Discriminative Ensemble Averaging for Accelerated Denoising

Corey McCall CMCCALL@STANFORD.EDU
Stanford University, 450 Serra Mall. Stanford, CA 94305 USA

Abstract

It is difficult to filter a signal with high-amplitude in-band noise if an independent noise reference cannot be obtained. If the signal is periodic relative to a known sequence of temporal markers, *Ensemble Averaging* can be used to reduce the noise by averaging the signal over several periods in the same way that artificial neural networks average many modest experts to achieve a higher overall accuracy. Although this method effectively reduces in-band noise for this type of signal, it costs several time periods of waiting and is subject to corruption by high-amplitude spikes that take take several periods to reduce. In this paper, we use k-means clustering to discriminate against corrupted periods in order to achieve faster convergence of the ensemble average in a ballistocardiography dataset recorded on NASA's Zero-G aircraft. By using a single clean training example from each subject to select the least noisy cluster, ensemble averaging is accelerated for all ten subjects in the dataset within the first 4 cycles.

1. Introduction

Ensemble Averaging (EA) is a technique used in artificial neural networks to reduce many moderately accurate classification models to a single averaged model that together outperforms any of its individual components. The same technique can be used to denoise signals with high-amplitude in-band noise that would be otherwise be difficult to filter using typical frequency-based techniques. Unlike the typical approach to this problem, adaptive filtering (Inan et al., 2009), EA does not require an independently referenced noise signal so long as the signal is periodic relative to a known sequence of temporal markers. This type of signal is commonly encountered in fields like electrophysiology where measurements are very noisy, but the underlying signal of interest is periodic with respect to heart beats or action potential firings.

One exemplary case exists in ballistocardiography (BCG), the measurement of the small ballistic force induced by the heart pumping a volume of blood across the chest. It can be measured with either an accelerometer mounted to a person's center of mass, or with a sensitive scale measuring the differential weight between the legs of a standing person. A clean BCG waveform contains many hemodynamic ("blood movement") parameters useful for assessing a person's cardiac health in terms of cardiac output, contractility, synchronization, and other metrics as shown in Figure 1 (Inan et al., 2008; Etemadi et al., 2009). The frequency range of the BCG waveform is 1-30Hz, and highly susceptible to in-band noise caused by body motions as small as chewing or balancing while standing. BCG can be obtained carefully in a controlled environment, but cannot be used in congested settings without a robust denoising strategy.

In this paper, a discriminative EA approach is described for measuring small signals with high-amplitude in-band noise. The motivating example is measuring the BCG of astronauts onboard a crowded space capsule as in (Wiard et al., 2013). The signal can be ensembled by synchronizing with the simultaneously obtained electrocardiogram (ECG), in which the ECG R peaks proceed each BCG waveform complex (Figure 1). Although each BCG beat can be effectively referenced to an ECG timing marker, the congested setting induces frequent high-amplitude motion noise spikes from collisions with weightless objects and other astronauts, each of which increases the convergence time of the average. In order to solve this issue, the convergence is accelerated by discriminating against corrupted beats. This is done by clustering each sample of the ensemble at each step in order to filter out abnormal samples that negatively affect the average and increase the convergence time. Unlike naive EA,

