Moment Invariants in Image Analysis

Jan Flusser

Abstract—This paper aims to present a survey of object recognition/classification methods based on image moments. We review various types of moments (geometric moments, complex moments) and moment-based invariants with respect to various image degradations and distortions (rotation, scaling, affine transform, image blurring, etc.) which can be used as shape descriptors for classification. We explain a general theory how to construct these invariants and show also a few of them in explicit forms. We review efficient numerical algorithms that can be used for moment computation and demonstrate practical examples of using moment invariants in real applications.

Keywords—Object recognition, degraded images, moments, moment invariants, geometric invariants, invariants to convolution, moment computation.

I. INTRODUCTION

Analysis and interpretation of an image which was acquired by a real (i.e. non-ideal) imaging system is the key problem in many application areas such as remote sensing, astronomy and medicine, among others. Since real imaging systems as well as imaging conditions are usually imperfect, the observed image represents only a degraded version of the original scene. Various kinds of degradations (geometric as well as radiometric) are introduced into the image during the acquisition by such factors as imaging geometry, lens aberration, wrong focus, motion of the scene, systematic and random sensor errors, etc.

In the general case, the relation between the ideal image \( f(x, y) \) and the observed image \( g(x, y) \) is described as \( g = D(f) \), where \( D \) is a degradation operator. In the case of a linear shift-invariant imaging system, \( D \) has a form of

\[
g(\tau(x, y)) = (f * h)(x, y) + n(x, y),
\]

where \( h(x, y) \) is the point-spread function (PSF) of the system, \( n(x, y) \) is an additive random noise, \( \tau \) is a transform of spatial coordinates due to projective imaging geometry and \( * \) denotes a 2-D convolution. Knowing the image \( g(x, y) \), our objective is to analyze the unknown scene \( f(x, y) \).

By the term "scene analysis" we usually understand a complex process consisting of three basic stages. First, the image is segmented in order to extract objects of potential interest. Secondly, the extracted objects are "recognized", which means they are classified as elements of one class from the set of pre-defined object classes. Finally, spatial relations among the objects can be analyzed. In this tutorial, we focus on object recognition.

II. BRIEF HISTORY

The history of moment invariants began many years before the appearance of first computers, in the 19th century under the framework of the theory of algebraic invariants. The theory of algebraic invariants probably originates from famous German mathematician David Hilbert [1] and was thoroughly studied also in [2], [3].

Moment invariants were firstly introduced to the pattern recognition community in 1962 by Hu [4], who employed the results of the theory of algebraic invariants and derived his seven famous invariants to rotation of 2-D objects. Since
that time, numerous works have been devoted to various improvements and generalizations of Hu’s invariants and also to its use in many application areas.

Dudani [5] and Belkasim [6] described their application to aircraft silhouette recognition, Wong and Hall [7], Goshtasby [8] and Flusser and Suk [9] employed moment invariants in template matching and registration of satellite images, and many other authors used moment invariants for character recognition [6], [10]. Maitra [11] and Hupkens [12] made them invariant also to contrast changes, Wang [13] proposed illumination invariants particularly suitable for texture classification. Li [14] and Wong [15] presented the systems of invariants necessary to resolve practical object recognition tasks because of their descriptors were invariant to translation only. Despite this, the invariants have found successful applications in face recognition on out-of-focused photographs [34], in normalizing blurred images into the canonical forms [35], [36], in template-to-scene matching of satellite images [33], in blurred digit and character recognition [37], [13], in registration of images obtained by digital subtraction angiography [38] and in focus/defocus quantitative measurement [39]. Other sets of blur invariants (but still only shift-invariant) were proposed for some particular kinds of PSF — axisymmetric blur invariants [40] and motion blur invariants [41], [42]. A significant improvement motivated by a problem of registration of blurred images was made by Flusser et al. They introduced so-called combined blur-rotation invariants [43] and combined blur-affine invariants [44] and reported their successful usage in satellite image registration [45] and in camera motion estimation [46].

III. Basic Terms

First we define basic terms which will be then used in the construction of the invariants.

**Definition 1:** By image function (or image) we understand any real function \( f(x, y) \) having a bounded support and a finite nonzero integral.

