A Survey of Moment-Based Techniques

For Unoccluded Object Representation and Recognition

Richard J. Prokop
Anthony P. Reeves
School of Electrical Engineering
Cornell University
Ithaca NY

Abstract

The recognition of objects from imagery in a manner that is independent of scale, position, and orientation may be achieved by characterizing an object with a set of extracted invariant features. Several different recognition techniques have been demonstrated that utilize moments to generate such invariant features. These techniques are derived from general moment theory that is widely used throughout statistics and mechanics. In this paper, basic Cartesian moment theory is reviewed and its application to object recognition and image analysis is presented. The geometric properties of low-order moments are discussed along with the definition of several moment-space linear geometric transforms. Finally, significant research in moment-based object recognition is reviewed.

1. Introduction

The recognition of objects from imagery may be accomplished for many applications by identifying an unknown object as a member of a set of well-known objects. Various object recognition techniques utilize abstract characterizations for efficient object representation and comparison. Such characterizations are typically defined by measurable object features extracted from various types of imagery and any a priori knowledge available. Similarity between characterizations is interpreted as similarity between the objects themselves, therefore, the ability of a given technique to uniquely represent the object from the available information determines the effectiveness of the technique for the given application. Since no one representation technique will be effective for all recognition problems, the choice of object characterization is driven by the requirements and obstacles of a specific recognition task.

Several important issues may be identified that distinguish recognition tasks. One fundamental characteristic is whether or not the objects are occluded. In this paper, we are primarily interested in the class of tasks that involve strictly unoccluded (segmented) objects and, consequently, may be solved utilizing global feature techniques. Furthermore, many tasks require that objects be recognized from an arbitrary viewing position for a given aspect. This requirement necessitates the extraction of object features that are invariant to scale, translation, and/or orientation. The type of imagery will also determine the utility of a given representation technique. For example, techniques based solely on object boundaries or silhouettes may not be appropriate for applications where range imagery is collected. Another important issue is the presence of image noise and robustness of object features to such corruption. Finally, space and time efficiency of a representation technique is an issue for applications where the compactness of the object characterization and speed of classification is critical.

Research has been performed investigating the use of moments for object characterization in both invariant and non-invariant tasks utilizing 2-dimensional, 3-dimensional, range and/or intensity imagery. The principal techniques demonstrated include *Moment Invariants*, *Rotational Moments*, *Orthogonal Moments*, *Complex Moments*, and *Standard Moments*. Schemes for fast computation of image moments have been explored including optical, VLSI, and parallel architectures. Performance comparisons of the principal moment and other competing global-feature techniques have also been presented based on both theoretical analysis and experimental results.

The first section of this paper is a review of general moment concepts. The applicability of moments to object image analysis is presented along with a description of the geometric properties of the low order moment values. Several moment-space geometric transforms are also described. The following two sections are a survey of research exploring the principal moment techniques for object recognition (as outlined above). A brief explaination of each technique is presented along with subsequent improvements, applications and relationships to other techniques. In section 4, some novel techniques for the fast computation of moments are considered. Special purpose hardware and optical architectures are discussed. Section 5 is a summary of moment performance comparisons that have been performed.

2. Moment Theory

In general, moments describe numeric quantities at some distance from a reference point or axis. Moments are commonly used in statistics to characterize the distribution of random variables, and, similarly, in mechanics to characterize bodies by their spatial distribution of mass. The use of moments for image analysis is straightforward if we consider a binary or grey level image segment as a two-dimensional density distribution function. In this way, moments may be used to to characterize an image segment and extract properties that have analogies in statistics and mechanics.

2.1. Cartesian Moment Definition

The two-dimensional Cartesian moment, m_{pq}, of *order* $p+q$, of a density distribution function, $f(x,y)$, is defined as

$$m_{pq} \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x,y)\, dx\; dy \tag{2.01}$$

The two-dimensional moment for a $(N \times M)$ discretized image, $g(x,y)$, is

$$m_{pq} \equiv \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} x^p y^q \; g(x,y) \tag{2.02}$$

A complete moment set of order n consists of all moments, m_{pq}, such that $p+q \le n$ and contains $\frac{1}{2}(n+1)(n+2)$ elements. Note that the monomial product $x^p y^q$ is the *basis function* for this moment definition.

The use of moments for image analysis and object representation was inspired by Hu [1]. Hu's *Uniqueness Theorem* states that if $f(x,y)$ is piecewise continuous and has non-zero values only in the finite region of the (x,y) plane, then the moments of all orders exist. It can then be shown that the moment set $\{m_{pq}\}$ is uniquely determined by $f(x,y)$ and conversely, $f(x,y)$ is uniquely determined by $\{m_{pq}\}$. Since an image segment has finite area and, in the worst case, is piecewise continuous, moments of all orders exist and a moment set can be computed that will uniquely describe the information contained in the image segment. To characterize all of the information contained in an image segment requires a potentially infinite number of moment values. The goal is to select a meaningful subset of moment values that contain sufficient information to uniquely characterize the image for a specific application.

2.2. Properties of Low-Order Moments

The low-order moment values represent well-known, fundamental geometric properties of a distribution or body. To illustrate these properties and show the applicability to object representation, we can consider the moment values of a distribution function that is binary and contiguous, i.e. a silhouette image of a segmented object. The moment values for this distribution may be easily explained in terms of simple shape characteristics of the object.

2.2.1. Zeroth Order Moments : Area

The definition of the zeroth order moment, $\{m_{00}\}$, of the distribution, $f(x, y)$

$$m_{00} \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, dx\ dy \tag{2.03}$$

represents the total *mass* of the given distribution function or image. When computed for a silhouette image of a segmented object, the zeroth moment represents the total object area.

2.2.2. First Order Moments : Center of Mass

The two first order moments, $\{m_{10}, m_{01}\}$, are used to locate the *center of mass (COM)* of the object. The coordinates of the COM, $(\bar{x}, \bar{y})$, is the intersection of the lines, $x = \bar{x}$ and $y = \bar{y}$, parallel to the x and y axis respectively, about which the first order moments are zero. Alternatively, $x = \bar{x}$ and $y = \bar{y}$ represent lines where all the mass may be concentrated without change to the first order moments about the x and y axes respectively. In terms of moment values, the coordinates of the COM are

$$\bar{x} = \frac{m_{10}}{m_{00}} \qquad \bar{y} = \frac{m_{01}}{m_{00}} \tag{2.04ab}$$

The COM defines a unique location with respect to the object that may be used as a reference point to describe the position of the object within the field of view. If an object is positioned such that its COM is coincident with the origin of the field of view, i.e. $(\bar{x} = 0)$ and $(\bar{y} = 0)$, then the moments computed for that object are referred to as *central moments* and are designated by μ_{pq}. (Note that $\mu_{10} = \mu_{01} = 0$)

2.2.3. Second Order Moments

The second order moments, $\{m_{02}, m_{11}, m_{20}\}$, known as the *moments of inertia*, may be used to determine several useful object features. A description of each feature follows.

Principal Axes

The second order moments are used to determine the *principal axes* of the object. The principal axes may be described as the pair of axes about which there is the minimum and maximum second moment (major and minor principal axes respectively). In terms of moments, the orientation of the principal axes, ϕ, is given by

$$\phi = \frac{1}{2} \tan^{-1}\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right) \tag{2.05}$$

Note that in equation (2.05), ϕ is angle of the principal axis nearest to the x axis and is in the range $-\pi/4 \leq \phi \leq \pi/4$. The angle of either principal axis specifically may be determined from the specific values of μ_{11} and $(\mu_{20} - \mu_{02})$. Table 2.1 illustrates how the angle of the *major principal axis*, θ, may be determined by the second moments and the angle ϕ.

Table 2.1. Orientation of the Major Principal Axis.

μ_{11}	$\mu_{20} - \mu_{02}$	ϕ	θ
0	–	0	$+\pi/2$
+	–	$0 > \phi > -\pi/4$	$+\pi/2 > \theta > +\pi/4$
+	0	0	$+\pi/4$
+	+	$+\pi/4 > \phi > 0$	$+\pi/4 > \theta > 0$
0	0	0	0
–	+	$0 > \phi > -\pi/4$	$0 > \theta > -\pi/4$
–	0	0	$-\pi/4$
–	–	$+\pi/4 > \phi > 0$	$-\pi/4 > \theta > -\pi/2$

The angle of the principal axis of least inertia may be used as a unique reference axis to describe the object orientation within the field of view (*in-plane rotation*). Note that θ alone does not guarantee a unique orientation since a 180 degree ambiguity still exists. The third order central moments may be used to resolve this ambiguity (described below).

Image Ellipse

The first and second order moments also define an inertially equivalent approximation of the original image, referred to as the *image ellipse* [2]. The image ellipse is a constant intensity elliptical disk with the same mass and second order moments as the original image. If the image ellipse is defined with semi-major axis, α, along the x axis and semi-minor axis, β, along the y axis, then α and β may be determined from the second order moments using

$$\alpha = \left(\frac{2\left[\mu_{20} + \mu_{02} + \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}\,\right]}{\mu_{00}} \right)^{1/2} \tag{2.06a}$$

$$\beta = \left(\frac{2\left[\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2}\,\right]}{\mu_{00}} \right)^{1/2} \tag{2.06b}$$

The intensity of the image ellipse is then given by

$$I = \frac{\mu_{00}}{\pi \alpha \beta} \tag{2.07}$$

If we additionally require that all the moments through order two to be the same as the original image, we can center the ellipse about the image COM and rotate it by θ so that the major axis is aligned with the principal axis. the image ellipse for a silhouette image of a space shuttle is shown in figure 2.1.

Radii of Gyration

Another property that may be determined from the second order moments are the *radii of gyration (ROG)* of an image. The radius of gyration *about an axis* is the distance from the axis to a line where all the mass may be concentrated without change to the second moment about that axis. In terms of moments, The radii of gyration ROG_x and ROG_y about the x and y axes respectively are given by

$$ROG_x = \sqrt{\frac{m_{20}}{m_{00}}} \qquad ROG_y = \sqrt{\frac{m_{02}}{m_{00}}} \tag{2.08ab}$$

The radius of gyration *about the origin* is the radius of a circle centered at the origin where all the mass may be concentrated without change to the second moment about the origin. In terms of second order central moments, this value is given by

$$ROG_{com} = \sqrt{\frac{\mu_{20} + \mu_{02}}{\mu_{00}}} \tag{2.09}$$

The ROG_{com} has the property that it is inherently invariant to image orientation and, consequently, has been used as a rotationally invariant feature for object representation.

2.3. Moments of Projections

An alternative means of describing image properties represented by moments is to consider the relationship between the moments of an image segment and the moments of the projections of that image segment. Specifically, the moments in the sets $\{m_{p0}\}$ and $\{m_{0q}\}$ are equivalent to the moments of the image projection onto the x axis and y axis respectively. To illustrate this, consider the vertical projection, $v(x)$, of an image segment, $f(x, y)$, onto the x axis given by

$$v(x) = \int_{-\infty}^{\infty} f(x, y)\, dy \tag{2.10}$$

The one-dimensional moments, m_p, of $v(x)$ are then given by

$$m_p = \int_{-\infty}^{\infty} x^p\, v(x)\, dx \tag{2.11}$$

substituting (2.10) in (2.11) gives

$$m_p = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p f(x, y)\, dx\ dy = m_{p\,0} \tag{2.12}$$

The moment subsets corresponding to the x and y axis projections are shown in figure 2.2.

Now, if we consider the projection of an image segment onto an axis as a probability distribution, properties of central moments of an image segment may be described using classical statistical measures of this distribution. For example, the second central moments of a projection of an image segment onto the x axis are given by

$$\mu_{20} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 f(x, y)\, dx \tag{2.13}$$

which is proportional to the *variance* of the distribution.

