



**HAL**  
open science

# Fault Detection and Diagnosis in a Set “Inverter–Induction Machine” Through Multidimensional Membership Function and Pattern Recognition

Olivier Ondel, Guy Clerc, Emmanuel Boutleux, Eric Blanco

► **To cite this version:**

Olivier Ondel, Guy Clerc, Emmanuel Boutleux, Eric Blanco. Fault Detection and Diagnosis in a Set “Inverter–Induction Machine” Through Multidimensional Membership Function and Pattern Recognition. *IEEE Transactions on Energy Conversion*, 2009, 24 (2), pp.431 - 441. 10.1109/TEC.2008.921559 . hal-00372262

**HAL Id: hal-00372262**

**<https://hal.science/hal-00372262>**

Submitted on 21 May 2009

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Fault Detection and Diagnosis in a Set “Inverter–Induction Machine” Through Multidimensional Membership Function and Pattern Recognition

Olivier Ondel, Guy Clerc, *Member, IEEE*, Emmanuel Boutleux, and Eric Blanco

**Abstract**—Nowadays, electrical drives generally associate inverter and induction machine. Thus, these two elements must be taken into account in order to provide a relevant diagnosis of these electrical systems. In this context, the paper presents a diagnosis method based on a multidimensional function and pattern recognition (PR). Traditional formalism of the PR method has been extended with some improvements such as the automatic choice of the feature space dimension or a “nonexclusive” decision rule based on the  $k$ -nearest neighbors. Thus, we introduce a new membership function, which takes into account the number of nearest neighbors as well as the distance from these neighbors with the sample to be classified. This approach is illustrated on a 5.5 kW inverter-fed asynchronous motor, in order to detect supply and motor faults. In this application, diagnostic features are only extracted from electrical measurements. Experimental results prove the efficiency of our diagnosis method.

**Index Terms**—Data standardization, diagnosis, induction machine, inverter, membership function, nonexclusive decision rule, pattern recognition (PR), reliability index.

## I. INTRODUCTION

THE MONITORING and diagnosis of induction machines have been under focus for at least 20 years [1]–[3].

These motors present numerous advantages due to their robustness and their power–weight ratio. Thus, they are widely used in industry. Most electric motor failures interrupt a process, reduce production, and can damage other related machinery. In some factories, a very expensive scheduled maintenance is performed in order to prevent sudden motor failures. Therefore, there is a great demand to reduce maintenance costs and prevent unscheduled downtimes for electrical drive systems. In the past two decades, there has been research in order to provide new monitoring techniques for ac induction motors based on analyzing vibration signals [4]–[8], or signals other than cur-

Manuscript received June 5, 2007; revised December 12, 2007. First published May 7, 2008; current version published May 19, 2009. Paper no. TEC-00180-2007.

O. Ondel, E. Boutleux, and E. Blanco are with the Ecole Centrale de Lyon, Ecully F-69134, France. They are also with the Centre National de la Recherche Scientifique (CNRS), Unités Mixtes de Recherche (UMR) 5005, Laboratoire AMPERE, Villeurbanne F-69622, France (e-mail: olivier.ondel@ec-lyon.fr; emmanuel.boutleux@ec-lyon.fr; eric.blanco@ec-lyon.fr).

G. Clerc is with the Université de Lyon, Lyon F-69622, France. He is also with the University Claude Bernard Lyon I, Lyon F-69622, France and also with the Centre National de la Recherche Scientifique (CNRS), Unités Mixtes de Recherche (UMR) 5005, Laboratoire AMPERE, Villeurbanne F-69622, France (e-mail: guy.clerc@univ-lyon1.fr).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TEC.2008.921559

rents [9], [10]. However, the vibration sensors are delicate and expensive. Most of the recent research has been directed toward electrical monitoring of the motor, in particular, on inspecting the stator current. At this time, the motor current signature analysis (MCSA) was the first attempt to detect electrical and mechanical faults with the so-called semiinvasive and low-cost sensors [11]–[13].

These methods are rather effective when the motor is supplied by the three-phase main network. However, in more and more industrial applications, the asynchronous motor supply is done by converters. Thus, the currents are affected by the multiple harmonics of the switching frequency. Thus, it becomes very difficult to detect faulty modes by current signature analysis.

To our knowledge, there is no recent statistical study on the static converters breakdowns. It was shown that the components showing the highest failure rate are generally the electrolytic capacitors [14]. This study had been carried out for a dc/dc forward converter. The author had shown that the capacitors were responsible for three-fourth of the breakdowns.

For these reasons and assuming that this study is applicable to others static converters, we limited ourselves, during our application, with the influence of the capacitors failures on the converter itself and also on the possible interactions with the induction machine.

In this context, this paper presents a diagnosis method, used to detect and localize failures in a set inverter–induction motor. This method is based on a pattern recognition (PR) approach. Starting from the traditional formalism of PR, we have developed an automatic supervision system for the diagnosis of process (inverter–induction machine). To improve this technique, we integrated some specificities as the choice of feature dimension space or a new decision rule that allows associating a membership degree to the classification realized.

The aim is to automatically identify the operating conditions (faulty or not) of the system and differentiate if the failing operation is due to the inverter (power supply) or the induction machine.

In Section II, an introduction to diagnosis by PR is done. The PR method is divided into two phases. The first one, named *training phase*, consists in determining the pattern vector (determination of the features sensitive to the faults), the feature space by features selection methods, the decision space (the classes), and in developing a decision rule. The classification varies according to the choice of feature subset and also with the choice

of the number of features, which makes the optimal pattern vector. To determine this number of features, i.e., the feature space dimension, we optimize a criterion based on separability and compactness of the various classes. This technique is developed in Section II and the results are shown in Section VI. The second step, named *decision phase*, consists in associating an unknown pattern with one of the defined classes, according to the decision rule.

Sections III and IV are devoted to the development of a new decision rule being based on a well-known decision rule, the  $k$ -nearest neighbors ( $k$ -nn). Thus, a membership function specifying the membership degree of a new observation to each class is implemented. In order to minimize the errors of classification, we used two reject options (distance and ambiguity), which allow avoiding automatic assigning of an unknown pattern to one of the classes.

For the necessity of the application, in Section V, data standardization according to the value of the first harmonic of each measured signal is developed. It will allow freeing the data from the machine load level. Thus, one operating mode is represented by only one class.

In Section VI, we apply our diagnosis method on an asynchronous motor supplied by an inverter in order to differentiate the appearance of the faults occurring on the inverter from these on the machine.

## II. DIAGNOSIS BY PATTERN RECOGNITION

The aim of statistical PR is to classify objects (patterns) by comparing with reference patterns gathered into classes (clusters) [15]. Such a decision system based upon PR requires an *a priori* knowledge of the studied process in order to define objects and classes. In statistical PR, an object is a set of  $d$  features  $(x_1, x_2, \dots, x_p, \dots, x_d)$  represented as a point in the  $d$ -dimensional real space issued from these features. This space is named feature space. Thus, a pattern  $i$  ( $i = 1$  to  $n$ ,  $n$  being the total number of objects composing the initial database), is characterized by a pattern vector (or a signature)  $\underline{X}_i = [x_{i1}, x_{i2}, \dots, x_{id}]$  belonging to  $\mathfrak{R}^d$ . The classes or clusters  $(\Omega_1, \Omega_2, \dots, \Omega_e, \dots, \Omega_M)$  are subspaces of  $\mathfrak{R}^d$  including similar reference patterns. The principle of recognition is to determine with which class, among the  $M$  known classes, to associate a new observed pattern  $\underline{X}_u = [x_{u1}, x_{u2}, \dots, x_{ud}]$ .

The PR method is made in two phases. The first one, named training phase, consists in determining the feature space (a pattern vector), the decision space (the clusters), and developing a decision rule that produces borders between classes. The second one, named decision phase, consists in associating an unknown pattern with one of the defined clusters, according to the decision rule. This method is more precisely described in [16] and [17]. The accuracy of PR is based on the choice of features contained in this vector. Thus, in the following section, the features used to detect failures on our system are described.

