Proceedings of the 12th
KACC, October 1997

Extraction of Voice Signal Embedded in 1/f noise using Wavelet

°Naoki Toyama Takashi Sasaya Jaka Sembiring Kageo Akizuki

Department of Electrical, Electronics and Computer Engineering,
Waseda University, 3-4-1, Ohkubo, Shinjuku-ku, Tokyo 169, Japan
Tel: +81-3-3207-7143;Fax: +81-3-3207-7143
E-mail: toyama@akizuki.elec.waseda.ac.jp

Abstract: This paper deals with the problem of extraction of voice signal embedded in 1/f noise. We propose the extraction method using wavelet. This method is based on Wornell's modelling which can construct 1/f process in terms of uncorrelated variables and is well suited on treating 1/f process. Finally, we show further describe our method through simulation.

Keywords: Wavelet, Multiresolution analysis, 1/f process, ML estimation

1. Introduction

A curent topic of great interest is the multiresolution analysis (MRA) of signals and the development of multiscale signal processing algorithms. In such trend, the modeling for $1/f$ processes using wavelet was introduced by Wornell [1][2]. As yet no entirely satisfactory framework had been described for the analysis of 1/f processes. Orthonormal wavelet bases are used to provide a new construction for nearly $1/f$ processes from a set of uncorrelated random variables. In this paper, we propose the method using wavelet on the extraction of voice signal embedded in $1/f$ noise. Our method consists of prameter estimation of $1/f$ noise followed by thresholding. In the parameter estimation step we use Wornell's model ,and then the estimated parameters are used in the next step. Finally, we illustrate our result with simulation which shows the effectieness of our method.

2. Orthonormal Wavelet Bases

In this section,we review some results from orthonormal wavelet theory required subsequently. For more complete discussions of orthonormal wavelet, see [4][5].

An orthonormal wavelet transform of a signal $f(t)$

$$f(t) \leftrightarrow d_n^m$$

is defined through the synthesis/analysis equation (inverse wavelet transform / wavelet transform)

$$f(t) = \sum_m \sum_n d_n^m \, \psi_n^m(t) \qquad (2.1a)$$

$$d_n^m = \int_{-\infty}^{\infty} f(t)\, \psi_n^m(t)\, dt \qquad (2.1b)$$

The basis functions $\psi_n^m(t)$ have the property that they are dilations and translations of a single function $\psi(t)$:

$$\psi_n^m(t) = 2^{m/2}\, \psi(2^m t - n) \qquad (2.2)$$

where m and n are the dilation and translation indices respectively.

It is also possible to view the wavelet transformation in the context of a multiresolution analysis (MRA).

At first, approximation of a signal $f(t)$ at level m (resolution-2^m) is defined as

$$f^m(t) = \sum_n a_n^m \, \phi_n^m(t) \qquad (2.3)$$

with the coefficients a_n^m obtained by projection

$$a_n^m = \int_{-\infty}^{\infty} f(t)\, \phi_n^m(t)\, dt \qquad (2.4)$$

The basis function $\phi_n^m(t)$ also have the property that they are dilations and translations of a single function $\phi(t)$ called "scaling function":

$$\phi_n^m(t) = 2^{m/2}\phi(2^m t - n) \qquad (2.5)$$

Having expressed $f^m(t)$ as (2.3), by using $\psi(t)$ which has a certain relationship with $\phi(t)$, $f^m(t)$ can also be expressed as follows

$$f^m(t) = f^{m-1}(t) + \sum_n d_n^{m-1}\psi_n^{m-1}(t) \qquad (2.6)$$

We can view the second term of the right hand side of (2.6) as the additional information or detail in going from a resolution-2^{m-1} approximation $f^{m-1}(t)$ to a resolution-2^m approximation $f^m(t)$.
Accumulating this information over all scales m leads to the synthesis formula (2.1a)

3. 1/f process

3.1 Definition of 1/f process

In a large number of physical phenomina, $1/f$ -type spectral behaviors are observed over wide ranges of frequencies. The $1/f$ processes are generally defined as processes whose empirical power spectral are of the form [3]

$$S(f) \propto \frac{1}{f^\gamma} \qquad (3.1)$$

$S(f)$: power spectral

f : frequency

γ : parameter ($0 < \gamma < 2$)

3.2 Wavelet Representation of 1/f Processes

Although $1/f$ -type process is of great importance, we had no adequate model of such processes. However, recently Wornell [1][2] provided a well-suited model of such processes using wavelet expansions in terms of uncorrelated coefficients. He constructed $1/f$ process as follows,

$$r(t) = \sum_m \sum_n r_n^m \psi_n^m(t) \qquad (3.2)$$

$r(t)$: $1/f$ process

r_n^m : white process having the variance $\sigma^2 2^{-\gamma m}$

$\psi_n^m(t)$: orthonormal wavelet base

4. Extraction of Voice Signal Embedded in 1/f Noise

This section will consider the extraction of voice signal embedded in $1/f$ noise.

