

Image Normalization by Complex Moments

YASER S. ABU-MOSTAFA AND DEMETRI PSALTIS, MEMBER, IEEE

Abstract—The role of moments in image normalization and invariant pattern recognition is addressed. The classical idea of the principal axes is analyzed and extended to a more general definition. The relationship between moment-based normalization, moment invariants, and circular harmonics is established. Invariance properties of moments, as opposed to their recognition properties, are identified using a new class of normalization procedures. The application of moment-based normalization in pattern recognition is demonstrated by experiment.

Index Terms—Circular harmonics, complex moments, invariance, moment invariants, normalization, pattern recognition, principal axes, template matching.

I. INTRODUCTION

A RECURRING problem in pattern recognition is the question of between-class discrimination versus within-class invariance. Ideally, classification features should be sensitive to the identity of the class, but not to the variations within the class. In general, however, features that are invariant to within-class variations tend to be insensitive to between-class differences, and vice versa.

Moments have been proposed as a solution to this problem in two ways [1]:

- 1) by using moment invariants as features of the image that retain their values when the image is shifted, scaled, or rotated;
- 2) by using moments through the second order to normalize the image, i.e., transform it into a standard form, and then extracting suitable features to classify the normalized image.

We have recently introduced complex moments [2] to characterize the ability of moment invariants to discriminate between classes. We use the same formalism here to define a general class of normalization procedures using higher order moments. The establishment of normalization procedures based on higher order moments allows us to directly compare the invariance properties of moment invariants versus those of images that have been normalized through moments. In particular, we show that an image normalized in the manner we propose retains all the invariance properties of moment invariants in the presence of noise. The significance is that pattern recognition algorithms can be devised that operate on normalized images and, in principle, have all the within-class insensitivity of moment invariants but not their shortcomings with regard to distinguishing between different classes.

In Section II, we define the classical moment-based normalization in terms of complex moments. In Section III, we generalize the classical method and interpret the normalization

process geometrically. The invariance properties of moment invariants are the subject of Section IV. In Section V, we present experimental results in support of using moments for normalization, but not for classification.

II. MOMENT-BASED NORMALIZATION

In this section, we define image normalization and investigate the classical moment-based normalization procedure, extend it to admit scale and contrast normalization and redefine it in terms of complex moments as a basis for its generalization in the next section.

A. Introduction

Normalization is the process of transforming the image function $g_1(x, y)$ into the function $g_2(x, y)$ so that it retains all the relevant information of the original image and also satisfies a set of conditions which we call *the normalization criteria*. Therefore, $g_2(x, y)$ can be considered as a "standard" version of the original image $g_1(x, y)$.

Throughout this work, we shall consider all image functions to be nonnegative, vanishing outside a bounded domain and possessing positive integrals. A class of normalization processes of special interest is described by the following relations between g_1 and g_2 :

$$g_2(x_2, y_2) = Gg_1(x_1, y_1) + B$$

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} X \\ Y \end{bmatrix} + \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} \quad (2.1)$$

where G, B, a, b, c, d, X, Y are real constants. In most cases, conditions (2.1) are further restricted by the following relations:

$$G > 0, \quad B = 0$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} = (1/D) \begin{bmatrix} \cos \Phi & -\sin \Phi \\ \sin \Phi & \cos \Phi \end{bmatrix} \quad (2.2)$$

$$D > 0.$$

The conditions on G (contrast level) and B (bias) maintain the nonnegative nature of the image function. The condition on D (dimensional scale) excludes reflection of the image, and the form of the matrix given by (2.2) implies that the scaling is isotropic and that g_1 is rotated by a clockwise angle Φ (or the image plane axes rotated by a counterclockwise angle Φ). From this point on, we shall consider only this type of transformation for normalization.

The values of the parameters in (2.1), (2.2) are specified by requiring that g_2 satisfies a set of normalization criteria. By "moment-based normalization," we mean a normalization pro-

Manuscript received March 25, 1983; revised July 23, 1984. This work was supported in part by the IBM Corporation and in part by the Army Research Office. This paper was presented in part at the IEEE Conference on Computer Vision and Pattern Recognition, June 1983.

The authors are with the Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125.

cess for which the normalization criteria are conditions imposed on the *moments* of $g_2(x, y)$. We shall use complex moments because the normalization can be conveniently defined in terms of them. The fact that the kernels of ordinary moments and complex moments are both bases for the space of polynomials with complex coefficients guarantees the equivalence of this approach and the approach using ordinary moments.

It should be pointed out that the normalization need not be unique, i.e., the function $g_2(x, y)$ which satisfies the normalization criteria and is obtained from $g_1(x, y)$ by means of relations (2.1), (2.2) is not necessarily unique.

B. The Classical Normalization

In classical mechanics, a moment-based normalization procedure tackling the translation and rotation of two-dimensional objects has been introduced through the definition of *centroid* and *principal axes*. The same procedure has been applied to the image function $g(x, y)$. This procedure, which we denote by $N_{2,0}$, will be generalized in the following section to a class of normalization procedures denoted by $\{N_{p,q}\}$ which are based on complex moments of the image function $g(x, y)$. This approach will also admit normalization against scale and contrast changes. We start here by redefining the classical normalization in terms of complex moments.

The ordinary moment of order (p, q) is defined as

$$M_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q g(x, y) dx dy \quad (2.3)$$

whereas the complex moment of order (p, q) is defined by the following equation:

$$C_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x + iy)^p (x - iy)^q g(x, y) dx dy \quad (2.4)$$

where $i = \sqrt{-1}$. Complex moments can be expressed as a linear combination of ordinary moments and vice versa, and any condition involving ordinary moments can be expressed in terms of complex moments. For example, the classical definition of the centroid can be restated as the new origin of the image plane with respect to which the following condition holds [for $g_2(x, y)$]:

$$C_{10} = 0. \quad (2.5)$$

Since $C_{10} = M_{10} + iM_{01}$, the coordinates of the centroid (X, Y) can be expressed in terms of the ordinary moments of $g_1(x, y)$. They are given by

$$X = \frac{M_{10}}{M_{00}}, \quad Y = \frac{M_{01}}{M_{00}}. \quad (2.6)$$

Furthermore, after the image is centralized around its centroid, the orientation of the principal axes can be defined as the angle through which the image plane axes are rotated to make the condition

$$C_{20} \text{ is real} \quad (2.7)$$

hold. Indeed, C_{20} can be expressed in terms of ordinary moments as

$$C_{20} = M_{20} - M_{02} + 2iM_{11} \quad (2.8)$$

and the condition (2.7) applied to (2.8) is equivalent to $M_{11} = 0$, which is the classical definition of the principal axes. However, there is ambiguity in this definition since rotation of these principal axes by any integer multiple of $\pi/2$ radians yields yet another valid pair of principal axes. Most attempts to get rid of this ambiguity impose further *inequality* conditions on other moments to define a unique orientation of the principal axes (e.g., [1]). A unified treatment for the ambiguity question will be given in the next section.

