Electric Power Systems Research 118 (2015) 37–46

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

Comparison between Principal Component Analysis and Wavelet Transform 'Filtering Methods for Lightning Stroke Classification on Transmission Lines

John A. Morales [a,b,*], E. Orduña [a], C. Rehtanz [c], R.J. Cabral [d], A.S. Bretas [d]

[a] Instituto de Energía Eléctrica, Universidad Nacional de San Juan, Av. Lib. Gral. San Martín Oeste 1109, (J5400ARL), San Juan, Argentina
[b] Facultad de Ingeniería Eléctrica, Universidad Politécnica Salesiana, Calle vieja 12-30 y Elia Liut, 010150 Cuenca, Ecuador
[c] Institute of Energy Systems, Energy Efficiency and Energy Economics of TU Dortmund University, Emil-Figge-Straße 70, 44227 Dortmund, Germany
[d] Electrical Engineering Department, Federal University of Rio Grande do Sul, Av. Osvaldo Aranha, 103, Room 116, Porto Alegre, Rio Grande do Sul 90035-190, Brazil

ARTICLE INFO

Article history:
Received 7 February 2014
Received in revised form 23 May 2014
Accepted 27 May 2014
Available online 14 June 2014

Keywords:
Orthogonal functions
Lightning disturbances
Comparative evaluation
Filters

ABSTRACT

This paper presents an assessment between Principal Component Analysis (PCA) and Wavelet Transform (WT) signal processing techniques applied for Transmission Lines (TLs) lightning stroke classification. In this work, the atmospherics discharges signals are analyzed in two steps. The first step objective is patterns extraction, which is developed through Principal Component Analysis and the Wavelet Transform. The second step objective is pattern classification, which is developed using three different techniques: Artificial Neural Network (ANN), k-Nearest Neighbors (k-NN) and Support Vector Machine (SVM).

This work presents as assessment of lightning stroke classification, providing useful information, especially in extraction and selection of mother functions and the use of PCA. Both methodologies are assessed under different lightning stroke conditions. Features as extraction, speed, orthogonal functions and others are comparatively assess.

Results show that by using PCA, optimal mother functions can be extracted, presenting a new alternative for relaying protection.

1. Introduction

Electric Power Systems (SEP) are daily exposed to disturbances. Events as faults produced by trees, birds, weather and others affect the system continuity operation [1–3]. However, between all possible disturbances, lightning is the severe, affecting specially transmission line protection relays [4–10].

Lightning effects on system devices can be divided in two: first when their overvoltages do not exceed the Basic Insulator Level (BIL), a fault is not produced and therefore the protection relay must not send a trip order. Second, when the generated overvoltage exceeds the Basic Insulator Level, a fault is produced and the relay must send the trip signal [11]. In this context, it is clear that protection devices must correctly classify these phenomena, but currently, traditional relays performance under lightning conditions still has room for improvement [12–14]. Therefore, it is crucial for relaying performance, correct lightning stroke classification.

Analyzing lightning induced overvoltages, it becomes clear that there is a difference between lightning stroke signals that generate or not faults, especially in their high frequency transient waveforms. However, the adequate extraction of their features depends of the Signal Processing Technique applied. An essential requirement in protection relays is to accurately classify different patterns corresponding to lightning stroke. For this reason, a very effective feature extraction is the most important step to improve the accuracy of lightning patterns classification.

Based on the above stated, for many years wavelets transforms have been widely used for the lightning stroke classification, especially because of their performance and applications [13,17]. On the other side [18,19], presented a new approach for the lightning stroke classification. This proposed approach uses Principal Component Analysis to extract different patterns, which are useful for signal classification.

* Corresponding author at: Instituto de Energía Eléctrica, Universidad Nacional de San Juan</org>, Av. Lib. Gral. San Martín Oeste 1109, J5400ARL San Juan, Argentina. Tel.: +54 264 4226444; fax: +54 264 4210299.Universidad Politécnica Salesiana, Calle Vieja y Elia Liut, Cuenca, Ecuador.

E-mail addresses: moralesg@ups.edu.ec, jmorales@iee.unsj.edu.ar (J.A. Morales).

http://dx.doi.org/10.1016/j.epsr.2014.05.018
0378-7796/

Since the progress and development of new samplers, computational efficiency, new microprocessors technologies and others [15,16], new Signal Processing Techniques can be applied to develop more dependable protection algorithms. Thus, currently, several Signal Processing Techniques can be considered for new protection methodologies. However, in order to determine if a technique can be considered as an option in protection algorithms, it is necessary to assess this technique. In the work, a comparison between the Wavelet Transform and a new technique based on Principal Component Analysis is presented.

