Proc. of the 5th WSEAS/IASME Int. Conf. on Electric Power Systems, High Voltages, Electric Machines, Tenerife, Spain, December 16-18, 2005 (pp177-181)

Investigation of neighbor multi-wavelet denoising in partial discharge detection

ZENG YI
Department of Electrical Engineering
Shanghai Jiaotong University
800#, Dongchuan Road, Minhang district, Shanghai City
CHINA

QIAN YONG
Department of Electrical Engineering
Shanghai Jiaotong University
800#, Dongchuan Road, Minhang district, Shanghai City
CHINA

HUANG CHENG-JUN
Department of Electrical Engineering
Shanghai Jiaotong University
800#, Dongchuan Road, Minhang district, Shanghai City
CHINA

XU BING-YAN
Department of Electrical Engineering
Shanghai Jiaotong University
800#, Dongchuan Road, Minhang district, Shanghai City
CHINA

JIANG XIU-CHEN
Department of Electrical Engineering
Shanghai Jiaotong University
800#, Dongchuan Road, Minhang district, Shanghai City
CHINA

Abstract:- In this contribution, we extend the work poineered by Chen[4], investigating the bigger neighborhood, and further employ the neighbor multi-wavelet denoising scheme to detect partial discharge signals, which are usually overwhelmed by excessive on-site noise. By massive simulation, the relation curve among mean square error (MSE), neighborhood and exponent of neighboring coefficient is derived. Based on the relation curve, the optimal neighborhood and index of neighboring coefficients are obtained. What is more, we apply the neighbor multi-wavelet denoising, together with the optimal parameters, to process partial discharge. The obtained result is promising and the derived parameters are justified.

Key-words: - Multi-wavelet,Neighboring coefficient, Neighbourhood, Partial discharge, White noise, Denoising

1 Introduction
Detection of partial discharge is an important means of monitoring the insulation condition of large power apparatus; while in the detecting process, denoising always plays the significant role, because partial discharge is weak electric signal, usually overwhelmed by the noise existing on site. Without denoising, we can hardly analyze the partial discharge with any accurate result [1].

In the past decade, the most commonly used method for denoising is wavelet transform. Wavelet transform can provide multi-resolution analysis, and outperforms other methods in processing nonstationary signals, in which partial discharge is typical. However, with the development of wavelet theory, there appears multi-wavelet. According to the massive results reported, multi-wavelet is superior to wavelet in signal processing [2-4]. In particular, in literature [4], Chen claimed that neighbor multi-wavelet denoising could be used in place of wavelet denoising. And this evokes our great interest and furthermore leads to the investigation into the method.

Based on the work of literature [4], we extend neighbor multi-wavelet denoising, investigating the neighborhood as suggested by Chen. Besides, we study the exponent used in the method. The objective of our work is to search for the relatively optimal parameters by massive simulation. Further, we apply the method, together with derived optimal parameters, to detect partial discharge overwhelmed by excessive noise and evaluate its performance.

2 Basic theory of multi-wavelet
For a multi-resolution of multiplicity $r>1$, there are r scaling functions $\phi_1(t)$, $\phi_2(t)$, …, $\phi_r(t)$, and r scaling functions $\psi_1(t)$, $\psi_2(t)$, …, $\psi_r(t)$, usually written as vectors $\mathbf{\Phi}(t)=[\phi_1(t),...,\phi_r(t)]^T$ and $\mathbf{\Psi}(t)=[\psi_1(t),...,\psi_r(t)]^T$ respectively, which satisfy the following matrix equations[5-7]

$$\Phi(t)=\sum_{k\in\mathbf{Z}}\boldsymbol{h}_k\Phi(2t-k) \tag{1}$$

$$\psi(t)=\sum_{k\in\mathbf{Z}}\boldsymbol{g}_k\Phi(2t-k) \tag{2}$$

where, $\boldsymbol{h}_k$、 $\boldsymbol{g}_k$ are $r\times r$ matrix low-pass filter and high-pass filter, respectively.

