Pattern Recognition Using Invariants Defined from Higher Order Spectra: 2-D Image Inputs

Vinod Chandran, *Member, IEEE*, Brett Carswell, Boualem Boashash, *Senior Member, IEEE*, and Steve Elgar, *Member, IEEE*

***Abstract*—A new algorithm for extracting features from images for object recognition is described. The algorithm uses higher order spectra to provide desirable invariance properties, to provide noise immunity, and to incorporate nonlinearity into the feature extraction procedure thereby allowing the use of simple classifiers. An image can be reduced to a set of one-dimensional (1-D) functions via the Radon transform, or alternatively, the Fourier transform of each 1-D projection can be obtained from a radial slice of the two-dimensional (2-D) Fourier transform of the image according to the Fourier slice theorem. A triple product of Fourier coefficients, referred to as the deterministic bispectrum, is computed for each 1-D function and is integrated along radial lines in bifrequency space. Phases of the integrated bispectra are shown to be translation- and scale-invariant. Rotation invariance is achieved by a regrouping of these invariants at a constant radius followed by a second stage of invariant extraction. Rotation invariance is thus converted to translation invariance in the second step. Results using synthetic and actual images show that isolated, compact clusters are formed in feature space. These clusters are linearly separable, indicating that the nonlinearity required in the mapping from the input space to the classification space is incorporated well into the feature extraction stage. The use of higher order spectra results in good noise immunity, as verified with synthetic and real images. Classification of images using the higher order spectra-based algorithm compares favorably to classification using the method of moment invariants.**

I. Introduction

HIGHER ORDER spectra [1]–[3] were introduced as spectral representations of cumulants or moments of ergodic processes, and are useful in the identification of nonlinear and non-Gaussian random processes, as well as deterministic signals [4]–[6]. Higher order spectral representations of real-valued deterministic signals may be expressed as products of Fourier coefficients at component frequencies and the conjugate of the Fourier coefficient at the sum frequency. For example, the bispectrum, $B(f_1, f_2)$, of a real-valued sequence, $x(n)$, may be defined by

$$B(f_1, f_2) = X(f_1)X(f_2)X^*(f_1 + f_2) \tag{1}$$

where $X(f)$ is the discrete-time Fourier transform of the sequence at frequency f. If the sequence $x(n)$ is random, an average of the deterministic bispectra for the set of realizations serves as an estimate of the bispectrum for the random sequence. The term *higher order spectrum* is used here to represent the deterministic or the averaged estimate, depending on the context. Higher order spectra have been used in image reconstruction [5], pattern recognition [6]–[8], image restoration [9], and edge detection [10]. The use of higher order spectra for feature extraction is motivated by the following.

1) Higher order spectra retain both amplitude and phase information from the Fourier transform of a signal, unlike the power spectrum. The phase of the Fourier transform contains important shape information [11].
2) For one-dimensional (1-D) patterns, the Fourier phase is a shape-dependent nonlinear function of the frequency, and higher order spectra can extract this information.
3) Finite-length patterns that are symmetric about their centers of support have a Fourier phase that is a linear function of frequency. Therefore, the shape information resides in the Fourier magnitudes. However, the Fourier magnitude for positive frequencies is an asymmetric function whose shape is related to the shape of the original input. Therefore, higher order spectra can still be used to extract features indirectly from Fourier magnitudes.
4) Higher order spectra are translation invariant because linear phase terms are cancelled in the products of Fourier coefficients that define them. Functions that can serve as features for pattern recognition can be defined from higher order spectra that satisfy other desirable invariance properties such as scaling, amplification, and rotation invariance.
5) Higher order spectra are zero for Gaussian noise and, thus, provide high noise immunity to features.
6) Multidimensional signals can be decomposed into 1-D projections. Transformations such as shift, scaling, or rotation of the multidimensional signal can be related to shift or scaling of the projections. Higher order spectral features derived from the projections can be used to

Manuscript received January 4, 1995; revised April 23, 1996. This work was supported by the Signal Processing Research Centre, Queensland University of Technology, Washington State University, and the U.S. Office of Naval Research. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Charles A. Bouman.

V. Chandran, B. Carswell, and B. Boashash are with the Signal Processing Research Centre, School of Electrical and Electronic Systems Engineering, Queensland University of Technology, Brisbane, Qld. 4001, Australia (e-mail: v.chandran@qut.edu.au).

S. Elgar is with the School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99164-2752 USA.

Publisher Item Identifier S 1057-7149(97)03084-4.

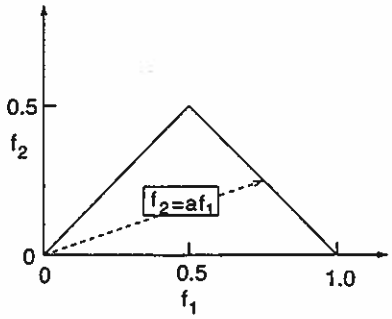

Fig. 1. Region of computation of the bispectrum. Features are obtained by integrating the complex bispectrum along a radial line with slope $= a$ (dashed line). The phase of this integral is translation and scale invariant.

derive invariant features for the multidimensional signal as demonstrated in the present study.

A global feature extraction method based on higher order spectral information that extracts shift, scale, and rotation invariant features is described in this paper. It permits the use of simple classifiers because most of the nonlinearity required in the mapping from the input space to the classification space is incorporated into the feature extraction stage. The new algorithm uses information in the deterministic bispectrum [4] or, equivalently, third-order correlations of the input, and is robust to additive Gaussian noise.

Other correlation-based [12]–[14] and higher-order-correlation-based methods [7] have been proposed that use matched filtering in the correlation domain or normalized energies in a higher order spectral domain representation. Matched filters fail to satisfy all the desired invariance properties. Normalized energies in higher order spectral representations discard shape-dependent information that is potentially useful for classification. The algorithm proposed here is designed to obtain greater sensitivity to shape by exploiting the structure in the bispectral domain.

The paper is organized as follows. Section II presents the algorithm for invariant feature extraction from 1-D sequences and Section III extends the algorithm to two-dimensional (2-D) images. Test results on bilevel synthetic and grey-scale images of simple machine tools and parts are presented in Section IV. The new higher order spectra-based image classification algorithm is compared to both the method of moment invariants and the Universida de Nova de Lisboa (UNL) [15] transform in Section V, followed by conclusions in Section VI.

