Reinforcement Learning for Dynamic Filter Design: Application to ECG Baseline Wander Removal

Masoud Nateghi[1]

[1] Department of Biomedical Informatics, Emory University, USA

E-mail: masoud.nateghi@emory.edu

Background

Electrocardiogram (ECG) is a vital diagnostic tool that provides critical insights into cardiac health by recording the electrical activity of the heart. However, the acquisition of clean ECG signals is often challenged by various types of noise, including power line interference, muscle artifacts, baseline wander, and motion artifacts, which can mask important diagnostic features and potentially lead to misdiagnosis. This makes ECG denoising a crucial preprocessing step in both clinical settings and remote health monitoring systems. Effective denoising techniques help preserve the morphological characteristics of key ECG components like P waves, QRS complexes, and T waves while removing unwanted noise, enabling healthcare providers to make more accurate diagnoses of cardiac conditions.

Baseline wander (Figure 1) is one of the most significant sources of noise in ECG recordings, primarily caused by patient respiration, body movements, and poor electrode contact. Its removal is crucial for accurate ECG interpretation since baseline drift can mask subtle but clinically significant ECG features and complicate both manual analysis and automated detection algorithms. However, removing baseline wander presents a challenging trade-off: while baseline wander typically occupies the very low frequency spectrum (usually below 0.5 Hz), it partially overlaps with some clinically important low-frequency ECG components. Conventional approaches using fixed low-pass or high-pass filters can inadvertently distort critical diagnostic

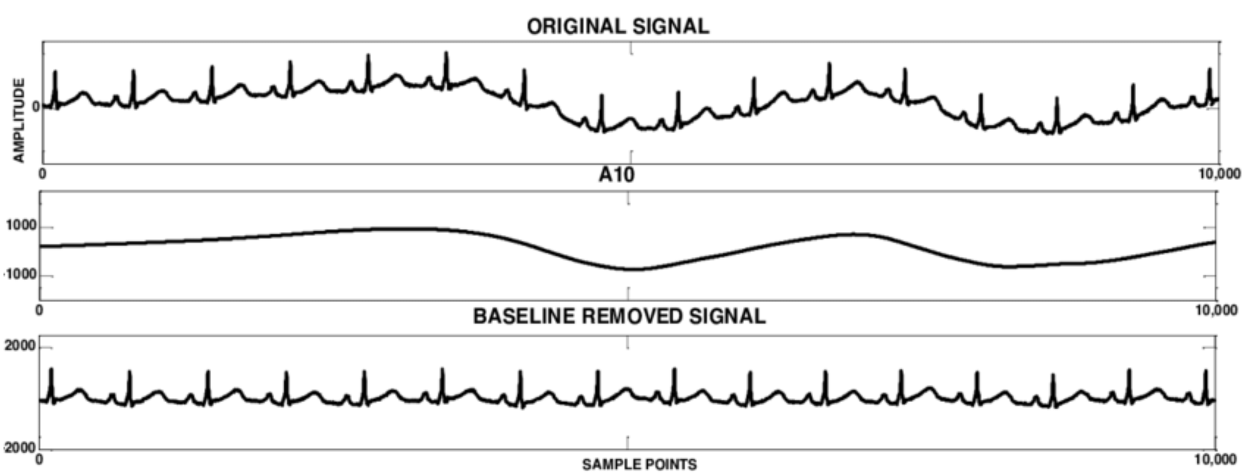

Figure 1: From top to bottom: ECG signal contaminated by baseline wander, baseline wander signal, and the denoised ECG signal [1].

features of the ECG, particularly the ST segment and QT interval, which contain valuable low-frequency components. This is especially problematic because these intervals are vital markers for various cardiac conditions - ST segment changes can indicate myocardial ischemia or infarction, while QT interval abnormalities may signal increased risk of life-threatening arrhythmias. Therefore, we propose an RL-based adaptive filtering approach that can dynamically adjust its cutoff frequency based on the specific characteristics of different ECG segments appears essential. Such an approach would provide more precise baseline wander removal while preserving the integrity of crucial low-frequency diagnostic features in the ECG signal.

