A LightWAVE client for Semi-automated Annotation of Heart Beats from ECG Time Series

Luca Citi[1], Claudia Olariu[1], Riccardo Barbieri[2,3]

[1] School of Computer Science and Electronic Engineering, University of Essex, Colchester, UK
[2] Dept. Anesthesia, Massachusetts General Hospital – Harvard Medical School, Boston, MA, USA
[3] Dept. Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

LightWAVE is an open-source web-based software for viewing ECGs and other physiologic waveforms and associated annotations (such as heart-beat markers). At present, most users run the raw ECG through an automated QRS detector and later use LightWAVE to review and correct the detected heart beats. Although this 2-stage procedure may work well with clean signals, it is inefficient and time consuming when the recordings are contaminated by noise, artefacts or recurring ectopic events. To overcome this limitation, we customized the LightWAVE client to allow automated and semi-automated annotation of heart beats from ECG time series. In semi-automatic mode, the algorithm automatically identifies most QRS complexes and stops — asking for manual intervention — whenever the confidence of a detection falls below a given threshold. Additionally, the software now shows the series of inter-beat intervals, which is an invaluable tool to easily spot R-wave misdetections and genuine arrhythmias. The new client introduces further additional features compared to the standard version, for example the possibility of importing raw signals from local CSV files and of exporting the current plot in SVG format. Overall, our customized client extends the functionality of LightWAVE and brings it closer to one of its design goals, i.e. to provide a comfortable and efficient method of annotating physiologic data.

1. Introduction

LightWAVE (http://physionet.org/lightwave/) [1] is an open-source web-based software for viewing ECGs and other physiologic waveforms and associated annotations (such as heart-beat markers) created by George Moody as a replacement for the WAVE waveform and annotation viewer and editor. LightWAVE is part of PhysioNet (http://physionet.org/) [2], a research resource intended to stimulate current research and new investigations in the study of complex biomedical and physiologic signals. Each month, about 45000 visitors worldwide use PhysioNet, retrieving about 4 terabytes of data. LightWAVE allows PhysioNet's users to view physiologic waveform and their associated annotations. It also allows PhysioNet administrators and users to add and revise annotations in the course of creating new data collections and curating existing ones. Most often, users are interested in marking the times of occurrence of heart beats, usually by locating the QRS complexes in ECG signals. Automated QRS detectors can do this task quite well in clean signals, but require supervision since they can be misled by noise and artefacts. This paper presents a customized LightWAVE client that implements an automated QRS detection algorithm as well as a semi-automated one to assist the user in the challenging and time consuming task of annotating physiologic data corrupted by noise or frequent artefacts.

1.1. Automatic heart beat detection

Electrocardiogram (ECG) signals are essential in the diagnosis of heart diseases, as they convey information regarding the morphology of the heart beats as well as the heart rate. Heart-beat detection and analysis is an essential part of this process and can reveal, for example, tachycardias, bradycardias, extra systoles and other types of arrhythmias. Several heart-beat detection algorithms have been implemented, such as SQRS [3], WQRS [4], and ECGPUWAVE [5]. They normally consist of a noise reduction part (e.g., FIR filters), a threshold adaptation part (adjusting a threshold to the data in order to improve the accuracy of the detections, and in particular, to reduce the number of false positives), and a decision rule based on feature extraction.

The WQRS algorithm consists of a three stage process, with much focus set on noise reduction. It first applies a low-pass filter at 16 Hz, followed by a curve-length transformation for multi-channel ECG [4], and selection of a time window. The process is then completed by the final stage, the decision rule, based on adaptive thresholding and local backward search strategies.

Pan and Tompkins proposed another algorithm for QRS detection, ECGPUWAVE [5]. This algorithm follows the three-stage pattern mentioned above, however it integrates a complex digital bandpass filter with the aim of increasing sensitivity and decreasing the number of false positive detections. The adaptive thresholding system is also designed to accommodate changes in the heart rate and QRS morphology. The digital filtering comprises of a $5 - 15$ Hz bandpass filter (implemented as the cascade of a high-pass and a low-pass filter with integer coefficients) followed by a derivative filter. The filtered signal is used in the final stage, where a squaring function and a set of rules are applied to the signal to localize the fiducial points.

The current state of the art algorithm in terms of QRS detection accuracy is a processor based on the Quadratic Spline Wavelet Transform (or QSWT). Its detection rates reach up to 99.31 - 99.70% across all recordings from the MIT-BIH arrhythmia database. The details of the methodology used are described in detail in [6].