II. Bispectral Invariants from 1-D Sequences

The bispectrum [4] of a 1-D *deterministic* sequence may be defined as in (1). It will be assumed that the sequence is oversampled for all scales of interest such that $X(f) = 0$ for $f \leq \frac{1}{2}$, with the frequency f normalized by the Nyquist frequency. The bispectrum is then uniquely defined within the triangle $0 \leq f_2 \leq f_1 \leq f_1 + f_2 \leq 1$, as shown in Fig. 1, by virtue of its symmetry properties.

The bispectrum (1) is a triple product of Fourier coefficients, and is a complex-valued function of two frequencies, similar to the power spectrum, $P(f) = X(f)X^*(f)$, which is a second-order product of Fourier coefficients and a function of only one frequency. Unlike the power spectrum, the bispectrum retains some information about the phase of the Fourier transform of the sequence. A symmetric sequence of finite extent (or a shifted version of it) has a Fourier transform whose phase is a linear function of frequency and, thus, the biphase (the phase of the bispectrum) is zero. For an asymmetric sequence, the phase of the Fourier transform is a nonlinear function of frequency and this nonlinearity is extracted by the biphase. For example, left and right asymmetric sequences will have opposite signs for the biphase. These properties form a basis for the use of the bispectrum in extracting features from 1-D patterns. Additionally, parameters can be defined from the bispectrum that are invariant to translation, scaling, DC-shifting, and amplification. In particular, the phase of the integrated bispectrum along a radial line of slope a (see Fig. 1) satisfies these properties. Parameters

$$P(a) = \arctan\left[\frac{I_i(a)}{I_r(a)}\right] \tag{2}$$

where

$$I(a) = I_r(a) + jI_i(a) = \int_{f_1=0+}^{1/(1+a)} B(f_1, af_1)\, df_1 \tag{3}$$

for $0 < a \leq 1$, can serve as features for 1-D patterns. The variables I_r and I_i refer to the real and imaginary parts of the integrated bispectrum, respectively, and $j = \sqrt{-1}$. It has been shown [6] that these parameters satisfy the desired invariance properties. Specifically, DC-level invariance is trivially satisfied because the zero-frequency coefficient is ignored. Translation invariance is satisfied by the bispectrum itself because a linear frequency-dependent phase shift in the Fourier transform coefficients as a result of translation is cancelled in the triple product in (1). Both the real and the imaginary parts of the integrated bispectrum are multiplied by the same real-valued constant upon amplification of the original sequence, leaving the phase unchanged. Scaling the original 1-D sequence results in an expansion or contraction of the Fourier transform that is identical along the f_1 and f_2 directions. Thus, the bispectral values along a radial line in bifrequency space map onto the same line upon scaling. The real and imaginary parts of the integrated bispectrum along a radial line are multiplied by identical real-valued constants upon scaling and therefore the phase of the integrated bispectrum is unchanged. Different features can be obtained for different values of a, provided the bispectrum has sufficient structure to it. Fig. 2 shows the procedure for extraction of invariant parameters from 1-D sequences.

Inputs containing finite-width patterns will not be strictly bandlimited. The discrete Fourier transforms (DFT's) of scaled versions of the input will then no longer be scaled versions in the frequency domain. Aliasing introduces error into the parameters $P(a)$, which are then no longer strictly scale invariant. One method of reducing such error is to use a small enough sampling interval to reduce aliasing. Alternatively,

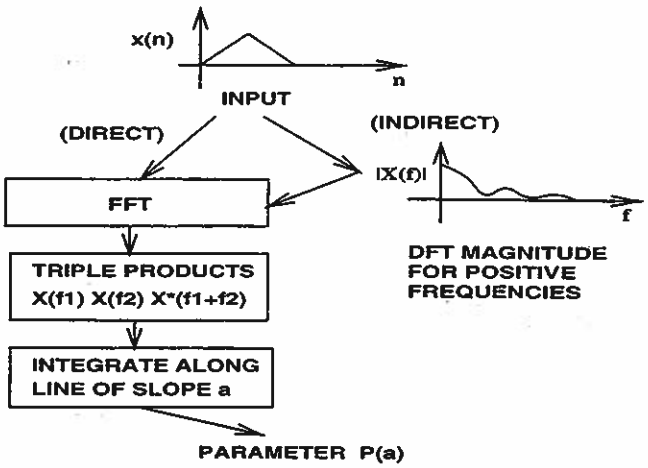

Fig. 2. Direct and indirect methods of computation of invariant parameters for a 1-D sequence. The indirect method discards phase information but yields better scale invariance.

a transformation of the input sequence that will be better bandlimited is sought. One such transformation is the power spectrum. The Fourier transform of the power spectrum of the sequence will be the convolution of the sequence with a reversed version of itself, and will therefore be bandlimited for sequences with finite number of nonzero values. The power spectrum, however, cannot be directly used to compute bispectral invariant parameters because it is an even function and the biphase is theoretically zero. Therefore, the positive frequency half of the power spectrum is used to compute bispectral invariants for the input sequence in an indirect procedure. It has been shown [6] that this procedure yields better scale-invariant parameters. DFT magnitudes can be used in place of the power spectrum [see Fig. 2]. Note that invariance to translation and DC-shift is provided by this transformation itself. This indirect procedure discards all the phase information from the original sequence and, thus, is a tradeoff to achieve better scale invariance. The consequent loss of uniqueness in classifying different sequences that have the same DFT magnitude can be resolved by also computing parameters by the direct procedure. The DFT magnitude itself will suffer from aliasing, and DFT magnitudes of scaled versions of the same input sequence are not exactly scaled versions of each other. However, this error is much smaller than the error introduced in the parameters $P(a)$ because of aliasing in the bifrequency domain. The loss of uniqueness caused by the loss of phase information in the indirect procedure is not a serious drawback for recognition of objects from 2-D images because images will be classified from several 1-D projections and hence, features derived from different classes can be expected to be different.

