

A Combined fBm and PPCA Based Signal Model for On-Line Recognition of PD Signal

Pradeep Kumar Shetty

Honeywell Technology Solutions Lab, Bangalore, India
pradeep.shetty@honeywell.com

Abstract. The problem of on-line recognition and retrieval of relatively weak industrial signal such as Partial Discharges (PD), buried in excessive noise has been addressed in this paper. The major bottleneck being the recognition and suppression of stochastic pulsive interference (PI), due to, overlapping broad band frequency spectrum of PI and PD pulses. Therefore, on-line, on-site, PD measurement is hardly possible in conventional frequency based DSP techniques. We provide new methods to model and recognize the PD signal, on-line. The observed noisy PD signal is modeled as linear combination of systematic and random components employing probabilistic principal component analysis (PPCA). Being a natural signal, PD exhibits long-range dependencies. Therefore, we model the random part of the signal with fractional Brownian motion (fBm) process and pdf of the underlying stochastic process is obtained. The PD/PI pulses are assumed as the mean of the process and non-parametric analysis based on smooth FIR filter is undertaken. The method proposed by the Author found to be effective in recognizing and retrieving the PD pulses, automatically, without any user interference.

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, 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 with varying degree of success [1]. Due to the inherent difficulties involved in signal modeling and pattern analysis of PD signal, on-line, on-site PD measurement is still elusive which forms the subject matter of this paper.

1.1 Problem Enunciation

Discrete spectral interferences can be identified and eliminated in frequency domain as they have a distinct narrow-band frequency spectrum concentrated around the dominant frequency, whereas, PD pulses have relatively a broad band 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 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, which aggravates the process of separation. Thus, PI continues to pose serious problems for reliable on-line, on-site PD measurement.

In this paper, we analyze the PD time series in statistical perspective to model and classify the PD data. Locating the PD/PI pulses are the first step in further analysis of the signal. In this regard, we enhance the observed noisy signal using wavelet based soft thresholding method and the pulses are detected using simple peak-detector. Windowed signal of appropriate length is then taken around the detected location and further analysis is undertaken on the windowed signal. Considering the signal structure of the observed PD signal, we incorporate a combination of PPCA and fBM process based methodology, which effectively models the systematic component and random component in the observed PD signal. Since the PD signal is combination of different sinusoidal and random noises, this method is quite realistic, also, the long-range dependence of this natural signal is effectively modeled by fBM process. The pure PD/PI pulses are modeled as mean of the process and Non-parametric analysis based on smooth FIR filters is undertaken to extract the features of the pulse. The procedure adopted in this work is completely different from the research work reported in the literature, which is generally based on deserved signal frequency and noise frequency.

1.2 PD/PI Pulse Detection

It has been observed that, PD and PI pulses randomly occur in time. Therefore detection of the pulses is a primary requirement in further analysis of the signal. The signal-to-noise ratio of the PD signal is generally less (around -25dB) and it is difficult to visualize the location and the form of pulses in the observed noisy signal. In this regard, we denoise the noisy observed signal using wavelet based *sureshrink* soft thresholding method and make use of a simple peak detector to detect the location of pulsive activity. A minimum of S scale discrete wavelet transform is taken, where, $S = \lfloor \frac{\log(F_s) - \log(F_d)}{\log(2)} - 1 \rfloor$. Here, F_s is the sampling frequency and F_d is the upper cutoff frequency of the PD detector. A windowed signal of appropriate size is taken around the detected location for further analysis.

2 Probabilistic PCA

Principal Component Analysis (PCA) is a widely used tool for data analysis. Given a set of d -dimensional data vector y , the q principal axes $U_j, j = 1, 2, \dots, q$, are those onto which the retained variance under projection is maximal. These principal axes are the q eigenvectors corresponding to the q dominant eigenvalues of the sample covariance matrix of the data y . The analysis using PCA does not involve any probability model for

the data. Tipping and Bishop [2] showed that by assuming a latent variable model for the data vectors, the data vectors can be represented in terms of its principal components. This approach is very useful because, we not only represent the data in terms of its principal components, but also a probability model for the data can be derived. This model in turn can be used for the tasks like estimation and detection of signals.

