Visual Pattern Recognition by Moment Invariants*

MING-KUEI HU† SENIOR MEMBER, IRE

Summary—In this paper a theory of two-dimensional moment invariants for planar geometric figures is presented. A fundamental theorem is established to relate such moment invariants to the well-known algebraic invariants. Complete systems of moment invariants under translation, similitude and orthogonal transformations are derived. Some moment invariants under general two-dimensional linear transformations are also included.

Both theoretical formulation and practical models of visual pattern recognition based upon these moment invariants are discussed. A simple simulation program together with its performance are also presented. It is shown that recognition of geometrical patterns and alphabetical characters independently of position, size and orientation can be accomplished. It is also indicated that generalization is possible to include invariance with parallel projection.

I. Introduction

RECOGNITION of visual patterns and characters independent of position, size, and orientation in the visual field has been a goal of much recent research. To achieve maximum utility and flexibility, the methods used should be insensitive to variations in shape and should provide for improved performance with repeated trials. The method presented in this paper meets all these conditions to some degree.

Of the many ingeneious and interesting methods so far devised, only two main categories will be mentioned here: 1) The property-list approach, and 2) The statistical approach, including both the decision theory and random net approaches.[1] The property-list method works very well when the list is designed for a particular set of patterns. In theory, it is truly position, size, and orientation independent, and may also allow for other variations. Its severe limitation is that it becomes quite useless, if a different set of patterns is presented to it. There is no known method which can generate automatically a new property-list. On the other hand, the statistical approach is capable of handling new sets of patterns with little difficulty, but it is limited in its ability to recognize patterns independently of position, size and orientation.

This paper reports the mathematical foundation of two-dimensional moment invariants and their applications to visual information processing.[2] The results show that recognition schemes based on these invariants could be truly position, size and orientation independent, and also flexible enough to learn almost any set of patterns.

In classical mechanics and statistical theory, the concept of moments is used extensively; central moments, size normalization, and principal axes are also used. To the author's knowledge, the two-dimensional moment invariants, absolute as well as relative, that are to be presented have not been studied. In the pattern recognition field, centroid and size normalization have been exploited[3-5] for "preprocessing." Orientation normalization has also been attempted.[5] The method presented here achieves orientation independence without ambiguity by using either absolute or relative orthogonal moment invariants. The method further uses "moment invariants" (to be described in III) or invariant moments (moments referred to a pair of uniquely determined principal axes) to characterize each pattern for recognition.

Section II gives definitions and properties of two-dimensional moments and algebraic invariants. The moment invariants under translation, similitude, orthogonal transformations and also under the general linear transformations are developed in Section III. Two specific methods of using moment invariants for pattern recognition are described in IV. A simulation program of a simple model (programmed for an LGP-30), the performance of the program, and some possible generalizations are described in Section V.

II. Moments and Algebraic Invariants

A. A Uniqueness Theorem Concerning Moments

In this paper, the two-dimensional $(p + q)$th order moments of a density distribution function $\rho(x, y)$ are defined in terms of Riemann integrals as

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q \rho(x, y)\, dx\, dy, \qquad p, q = 0, 1, 2, \cdots. \tag{1}$$

If it is assumed that $\rho(x, y)$ is a piecewise continuous therefore bounded function, and that it can have nonzero values only in the finite part of the xy plane; then moments of all orders exist and the following uniqueness theorem can be proved.

Uniqueness Theorem: The double moment sequence $\{m_{pq}\}$ is uniquely determined by $\rho(x, y)$; and conversely, $\rho(x, y)$ is uniquely determined by $\{m_{pq}\}$.

It should be noted that the finiteness assumption is important; otherwise, the above uniqueness theorem might not hold.

* Received by the PGIT, August 1, 1961.

† Electrical Engineering Department, Syracuse University, Syracuse, N. Y.

[1] M. Minsky, "Steps toward artificial intelligence," Proc. IRE, vol. 49, pp. 8–30; January, 1961. Many references to these methods can be found in the Bibliography of M. Minsky's article.

[2] M-K. Hu, Pattern recognition by moment invariants," Proc. IRE (Correspondence), vol. 49, p. 1428; September, 1961.

[3] W. Pitts and W. S. McCulloch, "How to know universals," *Bull. Math. Biophys.*, vol. 9, pp. 127–147; September, 1947.

[4] L. G. Roberts, "Pattern recognition with an adaptive network," 1960 IRE International Convention Record, pt. 2, pp. 66–70.

[5] Minsky, *op. cit.*, pp. 11–12.

B. Characteristic Function and Moment Generating Function

The characteristic function and moment generating function of $\rho(x, y)$ may be defined, respectively, as

$$\phi(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp (iux + ivy) \rho(x, y) \, dx \, dy, \quad (2)$$

$$M(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp (ux + vy) \rho(x, y) \, dx \, dy. \quad (3)$$

In both cases, u and v are assumed to be real. If moments of all orders exist, then both functions can be expanded into power series in terms of the moments m_{pq} as follows:

$$\phi(u, v) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} m_{pq} \frac{(iu)^p}{p!} \frac{(iv)^q}{q!}, \quad (4)$$

$$M(u, v) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} m_{pq} \frac{u^p}{p!} \frac{v^q}{q!}. \quad (5)$$

Both functions are widely used in statistical theory. If the characteristic function $\phi(u, v)$ which is essentially the Fourier transform of $\rho(x, y)$ is known, $\rho(x, y)$ may be obtained from the following inverse Fourier transform,

$$\rho(x, y) = \frac{1}{(2\pi)^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp (-iux - ivy) \phi(u, v) \, du \, dv. \quad (6)$$

The moment generating function $M(u, v)$ is not as useful in this respect, but it is convenient for the discussion in Section III. The close relationships and differences between $\phi(u, v)$ and $M(u, v)$ may be seen much more clearly, if we consider both as special cases of the two-sided Laplace transform of $\rho(x, y)$,

$$\mathcal{L}[\rho(x, y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp (-sx - ty) \rho(x, y) \, dx \, dy, \quad (7)$$

where s and t are now considered as complex variables.

