphic filtering; Shannon energy in the time domain; and the entropy gradient. Figure 2 presents the signal and the extracted Homomorphic envelopegram.

Shannon energy in the frequency domain which we compute as in [13], accentuates the pressure differences found across heart valves, which leads to distinct frequency signatures of the valve closing sounds.

Homomorphic filtering, the signal is viewed as a product of slowly varying components (heart sounds) with fast oscillatory components (noise). These fast components are rejected by applying a non-linear transformation and is compute as in [8].

Shannon energy in the time domain which is computed as in [2], is used to emphasize the medium intensity of the signals and attenuate high intensity. This tends to make medium and high intensity signals similar in amplitude.

Finally, entropy gradient envelopegram measures the predictability of the heart sound components in a signal by looking to the total entropy fluctuation in the “expanded region” as the original time series is shifted in a circular motion and is compute as in [4].

C. Experimental Setup

In order to optimize the HMM and HSMM parameters we used the EM algorithm [11] also known as the Baum-Welch algorithm.

In our experimental setup all states have equal starting probabilities (π_1). The Γ parameters are fixed to:

$$\Gamma = \begin{matrix} & \begin{matrix} S1 & siSys & S2 & siDia \end{matrix} \\ \begin{matrix} S1 \\ siSys \\ S2 \\ siDia \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \end{matrix}, \quad (9)$$

because in the normal cardiac system the state sequence $\{S1 \rightarrow siSys \rightarrow S2 \rightarrow diSys \rightarrow S1\}$ is fixed as is shown in Figure 1.

To compute the initial parameters P ($\mu_s, \sigma_s \forall s \in S$) we use an envelopegram segment around the corresponded annotated state s . To compute the initial parameters D ($\lambda_s \forall s \in S$) we use the annotated time lapse between the beginning and the end of the corresponding state s . We use from $\frac{1}{3}^{rd}$ to $\frac{3}{4}^{th}$ of the first part of each subject signal to initialize P, D and use the rest as a test dataset.

Since insufficient annotated heart beats are used for computing the initial parameters, the initial standard deviations σ of the Gaussian distributions are biased. In order to reduce the bias, we widen the Gaussian distribution by multiplying σ by a factor of 10^2 .

We use this methodology because the true parameters diverge from subject to subject. Attempting to generalize over all subjects, leads to inaccurate parameter initialization and the algorithm does not converge to an optimum solution.

The signal features are extracted in Matlab and the experiments are conducted using mhsmm package for R [11]. For more implementation details on the Viterbi or the Baum-Welch algorithm see [11].

IV. RESULTS

The performance of the HSMM and HMM was measured as the model’s capacity to recreate the continuous state sequence annotated by the cardiapulmonologists. The mean error rate per sample is calculated by:

$$\bar{\mu}_E = \frac{\# \text{ False labeled Samples}}{\# \text{ Samples}} \quad (10)$$

A sample in the instance t is labeled falsly, when the predicted state of the sample and the annotated state of the sample are not equal.

As can be seen from the Figure 3, the best results with the standard HMM has a $\bar{\mu}_E = 0.57$, using the Homomorphic filtering as observational input and a training dataset fraction size of $\frac{3}{4}^{th}$. The best HSMM results were achieved using Shannon energy in the frequency domain with a $\bar{\mu}_E = 0.23$, for a training dataset fraction size of $\frac{1}{3}^{rd}$. For all types of observational input, the HSMM performed better than the equivalent HMM.

Figure 4(a) illustrates the true state sequence of the signal. The HMM is not capable to detect the right sequence of events and not even the state duration in each state as it can be seen in Figure 4(b), this might be a consequence of using

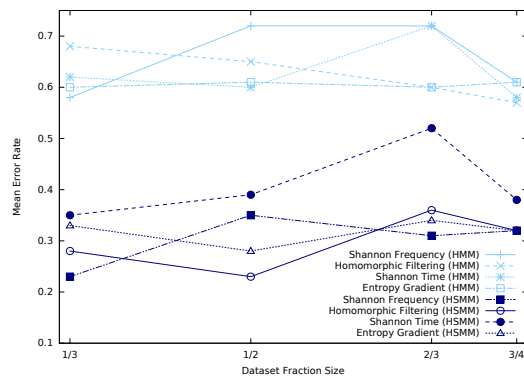


Fig. 3. Mean Error rate per sample for HMM and HSMM.

static state transition matrix. In the HSMM, the assumption that the Markov chain is homogeneous is dropped, instead it is assumed that the state transition matrix is dependent on time (following a Poisson distribution in the present case), and this ultimately leads to a model capable of describing the non-stationary events in the heart sound signal as depicted in Figure 4(c) with more accuracy than the standard HMM.

Finally, we noticed in some case that the starting state was misclassified and as a result the signal has classified reversely by the HSMMs. These signals are properly classified if we set an observed starting state.

V. CONCLUSION AND FUTURE WORK

In this paper, a heart sound classification algorithm is proposed using HMMs or HSMMs, furthermore we used four different type of features as input to the system. These features are very descriptive and sensitive to $S1$ and $S2$ events. Our experiments shows that HSMM outperformed HMM regardless the observational features tested, this suggests that using information concerning the duration probability distribution in each state is a requirement step in modeling heart sound signals. We approximated the duration probability distribution by the Poisson distribution. Furthermore, from all the observation features tested Homomorphic filtering and Shannon energy in the frequency domain showed the best performance for both HMMs and HSMMs.

