Emission Modelling for Supervised ECG Segmentation using Finite Differences

B. Frénay[1], G. de Lannoy[1] and M. Verleysen[1]

[1]Machine Learning Group, EPL/ELEC/DICE, Université catholique de Louvain, Place du Levant, 3, 1348 Louvain-la-Neuve, Belgium

Abstract — **The segmentation of ECG signals into P waves, QRS complexes, T waves and baselines is an important practical problem for physicians diagnosing cardiac diseases. The duration of the signal and the number of beats to segment are often too large for a manual annotation, so that automatic segmentation is a challenging and useful tool. State-of-the-art algorithms use hidden Markov models with wavelet transform encoding and represent the ECG in multidimensional spaces using Gaussian mixtures models. The main problem of this approach is its computational cost due to the number of free parameters, the choice of the wavelet transform parameters and the high failure rate of the EM algorithm. In this work, we propose an alternative emission encoding for hidden Markov models using both the ECG signal and its derivative in order to better model the dynamics of the signal in a lower dimensional space. We show that this method achieves similar performances with much less model parameters and is less subject to failures.**

Keywords — **ECG segmentation, hidden Markov models, emission modelling, finite differences, QT database.**

I. Introduction

Physicians use the electrocardiogram (ECG) signal to diagnose many cardiac diseases [1] such as the long QT syndrome or the torsades de pointes syndrome. The recording of the heart electrical activity follows a well-know dynamics in normal situations so that time distances between specific peaks contain useful information for the diagnosis of cardiac anomalies.

In drug studies, ECG signals are used to determine whether the heart behaviour is modified or not. However their analysis is problematic: they often last several hours and contain thousands of beats. Human specialists simply cannot systematically annotate (i.e. identify the peaks) the whole signal; an automated approach is required.

There exists an important literature to help the practitioner annotating the signals, but most approaches are based on ad-hoc heuristics [2] using the tangent rule, thresholds, etc. The more general approach [3,4] adopted in this paper uses hidden Markov models [5] to model the whole beat by a sequence of states, each associated with a particular peak.

The ECG signal contains an important noise component so that a preprocessing is applied before the segmentation effectively takes place. The state-of-the-art wavelet transform encoding [6] has been studied in details [7,8,9], but the derivative of the ECG signal has, to our knowledge, never been used. In this paper, we show that the derivative contains useful information and allows achieving similar performances with much less complex models.

The following of this paper is organised as follows. Section II describes ECG signals and the segmentation problem and introduces a probabilistic model. Section III analyses the appearance of the ECG signal derivatives, proposes a new bidimensional representation and quickly reviews Gaussian mixture models. Eventually, Section IV presents the experimental results discussed in Section V.

II. ECG segmentation

A. Definitions and problem statement

An ECG signal is a measure of the electrical activity of the heart and consists of successive beats. Fig. 1 shows the structure of a human beat which can be divided into three waves (P wave, QRS complex and T wave) separated by three baselines (B1, B2 and B3). Note that measured ECG signals are discrete: for example, those used in our experiments are sampled at 250 Hz.

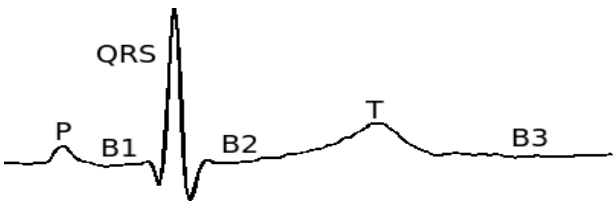

Fig. 1 A human beat extracted from a filtered ECG signal

In practice, ECG signals barely look like the signal of Fig. 1: they contain noise and are unique to each patient. This is the reason why a probabilistic framework is interesting here: rather than designing ad-hoc non-robust heuristics, we can learn a complete probabilistic model of the ECG signal and then choose the most probable annotations. This process is called segmentation.

More specifically, the problem is here to learn a model from a small manually annotated part of the ECG signal

J. Vander Sloten, P. Verdonck, M. Nyssen, J. Haueisen (Eds.): ECIFMBE 2008, IFMBE Proceedings 22, pp. 1212–1216, 2008
www.springerlink.com

(typically about 30 beats) in order to annotate the remaining signal. In others words, the physician only annotates a few beats and then leaves the rest to an automated algorithm.

