Orthogonal Fourier–Mellin moments for invariant pattern recognition

Yunlong Sheng and Lixin Shen

Centre d'Optique Photonique et Laser, Département de Physique, Université Laval, Ste-Foy, Québec G1K 7P4, Canada

Received July 7, 1993; revised manuscript received November 22, 1993; accepted December 14, 1993

We propose orthogonal Fourier–Mellin moments, which are more suitable than Zernike moments, for scale- and rotation-invariant pattern recognition. The new orthogonal radial polynomials have more zeros than do the Zernike radial polynomials in the region of small radial distance. The orthogonal Fourier–Mellin moments may be thought of as generalized Zernike moments and orthogonalized complex moments. For small images, the description by the orthogonal Fourier–Mellin moments is better than that by the Zernike moments in terms of image-reconstruction errors and signal-to-noise ratio. Experimental results are shown.

1. INTRODUCTION

The geometrical moments of the image are integrals of the image function over space. The low-order moments are related to global properties of the image and are commonly used to determine the position, orientation, and scale of the image. Optical generation of the geometrical moments has been demonstrated. The moment invariants[1] are linear combinations of the geometrical moments, used as image descriptors for scale- and rotation-invariant pattern recognition. In general, the moment invariants are not good image features. They are sensitive to noise and suffer from information suppression, loss, and redundancy.[2]

From the uniqueness theorem of moments, an image is uniquely determined by its geometrical moments of all orders. Low-order moments contain less information about image detail. High-order moments are vulnerable to noise. Interesting questions are how a finite number of moments can work well to describe an image for a particular application and what is gained by inclusion of higher-order moments. The answers to these questions can be found with the use of the orthogonal moments.[3] Teh and Chin recently evaluated various types of image moments, including geometrical moments, Legendre moments, Zernike moments, Pseudo-Zernike moments, Fourier–Mellin moments, and complex moments, in terms of noise sensitivity, information redundancy, and capability of image description. They found that Zernike moments have the best overall performance.[4]

Zernike moments (ZM's) play a fundamental role in the theories of optical aberration and diffraction. Teague proposed ZM's for generating the moment invariants.[5] ZM's are orthogonal and rotation invariant. But, when they are used for scale-invariant pattern recognition, ZM's have difficulty in describing images of small size, as we show in this paper.

Bhatia and Wolf pointed out that there exist an infinite number of complete sets of polynomials that are rotation invariant in form and are orthogonal for the interior of the unit circle.[6] In this paper we introduce orthogonal Fourier–Mellin moments (OFMM's), based on a new set of radial polynomials. The new moments belong to generalized ZM's and to orthogonalized complex moments. We show that OFMM's have better performance than ZM's in terms of image description and noise sensitivity. In Section 2 we introduce OFMM's and show their relations to ZM's and to complex moments. In Section 3 we compare the performance of OFMM's and ZM's. In Section 4 we present an experiment of rotation, scale, intensity, and shift-invariant pattern recognition, using OFMM's.

2. ORTHOGONAL FOURIER–MELLIN MOMENTS

A. Definition

The circular Fourier and radial Mellin transform defined in a polar coordinate system (r, θ) is useful for scale- and rotation-invariant pattern recognition:

$$M_{s,m} = \int_0^{2\pi} \int_0^{\infty} r^s f(r,\theta) \exp(-jm\theta) r \mathrm{d}r \mathrm{d}\theta , \qquad (1)$$

where $f(r,\theta)$ is the image and $m = 0, \pm 1, \pm 2, \ldots$ is the circular harmonic order. By definition, the Mellin transform order s is complex valued. With integer $s \geq 0$, Eq. (1) defines Fourier–Mellin moments (FMM's), also called rotational moments.[7] We now define the OFMM's as

$$\Phi_{nm} = \frac{1}{2\pi a_n} \int_0^{2\pi} \int_0^{1} f(r,\theta) Q_n(r) \exp(-jm\theta) r \mathrm{d}r \mathrm{d}\theta , \qquad (2)$$

where a_n is a normalization constant and $Q_n(r)$ is a polynomial in r of degree n. The set of $Q_n(r)$ is orthogonal over the range $0 \leq r \leq 1$:

$$\int_0^1 Q_n(r) Q_k(r) r \mathrm{d}r = a_n \delta_{nk} , \qquad (3)$$

where δ_{nk} is the Kronecker symbol and $r = 1$ is the maximum size of the objects that can be encountered in a particular application. Hence the basis functions $Q_n(r)\exp(-jm\theta)$ of the OFMM's are orthogonal over the interior of the unit circle.

0740-3232/94/061748-10$06.00

In OFMM's we use the radial polynomials of low degrees of n, which are independent of the circular harmonic order m. We obtain the polynomials $Q_n(r)$ by applying the Gram–Schmidt orthogonalization process to the sequence of natural powers of r over the range $0 \le r \le 1$:

$$1, r, r^2, \ldots r^n . \tag{4}$$

In Appendix A we show that the polynomials $Q_n(r)$ are equal to

$$Q_n(r) = \sum_{s=0}^{n} \alpha_{ns} r^s , \tag{5}$$

with

$$\alpha_{ns} = (-1)^{n+s} \frac{(n+s+1)!}{(n-s)!\, s!(s+1)!} , \tag{6}$$

and the normalization constant in Eq. (3) is $a_n = 1/[2(n+1)]$. The explicit expressions of $Q_n(r)$ are

$$Q_0(r) = 1 ,$$

$$Q_1(r) = -(2 - 3r) ,$$

$$Q_2(r) = 3 - 12r + 10r^2 ,$$

$$Q_3(r) = -(4 - 30r + 60r^2 - 35r^3) ,$$

$$Q_4(r) = 5 - 60r + 210r^2 - 280r^3 + 126r^4 ,$$

$$Q_5(r) = -(6 - 105r + 560r^2 - 260r^3 + 1260r^4 - 462r^5) ,$$

$$\ldots .$$

The OFMM's are integrable when the degree of $Q_n(r)$ $n \ge 0$. On substituting from Eqs. (5) and (1) into Eq. (2), we express the OFMM's as linear combinations of FMM's:

$$\Phi_{nm} = \frac{n+1}{\pi} \sum_{s=0}^{n} \alpha_{ns} M_{sm} . \tag{7}$$

OFMM's can be thought of as generalized ZM's. Bhatia and Wolf have shown[6] that a polynomial that is invariant in form for any rotation of axes about the origin must be of the form

$$V(r \cos\theta, r \sin\theta) = R_n(r) \exp(jm\theta) , \tag{8}$$

where $R_n(r)$ is a radial polynomial in r of degree n. If $V(x,y)$ must be polynomials in x and y, it follows that there is one and only one orthogonal set, the set of Zernike radial polynomials $R_n^{|m|}(r)$, with $n \ge |m|, |m|+2, |m|+4, \ldots$. Bhatia and Wolf also derived another orthogonal set, the pseudo-Zernike polynomials $P_n^{|m|}(r)$, which are the radial polynomials in r containing no powers of r lower than $|m|$. The radial polynomials of the OFMM's, the ZM's, and the pseudo-Zernike moments (PZM's) all belong to the Jacobi polynomials. In Appendix B we show the relations among the three radial polynomials.

By definition, the degree n of $Q_n(r)$ is independent of the circular harmonic order m. As a result, the basis function $Q_n(r)\exp(-jm\theta)$ is no longer polynomials in x and y but may still be expressed as complex polynomials in $(x + jy)$ and $(x - jy)$, because

$$Q_n(r)\exp(-jm\theta) = \sum_{s=0}^{n} \alpha_{n,s} r^s \exp(-jm\theta)$$
$$= \sum_{s=0}^{n} \alpha_{n,s} (x+jy)^p (x-jy)^q , \tag{9}$$

where

$$p = \frac{s-m}{2} , \qquad q = \frac{s+m}{2} . \tag{10}$$

The OFMM's can then be expressed as linear combinations of complex moments (CM's):

$$\Phi_{nm} = \frac{n+1}{\pi} \sum_{s=0}^{n} \alpha_{ns} C_{(s-m/2),(s+m)/2} , \tag{11}$$

where the CM's are defined by Abu-Mostafa and Psaltis as[2]

$$C_{pq} = \iint_{-\infty}^{\infty} f(x,y) (x+jy)^p (x-jy)^q \, dx\, dy . \tag{12}$$

By definition, the orders p and q are nonnegative integers. In Eq. (11) the CM's are modified so that the orders $p = (s-m)/2$ and $q = (s+m)/2$ are real valued and can be negative. The OFMM's are therefore the orthogonal CM's. When both $p + q = s$ and $q - p = m$ are integers and $s = p + q \ge 0$, the modified CM's are convergent and integrable. The OFMM's therefore may be computed in the Cartesian coordinate system according to Eqs. (11) and (12).