### A. Relevant Signatures for Induction Machine Fault Monitoring

A complete study made by [18] on failures location in a medium-power induction machine has given the following re-

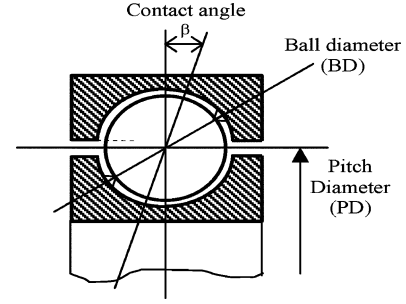


Fig. 1. Bearing sizes marked.

sults: more than 50% of faults affect the bearings; about 16% of faults are localized on the stator winding; finally, 5% and 2.5% of faults affect the rotor and the shaft, respectively.

Currents can provide very rich information on rings, misalignment, and eccentricities or broken bars [19], [20]. The following paragraphs present the main signatures, which can be observed on the current spectrum.

1) *Eccentricities*: The stator currents are affected by the variation of the air gap length. Therefore, they present harmonics at the frequencies given by

$$f_{ecc} = f_s \cdot \left[ (N_r \pm n_d) \left( \frac{1-g}{p} \right) \pm n_w \right] \quad (1)$$

where  $f_s$  is the supply frequency,  $N_r$  is the number of rotor bars,  $p$  is the number of pole pairs, and  $n_w$  is an odd integer. The quantity  $g$  is called per unit slip. The integer  $n_d$  equals zero for the static eccentricity, and one for the dynamic eccentricity.

2) *Bearing Faults*: The current harmonics due to the bearing faults are given in [19] and [21]

$$f_{bng} = f_s \pm N_b \cdot f_{i,o} \quad (2)$$

where  $N_b$  is the number of balls. The frequencies  $f_o$  and  $f_i$  (3) are related to the inner and outer race, respectively,

$$f_o = \frac{f_r}{2} \left[ 1 - \frac{BD}{PD} \cos \beta \right] \quad \text{and} \quad f_i = \frac{f_r}{2} \left[ 1 + \frac{BD}{PD} \cos \beta \right] \quad (3)$$

where BD and PD are, respectively, the ball diameter and the pitch diameter, and  $\beta$  is the contact angle between the balls and the ball bearing rings (see Fig. 1).

3) *Rotor Faults*: Broken bars and end-rings induce the same harmonic in the stator currents. In presence of these faults, the stator currents are modulated in amplitude by the slip frequency ( $2gf_s$ ).

This modulation increases according to the severity of the fault. In the current spectrum, the harmonics due to the fault can be found at the following characteristic frequencies:

$$f_b = f_s \cdot (1 \pm 2 \cdot n \cdot g). \quad (4)$$

4) *Stator Winding Faults*: Stator winding failures are dangerous damages. By using a three-phase models and experimental studies, it has been shown by [22] and [23] that turn-to-turn short circuits create electrical unbalance on the currents. This asymmetry was quantified through the inverse impedance by [24].

Throughout this synthesis, it appears that most of the faults produce different sensitive frequency harmonics. Thus, 25 harmonics are computed and integrated in the feature vector.

However, some faults (stator faults) require more investigations. Some other features are based on:

- 1) the Park's transformation of currents  $I_{s\alpha}$ ,  $I_{s\beta}$  and voltages  $V_{s\alpha}$ ,  $V_{s\beta}$ ;
- 2) the estimation of the positive sequence components of the line voltages and currents.

a) *Park vector components*: Currents and the voltages modulus,  $I_{s\alpha,\beta}$  and  $V_{s\alpha,\beta}$ , can be expressed by

$$\begin{cases} I_{s\alpha,\beta} = \left\| \overrightarrow{I_{s\alpha,\beta}} \right\| = \sqrt{I_{s\alpha}^2 + I_{s\beta}^2} \\ V_{s\alpha,\beta} = \left\| \overrightarrow{V_{s\alpha,\beta}} \right\| = \sqrt{V_{s\alpha}^2 + V_{s\beta}^2}. \end{cases} \quad (5)$$

The active and reactive powers ( $P$  and  $Q$  respectively) are directly calculated with these currents and voltages [25]:

$$\begin{aligned} P &= V_{s\alpha} \cdot I_{s\alpha} + V_{s\beta} \cdot I_{s\beta} \\ Q &= V_{s\beta} \cdot I_{s\alpha} - V_{s\alpha} \cdot I_{s\beta}. \end{aligned} \quad (6)$$

Reactive and active powers are normalized by the rms value of apparent power modulus  $S = \sqrt{P^2 + Q^2}$  and are denoted  $Q'$  and  $P'$ .

Three features are extracted from these analyses.

- 1) The variation of the apparent power called  $\zeta$ .
- 2) The mean values of  $P'$  and  $Q'$ , called  $m_p$  and  $m_q$ .

Let  $\mathbf{m}_{p,q}$  be the center of the characteristic  $[P'Q']$  and  $s_k$  one point of the set, respectively,

$$s_k = [P'(k) \quad Q'(k)] \quad (7)$$

$$\mathbf{m}_{p,q} = [m_p \quad m_q]. \quad (8)$$

Thus, the variation of the apparent power is the scalar defined by

$$\zeta = \sum_{k=1}^{N_p} (\underline{s}_k - \underline{\mathbf{m}}_{p,q})(\underline{s}_k - \underline{\mathbf{m}}_{p,q})^T \quad (9)$$

where  $N_p$  is the number of points in the signals.

Five more features are calculated from the currents.

- 1) The peak-to-peak values of  $I_{s\alpha}$  and  $I_{s\beta}$ :

$$\begin{aligned} \delta_\alpha &= |\max(I_{s\alpha}) - \min(I_{s\alpha})| \text{ and } \delta_\beta \\ &= |\max(I_{s\beta}) - \min(I_{s\beta})|. \end{aligned} \quad (10)$$

- 2) The standard deviations of  $I_{s\alpha}$ ,  $I_{s\beta}$ , and the modulus  $I_{s\alpha,\beta}$ , called  $\sigma_\alpha$ ,  $\sigma_\beta$ , and  $\sigma_s$ , respectively.

These are the eight first features calculated for the pattern vector building.

b) *Positive sequence components*: Other features are calculated from a unique expression of the three-phase voltages and currents: their positive sequence components. The complex positive sequence components for the currents and voltages are

$$\underline{I}_1 = \frac{1}{3} \cdot (I_{sa} + \underline{a}I_{sb} + \underline{a}^2I_{sc}) \quad (11)$$

$$\underline{V}_1 = \frac{1}{3} \cdot (V_{sa} + \underline{a}V_{sb} + \underline{a}^2V_{sc}), \quad \text{with } \underline{a} = e^{j(2\pi/3)}.$$

Furthermore, two other features are defined with the help of the positive sequence approach.

- 1) The mean power of the current's ( $P_1$ ):

$$P_1 = \frac{1}{N_p} \sum_{k=1}^{N_p} \|\underline{I}_1(k)\|^2. \quad (12)$$

- 2) The positive sequence impedance ( $Z_1$ ) at the supply frequency [24] and [26]:

$$Z_1 = \frac{\|\underline{V}_1\|}{\|\underline{I}_1\|} \quad f = f_s = 50 \text{ Hz}. \quad (13)$$

Finally, 35 features are determined. This set represents a list of indicator of multiple faults (electrical and mechanical faults). They are summarized in [16].

Unfortunately, either some of these features are not really relevant, or they are correlated. Furthermore, the high number of features involves an increasing of CPU time during the classification.

To correct these drawbacks and to obtain the appropriate signature for the induction motor, we must keep only the most representative features for the studied faults [27]–[31].

## B. Reducing the Feature Space

To reduce the number of features without losing any information, selection methods must be used. The objective is to seek, among  $d$  initial features, a subset of  $d'$  features giving most of the information. The features subset will have to maximize a criterion taking into account:

- 1) a weak dispersion of the points in each class, which suggests minimal intraclass dispersion;
- 2) a maximum distance between the various classes; therefore, maximum interclasses dispersion.

The criterion is based on within-class scatter  $S_W$  and the between-class scatter  $S_B$  matrices whose expressions are as follows, respectively,

$$S_W = \frac{1}{n} \sum_{c=1}^M \sum_{v=1}^{n_c} (\underline{X}_{cv} - \underline{m}_c)(\underline{X}_{cv} - \underline{m}_c)^T \quad (14)$$

$$S_B = \frac{1}{n} \sum_{c=1}^M (\underline{m}_c - \underline{m})(\underline{m}_c - \underline{m})^T \quad (15)$$

where  $M$  is the number of classes,  $n_c$  number of samples in the class  $\Omega_c$ ,  $n$  the total number of samples,  $\underline{m}_c$  the gravity center of class  $\Omega_c$ ,  $\underline{X}_{cv}$  the  $v$ th sample of class  $\Omega_c$ , and  $\underline{m}$  the general gravity center of the initial training set.