2.1 Probability Model for PCA

Any d -dimensional data vector y can be related to q -dimensional ($q < d$) latent variables z as:

$$y = h + Lz + \gamma \tag{1}$$

where, γ and z are independent random processes. h is the mean of the data vectors. By defining a prior pdf to z , the above equation induces a corresponding pdf to y . If we assume $z \sim N(0, I_q)$ and $\gamma \sim N(0, C_\gamma)$, then, y is also a Gaussian with, $y \sim N(h, LL^T + C_\gamma)$, where, I_q and I are $q \times q$ and $d \times d$ identity matrices. With the above pdf's for z and γ , we can show that the columns of L are the rotated and scaled principal eigenvectors of the covariance matrix of the data vector y . In the above model, the observed vector y is represented as the sum of systematic component (Lz) and random noise component (γ). It is shown in [2] that the ML estimate of L and σ^2 are given by, $L = U_q(A_q - \sigma^2 I)^{1/2} R$, where, the q column vectors in U_q are the eigenvectors of the covariance matrix of the data with the corresponding eigenvalues in the diagonal matrix A_q . R is an arbitrary rotation matrix. The energy in the remaining $(d - q)$ eigen vectors is given by σ^2 . The model order (q) is estimated using Akaike information criterion (AIC), which is found to be two.

3 A Combination of PPCA and fBm Time-Series Model

We propose a model for the analysis of PD/PI pulses buried in noise, as:

$$y(t) = \sum_{k=0}^d x(t - k)h(k) + w(t), \quad t = 0 \dots d - 1 \tag{2}$$

where, y , x and w are observed time series, system impulse response and noise component respectively. We model h by non parametric model based on smooth FIR filter. The eqn. 2 can be written in matrix form as, $y = Xh + w$, where, X is the convolution matrix, which is identity matrix I_d . In this, the noise w can be represented using the latent variable model as explained in section 2. i.e, $w = Lz + \gamma$, where z is a q -dimensional ($q < N$) latent variable and γ is a random noise component. This method is quite realistic in modeling the PD signal, since the observed PD signal is combination of pulses, DSI and other random components. The matrix L is called as systematic noise matrix, which characterizes the systematic noise component by considering q principal components of w , corresponding to the first q dominant eigenvalues. Assuming Gaussian pdf models as described in the section 2, the pdf of noise can be given as, $w \sim N(0, C_y)$, where, $C_y = LL^T + C_\gamma$. Finally, the observed time series can be represented as,

$$y = Xh + Lz + \gamma \tag{3}$$

Therefore, the probability model for the observed PD/PI time-series y for a given h is distributed as,

$$y|h \sim N(Xh, C_y) \tag{4}$$

The physical phenomena like PD exhibits long-term dependencies and $1/f$ type of behaviour over wide range of frequencies [3]. Also, the natural signals are non-stationary in nature. Therefore, the standard assumption of independence of noise random variables are not valid in modeling the natural signals. We model the noise component (γ) as, $\gamma = \gamma_d + \gamma_i$, where, γ_d describe the long-memory process and γ_i represent the independent random noise. One well-known model of long-memory processes proposed by Mandelbrot and VanNess [4] is fractional Brownian motion. Among others, self-similarity property makes wavelet transform, a preferred tool for analysis of the fBm processes. Thus, By taking DWT of scale m , we have $W\gamma = W\gamma_d + W\gamma_i$. Exploiting the Karhunen-Loeve-like properties of the wavelet decomposition for $1/f$ type of process, we have, $W\gamma_d$ and $W\gamma_i$ uncorrelated and independent of each other. The variance of the wavelet coefficients for each scale is: $var(W\gamma) = \sigma_\gamma^2 = \sigma_d^2\beta^{-m} + \sigma_i^2$. The parameter β is related to Hurst component (H), which describes the fBm process completely. The Hurst component is given by, $H = (\log_2\beta - 1)/2$. The covariance function of the self-similar process with H is: $R_d(t, s) = \frac{\sigma_H^2}{2} (|s|^{2H} + |t|^{2H} - |t-s|^{2H})$, where, $\sigma_H^2 = \Gamma(1 - 2H)\cos(\pi H)/(\pi H)$. The parameter set $\Theta = [H, \sigma_d^2, \sigma_i^2]$ are estimated using ML technique in wavelet domain. The likelihood function L is given by,

$$L(\Theta) = p(W\gamma ; \Theta) = \prod_{m,n \in R} \frac{1}{\sqrt{2\pi\sigma_\gamma^2}} \exp\left(-\frac{(W\gamma_n^m)^2}{2\sigma_\gamma^2}\right) \tag{5}$$

where, n represents the number of wavelet coefficients in scale m . The covariance matrix (C_γ) is obtained as, $C_\gamma = C_d + C_i$. The covariance structure (C_d) is estimated using H and C_i is given by, $C_i = \sigma_i^2 I_d$.

A set of training data (y_t) for noise (i.e. non-pulsive region of the data) is required to estimate C_y , which requires the prior knowledge of absence of PD/PI pulse in the observed data. To extract the random noise component from y_t , we project y_t on to DSI space (given by $U_q U_q^T$) and γ is obtained using the additive model as given in Eqn. 3. Finally, the overall covariance matrix C_y is obtained using, $C_y = LL^T + C_\gamma$.

