Project Report
Expert Advice: Signal Denoising

Ce Fang, Fangfang Sun, Junlin Liu, Lin Qi, Tianpei Cai
cf1883, fs1337, jl7659, lq394, tc1996

April 28, 2017

Contents

1 Introduction

To denoise digital signal is to filter out the noise, where the challenge is how to preserve the important feature during the denoising process. In the real world, clean signal and noise may be time-variant. However, our filter is fixed. Therefor, to improve the performance of denoising for time-variant signal, we use Expert Advice Algorithm[1] [2] to cope with non-stationary sequence denoising problem.

To simplify the problem, we use two kinds of clean signal, blocky signal and non-blocky signal, and two kinds of noise which are Gaussian noise with different covariances. Each expert is specifically designed to deal with a certain type of signal, which includes both clean signal and the noise.

For Expert Advice algorithm, each expert can be considered as a black box. The result of each expert has been placed in the Appendix.

2 Expert Advice

2.1 The Model of Fixed-α algorithm

As our input signal is non-stationary sequence, we choose fixed-alpha algorithm to denoise the signal. The model for our system is shown in the following figure.

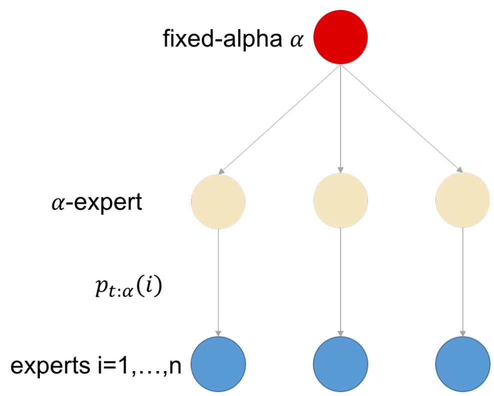


Figure 1: model of Fixed-α

Our system calculates the current loss for each expert and updates their corresponding weights $p_{t;\alpha(i)}$.

We compare the cumulative loss for every experts and prediction. To evaluate the performance of our system, we see whether or not the cumulative loss of prediction is less than that of each expert. we change the value of the alpha for each expert and see how the cumulative loss change with it on some training data, and then select the value with the least cumulative loss.

2.2 Fixed-α Algorithm

Inputs: Signals $(a_1, a_2, a_3, ..., a_n)$;
Initialization: $p_1(i) = \frac{1}{k}$, for all $i \in \{1, 2, ..., k\}$,k is the number of experts;
for each time $t = 1, 2, ...$
    received data point $y_i(t)$ in the stream
    for each expert $i \in \{1, 2, ..., k\}$
        compute the loss for each expert: $L_{t-1}(i)$
        update $p_{t;\alpha(i)}$

Notice:

$$p_t(i;\alpha) = \frac{1}{Z_t}\sum_{j=1}^{n} p_{t-1}(j;\alpha)e^{-\eta L_{t-1}(j)}p(i|j;\alpha)$$

$$\Theta_{i,j} = p(i|j;\alpha) = \begin{cases} 1-\alpha & i=j \\ \frac{\alpha}{k-1} & i \neq j \end{cases}$$

$$Z_t = \sum_{j=1}^{k} p_{t;\alpha}(i)$$

$L_t(i) = L_t(y_i(t), x(t)) = \frac{1}{N}\sum_{j=1}^{k}(y_i(t) - x(t))^2$ is the loss of expert i at time t.

2.3 Expert of Denoising

2.3.1 TVD1

Total variation denoising (TVD)[3] is an approach for noise reduction developed to preserve sharp edges in the underlying signal.

Total variation denoising assumes that the noisy data $y(n)$ is of the form:

$$y(n) = x(n) + w(n), n = 0, 1, ..., N-1$$

where $x(n)$ is a (approximately) piecewise constant signal and $w(n)$ is a white Gaussian noise.

