Generalized Affine Invariant Image Normalization

Dinggang Shen and Horace H. S. Ip

Abstract—We provide a generalized image normalization technique which basically solved all problems in image normalization. The orientation of any image can be uniquely defined by at most three non-zero generalized complex (GC) moments. The correctness of our method is demonstrated theoretically as well as in practice by applying them to a number of "degenerate" images which have failed other previously reported techniques for image normalization.

Index Terms—Image normalization, image orientation, invariant image matching, symmetry detection, fold detection, complex moment, rotationally symmetric image.

1 Introduction

Techniques on invariant pattern recognition include integral transformations, construction of algebraic moments and the use of structured neural networks [1]. The image normalization method, developed as an elegant preprocessing technique, transforms the distorted input pattern into its corresponding normal form such that it is invariant under translation, scaling, skew, and rotation.

The following overview will clarify that the existing techniques do not provide for a complete solution to the problem in the sense that they are not able to produce a unique normalization representation for any type of images, particularly for the rotationally symmetric images (RSI). More importantly, it becomes clear to us that the detection of the orientation and the fold number of the compact image are the key issues to be tackled in order to provide a complete solution for generalized image normalization.

Abu-Mostafa and Psaltis [2] used complex moment and circular harmonic coefficient function to analyze RSI, while the technique is not applicable to the skewed patterns. Leu [3] proposed a *compact algorithm* to handle skew distortion efficiently. However, the problem of rotational invariance remained unsolved. Pei [4] later extended the method to achieve rotation invariance. Based on the tensor theory, he derived simple equations based on the third-order central moments of the compact image. From these equations an orientation could be calculated to make the pattern invariant to rotation. However, as noted in [5], the method fails when all the third-order central moments of the tested pattern are zero, which happens frequently for the RSIs. For such cases, Pei [5] further introduced a method called the modified Fourier descriptor to normalize the RSI. There are certain limitations on the method. For example, based on the observation that the first non-zero Fourier coefficient is located on the fundamental frequency, the method was able to detect the fold number and the rotation angle simultaneously. Unfortunately, this observation is not always true for most cases, particularly for patterns in Fig. 1c through Fig. 1g and Fig. 1i. Pei [6] further proposed to reduce the 2D pattern in the radial direction to several 1D discrete data sequences, not to a single 1D sequence as done in [5]. Since these 1D discrete data sequences are locally extracted, this method is unable to detect the fold number and rotation angle of the pattern where the internal fold number is different from the external fold number. Examples of such patterns are shown in Figs. 1e and Fig. 1i. Also, this method does not solve other limitations in [5].

There are many techniques developed for detecting a pattern's orientation, such as principal axes [7], shape matrices [8], mirror-symmetry axes [9], [10, [11], the line through the centroid, and radius weighted mean [12]. However, these methods are usually inapplicable when the input pattern is an RSI, or not mirror symmetric. To remedy this, GPA [13], FPA [14], and FISSP [15] were proposed for defining the orientation of an RSI. However, a disadvantage respectively related to GPA and FPA is that "double-matching" or "multi-matching" is required in real applications. In FISSP, two shape-specific points often do not exist for many kinds of shapes. Furthermore, since all the above methods require the fold number of an RSI be a known priori, they are neither efficient nor generalized.

Leou and Tsai [17] solved the fold number of a simple rotationally-symmetric closed contour by finding the number of crossing points between the average radius and the contour. In general, most RSIs cannot be easily represented by only one single closed contour, particularly for RSIs with complicated boundary and inside hole. In [18], Lin detected the fold number by a simple mathematical property, and the orientation of an RSI was viewed as a by-product of his method. However, this by-product fails when the pattern's fold number is a factor of the smallest order making the complex moment non-zero. Examples of RSIs making the by-production failed are Figs. 1c through 1g and 1i. Lin [19] developed a convenient tool to define universal principal

• *The authors are with the Image Computing Group, Department of Computer Science, City University of Hong Kong, Tat Chee Ave., Kowloon, Hong Kong. E-mail: {csshen, cship}@cityu.edu.hk.*

Manuscript received 13 June 1996; revised 27 Mar. 1997. Recommended for acceptance by V.S. Nalwa.
For information on obtaining reprints of this article, please send e-mail to: transpami@computer.org, and reference IEEECS Log Number 104800.

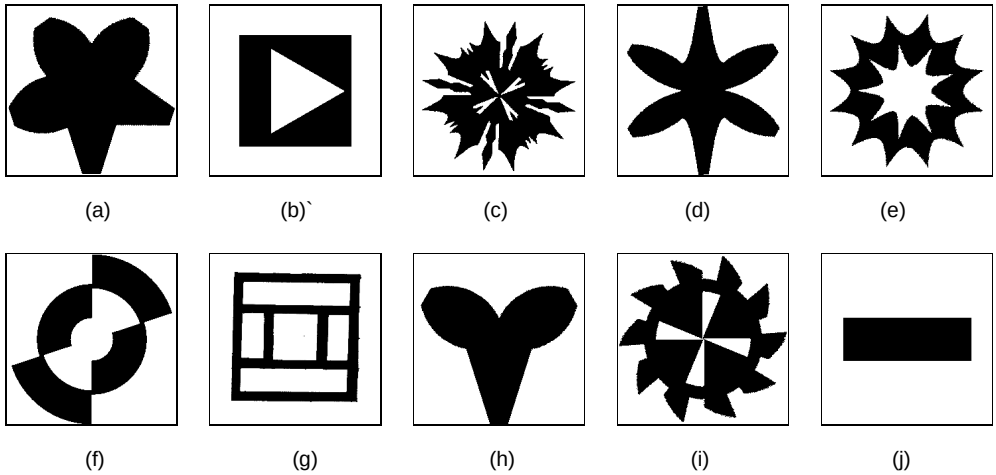


Fig. 1. All image shapes used in this paper.

axes of pattern. Its advantage is that no preprocessing is required to judge whether the input pattern is RSI or not, while its disadvantage is that it usually uses more than one axes to represent the pattern's orientation. In fact, the orientation of almost every pattern can be uniquely defined.

The main contribution of this paper is the development of a generalized image normalization method which is applicable to every kind of patterns. Additionally, the determination of whether the input pattern is an RSI and the detection of its fold number are integrated into an algorithm which simultaneously defines a unique orientation for the input pattern. Also, we show that the Hamming distance between any two normalized images can be increased by using the magnitude of the first non-zero GC moment to specify the size and orientation of the normalized image.

2 A Compact Algorithm Based on Regular Moments

A prerequisite for making an image compact is to know the affine relationship between the normalized image and the affine-transformed image. The affine relationship using corresponding points is described in Definition 1, and a new form of affine relationship using regular moments is given in Theorem 2.1. In the end of this section, an algorithm (Theorem 2.2) is proposed to obtain the affine coefficients from the regular moments of the affine-transformed image.

