Arc fault detection and appliances classification in AC home electrical networks using recurrence quantification plots and image analysis

Arc fault detection and appliances classification in AC home electrical networks using Recurrence Quantification Plots and Image Analysis

Francesco Ferracuti[a], Patrick Schweitzer[b,*], Andrea Monteriù[a]

[a] *Università Politecnica delle Marche, Department of Information Engineering via Brecce Bianche, 60131 Ancona, Italy*

[b] *Université de Lorraine, Institut Jean Lamour (IJL), CNRS, F-54000 Nancy, France*

Abstract

This paper presents a method for the detection of series arc faults in electrical circuits, which has been developed starting from the recurrence quantification plots that allow to quantify the periodic behavior of time-series and to analyze the recurrences of a dynamical system presented by its phase space trajectory. Starting from this, the authors have found that it is possible to exploit recurrence quantification plots by using the gray-level co-occurrence matrix from which the extracted textural image features represent a proper set of indicators for suitably detecting arc faults. The database of this research is collected from 13 different types of load according to IEC 62606 standard. The proposed method's effectiveness is shown by means of experimental tests, which were carried out in both arcing and non-arcing conditions and in the presence of different loads.

Keywords: arc fault, recurrence quantification plot, gray-level co-occurrence matrix, image analysis, series arc

Nomenclature

*Corresponding author

Email addresses: f.ferracuti@univpm.it (Francesco Ferracuti), patrick.schweitzer@univ-lorraine.fr (Patrick Schweitzer), a.monteriu@univpm.it (Andrea Monteriù)

μ_x, μ_y	Mean deviations of $P_x(i)$ and $P_y(j)$
σ_x, σ_y	Standard deviations of $P_x(i)$ and $P_y(j)$
τ	Time delay
Θ	Similarity function
AC	Alternating current
ACF	AutoCorrelation Function
$C(i,j)$	Co-occurrence probabilities between gray level i and j
d, θ	Trigonometric functions
f_0	Fundamental frequency
F_d	Sampling frequency after downsampling
F_s	Sampling frequency
$GLCM$	Grey Level Co-Occurence Matrix
HC	Number of Half Cycle
$HXY1$	GLCM entropy 1
$HXY2$	GLCM entropy 2
$I(x,y)$	gray image of size N_x pixels by N_y pixels
L	Quantized gray levels
m	Embedding dimension in the phase space
$P(i,j)$	Number of co-occurrences of gray level i and j
RQ	Recurrence Quantification
RQA	Recurrence Quantification Analysis
RQP	Recurrence Quantification Plots

S Similarity Matrix

T_W Time window duration

$TDV-CNN$ Temporal Domain Visualization Convolutional Neural Network

1. Introduction

Arc faults are electric arcs that originated in different home areas or virtually in any electrical fixture that occur accidentally in series or parallel and generally caused by wire's long-term work with load or overload, junction's loose connection, external intrusion raised insulation damage, etc. [1]. If an arc fault is maintained for a long time, the energy produced by the arc may lead to the ignition of a fire accident. According to the Forum for European Electrical Safety (FEEDS), an estimated 270000 (reported or not) domestic fires of electrical origin (about 20 to 30% of all domestic fires) still occur every year throughout the EU, estimated to cause an average of 1000 fatalities and 20000 injuries each year and annual property damage of 6.25 billion euro [2]. Whereas according to the National Fire Protection Association (NFPA), 44880 home fires involving electrical failure or malfunction are reported each year in 2012-2016, with 440 civilian deaths and 1250 civilian injuries each year, as well as an estimated 1.3 billion in direct property damage a year [3].

Arc faults usually present as signature an indistinctive stochastic change of current magnitude [4], which makes difficult to detect this kind of fault with the common circuit protectors. In addition, in the literature, it is put in evidence that the features of arc faults are various and complex, especially in the presence of nonlinear or switching loads [4], therefore arc fault detection and appliances classification require customized and advanced techniques which are able to extract arc fault signatures [5]. Normally, the current time-series vary according to loads and circuit state (normal and arcing). The wave shapes of normal current behave variously on different load types: some are almost sinusoidal, others have

flat shoulders or catastrophe points. When an arc fault is occurring, many abnormal behaviors arise, such as impulse, or spiking, increment of high-frequency harmonic component, decrement of the conduction angle, abrupt change of AC signal, time-domain waveform distortion and amplitude decreases. The variety, complexity and chaotic nature of the current time-series during arc fault and the number of possible scenarios impose a great challenge on designing reliable arc fault detection methods [6, 7].

Different arcing fault detection based on single or multi-criteria approaches have been presented in literature. The widely used methods are correlation analysis, crest factor, Kalman filter, inter-period correlations of current, algebraic derivative [4, 8, 9, 10, 11]. Others are based on time-frequency analysis and mainly the wavelet transform. In this case, the detection performances are strongly linked to the sampling frequency, the level of signal composition and the mother wavelet according to the domestic appliance in the circuit [12, 13, 14, 15, 16]. With the rise of processing power for the embedded system, internet of things and the promising results of machine learning in many fields, several works propose arc fault detection methods based on machine learning or deep learning techniques [17, 18, 19, 20, 21]. Machine learning-based algorithms in conjunction with classical methods can bring better performance for arc fault detection. Machine learning approaches are mainly based on pattern classification in order to detect the arc fault and, consequently, identify the appliance involved in the fault [22, 23]. Duc Vu *et al.* [20] proposed an original methodology for the choice of optimal arc fault features (transforms or direct feature extraction) in a first step and then for combining them using a machine learning technique. The method consists of creating an arc fault feature pool and finding a combination of those features which satisfy the desired performance. The method is tested efficiently with twenty-one specific transforms associated with 10 different descriptors. A major interest would be to be able to also locate the arc fault on the line, and more precisely identify the load which presents a fault for electrical troubleshooting for example. Many methods related to non-intrusive load monitoring have been presented in the literature in order to

appliance recognition [24, 25, 26], and the most recent ones are based on a machine learning approach [27, 28, 29, 30] but none is associated with a detection method. There are few methods in the literature that combine both arc detection and load classification. Wang *et al.* [6] proposed a processing procedure with a sparse representation with a fully connected neural network methodology for series arc fault identification. The general classification reaches 94.3% with 10 tested classes. Chu *et al.* [16] proposed a temporal domain visualization step associated with a convolutional neural network (TDV-CNN). Every half-cycle of the measured current is transformed into a grey image. The detection accuracy reaches 98.7% and can be used to distinguish various load types. Yang *et al.* [12] developed a signal processing method to extract the high frequency radiation characteristics, the current waveform periodic integral characteristics, and the current waveform slope characteristics of the arc faults. Then the recognition algorithm is based on multi-information fusion and support vector machines. Arc fault identification is made on single and combined loads.

In this work, a method for the detection of series arc faults in electrical circuits and appliances classification is proposed by exploiting the Recurrence Quantification (RQ) plots. The Recurrence Quantification Analysis (RQA) is a methodology that allows quantifying the periodic behavior of time-series and analyzing the recurrences of a dynamical system presented by its phase space trajectory. RQ plots can be analyzed by using the gray-level co-occurrence matrix from which the extracted textural image features represent a proper set of indicators for suitably detecting arc faults. The effectiveness of the proposed method is shown by means of experimental tests, which were carried out in both arcing and non-arcing conditions and in the presence of different loads. The main contribution and novelty of this paper consists of the exploitation of the recurrence plots to discriminate the occurrence of arc faults and classify the load types. The recurrence quantification analysis demonstrated to be a reliable tool to analyze AC signals for electric motor fault detection [31, 32] and power system analysis [33]. The recurrence plots are analyzed by using an efficient image analysis technique such as Gray-Level Co-Occurrence Matrix (GLCM).

Moreover, the authors propose a solution based on the physical knowledge of such arc signals, in order to set the time delay that usually is found by the application of mutual information [34].

The rest of this paper is organized as follows. Section 2 introduces the preliminary theory of RQA, and details of textural features extraction and classification for detecting arc faults. The general experimental test bench used to generate a series arc faults in an electrical network in order to record the current electrical signatures is described in Section 3. In Section 4, the proposed methodology is evaluated with real data of various load types and work states; the comparison with other methods is also provided. Section 5 presents the conclusion.

