Fuzzy Integral Based Multi-Sensor Fusion for Arc Detection in the Pantograph-Catenary System

Ilhan Aydin[1*], Selahattin B Celebi[1], Sami Barmada[2] and Mauro Tucci[2]
[1] Department of Computer Engineering, University of Firat, Turkey
[2] DESTEC, University of Pisa, Italy

Abstract

The pantograph - catenary subsystem is a fundamental component of a railway train since it provides the traction electrical power. A bad operating condition or, even worse, a failure can disrupt the railway traffic creating economic damages and, in some cases, serious accidents. Therefore, the correct operation of such sub-systems should be ensured in order to have an economically efficient, reliable and safe transportation system.

In this study, a new arc detection method was proposed and is based on features from the current and voltage signals collected by the pantograph. A tool named mathematical morphology is applied to voltage and current signals to emphasize the effect of the arc, before applying the Fast Fourier Transform (FFT) to obtain the power spectrum.

Afterwards, three support vector machines based classifiers are trained separately to detect the arcs, and a fuzzy integral technique is used to synthetize the results obtained by the individual classifiers, therefore implementing a classifier fusion technique.

The experimental results show that the proposed approach is effective for the detection of arcs, and the fusion of classifier has a higher detection accuracy than any individual classifier.

Corresponding author: Department of Computer Engineering, University of Firat, 23119, Elazig, Turkey.
E-mail: iaydin@firat.edu.tr

Keywords

Pantograph-catenary system, mathematical morphology, Fourier transform, support vector machines, fuzzy integral.

1. Introduction

Reliability and safety of railway systems have always been an important issue, and the research dedicated to their improvement is constantly increasing also due to the worldwide development of high speed trains. Electric trains obviously require a great amount of power which is often provided by the pantograph – catenary sub-system. Since power collection (through voltage and current) is ensured by a sliding device a continuous contact between the pantograph and the overhead contact wire must been maintained. When the contact is not perfect (due to defects on the overhead line or on the contact strip) or completely absent an electric arc is generated which can further degrade the components (depending on its duration and magnitude) [1], [2]. A common procedure is to periodically replace the contact strip as a rudimental form of preventive maintenance, but the replacement is often operated before its useful life was completed. Another procedure often adopted is to periodically inspect the overhead line by using a vehicle; however, this method implies a reduction of the traffic load since a path is kept busy by the vehicle itself.

To this aim different signal and image processing based methods were utilized for contactless monitoring of the pantograph – catenary sub-system. In image processing

based research efforts, most utilized technique employs the edge detection and Hough transform (applied to thermal images) to locate the position of the pantograph. The position of the contact wire is investigated by using thermal images [3]. Despite its advantages (thermal imaging is not affected by environmental conditions such as trees, illumination, and resolution) this technique can be both expensive and difficult to implement in real-time because of the amount of data do be processed.

Many other researchers have focused their attention in developing efficient algorithms for the detection of faults in the pantograph – catenary subsystem by using different image processing techniques (see for instance [4] – [12]): the results are in general promising but limited by the need of a recording camera set on top of the train and a big amount of data to be processed.

Other method to detect anomalies are based on vibration measurements [13]: signal processing of data coming from vibration sensors mounted on the railway infrastructure can detect macroscopic defects.

A less expensive device, if compared to thermal cameras, is a photosensitive device (photo diode), giving as output a continuous signal which is related to the presence of an electric arc (when the signal is higher than a predefined threshold): [14] and [15] show its application with the aim of detecting defects in the catenary.

Since the main goal would be to detect anomalies and contact strip wear with no additional equipment mounted on the train, [14] - [16] are dedicated to the analysis of

current characteristics (of DC trains) with the aim of detecting the presence of arcs using the phototube signal only as a validation data set. While the technique proposed in [16] works well for DC currents, good results could not be obtained in case of AC trains (most modern high speed trains). For this reason, in [17] a procedure based on Support Vector Machines (SVM) shows promising result identifying a high percentage of electric arcs from voltage and current measurements (when phototube signals are only used as validation).

In this study the authors propose a new approach which is capable of detecting electric arcs in high speed train with a higher accuracy if compared to [17]. The proposed method uses the Fourier transform and mathematical morphology for extraction of features from current and voltage signals. In particular, mathematical morphology is applied to each signal (voltage and current) and the spectrum of new signal is constructed by applying Fourier transform. The features extracted from Fourier spectrum of new signal are characterized by a higher information content (from the arc detection point of view) than those extracted from original signal. The obtained features for each signal are given to individual classifier and the results of the classifiers are combined by using the fuzzy integral method. There are two main contributions in this paper: the first one is a new feature extraction method while the second is the fusion of the classifiers by using fuzzy integral.

The rest of the paper is organized as follows: section 2 gives a description of the data available from direct measurements; section 3 introduces the data preprocessing and feature extraction by the use of the so-called mathematical morphology technique; the detailed explanation of the SVM classifier is given in section 4. In section 5 the theoretical background of fuzzy integral is given, and the results are shown in section 6.

2. Description of the data recorded on test runs

The data available for the analysis are relative to 6 test runs of an a.c. high speed train, operated on regular passengers railway tracks. The train is equipped with voltage and current recording instruments (usually always present on high speed trains); in addition, two phototube sensors are located at the front and the back of the pantograph and are used to validate the occurrence of an arc (the data used in this contribution are the same ones used for paper [23]). The quantities are sampled at 20 kHz and for each test run the data available are: voltage, current, train velocity, phototube output.

In each test run the above described quantities were recorded for approximately 25 minutes, during a regular train ride. Figure 1 shows the typical velocity profile of a test run and the envelope of the current collected by the pantograph: it is evident that the train is accelerated at the beginning of the run and decelerated at the end of the run while the velocity is kept approximately constant for about 15 minutes. During this time different

values of the current are used to simulate different operating conditions.

It is important to underline that, due to the sampling rate, the amount of data is in the order of 35 x 10^6 samples, i.e. a considerably high number considering that the main goal is to find useful information at reasonably small computation times. This work mainly addresses the problem of arc detection. The problem of correlating the detected arcs to defects in the catenary or in the pantograph comes in second place, and it can be addressed by analyzing the signal of the detected arcs events. For example, if repeated arcs at specific positions are detected then the line is probably defective at that point, while a global degradation could be ascribed to a pantograph deterioration.