DEFINITION 1. *Let $f(u,v)$ be the affine-transformed image, and $f(x,y)$ the standard normalized image. The affine relationship between $f(u,v)$ and $f(x,y)$ are as follows.*

$$f(u,v) = f(x,y),\ \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = A \begin{bmatrix} x \\ y \\ 1 \end{bmatrix},\ A = \begin{bmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ 0 & 0 & 1 \end{bmatrix}$$

where $[u \quad v]^T$ is the affine-transformed position corresponding to point $[x \quad y]^T$, and A is the affine coefficient matrix.

Although the above definition gives the relationship between two corresponding points by affine matrix, it is usually impossible to obtain the affine coefficients since finding the corresponding points in two real images is not trivial and is also sensitive to noise. Hence, in Theorem 2.1 we present a new form of affine relationship, which is not based on points but on regular moments.

THEOREM 2.1. *The affine relationship between regular moments of the normalized image $f(x,y)$ and regular moments of the affine-transformed image $f(u,v)$ are as follows.*

$$\begin{bmatrix} a_{11} & a_{12} & b_1 \\ a_{21} & a_{22} & b_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} m_{20} & m_{11} & m_{10} \\ m_{11} & m_{02} & m_{01} \\ m_{10} & m_{01} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{21} & 0 \\ a_{12} & a_{22} & 0 \\ b_1 & b_2 & 1 \end{bmatrix} = \begin{bmatrix} m'_{20} & m'_{11} & m'_{10} \\ m'_{11} & m'_{02} & m'_{01} \\ m'_{10} & m'_{01} & 1 \end{bmatrix}$$

where m_{pq} is the pq-order regular moment of the normalized image $f(x,y)$, and m'_{pq} the pq-order regular moment of the affine-transformed image $f(u,v)$. Their definitions are

$$m_{pq} = E\left(x^p y^q\right) = \sum_x \sum_y x^p y^q f(x,y) \,/\, \left(\sum_x \sum_y f(x,y)\right)$$

$$m'_{pq} = E\left(u^p v^q\right) = \sum_u \sum_v u^p v^q f(u,v) \,/\, \left(\sum_u \sum_v f(u,v)\right)$$

For the purpose of image normalization, we need to know the mapping from point $[u \quad v]^T$ to point $[x \quad y]^T$. Theorem 2.2 gives an algorithm for constructing the compact image $f(x', y')$ from the affine-transformed image $f(u,v)$. Using this theorem, patterns can be normalized under translation, scaling and skew, leaving us to concentrate on detecting the orientation of the compact image in Section 3. In fact, this is the key problem in the image normalization.

THEOREM 2.2. *The relation between the compact image* $f(x', y')$ *and the affine-transformed image* $f(u, v)$ *is*

$$f(x', y') = f(u, v), \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} d / \sqrt{\lambda_1} & \\ & d / \sqrt{\lambda_2} \end{bmatrix} E^T \begin{bmatrix} u - m'_{10} \\ v - m'_{01} \end{bmatrix}$$

where λ_1 *and* λ_2 *are the two eigenvalues of the matrix*

$$U' = \begin{bmatrix} m'_{20} - m'^2_{10} & m'_{11} - m'_{10}m'_{01} \\ m'_{11} - m'_{10}m'_{01} & m'_{02} - m'^2_{01} \end{bmatrix}$$

and $E = \begin{bmatrix} e_1 & e_2 \end{bmatrix}$, *where* e_1 *and* e_2 *are the corresponding eigenvectors of* U' *associated with* λ_1 *and* λ_2.

3 Unique Orientation Determination for a Compact Image

This section will show that the lower-order non-zero GC moments can be used to determine a unique orientation for a pattern. We do this by first giving the necessary and sufficient condition for making GC_{pq} non-zero, then showing that for $GC_{pq} \neq 0$, there exists exactly q half-lines to express the pattern's orientation (Theorems 3.1 and 3.2). Finally, we prove that a unique orientation of pattern can be defined by the fold number detection (Theorem 3.4) and a third appropriate non-zero GC moment (Theorem 3.5). The technique is applicable to any type of centralized images.

3.1 Basic Property of Image Function

Suppose that $f(r, \theta)$ is the corresponding function of the centered compact image $f(x', y')$ in the polar coordinate. In general, any image $f(r, \theta)$ can be expanded by Fourier series.

$$f(r, \theta) = \sum_{m=-\infty}^{\infty} f_m(r) e^{-jm\theta}$$

where $f_m(r) = \frac{1}{2\pi} \int_0^{2\pi} f(r, \theta) e^{jm\theta} d\theta$. If the given image $f(r, \theta)$ is a K-RSI, its Fourier expansion is

$$f(r, \theta) = \sum_{l=-\infty}^{\infty} f_{K\times l}(r) e^{-j(K\times l)\theta}$$

where $f_{K\times l}(r) = \frac{K}{2\pi} \int_0^{\frac{2\pi}{K}} f(r, \theta) e^{j(K\times l)\theta} d\theta$.

3.2 First Two Non-Zero GC Moments Used for Determining Unique Orientation of Almost Every Kind of Images

The pq-order GC moment of image $f(r, \theta)$ is defined as follows:

$$GC_{pq} = R_{pq} e^{j\varphi_{pq}} = \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} f(r, \theta) \left(r^p e^{jq\theta} \right) r dr d\theta$$

where p is a non-negative integer, and q a positive integer. Since the lower-order non-zero GC moments will be used to detect the orientation of the given image, it is helpful to know the condition for making GC moments non-zero. Using the basic property of image function, we can prove that the necessary and sufficient condition for making GC_{pq} non-zero is $\int_0^{\infty} f_q(r) r^p r dr \neq 0$. For K-RSI, $f_q(r)$ must be zero when q is not a multiple of K. Based on this property, we get the following three theorems for determining unique orientation of almost every kind of images.

THEOREM 3.1. *If* $GC_{pq} \neq 0$, *then there exist* q *half lines starting from the origin* O *and having directional angles* $\theta_i = \frac{\varphi_{pq} + (i-1) \times 2\pi}{q}$ *with* $i = 1, 2, \ldots, q$, *which are invariant to the translation, scaling and rotation of the identical image.*

THEOREM 3.2. *Let* GC_{Pq_1} *and* GC_{Pq_2} *be the first and second non-zero GC moments encountered in the sequence* GC_{Pq} *with* $q = 1, 2, \ldots$, *respectively. The phase of the combined moment* $\left(GC_{Pq_1}\right)^x \left(GC_{Pq_2}\right)^y$ *is* $x\varphi_{Pq_1} + y\varphi_{Pq_2}$. *Then there exist* $xq_1 + yq_2$ *half lines starting from the origin* O *and having directional angles*

$$\theta_i = \frac{\left(x\varphi_{Pq_1} + y\varphi_{Pq_2}\right) + (i-1) \times 2\pi}{xq_1 + yq_2}$$

with

$$i = 1, 2, \ldots, \left(xq_1 + yq_2\right)$$

which are invariant to the translation, rotation and scaling of the identical image.