Another source of error in the computation of the parameters $P(a)$ is the replacement of the integral in (3) by a summation. This error can be reduced by zero padding the input before computing its bispectrum. Although zero padding is not very useful in power spectral analysis because it only provides additional values that depend on the window function that truncates the data, it is useful in recognizing finite-width patterns because the window function is the pattern itself. By zero padding the input, the range of scale variation relative to the total length of the sequences is reduced. Consequently, the variation in the parameters $P(a)$ owing to scaling is also reduced.

By considering the indirect procedure for invariant feature extraction (Fig. 2), it can be shown that $P(a)$ are sensitive to changes in the shape of the input 1-D pattern, which is essential for object recognition. Because any change in the DFT magnitude affects the power spectrum and vice-versa, and the relationship is one-to-one, without loss of generality the power spectrum is used in place of the DFT magnitude in the following. The process of taking an asymmetric half of the power spectrum that includes only the positive frequencies is equivalent to convolving the autocorrelation of the input with the Fourier inverse of a unit-step function in the time domain. The second Fourier transform operation in the indirect procedure yields a complex-valued function with real and imaginary parts equal to the positive and negative halves of the autocorrelation of the input and of its Hilbert transform, respectively. The autocorrelation is the output of a matched filter to detect the particular input. For deterministic signals in white Gaussian noise, the matched filter is the optimal filter for detection. Let $x(t)$ be the input pattern with autocorrelation $R_{xx}(k)$. A change in the input can be modeled as $x(t)+\alpha y(t)$. The autocorrelation becomes $R_{xx}(k)+\alpha^2 R_{yy}(k)+\alpha R_{xy}(k)+\alpha R_{yx}(k)$. This is a significant change if α is large or if the energy of $y(t)$ is comparable to that of $x(t)$. Therefore, the bispectrum (a function of triple products of the autocorrelation and its Hilbert transform) will be sensitive to changes in the shape of the input. If the input pattern is nonnegative, its autocorrelation will also be nonnegative. Parameters $P(a)$ obtained from the phase of integrals of the bispectrum along radial lines in bifrequency space (excluding the origin) are guaranteed to retain this sensitivity provided the changes are all of the same sign. Changes that add to some parts of the input and subtract from others may be lost in the integration step. However, robustness to such changes (often resulting from noise or degradations) may be desirable. A detailed analysis of the discriminating power and optimal selection of the features is under investigation. The feature extraction procedure for 2-D patterns relies on 1-D projections, and will retain sensitivity to changes in 2-D shape because they result in changes to the 1-D patterns.

Features must also be robust to noise. If the noise is uncorrelated with itself and with the input, its auto- and cross-correlations will be impulses at the origin. Because the integration step excludes the origin, the features $P(a)$ are insensitive to such noise. The results presented here are for the special case of white Gaussian noise. The features incorporate additional noise immunity owing to the integration in the bispectral domain as is demonstrated in the simulations. An analytical treatment including other types of noise will be presented in another report.

III. Feature Extraction from 2-D Images

The computation of parameters that satisfy desired invariance properties from one dimensional inputs can be extended

to two dimensional inputs by using the Radon transform [16] or parallel beam projections. Let $g(u, v)$ be an $N \times N$ image. Let $\{x_\theta(n)\}$ be the Radon transform of the image, that is, a set of parallel beam projections of the image at angles θ with respect to the horizontal axis. Assume that the projections are computed at equal increments of angle θ, or $\theta_i = i\pi/N_\theta$ for $i = 0$ to $N_\theta - 1$. The 2-D image is thus reduced to a set of 1-D projections.

It can be readily deduced that

1) a shift in the 2-D image results in a shift in every projection except for the one parallel to the direction of the shift;
2) a scale change of the 2-D image in the direction of projection results in multiplication of the 1-D projection by a constant;
3) a scale change of the 2-D image perpendicular to the direction of projection results in a scale change of the 1-D projection; and
4) a rotation of the 2-D image results in a cyclic shift in the set of projection functions.

Features $P(a)$ that are invariant to shift, scaling, or amplification of the 1-D functions, $x(n)$, will therefore provide a set $\{P(a)(\theta), a = 1, N_a\}$ of 1-D functions of θ, which are cyclically shifted when the image is rotated. Note that the slope of integration in Fig. 1 is a, whereas in the discussion that follows, a denotes an integer and the slope is a/N_a where N_a is the number of lines of integration (see Fig. 1). It is assumed here that the images have a *zero background* such that there are no effects at the edges upon rotation or translation.

The Fourier slice theorem [16] states that the Fourier transform of the projection of an image on to a line is the 2-D Fourier transform of the image evaluated along a radial line. Slices of the 2-D Fourier transform can be used to extract features instead of the Radon transform. This is computationally more efficient since raster rotations are avoided. The DC component is ignored and only the positive frequency half of the DFT magnitude is used to extract the bispectral features.

The procedure for feature extraction is as follows. An $N \times N$ image, $g(u, v)$, is Fourier transformed using a 2-D fast Fourier transform (FFT) routine and the Fourier transform is mapped onto a polar grid using bilinear interpolation to yield $G(r, \theta)$, where $r = 0$ to $N/2$ and $\theta = i\pi/N_\theta$ for $i = 0$ to $2N_\theta - 1$. The Fourier transform magnitude along a radial line is zero padded to form the sequence

$$y_\theta(n) = \begin{cases} |G(n+1, \theta)| & n = 0, 1, \cdots, \dfrac{N}{2} - 1 \\ 0 & \dfrac{N}{2} \le n < N. \end{cases} \quad (4)$$

Bispectral invariants are computed for $y_\theta(n)$ to yield $P_\theta(a)$ for $a = 1, 2, \cdots, N_a$.The procedure is repeated for different angles, θ. To obtain rotation invariance the parameters $P_\theta(a)$ are then considered as 1-D sequences for each value of parameter a. An N_θ-point DFT is computed for each such sequence and the magnitude of the Fourier transform for positive frequencies, including the DC value, is zero padded to N_θ points. Bispectral invariants are then computed for the DFT magnitude sequences, yielding the set of features, $P(a)(b)$, for $a = 1, \cdots, N_a$ and $b = 1, \cdots, N_b$ where N_b is the number of lines of integration in the second step, after regrouping invariants for each value of a, $P_\theta(a)$. These features are also rotation invariant because of the cyclic shift invariance of the DFT.