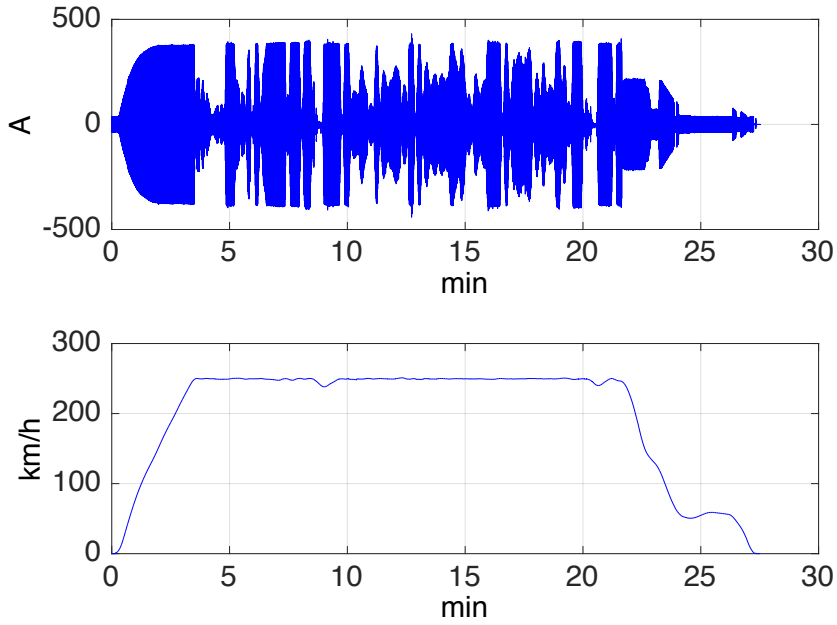

Figure 1: Velocity and current profile of a test run.

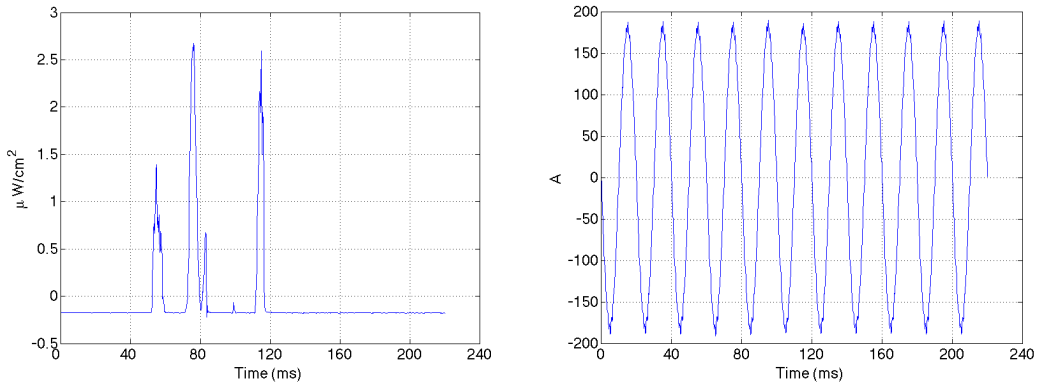

Figure 2: Phototube output and corresponding current

Figure 2 shows a portion of the phototube signal in presence of arcs, and the corresponding recorded current: a simple visual analysis of the current in Figure 2 would not lead to any conclusion.

3. Feature Extraction using Mathematical Morphology

Mathematical morphology

Mathematical morphology is a technique which was originally applied to binary images and then extended to grayscale images. This method is an essential tool for applications such as identifying and removing the skeletal and limit structure in an image, noise reduction, and segmentation [18]. Two morphological operators lie at the basis of this method: erosion and dilation, which are equivalent to the Minkowski set subtraction and addition, respectively. Roughly speaking, thinking about operating dilation to a binary image, the holes tend to be filled and the corners smoothed. On the contrary an image which has undergone an erosion procedure, is characterized by wider holes and thinner objects [19].

Mathematical morphology can be also applied to signals as nonlinear digital signal processing procedures: the information of a known signal is modified through the interaction with another function called structuring element.

The dilation of *f(n)* by *μE(n)* is defined as

$$f(n)\oplus\mu E(n) = \max_{m=-\mu,..,\mu}\{f(n+m)+\mu E(m)\} \quad (1)$$

where *f(n)* and *μE(n)* are respectively a discrete signal of length *N* and a symmetric function of length *M* representing the structuring element of size *μ*, with *M=2μ+1* (typically *N>>M*). The erosion operator corresponds to the dual form of the dilation operator. It is represented as

$$f(n)\ominus\mu E(n) = \min_{m=-\mu,..,\mu}\{f(n+m)-\mu E(m)\} \quad (2)$$

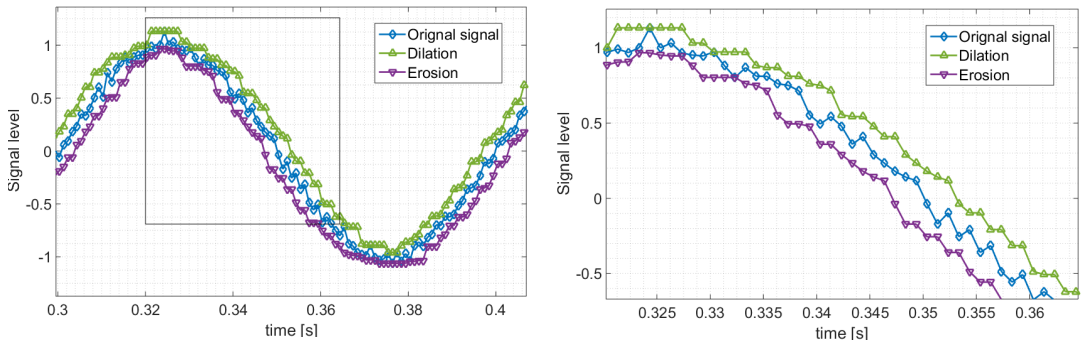

Figure 3. The illustration of the mathematical morphology on a sinusoidal signal. Left side: one period of the signal; right side: zoomed portion of the signal.

In this work, we use a flat structuring element of the form:

$$E(m)=\begin{cases}0, & -\mu \le m \le \mu \\ -\infty, & \text{otherwise}\end{cases}$$

Figure 3 represents an illustration of the two mathematical morphology operators, by using a structuring element of size $\mu=3$, on a synthetic 10 Hz sinusoidal signal with a superimposed white Gaussian noise. If the position of the structuring element is located at sample *i*, a window of size $M=2\mu+1=7$ is analyzed to extract the maximum value among the samples *i-3, i-2, i-1, i, i+1, i+2 i+3*; the dilated signal at sample *i* then assumes the above mentioned maximum. The window is then shifted one step (sample *i+1*) and the same process is iterated in order to calculate the dilated signal at sample *i+1*. The erosion operator works in a dual way as described in (2). In this way, dilation and erosion operators are used to eliminate or to emphasize some specific components in the signal. In particular, both dilation and erosion operate a low-pass nonlinear transformation of the signal, with a global amplification (dilation) or attenuation (erosion) effect on the low frequencies. Determining their exact effect in the frequency domain is not trivial, as they are nonlinear operators, but one can easily observe a low-pass effect from Fig. 3. In recent years, the mathematical morphology was applied to many areas such as power quality monitoring [20], fault diagnosis on power transformers [21] and broken rotor bar detection in induction motors [22].