THEOREM 3.3. *There exist appropriate* x *and* y *making* $1 \leq \left(xq_1 + yq_2\right) \leq q_1$. *The appropriate values* x *and* y *can be obtained by solving the linear programming problem and selecting the smaller* $|x|$ *and* $|y|$ *in the solution set. The objective function is* $\min_{x,y} (xq_1 + yq_2)$ *and the constraint condition is* $\left(xq_1 + yq_2\right) \geq 1$.

Theorem 3.1, first presented in [16], tell us that for certain P, if the first non-zero GC moment encountered in the sequence GC_{Pq} with $q = 1, 2, \ldots$, is GC_{P1}, then the image orientation is unique. Otherwise, more than one half lines are needed to express the orientations of the image. To reduce the number of orientations, Theorems 3.2 and 3.3 are presented. Theorem 3.2 indicates the possibility that, if number $\left(xq_1 + yq_2\right)$ is smaller than q_1, then the number of orientations will be reduced to $\left(xq_1 + yq_2\right)$. The best choice is $\left(xq_1 + yq_2\right) = 1$, which means only a single half line is needed to represent the image's orientation. Theorem 3.3 gives the method of obtaining appropriate x and y. The result is that, if the common factor of q_1 and q_2 is one, then the resulting $\left(xq_1 + yq_2\right)$ is one. Otherwise, the resulting $\left(xq_1 + yq_2\right)$ is equal to the biggest common factor of q_1 and q_2. Accordingly, we obtained Table 1 which gives the ap-

TABLE 1
APPROPRIATE x AND y MAKING POSITIVE INTEGER $(xq_1 + yq_2)$ NEAR TO ONE.
VALUES IN EACH CELL ARE (x, y) AND $(xq_1 + yq_2)$.

q1 \ q2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
2	(-1, 1) 1	(1, 0) 2	(-2, 1) 1	(1, 0) 2	(-3, 1) 1	(1, 0) 2	(-4, 1) 1	(1, 0) 2	(-5, 1) 1	(1, 0) 2	(-6, 1) 1	(1, 0) 2	(-7, 1) 1	(1, 0) 2
3		(-1, 1) 1	(2,-1) 1	(1, 0) 3	(-2, 1) 1	(3,-1) 1	(1, 0) 3	(-3, 1) 1	(4,-1) 1	(1, 0) 3	(-4, 1) 1	(5,-1) 1	(1, 0) 3	(-5, 1) 1
4			(-1, 1) 1	(-1, 1) 2	(2,-1) 1	(1, 0) 4	(-2, 1) 1	(-2, 1) 2	(3,-1) 1	(1, 0) 4	(-3, 1) 1	(-3, 1) 2	(4,-1) 1	(1, 0) 4
5				(-1, 1) 1	(3,-2) 1	(-3, 2) 1	(2,-1) 1	(1, 0) 5	(-2, 1) 1	(5,-2) 1	(-5, 2) 1	(3,-1) 1	(1, 0) 5	(-3, 1) 1
6					(-1, 1) 1	(-1, 1) 2	(-1, 1) 3	(2,-1) 2	(2,-1) 1	(1, 0) 6	(-2, 1) 1	(-2, 1) 2	(-2, 1) 3	(3,-1) 2
7						(-1, 1) 1	(4,-3) 1	(3,-2) 1	(-3, 2) 1	(-5, 3) 1	(2,-1) 1	(1, 0) 7	(-2, 1) 1	(7,-3) 1
8							(-1, 1) 1	(-1, 1) 2	(-4, 3) 1	(-1, 1) 4	(5,-3) 1	(2,-1) 2	(2,-1) 1	(1, 0) 8
9								(-1, 1) 1	(5,-4) 1	(-1, 1) 3	(3,-2) 1	(-3, 2) 1	(2,-1) 3	(-7, 4) 1
10									(-1, 1) 1	(-1, 1) 2	(4,-3) 1	(3,-2) 2	(-1, 1) 5	(-3, 2) 2
11										(-1, 1) 1	(6,-5) 1	(-5, 4) 1	(-4, 3) 1	(3,-2) 1
12											(-1, 1) 1	(-1, 1) 2	(-1, 1) 3	(-1, 1) 4

propriate values of x and y for various q_1 and q_2. In Table 1, numbers in the first column represent possible q_1 and numbers in the first row represent possible q_2. Numbers in the other cells correspond to the solutions, (x, y) and $(xq_1 + yq_2)$, for various of q_1 and q_2.

In this way, we can conclude:

1) For almost every kind of non-RSIs, unique orientation can be detected.
   Fig. 2 is the typical example of non-RSIs. Since the first and second non-zero GC moments of Fig. 2b are $GC_{0,3}$ and $GC_{0,4}$, its orientation can be uniquely defined by the directional angle $(-\varphi_{0,3} + \varphi_{0,4})$.
2) For almost every kind of RSIs, the detected number of half lines is equal to its fold number.
   Using our method, even the fold number of the strange pattern given in [18], now shown in Fig. 3, can be directly detected. The first and second non-zero GC moments of this strange shape are $GC_{0,8}$ and $GC_{0,12}$. Therefore, the number of the half lines is $12 - 8 = 4$.

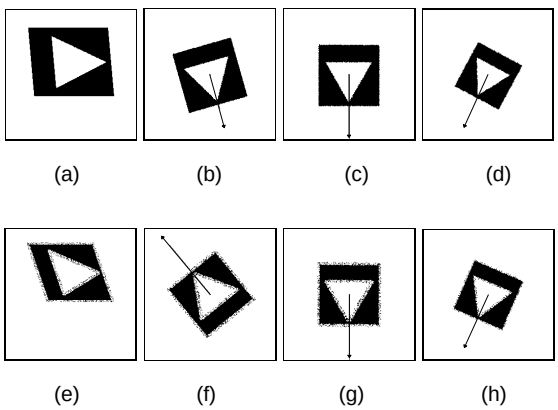


Fig. 2. Non-RSI. Its first and second non-zero GC moments are $GC_{0,3}$ and $GC_{0,4}$, and its number of half lines is one.

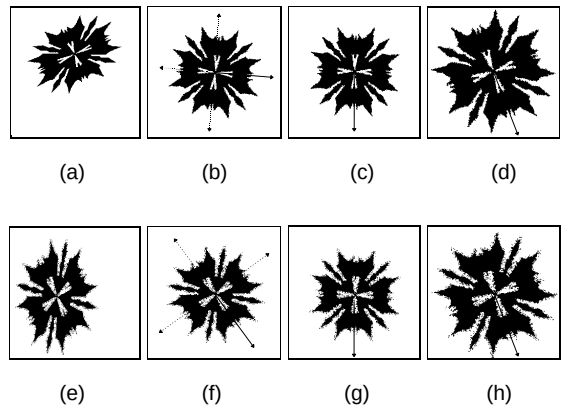


Fig. 3. 4-RSI. Its first and second non-zero GC moments are $GC_{0,8}$ and $GC_{0,12}$, and its number of half lines is 4.