One of the QRS detection algorithms used as a base for this project is the SQRS algorithm [3]. It represents a good compromise between detection performance, computational cost and ease of implementation. It follows the pattern presented above, using a simple FIR filter and an adaptive threshold. It will be thoroughly described in Section 2.1.

Along with SQRS, this paper presents a new approach to QRS complex detection, in the form of a semi-automated heart beat detection algorithm, which was previously developed by the authors and has now been adapted and integrated in a customized LightWAVE client.

1.2. The LightWAVE architecture

LightWAVE is made of three components: the server, the scribe, and the client. The server is a CGI application that retrieves the recordings from the data repository (normally PhysioNet's PhysioBank) and delivers them in response to JSON requests generated by the client. The scribe (lw-scribe) is an independent server-side component that receives edit logs uploaded by the LightWAVE client and allows the user to permanently store them on the server. The client is a waveform and annotation viewer and editor that communicates with the server to obtain raw data, and with the scribe to store the local changes made to the annotations. It runs within the user's web browser and is portable across all popular browsers and platforms. It is a web application written in JavaScript using the jQuery library and does not require compiling. To customize the LightWAVE client, one simply needs to download its source files and edit the Javascript code to implement additional features. In the following, we present a customized version of the LightWAVE client specifically created with the aim of making the tedious process of annotating long and noisy ECG recordings more efficient.

2. Customized client

Our customized LightWAVE client presents several improvements over the standard version. Most notably, it implements an automated (SQRS) and a semi-automated beat-detection algorithm. The latter also comes with visual hints to help the user make the best guess when the semi-automated procedure requires manual intervention. This section explains the implementation of these improvements in detail.

2.1. SQRS automated algorithm

When the "New sqrs annotation set" button is pressed, the software creates a new annotation set and then executes the SQRS algorithm [3] for automated QRS detection. Our Javascript implementation follows the C code available on Physionet quite closely, also owing to the similar syntax of the two languages. SQRS bears the shape of a standard algorithm of its class: a noise filter, an adaptive thresholding system and a decision rule. The noise filter is based on a 10-tap FIR filter with pre-defined weights optimized for signals sampled at approximately 250 Hz.

As in the original paper [3], the filtered signal is analysed in windows of 2 seconds, a length that was chosen empirically by the authors under the assumption that at least one and at most five QRS complexes (or heart contraction indicatives) can appear within this amount of time in the case of adult human ECGs. The threshold adaptation in this algorithm happens according to the number of potential QRS detections within the 2-second window. Should the number of slopes be zero, the threshold is lowered by a 16^{th} of its value. In case five or more slopes are detected, the threshold is risen by a 16^{th} of its value. The threshold has a minimum and a maximum level: 0.5 mV and 5 mV, respectively, which correspond to roughly a 5–15 Hz range in terms of frequency.

The decision rule is rather complex. Briefly, if between two and four slopes have been found, they are annotated as normal QRS complexes using LightWAVE's "edlog" method. If there are 5 or more slopes, it is deemed that an artefact has been found and the signal will be annotated accordingly.

If the signal's sampling frequency is below 200 Hz, the algorithm uses a 5-tap FIR filter and thresholds between 0.25 mV and 2.5 mV, like in Physionet's "SQRS125" algorithm.

2.2. Semi-automated detection algorithm

Our semi-automated detection algorithm is based on a Matlab® tool that was developed by the authors in 2007