Data pre-processing based on Fourier transform

Based on the experience of previous works by the authors, in which voltage and currents data are analyzed by a clustering procedure [23], Fourier Transform (FT) [24] is an

efficient pre-processing technique to extract the desired feature (presence or absence of electric arc) from a raw signal.

In [23] FT is used both to extract features and also to reduce the dimension of the problem, since time domain signals resulting from real time data acquisition are of high dimension, and an efficient use of the frequency spectrum (in this case only low frequencies are taken into account) can reduce complexity. However, the use of FT alone is not straightforward, and additional techniques (such as clustering in [23]) need to be used.

In this paper FT is applied after the original signal (voltage or current) is modified by the dilation operation. The so obtained spectrum is the input to the SVM and subsequent fusion technique.

Figure 2 represents the power spectrums of four different signals related to time windows with a duration of 80 ms: a raw time window with no electric arc, a raw time window with electric arc, and the above two signals modified by the dilation operator.

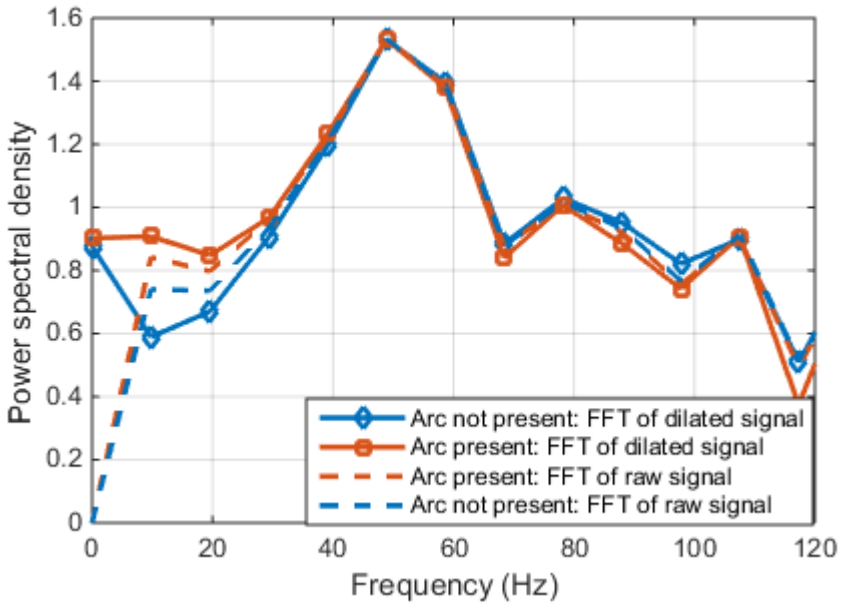

Figure 4. Power spectrums of time segments in presence and absence of electric arcs.

As shown in Figure 4, it is in general not easy to distinguish between the arc and no-arc condition even though differences can be noticed. However, the difference between the two conditions are enhanced in the frequency domain if the dilation operator is applied. Based on this observation the pre-processing of the raw signal is composed by dilation and FT.

4. Support Vector Machine

The support vector machine was proposed as a binary classification technique that uses the structural risk minimization [25]: it finds the best hyperplane which separates data in classes, maximizing the distance between them. For a generic multiclass problem in N dimensions, a data set is represented as $\{(\bar{x}_1, y_1), (\bar{x}_2, y_2), \dots, (\bar{x}_n, y_n)\}$, where the inputs of the classifier are given by the column vectors $\bar{x}_i \in \mathbb{R}^N$, and the targets $\bar{y}_i \in \mathbb{R}$

represent the class label associated to the N dimensional input. Figure 5 shows an example of a binary support vector classifier for linearly separable data in two dimensional space.

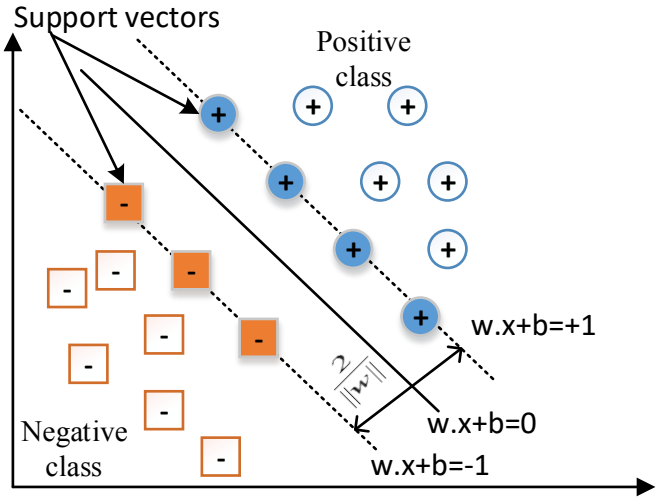

Figure 5. Support vector classifier for linearly separable data

The aim of the support vector machine is to find the maximum margin hyperplane, where the margin is defined as the distance between the separation hyperplane and the nearest training samples, which are called support vectors as shown in Figure 5. We describe the hyperplane as $\bar{w} \cdot \bar{x} = b$, where $\bar{w} \in \mathbb{R}^N$ is the normal vector, and $b \in \mathbb{R}$ is the bias term. It can be shown that the margin is given by $2/\|\bar{w}\|$. The SVM theory solves the problem of maximizing the margin by minimizing $\|w\|$. This leads to a quadratic optimization problem based on inequality constraints, given by (3).

$$\left\{\begin{array}{ll} Minimize & \frac{1}{2}\overline{w}^T\overline{w}+C\sum_{i=1}^{n}\varepsilon_i \\ Subject\ to & y_i(\overline{w}\cdot\phi(\overline{x}_i)+b)\geq 1-\varepsilon_i \end{array}\right\} \tag{3}$$

The function $\phi(\overline{x}_i)$ maps the input vectors to a higher dimensional space, with the aim to make the dataset linearly separable. In (3), the so called slack variables $\varepsilon_i \geq 0$, represent the misclassification error, and they are introduced to obtain a solution if the mapped data set is not linearly separable. The user can define the parameter C, which controls the tradeoff between margin maximization and slack variable minimization. The problem in (3) can be solved by using a Lagrange representation [26], which is given by

$$\left\{\begin{array}{ll} Maximize & \sum_{i=1}^{n}\alpha_i-\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j K(\overline{x}_i,\overline{x}_j) \\ Subject\ to & \sum_{i=1}^{n}\alpha_i y_i=0,\ 0\leq\alpha_i\leq C,\ i=1,...n \end{array}\right\} \tag{4}$$