3.3 Determining the Unique Orientation of the Degenerate Images by Fold Number Detection and a Third Appropriate Non-Zero GC Moment

Based on the above analysis, it appears that, if the number of the detected half lines for an image is bigger than one, then the image is an RSI and its fold number is equal to the number of the detected half lines. However, this may not be the case for some degenerate images, such as those shown in Figs. 5 and 6. Although Fig. 5 is a non-RSI, its number of the detected half lines is three. Fig. 6 is a 2-RSI, but its number of the detected half lines is four, i.e., a multiple of two. For such cases, if we still wish to uniquely express the image orientation, we have to detect the fold number of the input image. Theorem 3.4 gives the method of detecting the fold number of image, which is the necessary condition for determining the unique orientation of the degenerate image by first two non-zero GC moments plus a third appropriate non-zero GC moment in Theorem 3.5.

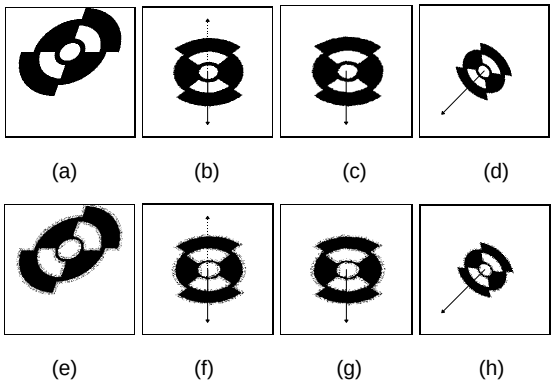


Fig. 4. 2-RSI. Its first and second non-zero GC moments are $GC_{2,4}$ and $GC_{2,6}$, and its number of half lines is 2.

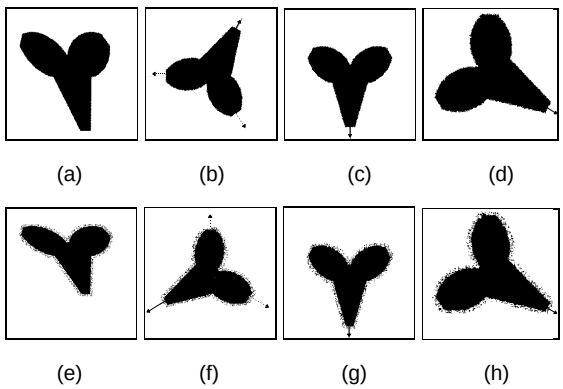


Fig. 5. Non-RSI. Its first and second non-zero GC moments are $GC_{0,3}$ and $GC_{0,6}$, and its number of half lines is three. However, it is a non-RSI.

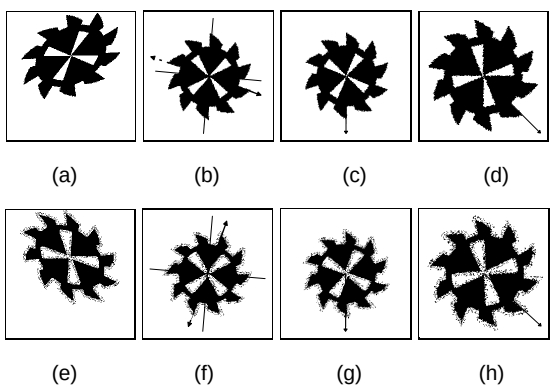


Fig. 6. 2-RSI. Its first, second, and third non-zero GC moments are $GC_{0,4}$, $GC_{0,8}$, and $GC_{0,10}$, respectively. Four fine half lines (in Fig. 6b or Fig. 6f) are computed by $GC_{0,4}$ and $GC_{0,8}$. Two coarse arrows are calculated by $GC_{0,4}$, $GC_{0,8}$, and $GC_{0,10}$, and the solid one is the final direction.

Let the number of the detected half lines be $N = x_1 q_1 + y_1 q_2$, and the original phase $\phi = x_1 \varphi_{Pq_1} + y_1 \varphi_{Pq_2}$. The ith fold of an image is defined in the area

$$\theta_i - \frac{\pi}{N} \le \theta < \theta_i + \frac{\pi}{N},$$

where

$$\theta_i = \frac{\phi + (i-1) \times 2\pi}{N}, \quad i = 1, 2, \ldots, N$$

Let $FC_{N \times N}$ be a matrix, whose element FC_{ij} describes the fold difference between the fold images of fold i and fold j, and is defined as

$$FC_{ij} = \frac{\int_{\theta_i - \frac{\pi}{N}}^{\theta_i + \frac{\pi}{N}} \int_0^{\infty} \left| f(r,\theta) - f\left(r, \theta + (j-i) \times \frac{2\pi}{N}\right) \right| r dr d\theta}{\frac{2\pi}{N} \times R}$$

Here, R is the maximum radius.

From matrix $FC_{N \times N}$, we can find out which folds are similar. The similar folds can then be grouped together. Let $group_i$ be the ith group with GM_i members, and GN be the total number of groups. From these, there are then two possibilities related to this image.

1) If the biggest common factor of GM_1, GM_2, ..., and GM_{GN} is one, then this image is non-RSI.
   In this case, if there exists one and only one group and the number of member in the group is one, then the orientation of this non-RSI can be determined by a unique half line. Otherwise, a third appropriate non-zero GC moment has to be used in order to describe this image by a unique half line. For example, the image shown in Fig. 5 consists of three half lines, which indicates three folds. Since two of the folds are identical, the half line for the third fold can then be regarded as the normalized orientation of this image.
2) If the biggest common factor of GM_1, GM_2, ..., and GM_{GN} is bigger than one, then this image may be an RSI.

THEOREM 3.4. *Let L be the biggest common factor of GM_1, GM_2, ..., and GM_{GN}. Clearly, L is the factor of N. Assume that $\{b_i | 1 \le i \le I\}$ be the set of all positive factors of L with $L = b_1 > b_2 > \ldots > b_I \ge 1$. Then all possible fold numbers of this image are included in the set $\{b_i | 1 \le i \le I\}$. Suppose b_i be the actual fold number of this image, and $FR_{b_i \times b_i}$ its corresponding fold-difference matrix. Matrix $FR_{b_i \times b_i}$ can be obtained from matrix $FC_{N \times N}$.*

$$FR_{mn} = \frac{b_i}{N} \sum_{l=0}^{N/b_i - 1} FC_{\left(m \times \frac{N}{b_i} + l\right)\left(n \times \frac{N}{b_i} + l\right)}$$

The order for detecting the fold number of the image is from b_1 to b_I, until the actual fold number of the image is detected.

Although the fold number of any type of images can be detected using the above algorithm, for the image in Fig. 6, however, since its actual fold number is two which is a factor of $N = 4$, it is thus impossible to uniquely express the orientation of this image. Even for such images, a

unique half line can still be found by using a third appropriate non-zero GC moment (Theorem 3.5).