Suppose we have an observation $y(t)$ of a voice signal embedded in $1/f$ noise $r(t)$.

$$y(t) = x(t) + r(t) \qquad (4.1)$$

where,

$y(t)$: observation

$x(t)$: voise signal

$r(t)$: $1/f$ noise

Our method for extracting voice signal $x(t)$ consists of 4 steps, as in **Fig.1**.

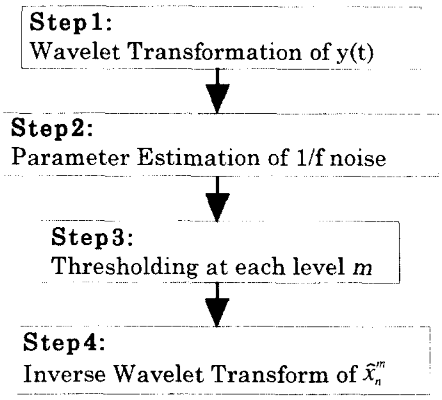

Fig.1 Extraction of Voice Signal

Step1.Wavelet Transformation of y(t)

At first, $y(t)$ is transformed using wavelet expansion.

$$y(t) = x(t) + r(t)$$

$$\downarrow \text{ wavelet transform}$$

$$y_n^m = x_n^m + r_n^m$$

Step2.Parameter Estimation of 1/f noise

In this step, we estimate $1/f$ noise parameters. The parameter set we wish to estimate is

$$\Theta = (\beta, \sigma^2) \qquad (\beta = 2^\gamma)$$

We may express the likelihood as a function of the parameters by

$$L(\Theta) = p(r|\Theta) = \prod_{m,n} \frac{1}{\sqrt{2\pi\sigma^2\beta^{-m}}} \exp\left[-\frac{(r_n^m)^2}{2\pi\sigma^2\beta^{-m}}\right]$$

And log-likelihood function as

$$L(\Theta) = \ln p(r|\Theta) = -\frac{1}{2}\sum_m N(m)\left(\frac{\hat{\sigma}_m^2}{\sigma^2\beta^{-m}} + \ln(2\pi\sigma^2\beta^{-m})\right) \quad (4.2)$$

where,

$$\hat{\sigma}_m^2 = \frac{1}{N(m)}\sum_n (r_n^m)^2$$

$N(m)$: data length of r_n^m

Differenciating $L(\Theta)$ with respect to $\beta \cdot \sigma^2$, we have $\frac{\partial L}{\partial \beta}, \frac{\partial L}{\partial \sigma^2}$. With $\frac{\partial L}{\partial \beta} = 0, \frac{\partial L}{\partial \sigma^2} = 0$ and some algebraic computation, we obtain

$$\frac{\partial L}{\partial \beta} = 0 \rightarrow \sum_m m\beta^{-m}\frac{N(m)}{\sigma^2\beta^{-m}}\left(1 - \frac{\hat{\sigma}_m^2}{\sigma^2\beta^{-m}}\right) = 0$$

$$\frac{\partial L}{\partial \sigma^2} = 0 \rightarrow \sum_m \beta^{-m}\frac{N(m)}{\sigma^2\beta^{-m}}\left(1 - \frac{\hat{\sigma}_m^2}{\sigma^2\beta^{-m}}\right) = 0$$

Eliminating σ^2 from above two equations, we have the following polynomial equation.

$$\sum_m \left(\frac{m}{\sum_m mN(m)} - \frac{1}{\sum_m N(m)}\right) N(m)\hat{\sigma}_m^2\beta^m = 0 \quad (4.3)$$

Because this equation has only one positive real solution, we make the solution β_{ML} which is ML estimate β.

