

An Undecimated Wavelet Transform Based Denoising, PPCA Based Pulse Modeling and Detection-Classification of PD Signals

Pradeep Kumar Shetty, Ramu. T.S
Indian Institute of Science, Bangalore
Department of HVE, India-560012
pradeep@hve.iisc.ernet.in

Abstract

Authors Address the problem of recognition and retrieval of relatively weak industrial signal such as Partial Discharges (PD) buried in excessive noise. The major bottleneck being the recognition and suppression of stochastic pulsive interference (PI) which has similar frequency characteristics as PD pulse. Also, the occurrence of PI is random like PD pulses. In this paper we provide techniques to de-noise, detect, estimate and classify the PD signal in a statistical perspective. To avoid aliasing due to interference of high frequency noise, PD signals are generally digitized in much higher sampling rates (in terms of tens of MHz), than actually required. A multi-resolution analysis based technique is incorporated to discard the huge amount of redundant data in acquired signal. A scale dependent MMSE based estimator is implemented in undecimated wavelet transform (UDWT) domain to enhance the noisy signal, due to its inherent advantages offered in the analysis of PD signal. The probability density function of the enhanced signal is derived using probabilistic principal component analysis (PPCA) in which PD/PI pulses are modeled as mean of the distribution. The parameters of the pulses are estimated using maximum a posteriori probability (MAP) based technique. A statistical test known as generalized log likelihood ratio test (GLRT) was incorporated to ensure the existence of the pulse. The decision as to whether a pulse is a noise or a desired signal has been made based on a weighted-nearest neighbor methodology.

1. Introduction

In spite of advances in the areas of manufacturing, processing optimal design and quality control, the high voltage (HV), high power apparatus have continued to fail in service prematurely. Investigations reveal that, in most cases insulation failure is the primary cause. In this context, power utilities are increasingly resorting to on-line, on-site diagnostic measurements to appraise the condition of insulation system. Amongst others, the PD measurement is emerged as

an indispensable, non-destructive, sensitive and most powerful diagnostic tool.

A major constrain encountered with on-line digital PD measurements is the coupling of external interferences (usually very high amplitude compared to PD signal) that directly affect the sensitivity and reliability of the acquired PD data. The more important of them being, discrete spectral interferences (DSI), periodic pulse shaped interferences, external random pulsive interferences and random noise generic to measuring system itself. In most of the cases, external interferences yield false indications, there-by reducing the credibility of the PD as a diagnostic tool. Many researchers, have proposed signal processing techniques to suppress the different noise component such as, FFT thresholding, adaptive digital filter, IIR notch filter, wavelet based method [3] with varying degree of success. But acquisition of pure PD signal on-site and on-line is still elusive which forms the subject matter of this paper.

1.1. Problem Definition

Of all external interferences mentioned, DSI can be identified and eliminated in frequency domain as they have a narrow-band frequency spectrum concentrated around the dominant frequency, whereas, PD pulses have relatively a broad band frequency spectrum. Periodic pulse shaped interferences can be gated-off in time domain any PD occurring in that time interval is lost). But, it is very difficult to identify and suppress the stochastic pulse-shaped interferences (PI) as they have many characteristics in common (both in time and frequency domain) with PD pulses. Also, PI is a random occurrence like PD pulse which aggravates the process of separation [3]. Thus, PI continues to pose problems for reliable on-line, on-site PD measurement.

In this paper, the Authors are approaching the problem in a stochastic perspective to model and classify the PD as well PI. Generally, PD signal is sampled at a very high sampling frequency (in terms of MHz), to avoid aliasing. This in part is due to, coupling of high frequency noise component in measuring circuit. Therefore, a large amount of data points contain redundant information. In this regard,

Authors adopt a wavelet based multi-resolution analysis (MRA) scheme to discard the redundant data. Then, signal is decomposed into different frequency scales using an undecimated wavelet transform (UDWT). The advantage of UDWT being, availability of time as-well the frequency information in the wavelet domain, which is very useful in suppressing PI. A scale dependent FIR Wiener filter is implemented to obtain wavelet coefficient of clean signal.

The probability density function of the enhanced signal is derived using PPCA in which PD/PI pulses are characterized as mean of the distribution. The parameters (features) of the pulses are estimated using MAP based technique. A GLRT based on Neyman-Pearson (NP) criterion is performed, to doubly ensure the existence of the pulse. Finally the pulsive noise and the PD pulses are discriminated by weighted nearest neighbor classifier which utilizes the knowledge of estimated parameters of the pulses.