In (4), α_i represents a positive Lagrange multiplier and it cannot be larger than C. $K(\overline{x}_i,\overline{x}_j)$is the kernel function and it represents the dot product in the higher dimensional space. The selection of the kernel function is very important for the accuracy of classification, and some common kernels include linear, Gaussian, and hyperbolic tangent. The Gaussian kernel is perhaps the most used for the classification of non-linear datasets, and it is defined as

$$K(\overline{x}_i,\overline{x}_j)=\exp\left(\left\|\overline{x}_i-\overline{x}_j\right\|/2\sigma^2\right) \tag{5}$$

After solving the quadratic optimization problem (4), and obtaining the Lagrange multipliers α_i, the following decision boundary is applied to a new sample $\bar{x}$ and its class is determined accordingly:

$$Class(\bar{x}) = \text{sgn}(\sum_{i=1}^{n} \alpha_i y_i K(\bar{x}_i, \bar{x}) + b) \quad (6)$$

5. A Fusion of Classifiers by Using Fuzzy Integral for Arc Detection

The Fuzzy integral is a nonlinear function that combines the results of multiple sources of information; it was first introduced by Sugeno [27-29] and the main characteristics are summarized in this paragraph. Let X be a finite set of elements, and 2^X indicates the family of all subsets of X.

A fuzzy measure is defined as a set function $g: 2^X \rightarrow [0,1]$ with the following properties

$$\begin{aligned} &g(0) = 0 \\ &g(X) = 1 \\ &g(A) \leq g(B) \text{ if } A \subset B \end{aligned} \quad (7)$$

A g_λ fuzzy measure, introduced by Sugeno, satisfies the following additional property

$$g(A \cup B) = g(A) + g(B) + \lambda g(A) g(B) \quad (8)$$

for $\lambda > -1$, $A, B \subset X$ and A $\cap$ B=0.

Starting by the above defined fuzzy measure, the definition of fuzzy integral is given, with $h : X \rightarrow [0,1]$ being a fuzzy subset of X:

$$h(x) \circ g(\cdot) = \max_{E \subseteq X} \left[\min \left(\min_{x \in E} h(x), g(E) \right) \right] = \max_{\alpha \in [0,1]} \left[\min \left(\alpha, g(h_\alpha) \right) \right] \tag{9}$$

where h_α is the α level set of h

$$h_\alpha = \{ x : h(x) \geq \alpha \} \tag{10}$$

As a matter of fact the fuzzy integral (a nonlinear functional defined with respect to a fuzzy measure) is defined over the support X of the function h with respect to a fuzzy measure g. Additional properties of the fuzzy integral can be found in [27-29].

Trying to give an insight on the practical application of the fuzzy integral, we can say the following: *h(x)* measures the degree to which h is satisfied by x, and $\min_{x \in E} h(x)$ is referred to the satisfaction relative to all the elements of E; on the other hand the value *g(E)* is a measure of the degree of satisfaction of the objects of E with respect to the measure g. Equation (9) (with the max and min operators) selects the set E which best satisfies both the measure criteria g and $\min_{x \in E} h(x)$.

From a practical point of view the calculation of the fuzzy integral can be operated as follows: let $Y = \{y_1, y_{2,\ldots,} y_n\}$ be a finite set and let $h : Y \rightarrow [0,1]$ be a function. Ordering the set such as $h(y_1) \geq h(y_2) \geq \ldots \geq h(y_n)$, a fuzzy integral e, with respect to a fuzzy measure g over Y is computed as

$$e = \max_{i=1..n} \left[\min \left(h(y_i), g(A_i) \right) \right] \tag{11}$$

where $A_i = \{y_1, y_{2,\ldots,} y_i\}$. The use of a fuzzy measure $g\lambda$ allows an easy calculation of the $g(A_i)$ which can be determined recursively as

$$\begin{aligned} g(A_1) &= g(\{y_1\}) = g^1 \\ g(A_i) &= g^i + g(A_{i-1}) + \lambda g^i g(A_{i-1}) \end{aligned} \tag{12}$$

in which

$$\lambda + 1 = \prod_{i=1}^{n} (1 + \lambda g^i) \tag{13}$$

Is calculated by a solving a $(n-1)$ degree polynomial. In (11), $h(y_i)$ represents the output of one classifier (SVM in this case), while the degree of importance g^i of how important y_i is in the recognition of a specific class must be defined. Once this is defined, λ is calculated by (13), then using (11) and (12) the fuzzy integral is calculated and the largest integral value is taken as the output class.

In this paper, fuzzy integral is applied to the output of three different classifiers (the SVM signals) which operate (each of them) on a different physical quantity; Figure 6 shows the flow chart of the proposed approach.

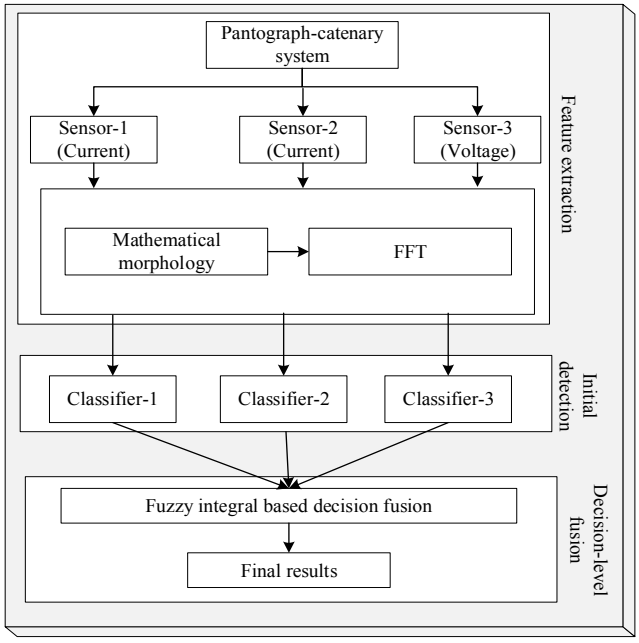

Figure 6. The block scheme of the proposed method

The detailed processes of the proposed approach is described as follows:

Step 1: Two currents, one voltage, and two photo tube signals are acquired.

Step 2: The signals relative to voltage and currents are pre-processed as described before (dilation and FT); then the so-obtained signal, together with the corresponding phototube data are used to create the training and testing data sets.