2.4. Moments of Order Three and Greater

Moments of order three and greater are most easily described using properties of the projection of the image onto the x or y axis rather than properties of the image itself.

2.4.1. Third Order Moments : Projection Skewness

The two third order central moments, $\{\mu_{30}, \mu_{03}\}$, describe the *skewness* of the image projections. Skewness is a classical statistical measure of a distribution's degree of deviation from symmetry about the mean. The coefficient of skewness for image projections onto the x and y axes are given by

$$Sk_x = \frac{\mu_{30}}{\mu_{20}^{3/2}} \qquad Sk_y = \frac{\mu_{03}}{\mu_{02}^{3/2}} \tag{2.14ab}$$

The signs of the coefficients are an indication as to which side of an axis the projection is skewed as shown in table 2.2.

Table 2.2. Skewness of Projections based on signs of Sk_x and Sk_y.

Sk_x	X Projection Skewed
+	*left of y axis*
0	*symmetric about y axis*
–	*right of y axis*

Sk_y	Y Projection Skewed
+	*below x axis*
0	*symmetric about x axis*
–	*above x axis*

Note that $Sk_x = 0$ or $Sk_y = 0$ does not guarantee that the object is symmetric.

As mentioned previously, the third order moments may be used to resolve the 180 degree ambiguity of the principal axis rotation. This is based on the fact that the rotation of an image by 180 degrees changes the sign of the skewness of the projection on either axis. Additionally, the sign of the coefficient of skewness dependents only on the sign of μ_{30} or μ_{03} since μ_{20} and μ_{02} are always positive. Specifically, if the image is rotated by the negative of angle θ so that the major principal axis is coincident with the x axis, then the sign of μ_{30} may be used to distinguish between the two possible orientations.

2.4.2. Fourth Order Moments : Projection Kurtosis

Two of the fourth order central moments, $\{\mu_{40}, \mu_{04}\}$, describe the *kurtosis* of the image projections. Kurtosis is a classical statistical measurement of the "peakedness" of a distribution. The coefficient of kurtosis for projection of the image onto the x and y axes is given by

$$K_x = \frac{\mu_{40}}{\mu_{20}^2} - 3 \qquad K_y = \frac{\mu_{04}}{\mu_{02}^2} - 3 \tag{2.15ab}$$

A kurtosis of zero is the value for a Gaussian distribution, values less than zero indicate a flatter and less peaked distribution, while positive values indicate a narrower and more peaked distribution.

2.5. Transformations of Moments

In addition to providing a concise representation of fundamental image geometric properties, basic geometric transformations may be performed on the moment representation of an image. These transformations are more easily accomplished in the moment domain than the original pixel domain. A complete derivation of each of the following transforms may be found in [3].

2.5.1. Scale Transformation

A scale change of α in the x dimension and β in the y dimension of an image, $f(x, y)$, results in a new image, $f'(x, y)$, defined by

$$f'(x, y) = f(x/\alpha, y/\beta) \tag{2.16}$$

The transformed moment values $\{m'_{pq}\}$ are expressed in terms of the original moment values $\{m_{pq}\}$ of $f(x, y)$ as

$$m'_{pq} = \alpha^{1+p} \beta^{1+q} m_{pq} \qquad \alpha \neq \beta \tag{2.17}$$

$$m'_{pq} = \alpha^{2+p+q} m_{pq} \qquad \alpha = \beta \tag{2.18}$$

2.5.2. Translation Transformation

A translation of α in the x dimension and β in the y dimension of an image, $f(x, y)$, results in a new image, $f'(x, y)$, defined by

$$f'(x, y) = f(x-\alpha, y-\beta) \tag{2.19}$$

The transformed moment values $\{m'_{pq}\}$ are expressed in terms of the original moment values $\{m_{pq}\}$ of $f(x, y)$ as

$$m'_{pq} = \sum_{r=0}^{p} \sum_{s=0}^{q} \binom{p}{r} \binom{q}{s} \alpha^{p-r} \beta^{q-s} m_{rs} \tag{2.20}$$

2.5.3. Rotation Transformation

A rotation of θ about the origin of $f(x, y)$ results in a new image, $f'(x, y)$, defined by

$$f'(x, y) = f(x \cos\theta + y \sin\theta, -x \sin\theta + y \cos\theta) \tag{2.21}$$

The transformed moment values $\{m'_{pq}\}$ are expressed in terms of the moment values $\{m_{pq}\}$ of $f(x, y)$ as

$$m'_{pq} = \sum_{r=0}^{p} \sum_{s=0}^{q} \binom{p}{r} \binom{q}{s} (-1)^{q-s} (\cos\theta)^{p-r+s} (\sin\theta)^{q+r-s} m_{p+q-r-s,r+s} \tag{2.22}$$

Note that the transformed moments are a combination of the original moments of the same order or less.

2.5.4. Reflection Transformation

A reflection about the x axis of $f(x, y)$ results in a new image, $f'(x, y)$, defined by

$$f'(x, y) = f(-x, y) \tag{2.23}$$

The transformed moment values $\{m'_{pq}\}$ are expressed in terms of the original moment values $\{m_{pq}\}$ of $f(x, y)$ as

$$m'_{pq} = (-1)^p m_{pq} \tag{2.24}$$

The analogous result holds for reflection about the y axis. Note that reflection about an arbitrary axis is achieved by first rotating the reflection axis to be aligned with the x or y axis, performing the reflection, and then rotating the moments back to the original orientation.

2.5.5. Intensity Transformation

A uniform intensity (contrast) change α on $f(x, y)$ results in a new image, $f'(x, y)$, defined by

$$f'(x, y) = \alpha f(x, y) \tag{2.25}$$

The transformed moment values $\{m'_{pq}\}$ in terms of $\{m_{pq}\}$ are simply

$$m'_{pq} = \alpha\, m_{pq} \tag{2.26}$$

2.5.6. Discrete Convolution

The convolution of an image, $f(x, y)$, with a discrete $N \times M$ kernal, $w(i, j)$, may be considered the sum of a series of translations and scalings [4]. The convolved moment values $\{m'_{pq}\}$ are expressed in terms of the original moment values $\{m_{pq}\}$ of $f(x, y)$ as

$$m'_{pq} = \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \sum_{r=0}^{p} \sum_{s=0}^{q} \binom{p}{r} \binom{q}{s} w(i, j)\, \alpha^{p-r} \beta^{q-s} m_{rs} \tag{2.27}$$

$$\alpha = \left\lfloor \frac{N}{2} - i \right\rfloor \quad \beta = \left\lfloor \frac{M}{2} - j \right\rfloor \tag{2.28}$$

This may be rewritten as

$$m'_{pq} = \sum_{r=0}^{p} \sum_{s=0}^{q} \Omega_{pq}(r, s) m_{rs} \tag{2.29}$$

$$\Omega_{pq}(r, s) = \sum_{k=0}^{p-r} \sum_{l=0}^{q-s} \binom{p}{r} \binom{q}{s} \binom{p-r}{k} \binom{q-s}{l} (-1)^{k+1} \left\lfloor \frac{N}{2} \right\rfloor^{p-r-k} \left\lfloor \frac{M}{2} \right\rfloor^{q-s-l} \sum_{l=0}^{N-1} \sum_{j=0}^{M-1} i^k j^l w(i, j) \tag{2.30}$$

Note that for a given convolution kernal, w, the set of coefficients, Ω, need only be calculated once and may then be reapplied using equation (2.29).

3. Moment Techniques for Object Representation

Several techniques have been demonstrated that derive invariant features from moments for object representation. These techniques are distinguished by their moment definition, the type of image data exploited, and the method for deriving invariant values from the image moments. Various moment definitions are characterized by the choice of basis functions, which may be orthogonal or non-orthogonal polynomials, and the sampling of the image, which may be rectangular or polar. Moments have been defined for 2-dimensional (silhouette and boundary), 2½-dimensional (range), 3-dimensional, and grey-level (brightness) imagery. Most invariant characterizations achieve object scale and translation invariance through feature normalization since this is easily accomplished based on the low-order moments. The difficulty in achieving object rotation invariance has inspired much of the moment research.

Five principal moment-based invariant feature techniques may be identified from the research to date. The earliest method, *Moment Invariants*, is based on non-linear combinations of low-order two-dimensional Cartesian moments that remain invariant under rotation. Alternative moment definitions based on polar image representations, *Rotational Moments*, were also proposed as a solution for their simple rotation properties. Moment definitions utilizing uncorrelated basis functions, *Orthogonal Moments*, were developed to reduce the information redundancy that existed with conventional moments. Furthermore, orthogonal moments have more simply defined inverse transforms, and may be used to determine the minimum number of moments required to adequately reconstruct, and thus uniquely characterize, a given image. Related to orthogonal moments, *Complex Moments*, provide straightforward computation of invariant moments of an arbitrary order. Finally, *Standard Moments* are unique in that they achieve invariance completely through image feature normalization in the moment domain rather than relying on algebraic invariants.

3.1. Moment Invariants

The first significant work considering moments for pattern recognition was performed by Hu [1]. Hu derived relative and absolute combinations of moment values that are invariant with respect to scale, position, and orientation based on the theories of invariant algebra that deal with the properties of certain classes of algebraic expressions which remain invariant under general linear transformations.

Size invariant moments are derived from algebraic invariants but can be shown to be the result of a simple size normalization. Translation invariance is achieved by computing moments that have been translated by the negative distance to the centroid, thus normalized so that the center of mass of the distribution is at the origin (central moments). Hu recognized that rotation invariance was the most difficult to achieve and proposed two different methods for computing rotation invariant moments.

The first method, the method of *principal axes*, is based on the observation that moments may be computed relative to a unique set of principal axes of the distribution and will therefore be invariant to the orientation of the distribution. It was noted, however, that this method breaks down for rotationally symmetric objects, i.e. objects with no unique set of principal axes. Principal axes were utilized in early character recognition experiments performed by Giuliano, et.al [5]. However, very little research followed based on this method. The second proposed technique for rotation invariance is the method of *absolute moment invariants*. This technique, and its subsequent variations, proved to be basis for the majority of the moment research to date.

3.1.1. Two-Dimensional Moment Invariants

The method of moment invariants is derived from algebraic invariants applied to the moment generating function under a rotation transformation. The set of absolute moment invariants consists of a set of non-linear combinations of central moment values that remain invariant under rotation. Hu defines seven values, computed from central moments through order three, that are invariant to object scale, position, and orientation. In terms of the central moments, the seven moment invariants are given by

$$M_1 = \mu_{20} + \mu_{02} \tag{3.01a}$$

$$M_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \tag{3.01b}$$

$$M_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \tag{3.01c}$$

$$M_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \tag{3.01d}$$

$$\begin{aligned} M_5 = {} & (\mu_{30}-3\mu_{12})(\mu_{30}+\mu_{12})[(\mu_{30}+\mu_{12})^2-3(\mu_{21}+\mu_{03})^2] \\ & + (3\mu_{21}-\mu_{03})(\mu_{21}+\mu_{03})[3(\mu_{30}+\mu_{12})^2-(\mu_{21}+\mu_{03})^2] \end{aligned} \tag{3.01e}$$

$$M_6 = (\mu_{20}-\mu_{02})[(\mu_{30}+\mu_{12})^2-(\mu_{21}+\mu_{03})^2+4\mu_{11}(\mu_{30}+\mu_{12})(\mu_{21}+\mu_{03}) \tag{3.01f}$$

One skew invariant is defined to distinguish mirror images and is given by

$$\begin{aligned} M_7 = {} & (3\mu_{21}-\mu_{03})(\mu_{30}+\mu_{12})[(\mu_{30}+\mu_{12})^2-3(\mu_{21}+\mu_{03})^2] \\ & - (\mu_{30}-3\mu_{12})(\mu_{21}+\mu_{03})[3(\mu_{30}+\mu_{12})^2-(\mu_{21}+\mu_{03})^2] \end{aligned} \tag{3.01g}$$

It should be noted that, just as for the method of principal axes, this method breaks down for objects that are n-fold symmetric since the seven moment invariants for such an object are all zero.

