Classification of Induction Machine Faults by Optimal Time-Frequency Representations

Abdesselam Lebaroud, Guy Clerc

Classification of Induction Machine Faults by Optimal Time–Frequency Representations

Abdesselam Lebaroud and Guy Clerc, *Member, IEEE*

***Abstract*—This paper presents a new diagnosis method of induction motor faults based on time–frequency classification of the current waveforms. This method is based on a representation space, a selection criterion, and a decision criterion. In order to define the representation space, an optimized time–frequency representation (TFR) is designed from the time–frequency ambiguity plane. The selection criterion is based on Fisher's discriminant ratio, which allows one to maximize the separability between classes representing different faults. A distinct TFR is designed for each class. The following two classifiers were used for decision criteria: the Mahalanobis distance and the hidden Markov model. The flexibility of this method allows an accurate classification independent from the level of load. This method is validated on a 5.5-kW induction motor test bench.**

***Index Terms*—Diagnosis, hidden Markov model (HMM), induction motor, time–frequency classification.**

I. Introduction

MANY CRITICAL industrial processes require reliability and safety operation of electric motors. However, unexpected machinery failures provide loss of production, high emergency maintenance costs, damages to other related machinery, and extended process downtime. Thus, very expensive scheduled maintenance is performed in order to detect machine problems before they can result in catastrophic failure [1]. Therefore, there is a considerable demand to reduce maintenance costs and prevent unscheduled downtimes for electrical drive systems. The major faults of electrical machines can broadly be classified as the following [2]:

1) stator faults resulting in the opening or shorting of one or more of stator phase windings;
2) abnormal connection of the stator windings;
3) broken rotor bar or cracked rotor end-rings;
4) static and/or dynamic air-gap irregularities;
5) bent shaft (dynamic eccentricity . . .) which can result in a rub between the rotor and stator, causing serious damage to stator core and windings;
6) shorted rotor field winding;
7) bearing and gearbox failures.

Manuscript received December 18, 2007; revised August 5, 2008. First published September 12, 2008; current version published December 2, 2008.

A. Lebaroud is with the Laboratory LEC, University of Constantine, Constantine, 25017 Algeria and the University of Skikda, Skikda 21000, Algeria (e-mail: lebaroud80@gmail.com).

G. Clerc is with the Ampère Laboratory, Université Claude Bernard Lyon 1, 69622 Villeurbanne, 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/TIE.2008.2004666

Of the aforementioned types of faults, bearing faults, stator faults, and broken bars are the most prevalent, although almost 40%–50% of all failures are bearing related. Bearing faults might manifest themselves as rotor asymmetry faults [3] which are usually covered under the category of eccentricity-related faults. Stator faults are usually related to insulation failure; they manifest themselves through phase-to-ground connections or phase-to-phase faults. Also, stator faults manifest themselves by abnormal connection of the stator windings. Almost 30%–40% of all reported induction motor failures fall in this category [4]. The rotor fault now accounts for around 5%–10% of the total induction rotor failures [4]. The principal reasons for rotor bar and ring breakage are thermal stresses due to thermal overload and unbalance. On the other side, magnetic stresses are caused by electromagnetic forces, unbalanced magnetic pull, electromagnetic noise, and vibration. In recent years, many research works have been carried out for the study and development of fault detection and diagnosis methods for electric machines. Recent advances of signal processing techniques, such as artificial neural networks [5], wavelets [6], etc., have provided more powerful tools for fault diagnosis. Several diagnostic strategies such as model-based approaches [7], pattern recognition [8] and artificial intelligence [9] have been recently presented in technical literature. Recently proposed approaches for automated detection and classification of induction machine faults are based on wavelet analysis, artificial neural networks, and hidden Markov models (HMMs) [10], [11]. As input parameters of diagnosis system, signals (current, voltage, torque, vibration, etc.) are most often used. The problem of diagnosis systems is that they use signals either in the time or frequency domain. In our approach, instead of using a time or a frequency approach, it is potentially more informative to use both time and frequency. Time–frequency analysis of the motor current makes signal properties, related to fault detection, more evident in the transform domain [12].

In many classification applications [13]–[15], features are traditionally extracted from standard time–frequency representations (TFRs). This assumes that the implicit smoothing is appropriate for the classification task. Making such assumptions can degrade classification performance. TFRs can be uniquely characterized by an underlying function called a kernel. In previous time–frequency research, kernels have been derived in order to fulfill properties, such as minimizing quadratic interference, although some of the resulting TFRs can offer advantages for classification of certain types of signals. The goal of sensitive detection or accurate classification is rarely an explicit goal of kernel design. Those few methods that optimize

the kernel for classification purposes constrain the form of the kernel to predefined parametric functions with symmetries that cannot be suitable to detection or classification [15]. Traditionally, the objective of time–frequency research is to create a function that will describe the energy density of a signal simultaneously in time and frequency. For explicit classification, it is not necessarily desirable to accurately represent the energy distribution of a signal in time and frequency. In fact, such a representation may conflict with the goal of classification, generating a TFR that maximizes the separability of TFRs from different classes. It may be advantageous to design TFRs that specifically highlight differences between classes [15]. The use of TFR includes the following two sequential processes: feature extraction and rule decision. This technique has been successfully applied for classification of induction machine faults [16], tool-wear monitoring, and radar transmitter identification [17]. For classification, the optimization procedure of TFR (1) via parameter kernel is very computationally prohibitive. It would be better to use the optimal TFR that can be classified directly in the ambiguity plane. We propose to design and use the classifier directly in the ambiguity Doppler-delay plane. Since all TFRs can be derived from the ambiguity plane, no *a priori* assumption is made about the smoothing required for accurate classification. Thus, the smoothing quadratic TFRs retain only the information that is essential for classification.

In this paper, we propose a classification technique based on the design of optimized TFR from a time–frequency ambiguity plane in order to extract the feature vector. A decision criterion based on a "Mahalanobis distance" was used to ensure assignment to the appropriate class, but given these limitations, it has been replaced by a new criterion, namely, the HMM. The goal is the realization of an accurate diagnosis system of motor faults such as bearing faults, stator faults, and broken bars.