Similarly, from the MRA of multi-wavelet, decomposition and reconstruction equation can be derived as follows [8]:

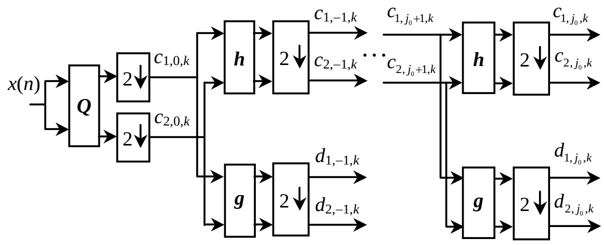
$$\boldsymbol{C}_{j-1,k}=\sqrt{2}\sum_m\boldsymbol{h}_m\boldsymbol{C}_{j,m+2k},\ \ j,k\in\mathbf{Z} \tag{3}$$

$$\boldsymbol{D}_{j-1,k}=\sqrt{2}\sum_m\boldsymbol{g}_m\boldsymbol{D}_{j,m+2k},\ \ j,k\in\mathbf{Z} \tag{4}$$

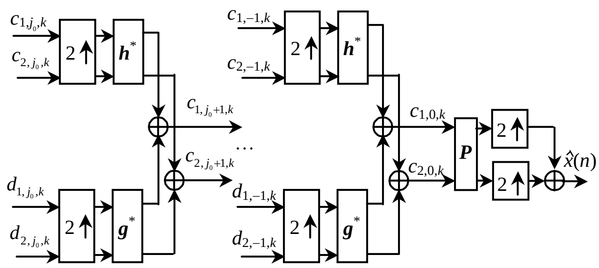
$$\boldsymbol{C}_{j,m}=\sqrt{2}\sum_k(\boldsymbol{h}_k^*\boldsymbol{C}_{j-1,m+2k}+\boldsymbol{g}_k^*\boldsymbol{D}_{j-1,m+2k}) \tag{5}$$

where $\boldsymbol{C}_{j,k}=[c_{1,j,k},c_{2,j,k},...,c_{r,j,k}]^{\mathrm{T}}$, $\boldsymbol{D}_{j,k}=[d_{1,j,k},d_{2,j,k},...,d_{r,j,k}]^{\mathrm{T}}$, whereas $*$ denotes the complex conjugate transpose. In our work, the multi-wavelet adopted is the most commonly used GHM（Geronimo-Hardin-Massopust）multi-wavelet developed by Geronimo et al[5].

The decomposition and reconstruction process of multi-wavelet transform are illustrated in Fig.1, in which $\boldsymbol{Q}$ and $\boldsymbol{P}$ denote prefilter and post-filter respectively, for more details of prefilter and post-filter, the reader is referred to literatures [8-10]. The prefilter and post-filter, together with the matrix filter coefficients, used in our study can be found in the appendix.



(a) decomposition



(b)reconstruction

Fig.1 Decomposition and reconstruction of multiwavelet transform

3 Neighbor multi-wavelet denoising
Neighbor multi-wavelet denoising scheme considers several adjacent coefficients as a block, and thresholds them as a whole. As the relation between coefficients has been taken into consideration, better performance can be obtained. Applying multi-wavelet transform with an appropriate prefilter, we get r stream coefficients in the form of $\boldsymbol{D}_{j,k}=\boldsymbol{D}^*_{j,k}+\boldsymbol{E}_{j,k}$, where $\boldsymbol{D}^*_{j,k}$ are the signal coefficients, and $\boldsymbol{E}_{j,k}$ have multivariate normal distribution $N(0,\boldsymbol{V}_j)$. The matrix $\boldsymbol{V}_j$ is the covariance

matrix for the error term that depends on the resolution level *j*. In the absence of any signal component, the quantity $\theta_{j,k} = \boldsymbol{D}_{j,k}^{\mathrm{T}} \boldsymbol{V}_j^{-1} \boldsymbol{D}_{j,k}$ will have a χ_n^2 distribution. The threshold rule used for processing coefficients is based on the values of $\theta_{j,k}$. And $\boldsymbol{V}_j$ can be obtained directly by using robust covariance estimation suggested in[3].

The idea of neighbor thresholding lies in the fact that: if the current coefficient contains some signal, then it is likely that the closely adjacent coefficients also do. For this reason, define a variable $S_{j,k}$, if only two immediate neighbor coefficients are considered, $S_{j,k} = \theta_{j,k-1}^r + \theta_{j,k}^r + \theta_{j,k+1}^r$, the thresholding algorithm of literature [4] can be represented as $\hat{\boldsymbol{D}}_{j,k} = \boldsymbol{D}_{j,k} \cdot f(S_{j,k}, \mu^r)$, where $f(\cdot)$ denotes threshold rule and μ^r is the corresponding threshold (*r* is an nonnegative integer).

We extend the neighbor thresholding by representing the coefficient block as

$$S_{j,k}(q,r) = \sum_{i=k-q}^{k+q} \theta_{j,i}^r \tag{6}$$

where *q* denotes the neighborhood, an nonnegative integer (*q*=1 corresponds to the case in [4]); *r* is a positive exponent; with regard to the marginal problem encountered in calculating $S_{j,k}$, we process it by periodic extension.

