

COUPLED HIDDEN MARKOV MODEL FOR AUTOMATIC ECG AND PCG SEGMENTATION

¹Jorge Oliveira, ²Catarina Sousa ¹Miguel.T.Coimbra

¹Instituto de Telecomunicações, Department of Computer Science, Faculty of Science, Porto University
²Faculty of Medicine, Faculty of Science, Porto University

ABSTRACT

Automatic and simultaneous electrocardiogram (ECG) and phonocardiogram (PCG) segmentation is a good example of current challenges when designing multi-channel decision support systems for healthcare. In this paper, we implemented and tested a Montazeri coupled hidden Markov model (CHMM), where two HMM's cooperate to recreate the "true" state sequence. To evaluate its performance, we tested different settings (two fully connected and two partially connected channels) on a real dataset annotated by an expert. The fully connected model achieved 71% of positive predictability (P^+) on the ECG channel and 67% of P^+ on the PCG channel. The partially connected model achieved 90% of P^+ on the ECG channel and 80% of P^+ in the PCG channel. These results validate the potential of our approach for real world multichannel application systems.

Index Terms— Coupled Hidden Markov Models (CHMM), Hidden Markov Models (HMM), Phonocardiogram (PCG), Electrocardiogram (ECG), Heart Sounds.

1. INTRODUCTION

The PCG and the ECG are valuable vital signals that directly encode the electro-mechanical status of the heart (Figure 1).

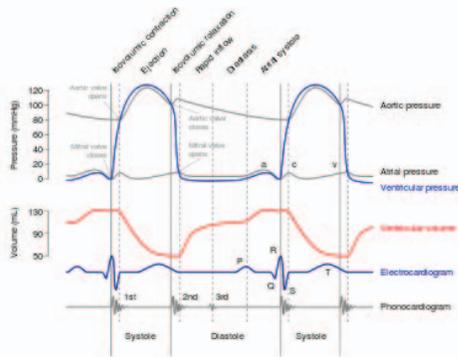


Figure 1: The Wiggers diagram, including the ECG and the PCG at bottom (adapted from [1]).

The ECG records the heart electrical activity via electrodes placed over the skin, mapping "commands" for the myocardium to contract and relax [1]. The PCG, gives us the heart physiological "response" to the aforementioned "commands". Therefore, an electro-mechanical model that simultaneously uses both ECG and PCG channels, can give us a general portrait over the different cardiac phases, and give us valuable information for pathology screening and assessment.

Coupled HMM appears as an interesting model to study the highly dynamic and non-stationarity nature of the cardiac system. It is assumed that the state sequence happens sequentially, and the channels are co-dependent through their past states and observations [2].

CHMM have been proposed by Brand in [3] as a generalization of HMM. He observed that CHMM outperformed HMM when classifying visual tasks (two-handed actions). Montazeri [2] presented a CHMM, where the state inter-dependencies are modeled using a stochastic matrix. They proposed a novel apnea-bradycardia detection method for preterm infants, integrating a phase of multivariate feature extraction from the ECG, and a phase of time-series characterization through the proposed CHMM.

In this paper, our main goal is to segment the PCG channel using a HMM. We do know from past experiments that using a simple HMM on a single channel is not enough, the model is too static and fails to model the dynamic events in the heart sound signal [4]. But what happen if someone disposes not only the PCG channel but also the ECG channel, *how can we use the ECG channel to enhance the PCG segmentation task?* Motivated by current CHMM success in forensic electronic [5], genetics [6], audio-visual speech recognition system [7] and target tracking [8] and more recently apnea-bradycardia detection [2], we implemented and evaluated Montazeri's CHMM capacity in recreating the "true" state sequence on the ECG and PCG coupled channels. These are tested over different settings, initializations and architectures.

The paper is organized as follows: in the second section, the CHMM are explained. In the third sections, the methodology is presented. In the fourth and fifth sections, results are presented and conclusions are withdrawn.

2. CHMM

HMMs are stochastic finite state automata, where an observation sequence $\mathbf{O} = \{o_1, \dots, o_T\}$ depends on the underlying hidden state sequence $\mathbf{Q} = \{q_1, \dots, q_T\}$ and the unobserved Markov process. CHMM are an extension of the typical hidden Markov model to a multichannel system. These channels have their unique underlying generation process, and the transitions depend on the current occupied states (Figure 2).