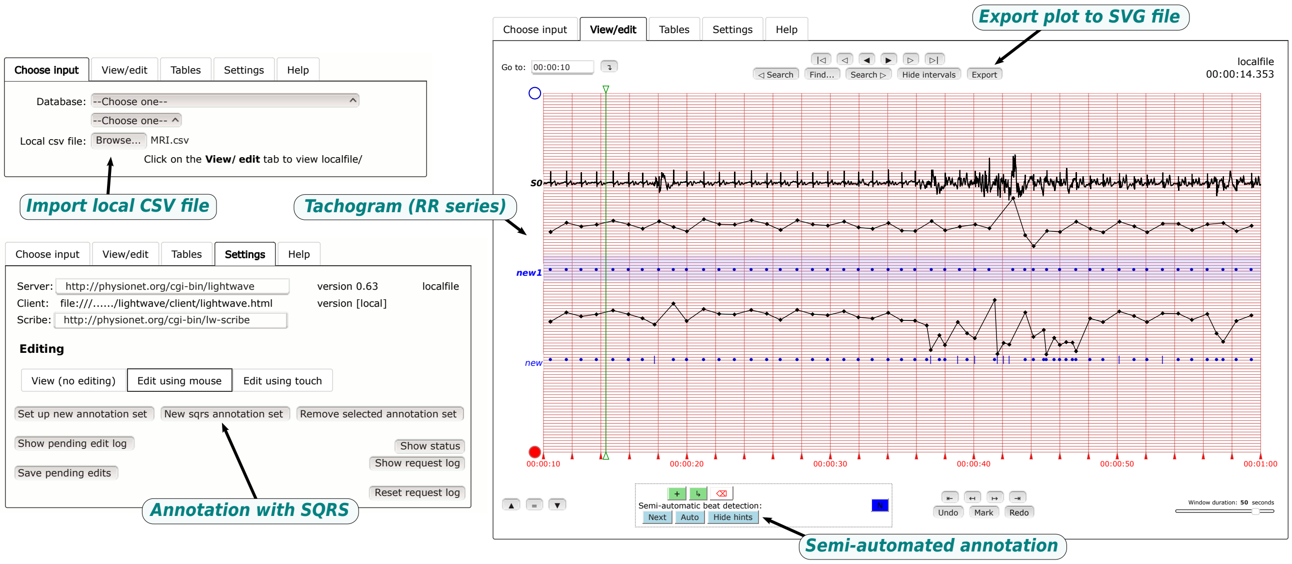

Figure 1. Overview of the additional features introduced in the customized LigthWAVE client.

and has been successfully used internally ever since to process ECG time series, including several 52-hour recordings from a sleep-deprivation dataset [7].

First, the ECG signal following the last R-wave is convoluted with a Mexican hat wavelet, i.e. the second derivative of a Gaussian function. This step can be interpreted both as a band-pass filter and as a matched filter for the QRS complex. In the following, we will refer to the result of this convolution as the shape predictor $w(t)$. Second, the information about the most recent successful identifications is combined to provide a probability distribution of the location of the current heartbeat. Like the "NGUESS" algorithm from PhysioNet, the prediction uses the observation that most of the short-term variability in normal sinus rhythm is due to respiratory sinus arrhythmia (RSA), and therefore that past beats having the same "phase" within the respiratory cycle may be better predictors. Assuming a conventional respiratory cycle lasting B interbeat intervals, the prediction of the n-th beat, where $n = \hat{k}\,B + j$, uses past intervals with indices in the set $I = \{kB + j\}_{k=k_0}^{\hat{k}-1}$. For each $i_k \in I$, an unnormalised Gaussian of width σ and amplitude $g\rho^{\hat{k}-k}$ is placed at time $\mathrm{R}_{i_k} - \mathrm{R}_{i_k-1}$, where $\rho \leq 1$ is a forgetting factor. We sum all these scaled and shifted Gaussian functions and increment the result by one to obtain the history predictor $h(t)$ (drawn in light blue in Fig. 2).

The two predictors are combined together as $p(t) = w(t)\,h(t)$ (drawn in green in Fig. 2) for final determination. The signals $w(t)$ and $p(t)$ are scanned in chronological order looking for local maxima of $w(t)$ with the highest value of $p(t)$. If a candidate at time t_1 is found, a further candidate at time $t_2 > t_1$ is only accepted if $p(t_2) > \nu p(t_1)$ with $\nu > 1$, in order to avoid skipping a beat and detecting the second next beat instead.

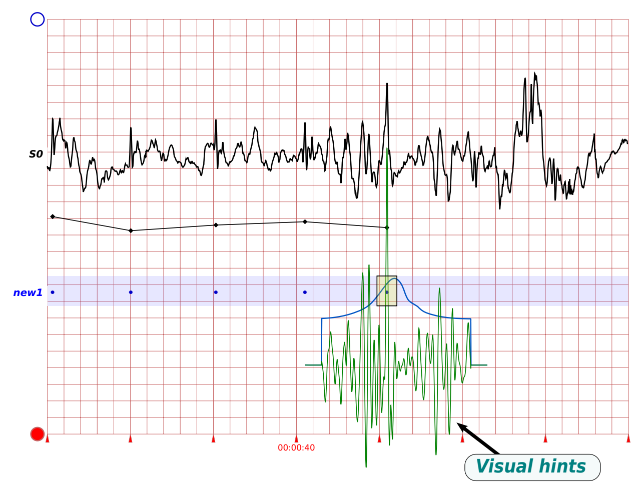

Figure 2. User interface presented to the user when annotating the raw ECG signals. As visual aids, the tachogram (in black) as well as the predictors, $p(t)$ (green) and $h(t)$ (light blue) used by the semi-automated algorithm are shown to the user.

