BEST BASIS SEGMENTATION OF ECG SIGNALS USING NOVEL OPTIMALITY CRITERIA

Dana H. Brooks[1] *Hamid Krim*[2] *J.C. Pesquet*[3] *Robert S. MacLeod*[4]

[1] CDSP Center, ECE Dept., Northeastern University, Boston, MA USA
[2] LIDS, MIT, Cambridge, MA, USA 02139
[3] CNRS/UPS & GDR TdSI, Gif sur Yvette, Cedex, France 91192
[4]CVRTI, University of Utah, Salt Lake City, UT, USA

ABSTRACT

Automatic segmentation of the ECG is important in both clinical and research settings. Past algorithms have relied on incorporation of detailed heuristics. In this paper we avoid heuristics by employing a best-basis algorithm. As large variability of the local SNR causes the standard entropy criterion to produce an overly-fine segmentation, we introduce a novel optimality criterion which is based on a linear combination of the entropy measure and a function of a smoothness measure, and is quite general in form. We tested the algorithm on the MIT-BIH arrythmia database and body surface potential maps.

1. INTRODUCTION

Automatic segmentation of the electrocardiogram (ECG) using a minimum of heuristic *a priori* information is an important problem in many clinical and research application areas. The various segments of the ECG have different physiological meaning, and the presence, timing, and duration of each of these segments have diagnostic and biophysical importance. The problem is made considerably more difficult because the shape of the ECG is quite variable both within and across patients. This intra-subject variability depends on a variety of factors; the most important is the location of the electrode relative to the cardiac sources which are active at a particular time. Other causes include the body position, measurement noise, muscle artifacts, and poor electrode contact. Variability is considerable even across healthy subjects; in individuals suffering from ischemia (insufficiency of oxygen supply to the cardiac tissue) or arrhythmias (irregular heart rhythms) it is dramatic, even changing from beat to beat. Thus heavy reliance on heuristics or the tuning of algorithms to any particular ECG database is problematic. And yet one can usually segment these waveforms easily by eye. Thus the challenge is to develop automatic segmentation tools that are robust to inherent variability in the signal with minimal reliance on explicit incorporation of heursitics. The approach we describe below attempts to accomplish this goal by maximizing the algorithmic incorporation of the type of intrinsic signal features used by the human observer.

This material is based upon work supported by a Biomedical Engineering Research Grant from the Whitaker Foundation. The work of the second author was in part supported by grants from ARO(DAAL03-92-G-0115) (Center for Intelligent Control) and AFOSR (F49620-95-1-0083.)

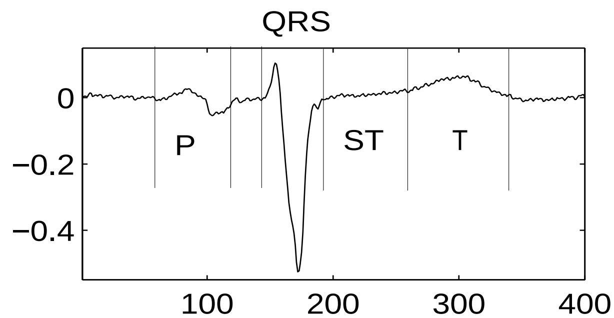

Figure 1. Annotated typical electrocardiogram showing the desired segmentation into the four phases. Vertical and horizontal axes are in mV and msec.

1.1. Segments of the ECG

Electrocardiograms are traditionally divided into four main electrical events, each reflecting the electrical activity associated with a particular phase of the cardiac cycle. These four events are named the P wave, the QRS complex, the ST segment, and the T wave, and are illustrated for a "typical" ECG by the annotations on the waveform in Figure 1. In a normal heartbeat, the P and QRS segments reflect the electrical excitation of the heart muscle via temporally rapid and spatially localized depolarization wavefronts in the cardiac tissue. The QRS complex is followed immediately by an interval of low-level electrical activity known as the ST segment, during which there is normally little potential difference across the heart and thus small amplitudes and little temporal change in the ECG. The T wave, caused by the return of the ventricular tissue to its initial resting state, reflects a much more temporally and spatially diffuse wavefront than P and QRS.