B. Properties

OFMM's have a single orthogonal set of the radial polynomials $Q_n(r)$ for all the circular harmonic orders m. ZM's and PZM's have a set of orthogonal radial polynomials for each circular order m. Only the polynomials $R_n^{|m|}(r)$ and $P_n^{|m|}(r)$ with the same m are orthogonal. The degree n of $Q_n(r)$ in the OFMM's can be much lower than that of the Zernike and the pseudo-Zernike polynomials. OFMM's can have $n = 0, 1, 2, \ldots$, which are not permissible in ZM's and PZM's if $n < |m|$.

The number of zeros of the radial polynomials corresponds to the capability of the polynomials to describe high-spatial-frequency components of the image. Since the $Q_n(r)$ contains natural powers $1, r, r^2, \ldots, r^n$, the equation $Q_n(r) = 0$ has n real and distinct roots in the interior of the interval $0 \le r \le 1$.[8] It is easy to see that the equation $R_n^{|m|}(r) = 0$ has $(n-m)/2$ duplicated roots in the interior of the interval $0 \le r \le 1$, apart from $R_n^{|m|}(0) = 0$. For a given degree n and circular harmonic order m, the $Q_n(r)$ has n zeros and the $R_n^{|m|}(r)$ has $(n - |m|)/2$ zeros. For one to have the same number n_0 of zeros, the degree of the $R_n^{|m|}(r)$ has to be as high as $2n_0 + |m|$, much higher than the degree n_0 of the OFMM's. Note that the description of an image usually requires the circular harmonic orders $|m|$ up to higher than 10.[9]

Comparing the $Q_n(r)$ and the $R_n^{|m|}(r)$ with the same number of zeros, we find that the zeros of the $Q_n(r)$ are nearly uniformly distributed over the interval $0 \le r \le 1$, whereas the zeros of the $R_n^{|m|}(r)$ are located in the region of large radial distance r from the origin. Figure 1(a) shows a plot of the radial polynomials $Q_n(r)$ with $n = 0, 1, \ldots, 9$, and Fig. 1(b) shows the Zernike polynomial

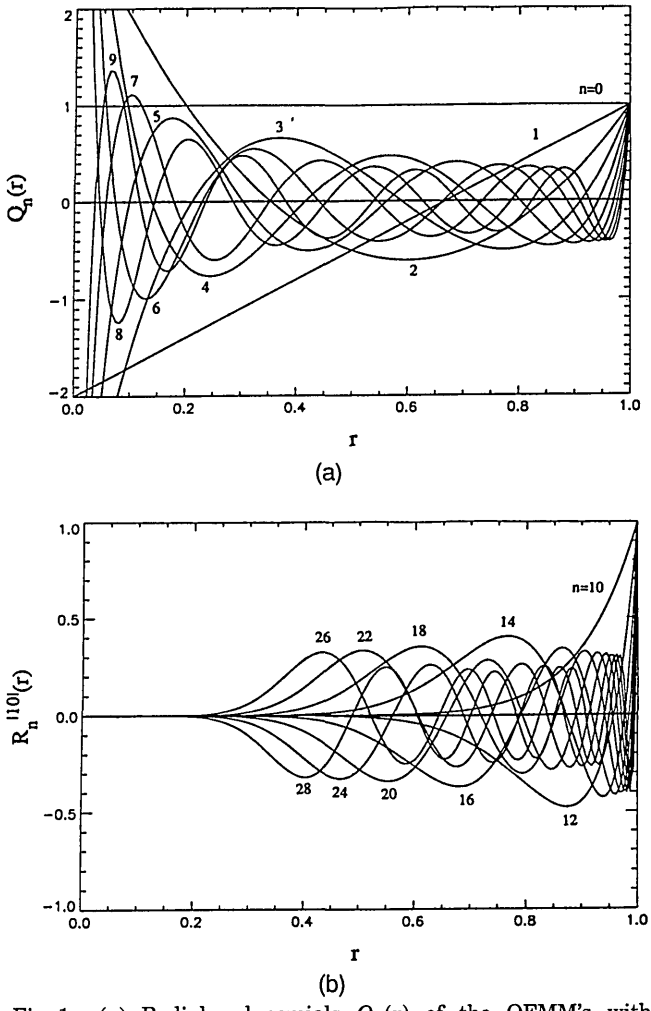


Fig. 1. (a) Radial polynomials $Q_n(r)$ of the OFMM's with $n = 0, 1, 2, \ldots, 9$; (b) Zernike radial polynomials $R_n^{|m|}(r)$ with $m = 10$ and $n = 10, 12, \ldots, 28$.

Table 1. Positions of the First Zeros of $Q_n(r)$ and $R_n^{|10|}(r)$

$Q_n(r)$		$R_n^{\|10\|}(r)$	
n	r	n	r
0	—	10	—
1	0.67	12	0.96
2	0.35	14	0.87
3	0.21	16	0.79
4	0.14	18	0.72
5	0.10	20	0.66
6	0.072	22	0.61
7	0.059	24	0.56
8	0.047	26	0.52
9	0.034	28	0.49

$R_n^{|m|}(r)$ with $|m| = 10$ and $n = 10, 12, \ldots, 28$. Both $Q_9(r)$ and $R_{28}^{10}(r)$ have nine zeros. The zeros of the $R_n^{10}(r)$ are in the region of large r. They are shifted to small r with increasing degree n. The first zero, which is closest to the origin, of the $R_{28}^{10}(r)$ is at $r = 0.49$. The first zero of $Q_9(r)$ is at $r = 0.034$, as shown in Table 1. This difference is important, since in scale-invariant pattern recognition the object sizes are unknown *a priori* and the moments of objects of different sizes should be calculated with the same basis functions. Hence, the ZM's have difficulty in describing small images, as we show in Section 3. The OFMM's are more suitable than the ZM's for scale- and rotation-invariant pattern recognition.

3. PERFORMANCE

Teh and Chin have recently evaluated the image's geometrical moment, Legendre moment, ZM, PZM, FMM, and CM. They found that the ZM and the PZM outperform the others in terms of overall performance for image description and robustness to noise.[4] In this section we compare the performance of the proposed OFMM with that of the ZM. We use measures of performance similar to those used by Teh and Chin. The expressions of the values are given, however, only for the OFMM.

A. Image Reconstruction

Pattern classification uses a finite number of OFMM's, Φ_{nm}, as image features, whose sum gives a reconstructed image:

$$\hat{f}(r,\theta) \simeq \sum_{n=0}^{N} \sum_{m=-M}^{M} \Phi_{nm} Q_n(r) \exp(jm\theta). \tag{13}$$

Because of the orthogonality of the set $Q_n(r)\exp(jm\theta)$, each OFMM makes an independent contribution to the reconstructed image. It can be shown that the partial sum, given by relation (13), converges to the image when the image is nonzero in a limited region, is continuous, or has a finite number of finite jumps so that $f(r,\theta)$ and $|f(r,\theta)|^2$ are integrable and the polynomials $Q_n(r)$ are bounded over $(0,1)$ as n becomes infinite.[8] These conditions are, in general, satisfied.