Step 3: The parameters of each support vector classifier are adjusted and each classifier is trained based on a training data set; in particular a radial basis kernel function is used because of its good performance on non-linear data. As index to evaluate the performance of each classifier the Accuracy rate defined below is introduced

$$Accuracy\ rate = \frac{TP_N}{(TP_N + FP_N)} \tag{14}$$

where subscript *N* represents the size of the data set. *TP* and *FP* show the number of correctly and incorrectly classified samples.

Step 4: Success rates of three classifiers are combined by applying fuzzy integral. After the final results were obtained from fuzzy integral based fusion approach, occurred arcs are detected according to the output of fuzzy integral.

6. Experimental Results

Let us suppose that some arc events occur at certain instants during train run; the time positions relative to all the occurrences is recorded as k_i ($i=1...N$) by analyzing the signals of the phototube sensors. This allows the construction of the training set relative to the arc events considering time windows of length $2m$ $[k_i - m, k_i + m]$ of current and voltage signals samples. Figure 7 shows the output of one phototube signal and a threshold signal obtained using a moving average approach.

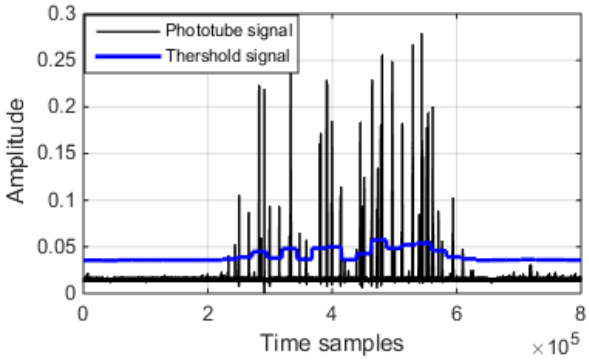

Figure 7. Phototube signal and threshold signal

As shown in Figure 7, the phototube signal is affected by noise; for this reason, in order to correctly detect the presence of an arc a thresholding procedure should be set up. In order to do so, a moving average is performed over a sliding window of 24000 samples. Afterwards, a threshold value is added to this average value in order to remove the noise. When the phototube signal exceeds the threshold value the time sample is selected and, as said before, a corresponding vector of dimension *2m* (for all the recorded physical quantities) is included in the arc event data set (AI); on the contrary if a time window of *2m* samples is not related to an arc event, it is inserted in the healthy event data set (HI). It is worth underlining that the phototube signal is no longer used during the test stage.

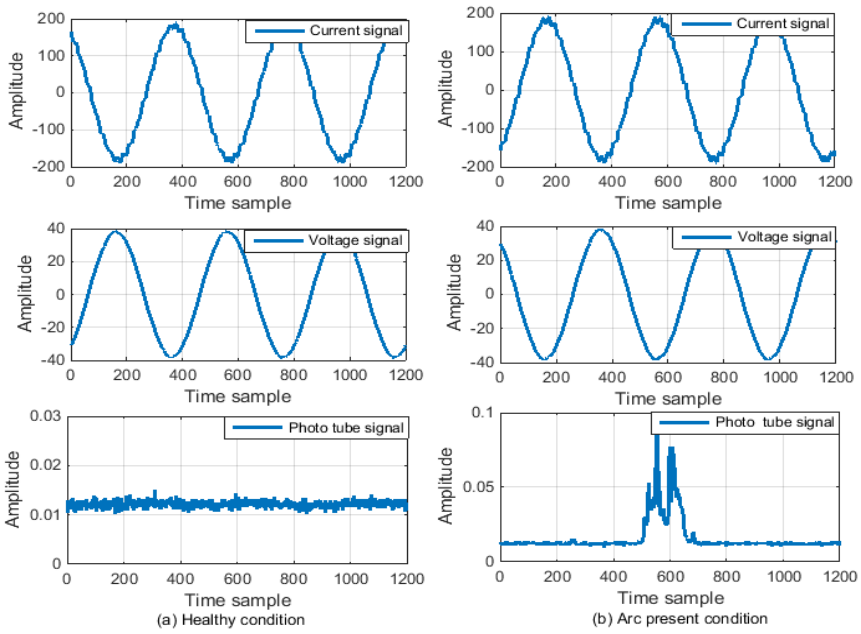

Figure 8. A sample for current, voltage, and phototube signals for healthy and arc present condition

Arc detection results

Figure 8 shows the current, voltage and phototube signals for a healthy condition and an arc present condition. In Figure 8 (a), the level of the phototube signal is near to the zero for a healthy condition. A peak occurs in the middle of phototube signal when an arc occurs as shown in Figure 8 (b): it is evident that the arc and no arc conditions cannot be separated by a direct analysis of the time domain signals. The dilation operation is applied to all the signals and the relative spectrum is calculated as shown in Figure 6.

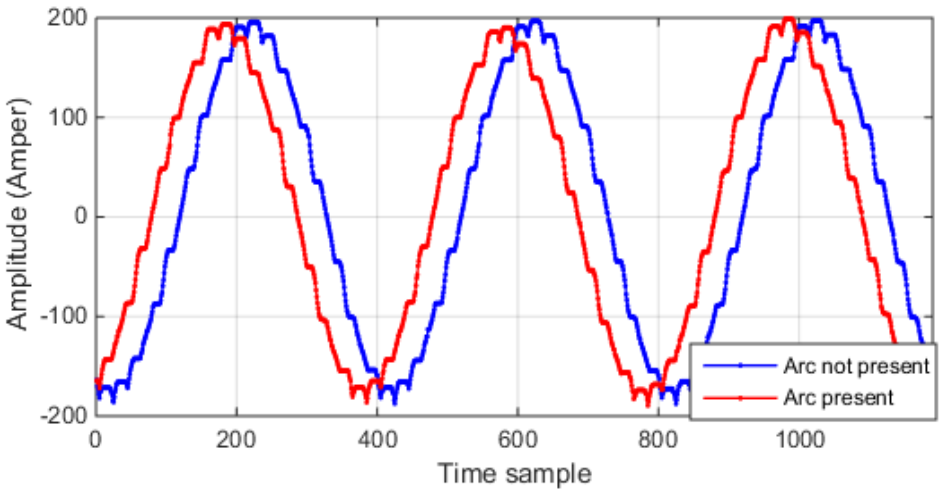

Figure 9. A sample of dilated current signal for healthy and arc occurred conditions

As shown in Figure 9, current signal is distorted by applying dilation operator. The window size of dilation operator in this figure is set to 10 and was selected by a heuristic trial and error procedure. A low window size would lead to spectrums of a healthy and an arc condition too close to each other, on the contrary if the window size is too large the information contained by original signal may be lost.

