

Current Transformer Saturation Detection Using Gaussian Mixture Models

M. Moghimi Haji, B. Vahidi*, S. H. Hosseini

Department of Electrical Engineering
Amirkabir University of Technology
Tehran, Iran
*vahidi@aut.ac.ir

ABSTRACT

This paper presents a novel current transformer (CT) saturation detection approach based on Gaussian Mixture Models (GMMs). High accuracy is the advantage of this method. GMMs are trained with secondary current of CT. The appropriate performance of the proposed method is tested by simulation of different fault conditions in PSCAD/EMTDC software. The results show that the trained GMMs can successfully detect CT saturation with high accuracy.

Keywords: CT saturation, GMM, protective relaying, transient analysis.

1. Introduction

Current transformer (CT) saturation distorts the secondary current and in consequence leads to operating delay or malfunction of protection relays (e. g. under reaching of over current relays, overestimation of fault loop impedance in distance relays) [1]. Therefore an appropriate saturation detection method is necessary to maintain the protection system reliability.

A method for detecting CT saturation onset based on the abrupt change in the current when CT saturates is suggested in [2]. This method can detect the saturation successfully only if the current collapses to zero after inception of saturation; however, it may operate incorrectly when the current does not change instantly when an anti-aliasing low-pass filter is used.

In [3], an algorithm for calculating the core flux from the secondary current in order to compensate the saturation is proposed. This algorithm calculates the core flux and detects saturation based on given CT parameters.

Another approach proposed in [4] is based on evaluating the mean of the error and the mean and variance of the current magnitude. The error is derived on the following assumption: If the current

is a perfect sinusoid, the summation of the current and its second-order derivative is zero over time.

In [5] an impedance-based CT saturation detection algorithm for bus-bar differential protection is suggested. Calculation of power system source impedance at the relay position is based on a first-order differential equation. This method uses voltage and current signals of the bus-bar to detect CT saturation.

Another CT saturation detection algorithm based on the third difference of the secondary current is proposed in [6]. The effects of remanence flux in the core and a low-pass filter on the saturation are included in this method.

A method based on symmetrical components is suggested in [7]. The zero-sequence differential current gradient with respect to a bias current is utilized to detect saturation in a numerical current differential feeder protection relay.

Another approach for detecting CT saturation is the use of artificial intelligence (AI) techniques such as artificial neural networks (ANNs) (e.g., [8]). In [9], it is proved that considerable improvement of the operation and quite simple achievement of adaptive features of protection function may be

obtained by utilizing various AI techniques [1]. Neural computing methodologies have some advantages over conventional methods. However there are no general rules for choosing the type of ANN structure and its further parameters (such as number of layers and neurons, neuron activation function, and input signals) which depend on designer experiences with ANN usage. To overcome this drawback, in [1] an optimization approach based on the genetic algorithm is proposed.

GMMs have been used as classifiers in a lot of applications such as speaker identification (SI) [10], multiple limb motion classification [11], image processing and classification [12,13], control engineering [14], simplification of controller design [15], and rotating machinery fault diagnosis [16]. In [17] wavelet based GMMs are used for magnetizing inrush current Identification. Recently, a time-domain analyzer based on GMMs has been introduced to discriminate permanent and transient faults [18].

The ability of GMMs to solve nonlinear and high dimensional pattern problems makes them an appropriate choice in power system transient classifications [19]. In this paper, a new approach for CT saturation detection using GMMs is presented. Two GMMs are trained with the secondary current of CT. One of them is trained with CT secondary current when the fault inception angle is between 0–90 degrees and the other one is trained when the fault inception angle is between 90-180 degrees. This method of training GMMs for different fault inception angle improves the accuracy of the algorithm.

2. Gaussian Mixture Models

The mathematical formulation of GMM is presented in this section [17].