The TV denoising problem can be written compactly as

$$\arg\min_{x\in\mathbb{R}^N}\{F(x) = \frac{1}{2}\|y-x\|_2^2 + \lambda\|Dx\|_1\}$$

which can be written as,

$$\arg\min_{x\in\mathbb{R}^N}\{F(x) = \frac{1}{2}\sum_{n=0}^{N-1}|y(n)-x(n)|^2 + \lambda\sum_{n=0}^{N-1}|x(n)-x(n-1)|\}$$

The matrix D is defined as

$$D = \begin{bmatrix} -1 & 1 & & & & \\ & -1 & 1 & & & \\ & & & \ddots & \ddots & \\ & & & & -1 & 1 \end{bmatrix}$$

An upper bound (majorizer) of $f(t) = |t|$ that agrees with $f(t)$ at $t = t_k$ is:

$$g(t) = \frac{1}{2|t_k|}t^2 + \frac{1}{2}|t_k|$$

Using MM algorithm, we can obtain:

$$x_{k+1} = y - D^\top(\frac{1}{\lambda}diag(Dx_k)DD^\top)^{-1}Dy$$

The regularization parameter $\lambda > 0$ controls the degree of smoothing.

Total variation (TV) denoising is a method to smooth signals based on a sparse-derivative signal model. TV denoising is formulated as the minimization of a non-differentiable cost function. Unlike a conventional low-pass filter, the output of the TV denoising filter can only be obtained through a numerical algorithm. Total variation denoising is most appropriate for piecewise constant signals. However, in the feature, it can be modified and extended to be effective for more general signals. Here, we generate a second order TVD filter for non-blocky signal.

2.3.2 TVD2

A modified TVD1 filter to a second-ordered TVD filter, which has better performance on non-blocky signal.

2.3.3 Wavelet Denoised Filter

For wavelet denoised filter [4]:

1. Firstly, the received signal levels are separated by wavelet transform. Then, the received signal wavelet coefficients are calculated up to the desired level.
2. the variance (σ^2) of the noise is calculated using the wavelet coefficients:

$$\hat{\sigma} = \frac{med(|W_{j,k}|)}{0.6745}$$

where med(.) denotes the median.

3. The threshold value is calculated using the variance.

$$T = \sigma\sqrt{2\log(n)}$$

4. Performing the hard-thresholding.
5. The original signal is reconstructed using the inverse wavelet transform and retained coefficients.

There are two experts in our system for wavelet denoised filter: wavelet denoised filter-harr and wavelet denoised filter-symlets. Harr wavelet is used for the blocky, while symlets wavelet is used for the non-blocky signal.

We set the level = 3.

The performance of these filters are attached in the appendix.

3 Results and Discussion

3.1 Finding the best alpha

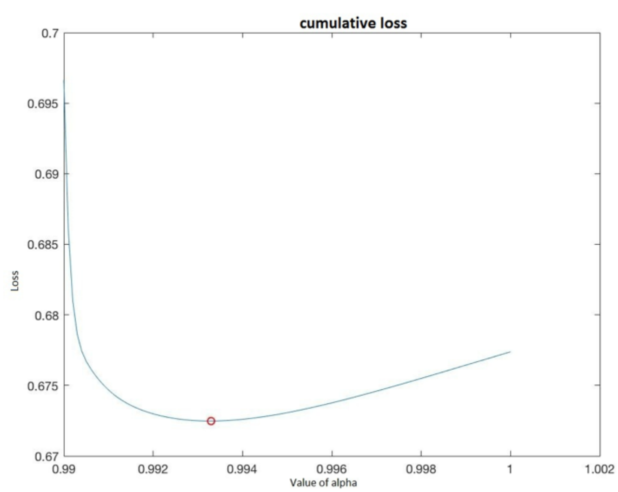


Figure 2: curve of the cumulative loss function in our system

Figure 2 indicates how the cumulative loss function changes with the value of α. The cumulative loss function is convex, in which the global minimum is obtained as 0.6731 when alpha equals to 0.9932.