THEOREM 3.5. *The number of the detected half lines is* $N = x_1q_1 + y_1q_2$, *where* x_1 *and* x_2 *are obtained for* q_1 *and* q_2 *based on Theorem 3.3.* RN, *the factor of* N, *is the actual fold number of the image. Then there must exist a third non-zero GC moment* GC_{Pq_3} *which makes the common factor of* N *and* q_3 *equal to* RN, *that is*

$$RN = x_2N + y_2q_3 = x_2\left(x_1q_1 + y_1q_2\right) + y_2q_3$$

Therefore, the orientations of the image can be expressed by RN *half lines, with directional angles*

$$\theta_i = \frac{x_2\left(x_1\varphi_{Pq_1} + y_1\varphi_{Pq_2}\right) + y_2\varphi_{Pq_3} + (i-1)\times 2\pi}{RN}, \quad 1 \le i \le RN$$

Any one of these half lines can be regarded as the unique orientation of the image because this image isRN*-RSI.*

3.4 Using Alternating Energy to Control Orientation Detection

Since the order p in GC_{pq} is always fixed when GC_{pq} is applied in the image orientation detection, it is necessary to propose a selection rule for determining the appropriate p. We do this by using the alternating energy of the Fourier spectrum of the image.

Suppose that $h_p(\theta)$ is a 1D function obtained from $f(r,\theta)$, $h_p(\theta) = \int_0^\infty f(r,\theta)r^p rdr$. Although the 2D image has very strong periodicity, however, periodicity may become weak in some $h_p(\theta)$ pictures. We thus have to suggest a rule so as to select a $h_p(\theta)$ with strong periodicity. The Fourier transform of $h_p(\theta)$ is

$$H_p(q) = \frac{1}{2\pi}\int_0^{2\pi} h_p(\theta)e^{jq\theta}d\theta$$

which leads to $\int_0^{2\pi}\left[h_p(\theta)\right]^2 d\theta = 2\pi\sum_{q=-\infty}^{\infty}\left\|H_p(q)\right\|^2$. Notice that $GC_{pq} = H_p(q)$, which implies that GC_{pq} can be directly obtained from the Fourier transform of the 1D function $h_p(\theta)$ when the order p has been fixed. Then the ratio

$$a_p = 1 - \frac{2\pi\left\|H_p(0)\right\|^2}{\int_0^{2\pi}\left[h_p(\theta)\right]^2 d\theta}$$

can be used to determine whether the 1D function $h_p(\theta)$ has a strong periodicity. Large value, a_p, means strong alternating energy in the function $h_p(\theta)$. In our study, we select 1D function whose corresponding a_p is over 5 percent. Fig. 8 shows the plots of $h_i(\theta) / H_i(0)$ with $i = 0, \ldots, 5$, all of which are the radial projections of Fig. 4b. Among these plots, Fig. 8c is chosen, that is $p = P = 2$. As for circular disks and rings, all of their a_p are equal or near to zero. The value of a_p also serves to determine whether the image is a circular disk or a ring, and hence no orientation is needed. It should be noted that for almost all existing algorithms including [19], they are not able to do so.

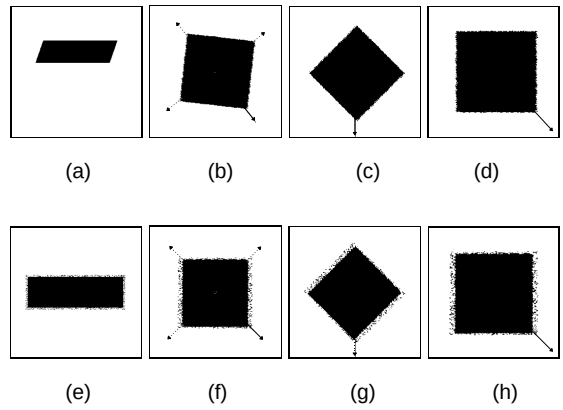


Fig. 7. 4-RSI. Only one non-zero GC moment of Fig. 7b can be distinctly detected.

Besides applying function energy to determine the order p, function energy is also suggested here to determine whether the remaining non-zero GC moments can be detected, or not. Let

$$b_q = 1 - \frac{2\pi\left(\left\|H_P(0)\right\|^2 + 2\sum_{m=1}^{q}\left\|H_P(m)\right\|^2\right)}{\int_0^{2\pi}\left[h_P(\theta)\right]^2 d\theta}$$

If b_q is small, it means that there is little alternating energy left, and no more non-zero GC moments can be detected. In our study, the threshold for b_q has been set to one percent. This rule is useful for special images, such as Fig. 7, where only the first non-zero GC moment can be distinctly detected.

4 Improving Hamming Distance-Based Matching of the Normalized Images

While the goal of image normalization is to reduce different views of the same image objects to its canonical size and principal orientation, it also causes different image objects when normalized to have similar size and principal orientation. Consequently the Hamming distance between different normalized image objects becomes small, which in turn affects the efficiency and the sensitivity of Hamming distance based pattern matching.

To improve the sensitivity of Hamming distance based matching of normalized images, we need find a way to make different normalized objects to have different sizes and different principal orientations. Previous normalization methods make use of the phase of moment only. However, the magnitude of the moment also contains much information about pattern. In the following, the magnitude of moment is proposed to specify the size and the orientation of the nor-

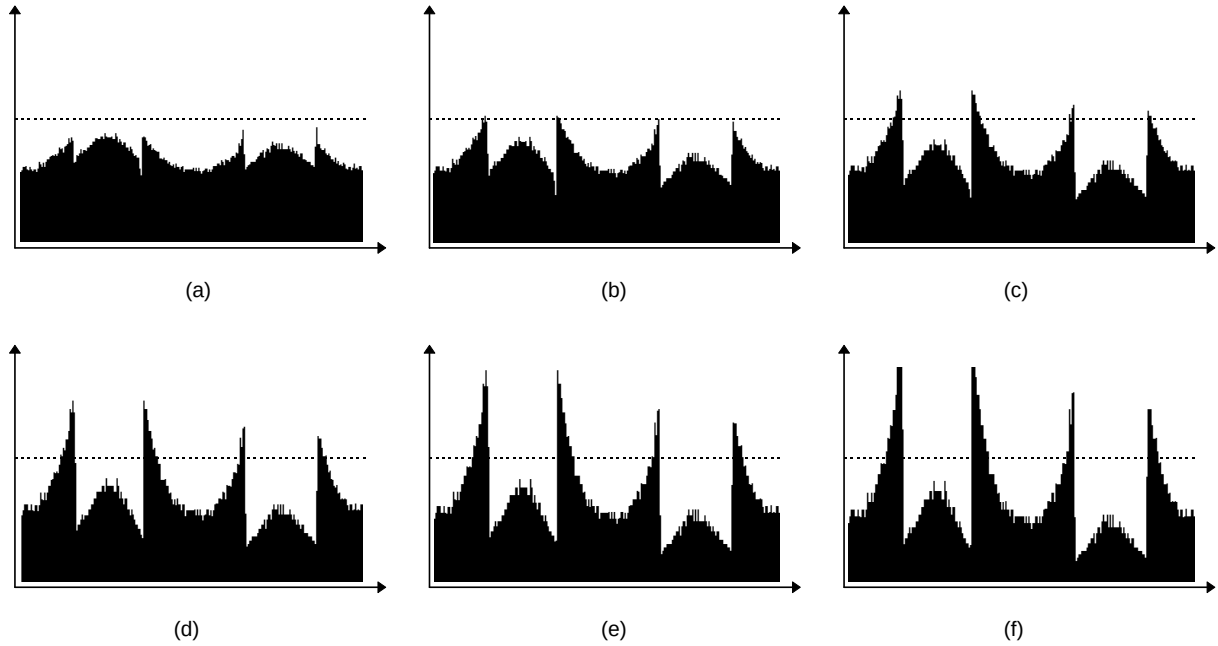


Fig. 8. Radial projections of Fig. 4b. Figs. 8a to 8f represent functions $h_0(\theta) \,/\, H_0(0)$ to $h_5(\theta) \,/\, H_5(0)$, respectively.

