AN UNDECIMATED WAVELET TRANSFORM BASED ENHANCEMENT, STATISTICAL FEATURE EXTRACTION AND DETECTION-CLASSIFICATION OF PD SIGNALS

Pradeep Kumar Shetty, T.S.Ramu

Department of HVE
Indian Institute of Science
Bangalore-560012, INDIA
E-mail: pradeep@hve.iisc.ernet.in

ABSTRACT

We address the problem of recognition and retrieval of relatively weak industrial signal such as Partial Discharges (PD) burried in excessive noise. The major bottleneck being the recognition and supression of stochastic pulsive interference (PI) which has similar frequency characteristics as PD pulse. In this paper, we provide techniques to de-noise, detect, estimate and classify the PD signal in a statistical perspective. 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. We characterize the PD and PI pulses using a statistical model as the first moment of multi variate Gaussian distribution and its parameters are estimated using maximum likelihood (ML) and maximum aposteriroi probability (MAP) based techniques. 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

Inspite 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 [3], adaptive digital filter [6], IIR notch filter [7], wavelet based method [2] 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 occuring 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, pulsive noise is a random occurrence like PD pulse which agravates the process of separation. Thus, pulse shaped interference continues to pose problems for reliable on-line, on-site PD measurement.

In this paper, we 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 (interms 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, we 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 feature of the PD and PI are obtained by modeleling them as first moment of multi variate Gaussian distribution with a double exponential function and its parameters are estimated using ML and Bayesian (MAP) techniques. A generalized log-likelihood ratio test (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 methodology, 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 Compression

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

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,

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

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 implimentational ease. MRA proposed by Mallat [9] relies on decomposition of the frequency space of the signal into two complementary subspaces and further decomposition is done on the vector space containing lower frequency components.

Based on similar lines, we decompose the observed noisy PD data into predetermined levels which is determined *apriori* 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,

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

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 greately 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 compression 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. We implement an scale dependent MMSE based estimator in the UDWT domain for two reasons. First one being, the availability of time index in the transformed domain and second one being the better denoising performance of non-stationary signals compared to DWT [1]. We rely on $a'trous\ algorithm$ scheme [9] with Daubeches wavelet ($db16$) as basis function to implement UDWT.

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. An spatially selective Wiener filter is used to modify the wavelet coefficients. We 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 $x' = y' * h$, where h represents Wiener filter coefficients, which is given by,

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

$R_{x'}^{-1}$ is the autocorrelation matrix and $r_{x'n'}$ is vector of cross-correlations. We take UDWT of pulse-free noise region from which autocorrelation matrix in each scale for pure noise is obtained and $R_{x'}^{-1}$ is obtained using additive signal model in UDWT domain. The number of filter coefficients in each scale is determined using the correlogram.

3. PARAMETRIC CHARACTERIZATION OF PD AND PI PULSES

The segregation of PD and noise pulse is often difficult process and off-line processing procedure is normally adopted. In this section we develop techniques to remove PI, on-line. The experimental evidences suggest that, some of the properties, such as, rise time, pulse height are useful in discriminating the PD pulses from PI. Thus, we characterize the pulses in a functional form and try to incorporate these information. It was found that, by the adopted enhancement technique, the location of the pulses were quite clear. Therefore a simple peak detector was employed to locate the pulses and rectangular window was used to extract the pulses from detected location. A GLRT based test explained in the next section ensures the exclusion of non-pulsive noise in the window.