2. Methodology

As a nonlinear time-series analysis technique, the proposed one has been developed from a phase space point of view [35, 36]. It is well-known that time-delay coordinates allow to synthesize additional dynamic variables using a time-series measurement from a single variable. Considering a time-series of N samples obtained from the sensor measurement $[y(1), y(2), \ldots, y(N)]$, its corresponding phase space can be constructed through a time-delay approach according to Takens' embedding theorem [37]. The dynamics in these time-delay coordinates produce a new attractor with the same topology. In particular, the following matrix is constructed by stacking delayed time-series of y as rows:

$$\begin{cases} Y(1) = [y(1), y(1+\tau), \ldots, y(1+(m-1)\tau)] \\ Y(2) = [y(2), y(2+\tau), \ldots, y(2+(m-1)\tau)] \\ \ldots \\ Y(N-(m-1)\tau) = \\ \qquad [y(N-(m-1)\tau), y(N-(m-2)\tau), \ldots, y(N)] \end{cases} \tag{1}$$

where τ is the time delay, and m is the embedding dimension in the phase space. The representation of time-series in phase is based on dynamical systems theory,

and the time-delay embedding process is able to reconstruct a topologically identical state space to the original state space of the system [37, 38, 39]. Hence, the features that capture the similarities and differences of the reconstructed state space are capturing similarities and differences of the underlying state structure of the system [40].

In this paper, a novel features extraction based on similarity matrix and statistical properties of the image derived from GLCM is proposed. This theoretical relationship between the reconstructed state space and the original state space to detect arc faults is then exploited. In order to investigate the m-dimensional phase space trajectory through a two-dimensional representation, the similarity matrix is exploited:

$$S_{i,j} = \Theta\left(\|Y(i) - Y(j)\|\right) \tag{2}$$

where $i, j \in [1, N-(m-1)\tau]$, Θ is a similarity function, and $\|\cdot\|$ is a norm. The plot generated by Eq. (2), is a graphical description of the dynamic properties of a time-series and it can be exploited to extract features to detect arc faults. RQA is a related approach that extracts the recurrence matrix from the similarity matrix as follows:

$$R^{\epsilon}_{i,j} = \Theta\left(\epsilon - \|Y(i) - Y(j)\|\right) = \begin{cases} 1: & \|Y(i) - Y(j)\| \leq \epsilon \\ 0: & \|Y(i) - Y(j)\| > \epsilon \end{cases} \tag{3}$$

where ϵ is a threshold parameter. Recurrence quantification analysis converts the similarity matrix into a binary image (i.e., the Recurrence Quantification Plots), depending on whether the value of each pixel is lower than the predefined threshold ϵ. Recurrence quantification plots have demonstrated to be a valuable data visualization and analysis tool for time-varying dynamical systems in a wide group of application areas, e.g. for engineering, finance, biology and other disciplines [41, 42, 43]. The characteristics of the system, which can be represented by some statistical parameters, such as recurrence rate, determinism, divergence, Shannon entropy, laminarity, and trapping time, extracted from line structure and point density of the recurrence plot in a quantitative

manner, are referred as RQA. The extracted features from RQA can be exploit to detect faults and anomaly by means of a classifier [32]. However, the setting of the threshold ϵ is not trivial, and the binary transformation of the similarity matrix (i.e., the RQP), consequently, it leads to a loss of information that could instead be exploited for fault detection. For this reason, in this work, the similarity matrix has been directly exploited to extract features to detect arc faults through its graphical description.

Feature extraction plays a crucial role in fault diagnosis. The different types of fault can generate different characteristics in the current measurements. Here, GLCM is investigated and its based features are extracted to capture spatial distribution information underlying the reconstructed state space. As described in [44], GLCM represents the conditional joint probabilities of all pair combinations of gray levels in the spatial window of interest with respect to two parameters: inter-pixel distance d, and direction θ. Given a gray image $I(x, y)$ of size N_x pixels by N_y pixels, the spatial relationship of the pixel pair (i, j) with specified pair (d, θ) is defined by trigonometric functions. The joint probabilities measurement is expressed as $P(i, j, d, \theta) = [C(i, j)|(d, \theta)]$, where $C(i, j)$, the co-occurrence probabilities between gray level i and j, can be given by:

$$C(i,j) = \frac{P(i,j)}{\sum_{i=1}^{L}\sum_{j=1}^{L} P(i,j)} \tag{4}$$

where L is the quantized gray level, and $P(i, j)$ represents the number of co-occurrences of gray level i and j, given by:

$$\begin{aligned} P(i,j) =& \#[((x_1,y_1),(x_2,y_2)) \in (L_x \times L_y) \\ & \times (L_x \times L_y) | I(x_1,y_1) = i, I(x_2,y_2) = j] \end{aligned} \tag{5}$$

where $x_2 = x_1 + d\cos(\theta)$, $y_2 = y_1 + d\sin(\theta)$, $i, j \in [1, \ldots, L]$, # denotes the number of the pixel pairs (i, j) satisfying the conditions, and (x_1, y_1) is the coordinate with gray level i, (x_2, y_2) the coordinate with gray level j, L_x the horizontal spatial domain $(1, \ldots, N_x)$, L_y the vertical spatial domain $(1, \ldots, N_y)$.

From the gray-tone spatial-dependence matrices, a set of 21 measures of textural features has been defined. Some of these measures relate to specific

textural characteristics of the image such as homogeneity, contrast, and the presence of organized structure within the image. Other measures characterize the complexity and nature of gray-tone transitions which occur in the image. The considered textural features are defined in the Table 1, where L is the number of distinct gray levels in the quantized image, and

$$P_{x+y}(k) = \sum_{\substack{i \\ i+j=k}}^{L} \sum_{j}^{L} P(i,j), \;\; k = 2, \ldots, 2L \tag{6}$$

$$P_{x-y}(k) = \sum_{\substack{i \\ |i-j|=k}}^{L} \sum_{j}^{L} P(i,j), \;\; k = 1, \ldots, L-1 \tag{7}$$

$$P_x(i) = \sum_{j}^{L} P(i,j) \tag{8}$$

$$P_y(j) = \sum_{i}^{L} P(i,j) \tag{9}$$

$$HXY1 = -\sum_{i}^{L} \sum_{j}^{L} P(i,j) \log(P_x(i) P_y(j)) \tag{10}$$

$$HXY2 = -\sum_{i}^{L} \sum_{j}^{L} P_x(i) P_y(j) \log(P_x(i) P_y(j)) \tag{11}$$

where μ_x, μ_y are the means and σ_x, σ_y the standard deviations of $P_x(i)$ and $P_y(j)$, respectively. More details about the significance of the previous equations can be found in [44].

The proposed method is shown in Figure 1, in particular the flowchart shows the steps of the proposed arc fault detection and classification algorithm. The input of the proposed methodology is the acquisition of a time window of AC measurement, the second step regards the transformation of the acquired signal in phase space through a time-delay τ and an embedding dimension m (see Eq. (1)). From phase space, a similarity matrix is constructed using in this work the Euclidean distance. By the previous steps, the original AC signal is transformed into two-dimensional gray-scale image. The features are extracted from the image by GLCM and to perform this task the parameters d, θ and L

Table 1: Texture Features Extracted from GLCM

#	Feature	Equation
f1	Autocorrelation [45]	$\sum_{i}^{L}\sum_{j}^{L}(i \cdot j)P(i,j)$
f2	Cluster Prominence [45]	$\sum_{i}^{L}\sum_{j}^{L}(i+j-\mu_x-\mu_y)^4 P(i,j)$
f3	Cluster Shade [45]	$\sum_{i}^{L}\sum_{j}^{L}(i+j-\mu_x-\mu_y)^3 P(i,j)$
f4	Dissimilarity [45]	$\sum_{i}^{L}\sum_{j}^{L}\lvert i-j\rvert P(i,j)$
f5	Entropy [45, 44]	$-\sum_{i}^{L}\sum_{j}^{L}P(i,j)\log(P(i,j))$
f6	Maximum probability [45]	$\max_{i,j}P(i,j)$
f7	Inverse Difference Moment [44], or Homogeneity II [45]	$\sum_{i}^{L}\sum_{j}^{L}\frac{P(i,j)}{1+\lvert i-j\rvert^2}$
f8	Inverse difference [46] or Homogeneity I [45]	$\sum_{i}^{L}\sum_{j}^{L}\frac{P(i,j)}{1+\lvert i-j\rvert}$
f9	f7 Normalized	$\sum_{i}^{L}\sum_{j}^{L}\frac{P(i,j)}{1+(\lvert i-j\rvert/L)^2}$
f10	f8 Normalized	$\sum_{i}^{L}\sum_{j}^{L}\frac{P(i,j)}{1+\lvert i-j\rvert/L}$
f11	Contrast [45]	$\sum_{i}^{L}\sum_{j}^{L}\lvert i-j\rvert^2 P(i,j)$
f12	Energy [45]	$\sum_{i}^{L}\sum_{j}^{L}P(i,j)^2$
f13	Correlation [44], or Correlation II [45]	$\sum_{i}^{L}\sum_{j}^{L}\frac{(i\cdot j)P(i,j)-\mu_x\mu_y}{\sigma_x\sigma_y}$
f14	Difference entropy [45, 44]	$-\sum_{i=0}^{L-1}P_{x-y}(k)\log(P_{x-y}(k))$
f15	Difference variance [45, 44]	$\sum_{i=0}^{L-1}i^2 P_{x-y}(k)$
f16	Information measure of correlation I [45, 44]	$\frac{HXY-HXY1}{\max(HX,HY)}$
f17	Information measure of correlation II [45, 44]	$(1-\exp[-2(HXY2-HXY)])^{1/2}$
f18	Sum average [45, 44]	$\sum_{i=2}^{2L}iP_{x+y}(k)$
f19	Sum entropy [45, 44]	$-\sum_{i=2}^{2L}P_{x+y}(k)\log(P_{x+y}(k))$
f20	Sum of squares [45, 44]	$\sum_{i}^{L}\sum_{j}^{L}(i-\mu)^2 P(i,j)$
f21	Sum variance [45, 44]	$\sum_{i=2}^{2L}(i-\text{f19})^2 P_{x+y}(k)$

have to be set. The extracted features are exploited to train a classifier that predicts the presence or absence of arc faults and and the type of appliance.