The spectrums of two current and one voltage signals are given in Figure 10 for three healthy and arc related conditions.

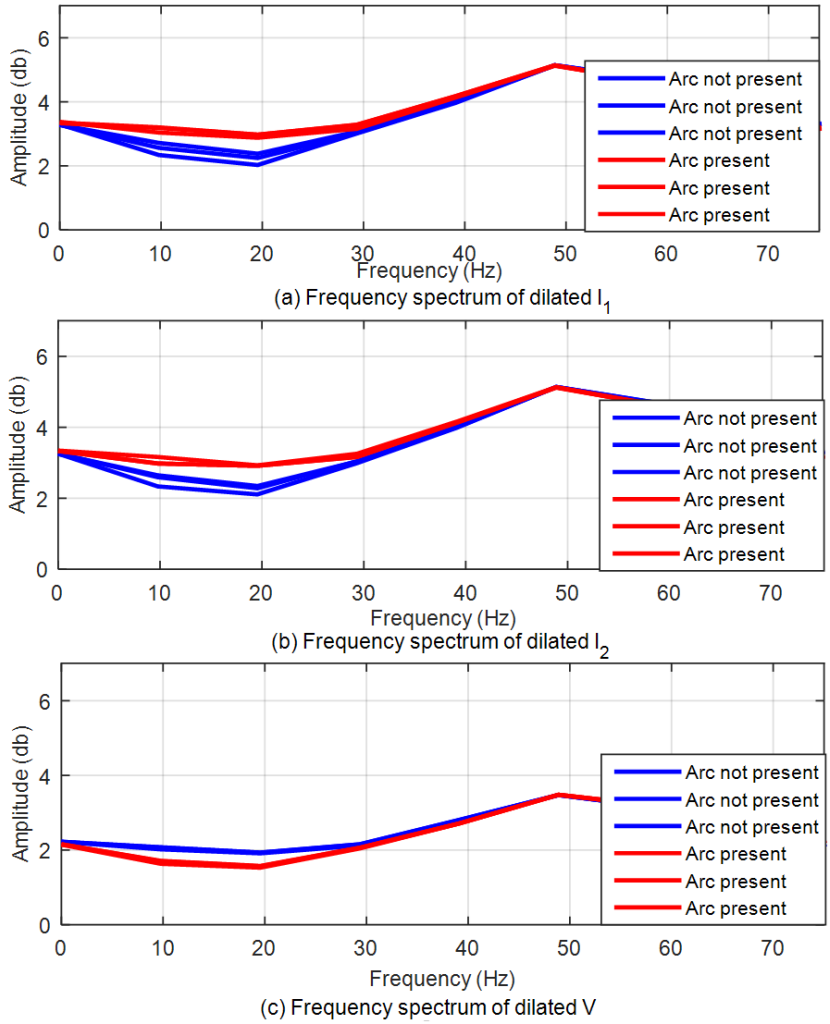

Figure 10. Fourier spectrum of dilated signals for three healthy and arc occurred conditions

In Figure 10, the spectrums of two current and one voltage signals are depicted for three samples of each condition. As it is evident in Figure 10 (a) and (b), the spectrums of the two current signals have distinctive features because the distance between spectrums of two conditions are large, while the same does not hold for the voltage current, in which the frequency spectrums are relatively closer to each other. Moreover for the current

signals in Figure 10(a) and 10(b) the spectrum signals relative to the presence of arc have higher values than the spectrum signals of the no-arc condition, which is not the case for the voltage signal in figure 10(c). This heuristic analysis confirm that the current signal is more informative, as observed also in [17]

As additional information which can be extracted from Figure 10 is that the frequency range of interest is in the range [0 30 Hz], in which the characteristics of an arc and no-arc condition are different.

After constructing the three training datasets, that contain a balanced number of spectrum samples of the dilated signals for the two conditions, the spectrum samples are used to train three support vector machines, one for each different input signal (current I_1, current I_2 and voltage V).

As for the characteristics of the Support Vector Machines, radial basis kernel functions where used, while the selection of the parameters C and σ was chosen based on previous authors' experience, in particular they were respectively set to $C = 22.80$ and $\sigma = 0.95$.

Once the above mentioned parameters are found, each support vector machine is trained and an optimal separating hyperplane with its relative support vectors are obtained. The next step is the fusion of the results obtained by the three classifiers; to this aim a set of parameters relative to the fuzzy integral should be defined, and they are shown in Table 1.

Table 1. The parameters of fuzzy integral

Parameter	Value
g^1	0.34
g^2	0.32
g^3	0.33
λ	0.0305

In Table 1, the parameters $g^{1, 2, 3}$ represents the weight of each classifier, while λ is calculated by solving (n-1)st degree polynomial in (13). The unique root (λ_t) of this polynomial take a value in the range of -1 and $+\infty$ and should be different from 0. In (13), the unique root that satisfy this condition is calculated as 0.0305. The calculation of fuzzy integral is based on the knowledge of each fuzzy density g^i. Fuzzy densities are interpreted as the weight of each classifier and they are determined by an expert knowledge. In our study, the same weight is evenly selected for each classifier

The accuracy of the fuzzy integral based method was proved by applying n-fold cross-validation technique: in all experiments, tenfold cross-validation was used. In a cross-validation method, the entire data set is divided into ten partitions: nine partitions are used for training and the remaining one is used to test the classifier. A classification performance is obtained by using a test partition, afterwards, another partition is used for the test and remaining nine partitions are used for the training purpose. An average performance and a standard deviation were obtained over ten-folds. Figure 11 shows the accuracy rate of the classification procedure obtained by using different time windows of the dilation procedure. Previously we mentioned that there is tradeoff between shorter

and longer window. For this reason, we repeated the whole procedure using four different time windows, and the results clearly show that a dilation window dimension of 10 samples gives the best performances.