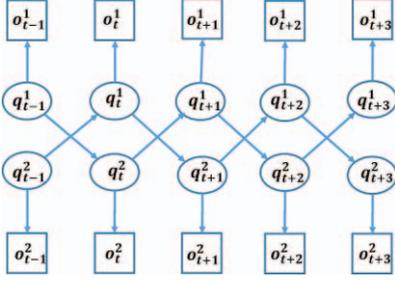


Figure 2: The CHMM scheme for a particular case when two channels are completely linked. The probabilistic relationship among states and observations are denoted by arrows [2].

A CHMM assumes that its state transition probability matrix \mathbf{A} is homogenous:

$$a_{nm}^{c'c} = P(q_t^c = S_m^c | q_{t-1}^{c'} = S_n^{c'}) \quad (1),$$

the $a_{nm}^{c'c}$ denotes the transition probability from the state n channel c' to the state m channel c at the next time instant. The emission distribution matrix \mathbf{B} is assumed to be a continuous Gaussian function:

$$b_m(o_t^c) = P(o_t^c | q_t^c = S_m^c) \quad (2),$$

the $b_m(o_t^c)$ is the probability of observing o_t^c in the state m channel c . Finally $\boldsymbol{\pi}$ is the initial state distribution matrix:

$$\pi_m^c = P(q_1^c = S_m^c) \quad (3),$$

the π_m^c is the probability that the process in the channel c starts in the state m . Therefore a CHMM is specified by $\boldsymbol{\lambda} = \{\mathbf{A}, \mathbf{B}, \boldsymbol{\pi}\}$ [2].

2.1. Computing the complete data likelihood

The likelihood $L_\lambda(O_{1:T})$ is the metric used to evaluate how likely the data sequence $\mathbf{O}_{1:T}$ was generated by our model?

$$L_\lambda(O_{1:T}) = \sum_{\mathbf{S}} L_\lambda(O_{1:T}, \mathbf{S}), \quad (4),$$

the $L_\lambda(O_{1:T}, \mathbf{S})$ can be written as:

$$L_\lambda(O_{1:T}, \mathbf{S}) = \prod_{c=1}^C \pi_1^c b_1(o_1^c) \times \prod_{t=2}^T \prod_{c'=1}^C a_{m_{t-1}m_t}^{c'c} b_{m_t}(o_t^c) \quad (5),$$

in which \mathbf{S} is the set of the state sequence in all channels. Our objective is to maximize $L_\lambda(O_{1:T})$ over all possible $\boldsymbol{\lambda}$. This is

achieved using the Expectation Maximization algorithm (EM), or an EM variant called stepwise EM (sEM). In each step of the sEM, the updated parameter set $\boldsymbol{\lambda}_{\text{updated}}$ is interpolated over the previous $\boldsymbol{\lambda}_{\text{prev}}$ and the current $\boldsymbol{\lambda}_{\text{current}}$ parameter set, using the stepsize Ω . Therefore in each step $\boldsymbol{\lambda}_{\text{updated}}$ is:

$$\lambda_{\text{updated}} = (1 - \Omega) \times \lambda_{\text{current}} + \Omega \times \lambda_{\text{prev}} \quad (6),$$

The algorithm starts with an initial guess λ_0 and iterates until some stop criteria is satisfied ($N_{\text{iteration}} \geq 100$) or a global or local likelihood minimum is accomplished.

These critical points are detected, when the new likelihood $L_\lambda^{\text{new}}(O_{1:T})$ estimated increases only 10^{-8} when compared to the previous likelihood estimation $L_\lambda^{\text{prev}}(O_{1:T})$.

2.2. Optimizing CHMM parameters

The \mathbf{A} and $\boldsymbol{\pi}$ matrix are re-estimated as in [2]. The emission distribution matrix \mathbf{B} (expected mean $\bar{\boldsymbol{\mu}}$ and the expected variance $\bar{\boldsymbol{\sigma}^2}$) are re-estimated as:

$$\bar{\mu}_m^c = \frac{\sum_{t=1}^T \alpha_{t|T}^c(m) o_t^c}{\sum_{t=1}^T \sum_{m'} \alpha_{t|T}^c(m')} \quad (7),$$

$$\bar{\sigma}_m^c = \frac{\sum_{t=1}^T \alpha_{t|T}^c(m) (o_t^c - \bar{\mu}_m^c)^2}{\sum_{t=1}^T \sum_{m'} \alpha_{t|T}^c(m')} \quad (8),$$

the forward parameters $\boldsymbol{\alpha}$ are:

$$\alpha_{t|x}^c(m) = P(S_m^c | o_{1:x}) \quad x = t - 1, t, T \quad (9),$$

the above quantity is the predicted, filtered and smoothed probability respectively. The backward parameters $\boldsymbol{\beta}$ are defined as:

$$\beta_t^c(m) \triangleq \frac{P(S_m^c | o_{1:T})}{P(S_m^c | o_{1:t-1})} \quad (10).$$

So it's easy to define the $\mathbf{L}_T^c = P(\mathbf{S}_m^c | \mathbf{o}_{1:T})$. The recursive $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are derived as in [2].