malized image. The magnitude-based scaling and specification of orientation are defined respectively as below.

$$s = g\left(\frac{R_{Pq_1}}{0.7}\right) \text{ here } g(x) = \begin{cases} 0.7, & x < 0.7 \\ x, & other \\ 1.3, & x > 1.3 \end{cases}$$

$$\psi = \frac{\pi}{RN} \times \frac{s - 1.0}{0.3}$$

where R_{Pq_1} is the magnitude of the first non-zero GC moment. RN is the actual fold number of the image. Notice for non-RSI, $RN = 1$. According to the definition of s, the possible distribution of s ranges from 0.7 to 1.3. The possible range of ψ is from $-\frac{\pi}{RN}$ to $+\frac{\pi}{RN}$.

5 Generalized Image Normalization Algorithm and Experiments

5.1 A Generalized Image Normalization Algorithm

The complete algorithm for generalized image normalization can be summarized as follows:

1) Calculate the moments of the affine-transformed image $f(u, v)$, m'_{pq} with $1 \le p + q \le 2$. Then obtain the covariance matrix U' defined in Theorem 2.2.
2) Compute the two eigenvalues and the two eigenvectors of U'.
3) Obtain the compact image $f(x', y')$ according to Theorem 2.2.
4) Select an appropriate P to make 1D function $h_P(\theta)$ with strong alternating energy. If all of their a_p are equal or near to zero, then this object must be a circular disk or ring.
5) Detecting the first non-zero GC moment GC_{Pq_1}.
   a) If $q_1 = 1$, a unique orientation is obtained. Its angle is φ_{Pq_1}. Go to step 10.
   b) If $b_{q_1} < 1$ percent, go to step 8.
6) Detecting the second non-zero GC moment GC_{Pq_2} as long as $b_{q_1} \ge 1$ percent.
7) Get x_1, y_1 and $N = x_1 q_1 + y_1 q_2$ from Table 1 for the values of q_1 and q_2.
   a) If $N = 1$, then this compact image is a non-RSI. Go to step 10.
   b) If $N > 1$, go to next step.
8) Determine the actual fold number RN of the compact image, here $1 \le RN \le N$.
   a) If $RN = N$, then this compact image is RSI. Any half line, given in Theorem 3.2, can be regarded as the unique orientation of this image. Go to step 10.
   b) Otherwise, go to next step.
9) Detect a third appropriate non-zero GC moment GC_{Pq_3}, and compute the unique orientation according to Theorem 3.5.
10) Rotate the compact image, making its orientation match the x axis.
11) Scale and rotate the normalized image based on the method described in Section 4.

Obviously, step 10 and step 11 can be merged for lessening computation time.

5.2 Experiments

We design a series of examples (four cases) to demonstrate clearly that the generalized image normalization can be successfully achieved by our method. Notice that the method in [4] will fail in all of our examples.

5.2.1 Image Normalization

In Figs. 2 to 7, Subfig. a is the affine-transformed images; Subfig. b is, respectively, the compact images of Subfig. a;

Subfig. c is the normalized images; and Subfig. d is the post-normalized images. The corrupted images in Subfig. e, f, g, and h are separately similar to images in Subfig. a, b, c, d. Notice that Subfig. c and Subfig. g are almost identical, and Subfig. d and Subfig. h are almost identical.

CASE 1: Non-RSI Under Random Transformation (Fig. 2)

Fig. 2a shows a non-RSI under random affine-transformation, and its compact image is shown in Fig. 2b. Since the radial projection function $h_0(\theta)$ of Fig. 2b meets the requirement that $a_0 > 5$ percent, then $p = 0$. The first and second non-zero moments of the compact image are separately $GC_{0,3}$ and $GC_{0,4}$. Hence, its orientation can be uniquely defined by a half line with directional angle $\left(-\varphi_{0,3} + \varphi_{0,4}\right)$. That is the solid arrow in Fig. 2b. Rotating the orientation of the compact image to x-axis leads to Fig. 2c. The magnitude of $GC_{0,3}$ is smaller than 0.7, which leads to $s < 1$. Thus, the size of the normalized image is lessened, and its orientation is clockwise rotated by an angle.

If the method in [6] is used to normalize this compact image, this compact image will be regarded as a three-RSI, which is incorrect. If algorithm in [19] is used to define the orientations of Fig. 2b, it will need three universal principal axes, which is obviously redundant.

Case 2: N-RSI Under Random Affine Transformation (Figs. 3 and 4)

The original image of Fig. 3 was obtained from the Fig. 3 shown in [18]. Fig. 3a is a randomly affine-transformed four-RSI. The first non-zero GC moment of Fig. 3b is $GC_{0,8}$.

Note that the methods in [5] and [6] will regard the compact image as eight-RSI, whilst the method in [15] will give a zero for FISSP, and the method in [19] will require eight universal principal axes. All of these results are unreasonable. Although the method in [18] is able to detect the fold number of Fig. 3b correctly, but it is unable to define the orientation for Fig. 3b. The second non-zero GC moment of figure is $GC_{0,12}$, then the number of half lines is four. Using our algorithm for detecting the fold number of the compact image, the fold number of Fig. 3b is actually four. Hence, the half line whose directional angle is counterclockwise the smallest is selected as the orientation of Fig. 3b. That is the solid arrow shown in Fig. 3b. The magnitude of $GC_{0,8}$ is bigger than 0.7, thus Fig. 3c is enlarged and rotated counterclockwise, as shown in Fig. 3d.

Another example shown in Fig. 4 is a two-RSI. This is an excellent example to illustrate the necessity of selecting an appropriate parameter p for GC_{pq}. The radial projections of Fig. 4b is shown in Fig. 8, and Fig. 8c, which shows $h_2(\theta) / H_2(0)$, meets the requirement of strong periodicity. Thus $p = 2$.