B. Hidden Markov models for ECG signal segmentation

ECG signals can easily be modeled by hidden Markov models (HMMs) which are intensively used in speech recognition [5]. Let be a (finite) HMM $\mathcal{H} = <\mathcal{S},A,B,q>$ where $\mathcal{S}$ is the set of hidden states $\{s_1,..,s_n\}$. A is the state transition probability matrix defined by

$$a_{ij} = Pr(S_{t+1} = s_j \,|S_t = s_i\,), \quad (1)$$

B contains the emission probability distributions $b_1,..,b_n$ such that

$$b_i(o_j\,) = Pr(O_t = o_j \,|S_t = s_i\,) \quad (2)$$

and q is the initial distribution such that $q_i = Pr(S_1=s_i)$. $\mathcal{H}$ is thus a description of the relationship between an unobservable sequence of hidden states S and an observable sequence O. In our case, S corresponds to the sequence of waves and baselines whereas O is the measured ECG signal.

Fig. 2 shows the architecture of the HMM used in this paper [4,7]. Each wave and baseline is modeled by an individual state so that only transitions between successive beat segments are allowed.

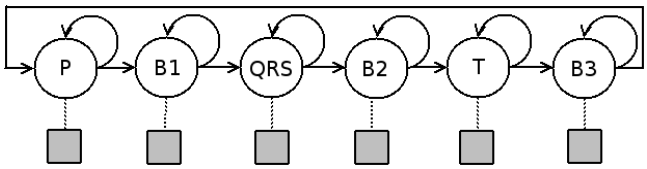

Fig. 2 HMM architecture for ECG signal segmentation

Using HMMs, the segmentation is divided into two steps. Firstly, the transition probability matrix A is built and an emission model b_i is fit to each state using the manual annotations. The model is assumed to start in P state so that q = (1 0 0 0 0 0). Secondly, the Viterbi algorithm [10] is used to find the most probable sequence of states (or annotations) for the rest of the ECG signal.

III. Emissions Modeling using Finite Differences

When using HMMs for ECG segmentation, an important issue is the proper modeling of the emissions. Firstly, the signal is encoded into a (multidimensional) representation which is easier to work on. Secondly, probability density estimators b_i are built from the new ECG signal.

State-of-the-art algorithms [7,8,9] use the wavelet transform (WT) [6] in order to obtain a time-frequency multidimensional representation. However, this solution has an important computational cost due to the model complexity. This section introduces the contribution of this paper and shows that finite differences allow modeling the emissions for ECG signals with fewer model parameters.

A. Finite differences

An important limitation of HMMs is that the emissions are considered to be independent. This limitation often produces additional beats lasting only a few milliseconds that are inserted because they statistically better explain the signal. Fig. 3 shows a example of bad segmentation achieved by an HMM using a separate univariate Gaussian distribution for each state. The annotations incorrectly inserted during the T wave are symptomatic: they explain the data but have no biological meaning.

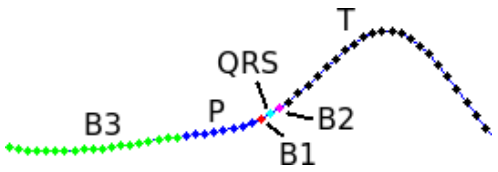

Fig. 3 An example of degenerated beat inserted in a T wave

In order to avoid adding additional beats, a good emission model should take the context, i.e. the time-contiguous values, into account. State-of-the-art solutions use WT encoding, but the choice of the WT parameters requires building several models. Moreover, these solutions use Gaussian mixture models whose computational cost increases quickly with the number of wavelet scales: a k-components d-dimensional Gaussian mixture model has kd(d+1)/2+kd+k parameters. The EM is also likely to fail because of ill-conditioned covariance matrices [11].

Alternatively, the derivative of the ECG signal also gives useful information about the context. Let us define the first-order derivative of the ECG signal as the finite difference (FD) $O'_t = f(O_t–O_{t-1})$ where f is the sampling frequency. Fig. 4 shows the scatter plots of the ECG signal and its derivative.

Waves are characterised by circular probability distributions whereas baselines are rather concentrated; B1 and B2 also have long tails. Interestingly, each distribution has a very different shape and scale which is unique to each patient. It suggests that an HMM using FD encoding should easily discriminate the segments of a given ECG signal.