II. Classification Algorithm

The classification algorithm is composed of the following two parts: extraction features and decision making. In the training stage, in order to build the extraction features, three optimal kernels are designed for separating four classes. The kernel design process selects, for each class, three locations from the time–frequency ambiguity plane. In the decision making stage, the following two different classifiers are proposed: the Mahalanobis distance and the HMM.

The classification algorithm is applied to detection of three kinds of induction machine faults, which are bearing fault, stator fault, and rotor fault. Thus, the following four classes are considered:

1) class of healthy motor;
2) class of bearing fault;
3) class of stator fault;
4) class of broken bars.

The classification algorithm of these three classes of defects and the healthy class is shown in the Fig. 1.

Two steps are necessary for optimal classification. The details of each step are described in the following paragraphs.

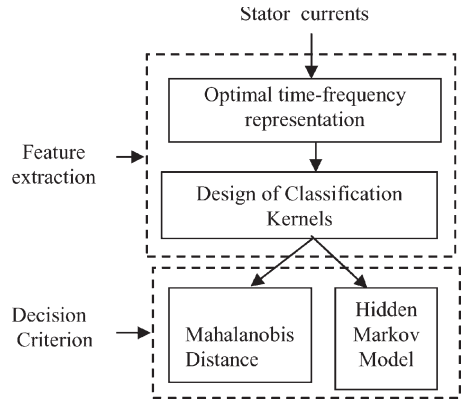


Fig. 1. Classification procedure.

III. Feature Extraction

A. Optimal TFR

The important connection between the ambiguity plane and TFRs has been well established for a long time. Any bilinear (Cohen class) TFR of signal $x(t)$ can be expressed as the 2-D Fourier transform of the product of the ambiguity plane $A(\eta,\tau)$ of the signal and a kernel function $\varphi(\eta,\tau)$ [13]

$$\mathrm{TFR}_x^{\varphi}(t,f)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}A(\eta,\tau)\varphi(\eta,\tau)e^{j2\pi\eta t}e^{-2\pi jf\tau}d\eta d\tau \quad (1)$$

where t represents the time, f represents the frequency, η represents the continuous frequency shift, and τ represents the continuous time lag. The ambiguity plane $A(\eta,\tau)$ for a given signal $x(t)$ is defined as

$$A_x(\eta,\tau)=\int_{-\infty}^{\infty}x(t)x^*(t+\tau)e^{j2\pi\eta t}dt \quad (2)$$

$x(t+\tau)$ represents the signal at a future time $(t+\tau)$ and $x^*(t+\tau)$ is the complex conjugate of $x(t+\tau)$.

The discrete versions of ambiguity plane (2) is [17]

$$A[\eta,\tau]=F_{n\to\eta}\left\{\Re[n,\tau]\right\}=\sum_{n=0}^{N-1}\Re[n,\tau]e^{-j\frac{2\pi}{N}n\eta} \quad (3)$$

where F represents the Fourier transform, η represents the discrete frequency shift, τ represents the discrete time lag, and n represents the sample. The instantaneous autocorrelation function $\Re[n,\tau]$ is defined as

$$\Re[n,\tau]=x^*[n]\cdot x\left[(n+\tau)_N\right]. \quad (4)$$

Thus, the class-dependent TFR [15] is defined by

$$\mathrm{TFR}[n,k]=F_{\eta\to n}^{-1}\left\{F_{\tau\to k}\left\{\varphi[\eta,\tau]A[\eta,\tau]\right\}\right\} \quad (5)$$

where k is the discrete frequency.

The characteristic function for each TFR is $A(\eta,\tau)\varphi(\eta,\tau)$. In other words, for a given signal, a TFR can be uniquely mapped from a kernel. The classification-optimal representation TFR_i can be obtained by smoothing the ambiguity plane with an appropriate kernel φ_{opt}, which is a classification-optimal kernel. The problem of designing the TFR_i becomes equivalent to designing the classification-optimal kernel $\varphi_{\text{opt}}(\eta,\tau)$.

This method, used to design kernels (and thus TFRs), optimizes the discrimination between predefined sets of classes. The resulting kernels are not restricted to any predefined function but are rather arbitrary in shape. This approach ascertains the necessary smoothing to achieve the best extraction features.

The kernel determines the representations and its properties. A kernel function is a generating function that operates upon the signal to produce the TFR. Gillespie and Atlas [17] have proposed feature extraction methods based on designing class-dependent TFRs from the time–frequency ambiguity plane. Features can be extracted directly from $A(\eta,\tau)\varphi_{\text{opt}}(\eta,\tau)$ instead of the classification-optimal TFR_i. This shortcut simplifies the computation complexity of feature extraction by reducing the calculations described in (1).

B. Design of Classification Kernels

The kernel $\varphi_{\text{opt}}(\eta,\tau)$ is designed for each specific classification task. We determine N locations from the ambiguity plane, in such a way that the values in these locations are very similar for signals from the same class, while they vary significantly for signals from different classes. The notation $A_{ij}[\eta,\tau]$ represents the ambiguity plane of the jth training example in the ith class. We design and use Fisher's discriminant ratio (FDR) to get those N locations.

In our classification procedure, $C-1$ kernels must be designed for a C-class classification system. As we have four classes (three fault cases and one healthy case of the machine), we must design the following three kernels: bearing fault kernel, stator fault kernel, and rotor fault kernel. Each kernel separates the healthy case from the fault case.

The discrimination between different classes [10] is made by separating the class i from all remaining classes $\{i+1,\ldots,N\}$ (Fig. 2). In this case, the stator fault kernel is designed to discriminate the stator fault class from the other classes (rotor fault, bearing fault, and healthy motor). The rotor fault kernel is designed to discriminate the rotor fault class from the remaining classes (bearing fault and healthy motor). The bearing fault kernel is designed to discriminate the bearing fault class from the healthy motor class. The advantage of the method lies in the optimum separation between the different classes.

Regarding the degree of severity from machine faults by TFR, a method of calculation is presented in the paper published in the IECON 06 conference [16].