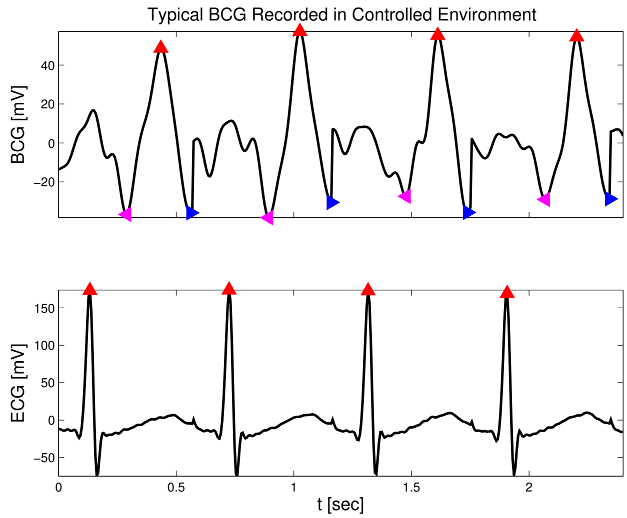

Figure 1. In a controlled environment, BCG can be easily obtained by standing still on a weight scale as shown (top). Various useful physiological parameters can be derived from the timing differences between the simultaneously recorded ECG (bottom) R peaks (red triangles), and the IJK complex of the trailing BCG waveform marked by the magenta (I), red (J), and blue (K) triangles. Signal segments are truncated to emphasize the window of interest.

which blindly averages each beat together, the algorithm described in this paper constructs the average from a subset of of the recorded beats determined at each sample by clustering. The output is a composite signal consisting of any number of parts from each of the recorded beats.

A previously used simpler approach in (McCall, 2013) clustered entire beats in n dimensional feature space instead of each sample point. This approach was less effective because high frequency noise would cause an entire beat to be filtered out even if the uncorrupted parts were useful.

This paper proceeds as follows: In Section 2, the BCG dataset is described. In Section 3, the discriminative EA approach is described. In Section 4, the results of the discriminative EA approach on the BCG dataset are described. And finally, we conclude in Section 5 with a discussion of the results.

2. BCG Dataset

A dataset of $N = 10$ subjects was recorded aboard NASA's Zero-G aircraft, with an average of 155 heart beats recorded from each subject. A typically noisy recording from the dataset is shown in Figure 2, which has no distinguishable BCG waveforms when compared to the controlled environment recording shown in Figure 1. The BCG signal was pre-processed using a 1-30Hz digital bandpass filter and subtracted from a 0.5-second windowed third-order polynomial fit (Savitzky & Golay, 1964) in order to reduce baseline wander before averaging. The m training example will be referred to as $\{x^{(i)} \in \mathbb{R}^n; i = 1, \ldots, m\}$, where each example is $n = 102$ samples long (0.4 seconds at 256Hz). The full matrix of examples $X \in \mathbb{R}^{m \times n}$ is defined as:

$$\mathbf{X} = \begin{bmatrix} \text{——} & (x^{(1)})^T & \text{——} \\ \text{——} & (x^{(2)})^T & \text{——} \\ & \vdots & \\ \text{——} & (x^{(m)})^T & \text{——} \end{bmatrix}$$

And the associated subject labels $\{y^{(i)} \in \{1, \ldots, N\}; i = 1, \ldots, m\}$ are defined in $Y \in \mathbb{R}^m$ as:

$$\mathbf{Y} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(m)} \end{bmatrix}$$

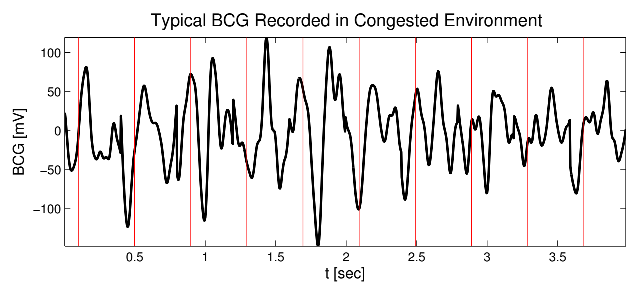

Figure 2. In a congested environment, BCG is difficult to distinguish because of the in-band noise that is much larger than the BCG component. The vertical red bars indicate ECG R peaks used to synchronize the ensemble.

3. Discriminative Ensemble Averaging

In the following sections, EA and the strategy to reduce its convergence time with clustering is discussed.