Hu demonstrated the utility of moment invariants through a simple pattern recognition experiment. The first two moment invariants were used to represent several known digitized patterns in a two-dimensional feature space. An unknown pattern could be classified by computing its first two moment values and finding the minimum Euclidean distance between the unknown and the set of well-known pattern representations in feature space. If the minimum distance was not within a specified threshold, the unknown pattern was considered to be of a new class, given an identity, and added to the known patterns. A similar experiment was performed using a set of twenty-six capital letters as input patterns. When plotted in two-dimensional space, all the points representing each of the characters were distinct. It was observed, however, that some characters that were very different in image shape were close to each other in feature space. In addition, slight variations in the input images of the same character resulted in varying feature values that in turn lead to overlapping of closely spaced classes. Hu concluded that increased image resolution and a larger feature space would improve object distinction.

3.1.2. Three-Dimensional Moment Invariants

Sadjadi and Hall [6] have extended Hu's two-dimensional moment invariants to objects defined in three dimensional space. The definition of three dimensional moments is given by

$$m_{pqr} \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p \; y^q \; z^r \; f(x, y, z) \, dx \; dy \; dz \tag{3.02}$$

Using the theory of invariant algebra and properties of ternary quantics, Sadjadi and Hall presented a derivation of moment invariants that are analogous to Hu's two-dimensional moment invariants. Three relative moment invariants values are derived from second order central moments and are given by

$$J_1 = \mu_{200} + \mu_{020} + \mu_{002} \tag{3.03a}$$

$$J_2 = \mu_{020}\mu_{002} - \mu_{011}^2 + \mu_{200}\mu_{002} - \mu_{101}^2 + \mu_{200}\mu_{020} - \mu_{110}^2 \tag{3.03b}$$

$$\Delta_2 = \det \begin{bmatrix} \mu_{200} & \mu_{110} & \mu_{101} \\ \mu_{100} & \mu_{020} & \mu_{011} \\ \mu_{101} & \mu_{011} & \mu_{002} \end{bmatrix} \tag{3.03c}$$

Two absolute moment invariants are then defined by

$$I_1 = \frac{J_1^2}{J_2} \qquad I_2 = \frac{\Delta_2}{J_1^3} \tag{3.04ab}$$

Experiments were conducted to confirm the invariance of these values. Three-dimensional moment invariants were calculated for a rectangular solid, a cylinder, and a pyramid in several different orientations. The computed values were shown to be invariant for each object.

3.1.3. Boundary Moment Invariants

Dudani, Breeding, and McGhee [7] applied moment invariants to a model-based three-dimensional object recognition system. The system was developed to perform automatic classification of aircraft from television images using moment invariant feature vectors computed from silhouette and boundary information. Calculation of the moment invariants was based on Hu's seven invariants with the exception of size normalization. Size normalization was based on the object to sensor distance and the radius of gyration of the object. It was claimed that high frequency details in the image are best characterized by moments derived from the object boundary while overall shape characteristics are best represented by silhouette moments. Moment invariants were therefore calculated for both the silhouette and the boundary of each object to create a feature vector. Object classification was based on a distance-weighted k-nearest-neighbor rule between the object feature vector and all the feature vectors of the model database. Their results showed the moment based classification to be more accurate than several qualified human observers.

Sluzek [89] proposed a method for using moment invariants to identify objects from local boundaries. If the object boundary is represented by a closed curve, $x(t)$ and $y(t)$, a fragment of this curve may be specified by a starting point, $t = \alpha$, and a length, β. The moment definition for this fragment is then

$$m_{pq}(\alpha,\beta) = \int_{\alpha}^{\alpha+\beta} x(t)^p \; y(t)^q \left(\frac{dx^2}{dt^2} + \frac{dy^2}{dt^2} \right)^{1/2} dt \tag{3.05}$$

The basis for Sluzek's technique is the notion that these moments and, subsequently, moment invariants derived from these moments, are continuous functions of α and β and that these functions may be determined for each object. A complete object is then represented by analytical descriptions of the functions of the first two moment invariants designated by $I_1(\alpha, \beta)$ and $I_2(\alpha, \beta)$. To determine a match between the moment invariants of a fragment, I'_1 and I'_2, and an object, one attempts to solve the following system of equations for α and β

$$I'_1 = I_1(\alpha, \beta) \qquad I'_2 = I_2(\alpha, \beta) \tag{3.06ab}$$

The existence of a solution indicates a match. Additionally, the determined α and β indicates which segment of the object boundary matched the fragment. Sluzek, however, identifies that analytic descriptions of the moment invariants $I_1(\alpha, \beta)$ and $I_2(\alpha, \beta)$ are complex and a unique solution to equations (3.06a) and (3.06b) is not guaranteed.

3.1.4. Other Applications of Moment Invariants

Gilmore and Boyd [10] utilized Hu's seven moment invariants to identify well-known building and bridge targets with infrared imagery. In their application, the orientation and range of the image sensor was known so the expected shape and size of the target could be calculated based on a target model. First, the scene was segmented and thresholded into several silhouette regions. A preprocessing step was then used to disqualify regions that greatly differ from the expected target. The seven moment invariants were then computed from silhouettes of each of the potential target regions. Since the sensor to scene geometry was known, the actual region area was determined and used for size normalization. Classification was based on a weighted difference between the region moments and the expected target moments. Correct classification of targets was demonstrated with this technique.

Sadjadi and Hall [11] investigated the effectiveness of moment invariants for scene analysis. Through a simple experiment, they showed that moment theory was consistent with empirical results when applied to grey-level imagery. The moment values were computed from a grey-level image subject to various size and rotation transformations. The seven invariant values were found to be similar for all the transformed images.

Wong and Hall [12] used moment invariants to match radar images to optical images. Square sub-regions of the optical image were compared to sub-regions in the radar image using a correlation based on the log of the moment values. The log was used to reduce the dynamic range of the moment values. The moment invariants were shown to be useful features for matching the images, however, it was assumed that radar and optical images were of the same scale and orientation.

3.1.5. Alternative Moment Invariant Techniques

Maitra [13] presented a variation of Hu's moment invariants that are additionally invariant to contrast change. These new moments are also inherently size invariant and thus do not require size normalization. In terms of Hu's moment invariants, Maitra's invariants are defined by

$$\beta_1 = \frac{\sqrt{M_2}}{M_1} \tag{3.07a}$$

$$\beta_2 = \frac{M_3\,\mu_{00}}{M_1\,M_2} \tag{3.07b}$$

$$\beta_3 = \frac{M_4}{M_3} \tag{3.07c}$$

$$\beta_4 = \frac{\sqrt{M_5}}{M_4} \tag{3.07d}$$

$$\beta_5 = \frac{M_6}{M_1\,M_4} \tag{3.07e}$$

$$\beta_6 = \frac{M_4}{M_3} \tag{3.07f}$$

Maitra demonstrated moment invariance with two digitized images of the same scene each taken with a different camera position to provide a difference in scale, illumination, position, and rotation. The six invariants are computed for each image and compared. Maitra claimed that the variation in invariant values is an improvement over previous results.

Abo-Zaid, Hinton, and Horne [14] also suggest a variation of Hu's moment invariants by defining a new moment normalization that is used to cancel scale and contrast changes before the computation of the moment invariants. In terms of central moments, the new normalization factor is defined by

$$\mu'_{pq} = \mu_{pq} \frac{1}{\mu_{00}} \left(\frac{\mu_{00}}{\mu_{20} + \mu_{02}} \right)^{\frac{p+q}{2}} \tag{3.08}$$

Abo-Zaid, et.al. claim that in addition to being position, contrast, and size invariant, these moments have decreased dynamic range when compared to moments that have been size normalized using equation (2.18). Decreased dynamic range allows higher order moments to be represented without resorting to logarithmic representation and without loss of accuracy.

3.2. Rotational Moments

Rotational moments are an alternative to the conventional Cartesian moment definition. These moments are based on a polar coordinate representation of the image and have well defined rotation transform properties. The (complex) rotational moment D_{nl} of order n is defined by [15]

$$D_{nl} = \int_0^{2\pi} \int_0^{\infty} r^n \, e^{il\theta} f(r,\theta) \, r \, dr \, d\theta \qquad |l| \le n \qquad n - l = even \tag{3.09}$$

Rotational moments may be derived from conventional moments by

$$D_{nl} = \sum_{j=0}^{\frac{1}{2}(n-l)} \sum_{k=0}^{l} (-i)^k \binom{\frac{1}{2}(n-l)}{j} \binom{l}{k} m_{n-l+k-2j,\, l-k+2j} \qquad 0 \le l \le n \tag{3.10}$$

To illustrate the simplicity of a rotation transformation, consider an image, $f(r,\theta)$, rotated by an angle ϕ. The transformed rotational moments are defined by

$$D_{nl} = \int_0^{2\pi} \int_0^{\infty} r^n \, e^{il\theta} f(r, (\theta-\phi)) \, r \, dr \, d\theta \tag{3.11}$$

In terms of the original rotational moments, the transformed moments are

$$D'_{nl} = e^{(il\phi)} D_{nl} \tag{3.12}$$

A rotation of ϕ is thus achieved by a phase change of the rotational moments. Another transform easily accomplished with rotational moments is dilatation or radial scale change. In terms of the original radial moments, a radial scale of α results in transformed moments given by

$$D'_{nl} = \alpha^{n+2} D_{nl} \tag{3.13}$$

Intensity (contrast) change is also easily defined. In terms of the original radial moments, an intensity change of β results in transformed moments given by

$$D'_{nl} = \beta \, D_{nl} \tag{3.14}$$

Rotational moments, however, have complicated translation transformations. Consequently, rotational moment techniques typically rely on Cartesian moments to find the center of mass and then compute the rotational moments about that point. (*i.e. central rotational moments*)

3.2.1. Rotational Moment Invariants

Smith and Wright [16] used a simplified rotational moment technique to derive invariant features for characterizing noisy, low resolution images of ships. The given image function, $f(x, y)$, was considered in polar coordinates with the polar origin located at the image COM, $(\bar{x}, \bar{y})$, to provide position invariance. Two new moment values $\hat{C}_{nl}$ and $\hat{S}_{nl}$ were defined as

$$\hat{C}_{nl} = \int\int r^n \cos l\theta \, f(r,\theta) \, r \, dr \, d\theta \tag{3.15}$$

$$\hat{S}_{nl} = \int\int r^n \sin l\theta \, f(r,\theta) \, r \, dr \, d\theta \tag{3.16}$$

These moment definitions are the real-valued parts of the rotational moments.

Intensity invariance was achieved by normalizing the moment values with the zeroth order moment m_{00}. Rotation invariance was achieved by measuring θ relative to the angle of the principal axis θ_p. The resulting invariant moments were given by

$$C_{nl} = \frac{\hat{S}_{nl} \sin l\theta_p + \hat{C}_{nl} \cos l\theta_p}{m_{00}} \tag{3.17}$$

$$S_{nl} = \frac{\hat{S}_{nl} \cos l\theta_p - \hat{C}_{nl} \sin l\theta_p}{m_{00}} \tag{3.18}$$

which are the real-valued rotational moments rotated through angle θ_p.