2.3. Initializing the CHMM

The ECG is modelled using 4 single states $\{\mathbf{QRS}, \mathbf{ST}, \mathbf{T}, \mathbf{TQ}\}$, where \mathbf{QRS} is the QRS complex wave, \mathbf{ST} denotes the event agglomerated between the QRS complex and the T wave, \mathbf{T} is the T wave, \mathbf{TQ} is the events agglomerated between the T wave and the QRS complex wave. The PCG is modelled using 4 single states $\{\mathbf{S1}, \mathbf{Sys}, \mathbf{S2}, \mathbf{Dias}\}$. The $\mathbf{S1}$ is the first heart sound event, \mathbf{Sys} is the systole event, $\mathbf{S2}$ is the second heart sound event and \mathbf{Dias} is the diastole event. The $\boldsymbol{\pi}$ matrix is initialized with equal starting probabilities. The \mathbf{A} matrix is partitioned, and each entry $A_i^j \in \mathcal{R}^{4 \times 4}$ defines the transition rules from the system i to j as:

$$\mathbf{A} = \begin{pmatrix} A_{ECG}^{ECG} & A_{ECG}^{PCG} \\ A_{PCG}^{ECG} & A_{PCG}^{PCG} \end{pmatrix} \quad (11).$$

The \mathbf{A} matrix (Figure 3) is well initialized if we start by assuming that S1 and the QRS-complex happen around the same time (although S1 occurs shortly after the beginning of

the QRS-complex) and the S2 and T-wave also appear around the same time too (although the S2 appears slightly after the end of the T-wave).

A_{ECG}^{ECG}	QRS	ST	T	TQ	A_{PCG}^{PCG}	S1	Sys	S2	Dias
QRS	.9	.1	.0	.0	QRS	.9	.1	.0	.0
ST	.0	.9	.1	.0	ST	.0	.9	.1	.0
T	.0	.0	.9	.1	T	.0	.0	.9	.1
TQ	.1	.0	.0	.9	TQ	.1	.0	.0	.9
A_{PCG}^{ECG}	QRS	ST	T	TQ	A_{PCG}^{PCG}	S1	Sys	S2	Dias
S1	.9	.1	.0	.0	S1	.9	.1	.0	.0
Sys	.0	.9	.1	.0	Sys	.0	.9	.1	.0
S2	.0	.0	.9	.1	S2	.0	.0	.9	.1
Dias	.1	.0	.0	.9	Dias	.1	.0	.0	.9

Figure 3: State transition matrix A for the fully connected model, the partially connected model is identical, except that each entry in A_{PCG}^{PCG} is zero.

To compute the initial parameters $B(\mu_m^c, \sigma_m^c, \forall c \in C, \forall m \in S)$, we use a segment around the correspondent annotated state m in the channel c .

2.4. Model 1 - Fully Connected Model

It is important to stress out that ECG channel does not leverage instantaneously the PCG channel (no vertical links), and there is a certain time lag (diagonal links) when carrying out state information's to the PCG channel (the opposite is also true). This model is a realistic approximation to healthy subjects and it is also our most completed model (A contains 48 independent variables). It is ideal in cases when the ECG and PCG are very clean signals, where the information inflow and outflow among the 2 channels are reliable (Figure 4).

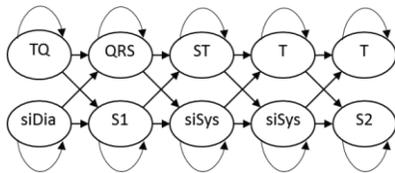


Figure 4: A correct CHMM state sequence for a particular case when two channels are fully connected and the ECG is modelled using 4 single states.

2.5. Model 2 - Partially connected model

This model is simpler than our first model (A contains 36 independent variables). It is ideal in cases when the PCG is corrupted and the ECG is very clean, therefore only the information outflow from the ECG channel is reliable (Figure 5).