C. Central Moments

The central moments μ_{pq} are defined as

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q \rho(x, y) \, d(x - \bar{x}) \, d(y - \bar{y}), \quad (8)$$

where

$$\bar{x} = m_{10}/m_{00}, \qquad \bar{y} = m_{01}/m_{00}. \quad (9)$$

It is well known that under the translation of coordinates,

$$\begin{aligned} x' &= x + \alpha, \\ y' &= y + \beta, \end{aligned} \qquad \alpha, \beta - \text{constants} \quad (10)$$

the central moments do not change; therefore, we have the following theorem.

Theorem: The central moments are invariants under translation.

From (8), it is quite easy to express the central moments in terms of the ordinary moments. For the first four orders, we have

$$\begin{aligned} \mu_{00} &= m_{00} \equiv \mu, \qquad \mu_{10} = \mu_{01} = 0, \\ \mu_{20} &= m_{20} - \mu\bar{x}^2, \qquad \mu_{11} = m_{11} - \mu\bar{x}\bar{y}, \\ \mu_{02} &= m_{02} - \mu\bar{y}^2, \\ \mu_{30} &= m_{30} - 3m_{20}\bar{x} + 2\mu\bar{x}^3, \\ \mu_{21} &= m_{21} - m_{20}\bar{y} - 2m_{11}\bar{x} + 2\mu\bar{x}^2\bar{y}, \\ \mu_{12} &= m_{12} - m_{02}\bar{x} - 2m_{11}\bar{y} + 2\mu\bar{x}\bar{y}^2, \\ \mu_{03} &= m_{03} - 3m_{02}\bar{y} + 2\mu\bar{y}^3. \end{aligned} \quad (11)$$

From here on, for the simplicity of description, all moments referred to are central moments, and μ_{pq} will be simply expressed as

$$\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q \rho(x, y) \, dx \, dy, \quad (12)$$

and the moment generating function $M(u, v)$ will also be referred to central moments.

D. Algebraic Forms and Invariants

The following homogeneous polynomial of two variables u and v,

$$f = a_{p0}u^p + \binom{p}{1} a_{p-1,1} u^{p-1} v + \binom{p}{2} a_{p-2,2} u^{p-2} v^2 + \cdots + \binom{p}{p-1} a_{1,p-1} u v^{p-1} + a_{0p} v^p, \quad (13)$$

is called a binary algebraic form, or simply a binary form, of order p. Using a notation, introduced by Cayley, the above form may be written as

$$f \equiv (a_{p0}; a_{p-1,1}; \cdots ; a_{1,p-1}; a_{0p})(u, v)^p. \quad (14)$$

A homogeneous polynomial $I(a)$ of the coefficients $a_{p0}, \cdots, a_{0p}$ is an algebraic invariant of weight w, if

$$I(a'_{p0}, \cdots, a'_{0p}) = \Delta^w I(a_{p0}, \cdots, a_{0p}), \quad (15)$$

where $a'_{p0}, \cdots, a'_{0p}$ are the new coefficients obtained from substituting the following general linear transformation into the original form (14).

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix} \begin{bmatrix} u' \\ v' \end{bmatrix}, \qquad \Delta = \begin{vmatrix} \alpha & \gamma \\ \beta & \delta \end{vmatrix} \neq 0. \quad (16)$$

If $w = 0$, the invariant is called an absolute invariant; if $w \neq 0$ it is called a relative invariant. The invariant defined above may depend upon the coefficients of more than one form. Under special linear transformations to be discussed in Section III, Δ may not be limited to the determinant of the transformation. By eliminating Δ between two relative invariants, a nonintegral absolute invariant can always be obtained.

In the study of invariants, it is helpful to introduce another pair of variables x and y, whose transformation with respect to (16) is as follows:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}. \tag{17}$$

The transformation (17) is referred to as a cogredient transformation, and (16) is referred to as a contragredient transformation. The variable x, y are referred to as covariant variables, and u, v as contravariant variables. They satisfy the following invariant relation

$$ux + vy = u'x' + v'y'. \tag{18}$$

The study of algebraic invariants was started by Boole, Cayley and Sylvester more than a century ago, and followed vigorously by others, but interest has gradually declined since the early part of this century. The moment invariants to be discussed in Section III will draw heavily on the results of algebraic invariants. To the authors knowledge, there was no systematic study of the moment invariants in the sense to be described.