Problem Formulation

The ECG denoising problem can be formulated as a reinforcement learning task with the following components:

Environment: The environment consists of the noisy ECG recordings, containing both cardiac electrical activity (the signal of interest) and various noise components, particularly baseline wander.

States: The observable states are the raw ECG measurements, which can be represented as:

$$f_i = g_i + b_i$$

where $\boldsymbol{g_i}$ is the clean ground-truth ECG signal and $\boldsymbol{b_i}$ is the baseline wander component of the i-th segment (and are clearly a function of time). States are defined as different segments of the noisy ECG signal $\boldsymbol{f_i}$ (with or without overlap).

Reward: The reward function $\boldsymbol{r_i}$ can be computed through two alternative approaches. First, using a simple negated mean-squared error (MSE) between the ground-truth clean signal g_i and the denoised input signal $\hat{g}_i$ for the i-th segment:

$$r_i = -\|g_i - \hat{g}_i\|_2^2$$

Alternatively, we can use a periodicity measure calculated with the Rayleigh quotient for the whole segment [2, 3]. For a given signal y(t), this measure is calculated as:

$$\epsilon(y(t), \tau) = \frac{E_t\{y(t)y(t+\tau)\}}{E_t\{y(t)^2\}}$$

Where τ is the period of the estimate which can be estimated by peak detection and RR-interval.

This approach leverages the fundamental difference between the quasi-periodic nature of the ECG signal and the aperiodic characteristics of baseline wander. Therefore, the reward can be formulated as:

$$R = \frac{\epsilon(\hat{g}(t), \tau)}{\epsilon\left(\hat{b}(t), \tau\right)}$$

Where $\hat{g}(t)$ is the whole estimated clean ECG signal and $\hat{b}(t)$ is the estimated baseline wander for the whole period. This ratio should be maximized, as we expect high periodicity in the clean ECG signal (numerator) and low periodicity in the baseline wander component (denominator).

Action: The action on environment is generating the optimal parameters $\boldsymbol{s_i}$ for the filtering algorithm which will be applied on the current segment of the noisy ECG signal. Through this process, the filter produces two estimates:

$$\hat{b}_i = LPF(f_i, s_i)$$

$$\hat{g}_i = f_i - \hat{b}_i$$

RL Agent Design: The agent receives rewards and states while attempting to predict optimal filtering algorithm parameters that preserve the main signal components while removing baseline wander. The parameter can be a simple input to a filtering function implemented in Python or MATLAB, or it can be the coefficients of a filter. Multiple filters with different parameter sets can be evaluated and compared. For simplicity, we will begin with a simple first-order low-pass filter developed in the OSET package (lp_filter_zero_phase) [4]. The figure below can describe the overall architecture of the RL-agent which will be used for this project [5].

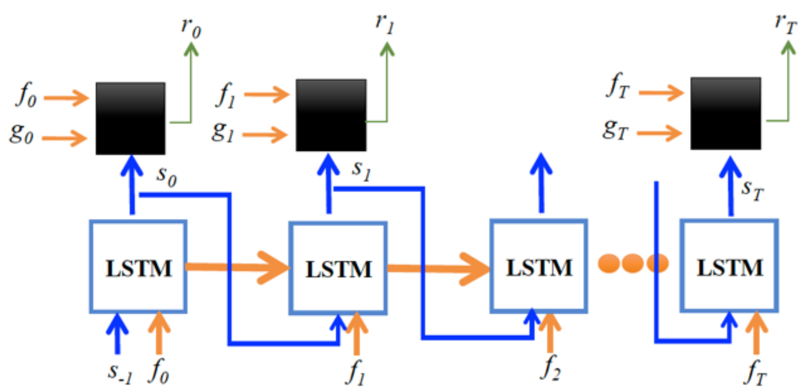

Figure 2: Proposed architecture. g_t and f_t are clean and noisy input frames, while s_t and r_t are action and reward values at time step t, respectively. At each time step, our model utilizes s_t and f_t to find the best set of control parameters that maximize r_t of the ECG denoising black-box algorithm [5].