4 Feature Extraction Based on Smooth FIR Filter

In this section we attempt to extract the feature of PD and PI pulses by completely representing them using a smooth FIR filter, irrespective of the shape of the pulse. This method aspire to estimate the function at each time sample making no assumption of the PD/PI shape and robust enough to withstand low SNR, without over constraining the problem. The basic approach to model pulses by a FIR filter is taken from Marrelec [6] in which, the noise space is assumed to be spanned by a third order polynomial functions. In our work, we evaluate the basis vectors for the noise subspace at each windowed section using PPCA and fBm combination model.

In this methodology, the coefficients of the filter emulate the desired pulse. Therefore, estimating the filter coefficients at each windowed signal is the task at hand. The estimation problem is ill posed, since, the number of coefficients to be estimated are large. We assume that, PD/PI pulses are smooth and regularize the problem by introducing smoothing priors. Therefore, we choose to work in a Bayesian framework. The smoothing constraint is imposed by setting up a Gaussian prior for the norm of second derivative of the pulse. The relative weight of this prior to data is controlled by the hyperparameter ϵ .

The pdf of observed data (y) conditioned on h can be written as:

$$p(y|h) = \frac{1}{(2\pi)^{\frac{k}{2}} |C_y|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(y - Xh)^T C_y^{-1} (y - Xh)\right] \quad (6)$$

The smoothing prior for h is given by, $p(h|\epsilon) \propto (\epsilon^2)^{\frac{k-1}{2}} \exp(-\frac{\epsilon^2}{2} h^T M_p h)$, where, ϵ is the hyper parameter which is a measure of relative weight of the prior. The higher ϵ , the more the prior constraint taken into account. M_p is called as concentration matrix of the dimension $k - by - k$. A Jeffrey's non-informative prior [7] is used for ϵ , which is given by, $p(\epsilon) = \frac{1}{\epsilon}$. Using Bayes rule, the joint posterior pdf for h and ϵ can be written as, $p(h, \epsilon|y) \propto p(y|h, \epsilon)p(h|\epsilon)p(\epsilon)$. Therefore, the required marginal posterior pdf for h is given by, $p(h|y) = \int p(h, \epsilon|y) d\epsilon$. The closed form solution to this marginal pdf is not possible and therefore numerical optimization is required to find the maximum a posteriori probability (MAP) estimate of h . To overcome this problem, we first estimate ϵ by MAP and approximate the a posteriori pdf for h as, $p(h|y) \approx p(h|y, \epsilon = \hat{\epsilon})$, where $\hat{\epsilon}$, the MAP estimate of ϵ which is given by, $\hat{\epsilon} = \arg \max_{\epsilon} p(\epsilon|y)$, The marginal posterior probability for ϵ is given by:

$$p(\epsilon|y) \propto |M(\epsilon)|^{-1/2} \epsilon^{k-2} \exp\left(\frac{-1}{2} y^t C_y^{-1} y - \hat{\mu}_B^t M(\epsilon) \hat{\mu}_B\right) \quad (7)$$

wherein,

$$M(\epsilon) = X^t C_y^{-1} X + \epsilon^2 M_p$$

$$\hat{\mu}_B = M(\epsilon)^{-1} (X^t C_y^{-1} \mathbf{y})$$

Equation 7 is maximized w.r.t ϵ , to get the MAP estimate. Now the task is to estimate the filter coefficients h . We have, $p(h|y, \epsilon = \hat{\epsilon}) \propto p(y|h, \epsilon = \hat{\epsilon})p(h|\epsilon = \hat{\epsilon})$. i.e $p(h|y, \epsilon = \hat{\epsilon}) \propto \exp[-\frac{1}{2}(y - Xh)^T C_y^{-1} (y - Xh)] + \hat{\epsilon}^2 h^T M_p h$. The filter coefficient h is obtained by maximizing the posterior probability of h , which is given by,

$$\hat{h} = [X^T C_y^{-1} X + \hat{\epsilon}^2 M_p]^{-1} X^T C_y^{-1} Y \quad (8)$$

Estimates of a PD and PI pulses considering both real and simulated signal is shown in Fig. [1].