A possible extension of the classical normalization uses further conditions on C_{00} ($=M_{00}$) and C_{11} ($=M_{20} + M_{02}$) to normalize $g(x, y)$ against scale and contrast changes. For example, the conditions $C_{00} = \alpha$ and $C_{11} = \beta$ can be used, where α and β are positive real constants. These conditions together provide two equations in two unknowns, D (dimensional scale) and G (contrast level), which can be solved for the two unknowns:

$$D = \sqrt{\beta/\alpha} \times \left(\frac{M_{00}}{M_{20} + M_{02}} \right)^{1/2}$$

$$G = (\alpha^2/\beta) \times \frac{M_{20} + M_{02}}{M_{00}^2}. \quad (2.9)$$

These parameters are then used to transform $g_1(x, y)$ into the normalized version $g_2(x, y) = Gg(x/D, y/D)$. The value of D in (2.9) is independent of the image contrast and the value of G is independent of the image scale. Adjusting the scale and contrast of an image using D and G is equivalent to adjusting moment invariants as described in [4] to achieve independent scale and contrast invariance.

In summary, in the classical normalization procedure $N_{2,0}$, the centroid (X, Y) is determined by the condition on C_{10} , the scales D, G are determined by the conditions on C_{00}, C_{11} , and the rotation angle Φ is determined by the condition on C_{20} (which is the rationale for denoting the procedure by $N_{2,0}$). We are now in a position to introduce the broader class $\{N_{p,q}\}$ of normalization procedures.

III. THE (p, q) NORMALIZATION

Guided by the classical normalization discussed in the previous section, we introduce a class of normalization procedures $\{N_{p,q}\}$ based on complex moments of different orders (p, q) and interpret these procedures geometrically.

A. Definition

We define the normalization procedure $N_{p,q}$ (where $p > q, p > 1$) of the image function $g_1(x, y)$ as the transformation defined in (2.1) and (2.2) which has the following normalization criteria (independent of p and q):

$$\begin{aligned} C_{00} &= \alpha \\ C_{10} &= 0 \\ C_{11} &= \beta \end{aligned} \quad (3.1)$$

where α and β are positive real constants, and the following normalization criterion which is dependent on p and q :

$$C_{pq} \text{ is positive real.} \quad (3.2)$$

Significantly, the extension of the classical normalization procedure involves imposing *phase* constraint on higher order complex moments [the condition on C_{pq} in (3.2)] which is completely analogous to the derivation of higher order moment invariants [2]. We observe that since C_{pq} is the complex conjugate of C_{qp} , $N_{p,q}$ is equivalent to $N_{q,p}$ (we get rid of the redundancy by requiring that $p > q$). By simple enumeration, we find that the number of distinct normalization procedures $N_{p,q}$ is $N(N+1) - 1$ for $p+q \leq 2N$.

We now consider the question of whether for a general image $g_1(x, y)$ there exists a normalized image $g_2(x, y)$ obtained by the transformation (2.1) and (2.2) and satisfying (3.1) and (3.2). We also examine the uniqueness question. As discussed in the previous section, conditions (3.1) determine X, Y, D, G for any image $g_1(x, y)$. These values, given by (2.6) and (2.9), are unique and independent of Φ [the rotational parameter of (2.2)]. Therefore, for a given image $g_1(x, y)$, all the (p, q) normalization procedures $N_{p,q}$ have the same centroid (X, Y) , the same dimensional scale D , and the same contrast level G .

The existence of the $N_{p,q}$ -normalized image $g_2(x, y)$ now amounts to satisfaction of condition (3.2). Denoting by $C_{pq}(\varphi)$ the value of C_{pq} when the axes of the original image are rotated through a counterclockwise angle φ , we have the following relation:

$$C_{pq}(\varphi) = C_{pq}(0) e^{-i(p-q)\varphi} \quad (3.3)$$

which follows by substitution in (2.4). Equation (3.3) says that as the image is rotated, each complex moment goes through all possible phases of a complex number (recall that $p > q$ in the definition of $N_{p,q}$) while its magnitude $|C_{pq}|$ remains unchanged. Therefore, unless the magnitude happens to be zero, C_{pq} will become positive real at some angle $\varphi = \Phi$. In fact, there will be exactly $p - q$ values of the angle φ , equally spaced at $2\pi/(p - q)$ radians apart, for which C_{pq} is positive real, as follows from (3.3). This means that the normalization procedure $N_{p,q}$ has a degeneracy of $p - q$ and is not unique unless $p - q = 1$.

For example, $N_{2,0}$ has $p - q = 2$ possible normalized versions $g_2(x, y)$ for an image $g_1(x, y)$. These two versions are rotations of one another with angular separation $2\pi/2 = \pi$ radians. For each of them, condition (3.2) is equivalent to

$$M_{11} = 0 \text{ and } M_{20} > M_{02}. \quad (3.4)$$

Without the second part of (3.4), it becomes the classical definition of principal axes. The condition $M_{20} > M_{02}$ has been traditionally added to reduce the ambiguity of Φ from four possible angles to two possible angles.

On the other hand, $N_{2,1}$ has degeneracy $p - q = 1$ and, hence, is unique. In this case, condition (3.2) reduces to

$$M_{03} + M_{21} = 0 \text{ and } M_{30} + M_{12} > 0 \quad (3.5)$$

which will be satisfied for only one angle Φ .

In conclusion, we point out that the normalization scheme we propose here is one of several possible schemes that can be

defined in a similar way. For example, another set of conditions can be used instead of (3.1) to define the centroid and normalization scales. Also, condition (3.2) can be replaced, for instance, by a condition on a linear combination of C_{pq} 's having the same value of $p - q$ (the value $p - q$ is called the *repetition* of C_{pq}) like the Zernike moments [3]. We now give a geometrical interpretation of $N_{p,q}$ in order to develop more insight into moment-based normalization.