In practice, the truncation of the discrete-time Fourier transform and aliasing involved in the computation of the DFT result in some variation of the features with scale. It has been shown [14] that essential information is retained in the DFT despite such effects, provided the size of the object and details of interest lie between about 4 and $N/4$ points.

IV. Results

The algorithm was tested using i) synthetic binary images of regular geometric shapes, ii) preprocessed images of 2-D views of simple machine parts, and iii) preprocessed images of 2-D views of some tools. The objective was to test the performance on ideal binary images and then in the presence of errors introduced by nonidealities in lighting and preprocessing. For the first test, the images are ideal binary shapes and the only error is owing to rotations on a sampled surface. For the second test, some error is residual in the images after preprocessing because of lighting and perspective variations in the prototype images acquired using a video camera. For the third test, the images after preprocessing have broken contours and holes within the objects, because the objects had translucent surfaces. These tests are described below, followed by tests to quantitatively assess the performance of the algorithm in the presence of Gaussian noise. In all cases, good separation of the images in feature space was obtained.

A. Synthetic Images

Bilevel, 256 × 256 pixel, 8-b images of four regular geometric shapes (box, circle, the letter H, and a plus sign) were chosen. The inputs were zero padded to 512 × 512 points before feature extraction. Cross sections of the DFT magnitudes of these images are shown in Fig. 3, and reveal that they retain information that can be used to classify these shapes.

Thirty-five rotated (uniform random in $[0, \pi]$) and translated (uniform random in $[-16, 16]$) versions of each prototype were generated. Two-hundred fifty-six projections were considered, and each image was reduced to a set of 8 × 8 features. A scatter plot of features derived from the randomly rotated and translated images (see Fig. 4) shows that the images are well separated in feature space, and thus 100% classification accuracy is possible.

B. 2-D Views of Simple Machine Parts

Grey-level (8-b), 256 × 256 pixel images of a hexagonal nut, a square nut, and a washer were acquired using a video camera and digitizer. A thresholding operation was performed to remove any texture in the background. The prototype images

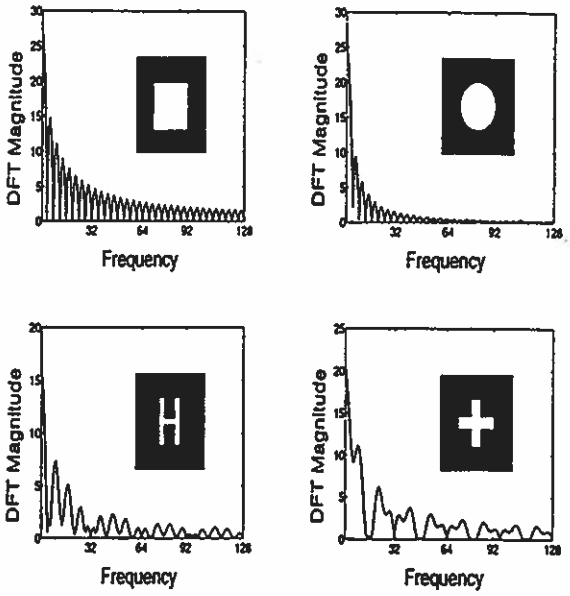

Fig. 3. Cross sections of DFT magnitudes of synthetic images of a box, a circle, the letter H, and a plus sign. The shapes are shown alongside the corresponding spectra. The frequency is in cycles per 256 pixels.

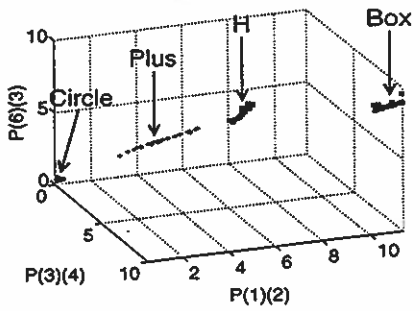

Fig. 4. Scatter plot of features for the synthetic images.

after preprocessing are shown in Fig. 5. The inputs were zero padded to 512 × 512 pixels before feature extraction.

Each object was imaged at three different scales in the ratio 0.7 : 0.9 : 1.1. Additional images were generated by random translation and rotation to yield 144 images from each class. Two hundred fifty six projections were considered and an 8 × 8 set of features was extracted. A scatter plot using three features (Fig. 6) shows the well isolated clusters in feature space. Any standard classifier [17] can be used to obtain 100% accuracy given these features.

C. Images of Tools

The algorithm was also tested on grey-level images of four machine tools (screw driver, pair of pliers, hammer, and saw shown in Fig. 7). Each object was imaged twice at three different scales and 48 randomly rotated and translated versions were created. The processing of these images was identical to that for the machine parts. A scatter plot using three features (Fig. 8) reveals separate clusters for each object. The dispersion of the clusters is owing to degraded inputs (especially the saw) and some impulse noise in the background and can be reduced by preprocessing.

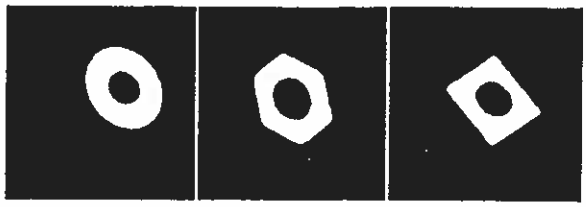

Fig. 5. Prototype images of machine parts.

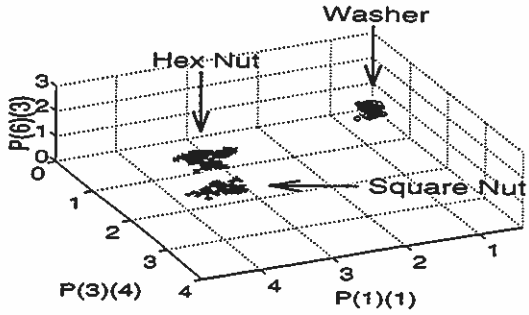

Fig. 6. Scatter plot of features for randomly translated and rotated images of the machine parts, showing well-isolated clusters.