1. *Reconstruction Error for Deterministic Images*

We did an experiment of reconstruction of a deterministic image. The image was a capital letter E in a 64×64 matrix. The normalized image reconstruction error (NIRE) is defined as

$$\varepsilon^2 = \frac{\iint_{-\infty}^{\infty} [f(x,y) - \hat{f}(x,y)]^2 \mathrm{d}x\mathrm{d}y}{\iint_{-\infty}^{\infty} f^2(x,y)\mathrm{d}x\mathrm{d}y}, \tag{14}$$

where $\hat{f}(x,y)$ is the reconstructed image. The NIRE is shown in Fig. 2 as a function of the highest degree N of the radial polynomials. In the reconstruction with OFMM's, all the circular harmonic orders $m = 0, 1, 2, \ldots M$ with $M = N$ were used. In the reconstruction with ZM's, all the permissible orders m satisfying the condition $N \geq |m|$ and $N - |m| =$ even were used. With the same highest-degree N, OFMM's yield much lower reconstruction error than do ZM's.

Figure 3(a) shows the NIRE as a function of the total number of moments used in the reconstruction. The total number of OFMM's is equal to $(2M + 1)(N + 1)$, and we used $M = N$ in the experiment. The total number of ZM's is equal to $(N + 1)(N + 2)/2$. In fact, only approximately half of the moments are independent, because $\Phi_{n,-m} = \Phi_{n,m}{}^*$. The NIRE decreases with increasing number of moments. The NIRE's are almost the same in the reconstruction of a large image E6 with the

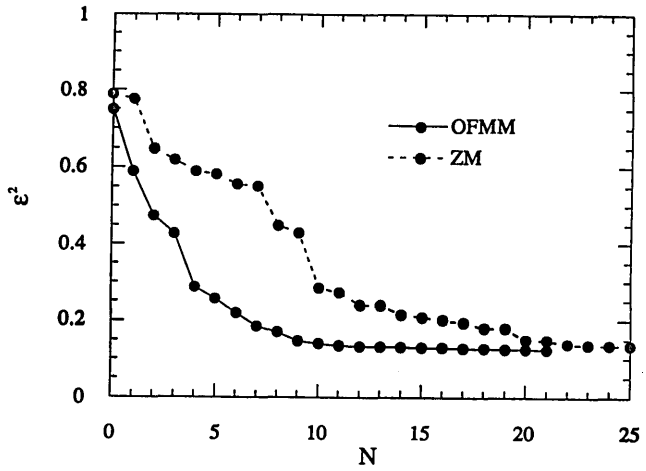
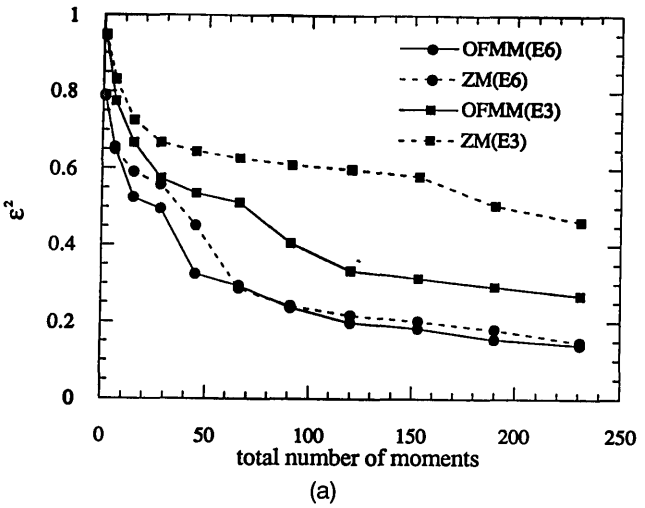


Fig. 2. Normalized image reconstruction error ε^2 for a letter E with OFMM's and ZM's, as a function of the highest degree N of the radial polynomials.



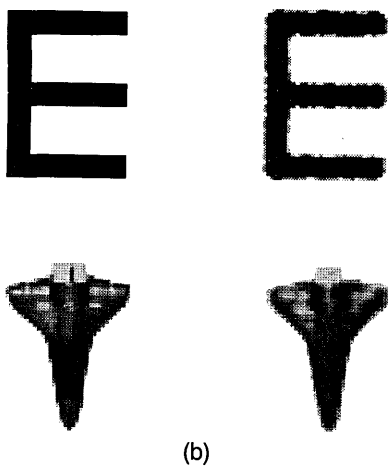


Fig. 3. (a) NIRE ε^2 for a letter E6 and a half-size letter E3 with OFMM's and ZM's, as a function of the total number of moments used in the reconstruction; (b) reconstructed binary and gray-level images with OFMM's of up to and including $n = 20$. Left, original images; right, reconstructed images.

same number of OFMM's and ZM's. But, for image E3, which was reduced to half the size of E6, the NIRE with ZM's is much higher than that of OFMM's, as can be seen from Fig. 3(a).

Figure 3(b) shows the images reconstructed with the use of 861 OFMM's, which are 431 independent OFMM's. The shapes of the images may be reconstructed with many fewer OFMM's. The same OFMM's describe equally all the continuously rotated and scaled versions of the original image. In this sense the OFMM's are highly condensed image features.

Eight reconstructed images of E3 from as many as 64 independent OFMM's and from the same numbers of ZM's are shown in Figs. 4 and 5. The images reconstructed

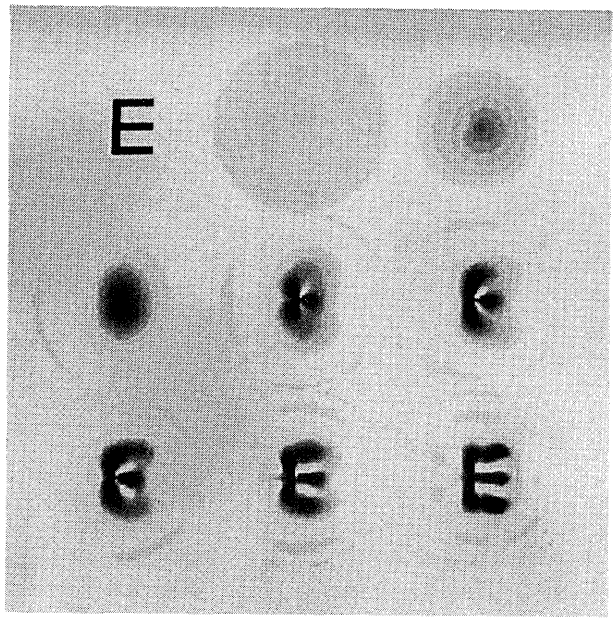


Fig. 4. Reconstructed images for E3 with OFMM's. From top left to bottom right: original image and reconstructed images with OFMM's of $N = M = 0, 1, 2, \ldots, 7$, with $N = M = 7$ corresponding to a total of 64 independent OFMM's. The reconstructed images were not thresholded.

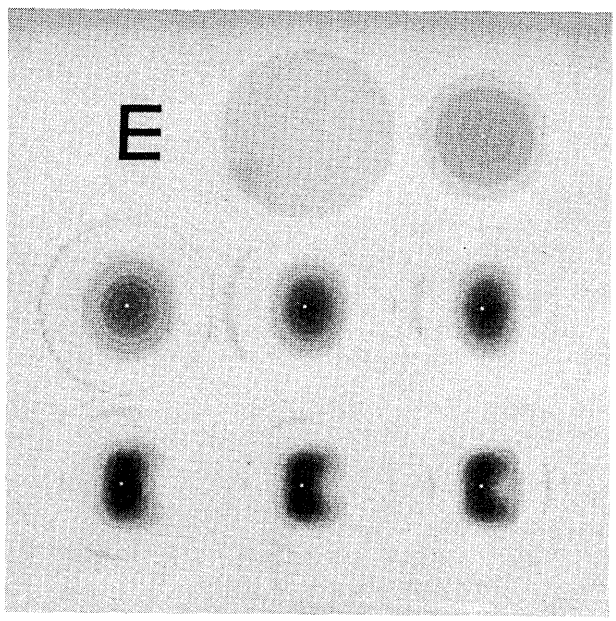


Fig. 5. Reconstructed images for E3 with ZM's. From top left to bottom right: original image and reconstructed images with ZM's of $N = 0, 2, 4, \ldots, 14$, with $N = M = 7$ corresponding to a total of 64 independent ZM's. The reconstructed images were not thresholded.

from the OFMM's have much better quality than those from the ZM's.