3.1. Stochastic Analysis of PD/Interference Pulses.

Let,
$$x = h + \omega \quad (4)$$
be the signal model, where, x is the k-dimensional windowed signal vector. Complete denoising of non-pulsive interference is rarely possible by known enhancement techniques. Therefor, ω represents the remanant noise. Let, $\omega \sim N(0, C)$ and $x \sim N(h, C)$, where h describes the noise-free PD/PI. Therefore the pulses are characterized as the first moment of the multi variate Gaussian distribution (MVGD) and assumed to follow a piecewise continuous in time. Thus:

$$\begin{aligned} h(t) &= \eta\left(e^{\alpha_1(t-t_0)}\right) \qquad 0 \le t \le t_0 \\ &= \eta\left(e^{-\alpha_2(t-t_0)}\right) \qquad t > t_0 \end{aligned} \quad (5)$$

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. Parameter Estimation

We propose ML and MAP estimation procedure for the estimation of pulse parameters [5].

3.2.1. ML Estimation

Let $p(x;\theta)$ represent the probability density function (pdf) of x. Then,
$$p(x;\theta) = \frac{1}{(2\pi)^{\frac{k}{2}}|C|^{\frac{1}{2}}} exp[-\frac{1}{2}(x-h)^T C^{-1}(x-h)] \quad (6)$$
Where, $|C|$ represents determinent of C. To find out the parameters, following equation is minimized.
$$L = \frac{1}{2}(x-h)^T C^{-1}(x-h) \quad (7)$$

3.2.2. MAP Estimation

We know that, $p(\theta|x) \propto p(x|\theta)p(\theta)$
where, $p(\theta|x)$, $p(x|\theta)$, $p(\theta)$ are respectively posteriori probability of θ, conditional density of x and prior pdf of θ. We assume independence of the parameters $[\eta, \alpha_1, \alpha_2]$, thus $p(\theta) = p(\eta).p(\alpha_1).p(\alpha_2)$. wherein $\eta \backsim N(\mu_\eta,\ \sigma_\eta^2)$, $\sigma_1 \backsim N(\mu_{\sigma_1}, \sigma_{\sigma_1}^2)$ and $\sigma_2 \backsim N(\mu_{\sigma_2}, \sigma_{\sigma_2}^2)$. Let,
$\mu_\theta = [\mu_\eta, \mu_{\sigma_1}, \mu_{\sigma_2}]$, and $V_\theta = diag\ [\sigma_\eta^2, \sigma_{\sigma_1}^2, \sigma_{\sigma_2}^2]$. 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, $K = [(2\pi)^{\frac{k+3}{2}}\sigma_\eta.\sigma_{\sigma_1}.\sigma_{\sigma_2}.|C|^{\frac{1}{2}}]^{-1}$. The MAP estimation of the parameter set θ is obtained by maximizing the posteriori probability, which 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,
θ_{PD1} = [0.47, 0.05, 0.48], θ_{PI1} = [1.21, 0.11, 0.08]
θ_{PD2} = [0.94, 0.04, 0.7] θ_{PI2} = [1.41, 0.15, 0.13]

4. DETECTION AND CLASSIFICATION OF PD PULSES

In this section, we introduce 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 [4], 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:
$$\begin{aligned} H0 &: x = \omega \\ H1 &: x \neq \omega \end{aligned} \quad (9)$$
be the two hypotheses. Here, H_0 is the null hypotheses and H_1 is the alternative. Under $H_0 : x \backsim N(0, C)$, is known, under $H_1 : x \backsim N(h, C)$, 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 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 [4] that,
$$\begin{aligned} 2\,ln(L_G(x)) &\sim \chi_3^2 \qquad under\ H_0 \\ &\sim \chi_3^2(\lambda) \quad under\ H_1 \end{aligned}$$
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 fully 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,
$$2\,ln(L_G(x)) = 2x^T C^{-1}\hat{h} - \hat{h}^T C^{-1}\hat{h} \quad (10)$$

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 binary classifier known as weighted nearest neighbor classifier, based on the Eucledian 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.

$$d_m(p, q_m) = [w_1(\alpha_{1p} - \alpha_{1qm})^2 + w_2(\alpha_{2p} - \alpha_{2qm})^2 + w_3(\eta_p - \eta_{qm})^2]^{\frac{1}{2}} \quad (11)$$