III. Moment Invariants

A. *A Fundamental Theorem of Moment Invariants*

The moment generating function with the exponential factor expanded into series form is

$$M(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{p=0}^{\infty} \frac{1}{p!} (ux + vy)^p \rho(x, y)\, dx\, dy. \tag{19}$$

Interchanging the integration and summation processes, we have

$$M(u, v) = \sum_{p=0}^{\infty} \frac{1}{p!} (\mu_{p0}, \cdots, \mu_{0p})(u, v)^p. \tag{20}$$

By applying the transformation (17) to (19), and denoting the coefficients of x' and y' in the transformed factor $(ux + vy)$ by u' and v', respectively, or equivalently by applying both (16) and (17) simultaneously to (19), we obtain

$$M_1(u', v') = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \sum_{p=0}^{\infty} \frac{1}{p!} \cdot (u'x' + v'y')^p \rho'(x', y') \frac{1}{|J|}\, dx'\, dy' \tag{21}$$

where $\rho'(x', y') \equiv \rho(x, y)$, $|J|$ is the absolute value of the Jacobian of the transformation (17), and $M_1(u', v')$ is the moment generating function after the transformation. If the transformed central moments μ'_{pq} are defined as

$$\mu'_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x')^p (y')^q \rho'(x', y')\, dx'\, dy', \quad p, q = 0, 1, 2, \cdots \tag{22}$$

then we have

$$M_1(u', v') = \frac{1}{|J|} \sum_{p=0}^{\infty} \frac{1}{p!} (\mu'_{p0}, \cdots, \mu'_{0p})(u', v')^p. \tag{23}$$

In the theory of algebraic invariants, it is well known that the transformation law for the a coefficients in the algebraic form (14) is the same as that for the monomials, $x^{p-r}y^r$, in the following expression:

$$(ux + vy)^p = (x^p, x^{p-1}y, \cdots, y^p)(u, v)^p. \tag{24}$$

From (19), (20), (21) and (23), it can be seen clearly that the same relationship also holds between the pth order moments and the monomials except for the additional factor $1/|J|$. Therefore, the following fundamental theorem is established.

Fundamental Theorem of Moment Invariants: If the algebraic form of order p has an algebraic invariant,

$$I(a'_{p0}, \cdots, a'_{0p}) = \Delta^w I(a_{p0}, \cdots, a_{0p}), \tag{25}$$

then the moments of order p have the same invariant but with the additional factor $|J|$,

$$I(\mu'_{p0}, \cdots, \mu'_{0p}) = |J| \Delta^w I(\mu_{p0}, \cdots, \mu_{0p}). \tag{26}$$

This theorem holds also between algebraic invariants containing coefficients from two or more forms of different orders and moment invariants containing moments of the corresponding orders.

B. *Similitude Moment Invariants*

Under the similitude transformation, *i.e.*, the change of size,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \alpha & 0 \\ 0 & \alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad \alpha - \text{constant}, \tag{27}$$

each coefficient of any algebraic form is an invariant

$$a'_{pq} = \alpha^{p+q} a_{pq}, \tag{28}$$

where α is not the determinant. For moment invariants we have

$$\mu'_{pq} = \alpha^{p+q+2} \mu_{pq}. \tag{29}$$

By eliminating α between the zeroth order relation,

$$\mu' = \alpha^2 \mu \tag{30}$$

and the remaining ones, we have the following absolute similitude moment invariants:

$$\frac{\mu'_{pq}}{(\mu')^{(p+q)/2} + 1} = \frac{\mu_{pq}}{\mu^{(p+q)/2} + 1}, \quad p + q = 2, 3, \cdots \tag{31}$$

and $\mu'_{10} = \mu'_{01} \equiv 0$.

C. Orthogonal Moment Invariants

Under the following proper orthogonal transformation or rotation:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{32}$$

we have

$$J = \begin{vmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{vmatrix} = +1. \tag{33}$$

Therefore, the moment invariants are exactly the same as the algebraic invariants. If we treat the moments as the coefficients of an algebraic form

$$(\mu_{p0}, \cdots, \mu_{0p})(u, v)^p \tag{34}$$

under the following contragredient transformation:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} u' \\ v' \end{bmatrix}, \tag{35}$$

then we can derive the moment invariants by the following algebraic method. If we subject both u, v and u', v' to the following transformation:

$$\begin{bmatrix} U \\ V \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}, \quad \begin{bmatrix} U' \\ V' \end{bmatrix} = \frac{1}{2}\begin{bmatrix} 1 & i \\ 1 & -i \end{bmatrix} \begin{bmatrix} u' \\ v' \end{bmatrix}, \tag{36}$$

then the orthogonal transformation is converted into the following simple relations,

$$U' = Ue^{-i\theta}, \qquad V' = Ve^{i\theta}. \tag{37}$$

By substituting (36) and (37) into (34), we have the following identities:

$$\begin{aligned} (I_{p0}, \cdots, I_{0p})(U, V)^p &\equiv (\mu_{p0}, \cdots, \mu_{0p})(u, v)^p \\ &\equiv (\mu'_{p0}, \cdots, \mu'_{0p})(u', v')^p \\ &\equiv (I'_{p0}, \cdots, I'_{0p})(Ue^{-i\theta}, Ve^{i\theta})^p, \end{aligned} \tag{38}$$

where $I_{p0}, \cdots, I_{0p}$ and $I'_{p0}, \cdots, I'_{0p}$ are the corresponding coefficients after the substitutions. From the identity in U and V, the coefficients of the various monomials $U^{p-r}V^r$ on the two sides must be the same. Therefore,

$$\begin{aligned} &I'_{p0} = e^{ip\theta} I_{p0}; \qquad I'_{p-1,1} = e^{i(p-2)\theta} I_{p-1,1}; \cdots; \\ &\qquad I'_{1,p-1} = e^{-i(p-2)\theta} I_{1,p-1}; I'_{0p} = e^{-ip\theta} I_{0p}. \end{aligned} \tag{39}$$