2. Reconstruction Error of Homogeneous Random Fields

We consider the image $f(x,y)$ a homogeneous (wide-sense stationary) stochastic process with zero mean:

$$E\{f(x,y)\} = 0, \tag{15}$$

where $E\{.\}$ is the expectation in ensemble averaging.[4] The normalized mean-square reconstruction error is defined as

$$\overline{\varepsilon^2} = \frac{E\left\{\iint_{-1}^{1}[f(x,y) - \hat{f}(x,y)]^2 \mathrm{d}x\mathrm{d}y\right\}}{E\left\{\iint_{-1}^{1}[f(x,y)]^2 \mathrm{d}x\mathrm{d}y\right\}}. \tag{16}$$

On substituting Eq. (13) for $\hat{f}(x,y)$ into Eq. (16) and considering that the object $f(x,y)$ can be scaled by a factor k to $f(x/k, y/k)$ and that the maximum size of the scaled object is equal to $k \le 1$, we obtain the theoretical or statistical NIRE as

$$\overline{\varepsilon^2}(N,M) = 1 - \frac{1}{\pi k^2}\sum_{n=0}^{N}\sum_{m=-M}^{M}\frac{n+1}{\pi} \times \int_0^{2\pi}\int_0^{k}\int_0^{2\pi}\int_0^{k}\frac{C_{ff}(x,y,u,v)}{C_{ff}(0,0)} \times Q_n(r)Q_n(\rho)\cos[m(\theta-\phi)]r\mathrm{d}r\mathrm{d}\theta\rho\mathrm{d}\rho\mathrm{d}\phi, \tag{17}$$

where $u = \rho\cos\phi$, $v = \rho\sin\phi$. The autocorrelation $C_{ff}(x,y;u,v)$ of the stochastic process $f(x,y)$ is expressed as

$$C_{ff}(x,y,u,v) = C_{ff}(0,0)\exp(-\alpha_1|x-u| - \alpha_2|y-v|), \tag{18}$$

according to the experimental evidence,[10] where α_1 and α_2 are positive constants and

$$C_{ff}(0,0) = E\{[f(x,y)]^2\} = \frac{1}{\pi k^2}\int_0^{2\pi}\int_0^{k}[f(r,\theta)]^2 r\mathrm{d}r\mathrm{d}\theta \tag{19}$$

is the average energy per unit area. The experiments that led to the autocorrelation function of Eq. (18) were made with television images. We modify Eq. (18) for the images having no privilege-shift directions as

$$C_{ff}(x,y,u,v) = C_{ff}(0,0)\exp\{-\alpha[(x-u)^2 + (y-v)^2]^{1/2}\}, \tag{20}$$

where the constant α is determined by experiments. Figure 6 shows the function $C_{ff}(x,y;u,v)$ with different values of α. The dashed curve is the autocorrelation of a uniform unit circle. The autocorrelation of an image that does not fill the unit circle should be below this dashed curve. We choose $\alpha = 3$ in the experiments that follow. When the image is scaled by a factor k, the scaled autocorrelation becomes

$$C_{ff}(x,y,u,v) = \begin{cases} C_{ff}(0,0)\exp\left[-\dfrac{\alpha}{k}R\right] & R \le 2k \\ 0 & R > 2k \end{cases}, \tag{21}$$

where $R = [(x-u)^2 + (y-v)^2]^{1/2}$.

Figure 7 shows the statistical NIRE as a function of the total number of moments used in the reconstruction. The OFMM and the ZM have similar performance for full-size objects when $k = 1.0$. But for a smaller object with $k = 0.5$, OFMM's yield a much smaller reconstruction error than do ZM's. This result is in agreement with the experimental result for reconstruction of deterministic images shown in Subsection 3.A.1.

B. Noise Sensitivity

Sensitivity to noise is a critical issue for image moments. We compare the OFMM and the ZM for their signal-to-noise ratios (SNR's) and their performance for image description in the presence of noise.

1. Signal-to-Noise Ratio

Assume that the image $f(x,y)$ is a homogeneous random field with zero mean and is corrupted by a zero-mean additive white noise with the autocorrelation of the noise as

$$C_{nn}(x,y,u,v) = \sigma^2\delta(x-u, y-v), \tag{22}$$

where $\delta(x-u, y-v)$ is the delta function and σ^2 is the spectral density of the noise. The statistical SNR of the

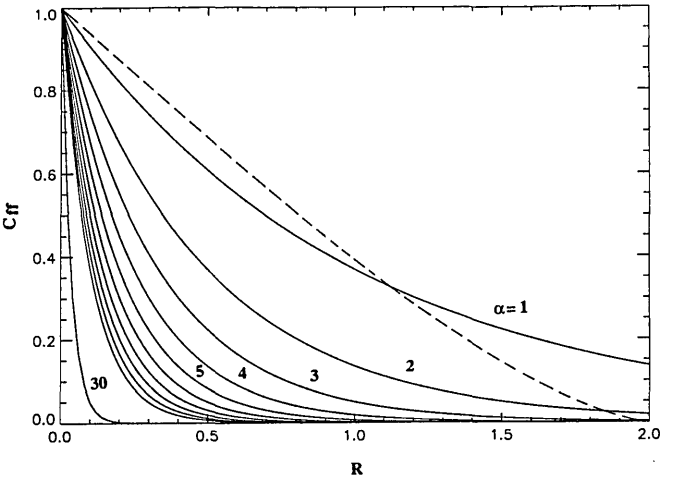


Fig. 6. Autocorrelation $C_{ff}(x,y;u,v)$ as a function of $R = [(x-u)^2 + (y-v)^2]^{1/2}$ with different constants α. The dashed curve is the autocorrelation of a uniform unit circle.

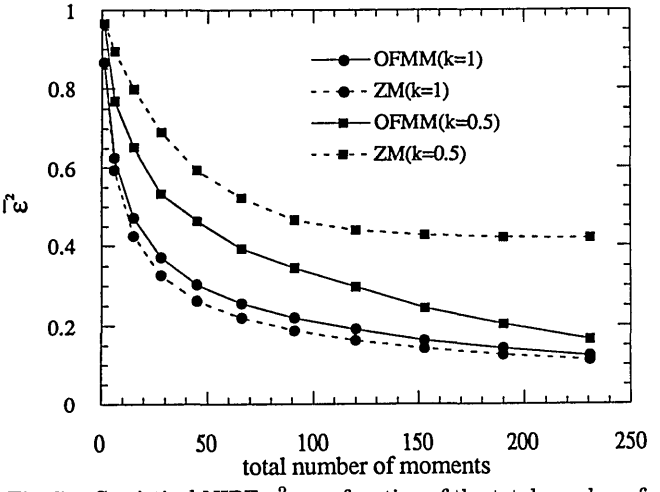


Fig. 7. Statistical NIRE ε^2 as a function of the total number of the OFMM's and ZM's used in the reconstruction.

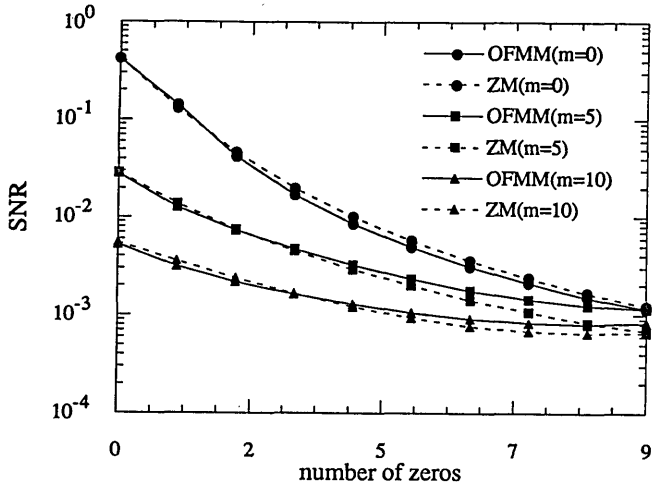


Fig. 8. Statistical SNR of the OFMM and the ZM with constant $\alpha = 3$, image size $k = 1$, and order $m = 0, 5, 10$ as a function of the number of zeros of their radial polynomials.