2. WAVELET ANALYSIS

The SNR of the PD signal is considerably low (around -40dB) and it is quite difficult to visualize the location and the form of PD pulses in the observed noisy non-stationary signal. Therefore denoising DSI, random noises and other non-pulsive interferences is a primary requirement for further analysis of the signal.

2.1. Signal Resampling

PD signal is generally sampled at higher sampling rate (nano-seconds), although, the pass-band of the PD detector happens to be around 100 to 300 kHz. This is done to avoid the aliasing effect due to interference of high frequency components in the measuring circuit. Therefore an enormous amount of redundant data is encountered in the acquired PD signal which is a great computational burden. In this section Authors provide an simple intuitive wavelet based MRA scheme to dispose the redundant data points in acquired PD signal without losing useful PD information.

Wavelet transformation is a linear transformation which is very useful in analyzing non-stationary signals. This relies on breaking up of a signal into shifted and scaled version of mother wavelet or the wavelet basis function. The continuous wavelet transform of a signal $x(t)$ is given by,

$$CWT_x^\psi(\tau, s) = \frac{1}{\sqrt{|s|}} \int x(t) \psi^* \frac{(t-\tau)}{s} dt \quad (1)$$

Where, $\psi(t)$ is the mother wavelet. s & τ are respectively, dilation and translation parameters. Generally for most of the signal analysis, discrete wavelet transforms (DWT) based on MRA are used, because of computational and implementational ease. MRA proposed by Mallat [6] relies on breaking of the frequency space of the signal into two complementary subspaces and further decomposition is done on the space having lower frequency components.

Based on similar lines, Authors decompose the observed noisy PD data into predetermined levels which is determined *a priori* depending on the sampling frequency and

bandwidth of the PD detector. Let F_s be the sampling frequency and F_d be the upper cutoff frequency of the PD detector. Then, the noisy signal is decomposed into L levels using MRA scheme with Daubeches wavelet (*db16*) as basis function. where,

$$L = \lfloor \frac{\log(F_s) - \log(F_d)}{\log(2)} - 1 \rfloor \quad (2)$$

It can be easily verified that, the wavelet coefficients across all the scales correspond to frequencies higher than the pass-band of PD detector and scaling coefficients correspond to frequency range $0 - F_d$. Therefore, wavelet coefficients across all the scales were nulled and signal is reconstructed considering only the scaling coefficients. By doing this, we have effectively removed the undesired higher frequency components in the observed signal. Then, signal is down-sampled by a factor M , where, $M = 2^L$. The Computation labor of the proposed algorithms in the following sections are greatly reduced by implementing this step and all the future analysis are carried out using resampled data.

2.2. Signal Enhancement

Most of the high frequency component of observed noisy PD signal was removed by implementation of MRA based resampling scheme. Still considerable noise component remains in the signal. Experimental evidences reveal that, PD pulses commonly occur around the phase angle $\{0 - \frac{\pi}{2}\}$ and $\{\frac{3}{2}\pi - 2\pi\}$ of the power frequency cycle. This prior information can be effectively used, if we have the control over time in the transformed domain. Authors implement an scale dependent MMSE based estimator in the UDWT domain. The important advantage of UDWT is that, not only the denoising performance is better than DWT [2], but also the availability of time index in transformed domain makes it an attractive and most suitable tool for PD denoising. Authors employ *a'trous algorithm* scheme with Daubeches wavelet (*db16*) as basis function to implement UDWT [6].

Let, $y = x + n$ be the signal model, wherein y , x and n are respectively k dimensional observed noisy signal, clean pulsive signal and non-pulsive noise respectively. By taking UDWT, we have, $Wy = Wx + Wn$ which is denoted as $y' = x' + n'$. Let H be the operator to modify the wavelet coefficients to obtain the enhanced clean signal. Then, $\hat{x} = W^{-1}H(y')$, where W^{-1} represents the inverse wavelet transform. A spatially selective Wiener filter is adopted to modify the wavelet coefficients. Authors treat the wavelet coefficients in each scale as a time series and a scale dependent Wiener filter is implemented. The wavelet coefficients of the clean signal $x' = y' * p$, where, p represents Wiener filter coefficients, which is given by,

$$p = R_{x'}^{-1} r_{x'n'} \quad (3)$$

$R_{x'}^{-1}$ is the autocorrelation matrix and $r_{x'n'}$ is vector of cross-correlations. Authors take UDWT of pulse-free noise region from which autocorrelation matrix in each scale for pure noise is obtained and $R_{x'}^{-1}$ is computed using additive

signal model in UDWT domain. The number of filter coefficients in each scale is determined using the correlogram.