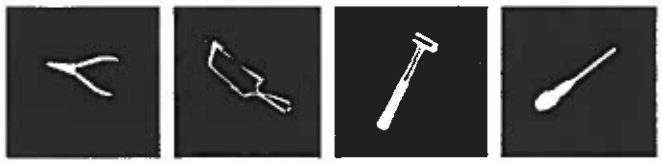

Fig. 7. Prototype images of machine tools.

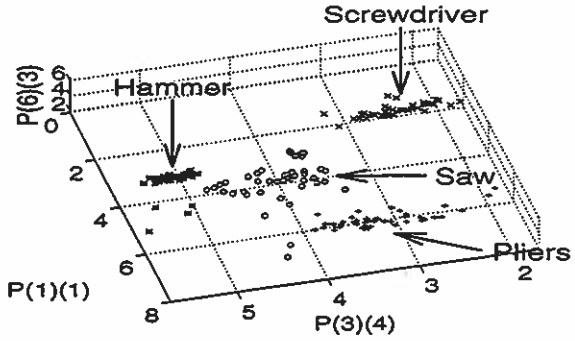

Fig. 8. Scatter plot of features for randomly rotated and translated images of the machine tools.

V. Comparison with Other Algorithms

To quantify the discrimination performance of bispectral features, a comparison was made with the well established method of moment invariants [18]–[21]. Comparisons were made using i) synthetic binary images in Gaussian noise; and ii) preprocessed images with background impulses in Gaussian noise.

The objective here is to compare the immunity of the classification methods to Gaussian noise, and also to test the effect of impulses in the background that may remain in the image after preprocessing steps such as filtering and thresholding.

Fig. 9. Synthetic binary image prototypes used for the classification accuracy tests.

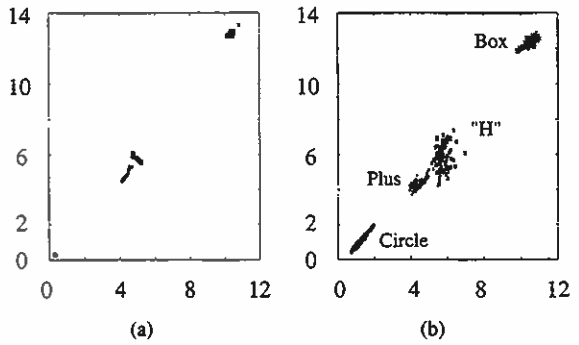

Fig. 10. Cluster plots for two bispectral features. (a) SNR = ∞. (b) SNR = 10 dB.

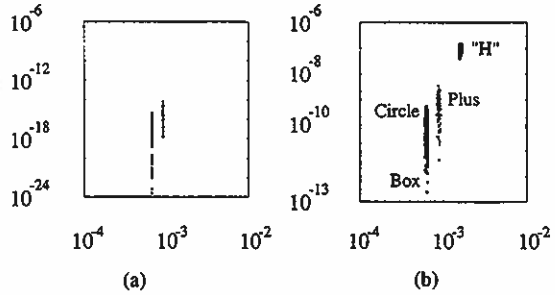

Fig. 11. Cluster plots for two moment invariants. (a) SNR = ∞. (b) SNR = 10 dB.

The moment invariant method is especially vulnerable to such impulses because they can shift the centroid of the object considerably.

A. Synthetic Binary Images in Gaussian Noise

Binary image sets were generated from the four prototypes shown in Fig. 9. Two sets of images were generated, each with 100 randomly translated and rotated instances of each of the four prototypes. One set of images was used to train the classifier and the other was used for testing. Both algorithms were tested with clean images as well as those with Gaussian noise, with signal-to-noise ratios (SNR's) ranging from 20 to −10 dB.

The effect of noise on the feature clusters was observed to be both an increased variance as well as a shift in the cluster mean. Scaling also introduces shifts in the cluster means. As a result, the feature clusters are not Gaussian distributed (see Fig. 10). Therefore, a nearest neighbor classifier [22] which works well when the noise is low and the clusters are well separated, was used.

If the training set is large, the clusters will be dense. Then the performance of a one nearest neighbor (1−NN) classifier will not be different from that of a $k-$NN classifier for $k >$ 1. This was observed to be the case. The results for a one nearest-neighbor classifier are given below.

The "optimal" training sets for nearest neighbor classifiers are those whose feature clusters are maximally spread in feature space. This ensures that feature vectors distant from their class cluster center, but not overlapping into other clusters, will be correctly classified. The optimal SNR's for training data were determined by training with increasingly noisy data until that value beyond which large numbers of classification errors started to occur, indicating that the clusters in the training data set were overlapping.

TABLE I
CLASSIFICATION ERROR PERCENTAGES USING ALL BISPECTRAL FEATURES WHEN THE TRAINING SET WAS NOISE FREE

Training set - ∞ SNR				
SNR	'H'	Box	Circle	'+'
∞	0	0	0	0
20	0	0	0	0
10	0	0	0	1
0	33	17	17	28
-10	44	33	69	27

Cluster plots[1] for selected bispectral and moment invariant features are shown in Figs. 10 and 11, respectively, for noise-free [Figs. 10(a) and 11(a)] and noisy [10 dB SNR, Figs. 10(b) and 11(b)] images. The features from both methods are quite insensitive to noise, as the clusters in Figs. 10(b) and 11(b) are reasonably compact. Classification error percentages as a function of SNR for different training schemes for both the higher order spectra- and moment-based schemes are given in Tables I–IV. The two methods have similar success for all SNR's.

B. Images with Background Impulses and Gaussian Noise

Grey-level images are often thresholded at the preprocessing stage, leaving impulse noise in the background. Some of this noise is likely to remain even after filtering (such as median

[1]Feature selection criterion is to find the features for which $(mean_i - mean_j)/\sigma_i^2\sigma_j^2$ is maximum, where $mean_{[i|j]}$ and $\sigma_{[i|j]}$ are mean and standard deviation for class i and j, respectively.

Fig. 12. Example prototypes for machine parts used for the classification accuracy tests.

