Supervised ECG Delineation Using the Wavelet Transform and Hidden Markov Models

G. de Lannoy[1], B. Frenay[1], M. Verleysen[1] and J. Delbeke[2]

[1]Machine Learning Group, Université catholique de Louvain, pl. du Levant 3, 1348 Louvain-la-Neuve, Belgium
[2]Department of physiology and pharmacology, Université catholique de Louvain, av. Hippocrate 54, 1200 Bruxelles, Belgium

Abstract — **Clinical monitoring and pharmaceutical phase-one studies require feature extraction from the ECG signal in order to evaluate the state of a patient's heart. Automatic annotation of the characteristic ECG waveforms (or so-called *delineation*) is therefore of great interest. Hidden Markov Models (HMM) coupled to wavelet transforms (WT) of the ECG signal offer significant improvements over standard heuristic delineation methods. Nevertheless, the choice of the WT parameters remains empirical rather than data-driven. In these conditions, suboptimal parameters for the WT may degrade the results very much. In this paper, an algorithm for the optimal selection of the WT parameter values is introduced. The model complexity is strongly reduced and the algorithm can adapt itself to the specificities of each ECG signal while avoiding redundancy, noise and useless information. Evaluation on recordings from the public MIT-QT database leads to results higher than with state of the art methods.**

Keywords — **ECG Delineation, Hidden Markov Models, Wavelet Transform, Stepwise forward selection, MIT QT Database.**

I. INTRODUCTION

The electrocardiogram (ECG) signal is a powerful non-invasive source of information on the clinical condition of patients. It is a measure of the electrical activity associated with the heart; the signal is characterized by a time-variant cyclic occurrence of patterns with different frequency contents (QRS complexes, P and T waves). The P wave corresponds to the contraction of the atria, the QRS complex to the contraction of the ventricles and the T wave to their repolarization. During clinical monitoring or phase-one evaluation studies of new drugs, measurements pertaining to these characteristic ECG waveforms are used to assess the state of a patient's heart. For example, an abnormal prolongation of the time between the Q point and the end of the T wave (*QT interval*) indicates an increased risk of sudden cardiac death [1].

The annotation (or so-called *delineation*) of the characteristic ECG waveforms on the digitalized signal is therefore of great interest. On the other hand, manually analyzing long-term ECGs is a time consuming process and it can lead to errors and misinterpretations. Automatic computer-aided annotation of the ECG signal can thus greatly help physicians in their diagnosis. However, it is a difficult task in real situations. First, the physiological variations due to the patient and his disease make of the ECG a non-stationary signal. Second, the time interval between successive waveforms is varying with time. Third, many sources of noise pollute the ECG signal, such as power line interferences, muscular artifacts, poor electrode contacts and baseline wanderings due to respiration. These problems highly compromise the effective delineation of the signal, especially with real-life signals such as ambulatory recordings.

In this paper, a robust algorithm for automatic ECG delineation is introduced. In order to deal with the issues listed above, it is based on a strong signal pre-filtering step by wavelet transform. Subsequent detection of the characteristic ECG waveforms is then achieved by a probabilistic modeling approach learning a model from expert annotated data. Our contribution consists in the development and testing of an automatic supervised procedure for the optimal settings of the wavelet transform parameters.

The remainder of this paper is structured as follows. After this introduction, Section 2 gives a brief account of the literature about the state of the art in ECG delineation. Section 3 provides a theoretical summary of the machine learning methods used in this work. Section 4 introduces the methodology followed by the algorithm and Section 5 shows the results obtained on a real public database.

II. STATE OF THE ART

Standard approaches to automatic ECG delineation attempt to find the characteristic points in a number of successive steps by using threshold methods and heuristic rules such as slope criterions. These methods require the setting of many empirical parameters and are very sensitive to noisy waveforms or changes in wave morphology [2]. Furthermore, these methods have been proven unreliable even on clean recordings [3]. As a result, interval measurements are usually performed manually by experts.

In order to circumvent these issues, probabilistic modeling approaches and especially hidden Markov models (HMMs) have been introduced for ECG delineation [4,5]. The ECG signal can be viewed as the result of a generative process, where each waveform is generated by a particular state of the heart, this state being hidden to the observer. The cardiological process is sequential, and each state is solely dependent on the previous state. HMMs are naturally appropriate for modeling this kind of process [6].