**Definition 2:** Geometric moment \( m_{pq} \) of image \( f(x, y) \), where \( p, q \) are non-negative integers and \( (p + q) \) is called the order of the moment, is defined as

\[
m_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x, y) dx dy.
\]

Corresponding central moment \( \mu_{pq} \) and normalized moment \( \nu_{pq} \) are defined as

\[
\mu_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - x_c)^p (y - y_c)^q f(x, y) dx dy,
\]

\[
\nu_{pq} = \frac{\mu_{pq}}{\mu_{00}},
\]

respectively, where the coordinates \((x_c, y_c)\) denote the centroid of \( f(x, y) \), and \( \omega = (p + q + 2)/2 \).

**Definition 3:** Complex moment \( c_{pq} \) of image \( f(x, y) \) is defined as

\[
c_{pq} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q e^{i \omega} f(x, y) dx dy
\]

where \( i \) denotes imaginary unit. Definitions of central and normalized complex moments are analogous to (3) and (4).

Geometric moments and complex moments carry the same amount of information. Each complex moment can be expressed in terms of geometric moments as

\[
c_{pq} = \sum_{k=0}^{p} \sum_{j=0}^{q} \binom{p}{k} \binom{q}{j} (-1)^{p-j} \cdot i^{p-q+k-j} \cdot m_{k+j, p+q-k-j}
\]
and vice versa:

\[ m_{pq} = \frac{1}{2^{p+q+1}i} \sum_{k=0}^{p} \sum_{j=0}^{q} \left( \frac{p}{k} \right) \left( \frac{q}{j} \right) (-1)^{q-j} \cdot c_{k+j,p+q-k,j}. \]  

(7)

The reason for introducing complex moments is in their favorable behavior under image rotation, as will be shown later.

IV. INVARIANTS TO ROTATION, TRANSLATION, AND SCALING

Invariants to similarity transformation group were the first invariants that appeared in pattern recognition literature. It was caused partly because of their simplicity, partly because of great demand for invariant features that could be used in position-independent object classification. In this problem formulation, degradation operator \( D \) is supposed to act solely in spatial domain and to have a form of similarity transform. Eq (1) then reduces to

\[ g(\tau(x,y)) = f(x,y), \]

(8)

where \( \tau(x,y) \) denotes arbitrary rotation, translation, and scaling.

Invariants to translation and scaling are trivial – central and normalized moments themselves can play this role. As early as in 1962, M.K. Hu [4] published seven rotation invariants, consisting of second and third order moments (we present first four of them):

\[
\begin{align*}
\phi_1 &= \mu_{20} + \mu_{02}, \\
\phi_2 &= (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2, \\
\phi_3 &= (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2, \\
\phi_4 &= (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2.
\end{align*}
\]

(9)

The Hu’s invariants became classical and, despite of their drawbacks, they have found numerous successful applications in various areas. Major weakness of the Hu’s theory is that it does not provide for a possibility of any generalization. By means of it, we could not derive invariants from higher-order moments and invariants to more general transformations. These limitations were overcome thirty years later.

After Hu, there have been published various approaches to the theoretical derivation of moment-based rotation invariants. Li [14] used Fourier-Mellin transform, Teague [18] and Wallin [19] proposed to use Zernike moments, Wong [15] used complex monomials which originate from the theory of algebraic invariants, and Mostafa and Psaltis [24] employed complex moments. Here, we present a scheme introduced by Flusser [16], [17], which is based on the complex moments.

In polar coordinates, (5) becomes the form

\[ c_{pq} = \int_{0}^{2\pi} \int_{0}^{\pi} e^{i(p+q+1)(\alpha - \phi)} f(r, \theta) r dr d\theta. \]

(10)

It follows from the definition that \( c_{pq} = c_{qp}^* \) (the asterisk denotes complex conjugate). Furthermore, it follows immediately from (10) that the moment magnitude \( |c_{pq}| \) is invariant to rotation of the image while the phase is shifted by \((p-q)\alpha\), where \( \alpha \) is the angle of rotation. More precisely, it holds for the moment of the rotated image

\[ c_{pq}' = e^{-i(p-q)\alpha} \cdot c_{pq}. \]

(11)

Any approach to the construction of rotation invariants is based on a proper kind of phase cancellation. The simplest method proposed by many authors is to use the moment magnitudes themselves as the invariants. However, they do not generate a complete set of invariants. In the following Theorem, phase cancellation is achieved by multiplication of appropriate moment powers.