3. Feature Extraction

In section II, Authors employed a UDWT based approach to suppress large portion of non-pulsive noise from the observed signal (y). Further analysis of the enhanced signal, is made by windowing the relevant part of the enhanced signal (x). A simple peak detector is used to extract the pulses and further analysis is undertaken on this. Authors incorporate a probabilistic principal component analysis (PPCA) technique to model the enhanced signal.

3.1. PPCA Based Data Modeling

A k -dimensional data vector x can be related to a q -dimensional ($q < k$) latent variable z as,

$$x = Pz + h + \omega \quad (4)$$

where, z and ω are independent random processes [1] and h is the mean of the data vector x . By defining a prior pdf for to z , the model induces corresponding pdf to x .

Let, $z \sim N(0, C)$ and $\omega \sim N(0, \sigma^2 I)$. Then, x is also Gaussian. i.e $x \sim N(h, PP^T + \sigma^2 I)$. The columns of P are scaled and rotated principal eigenvectors of covariance matrix [4] of x . In this model, the remnant noise (posterior to UDWT enhancement) is modeled as combination of systematic component (Pz) and independent component (ω). The systematic component captures the trend and ω models the random part in the data vector x . The important advantage of the above model is that, the observed data is not only represented in terms of principal components, but also resulted in a probability model for the observed data.

The ML estimate of of P and σ^2 are given by.

$$P = U_q (\Lambda_q - \sigma^2 I)^{\frac{1}{2}} R$$

$$\sigma^2 = \frac{1}{k-q} \sum_{j=q+1}^k \lambda_j \quad (5)$$

where q column vectors in U_q are eigen vectors of covariance matrix of the data, with the corresponding eigen values in the diagonal matrix Λ_q and R is an arbitrary rotation matrix. The estimate for σ^2 has the interpretation of lost variance in the projection, averaged over the lost dimension. The mean vector h is modeled as a function to describe the noise-free PD/PI pulse in the windowed signal.

$$h(t) = \eta \begin{cases} e^{\alpha_1(t-t_0)} & 0 \leq t \leq t_0 \\ e^{-\alpha_2(t-t_0)} & t > t_0 \end{cases} \quad (6)$$

Where η , is the pulse height, α_1, α_2 are the parameters which are functions of rise time and fall time respectively. The time t_0 is the time at which the value of h is η . The autocovariance matrix C is estimated from the part of the signal x , outside the window. The value of the parameter set $\theta = [\eta, \alpha_1, \alpha_2]$ is estimated, as explained below.

3.2. MAP Parameter Estimation

We know that, $p(\theta|x) \propto p(x|\theta)p(\theta)$, where, $p(\theta|x)$, $p(x|\theta)$, $p(\theta)$ are respectively posterior probability of θ , conditional density of x and prior pdf of θ .

$$p(x|\theta) = \frac{1}{(2\pi)^{\frac{k}{2}} |C|^{\frac{1}{2}}} \exp[-\frac{1}{2}(x-h)^T C^{-1}(x-h)] \quad (7)$$

where $C = PP^T + \sigma^2 I$, which is PPCA based model. We assume independence of the parameters $[\eta, \alpha_1, \alpha_2]$, thus $p(\theta) = p(\eta).p(\alpha_1).p(\alpha_2)$. wherein $\eta \sim N(\mu_\eta, \sigma_\eta^2)$, $\sigma_1 \sim N(\mu_{\sigma_1}, \sigma_{\sigma_1}^2)$ and $\sigma_2 \sim N(\mu_{\sigma_2}, \sigma_{\sigma_2}^2)$. Let, $\mu_\theta = [\mu_\eta, \mu_{\sigma_1}, \mu_{\sigma_2}]$, and $V_\theta = \text{diag} [\sigma_\eta^2, \sigma_{\sigma_1}^2, \sigma_{\sigma_2}^2]$. Thus, $p(\theta|x) = K.e^{[-\frac{1}{2}[(x-h)^T C^{-1}(x-h) + (\theta-\mu_\theta)^T V(\theta)^{-1}(\theta-\mu_\theta)]]}$, where, $K = [(2\pi)^{\frac{k+3}{2}} \sigma_\eta \cdot \sigma_{\sigma_1} \cdot \sigma_{\sigma_2} \cdot |C|^{\frac{1}{2}}]^{-1}$. The MAP estimation of the parameter set θ is obtained by maximizing the posterior probability or can be shown to be minimizing the cost function,