Having obtained β_{ML}, $\hat{\sigma}_{ML}^2$ is calculated through

$$\hat{\sigma}_{ML}^2 = \frac{\sum_m N(m)\hat{\sigma}_m^2[\beta_{ML}]^m}{\sum_m N(m)} \quad (4.4)$$

Step3. Thresholding step

Having estimated $1/f$ noise parameters $\beta_{ML} \cdot \hat{\sigma}_{ML}^2$, we can express the mean and variance of $1/f$ noise at each level m as follows

$$E[r_n^m] = 0 \quad (4.5a)$$

$$V[r_n^m] = \beta_{ML}^{-m}\hat{\sigma}_{ML}^2 \quad (4.5b)$$

Now we consider the avarage power over short times k defined as

Average power

$$\overline{(r^m)}^2 = \frac{1}{k}\left[(r_n^m)^2 + (r_{n+1}^m)^2 + \cdots + (r_{n+k}^m)^2\right] \quad (4.6)$$

According to χ^2-distribution, the mean and variance of average power are calculated as follows.

Mean of the average power

$$E\left[\overline{(r^m)}^2\right] = \beta_{ML}^{-m}\hat{\sigma}_{ML}^2 \quad (4.7)$$

Variance of the average power

$$V\left[\overline{(r^m)}^2\right] = \beta_{ML}^{-2m}\hat{\sigma}_{ML}^4\frac{2}{k} \quad (4.8)$$

The average power in the part that includes voice signal is much bigger than the one that not including it. So, by determining the threshold as

Threshold value at level m

$$E\left[\overline{(r^m)}^2\right] + 4\sqrt{V\left[\overline{(r^m)}^2\right]} \quad (4.9)$$

and cutting out coefficients which are smaller than the threshold value, we can extract voice signal coefficients $\hat{x}_n^m$ at level m.

Step4. Inverse Wavelet Transform of $\hat{x}_n^m$

Finally, by inverse wavelet transfomation of $\hat{x}_n^m$, we can obtain the extracted voice signal $\hat{x}(t)$.

$$\hat{x}(t) = \sum_m\sum_n \hat{x}_n^m\psi_n^m(t) \quad (4.10)$$

5. Simulation

In this section, we show a simulation result obtained by using the method proposed in the previous section.

Data condition is as follows.

record data : 4.096 *sec*
sampling time : 1/8000 *sec*
number of sample data : 32768 *samples*

The original voice signal embedded in $1/f$ noise is depicted in **Fig.2**. The extracted voice signal after applying our method is shown in **Fig.3**.

Simulation result demonstrated the effectiveness of our method.

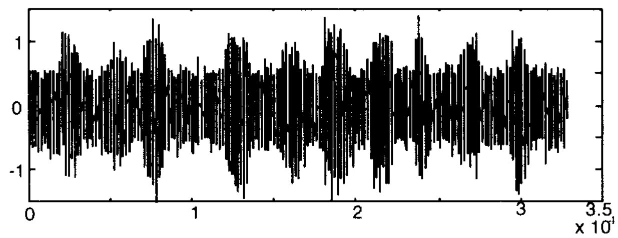

Fig.2 The original signal

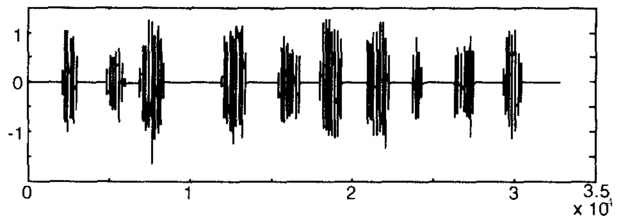

Fig.3 Extracted voice signal

6. Conclusion

In this paper, we have dealt with the problem of extraction of voice signal embedded in $1/f$ noise, and proposed the extaction method using wavelet.

We have some reasons to convince that the proposed method is useful. First, on treating $1/f$ process, the model using wavelet is well suited. Second, on treating nonstationary signal in stationary noise, the wavelet analysis is effective. Furthermore, in this application, the only known information is that the additive noise is $1/f$ -type. Nevertheless, we can extract the voice signal. This also proves the effectiveness of our method.

References

[1]Wornell.G.W, "Estimation of fractal signals from noisy measurements using wavelet", IEEE Trans. on Signal Processing, 40(3), pp.611-623, 1992

[2]Wornell.G.W, "Signal Processing with Fractals: A Wavelet-based Approach", Prentice Hall,1996

[3]M.S.Keshner, "1/f noise", Proc.IEEE, vol.70, pp.212-218, Mar.1982

[4]Daubechies.I, "Orhonormal bases of compactly supported wavelets", Commun.Pure Appl.Math., vol.41, pp.909-996, Nov.1988

[5]S.Sakakibara, "Wavelet beginner's guide", Tokyo Denki Pub., 1995

[6]H.Asano, K.Akizuki, "Extraction of Voice Signal Embedded in Stationary Noise by Wavelet Transform", National Convention Record I.E.E. Japan, 1994