In this context, in previous authors publications, initial evaluation of the advantages and disadvantages of Principal Component Analysis and Wavelet Transform applied for the lightning stroke classification was presented [4]. Both Signal Processing Techniques were used for extract different features corresponding to atmospherics discharges, and an Artificial Neural Network (ANN) classifier was used for classify those signals.

Instead, in this paper a more detailed comparison by using other classification techniques as k-Nearest Neighbors (k-NN) and Support Vector Machine (SVM) is presented. On the other hand, in order to illustrate the potential of the methodology based on Principal Component Analysis, different Transmission Lines are analyzed.

In Ref. [4], different disadvantages and advantages useful for protection relays as methodology, percent of classification using ANN, pattern extraction, mother function and others were considered. However, in this research, not only these characteristics but also other crucial topics in protection relays are analyzed and discussed as follows:

- Viability using different patterns classification techniques (k-Nearest Neighbors, Support Vector Machine).
- Methodology for mother functions extraction.
- Optimization for selection of mother functions.

2. Application of Signal Processing Techniques for Lightning Stroke Classification

2.1. Wavelet Transform

By using the Wavelet Transform any signal can be analyzed as an infinite series of wavelets, expressing this signal as a linear combination of a set of function though of translations (τ) and dilations (s) of a mother wavelet [20]. The methodology for lightning stroke classification based on MRA is developed through the following processing:

- Data collection through simulation in ATP program.
- Processing through Multiresolution Analysis, the function is expressed as follows:

$$f(t) = \sum m\psi(t) \tag{1}$$

where m is an integer and ψ represents a set of functions, which is expressed as:

$$\psi_{\tau,s}(t) = \frac{1}{\sqrt{|s|}}\psi\left(\frac{t-\tau}{s}\right) \tag{2}$$

where τ and s are translation and scale, ψ is the transformation function called mother wavelet.

The function is represented as:

$$f(t) = \sum_{p} a_{qp}\varsigma_{qp}(t) + \sum_{q=q_0}\sum_{p} d_{p,q}\upsilon_{p,q}(t) \tag{3}$$

where, a_{qp} and d_{pq} are the approximation and detail coefficients, respectively.

From Eq. (3) it is possible to see that the function is a linear combination of wavelet approximation and detail coefficients a_{qp} and d_{pq}. These coefficients are calculated as follows:

$$a^j(n) = \sum_{i=0}^{L-1} l(i).a^{j-1}(2n-i), \quad 0 \le n < N_j \tag{4}$$

$$d^j(n) = \sum_{i=0}^{L-1} h(i).d^{j-1}(2n-i), \quad 0 \le n < N_j \tag{5}$$

where j represents the decomposition level, $l(i)$ and $h(i)$ correspond to low-frequency and high-frequency filters, respectively. They divide the function $f(s)$ in two parts, the first contains highest frequencies than $fs/2$, and the second contains lower frequencies than $fs/2$. This last part is used in order to obtain a second level, thus the process is continually repeated to obtain different levels. In this context, in this paper, five mother wavelets are used, developing all their decomposition levels.

- The patterns calculated with every mother wavelet are used as input to the classifiers. They are trained and verified for every pattern, and the best result is chosen. This process is repeated with each classifier (Artificial Neural Network, k-Nearest Neighbors and Support Vector Machine). A more detailed explication is presented in Ref. [17].

2.2. Principal Component Analysis

By employing PCA, lightning strokes can be represented as a linear combination of original variables. These new variables are extracted through their eigenvectors, which have orthogonal directions among them [21]. The methodology for lightning stroke classification based on PCA is developed through the following processing:

- Data collection through simulation in ATP program.
- Patterns extraction through PCA, which are based on the variance-covariance matrix as follows:

$$S = \frac{\sum_{i=1}^{n}(f_n - \bar{f})(f_n - \bar{f})}{n-1} \tag{6}$$

where f represents the vector corresponding to lightning stroke signals, which is a signal of 3000 points, and $\bar{f}$ is the mean vector.

In order to extract patterns, their eigenvectors must be calculated. They are selected based on the value of their eigenvalues. For example, the first eigenvector correspond to the eigenvalue, which has the higher variance percentage. The second eigenvector correspond to the eigenvector with the second higher variance percentage. The matrix corresponding to eigenvectors and eigenvalues are presented as follows:

$$U = [eigvector_1 \quad eigvector_2 \quad eigvector_3 \quad eigvector_4 \ldots eigvector_p] \tag{7}$$

$$V = \begin{bmatrix} \lambda_1 & 0 & . & . & 0 \\ 0 & \lambda_2 & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & 0 & 0 & \lambda_p \end{bmatrix} \tag{8}$$

where, λ represents the eigenvalues of the variance-covariance matrix. Finally, the new patterns are represented as vectors projected on the new base corresponding to principal components.