The diagram shows several key components. Here, f_i represents different segments of the noisy ECG signal, which constitutes the current state. The parameter s_i represents the action chosen by the RL agent and serves as the input parameter for the filtering algorithm. The central black box represents the filtering algorithm, which processes the noisy ECG using the parameter s_i. Using the ground truth signal g_i, the reward r_i is calculated to evaluate the filtering performance. Importantly, g_i is only used for reward calculation and is not used in the black box algorithm in any other way. Through this setup, the RL agent learns to maximize the reward by interacting with the environment across multiple episodes.

We can summarize the whole agent-environment interaction in the figure below:

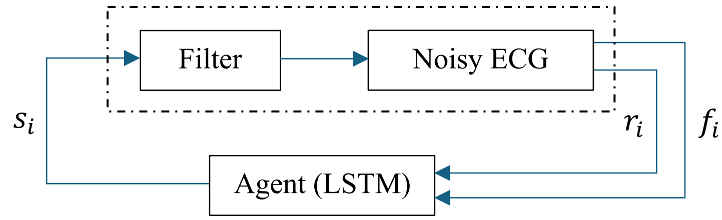

Figure 3: Block diagram of the agent-environment interaction. The dashed line denotes the environment.

RL algorithm

As detailed in the preceding sections, we adopted a purely data-driven approach to adaptively tune the filtering algorithm's parameters, without relying on any model for environment. Given that our action space is continuous—the agent determines the optimal input parameter for the filtering algorithm along a continuous spectrum—a model-free reinforcement learning approach was deemed most appropriate. We selected the REINFORCE algorithm for this task for several reasons. First, it can optimize the ECG baseline wander removal algorithm's parameters without requiring knowledge of the system's underlying mechanics. Second, it naturally accommodates continuous parameter adjustments in real-time. Finally, its ability to learn directly from outcome feedback makes it particularly suitable for the sequential decision-making process inherent in frame-by-frame ECG denoising.

Experiments and Benchmarks

To validate our proposed reinforcement learning-based adaptive filtering approach, we will conduct comprehensive experiments using the PTB-XL dataset [6]. PTB-XL is one of the largest publicly available electrocardiogram datasets, containing 21,799 clinical 12-lead ECG records from 18,869 patients. This dataset is particularly suitable for our study as it includes ECGs of varying quality and contains detailed annotations of ECG characteristics and diagnostic findings.

References:

[1] Sadhukhan D, Mitra M. ECG noise reduction using linear and non-linear wavelet filtering. InProc. of Int. Conf. on Computing, Communication & Manufacturing (ICCCM), Howrah, India 2014 Dec 22 (pp. 252-258).

[2] Kharabian S, Shamsollahi MB, Sameni R. Fetal R-wave detection from multichannel abdominal ECG recordings in low SNR. In2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society 2009 Sep 3 (pp. 344-347). IEEE.

[3] Sameni R, Jutten C, Shamsollahi MB. A deflation procedure for subspace decomposition. IEEE Transactions on Signal Processing. 2009 Nov 24;58(4):2363-74.

[4] Kazemnejad A, Karimi S, Gordany P, Clifford GD, Sameni R. An open-access simultaneous electrocardiogram and phonocardiogram database. Physiological Measurement. 2024 May 15;45(5):055005.

[5] Fakoor R, He X, Tashev I, Zarar S. Reinforcement learning to adapt speech enhancement to instantaneous input signal quality. arXiv preprint arXiv:1711.10791. 2017 Nov 29.

[6] Wagner P, Strodthoff N, Bousseljot RD, Kreiseler D, Lunze FI, Samek W, Schaeffter T. PTB-XL, a large publicly available electrocardiography dataset. Scientific data. 2020 May 25;7(1):1-5.