Comparative Study of Time-Frequency Methods for the Detection and Categorization of Intermittent Faults in Electrical Drives

S. Sajjad H. Zaidi, Wesley G. Zanardelli, Selin Aviyente and Elias G. Strangas

Abstract— **The detection of non-catastrophic faults in conjunction with other factors can be used to determine the condition and remaining life of an electric drive. As the frequency and severity of these faults increase, the working life of the drive decreases, leading to eventual failure. In this work, four methods to identify developing electrical faults are presented and compared. They are based on the Short-Time Fourier transform, Undecimated Wavelet analysis, Wigner and Choi Williams Distribution of the field oriented currents in PMAC drives. The different fault types are classified by developing a linear discriminant classifier based on the transform coefficients. The comparison is based on the number of correct classifications and Fisher's discriminant ratio.**

Index Terms— **Intermittent Fault Detection, Permanent Magnet AC Drives, Undecimated Discrete Wavelet Transform, Short-Time Fourier Transform, Choi-Williams, Wigner Distribution.**

I. Introduction

The early, efficient, and correct diagnosis of a failure in an electrical drive can lead to mitigation and continued operation, or, at least, notification of the operator before a fault becomes catastrophic. Prognosis of a failure often involves the detection and recognition of signs of early, minor faults; these early faults are intermitted before they become severe and permanent.

In previous work we have discussed such fault recognition using standard methods [1]. We recognize, though, the fact that although successful methods have been presented, only few studies have compared the efficiency and accuracy of the proposed methodologies. Some of the differences between methods are:

1) Analysis techniques. In this work as well as in work by others e.g. [2], a number of time-frequency methods were used to to analyze signals from faulted systems,
2) Analysis windows [3],
3) Classification methods, e.g. nearest neighbor rule or linear discriminant function [2].

In the approach presented in this work we study the effect of the analysis methodology on the overall accuracy of the fault detection and categorization. We selected as the four analysis methods to compare the Short-Time Fourier Transfor, Undecimated Wavelet Transformation, Wigner Distribution, and Choi-Williams Distribution. We chose rectangular windows and the k-means as the sole categorization methods. We also limited the work to using the "best" common parameters (data sample, levels of analysis etc.). Determining the fault type and severity requires the collection of signals from both healthy and faulted drives, processing of the signals to extract their salient characteristics, and finally the detection and classification of faults. For calculating the coefficients of the linear discriminant function and as well as test signals, we used the data from a set of previous, extensive experiments on a PMAC motor, discussed in [1].

W. G. Zanardelli is with the Advanced Electrical Group, Delphi Steering, Saginaw, MI 48601, USA (e-mail: wes.zanardelli@ieee.org)

S. Sajjad H. Zaidi, S. Aviyente and E. G. Strangas are with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA (e-mail: strangas@egr.msu.edu; aviyente@egr.msu.edu).

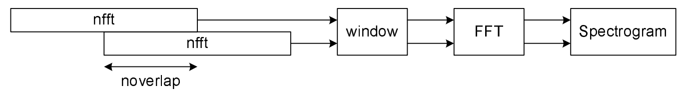

Fig. 1. STFT Block Diagram

II. Analysis methods

In this work, fault detection and classification results based on four time-frequency distributions are compared. These distributions are the Short-Time Fourier Transform (STFT), Undecimated Discrete Wavelet Transform (UDWT), Wigner (WVD), and Choi-Williams Distribution (CWD).

A. Short-Time Fourier SeriesTransform [4]

The STFT [4] is an extension of the Fourier transform, allowing for the analysis of non-stationary signals. Here, a signal is divided into small time windows, and each is analyzed using the Fast Fourier Transform as follows:

$$STFT(t,f) = \int h(t-\tau)s(\tau)e^{-j2\pi f\tau}d\tau, \tag{1}$$

where h is the window function. This formulation provides localization in time, while simultaneously capturing frequency information. The time-frequency tiling for the STFT is uniform across time and frequency. In the implementation of the STFT, a design tradeoff must be made between time and frequency resolution. A block diagram for the STFT algorithm is shown in Fig. 1, where *nfft* is the length of the FFT, *noverlap* is the number of samples the two frames overlap, and *window* is a weighting vector applied to the FFT input.

B. Undecimated Wavelet analysis [5]

This is another common tool used for non-stationary signals. The Discrete Wavelet transform (DWT) has greater flexibility than the STFT. Different basis functions, or mother wavelets, may be used in wavelet analysis while the basis functions for Fourier analysis are always complex exponentials. Unlike sinusoids, wavelets have finite duration and their energy is localized around a point in time. One can choose, or design a wavelet to achieve the best results for a specific application. The UDWT [6], [7] is a shift-invariant representation of the DWT. Determining the fault type and severity requires the collection of signals from both healthy and faulted drives, processing of the signals to extract their salient characteristics, and finally the detection and classification of faults.