For future work, we intend to conduct extensive experiments with different distributions in order to approximate the duration probability distribution of the HSMM. Furthermore, we want to conduct experiments where we test the influence of heart rate variability (such as the heart rate of infants and subjects with arrhythmia) in HMMs or HSMMs. Finally, the proposed algorithm can be adapted and applied to different situations, such as the detection of $S3$ and $S4$, murmurs and valves defects.

ACKNOWLEDGMENT

The authors would like to thank the two cardiopulmonologists Christina Oliveira and Bruna Cardoso for their contribution in annotating the heart signals.

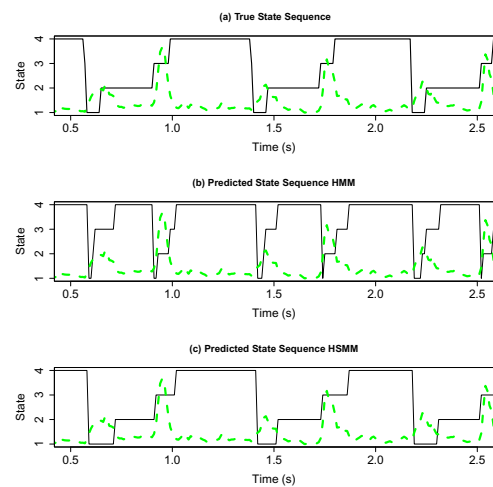


Fig. 4. Classification results of heart sound recordings from a normal subject. The states are labeled from 1 to 4 where: 1 is $S1$; 2 is $siSys$; 3 is $S2$; and 4 is $siDia$. The solid line presents the classification of the signal and the dashed line presents the observation input to the model.

REFERENCES

- [1] J. E. Hall and A. C. Guyton, *Textbook of medical physiology*. Philadelphia, Pa.: Saunders/Elsevier, 12th ed., 2011.
- [2] H. Liang, S. Lakkari, and I. Hartimo, "Heart sound segmentation algorithm based on heart sound envelopment," in *Computers in Cardiology*, pp. 105–108, 1997.
- [3] C. N. Gupta, R. Palaniappan, S. Swaminathan, and S. M. Krishnan, "Neural network classification of homomorphic segmented heart sounds," *Appl. Soft Comput.*, vol. 7, no. 1, pp. 286–297, 2007.
- [4] J. Oliveira, A. C. Castro, and M. Coimbra, "Exploring embedding matrices and the entropy gradient for the segmentation of heart sounds in real noisy environments," in *Engineering in Medicine and Biology Society, IEEE Conference*, vol. 1, pp. 3244–3247, 2014.
- [5] T. Leung, P. White, W. Collis, E. Brown, and A. Salmon, "Classification of heart sounds using time-frequency method and artificial neural networks," in *Engineering in Medicine and Biology Society, IEEE Conference*, vol. 2, pp. 988–991, 2000.
- [6] J. Oliveira, C. Oliveira, B. Cardoso, M. Sultan, and M. Tavares Coimbra, "A multi-spot exploration of the topological structures of the reconstructed phase-space for the detection of cardiac murmurs," in *Engineering in Medicine and Biology Society, IEEE Conference*, pp. 4194–4197, 2015.
- [7] J. Vepa, "Classification of heart murmurs using cepstral features and support vector machines," in *Engineering in Medicine and Biology Society, IEEE Conference*, pp. 2539–2542, 2009.
- [8] D. Gill, N. Gavriel, and N. Intrator, "Detection and identification of heart sounds using homomorphic envelopment and self-organizing probabilistic model," in *Computers in Cardiology*, pp. 957–960, 2005.
- [9] Y.-J. Chung, *Pattern Recognition and Image Analysis, Iberian Conference*, ch. Classification of Continuous Heart Sound Signals Using the Ergodic Hidden Markov Model, pp. 563–570. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007.
- [10] S. Schmidt, E. Toft, C. Holst-Hansen, C. Graff, and J. Struijk, "Segmentation of heart sound recordings from an electronic stethoscope by a duration dependent hidden-markov model," in *Computers in Cardiology*, pp. 345–348, 2008.
- [11] J. O'Connell and S. Højsgaard, "Hidden semi markov models for multiple observation sequences: The mhsmm package for r," *Journal of Statistical Software*, vol. 39, no. 1, pp. 1–22, 2011. <https://cran.r-project.org/web/packages/mhsmm/mhsmm.pdf>.
- [12] D. Pereira, F. Hedayioglu, R. Correia, T. Silva, I. Dutra, F. Almeida, S. Mattos, and M. Coimbra, "DigiScope - Unobtrusive collection and annotating of auscultations in real hospital environments," in *Engineering in Medicine and Biology Society, IEEE Conference*, pp. 1193–1196, 2011.
- [13] D. Kumar, P. Carvalho, M. Antunes, R. P. Paiva, and J. Henriques, "Noise detection during heart sound recording using periodicity signatures," *Physiological Measurement*, vol. 32, no. 5, p. 599, 2011.