2.1 Definition

Mixture models are types of probabilistic density models that comprise a number of component functions. These component functions, which are usually, Gaussian are combined to provide a multimodal density. These models are semi-parametric alternatives to nonparametric histograms (which can also be used as densities)

and provide better flexibility and precision in modeling the underlying statistics of sample data. For an S-class pattern recognition system, a set of GMMs $\{\mathfrak{I}_1, \mathfrak{I}_2, \dots, \mathfrak{I}_S\}$ are associated with S classes. A random variable x with D-dimensions is said to follow a Gaussian mixture model, when its probability density function can be formulated by Equation 1 following the constraints presented in Equation 2.

$$P(x|\mathfrak{I}) = \sum_{k=1}^m \alpha_k P(x|\theta_k) \quad (1)$$

$$\sum_{k=1}^m \alpha_k = 1, \quad \alpha_k > 0 \quad (2)$$

where α_k is the mixture weight for the k th component, $\mathfrak{I} = \{\alpha_1, \dots, \alpha_m, \mu_1, \dots, \mu_m, \Sigma_1, \dots, \Sigma_m\}$ is the complete set of parameters to define the model, and $\theta_k = \{\mu_k, \Sigma_k\}$ is the mean and covariance of the k th component, respectively. The Gaussian mixture density is a weighted linear combination of m component Gaussian density functions, $\theta_1, \theta_2, \dots, \theta_m$. Each component density is a D-varient Gaussian function parameterized by a $D \times 1$ mean vector and a $D \times D$ covariance matrix. The component's density $P(x|\theta_k)$ is a normal probability distribution defined by Equation 3:

$$P(x|\theta_k) = \frac{1}{(2\pi)^{D/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(x-\mu_k)\Sigma_k^{-1}(x-\mu_k)^T} \quad (3)$$

2.2 Training Process

The goal of model training in a GMM-based classification system is to estimate the parameters of the GMM, \mathfrak{I}_g so that the Gaussian mixture density can best match the distribution of the training feature vectors. For a set of n independent and identically distributed vectors $X = \{x_1, x_2, \dots, x_n\}$, the corresponding likelihood of a mixture is

$$P(X|\mathfrak{I}) = \prod_{i=1}^n P(x_i|\mathfrak{I}) \quad (4)$$

where P is the likelihood of the data X given the distribution parameters of \mathfrak{I} . Specifically, by having

the distribution parameters in \mathfrak{Z} , the goal is to find $\hat{\mathfrak{Z}}$ that maximizes the following likelihood:

$$\hat{\mathfrak{Z}} = \arg \max_{\mathfrak{Z}} P(X | \mathfrak{Z}) \quad (5)$$

Usually, direct maximization of this function is difficult. The logarithm of the above probability, which is called the log-likelihood function given by Equation 6, is easier to calculate.

$$\log P(X | \mathfrak{Z}) = \sum_{i=1}^n \log \sum_{k=1}^m \alpha_k P(x_i | \theta_k) \quad (6)$$

The logarithm of maximum likelihood estimation of each model \mathfrak{Z} cannot be solved analytically. The expectation maximization (EM) algorithm presented in [19] is widely used to estimate the parameters of GMM. EM is an iterative algorithm which maximizes the likelihood probability generated by each GMM, $P(X | \mathfrak{Z}_g)$, given the data for that class. The EM algorithm has two main steps:

2.2.1 E-step

In this step, the posterior probability of sample x_i in the t^{th} step is calculated using the following equation:

$$P(l | x_i, \mathfrak{Z}^{(t)}) = \frac{\alpha_l^t P(x_i | \theta_l^{(t)})}{\sum_{k=1}^m \alpha_k^t P(x_i | \theta_k^{(t)})} \quad (7)$$

2.2.2 M-step

This re-estimation process will continue by replacing

$$\hat{\mathfrak{Z}}^{(t+1)} = \left\{ \alpha_1^{(t+1)}, \dots, \alpha_m^{(t+1)}, \mu_1^{(t+1)}, \dots, \mu_m^{(t+1)}, \Sigma_1^{(t+1)}, \dots, \Sigma_m^{(t+1)} \right\}$$

instead of $\mathfrak{Z}^{(t)} = \left\{ \alpha_1^{(t)}, \dots, \alpha_m^{(t)}, \mu_1^{(t)}, \dots, \mu_m^{(t)}, \Sigma_1^{(t)}, \dots, \Sigma_m^{(t)} \right\}$, by using the following equations:

$$\alpha_k^{(t+1)} = \frac{1}{n} \sum_{i=1}^n P(k | x_i, \mathfrak{Z}^{(t)}) \quad (8)$$

$$\mu_k^{(t+1)} = \frac{\sum_{i=1}^n x_i P(k | x_i, \mathfrak{Z}^{(t)})}{\sum_{i=1}^n P(k | x_i, \mathfrak{Z}^{(t)})} \quad (9)$$

$$\Sigma_k^{t+1} = \frac{\sum_{i=1}^n P(k | x_i, \mathfrak{Z}^{(t)}) \left[(x_i - \mu_m^{(t+1)})(x_i - \mu_m^{(t+1)})^T \right]}{\sum_{i=1}^n P(k | x_i, \mathfrak{Z}^{(t)})} \quad (10)$$

The above-mentioned method is discussed in more details in [20]. It should be noted that it is preferred to assume the S GMMs independent in calculations. In the case of independent classes, the estimation problem of S class pdfs can be divided into S separate estimation problems.