These are $(p + 1)$ linearly independent moment invariants under proper orthogonal transformations, and $\Delta = e^{i\theta}$ which is not the determinant of the transformation. From the identity of first two expressions in (38), it can be seen that $I_{r,p-r}$ is the complex conjugate of $I_{p-r,r}$,

$$\begin{aligned} I_{p0} &= \mu_{p0} - i\binom{p}{1}\mu_{p-1,1} - \binom{p}{2}\mu_{p-2,2} \\ &\quad + i\binom{p}{3}\mu_{p-3,3} + \cdots + (-i)^p \mu_{0p}, \\ I_{p-1,1} &= (\mu_{p0} + \mu_{p-2,2}) - i(p-2)(\mu_{p-1,1} + \mu_{p-3,3}) \\ &\quad + \cdots + (-i)^{p-2}(\mu_{2,p-2} + \mu_{0p}), \\ I_{p-2,2} &= (\mu_{p0} + 2\mu_{p-2,2} + \mu_{p-4,4}) \\ &\quad - i(p-4)(\mu_{p-1,1} + 2\mu_{p-3,3} + \mu_{p-5,5}) \\ &\quad + \cdots + (-i)^{p-4}(\mu_{4,p-4} + 2\mu_{2,p-2} + \mu_{0p}), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ I_{p-r,r} &= [(\mu_{p0}; \mu_{p-2,2}; \cdots; \mu_{p-2r,2r})(1, 1)^r; \\ &\quad (\mu_{p-1,1}; \mu_{p-3,3}; \cdots; \mu_{p-2r-1,2r+1})(1, 1)^r; \cdots; \\ &\quad (\mu_{2r,p-2r}; \mu_{2r+2,p-2r-2}; \cdots; \mu_{0p})(1, 1)^r](1, -i)^{p-2r}, \\ &\qquad p - 2r > 0. \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned} \tag{40}$$

and

$$I_{p/2,p/2} = \mu_{p0} + \binom{p/2}{1}\mu_{p-2,2} + \binom{p/2}{2}\mu_{p-4,4} + \cdots + \mu_{0p}, \qquad p = \text{even}.$$

It may be noted that these $(p + 1)$ I's are linearly independent linear functions of the μ's, and vice versa.

For the following improper orthogonal transformation, *i.e.*, rotation and reflection:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad \begin{vmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{vmatrix} = -1. \tag{41}$$

Similarly, we have

$$U' = Ve^{i\theta} \qquad V' = Ue^{-i\theta} \tag{42}$$

and

$$\begin{aligned} &I'_{p0} = e^{-ip\theta} I_{0p}; \qquad I'_{p-1,1} = e^{-i(p-2)\theta} I_{1,p-1}; \cdots; \\ &\qquad I'_{1,p-1} = e^{i(p-2)\theta} I_{p-1,1}; I'_{0p} = e^{ip\theta} I_{p0}, \end{aligned} \tag{43}$$

where $I_{p0}, \cdots, I_{0p}$ and $I'_{p0}, \cdots, I'_{0p}$ are the same as those given by (40).

The orthogonal invariants were first studied by Boole, and the above method was due to Sylvester.[6]

[6] E. B. Elliott, "Algebra of Quantics," Oxford Univ. Press, New York, N. Y., 2nd ed., ch. 15; 1913.

D. A Complete System of Absolute Orthogonal Moment Invariants

From (39) and (43), we may derive the following system of moment invariants by eliminating the factor $e^{i\theta}$:

For the second-order moments, the two independent invariants are

$$I_{11}, \qquad I_{20}I_{02}. \tag{44}$$

For the third-order moments, the three independent invariants are

$$\begin{aligned} &I_{30}I_{03}, \qquad I_{21}I_{12}, \\ &(I_{30}I_{12}^3 + I_{03}I_{21}^3). \end{aligned} \tag{45}$$

A fourth one depending also on the third-order moments only is

$$\frac{1}{i}(I_{30}I_{12}^3 - I_{03}I_{21}^3) \tag{46}$$

There exists an algebraic relation between the above four invariants given in (45) and (46). The first three given by (45) are absolute invariants for both proper and improper rotations but the last one given by (46) is invariant only under proper rotation, and changes sign under improper rotation. This will be called a skew invariant. Therefore it is useful for distinguishing "mirror images."

One more independent absolute invariant may be formed from second and third order moments as follows:

$$(I_{20}I_{12}^2 + I_{02}I_{21}^2). \tag{47}$$

For pth order moments, $p \geq 4$ we have $[p/2]$, the integral part of $p/2$, invariants

$$I_{p0}I_{0p}; \qquad I_{p-1,1}I_{1,p-1}; \cdots ; \qquad I_{p-r,r}I_{r,p-r}; \cdots . \tag{48}$$

If p is even, we also have

$$I_{p/2,p/2}. \tag{49}$$

And also combined with $(p - 2)$th order moments, we have $[p/2 - 1]$ invariants

$$\begin{aligned} &(I_{p-1,1}I_{0,p-2} + I_{1,p-1}I_{p-2,0}), \\ &(I_{p-2,2}I_{1,p-3} + I_{2,p-2}I_{p-3,1}), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ &(I_{p-r,r}I_{r-1,p-r+1} + I_{r,p-r}I_{p-r+1,r-1}), \quad p - 2r > 0 \\ &\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned} \tag{50}$$

combined with second-order moments, if $p =$ odd we have

$$(I^2_{[p/2],[p/2]+1}I_{20} + I^2_{[p/2]+1,[p/2]}I_{02}), \tag{51}$$

if $p =$ even,[7]

$$(I_{p/2-1,p/2+1}I_{20} + I_{p/2+1,p/2-1}I_{02}). \tag{52}$$

Therefore we always have $(p + 1)$ independent absolute invariants. By changing the above sums into differences, we can also have the skew invariants.