TABLE II
CLASSIFICATION ERROR PERCENTAGES USING ALL BISPECTRAL FEATURES WHEN THE TRAINING SET HAD 0-dB SNR. THIS SNR PRODUCED THE LEAST ERRORS IN THE BISPECTRAL FEATURES EXTRACTION METHOD FOR THE GIVEN DATA

Training set - 0dB SNR				
SNR	'H'	Box	Circle	'+'
∞	0	0	0	0
20	1	0	0	0
10	1	0	0	0
0	2	0	0	1
-10	30	32	8	21

TABLE III
CLASSIFICATION ERROR PERCENTAGES USING SEVEN INVARIANT MOMENTS AS FEATURES WHEN THE TRAINING SET WAS NOISE FREE

Training set - ∞ SNR				
SNR	'H'	Box	Circle	'+'
∞	0	0	0	0
20	0	0	0	0
10	0	0	2	0
0	0	8	27	2
-10	16	34	50	26

TABLE IV
CLASSIFICATION ERROR PERCENTAGES USING SEVEN INVARIANT MOMENTS AS FEATURES WHEN THE TRAINING SET HAD −10 DB SNR. THIS SNR PRODUCED THE LEAST ERRORS IN THE MOMENT-INVARIANT METHOD FOR THE GIVEN DATA

Training set - 10dB SNR				
SNR	'H'	Box	Circle	'+'
∞	0	0	0	0
20	0	0	0	0
10	0	0	0	0
0	0	15	22	1
-10	15	40	50	24

filtering). These impulses can be located with reference to the object of interest in the following three ways.

1) They maintain a fixed spatial relationship to the object.
2) They maintain a fixed spatial relationship to the frame.
3) They occur randomly with no fixed spatial relationship to either the object or the frame.

Cases 2) and 3) will affect adversely both techniques for rotation-invariant feature extraction. The moment-invariant method will be affected especially by the random shifts in the centroid owing to the impulse noise. However, case 1) is not really impulse noise because the impulses rotate with the object. The feature extraction method will work poorly if the features are more sensitive to the impulses than to the shape of the object. The moment-invariant method is likely to be more sensitive to pixels far away from the centroid and, hence, to be affected by the impulses in case 1). The higher order spectral feature extraction method proposed here uses the 2-D FFT magnitude (asymmetric halves of radial slices of it). Impulses, as in case 1), will introduce a modulation distortion in the 2-D FFT, which does not change with rotation. The higher order spectral features will therefore still be sensitive to the shape of the object, but they may be more vulnerable to noise in the image, such as additive Gaussian noise. Tests were performed to compare the performance of the two methods for images with impulses in the background of the type in case 1), with Gaussian noise added.

Grey-level images of machine parts were acquired and thresholded to remove a common texture in the background. The training data set consisted of ten randomly rotated and translated versions of 12 prototypes of each of three classes of objects. The prototype images were acquired at three different scales and arbitrary orientations. The testing data set consisted of ten randomly rotated and translated versions of 15 prototypes from each of the three classes. An example prototype from each class is shown in Fig. 12. These images have been thresholded to remove a background texture common to each image, leaving some impulse noise in the images. Gaussian noise with SNR from 20 dB to −10 dB was added to the images.

Classification accuracy results were obtained for the machine parts images for both bispectral features and moment invariants. A nearest-neighbor classifier [17], [22] was used, and results were obtained using both noiseless and noisy training sets as in the previous section. For the purpose of these experiments, the images as shown in Fig. 12 will be termed to be noiseless, and noisy images will be those that have had Gaussian noise added in testing. The results are given in Tables V–VIII.

Cluster plots show that the bispectral features (Fig. 13) form compact isolated clusters in feature space, whereas there is

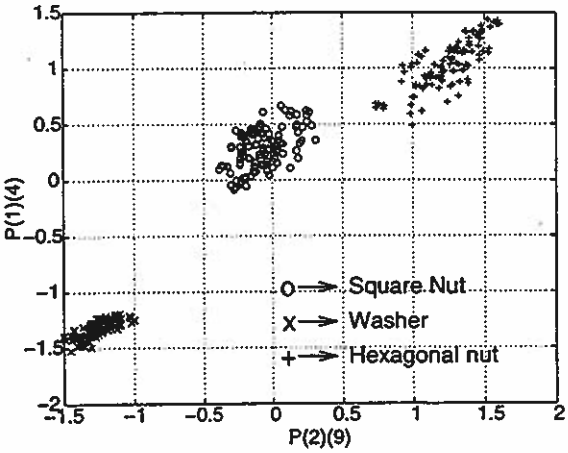

Fig. 13. Cluster plot for two bispectral features—machine parts input.

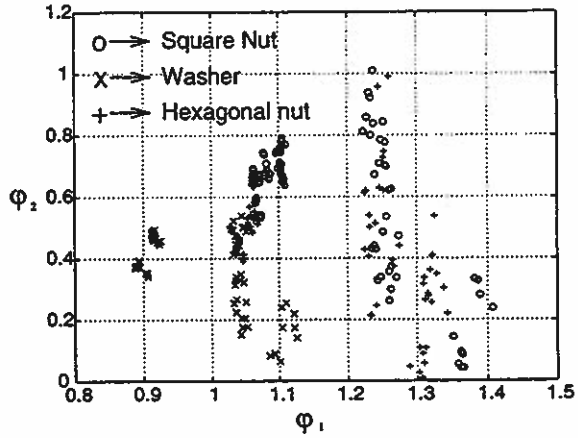

Fig. 14. Cluster plot for two moment invariants—machine parts input.

TABLE V
CLASSIFICATION ERROR PERCENTAGES USING ALL BISPECTRAL FEATURES AND A "CLEAN" TRAINING SET FOR THE MACHINE PARTS

Training set - ∞SNR			
SNR	Washer	Hexagonal Nut	Square Nut
∞	3.33	0.00	0.00
20	2.00	0.00	0.00
10	6.67	6.00	10.67
0	38.00	50.00	61.33
-10	24.00	44.67	72.00

considerable overlap in the clusters for the moments (Fig. 14). For the images tested here, bispectral features (Tables V, VI, Fig. 13) yield better classification results than the moment-invariance features (Tables VII, VIII, and Fig. 14).