C. Wigner Distribution

[4] Wigner distribution of a signal $s(t)$ is defined as:

$$W(t,\omega) = \int s(t+\frac{\tau}{2})s^*(t-\frac{\tau}{2})d\tau \tag{2}$$

Wigner distribution gives the energy distribution of the signal as a function of time and frequency. It is known to have high time-frequency resolution and does not suffer from the temporal vs. frequency resolution tradeoff encountered in STFT. Wigner distribution satisfies both the time and frequency marginals, and preserves the energy of the underlying signal. The major shortcoming of Wigner distribution occurs for multicomponent signals in terms of the cross-terms. The cross-terms occur due to the bilinear nature of the Wigner distribution and can sometimes obstruct the actual energy distribution.

D. Choi-Williams Distribution

Choi-Williams distribution of a signal $s(t)$ is defined as:

$$\begin{aligned} C(t,\omega) = \quad & \iiint \phi(\theta,\tau)s(u+\tfrac{\tau}{2})s^*(u-\tfrac{\tau}{2}) \\ & e^{j(\theta u - \theta t - \tau\omega)} du\, d\theta\, d\tau, \end{aligned} \tag{3}$$

where $\phi(\theta,\tau) = \exp(-\frac{(\theta\tau)^2}{\sigma})$ is the kernel function that acts as a filter on the signal's autocorrelation function. This distribution can be thought of as a filtered/smoothed version of the Wigner distribution and the amount of smoothing is controlled by σ. This smoothing removes the cross-terms seen in the Wigner distribution at the expense of reduced resolution.

III. Pattern Recognition

Once the analysis coefficients have been computed, the presence of a fault is detected based on thresholding the energy of the coefficients across the different frequency bands. These energy values are formed into vectors that provide the input features for the classification algorithm. The next step is to categorize these feature vectors, with the categories corresponding to one of several known fault types.

In this work, we consider two different classification methods, linear discriminant classifier and k-means classifier.

A. Linear Discriminant Classifier

The linear discriminant classifiers are trained on the input feature vectors since no *a priori* knowledge of a probability distribution for the sample points is assumed. The feature space is divided into K regions, each having its own weighting coefficients. Based on the training data, the linear discriminant functions (4) [8] are defined as:,

$$\begin{aligned} D_k(\mathbf{x}) = x_1\alpha_{1k} + x_2\alpha_{2k} +, \ldots, + x_N\alpha_{Nk} + \alpha_{N+1,k} \\ k = 1, 2, \ldots, K \end{aligned} \tag{4}$$

where $\mathbf{x}$ is the N-dimensional feature vector and α are the normalized weighting coefficients for the k-th class. Linear discriminant functions were chosen for the algorithm since they are computationally efficient. A sample vector belongs to a particular class if the discriminant function is greater for that class than for any other class, i.e., $\mathbf{x}_i$ belongs to class C_j if

$$D_j(\mathbf{x}) > D_k(\mathbf{x}) \quad \text{for every } k \neq j.$$

The weighting coefficients are adjusted from their initial guess through a training procedure using sample vectors for which the proper classification is known. The algorithm for this procedure makes adjustments to the weighting coefficients until each sample vector is correctly classified.

Young and Calvert [8] show that this training algorithm will converge in a finite number of steps. When a sample vector is correctly classified, no adjustment to the weighting coefficients is made. When a sample vector is incorrectly classified, or

$$D_j(\mathbf{x}) \leq D_l(\mathbf{x}),$$

where

$$D_l(\mathbf{x}) = \max_{l \neq j}\left[D_1(\mathbf{x}), \ldots, D_K(\mathbf{x})\right],$$

adjustments are made to $\boldsymbol{\alpha}_j$ (5) and $\boldsymbol{\alpha}_l$ (6) only,

$$\boldsymbol{\alpha}_j(i+1) = \boldsymbol{\alpha}_j(i) + a\mathbf{x}_i \tag{5}$$

$$\boldsymbol{\alpha}_l(i+1) = \boldsymbol{\alpha}_l(i) - a\mathbf{x}_i, \tag{6}$$

where a is a gain constant.