Because of the close tie between the segments of the ECG signal and the underlying physiological states, there has been considerable interest in the development of signal processing techniques to automatically perform this segmentation. In particular, automatic segmentation is important in applications such as patient monitoring and body surface potential mapping (BSPM) (when up to 200 electrodes are used simultaneously to better measure the spatial pattern of cardiac activity. In the first of these applications, the segmentation is used for purposes ranging from determining the presence or absence of a heartbeat to detecting, counting and characterizing unusual heartbeat waveforms. The output algorithm may sound an alarm, control a pacemaker, or even trigger an implantible defibrillator. In BSPM, automatic segmentation is a practical requirement for extracting spatio-temporal information from the data about spatial variation and temporal change in the timing, duration, or shape of the ECG segments.

1.2. Current Approaches to ECG Segmentation

Most previous approaches to this problem have combined simple signal processing techniques such as discrete differentiation and filtering with a large body of heuristic rules, such as minimum and maximum lengths for each segment, complicated adaptive threshold criteria, frequency range criteria, signal templates, *etc.* (see, for example, [1, 2]). In recent years there has been interest in the application of wavelet techniques to this problem [3, 4, 7]-[9]. Most of this work has used linear phase, spline-based wavelets [10] to detect peaks in the signal, combined with a complicated set of heuristics. The problem with this reliance on heuristics, as mentioned earlier, is the wide variability of ECG signals, both among different subjects and among different sensing locations on the body surface. For instance, the waveforms in Fig. 3 vary considerably despite the fact that they were recorded simultaneously from the same subject. Reliance on heuristics can cause an algorithm to be too sensitive to the particular data-base on which it was developed,

1.3. Segmentation by Best Basis Selection

The notion of the *best basis* for a given signal has classically relied on the parsimony of its resulting representation in this basis; the basis is typically chosen using a metric which reflects an energy concentration measure (ECM). This approach has been successfully applied for best wavelet packet representation (*i.e.* best spectral segmentation) and has also been proposed for time/spatial segmentation using a *local trigonometric basis* [5] The aim of this approach is to capture the dominant localized coherent structures and assign them to a corresponding segmentation.

Experiments have shown that such ECM criteria can indeed achieve an adequate and acceptable delineation/segmentation for a number of signal classes. These criteria, however, are not universal, in the sense that a *desirable* segmentation may need to reflect the morphology of a signal in addition to simply optimizing its parsimonious representation. It is thus of interest to construct criteria which also reflect such morphological structures.

2. MATHEMATICAL PRELIMINARIES

In this section we specify the local trigonometric bases we use and the best-basis approach to segmentation as introduced by Coifman and Meyer [5].

2.1. Local Trigonometric Bases

In any application of wavelet transforms one important issue is the type of wavelet to be applied. It can be seen from Fig. 1 that the three "active" intervals, P, QRS, and T, have sinusoidal or raised-cosine-like shapes although with quite different "periods" (different frequencies), while the ST segment is marked by its slowly varying shape. These features prompted us to choose local cosine basis functions [11].

A local trigonometric basis is constructed as the product of an oscillating waveform, such as a sinusoid, and a windowing function $g(t)$. This window is chosen to localize the support of the basis while preserving the orthogonality and complementarity properties between adjacent elements of the basis. These functions satisfy the same hierarchical properties as classical wavelets and can be generated in a systematic way starting with a basis mother function:

$$W_{j,k}(t) = \sqrt{\frac{2}{\mid I_j \mid}} g_j(t) \cos\left[\pi(k+1/2)\frac{t-t_j}{\mid I_j \mid}\right],$$

with $I_j = [c_j, c_{j+1})$ and $t_j = \frac{c_j + c_{j+1}}{2}$. The parameter k controls the frequency of oscillation and the subscript of $I_{(\cdot)}$ the extent and position of the support. Determining the best set of $I_{(\cdot)}$ is the goal of the *optimal* segmentation problem using a local trigonometric basis.