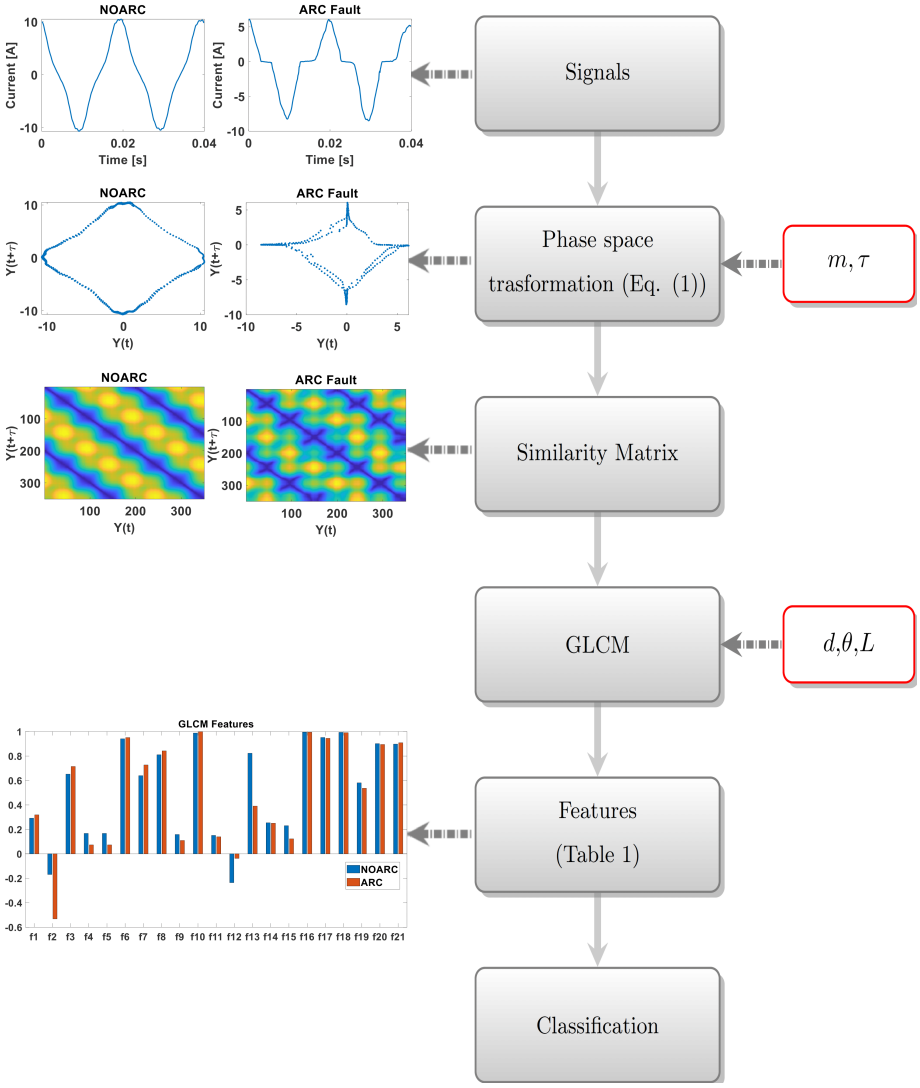


Figure 1: Flowchart of the proposed fault classification algorithm.

2.1. Linear Discriminant Analysis

The regularization operation allows overcoming the issues related to high dimensionality, where the applicability of the classifier is a demanding task

since, in this situation, direct high dimension matrix operations are required [47]. To overcome these issues, the unbiased estimate of the pooled-in covariance matrices considered in the classifier were replaced with:

$$\Sigma_\alpha = (1-\alpha)\,\Sigma + \alpha \mathrm{diag}(\Sigma) \tag{12}$$

where Σ is the empirical, pooled covariance matrix and α is the amount of regularization. For training the classier, the hyperparameter δ has to be set. δ is a non-negative scalar shrinkage parameter that specifies a threshold: if a coefficient of the model has magnitude smaller than δ, the coefficient is set to 0. Then $\alpha \in [0, 1)$ and $\delta \in [0, \infty)$ are two parameters to optimize over a 2-dimensional grid using crossvalidation on the training samples.

2.2. Setting the hyperparameters

The time delay τ is set equal to the delay corresponding to the first minimum point the of AutoCorrelation Function (ACF) of the AC time series. Considering F_s the sampling frequency, F_d the sampling frequency after downsampling, f_0 the fundamental frequency (50 Hz), n the number of samples and HC the number of half-cycles, the minimum point of the autocorrelation function for the monitored appliances is $\tau = n/HC/2$ or equivalently $\tau = F_d/f_0/4$. This is explained by the fact that a sinusoidal signal and its autocorrelation shows that optimal delay should be a quarter of its period since it provides for $\pi/2$ shift in phase. Table 2 summarizes the time delay for different values of sampling frequency, acquisition time and appliances. The last column of the table shows the time delay τ considered in the experimental analysis.

As an example, the ACFs of AC signal regarding blow heater for different sampling frequencies F_d are shown in Figure 2 and the minimum points of the autocorrelation functions correspond to those defined in the Table 2. Figures 2(a), 2(b), 2(c), 2(d), 2(e) and 2(f) show the sample autocorrelation function at sampling frequency F_d=25, 5, 4, 3.125, 2.5 and 2 kHz, respectively.

The hyperparameters m, L and d are considered as tradeoff between improved prediction accuracy of the classification and increased computation bur-

Table 2: Settings of the algorithm

F_d	**Time**	**Samples** (n)	**Half-Cycles** (HC)	**Time delay** (τ)
25 kHz	0.02 s	500	2	125
5 kHz	0.1 s	500	10	25
5 kHz	0.08 s	400	8	25
5 kHz	0.05 s	250	5	25
4 kHz	0.1 s	400	10	20
4 kHz	0.08 s	320	8	20
4 kHz	0.05 s	200	5	20
3.125 kHz	0.1 s	313	10	16
3.125 kHz	0.08 s	250	8	16
3.125 kHz	0.05 s	157	5	16
2.5 kHz	0.1 s	250	10	12
2.5 kHz	0.08 s	200	8	12
2.5 kHz	0.05 s	125	5	12
2 kHz	0.1 s	200	10	10
2 kHz	0.08 s	160	8	10
2 kHz	0.05 s	100	5	10

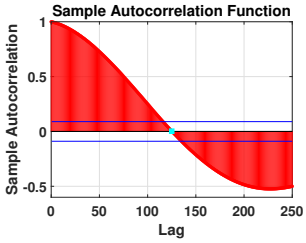
den of the algorithm. In particular, for the experimental analysis, the values $m = 3$, $L = 2^5$, and 16 inter-pixel distances (d) have been considered and set by using the training dataset. The parameters θ has been fixed to 45° because of the specific pattern of the images in the shape of diagonal stripes; as an example, the unthresholded RQ plots for blow heater, Dell desktop computer, Titan 710 W drill and for Philips vacuum cleaner, are displayed in Figure 7, showing the particular shape of diagonal stripes.

3. Test setup

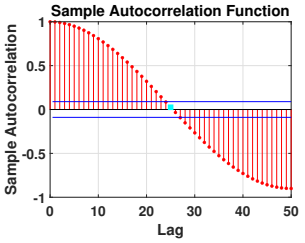
Figure 3 shows the general experimental test bench used to generate a series arc fault in an electrical network in order to record the current and voltage electrical signatures and to build the database.

The AC mains power source (230 V - 50 Hz) supplies a load which may be a single appliance or is composed of several associated appliances connected in parallel. The various household appliances considered in this study are resistive (Boiler), with switching units (PC computer), include universal motor (Vacuum cleaner, drill) or dimmer-based (Halogen).