The objective is to define a new space maximizing the interclasses dispersion and maintaining intraclass dispersion. The eigenvectors of  $S_W^{-1}S_B$  can be used, with  $(n - M)$  and  $M$  higher than  $d$ , in order to define the features selection criterium:

$$J_1 = \text{trace}(S_W^{-1} \cdot S_B). \quad (16)$$

The Fisher ratio can also be applied as criterion, for the features selection. The idea is successively to consider each feature contained in the initial pattern vector (dimension  $d$ ) and to calculate the criterion  $J_2$  for this feature.

For a problem with  $M$  classes, the criterion is expressed, for the feature  $\alpha$ , as

$$J_2(\alpha) = \sum_{c=1}^M \sum_{r=1}^{M-1} \frac{m_c(\alpha) - m_r(\alpha)}{n_c \sigma_c^2(\alpha) - n_r \sigma_r^2(\alpha)} \quad (17)$$

with  $m_c(\alpha)$ , gravity center of the class  $\Omega_c$  by considering only the feature  $\alpha$

$$m_c(\alpha) = \frac{1}{n_c} \sum_{v=1}^{n_c} X_{cv}(\alpha) \quad (18)$$

with  $X_{cv}$  the  $v$ th sample of the class  $\Omega_c$ ,  $n_c$  the number of samples in the class  $\Omega_c$ , and  $\sigma_c(\alpha)$  the variance of the component  $\alpha$  of the sample of the class  $\Omega_c$ :

$$\sigma_c^2(\alpha) = \frac{1}{n_c} \sum_{v=1}^{n_c} [X_{cv}(\alpha) - m_c(\alpha)]^2. \quad (19)$$

The numerator of expression (17) represents the separation of the classes (interclasses dispersion). The denominator reflects the compactness of the classes (by analogy with the intraclass dispersion).

In the experimental part (Section VI), the Fisher ratio (17) is used to define the capacity of discrimination of each feature compared to the studied training set and the expression  $J_1$  (16) is used as criterion in the feature selection method. It will be noted as  $J$ .

The maximization of the criterion gives the optimal features subset. Let us note  $V = (\alpha_1, \alpha_2, \dots, \alpha_j, \dots, \alpha_d)$  the initial pattern vector, composed of  $d$  features. Thus, it is possible to seek a subset  $V'$  composed of  $d'$  relevant features so that

$$J(V') = \text{Max}_{h=1, \dots, n_p} [J(V_{d'}^{(h)})] \quad (20)$$

with  $n_p$  the number of studied combinations (of  $d'$  features among  $d$ ) and  $V_{d'}^{(h)}$  the  $h$ th combination of  $d'$  features for which the selective function value is calculated.

In [32], the authors present the advantages and drawbacks of the selection methods. According to the author, the sequential backward selection (SBS) is one of the most simple feature selection methods. This method consists in deleting, at each step, the feature that penalizes the aforementioned criterion. This method requires knowing the partitioning of patterns contained in the initial training set in various clusters. The SBS algorithm is based on the following process.

At step  $k$ , we calculate the criterion for the subsets  $V_{k,j}$  of dimension  $d - k$ . These subspaces result from the subspace  $V_{k-1}$  by eliminating the feature  $\alpha_j$

$$V_{k-1} \supset V_{k,j} = \{V_{k-1} - \alpha_j\}. \quad (21)$$

Thus, the subspace  $V_k$  that is taken into account for the following step optimizes the criterion:

$$J(V_k) = \text{Max}_{j=1, d-k} [J(V_{k,j})]. \quad (22)$$

Fig. 2 illustrates this method with  $d = 5$  and  $d' = 2$ . At each stage, we select the features subset for which the criterion is maximum. The numbers represent the position of the features in the initial pattern vector. The selected features are  $\alpha_1$  and  $\alpha_4$ .

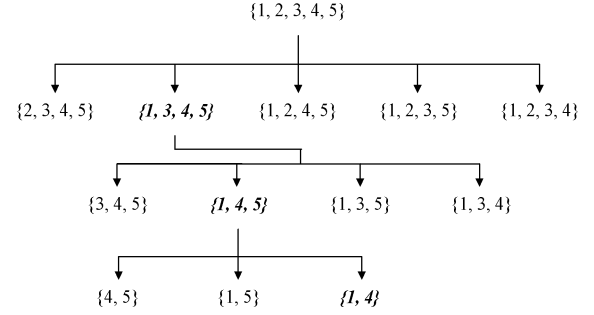


Fig. 2. General diagram of the features selection by method SBS, with  $d = 5$  and  $d' = 2$ .

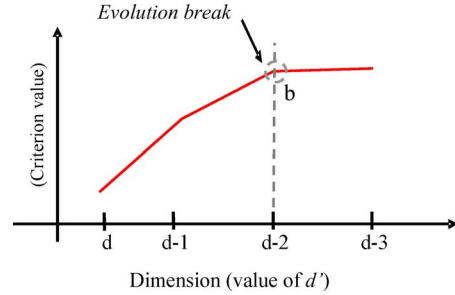


Fig. 3. Choice of the feature space dimension.

Thereafter, the classification of a new observation, i.e., the diagnosis of the system state, strongly depends on the selected features and can also strongly depend on the feature space dimension. The ideal would be to simultaneously determine the best features and also the best dimension. For that, the criterion evolution used in SBS method according to the different values of  $d'$  is looked at.

Fig. 3 shows an example of the criterion evolution. It allows determining the feature space dimension. Indeed, starting from a given dimension, the addition of a feature does not bring any more information. Clearly, on Fig. 3, “b” appears as the best dimension because the slope of the criterion evolution became almost equal to zero.

### C. Decision Phase: “ $k$ -Nearest Neighbors” ( $k$ -nn) Rule

The feature space being now defined (the choice of relevant features and the definition of clusters), it is necessary to decide in which cluster to assign a new observation  $\underline{X}_u$ , collected at a given moment on the system. The decision rule developed is based on the  $k$ -nearest neighbors ( $k$ -nn) rule. It is easy to implement and offers good results of classification. The principle is to assign a new observation  $\underline{X}_u$  in the best-represented class among its  $k$ -nearest neighbors. The simplest way to work out this rule is to measure the Euclidean distance  $d_E$  (23) between the new observation  $\underline{X}_u$  and each sample of the training set  $\underline{X}_i$  ( $\underline{X}_1, \underline{X}_2, \dots, \underline{X}_n$ ) (with  $i = 1$  to  $n$ ,  $n$  being the total number of samples composing the training set):

$$d_E(\underline{X}_u, \underline{X}_i) = \left[ (\underline{X}_u - \underline{X}_i) \cdot (\underline{X}_u - \underline{X}_i)^T \right]^{1/2}. \quad (23)$$

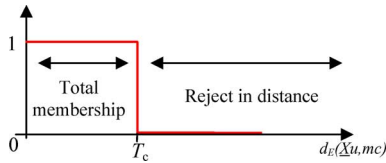
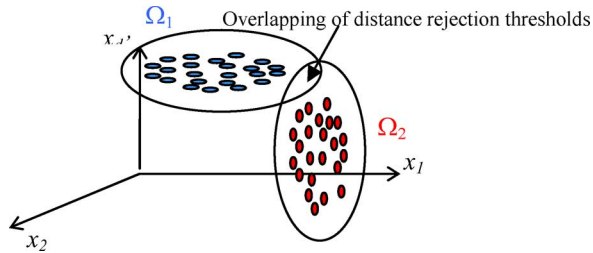

 Fig. 4. Classification according to the distance by the traditional  $k$ -nn rule.


Fig. 5. Distance rejection threshold from not very distant classes.

In order to reduce the error of classification, two rejects options [33], which allow avoiding the systematic classification of unknown patterns, are defined.

The ambiguity reject threshold can be related to the value of  $k$  [34], [35] by fixing a minimal number of neighbors ( $k'$ ) for observation  $\underline{X}_u$

$$\begin{cases} \underline{X}_u \rightarrow \Omega_o, & \text{if } k_c = \max_{r=1, M} k_r < k' \\ \underline{X}_u \rightarrow \Omega_c, & \text{if } k_c = \max_{r=1, M} k_r \geq k' \end{cases} \quad (24)$$

where  $k_c$  is the number of neighbors of  $\underline{X}_u$ , which are found in the class  $\Omega_c$ . This minimal value can be defined especially for each class [36].