2.2. Segmentation by Best Basis Search

As first described by Coifman and Meyer [5] the search for the best segmentation of a given signal can be cast as a search of the leaves of a dyadic (binary) tree whose first resolution level represents the finest segmentation conjectured as reasonable for the signal. A cost function is defined in terms of criteria such as the ECM which determines the cost of a given segment I_j based on the coefficients $W_{j,k}(t)$. The tree pruning (best basis search) then proceeds upward from all leaves (*i.e.* from the finest segmentation) by comparing the combined cost of two "child" segments to that of their "parent" segment. A parent segmentation consists of a twinned (or merged) version of two child segments. The cost comparison is performed, following the development in [6, 11], by defining the functions $W_{ik}(t)$ in an interval I_i for $i = 1, \ldots, J$ and $k = 1, 2, \cdots$ as above. For purposes of search efficiency, we chose both the intervals and the total support (*i.e.* the observation interval of the signal $x(t)$) to be dyadic. Then if we denote the subspace spanned by $\{W_{ik}(t)\}$ by $\mathcal{H}_i$, $\mathcal{H}_i \oplus \mathcal{H}_{i+1}$ is spanned by

$$\begin{aligned} W_{i',k}(t) &= \sqrt{\frac{2}{\mid I_i \mid + \mid I_{i+1} \mid}} g_{i'}(t) \cdot \\ &\quad \cdot \cos\left[\pi(k+1/2)\frac{t-t_{i'}}{\mid I_i \mid + \mid I_{i+1} \mid}\right]. \end{aligned}$$

Thus, we can determine if the cost of the parent is smaller than the combined cost of the children. If yes, the segments are joined into a larger one; if not, the child segmentation is retained. The search proceeds in parallel up the branches until all branch searches have terminated. The resulting segmentation is then optimal for the given criterion.

3. NOVEL MODIFIED CRITERIA FOR BEST BASIS SEGMENTATION

3.1. Limitations of the Standard Criterion

The usual entropy based parsimonious ECM representation criterion [11] for segmentation is not adequate for ECG segmentation. Two examples of the failure of this approach can be seen in the top two graphs in Fig. 2, where the segmentation was performed using the standard best-basis algorithm. The data is from the MIT-BIH arrythmia database and shows some typical normal beats. The difference between the two graphs is due to different initial fine segmentations. A heuristic explanation of the failure of the entropy criterion is that the dynamic range of an ECG, and thus the variability of the local SNR, is quite large. Therefore, in regions where either the noise dominates or the signal is relatively static (*e.g.* the ST segment), there is no "payoff" of decreased entropy if adjacent segments are concatenated (i.e. choosing the parent node of the best basis tree over its children) and the initial fine (over-) segmentation survives almost intact.

3.2. Novel Segmentation Criterion

To mitigate this undesired effect, we conjecture that visual discrimination of the ECG segments employs not only amplitude but also smoothness and curvature (at least first and second derivative) information. Thus we propose a criterion which reflects not only the parsimony of the representation of a signal but also its smoothness. Our new criterion combines an appropriately constructed function of the expected smoothness of the signal with an entropy-like measure that reflects the "best" (most parsimonious) representation. Thus our combined criterion is constructed to

1 Account for the simplest representation as suggested by the ECG signal via an entropy function $\phi_1(\mathcal{W}_{(\cdot,\cdot)})$

2 Account for the morphology of the signal smoothness as well as its rate of variation by defining a function $\kappa(\mathcal{W}_{(\cdot,\cdot)})$ which is in turn used in the additional cost function $\phi_2\left(\kappa\left(\mathcal{W}_{(\cdot,\cdot)}\right)\right)$ for the search criterion, where $\phi_2(\cdot)$ is monotonically non-increasing.

Thus this criterion not only penalizes this tendency to over-segment, but also reflects the fact that the features of an ECG signal are expected to be relatively smooth by penalizing for increased curvature.

In general, κ can be chosen to be any discrete estimate of the relevant features, such as smoothness, that reflect the expected inherent signal qualities. To preserve the efficiency of the search, the function $\phi_2(t)$ is chosen to be convex and order-preserving (i.e. for weak majorizations of sequences). This function, which can be viewed as a penalizing function for lack of smoothness, when combined with $\phi_1(\mathcal{W}_{(\cdot,\cdot)})$, forms the basis of our search criterion,

$$\phi_s(\mathcal{W}_{(\cdot,\cdot)}) = \phi_1(\mathcal{W}_{(\cdot,\cdot)}) + \lambda\phi_2\left(\kappa\left(\mathcal{W}_{(\cdot,\cdot)}\right)\right), \qquad (1)$$

where λ is a parameter selected to adjust the penalty enforcement. Thus the comparison/selection process depends on the significant intrinsic features of the signal itself.