B. Phasor Diagrams

The normalization procedure $N_{p,q}$ has an interesting interpretation that becomes apparent when we rewrite (2.4) after substituting the circular harmonic expansion [6] for the function $g(x, y)$:

$$g(r \cos \vartheta, r \sin \vartheta) = \sum_{n=-\infty}^{+\infty} c_n(r) e^{in\vartheta}. \quad (3.6)$$

Complex moments can then be written in terms of the coefficients of this expansion as follows:

$$C_{pq} = 2\pi \int_0^{+\infty} r^{p+q+1} c_{q-p}(r) dr. \quad (3.7)$$

Equation (3.7) shows that the complex moment of order (p, q) depends only on the $(q - p)$ th circular harmonic coefficient function $c_{q-p}(r)$ of the image function $g(x, y)$. This accounts for the $(p - q)$ -fold ambiguity of $N_{p,q}$ (which is based on C_{pq}) since the component of $g(x, y)$ in (3.6) that depends on $c_{q-p}(r)$ is of the form $c_{q-p}(r) e^{-i(p-q)\vartheta}$ and repeats itself every $2\pi/(p - q)$ radians. Also, for images with n -fold rotational symmetry, all C_{pq} 's for which $p - q$ is not divisible by n are identically zero, as their corresponding circular harmonic coefficient functions are zeros. Interestingly, a normalization procedure, which in our notation is $N_{3,0}$, was suggested in [1] as an alternative condition for defining the principal axes for images having threefold rotational symmetry, because $N_{2,0}$ fails since $|C_{20}| = 0$ for such images.

Assume now that the image $g(x, y)$ has been centered and properly scaled according to conditions (3.1). We fix the coordinate system and consider the complex moments

$$C_{pq} = A_{pq} e^{i\Psi_{pq}}. \quad (3.8)$$

For all C_{pq} 's having the same repetition $r = p - q$, we plot a *phasor* diagram, denoted by P_r , in which each C_{pq} is represented by a vector or phasor of length A_{pq} and orientation Ψ_{pq} . Thus, we get phasor diagrams P_1, P_2, \dots, P_R which represent subsets of C_{pq} 's having repetitions $1, 2, \dots, R$.

According to (3.3), if the image plane axes are rotated through an angle $\Delta\Phi$, each phasor diagram P_r will rotate through an angle $-r\Delta\Phi$. Therefore, phasors in the same diagram maintain their *relative* phases upon rotation of the image. Furthermore, as the image plane axes are rotated one complete cycle, the phasor diagram P_r rotates r cycles and each phasor coincides with the positive real axis r times.

As a result, there is an r -fold ambiguity in the orientation of the $N_{p,q}$ principal axes in the image plane, which corresponds to the repetition $p - q = r$ of the complex moment C_{pq} . Denote

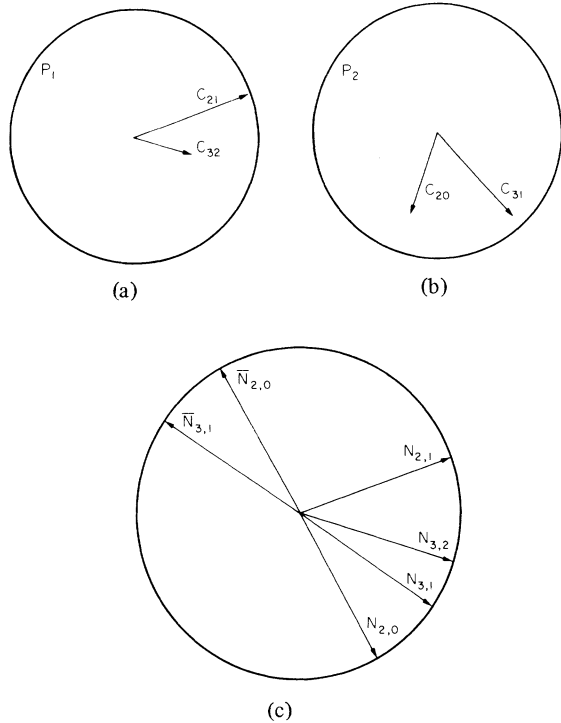


Fig. 1. (a) Two phasors in the phasor diagram with repetition = 1. (b) Two phasors in the phasor diagram with repetition = 2. (c) The corresponding x -axes of the (p, q) normalization procedures drawn on the image plane.

by $\Phi_{p,q}$ the orientation of these principal axes with respect to (w.r.t.) the fixed coordinate system in the image plane ($\Phi_{p,q}$ is ambiguous for $p - q \neq 1$). The relation between $\Phi_{p,q}$ (an angle in the image plane) and Ψ_{pq} [the phase of C_{pq} as defined in (2.4)] is

$$\Psi_{pq} = (p - q) \Phi_{p,q} \quad (3.9)$$

as follows from (3.2) and (3.3). Using (3.9), we can draw the principal axes of $N_{p,q}$ in terms of the phasor diagram P_{p-q} .

Fig. 1(a) shows two phasors in P_1 for a sample image. Fig. 1(b) shows two phasors in P_2 for the same image. These phasor diagrams do not have the same repetition (speed of rotation). Fig. 1(c) shows the image plane with the x -principal axes of $N_{p,q}$ corresponding to the phasors in (a) and (b). Observe that the phasors in Fig. 1(a) and (b) relate to the principal axes in Fig. 1(c) according to (3.9).

Since $N_{p,q}$'s corresponding to the phasor diagram P_1 , i.e., having degeneracy $p - q = 1$, are not ambiguous, they can be used to remove the ambiguity in all other $N_{p,q}$'s. For example, let the principal axes of $N_{3,2}$ (unambiguous) be the fixed coordinate system. For any $N_{p,q}$ with $p - q > 1$, one can unambiguously define the $N_{p,q}$ principal axes by requiring the orientation of the x -axis of the $N_{p,q}$ -normalized image to fall within $(-\pi/(p - q), +\pi/(p - q))$ from the fixed x -axis of $N_{3,2}$. This will specify only one of the $p - q$ possible values of the angle $\Phi_{p,q}$ as the unique orientation of the $N_{p,q}$ principal axes. For example, this will exclude the two LHS axes, denoted by $\bar{N}_{2,0}$ and $\bar{N}_{3,1}$, in Fig. 1(c) from consideration because they are not within $(-\pi/2, \pi/2]$ from the fixed x -axis denoted by $N_{3,2}$ in this figure.

IV. INVARIANCE ANALYSIS

In this section, we relate the class $\{N_{p,q}\}$ of normalization procedures to moment invariants. Through this relation, moment invariants are shown to be *variant* features of a normalized image. In other words, if the image is normalized and, due to noise and quantization, the normalized version is offset from the standard form, this error is equally sensed by a moment invariant computed in terms of the original image. This result puts moment invariants on an equal footing with other possible pattern recognition features of the normalized image as far as invariance is concerned.