$$L = [(x-h)^T C^{-1}(x-h) + (\theta-\mu_\theta)^T V(\theta)^{-1}(\theta-\mu_\theta)] \quad (8)$$

Sample values of the estimated parameters are given by,

$$\theta_{PD1} = [0.47, 0.05, 0.48], \quad \theta_{PI1} = [1.21, 0.11, 0.08]$$

$$\theta_{PD2} = [0.94, 0.04, 0.7] \quad \theta_{PI2} = [1.41, 0.15, 0.13].$$

Classification of the pulses (ie PD or PI) is accomplished based on these parameter values (features).

4. Detection and Classification of PD/PI Pulses

In this section the Author introduces a detection/classification technique based on generalized log-likelihood ratio test and weighted nearest-neighbor methodology.

4.1. Generalized Log-Likelihood Ratio Test

An GLRT is adopted [5], to doubly ensure the existence of a pulse in the window. This is an important step prior to a binary classification detailed in the following section, which ascertains whether a PD pulse is included in the window or not. A hypotheses test is instituted to verify this as under:

$$H_0 : x = \omega$$

$$H_1 : x \neq \omega \quad (9)$$

be the two hypotheses. Here, H_0 is the null hypotheses and H_1 is the alternative. Under $H_0 : x \sim N(0, C)$, is known, under $H_1 : x \sim N(h, C)$, is unknown. Therefore a GLRT based statistical testing has been employed for making decision on the existence of the pulse. In GLRT the unknown parameters are replaced by their maximum likelihood (ML) estimates and the asymptotic performance of the *GLRT - statistic* is deduced. Asymptotically, this statistic tends to a chi-square distribution. It can be shown [5] that,

$$2 \ln(L_G(x)) \sim \chi_3^2 \quad \text{under } H_0$$

$$\sim \chi_3^2(\lambda) \quad \text{under } H_1 \quad (10)$$

Where, χ_3^2 represents a chi-square distribution with three degrees of freedom. The non-centrality parameter of the chi-square distribution is given by λ . The likelihood ratio $L_G(x)$ is given by,

$$L_G(x) = \frac{p(x, \hat{\theta}_1; H_1)}{p(x, \theta_0; H_0)}$$

Under H_0 , pdf is completely defined by the parameter C . Under H_1 , the pdf is a joint function of C and the ML estimate of $\theta = \hat{\theta} = [\hat{\eta}, \hat{\alpha}_1, \hat{\alpha}_2]$. Thus, the GLRT-statistic is given by, $2 \ln(L_G(x)) = 2x^T C^{-1} \hat{h} - \hat{h}^T C^{-1} \hat{h}$ (11) Now, the pulse is believed to exist, implying that H_1 is true, if, $2 \ln(L_G(x))$ greater than a threshold, γ , where $\gamma = Q^{-1}(1 - P_{fa})$. The term, Q^{-1} is the right-tail probability of chi-square distribution with three degrees of freedom, P_{fa} is termed as probability of false alarm, which is, probability that GLRT-statistic exceeds γ under H_0 .

4.2. Classification of PD Pulses

A weighted nearest neighbor classifier, based on the Euclidean distance from the unlabeled instance to the training set has been implemented, wherein, one class represents the PD pulses and the other one represents the PI.

$$d_m(p, q_m) = [w_1(\alpha_{1p} - \alpha_{1q_m})^2 + w_2(\alpha_{2p} - \alpha_{2q_m})^2 + w_3(\eta_p - \eta_{q_m})^2]^{\frac{1}{2}} \quad (12)$$

Where, p is the unlabeled testing parameter set. q_m represents m^{th} training parameter set and $w = [w_1, w_2, w_3]$ is the weight matrix. which is carefully chosen for effective classification (since the parameter α_1 has the more discriminant feature w_1 is relatively higher than other weights).