Case 3: Degenerate Non-RSI (Fig. 5)

Since the first and second non-zero GC moments of Fig. 5b are $GC_{0,3}$ and $GC_{0,6}$, respectively, three half lines are needed to express the orientation of Fig. 5b, which are the two dotted arrows and the one solid arrow. However, by using our fold number determining algorithm, it was found that the fold number of Fig. 5b is one. That is, Fig. 5b is a non-RSI. Based the fold-difference matrix $FC_{3\times3}$, the two folds associated with two dotted arrows are in fact identical, while the fold associated with the solid arrow is unique. Thus the solid arrow is chosen as the orientation of Fig. 5b. Fig. 5d is the enlarged and counterclockwise rotated form of Fig. 5c based on the magnitude of $GC_{0,3}$.

Case 4: Degenerate N-RSI (Fig. 6)

The compact image is shown in Fig. 6b, where the internal fold number is four and the external fold number is six. The number of the directional half lines for Fig. 6b is four, multiple of two, since the first and second non-zero GC moments of Fig. 6b are $GC_{0,4}$ and $GC_{0,8}$, respectively. The four lines in Fig. 6b define these four directions. The result of our fold number determining algorithm indicates that this compact image is in fact a two-RSI. In this case, Theorem 3.5 is used and the orientation of this compact image is defined by $GC_{0,4}$, $GC_{0,8}$, and $GC_{0,10}$, where $GC_{0,10}$ is the third appropriate non-zero GC moment. The two coarse arrows shown in Fig. 6b, with directional angle

$$\frac{\left(-2\times\left(-\varphi_{0,4}+\varphi_{0,8}\right)+\varphi_{0,10}\right)+(i-1)\times 2\pi}{2},\quad i=1,\,2$$

define the orientation of this compact image. Finally, the solid arrow is chosen as the orientation of Fig. 6b.

5.3 Increasing Hamming Distance

To demonstrate that the Hamming distance between two normalized images can be increased by adjusting the scale and the orientation of the normalized images based on the magnitude of the first non-zero GC moment of the compact image using equations in Section 4, Table 2 shows the increase, in percentages, of the Hamming distance between every pair of the image objects used in our experiments. The numbers in the first row and first column of the table represent 10 objects in Fig. 1. It can be seen from Table 2, that the proposed algorithm for most cases increases the Hamming distance between the objects. There are however four cases when the post-normalized images constructed by this method do not differ significantly from their normalized images.

5.4 Robustness of the Algorithm Against Noise

In our experiments, Gaussian noise is added to all of our

TABLE 2
INCREASED PERCENT FOR HAMMING DISTANCE BETWEEN EVERY COUPLE OF OBJECTS AFTER THE NORMALIZED IMAGES ARE RESCALED AND REROTATED

Objects	2	3	4	5	6	7	8	9	10
1	31	2	37	38	59	17	**-8**	1	1
2		48	5	1	8	12	9	63	37
3			15	53	88	25	3	**-6**	**-9**
4				**-3**	12	8	2	13	13
5					15	0	14	59	55
6						15	27	89	82
7							15	24	17
8								2	0
9									2

TABLE 3
DEVIATION VALUES OF THENORMALIZED ORIENTATION WITH INCREASING NOISE LEVEL

Figs / noise level	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	(mean, std. dev.)
fig 1(a)	0.01	-8.5	-0.69	-1.80	-5.27	-4.22	**-8.5**	-5.06	0	-8.06	-3.72	-0.87	-3.05	-3.72	-1.08	-2.58	-3.05	-8.19	**(-3.8, 2.86)**
fig 1(b)	0.23	-0.9	-0.28	-1.22	1.28	1.4	-2.25	-2.84	-0.78	-1.01	-0.93	-3.33	-0.12	1.55	-0.3	-0.81	0.3	-2.51	(-0.69, 1.36)
fig 1(c)	0	-1.20	-1.34	-2.50	-3.25	-4.14	-2.93	2.62	0.96	2.82	0.90	2.64	1.89	-0.46	0.70	2.63	1.33	2.26	(0.16, 2.2)
fig 1(d)	0	-0.51	-2.02	-3.53	-0.03	0.09	-1.02	-1.52	-1.52	-1.52	-2.01	-0.95	1.44	-4.43	-2.51	1.96	-2.52	3.3	(-0.96, 1.86)
fig 1(e)	0	0.15	-0.26	-0.51	-0.33	-0.54	0.07	-0.63	1.1	-1.22	-1.06	-1.5	-1.03	1.01	0.09	1.22	1.08	2.55	(0.01, 1.01)

examples to show that the proposed method is able to endure certain degree of noise. Gaussian noise is added around the edges of the objects because these areas are the most sensitive areas to be corrupted as a result of image segmentation. The method used for adding noise is that, random noise is first added onto every edge-point's $W \times W$ neighboring pixels, then the noisy image is converted into a binary image. The size of neighbor window W determines the noise level.

In our experiments, different versions of the same object are obtained by adding noise to the original object with different window sizes (W). The window size (W) changes from one to 18, which leads to 18 corrupted versions. Fig. 9a shows Fig. 1d and its three corrupted versions, corresponding to three window sizes, six, 12, and 18. The "eye"-like feature area shown with each of the input objects in Fig. 9a is the fixed reference point for orientation comparisons. Notice that this feature area will be transformed similarly during image normalization, but it will not affect image object's normalization. The position of the final transformed feature area can be used to represent the orientation of the normalized image. Suppose that the normalized orientation for the normalized image of the noise-free version is O_0, and the normalized orientations for the corrupted versions are O_i, $i = 1, 2, \ldots, 18$. Fig. 9b shows the normalized orientations for four normalized images. The values $(O_i - O_0)$, $i = 1, 2, \ldots, 18$, represent the deviations of the normalized orientations under normalization procedure with different noise levels. Table 3 shows these deviations for the first five objects in Fig. 1. The first row in this table represents the noise levels from one to 18, while the first column represents five objects and the last column gives the means and the standard deviations. Values in the other cells are the deviations of the normalized orientations compared with their corresponding noise-free versions. The unit for these values is *degree*. It can be seen that the deviations are not too significant under these noise levels. And, the normalized orientation of Fig. 1a is affected most, compared with other figures.

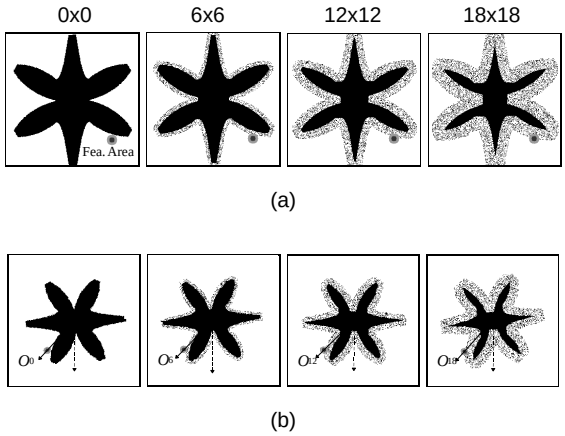


Fig. 9. Robustness of the proposed method against noise. (a) Fig. 1d and its three corrupted versions under different noise levels. (b) Corresponding normalized objects and their normalized orientations O_i.