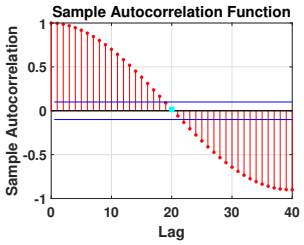
The measurements are performed with a Lecroy HDO 6104 oscilloscope (2



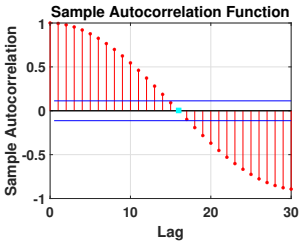
(a) ACF - F_d =25 kHz



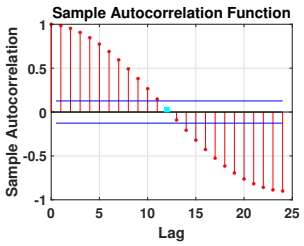
(b) ACF - F_d =5 kHz



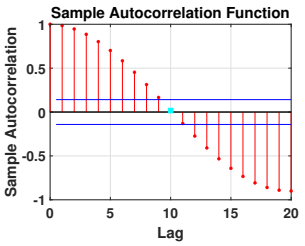
(c) ACF - F_d =4 kHz



(d) ACF - F_d =3.125 kHz



(e) ACF - F_d =2.5 kHz



(f) ACF - F_d =2 kHz

Figure 2: ACF of blow heater a) F_d =25 kHz, b) F_d =5 kHz, c) F_d =4 kHz, d) F_d =3.125 kHz, e) F_d=2.5 kHz, f) F_d=2 kHz.

GHz bandwidth). A sub-sampling is performed if the method of analysis requires to operate at a lower frequency. The line current is measured at the output of the power source (Lecroy AP 30 probe – 100 MHz bandwidth). To ensure that the fault is continuously present in the circuit, the measurement of the arc voltage across the fault is also recorded (TT-SI 9010, 70MHz bandwidth). The data files are then transferred to a computer for a Matlab analysis. The link given in [48] gives access to the database of electrical signatures which can be

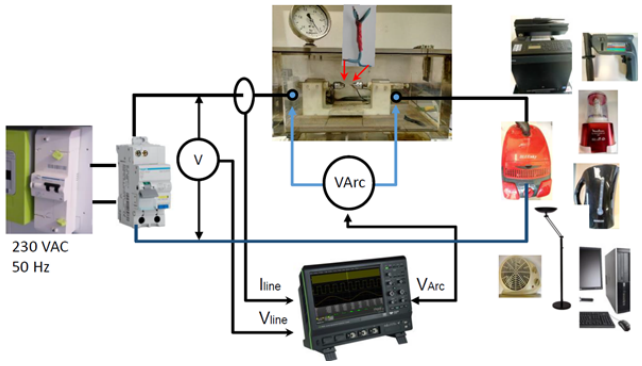


Figure 3: Experimental test bench.

used under certain conditions. The arc fault is created according to the protocol defined in the UL 1699 or IEC 62606 standards [49, 50].

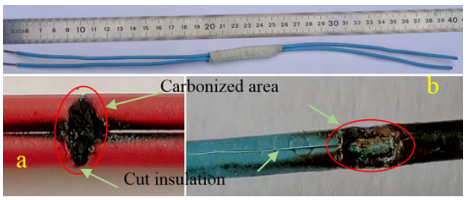


Figure 4: Cables preparation.

Two copper cables (20 cm length), with a slit across the insulation (5 cm length), are closely tied together and then wrapped. A carbonized conductive path is created between the two conductors (see Figure 4). The prepared sample cable can be inserted directly into the circuit to reproduce the arcing fault. The series arc fault is ignited into the circuit at two different times via the switching of a relay. The various household appliances considered in this study are listed in Table 3 along with the labels associated to each appliance or load type. The dataset description can be found in [48].

3.1. Data Preprocessing

The primary database is created by the raw data collection of 5 s. Besides having the primary database of 1 MHz sampling rate, to check the accuracy of

Table 3: List of appliances and corresponding labels

APPLIANCE	LOAD	APPLIANCES CLASSES	LOAD CLASSES
Pump	Universal motor	1	1
Dell Desktop Computer	Converter: switching power supply	2	2
Dolce Gusto Espresso Machine	Resistive	3	3
Titan 710 W Drill	Universal motor	4	1
Peugeot 550 W Drill	Universal motor	5	1
Mewal Halogen Lamp	Converter: single-phase rectifiers	6	4
Blow Heater	Resistive	7	3
Severin 2200 Kettle	Resistive	8	3
Moulinex Mini Food Processor	Universal motor	9	1
Multi-functional Printer	Universal motor	10	1
Sabre Electric Jigsaw	Universal motor	11	1
Bluesky Optimo 1200 W Vacuum Cleaner	Universal motor	12	1
Phillips 1250 W Vacuum Cleaner	Universal motor	13	1

arc detection with a different sampling frequency, six databases were created by sampling at 25 kHz, 5 kHz, 4 kHz, 3.125 kHz, 2.5 kHz, 2 kHz, respectively. In each of the databases, the training and testing split is the first half part of the raw signal for the training (400 samples for each appliance) and the last part for testing (400 samples for each appliance). The dataset is composed of 5200 observations of arcing and 4800 observations of normal load current resulting in a total of 10400 observations. All data are labeled in such a way so that the proposed algorithm can not only able to identify the arc and normal load currents but also it can specify the load type (i.e., resistive, inductive, switching, dimmer). The load categories, their corresponding labels, and the number of respective training and test samples are depicted in Table 3. The primary database is not normalized. Each observation of data having from 1 to 10 half-cycles of the current wave is converted into an image from which a vector of features is extracted. The 21 descriptors considered to describe the textural characteristics of the images are summarized in Table 1, for a total of 336 features, which is each descriptor for each inter-pixel distances d. Others six statistical features are extracted by the raw AC signal, which are the maximum, the minimum, the mean, the median, the standard deviation and the kurtosis of the AC signal.

4. Experimental Results and Analysis

In this section the experimental analysis of the proposed approach and the comparison with the classical RQ analysis are presented.

4.1. Experimental results

Figure 5 shows the non-stationary trends of the arc voltage and line current signatures of vacuum cleaner with 1200 W nominal power when an arc fault occurs over time and Figure 6 shows the corresponding downsampled current signal and the monitoring of two different GLCM features. In particular, Figure 6(a) shows the feature f17 (information measure of correlation II), whereas Figure 6(b) shows the feature f9 (inverse difference moment normalized). Features are calculated with a window of length 0.1 s, overlap of 0.99 s and decimation to 4 kHz. It is worth to note as the features change the mean value when the arc faults occur.

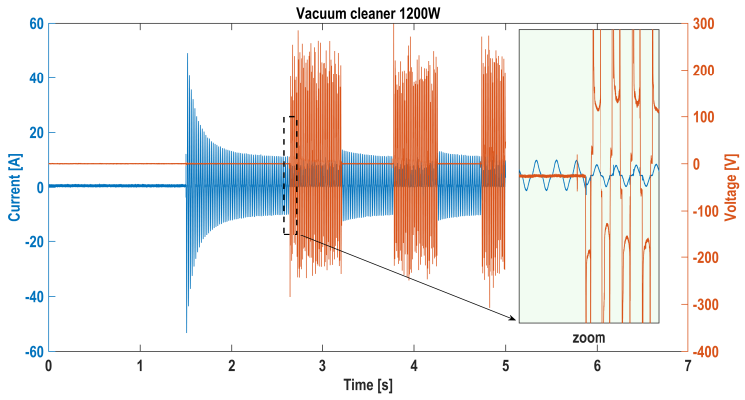
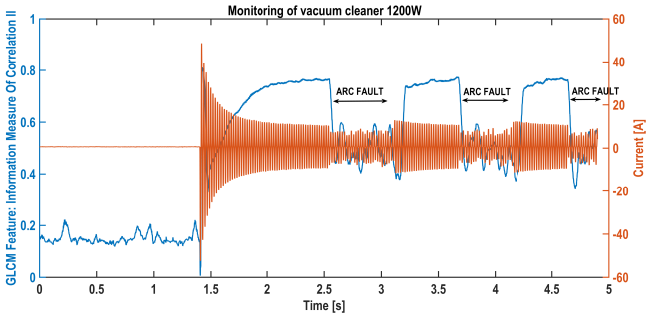
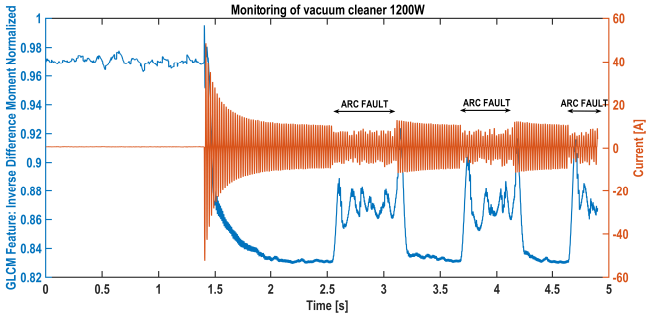


Figure 5: Voltage and current signals of vacuum cleaner 1200 W

A preliminary visual inspection of the similarity matrices for different appliances without arc faults is shown in Figure 7, where unthresholded recurrence plots of blow heater, desktop computer, drill and vacuum cleaner are shown in Figures 7(a), 7(b), 7(c), 7(d), respectively. It is worth to note as unthresholded Recurrence Plots (RPs) of different loads presents a different image pattern



(a) GCLM feature: Information Measure of Correlation II



(b) GCLM feature: Inverse Difference Moment Normalized

Figure 6: GLCM features monitoring for vacuum cleaner 1200 W

highlighting the possibility of using this information to detect arc faults and load type. The corresponding time series of AC signals are shown in Figures 7(e), 7(f), 7(g), 7(h), respectively. The figures are obtained with decimation of 5 kHz and 0.1 s of time acquisition.