Polynomials of these moments, through order three, derived using a linear regression technique, were used to estimate the length and aspect ratio of the ship for classification. Although moments through order five were used, it was observed that moments through order three were most useful as they were less sensitive to noise.

Boyce and Hossack [17] derived rotational moments of arbitrary order that are invariant to rotation, radial scaling (dilatation), and intensity change. Based on the rotation transform for rotational moments, as given in equation (2.22), it follows that the product of rotational moments

$$\prod_i D(n_i, l_i) \text{ for } which \sum_i l_i = 0$$

will be invariant under rotation. (Note that $D(n, l) = D_{nl}$) Dilatation invariance is achieved by choosing quotients of the above products such that the sum

$$\sum_i (n_i + 2)$$

is the same for the numerator and denominator, thus canceling out the radial scale factor. Finally, intensity invariance is achieved by ensuring that the number of terms in the numerator and denominator are equal. These rotational moment invariants are defined in terms of rotational moments, $D(n, l)$, for a given order, n, with $|l| \le n$ and $n - l = even$.

For n $even$, the moment invariants are given by

$$\frac{D(n, n-2m)\, D(n, -n+2m)}{D(n, 0)^2} \qquad 0 \le m < \tfrac{1}{2}(n-2) \tag{3.19a}$$

$$\frac{D(n, n-2m)\, D(n-2, -n+2m)}{D(n, 0)\, D(n-2, 0)} \qquad 2 \le m < \tfrac{1}{2}(n-2) \tag{3.19b}$$

$$\frac{D(n, 0)\, D(0, 0)}{D(n-2, 0)\, D(2, 0)} \tag{3.19c}$$

$$\frac{D(n, n)\, D(n-2, -n+2)\, D(2, -2)}{D(n, 0)\, D(n-2, 0)\, D(2, 0)} \tag{3.19d}$$

and for n odd, the invariants are

$$\frac{D(n, n-2m)\, D(n, -n+2m)\, D(0, 0)}{D(n-1, 0)^2\, D(2, 0)} \qquad 0 \le m \le \tfrac{1}{2}(n-1) \tag{3.19e}$$

$$\frac{D(n, n-2m)\, D(n-2, -n+2m)}{D(n-1, 0)^2} \qquad 1 \le m \le \tfrac{1}{2}(n-1) \tag{3.19f}$$

$$\frac{D(n, n)\, D(n-2, -n+2)\, D(2, -2)}{D(n-1, 0)^2\, D(2, 0)} \tag{3.19g}$$

Two special-case definitions are provided for the last two invariants when $n = 3$. In these invariants, the term, $D(n-2, l)$, will always be zero causing the invariant to always evaluate to zero. The special invariant definitions are given by

$$\frac{D(3, 1)^2\, D(2, -2)\, D(0, 0)}{D(2, 0)^4} \tag{3.19h}$$

$$\frac{D(3, 3)\, D(3, 1)\, D(2, -1)\, D(0, 0)}{D(2, 0)^4} \tag{3.19i}$$

3.2.2. Radial and Angular Moment Invariants

Reddi [18] presented an alternative formulation of moment invariants based on the image representation in polar coordinates. The definition of the radial and angular central moments is given by

$$\psi_r(k, f) = \int_0^{\infty} r^k f(r, \theta) dr \tag{3.20}$$

$$\psi_\theta(p, q, f) = \int_{-\pi}^{\pi} \cos^p\theta \sin^q\theta \, f(r, \theta) d\theta \tag{3.21}$$

$$\psi(k, p, q, f) = \int_{-\pi}^{\pi} \int_0^{\infty} r^k \cos^p\theta \sin^q\theta \, f(r,\theta) dr d\theta \tag{3.22}$$

$$\mu_{pq} = \psi(p+q+1, p, q, f) \tag{3.23}$$

Hu's moment invariants based on radial and angular moments of order three are defined as follows

$$M_1 = \psi_r(3, \psi_\theta(f)) \tag{3.24a}$$

$$M_2 = \left| \psi_r(3, \psi_\theta(fe^{j2\theta})) \right|^2 \tag{3.24b}$$

$$M_3 = \left| \psi_r(4, \psi_\theta(fe^{j3\theta})) \right|^2 \tag{3.24c}$$

M_4 is derived from Hu's moment invariant for illustration

$$M_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2$$

$$= \left[\psi(4, 3, 0, f) + \psi(4, 1, 2, f)\right]^2 + \left[\psi(4, 2, 1, f) + \psi(4, 0, 3, f)\right]^2$$

$$= \left\{ \int_{-\pi}^{\pi} \int_0^{\infty} r^4 \cos\theta(\cos^2\theta + \sin^2\theta) f(r,\theta) dr d\theta \right\}^2 + \left\{ \int_{-\pi}^{\pi} \int_0^{\infty} r^4 \sin\theta(\cos^2\theta + \sin^2\theta) f(r,\theta) dr d\theta \right\}^2$$

$$= \left\{ \int_{-\pi}^{\pi} \int_0^{\infty} r^4 \cos\theta f(r,\theta) dr d\theta \right\}^2 + \left\{ \int_{-\pi}^{\pi} \int_0^{\infty} r^4 \sin\theta f(r,\theta) dr d\theta \right\}^2$$

$$= \left[\psi_r(4, \psi_\theta(1, 0, f))\right]^2 + \left[\psi_r(4, \psi_\theta(0, 1, f))\right]^2$$

$$= \left| \psi_r(4, \psi_\theta(1, 0, f)) + j\psi_r(4, \psi_\theta(0, 1, f)) \right|^2$$

$$= \left| \int_{-\pi}^{\pi} \int_0^{\infty} r^4 (\cos\theta + j\sin\theta) f(r, \theta) dr d\theta \right|^2$$

$$= \left| \psi_r(4, \psi_\theta(fe^{j\theta})) \right|^2 \tag{3.24d}$$

$$M_5 = \mathrm{Re}[\psi_r(4, \psi_\theta(fe^{j3\theta})) \cdot \psi_r^3(4, \psi_\theta(fe^{-j\theta}))] \tag{3.24e}$$

$$M_6 = \mathrm{Re}[\psi_r(3,\ \psi_\theta(fe^{j2\theta}))\cdot\psi_r^2(4,\ \psi_\theta(fe^{-j\theta}))] \tag{3.24f}$$

$$M_7 = \mathrm{Im}[\psi_r(4,\ \psi_\theta(fe^{j3\theta}))\cdot\psi_r^3(4,\ \psi_\theta(fe^{-j\theta}))] \tag{3.24g}$$

A general form of radial and angular invariant is presented as

$$I_{kl} = \left|\psi_r(k\,,\ \psi_\theta(e^{jl\theta}))\right|^2 \quad \text{for } any\ k \text{ and } l \tag{3.25}$$

Using this definition, absolute moment invariants can be derived without the use of algebraic invariants. Furthermore, it is noted that if the image, $f(r,\theta)$, is radially scaled by a factor α, the resulting radial moments are given by

$$\psi_r(k\,,\ f(\alpha r,\theta)) = \alpha^{-(k+1)}\ \psi_r(k\,,\ f(r,\theta)) \tag{3.26}$$

This allows size invariant moments to be derived by choosing fractions of radial moments that cancel α.

Yin and Mack [19] compared the effectiveness of radial and angular moment invariants with Hu's (Cartesian) moment invariants for object classification from both silhouette and grey-level imagery. Moment based feature vectors were computed for objects from video and FLIR imagery. Classification was based on a weighted k-nearest neighbor approach. They found that both moment techniques provided similar results. It was observed, however, Hu's moment invariants require less computation time than radial and angular moments.

3.3. Orthogonal Moments

Teague [2] presented two inverse moment transform techniques to determine how well an image could be reconstructed from a small set of moments. The first method, *moment matching*, derives a continuous function

$$g(x\,,y) = g_{00} + g_{10}\,x + g_{01}\,y + g_{20}\,x^2 + g_{11}\,xy + g_{02}\,y^2 + \cdots \tag{3.27}$$

whose moments exactly match the moments, $\{m_{pq}\}$, of $f(x\,,y)$ through order n. However, this method is shown to be impractical as it requires the solution to an increasing number of coupled equations as higher order moments are considered.

The second method for determining an inverse moment transform is based on *orthogonal moments*. Teague observed that the Cartesian moment definition

$$m_{pq} \equiv \int\int x^p y^q f(x\,,y)\ dx\ \ dy \tag{3.28}$$

has the form of the projection of $f(x\,,y)$ onto the non-orthogonal, monomial basis set, $x^p y^q$. Replacing the monomials with an orthogonal basis set (e.g. Legendre and Zernike polynomials), results in an orthogonal moment set with an approximate inverse moment transform.

3.3.1. Legendre Moments

The Legendre polynomials, $P_n(x)$, are defined by

$$P_n(x) = \frac{1}{2^n} \sum_{m=0}^{n/2} (-1)^m \frac{(2n-2m)!}{m!\,(n-m)!\,(n-2m)!} x^{n-2m} \tag{3.29}$$

or more simply

$$P_n(x) = \sum_{k=0}^{n} C_{nk}\, x^k \tag{3.30}$$

where the Legendre coefficients, C_{nk}, are given by

$$C_{nk} = (-1)^{(n-k)/2} \frac{1}{2^n} \frac{(n+k)!}{[(n-k)/2]!\,[(n+k)/2]!\,k!} \qquad n-k = even \tag{3.31}$$

The Legendre polynomials are orthogonal over the interval $-1.0 \le x \le 1.0$.

The nature of the monomial basis functions and Legendre polynomials are illustrated in figures 3.1-3.3. In figure 3.1, monomials up to order 5 are shown for the interval $-100 \le x \le 100$. These monomials increase very rapidly in range as the order increases, but they do have the advantage that simple integer data representation may be used with discrete digitized imagery. Figure 3.2 shows the monomials up to order 5 for $-1.0 \le x \le 1.0$. The range of the monomials is now $-1.0 \le f(x) \le 1.0$, however, a precision problem still remains and a floating point or scaled data format is necessary. The monomials are highly correlated and the important information is contained within the small differences between them. As the order increases, the precision needed to accurately represent these differences also rapidly increases. The Legendre polynomials through order 5 for $-1.0 \le x \le 1.0$ are shown in figure 3.3. Since these polynomials are orthogonal over this range, less precision is needed to represent differences between the polynomials to the same accuracy as the monomials.

Teague utilized Legendre polynomials $P_n(x)$ as a moment basis set and defined the orthogonal Legendre moment, L_{pq}, as

$$L_{pq} = \frac{(2p+1)(2q+1)}{4} \int_{-1}^{1}\int_{-1}^{1} P_p(x)\, P_q(y)\, f(x,y)\, dx\, dy \tag{3.32}$$

Note that for the moments to be orthogonal, the image $f(x,y)$ must be scaled to be within the region $-1.0 \le x, y \le 1.0$.

If the Legendre polynomials are expressed in terms of their coefficients, C_{nk}, then the relationship between conventional and Legendre moments is defined by

$$L_{pq} = \frac{(2p+1)(2q+1)}{4} \sum_{r=0}^{p} \sum_{s=0}^{q} C_{pr}\, C_{qs}\, \mu_{rs} \tag{3.33}$$

Teague derived a simple approximation to the inverse transform for a set of moments through order N given by

$$f(x, y) \approx \sum_{n=0}^{N} \sum_{m=0}^{n} P_{n-m}(x)\, P_m(y)\, L_{n-m,m} \tag{3.34}$$

The Legendre based inverse transform has an advantage over the method of moment matching in that there are no coupled algebraic equations to solve. Furthermore, the Legendre moments are easily computed from the conventional moments and the well-defined polynomial coefficients.