The kernels are designed by I training example signals from each class with the equation as follows:

$$\text{FDR}_i(\eta,\tau)=\frac{(m_i[\eta,\tau]-m_{i-\text{remain}}[\eta,\tau])^2}{V_i^2[\eta,\tau]+V_{i-\text{remain}}^2[\eta,\tau]} \tag{6}$$

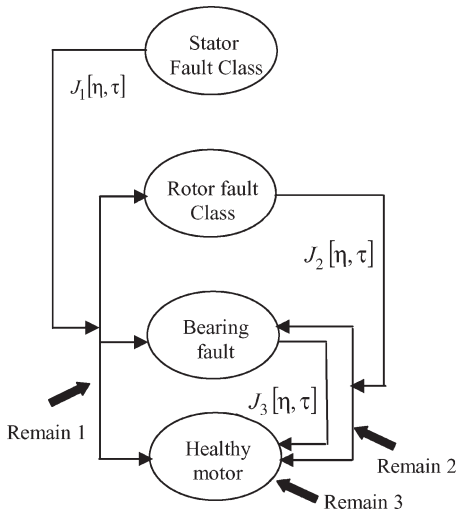


Fig. 2. Algorithm of separation between classes.

where $m_i[\eta,\tau]$ and $m_{i-\text{remain}}[\eta,\tau]$ represent two means of location (η,τ)

$$m_i[\eta,\tau]=\frac{1}{N_i}\sum_{j=1}^{N_i}A_{ij}[\eta,\tau]. \tag{7}$$

In order to avoid unnecessary computation to separate classes, we proposed the principle of the remaining classes (Fig. 2)

$$m_{i-\text{remain}}[\eta,\tau]=\frac{\sum_{k=i+1}^{4}\sum_{j=1}^{N_k}A_{kj}[\eta,\tau]}{\sum_{k=i+1}^{4}N_k} \tag{8}$$

$$V_i^2[\eta,\tau]=\frac{1}{N_i}\sum_{j=1}^{N_i}(A_{ij}[\eta,\tau]-m_i[\eta,\tau])^2 \tag{9}$$

$$V_{i-\text{remain}}^2[\eta,\tau]=\frac{\sum_{k=i+1}^{4}\sum_{j=1}^{N_k}(A_{kj}[\eta,\tau]-m_{i-\text{remain}}[\eta,\tau])^2}{\sum_{k=i+1}^{4}N_K}. \tag{10}$$

$V_i^2[\eta,\tau]$ and $V_{i-\text{remain}}^2[\eta,\tau]$ Two variances of location (η,τ).

We transform the FDR to φ_{opt} kernel in a binary matrix by replacing the maximum N points with one and the other points with zero (Fig. 3). Features can be extracted directly from $\varphi_{\text{opt}}[\eta,\tau]oA[\eta,\tau]$, where o is an element-by-element matrix product. The kernel has the same dimensions as the ambiguity plane. By multiplying the φ_{opt} kernel with a given signal ambiguity plane, we will find k feature points for this signal. We put them into a vector in order to create the training feature vector $\text{FV}_{\text{train}}^{(c)}(k)$ of class C

$$\varphi_{\text{opt}}^{(c)}[\eta,\tau]o\overline{A}^{(c)}[\eta,\tau]=\begin{cases}\overline{A}^{(c)}[\eta,\tau], & \text{if } \varphi_{\text{opt}}^{(c)}[\eta,\tau]=1\\ 0, & \text{if } \varphi_{\text{opt}}^{(c)}[\eta,\tau]=0\end{cases} \tag{11}$$

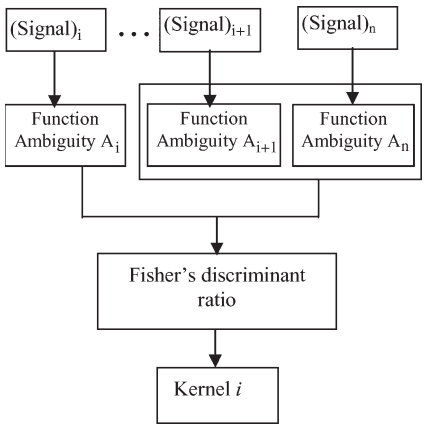


Fig. 3. Kernel design.

where

$\varphi_{\text{opt}}^{(c)}[\eta,\tau]$ training optimal kernel;

$\overline{A}^{(c)}[\eta,\tau]$ mean class of the ambiguity plane.

According to the symmetries of the ambiguity plane, only points on a quarter planes are considered.

The goal of feature extraction is to generate N-point feature vector from the original 10 000-point current signal.

IV. Decision Criteria

A. Mahalanobis Distance

Introduced in 1936 by Mahalanobis [18], the Mahalanobis distance is a distance measure. It is based on correlations between variables by which different patterns can be identified and analyzed. It is a useful way of determining similarity of an unknown sample set to a known one.

In order to reduce the size of the signal, while preserving the relevant information, the original signal is resampled. Only the range of the required frequencies is preserved. By resampling, the signal dimension has been reduced greatly. This leads to a great reduction of the computation complexity.

TFR Mahalanobis-distance-based fault diagnosis procedure (Fig. 4) permits one to overcome the problem of load level to take only two levels, namely, without load (condition 1, represented by $A_1(\eta,\tau)o\varphi_1$) and rated load (condition 2, represented by the characteristic function $A_2(\eta,\tau)o\varphi_2$). The characteristic function of the fault class is the average of the two previous function characteristics. Extraction of relevant points of the characteristic function provides an average feature vector FV_{AV}. A test signal (unclassified, condition 3) is represented after feature extraction by FV_{TEST}. The Mahalanobis distance is computed between FV_{AV} and FV_{TEST}.

After designing the kernels on the basis of a set of samples of the C classes, actual classification is performed.