5. Results and Discussion

The results shown in this paper pertains to both simulated as well the real data. Noisy PD data is simulated on the similar lines as given by Satish [3]. The real PD data was obtained in laboratory conditions, using a point-plane configuration for PD source [3]. The bandwidth of the PD detector was chosen to be 30kHz-300kHz. A 10 bit digitizer with sampling time of 100ns was employed to obtain PD data. PI pulses seen in Fig. I(c) are completely removed

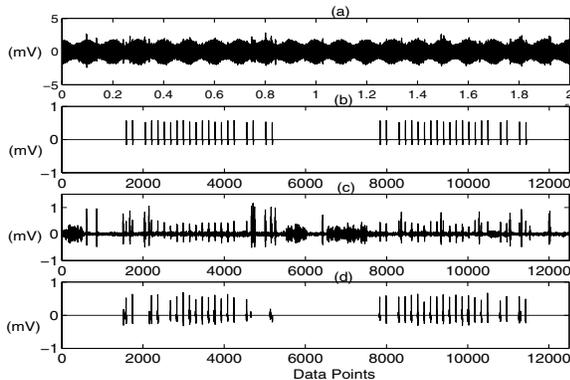


Figure 1: (a) Simulated noisy signal (b) Location of constant height PD pulses (c) Enhanced signal using FIR Wiener filter in UDWT domain (d) Retrieved PD pulses.

and only PD pulses are retrieved without any time shift as shown in Fig. I(d). About 95% of the total PD pulses are retrieved and the loss of the remaining pulses are mainly attributed to the overlapping of PD and PI pulses in time. The

performance of the technique on real data is shown in Fig. 2. The MMSE based UDWT domain enhancement is most suitable for PD signal analysis and performance was found to be better than other techniques listed in the literature. An

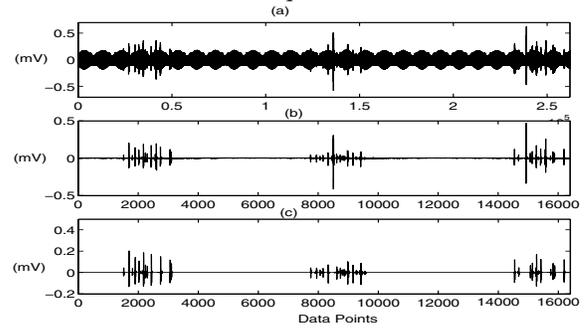


Figure 2: (a) Real noisy signal (b) Enhanced signal by UDWT FIR Wiener filter. (c) Retrieved PD pulses.

energy detector was incorporated to detect the ringing component of the pulses and noise at non-PD location is completely nullified.

6. Conclusion

A huge amount of redundant data was removed from observed signal, there-by decreasing the computation time, which is very important for on-line applications. The simultaneous availability of time and frequency band knowledge in UDWT domain was exploited for the first time in PD signal analysis and enhancement technique found to be better than most of the methods reported in the literature. The PPCA based method is innovative and models the enhanced signal effectively. For the first time, the functional characterization for pattern recognition of PD signal is undertaken and PD/PI pulses are efficiently detected and classified. Thus a new, complete, mathematically robust pattern recognition methodology for on-site, on-line retrieval the PD pulses buried in excessive noise has been implemented and tested.

References

- [1] M. Svensen, F. Kruggel, D. Y. Von Cramon, Probabilistic Modeling of Single-Trial fMRI Data, IEEE trans. Medical Imaging, vol, 19, no. 12, pp. 1188-1201, 2000
- [2] H. Zhang et. al. Image Denoising Via Wavelet-Domain Spatially Adaptive FIR Wiener Filtering, Proc. IEEE, ICASSP-2000, pp. 2179-2182.
- [3] L. Satish, B. Nazneen, Wavelet Denoising of PD Signals buried in Excessive Noise and Interference, IEEE, Transaction, on, DEI, pp. 354-367, 2003.
- [4] M.E. Tipping, C.M Bishop, A Hierarchical Latent Variable Model for Data Visualization, IEEE trans. PAMI. Vol. 20, no-3, pp.25-35, 281-293, 1998.
- [5] S. M. Kay Fundamentals of Statistical Signal Processing. Estimation-Detection Theory, Prentice Hall PTR, Newjersy, pp.186-247,1998.
- [6] S.Mallat, A Wavelet Tour of Signal Processing, Academic Press, 1999.