Teague performed image reconstruction on increasing order moment sets (through 15th order) and computed the pixel error between the original and reconstructed images. It was found that the pixel error image steadily decreased as higher order moments were used. Teague demonstrated that higher order moments (greater than order three) contain significant information and may be necessary to sufficiently characterize an image for a given application. He notes, however, that although higher order moments may be required, the set of moment values is still small when compared to the pixel representation of the image.

Reeves and Taylor [4] identify that the problem of perfectly reconstructing a binary valued, discretely sampled image directly using the method described above by Teague is difficult, since the original image violates the necessary continuity assumptions. Consequently, even increasing the order of the moment set used in the reconstruction will not guarantee a good result. In an effort to compensate for this problem, an iterative scheme using error feedback was devised to help reconstruct silhouette images.

The fundamental approach in the iterative scheme was based on the fact that the moment transform is a linear operation. For example, moments of the difference of two images are the same as the difference of the two images moments. Using this property, an error image can be constructed from the moment set error, and then subtracted from the current reconstruction to enhance its accuracy. When compared with Teague's approach, the iterative scheme demonstrated substantially improved results. Results for simple geometric shapes indicate that moment sets as small as order 4 may produce good reconstructions. In general, objects showed their best results for 12th order moment reconstructions. Additionally, complex shapes required lower feedback than simpler shapes, to produce stable iterations that would result in an improved image reconstruction.

3.3.2. Zernike Moments

To derive orthogonal, rotationally invariant moments, Teague used the complex Zernike polynomials as the moment basis set. The Zernike polynomials, $V_{nl}(x, y)$, of order n, are defined by

$$V_{nl}(x, y) = R_{nl}(r)\, e^{il\theta} \qquad 0 \le l \le n \quad n - l = even \tag{3.35}$$

where the real-valued radial polynomial is given by

$$R_{nl}(r) = \sum_{m=0}^{(n-l)/2} (-1)^m \frac{(n-m)!}{m!\,[(n-2m+l)/2]!\,[(n-2m-l)/2]!}\, r^{n-2m} \tag{3.36}$$

or more simply

$$R_{nl}(r) = \sum_{k=l}^{n} B_{nlk}\, r^k \tag{3.37}$$

where the Zernike coefficients, B_{nlk}, are given by

$$B_{nlk} = (-1)^{(n-k)/2} \frac{[(n+k)/2]!}{[(n-k)/2]!\,[(k+l)/2]!\,[(k-l)/2]!} \qquad n - k = even \tag{3.38}$$

The Zernike polynomials are orthogonal within the unit circle $x^2 + y^2 = 1$. Figure 3.4 shows the Zernike polynomials through order 5 in the interval $0.0 \le r \le 1.0$ for various values of l. Notice that these polynomials have desirable dynamic range characteristics but become more correlated as the radius approaches 1.

The complex Zernike moment Z_{nl} is defined as

$$Z_{nl} = \frac{(n+1)}{\pi} \int_0^{2\pi}\int_0^{\infty} V_{nl}(r, \theta)^*\, f(r, \theta)\, r\, dr\, d\theta \tag{3.39}$$

where * indicates the complex conjugate. Note that for the moments to be orthogonal, the image must be scaled to be within a unit circle centered at the origin.

Zernike moments may be derived from conventional moments μ_{pq} by

$$Z_{nl} = \frac{(n+1)}{\pi} \sum_{k=l}^{n} \sum_{j=0}^{q} \sum_{m=0}^{l} (-i)^m \binom{q}{j} \binom{l}{m} B_{nlk}\, \mu_{k-2j-l+m,\, 2j+l-m} \tag{3.40}$$

Zernike moments may be more easily derived from rotational moments [2] , D_{nk}, by

$$Z_{nl} = \sum_{k=l}^{n} B_{nlk}\, D_{nk} \tag{3.41}$$

An approximate inverse transform for a set of moments through order N is given by

$$f(x, y) \approx \sum_{n=0}^{N} \sum_{l} Z_{nl}\, V_{nl}(x, y) \tag{3.42}$$

To illustrate the rotational properties of Zernike moments, Teague showed that a distribution, $f(r, \theta)$, rotated through an angle ϕ, results in the transformed moments

$$Z'_{nl} = \frac{(n+1)}{\pi} \int_0^{2\pi}\int_0^{\infty} R(r)_{nl}\, e^{(-il}\phi) f(r, \theta - \phi)\, r\, dr\, d\theta \tag{3.43}$$

which is equivalent to

$$Z'_{nl} = Z_{nl}\ e^{-il\phi} \tag{3.44}$$

Under a rotation transformation, the angle of rotation of the Zernike moments is simply a phase factor. Like rotational moments, however, the disadvantage of Zernike moments is the complex translation transformation.

Boyce and Hossack [17] demonstrated the effectiveness of image construction using Zernike moments. A 64×64 by 256 grey-level image was reconstructed using Zernike moments of increasing order. The normalized squared error between the original, $f(x, y)$, and then reconstructed, $f'(x, y)$, images was computed using

$$error = \frac{\sum_{x,y} \left| f(x, y) - f'(x, y) \right|^2}{\sqrt{\sum_{x,y} f(x, y)^2 \sum_{x,y} f'(x, y)^2}} \tag{3.45}$$

It was shown that Zernike moments of order 6 were sufficient to reconstruct the image with an error of 10%. Utilizing Zernike moments through order 20 resulted in a reconstruction with an error of 6%.

Khotanzad and Hong [2021] present a set of rotationally invariant features based on the magnitudes of the Zernike moments. As shown in (3.44), the rotation of Zernike moments only causes a phase shift. Therefore, the magnitudes of the Zernike moments remain invariant under rotation.

To determine the order of Zernike moment required for object representation, increasing order moments were used to reconstruct an object until the error between the original and reconstructed object images was below a preselected threshold. The Hamming distance was used as the dissimilarity measure. Additionally, this technique can be used to identify the contribution of the ith order moments to object representation. It was shown that the information content of the moments may be inferred by comparing the reconstructions from moments inclusive and exclusive of a specific moment order.

Experimental results demonstrated 99% recognition rate on a set of 24 English characters using 23 Zernike features and nearest neighbor classification. In comparison, moment invariants allowed only 87% accuracy. A second experiment utilized 10 classes of hand printed numerals. Using 47 features, an 84% classification accuracy was achieved. With noisy data, Zernike moments are described as good for SNR of 25 dB.

In other work, Khotanzad and Lu [22] utilized Zernike moment based features with a neural network classifier. The neural network was a multi-layer perceptron with one hidden layer. Back projection was used for network training. The neural network was compared with nearest-neighbor, Bayes, and minimum-mean distance. Additionally, moment invariants and Zernike moments of varying order were compared. Experimental results demonstrated that the neural net outperforms the competing classifiers, especially for low

SNR images. Additionally, Zernike moments are shown to outperform moment invariants.

Belkasim, Shridhar, and Ahmadi [23] derived a generalized form of Zernike moment invariant (ZMI) for the nth order. These invariants are based on the rotational properties of Zernike moments shown in (3.44).

The *primary* invariants are given by

$$ZMI_{n0} = Z_{n0} \tag{3.46a}$$

$$ZMI_{nL} = \left| Z_{nL} \right| \tag{3.46b}$$

And the *secondary* invariants are given by

$$ZMI_{n,n+z} = \left[Z^{*}_{mh} Z^{p}_{nL} \right] \pm \left[Z^{*}_{mh} Z^{p}_{nL} \right]^{*} \tag{3.46c}$$

$$where \quad h \le L, \ m \le n, \ p = h/L, \ 0 \le p \le 1, \ z = L/H$$

The number of independent ZMI of order n is $n+1$ and are defined for odd and even n as

for n even :

$$ZMI_{n0} \text{ and } ZMI_{nL} \quad \text{for } L = 2,4,6,\ldots,n \tag{3.47a}$$

$$ZMI_{n,n+z} = 2\left|Z_{n2}\right|\left|Z_{nL}\right|^{p} \cos(p\,\phi_{nL} - \phi_{n2}) \text{ for } L = 4,6,8,\ldots,n \ \ p = 2/L \ \ z = L/2 \tag{3.47b}$$

$$ZMI_{n,n+1} = 2\left|Z_{n-2,2}\right|\left|Z_{n2}\right| \cos(\phi_{n-2,2} - \phi_{n2}) \tag{3.47c}$$

for n odd :

$$ZMI_{nL} = \left|Z_{nL}\right. \quad \text{for } L = 1,3,5,\ldots,n \tag{3.48a}$$

$$ZMI_{n,n+L} = 2\left|Z_{n1}\right|\left|Z_{nL}\right|^{p} \cos(p\,\phi_{nL} - \phi_{n1}) \text{ for } L = 3,5,7,\ldots,n \ \ p = 1/L \tag{3.48b}$$

$$ZMI_{n,n+1} = 2\left|Z_{n-2,1}\right| \cos(\phi_{n-2,1} - \phi_{n1}) \tag{3.48c}$$

Analogous invariants were also derived for pseudo-Zernike moments.

A normalization technique is described that is claimed to reduce dynamic range and information redundancy. The normalized Zernike moments (NZM) are given by

$$NZM_{nL} = \frac{Z_{nL}}{Z_{n-2,L}} \quad \text{for } Z_{n-2,L} \ne 0 \text{ and } L < n \tag{3.49a}$$

$$NZM_{nL} = Z_{nL} \quad \text{for } Z_{n-2,L} = 0 \text{ or } L = n \tag{3.49b}$$

Experimental results showed normalized Zernike moment invariants outperform Zernike, pseudo-Zernike, Teague-Zernike [2], and moment invariants.

3.3.3. Pseudo-Zernike Moments

Teh and Chin [24] presented a modification of Teague's Zernike moment based on a related set of orthogonal polynomials that have properties analogous to Zernike polynomials. These polynomials, called *pseudo-Zernike* polynomials, differ from the conventional Zernike in definition of the radial polynomial R_{nl}. The pseudo-Zernike radial polynomials are defined by

$$R_{nl}(r) = \sum_{m=0}^{n-l} (-1)^m \frac{(2n+1-m)!}{m!\,(n-l-m)!\,(n+l+1-m)!} r^{n-m} \quad n = 0, 1, 2, \ldots, \infty \quad 0 \le l \le n \tag{3.50}$$

The set of pseudo-Zernike polynomials contains $(n+1)^2$ linearly independent polynomials of degree $\le n$, while the set of Zernike polynomials contain only $\frac{1}{2}(n+1)(n+2)$ linearly independent polynomials. Figure 3.5ab shows the pseudo-Zernike polynomials through order 5 for $l = 0,1$. These polynomials exhibit a wider dynamic range than conventional Zernike polynomials and similarily become more correlated as the radius approaches 1. Moments based on pseudo-Zernike polynomials were theoretically shown to be less sensitive to noise than the conventional Zernike moments.

Belkasim, Shridhar, and Ahmadi [23] derived a generalized form of an nth-order pseudo-Zernike moment invariant. These moment invariants are analogous to invariants derived for Zernike moments described in the previous section. As with the Zernike moments, a normalization scheme that reduces dynamic range and information redundancy was also described.

3.4. Complex Moments

The method of Complex moments, presented by Abu-Mostafa and Psaltis [25] is based on yet another alternative to the moment definition and provides a simple and straightforward technique for deriving a set of invariant moments.