Given a particular unknown test signal vector (the classifier is not trained on this example), the classifier estimates the class membership of this example. The classification of x in one of the C classes can be realized via a Mahalanobis distance

$$d_M(x,y)=\sqrt{(x-y)^{\text{T}}\sum^{-1}(x-y)} \tag{12}$$

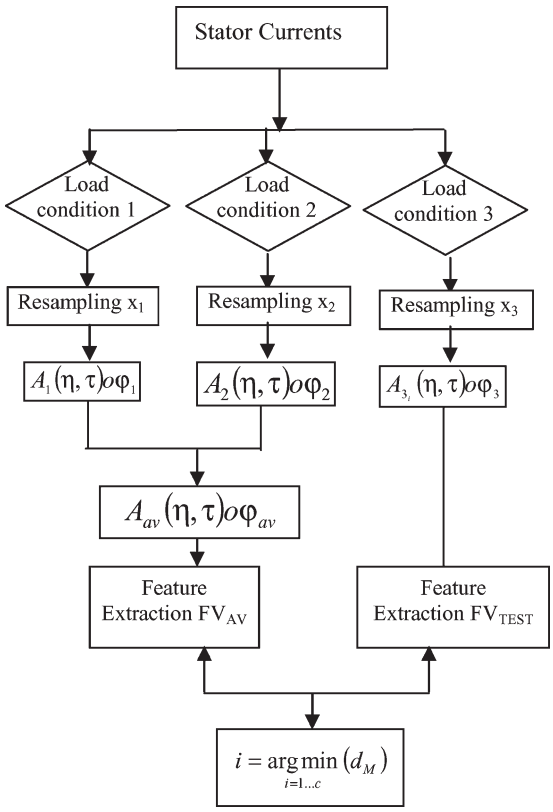


Fig. 4. Scheme of the TFR Mahalanobis-distance-based fault diagnosis.

where x is affected to the class $c_i \Leftrightarrow i = \arg\min_{i=1,\ldots,c}\{d_M(x,y)\}$. In TFR, planes x and y become

$$x=\varphi_{\text{opt}}^{(c)} o A_x$$

$$y=\varphi_{\text{opt}} o A^{(c)}$$

where

A_x ambiguity plane computed with the signal x;

$A^{(c)}$ ambiguity plane computed with the training set.

The Mahalanobis distance of (12) becomes

$$d_M\left((\varphi_{\text{opt}} o A_x),(\varphi_{\text{opt}} o A^{(c)})\right) = \left((\varphi_{\text{opt}} o A_x-\varphi_{\text{opt}} o A^{(c)})^{\text{T}} \cdot \sum_c^{-1}(\varphi_{\text{opt}} o A_x-\varphi_{\text{opt}} o A^{(c)})\right)^{1/2} \tag{13}$$

where $(.)^{\text{T}}$ denotes the matrix transpose. The covariance matrix $\sum_c$ is estimated from the training data.

A rejection decision is taken when the signal x to be classified is far from any class

$$\begin{cases} x \text{ is affected}, & \text{if } d_M(FV_x) \prec \beta \\ x \text{ is rejected}, & \text{otherwise} \end{cases} \tag{14}$$

where β is a given rejection threshold.

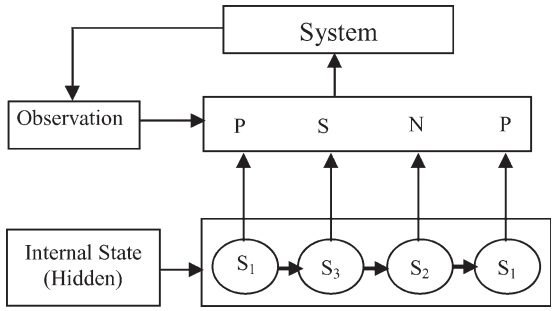


Fig. 5. Principle of an HMM.

The rejection threshold is obtained from half of the smallest distance between two classes

$$\beta = d_{\min}/2. \tag{15}$$

The value $\beta = 0.4$ is calculated from three classes constituting the whole learning of the used machine.

The error of the badly classified points of feature vectors FV_x is calculated by

$$e(i,j)\% = \frac{1}{I}\sum_{i=1}^{N_1}\sum_{j=1}^{N_2}\frac{\mathrm{FV}_{\mathrm{train}_{i,j}} - \mathrm{FV}_{x_{i,j}}}{\mathrm{FV}_{\mathrm{train}_{i,j}}} \cdot 100 \tag{16}$$

where

$I = N_1 \cdot N_2$ number of examples per class;
N_2 number of current examples of the same load level;
N_1 number of load levels.

B. HMM

A finite-state HMM consists of a finite number of states where transitions among the states are governed by a Markov chain with a set of transition probabilities [19]. HMMs are an extension of Markov models that include the case where the observations are probabilistic functions of the states rather than the states themselves. The underlying stochastic process, the state sequence, is not observable and can only be estimated through another set of stochastic processes that produce observation sequences (Fig. 5).

1) Elements of an HMM: The following parameters are needed to define a HMM [19].

Notation:

$$\mathrm{HMM} = (N, M, A, B, \pi); \tag{17}$$

with N states, $S = \{S_1, \ldots, S_N\}$, with q_t at time t; M observations, v_1 to v_M, with O_t at time t; and $A = \{a_{ij}\}$ is the transition probabilities.

Then

$$a_{ij} = P(q_{t+1} = S_j | q_t = S_i) \tag{18}$$

with q_t denoting the current state, i.e., the probability of being in state S_j at time $t+1$, provided that the state at time t is S_i and

$B = \{b_j(k)\}$ is the observation probabilities

with

$$b_j(k) = P(v_t = O_k | q_t = S_j) \tag{19}$$

where O_k and M denote the kth observation and the number of distinct observations, respectively

$\pi = \{\pi_i\}$ is the initial state probabilities

where

$$\pi_i = P(q_1 = S_i) \tag{20}$$

i.e., the probability of the lth state being the initial state.

The compact notation $\lambda = (A, B, \pi)$ is generally used to represent an HMM.

The following are the three basic problems to solve for an HMM.

Problem 1) Given the observation sequence $0 = 0_1 0_2, \ldots, O_T$ and a model $\lambda = (A, B, \pi)$, how do we efficiently compute $P(0_1)$, the probability of the observation sequence, given the model?

Problem 2) Given the observation sequence $0 = 0_1 0_2, \ldots, O_T$ and a model $\lambda = (A, B, \pi)$, how do we choose a corresponding state sequence $Q = q_1 q_2, \ldots, q_T$, which is optimal to generate the observation sequence? The optimal measure can be maximum likelihood (ML).

Problem 3) How do we adjust the parameters $\lambda = (A, B, \pi)$ to maximize the likelihood of all observation sequences?