The distance reject rule can be established by comparing the distance between the new sample and its assigned class with a threshold  $T_c$ . In this study, this threshold is

$$T_c = 2 \cdot \max_{v=1, n_c} [d_E(\underline{X}_{c_v}, \underline{m}_c)] \quad (25)$$

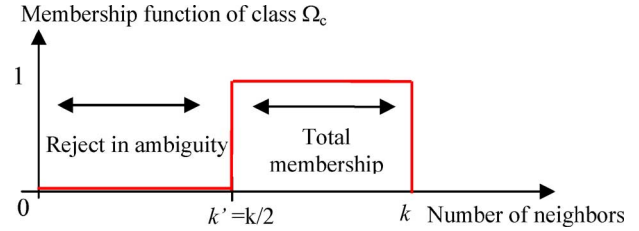
where  $\underline{m}_c$ ,  $\underline{X}_{c_v}$ , and  $n_c$  are, respectively, the center of gravity, the  $v$ st sample, and the total number of samples of the class  $\Omega_c$ .

This traditional  $k$ -nn rule (exclusive) assigns a new observation to the class mainly represented among its nearest neighbors. This can be summarized in the following way (cf. Fig. 4).

However, two or several classes contained in our initial training set  $X_a$  may not be sufficiently distant. Their distance rejection thresholds can intersect as Fig. 5 shows it.

A new observation  $\underline{X}_u$  located in the membership area of two classes (cf. Fig. 5) will be assigned to the class mainly represented among its nearest neighbors. This can be summarized in the following way (cf. Fig. 6).

Using an exclusive classification (0 or 1) does not allow defining a membership degree of the new observation in two (or several) classes. Thus, we will consider that  $\underline{X}_u$  belongs completely to one class. However, when a new observation is located


 Fig. 6. Classification according to the number of neighbors by the traditional  $k$ -nn rule.

between two or several classes, it is often due to the appearance of an unknown operating mode of  $X_a$ .

For these reasons, we create a “nonexclusive” decision rule. Thus, the membership to a class could be modulated and the proximity or distance concepts quantified.

### III. MEMBERSHIP FUNCTION

In order to take into account the uncertainty of classification, a membership function, which characterizes the degree of compatibility between the observation  $\underline{X}_u$  and the class  $\Omega_c$  ( $c = 1$  with  $M$ ) must be introduced [37], [38]. This function is noted as  $\mu_{\Omega}(\underline{X}_u)$ . Compared to the traditional set theory, where the membership function, named characteristic function, can take only values 0 or 1,  $\mu_{\Omega}(\underline{X}_u)$  introduces the concept of gradual membership since it is a continuous function with values in the interval  $[0, 1]$ . This gradation of the membership allows defining classes where the borders are not so stiff.

There is not a general method for building the membership function associated with a given set. The pattern of this function is closely related to the context.

In our case, we must determine a function that applies the following.

- 1) A total membership when the new observation  $\underline{X}_u$  is located between the center of gravity  $\underline{m}_c$  of the class  $\Omega_c$  and the sample, which is the most far away from this center.
- 2) A soft decrease when  $\underline{X}_u$  is located between the previous limit and the rejection threshold  $T_c$  (cf. Fig. 7).

To fulfill these constraints, the adopted solution defines a membership function using a multidimensional function  $\pi$ . This membership function is described by the following formulation:

$$\begin{aligned} & \text{if } 0 \leq d(\underline{X}_u, \underline{m}_c) \leq d_{s_c} \\ & \mu_{\Omega_c}(\underline{X}_u) = \left[ 1 - (1 - \eta) \left( \frac{d(\underline{X}_u, \underline{m}_c)}{d_{s_c}} \right)^g \right] \\ & \text{if } d_{\text{seuil}_c} \leq d(\underline{X}_u, \underline{m}_c) \leq \beta \cdot d_{\text{seuil}_c} \\ & \mu_{\Omega_c}(\underline{X}_u) = \eta \left( \frac{\beta}{\beta - 1} \right)^g \left( \frac{\beta \cdot d_{s_c} - d(\underline{X}_u, \underline{m}_c)}{\beta \cdot d_{s_c}} \right)^g \\ & 0, \quad \text{otherwise.} \end{aligned} \quad (26)$$

- 1) The term  $d(\underline{X}_u, \underline{m}_c)$  is the distance between an observation  $\underline{X}_u$  and the center of gravity  $\underline{m}_c$  of the class  $\Omega_c$ .
- 2) The term  $d_{s_c}$  is the parameter that defines the position of the inflexion points. It corresponds to the maximum

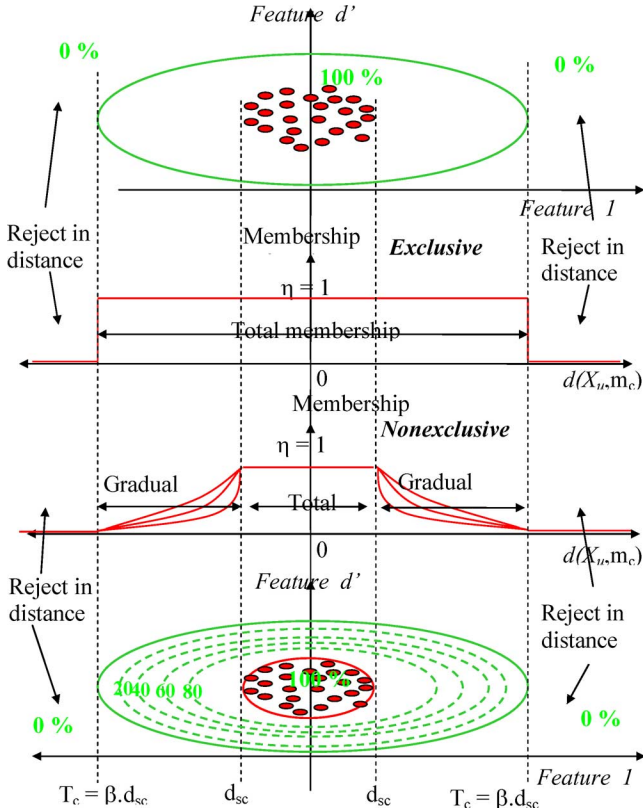


Fig. 7. Schematic representation of the membership function in the exclusive and nonexclusive case.

distance between the center of gravity of the class  $\Omega_c$  and the sample further from this same class  $d_{s_c} = \max_{v=1, n_c} [d(\underline{X}_{cv}, \underline{m}_c)]$ .

- 3)  $\mu_{\Omega_i}(\underline{X}_u) = \eta$  if  $d(\underline{X}_u, \underline{m}_c) = d_{s_c}$ .
- 4) The variable  $\eta$  is the parameter being used to fix the constant part of the  $\pi$  function.
- 5) The variable  $\beta$  is a parameter ( $\beta \in \mathbb{R}^{+*} > 1$ ) that allows defining the bandwidth of the  $\pi$  function (stop points of the  $\pi$  function).
- 6) The variable  $g$  is a parameter allowing to adjust the slope of the function.

With this  $\pi$  function, we take into account the position of an observation in relation with the classes defined in the training set. It brings information about the level of the gradual membership. This estimation decreases as we move away from the center of gravity of the classes. It allows obtaining a classification with a certain degree of membership.

Moreover, it would be worth, for a new observation, to take into account not only the distance with different center of gravity from each class but also the number of nearest neighbors within the classes.

For example, if we suppose that  $k$  nearest neighbors are all resulting from the same class, the degree determined by the membership function must be equal to the maximum 1 if  $d(\underline{X}_u, \underline{m}_c) \leq d_{s_c}$  and ranging between  $[0,1]$  if  $d_{s_c} \leq$

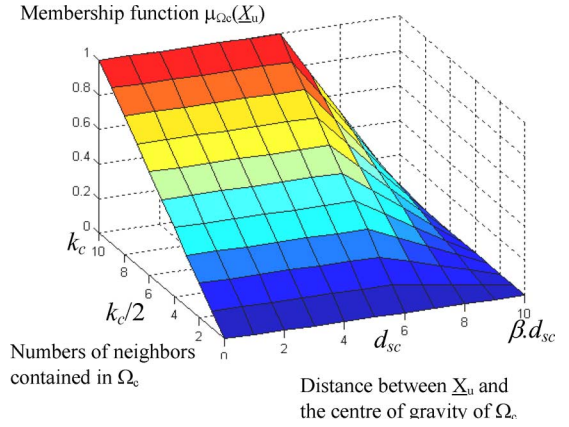


Fig. 8. Membership function of  $\underline{X}_u$  according to the distance and the nearest neighbors.