3. Implementation Details

3.1 CT Model

A 132 kV power system including a CT is simulated to study CT saturation. The modeled CT is based on the Jiles-Atherton theory of ferro magnetic hysteresis [21]. In this model, the effects of saturation, hysteresis, remanence, and minor loop formation are modeled on the basis of the physics of the magnetic material [22]. In [23], the accuracy of this model has been checked. In [24], ANN is used to enhance the accuracy of this model. This model has been previously used to study the CT saturation compensation in [25].

In this model, the total magnetization (M) has been split in to two components [26]:

$$M = M_{rev} + M_{irr} \quad (11)$$

where M_{rev} and M_{irr} are reversible and irreversible magnetization components, respectively. M_{irr} is described as follows:

$$\frac{dM_{irr}}{dH} = \frac{M_{an} - M_{irr}}{k\delta - \alpha(M_{an} - M_{irr})} \quad (12)$$

where δ is a directional parameter to distinguish between the ascending and the descending part of

the hysteresis loop and M_{an} is the unhysteretic magnetization provided by the Langevin function.

$$\delta = \text{sign}(dH / dt) \quad (13)$$

$$M_{an} = M_s \left[\coth\left(\frac{H + \alpha M}{a}\right) - \frac{a}{H + \alpha M} \right] \quad (14)$$

Finally, M_{rev} is given by

$$M_{rev} = c(M_{an} - M_{irr}) \quad (15)$$

Among the five parameters of the JA model, M_s and a are the parameters of the unhysteretic curve which determine the saturation behavior of the magnetic core according to the Langevin function. The simulated CT in the present study is similar to the model used in [25], with the exception of the rated voltage, which is 132 KV in the present study. The model of the used CT is shown in Figure 1. Also, the parameters of the simulated CT are given in Table 1.

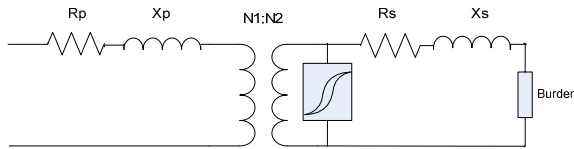


Figure 1. CT model.

Rated voltage	132 kV
Turn ratio	2000:1
Area	6.5×10 ⁻⁴ m ²
Path length	0.5 m
Frequency	50 Hz
Secondary winding resistance and inductance	Rs=0.5 Ω Ls=8×10 ⁻⁴ H

Table 1. Parameters of simulated CT.

3.2 Training Patterns

The training data set of a GMM should contain the information needed to generalize the problem. For this purpose, different fault types with different fault conditions are modeled using an EMTDC electromagnetic transient program [22]. A simple 132 kV power system is considered and several simulations for generating training patterns are done by changing fault current, fault type, fault inception angle, and burden. A combination of different fault conditions is shown in Table 2.

The saturation of CT is confirmed when the relative difference between the primary current i_1' (related to the secondary) and the secondary current i_2 is higher than 10% of the primary current [1]. Therefore, the time dependant training function for detecting the saturation is defined as follows:

$$\text{Sat}(n) = \begin{cases} 1 & \text{if } \left| \frac{i_1'(n) - i_2(n)}{i_1'(n)} \right| > 0.1 \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

fault current	Different values up to 33 kA
fault type	Single phase, phase to phase, three phase
fault inception angle	Different values between (0,180)
burden	5Ω< Zload <20Ω 0.5<cosφ<1

Table 2. Data for generating training patterns.