3.4.1. Two-Dimensional Complex Moment Invariants

The two-dimensional complex moment, C_{pq}, of order, (p, q), is defined by

$$C_{pq} \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x+iy)^p \, (x-iy)^q \, f(x,y)\, dx\, dy \tag{3.51}$$

If $f(x,y)$ is non-negative real then C_{pp} is non-negative real and C_{pq} is the complex conjugate of C_{qp}. Complex moments may be expressed as a linear combination of conventional moments by

$$C_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sum_{r=0}^{p} \binom{p}{r} x^r \ (iy)^{p-r} \sum_{s=0}^{q} \binom{q}{s} (-1)^{q-s} \ x^s \ (iy)^{q-s} \ f(x,y) \ dx \ dy \tag{3.52a}$$

$$C_{pq} = \sum_{r=0}^{p} \sum_{s=0}^{q} \binom{p}{r} \binom{q}{s} x^r \ (iy)^{p-r} \ (-1)^{q-s} \ (i)^{(p+q)-(r+s)} \ m_{r+s,(p+q)-(r+s)} \tag{3.52b}$$

When considered in polar form

$$C_{pq} \equiv \int_{0}^{2\pi}\int_{0}^{\infty} r^{p+q} \ e^{i(p-q)\theta} f(r,\theta) \ r \ dr \ d\theta \tag{3.53}$$

the complex moments are related to rotational moments by

$$C_{pq} = D_{p+q,p-q} \tag{3.54}$$

and may also be shown to be related to Zernike moments. Like rotational and Zernike moments, the result of a rotation of angle ϕ is defined as

$$C'_{pq} = C_{pq} e^{-i(p-q)\phi} \tag{3.55}$$

Moment invariants may be derived from complex moments using the following formula

$$(C_{rs} \ C_{tu}{}^{k} + C_{sr} \ C_{ut}{}^{k}) \quad where \quad (r-s)+k(t-u)=0 \tag{3.56}$$

This combination of complex moments cancels both the imaginary moment and the rotational phase factor thus providing real-valued rotation invariants.

Abu-Mostafa and Psaltis [25] utilized complex moments to analyze the informational properties of moment invariants to arrive at a theoretical measure of moments ability to distinguish between patterns. Information loss, suppression, and redundancy in moment invariants were considered and compared with Zernike moments. It was determined that moment invariants suffer from all of the above, while Zernike moments mainly suffer only from information loss. From this, they concluded that moment invariants are not good image features, in general. They note, however, there are specific instances when performance is not degraded by these informational properties.

In other work, Abu-Mostafa and Psaltis [26] investigated the use of moments in a generalized image normalization scheme for invariant pattern recognition. They first redefined the classical image normalizations of size, position, rotation (principal axis), and contrast, in terms of complex moments. They then systematically extended the normalization procedures to higher orders of complex moments. Moment invariants were shown to be derivable from complex moments of the normalized imagery.

Abo-Zaid, Hinton, and Horne [14] presented an alternate technique for computing normalized complex moments based on linear combinations of normalized conventional moments. They derived normalized complex moments from

$$C_{pq}^{norm} = \frac{1}{\mu_{00}} \left(\frac{\mu_{00}}{\mu_{02} + \mu_{20}} \right)^{\frac{p+q}{2}} C_{pq}^{central} \tag{3.57}$$

where $C_{pq}^{central}$ are complex moments generated using equation (3.52b) with central moments.

3.4.2. Three-Dimensional Complex Moment Invariants

Lo and Don [27] presented a derivation of three-dimensional complex moments using group representation theory. A vector of computing complex moments is computed from a vector of conventional moments via a complex matrix that transforms between the monomial basis and the harmonic polynomial basis. A group-theoretic approach is then used to construct three-dimensional moment invariants. Complex moments and moment invariants are derived for second and third order moments using this technique.

3.5. Standard Moments

The first technique not based on algebraic invariants, *standard moments*, was introduced by Reeves and Rostampour [28]. In general, this technique takes advantage of the simple linear transform properties of moments and achieves invariance through image feature normalization in the moment domain.

3.5.1. Two-Dimensional Standard Moments

Two-dimensional standard moments are based on robust normalization criteria for scale, position, orientation, and aspect ratio. Initially, a *raw* moment set, $\{m_{pq}\}$, of desired order is computed from the given image silhouette using equation (2.02). The normalization transformations are then performed on the raw moment set to derive the standard moment set, $\{M_{pq}\}$. A description of each normalization follows.

Size normalization is achieved by transforming the moment set so that the resulting moments represent the object image at a scale that makes the object area 1.

Translation normalization is achieved by transforming the moment set so that the resulting moments represent an object whose origin is at a unique point within the image, specifically, the center of gravity (central moments). This normalization results in a new moment set with ($M_{10} = M_{01} = 0$).

Rotation normalization is performed by rotating the moment set so that the moments represent an object with its principal axes aligned with the coordinate axes. This is based on Hu's original idea of rotation normalization by *Principal Axes*. There are four possible rotation angles that align the principal axes with the coordinate axes ($\phi + \frac{1}{2} n \ \pi$). To determine a unique orientation of the principal axes, n is chosen such that the major principal axis is aligned with the x-axis and the projection of the object onto the major axis is skewed to the left. This is accomplished by constraining the rotationally transformed moments to $M_{20} \geq M_{02}$ and $M_{30} \geq 0$ respectively. In addition to the above constraints, the normalized moment set has $M_{11} = 0$ since $\theta = 0$. If reflection normalization is desired, an additional constraint, $M_{03} \geq 0$, may be imposed. This constraint causes the projection of the object onto the minor principal axis to be skewed towards the bottom.

Reeves and Rostampour [1528] utilized standard moments for global generic shape analysis. Four "ideal" symmetric generic shapes were selected; a rectangle, ellipse, diamond, and concave object. The kurtosis of the major axis (x-axis) projection, K_x, of each of these shapes was computed from standard moments using

$$K_x = \frac{M_{40}}{{M_{20}}^2} - 3 \tag{3.58}$$

Additionally, the *normalized length* and *width* was determined for the first three shapes as a function of the second order moments. These values are given in table 3.1.

Table 3.1. Kurtosis and Normalized Dimensions of Generic Shapes

Shape	K_x	Length	Width
rectangle	–1.2	$\sqrt{12M_{20}}$	$1/\sqrt{12M_{20}}$
ellipse	–1.0	$\sqrt{16M_{20}}$	$1/\pi \sqrt{M_{20}}$
diamond	–0.6	$\sqrt{24M_{20}}$	$1/\sqrt{6M_{20}}$
concave	>–0.6		

Standard moments, $\{M_{pq}\}$, were computed for segmented test input images. The general shape of the input object was determined from the kurtosis of the major axis projection. The length or width of an object was then estimated by multiplying the calculated normalized values by $\sqrt{m_{00}}$. This technique was used successfully to distinguish between low resolution aerial views of buildings, a storage tank, and an airplane.

Reeves, Prokop, et.al [29]. demonstrated the technique of *aspect ratio normalization* in order to improve the behavior of standard moments. It was observed that if the object image coordinates were constrained to ($-1.0 \leq x, y \leq 1.0$) (i.e. a 2×2 square centered at the origin), the magnitudes of the moments decrease as their order increased. Additionally, if the moment set is size normalized with ($M_{00} = 1$) then all the moments have a magnitude ≤ 1. Aspect ratio normalization is an attempt to meet these constraints by changing the ellipsoid of inertia of the object to a circle while leaving the object area unchanged. This is equivalent to differentially scaling the object so that the transformed moments are $M_{20} = M_{02}$ and $M_{00} = 1$ respectively. Improved representation performance was demonstrated through aspect ratio normalization. Additionally, the aspect ratio was utilized as a highly discriminating object feature. In summary, the low-order moments of a Standard Moment Set have the values given in table 3.2.

Table 3.2. Two-Dimensional Standard Moments

Standard Moment	Normalization
$M_{00} = 1$	*area*
$M_{10} = 0$	*x - translation*
$M_{01} = 0$	*y - translation*
$M_{11} = 0$	*rotation*
$M_{20} \geq M_{02}$	*rotation*
$M_{30} \geq 0$	*rotation*
$M_{20} = M_{02}$	*aspect ratio*

3.5.2. Grey-Level Standard Moments

Reeves [15] defines the grey-level moments, $\{m_{pqr}\}$, of *order* $(p+q+r)$, of an image, $f(x, y)$, as

$$m_{pqr} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y)^r \, dx \, dy \tag{3.59}$$

A complete moment set of order n consists of all moments, m_{pqr}, such that $p+q+r \leq n$ and contains 1/6 $(n+1)(n+2)(n+3)$ elements. (Note that the set $\{m_{pq0}\}$ are the silhouette moments and the set $\{m_{pq1}\}$ are the moments of the grey-levels.)

Since the grey-levels may have an arbitrary mean and variance due to the illumination and sensor characteristics, they must be normalized with respect to these values. Normalization of the moments requires operations to offset and scale the grey-levels. The addition of a bias, α, to the grey-levels (translation in the z direction) is defined by

$$m'_{pqr} = \int\int x^p y^q (\alpha + f(x, y))^r \, dx \, dy \tag{3.60}$$

$$m'_{pqr} = \sum_{s=0}^{r} \binom{r}{s} \alpha^{r-s} m_{pqs} \tag{3.61}$$

A scale change of the grey-levels by a factor β (scaling in the z dimension) is defined by

$$m'_{pqr} = \int\int x^p y^q (\beta f(x, y))^r \, dx \, dy \tag{3.62}$$

$$m'_{pqr} = \beta^r \, m_{pqr} \tag{3.63}$$

The grey-level mean of the image segment, $\overline{M}_g$, is given by

$$\overline{M}_g = \frac{m_{000}}{m_{001}} \tag{3.64}$$

and the variance, V_g, is given by

$$V_g = \frac{m_{002}}{m_{000}} \tag{3.65}$$

The grey-level moments are normalized with $\alpha = -\overline{M}_g$ and $\beta = \sqrt{V_g}$ using equations (2.57) and (2.59) respectively. Grey-level standard moments have values shown in table 3.3.

Table 3.3. Grey-Level Standard Moment Values

Standard Moment	Normalization
$M_{001} = 0$	*mean*
$M_{002} = 1$	*variance*

Taylor and Reeves [4] extended the grey-level moment transforms to include rotations about the x and y axes. In terms of grey-level standard moments, M_{pqr}, a positive rotation, θ_x, of the coordinate system about the x axis is given by the transform

$$M'_{pqr} = \sum_{s=0}^{q} \sum_{w=0}^{r} \binom{q}{s} \binom{r}{w} (-1)^s \ (\cos\theta_x)^{q-s+w} \ (\sin\theta_x)^{s+r-w} \ M_{p,q-s+r-w,s+w} \tag{3.66}$$

A positive rotation, θ_y, of the coordinate system about the y axis is given by the transform

$$M'_{pqr} = \sum_{s=0}^{p} \sum_{w=0}^{r} \binom{p}{s} \binom{r}{w} (-1)^{r-w} \ (\cos\theta_y)^{p-s+w} \ (\sin\theta_y)^{s+r-w} \ M_{p-s+r-w,q,s+w} \tag{3.67}$$

3.5.3. Range Standard Moments

Reeves and Wittner [30] used standard moments to represent 2½-dimensional imagery derived from a range sensor. The technique requires computation of two sets of moments, one for the range image and a second for a silhouette of the image. First, a standard moment set, $\{S_{pq}\}$, is computed from the raw silhouette moments. The transform parameters computed for the silhouette moments are then used to normalize the raw range moments. In this way, the object represented by the normalized range moment set, $\{R_{pq}\}$, will be consistent with the object represented by the normalized silhouette moment set. Additionally, it is assumed that the depth dimension of the range image is of the same scale as the width and height of the image. Therefore, to keep the image depth consistent with the width and height, the scale factor used for size normalization is also used to scale the depth dimension of the range moments. Scaling the depth is equivalent to an intensity change and is accomplished using equation (2.18). The range data, however, requires the further normalization of image volume and position in the depth (z) dimension.