$d(\underline{X}_u, \underline{m}_c) \leq \beta \cdot d_{s_c}$ . Equation (26) then becomes

if  $0 \leq d(\underline{X}_u, \underline{m}_c) \leq d_{s_c}$

$$\mu_{\Omega_c}(\underline{X}_u) = \left[ \frac{1}{N_c} - \left( \frac{1}{N_c} - \eta \right) \left( \frac{d(\underline{X}_u, \underline{m}_c)}{d_{s_c}} \right)^g \right] \cdot k_c$$

if  $d_{seuil_c} \leq d(\underline{X}_u, \underline{m}_c) \leq \beta \cdot d_{seuil_c}$

$$\mu_{\Omega_c}(\underline{X}_u) = \eta \left( \frac{\beta}{\beta - 1} \right)^g \left( \frac{\beta \cdot d_{s_c} - d(\underline{X}_u, \underline{m}_c)}{\beta \cdot d_{s_c}} \right)^g \cdot k_c$$

0, otherwise (27)

where  $n_c$  is the number of samples contained in the class  $\Omega_c$  and  $k_c$  is the number of  $k$  nearest neighbors resulting from the class  $\Omega_c$ .

Fig. 8 represents the evolution of the membership function of  $\underline{X}_u$  at the class  $\Omega_c$  according to the distance between  $\underline{X}_u$  and the center of gravity of this class and of the number of neighbors among  $k$  nearest neighbors contained in this class.

In contrary to fuzzy methods, the sum of the membership functions of a new observation to the classes defined in  $X_u$ ,  $\mu_{\Omega_c}(\underline{X}_u)$  ( $c = 1$  to  $M$ ) is not necessarily equal to 1. Indeed, only the classes having samples among the nearest neighbors are taken into account allowing reducing the computing times. Thus, the membership degree to the other classes is null. The membership degree of  $\underline{X}_u$  varies within  $[0, 1]$  for each selected classes.

#### IV. EXTENSION OF THE $k$ -nn RULE BY ASSOCIATION OF A NONEXCLUSIVE MEMBERSHIP FUNCTION

The new membership function is associated to the traditional  $k$ -nn rule. Now, the decision problem must be considered in terms of nonexclusive membership to the classes. The membership function values  $\mu_{\Omega_c}(\underline{X}_u)$  relating to the various classes  $\Omega_c$  ( $c = 1, M$ ) can be used for this purpose. Moreover, to decrease the risk of bad classification, it is necessary to define new rules of nonexclusive decision integrating the various reject options (ambiguity and distance).

### A. Membership Reject

A sample being located far from a class must have a small membership degree to this class. The concept of distance rejection can be expressed in term of membership.

An observation is rejected in membership if its membership to all the possible classes is lower than a membership threshold noted as  $T_{ra}$ :

$$\underline{X}_u \in \Omega_d, \quad \text{if } \max_{c=1,M} (\mu_{\Omega_c}(\underline{X}_u)) \leq T_{ra}. \quad (28)$$

### B. Ambiguity Reject

This rejection is carried out only if the  $k$  nearest neighbors results from at least two different classes. Thus, a new observation is rejected in ambiguity.

If the difference between the membership degrees of the physically closest classes (i.e., containing closer neighbors) is lower or equal to an ambiguity threshold  $T_{ne}$  (for example, for  $T_{ne} = 0.05$ , if  $\underline{X}_u$  has a membership of 0.25 to the class  $\Omega_c$  and a membership of 0.22 to the class  $\Omega_r$ ), then this observation is rejected because  $0.25 - 0.22 < T_{ne}$ .

The equations governing this rejection are

If  $\mu_{\Omega_c}(\underline{X}_u) \neq 0$ , then

$$\underline{X}_u \in \Omega_o, \quad \text{if } \mu_{\Omega_c}(\underline{X}_u) - \mu_{\Omega_r}(\underline{X}_u) \leq T_{ne}. \quad (29)$$

Finally, the nonexclusive  $k$ -nn rule including the two reject options can be formulated in the following way.

- 1)  $\underline{X}_u \in \Omega_c$ , if  $\mu_{\Omega_c}(\underline{X}_u) = \text{Max}_{r=1,M} (\mu_{\Omega_r}(\underline{X}_u))$  and  $\mu_{\Omega_c}(\underline{X}_u) \geq T_{ra}$  and (29) is right.
- 2)  $\underline{X}_u \in \Omega_d$ , if  $\max_{c=1,M} (\mu_{\Omega_c}(\underline{X}_u)) \leq T_{ra}$ .
- 3)  $\underline{X}_u \in \Omega_0$ , if  $\mu_{\Omega_c}(\underline{X}_u) - \mu_{\Omega_r}(\underline{X}_u) \leq T_{ne}$ .

The gradation of the membership function allows quantifying the reliability of classification and following the evolution of the new observations in the feature space.

The operating modes can depend on several variables that naturally vary. For instance, the representation of an operating mode can depend of the load level. Thus, if some of these variables are used to represent these modes in the feature space, the same operating mode can be represented by several classes. So, a new observation located between two classes can correspond in an intermediate state of this mode (variation of load level). By using the nonexclusive decision rule with the reject options, the observation can be rejected. It is thus necessary to take into account the evolution of various operating modes from the system due to the load variation. This is developed in the following section.

## V. DATA STANDARDIZATION

To make the decision phase easier and more effective, we standardized the raw data. By standardizing operating modes,

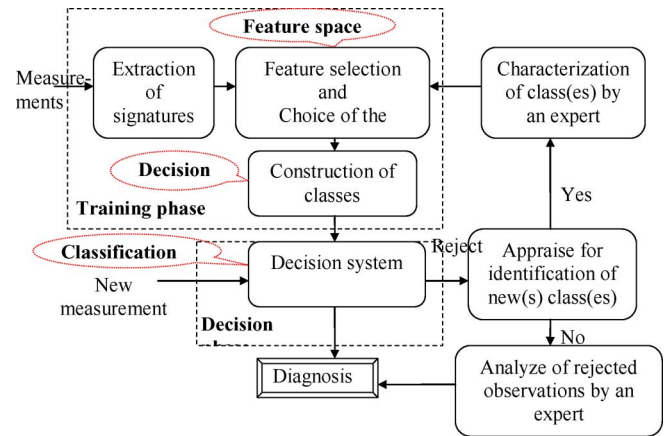


Fig. 9. General diagram of the diagnosis procedure by PR.

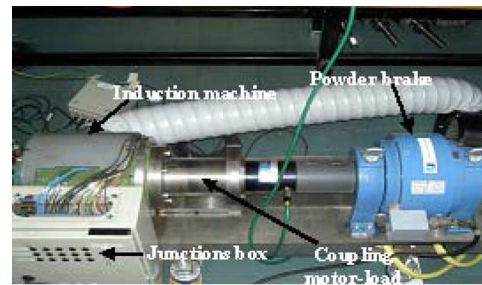


Fig. 10. Motor 5.5 kW coupled with load (powder brake).

i.e., by freeing them from their level of load, they will be represented by only one class, gathering all the load variations. For that, the data are standardized starting from the value of the fundamental amplitude of each measured signal. This standardization allows avoiding classification errors, and thus, improving the diagnosis. Indeed, the training set cannot be exhaustive. If we consider an operating mode, it is impossible to have measurements for all the levels of load. So without standardization, when a new measurement is carried out on the system for this operating mode and the level of load is not defined in  $X_a$ , then this measurement is rejected. The advantage of standardization becomes very clear here. For any level of load, this operating mode is just represented by only one class, new measurement will be affected and not rejected.

Finally, the various phases making a PR system, the training and the decision phases, can be represented in Fig. 9.