3.3 GMM Classification

In this stage, the objective is to find the model which has the maximum probability P_{\max} for a given observation sequence. The test patterns are provided for each of the GMM models ($P_{\mathfrak{S}_1}, P_{\mathfrak{S}_2}, \dots, P_{\mathfrak{S}_S}$), and S probabilities are computed and compared, i.e. P_i , $i = 1, 2, \dots, S$. Class f associated with the GMM that has the

highest Gaussian mixture likelihood probability P_{\max} is chosen as the predicted class for the given feature set. A feature vector x produced by one type of signals (saturated or unsaturated secondary current of CT) is assumed. The identity of the signal is determined by finding signal model \mathfrak{T}_f from S models which maximizes the probabilities across the signal set $\mathfrak{T}_g : g \in (1, S)$:

$$\mathfrak{T}_f = \arg \max_{1 \leq g \leq S} P(\mathfrak{T}_g | x) \quad (17)$$

By using the Bayes rule, Equation 17 can be expressed as

$$\mathfrak{T}_f = \arg \max_{1 \leq g \leq S} \frac{P(x | \mathfrak{T}_g)P(\mathfrak{T}_g)}{P(x)} \quad (18)$$

By assuming equal probabilities for all models and noting that $P(x)$ is the same for them, the identification task can be summarized as finding the logarithm of the following equation:

$$\mathfrak{T}_f = \arg \max_{1 \leq g \leq S} P(x | \mathfrak{T}_g) \quad (19)$$

Equation 19 can be expressed as the following equation for simplicity.

$$\mathfrak{T}_f = \arg \max_{1 \leq g \leq S} [\log \sum_{k=1}^m \alpha_{kg} P(x | \theta_{kg})] \quad (20)$$

where α_{kg} and θ_{kg} are mixture weight, mean, and covariance of the g th signal model, respectively.

3.4 GMM Parameters

In order to achieve better performance of the Gaussian mixture model, an effort to investigate an optimum configuration of algorithmic issues must be considered [11]. These issues include convergence threshold, model order selection, and the form of the covariance matrix.

Convergence threshold of the EM Algorithm: The convergence threshold of the EM algorithm is defined as the difference between probabilities $P(X | \mathfrak{T}_g)$ between two consecutive iterations. The convergence threshold and the maximum iteration

number are two conditions for stopping the EM algorithm. Different values of the convergence threshold were carefully examined and a value of 0.0005 was found to be the best for achieving good performance.

GMM Model Order: The GMM order selection is the most crucial factor in the performance of this classifier [11]. The objective is to choose the number of mixture components that yields the highest discriminative accuracy for the test data set. Theoretically, too few mixture components can yield a GMM which does not accurately model the distinguishing characteristics of classes. However, too many components can also reduce the performance, while increase the computation burden. It also makes the classification more complex. This is especially important for situations with smaller amounts of training data [11]. A larger model order should be considered when larger amounts of training data are available. Therefore, the optimal number of mixture components for the GMM must be carefully examined to achieve high classification performance. The optimal number of mixture is achieved by computing the classification rate presented in Equation 21 for test data set and varying the number of mixtures.

$$\text{Classification rate} = \frac{\text{Correct classified transient events}}{\text{Total transient events}} \quad (21)$$

Form of the covariance matrix: Another effective issue on the performance of the classifier is the form of the covariance matrix. The form of the covariance matrix may be full or diagonal. In this work, diagonal covariance matrices are selected. This choice is based on empirical evidence that shows that diagonal matrices outperform full matrices. The diagonal matrix GMMs are more computationally efficient than full covariance GMMs for training since repeated inversions of a $D \times D$ matrix are not required [11].

In the saturation detection problem, two GMMs should be trained. One for saturated and the other for unsaturated secondary current signals. The process of finding optimal GMMs is shown in the flowchart in Figure 2. After these two GMMs are found, the identification process for a new signal is as presented in the flowchart in Figure 3.

4. Simulation Results

In this section, the training and testing process of GMMs for CT saturation detection is presented. As said in the previous section, a few hundred simulations have been performed. 70% of them are used for training and the remaining 30% are used for testing. Fault inception angles within $[0,180]$ are considered in this study and two GMMs are trained. One of them for detecting the saturation when fault inception angle is between 0-90 degrees, and the other one for fault inception angles between 90-180 degrees. Fault inception angles between 180-360 degrees are not considered because of similarity. When a fault occurs, a switch can easily select one of the GMMs according to the fault inception angle. Training multiple GMMs for different fault inception angle ranges improves the accuracy of the method. Moreover, it does not need additional equipment or additional computation time. The primary and secondary currents of the used CT for fault inception angles equal to 45 and 135 degrees are illustrated in Figure 4(a) and Figure 4(b), respectively.