The experimental results are summarized in Table 4, where the results are shown in terms of F1 score, a commonly used criterion measuring the performance of a classification method [51]. The tests are reported for different cases varying the decimation factor and length of time window as shown in Table 2. The considered time window duration is $T_W = [0.05, 0.08, 0.1]$ s, whereas the decimation by a factor of $r = [200, 250, 320, 400, 500]$, i.e., $F_d = [5, 4, 3.125, 2.5, 2]$ kHz. In addition, the tests were carried out considering also one complete cycle of the current wave (i.e, 2 half-cycle) at 25 kHz of sampling

frequency.

The first two columns show the false positive and false negative rates related to the arc faults detection (i.e., 2 classes, regardless the identification of the appliance). The third column of the Table 4 shows the detection of arc fault (i.e., 2 classes, regardless the identification of the appliance), the F1 score is 0.98/0.99 for most of considered time window duration and decimation factor. The fourth column of the Table 4 shows the appliances classification during the occurrence of an arc fault (i.e., 13 classes), with a F1 score varying from almost 0.82 to 0.96. The F1 score is always more than 0.89 when the sampling frequency is bigger than 4 kHz. The fifth column of the Table 4 shows the appliances classification without the occurrence of an arc fault (i.e., 13 classes), the F1 score is close to 1.00. The last column of the Table 4 shows the appliances classification with and without the occurrence of an arc fault (i.e., 26 classes), with a F1 score varying from 0.89 to 0.96. The F1 score is always more than 0.94 when the time window is longer than 0.1 s. It is worth noting as the case 25 kHz does not give the best results, this is due to a small number of half-cycle, i.e. 2, highlighting as the method is robust in the case of several half-cycles analyzed.

Table 4: Arc fault detection and appliances classification: F1 score (26 classes)

	False Positive %	False Negative %	ARC vs NORMAL (2 CLASSES)	ARC (13 CLASSES)	NORMAL (13 CLASSES)	ALL APPLIANCES ARC vs NORMAL (26 CLASSES)
0.05 s - 2 kHz	0	0.31	0.99	0.82	0.98	0.89
0.05 s - 2.5 kHz	0.02	0.33	0.99	0.85	0.98	0.91
0.05 s - 3.125 kHz	0	0.63	0.99	0.87	0.98	0.91
0.05 s - 4 kHz	0	1.83	0.98	0.89	0.98	0.92
0.05 s - 5 kHz	0	2.08	0.98	0.90	0.99	0.93
0.08 s - 2 kHz	0	0.10	0.99	0.88	0.99	0.92
0.08 s - 2.5 kHz	0.12	0.13	0.99	0.90	0.99	0.93
0.08 s - 3.125 kHz	0	1.21	0.98	0.92	0.99	0.93
0.08 s - 4 kHz	0	0.85	0.98	0.93	0.99	0.95
0.08 s - 5 kHz	0	2.69	0.97	0.93	0.99	0.94
0.1 s - 2 kHz	0	0.04	0.99	0.89	0.99	0.94
0.1 s - 2.5 kHz	0.63	0.04	0.99	0.92	0.99	0.94
0.1 s - 3.125 kHz	0.83	0.56	0.99	0.93	0.99	0.95
0.1 s - 4 kHz	0.06	0.08	0.99	0.94	0.99	0.96
0.1 s - 5 kHz	0	1.85	0.97	0.96	0.99	0.96
0.02 s - 25 kHz	2.27	3.29	0.97	0.89	0.99	0.93
AVG±STD	**0.25±0.59**	**1.00±1.04**	**0.98±0.01**	**0.90±0.04**	**0.99±0.01**	**0.93±0.02**

The confusion matrix for 4 kHz and 0.1 s of time acquisition is shown in Figure 8. The first 13 labels refer to the appliances with arc faults ordered as listed in Table 3, whereas the labels from 14 to 26 refer to the normal appliances. Here, it is possible to notice only three missed detection (type II error) for fault class 2 (Dell Desktop Computer). The error is highlighted in red color in the confusion matrix. In addition, the classification of the appliances heater and kettle (i.e., resistive loads) in arc fault condition finds to be a difficult task. Considering only one class for appliances heater and kettle (i.e., a total of 24 classes), it is possible to reach better classification results as shown in Table 5. For the sake of simplicity only the case of 4 kHz sampling frequency and 0.1 s of time acquisition is reported and, for this case, F1 score reaches 0.98.

Table 5: Arc fault detection and appliances classification: F1 score (24 classes)

	ARC (12 CLASSES)	NORMAL (12 CLASSES)	ALL APPLIANCES ARC vs NORMAL (24 CLASSES)
0.05 s - 4 kHz	0.92	0.98	0.95
0.08 s - 4 kHz	0.96	0.99	0.97
0.1 s - 4 kHz	0.98	0.99	0.98

The experimental results, summarized in Table 6, shows the F1 score by grouping the appliances according to the load types as defined in Table 3. The first column of the Table 6 shows the load type classification during the occurrence of an arc fault (i.e., 4 classes), with a F1 score varying from almost 0.94 to 0.99. The second column of the Table 6 shows the load type classification without the occurrence of an arc fault (i.e., 4 classes), the F1 score is 0.99 for all cases. The last column of the Table 6 shows the load type classification with and without the occurrence of an arc fault (i.e., 8 classes), with a F1 score varying from 0.96 to 0.99.

The confusion matrix for 4 kHz and 0.1 s of time acquisition of different types of load type is provided in Figure 9. The first 4 labels refer to the loads type with arc faults ordered as listed in Table 3, whereas the labels from 5 to

Table 6: Arc fault detection and load type classification: F1 score (8 classes)

	ARC (4 CLASSES)	NORMAL (4 CLASSES)	ALL LOAD TYPES ARC vs NORMAL (8 CLASSES)
0.05 s - 2 kHz	0.97	0.99	0.98
0.05 s - 2.5 kHz	0.97	0.99	0.98
0.05 s - 3.125 kHz	0.97	0.99	0.97
0.05 s - 4 kHz	0.94	0.99	0.97
0.05 s - 5 kHz	0.97	0.99	0.97
0.08 s - 2 kHz	0.99	0.99	0.99
0.08 s - 2.5 kHz	0.98	0.99	0.99
0.08 s - 3.125 kHz	0.99	0.99	0.99
0.08 s - 4 kHz	0.98	0.99	0.98
0.08 s - 5 kHz	0.98	0.99	0.96
0.1 s - 2 kHz	0.98	0.99	0.99
0.1 s - 2.5 kHz	0.99	0.99	0.99
0.1 s - 3.125 kHz	0.99	0.99	0.98
0.1 s - 4 kHz	0.99	0.99	0.99
0.1 s - 5 kHz	0.99	0.99	0.97
0.02 s - 25 kHz	0.98	0.99	0.97
AVG±STD	**0.98±0.01**	**0.99±0.001**	**0.98±0.01**

8 refer to the normal load types. Here, it is possible to notice only four missed detection (type II error) for fault class 2 (Dell Desktop Computer). The errors are highlighted in red color in the confusion matrix.