All the independent moment invariants together form a complete system, *i.e.*, for any given invariant; it is always possible to express it in terms of the above invariants. The proof is omitted here.

E. Moment Invariants Under General Linear Transformations

From the theory of algebraic invariants under the general linear transformations (17), it is known that the factor Δ is the determinant of the transformation. For linear transformations, J is also the determinant. For simplicity, let A, B, C and a, b, c, d denote the second and third order moments, then we may write the following two binary forms in terms of these moments as

$$\begin{aligned} &(A, B, C)(u, v)^2 \\ &(a, b, c, d)(u, v)^3. \end{aligned} \tag{53}$$

From the theory of algebraic invariants, we have the following four algebraically independent invariants,

$$\begin{aligned} I_1 &\equiv AC - B^2, \\ I_2 &\equiv (ad - bc)^2 - 4(ac - b^2)(bd - c^2), \\ I_3 &\equiv A(bd - c^2) - B(ad - bc) + C(ac - b^2), \\ I_4 &\equiv a^2C^3 - 6abBC^2 + 6acC(2B^2 - AC) \\ &\quad + ad(6ABC - 8B^3) \\ &\quad + 9b^2AC^2 - 18bcABC + 6bdA(2B^2 - AC) \\ &\quad + 9c^2A^2C - 6cdBA^2 + d^2A^3, \end{aligned} \tag{54}$$

of weight $w = 2, 6, 4$ and 6, respectively.

For the zeroth order moment, we have

$$\mu' = |J| \mu. \tag{55}$$

With the understanding that $\Delta^2 = |J|^2$, the following four absolute moment invariants are obtained,

$$\frac{I_1}{\mu^3}, \quad \frac{I_2}{\mu^7}, \quad \frac{I_3}{\mu^5}, \quad \text{and} \quad \frac{I_4}{\mu^7}. \tag{56}$$

There also exists a skew invariant,[8] I_5, of weight 9 depending on the moments A, B, C and a, b, c, d. This also may be normalized as

$$\frac{\Delta}{|J|}\left(\frac{I_5}{\mu^{10}}\right), \tag{57}$$

where $\Delta/|J|$ indicates the sign of the determinant. This invariant contains thirty monomial terms, and it is not algebraically independent.

By counting the number of relations among the moments and the number of parameters involved, it can be shown

[7] For $p = 4$, (52) is the same as the one given by (50); instead of (52), $(I_{40}I_{02}^2 + I_{04}I_{20}^2)$ may be used.

[8] G. Salmon, "Modern Higher Algebra," Stechert, New York, N. Y., 4th ed., p. 188; 1885. (Reprinted 1924.)

that four is the largest number of independent invariants possible for this case. Various methods have been developed in finding algebraic invariants, and many invariants have been worked out in detail. In case extension to higher moment invariants are required, the known results for algebraic invariants will be of great help.

IV. Visual Information Processing and Recognition

A. *Pattern Characterization and Recognition*

Any geometrical pattern or alphabetical character can always be represented by a density distribution function $\rho(x, y)$, with respect to a pair of axes fixed in the visual field. Clearly, the pattern can also be represented by its two-dimensional moments, m_{pq}, with respect to the pair of fixed axes. Such moments of any order can be obtained by a number of methods. Using the relations between central moments and ordinary moments, the central moments can also be obtained. Furthermore, if these central moments are normalized in size by using the similitude moment invariants, then the set of moment invariants can still be used to characterize the particular pattern. Obviously, these are independent of the pattern position in the visual field and also independent of the pattern size.

Two different ways will be described in the next two sections to accomplish orientation independence. In these cases, theoretically, there exist either infinitely many absolute moment invariants or infinitely many normalized moments with respect to the principal axes. For the purpose of machine recognition, it is obvious that only a finite number of them can be used. In fact, it is believed only a few of these invariants are necessary for many applications. To illustrate this point, a simple simulation program, using only two absolute moment invariants, and its performance will be described in Section V.

B. *The Method of Principal Axes*

In (39) and (40), let $p = 2$, we have the following moment invariants,

$$\begin{aligned} (\mu_{20}' - \mu_{02}') - i2\mu_{11}' &= e^{i2\theta}[(\mu_{20} - \mu_{02}) - i2\mu_{11}], \\ (\mu_{20}' - \mu_{02}') + i2\mu_{11}' &= e^{-i2\theta}[(\mu_{20} - \mu_{02}) + i2\mu_{11}], \\ \mu_{20}' + \mu_{02}' &= \mu_{20} + \mu_{02}. \end{aligned} \tag{58}$$

If the angle θ is determined from the first equation in (58), to make $\mu_{11}' = 0$, then we have

$$\tan 2\theta = \frac{+2\mu_{11}}{\mu_{20} - \mu_{02}}. \tag{59}$$

The x', y' axes determined by any particular values of θ satisfying (59) are called the principal axes of the pattern. With added restrictions, such as $\mu_{20}' > \mu_{02}'$ and $\mu_{30}' > 0$, θ can be determined uniquely. Moments determined for such a pair of principal axes are independent of orientation. If this is used in addition to the method described in the last section, pattern identification can be made independently of position, size and orientation.