HMMs offer significant improvements over heuristic methods. First, they are data-driven models, thus able to learn the statistical properties of each waveform by using annotations provided by experts on a small signal sample. Second, they provide the ability to incorporate prior knowledge about the problem such as the sequential occurrence of waveforms but also about their statistical properties. Third, efficient algorithms exist for HMMs training and testing.

However, it is only recently that the choices of the ECG representation and of the HMMs architecture have been investigated in order to obtain high delineation performances [2]. In particular, the multi-dimensional signal decomposition by wavelet transform (WT) highly improves the delineation performances [2,7].

Until now, the choice of the WT parameter values was made empirically and was not automatically adapted to each dataset's characteristics. Nevertheless, each patient and each recording situation may lead to different signal morphologies and suboptimal parameter values for the WT may degrade the results very much [8,9].

In this paper, a supervised optimal selection of the WT parameters for subsequent robust modeling of the ECG sequence by HMMs is introduced.

III. Theoretical background

This section provides a brief theoretical introduction to HMMs and the continuous wavelet transform (CWT). A more detailed description can be found respectively in [6] and [10] for example.

A. Hidden Markov models

Since their development in the late 1960s, HMMs have proven to be a powerful and flexible class of statistical models for describing many different kinds of sequential data. Let $H = <A,B,s,q>$ be a finite HMM describing the statistical relationship between an observable sequence O and an unobservable or hidden sequence S. Parameter s is the set of hidden states $\{s_1, s_2, ..., s_n\}$. Parameter A is the state transition matrix with $a_{ij}=P(S_{t+1}=s_j|S_t=s_i)$. Parameter B

contains the emission probability distributions $\{b_1, b_2, ..., b_n\}$ for each state, such that $b_i(o_j) = P(O_t=o_j|S_t=s_i)$. Parameter q is the initial state probabilities, with $q_i = P(S_1=s_i)$. The standard assumption in HMMs is that observation values within a given state are defined as being independent and identically distributed (i.i.d.). In the case of supervised learning, the annotated sample can be used to estimate the parameters.

H thus defines a joint probability distribution $P(O,S|\lambda)$ with $\lambda=\{S, A, B, q\}$ over the hidden state sequences S and the observation sequences O. One of the most important tasks in HMMs is to infer the most likely sequence of states given a sequence of observation,

$$S* = \arg\ \max\ \{P(S \mid O, \lambda)\}.\qquad(3)$$

Using Bayes' rule, equation 3 simplifies to finding the state sequence $S*$ which maximizes the joint distribution $P(O,S|\lambda)$. This problem is solved by the Viterbi algorithm [11].

B. The continuous wavelet transform

The continuous wavelet transform (CWT) is a time-frequency decomposition of a signal by the convolution of this signal with a so-called *wavelet function*. A wavelet function $\psi(t)$ is a function with several properties: it must have a zero mean and square norm one. From a wavelet function, one can obtain a family of time-scale waveforms by translation b and scaling a,

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right), a = 2^k, k \in \aleph, b \in \Re.\qquad(4)$$

When $a=1$ and $b=0$, the wavelet is called the *mother wavelet*. The wavelet transform of a function $x(t)$ is a projection of this function on the wavelet basis $\{\psi_{a,b}\}$

$$T(a,b) = \int_{-\infty}^{+\infty} x(t)\Psi_{a,b}(t)dt.\qquad(5)$$

The CWT is a suitable tool for ECG analysis because of this time-frequency representation of the signal [2]. With the multi-scale feature of WTs, the different characteristic ECG waveforms can be distinguished over noise, baseline drift, and artifacts. The important time course of the non-stationary ECG signal is preserved. Moreover, efficient implementations of the algorithm exist and a low computational complexity is required, allowing real-time analysis.