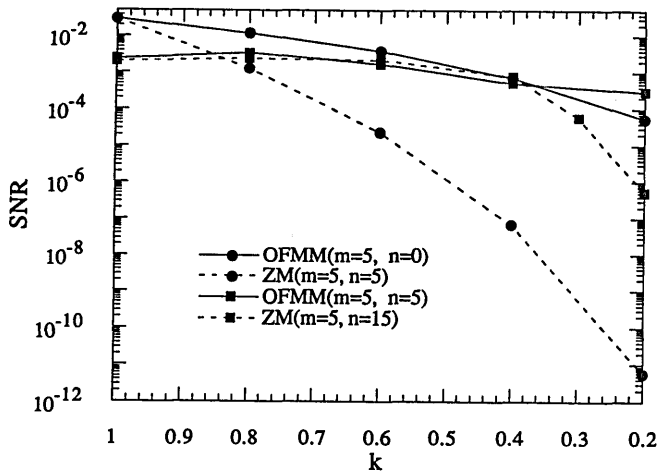


Fig. 9. Statistical SNR of OFMM's with $m = 5$, $n = 0, 5$ and of ZM's with $m = 5$, $n = 5, 15$, with the number of zeros equal to 0 and 5 respectively, as a function of image size k.

OFMM may be defined as

$$\mathrm{SNR}_{nm} = \frac{\mathrm{var}\{(\Phi_{nm})_f\}}{\mathrm{var}\{(\Phi_{nm})_{\mathrm{noise}}\}} = \frac{1}{\sigma^2}\,\mathrm{var}\{(\Phi_{nm})_f\}, \tag{23}$$

where the variance of the OFMM's for the random signal $f(x,y)$ is

$$\mathrm{var}\{(\Phi_{nm})_f\} = \int_0^{2\pi}\int_0^{k}\int_0^{2\pi}\int_0^{k} C_{ff}(x,y,u,v)Q_n(r)Q_n(\rho) \times \cos[m(\theta-\phi)]r\,dr\,d\theta\,\rho\,d\rho\,d\phi\,, \tag{24}$$

with the autocorrelation of the signal C_{ff} given by Eq. (21).

Figure 8 shows the statistical SNR's normalized by $C_{ff}(0,0)/\sigma^2$ of the OFMM and the ZM as a function of the number of zeros of the radial polynomials $Q_n(r)$ and $R_n^{|m|}(r)$, which is equal to the degree n for the OFMM and to $(n - |m|)/2$ for the ZM. The SNR decreases with increasing n and m. When the object size $k = 1$, the OFMM and the ZM have almost the same SNR's. But when the object size k decreases, the SNR of the ZM's decreases much faster than that of the OFMM's, as shown in Fig. 9. This is because the ZM decreases with decreasing object size k faster than does the OFMM; this can be understood from Fig. 1.

2. Noisy Image Reconstruction

Assume that an object of size $k \le 1$ is degraded by an additive random noise $n(x,y)$ and that the unit circle in the input scene is full of the noise. The OFMM's of the noisy image are additions of the OFMM's of the object, Φ_{nm}, and of the OFMM's of the noise, $(\Phi_{nm})_n$. The OFMM's of the noisy image then are used for image description and classification, and the error of the images reconstructed by the OFMM's of the noisy image is a measure of the effect of the noise on the pattern-recognition performance of the OFMM.

If we assume again that the signal $f(x,y)$ is a zero-mean homogeneous random process and that the noise is zero mean with variance σ^2, the statistical NIRE of the noisy image is[4]

$$\begin{aligned}\overline{\varepsilon_n^2}(N,M) &= \frac{E\left\{\iint_{-1}^{1}[f(x,y)-\hat{f}(x,y)-\hat{n}(x,y)]^2\,dx\,dy\right\}}{E\left\{\iint_{-1}^{1}[f(x,y)]^2\,dx\,dy\right\}} \\ &= \overline{\varepsilon^2}(N,M) + \frac{E\left\{\int_0^{2\pi}\int_0^{1}[\hat{n}(r,\theta)]^2 r\,dr\,d\theta\right\}}{E\left\{\int_0^{2\pi}\int_0^{1}[f(r,\theta)]^2 r\,dr\,d\theta\right\}} \\ &= \overline{\varepsilon^2}(N,M) + \frac{N_{\mathrm{total}}}{\pi\mathrm{SNR}_{\mathrm{input}}}\,,\end{aligned} \tag{25}$$

where $\hat{f}(r,\theta)$ is the partial sum of the Φ_{nm} given by Eq. (13), $\hat{n}(r,\theta)$ is the partial sum of the $(\Phi_{nm})_n$, and $\overline{\varepsilon^2}$ is the statistical NIRE of the OFMM's of the image without noise given by Eq. (17). The second term on the right-hand side of Eq. (25) is the normalized error caused by the noise. N_{total} is the total number of moments used in the reconstruction. $\mathrm{SNR}_{\mathrm{input}}$ is defined as

$$\mathrm{SNR}_{\mathrm{input}} = \frac{k^2 C_{ff}(0,0)}{\sigma^2}\,. \tag{26}$$

Figure 10 shows the statistical NIRE of the noisy image $\overline{\varepsilon_n^2}$ of OFMM's and ZM's as a function of the total number of moments for the object sizes $k = 0.5$ and $k = 1$ with $\mathrm{SNR}_{\mathrm{input}} = 100$. The second term in Eq. (25) is the same for OFMM's and ZM's. Only the statistical NIRE's, $\overline{\varepsilon^2}$, of the image without noise are different. For a large object with $k = 1$, the OFMM and the ZM have similar performance. For a small object with $k = 0.5$, the reconstruction error with OFMM's is smaller than that with ZM's.

For a given input SNR there is an optimum number of moments that generate the minimum NIRE for a noisy image. The high-order moments are sensitive to noise. Above a certain order the contribution of the noise can exceed the contribution of the moments to the image reconstruction. The NIRE cannot be further reduced by addition of more moments.

Figure 11 shows the experimental NIRE of the OFMM's and ZM's for two deterministic objects, E6 and E3, with noise. The objects E6 and E3 have binary intensities 0 and 255. The input noise was additive Gaussian noise with zero mean and a standard deviation equal to 345

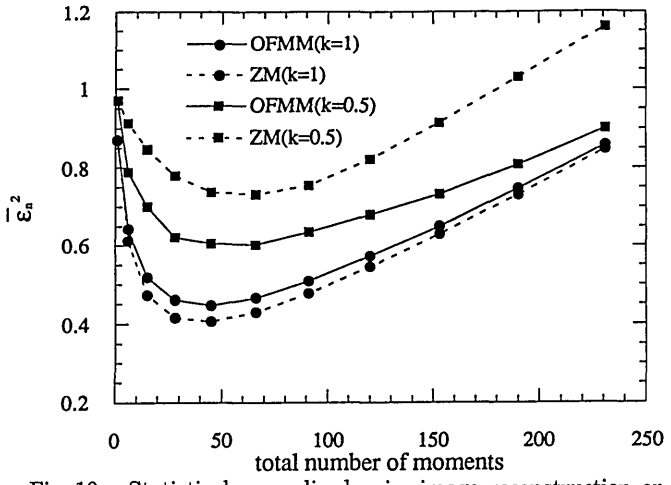


Fig. 10. Statistical normalized noisy-image reconstruction error $\bar{\varepsilon}_n^2$ with input SNR = 100 as a function of the total number of the OFMM's and ZM's used in the reconstruction.