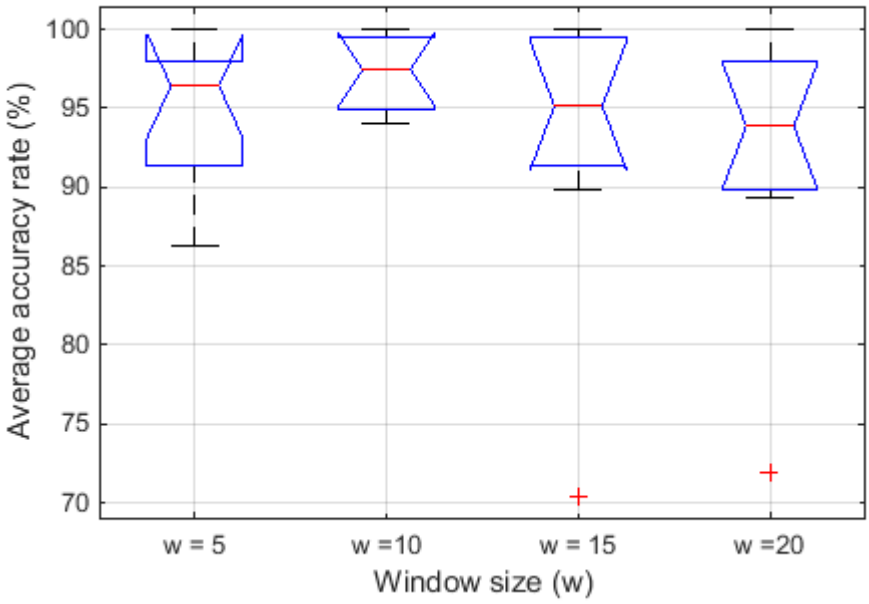

Figure 11. The average accuracy rates for different window sizes for dilation transform

To validate the fuzzy integral based fusion method, ten-fold cross validation was applied to each condition. The five thousand samples dataset was divided into ten parts and the nine parts were used for training and remaining one was used for testing in each step. Fuzzy fusion based method was also tested applying the Fourier transform to the original signal, i.e. not using the dilation operator. Comparison results are given in Table 2.

Table 2. Accuracy rates of individual classifier versus fusion approach

Classifier	Used signals	Accuracy rate (%)					
		Data set-1		Data set-2		Data set-3	
		Original	Dilated	Original	Dilated	Original	Dilated
SVM_1	V	71.19±17.23	90.14±6.74	87.82±11.76	91.83±11.42	87.14±17.85	90.48±11.04
SVM_2	I_1	70.16±16.52	88.56±18.65	90.35±6.28	89.90±13.54	84.49±9.67	90.50±10.31
SVM_3	I_2	90.13±6.76	91.71±16.99	57.48±29.45	87.28±13.71	84.70±6.36	87.90±16.83
FF	V, I_1, I_2	92.03±10.42	97.92±3.93	91.43±13.11	96.84±4.01	88.2±9.20	94.18±9.85

In particular, Table 2 contains the accuracy values of the results obtained by the single classifiers, with and without dilation and the results of the fusion method.

By combining three classifiers, the obtained overall performance of the fuzzy fusion is 96.31%. However, the average accuracy rate of an individual SVM classifier is less than 90%. Results indicate that the fusion of classifier outperforms each SVM classifier. Moreover the use of the dilation nonlinear filtering loads to substantial superior performance in almost all the cases.

In the literature, limited studies were proposed to detect arcs by using voltage and current signals in the pantograph-catenary system. The performance of the proposed approach was also compared to other methods. Regarding the preprocessing step, methods such as periodogram, wavelet analysis, and empirical mode decomposition [30] were used to obtain arc related features from current or voltage signals. Obtained features are classified as healthy or anomalous by applying an intelligent method such as multi-layer neural

networks, support vector machines, or radial based networks. Comparison results are given in Table 3, relative to a 5-fold cross validation evaluation.

Table 3. Comparison results of different methods

The method	Used signals	Pre-processing	Classifier	Accuracy Rate (%)
[17]	Current and voltage signals	Periodogram	Support vector machines	90.00
	Current and voltage signals	Periodogram	Radial based function network	85.00
[30]	Current signals	Empirical mode decomposition	Feed-forward neural networks	82.9
Our study	**Current and voltage signals**	**Mathematical morphology based Fourier spectrum**	**Fuzzy integral based fusion**	**95.64**

As shown in Table 3, it can be seen that the accuracy rate of the fuzzy integral based fusion method is higher than that of other methods. In [17], different combinations of the current and voltage signals were used for the detection of arcs. The best performance was obtained when the combination of three signals (voltage appended to the sum between two currents) was used as the input of the support vector machine classifier. The power spectrum of each signal is computed and truncated spectrum signals are taken as inputs for support vector machines. However, the input size of support vector machine is large because the window size of each spectrum is selected as 48 samples. This increases both the computational complexity and processing time. In our method, ten samples from each spectrum are given to each individual SVM classifier. Afterwards, the results of each individual classifier are combined to improve detection performance by using fuzzy integral. In [17], the

parameters of radial based function are very important to obtain a good performance. So, the number of neurons in hidden layer and the width of radial basis functions are selected as 120 and 6, respectively. The number of neurons in the hidden layer is relatively large because of the high dimensionality of input size. However, neural networks have some drawbacks such as slow convergence rate to different local minima, and use of many neurons in the hidden layer.

By comparing the accuracy rate shown in Table 3, it can be seen that the performance of our method is superior to the other methods. This performance is mainly achieved by applying a new feature extraction and a fusion approach of individual classifiers.

7. Conclusions

This paper presents a new approach for the detection of electric arcs that uses the fuzzy integral based fusion method. The proposed fusion approach combines the classification results of three SVM classifiers by using the fuzzy integral. The performance of each individual classifier is increased by applying the dilation transform to the time domain signals as a preprocessing step. Two current signals, two photo tube signals, and one voltage signal are recorded during three different travels. Photo tube signals show whether an arc occurs or not. So, these signals are used to separate healthy and arc occurred signals in the training stage. Mathematical morphology was demonstrated to improve the separation of healthy and arc occurred signals, as arc related frequencies are

emphasized by applying dilation transform to the time domain signals. The main contributions of the proposed method are summarized as follows:

- The presented dilation preprocessing method is shown to emphasize arc related frequencies,
- The dimension of the data vectors is reduced with respect to other approaches
- The fuzzy integral classification fusion technique outperforms single SVM classifiers.

In conclusion, the experimental results show that the proposed fusion approach gives better results than other methods, leading to a state of the art approach for electric arc detection in real time railway applications. The fusion of SVMs method used in this paper is a universal approximator, and we have shown that it can be appropriately trained to provide a good generalization performance on new test data. In general, it is no possible to guarantee that the same method maintains the same performance on a completely different dataset or application. However, if the application is similar, because the underlying physics represents the same phenomenon, it is expected that for similar data we could obtain similar performance. The six runs of data analyzed contain a huge number of observations, consequently these data can be considered statistically significant for representing the phenomenon of electrical arcing in high speed trains. Therefore, the method adopted for revealing the electric arc presence possess a general validity for applications in similar conditions.

Acknowledgments

This work was supported by the TUBITAK (The Scientific and Technological Research Council of Turkey) under Grant No: 112E067.

References

[1] **Swift, M, Aurisicchio G., Pace P**. New practices for railway condition monitoring and predictive analysis. *IET Conference on Railway Condition Monitoring and Non-Destructive Testing*, *RCM 2011,* Derby, UK, 29–30 November 2011, pp.1–6.