## VI. APPLICATION TO THE CLASSIFICATION

A specific experimental setup has been designed in order to perform the methodology of rotor and stator faults detection (cf. Fig. 10). Three voltage sensors and three current sensors are used to monitor the set inverter–induction machine operation. A powder brake, which can be tuned by means of a control unit has been used to simulate the shaft load.

In order to test the validity of the new decision rule, we used the measurements carried out on a three-phase asynchronous squirrel-cage motor 5.5 kW, 11.4 A, four poles, and 28 bars with the rotor. The machine is supplied through a converter



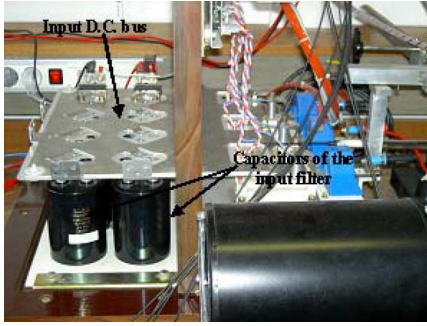


Fig. 11. Inverter used for the experiment.

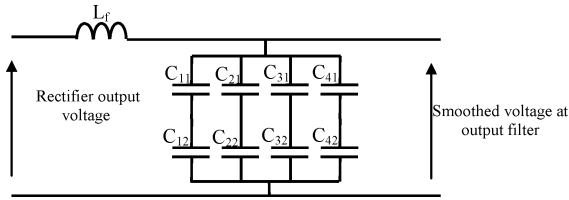


Fig. 12. Input filter of the inverter.

(three-phase inverter). The wide spectrum of the signal makes the detection of the faulty modes more difficult. The experiments are realized for open-loop drives. The control of inverter is a symmetric vector pulse width modulation (PWM).

For each functional state, 20 acquisitions of 5 s, at 20 kHz, have been made in order to check the signature robustness (10+5 for the training step and test set, 5 for the validation). The total number of operating conditions is 4: one for the healthy machine, one for the rotor having three broken bars, one for healthy inverter, and one for faulty inverter (ageing of capacitors). The rotor bars were broken by boring the conductors of the squirrel cage.

The ageing of the capacitors constituting the input filter  $LC$  is characterized by the increase of the equivalent series resistance (ESR) [40]. This increase is due to internal heating of capacitors.

To carry out a failure, we removed part of these capacitors. This step simulates an increase of the ESR and causes the increase of the input voltage ripple. Fig. 11 shows the inverter and the capacitors of the input filter.

The input filter voltage is obtained from a full-wave rectifier. This rectifier is supplied by the three-phase main network. The input filter consists of a smoothing inductor of 100 mH and eight capacitors assembled in series-parallel. The characteristics of these capacitors are 3300  $\mu\text{F}$ , 400 V, 10.1 A, and 85  $^{\circ}\text{C}$ . Fig. 12 schematically represents the input filter of the inverter.

The ESR of each capacitor was characterized using a bridge resistive-inductive-capacitive (RLC) at a frequency 10 KHz. The values are presented in Table I.

From these values, it is possible to define the value of the total ESR,  $\text{ESR}_{\text{tot}}$ , formed by the set of these capacitors:

$$\text{ESR}_{\text{tot}}(\text{healthy}) = 14.6 \text{ m}\Omega. \quad (31)$$

TABLE I  
ESR VALUE FOR EACH CAPACITORS OF THE INPUT FILTER

	$C_{11}$	$C_{12}$	$C_{21}$	$C_{22}$	$C_{31}$	$C_{32}$	$C_{41}$	$C_{42}$
ESR value (m $\Omega$ )	26,06	30,7	28,6	31,2	28,5	29,4	31	27,4

TABLE II  
COMPOSITION OF TRAINING SET  $X_a$  AND TEST SETS WITH STANDARDIZED DATA

Cluster	Operating mode	Load level	Numbers of samples for training	Numbers of samples for test
$\Omega_1$	Healthy inverter + Healthy machine	0 %	10	5
		50 %	10	5
$\Omega_2$	Healthy inverter + 3 broken bars	0 %	10	5
		50 %	10	5
$\Omega_3$	Faulty inverter ( ageing of capacitors) + Healthy machine	0 %	10	5
		50 %	10	5
$\Omega_4$	Faulty inverter ( ageing of capacitors) +3 broken bars	0 %	10	5
		50 %	10	5

Generally, the manufacturer recommends a change of the capacitors when the value of the ESR is twice the initial value. In practice, to accelerate this ageing, it is necessary to place the capacitors in a climatic chamber at high temperatures but this operation is long before obtaining a representative ageing. Consequently, to simulate this fault, we removed four capacitors over the eight present (only remainder  $C_{11}$ ,  $C_{12}$ ,  $C_{21}$ , and  $C_{22}$ ). So, the total ESR is increased by a factor 2:

$$\text{ESR}_{\text{tot}}(\text{aged}) = 29.12 \text{ m}\Omega. \quad (32)$$

Finally,  $X_a$  is constituted by four operating modes, represented by four classes. Each one made up of 20 samples (ten by load level) distributed in the following way.

Each test set is composed of five samples by level of load. Finally, there are ten samples by operating mode; so, 40 test samples (Table II). These 40 samples were not used to carry out the training set. In this way, the training set and test sets are always disjointed.

The first and the second steps are, respectively, the pattern vector determination and the creation of the decision space, i.e., the set of all possible clusters. This second step is realized by an expert because the SBS method used implies it being in supervised mode (knowledge of the distribution of the samples in classes). The pattern vector is composed initially of 35 features ( $d = 35$ ) [16].

Before using the selection method SBS, it is interesting to determine the capacity of discrimination of each feature providing the initial pattern vector compared to the training set  $X_a$ . For that, the Fisher ratio (17) is calculated for each feature. In order to illustrate the capacity of each one, the following histogram shows the criterion variation for the 35 calculated features.

The Fisher ratio shows that there are three features [ $\alpha_8, \alpha_{33}, \alpha_{31}$ ], which, taken separately, have a discriminating

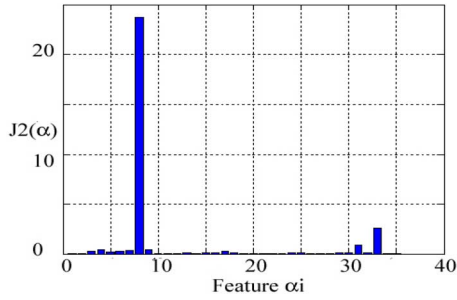


Fig. 13. Variation of the Fisher ratio for each feature.

TABLE III  
NUMBERS OF TEST SAMPLES BADLY AFFECTED ACCORDING TO VARIOUS DIMENSIONS OF PATTERN VECTOR

Features constituting the pattern vector	Numbers of samples badly affected
$[\alpha_8]$	6
$[\alpha_8, \alpha_{29}]$	5
$[\alpha_8, \alpha_{29}, \alpha_{35}]$	1
$[\alpha_8, \alpha_{29}, \alpha_{30}, \alpha_{35}]$	3
$[\alpha_1, \alpha_8, \alpha_{29}, \alpha_{30}, \alpha_{35}]$	1
$[\alpha_1, \alpha_8, \alpha_9, \alpha_{29}, \alpha_{30}, \alpha_{35}]$	3

capacity. Fig. 13 clearly shows that the feature  $\alpha_8$  has very strong capacity of discrimination compared to the others. However, this result does not mean that it is necessary to limit the size of the pattern vector to one dimension. Indeed, this one must be defined on the basis of collective contribution of the discriminating features.

Thus, according to the result of Fig. 13, the pattern vector dimension and thus feature space can be made from 1 to 3 elements.

In another way, the optimal features can be selected by the SBS algorithm, which used the eigenvalues of  $S_w^{-1}S_B$  as the criterion. It represents an indicator of the ratio between the interclasses and intraclass dispersion (16).

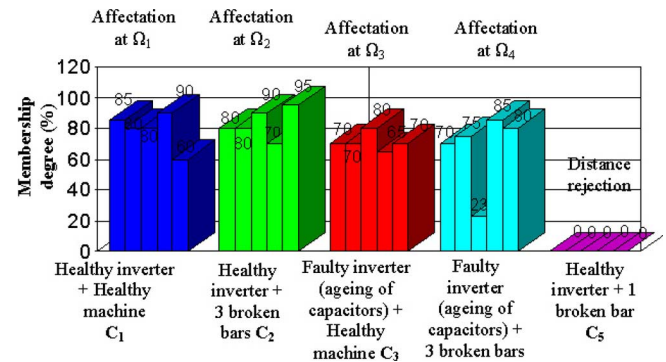
The initial space is the set of the calculated features. At each step of the procedure, the subspace that maximizes the criterion is taken into account. We preserve the subset that provides the best feature space according to the desired dimension. For that, at each step  $k$  of the selection, the subspace  $V_k$ , taken into account for the following step, is the one with the optimal criterion.