C. Bispectral Features versus UNL Fourier Features

A further comparison of the performance of bispectral features was made with that of UNL Fourier features [15], [23], which is also a Fourier-based global feature extraction method. To obtain UNL Fourier features, the image contour is extracted and then mapped from Cartesian space to the polar coordinate (r, θ) system. The centroid of the contour representation is used as the origin of the polar coordinates. This makes the features translation invariant. Radial distances are scaled by the largest value, making the features scale invariant. Angles (θ) are measured from an arbitrary line passing through the centroid. Rotation of the image results in a cyclic shift along the θ-axis. The 2-D Fourier transform of the polar space is calculated, and the features are selected from the magnitudes of the Fourier coefficients normalized by the average value of the image. Because the 2-D Fourier magnitude is translation invariant, the UNL features are also rotation invariant. They are not tolerant of occlusion, but work for open and nested curves.

The test data in this case were the machine parts described above (Fig. 12). Clusters for two features from each set of bispectral and UNL Fourier features for test data sets containing 120 randomly rotated and translated instances of the prototypes are shown in Figs. 15 and 16, respectively, and illustrate the problems associated with using contour-based methods on noisy images.

TABLE VI
CLASSIFICATION ERROR PERCENTAGES USING ALL BISPECTRAL FEATURES AND A NOISY TRAINING SET—MACHINE PARTS

Training set - 10dB SNR			
SNR	Washer	Hexagonal Nut	Square Nut
∞	2.00	1.33	0.00
20	5.33	2.00	0.00
10	8.00	0.00	0.00
0	46.00	44.00	49.33
-10	38.67	47.33	60.00

TABLE VII
CLASSIFICATION ERROR PERCENTAGES USING MOMENT INVARIANTS AND A "CLEAN" TRAINING SET—MACHINE PARTS

Training set - ∞ SNR			
SNR	Washer	Hexagonal Nut	Square Nut
∞	37.33	47.33	20.00
20	41.33	56.66	16.67
10	44.00	52.00	16.00
0	53.33	67.33	28.67
-10	51.33	82.67	36.00

TABLE VIII
CLASSIFICATION ERROR PERCENTAGES USING MOMENT INVARIANTS AND NOISY TRAINING SET—MACHINE PARTS

Training set - 10dB SNR			
SNR	Washer	Hexagonal Nut	Square Nut
∞	72.67	18.67	28.00
20	69.33	23.33	20.00
10	61.33	30.67	19.33
0	70.66	31.33	32.67
-10	72.673	42.00	69.33

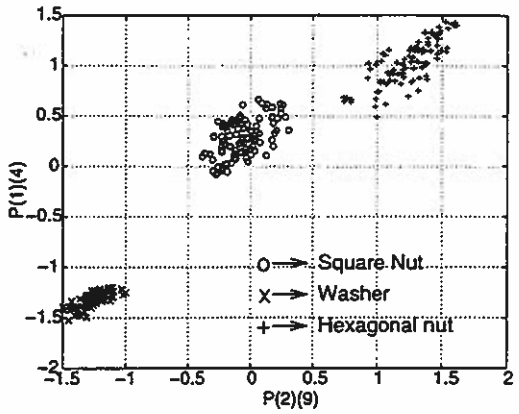

Fig. 15. Cluster plot for two bispectral features—machine parts.

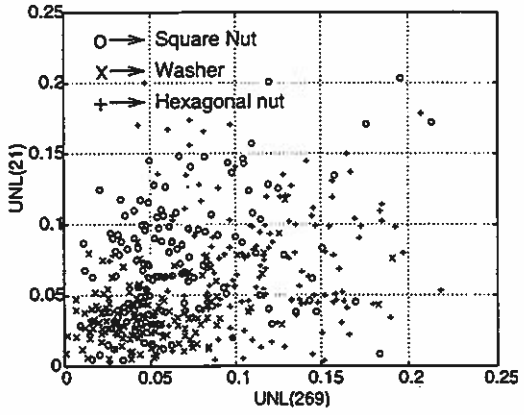

Fig. 16. Cluster plot for two UNL Fourier features—machine parts.

VI. Conclusions

A new algorithm for object recognition from 2-D images based on higher order spectral features that are invariant to translation, rotation, and scaling is shown to accurately distinguish between similar 2-D shapes. The feature extraction procedure incorporates most of the nonlinearity required in mapping from input to class, permitting the use of linear classifiers. The higher order spectral features are shown to be as immune to Gaussian noise as features based on moment invariants, but are superior in their immunity to background impulses that cause problems for methods based on accurate calculation of the centroid.