Similar to classical multi-wavelet thresholding, neighbor thresholding can be implemented in two ways, either in hard thresholding or in soft thresholding. Soft thresholding can be expressed as

$$\hat{\boldsymbol{D}}_{j,k} = \begin{cases} \boldsymbol{D}_{j,k}\left(1 - \dfrac{\mu_j^r}{S_{j,k}}\right), S_{j,k} \ge \mu_j^r \\ 0, \text{otherwise} \end{cases} \tag{7}$$

and hard thresholding corresponds to

$$\hat{\boldsymbol{D}}_{j,k} = \begin{cases} \boldsymbol{D}_{j,k}, S_{j,k} \ge \mu_j^r \\ 0, \text{otherwise} \end{cases} \tag{8}$$

where $\mu_j = 2\lg N_j$, N_j is the length of the multi-wavelet coefficients at resolution level j; μ_j^r is the threshold for resolution level j; $\hat{\boldsymbol{D}}_{j,k}$ is the estimate of coefficient corresponding to signal. In the successive section, only hard threshold is considered.

4 Analysis of simulation signals

Partial discharge detected on site usually turn out to be resonant damped pulses. In theory research, it can be simulated with the following analytical expression [11]

$$f(t) = V_{\mathrm{m}} \mathrm{e}^{-t/\tau} \sin(f_c \times 2\pi t) \tag{9}$$

where V_{m} denotes the peak value; τ time constant; f_c resonant frequency. In simulation, let V_{m}=1mV, τ=0.1μs, 0.5μs, 1μs, 2μs, 3μs or 4μs, f_c=1MHz. White noise superimposed has a $N(0, 0.25^2)$ distribution. With these parameters, we can obtain original partial discharge pulses and corrupted data, which are illustrated in Fig.2 and Fig.3 respectively.

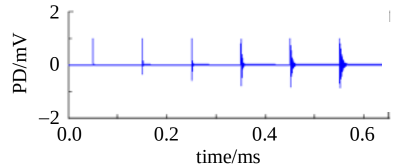


Fig.2 Simulation signal of partial discharge

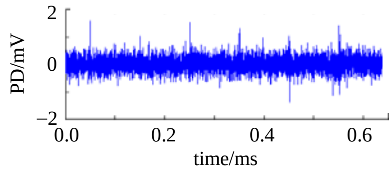


Fig.3 Simulation signal of partial discharge mixed with white noise

In signal processing, Root Mean Square Error (MSE) is an important criterion for evaluating the denoising performance [3], and in our work, it is adopted.

Based on (6) and partial discharge shown in Fig.2, by massive simulation, which was conducted 300 times independently, we get the relation between MSE and parameters *r*、*q*(see table 1 and Fig.4). Fig.4 shows that, with the increase of *r*, MSE decreases all the while, and at last approximates some constant (see table 1); with the increase of *q*, MSE increases consistently. As for term-by-term denoising, the mean of 300 results is MSE=0.0047mV. Compare 0.0047 with the results in table 1, it can be seen that, with appropriate parameters *r*、*q*, neighbor denoising can give better results.

Table 1 Relation among MSE、*q* and *r*

	MSE/μV						
	r =1	*r*=1.5	*r*=2	*r*=2.5	*r*=3	*r*=3.5	*r*=4
q=0	4.5	4.5	4.5	4.5	4.5	4.5	4.5
q=1	6.1	4.7	4.2	4.2	4.2	4.2	4.2
q=2	14.0	5.5	4.9	4.7	4.4	4.4	4.4
q=3	29.3	7.8	5.7	5.3	4.7	4.7	4.7
q=4	45.4	10.7	6.6	5.8	5.1	5.1	5.1

q=5	56.2	14.9	8.0	6.5	5.6	5.5	5.5

From Fig.4 and table 1, it can be concluded that, when q keeps constant (q=1 is preferred), if r>3, the change of MSE is trivial, but the computing complexity becomes intensive. Therefore, the selection of q 1, $r \in [2,3]$ may achieve good performance. This coincides with the results reported in literature [4], where q=1, r=2. In the successive calculation, only r=2.5 is considered.

With the optimal parameters r=2.5 and q=1, we give an instance of denoising the data shown in Fig.3 (see Fig.5). For convenience of comparison, results by using term-by-term denoising are provided together (see Fig.6).