At each step  $k$ , a feature subset is selected, constituted of  $d - k$  features. To define the best pattern vector (feature subset), and consequently, the feature space dimension, we analyze the vector that gives the best performances according to the numbers of test samples (see Table III) badly affected.

The Table III gives the number of samples badly affected for various configurations of pattern vector. These vectors are obtained according to the results of Fig. 14. This analysis was carried out using the  $k$ -nn rule with  $k = 5$ . The choice to take  $k = 5$  is arbitrary. The distance reject used is defined by (25). The minimal number of neighbors  $k'$  is fixed to  $k/2$  ( $k$  equals 5, consequently  $k' = 3$ ).

TABLE IV  
PRESENTATION OF THE VARIOUS OBSERVATIONS USED FOR THE VALIDATION OF DECISION RULE

Validation	Operating mode	Load level	Number of samples
$C_1$ : $X_{n1}$ at $X_{n5}$	Healthy inverter + Healthy machine	0 %	5
$C_2$ : $X_{n6}$ at $X_{n10}$	Healthy inverter + 3 broken bars	0 %	5
$C_3$ : $X_{n11}$ at $X_{n15}$	Faulty inverter ( ageing of capacitors) + Healthy machine	0 %	5
$C_5$ : $X_{16}$ at $X_{n20}$	Faulty inverter ( ageing of capacitors) + 3 broken bars	50 %	5
$C_5$ : $X_{21}$ at $X_{n25}$	Healthy inverter + 1 broken bar	50 %	5

Fig. 14. Distribution of the new observations realized by the nonexclusive  $k$ -nn rule.

For a pattern vector composed of  $d' = 3$  features, the numbers of test samples badly affected is the weakest. Then, when dimension increases, this rate is constant and even tends to increase.

So, classification varies according to the pattern vector dimension. However, while following the evolution of the criterion value [cf. ([16])], it is possible to automatically determine the value of  $d'$ . In our case,  $d' = 3$  is the best compromise since for a greater dimension, the criterion value does not decrease any more to a significant degree. This break allows determining, without ambiguity, the choice of  $\langle\langle d' \rangle\rangle$ . Increasing the value of  $\langle\langle d' \rangle\rangle$  will not bring any additional information on classification.

This confirms the results obtained in Table III, the dimension optimal is 3. Thus, the three features that composed the optimal pattern vector are  $V_{1op} = [\alpha_8, \alpha_{29}, \alpha_{35}]$ .

The third step is the validation, i.e., the assignment given by the decision rule for all new observations  $X_n$ . These observations (see Table IV) are perfectly disjointed of the training set and test sets. Thus, they will be used to test the effectiveness of the decision rules.

Fig. 14 gathers the assignment results and the membership degree given by the nonexclusive  $k$ -nn rule (described in Sections III and IV) for all new healthy or faulty observations. The membership function of each class is defined by (28). The adjustable parameters have the following values:  $g = 1$ ,  $\beta = 2$ ,  $k_c = 20$ , and  $\eta = 1/n_c$  with  $n_c$ , the number of samples in class

$\Omega_c$ , here  $\eta = 0.05$ . The membership and ambiguity rejects are defined by (29) and (30), respectively.

All the observations are assigned to their operating mode contained in the training set  $X_a$ . The mode "healthy inverter+1 broken bar" not being present in  $X_a$ , the observations  $X_{u21}$  to  $X_{u25}$  are rejected in distance. These observations will have to be treated to detect the possible appearance of a new class (new operating mode). The membership degree for each observation varies between 65% and 95%. That gives a good reliability to classification carried out, no ambiguity is possible. Only the observation  $X_{u18}$  has a degree smaller than the others (23%). That is certainly due to an inaccuracy of measurements (equipment, noises, etc.).

This result shows that the features allow a good identification of the operating modes of the machine as well as a good differentiation between the machine faults and the inverter faults.

## VII. CONCLUSION

A diagnosis system based on PR was presented and applied with success to the detection of various faults being which can appear, either on the inverter or on the induction machine or even on both. Starting from the traditional formalism of PR, we added some specificity such as the automatic choice of dimension space, the data standardization, and a new decision rule allowing defining a reliability index in the diagnosis realized to improve the diagnosis system. The results show that the criterion evolution tracking based on separability and compactness of classes allows choosing effectively the feature space dimension, and thus, obtaining a better classification. The data standardization allows freeing the data from the level of load, and consequently, reducing the numbers of acquisitions necessary to form the initial training set. The classification method based on the  $k$ -nn rule was improved by developing a membership function associated with each class contained in the initial training set. It allows in identifying the evolution of an operating mode toward another or the appearance of an unknown operating mode. Thus, it is possible to predict the evolution of process toward a dangerous operating mode. Finally, the whole forms a supervision system dedicated to the diagnosis and allows realizing a preventive maintenance and ensuring the safety of the material and the person.

The PR approach is an original and effective way to provide the diagnosis of a system that takes into account the power supply used.