The discrimination property of the patterns is increased if higher moments are also used. The higher moments with respect to the principal axes can be determined with ease, if the invariants given by (39) and (40) are used. These relations are also useful in other ways. As an illustration, for $p = 3$ we have

$$\begin{aligned} &(\mu_{30}' - 3\mu_{12}') - i(3\mu_{21}' - \mu_{03}') \\ &\qquad = e^{i3\theta}[(\mu_{30} - 3\mu_{12}) - i(3\mu_{21} - \mu_{03})], \\ &(\mu_{30}' + \mu_{12}') - i(\mu_{21}' + \mu_{03}') \\ &\qquad = e^{i\theta}[(\mu_{30} + \mu_{12}) - i(\mu_{21} + \mu_{03})]. \end{aligned} \tag{60}$$

The two remaining relations, which are the complex conjugate of these two, are omitted here. If θ and the four third moments are known, the same moments with respect to the principal axes can be computed easily by using the above relations. There is no need of transforming the input pattern here.

In the above method, because of the complete orientation independence property, it is obvious that the numerals "6" and "9" can not be distinguished. If the method is modified slightly as follows, it can differentiate "6" from "9" while retaining the orientation independence property to a limited extent. The value of θ is still determined by (59), but it is also required to satisfy the condition $|\theta| < 45$ degrees. The use of third order moments in this case is also essential.

If the given pattern is of circular symmetry or of n-fold rotational symmetry, then the determination of θ by (59) breaks down. This is due to the fact that both the numerator and the denominator are zero for such patterns. As an example, assume that the pattern is of 3-fold rotational symmetry, *i.e.*, if the pattern is turned $2\pi/3$ radians about its centroid, it is identically the same as the original. In the first equation in (58), there are only two possible values for θ to make the imaginary part of $I_{20}' = 0$, *i.e.*, to make $\mu_{11}' = 0$. Under this symmetry requirement, there are more than two possible values to make the imaginary part of $I_{20}' = 0$; therefore, the only possibility is to have $I_{20}' \equiv 0$, and also $I_{20} \equiv 0$. In this 3-fold rotational symmetry case, the first equation in (60) can be used to determine the θ and the principal axes by requiring $3\mu_{21}' - \mu_{03}' = 0$. Based upon this example, we may state the following theorem.

Theorem: If a pattern is of n-fold rotational symmetry, than all the orthogonal invariants, I's, with the factor $e^{\pm iw\theta}$ and $w/n \neq$ integer must be identically equal to zero. For the limiting case of circular summetry, only I_{11}, I_{22}, $\cdots$ are not zero.

For patterns with mirror symmetry, a similar theorem may be derived.

C. *The Method of Absolute Moment Invariants*

The absolute orthogonal moment invariant described in Section III-D can be used directly for orientation independent pattern identification. If these invariants are combined with the similitude invariants of central moments, then pattern identification can be made independently of position, size and orientation. A specific example is given in Section V-A.

For the second and third order moments, we have the following six absolute orthogonal invariants:

$$\begin{aligned}
&\mu_{20} + \mu_{02}, \\
&(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2, \\
&(\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2, \\
&(\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2, \\
&(\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\
&\qquad + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03}) \\
&\qquad \cdot [3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2], \\
&(\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \\
&\qquad + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}),
\end{aligned} \tag{61}$$

and one skew orthogonal invariants,

$$\begin{aligned}
&(3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\
&\quad - (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2].
\end{aligned} \tag{62}$$

This skew invariant is useful in distinguishing mirror images.

This method can be generalized to accomplish pattern identification not only independently of position, size and orientation but also independently of parallel projection. In this generalization the general moment invariants are used instead of the orthogonal and similitude moment invariants.

V. Visual Pattern Recognition Models

A. The Simulation of a Simple Model

A simulation program of a simple pattern recognition model, using only two moment invariants, has been written for an LGP-30 computer. No information, properties, or features about the patterns to be recognized are contained in the simulation program itself; it learns. The visual field is a 16×16 matrix of small squares. A pattern is first projected onto the matrix and then each small square is digitalized to the values, 0, 2, 4, 6, or 8. After loading each pattern, the following two moment invariants[9]

$$\begin{aligned}
X &= \mu_{20} + \mu_{02} \\
Y &= \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}
\end{aligned} \tag{63}$$

are computed. The central moments, μ_{20}, μ_{11}, μ_{02} used above (normalized with respect to size) are obtained from the ordinary moments by (11). This point (X, Y) in a two dimensional space is used as a representation of the pattern.

[9] X and Y are, respectively, the sum and difference of the two second moments with respect to the principal axes, and may be interpreted as "spread" and "slenderness."

Assume the program or model has already learned a number of patterns, represented by (X_i, Y_i), $i = 1, 2, \cdots, n$, together with their names. If a new pattern is presented to the model, a new point (X, Y) and the distances

$$d_i = \sqrt{(X - X_i)^2 + (Y - Y_i)^2}, \quad i = 1, 2, \cdots, n \tag{64}$$

between (X, Y) and (X_i, Y_i) are computed. Let the minimum distance, $d_{\min}$, be defined as

$$d_{\min} = \min_i d_i. \tag{65}$$

The distance d_k satisfying $d_k = d_{\min}$ is selected (if more than one of the distances satisfying the condition, one d_k is selected at random). Then d_k is compared with a preselected recognition level L. 1) If $d_k > L$, the computer will type out "I do not know", then wait to learn the name of the new pattern. If a name is now entered, the computer then stores (X, Y) as (X_{n+1}, Y_{n+1}) together with the assigned name. Hence, a new pattern is learned by the program. 2) If $d_k \leq L$, the computer will identify the pattern with (X_k, Y_k) by typing out the name associated.