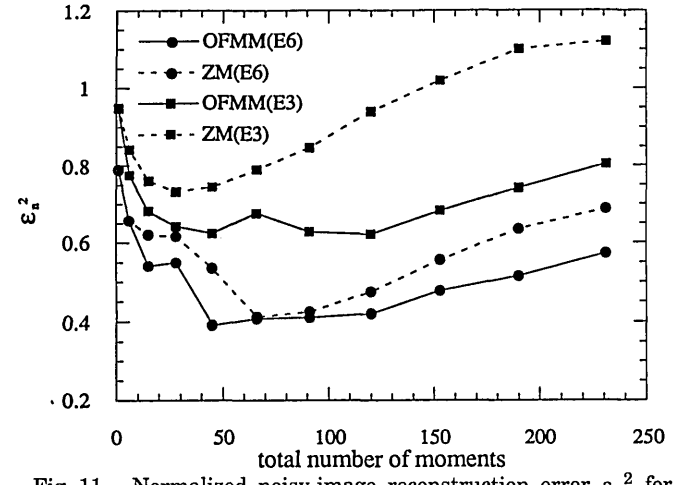


Fig. 11. Normalized noisy-image reconstruction error ε_n^2 for deterministic objects E6 and E3 with additive noise and input SNR = 100, as a function of the total number of the OFMM's and ZM's used in the reconstruction.

for E6 and 185 for E3. Those noise levels correspond to the $\mathrm{SNR}_{\mathrm{input}}$ defined in Eq. (26) as equal to 100 with the variance σ^2 calculated according to Eq. (22). These experimental results agree approximately with the theoretical results shown in Fig. 10. Figures 12 and 13 show reconstructed images E3 with up to 64 independent OFMM's and ZM's of the noisy image, respectively. The quality of the reconstructed images with the OFMM's is much better than that with the ZM's.

4. PATTERN RECOGNITION

A. Normalization of the Orthogonal Fourier–Mellin Moments

OFMM's may be normalized to be invariant to shift and rotation and to scale and intensity changes of the image. First we determine the center of image by using the first-order geometrical moments. The center of image is used as the origin of the coordinate system. All the moments calculated in this coordinate system are shift invariant; then we calculate the FMM's, M_{sm}, in the Cartesian coordinate system, using the corresponding modified complex moments, C_{pq}, with $p = (s - m)/2$ and $q = (s + m)/2$; we normalize the FMM's against changes in scale and intensity. When an image $f(r, \theta)$ is scaled by a factor k and its intensity is changed by a factor g, its FMM's become

$$M_{sm}' = \int_0^{2\pi} \int_0^1 gf(r/k, \theta) r^s r \mathrm{d}r \mathrm{d}\theta = gk^{s+2} M_{sm}, \quad (27)$$

where M_{sm} is the FMM of the original image. We obtain the scale and intensity factors k and g, respectively, using

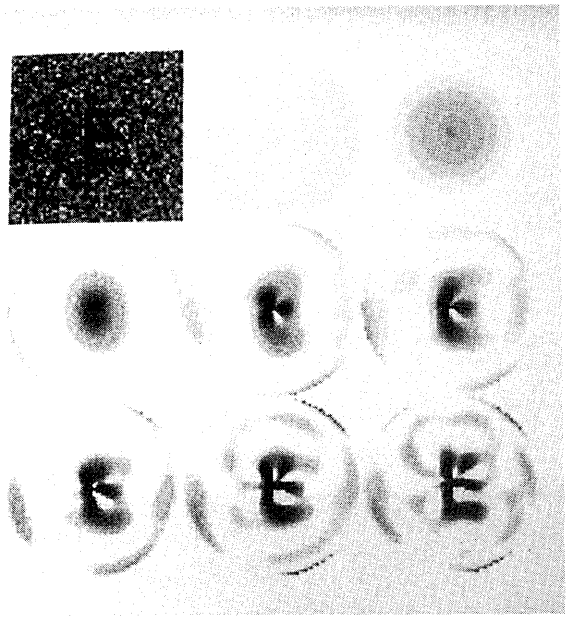


Fig. 12. Reconstructed images with the OFMM's of the E3 corrupted by an additive Gaussian noise of zero mean and $\sigma = 185$. From top left to bottom right: original image and reconstructed images with OFMM's of $N = 0, 1, 2, \ldots, 7$, with $N = 7$ corresponding to a total of 64 independent OFMM's. The reconstructed images were not thresholded.

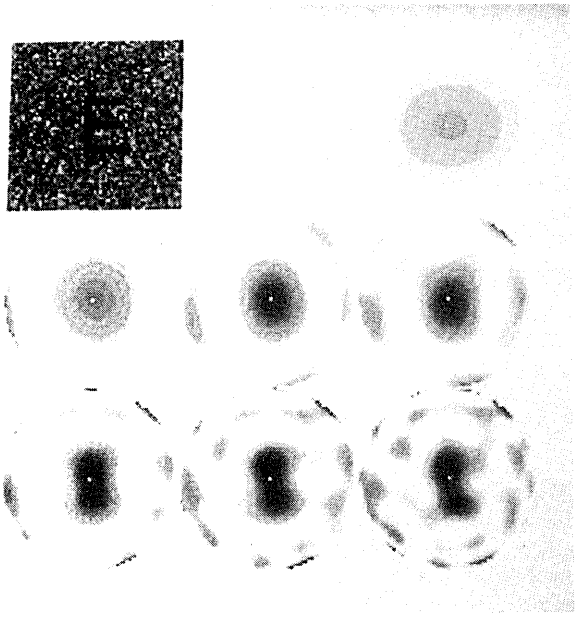


Fig. 13. Reconstructed images with the ZM's of the E3 corrupted by additive Gaussian noise of zero mean and $\sigma = 185$. From top left to bottom right: original image and reconstructed images with ZM's of $N = 0, 2, 4, \ldots, 14$, with $N = 14$ corresponding to a total of 64 ZM's. The reconstructed images were not thresholded.

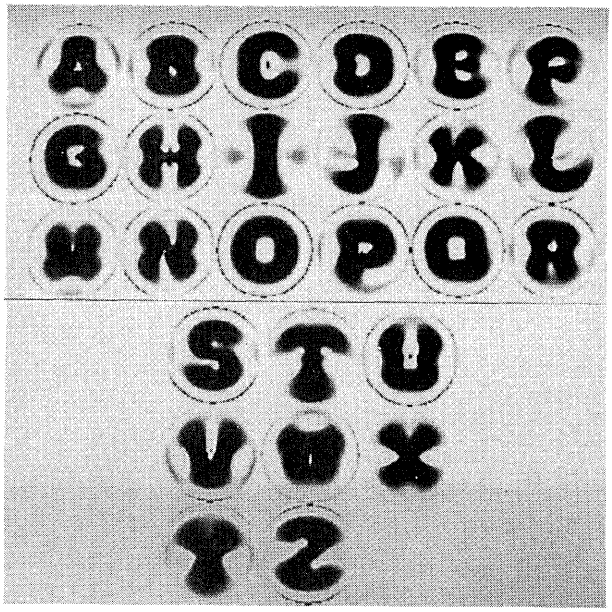


Fig. 14. Reconstructed images of the 26 English alphabet letters with 25 OFMM's of n and $m = 0, 1, 2, 3, 4$.

the low-order FMM's:

$$K = \left(\frac{M_{10}'}{M_{00}'}\right) \Big/ \left(\frac{M_{10}}{M_{00}}\right), \tag{28}$$

$$g = \left[\left(\frac{M_{10}}{M_{00}}\right) \Big/ \left(\frac{M_{10}'}{M_{00}'}\right)\right]^2 \left(\frac{M_{00}'}{M_{00}}\right). \tag{29}$$

The normalized FMM's are obtained as $M_{sm}'/(gk^{s+2})$, which are scale and intensity invariant.

When computing k and g we need to determine the values of M_{10} and M_{00} in Eqs. (28) and (29), which are the same for all the images in the training and testing sets. We choose simply $M_{00} = 1$. According to Eq. (28), the ratio M_{10}/M_{00} is related to the scale k of the image. We choose the ratio M_{10}/M_{00} slightly smaller than the minimum M_{10}'/M_{00}' of all the images in the training set to ensure that the normalized images remain inside the unit circle.