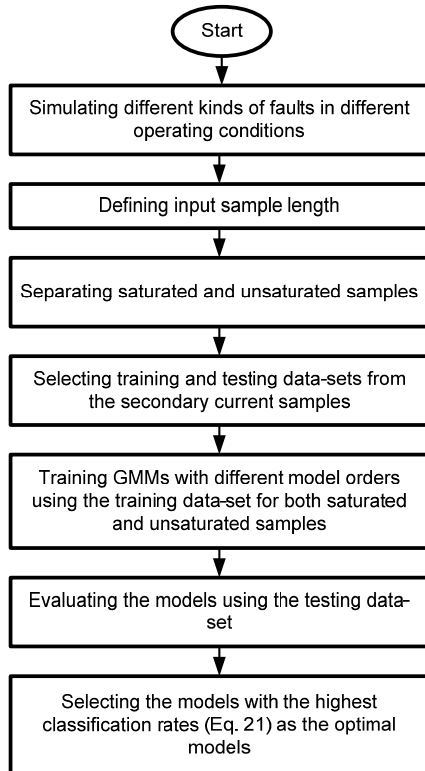


Figure 2. GMM training process.

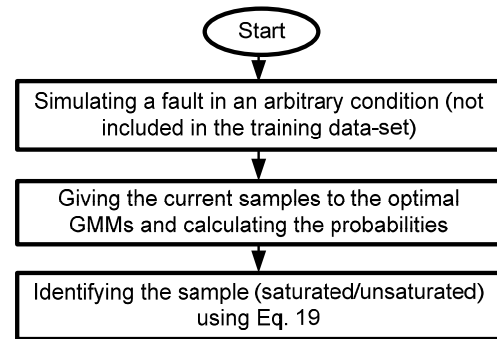


Figure 3. Signal identification process.

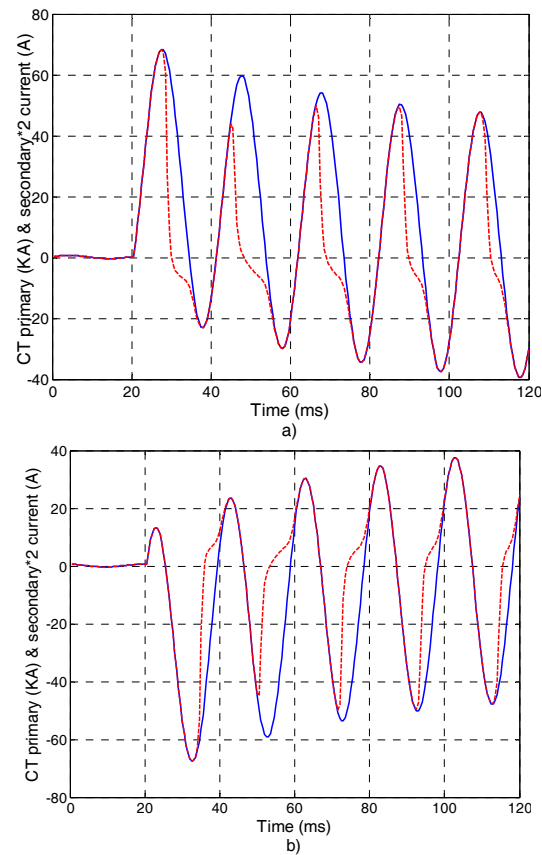


Figure 4. Primary and secondary currents of CT for fault inception angles equal to a) 45 degrees and b) 135 degrees.

As can be seen in Figure 4, the wave forms of the CT for these cases have a great difference; therefore, training only one GMM for all of inception angle deteriorates the accuracy of the method.

The implementation results are presented in Table 3 for various numbers of input samples. To show the advantage of using two GMMs, the results of using only one GMM for all inception angles within $[0, 180]$ is also presented in this table.

To show the ability of the proposed method to detect saturation in new situations (which were not included in the prepared data set), the trained GMM responses for two new cases are illustrated in Figures 5 and 6. Additionally, two examples of comparison between the developed GMM-based saturation detector and a chosen deterministic CT saturation identification scheme based on the calculation of the third derivative of the CT secondary current [27] are shown in these figures.

Input sample length	Classification rate (%)		
	GMM1 $[0-90^\circ]$	GMM2 $(90-180^\circ]$	GMM $[0-180^\circ]$
4	88.86	92.72	83.06
7	93.45	95.99	86.83
10	94.95	94.50	89.40
13	96.16	95.15	90.76
16	94.96	94.59	92.63

Table 3. Classification rates of the trained GMMs for various numbers of input samples.