Discriminant functions have minimal storage requirements after the training phase since, for each class, only a single vector of weighting coefficients need to be stored. Storage of training samples is no longer required during the classification phase. For multiclass problems $(K > 2)$ the classes are linearly separable if linear discriminant functions $D_1(\mathbf{x}), \ldots, D_K(\mathbf{x})$ exist, such that (7) is true.

$$D_j(\mathbf{x}) > D_k(\mathbf{x}) \quad \text{for every } \mathbf{x} \text{ in } C_j \text{ and all } k \neq j \tag{7}$$

B. k-means Classification

k-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. These centroids can be chosen in different ways such as choosing k random points or using the means from training data as centroids. The next step is to take each point belonging to a given data set and associate it to the nearest centroid. In pattern recognition, k-means is a method for classifying objects based on closest training examples in the feature space. The distance is usually measured using the Euclidean distance. When all of the data points have been assigned to clusters, this procedure is repeated by measuring the means of each cluster and using these as the centroids for clustering. This procedure is repeated until there are no changes in the k centroids.

In this paper the initial centroids are estimated based on the training data and each new fault sample is assigned to the class to which it is closest to, i.e. $C_j = \mathrm{argmin}_{i\in\{1,2,\ldots,C\}}\|x_j - m_i\|$, where m_i is the different class averages obtained from the training data.

C. Evaluation of Different Features: Fisher Discriminant Ratio

Fisher discriminant ratio is a measure that quantifies the discrimination capacity of features regardless of the classifier. In this paper, we are considering four different signal transforms, i.e. four different ways of extracting features from the fault signals. The goodness of the features can be quantified through the classification accuracy or through Fisher discriminant ratio, which is independent of the classifier. The classification results obtained by the different classifiers give information about which class each fault belongs to. However, these results are not necessarily informative of the separation of the different fault classes and how the time-frequency coefficients cluster with the different methods. In order to quantify the discrimination power of the different time-frequency features, we propose to use Fisher's discriminant ratio. In this paper, we will use both methods and show that the results are consistent.

Fisher's discriminant ratio [9] can be both used as a classification method and a class separability criterion. In this paper, we will adopt the latter approach and use it as an indicator of the class separability for the four different time-frequency analysis methods. For multi-class data, Fisher's discriminant ratio is defined as:

$$\begin{aligned} F(\mathbf{X}) &= \frac{S_B}{S_W} = \frac{\|\sum_{i=1}^{C} K_i(\mathbf{m}_i-\mathbf{m})(\mathbf{m}_i-\mathbf{m})^T\|_2^2}{\sum_{i=1}^{C} s_i^2} \\ \mathbf{m_i} &= \frac{1}{K_i}\sum_{j\in C_i}\mathbf{x}_j \\ s_i^2 &= \frac{1}{K_i}\sum_{j\in C_i}\|\mathbf{x}_j-\mathbf{m}_i\|_2^2 \\ \mathbf{m} &= \frac{1}{K}\sum_{i=1}^{K}\mathbf{x}_i \end{aligned} \tag{8}$$

Fisher discriminant ratio can be interpreted as the ratio of the inter-class distance to inner-class scatter. Fisher's criterion is motivated by the intuitive idea that the discrimination power is maximized when the centroids of different classes are spatially distributed as far away as possible from each other and the samples from the same class are as closely distributed to each other as possible.

In the application proposed in this paper, this criterion will be applied to the time-frequency features extracted by the four methods to compare the discrimination power of the different transforms. In this case, the number of classes $C = 4$, and $\mathbf{X}$ is the matrix of the time-frequency coefficients.

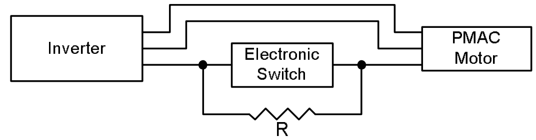

Fig. 2. Experimental Mimic of Increased Contact Resistance Fault

IV. Types of Faults

The test machine used in this analysis is a six-pole surface mounted PMAC machine for a $12V$ automotive application. Its rated power and no-load speed are approximately $1HP$ and $3000rpm$, respectively. The machine is operated in a vector drive with the torque angle set to $\pi/2$ [10]. This mode of operation minimizes losses in the machine, and is suitable for operating speeds up to the base speed. In all the tests in this paper the torque-producing component of the stator current command was $i_{qs}^* = 0.3pu$, and the flux-weakening component $i_{ds}^* = 0$; the speed is held constant by the dynamometer and is set to $400rpm$. The resultant load level is approximately $1N\cdot m$.

The electrical faults imposed in this work are not periodic. The time interval between successive faults is random. In the case of the missing gear teeth, the fault occurs once per mechanical revolution.

Intermittent Increased Contact Resistance. The first fault explored in this work is an intermittently increased contact resistance between the motor and the controller. An experiment was designed to simulate this. The fault is achieved by adding the parallel combination of a normally closed switch, and a resistance in series with one of the motor phases as shown in Fig. 2.

The fault is initiated by opening the switch for a short time interval, causing current to flow through the resistance. The switch is described in the Section on the Experimental Setup.