Reeves, Prokop, and Taylor [31] presented a range normalization method that accounts for both depth position and volume and is easily implemented with the information at hand. A robust representation for the entire object is contrived using a reasonable and consistent set of assumptions about the occluded part of object. First, it is assumed that the back of the object is flat and parallel to the image plane. In addition, it is assumed that the cross-section of the occluded part of the object has the same shape as the occluding boundary. Finally, it is assumed that the occluded part has a depth (intensity) represented by α. Figure 3.6 illustrates these assumptions. The top of this figure is an example of a ranged sensed object. For this view the range sensor is at $z = \infty$. Note that the coordinate axis are drawn only for direction reference, no assumption is made concerning the actual object position in space. The lower part of figure 3.6 is the assumed cross-section of the object at the x axis.

Based on the given assumptions, the moments for the visible part of the object are R_{pq} and the moments for the occluded part of the object are $\{\alpha S_{pq}\}$. The moments for the entire contrived object are given by

$$M_{pq} = R_{pq} + \alpha S_{pq} \tag{3.68}$$

Volume normalization is accomplished by computing α in the above expression to make $(M_{00} = 1)$. Depth position normalization is accomplished by setting the origin of the depth dimension to the assumed back of the object.

3.5.4. Three-Dimensional Standard Moments

Reeves and Wittner [30] also extended standard moments to represent objects defined in three-dimensional space. Note that this differs from range moments in that range moments represent the object shape and surface characteristics while three-dimensional moments represent internal information about an object. The three-dimensional analogies to the two-dimensional silhouette and range moments are referred to as *solid* and *density* moments, respectively.

The three-dimensional Cartesian moment, m_{pqr}, of order, $(p+q+r)$, is defined by

$$m_{pqr} \equiv \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p \, y^q \, z^r \, f(x, y, z) \, dx \, dy \, dz \tag{3.69}$$

Solid moments are generated when the object description is binary (i.e. $f(x, y, z) = 1$ within the object and $f(x, y, z) = 0$ outside the object). Density moments are generated when the object function represents an object with a varying internal density distribution.

The properties, transformations, and normalizations of three dimensional moments are completely analogous to the two-dimensional case. Consequently, a three dimensional standard moment may be defined for solid moments as for silhouette moments. The three dimensional standard moment set has the low order moment values given in table 3.4.

Table 3.4. Three-Dimensional Standard Moment Values

Standard Moment	Normalization
$M_{000} = 1$	*volume*
$M_{100} = 0$	*x - translation*
$M_{010} = 0$	*y - translation*
$M_{001} = 0$	*z - translation*
$M_{110} = M_{101} = M_{011} = 0$	*rotation*
$M_{200} \geq M_{020} \geq M_{002}$	*rotation*
$M_{300} \geq 0$ and $M_{300} \geq 0$	*rotation*

Note that aspect ratio normalization had not been defined for moments when this work was originally presented. However, the analogous 3-dimensional requirement for aspect ratio normalization would be

$$M_{200} = M_{020} = M_{002} = 0$$

As of this point, a normalization technique for density moments, the-three dimensional analogy to range normalization, has not yet been explored.

Reeves and Wittner also conducted three-dimensional generic shape analysis experiments similar to those presented in previous work [28]. Analogous to the two-dimensional case, the moment set $\{M_{p\,00}\}$ is the set of moments of a projection of a three-dimensional shape onto the x axis. The kurtosis in each dimension is defined by

$$K_x^{3d} = \frac{S_{400}}{S_{200}^2} - 3 \qquad K_y^{3d} = \frac{S_{040}}{S_{020}^2} - 3 \qquad K_z^{3d} = \frac{S_{004}}{S_{002}^2} - 3 \tag{3.70abc}$$

The generic shapes under consideration were a rectangular solid and an elliptical cylinder with a resolution of $32 \times 32 \times 32$. Standard moments and a weighted Euclidean classification scheme were used to distinguish between fifty random views of each object with 100% accuracy. Kurtosis values of the shapes were then used to estimate the object dimensions. It was noted that the relative deviation (standard deviation / mean) for M_{020} was more that ten times greater than for S_{200} and S_{002}. This is attributed to the fact that a rotation about the major or minor principal axis of a three dimensional rigid body is stable while rotation about an intermediate principal axis is unstable. This effect was also reflected in the kurtosis values. While kurtosis in the x and z dimensions was stable and a robust predictor of the basic shape of the object, the kurtosis in the y dimension was considerably smaller than the ideal values in all tests.

4. Fast Moment Computation

In each of the principal moment techniques, a significant amount of computation is required to generate the original moment values, $\{m_{pq}\}$, from the imagery. To allow moment techniques to be used in real-time image processing and object classification applications, various special purpose architectures have been proposed for the fast calculation of moments.

4.1. Optical Moments

Optical moment calculation takes advantage of the relationship between moments and the Fourier transform of a distribution. Specifically, the characteristic function of a distribution may be defined as

$$\Phi(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y)\, e^{-i\,2\pi(ux+vy)}\, dx\ dy \tag{4.01}$$

which is the Fourier transform of $f(x, y)$. Furthermore, if moments of all orders exist, then $\Phi(u, v)$ may be also expressed as a power series in terms of moments, m_{pq}, as

$$\Phi(u, v) = \sum_{p=0}^{\infty} \sum_{q=0}^{\infty} \frac{(-i\,2\pi)^{p+q}}{p\,!q\,!}\, u^p v^q\ m_{pq} \tag{4.02}$$

Teague [32] describes a system that calculates moments based on the derivatives of the optically computed Fourier transform of an image. Given the Fourier transform, $F(\xi, \eta)$, of an image plane irradiance distribution, $f(x, y)$, the moments, m_{pq}, may be computed using

$$m_{pq} = \frac{1}{(-i\,2\pi)^{p+q}} \left[\left(\frac{\partial}{\partial \xi} \right)^p \left(\frac{\partial}{\partial \eta} \right)^q F(\xi, \eta) \right]_{\xi=\eta=0} \tag{4.03a}$$

$$\approx \frac{1}{(-i\,2\pi)^{p+q}}\ \Delta_\xi^p \Delta_\eta^q F(\xi, \eta)_{\xi=\eta=0} \tag{4.03b}$$

Optical calculation of moments with this method requires a lens, phase plate and a network of mirrors, beam splitters, and detectors to determine the Fourier transform. The partial derivatives are then estimated by the method of finite differences and measured by strategic spacing of the detectors in the Fourier plane.

Casasent and Psaltis [33] describe a hybrid optical/digital processor that optically computes all the moments, $\{m_{pq}\}$ of a 2-dimensional image in parallel. A laser light passes through a transparency of an image, $f(x, y)$, then through a mask, $g(x, y)$, then through a Fourier transform lens and the final pattern is collected in a photodetector. In this arrangement, the amplitude at the photo detector, $u(\omega_x, \omega_y)$, is given by

$$u(\omega_x, \omega_y) = \int\int f(x, y)\, g(x, y)\, e^{-j(\omega_x x + \omega_y y)}\, dx\ dy \tag{4.04}$$

A simple case to consider is the on-axis amplitude, $u(0, 0)$,

$$u(0, 0) = \int\int f(x, y)\, g(x, y)\, dx\ dy \tag{4.05}$$

Proper selection of the mask will cause the on-axis output to be a specific moment value. For example, if $g(x, y) = xy$, then the on-axis value will be m_{11}. Calculation of moments in this manner, however, would require a different mask for each moment value. Casasent and Psaltis propose a single mask function

$$g(x, y) = e^{xe^{j\omega_0 x}}\ e^{ye^{j\omega_0 y}} \tag{4.06}$$

that results in a light pattern given by

$$u(p\,\omega_0, q\,\omega_0) = \frac{m_{pq}}{p\,!q\,!} \tag{4.07}$$

thus computing all the moments in parallel. Once the moments, m_{pq}, have been optically computed, digital processing is used to compute moment invariants.

Casasent, Pauly, and Fetterly [34], utilize the hybrid optical/digital moment generation for classification of ships from infrared imagery. A new estimation approach is presented motivated by statistical analysis that shows raw moments to be superior to moment invariants for the task at hand.

Casasent, Cheatham, and Fetterly [35] utilize the hybrid optical/digital computation of moments in a robotic pattern recognition system. In this work, a simpler mask function, $g(x, y)$, is presented that may be used to generate moments of finite order. This mask is based on translating the input function, $f(x, y)$, into the first quadrant so that all the moments are positive. The positive and real mask is given by

$$g(x, y) = \sum_{p=0}^{P} \sum_{q=0}^{Q} x^p y^q [B_p + \cos(\omega_1 + p\,\omega_0)x][B_q + \cos(\omega_1 + q\,\omega_0)y] \tag{4.08}$$

The optical processor using this mask is referred to as a *Finite-Order Cosine processor*. Note that the moments generated in this system are that of a shifted input function. The moments of the original unshifted input function, however, may be obtained by simple translation in the moment domain.

In other work, Cheatham, Casasent, and Fetterly [36] utilize the Finite-Order Cosine processor scheme and present a recognition system that is invariant to scale, translation, and in-plane rotational distortions.

4.2. Hardware Architectures

Reeves [37] has proposed a parallel, mesh-connected SIMD computer architecture for rapidly manipulating moment sets. This architecture is a triangular matrix of processing elements, one for each moment value in a complete moment set of a given order. Each processing element contains an ALU capable of both multiplication and addition, and some local memory. Performance is characterized by computational cost, speedup and processor utilization on the parallel moment computer for a host of moment operations including generation, scaling, translation, rotation, reflection, and superposition. The architecture offers a reasonable speedup over a single processor for high speed image analysis operations and may be implemented in VLSI technology.

Hatamian [38] , has proposed an algorithm and single chip VLSI implementation for generating raw moments at video rates. It is claimed that 16 moments, $m_{pq}(p = 0, 1, 2, 3,\ q = 0, 1, 2, 3)$ (a complete moment set of order 3 + additional higher order moments), of a $512 \times 512 \times 8$ bit image can be computed at 30 frames/sec. The moment algorithm is based on using the one dimensional discrete moment generating function as a digital filter. Z-transform analysis of the impulse response of this filter derives an implementation that is a 2 dimensional array of single-pole digital filters.

5. Moment Performance Comparisons

Teh and Chin [24] performed an extensive analysis and comparison of the most common moment definitions. Conventional, Legendre, Zernike, pseudo-Zernike, rotational, and complex moments were all examined in terms of noise sensitivity, information redundancy, and image representation ability. Both analytic and experimental methods were used to characterize the various moment definitions.

In terms of sensitivity to additive random noise, in general, high order moments are the most sensitive to noise. Among the explored techniques, it was concluded that complex moments are least sensitive to noise while Legendre moments are most severely affected by noise. In terms of information redundancy, orthogonal techniques (Legendre, Zernike, and pseudo-Zernike) are uncorrelated and thus have the least redundancy. In terms of overall performance, Zernike and pseudo-Zernike proved to be the best.

An experimental comparison of moment techniques was performed by Reeves, Prokop, et.al [29]. In this work, moment invariants, Legendre moments, standard moments, as well as Fourier descriptors [39] were compared based on their performance as invariant features for a standardized six airplane experiment. Note that the method of Fourier descriptors was provided as a representative non-moment technique. The task involved the classification of synthetically generated noiseless and noisy silhouette and/or boundary images of each of six aircraft viewed from 50 random angles as compared to a library of 500 views of each uniformly sampled over the entire viewing sphere. This experiment is considered to be representative of a difficult task since it involves a wide range of shapes (given all possible views of an aircraft) yet the basic three-dimensional shapes of the different objects are very similar. Feature vectors for each object image were generated utilizing the various techniques. A nearest-neighbor Euclidean distance classifier was used to compare the feature vectors. Varing feature vector lengths were tested to determine the minimum length for unique object representation. (Moment invariants were fixed at length 7.)