A. Moment Invariants

There are several ways to define moment invariants. Algebraic invariants [1] were the original basis of their definition. For our purposes, we define them in an equivalent way using complex moments [2] and normalized images.

Let $\hat{g}_1(x, y)$ be the normalized version of $g_1(x, y)$ obtained by shifting, contrast scaling, and dimensional scaling of $g_1(x, y)$ so that $\hat{g}_1(x, y)$ satisfies the normalization criteria given by (3.1), i.e.,

$$\hat{g}_1(x, y) = G g_1(X + x/D, Y + y/D) \quad (4.1)$$

where X, Y, D, G are given by (2.6) and (2.9). $\hat{g}_1(x, y)$ can be thought of as a preliminary step in obtaining the normalized version $g_2(x, y)$ of $g_1(x, y)$ using an $N_{p,q}$ procedure. We shall refer to $\hat{g}_1(x, y)$ as the prenormalized version of $g_1(x, y)$. Let C_{pq} be the complex moment of $\hat{g}_1(x, y)$, i.e., with $g(x, y)$ in (2.4) replaced by $\hat{g}_1(x, y)$. Consider the phasor diagrams P_r 's of $\hat{g}_1(x, y)$.

When the axes of $\hat{g}_1(x, y)$ are rotated by $\Delta\Phi$, the phasors of P_r rotate by $-r\Delta\Phi$. We can combine phasors with the appropriate relative speeds of rotation in such a way that the result does not rotate at all, i.e., the result becomes invariant under rotation of $\hat{g}_1(x, y)$. For example,

$$C_{(n+r)n} C_{(m+r)m}^* \text{ or } C_{(n+r)n}^k C_{(m+kr)m}^* \quad (4.2)$$

where $*$ denotes the complex conjugate are indeed invariant under rotation of $\hat{g}_1(x, y)$. The power k is meant to "speed up" the phasors in P_r so that they rotate at the same speed as the ones in P_{kr} . The forms in (4.2) are usually added to their conjugates to obtain real invariants.

Using simple algebra, this definition of moment invariants can be shown to be strictly equivalent to the definition based on scaling combinations of central moments. Similar definitions can be formulated for other forms of moment invariants.

B. Invariance

Before characterizing the invariance of moment invariants, one should define what is meant here by "invariance." If the image $g_1(x, y)$ is noise-free and no sampling or quantization errors occur, moment invariants as well as any valid normalization procedure provide *perfect* invariance. In this ideal case, the equivalence in terms of invariance of normalizing the image and then classifying it with general features and using moment invariants in the first place is obvious.

On the other hand, when noise and errors are present, we

have a different situation. The additional issue of *robustness* of the invariance scheme in these conditions becomes a crucial point in evaluating invariance. One should consider which invariance or normalization scheme is less vulnerable to noise. Another consideration is how tolerant the classification features are w.r.t. the expected errors in the normalization scheme.

Assume now that a noise process $\{n_1(x, y)\}$ is added to the image function $s_1(x, y)$ (where s_1 stands for the input signal) so that the input to the system $g_1(x, y)$ is given by

$$g_1(x, y) = s_1(x, y) + n_1(x, y). \quad (4.3)$$

Suppose now that we use some normalization procedure N to obtain the normalized version $g_2(x, y)$ from $g_1(x, y)$. By definition, the normalized image $g_2(x, y)$ will satisfy the normalization criteria. The two components of $g_1(x, y)$, namely $s_1(x, y)$ and $n_1(x, y)$, will be transformed according to (2.1) and (2.2) into two functions, $s_2(x, y)$ and $n_2(x, y)$. These functions are *not* normalized versions of $s_1(x, y)$ and $n_1(x, y)$, i.e., they do not necessarily satisfy the normalization criteria, but the image $g_2(x, y) = s_2(x, y) + n_2(x, y)$ does satisfy these criteria.

In this setup, robustness of a feature F extracted from the overall normalized image function $g_2(x, y)$ is twofold: how F tolerates the presence of the noise component $n_2(x, y)$, i.e., *noise tolerance*, and how F tolerates the erroneous normalization of the signal component $s_2(x, y)$, i.e., *invariance*. We now turn to the question of invariance characterization using these concepts. We shall carry out the discussion using moment invariants of the form $C_{pq} C_{pq}^*$ [defined on $\hat{g}_1(x, y)$].

C. Shift and Scale

According to our discussion in Section IV-A, the shift and scale invariance properties of moment invariants are derived by a method that is equivalent to computing the moments in terms of an image that has been normalized against scale and shift changes. We now investigate whether the equivalence is valid in the presence of noise. Suppose that the noisy image $g_1(x, y) = s_1(x, y) + n_1(x, y)$ is normalized using the prenormalization procedure described in Section IV-A into the prenormalized version $\hat{g}_1(x, y) = \hat{s}_1(x, y) + \hat{n}_1(x, y)$. Due to the presence of $n_1(x, y)$, the estimate for X, Y, D, G using (2.6) and (2.9) is offset from the nominal values which would be obtained were $n_1(x, y)$ absent. The offset values, say X_1, Y_1, D_1, G_1 , will misnormalize the signal part $s_1(x, y)$ such that $\hat{s}_1(x, y)$ is misplaced and misscaled.

In order to identify the shift and scale invariance properties of moment invariants, we consider, for example, the moment invariant whose value in terms of the *input* image $g_1(x, y)$ is the same as $C_{pq} C_{pq}^*$ in terms of the *prenormalized* image $\hat{g}_1(x, y)$ (as discussed in Section IV-A). A glance at (2.4), which defines C_{pq} in terms of an image $g(x, y)$, shows that if x and y are shifted or scaled, or if g is scaled, the value of $C_{pq} C_{pq}^*$ does change. For example, if the prenormalized image $\hat{g}_1(x, y)$ is obtained from the input image $g_1(x, y)$ using the wrong contrast level G_1 instead of the correct value G , the value of the moment invariant $C_{pq} C_{pq}^*$ will be changed by a factor of $(G_1/G)^2$. The value of the moment invariant will be

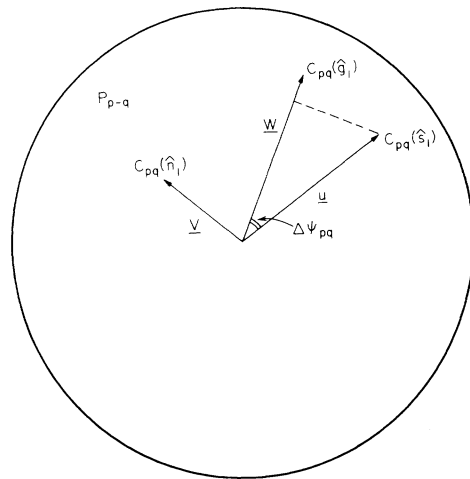


Fig. 2. Representation of signal and noise in the phasor diagram. The noise-free image contributes the vector \mathbf{u} while the noise contributes the vector \mathbf{v} in the overall value of the (p, q) complex moment represented by the vector \mathbf{w} . The projection of \mathbf{u} on \mathbf{w} depends on the orientation error $\Delta\Psi_{pq}$.

invariant only as long as the estimates for X, Y, D, G are correct. In other words, the shift and scale invariance of moment invariants is captured in the prenormalization procedure described in Section IV-A.