Turn-Terminal Short. The second fault explored in this work is an insulation failure in the stator windings of the motor. The machine used in this experiment has multiple parallel stranded-wire windings per phase, each with several coils in series. To simulate this fault in the experimental setup, at one point in a single strand of one of the windings, the insulation was removed, and a normally open switch was added between this point and the corresponding phase terminal of the motor. The fault was initiated by momentarily closing the switch, causing current to be split between the intended path and the switch.

The fault is initiated as the phase current command rises to 95% of its peak amplitude. Tests with fault durations of $5ms$ and $10ms$ have been performed to investigate invariance of the results with respect to this parameter.

V. Experiment and Experimental Results

To collect data from healthy and faulted drives, it is common to test a drive in its healthy state, and subsequently test it under a set of faulted conditions. This generally requires permanent

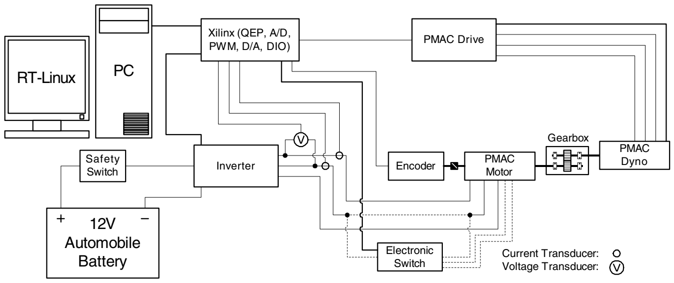

Fig. 3. Experimental Setup Block Diagram

and repeatable modification or damage to it, in order to change its characteristics Alternatives to controlled modifications include simply operating the drive until failure [11], or designing experiments where faults could be introduced and cleared in models during operation. The most acceptable method for data collection to develop and test fault detection strategies remains experimental [1], [11]–[21].

In this paper we are comparing analysis methods that are useful in detecting the transient phenomena associated with the inception and clearing of faults. In order to collect the data required to design and test the algorithms, an experiment was developed to mimic the electrical faults. Data from these faults were extracted from continuously sampled signals. Issues specific to transient fault detection include methods to make the detection invariant to the duration of the fault, and to the starting point of the sample with respect to the fault inception. The algorithms compared are based on the analysis of the torque producing component of the field oriented stator currents in AC Permanent Magnet motors. The energy of the transform coefficients are thresholded to extract low dimensional feature vectors. These vectors are then used in the fault detection algorithms. These algorithms can be implemented in an online system, using extra processing time in the motor controller. Such an online implementation mimics the laboratory setup, with certain variations. The laboratory controller is implemented on a PC rather than on a DSP, but the sensors, measured quantities, control algorithm, cycle duration and the control variables (e.g. i_q and i_d) are the same.

A. Experimental Setup

A PC running RT-Linux was used as the drive controller. A custom Xilinx FPGA based I/O board was developed for the interface with the drive hardware. A block diagram for the experimental setup is shown in Fig. 3.

Two phase currents were measured using current transducers with rated accuracy of 0.45% and bandwidth of $0-200kHz$. A quadrature encoder with 1024 counts per revolution (4096 for quadrature) and an index pulse was used to measure the rotor position. A bi-directionally conducting electronic switch was designed to initiate faults in the stator.

B. Fault Detection and Classification

The detection and classification of the faults described in Section IV are based on the analysis of the stator currents of the machine. Rather than analyzing the three phase stator currents independently of each other, the field oriented currents i_{qs} and i_{ds} are used. Together, i_{qs} and i_{ds} are a complete representation of the stator currents, however, it has been determined experimentally that through analysis of i_{qs} only, accurate fault detection and classification can be achieved.

The fault detection and classification algorithm used in this work was tested with two sets of data, or analysis coefficients. These coefficients were the distributions resulting fromthe analysis of a subset of the measured q-axis current, i_{qs}, respectively. For the STFT, WD and CWD, *nfft*=64, *noverlap*=48, and a 64-point rectangular window was used. The resultant STFT had 33 frequency bands. The two outermost bands corresponding to the DC and 10kHz components of i_{qs} were discarded. The algorithm was based on the remaining 31 frequency bands. For the UDWT, the Daubechies D4 mother wavelet was used and decomposition was performed for 6 levels.

The algorithm has two parts; a detection phase and a classification phase. The detection phase of the algorithm is based on thresholding the energy of the analysis coefficients.

The classification phase was based on linear discriminant analysis. This phase is implemented when the criterion for detection is met. Two distinct events were assigned to each intermittent fault, since the fault duration is random; one corresponding to the inception, and the other to the clearing of the fault. The advantage of two separate classes for each fault is invariance to the duration of the fault. Additionally, a postprocessing algorithm to verify subsequent classification of the beginning and end of the same fault type would further confirm existence of the fault.