IV. METHODOLOGY

In order to represent the dynamic of the ECG sequence, a purely sequential HMM is defined with the five following hidden states: (1) P wave, (2) baseline 1, (3) QRS wave, (4) T wave and (5) baseline 2. Figure 1 shows the graphical representation of the HMM.
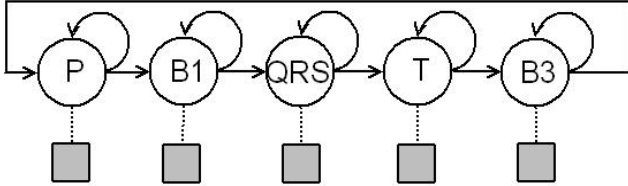

Fig. 1 Graphical representation of the HMM for ECG delineation

Due to the non-stationarity of the ECG signal and the presence of many sources of noise and artifacts, a strong pre-filtering step by CWT is required prior to modeling. State of the art performances are obtained by using a Coiflet mother wavelet with two vanishing moments and by using all the consecutive dyadic decomposition scales corresponding to the characteristic frequencies of the ECG signal (being scales 1 to 7 in this case) [2,7]. The sequence of observations O is thus made of the multi-dimensional WT coefficients and the hidden state sequence S is made of the state values associated to each point of O.

Until now, the choice of the mother wavelet and of the decomposition dimensions is left to empirical choice. However, these parameters may be of great influence on the results. First, the delineation performances degrades when too many decomposition dimensions start to be included in the model; wavelet scales corresponding to noise and baseline drift, which were intended to be filtered, start to be added to the model. Second, choosing too many dimensions may lead to a model complexity that is excessively high compared to the available training samples. Third, including the set of all consecutive dyadic scales from one to seven in the model may lead to redundancy or useless information. Finally, each patient and each recording situation may lead to different signal morphologies and may therefore require different mother wavelets [8,9].

As a solution to these issues, the available training samples can be used for a supervised selection of the optimal combination of mother wavelets and of their decomposition scales. More formally, the CWT is computed at a set of consecutive dyadic scales for several mother wavelets. Each item of the set is a signal corresponding to the WT at a specific scale of a particular mother wavelet. The full set of items is then provided as input to an iterative stepwise forward selection procedure to select the best subset of items. The selection procedure involves starting with an empty subset, trying out at each step the trial items one by one and including them to the model if a given criterion value is improved. The procedure stops when no item left in the set can improve the criterion value. In addition, at each step, the items previously selected are individually challenged: if their removal does not decrease the criterion value, the item has become useless and is therefore removed from the model. Multiple wavelets can thus be incorporated in the model, better reflecting each dataset's characteristics. Also, only their informative scales are included, avoiding redundancy, noise and useless information.

V. EXPERIMENTS AND RESULTS

The assessment of performances is achieved using the benchmark public QT database from Physionet, designed for evaluation of algorithms that detect waveform boundaries in the ECG [12]. Waveform boundaries for 30 beats have been manually determined by expert annotators in each recording of the database. A wide variety of ECG morphologies and pathologies are represented, and the ten recordings corresponding to normal sinus rhythm were used in this work, for a total of 300 annotated beats.

For each recording, the performances are evaluated using a 5-fold cross-validation procedure after range normalization of the signal and random permutation of the beats. Specifically, for each fold of the cross-validation procedure, the four other folds (defined as the training set) are used to learn the model's parameters and the current selected fold serves as an independent performance measure (the test set). Taking advantage of these available training samples, the criterion used by the forward procedure is the correct classification rate on this training set. It is defined as the average number of points of the training set correctly classified by the Viterbi algorithm.

The transition matrix A is estimated from the training set, and the assumption is made that the model always starts in state 1 (the P wave), therefore avoiding parameter q.

A Gaussian mixture model (GMM) for the probability distributions b_k has previously been found to be the most efficient emission choice [2]. Regardless of the great performances, a k-components mixture of d-dimensional Gaussian requires the estimation of $k-1+(k \times d)+(k \times d)/2 \times (1+d)$ free parameters. If k=5 and d=7, the number of free parameters is equal to 179. If k=10, this number raises to 359. In practice, it is rare to have enough annotated ECG samples for the estimation of so many free parameters, because manual annotation is an expensive process. If there are not enough training data, then the

model order should be reduced. For this reason, the observations probability models are set to multidimensional Gaussian models in this work, thus avoiding the complexity problem induced by GMMs and the reliability of the estimations when too few annotated samples are available.

The Viterbi algorithm is used to solve equation 3 and obtain the most likely sequence of states for the test set.