In this paper, we mainly discuss Problem 3) in order to learn the parameters of the model and Problem 1) in order to provide fault detection. For Problem 1), the ML algorithm is generally used for labeled training data, and the Baum–Welch algorithm is used for partially labeled or unlabeled training data.

2) Training Faults by HMM: The feature is used to train an HMM that represents the specific fault conditions. HMMs can be trained using multiple-feature vectors. Thus, multiple features can be used in constructing an HMM for a specific fault type. For instance, if we have several load conditions for the operation of the stator fault, we can train a single HMM using data for various load conditions (Fig. 6). Given the feature matrix, the probability of the HMM for the faulty condition is calculated. The HMMs are trained in order to represent the following faults: stator fault, rotor fault, and bearing fault.

3) Fault Detection and Diagnosis Using the HMM: We need to train HMMs to identify the motor faults that are of interest, e.g., stator fault, rotor fault, and bearing fault. Once the models are trained, a motor fault can be diagnosed by following the steps shown in Fig. 7. The first step is extracting the feature vector from the preprocessed signal (Fig. 6). Next, the probability of the feature vector is calculated, given all the HMMs in the previously constructed database. The HMM, for which the probability is maximum, determines the type of the fault.

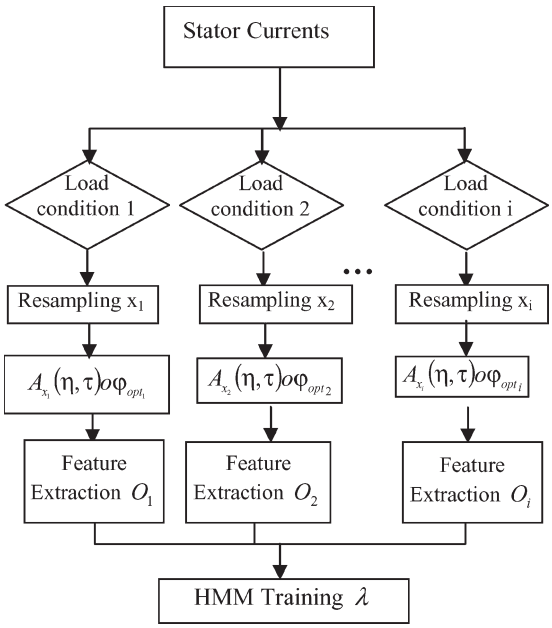


Fig. 6. TFR HMM training.

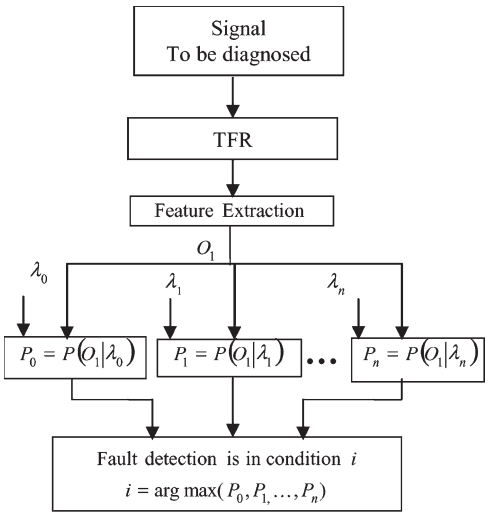


Fig. 7. Scheme of TFR HMM-based fault diagnosis.

V. Experiment Results

The experimental bench consists of a three-phase asynchronous-motor squirrel cage Leroy Somer LS 132S, IP 55, Class F, T °C standard $= 40$ °C. The motor is loaded by a powder brake. Its maximum torque (100 Nm) is reached at rated speed. This brake is sized to dissipate a maximum power of 5 kW. Fig. 8 shows the motor bench.

The wear obtained on the bearings is a real one [the bearings have been provided by SECCO (Fig. 9)]. For the rotor fault, the bar has been broken by drilling the bar of the cage squirrel (Fig. 10). The 5% of power imbalance for simulating the fault of imbalance stator is obtained with a variable autotransformer placed on a phase of the network (Fig. 8).

An acquisition of current signals was carried out on a test bench. The sampling rate is 10 kHz. The number of samples per signal rises to $N = 10\,000$ samples. The data acquisition

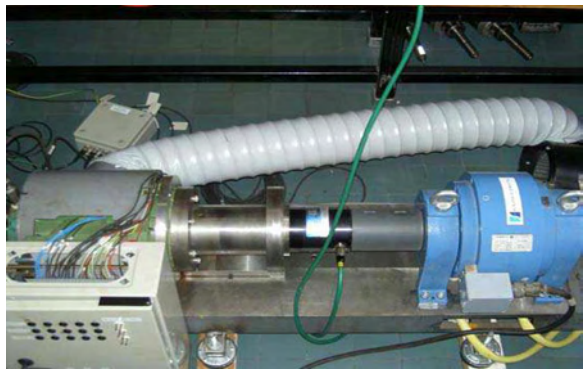


Fig. 8. Test bench of the induction motor.

Fig. 9. Bearing faults.

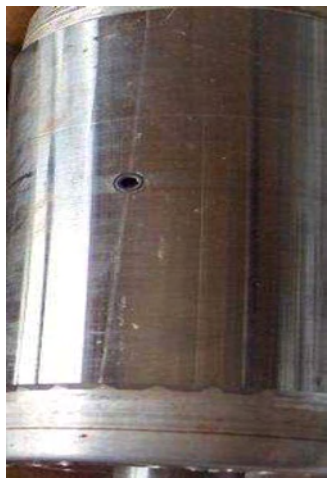


Fig. 10. Rotor with one broken bar.

set on the machine consists of 15 examples of stator current recorded on different levels of load (0%, 25%, 50%, 75%, and 100%). Different operating conditions from the machine were considered, namely, healthy, bearing fault, stator fault, and rotor fault. The training set is carried out on ten current examples. The last five current examples are used to test the classification.

The system is trained by using ten current signals at 0% and 100% load levels. We take $N_1 = 25$ (number of classes) in order to solve the problem of the load levels. The test is then performed on current signals collected at 25% and 70% load levels.