4.2. Comparison with standard RQA

Arc fault detection and appliances classification results are carried out by using the RQA metrics as predictors in a Discriminant Analysis classifier. The following results are obtained by a 10-fold cross-validation strategy. In the following results, the RQ parameters m and τ are set by false nearest neighbor algorithm and mutual information whereas ϵ is set as the value that gives the

best results in terms of classification accuracy. In this work, the characteristics of the system have been represented by the following statistical parameters: recurrence rate, determinism, divergence, Shannon entropy, laminarity, and trapping time and the features extracted from RQA have been exploited for arc fault detection. Others six statistical features are extracted by the raw AC signal, which are the maximum, the minimum, the mean, the median, the standard deviation and the kurtosis of the AC signal. The results obtained are shown in Figure 10, where the results are shown in terms of F1 score for the different cases varying the decimation factor and length of time window. Figure 10 shows the appliances classification during the occurrence of an arc fault (i.e., 13 classes), the F1 score never reaches 1.00 for each considered time window duration and decimation factor. The best value 0.86 is obtained in the case of T_W =0.1 s and F_d =5 kHz. RQA based approach achieves the lowest scores across all evaluation metrics. This is unsurprising as the binary transformation of the RQP, consequently, leads to a loss of information that could instead be exploited for fault detection. In addition, the features used for prediction are 12 with respect to 342 features of the proposed approach since only a limited number of features for image analysis are proposed in the literature through recurrence plots [52].

4.3. Comparison with prior methods and discussion

The results on arc fault detection and load classification have been compared with different methods from literature and resumed in Table 7. First of all, it is important to specify that few methods concern both the detection of arc fault and load classification. In [18] and [22], a frequency analysis is associated to a neural network. Furthermore, a genetic algorithm is necessary to optimize the initial value of the network in [22]. The method is nevertheless tested on a limited number of appliances type. In [18], the good results of classification have been obtained with a low number of category and the analysis require both the current and voltage measurement. The method presented in [16], tested on a large variety of different appliances, is based on the current analysis

measured with HCF sensor at a quite high sampling frequency equals to 1 MHz. The method presented in this paper give results of detection and classification comparable to other methods. However, the number of appliances and different types of load is higher. One other advantage is that the method is not based on a neural network which needs an important training process.

Reference	Method	Category	Loads	Remarks	Accuracy
Ref. [18]	Wavelet transform (db3); Fs=200 kHz; Category recognition: shoulder band phase analysis; Arcing: time and frequency indicators + BP neural Network	Resistive Re; Resistive-Inductive RI; RCCF	Resistance, Fluorescent lamp, Halogen lamp, Drill, Air conditioner, Computer, Tungsten lamp	Current and voltage analysis; Intel PC computer	Detection: 99.2%; Classification: 100% (7 appliances); 630 sets of data
Ref. [16]	HCF sensor; Fs=1 MHz; Gray image analysis + convolution neural network	7 load classes (no fault) associated to 7 classes (arc fault)	Electric heater, Vacuum cleaner, Computer, Dimmer Fluorescent lamp, Induction cooker, Drill	Current analysis; Implemented: Zynq-7020	Detection: 99.2%; Classification: 98.36 % (7 appliances); 1400 tests; Training set: 12600; Time response:11 ms
Ref. [22]	Frequency analysis (1st-3rd-5th harmonics) for category recognition and time domain features + fully connected neural network	Resistive RE; Capacitive-inductive CI; Switching SW	Resistive load, Inductive load, Capacitive load, Switching load	STM32F407ZG time response: 3 ms; Fs=20kHz	Detection accuracy: 99% (4 appliances); 3950 tests
Proposed approach	Recurrence quantification and Image analysis (gray level co-occurrence matrix)	Resistive, Motor, Switching Load, Rectfier load	Drill, Expresso machine, Halogen Lamp, Blow heater, Vacuum cleaner, Electric jigsaw, Food processor, Kettle, Printer	Current Analysis; Intel PC computer	Detection 99%; Classification: 98% (12 appliances); 5200 sets of data (Arc); 4800 sets of data (NoArc)

Table 7: Comparison with prior methods

The limitations of the proposed approach are related to the computational time and the misclassification that is often generated by transients. The algorithm computational time is evaluated as the average of 20 Monte Carlo runs. The simulations are reported considering the sampling frequency of 5 kHz, 4 kHz, 3.125 kHz, 2.5 kHz, 2 kHz, and the time window duration of 0.05, 0.08, 0.1

s. The platform used to compute the training time is a laptop with CPU Intel 7700HQ, 16GB RAM, Matlab 2019a. Figure 11 shows the training time versus the sampling frequencies and the time windows. The computational time is not negligible at high sampling frequencies and long time windows. The bottleneck of the computational time is related to the calculation of the GLCM matrix that requires 90% of the computational time shown in the Figure 11. This issue can be reduced by performing fast computation of GLCM matrix with parallel computation [53] or by the adoption of deep learning approaches for analyzing the recurrence images that require longer training computation, but less testing time [16]. The misclassification was analyzed in detail highlighting that most of the errors in the classification are generated by transients as shown in Figure 12. The figure shows the AC signal and the arc alarm index for desktop computer, coffee machine, kettle and printer. The AC signals are analyzed considering a time window of 0.05 s and a sampling frequency of 2 kHz. The red bar means that the proposed algorithm has detected arc fault and the green bar means no arc is detected. False detection samples are generated when the appliance moves from one steady-state to another for the cases of the coffee machine, printer and desktop computer. Then a drawback of the algorithm is related to the detection reliability of arc fault during the transients. This issue could generate false positive and false negative samples.

5. Conclusion

In this paper, a new method for series arc faults detection in a domestic AC network (220V, 50Hz) is presented. The main originality of the proposed method is the use of recurrence quantification plots and image analysis by using the gray-level co-occurrence matrix from which the extracted textural image features represent a proper set of indicators for suitably detecting arc faults. Thirteen different domestic appliances are considered in this article. Series electric arcs are produced on the single-phase AC power line using the carbonized path, which is a common arc situation. The analysis is performed on current line

measured signatures. The experimental results confirm the good performances of the proposed method. Using samples of 26 types of generalized household load currents, the proposed approach achieved F1 score of 0.99 for arc fault detection and F1 score of 0.96 for appliances classification along with arc fault detection. Using samples of 24 type of generalized household load currents the proposed approach achieved F1 score of 0.98 for appliances classification along with arc fault detection. Moreover, a higher number of load types have been considered for the experiment to make it more appropriate for practical applications. The proposed algorithm can not only detect arc fault but also can identify the load types for which the fault has occurred. The deployment of the algorithm and tests on embedded systems are currently the subject of this research.

References

[1] M. Rabla, E. Tisserand, P. Schweitzer, J. Lezama, Arc Fault Analysis and Localisation by Cross-Correlation in 270 V DC, in: 2013 IEEE 59th Holm Conference on Electrical Contacts (Holm 2013), 2013, pp. 1–6.

[2] G. Concas, B. i. Dôme, L. Lhardit, O. Tissot, WHITE PAPER - Residential electrical safety How to insure progress, Tech. rep. (2020).
URL `https://www.feedsnet.org/`

[3] R. Campbell, HOME ELECTRICAL FIRES, Tech. rep. (2019).
URL `https://www.nfpa.org//-/media/Files/News-and-Research/Fire-statistics-and-reports/US-Fire-Problem/Fire-causes/osHomeElectricalFires.pdf`

[4] J. Lezama, P. Schweitzer, E. Tisserand, J.-B. Humbert, S. Weber, P. Joyeux, An embedded system for AC series arc detection by inter-period correlations of current, Electric Power Systems Research 129 (2015) 227 – 234.

[5] J. J. Shea, Identifying causes for certain types of electrically initiated fires in residential circuits, Fire and Materials 35 (1) (2011) 19–42.

[6] F. Z. Y. Wang, S. Zhang, A New Methodology for Identifying Arc Fault by Sparse Representation and Neural Network, IEEE Transactions on Instrumentation and Measurement 67 (2018) 2526–2537.

[7] P. Qi, S. Jovanovic, J. Lezama, P. Schweitzer, Discrete wavelet transform optimal parameters estimation for arc fault detection in low-voltage residential power networks, Electric Power Systems Research 143 (2017) 130 – 139.

[8] P. S. E. Tisserand, J. Lezama, Y. Berviller, Series arcing detection by algebraic derivative of the current, Electric Power Systems Research 119 (2015) 91 – 99.

[9] J. L. S. Jovanovic, A. Chahid, P. Schweitzer, Shunt active power filter-based approach for arc fault detection, Electric Power Systems Research 141 (2016) 11 – 21.

[10] Y. Liu, F. Guo, Z. Ren, P. Wang, Tuan Nghia Nguyen, J. Zheng, X. Zhang, Feature analysis in time-domain and fault diagnosis of series arc fault, in: 2017 IEEE Holm Conference on Electrical Contacts, 2017, pp. 306–311.

[11] E. Calderon-Mendoza, P. Schweitzer, S. Weber, Kalman filter and a fuzzy logic processor for series arcing fault detection in a home electrical network, International Journal of Electrical Power & Energy Systems 107 (2019) 251 – 263.

[12] K. Yang, R. Zhang, S. Chen, F. Zhang, J. Yang, X. Zhang, Series arc fault detection algorithm based on autoregressive bispectrum analysis, Algorithms 8 (4) (2015) 929–950.