Classification results showed that moment invariants were the least effective for this task. Legendre moments performed better than moment invariants but not as well as Fourier descriptors. Fourier descriptors were shown to be adversely affected by noise. Feature vectors defined from standard moments of silhouette imagery outperformed all other tested methods for both uncorrupted and noisy imagery.

In other work, Reeves, Prokop, et.al [31]. revised the six airplane experiment to utilize a worst-case set of 252 unknown views that are evenly spaced about the viewing sphere as well as being intersitually located between the library views. In addition, synthetic 2½-dimensional (range) imagery was generated to evaluate moment techniques that exploit such information. A model of range noise was also developed to produce noisy range imagery.

Experimental results demonstrated that feature vectors comprised of a combination of silhouette and range standard moments provided the best classification results as well as being robust in the presence of noise.

Cash and Hatamian [40], performed an extensive comparison of the effectiveness of moment feature vector classification schemes including Euclidean distance, weighted Euclidean distance, cross correlation, and Mahanalobis distance. An optical machine-printed character recognition task was performed utilizing feature vectors of size-normalized, third order, central moments.

The highest classification rates were achieved using a cross correlation measure weighted by the reciprocal of the mean of the intra-class standard deviations. For several font classes the recognition rate was over 99%. Similar results were achieved for a Euclidean distance measure using the same weight. It was noted that the Euclidean method is probably more desirable of these two since it requires much less computation.

6. Conclusion

The method of moments provides a robust technique for decomposing an arbitrary shape into a finite set of characterisitic features. A major strength of this approach is that it is based on a direct linear transformation with no application specific "heuristic" parameters to determine.

The moment techniques have an appealing mathematical simplicity and are very versatile. They have been explored for a wide range of applications and image data types. A major limitation of the moment approach is that it cannot be "tuned" to be sensitive to the specific object features or constraints. Furthermore, it can only be directly applied to global shape identification tasks.

The principal moment techniques presented may be distinguished by four basic characteristics. The first is the basis function used in the moment definition. The presented techniques used a variety of orthogonal and non-orthogonal basis functions. Second is the type of image sampling used; i.e., rectangular or polar. The applicability of techniques to different forms of imagery is another important distinguishing feature. Finally, whether invariance is achieved through algebraic invariants or feature normalization may be considered. The characteristics of the principal techniques are summarized in table 6.7.

Table 6.7. Moment techniques.

Technique	Basis Polynomials	Sampling	Image Data
Moment Invariants	*monomials*	*rect*	*2-D, 3-D*
Rotational Moments	*circular harmonics*	*polar*	*2-D*
Orthogonal Moments	*Legendre*	*rect*	*2-D*
	Zernike	*polar*	*2-D*
	pseudo-Zernike	*polar*	*2-D*
Complex Moments	*circular harmonics*	*rect polar*	*2-D*
	spherical harmonics	*rect polar*	*3-D*
Standard Moments	*monomials*	*rect*	*2-D,* 2½-D, 3-D, grey

Note that all techniques utilize algebraic invariants except standard moments.

It is not clear from the studies conducted to date which technique is best for a given application. Some studies have implied [2] that important information may be contained in the higher order moments. Whereas, most practical experiments have shown little improvement in identification performance when moment orders are increased beyond order 4 or 5 [2931]. In general, high order moments are more sensitive to noise.

There are few image feature techniques that can be directly compared to the moment approach. One technique that may be directly compared to moments for binary shape identification is Fourier descriptors. The Fourier descriptors, which are based on the object boundary rather that the silhouette, may be shown to be more sensitive to more types of boundary variations. However, for a practical application involving a large number of shapes it is very difficult to predict which technique would provide the best performance without performing empirical experiments.

Object identification using the moment method involves two stages (1) object characterization and (2) object matching. This survey has focused on feature generation techniques. In many cases, object matching is achieved by a nearest neighbor approach after, possibly, some preconditioning of the moment features. Once again, the optimal matching technique is application dependent. In general, moment techniques have proved to be very effective for global recognition tasks involving rigid objects.

References

1. M. K. Hu, “Visual Pattern Recognition by Moment Invariants,” *IRE Transactions on Information Theory*, vol. IT-8, pp. 179-187, February 1962.

2. M. R. Teague, “Image Analysis via the General Theory of Moments,” *Journal of the Optical Society of America*, vol. 70, no. 8, pp. 920-930, August 1980.

3. R. J. Prokop, “The Technique of Standard Moments for Global Feature Object Representation,” *Cornell University Master of Science Thesis*, May 1990.

4. R. W. Taylor and A. P. Reeves, “Three-Dimensional Image Transforms in Moment Space,” *Proceedings of the IEEE Computer Society Workshop on Computer Vision*, pp. 366-368, 1987.

5. V. E. Giuliano, P. E. Jones, G. E. Kimball, R. F. Meyer, and B. A. Stein, “Automatic Pattern Recognition by a Gestalt Method,” *Information and Control*, vol. 4, no. 4, pp. 332-345, December 1961.

6. F. A. Sadjadi and E. L. Hall, “Three-Dimensional Moment Invariants,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-2, no. 2, pp. 127-136, March 1980.

7. S. A. Dudani, K. J. Breeding, and R. B. McGhee, “Aircraft Identification by Moment Invariants,” *IEEE Transactions on Computers*, vol. C-26, no. 1, pp. 39-46, January 1977.

8. A. Sluzek, “Using Moment Invariants to Recognize and Locate Partially Occluded 2D Objects,” *Pattern Recognition Letters*, vol. 7, no. 4, pp. 253-257, April 1988.

9. A. Sluzek, “Moment Based Methods of Identification and Localization of Partially Visible Objects,” pp. 281-291 in *Issues on Machine Vision*, ed. G.G. Pieroni,Springer Verlag (1989).

10. J. F. Gilmore and W. W. Boyd, “Building and Bridge Classification by Invariant Moments,” *SPIE*, August 1981.

11. F. A. Sadjadi and E. L. Hall, “Numerical Computation of Moment Invariants for Scene Analysis,” *Proceedings of IEEE Conference on Pattern Recognition and Image Processing*, pp. 127-136, 1978.

12. R. Y. Wong and E. L. Hall, “Scene Matching with Invariant Moments,” *Computer Graphics and Image Processing*, vol. 8, pp. 16-24, 1978.

13. S. Maitra, “Moment Invariants,” *Proceedings of the IEEE*, vol. 67, no. 4, pp. 697-699, April 1979.

14. A. Abo-Zaid, O. R. Hinton, and E. Horne, “About Moment Normalization and Complex Moment Descriptors,” *Proceeding of the 4th International Conference on Pattern*

Recognition, pp. 399-407, March 1988.

15. A. P. Reeves, “The General Theory of Moments for Shape Analysis and the Parallel Implementation of Moment Operations,” *Purdue University Technical Report* , no. TR-EE 81-37, October 1981.

16. F. W. Smith and M. H. Wright, “Automatic Ship Photo Interpretation by the Method of Moments,” *IEEE Transactions on Computers*, pp. 1089-1095, September 1971.

17. J. F. Boyce and W. J. Hossack, “Moment Invariants for Pattern Recognition,” *Pattern Recognition Letters*, vol. 1, no. 5,6, pp. 451-456, July 1983.

18. S. S. Reddi, “Radial and Angular Moment Invariants for Image Identification,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-3, no. 2, pp. 240-242, March 1981.

19. B. H. Yin and H. Mack , “Target Classification Algorithms for Video and FLIR Imagery,” *SPIE*, August 1981.

20. A. Khotanzad and Y. H. Hong, “Rotation Invariant Image Recognition Using Features Selected via a Systematic Method,” *Pattern Recognition*, vol. 23, no. 10, pp. 1089-1101, 1990.

21. A. Khotanzad and Y. H. Hong, “Invariant Image Recognition by Zernike Moments,” *IEEE Transactions on Pattern Analysis and Machine Intellegence*, vol. 12, no. 5, pp. 489-498, May 1990.

22. A. Khotanzad and J. H. Lu, “Classification of Invariant Image Representations Using a Neural Network,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 6, pp. 1028-1038, June 1990.

23. S. O. Belkasim, M. Shridhar, and M. Ahmadi, “Shape-Contour Recognition Using Moment Invariants,” *Proceedings of 10th ICPR, Atlantic City NJ*, pp. 649-651, 1990.

24. C. H. Teh and R. T. Chin, “On Image Analysis by The Method of Moments,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-10, no. 2, pp. 496-513, July 1988.

25. Y. S. Abu-Mostafa and D. Psaltis, “Recognitive Aspects of Moment Invariants,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, pp. 698-706, November 1984.

26. Y. S. Abu-Mostafa and D. Psaltis, “Image Normalization by Complex Moments,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-7, no. 1, pp. 46-55, January 1985.

27. C. H. Lo and H. S. Don, “Object Identification and Positioning via 3-D Moments : A Group-Theoretic Approach,” *State University of New York at Stony Brook Image*

Analysis and Graphics Laboratory Technical Report 87-06, October 12 1987.

28. A. P. Reeves and A. Rostampour, “Shape Analysis of Segmented Objects Using Moments,” *IEEE Computer Society Conference on Pattern Recognition and Image Processing*, pp. 171-174, August 1981.

29. A. P. Reeves , R. J. Prokop, and S. E. Andrews, “Three Dimensional Shape Analysis Using Moments and Fourier Descriptors,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 6, pp. 937-943, November 1988.

30. A. P. Reeves and B. S. Wittner, “Shape Analysis of Three Dimensional Objects Using The Method of Moments,” *Proceedings of 1983 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 20-26, June 1983.

31. A. P. Reeves , R. J. Prokop, and R. W. Taylor, “Shape Analysis of Three Dimensional Objects Using Range Information,” *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, June 1985.

32. M. R. Teague, “Optical Calculation of Irradiance Moments,” *Applied Optics*, vol. 19, no. 8, pp. 1353-1356, April 15, 1980.

33. D. Casasent and D. Psaltis, “Hybrid Processor to Compute Invariant Moments for Pattern Recognition,” *Optics Letters*, vol. 5, no. 9, pp. 395-397, September 1980.

34. D. Casasent, J. Pauly, and D. Fetterly, “Infrared Ship Classification Using a New Moment Pattern Recognition Concept,” *SPIE*, vol. 302, pp. 126-133, 1981.

35. D. Casasent, L. Cheatham, and D. Fetterly, “Hybrid Optical/Digital Moment-Based Robotic Pattern Recognition System,” *SPIE*, vol. 360, pp. 105-111, 1982.

36. L. Cheatham, D. Casasent, and D. Fetterly, *Distortion Invariant Recognition Using a Moment Feature Space*. 1983.

37. A. P. Reeves, “A Parallel Mesh Moment Computer,” *Proceedings of the 6th International Conference on Pattern Recognition*, pp. 465-467, October 1982.

38. M. Hatamian, “A Real-Time Two-Dimensional Moment Generating Algorithm and Its Single Chip Implementation,” *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. ASSP-34, pp. 546-553, June 1986.

39. T. P. Wallace and P. A. Wintz, “An Efficient Three-Dimensional Aircraft Recognition Algorithm Using Normalized Fourier Descriptors,” *Computer Graphics and Image Processing*, vol. 13, pp. 99-126, 1980.

40. G. L. Cash and M. Hatamian, “Optical Character Recognition by the Method of Moments,” *Computer Vision, Graphics, and Image Processing*, no. 39, pp. 291-310, 1987.