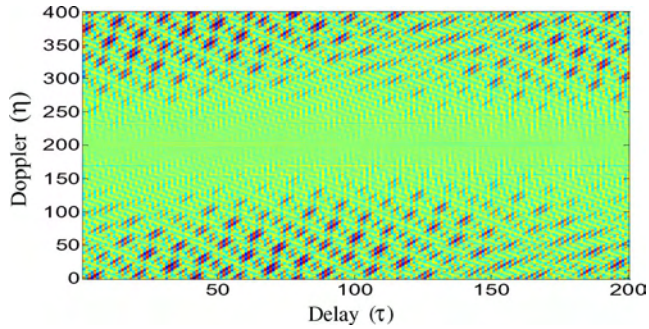


Fig. 11. Example of not-smoothed ambiguity plane of the fault of the stator imbalance class (5%).

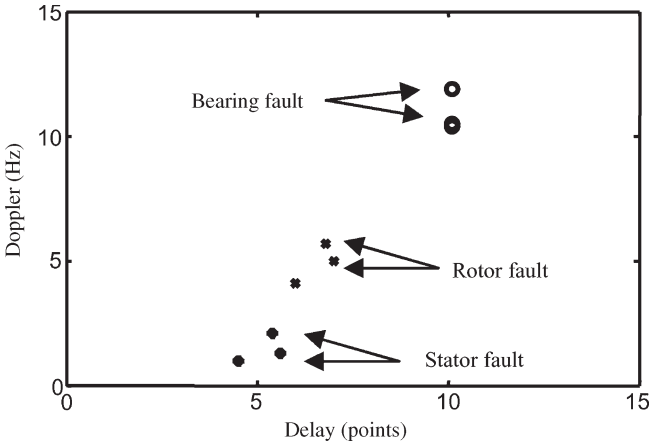


Fig. 12. Ambiguity plane smoothed by three kernels.

A. Training Set

Each class of the training set for the three faults and for the healthy machine is made of ten examples of no-load current and ten other examples for the full load. Consequently, we have 20 examples of training for each of the three faults and 20 examples of training also for the healthy machine. The dimension of ambiguity plane contains initially $200 \times 200 = 40\,000$ points; considering symmetry compared to the origin, we take the quarter of ambiguity plane, which corresponds to $N = 10\,000$.

Fig. 11 shows an example of ambiguity plane, which was not smoothed by the kernel, of the 5% imbalance voltage class. The plane represents all Doppler-delay positions to ensure the separation of class. However, this graphical representation, as we had previously assumed, does not provide useful information for diagnosis. However, it is difficult to provide more interpretation because the aim of the TFR design process is good separation between classes instead of accurate representation of a signal.

We designed the following three kernels: kernel stator fault, kernel rotor fault, and kernel bearing fault. Fisher's point locations are represented in the Doppler-delay plane (Fig. 12). We retained three-point location per kernel $\{(\eta,\tau)_1,\ldots,(\eta,\tau)_3\}$ of stronger contrast. These locations are ranged in the feature vector for training $\{\mathrm{FV}_1,\ldots,\mathrm{FV}_3\}$. This selection is made on the basis of contrast value.

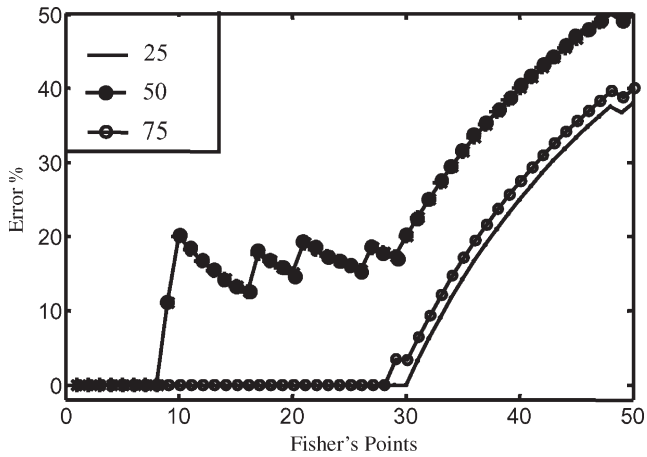


Fig. 13. Classification error for a bearing fault.

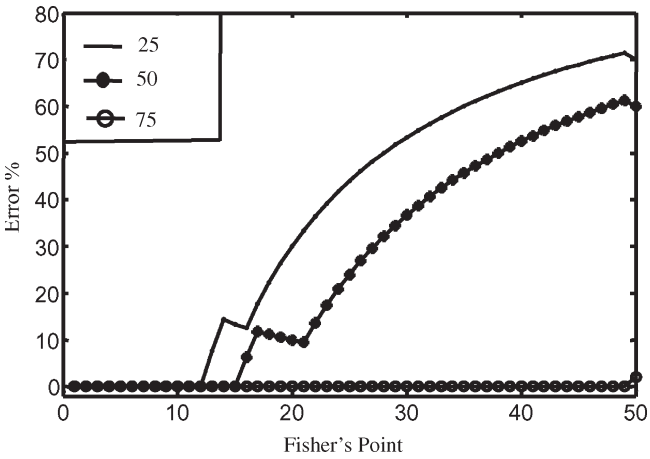


Fig. 14. Classification error for a stator fault.

B. Decision by Mahalanobis Distance

The feature vector of the signal to be classified, FV_x, was compared with the feature vectors of the training by using the Mahalanobis distance from (13). The decision rule of signal assignment is made by (14). The threshold $\beta = 0.4$ was tested successfully on several signals in order to obtain a correct classification. We tested the signals which have not been classified in the training set of the following three faults: bearing fault, stator fault, and rotor fault with various levels of load (25%, 50%, and 75%). Five signal examples are taken for each fault and for each load level. Thus, we will have $5 \times 3 = 15$-signal test for each fault. After extraction of feature vector of signal test, we took 50 points of each feature vector. The computation of the Mahalanobis distance $d_M(\mathrm{FV}_x)$ is done along these feature vectors. Fig. 13 shows that the error is null for the first eight points of the vector tests concerning the bearing fault. This is for the 25%, 50%, and 75% load levels. Finally, the bearing fault is only characterized by three points that are belonging to the first eight points. Consequently, the signals tested are identified with precision.