To train the algorithms, thereby determining the weighting coefficients, data were used from eight electrical fault experiments from each of the following operating conditions: Healthy, increased contact resistance, and turn-terminal short. The analysis coefficients begin at the index of the local maxima of the energy where the threshold was exceeded. The developed algorithm recognizes the fault inception and clearing as distinct events with different characteristics, while the duration of the fault plays no role in the recognition of the fault.

Following the training of the weighting coefficients, data that had not been used in the training algorithm were tested. Two data sets from each of the following electrical fault conditions were tested: Healthy, increased contact resistance ($2.14pu$, $2.80pu$, $4.03pu$, $6.33pu$, and $15.84pu$), and turn-terminal short.

C. Results

The results from two typical test cases for the electrical faults based on analysis of the coefficients resulting from each method are shown in Figs. 4–7.

The top rows of Figs. 4–7 show results from one of the increased contact resistance faults. The bottom rows have results from the turn-terminal short. The left columns show the measured torque-producing component of the current, i_{qs}, for each test case. The center columns show the corresponding coefficients (spectrogram, UDWT of the data etc.)to the left.

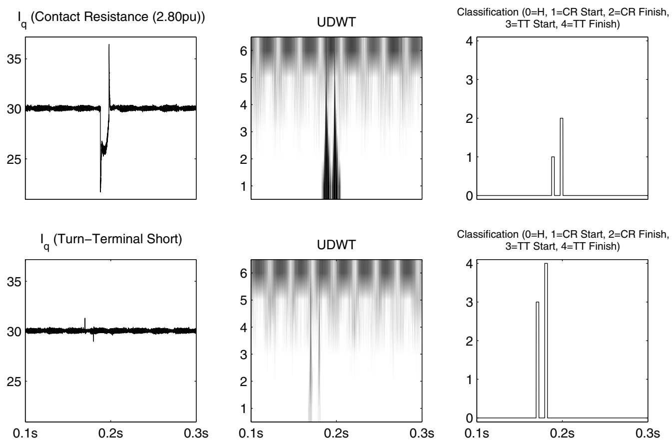

Fig. 4. Typical Results for Electrical Faults Based on Analysis of the STFT Coefficients

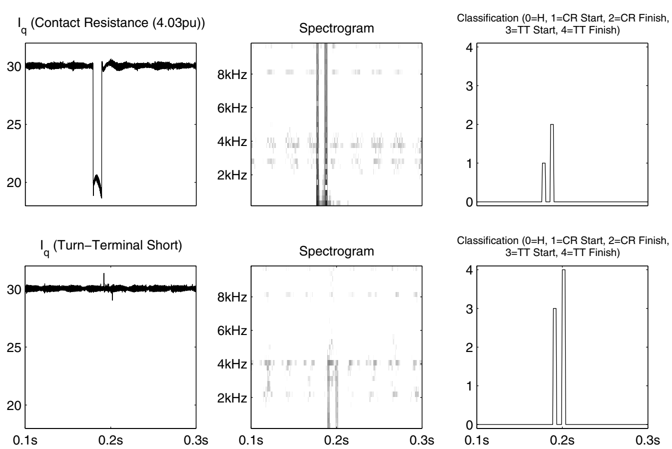

Fig. 5. Typical Results for Electrical Faults Based on Analysis of the UDWT Coefficients