3.2 Weight Change

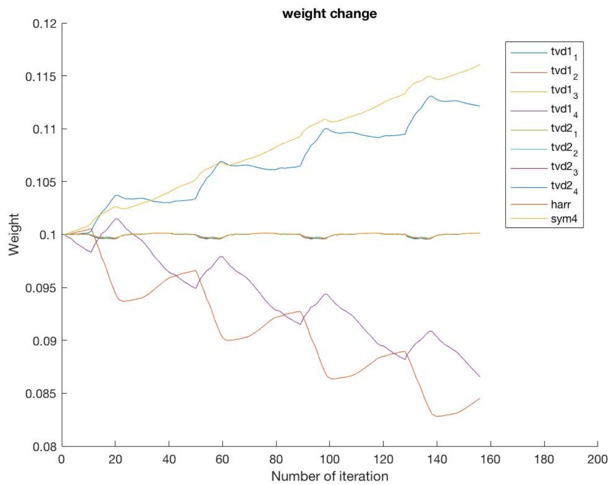


Figure 3: the evolution of weights at different experts over time

Figure 3 gives how the weight of each expert changes at different iterations. From this figure, we can clearly see that expert tvd1_1 and expert tvd2_4 always perform better than other experts, and expert tvd1_2 and tvd1_4 always perform worse than others, since our noise and clean signal is not abundant enough, those experts just don't meet the right noise. However, at different steps, the identity of best expert is changed, which means our algorithm has the ability to cope with the non-stationary input sequence by adjusting the weight of each expert at different times.

3.3 Cumulative Loss

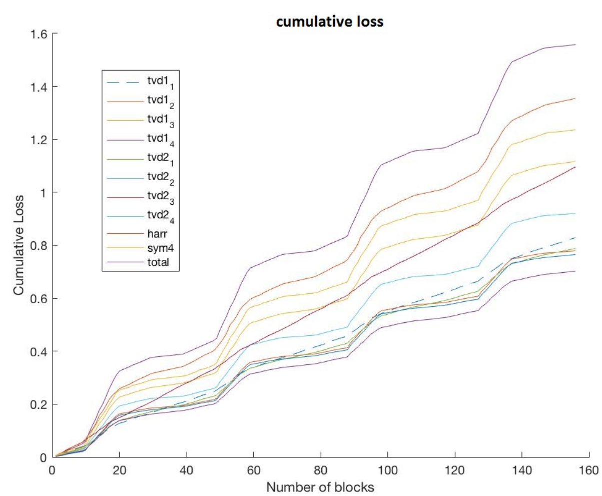


Figure 4: the evolution of cumulative loss over different experts over steps

Figure 4 shows the cumulative loss of each experts at each block. Our system has the minimal total loss almost at every step, which means the algorithm used in our system works well in this situation. It turns out that our algorithm can be potentially useful in signal denoising problem in our real world.

4 Summary and Future Work

4.1 Summary

From figures above we can see that, our expert advice algorithm can generate denoised signal with minimal loss, as it wins with each of this method run by itself.

4.2 Future Work

In our project, we just deal with one-dimension signal with Gaussian white noise. In the future, we would like to denoise two-dimensional signal, such as an image, where the TVD filter is widely used.

At present our fixed-α algorithm with switching rate works well on non-stationary sequence. In the future, we would like to work on learn-α algorithm, which can find the best expert to denoise signals even without signal type and distribution. Hopefully, it will save us amounts of time to find the best corresponding coefficient of the filter to denoise the input signal, since if we don't have the estimation of input distribution we may need a lot of experts to generate a satisfied result, which will lead to the increase of calculation.

References

[1] C. E. Monteleoni, "Online learning of non-stationary sequences," *NIPS*, pp. 1093–1100, May 2003.

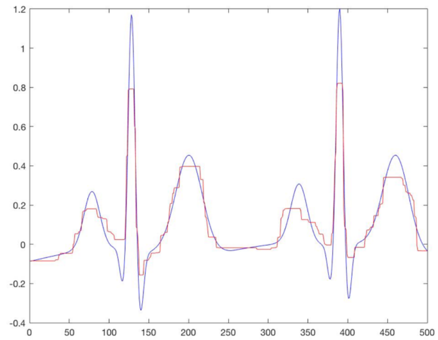
[2] M. HERBSTER and M. K. WARMUTH, "Tracking the best expert," *Machine Learning*, vol. 32, no. 151-178, 1998.

[3] I. Selesnick, "Total variation denoising (an mm algorithm)," 2012, nYU Polytechnic School of Engineering lecture note.