For the stator fault, Fig. 14 shows that the classification error is null for the first 12 points of the vector tests. Different load levels, i.e., 25%, 50%, and 75%, are considered, knowing that

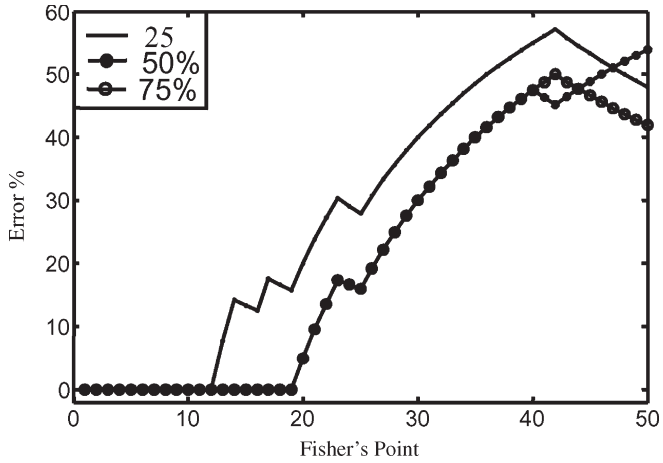


Fig. 15. Classification error for a rotor fault.

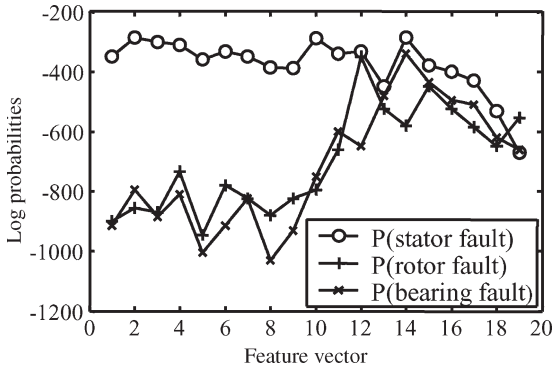


Fig. 16. HMM probabilities for a stator fault.

the stator fault is characterized by three points among the first twelve-points. Consequently, the signals tested are identified with precision.

Fig. 15 shows that for the rotor fault, the classification mean error is null for the first 13 points of the vector tests for different load levels (25%, 50%, and 75%), knowing that the rotor fault of the bars is characterized by three points among the first thirteen-points. Consequently, the signals tested are also identified with precision.

C. Decision by HMM

Three state left-to-right HMMs were used to model the fault features. Three Gaussian distributions were used in the output map for each state. We have tested the signals which do not belong to the training set of the following three faults: bearing fault, stator fault, and rotor fault with various levels of load (25%, 50%, and 75%). Five signal examples are taken for each fault and for each load level. Thus, we will have $5 \times 3 = 15$-signal test for each fault. After extracting the feature vectors of the signal to be classified, we took 20 points at each feature vector with various levels of load. The calculation of the HMM probabilities for the tested signals which do not belong to the training set of the three faults is done along these feature vectors.

Fig. 16 shows the logarithm of the probabilities for the first 20 points of the test feature vectors concerning the stator

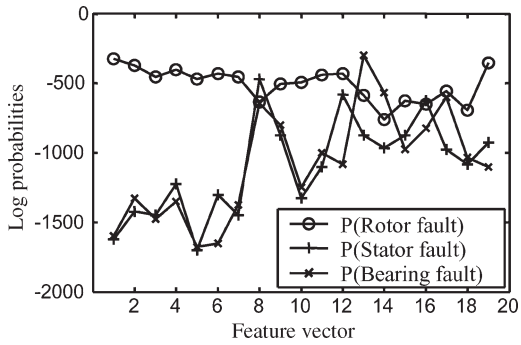


Fig. 17. HMM probabilities for a rotor fault.

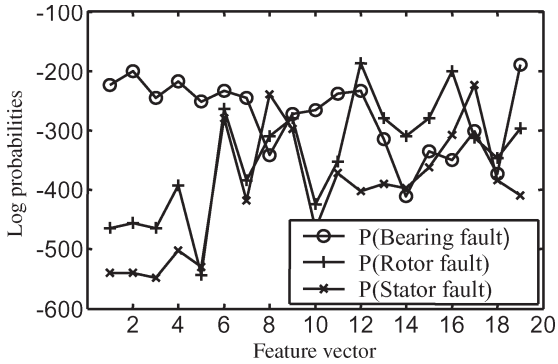


Fig. 18. HMM probabilities for a bearing fault.

fault. The probabilities of the training feature vector are clearly separable from the probabilities of the test feature vectors for only the first 11 points. Consequently, the signals tested are identified with precision.

For the rotor fault, different load levels (25%, 50%, and 75%) are considered. Fig. 17 shows that the probabilities of the training feature vector are clearly separable from the probabilities of the test feature vectors for the first seven points. Consequently, the rotor fault is identified with precision.

Concerning bearing fault, Fig. 18 shows that the logarithm of probabilities for the first five points from both training feature vector and test feature vectors is clearly separable. Consequently, the signals tested are identified with precision.

These results verify that the new scheme is able to detect and diagnose faults with very high accuracy, independent of the load condition and the type of the fault.

VI. Conclusion

In this paper, we proposed a new fault classification scheme of induction machine based on TFR and criterion decision [Mahalanobis distance and HMM]. We have shown that the classical TFRs which have a parametric kernel and *a priori* preset are not relevant for classification. Thus, we have based our classification on the ambiguity Doppler-delay plane where all the TFRs can be derived by a suitable choice of a kernel. In this application, the choice of the optimal kernel is crucial because it makes it possible to have an optimized TFR and, consequently, a precise classification of the signals. In this new scheme, the TFR is used to build the feature vectors for each type of fault. These faults were characterized by specific kernels. The

classification scheme was tested with experimental data collected from the stator current measurement from the induction machine drive. For the first criterion, the assignment of a not-classified signal was made by the Mahalanobis distance. The rejection threshold has been determined by tests. All faults were characterized by three points in the Doppler-delay plane. The mean error, of points badly classified, is null for the first eight points for the fault bearing, the first 13 points for the rotor fault, and the first 12 points for the stator fault. However, applying the criterion based on the "Mahalanobis distance" requires prior valuation of Doppler-delay positions, in the ambiguity plane, which makes the diagnostic process semiautomatic. In order to remedy this problem and make the process of diagnosis fully automatic, we have replaced the decision criterion based on the "Mahalanobis distance" by HHM. For the second criterion, the assignment of an unclassified signal was made with HMM models which have been trained to represent the feature vectors for different load conditions. These models were then used to classify the following three faults: stator, rotor, and bearing faults. It was shown that this method is very accurate in the classification of faults under different operating conditions. The classification by TFR provides results independent from the load level of stator current and very precise classification.