We conclude that the shift and scale invariance of moment invariants is provided entirely by the prenormalization procedure. If, due to the presence of noise, the normalization is imperfect, then their values are altered. The question becomes how much they are altered in comparison to other potential features defined on the prenormalized image that are *a priori* variant but perhaps tolerant to the expected errors in the prenormalization procedure.

D. Rotation

Rotation of the image $g_1(x, y)$ about any point is equivalent to a shift operation together with rotation about its centroid, and hence, $\hat{g}_1(x, y)$ will only rotate as a result (the shift invariance of $g_1(x, y)$ was treated in Section IV-C). Therefore, we need to investigate the invariance properties of $C_{pq} C_{pq}^*$ defined on $\hat{g}_1(x, y)$ as this image rotates.

The image $\hat{g}_1(x, y)$ is composed of the signal part $\hat{s}_1(x, y)$ and the noise part $\hat{n}_1(x, y)$. Therefore, the complex moment of the overall image $\hat{g}_1(x, y)$, denoted by $C_{pq}(\hat{g}_1)$, is given by

$$C_{pq}(\hat{g}_1) = C_{pq}(\hat{s}_1) + C_{pq}(\hat{n}_1) \quad (4.4)$$

where $C_{pq}(\hat{s}_1)$ and $C_{pq}(\hat{n}_1)$ are the corresponding complex moments of $\hat{s}_1(x, y)$ and $\hat{n}_1(x, y)$. Fig. 2 represents this relation in the phasor diagram P_{p-q} of the image $\hat{g}_1(x, y)$. Let us take $p-q = 1$ to fix the ideas without having to deal with the ambiguity associated with $p-q > 1$. The orientation of the x -axis of the principal axes in the normalization procedure $N_{p,q}$ will be in the same direction as $C_{pq}(\hat{g}_1)$, the vector \mathbf{w} in Fig. 1, according to relation (3.9).

Now, the value of $C_{pq}(\hat{g}_1) C_{pq}^*(\hat{g}_1)$, or equivalently $|C_{pq}(\hat{g}_1)|$, is the length of the vector \mathbf{w} . This length corresponds to the total value (signal + noise) of the moment invariant. It is the sum of the projections of the signal vector \mathbf{u} ($C_{pq}(\hat{s}_1)$) and the

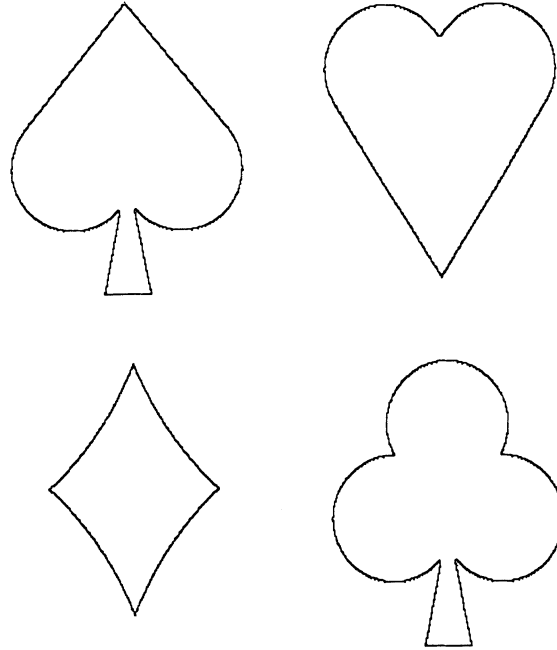


Fig. 3. The four objects of the experiment (extracted edges): Spade, Heart, Diamond, and Club.

noise vector $\mathbf{v}(C_{pq}(\hat{n}_1))$ on the direction of \mathbf{w} . Should noise have been absent, the value of $|C_{pq}|$ would be the length of the vector \mathbf{u} only.

Let us examine the invariance of $|C_{pq}(\hat{g}_1)|$ in view of Fig. 2. The component of the *noise-free* value $|C_{pq}(\hat{s}_1)|$ contributing to the total value $|C_{pq}(\hat{g}_1)|$ is $|C_{pq}(\hat{s}_1)| \cos \Delta\Psi_{pq}$ which is the same as the real part of $C_{pq}(\hat{s}_1)$ (referred to the $N_{p,q}$ axes). In other words, the *signal* part of the moment invariant $|C_{pq}(\hat{g}_1)|$ is precisely $\text{Re}(C_{pq}(\hat{s}_1))$. Let us define $R_{pq}(g)$ to be the real part of $C_{pq}(g)$ for any image $g(x, y)$, i.e.,

$$\begin{aligned}
 R_{pq}(g) &= \text{Re} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x+iy)^p (x-iy)^q g(x,y) dx dy \\
 &= \int_0^{2\pi} \int_0^{+\infty} r^{p+q} (\cos(p-q)\vartheta) \\
 &\quad \cdot g(r \cos \vartheta, r \sin \vartheta) r dr d\vartheta.
 \end{aligned}
 \tag{4.5}$$

The value of $R_{pq}(g_2)$ is the same as $|C_{pq}(\hat{g}_1)|$. Moreover, $R_{pq}(\hat{s}_1)$ is indeed the component of the moment invariant $|C_{pq}(\hat{g}_1)|$ coming from the signal part $\hat{s}_1(x, y)$.

Therefore, the invariance of this moment invariant, in the sense defined in Section IV-B, is captured in the normalization procedure $N_{p,q}$. The value of $R_{pq}(g)$ as a function of $g(x, y)$ is not rotation invariant. In fact, as $g(x, y)$ is rotated, $R_{pq}(g)$ assumes all values from $-|C_{pq}(g)|$ to $+|C_{pq}(g)|$. Comparison between the rotation invariance of this moment invariant and any other feature $F(g)$ defined on the $N_{p,q}$ -normalized image $g(x, y)$ now amounts to comparison between the tolerance of $R_{pq}(g)$ and $F(g)$ to the orientation error $\Delta\Phi_{p,q}$. Other forms of moment invariants can be treated in a similar way using hybrid normalization procedures instead of $N_{p,q}$.