4. DISCUSSION

Examples of segmentation resulting from the algorithm is shown in Figures 2 and 3. In the former, parts (c) and (d) show the segmentation achieved by the combined cost criterion under the same initial conditions as parts (a) and (b) respectively. In the latter figure, the graphs show four different ECG signals recorded simultaneously from four different locations on the torso of the same subject as part of BSPM studies at Dalhousie University in Canada[8]. Considerable variability in the waveforms is evident; in particular the morphology of the QRS complex and the amplitude of the T wave vary at the different electrode locations. Nonetheless, the performance of the algorithm is quite consistent. We believe that this reflects the fact that the algorithm responds to the intrinsic qualities of the signal without relying on signal-dependent heursitics.

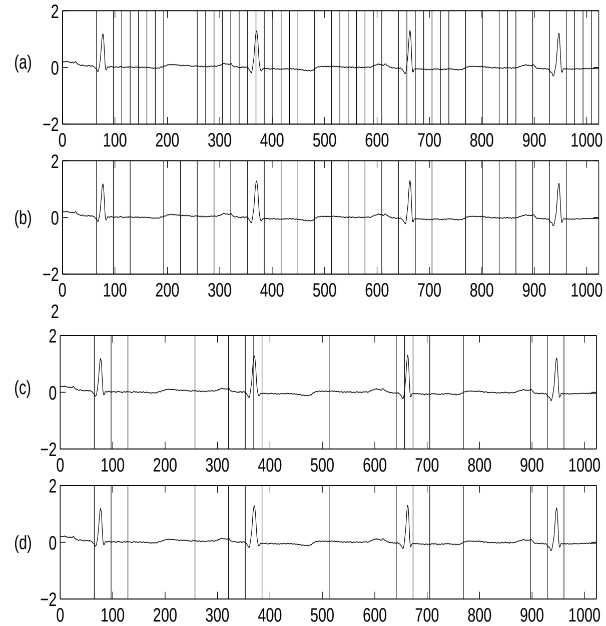

Figure 2. Example of segmentation using pure entropy criterion in parts (a) and (b) and using the combined criterion in parts (c) and (d). Parts (a) and (c) started from the one initial segmentation, parts (b) and (d) from another.

However it is clear that there are some problems remaining with the resulting segmentation. From both figures shown here it can be seen that the QRS segment endpoints are sometimes not as close to the actual QRS complex as might be desired, and the algorithm is not as reliable in discerning *both* beginning and endpoint of the P and/or T waves as it might be. There are several issues which remain in the design of the algorithm which relate to these limitations. Below we address each of these issues and propose some modifications.

- The search index set J is a control parameter which is used to determine the degree of refinement of the segmentation. The value of this parameter has a considerable effect on the final result. This suggests the need for a systematic determination of an appropraite value of J.

- The value of λ also influences the result. We are currently working on the implementation of an adaptive λ, $\lambda(x)$, to overcome this problem.

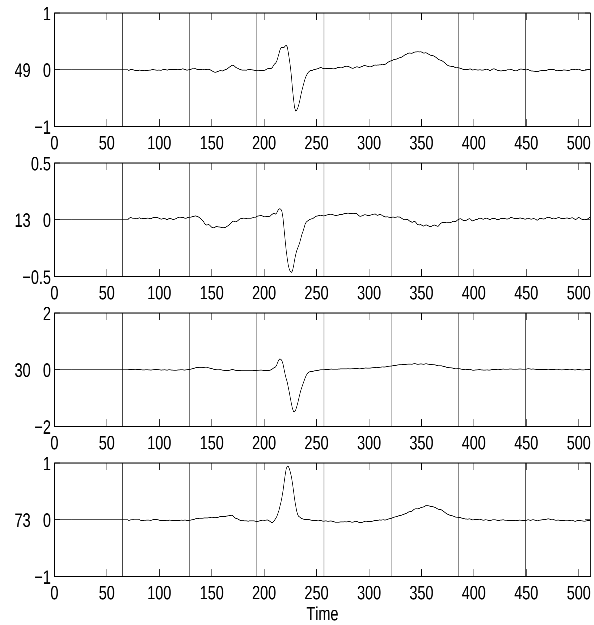

Figure 3. Examples of segmentation of simultaneous ECG signals from several locations on the torso of the same subject.