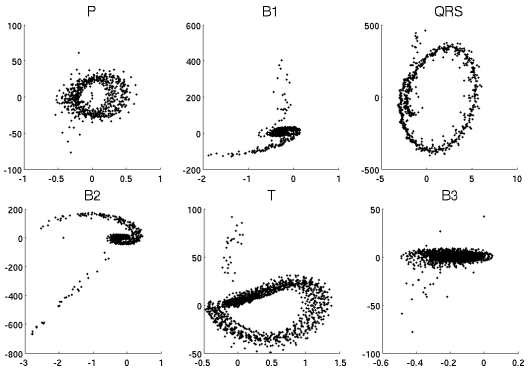

Fig. 4 Scatter plots of the ECG signal (x-axis) and its derivative (y-axis)

In the following of this paper, we will work on the bidimensional FD encoding of the ECG signal and show that it is a valuable alternative to the WT encoding.

B. Gaussian mixtures models for ECG signal modelling

In order to build the emission model b_i for each state, we must choose a probability density estimator. Except for B3, the representations in Fig. 4 cannot be easily modeled by standard unimodal distributions. In this paper, we therefore use semi-parametric probability density estimators called Gaussian mixture models (GMMs) [12].

A k-component d-dimensional GMM is a weighted sum of k separated Gaussian distributions with unknown means and covariance matrices. GMMs are flexible models widely used in machine learning, but their computational cost is important. Firstly, kd(d+1)/2+kd+k parameters have to be estimated by expectation-maximisation (EM). Secondly, the estimation has to be done many times with random initial conditions. Indeed, the EM algorithm can remain stuck in local maxima and covariances matrices often become ill-conditioned (which cause a failure of the EM procedure).

Since WT encoding uses up to 7 dimensions, we can expect an important computational benefit from our bidimensional FD encoding if it achieves similar performances.

Notice that the choice of the number of components of a GMM is a difficult problem. It can be automatically chosen by cross-validation, but the Akaike information criterion and the Bayesian information criterion [12] are much less computationally demanding and also give good results.

IV. Experiments

This section describes the setup and the results of the experiments on different discretised ECG signals sampled at 250 Hz.

A. Methodology

Firstly, a 3 Hz–30 Hz band-pass filter compatible with 250Hz sampling frequency is applied to remove low and high frequency noise. The whole signal is then separated into training and validation sets of beats for 10-fold cross-validation: 90% of the signal is used to build statistical models for automated annotation whereas 10% are used for annotation quality assessment.

Secondly, the multidimensional representation of the ECG signal is computed using FD and WT encoding. WT was done using a coiflet wavelet with two vanishing moments and contiguous dyadic scales $2^1,..,2^n$ [7] which give the best results. The learning and validation signals are processed separately.

Thirdly, a GMM is fit using EM on each wave and baseline of the learning signal in order to obtain the multimodal multidimensional probability density estimators b_i.

Eventually, the transition probability matrix A is computed and the Viterbi algorithm is run on the validation signal in order to obtain the corresponding most probable state sequence. The validation signal is then annotated according to the resulting state sequence.

B. Annotation quality assessment

The quality of automated annotation can be assessed using

- the **accuracy** which is the percentage of correct annotations;
- the **sensitivity** which is the ratio TP/TP+FN, i.e. the percentage of each segment which is recognised;
- the **single-beat segmentation percentage** which is the percentage of annotations which really correspond to actual annotations, i.e. a low value means that there are many degenerated beats like on Fig. 3;
- the **mean absolute error** which is the average absolute error in milliseconds between real and automated annotation for non-degenerated beats.

The number of free model parameters for each model and the percentage of failures during EM due to ill-conditioned matrices are also interesting information.

An efficient and fast algorithm should obtain a high accuracy, a high sensitivity, a high positive predictivity, a high percentage of single-beat segmentation, a low mean absolute error and compute only a few parameters with a low percentage of failures.

C. Datasets

The experiments use the QT database [13] sampled at 250 Hz. The QT database includes a wide variety of ECG morphologies; however only the two normal sinus rhythm ECG signals with no anomalies at all are used (sel16539 and 16786). Both signals contain 30 annotated beats and the automated annotations are forced to start in P wave and to end in B3 baseline.

D. Results

Three models were compared: a simple model using GMMs on the filtered ECG signal only, the state-of-the-art WT encoding with GMMs and the FD encoding with GMMs. The numbers of components for the GMMs were chosen using the Akaike information criterion: the Bayesian information criterion tends to choose too few components. The numbers of contiguous dyadic scales for the WT were respectively set to 5 and 4 for the first and second database.