In simulated data (as shown in first part of Fig. [1]), the estimated pulse (h) was found to be very close to the added pulse. This shows the power of the method in modeling the PD data. The estimated PD and PI pulse considering the real data is shown in second part of Fig. [1]. A PCA based binary classifier, one correspond to PD and other one corresponding to PI is employed to extract PD and reject PI. In this methodology,

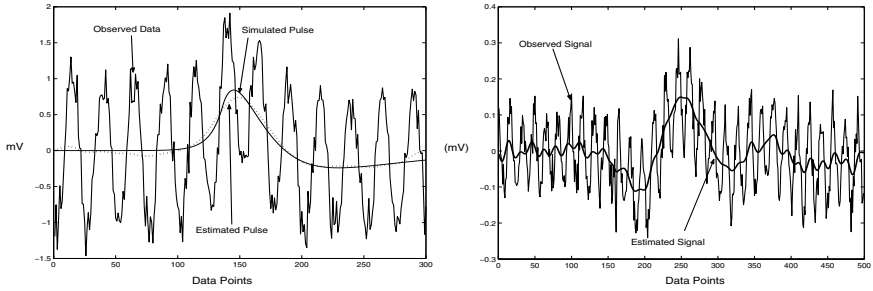


Fig. 1. Demonstration of pulse extraction using simulated and real data

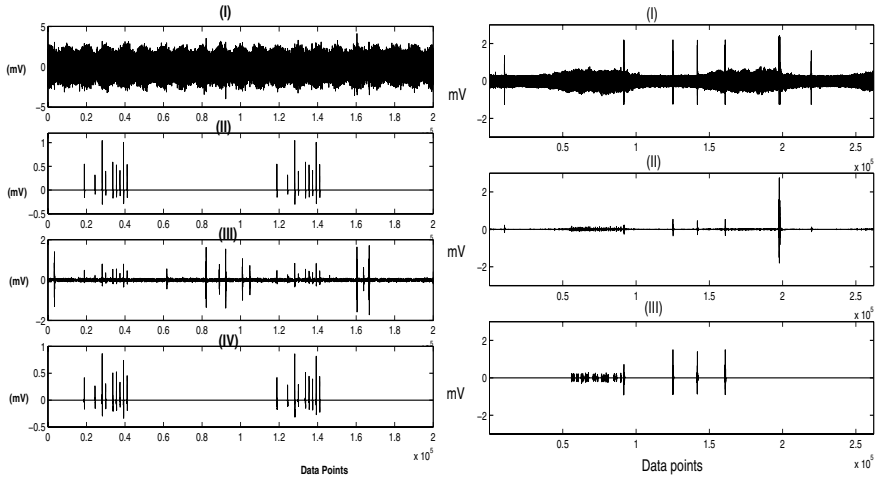


Fig. 2. Output of non-parametric method considering real and simulated data

subspaces are found for the training data of each class which are estimated from the using the non-parametric model as explained earlier. The overall output of the method on simulated data is shown in the first part of fig. [2]. The added PD pulses (we have this as we simulate the data), denoised signal and the retrieved signal are shown in (II), (III) and IV respectively. It can be seen that, all of the PD pulses have been retrieved rejecting all PI pulses. Also, the height of the retrieved pulses are close to that of added pulses, which is very important in making and diagnostic/prognostic decisions. The performance of the method in dealing with the real signal is shown in the second part of Fig. [2]. To the best of Authors knowledge, there is no general technique which can automatically retrieve the PD pulses rejecting PI. Therefore, we have not compared the performance of the method demonstrated in this paper with other techniques reported in the literature.

5 Conclusion

The problem of on-line recognition of PD signal is approached in a different perspective than conventional DSP techniques. A combination of PPCA and fBm process based

techniques have been developed and applied to the problem in hand. The methodology is found to be effective in PD signal modeling and recognition. Thus, Author has developed an statistical pattern recognition methodology for on-line, on-site measurements of PD signal.

References

1. Satish, L., Nazneen, B.: Wavelet denoising of PD signals buried in excessive noise and interference. IEEE Transaction on DEI. Vol. 10. No. 2. April 2003. pp 354–367.
2. Tipping, M.E., Bishop, C. M.: A hierarchical latent variable model for data visualization. IEEE trans. PAMI. Vol. 20. no-3. 1998. pp.25-35. 281-293.
3. Flandrin, P.: Wavelet analysis and synthesis of fractional Brownian motion. IEEE transaction on Information Theory. Vol. 38. no-2. 1992. pp.910-917.
4. Wornell, G: Signal Processing with Fractals: A Wavelet Based Approach. Prentice Hall PTR. Newjersy. 1996. Chap. 3. pp. 30-46.
5. Wornell, G. W.: A Karhunen-Loeve-like expansion for $1/f$ processes via wavelets. IEEE. Trans. Inform. Theory. vol. 36. July 1990. pp. 859-861.
6. Marrelec, G., Benali, H., Ciuciu, P., Poline, J: Bayesian estimation of haemodynamic response function in functional MRI, CP617, Bayesian Inference and Maximum Entropy methods in Science and Engineering: 21st International Workshop, 2002. pp. 229-247.
7. Bretthorst, G. L.,: Bayesian interpolation and deconvolution. The U.S Army Missile Command (1992). Tech. Rep. CR-RD-AS-92-4.