[13] G. Artale, A. Cataliotti, V. Cosentino, D. D. Cara, S. Nuccio, G. Tinè, Arc Fault Detection Method Based on CZT Low-Frequency Harmonic Current Analysis, IEEE Transactions on Instrumentation and Measurement 66 (2017) 888–896.

[14] Z. Z. H. Guan, Bowen Wang, S. Bimenyimana, Q. Wang, Arc Fault Current Signal's Power Spectrum Characteristics and Diagnosis Based on Welch Algorithm, International Journal of Engineering Science 2852 (2016).

[15] R. J. G. Bao, X. Gao, Novel Series Arc Fault Detector Using High-Frequency Coupling Analysis and Multi-Indicator Algorithm, IEEE Access 7 (2019) 92161–92170.

[16] P. S. R. Chu, R. Zhang, Series AC Arc Fault Detection Method Based on High-Frequency Coupling Sensor and Convolution Neural Network, Sensors 20 (17) (2020) 4910.

[17] Y. W. L. C. J. Wu, C. S. Hung, Intelligent detection of serial arc fault on low voltage power lines, Journal of Marine Science and Technology 25 (1) (2017) 43–53.

[18] X. Han, D. Li, L. Huang, H. Huang, J. Yang, Y. Zhang, X. Wu, Q. Lu, Series arc fault detection method based on category recognition and artificial neural network, Electronics (Switzerland) 9 (9) (2020) 1–21.

[19] J. Liu, J. Chen, J. Zuo, N. Qu, PSO–SOM Neural Network Algorithm for Series Arc Fault Detection, Advances in Mathematical Physics 2020 (2020) 1–8.

[20] H. D. Vu, E. Calderon, P. Schweitzer, S. Weber, N. Britsch, Multi criteria series arc fault detection based on supervised feature selection, International Journal of Electrical Power & Energy Systems 113 (2019) 23 – 34.

[21] E. Calderon-Mendoza, P. Schweitzer, S. Weber, A double ended AC series arc fault location algorithm for a low-voltage indoor power line using impedance parameters and a neural network, Electric Power Systems Research 165 (2018) 84 – 93.

[22] Y. Wang, F. Zhang, X. Zhang, S. Zhang, Series AC Arc Fault Detection Method Based on Hybrid Time and Frequency Analysis and Fully

Connected Neural Network, IEEE Transactions on Industrial Informatics 15 (12) (2019) 6210–6219.

[23] N. Qu, J. Zuo, J. Chen, Z. Li, Series Arc Fault Detection of Indoor Power Distribution System Based on LVQ-NN and PSO-SVM, IEEE Access 7 (2019) 184020–184028.

[24] P. Bilski, W. Winiecki, Generalized algorithm for the non-intrusive identification of electrical appliances in the household, in: 2017 9th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Vol. 2, 2017, pp. 730–735.

[25] T. Saitoh, T. Osaki, R. Konishi, K. Sugahara, Current Sensor Based Home Appliance and State of Appliance Recognition, SICE Journal of Control, Measurement, and System Integration 3 (2) (2010) 86–93.

[26] J. Z. Moro, L. F. Duarte, E. C. Ferreira, J. A. S. Dias, et al., A home appliance recognition system using the approach of measuring power consumption and power factor on the electrical panel, based on energy meter ICs, Circuits and Systems 4 (3) (2013) 245–251.

[27] D. Yan, Y. Jin, H. Sun, B. Dong, Z. Ye, Z. Li, Y. Yuan, Household appliance recognition through a Bayes classification model, Sustainable Cities and Society 46 (2019) 101393.

[28] L. Matindife, Y. Sun, Z. Wang, A Machine-Learning Based Nonintrusive Smart Home Appliance Status Recognition, Mathematical Problems in Engineering 2020 (2020).

[29] R. Bonfigli, S. Squartini, M. Fagiani, F. Piazza, Unsupervised algorithms for non-intrusive load monitoring: An up-to-date overview, in: 2015 IEEE 15th International Conference on Environment and Electrical Engineering (EEEIC), 2015, pp. 1175–1180.

[30] Y. Zhibin, C. Hong, Research on Household Appliances Recognition Method Based on Data Screening of Deep Learning, IFAC-PapersOnLine 52 (24) (2019) 140–144, 5th IFAC Symposium on Telematics Applications TA 2019.

[31] Y. Hsueh, V. R. Ittangihala, W.-B. Wu, H.-C. Chang, C.-C. Kuo, Condition Monitor System for Rotation Machine by CNN with Recurrence Plot, Energies 12 (17) (2019).

[32] F. Ferracuti, A. Freddi, S. Longhi, A. Monteriù, Recurrence Quantification Analysis of Stator-Current Measurements for Electric Motor Fault Classification, in: IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society, Vol. 1, 2019, pp. 3691–3696.

[33] P. Bhui, N. Senroy, Application of Recurrence Quantification Analysis to Power System Dynamic Studies, IEEE Transactions on Power Systems 31 (1) (2016) 581–591.

[34] A. M. Fraser, H. L. Swinney, Independent coordinates for strange attractors from mutual information, Phys. Rev. A 33 (1986) 1134–1140.

[35] R. J. Povinelli, M. T. Johnson, A. C. Lindgren, F. M. Roberts, Jinjin Ye, Statistical models of reconstructed phase spaces for signal classification, IEEE Transactions on Signal Processing 54 (6) (2006) 2178–2186.

[36] R. J. Povinelli, J. F. Bangura, N. A. O. Demerdash, R. H. Brown, Diagnostics of Bar and End-Ring Connector Breakage Faults in Polyphase Induction Motors through a Novel Dual Track of Time-Series Data Mining and Time-Stepping Coupled FE-State Space Modeling, IEEE Power Engineering Review 22 (2) (2002) 58–59.

[37] F. Takens, Detecting strange attractors in turbulence, in: D. Rand, L.-S. Young (Eds.), Dynamical Systems and Turbulence, Warwick 1980, Springer Berlin Heidelberg, Berlin, Heidelberg, 1981, pp. 366–381.

[38] H. D. I. Abarbanel, Analysis of Observed Chaotic Data, Springer, New York, NY, 1996.

[39] S. Ali, A. Basharat, M. Shah, Chaotic Invariants for Human Action Recognition, in: 2007 IEEE 11th International Conference on Computer Vision, 2007, pp. 1–8.

[40] J. Bangura, R. Povinelli, N. A. O. Demerdash, R. Brown, Diagnostics of eccentricities and bar/end-ring connector breakages in polyphase induction motors through a combination of time-series data mining and time-stepping coupled FE-state-space techniques, IEEE Transactions on Industry Applications 39 (4) (2003) 1005–1013.

[41] C.-Z. Yao, Q.-W. Lin, Recurrence plots analysis of the CNY exchange markets based on phase space reconstruction, The North American Journal of Economics and Finance 42 (2017) 584 – 596.

[42] N. C. Anderson, W. F. Bischof, K. E. W. Laidlaw, E. F. Risko, A. Kingstone, Recurrence quantification analysis of eye movements, Behavior Research Methods 45 (3) (2013) 842–856.

[43] Y. Qian, R. Yan, S. Hu, Bearing Degradation Evaluation Using Recurrence Quantification Analysis and Kalman Filter, IEEE Transactions on Instrumentation and Measurement 63 (11) (2014) 2599–2610.

[44] R. M. Haralick, K. Shanmugam, I. Dinstein, Textural Features for Image Classification, IEEE Transactions on Systems, Man, and Cybernetics SMC-3 (6) (1973) 610–621.

[45] W. Gomez, W. C. A. Pereira, A. F. C. Infantosi, Analysis of Co-Occurrence Texture Statistics as a Function of Gray-Level Quantization for Classifying Breast Ultrasound, IEEE Transactions on Medical Imaging 31 (10) (2012) 1889–1899.

[46] D. A. Clausi, An analysis of co-occurrence texture statistics as a function of grey level quantization, Canadian Journal of Remote Sensing 28 (1) (2002) 45–62.

[47] Y. Guo, T. Hastie, R. Tibshirani, Regularized linear discriminant analysis and its application in microarrays, Biostatistics 8 (1) (2006) 86–100.

[48] P. Schweitzer, F. Ferracuti, Arc fault detection and appliances classification in AC home electrical networks using Recurrence Quantification Plots and Image Analysis [Data set] (2020). doi:10.5281/zenodo.3931688.
URL https://zenodo.org/record/3931688

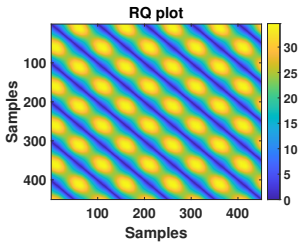
[49] IEC, General requirements for arc fault detection devices IEC62606, International Electrotechnical Commission (2013).