Figures 5 and 6 show that by using GMMs all starting and ending points of CT saturation intervals are properly detected. As can be seen in these figures, the third derivative shows high spikes at the beginning of the saturation, but there are no such spikes at the end of the saturation. Hence, this method detects the starting moments of the saturation properly but may have problems in detecting saturation endings.

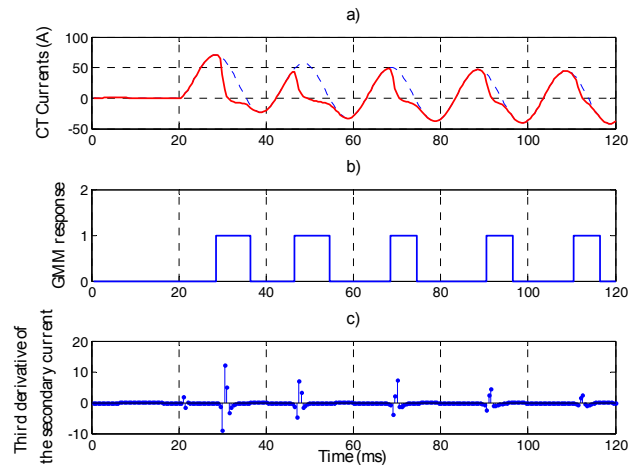


Figure 5. Comparison of CT saturation detection methods for fault inception angle equal to 27 degrees, a) CT signals, b) GMM1 response, c) third derivative of the secondary current.

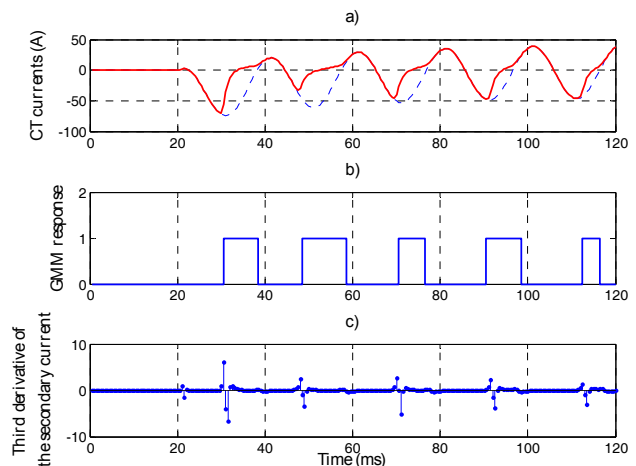


Figure 6. Comparison of CT saturation detection methods for fault inception angle equal to 162 degrees, a) CT signals, b) GMM1 response, c) third derivative of the secondary current.

5. Conclusion

A new approach for detecting CT saturation is presented in this paper which is based on a classification method called GMM. Secondary current samples are separated in two classes as saturated and unsaturated. GMMs are trained with the secondary current of CT in different conditions. To improve the accuracy of the method, two different GMMs are trained for fault inception angles between 0-90 degrees, and 90-180 degrees. Simulations are performed to confirm the effectiveness of the proposed GMM-based saturation detector. The results show that the proposed method can successfully detect the starting and ending points of saturation.