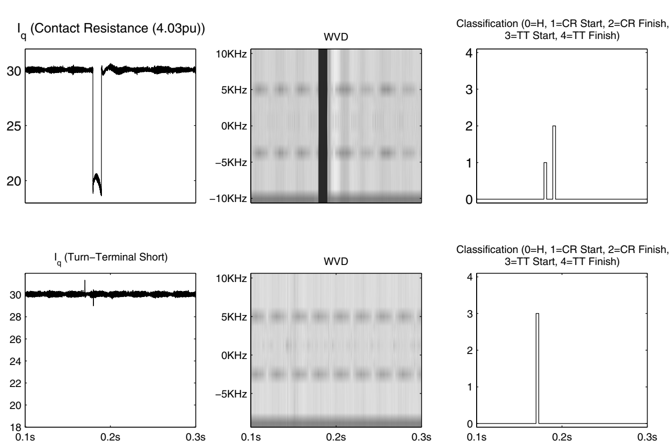

Fig. 6. Typical Results for Electrical Faults Based on Wigner distribution

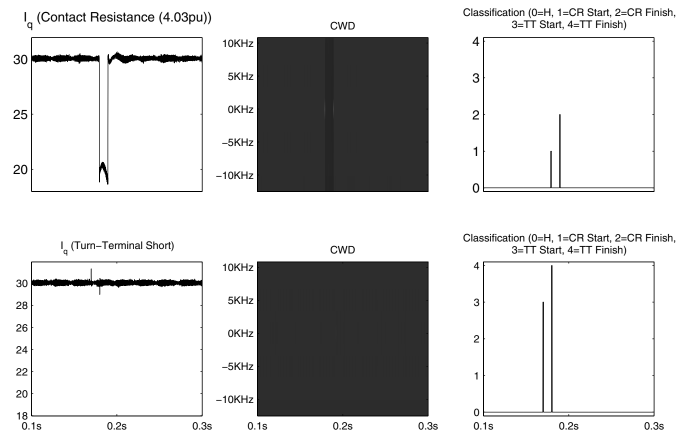

Fig. 7. Typical Results for Electrical Faults Based on He Choi-Williams Distribution

The right columns are the output of the detection and classification algorithm with 0=healthy, 1=beginning of increased contact resistance fault, 2=end of increased contact resistance fault, 3=beginning of turn-terminal short, and 4=end of turn-terminal short. The results presented are based on new data not used in the training set.

In Figs. 4–7, it can be seen that the amplitude of the analysis coefficients is increased at the points of inception and clearing of the faults. The increased energy at these points meets the criterion for detection. The results shown indicate correct classification of the inception and clearing of both faults explored in this work. The performance of the algorithm can be observed in Tables I–II.

In the case based on the STFT coefficients, there were 3 false detections. All fault events were detected, and 15/24 were classified correctly. In the case based on the UDWT coefficients, there were no false detections. All fault events were detected and classified correctly except for the clearing of one of the $2.14pu$ increased contact resistance faults. The $2.14pu$ increased contact resistance fault, however, is outside of the range of values for which the algorithm was trained.

The results indicate that the two classification algorithms, nearest neighborhood and linear discriminant analysis, give similar performance results, with linear discriminant analysis giving better results for the Wigner distribution. The linear discriminant analysis derives a set of coefficients that describe the high-dimensional surfaces that separate the different classes. These coefficients are saved for classifying the test samples. The nearest neighborhood (or k-means) as implemented in this paper only uses the average of each class from the training set for classifying the test samples. Since nearest neighborhood uses much less information from the training set, it performs slightly worse than linear discriminant analysis. The results of nearest neighborhood classifier can be improved by using an iterative approach for determining the clusters as each new test sample is made available. This iterative approach would improve the classification accuracy at the expense of computation time.

Table III, summarizes and compares the performance of the classification schemes. The classification results show that

Test Description	STFT Total / Detected / Classified Correctly	UDWT Total / Detected / Classified Correctly	Wigner Total / Detected / Classified Correctly	C-W Total / Detected / Classified Correctly
Contact Resistance ($2.14pu$) inception	2 / 2 /2	2 / 2 /2	2 / 2 /0	2 / 2 /0
Contact Resistance ($2.14pu$) clearing	2 / 2 / 0	2 / 2 / 1	2 / 2 /0	2 / 2 /0
Contact Resistance ($2.80pu$) inception	2 / 2 /2	2 / 2 /2	2 / 2 /2	2 / 2 /0
Contact Resistance ($2.80pu$) clearing	2 / 2 / 0	2 / 2 / 2	2 / 2 /0	2 / 2 /0
Contact Resistance ($4.03pu$) inception	2 / 2 /1	2 / 2 /2	2 / 2 /2	2 / 2 /2
Contact Resistance ($4.03pu$) clearing	2 / 2 / 2	2 / 2 / 2	2 / 2 /2	2 / 2 /2
Contact Resistance ($6.33pu$) inception	2 / 2 / 1	2 / 2 / 2	2 / 2 /2	2 / 2 /2
Contact Resistance ($6.33pu$) clearing	2 / 2 / 1	2 / 2 / 2	2 / 2 /2	2 / 2 /2
Contact Resistance ($15.84pu$) inception	2 / 2 / 0	2 / 2 / 2	2 / 2 /2	2 / 2 /2
Contact Resistance ($15.84pu$) clearing	2 / 2 / 2	2 / 2 / 2	2 / 2 /2	2 / 2 /2
Turn-Terminal Short inception	2 / 2 / 2	2 / 2 / 2	2 / 2 /2	2 / 2 /2
Turn-Terminal Short clearing	2 / 2 / 2	2 / 2 / 2	2 / 1 /0	2 / 2 /2