References

[1] K. Hasselman, W. Munk, and G. MacDonald, "Bispectra of ocean waves," in *Time Series Analysis*, M. Rosenblatt, Ed. New York: Wiley, 1963, pp. 48–69.
[2] D. Brillinger, "A introduction to polyspectra," *Ann. Math. Stat.*, vol. 36, pp. 1351–1347, 1965.
[3] D. Brillinger and M. Rosenblatt, "Computation and interpretation of *k*th order spectra," in *Spectral Anal. Time Series*, B. Harris, Ed. New York: Wiley, 1967, pp. 189–232.
[4] C. Nikias and M. Raghuveer, "Bispectrum estimation: A digital signal processing framework," in *Proc. IEEE*, vol. 75, pp. 869–889, 1987.
[5] B. Sadler, "Shift and rotation invariant object recognition using the bispectrum," in *Proc. Workshop on Higher Order Spectral Analysis*, Vail, CO, pp. 106–111, 1989.
[6] V. Chandran and S. Elgar, "Pattern recognition using invariants defined from higher order spectra-one-dimensional inputs," *IEEE Trans. Signal Processing*, vol. 40, pp. 205–212, Jan. 1991.
[7] M. Tsatanis and G. Giannakis, "Translation, rotation and scaling invariant object and texture classification using polyspectra," in *Proc. Society of Photo-Optical Instrumentation Engineers, SPIE'90*, San Diego, CA, vol. 1348, pp. 103–115.
[8] V. Chandran and S. Elgar, "Position, rotation, and scale invariant recognitions of images using higher order spectra," in *Proc. ICASSP'92*, San Francisco, CA.
[9] E. Sayrol, C. Nikias, and T. Gasuli, "Image restoration using HOS and the Radon transform," in *IEEE Signal Processing Workshop on Higher Order Statistics*, South Lake Tahoe, NV, June 1993, pp. 76–80.
[10] S. Carrato and G. Ramponi, "Edge detection using generalized higher-order statistics," in *IEEE Signal Processing Workshop on Higher-Order Stat.*, South Lake Tahoe, June 1993, pp. 66–70.
[11] A. Oppenheim and J. Kim, "The importance of phase in signals," in *Proc. IEEE*, vol. 69, pp. 529–541, May 1981.
[12] D. Casaasent and D. Psaltis, "Position, rotation and scale invariant optical correlation," *Appl. Opt.*, vol. 15, pp. 1795–1799, 1976.
[13] L. Capodiferro, R. Gusani, G. Jacoviti, and M. Vascotto, "A correlation based technique for shift, scale and rotation independent object identification," in *Proc. ICASSP*, Dallas, TX, 1987, pp. 221–224.
[14] J. Altmann and H. Reibock, "A fast correlation method for scale and translation invariant pattern recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, pp. 46–57, Jan. 1984.
[15] T. Rauber, "Two-dimensional shape description," Tech. Rep. GR UNINOVA-RT-10-94, Universade Nova de Lisboa, Lisbon, Portugal, 1994.
[16] A. Rosenfeld and A. Kak, *Digital Picture Processing*. New York: Academic, 1982, vol. 1.
[17] R. Duda and P. Hart, *Pattern Classification and Scene Analysis*. New York: Wiley, 1973.
[18] M. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inform Theory*, vol. IT-8, pp. 179–187, Feb. 1962.
[19] Y. Abu-Mostafa and D. Psaltis, "Recognitive aspects of moment invariants," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-6, no. 6, pp. 698–706, 1984.
[20] S. Dudani, K. Breeding, and R. McGhee, "Aircraft identification by moment invariants," *IEEE Trans. Comput.*, vol. C.26, pp. 39–45, Jan. 1977.
[21] R. Gonzalez and R. Woods, *Digital Image Processing*, 3rd ed. Reading, MA: Addison-Wesley, 1992.
[22] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inform. Theory*, vol. 13, pp. 21–27, 1967.
[23] T. Rauber and A. Steiger-Garcao, "Shape description by UNL Fourier features—An application to handwritten character recognition," in *Proc. 11th IAPR Int. Conf. Pattern Recognition*, The Hague, The Netherlands, 1992, pp. 466–469.

Vinod Chandran (S'85–M'90) received the B.Tech degree in electrical engineering from the Indian Institute of Techology, Madras, India, in 1982, the M.S. in electrical engineering from Texas Technical University, Lubbock, in 1985, and the M.S. degree in computer science, Ph.D. in electrical and computer engineering, both from Washington State University, Pullman, in 1990 and 1991, respectively.

From August 1990 to May 1993, he was a postdoctoral teaching associate at Washington State University. He is currently a Senior Lecturer in the School of Electrical and Electronic Systems Engineering, Queensland University of Technology (QUT), Brisbane, Australia, and a member of the Signal Processing Research Centre at QUT. His current research interests include pattern recognition, higher order spectra, and image compression.

Dr. Chandran is a member of the Optical Society of America, Tau Beta Pi and Phi Kappa Phi honor societies, and an associate member of Sigma Xi.

Brett Carswell received a double Bachelor's degree in electrical engineering and information technology from Queensland University of Technology (QUT), Brisbane, Qld., Australia, in 1993, and was a full-time Ph.D. student at the Signal Processing Research Centre, QUT, until 1995.

He is currently an Engineer with Tritronix Corporation.

Boualem Boashash (M'82–SM'89) received the Dipl. d'ing.-Physique–Electronique from the ICPI University of Lyon, France, in 1978. He received the M.S. degree from the Institut National Polytechnique de Grenoble, France, in 1979, and the Doct.-Ing. from the same university in May 1982.

In 1979, he joined Elf-Aquitaine Geophysical Research Centre, Pau, France, where he was a Research Engineer in the signal processing group. In May 1982, he joined the Institut National des Sciences Appliquées de Lyon, Lyon, France, where he was a Maitre-Assistant associé. In January 1984, he joined the Electrical Engineering Department, University of Queensland, Australia, as a Lecturer, Senior Lecturer (1986), and Reader (1989). In 1990, he joined Bond University, Graduate School of Science and Technology, as Adjunct Professor of signal processing. In 1991, he joined Queensland University of Technology as the foundation Professor of signal processing and Director of the Signal Processing Research Centre. His research interests are time–frequency signal analysis, spectral estimation, signal detection and classification, and higher order spectra. He is also interested in wider issues such as the effect of engineering developments on society. He is the editor of *Time-Frequency Signal Analysis: Methods and Applications* (Longman-Cheshire, 1991). He has contributed several chapters to advanced signal processing books including "Time-Frequency Signal Analysis," in *Advances in Spectrum Estimation*, S. Haykin, Ed. (Englewood Cliffs, NJ: Prentice-Hall, 1991). He has also written more than 200 technical publications, and has supervised 20 Ph.D. students and five Masters students.

Dr. Boashash was Technical Chair of ICASSP'94, the premium conference in signal processing. He is currently Chair of the International Symposium on Signal Processing and its Applications, organized regularly in Australia since 1987.

Prof. Boashash is fellow of IE Australia and a fellow of IREE.

Steve Elgar (M'86) received B.S. degrees in mathematics and civil engineering from the University of Idaho, Moscow, in 1980, and the Master's and Ph.D. degrees in oceanography from the Scripps Institution of Oceanography, University of California, San Diego, La Jolla, in 1981 and 1985, respectively.

In 1986, he joined the School of Electrical Engineering and Computer Science at Washington State University, Pullman, where he is a Professor. His current research interests involve the study of nonlinear waves, and in particular ocean surface gavity waves and signal processing techniques necessary for interpreting and modeling ocean field measurements. In collaboration with B. Werner, Dr. Elgar is also studying the impulse response of a natural beach.

He was an associate editor for IEEE TRANSACTIONS ON SIGNAL PROCESSING. He is ocean sciences editor for *Eos, Transactions of the AGU*. He is a member of the American Association for the Advancement of Science, the American Meteorological Society, and the American Geophysical Union.