6 CONCLUSIONS

In this paper, we have proposed an efficient and reliable new method for generalized 2D image normalization which provides the solutions for all the problems typically encountered in image normalization. The images can be any complex shapes irrespective of whether they are rotationally symmetric or not.

Defining image orientation is the key problem in image normalization. We have proved that, for almost all image, its orientation can be uniquely defined by the polar angles of the first two non-zero GC moments. For other degenerate objects, its orientation can be uniquely determined by the first two non-zero GC moments plus a third appropriate non-zero GC moment.

For the purpose of shape matching, the magnitude of the first non-zero GC moment is used to regulate the size and orientation of the normalized images, so that different image would be given different normalized size and orientation. Hence, increasing the Hamming distance computed for any pair of these images.

To demonstrate clearly the feasibility of the approach, in our experiments, a series of examples are given to show that the proposed method is applicable to a number of images which have been shown to fail other previous image normalized techniques. We also show that the approach is relatively insensitive to certain degree of Gaussian noise. It should be noted that whilst our approach is able to determine a unique normalized size and orientation for all of our test images, all other previously reported methods failed to do so at least for some of them. In summary, the proposed method provides a complete solution for image normalization.

REFERENCES

[1] J. Wood, "Invariant Pattern Recognition: A Review," *Pattern Recognition,* vol. 29, no. 1, pp. 1-17, 1996.

[2] Y.S. Abu-Mostafa and D. Psaltis, "Image Normalization by Complex Moments," *IEEE Trans. Pattern Analysis and Machine Intelligence* vol. 7, no. 1, pp. 46-55, Jan. 1985.

[3] J.G. Leu, "Shape Normalization Through Compacting," *Pattern Recognition Letters* vol. 10, pp. 243-250, 1989.

[4] S.C. Pei and C.N. Lin, "Image Normalization for Pattern Recognition," *Image and Vision Computing,* vol. 13, no. 10, pp. 711-723, 1995.

[5] S.C. Pei and C.N. Lin, "Normalization of Rotationally Symmetric Shapes for Pattern Recognition," *Pattern Recognition,* vol. 25, no. 9, pp. 913-920, 1992.

[6] S.C. Pei and L.G. Liou, "Automatic Symmetry Determination and Normalization for Rotationally Symmetric 2D Shapes and 3D Solid Objects," *Pattern Recognition,* vol. 27, no. 9, pp. 1,193-1,208, 1994.

[7] A. Rosenfeld and A.C. Kak, *Digital Picture Processing,* vol. 2, pp. 289-290. New York: Academic Press, 1982.

[8] A. Taza and C.Y. Suen, "Discrimination of Planar Shapes Using Shape Matrices," *IEEE Trans. Systems, Man and Cybernetics,* vol. 19, no. 5, pp. 1,281-1,289, 1989.

[9] S.A Friedberg, "Finding Axes of Skewed Symmetry," *Computer Vision Graphics Image Process,* vol. 34, pp. 138-155, 1986.

[10] M.J. Atallah, "On Symmetry Detection," *IEEE Trans. Computers,* vol. 34, pp. 663-666, 1985.

[11] G. Marola, "On the Detection of the Axes of Symmetry of the Symmetric and Almost Symmetric Planar Image," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 11, pp. 104-108, 1989.

[12] A. Mitiche and J.K. Aggarwal, "Contour Registration by Shape-Specific Points for Shape Matching," *Computer Vision Graphics Image Process,* vol. 22, pp. 396-408, 1983.

[13] W.H. Tsai and S.L. Chou, "Detection of Generalized Principal Axes in Rotationally Symmetric Shapes," *Pattern Recognition,* vol. 24, pp. 95-104, 1991.

[14] S.L. Chou, J.C. Lin, and W.H. Tsai, "Fold Principal Axis—A New Tool for Defining the Orientations of Rotationally Symmetric Shapes," *Pattern Recognition Letters,* vol. 12, pp. 109-115, 1991.

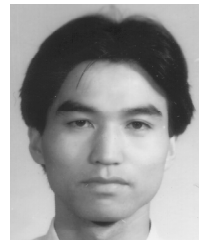
[15] J.C. Lin, S.L. Chou, and W.H. Tsai, "Detection of Rotationally Symmetric Shape Orientations by Fold-Invariant Shape-Specific Points," *Pattern Recognition,* vol. 25, pp. 473-482, 1992.

[16] D. Shen and F. Qi, "Detection of Rotationally Symmetric Shape Orientation by Fold Principal Orientation," *Proc. Int'l Symp. Information Theory and Its Applications,* Sydney, Australia, pp. 1,043-1,048, Nov. 1994.

[17] J.-J. Leou and W.-H. Tsai, "Automatic Rotational Symmetric Determination for Shape Analysis," *Pattern Recognition,* vol. 20, no. 6, pp. 571-582, 1987.

[18] J.C. Lin, W.H. Tsai, and J.A. Chen, "Detecting Number of Folds by a Simple Mathematical Property," *Pattern Recognition Letters,* vol. 15, pp. 1,081-1,088, 1994.

[19] J.C. Lin, "Universal Principal Axes: An Easy-To-Construct Tool Useful in Defining Shape Orientations for Almost Every Kind of Shape," *Pattern Recognition,* vol. 26, no. 4, pp. 485-493, 1993.



Dinggang Shen received his BS, MS, and PhD degrees in electronic engineering from Shanghai JiaoTong University in 1990, 1992, and 1995, respectively. He worked as a research assistant in the Department of Computer Science at the Honk Kong University of Science and Technology from December 1994 till June 1995. Since November 1995, Dr. Shen has been a lecturer of communication engineering at Shanghai Jiao-Tong University. Currently, he is a research fellow in the Department of Computer Science at City University of Hong Kong.

Dr. Shen's research interests are in the areas of computer vision, pattern recognition, image processing, neural networks, and image indexing and retrieval.

Horace H.S. Ip received his BSc (First Class Honours) in applied physics, and the PhD degree in image processing from University College, London, UK in 1980 and 1983, respectively. He was a research fellow of the research computing unit of the Imperial Cancer Research Fund Laboratories, London, England, from 1984 to 1987. He was with Cambridge Consultants Ltd. in the Cambridge Science Park, England, until he returned to Hong Kong in 1989.

There, Dr. Ip founded, and heads the Image Computing Group of the Department of Computer Science, City University of Hong Kong. He serves on the governing board of the IAPR, and co-chairs its technical committee (no. 12) on multimedia and communications. He is currently chairperson of the IEEE (HK) Section Computer Chapter.

Dr. Ip's research interests include image processing and analysis, pattern recognition, hypermedia computing systems, and computer graphics.

Dr. Ip is a member of the editorial board of the *Pattern Recognition* journal, the Chinese *Journal on CAD and Computer Graphics,* and a guest editor for the international journal of *Real-Time Imaging.* He has published more than 70 papers in international journals and conference proceedings.