REFERENCES

[1] P. J. Tavner, B. G. Gaydon, and D. M. Ward, "Monitoring generators and large motors," *Proc. Inst. Elect. Eng.—B*, vol. 133, no. 3, pp. 169–180, May 1986.

[2] P. Vas, *Parameter Estimation, Condition Monitoring and Diagnosis of Electrical Machines*. Oxford, U.K.: Clarendon, 1993.

[3] G. B. Kliman and J. Stein, "Induction motor fault detection via passive current monitoring," in *Proc. Int. Conf. Elect. Mach.*, Cambridge, MA, 1990, pp. 13–17.

[4] G. B. Kliman, W. J. Premerlani, R. A. Koegl, and D. Hoeweler, "A new approach to on-line turn fault detection in AC motors," in *Conf. Rec. IEEE IAS Annu. Meeting*, San Diego, CA, 1996, pp. 687–693.

[5] H. Su and K. T. Chong, "Induction machine condition monitoring using neural network modeling," *IEEE Trans. Ind. Electron.*, vol. 54, no. 1, pp. 241–249, Feb. 2007.

[6] A. Ordaz-Moreno, R. de Jesus Romero-Troncoso, J. A. Vite-Frias, J. R. Rivera-Gillen, and A. Garcia-Perez, "Automatic online diagnosis algorithm for broken-bar detection on induction motors based on discrete wavelet transform for FPGA implementation," *IEEE Trans. Ind. Electron.*, vol. 55, no. 5, pp. 2193–2202, May 2008.

[7] E. Schaeffer, E. Le Carpentier, E. H. Zaïm, and L. Loron, "Diagnostic des entrainements électriques: Détection de courts-circuits statoriques dans la machines asynchrone par identification paramétrique," in *17ième colloque GRETSI sur le traitement du signal et des images*, Vannes, France, Sep. 13–17, 1999, vol. 4, pp. 1037–1040.

[8] O. Ondel, E. Boutleux, and G. Clerc, "A method to detect broken bars in induction machine using pattern recognition techniques," *IEEE Trans. Ind. Appl.*, vol. 42, no. 4, pp. 916–923, Jul./Aug. 2006.

[9] F. Filippetti, G. Franceschini, C. Tassoni, and P. Vas, "Recent developments of induction motor drives fault diagnosis using AI techniques," *IEEE Trans. Ind. Electron.*, vol. 47, no. 5, pp. 994–1003, Oct. 2000.

[10] M. Wang, G. I. Rowe, and A. V. Mamishev, "Classification of power quality events using optimal time–frequency representations—Part 1: Application," *IEEE Trans. Power Del.*, vol. 19, no. 3, pp. 1496–1503, Jul. 2004.

[11] L. Abdesselam and C. Guy, "Diagnosis of induction machine by time frequency representation and hidden Markov modeling," in *Proc. IEEE SDEMPED*, Kraków, Poland, Sep. 7–9, 2007, pp. 272–276. CD-ROM.

[12] B. Yazıcı and G. B. Kliman, "An adaptive statistical time–frequency method for detection of broken bars and bearing faults in motors using stator current," *IEEE Trans. Ind. Appl.*, vol. 35, no. 2, pp. 442–452, Mar./Apr. 1999.

[13] M. Davy and C. Doncarli, "Optimal kernels of time–frequency representations for signal classification," in *Proc. IEEE-SP Int. Symp. Time-Freq. Time-Scale Anal.*, 1998, pp. 581–584.

[14] C. Heitz, "Optimum time–frequency representations for the classification and detection of signals," *Appl. Signal Process.*, vol. 2, no. 3, pp. 124–143, 1995.

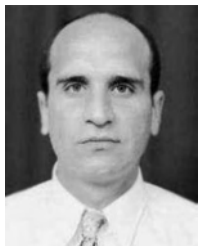
[15] B. W. Gillespie and L. Atlas, "Optimizing time–frequency kernels for classification," *IEEE Trans. Signal Process.*, vol. 49, no. 3, pp. 485–496, Mar. 2001.

[16] A. Lebaroud and G. Clerc, "Time–frequency classification applied to induction machine faults monitoring," in *Proc. IECON*, Paris, France, Sep. 6–9, 2006, pp. 5051–5056. CD-ROM.

[17] B. W. Gillespie and L. Atlas, "Data-driven time–frequency classification techniques applied to tool-wear monitoring," in *Proc. Acoust., Speech, Signal Process. Conf.*, 2000, vol. 2, pp. 2045–2048.

[18] P. C. Mahalanobis, "On the generalized distance in statistics," in *Proc. Nat. Inst. Sci. India*, 1936, vol. 12, pp. 49–55.

[19] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.



Abdesselam Lebaroud was born in 1969. He received the Engineer's degree in electrical engineering from Boumerdes University, Algeria, in 1993 and the Ph.D. degree from the Ampère Laboratory, Université Claude Bernard Lyon 1 Villeurbanne, France, in 2007.

Since 1999, he has been with the University of Skikda, Skikda, Algeria. He is also with the Laboratory LEC, University of Constantine, Constantine, Algeria. His current research interests include automatic diagnosis of electrical machinery faults.

Guy Clerc (M'00) was born in 1960. He received the Engineer's degree and the Ph.D. degree in electrical engineering from the Ecole Centrale de Lyon, Lyon, France, in 1984 and 1989, respectively.

He is currently a Professor in electrical engineering with the "Université Claude Bernard Lyon I," Villeurbanne, France, working on the control and diagnosis of induction machines at the Ampère Laboratory.