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 weight matrix. which is carefully choosen 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 [2]. The real PD data was obtained in laboratory conditions, using a point-plane configuration for PD source [2]. The bandwidth of the PD detector was choosen to be 30kHz-300kHz. A 10 bit digitizer with sampling time of 100ns was employed to obtain PD data. PI

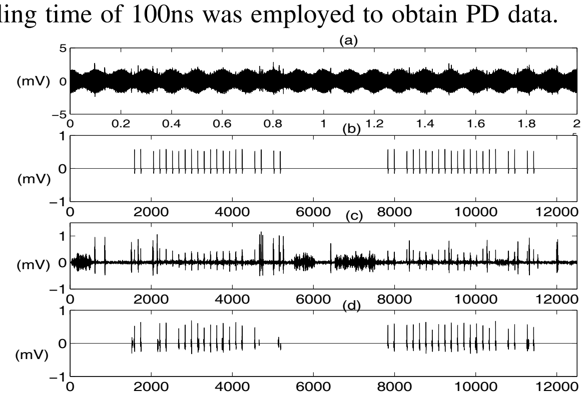

Fig. 1. (a) Simulated noisy signal (b) Location of constant height PD pulses (c) Enhanced signal using FIR Wiener filter in UDWT domain (d) Retrieved PD pulses.

pulses seen in Fig. I(c) are completely removed 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 energy detec-

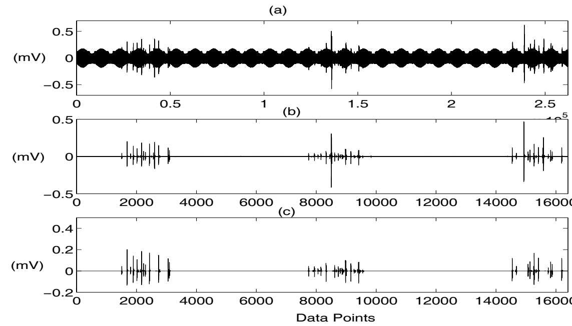

Fig. 2. (a) Real noisy signal (b) Enhanced signal using FIR Wiener filter in UDWT domain (c) Retrieved PD pulses.

tor 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 statistical methods presented here is innovative and found to be very useful in on-line, on-site PD signal analysis. Thus a new, complete, mathematically robust signal processing methodology for on-site, on-line retrieval the PD pulses buried in excessive noise has been implemented and tested.

7. REFERENCES

[1] H. Zhang et. al. "Image Denoising Via Wavelet-Domain Spatially Adaptive FIR Wiener Filtering". Proc IEEE, ICASSP 2000, pp. 2179–2182.

[2] L. Satish, B. Nazneen, "Wavelet Denoising of PD Signals burried in Excessive Noise and Interference", IEEE Transaction on DEI, pp 354–367, 2003.

[3] K. Feser, G. Konig, J. Ott and P. Setiz, "An Adaptive Filter Algorithm for On-site Partial Discharge Measurements". IISEI , Boston, USA, pp. 242–245, 1988.

[4] S. M. Kay " Fundamentals of Statistical Signal Processing-Detection Theory" Prentice Hall PTR, Newjersy, pp. 186–247, 446–472, 1998.

[5] S. M. Kay "Fundamentals of Statistical Signal Processing-Estimation Theory", Prentice Hall PTR,Newjersy, 1998.

[6] H. Borsi and M. Hartje, "New Methods to Reduce Disturbance Influence on On-site PD Measurement and Monitoring". 6^{th} ISH , New-Orleans USA, paper 15.10, 1989.

[7] V. Nagesh and B. I. Gururaj, "Evaluation of Digital Filter for Rejecting DSI in on-site PD measurements ". IEEE Transactions on EI, Vol. 28, pp. 73–84, 1993.

[8] S. Goswami, "Introduction to Wavelets". Academic Press, 1999.

[9] S. Mallat, "A Wavelet Tour of Signal Processing". Academic Press, 1999.