[50] U. L. Inc, UL Standard for Arc-Fault Circuit-Interrupters UL1699, 2nd ed., Underwriters Laboratories Inc. (2011).

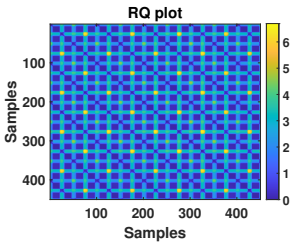
[51] D. M. W. Powers, Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation, International Journal of Machine Learning Technology 2 (1) (2011) 37–63.

[52] A. Facchini, C. Mocenni, A. Vicino, Generalized recurrence plots for the analysis of images from spatially distributed systems, Physica D: Nonlinear Phenomena 238 (2) (2009) 162 – 169.

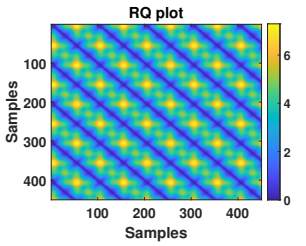
[53] A. Shahbahrami, T. A. Pham, K. Bertels, Parallel implementation of Gray Level Co-occurrence Matrices and Haralick texture features on cell architecture, The Journal of Supercomputing 59 (3) (2012) 1455–1477.



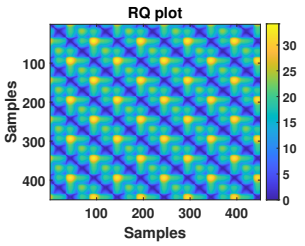
(a) Blow heater



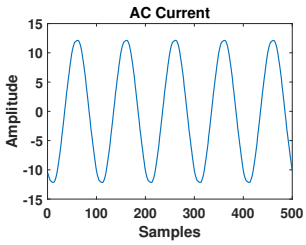
(b) Dell desktop computer



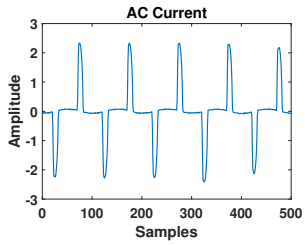
(c) Titan 710 W drill



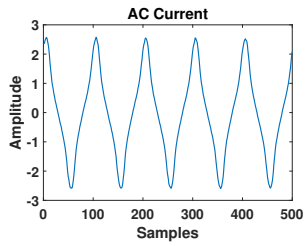
(d) Phillips vacuum cleaner



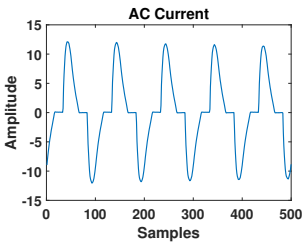
(e) Blow heater



(f) Dell desktop computer



(g) Titan 710 W drill



(h) Phillips vacuum cleaner

Figure 7: Unthresholded RQ plots for a) Blow heater, b) Dell desktop computer, c) Titan 710 W drill, d) Phillips 1250 W Vacuum Cleaner. AC time series for e) Blow heater, f) Dell desktop computer, g) Titan 710 W drill, h) Phillips 1250 W Vacuum Cleaner

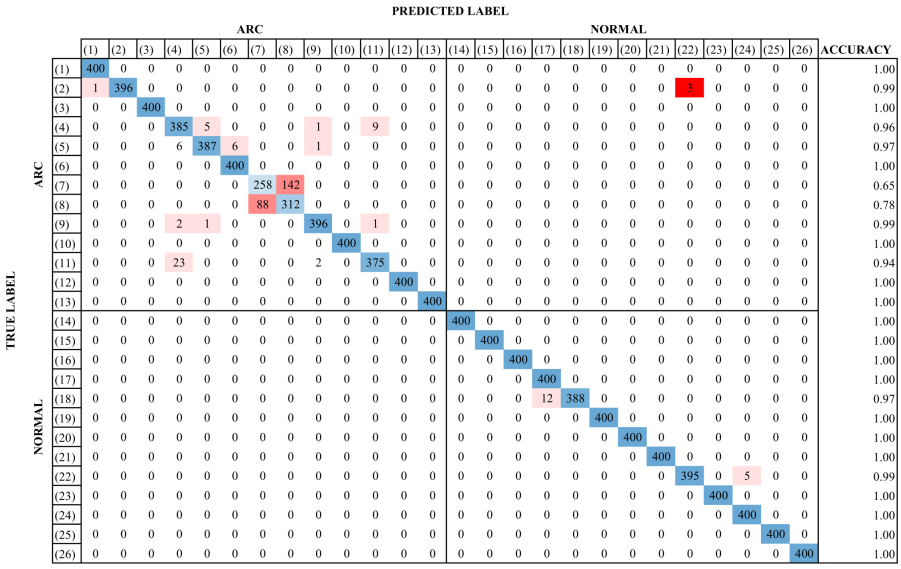


Figure 8: Appliances classification: confusion matrix for 4 kHz and 0.1 s of time acquisition

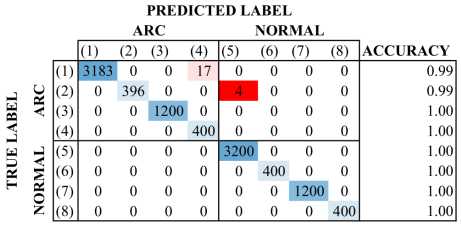


Figure 9: Load type classification: confusion matrix for 4 kHz and 0.1 s of time acquisition

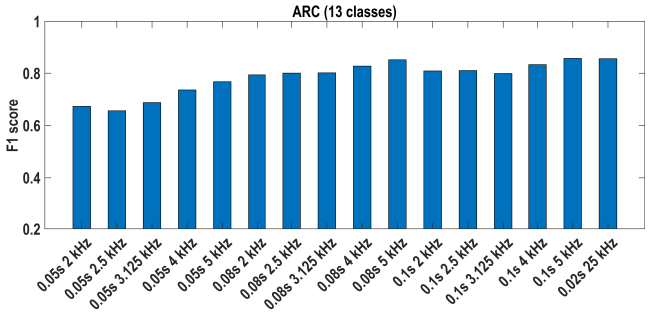


Figure 10: Classification results: F1 scores of the approach based on RQA

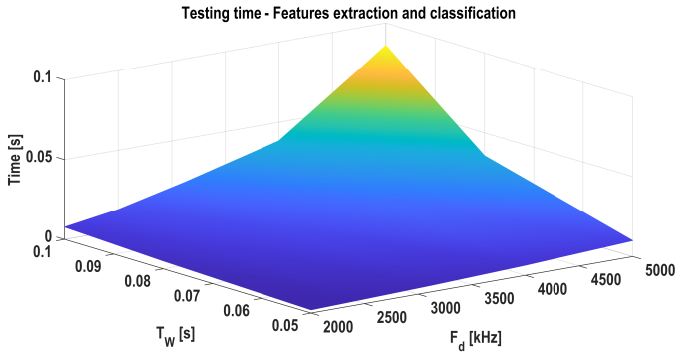


Figure 11: Testing time evaluated at different sampling frequencies and time windows

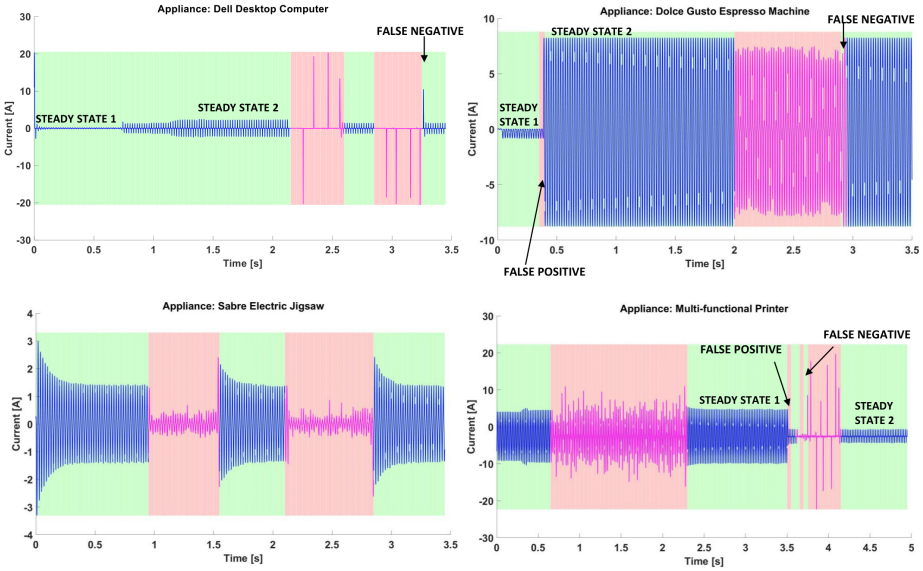


Figure 12: Arc detection