3.1. Ensemble Averaging

EA is a technique used to denoise periodic signals in which the in-band noise power is large compared to the signal of interest. Under the assumption that this noise is independently distributed with respect to the signal of interest, an estimate of a period of the underlying signal $\hat{x}^*$ can be calculated by averaging the signal over many periods. The following defines the naive EA estimate of subject a in the dataset:

$$\hat{x}^* \triangleq \frac{1}{\sum_{i=1}^{m} \mathbb{1}\{y^{(i)} = a\}} \mathbf{X}^\top \mathbb{1}\{\mathbf{Y} = a\}$$

where $\mathbb{1}$ is the componentwise indicator function: $\mathbb{1}\{z = a\} = 1$ if $z = a$, else 0.

An example of this procedure for half of the dataset is shown in Figure 3. It can be seen that in a realistic dataset, it can take many examples for $\hat{x}^*$ to converge. In many practical cases, such as busy astronauts checking into the BCG device for periodic checkups, it is desirable to minimize the measurement time required to estimate the underlying waveform x^*.

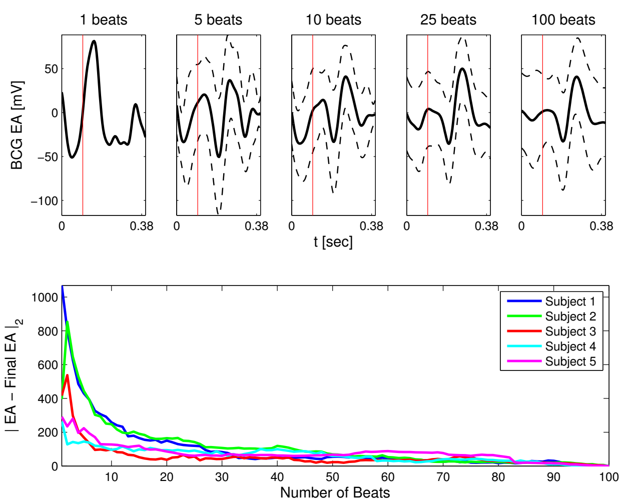

Figure 3. EA convergence for 5 subjects calculated naively by averaging each beat in sequence. The top row shows an example for Subject 1 as the average converges. The vertical red bars indicate the ECG R peaks used to synchronize the ensemble, and the thin dashed lines indicate the standard deviation of the average at each point.

3.2. Discriminative Ensemble Averaging

In discriminative EA, a single training example is added to the set of measured examples, then all $m+1$ examples are clustered at each point $j = 1, ..., n$. After clustering, only the points in each cluster containing the training example point are used to calculate the EA as described in Algorithm 1 and Figure 4. In this way, the clustering algorithm can select examples that look like the training example for the average, discriminating against examples corrupted by large motion noise or other artifacts in the other clusters. The final output if the algorithm is the naive EA or discriminative EA which is most similar to the training example. The algorithm is repeated until a target signal to noise ratio (SNR) is achieved. In the case of the BCG dataset, the algorithm would be repeated until the IJK complex is able to be distinguished, and the required training example would be collected in a controlled environment before the test measurements.

Algorithm 1 Discriminative Ensemble Averaging

Input: $X \in \{x^{(i)} \in \mathbb{R}^n; i = 1, \ldots, m\}$, $x_{train} \in \mathbb{R}^n$
Output: $\mu_{DEA} \in \mathbb{R}^n$
$\mu_{EA} = \text{EA}(X)$
$X := X \cup x_{train}$
for $j = 1$ **to** n **do**
  $z = X_j$
  $c =$CLUSTER(z)
  $z' = \{z_i; c_i = c_{m+1}\}$
  $z' := z' \setminus x_{train,j}$
  $\mu_{DEA,j} =$min(EA(z'),$\mu_{EA,j}$)
end for

The overall goal of the algorithm is to minimize the cost function defined as the 2-norm of the distance between the average and the training example:

$$\min_{\mathbf{X}} |\text{EA}(\mathbf{X}) - x_{train}|_2$$

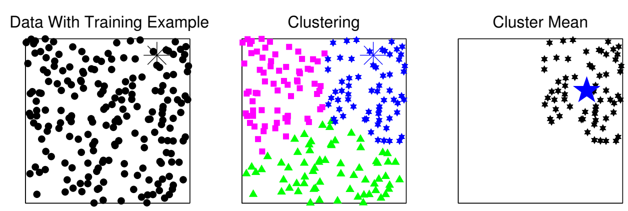

Figure 4. Graphical 2D example of Discriminative EA for $k = 3$ clusters. In the left figure, a single training example is added to the dataset. In the center figure, all of the points are clustered into k clusters. In the right figure, the training example is removed and the mean of its cluster is taken as the discriminative EA.