A very simple performance-improving program is also incorporated. When this program is used, it replaces the values of X_i, Y_i corresponding to the name now told, by

$$\frac{1}{\alpha}[(\alpha - 1)X_i + X], \quad \frac{1}{\alpha}[(\alpha - 1)Y_i + Y] \quad \alpha > 1. \tag{66}$$

This operation moves the point (X_i, Y_i) toward (X, Y).

B. Performance of the Simulation Program

Several experiments have been tried on the simulation program. For the convenience of description, two patterns are described as strictly similar, if one pattern can be transformed exactly into the other by a combination of translation, rotation and similitude transformations. In one experiment, patterns which are strictly similar after digitalization were fed to the program. If any one of such patterns is taught to the program just once, then it can identify correctly any other pattern of the same class. The number of different pattern classes capable of being learned is quite large, even with this simple program. There is no wrong identification except for specially constructed patterns.

Another experiment dealt with character recognition. A set of twenty-six capital letters from a 2-inch Duro Lettering Stencils were copied onto the 16×16 matrix and digitalized as inputs to the program. The values of X and Y, in arbitrary units, are given in Table I and Fig. 1, and two samples of the digitalized inputs for the letter M and W are shown in Fig. 2. The following may be noted:

1) Fig. 1 shows that the points for all the twenty-six letters are separated.

2) If inputs, prepared by using the same stencils but not strictly similar after digitalization are used, the corresponding points are not the same as those shown in Fig. 1. For a limited number of cases tried, the maximum

TABLE I

	X	Y		X	Y
A	6.2020	2.4986	N	5.7885	1.7933
B	6.1104	2.0853	O	8.2829	2.6246
C	10.4136	4.1818	P	7.0329	2.6456
D	8.2045	3.0911	Q	6.7674	1.9611
E	8.2147	4.3144	R	6.2707	1.9149
F	8.0390	4.5017	S	7.7501	3.3660
G	8.6096	3.0127	T	10.6216	7.1239
H	7.6243	1.1825	U	9.1728	2.1383
I	11.9780	11.2824	V	6.8761	3.2715
J	10.4118	6.6854	W	5.6667	0.1893
K	7.3278	2.5620	X	7.5714	3.5651
L	12.0662	8.3889	Y	8.3538	3.8612
M	5.7356	0.0540	Z	8.8843	5.1580

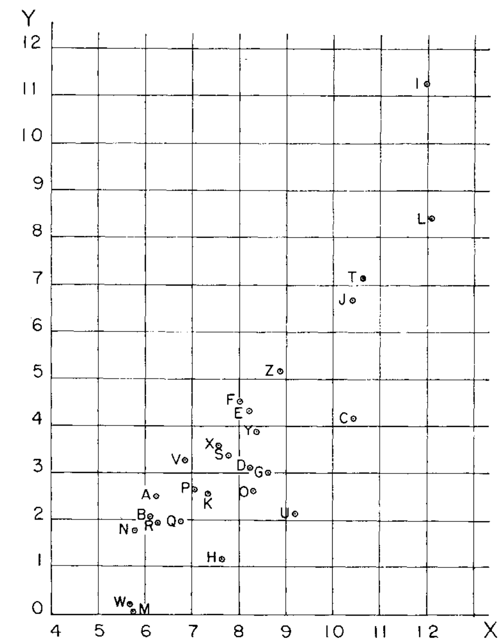

Fig. 1—Point representation of the twenty-six capital letters.

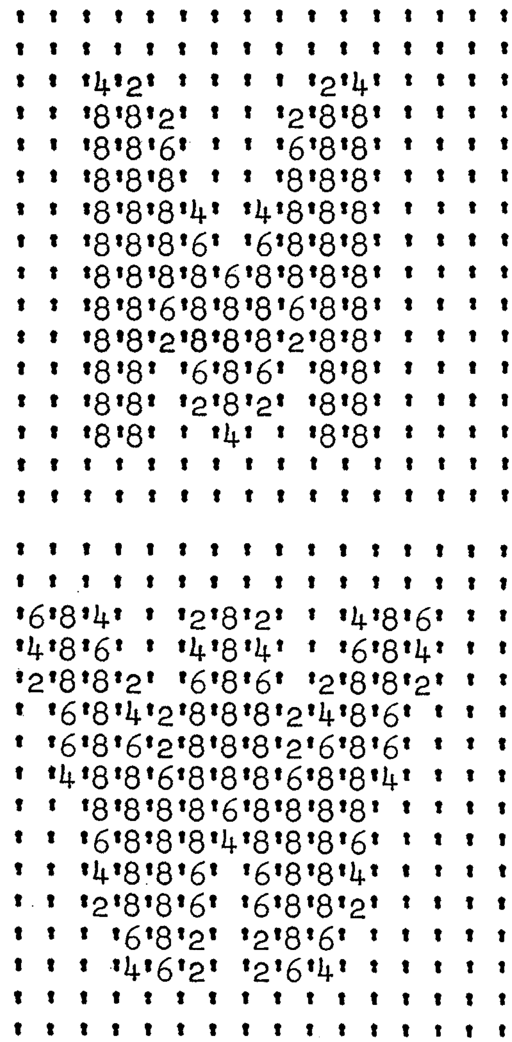

Fig. 2—Two samples of the digitalized inputs for the letters *M* and *W*.

variation in terms of distance between two points representing the same letter is of the order of 0.5. Compared with Fig. 1, it is obvious that overlapping of some classes will occur. If the resolution of the visual field is increased, the performance will definitely be improved.