E. Conclusion

We have traced moment invariants of the form $C_{pq}C_{pq}^*$ or $|C_{pq}|$ in order to characterize their invariance with respect to shift, dimensional scaling, contrast, and rotation. We have found that this invariant can be expressed as the feature $R_{pq}(g)$ in (4.5) applied to an $N_{p,q}$ -normalized image. It is simple to verify that $R_{pq}(g)$ is a considerably variant feature of the image function $g(x, y)$ by substituting the transformation (2.1), (2.2) in (4.5). Similar derivations can be carried out for other forms of moment invariants.

We conclude that moment invariants are variant features of a normalized image. After the ambiguity of the normalization procedure is removed, as discussed in Section III-B, one can utilize the invariance of moment invariants, which is captured in the corresponding normalization procedures, and use another feature of the normalized image for classification. It is a problem-oriented question to choose an optimal normalization procedure w.r.t. the noise conditions and a set of features which are tolerant w.r.t. the expected normalization errors and discriminant w.r.t. the classes in question.

V. EXPERIMENTAL RESULTS

In this section, we demonstrate an application of image normalization in pattern recognition. The experiment is intended to exemplify that normalizing the image using moments, and then classifying it using an alternative method, such as template matching, can be superior to direct classification using moment invariants in the case of high noise level or small object size (large sampling error).

A. The Experiment

In this experiment, we implemented a digital pattern recognition system to distinguish between the four playing card

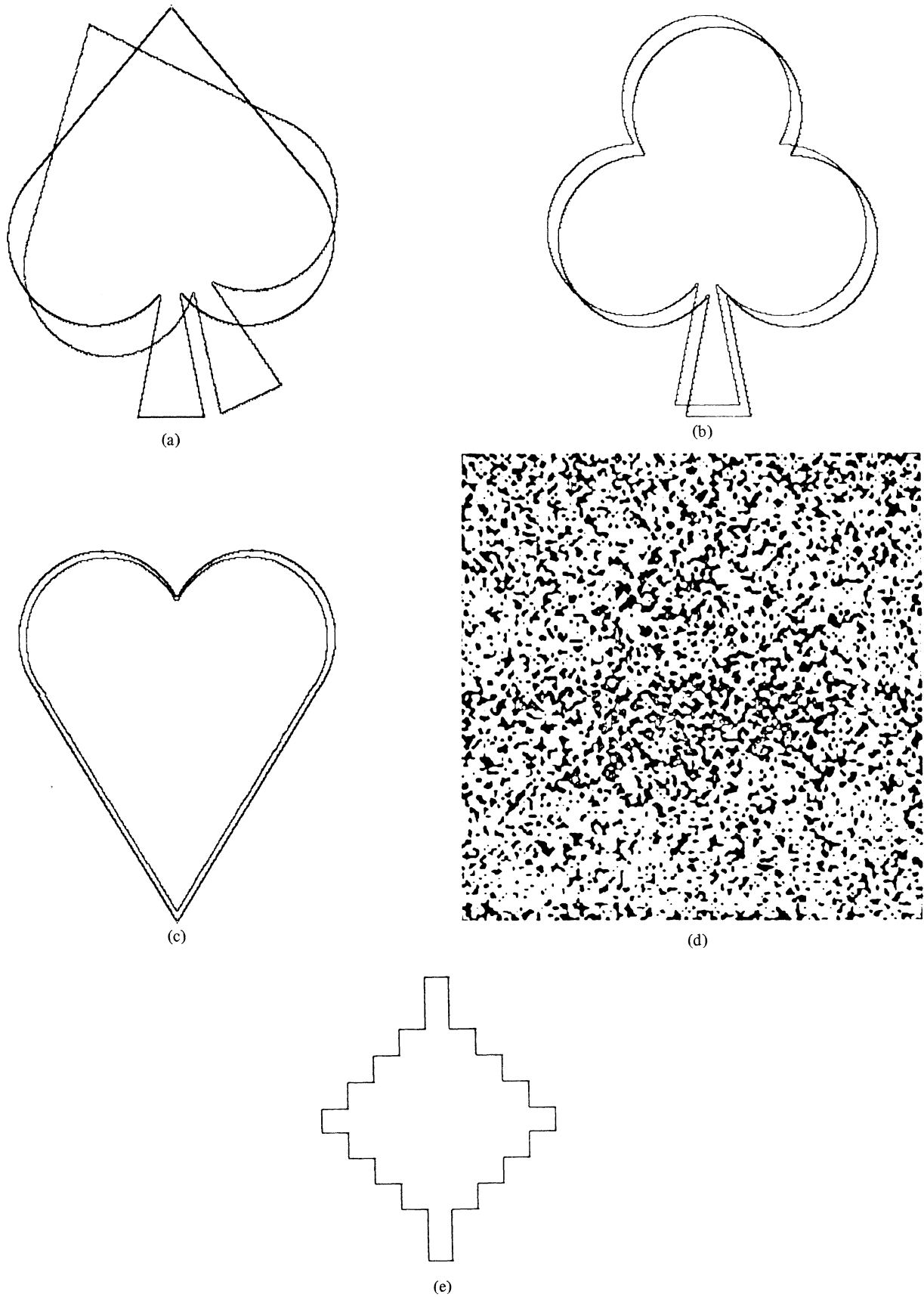


Fig. 4. (a) The maximum error in orientation of the Spade for correct classification. (b) Positioning error of the Club for 90 percent matching ratio. (c) Scaling error of the Heart for 90 percent matching ratio. (d) The maximum noise level for correct classification of a normalized Club (thresholded image). (e) The smallest grid for correct classification of the Diamond.

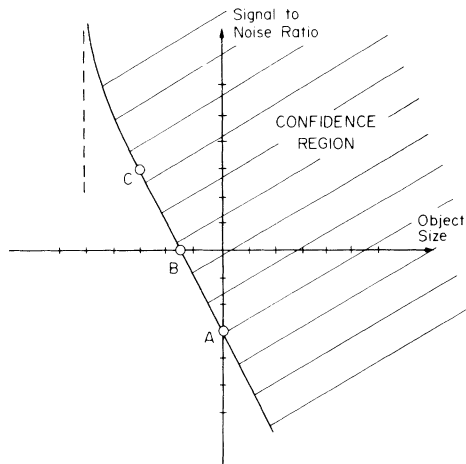


Fig. 5. A plot of the confidence region for correct classification on log scale of signal-to-noise ratio versus object size (one unit corresponds to 2 dB). The extremal points A , B , C are the subject of the verification experiment reported.

symbols, Spade, Heart, Diamond, and Club, irrespective of position, size, orientation, or contrast within a noisy image (Fig. 3 shows the *edges* of these objects for clarity, as in most figures of this section). The scheme utilized was to normalize the image using the procedure $N_{2,0}$ described in Section III, and then to classify the normalized image using *multiple* template matching (to make up for the ambiguity in $N_{2,0}$). To compare with the method of moment invariants, we used noise levels and object sizes for which the feature space of moment invariants, requiring approximately the same amount of computation, was used in a similar experiment and resulted in erroneous classification decisions.