Table 1 shows the results on the 10 recordings of the database. Three models are compared. The first one is an HMM on the raw data, with no WT pre-filtering step. The second one is the state of the art model, using a coif2 wavelet and dyadic scales from 1 to 7. The third one is our algorithm, with a stepwise forward selection of the optimal wavelets and scales. First column reports the classification rate, defined as the percentage of points being classified in the correct state. Column two shows the double beat detection rate, defined as the number of times that the model incorrectly infers two or more heartbeats where there is only one beat present in a particular region of the ECG signal. The results are averaged on each fold of the cross-validation procedure for the 10 datasets.

Table 1: Results on 10 recordings from the MIT QT database

Model	Delineation	Double beats
Raw data	70.1%	41%
Fixed parameters	93.4%	4.1%
Forward selection	94.7%	3.7%

From these results, the need of using a pre-filtering step by WT step is clearly demonstrated. The delineation results on raw data are very poor, and the double beat segmentation rate is problematic. Compared to a WT with a priori fixed parameters, the supervised selection improves the results and decreases the double beat segmentation rate. In addition, the improvements are expected to be much more pronounced on datasets with physiological variations. Still, it is important to note that the major drawback in HMMs is the lack of duration constraints, which causes most of the delineation errors and of the double beat segmentations.

VI. CONCLUSION

HMMs offer significant improvements over heuristic delineation methods in order to model the sequential cardiological process. Also, the multi-dimensional signal decomposition by WT can naturally be incorporated into HMMs and clearly improves the delineation performances. Nevertheless, wrong empirical parameters settings for the WT can degrade the results.

In this work, a stepwise forward selection procedure is introduced in order to automatically select the optimal WT parameters. Multiple mother wavelets can be incorporated in the model, and only the optimal decomposition scales are included. The model complexity is strongly reduced, and the algorithm can adapt itself to each dataset's specificities while avoiding redundancy, noise and useless information. Experiments on sinus rhythm recordings from the public MIT QT database yield an average correct delineation rate of 95% which is better than the results obtained with state-of-the-art methods.

Further work will include the use of semi hidden Markov models and include the modeling of durations.

REFERENCES

1. Laakso M, Aberg A, Savola J, Pentikäinen PJ and Pyörälä K (1987) Diseases and drugs causing prolongation of the QT interval. American journal of Cardiology, 59(8), pp 862-5
2. Clifford G. D., Azuaje F. and McSharry P. E. (2006) Advanced Methods and Tools for ECG Data Analysis. Artech House Publishing, Boston/London.
3. Malik M. (2004) Errors and misconceptions in ECG measurement used for the detection of drug induced QT interval prolongation. Journal of Electrocardiology, 37, pp 25-33
4. Coast D.A., Stern R.M., Cano G.G. and Briller S.A. (1990) An approach to cardiac arrhythmia analysis using hidden Markov models. IEEE Transactions on Biomedical Engineering, 37(9), pp 826-836
5. Koski A. (1996) Modeling ECG signals with hidden Markov models. Artificial Intelligence in Medicine, 8(5), pp 453-471
6. Rabiner L.R. (1989) A Tutorial on Hidden Markov Models and Selected Applications In Speech Recognition. Proceedings of the IEEE, 77(2), pp 257-286
7. Hughes N., Tarassenko L. and Roberts S. (2004) Markov Models for Automated ECG Interval Analysis, Advances in Neural Information Processing Systems (NIPS), 16
8. Singh B. N., Tiwari A. K. (2006) Optimal selection of wavelet basis function applied to ECG signal denoising. Digital Signal Processing, 16(3), pp 175-287
9. Addison P.D. (2005) Wavelet Transform and the ECG: A Review. Physiological Measurements, 25, 155-199
10. Mallat S. (1999) A Wavelet Tour Of Signal Processing (Wavelet Analysis And Its Applications). IEEE Press, San Diego
11. Forney, G. D. (1973) The Viterbi Algorithm. Proc. of the IEEE, 61, pp 268-278.
12. Goldberger A. L., Amaral L. A. N., Glass L., Hausdorff J. M., Ivanov P. Ch., Mark R. G., Mietus J. E., Moody G. B., Peng C.-K. and Stanley H. E. (2000) PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation, 101(23), pp 215