References

- [1] W. Rebizant and D. Bejمرت, "Current-Transformer Saturation Detection With Genetically Optimized Neural Networks," *IEEE Trans. Power Delivery*, vol. 22, no. 2, pp. 820-827, 2007.
- [2] A. G. Phadke and J. S. Thorp, *Computer Relaying for Power Systems*: Research Studies Press Ltd., 1988, pp. 185-186.
- [3] Y. C. Kang et al., "An algorithm for compensating the secondary current of current transformers," *IEEE Trans. Power Delivery*, vol. 12, pp. 116-124, 1997.
- [4] T. Bunyagul et al., "Overcurrent protection using signals derived from saturated measurement CTs," in *Proc. IEEE Power Eng. Soc. Summer Meeting*, Vancouver, BC, Canada, 2001.
- [5] C. Fernandez, "An impedance-based CT saturation detection algorithm for bus-bar differential protection," *IEEE Trans. Power Delivery*, vol. 16, pp. 468-472, 2001.
- [6] Y. C. Kang et al., "A CT saturation detection algorithm," *IEEE Trans. Power Delivery*, vol. 19, no. 1, pp. 78-85, 2004.
- [7] N. Villamagna and P. A. Crossley, "A CT saturation algorithm using symmetrical components for current differential protection," *IEEE Trans. Power Delivery*, vol. 21, no. 1, pp. 38-45, 2006.
- [8] M. M. Saha et al., "Application of ANN methods for instrument transformer correction in transmission line protection," in *Proc. 7th Int. Conf. Developments in Power System Protection*, Amsterdam, The Netherlands, 2001, pp. 303-306.
- [9] D. Neibur, "Artificial neural networks for power systems," Report by Task Force 38.06.06, Electra vol. 159, 1995.
- [10] B. Xiang and T. Berger, "Efficient Text-Independent Speaker Verification with Structural Gaussian Mixture Models and Neural Network," *IEEE Trans. Speech and Audio Processing*, vol. 11, pp. 447 - 456, 2003.
- [11] Y. Huang et al., "A Gaussian Mixture Model Based Classification Scheme for Myoelectric Control of Powered Upper Limb Prostheses," *IEEE Trans. Biomedical Engineering*, vol. 52, pp. 1801 - 1811, 2005.
- [12] X. Zhou and X. Wang, "Optimization of Gaussian mixture model for satellite image classification," *IEE Proc.-Vis. Image Signal Process*, vol. 153, pp. 349 - 356, 2006.
- [13] M. Hamouz et al., "Feature-based affine-invariant detection and localization of faces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, pp. 1490-1495, 2005.
- [14] J. C. Principe and M. Motter, "Identification and Control of Aircrafts using Multiple Models and Adaptive Critics," NASA Proposal No. 04-0617, 2007.
- [15] A. Ruano, "Intelligent Control Systems Using Computational Intelligence Techniques," *IEE Control Series*, 2005.
- [16] T. Marwala et al., "Hidden Markov Models and Gaussian Mixture Models for Bearing Fault Detection Using Fractals," *International Joint Conference on Neural Networks*, 2006, pp. 3237-3242.
- [17] S. Jazebi et al., "Magnetizing Inrush Current Identification Using Wavelet Based Gaussian Mixture Models," *Simulation Modeling Practice and Theory*, vol. 17, pp. 991-1010, 2009.
- [18] S. Jazebi et al., "Time domain single-phase reclosure scheme for transmission lines based on dual-Gaussian mixture models," *Engineering Applications of Artificial Intelligence*, In Press, 2012.
- [19] G. Xiong et al., "Dynamical Gaussian mixture model for tracking elliptical living objects," *Pattern Recognition Letters* 27, pp. 838-842, 2006.
- [20] P. Paalanen et al., "Feature representation and discrimination based on Gaussian mixture model probability densities—Practices and algorithms," *Pattern Recognition*, vol. 39, pp. 1346 - 1358, 2006.
- [21] D. C. Jiles and D. L. Atherton, "Theory of ferromagnetic hysteresis," *Journal of Magnetism and Magnetic Materials*, vol. 61, pp. 48-60, 1986.

[22] PSCAD/EMTDC User's Manual, Manitoba HVDC Research Center, Winnipeg, MB, Canada, 1988.

[23] U. D. Annakkage et al., "A Current transformer model based on the Jiles-Atherton theory of ferromagnetic hysteresis," *IEEE Trans. Power Del.*, vol. 15, no. 1, pp. 57–61, 2000.

[24] R. A. Naghizadeh et al., "Modeling of Inrush Current in Transformers Using Inverse Jiles-Atherton Hysteresis Model with a Neuro-Shuffled Frog-Leaping Algorithm Approach," accepted to *IET Electric Power Application*.

[25] H. Khorashadi-Zadeh and M. Sanaye-Pasand, "Correction of Saturated Current Transformers Secondary Current Using ANNs," *IEEE Trans. Power Delivery*, vol. 21, no. 1, pp. 73-79, 2006.

[26] R. A. Naghizadeh et al., "Parameter Identification of Jiles-Atherton Model using SFLA," *International Journal for Computation and Mathematics in Electrical and Electronic Engineering (COMPEL)*, vol. 31, no. 4, pp.1293 – 1309, 2012.

[27] Y. Kang et al., "An algorithm for detecting CT saturation using the secondary current third-derivative function," in *Proc. IEEE Bologna PowerTech Conf.*, 2003, pp. 320–326.