Table 1 Results with a 10-fold cross-validation on the signal sel16539 from the QT database using simple GMMs, WT (5 dyadic scales) and FD

	GMMs	WT + GMMs	FD + GMMs
Accuracy	76.58% 76.23-76.94%	93.74% 93.64-93.84%	91.69% 91.59-91.78%
Sensitivity	74.44% 74.03-74.86%	91.80% 91.59-92.00%	89.88% 89.64-90.13%
Percentage of single-beat segmentation	88.50% 88.50-88.50%	92.42% 91.93-92.91%	100% 100-100%
Consistency	7.74ms 7.53ms-7.94ms	2.21ms 2.17ms-2.25ms	3.26ms 3.23ms-3.30ms
Number of free parameters	54 53-55	1473 1467-1479	349 345-353
Percentage of failures	0.00% 0.00-0.00%	16.11% 15.86-16.36%	0.33% 0.28-0.37%

Table 2 Results with a 10-fold cross-validation on the signal sel16786 from QT database using simple GMMs, WT (4 dyadic scales) and FD

	GMMs	WT + GMMs	FD + GMMs
Accuracy	86.26% 85.91-86.61%	94.38% 94.31-94.45%	93.53% 93.46-93.61%
Sensitivity	82.93% 82.70-83.15%	93.34% 93.21-93.47%	91.57% 91.48-91.66%
Percentage of single-beat segmentation	83.47% 81.85-85.09%	99.42% 98.65-100.19%	100% 100-100%
Consistency	4.37ms 4.27ms-4.48ms	2.26ms 2.23ms-2.28ms	2.61ms 2.58ms-2.64ms
Number of free parameters	91 90-93	1065 1063-1068	396 395-398
Percentage of failures	0.00% 0.00-0.00%	15.33% 15.17-15.49%	0.62% 0.54-0.70%

Tab. 1 and Tab. 2 show respectively the results of our experiments for the two signals of the QT database. Each experiment was run 10 times on each database so that the tables contain the mean (black) and 95% confidence interval (grey) on the 10 runs for each measure.

Both encodings outperform the simple model using only GMMs. They achieve comparable results in terms of accuracy, sensitivity and consistency; the WT encoding is statistically significantly (but only slightly) better for these criteria. On the contrary, the number of free parameters and the percentage of failures are much smaller with FD encoding: WT encoding works with 4 or 5-dimensional GMMs whereas FD uses 2-dimensional GMMs. Moreover, the FD encoding reaches 100% of single-beat segmentation.

V. Conclusions

This paper introduces a novel encoding for ECG signals using finite differences. The experiments on the QT databases show that this encoding is a valuable alternative to the wavelet transform.

The main interest of the FD encoding approach is that it achieves comparable annotation qualities as the state-of-the-art encoding, while using fewer dimensions for ECG signal representation. The computational cost and the risk of failure during EM are therefore reduced.

We plan to test our approach on atypical ECGs and to improve the bidimensional encoding and modeling.

Acknowledgements

G. de Lannoy is funded by a Belgian F.R.I.A. grant.

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. 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
3. Koski A. (1996) Modeling ECG signals with hidden Markov models. Artificial Intelligence in Medicine, 8(5), pp 453-471
4. Clifford G. D., Azuaje F. and McSharry P. E. (2006) Advanced Methods and Tools for ECG Data Analysis. Artech House Publishers
5. Rabiner L.R. (1989) A Tutorial on Hidden Markov Models and Selected Applications In Speech Recognition. Proc. of the IEEE, 77(2), pp 257-286
6. Mallat S. (1999) A Wavelet Tour Of Signal Processing (Wavelet Analysis And Its Applications). IEEE Press, San Diego
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. Andreão R V and Boudy J (2007) Combining Wavelet Transform and Hidden Markov Models for ECG Segmentation, EURASIP Journal on Applied Signal Processing, Volume 2007 , Issue 1 Pages: 95 – 95
9. Addison P.D. (2005) Wavelet Transform and the ECG: A Review. Physiological Measurements, 25, 155-199
10. Forney, G. D. (1973) The Viterbi Algorithm. Proc. of the IEEE, 61, pp 268-278.
11. C. Archambeau, J.A. Lee and M. Verleysen (2003) On the convergence problems of the EM algorithm for finite Gaussian mixtures. Proc. of the ESANN'03, pp 99-106
12. Bishop C. M. (2007) Pattern Recognition and Machine Learning. Springer
13. 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