Preface

This book is intended for graduate students collecting and/or analyzing electrocardiogram (ECG) data, industrial researchers looking to develop, test, and apply new ECG analysis tools (both hardware and software), or simply students or teachers looking for signal processing examples involving an intuitive yet complex signal. Initially, this book was conceived primarily as a vessel for collecting together current research in ECG analysis that we considered to be important guides of how to apply complex techniques to this field. However, it is not our intention to simply present a collection of colleagues' papers with an index and contents. The articles we have selected have been specifically commissioned for this book with a request to the authors that they explain how they achieved their results, why they represent a significant improvement of existing techniques, and why the particular method they used is more appropriate for the task.

Not only do we wish to present an overview of many of the most interesting and useful advanced techniques currently available (together with an analysis of the validity of the assumptions these techniques require), we also wish to provide the reader with the relevant background on where distortion, noise, and errors can creep into an experiment or analysis. This includes not only the choice of (digital) mathematical processing operations, but also the design and choice of sensors, hardware, software, data storage, data transmission methods, databases on which the techniques are evaluated, and probably most importantly of all, the metrics by which we determine one technique is superior to another. There is, of course, one other source of noise or error: the subject itself. Choosing the correct database or population, and including or excluding particular data is an essential part of experimental design. Furthermore, the quantification of a section of data in terms of noise or signal quality depends largely on the type of analysis that is being performed. For example, muscle noise can be a signal that indicates activity, whereas abnormal beats may indicate a lapse in stationarity and invalidate Fourier methods.

The analysis of the electrocardiogram as a diagnostic tool is a relatively old field and it is therefore often assumed that the ECG is a simple signal that has been fully explored. However, there remain difficult problems in this field that are being incrementally solved with advances in techniques from the fields of filtering, pattern recognition, and classification, together with the leaps in computational power and memory capacity that have occurred over the last couple of decades.

The following is a nonexhaustive list of many of the key remaining challenges in ECG signal processing today:

- *Reliable P wave identification.* P waves are usually a low-amplitude feature that can often become subsumed by the baseline noise in a signal. Detection of the onset of atrial activity at the start of the P wave is important for heart rate variability studies. However, the enormity of this problem has led to a pervasive analysis of beat-to-beat intervals based upon the QRS complex as a fiducial marker.

- *Reliable QT interval estimation.* Similarly, despite the relatively large amplitude of the QRS complex and T wave, the *onset* of the Q wave and *offset* of the T wave are difficult features to measure (even for highly trained experts). Changes in the QRS complex, ST segment, and T wave morphology due to heart rate and sympathetic nervous system changes make this problem particularly acute.
- *Distinguishing ischemic from nonischemic ST changes.* Even in subjects who are known to have myocardial ischemia, ST changes are not considered a basis for definitive diagnosis of individual episodes of ischemia. This is because the ST segment changes with a subject's body position due to the movement of the ECG electrodes relative to the heart. Recent attempts to identify such artefacts are promising (see the Computers in Cardiology Challenge 2003), but the road to a full working system (particularly for silent ischemia) is not clear.
- *Reliable beat classification in Holter monitoring.* Ambulatory monitoring presents many challenges since the data collection process is essentially unsupervised, and evolving problems in the acquisition often go undetected (or at least are not corrected). One major problem is the high level of in-band noise encountered (usually muscle- or movement-related). Another problem for ambulatory monitoring is the degradation in electrode contact over time, leading to a lower signal-to-noise ratio. Since the P wave is usually a low-amplitude feature, reliable detection of the P wave is particularly problematic in this environment.
- *Robust, reliable in-band signal filtering or source separation* (such as muscle noise removal and fetal-maternal separation). Principal and independent component analysis has shown promise in this area, but problems persist with nonstationary mixing and lead position inaccuracies/changes. Model-based filtering methods have also shown promise in this area.
- *Identification of lead position misplacements or sensor shifts.* Most classifiers assume that the label for the clinical lead being recorded is known and make assumptions about waveform morphology, amplitude, and polarity based upon the lead label. If the lead is misplaced, or shifts during recording, then misclassifications or misdetections will probably occur. In particular, the maternal-fetal signal mixing is problematic in this respect. In this case two cardiac sources are moving with respect to each other, sometimes without correlation, and sometimes entrained on one or more scales.
- *Reliable confidence measures.* The ability of an algorithm to report back a "level of trust" associated with its parameter estimates is very important, particularly if the algorithm's output is fed to another data fusion algorithm.
- *The inverse problem.* It is well known that no unique solution exists for the inverse problem in ECG mapping. Attempts to reconstruct the dipole moment in the ECG have had some degree of success, and recent models for the ECG have proved useful in this respect.