Template matching was expected to be tolerant to normalization errors in this particular experiment, since the objects in question did not have a wide band of spatial frequencies [5]. We ran some preliminary tests to estimate this tolerance. For example, Fig. 4(a) shows the maximum error in the orientation of the Spade symbol for which template matching leads to correct classification. Fig. 4(b) and (c) shows shift and scale errors for the Club and Heart symbols which diminish the matching ratio to 90 percent. Fig. 4(d) shows the maximum noise level over a properly normalized Club symbol (the Club symbol looks very vague in the center of this noisy image) to be correctly classified, and Fig. 4(e) shows the worst sampling grid (minimum object size) for which a Diamond symbol is correctly classified.

Combining these factors, we derived the “confidence region” for reliable operation of our pattern recognition system. The coordinates in question are the object size and the signal-to-noise ratio. Fig. 5 shows a plot of the confidence region in the size-SNR plane using a logarithmic scale. At any point within the confidence region, the system was expected to operate reliably.

B. The Result

Guided by Fig. 5, we chose three extremal points, denoted by A , B , C in this figure, to test the moment normalization/template matching classification scheme. We chose the test object to be Diamond for point A , Club for point B , and Heart for point C . For the noise we used a pseudorandom process

which is periodically correlated. Fig. 6(a), (c), and (e) shows the input images in these experiments.

The results in all three experiments were positive. Each symbol was classified correctly with reasonable reliability. The matching ratios for the three symbols were 93, 89, and 94 percent, respectively. The closest potential misclassification was when the Spade symbol had a matching ratio of 83 percent with the Club symbol. Fig. 6(b), (d), and (f) show the normalized versions of the three input images just before carrying out the template matching. The correct positions should be central and either vertical or horizontal, so one can see that the tolerance of template matching in this case admitted notable error in the value of moments used for normalization.

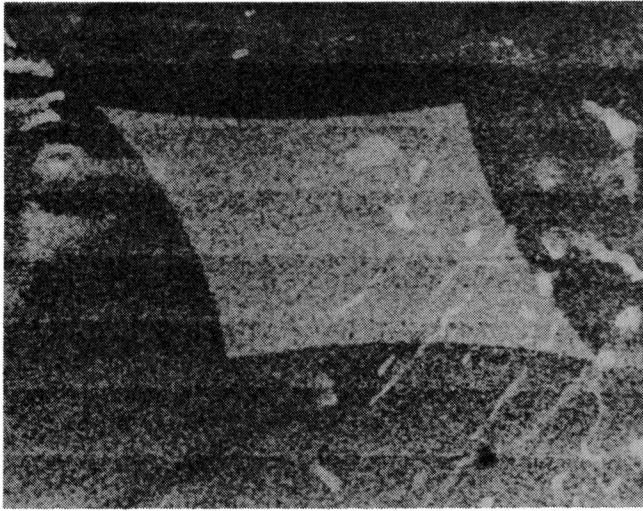
VI. CONCLUSION

We have been concerned with distinguishing between aspects of normalization and aspects of classification in moments, i.e., how moments behave in telling *where* the object is as opposed to *what* the object is, and analyzing moment-based normalization more closely. We found that the classical definition of centroid and principal axes is a special case of a class of normalization procedures directly related to moment invariants. From this relation, we found that moment invariants could be viewed as variant features of a moment-normalized image. This led to consideration of other features for recognition in place of moment invariants, and we reported an experiment where this proved to be favorable.

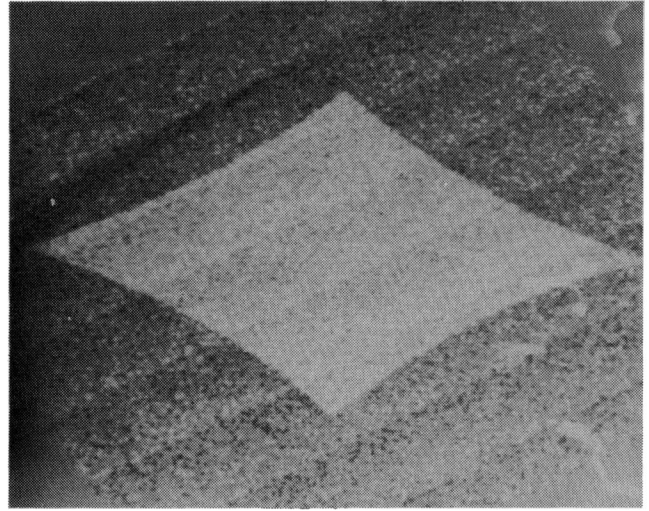
It was also pointed out that several variations of the moment-based normalization procedures introduced here were possible. It is suggested that the behavior of these normalization procedures in the presence of noise be analyzed, and specific questions as to how to combine them in an optimal way to produce an immune normalization procedure for given noise conditions be addressed.

ACKNOWLEDGMENT

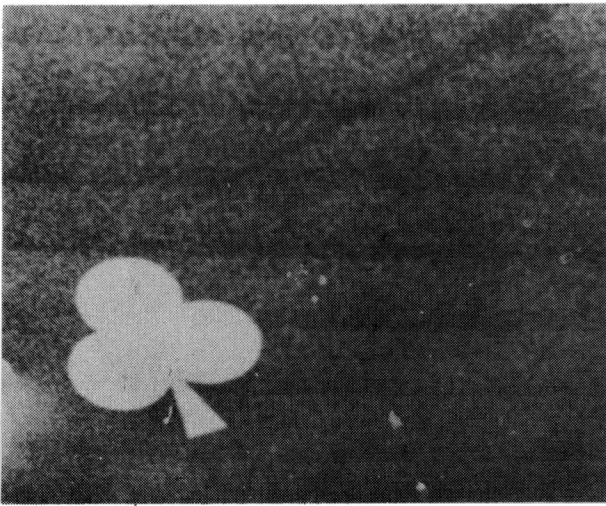
The reported experiment was a term project in the CS/EE 119 course at Caltech. Acknowledgment is due S.-A. El-Hamamsy and S. Venkatesh for their contribution.