4. Results

Algorithm 1 was tested on the dataset using k-means for the CLUSTER function with $k = 2$ in an attempt to discriminate noisy vs. non-noisy examples. A more complex clustering method such as a gaussian mixture model was not used because in this case our goal is to explicitly include outliers in the center calculation rather than to form tight clusters.

The results for two typical subjects are shown in Figures 5 and 6, where it can be seen that the discriminative EA approach converges faster than the naive

EA approach in both figures as expected. In Figure 5, discriminative EA performs better with a low number of examples because the second example is clearly corrupt as seen in the naive EA trace. In Figure 6, the first examples are not noisy, so the discriminative EA does not pay off until noisy examples start to show up and increase the naive EA cost around example 5.

A graphical depiction of performance on the entire dataset for $m = 60$ is given in Figure 7. This matrix of ratios of naive to discriminative cost functions highlights when the algorithm is most effective. Aligning with our goal, performance is increased within the first 4 cycles for every subject in the dataset. Performance gains for higher m values, such as in Subjects 3 and 4, are because the algorithm naturally skips examples which cause the cost function with respect to the training example to increase.

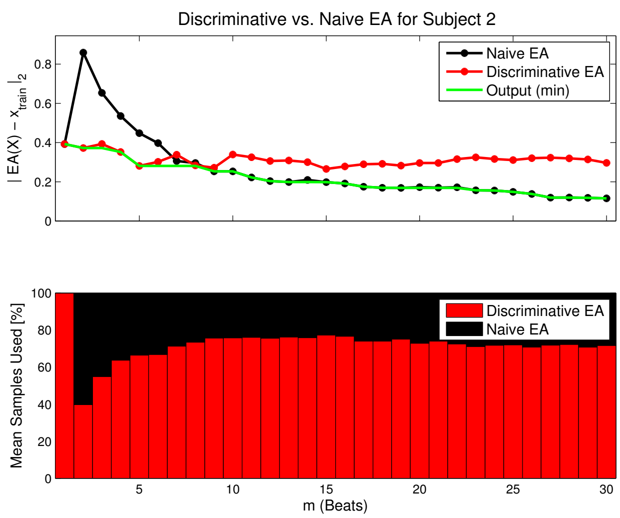

Figure 5. Discriminative EA increases the convergence rate compared to naive EA for typical Subject 2 (top). Mean percentage of data used at each point in the corresponding discriminative EA (Bottom).

5. Discussion

In this paper, a novel discriminative EA algorithm is described and tested on an appropriate dataset of ballistocardiogram data recorded in a very noisy environment. The algorithm, currently based on k-means clustering effectively increases the EA convergence rate in all subjects within 4 cycles.

Future work should include comparing this method to other non-frequency-based filtering techniques such as PCA dimensionality reduction before seeking peer-reviewed publication. It would also be valuable to test the algorithm on a dataset of simulated irregular noise for quantitative SNR analysis.