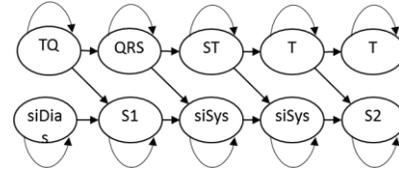


Figure 5: A correct CHMM state sequence for a particular case when two channels are partially connected and the ECG is modelled using 4 single states.

3. METHODOLOGY

3.1. Feature extraction

In the ECG channel, we choose to implement the Pan and Tompkins algorithm. In the pre-processing step, a cascade of filters is applied to attenuate the noise and to enhance the QRS complex in the signal (due to space limitation please consult [9]). Finally, the signal is properly scaled.

In the PCG channel, we implemented the entropy gradient algorithm (Figure 6). This measures the state predictability by looking to the total entropy fluctuation in the “expanded region” as the original time series is shifted in a circular motion and it is computed as in [10]. Finally, the signal is properly scaled.

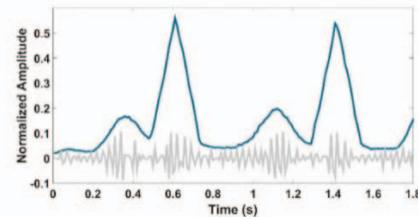


Figure 6. The $V_{entropy}$ (in blue) and the PCG signal (in gray).

3.2. Materials

The dataset we use is composed by samples from 16 healthy male adults (the average age is 30). The data acquisition was performed in a calm and relaxed environment without any ambient noises and under the supervision of a clinical technician. The PCG and the ECG are recorded at 44100Hz sampling rate, during at least 6 complete heartbeats. The PCG is recorded in the pulmonic spot and the ECG is measure in the Einthoven-II lead. One cardiacpulmonologist manually annotated the beginning and ending of the ECG and PCG states, using adequate software.

3.3. Performance Metric

The CHMM performance was measured as the model capacity to recreate the continuous state sequence in the PCG

channel. We compute the positive predictability per sample (P_{sample}^+) as:

$$P_{sample}^+ = \frac{TP_{sample}}{TP_{sample} + FP_{sample}} \quad (12)$$

where TP_{sample}, FP_{sample} is the count of true and false samples respectively. A sample at time t is a true positive when its predicted state sample and annotated state sample are the same ($S_t^{model} = S_t^{expert}$), otherwise is a false positive.

3.4. Model capacity in recreating the “true” state sequence

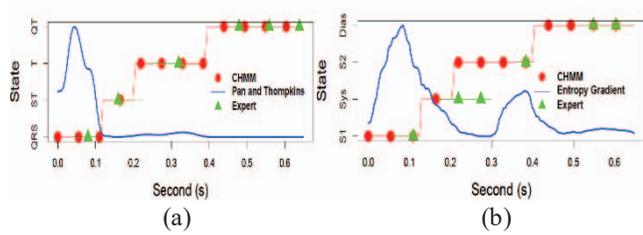


Figure 9. Classification results of the (a) ECG and (b) PCG channels. The red circles are the classification made by the CHMM, the green triangle are the classification made by the expert and the solid line represents the observation input.

In this section, we present an example of a very clean ECG signal (Figure 9). The PCG signal is corrupted although the heart sounds are still evident. The CHMM manages to decode the “true” state sequence in both channels, although because it uses a static state transition matrix A , it fails in estimating accurately the state duration in each state. This is quite evident in some cases, where the system stays only for a short time (milliseconds) in a specific state, infringing time constraints that exist in the cardiac muscle. Depending on the selected model, a channel might give negative feedback to the other channel (corrupted channel, arrhythmia, undetected state, etc), hurting our ability to decode the “true” state sequence in both channels.

4. RESULTS

The fully connected CHMM model is a very realistic model (for healthy subjects) and it is expected that this model outperforms the partially connected CHMM. In practice, this does not happen, the EM algorithm fails in converging to an interesting solution mainly because the searching space is indeed too large (48 independent parameters only when estimating A). In contrast, for the partially connected CHMM model, the EM algorithm succeeds in converging to an interesting solution, since the searching space is smaller (36 independent parameters when estimating A).

In Figure 10 a), the fully connected CHMM achieves the maximum P_{sample}^+ , when we do not re-estimate λ ($\Omega = 1$). The partial connected CHMM seems to behave differently

and the maximum P_{sample}^+ is achieved when using the typical EM algorithm to re-estimate λ ($\Omega = 0$).