- The dyadic nature of the search imposes two arbitrary limitations in the search; one is the need for a power-of-two length for the original interval and the other is that only odd-even indexed combinations of adjacent segmentations are possible. For example, we consider whether the first and second segments should be combined at any level, but not whether the second and third should be combined. These limitations inhibit the ability of the algorithm to accurately refine the segmentation, and we are currently exploring several ways of overcoming them.
- The lack of time invariance introduced by the decimations introduced in the tree structure to perserve orthogonality cause a sensitivity to the choice of starting point. For example, this explains in part the fact that in Fig. 2(c) the demarcations fall with QRS for some beats but not for the subsequent one. We are exploring several expansions, including the possibility of parallel trees producing parallel segmentations, out of which a best segmentation is then chosen. This can essentially be achieved by searching for the most time-invariant representation as discussed, for instance, in [12].

5. CONCLUSION

We are quite encouraged by the present state of our ECG segmentation algorithm. In particular, the combined criterion we have developed provides an effective and flexible tradeoff between the goal of a parsimonious segmentation and sensitivity to the intrinsic qualitative features of the signal. We have employed **no** heuristics based on a preconceived notion of an ECG signal, and can easily include a few of the most general such rules as a further refinement. Before we do this, there are still aspects of the generic algorithm currently under development. Among these are the problem of how best to choose the initial window size, *i.e.* the initial fine segmentation), and how to incoporate a non-dyadic sampling in scale which will allow us to further refine the segmentation according to the structure of the actual signal.

Acknowledgment The authors wish to thank Dr. George Moody of HST, MIT, for access to the MIT-BIH database.

REFERENCES

[1] J. Pan and W.J. Tompkins, "A real-time QRS detection algorithm", *IEEE Trans. Biomed. Eng.*, vol. 32, pp. 230-236, 1985

[2] P. Laguna, R. Jané, and P Caminal, "Automatic detection of wave boundaries in multilead ECG signals: validation with the CSE database", *Comp. and Biomed. Res.*, vol. 27, pp. 45-60, 1994

[3] G. Li, C. Zheng, and C. Tai, "Detection of ECG characteristic points using Wavelet Transforms", *IEEE Trans. Biomed. Eng.*, vol. 42, pp. 21-28, Jan. 1995

[4] D.H. Brooks, H. On, and R.S. MacLeod, "Multidimensional multiresolution analysis of array ECG signals during PTCA procedures", *Proc. IEEE Symposium on Time-Frequency and Time-Scale*, pp. 552-555, Philadelphia, PA, Oct. 1994

[5] R.R. Coifman and Y. Meyer, "Remarques sur l'analyse de Fourier à fenêtre, série I,", C. R. Acad. Sci., Paris, vol. 312, pp. 259-261, 1991.

[6] Y. Meyer, translated and Revised by R.D. Ryan, "Wavelets, Algorithms & Applications", SIAM, 1993.

[7] R. Murray, S. Kadambe, and G.F. Boudreaux-Bartels, "Extensive analysis of a QRS detector based on the Dyadic Wavelet Transform", *Proc. IEEE Symposium on Time-Frequency and Time-Scale*, pp. 540-543, Philadelphia, PA, Oct. 1994.

[8] D.H. Brooks, R. V.Chary, H. Krim, H. On, and R.S. MacLeod, "Wavelet-based temporal segmentation and analysis of Body Surface Potential Maps during PTCA-induced ischemia, accepted for *Proc. IEEE EMBS Conference*, Sept. 1995, Montreal, Quebec, Canada

[9] M. Nakashzuka, H. Kikuchi, H. Makino, and I. Ishii, "ECG data compression by multiscale peak analysis", *Proc. ICASSP-95*, pp. 1105-1108, Detroit, MI, USA, May 1995

[10] S.G. Mallat and S. Zhong, "Characterization of signals from multiscale edges", *IEEE Trans. on Pattern Analysis and Machine Intell.*, vol. 14, no. 7, pp 710-732, July 1992.

[11] R.R. Coifman and M.V. Wickerhauser, "Entropy-based algorithms for best basis selection", *IEEE Trans. Info. Theory*, vol. 38, pp. 713-718, 1992

[12] J.-C. Pesquet, H. Krim, and H. Carfantan, "Time invariant orthonormal wavelet representations" to appear in *IEEE Trans. on Sig.Proc.*