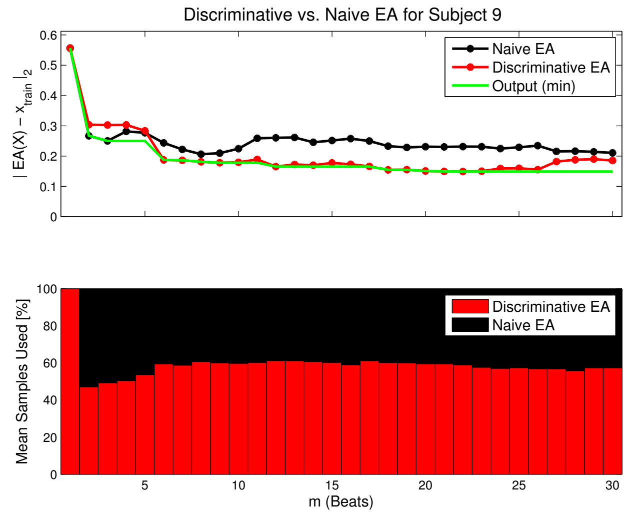

Figure 6. Discriminative EA increases the convergence rate compared to naive EA for typical Subject 9.

Acknowledgments

We acknowledge Professor Gregory Kovacs and Dr. Laurent Giovangrandi at Stanford University's Transducers Lab for discussing BCG signal processing techniques and sharing this dataset. We also acknowledge NASA's Reduced Gravity Office for facilitating the collection of this dataset. And finally, we acknowledge Professor Andrew Ng and the Stanford *CS229* lecture notes (Ng, 2013) for general instruction on machine learning.

References

Etemadi, M., Inan, O. T., R. M. Wiard, L. Giovangrandi, and Kovacs, G. T. A. Non-invasive assessment of cardiac contractility on a weighing scale. In *Proceedings of the 31st Annual IEEE Engineers in Medicine and Biology Conference*, Minneapolis, Minnesota, 2009.

Inan, O., Etemadi, M., Wiard, R., Kovacs, G. T. A., and Giovangrandi, L. Non-invasive measurement of valsalva-induced hemodynamic changes on a bathroom scale ballistocardiograph. In *Proceedings of the 30th Annual IEEE Engineers in Medicine and Biology Conference*, 2008.

Inan, O. T., Etemadi, M., Widrow, B., and Kovacs, G. T. A. Adaptive cancellation of floor vibrations in standing ballistocardiogram measurements using

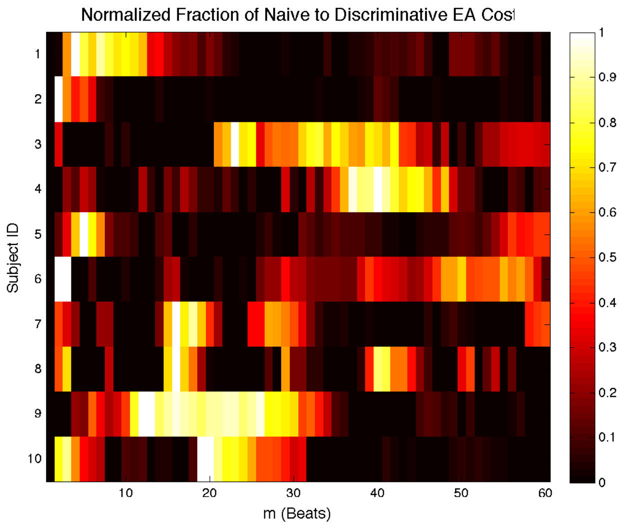

Figure 7. Algorithm performance across the entire dataset for $m = 60$ randomly sequenced beats. Each row is normalized to the maximum ratio of naive to discriminative EA cost functions (lighter = better). It can be seen that most subjects exhibit a performance increase in the first few cycles.

a seismic sensor as a noise reference. *IEEE Trans Biomed Eng*, 2009.

McCall, C. Discriminative ensemble averaging for accelerated denoising (milestone). *CS229 Final Project, Fall 2013*, 2013.

Ng, A. *CS229: Machine Learning Course Reader*. Stanford University, 2013.

Savitzky, A and Golay, MJE. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 1964.

Wiard, R. M., Inan, O. T., Giovangrandi, L., Cuttino, C. M., and Kovacs, G. T. A. Preliminiary results from standing ballistocardiography measurements in microgravity. In *Proceedings of the 35th Annual IEEE Engineers in Medicine and Biology Conference*, 2013.