- *ECG modeling and parameter fitting.* To date there exist many accurate representations for the ECG from a cellular level up to more phenomenological models. However, at best, each of these models can accurately reproduce the ECG waveform for only a short period of time. Although models also exist for variations in the beat-to-beat timings of the heart on both short and long scales, there are no known models that can reproduce realistic dynamic activity on all scales, together with an accurate or realistic resultant ECG. Recent work in the fitting of real data to ECG and pulsatile models is certainly promising, but much work is needed in order to ascertain whether the resultant parameters will yield any more information than traditional ECG metrics.
- *The mapping of diagnostic ECG parameters to disease classifications or predictive metrics.* Neural networks have shown promise in this area due to their ability to extrapolate from small training sets. However, neural networks are highly sensitive to outliers and the size and distribution of the training set. If only a few artifacts, or mislabeled patterns, creep into the training set, the performance of the classifier is significantly reduced. Conversely, over-restricting the training set to include too few patterns from a class results in over-training. Another known problem from the use of neural networks is that it is difficult to extract meaning from the output; a classification does not obviously map back to the etiology.
- *Global context pattern analysis.* Factoring in patient-specific history (from minutes, to hours, to years) for feature recognition and classification is required if a full emulation of the clinical diagnostic procedure is to be reproduced. Hidden Markov models, extended Kalman filters, and Bayesian classifiers are likely candidates for such a problem.
- *The development of closed-loop systems.* It is always problematic for a clinician to relinquish the task of classification and intervention for personal, legal, and ethical reasons. However, with the increasing accuracy of classifiers and the decreasing costs of machines relative to human experts, it is almost inevitable that closed-loop devices will become more pervasive. To date, few such systems exist beyond the classic internal cardioverters that have been in use for several decades. Despite promising advances in atrial fibrillation detection, which may soon make the closed-loop injection of drugs for this type of condition a reality, further work is needed to ensure patient safety.
- *Sensor fusion.* Information that can be derived from the ECG is insufficient to effectively solve many of the above problems. It is likely that the combination of information derived from other sensors (such as blood pressure transducers, accelerometers, and pulse oximeters) will be required. The paradigm of multidimensional signal analysis is well known to the ECG signal analyst, and parallel analysis of the ECG (or ECG-derived parameters) almost always enhances an algorithm's performance. For instance, blood pressure waves contain information that is highly correlated with the ECG, and analysis of these changes can help reduce false arrhythmia alarms. The ECG is also highly correlated with respiration and can be used to improve respiration rate estimates or to facilitate sleep analysis. However, when the associated signals do not

present in a highly correlated manner, the units of measurement are different (so that a 20-mmHg change does not mean the same as a 20-bpm change, for example) and their associated distributions differ, the task at hand is far more difficult. In fact, normalizing for these differences, and building *trust* metrics to differentiate artifactual changes from real changes in such signals, is one of the more difficult challenges in ECG signal processing today.

In order to address these issues, we have attempted to detail many of the key relevant advances in signal processing. Chapter 1 describes the physiological background and the specific autonomic mechanisms which regulate the beat-to-beat changes of timing and morphology in the ECG, together with the cause and effect of breakdowns in this mechanism. Chapter 2 presents an overview of the primary issues that should be taken into account when designing an ECG collection system.

Chapter 3 presents an overview of the relevant mathematical descriptors of the ECG such as clinical metrics, spectral characteristics, and beat-to-beat variability indices. Chapter 4 presents an overview of simple, practical ECG and beat-to-beat models, together with methods for applying these models to ECG analysis. Chapter 5 describes a unified framework for linear filtering techniques including wavelets, principal component analysis, neural networks, and independent component analysis. Chapter 6 discusses methods and pitfalls of nonlinear ECG analysis, with a practical emphasis on filtering techniques.

Chapter 7 provides an overview of T wave alternan methodologies, and Chapter 8 presents a comparative study of ECG derived respiration techniques. Chapter 9 presents advanced techniques for extracting relevant features from the ECG, and Chapter 10 uses these techniques to describe a robust ST-analyzer. Chapter 11 presents a wavelet and hidden Markov model–based procedure for robust QT-analysis. Chapter 12 describes techniques for supervised classification and hybrid techniques for classifying ECG metrics, where the data labels are already known, and Chapter 13 presents unsupervised learning techniques for ECG pattern discovery and classification.

Although many of the basics of ECG analysis are presented in Chapter 1, this is simply to draw the reader's attention the etiology of many of the problems we are attempting to solve. As a thorough grounding in the basics of ECG signal processing, the reader is referred to Chapters 7 and 8 in Sörnmo and Laguna's recent book *Bioelectric Signal Processing in Cardiac and Neurological Applications* (Elsevier, 2005). The reader is assumed to be familiar with the basics of signal processing and classification techniques. Furthermore, these techniques are necessarily implemented using a knowledge of computational programming. This book follows the open-source philosophy that the development of robust signal processing algorithms is best done by making them freely available, together with the labeled data on which they were evaluated. Many of the algorithms and data sets described in this book are available from the following URLs: http://www.ecgtools.org, http://www.physionet.org, and http://alum.mit.edu/www/gari/ecgbook.

Most of these algorithms have been written either in C or Matlab. Additionally, Java applet versions of selected algorithms are also available. Libraries for reading these databases are also freely available. We hope that through these URLs this book will continue to evolve and add to the growing body of open (repeatable) biomedical research.

It is important to acknowledge the giants whose shoulders we stand upon. We would therefore like to thank the many who offered advice and help along the way, including George Moody, Roger Mark, Rachel Hall Clifford, Emmeline Skinner McSharry, Raphael Schneider, Ary Goldberger, Julie Greenberg, Lionel Tarassenko, Lenny Smith, Christopher James, Steve Roberts, Ken Barnes, and Roy Sambles. We would also like to acknowledge the many organizations that have facilitated this project, including the National Institute of Biomedical Imaging and Bioengineering (under Grant R01 EB001659), the Royal Academy of Engineering, the Engineering and Physical Sciences Research Council, United Kingdom (grants GR/N02641 and GR/S35066), the European Union's Sixth Framework Programme (Grant 502299), and, of course, Artech House Publishing for all their editorial help. Finally, and above all, we thank all the authors who have contributed their valuable knowledge, time, and energy to this book.

Gari D. Clifford
Cambridge, Massachusetts

Francisco Azuaje
Jordanstown, United Kingdom

Patrick E. McSharry
Oxford, United Kingdom
Editors
September 2006