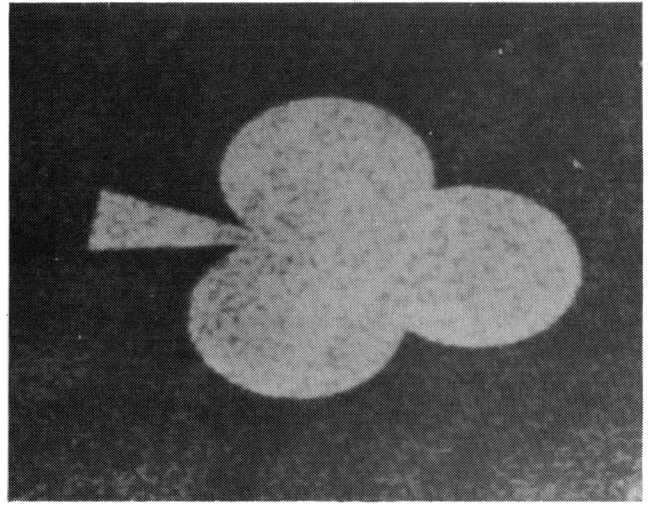
(a)



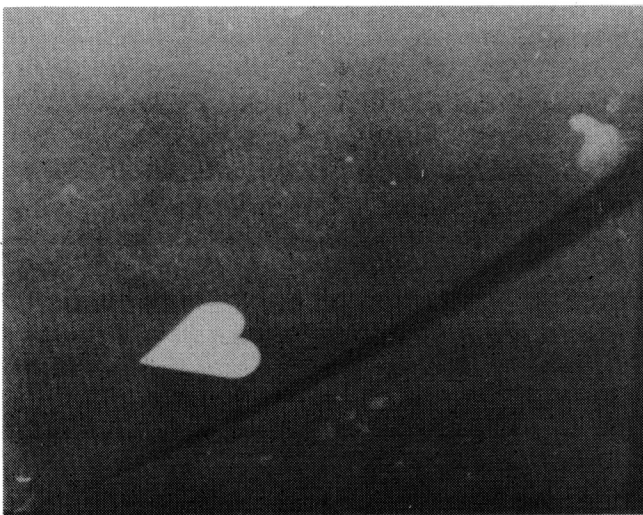
(b)



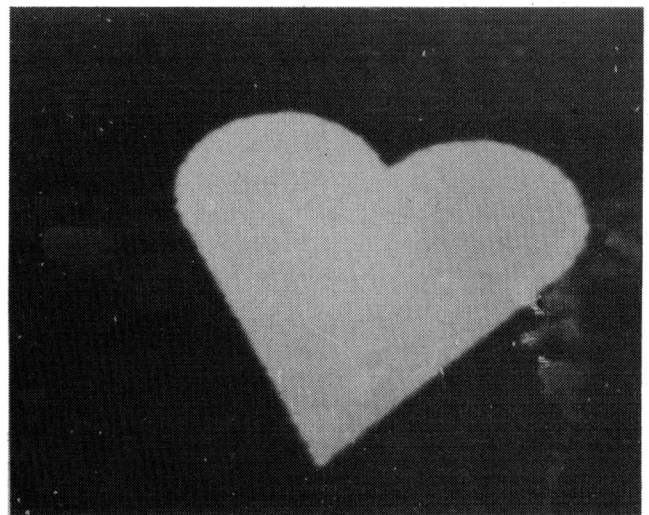
(c)



(d)



(e)

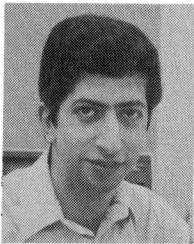


(f)

Fig. 6. (a) Input Diamond image. (b) Normalized Diamond image. (c) Input Club image. (d) Normalized Club image. (e) Input Heart image. (f) Normalized Heart image.

REFERENCES

- [1] M. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. Inform. Theory*, vol. IT-8, pp. 179-187, Feb. 1962.
- [2] Y. Abu-Mostafa and D. Psaltis, "Recognitive aspects of moment invariants," *IEEE Trans. Pattern Anal. Mach. Intell.*, pp. 698-706, Nov. 1984.
- [3] M. Teague, "Image analysis via the general theory of moments," *J. Opt. Soc. Amer.*, vol. 70, pp. 920-930, Aug. 1980.
- [4] S. Maitra, "Moment invariants," *Proc. IEEE*, vol. 67, pp. 697-699, Apr. 1979.
- [5] D. Casasent and D. Psaltis, "New optical transforms for pattern recognition," *Proc. IEEE*, vol. 65, pp. 77-84, Jan. 1977.
- [6] A. Cormack, "Representation of a function by its line integrals, with some radiological applications," *J. Appl. Phys.*, vol. 34, pp. 2722-2727, Sept. 1963.

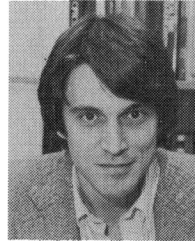


Yaser S. Abu-Mostafa was born in Giza, Egypt, on March 3, 1957. He received the B.Sc. degree in electrical engineering from Cairo University, Giza, in 1979, the M.S.E.E. degree from the Georgia Institute of Technology, Atlanta, in 1981, and the Ph.D. degree in electrical engineering and computer science from the California Institute of Technology, Pasadena, in 1983.

He was appointed Assistant Professor of Electrical Engineering and Computer Science at Caltech in 1983. His current research interests

include information theory, computational complexity, pattern recognition, and signal processing.

Dr. Abu-Mostafa was awarded the 1982-1983 Milton and Francis Clauser Prize for the most original doctoral dissertation at Caltech.



Demetri Psaltis (S'77-M'77) received the B.S. degree in electrical engineering and economics and the M.Sc. and Ph.D. degrees in electrical engineering from Carnegie-Mellon University, Pittsburgh, PA, in 1974, 1975, and 1977, respectively.

After the completion of his Ph.D. dissertation he remained at Carnegie-Mellon, as a Research Associate and later as a Visiting Assistant Professor, for three years. In 1980 he joined the faculty of the Department of Electrical Engineering,

California Institute of Technology, Pasadena, where he is now an Assistant Professor and a consultant to industry. His research interests are in the areas of optical information processing, image processing, pattern recognition, and optical devices. He has published over 80 technical papers in these areas. He is currently working on acousto-optic image processing architectures, adaptive optical processing methods, acousto-optic and photorefractive devices, and numerical optical computers.