TABLE I
ALGORITHM PERFORMANCE FOR ELECTRICAL FAULTS BASED ON LINEAR DISCRIMINANT CATEGORIZATION

Test Description	STFT Total / Detected / Classified Correctly	UDWT Total / Detected / Classified Correctly	Wigner Total / Detected / Classified Correctly	C-W Total / Detected / Classified Correctly
Contact Resistance ($2.14pu$) inception	2 / 2 /0	2 / 2 /2	2 / 2 /2	0 / 2 /0
Contact Resistance ($2.14pu$) clearing	2 / 2 / 0	2 / 2 / 0	2 / 2 /0	2 / 2 /0
Contact Resistance ($2.80pu$) inception	2 / 2 /0	2 / 2 /2	2 / 2 /2	1 / 2 /2
Contact Resistance ($2.80pu$) clearing	2 / 2 / 0	2 / 2 / 2	2 / 2 /0	2 / 2 /0
Contact Resistance ($4.03pu$) inception	2 / 2 /2	2 / 2 /2	2 / 2 /2	2 / 2 /2
Contact Resistance ($4.03pu$) clearing	2 / 2 / 1	2 / 2 / 2	2 / 2 /2	2 / 2 /2
Contact Resistance ($6.33pu$) inception	2 / 2 / 2	2 / 2 / 2	2 / 2 /2	2 / 2 /2
Contact Resistance ($6.33pu$) clearing	2 / 2 / 1	2 / 2 / 2	2 / 2 /2	2 / 2 /2
Contact Resistance ($15.84pu$) inception	2 / 2 / 2	2 / 2 / 2	2 / 2 /2	2 / 2 /2
Contact Resistance ($15.84pu$) clearing	2 / 2 / 0	2 / 2 / 2	2 / 2 /2	2 / 2 /2
Turn-Terminal Short inception	2 / 2 / 2	2 / 2 / 2	2 / 2 /1	2 / 2 /2
Turn-Terminal Short clearing	2 / 2 / 2	2 / 2 / 2	2 / 1 /0	2 / 2 /2

TABLE II
ALGORITHM PERFORMANCE FOR ELECTRICAL FAULTS BASED ON NEAREST NEIGHBORHOOD CATEGORIZATION

Fisher discriminant ratio is a good metric for evaluating the 'goodness' of different time-frequency transforms for extracting features. The Fisher discriminant ratios computed for the four transforms are consistent with the classification accuracy obtained by the two methods.

Based on the classification results, it is clear that UDWT gives the best features compared to the other transform methods. This is due to the fact that the faults are transient in nature and the wavelet transform at the right scales is more efficient in extracting this transient structure compared to the other transforms. We can improve the results of the other time-frequency methods by choosing a smaller range of frequencies and time points. The results obtained by the STFT are the worst, since STFT suffers from the time-frequency resolution tradeoff due to windowing.

VI. REAL-TIME IMPLEMENTATION

The proposed four different time-frequency analysis methods differ in their implementation. STFT and UDWT are linear transforms and thus require less number of computations, $O(N \log N)$ for STFT and $O(N)$ for UDWT. WVD and CWD, on the other hand, are bilinear transforms of the signal and thus require more computations, $O(N^2 \log N)$. Depending on the application and the limitations of the system, linear transform methods such as STFT and UDWT may be preferred over the bilinear ones.

VII. CONCLUSIONS

Four different methods for the time-frequency analysis are compared for use in an algorithm for the detection and categorization of faults in electrical drives. The algorithm consists of analysis and categorization of the inception and clearing of intermittent faults.

The categorization algorithm uses a linear discriminant function and is trained using a set of operating conditions which include healthy drives and samples of faulted drives. An exhaustive set of such conditions is necessary to develop a robust algorithm.

The four candidate methods are applied to the same training and test data sets and categorization method. The results are compared in terms of three measures: the number of correct and false categorizations, the confidence in their results

Distribution	Detection		Fisher's Coefficient	Linear Discriminant Analysis					Nearest Neighborhood Analysis				
	Correct	Incorrect		Inception		Clearing		percent correct	Inception		Clearing		percent correct
				correct	incorrect	correct	incorrect		correct	incorrect	correct	incorrect	
UDWT	28	0	2.98	14	0	13	1	96	14	0	12	2	93
Choi-Williams	28	0	1.79	10	4	12	2	79	12	2	10	4	79
Wigner	27	1	1.6788/ 1.6798*	12	2	9	5	75	10	4	9	5	68
STFT	25	3	1.01/ 1.4*	10	4	8	6	64	10	4	8	6	64

* the calculation for the Fisher coefficient for the first number used all events, categorized in their appropriate cluster; for the second number, incorrect categorizations were discarded.

TABLE III
COMPARISON OF CATEGORIZATION METHODS APPLIED TO FOUR ANALYSIS METHODS. FOURTEEN FAULT EVENTS ARE ANALYZED, EACH CONSISTING OF A FAULT INCEPTION AND CLEARING, TO A TOTAL OF 28 EVENTS.

as measured by Fisher's discriminant ratio, and the ease of implementation in hardware. For the particulars of the implementations presented here and the evaluation criteria used, the Undecimated Wavelet Transform performed best, while the Short-Time Fourier Transform performed the worst.