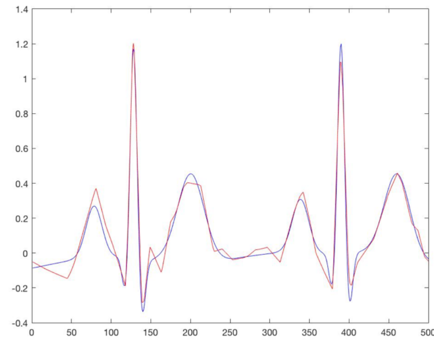
[4] M. ÜSTÜNDAĞ, A. ŞENGÜR, M. GÖKBULUT, and F. ATA, "Performance comparison of wavelet thresholding techniques on weak ecg signal denoising," *PRZEGLAD ELEKTROTECHNICZNY*, May 2013.

A Appendix

A.1 TVD filter for non-blocky signal



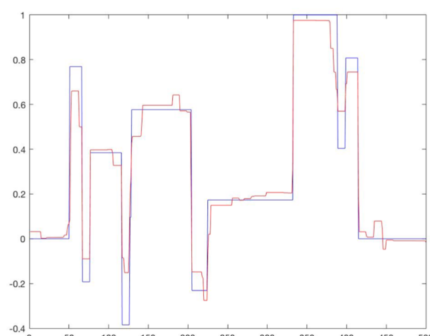
(a) tvd1 filter



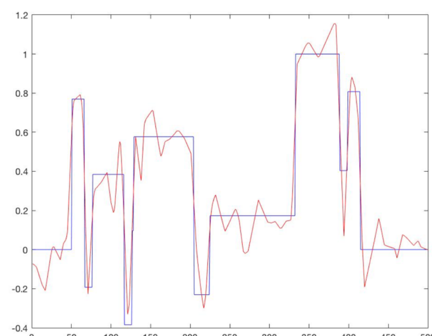
(b) tvd2 filter

Figure 5: tvd filter for non-blocky signal

A.2 TVD filter for blocky signal



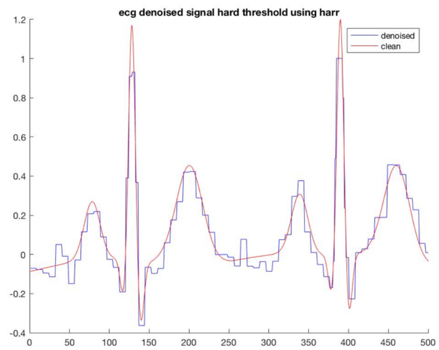
(a) tvd1 filter



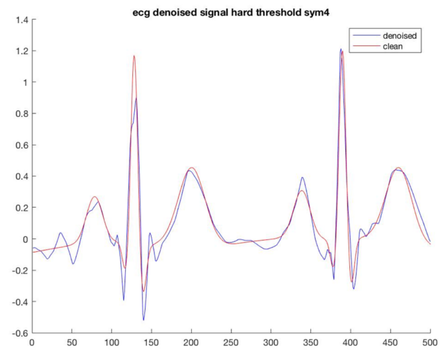
(b) tvd2 filter

Figure 6: tvd filter for blocky signal

A.3 Wavelet filter for non-blocky signal



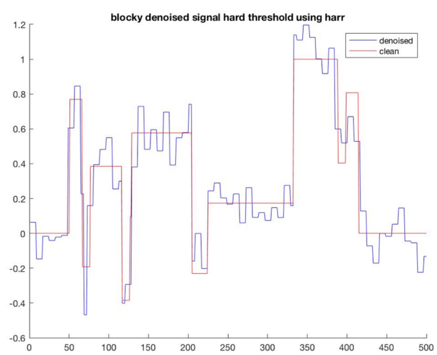
(a) wavelet filter-harr



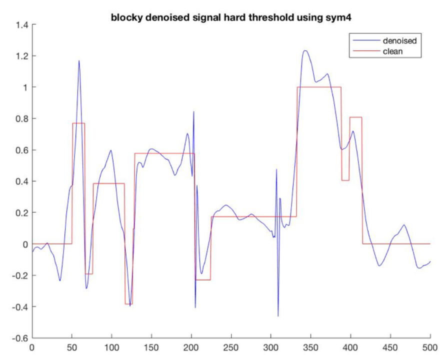
(b) wavelet filter-sym4

Figure 7: wavelet filter for non-blocky signal

A.4 Wavelet filter for blocky signal



(a) wavelet filter-harr



(b) wavelet filter-sym4

Figure 8: wavelet filter for blocky signal