In Figure 10 b), estimating the A matrix is indeed a crucial step. In the fully connected CHMM, it is mandatory to choose a large Ω , and the maximum P_{sample}^+ is achieved when we do not re-estimate A at all. The partially connected CHMM, again does not respond efficiently to the sEM and the maximum P_{sample}^+ is achieved when we use the EM algorithm ($\Omega = 0$).

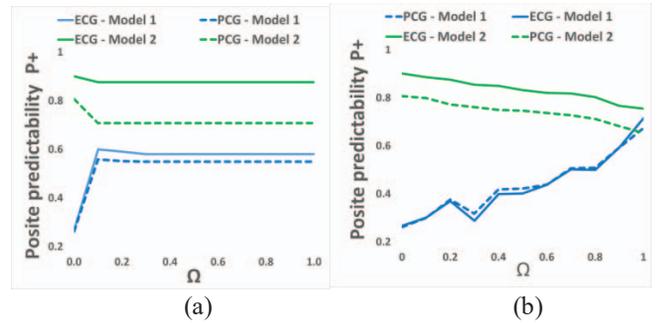


Figure 10. Positive predictability per sample (P_{sample}^+) as function of Ω .; a)The λ is re-estimated using the sEM algorithm; b) A is re-estimated using the sEM, B and π are re-estimated using the EM.

5. CONCLUSION

The fully connected CHMM did not perform as efficiently as it was expected, not because of the model design but because the EM algorithm gets very often stuck in a local optima (this happens more often as the searching space increases). The partially connected CHMM is a simplified alternative to the fully connected CHMM. It is ideal in cases when the PCG channel is corrupted, unreliable and therefore it should not leverage the ECG channel. The partially connected CHMM model outperforms the fully connected one, mainly because, the EM algorithm succeeds in converging to an interesting solution, since the searching space is smaller (36 independent parameters).

The partially connected CHMM performance is strictly dependent on the QRS and T-wave signatures. The QRS event is synchronized with the S1 event and the T-wave is also synchronized with the S2 event. Therefore, when the ECG is a trustful signal, the PCG uses the QRS and the T wave as a synchronizing point’s beat-by-beat.

7. ACKNOWLEDGEMENT

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.

12. REFERENCES

- [1] J. E. Hall and A. C. Guyton, *Textbook of medical physiology*. Philadelphia, Pa.: Saunders/Elsevier, 12th ed., 2011.
- [2] N. Montazeri Ghahjaverestan *et al.*, "Coupled Hidden Markov Model-Based Method for Apnea Bradycardia Detection," in *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 2, pp. 527-538, March 2016.
- [3] M. Brand, N. Oliver and A. Pentland, "Coupled hidden Markov models for complex action recognition", *Computer Vision and Pattern Recognition, 1997. Proceedings. 1997 IEEE Computer Society Conference on*, San Juan, 1997, pp. 994-999.
- [4] 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.
- [5] N. Brewer, N. Liu, O. D. Vel, and T. Caelli, "Using Coupled Hidden Markov Models to Model Suspect Interactions in Digital Forensic Analysis". In *Proceedings of the International Workshop on Integrating AI and Data Mining (AIDM '06)*. IEEE Computer Society, Washington, DC, USA, 2006, 58-64.
- [6] J.Zhao, F.Gonzalez, and D.Mu, "Apnea of prematurity: from cause of treatment," *European Journal of pediatrics*, vol.170, no.9, pp. 1097-1105, 2011.
- [7] A. V. Nefian, L. Liang, X. Pi, L. Xiaoxiang, C. Mao and K. Murphy, "A coupled HMM for audio-visual speech recognition," *Acoustics, Speech, and Signal Processing (ICASSP), 2002 IEEE International Conference on*, Orlando, FL, USA, 2002, pp. II-2013-II-2016.
- [8] J. Gai, Y. Li and R. L. Stevenson, "Coupled Hidden Markov Models for Robust EO/IR Target Tracking." *2007 IEEE International Conference on Image Processing*, San Antonio, TX, 2007, pp. I - 41-I - 44.
- [9] J. Pan and W. J. Tompkins, "A Real-Time QRS Detection Algorithm," in *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230-236, March 1985.
- [10] J. Oliveira, A. Castro, M. Coimbra, "Exploring Embedding Matrices and the Entropy Gradient for the Segmentation of Heart Sounds in Real Noisy Environments", in *Proc. IEEE EMBC 2014*, Chicago, USA, Aug 2014.