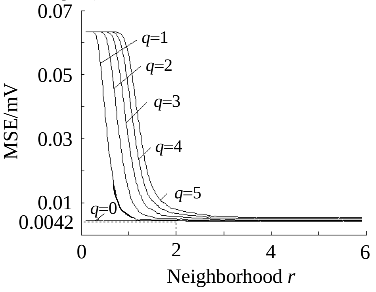


Fig.4 RMSE vs. r and q

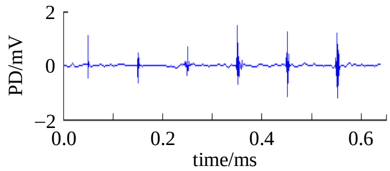


Fig.5 Denoised signals with neighbor multi-wavelet

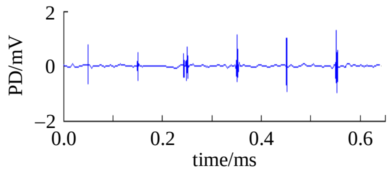


Fig.6 Denoised signals with term-by-term multi-wavelet

From Fig.5, Fig.6 and Fig.2, it can be seen that both neighbor multi-wavelet denoising and term-by-term multi-wavelet denoising can suppress the white noise efficiently. MSE of two methods are 0.0040mV, 0.0046mV respectively. It suggests that with appropriate parameters r and q, neighbor thresholding is superior to term-by-term thresholding.

5 Processing of on-site data

On-site data is derived from an online PD monitoring system installed on some power generator, and the sampling frequency of the system is 6.67MHz, while the length of the data is 262144 points. Ahead of denoising, FIR filtering is employed to the sample, and both methods are implemented. The ultimate result of processing is illustrated in Fig.8 and Fig.9.

From Fig.8 with Fig.9, it can be seen that, both methods can be used to suppress the white noise, extracting partial discharge pulses. Due to the original partial discharge is unknown; it is impossible to calculate out the performance of either method. For the sake of comparison, we apply the classical wavelet based method also (wavelet base function is db8 suggested in [12], and the denoising results is shown in Fig.10). For computing MSE, we choose the mean of three results as the reference original partial discharge. With the reference data, the final MSE of neighboring coefficient based method and classical method are calculated out, 1.0784mV and 1.2119mV respectively. The former outperforms the latter, and this coincides with the result obtained previously.

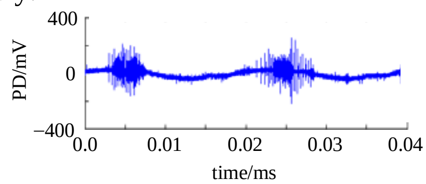


Fig.7 Partial discharge signals from power plant

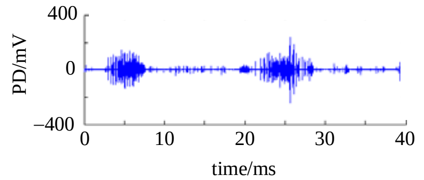


Fig.8 Denoised signals by multiwavelet with neighbouring coefficient

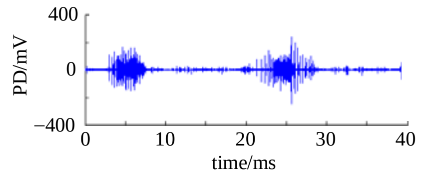


Fig.9 Denoised signals with multiwavelet (term-by-term)

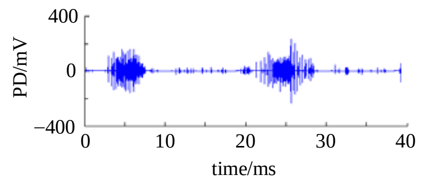


Fig.10 Denoised signals with wavelet(db8)

6 Conclusion

In this contribution, we extend the work of Chen, and furthermore apply it to denoise partial discharge signals. By massive simulation, together with on-site data processing, we conclude that:

(1) With proper q and r, neighbor multi-wavelet

denoising outperforms term-by-term multi-wavelet denoising. And this confirms the result reported by Chen.

(2) With the increase of q, MSE increases accordingly, while with the increase of r, MSE decreases, at last approximating some constant.

(3) With q=1 and $r \in [2,3]$, neighbor multi-wavelet denoising can give nearly optimal results.

Reference

[1] Huang chengjun, *Study of partial discharge on-Line monitoring system for large tubrine generators based on wavelet analysis* ,Shanghai Jiaotong University, 2000.

[2] Bui, T.D. and G. Chen, Translation-invariant denoising using multi-wavelets, IEEE Transactions on Signal Processing, Vol.46, No.12,1998,pp. 3414- 3420.