REFERENCES

[1] W. G. Zanardelli, E. G. Strangas, and S. Aviyente, "Identification of intermittent electrical and mechanical faults in permanent magnet ac drives based on time-frequency analysis," *IEEE Transactions on Industry Applications*, to appear.

[2] T. Lobos, J. Rezmer, and T. Sikorski, "Advanced signal processing methods for analysis of non-stationary signals in power systems," in *2003 IEEE International Symposium on Intelligent Signal Processing*, 4-6 Sept. 2003, pp. 43 – 47.

[3] I. Tsoumas, E. Mitronikas, G. Georgoulas, and A. Safacas, "A comparative study of induction motor current signature analysis techniques for mechanical faults detection," in *SDEMPED 2005 - International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives Vienna, Austria*, 7-9 September 2005.

[4] L. Cohen, *Time-Frequency Analysis*. Prentice Hall, 1995.

[5] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms, A Primer*. Prentice Hall, 1998.

[6] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. Academic Press, 1999.

[7] P. Dutilleux, "An implementation of the "algorithme à trous" to compute the wavelet transform," in *Wavelets: Time-Frequency Methods and Phase Space, Proceedings of the International Conference*. Springer-Verlag, 1987.

[8] T. Y. Young and T. W. Calvert, *Classification, Estimation and Pattern Recognition*. American Elsevier Publishing Co., Inc., 1974.

[9] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. Wiley-Interscience, 2000.

[10] R. Krishnan, *Electric Motor Drives: Modeling, Analysis, and Control*. Prentice Hall, 2001.

[11] H. S. Liu, B. Y. Lee, and Y. S. Tarng, "Monitoring of drill fracture from the current measurement of a three-phase induction motor," *International Journal of Machine Tools and Manufacture*, vol. 36, pp. 729–738, Jun 1996.

[12] R. R. Schoen, B. K. Lin, T. G. Habetler, J. H. Schlag, and S. Farag, "An unsupervised, on-line system for induction motor fault detection using stator current monitoring," *IEEE Transactions on Industry Applications*, vol. 31, pp. 1280–1286, Nov–Dec 1995.

[13] A. Cash, T. G. Habetler, and G. B. Kliman, "Insulation failure prediction in AC machines using line-neutral voltages," *IEEE Transactions on Industry Applications*, vol. 34, pp. 1234–1239, Nov.–Dec. 1998.

[14] H. Nejjari and M. E. Benbouzid, "Monitoring and diagnosis of induction motors electrical faults using a current park's vector pattern learning approach," *IEEE Transactions on Industry Applications*, vol. 36, pp. 730–735, May–Jun. 2000.

[15] A. Bellini, F. Filippetti, G. Franceschini, C. Tassoni, R. Passaglia, M. Saottini, G. Tontini, M. Giovannini, and A. Rossi, "On-field experience with online diagnosis of large induction motors cage failures using MSCA," *IEEE Transactions on Industry Applications*, vol. 38, pp. 1045–1053, Jul–Aug 2002.

[16] O. Moseler and R. Isermann, "Application of model-based fault detection to a brushless DC motor," *IEEE Transactions on Industrial Electronics*, vol. 47, pp. 1015–1020, Oct 2000.

[17] X. Q. Liu, H. Y. Zhang, J. Liu, and J. Yang, "Fault detection and diagnosis of permanent-magnet DC motor based on parameter estimation and neural network," *IEEE Transactions on Industrial Electronics*, vol. 47, pp. 1021–1030, Oct 2000.

[18] C. Combastel, S. Lesecq, S. Petropol, and S. Gentil, "Model-based and wavelet approaches to induction motor on-line fault detection," *Control Engineering Practice*, vol. 10, pp. 493–509, May 2002.

[19] S. Pöyhönen, M. Negrea, P. Jover, A. Arkkio, and H. Hyötyniemi, "Numerical magnetic field analysis and signal processing for fault diagnostics of electrical machines," in *International Conference on Electrical Machines*, Aug. 2002.

[20] S. Rajagopalan, J. A. Restrepo, J. M. Aller, T. G. Habetler, and R. G. Harley, "Wigner-ville distributions for detection of rotor faults in brushless dc (BLDC) motors operating under non-stationary conditions," *IEEE International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives*, pp. 313–319, Sep 2005.

[21] C. C. Yeh, B. Mirafzal, R. J. Povinelli, and N. A. O. Demerdash, "A condition monitoring vector database approach for broken bar fault diagnostics of induction machines," in *International Electric Machines and Drives Conference*, May 2005, pp. 29–34.