Finally, we calculate the OFMM's with the normalized FMM's, according to Eq. (7). A rotation of the image by angle θ_0 results in the same phase factor $\exp(jm\theta_0)$ for all the M_{sm} in Eq. (7). The modulus of the OFMM, $|\Phi_{nm}|$, is shift, scale, intensity, and rotation invariant.

B. Pattern-Recognition Experiment

The training set consisted of binary images of the 26 English alphabet letters with intensities of 0 and 255 in 64×64 pixel matrices. We used 25 OFMM's with $n = 0, 1, 2, 3, 4$, and $m = 0, 1, 2, 3, 4$ to describe the images. Figure 14 shows the images reconstructed with the 25 OFMM's. We estimated that the differences among the reconstructed images were large enough for image classification. Image reconstruction is not a necessary step of pattern recognition. We found that it was also possible to use fewer than 25 OFMM's in this particular pattern-recognition experiment. In general, one can describe the images by using a large number of OFMM's. The dimension of the feature space can be reduced further by use of statistical feature-selection algorithms such as the Karhunen–Loéve transform.[11]

Because of the scale and intensity normalization, the OFMM's Φ_{10} and Φ_{00} are equal for all the objects. We then classified the objects in the 23-dimensional feature space, using the weighted minimum-mean-distance rule.[12] The weighted distance was calculated as

$$d_i = \left\{ \sum_{\substack{n,m=0 \\ \text{when } m=0,\, n\neq 0,1}}^{4} \frac{[|\Phi_{nm}| - (|\Phi_{nm}|)_i]^2}{(\sigma_{nm})_i^2} \right\}^{1/2}, \tag{30}$$

where $|\Phi_{nm}|$ is the modulus of the OFMM of the testing object, $(|\Phi_{nm}|)_i$ is the OFMM of the reference object of class i, and $(\sigma_{nm})_i^2$ is the in-class variance of the $(|\Phi_{nm}|)_i$. The $(\sigma_{nm})_i^2$ is used as a weighting factor to the Euclidean distance. The values of $|\Phi_{nm}|$ change with the orders n and m. The $|\Phi_{nm}|$ with large values dominated the Euclidean distance, while the $|\Phi_{nm}|$ with small values had no effect on the classification. The weighting by $(\sigma_{nm})_i^2$ is to balance this effect. We rotated each reference object by 15, 30, and 45 deg and scaled it 1.5 and 2 times to determine the in-class variance of the OFMM's.

A testing image was classified to the class i for which the distance d_i was minimum. The testing set images were the rotated (20 and 130 deg), scaled (1.8 and 2.2 times), and intensity (3 times) changed versions of the training-set images and without or with zero-mean additive noise of $\sigma = 25.5$ and $\sigma = 255$. The testing set for the class of letter A is shown in Fig. 15. The misclassification rates were zero for all testing images, except for images of D, O, and Q with a high-noise-level noise $\sigma = 255$. We see from Fig. 14 that the reconstructed images D, O, and Q are similar to one another. More OFMM's have to be used to classify those images.

Figure 16 shows the misclassification rate of the OFMM's and ZM's for noisy images, as a function of

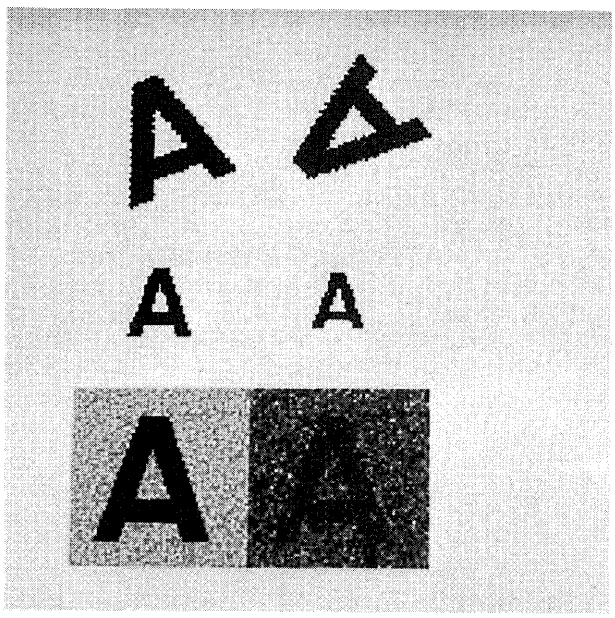


Fig. 15. Testing images for the letter A. From top left to bottom right: the letter A rotated by 20 and 130 deg, scaled by $k = 0.555$ and 0.455, and with zero-mean additive noise of $\sigma = 25.5$ and $\sigma = 255$.

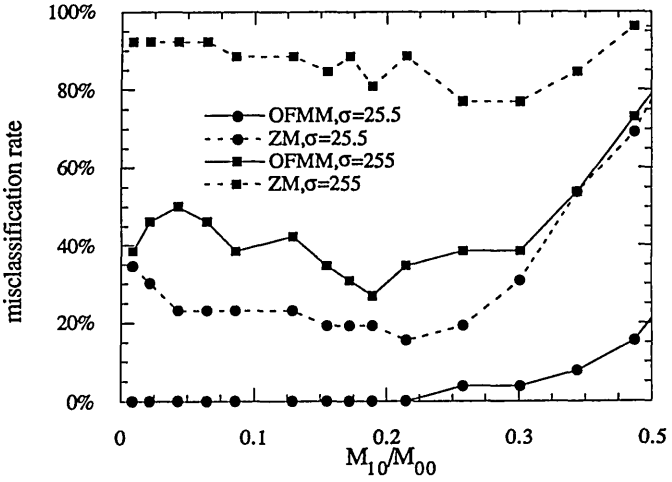


Fig. 16. Misclassification rate of the OFMM's and the ZM's with input noise $\sigma = 25.5$ and $\sigma = 255$ as a function of the ratio M_{10}/M_{00} that was used to calculate the scale ratio.

the ratio M_{10}/M_{00}. This ratio determines the size of the normalized images. The misclassification rate of the ZM's increases faster than that of the OFMM's with increasing M_{10}/M_{00}. This is because, with the influence of the input noise, the size of the normalized images can exceed the unit circle and cause significant errors. ZM's are more sensitive than OFMM's to noise. The misclassification rate of the ZM's also increases when the ratio $M_{10}/M_{00} < 0.05$, because ZM's are not good at describing small images. The performance of the OFMM's is not affected by a small value of M_{10}/M_{00}.

5. CONCLUSIONS

We have introduced the orthogonal Fourier–Mellin moments based on a new set of radial polynomials of low degree. We showed that the orthogonal Fourier–Mellin moments are the generalized Zernike moments and the orthogonalized complex moments. The new radial polynomials have more zeros than do the Zernike radial polynomials of the same degree. Compared with a high-degree Zernike polynomial that has the same number of zeros as the new polynomial, the new polynomial has more zeros in the region of small radial distance. The orthogonal Fourier–Mellin moments are therefore more suitable than the Zernike moments for scale- and rotation-invariant pattern recognition, in which small objects are described in the same way as large objects. We showed that for small objects the performance of the orthogonal Fourier–Mellin moments is superior to that of the Zernike moments in terms of image-reconstruction error, signal-to-noise ratio, and noisy-image-reconstruction error. A scale-, rotation-, shift-, and intensity-invariant pattern-recognition example with use of the orthogonal Fourier-Mellin moments has been shown.