3) In Fig. 1, it can be seen that some letters which are close to each other are of considerable difference in shape. A typical case is shown in Fig. 2, it is not difficult to conclude that the third order moments for the *M* and *W* examples shown will be considerably different.

From these results, it is clear that both the resolution and the number of invariants used should be increased but probably not greatly.

One additional experiment concerned the simple learning program. In this experiment, patterns belonging to the same class were generally represented by different points, clustered together, in the plane. As already described, a class represented by such a cluster was represented by a single point in this program, but this point together with the recognition level really form a circular recognition region for the class. For good performance, this region should be centered over the cluster of points representing the class. The point for the first sample of a class is not necessarily at the center of this region. Because of this fact, incorrect identifications may occur. The simple learning program, sometimes, is useful for such cases. If the clusters of points of different classes do not 'overlap,' generally, the program will improve the performance; otherwise, the performance may become worse. Another learning program will be described in the next section.

C. Other Visual Pattern Recognition Models

From the simulation program and the theoretical considerations described in IV, a considerably improved pattern recognition model is as follows: P absolute moment invariants or P normalized moments with respect to the principal axes, denoted by X^1, X^2, $\cdots$, X^P, are used; and the point $(X) \equiv (X^1, X^2, \cdots, X^P)$ in a P dimensional space is used as the representation for a pattern. It is believed that $P = 6$, (*i.e.*, using four more invariants related to the third order moments) and a 32×32 or

50×50 matrix as the visual field will be adequate for many purposes.

Let (X_i), $i = 1, 2, \cdots, n$ be the points representing the patterns already learned, and N_i be the number of samples of the ith pattern already learned. After each learning process for the ith pattern, N_i is replaced by $(N_i + 1)$, and (X_i) by

$$\left(\frac{N_i X_i + X}{N_i + 1}\right) \equiv \bar{X}_i \tag{67}$$

where (X) is the representation of the new sample. This new (X_i) is obviously equal to the average of all the $(N_i + 1)$ samples learned.

Instead of using a common recognition level, L, a separate one is determined for each pattern class in the learning process. After each sample is learned, L_i is replaced by the larger one of

$$L_i \quad \text{and} \quad \sqrt{\sum_{p=1}^{P} (\bar{X}_i^p - X^p)^2}. \tag{68}$$

The L_i thus determined, as the sample number increases, approaches the minimum radius of a hypersphere which includes most if not all the sample points in its interior. The center of the hypersphere is located at $(\bar{X}_i)$.

In this model, the following are stored for each class of patterns learned,

$$\text{Name}, \quad (X_i), L_i, N_i. \tag{69}$$

(X_i) and L_i form a spherical recognition region for the ith pattern. When a new pattern represented by (X) is entered, the distances

$$d_i = \sqrt{\sum_{p=1}^{P} (X_i^p - X^p)^2} \qquad i = 1, 2, \cdots, n \tag{70}$$

are computed. The distances d_j satisfying

$$d_j \leq L_j \tag{71}$$

are then selected. If no d_j is obtained, the pattern is considered as not yet learned, otherwise

$$D_j = \frac{d_j}{N_j} \tag{72}$$

is computed and

$$D_k = \min_j (D_j) \tag{73}$$

is selected, as in Section V-A, to identify the pattern. The use of N_i in the identification is believed to be useful when overlapping occurs.

If automatic input and digitalization equipment is used, there may be other types of noise introduced in addition to that due to digitalization. The well known local averaging process[10,11] can be used to reduce some of such noise, but the potential for discrimination possessed by such models is useful to combat whatever remains. In this connection, it seems worthwhile to point out the following two facts. 1) If two classes are separated, say, in two dimensions; they can never overlap when additional dimensions are introduced. 2) The use of moment invariants makes possible the derivation of models which may automatically generate additional dimensions—moment invariants—for the purpose of discrimination or combating noise.

The representation of a pattern by a point in a P dimensional space converts the problem of pattern recognition into a problem of statistical decision theory. Depending upon the particular decision method used, different statistical models may be devised. The work done by Sebestyen[12] is an example, his method can be used here directly.

The method of principal axes developed here has another application in connection with the statistical approaches mentioned at the beginning of this paper. It may be used as a preprocessor to normalize the inputs before the main processer is used. All the parameters necessary for translation, size and orientation normalizations can be obtained from some of the relations used in the method of principal axes. Such a preprocessor undoubtedly will increase the ability of the models based upon the statistical approach.

Acknowledgment

The author would like to express his deep appreciation to Dr. W. R. LePage for his constructive criticism and invaluable help during the preparation of this paper.

[10] G. P. Dineen, "Programming pattern recognition," *Proc. Western Joint Computer Conf.*, pp. 94–100; March, 1955.

[11] J. S. Bomba, "Alpha-numeric character recognition using local operations," *Proc. Eastern Joint Computer Conf.*, pp. 218–224; December, 1959.

[12] G. S. Sebestyen, "Recognition of membership in classes," IRE Trans. on Information Theory, vol. IT-7, pp. 44–50; January, 1961.