## REFERENCES

- [1] W. Deleroi, "Squirrel cage motor with broken bar in the rotor—Physical phenomena and their experimental assessment," in *Proc. Int. Conf. Electr. Mach. (ICEM 1982)*, Budapest, Hungary, , pp. 767–771
- [2] A. J. Penman, J. C. Tait, and W. E. Bryan, "A software-based approach for machine condition monitoring," presented at the Int. Conf. Electr. Mach. Design Appl., London, UK, Sep. 1985.
- [3] G. B. Kliman and R. A. Koegl, "Non-invasive detection of broken bars in operating induction motors," *IEEE Trans. Energy Convers.*, vol. 3, no. 4, pp. 873–879, Dec. 1988.
- [4] W. T. Thompson, "A review of on-line condition monitoring techniques for three phase squirrel induction motors—Past, present and future," in *Proc. IEEE Int. Symp. Diagn. Electr. Mach., Power Electron. Drives*, 1999, pp. 3–18.
- [5] P. Vas, *Parameter Estimation, Condition Monitoring and Diagnosis of Electrical Machines*. London, U.K.: Oxford Univ. Press, 1993.
- [6] B. G. Gaydon, "An instrument to detect induction motor rotor circuit defects by speed fluctuation measurements," in *Proc. Electr. Test Meas. Instrum.—Testmex 1979 Conf. Papers*, pp. 5–8.
- [7] P. J. Tavner and J. Penman, *Condition Monitoring of Electrical Machines*. Letchworth, U.K.: Research Studies Press, 1987.
- [8] S. C. Chang and R. Yacimini, "Experimental study of the vibrational behavior of machine stators," *Proc. Inst. Elect. Eng., Elect. Power Appl.*, vol. 143, no. 3, pp. 242–250, May 1996.
- [9] C. J. Dister and R. Schiferl, "Using temperature, voltage, and/or speed measurements to improve trending of induction motors RMS currents in process control and diagnostics," in *Proc. Ind. Appl. Conf., 1998. 33rd IAS Annu. Meeting*, vol. 1, pp. 312–318.
- [10] T. A. Lipo and K. C. Chang, "A new approach to flux and torque-sensing in induction machines," *IEEE Trans. Ind. Appl.*, vol. IA-12, no. 3, pp. 142–148, May 1986.
- [11] G. B. Kliman and R. A. Koegl, "Noninvasive detection of broken bars in operating induction motors," *IEEE Trans. Energy Convers.*, vol. 3, no. 4, pp. 873–879, Dec. 1988.
- [12] G. B. Kliman and J. Stein, "Methods of motor current signature analysis," *Elect. Mach. Power Syst.*, vol. 20, no. 5, pp. 463–474, Sep. 1992.
- [13] M. E. H. Benbouzid *et al.*, "Induction motor diagnostics via stator current monitoring," in *Proc. 1997 Int. Conf. Maintenance Rel.*, Knoxville, TN, vol. 1, pp. 36.01–36.10.
- [14] P. Venet, "Surveillance d'alimentation à découpage. Application à la maintenance prédictive," Thèse de Doctorat, Université Claude Bernard Lyon1, Lyon, France, Décembre 1993.
- [15] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. San Diego, CA: Academic, 1990.
- [16] R. Casimir, "Diagnostic des défauts des machines asynchrones par reconnaissance des formes," Thèse de l'Ecole Centrale de Lyon, Centre de Génie Électrique de Lyon (CEGELY), Lyon, France, 2003.
- [17] E. Blanco, O. Ondel, and A. Llor, "Using linear interpolation and Kalman prediction in pattern recognition: Application on an induction machine," presented at the Int. Conf. Symp. Diagn. Electr. Mach., Power Electron. Drives (SDEMPED 2005), Vienna, Austria, Sep. 2005.
- [18] O. V. Thorsen and M. Dalva, "A survey of faults on induction motors in offshore oil industry, petrochemical industry, gas terminal, and oil refineries," *IEEE Trans. Ind. Appl. (IAS)*, vol. 31, no. 5, pp. 1186–1196, Sep./Oct. 1995.
- [19] M. E. H. Benbouzid, M. Vieira, and C. Theys, "Induction motor's faults detection and localization, using stator current advanced signal processing techniques," *IEEE Trans. Power Electron.*, vol. 14, no. 1, pp. 14–22, Jan. 1999.
- [20] W. T. Thomson and M. Fenger, "Current signature analysis to detect induction motor faults," *IEEE Ind. Appl. Mag.*, vol. 7, no. 4, pp. 26–34, Jul./Aug. 2001.
- [21] R. R. Schoen, T. G. Habelter, F. Kamran, and R. G. Bartheld, "Motor bearing damage detection using stator current monitoring," *IEEE Trans. Ind. Appl.*, vol. 31, no. 6, pp. 1274–1279, Nov./Dec. 1995.
- [22] T. Boumegoura, G. Clerc, H. Yahoui, and G. Grellet, "Rotor resistance estimation by non linear observer for diagnostic and control systems," *JESA*, vol. 34, no. 8, pp. 1011–1024, Oct. 2000.
- [23] H. A. Toliyat and T. A. Lipo, "Transient analysis of cage induction machines under stator, rotor bar and end ring faults," *IEEE Trans. Energy Convers.*, vol. 10, no. 2, pp. 241–247, Jun. 1995.
- [24] T. Assaf, H. Henao, and G. A. Capolino, "Detection of voltage source dissymmetry using the measurement of signal components in working induction motors," in *Proc. IEEE Int. Symp. Diagn. Electr. Mach., Power Electron. Drives*, Grado, Italy, Sep. 1–3, 2001, pp. 441–447.
- [25] P. Vas, *Electrical Machines and Drives, A Space-Vector Theory Approach*. New York: Oxford Univ. Press, 1992.
- [26] M. G. Melero, M. F. Cabanas, C. Rojas, G. A. Orcajo, J. M. Cano, and J. Solares, "Study of an induction motor working under stator winding inter-turn short circuit condition," in *Proc. IEEE Int. Symp. Diagn. Electr. Mach., Power Electron. Drives*, Aug. 24–26, 2003, pp. 52–57.
- [27] A. K. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 2, pp. 153–158, Feb. 1997.
- [28] F. J. Ferri, P. Pudil, M. Hatef, and J. Kittler, "Comparative study of techniques for large-scale feature selection," in *Pattern Recognition in Practice IV, Multiple Paradigms, Comparative studies and Hybrid Systems*, E. S. Gelsema and L. S. Kanal, Eds. Amsterdam, The Netherlands: Elsevier, 1994, pp. 403–413.

- [29] A. Biem, S. Katagiri, and B.-H. Juang, “Pattern recognition using discriminative feature extraction,” *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 500–504, Feb. 1997.
- [30] T. W. S. Chow and D. Huang, “Estimating optimal features subset using efficient estimate of high dimensional mutual information,” *IEEE Trans Neural Netw.*, vol. 16, no. 1, pp. 213–224, Jan. 2005.
- [31] N. Kwak and C.-H. Choi, “Input feature selection for classification problems,” *IEEE Trans. Neural Networks.*, vol. 13, no. 1, pp. 143–159, Jan. 2002.
- [32] M. Kudo and J. Sklansky, “Comparison of algorithms to select features for pattern classifiers,” *Pattern Recog.*, vol. 33, no. 1, pp. 25–41, Jan. 2000.
- [33] B. Dubuisson, “Decision with reject options,” presented at the 5th Eur. Signal Process. Conf., Barcelona, Spain, 1990.
- [34] M. E. Hellman, “The nearest neighbour classification rule with a reject option,” *IEEE Trans. Syst., Man, Cybern.*, vol. 6, no. 3, pp. 179–185, Jul. 1970.
- [35] I. Tomek, “A generalisation of the kNN rule,” *IEEE Trans. Syst., Man, Cybern.*, vol. 6, no. 2, pp. 121–126, Feb. 1976.
- [36] P. A. Devijver and J. Kittler, “On the edited nearest neighbor rule,” in *Proc. 5th Int. Conf. Pattern Recog.*, Dec., 1980, pp. 72–80.
- [37] L. A. Zadeh, “Fuzzy sets,” *Inf. Control*, vol. 8, pp. 338–353, 1965.
- [38] J. C. Bezdek and S. K. Pal, *Fuzzy Models for Pattern Recognition: Methods That Search for Structures in Data*. Piscataway, NJ: IEEE Press, 1992.
- [39] O. Ondel, A. Yazidi, E. Boutleux, G. Clerc, H. Henao, R. Casimir, and G.-A. Capolino, “Comparative study of two diagnosis methods for induction machine,” in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Dec. 8–10, 2004, pp. 159–165.
- [40] A. Lahyani, P. Venet, G. Grellet, and P. J. Vivierge, “Failure prediction of electrolytic capacitors during operation of a switch mode power supply,” *IEEE Trans. Power Electron.*, vol. 13, no. 6, pp. 1199–1207, Nov. 1998.



**Olivier Ondel** was born in France, on October 16, 1978. He received the M.S. degree in electrical engineering from the University Claude Bernard, Lyon, France, in 2003 and the Ph.D. degree from the Department of Electrical Engineering, Ecole Centrale de Lyon, Ecully, France.

He is currently an Assistant Professor at the Centre National de la Recherche Scientifique (CNRS), Unités Mixtes de Recherche (UMR) 5005, Laboratoire AMPERE, Villeurbanne, France. His current research interests include fault diagnosis on a set

“converter-electrical machines” by means of pattern recognition methods.



**Guy Clerc** (M’00) was born in Libourne, France, on November 30, 1960. He received the HRD and Ph.D. degrees in electrical engineering from the Ecole Centrale de Lyon, Ecully, France, in 1984 and 1989, respectively.

He has been a Professor of many universities. He teaches electrical engineering at the “University Claude Bernard Lyon I,” Lyon, France. He is with the Université de Lyon, Lyon. He is also with the Centre National de la Recherche Scientifique (CNRS), Unités Mixtes de Recherche (UMR) 5005, Laboratoire AMPERE, Villeurbanne, France, where he is engaged in research on control and diagnosis of induction machines.



**Emmanuel Boutleux** was born on July 1970, in France. He received the Ph.D. degree in system control from the Heudiasyc Laboratory, Compiègne, France, in 1996.

He is currently an Assistant Professor in automatic at the Ecole Centrale de Lyon, Ecully, France. He is also with the Centre National de la Recherche Scientifique (CNRS), Unités Mixtes de Recherche (UMR) 5005, Laboratoire AMPERE, Villeurbanne, France, where he is engaged in diagnosis of complex systems by means of pattern recognition methods. He teaches

automatica to engineering students.



**Eric Blanco** was born in France, in 1975. He received the Ph.D. degree in automatic and signal processing from Claude Bernard University (Lyon1), Lyon, France, in 2002.

From 1998 to 2003, he was engaged in research on robust estimation problems in the Laboratory of Process Control and Chemical Engineering (LAGEP). Since 2003, he has been an Assistant Professor in automatic and signal processing in the Ecole Centrale de Lyon, Ecully, France. He is also with the Centre National de la Recherche Scientifique (CNRS),

Unités Mixtes de Recherche (UMR) 5005, Laboratoire AMPERE, Villeurbanne, France, where he is engaged in diagnosis, estimation, and prediction of complex systems.