APPENDIX A: POLYNOMIALS $Q_n(r)$

We obtain the well-known Jacobi polynomials $G_n(p,q,r)$ by orthogonalizing the sequence

$$1, r, r^2, \ldots r^n \tag{A1}$$

over the interval $0 \le r \le 1$ with a general weighting function,

$$w(r) = r^{q-1}(1-r)^{p-q}, \tag{A2}$$

such that

$$\int_0^1 w(r)G_n(p,q,r)G_{n'}(p,q,r)\mathrm{d}r = b_n(p,q)\delta_{nn'}, \tag{A3}$$

where the normalization constant $b_n(p,q)$ is

$$b_n(p,q) = \frac{n![(q-1)!]^2(p-q+n)!}{(q+n-1)!(p+n-1)!(p+2n)}. \tag{A4}$$

The Jacobi polynomials are expressed as

$$G_n(p,q,r) = \frac{n!(q-1)!}{(p+n-1)!} \times \sum_{s=0}^{n}(-1)^s \frac{(p+n+s-1)!}{(n-s)!\,s!(q+s-1)!} r^s. \tag{A5}$$

The orthogonal polynomials before normalization $Q_n(r)$ satisfy

$$\int_0^1 Q_n(r)Q_k(r)r\mathrm{d}r = a_n\delta_{nk}, \tag{A6}$$

where δ_{nk} is the Kronecker symbol and a_n is a normalization constant. From Eqs. (A3) and (A2) we have, with the Jacobi polynomials of $p = q = 2$,

$$\int_0^1 G_n(2,2,r)G_k(2,2,r)r\mathrm{d}r = \frac{1}{2(n+1)^3}\delta_{nk}. \tag{A7}$$

Hence

$$Q_n(r) = a_n{}^{1/2}[2(n+1)^3]^{1/2}G_n(2,2,r). \tag{A8}$$

We choose the normalization so that $Q_n(r) = 1$ for $r = 1$. The value $G_n(2,2,1)$ may be obtained from the generating function of the Jacobi polynomials:

$$\frac{\{z-1+[1-2z(1-2t)+z^2)]^{1/2}\}^m}{(2zt)^m[1-2z(1-2t)+z^2]^{1/2}} = \sum_{s=0}^{\infty}\binom{m+s}{s}G_s(m+1,m+1,t)z^s. \tag{A9}$$

If we let $t = r$ and $r = 1$, the left-hand side of Eq. (A9) reduces to $(1+z)^{-1}$. Setting $m = 1$, we obtain

$$G_n(2,2,1) = \frac{(-1)^n}{(n+1)}. \tag{A10}$$

Hence the normalization constant is

$$(a_n)^{1/2} = \frac{(-1)^n}{(2n+2)^{1/2}}. \tag{A11}$$

On substituting for a_n into Eq. (A8), we have

$$Q_n(r) = (-1)^n\binom{n+1}{n}G_n(2,2,r). \tag{A12}$$

But from Eq. (A5) the Jacobi polynomials are

$$G_n(2,2,r) = \frac{1}{n+1}\sum_{s=0}^{n}(-1)^s\frac{(n+s+1)!}{(n-s)!\,s!(s+1)!}r^s. \quad \text{(A13)}$$

Thus $Q_n(r)$ is obtained as

$$Q_n(2,2,r) = (-1)^n\sum_{s=0}^{n}(-1)^s\frac{(n+s+1)!}{(n-s)!\,s!(s+1)!}r^s. \quad \text{(A14)}$$

APPENDIX B: RELATIONS AMONG THE $Q_n(r)$, $R_n^{|m|}(r)$, AND $P_n^{|m|}(r)$

The generating function of the Jacobi polynomials is

$$\frac{\{z-1+[1-2z(1-2t)+z^2]^{1/2}\}^m}{(2zt)^m[1-2z(1-2t)+z^2]^{1/2}} = \sum_{s=0}^{\infty}\binom{m+s}{s}G_s(m+1,m+1,t)z^s. \quad \text{(B1)}$$

The orthogonal polynomials $Q_n(r)$ for the OFMM's are expressed in Eq. (A12):

$$Q_n(r) = (-1)^n\binom{n+1}{n}G_n(2,2,r). \quad \text{(B2)}$$

Setting $m = 1$ and $t = r$ and $z = -z$ in Eq. (B1), we obtain the generating function for $Q_n(r)$ as

$$\frac{z+1-[1+2z(1-2r)+z^2]^{1/2}}{2zr[1+2z(1-2r)+z^2]^{1/2}} = \sum_{s=0}^{\infty}Q_s(r)z^s. \quad \text{(B3)}$$

With $n = m + 2s$, the Zernike polynomials are related to the Jacobi polynomials by

$$R_{m+2s}^{|m|}(r) = (-1)^n\binom{m+s}{s}r^mG_s(m+1,m+1,r^2). \quad \text{(B4)}$$

Multiplying the two sides of Eq. (B1) by $r^{|m|}$ and setting $t = r^2$ and $z = -z$, we obtain the generating function for the Zernike polynomials:

$$\frac{\{z+1-[1+2z(1-2r^2)+z^2)]^{1/2}\}^m}{(2zr)^m[1+2z(1-2r^2)+z^2]^{1/2}} = \sum_{s=0}^{\infty}R_{m+2s}^{|m|}(r)z^s. \quad \text{(B5)}$$

With $n = m + s$, the pseudo-Zernike polynomials are related to the Jacobi polynomials by

$$P_{m+s}^{|m|}(r) = (-1)^s\binom{2m+s+1}{s}r^mG_s(2m+2,2m+2,r). \quad \text{(B6)}$$

Multiplying the two sides of Eq. (B1) by r^m and setting $t = r$ and $z = -z$, we obtain the generating function for the pseudo-Zernike polynomials:

$$\frac{\{z+1-[1+2z(1-2r)+z^2]^{1/2}\}^{2m+1}}{(2z)^{2m+1}r^{m+1}[1+2(1-2r)+z^2]^{1/2}} = \sum_{s=0}^{\infty}P_{m+s}^{|m|}(r)z^s. \quad \text{(B7)}$$

A comparison of Eqs. (B3) and (B5) shows the relation between the $Q_n(r)$ and the Zernike polynomials:

$$rQ_n(r^2) = R_{2n+1}^1(r), \quad \text{(B8)}$$

and a comparison of Eqs. (B3) and (B7) shows the relation between the $Q_n(r)$ and the pseudo-Zernike polynomials:

$$Q_n(r) = P_n^0(r). \quad \text{(B9)}$$

Also, a comparison of Eqs. (B5) and (B7) shows the relation between the Zernike and the pseudo-Zernike polynomials:

$$rP_n^{|m|}(r^2) = R_{2n+1}^{2|m|+1}(r). \quad \text{(B10)}$$

REFERENCES

1. M. K. Hu, "Visual pattern recognition by moment invariants," IEEE Trans. Inf. Theory **IT-8,** 179–187 (1962).
2. Y. S. Abu-Mostafa and D. Psaltis, "Recognition aspect of moment invariants," IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-6,** 698–706 (1984).
3. M. O. Freeman and B. E. A. Saleh, "Moment invariants in the space and frequency domains," J. Opt. Soc. Am. A **5,** 1073–1084 (1988).
4. C. H. Teh and R. T. Chin, "On image analysis by the methods of moments," IEEE Trans. Pattern Anal. Mach. Intell. **10,** 496–513 (1988).
5. M. R. Teague, "Image analysis via the general theory of moments," J. Opt. Soc. Am. **70,** 920–930 (1980).
6. A. B. Bhatia and E. Wolf, "On the circular polynomials of Zernike and related orthogonal sets," Proc. Cambridge Philos. Soc. **50,** 40–48 (1954).
7. Y. Sheng and J. Duvernoy, "Circular-Fourier–Radial-Mellin descriptors (FMD's) for pattern recognition," J. Opt. Soc. Am. A **3,** 885–888 (1986).
8. D. Jackson, *Fourier Series and Orthogonal Polynomials* (Mathematical Association of America, New York, 1941), Chaps. VII and XI.
9. H. H. Arsenault and Y. Sheng, "Properties of the circular harmonic expansion for rotation-invariant pattern recognition," Appl. Opt. **25,** 3225–3229 (1986).
10. E. R. Kretzmer, "Statistics of television signals," Bell Syst. Tech. J. **31,** 751–763 (1952).
11. K. Fukunaga, *Statistical Pattern Recognition*, 2nd ed. (Academic, Boston, 1990), Chap. 9.
12. A. Khotanzad and Y. H. Hong, "Invariant image recognition by Zernike moments," IEEE Trans. Pattern Anal. Mach. Intell. **12,** 489–497 (1990).