A threshold determines when the algorithm is unable to provide a single solution with sufficient confidence. In these cases, the procedure stops and prompts for user intervention. To assist the user in making the best guess about the position of a heart beat hidden in noise and artefacts, one or more predictors can be displayed as visual hints, as shown in Fig. 2. When the automated detection is resumed, the algorithm learns from the manual selection performed by the user by updating the history predictor $h(t)$ and by adaptively adjusting the scale of the wavelet using a greedy hill-climber strategy.

2.3. Other improvements

In addition to the beat detection algorithms, our customized client presents some further improvements over the standard version, as summarized in Fig. 1. For example, the current plot can be exported as SVG file that can be opened using one of the many vector graphics editors that support this file format, including the open-source software "inkscape" (http://inkscape.org). This procedure was also used to produce Fig. 2.

In the original client, the cursor (the vertical green line on the left of the large plot in Fig. 1) follows the mouse continuously when in the plotting area. This behaviour may be undesirable, especially when trying to reach the "nudge left/right" buttons to apply minor corrections to the position of the cursor before committing the edit. In the customized client, a right click of the mouse freezes the current position of the cursor temporarily, allowing the user to select the approximate position of a fiducial point and then click the "nudge left/right" to perform fine adjustments.

Currently, users willing to use LightWAVE on their own data must upload them to PhysioNetWorks or setup a web server and the LightWAVE server. With the new client, users can simply load signals from a local CSV file using the button shown on the top left panel of Fig. 1. At the moment this functionality is restricted to raw signals sampled at a common rate, with no support for importing annotations.

Finally, as shown in Fig. 1, the plot now includes a tachogram (i.e. the series of beat-to-beat RR-intervals), which is an invaluable tool to easily spot R-wave misdetections and genuine arrhythmias.

3. Real-data tests

We tested the software on time series with a wide range of different characteristics, including records from the MIT-BIH arrhythmia database and signals recorded during MR imaging sessions. Especially on noisy data, its use dramatically reduced the time needed to produce annotation files for previously unannotated records.

4. Conclusions

This paper presents a customized LightWAVE client for automated and semi-automated annotation of heart beats from ECG time series. The new client presents some further improvements over the standard version, for example the possibility of importing raw signals from local CSV files and of exporting the current plot in SVG format. Overall, our customized client extends the functionality of LightWAVE and brings it closer to one of its design goals, i.e. to provide a comfortable and efficient method of annotating physiologic data. Future work aims at further improving and polishing the interface so it can be used and tested by other researchers and eventually integrated in the main LightWAVE client.

References

[1] Moody GB. Lightwave: Waveform and annotation viewing and editing in a web browser. In Proceedings of Computing in Cardiology. 2013; 17–20.

[2] Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation 2000;101(23):e215–e220.

[3] Engelse W, Zeelenberg C. A single scan algorithm for QRS-detection and feature extraction. In Proceedings of Computers in Cardiology. 1979; 37–42.

[4] Zong W, Moody GB, Jiang D. A robust open-source algorithm to detect onset and duration of QRS complexes. In Proceedings of Copmuters in Cardiology. 2003; 737–740.

[5] Pan J, Tompkins WJ. A real-time QRS detection algorithm. IEEE Transactions on Biomedical Engineering 1985;BME-32(3):230–236.

[6] Ieong CI, Mak PI, Lam CP, Dong C, Vai MI, Mak PU, Pun SH, Wan F, Martins RP. A 0.83-uW QRS detection processor using quadratic spline wavelet transform for wireless ECG acquisition in 0.35-um CMOS. IEEE Transactions on Biomedical Circuits and Systems 2012;6(6):586–595.

[7] Citi L, Klerman EB, Brown EN, Barbieri R. Point process heart rate variability assessment during sleep deprivation. In Proceedings of Computing in Cardiology. 2010; 721–724.

Address for correspondence:

Luca Citi
School of Computer Science and Electronic Engineering
University of Essex
Wivenhoe Park
Colchester, CO4-3SQ, U.K.
E-mail: lciti@ieee.org