[2] **Betts, A. I., Hall, J. H., Keen, P. M.** Condition monitoring of pantographs. *International Conference on Main Line Railway Electrification*, 25 – 28 September 1989, pp. 129 – 133.

[3] **Landi, A., Menconi, L., Sani, L.** Hough transform and thermo-vision for monitoring pantograph-catenary system. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*,, 2006, 220 (4), 435 –444.

[4] **Zhang, H., Luo, L., Yang, K., Wang, L., Gao, X.** Improved multi-scale wavelet in pantograph slide edge detection. *Optik - International Journal for Light and Electron Optics*, 2014, 125 (19), 5681–5683.

[5] **Hamey, L. G. C., Watkins, T., Yen, S. W. T.** Pancam: in-service inspection of locomotive pantographs. *IEEE 9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications*, Glenelg, Australia, 3-5 Dec. 2007, pp. 493 – 499.

[6] **Xiao-heng, Z., Xiao-rong, G., Ze-yong, W., Li, W., Kai, Y.** Study on the edge detection and extraction algorithm in the pantograph slipper's abrasion. *IEEE International Conference on Computational and Information Sciences,* ICCIS2010, Chengdu, China, 17-19 Dec. 2010, pp. 474 - 477.

[7] **Stella, E., Mazzeo, P. L., Nitti, M., Distante, A.** Non-destructive analysis (NDT) of pantograph and catenary interaction. *IEEE International Conference on Pantograph Catenary Interaction Framework for Intelligent Control*, Amiens, France, 8 Dec. 2011, pp.1-6.

[8] **Kim, H., Park, Y., Cho, Y., Kim, I.** Development of image processing technology for interaction between pantograph and overhead contact wire. *Journal of the Korean Institute of Electrical and Electronic Material Engineers*, 2009, 22 (12), 1084-1088.

[9] **Boguslavskii, A.A., Sokolov, S. M.** Detecting objects in images in real-time computer vision systems using structured geometric models. *Journal of Programming and Software*. 2006, 32, 177-187.

[10] **Aydin, I., Karakose, M., Akin, E.** A new contactless fault diagnosis approach for pantograph-catenary system using pattern recognition and image processing methods. *Advances in Electrical and Computer Engineering*, 2014, 14 (3), 79-88.

[11] **Aydin, I., Karakose, M., Akin, E.** Anomaly detection using a modified kernel-based tracking in the pantograph-catenary system. *Expert Systems with Applications*, 2015, 42 (2), 938-948.

[12] **Cho, C. J., Ko, H.** Video-based dynamic stagger measurement of railway overhead power lines using rotation-invariant feature matching. *IEEE Transactions on Intelligent Transportation Systems*, 2015, 16(3), 1294 - 1304.

[13] **Mitsuo, A.** Precise measurement and estimation method for overhead contact line unevenness. *Electrical Engineering in Japan*, 2007, 160 (2), 77-85.

[14] **Bruno, O., Landi, A., Papi, M., Sani, L., Violi, A.G., Elettriche, S. P.** Pantograph-catenary monitoring: correlation between breack arcs and harmonics in the traction currents. *Proceedings of the World Conference on Railway Research*, Koln, Germany, 19-23 Oct. 2001, pp. 1-17.

[15] **Bruno, O., Landi, A., Papi, M., Sani, L.** Phototube sensor for monitoring the quality of current collection on overhead electrified railways. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, 2001, 215 (3), 231-241.

[16] **Barmada, S., A. Landi, M. Papi, and L. Sani.** Wavelet multiresolution analysis for monitoring the occurrence of arcing on overhead electrified railways. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit, 2003,* 217 (3), 177-187.

[17] **Barmada, S., Raugi, M., Tucci, M., Romano, F.** Arc detection in pantograph-catenary systems by the use of support vector machines-based classification. *IET Electrical Systems in Transportation,* 2014, 4 (2), 45–52. .

[18] **Soille, P.** Morphological image analysis: principles and applications. *Springer Science & Business Media*, 2013, p. 368.

[19] **Serra, J.** *Image analysis and mathematical morphology*. Academic Press, Inc., 1983, p. 411.

[20] **Radil, T., Ramos, P. M., Janeiro, F. M., Cruz-Serra, A.** PQ monitoring system for real-time detection and classification of disturbances in a single-phase power system. *IEEE Transactions on Instrumentation and Measurement*, 2008, 57 (8), 1725–1733.

[21] **Jing, M., Yan, X., Zengping, W., Haofang, L.** A novel adaptive scheme of discrimination between internal faults and inrush currents of transformer using mathematical morphology. IEEE International Conference of Power Engineering Society General Meeting, Montreal, Que, Jun. 2006, pp. 1–7.

[22] **de Jesus Rangel-Magdaleno, J., Peregrina-Barreto, H., Ramirez-Cortes, J. M., Gomez-Gil, P., Morales-Caporal, R.** FPGA-based broken bars detection on induction motors under different load using motor current signature analysis and mathematical morphology. *IEEE Transactions on Instrumentation and Measurement,* 2014, *63* (5), 1032-1040.

[23] **Barmada, S., Tucci, M., Menci, M., Romano, F.** Clustering techniques applied to a high-speed train pantograph–catenary subsystem for electric arc detection and classification. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, doi:10.1177/0954409714528486.

[24] **Hayes, M.,** Statistical Digital Signal Processing and Modeling. New York: Wiley, 2009, p. 624.

[25] **Cortes, C.,** Vapnik, V. N. Support vector networks, *Machine Learning*, 1995, 20 (3), 273–297.

[26] **Chang, C.-C., Lin, C.-J.** "LIBSVM: a library for support vector machines", ACM Trans. Intell. Syst. Technol., Vol. 2, No. 3, pp. 1–39, 2011.

[27] **Cho, S.B., Kim, J. H.** Multiple network fusion using fuzzy logic. *IEEE Transactions on Neural Networks*, 1995, 6 (2), 497–501.

[28] **Cho, S.B., Kim, J. H.** Combining multiple neural networks by fuzzy integral for robust classification. *IEEE Transactions on Systems, Man, and Cybernetics*, 1995, 25 (2), 380–384.

[29] **Temkoa, A., Macho, D., Nadeu, C.** Fuzzy integral based information fusion for classification of highly confusable non-speech sounds, *Pattern Recognition*, 2008, 41 (5), 1814–1823.

[30] **Barmada, S., Tucci, M.** Use of advanced signal processing techniques for arcing detection on AC pantograph catenary systems. *IEEE International Conference on International Conference on Pantograph-Catenary Interaction Framework for Intelligent Control*, pp. 75-81. 2011.