[3] Downie T R, Silverman B W, The discrete multiple wavelet transform and thresholding methods,*IEEE Trans on Signal Processing*, Vol.46, No.9,1998,pp.2558-2561.

[4] Chen G Y, Bui T D, Multiwavelets denoising using neighboring coefficients, *IEEE Signal Processing Letters*, Vol.10, No.7,2003, pp.211-214.

[5] Geronimo J Sm Hardin D P, Massopust P R, Fractal functions and wavelet expansions based on several scaling functions, *Journal of Approximation Theory*,Vol.78,No.3,1994,pp. 373 - 401.

[6] Bui T D,Chen G,Translation-invariant denoising using multiwavelets, *IEEE Trans on Signal Processing*,Vol.46, No.12, 1998, pp. 3414-3420.

[7] Chui C K,Lian J A, A study of orthonormal multi-wavelets, *Journal of Applied Numerical Mathematics*,Vol.20,No.3,1996,pp.273-298.

[8] Xia X G,Geronimo J S,Hardin D P *et al*,Design of prefilters for discrete multiwavelet transforms, *IEEE Trans on Signal Processing*, Vol.44, No.1, 1996,pp.25-35.

[9] Cotronei M,Montefusco L B,Puccio L, Multiwavelet analysis and signal processing, *IEEE Trans on Circuits and Systems-II:Analog and Digital Signal Processing*, Vol.45,No.8, 1998, pp.970- 987.

[10] Zeng Yida, Liu Zhigang, Qian Qingquan, Pre-processing methods of CL multiwavelet and their applications in power system fault data compression,*Power System Technology*, Vol.28, No.7,2004,pp.10-14.

[12] Huang Chengjun,Yu Weiyong,Study of adaptive filter algorithm based on wavelet analysis in suppressing PD's periodic narrow bandwidth noise,*Proceedings of the CSEE,* Vol.23, No.1,2003,pp.107-111.

[13] Ma X,Zhou C,Kemp I J,Interpretation of wavelet analysis and its application in partial discharge detection, IEEE Trans on Dielectrics and Electrical Insulation,Vol.9, No.3, 2002, pp. 446-457.

Appendix:

The matrix filter coefficients, prefilter and post-filter used in the paper are defined below

$$\boldsymbol{h}(0)=\begin{bmatrix} \frac{3}{5\sqrt{2}} & \frac{4}{5} \\ -\frac{1}{20} & -\frac{3}{10\sqrt{2}} \end{bmatrix}, \boldsymbol{h}(1)=\begin{bmatrix} \frac{3}{5\sqrt{2}} & 0 \\ \frac{9}{20} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\boldsymbol{h}(2)=\begin{bmatrix} 0 & 0 \\ \frac{9}{20} & -\frac{3}{10\sqrt{2}} \end{bmatrix}, \boldsymbol{h}(3)=\begin{bmatrix} 0 & 0 \\ -\frac{1}{20} & 0 \end{bmatrix}$$

$$\boldsymbol{g}(0)=\begin{bmatrix} -\frac{1}{20} & -\frac{3}{10\sqrt{2}} \\ -\frac{1}{10\sqrt{2}} & \frac{3}{10} \end{bmatrix}, \boldsymbol{g}(1)=\begin{bmatrix} \frac{9}{20} & -\frac{1}{\sqrt{2}} \\ -\frac{9}{10\sqrt{2}} & 0 \end{bmatrix}$$

$$\boldsymbol{g}(2)=\begin{bmatrix} \frac{9}{20} & -\frac{3}{10\sqrt{2}} \\ \frac{9}{10\sqrt{2}} & -\frac{3}{10} \end{bmatrix}, \boldsymbol{g}(3)=\begin{bmatrix} \frac{1}{20} & 0 \\ \frac{1}{10\sqrt{2}} & 0 \end{bmatrix}$$

$$\boldsymbol{Q}(0)=\begin{bmatrix} \frac{3}{8\sqrt{2}} & \frac{10}{8\sqrt{2}} \\ 0 & 0 \end{bmatrix}, \boldsymbol{Q}(1)=\begin{bmatrix} \frac{3}{8\sqrt{2}} & 0 \\ 1 & 0 \end{bmatrix}$$

$$\boldsymbol{P}(0)=\begin{bmatrix} 0 & 1 \\ 0 & -\frac{3}{10} \end{bmatrix}, \boldsymbol{P}(1)=\begin{bmatrix} 0 & 0 \\ \frac{4\sqrt{2}}{5} & -\frac{3}{10} \end{bmatrix}$$