Theorem 1: Let \( n \geq 1 \) and let \( k_i, p_i, \) and \( q_i \) \((i = 1, \cdots, n)\) be non-negative integers such that

\[ \sum_{i=1}^{n} k_i(p_i - q_i) = 0. \]

Then

\[ I = \prod_{i=1}^{n} c_{p_i, q_i}^{k_i} \]

(12)

is invariant to rotation.

Theorem 1 allows us to construct an infinite number of the invariants for any order of moments, but only few of them are mutually independent. The knowledge of their basis is a crucial point because dependent features do not contribute to the discrimination power of the system at all and may even cause object misclassifications due to the “curse of dimensionality”.

Fundamental theorem on how to construct an invariant basis for a given set of moments was firstly formulated and proven in [16] and later in more general form (which is shown below) in [17].

Theorem 2: Let us consider complex moments up to the order \( r \geq 2 \). Let a set of rotation invariants \( B \) be constructed as follows:

\[ B = \{ \Phi(p,q) \equiv c_{pq} c_{p_0, q_0}^{p_0+q_0} | p \geq q \land p + q \leq r \}, \]

where \( p_0 \) and \( q_0 \) are arbitrary indices such that \( p_0 + q_0 \leq r \), \( p_0 - q_0 = 1 \) and \( c_{p_0, q_0} \neq 0 \) for all images involved. Then \( B \) is a basis of a set of all rotation invariants created from the moments up to the order \( r \).

Theorem 2 has a very surprising consequence. We can prove that, contrary to common belief, the Hu’s system is dependent and incomplete, so in fact it does not form a good feature set. The same is true for invariant sets proposed by Li [14] and Wong [15]. This result firstly appeared in [16] and has a deep practical impact.

V. INVARIANTS TO AFFINE TRANSFORM

In practice we often face object deformations that are beyond the rotation-translation-scaling model. An exact model of photographing a planar scene by a pin-hole camera whose optical axis is not perpendicular to the scene is projective transform of spatial coordinates. Since the projective transform is not linear, its Jacobian is a function of spatial coordinates.
and projective moment invariants from a finite number of moments cannot exist [49], [50].

For small objects and large camera-to-scene distance is the perspective effect negligible and the projective transform can be well approximated by affine transform

\[ x' = a_0 + a_1 x + a_2 y, \]
\[ y' = b_0 + b_1 x + b_2 y. \]

(13)

Thus, having powerful affine moment invariants for object description and recognition is in great demand.

A pioneer work on this field was done independently by Reiss [21] and Flusser and Suk [20], [48], who introduced affine moment invariants (AMI’s) and proved their applicability in simple recognition tasks. They derived only few invariants in explicit forms and they did not study the problem capability in simple recognition tasks. They derived only few

Reiss [21] and Flusser and Suk [20], [48], who introduced Thus, having powerful affine moment invariants for object

Note that it is meaningful to consider only

where

\[ w \]

integrals can be expressed in terms of moments and, after

After an affine transform it holds \( T'_{12} = J \cdot T_{12}, \) where \( J \) is the Jacobian of the transform. The basic idea of the AMI’s generating is the following. We consider various numbers of points and we integrate their cross-products (or some powers of their cross-products) on the support of \( f. \) These integrals can be expressed in terms of moments and, after eliminating the Jacobian by proper normalization, they yield affine invariants.

More precisely, having \( N \) points \((N \geq 2)\) we define functional \( I \) depending on \( N \) and on non-negative integers \( n_{kj} \) as

\[ I(f) = \int_{-\infty}^{\infty} \prod_{k,j=1}^{N} T_{kj}^{n_{kj}} \cdot \prod_{i=1}^{N} f(x_i, y_i) dx_i dy_i. \]

(14)

Note that it is meaningful to consider only \( j > k, \) because \( T_{kj} = -T_{jk} \) and \( T_{kk} = 0. \) After an affine transform, \( I \) becomes

\[ I' = I^w |J|^N \cdot I, \]

where \( w = \sum_{k,j} n_{kj} \) is called the weight of the invariant and \( N \) is called the degree of the invariant.

If \( I \) is normalized by \( \mu_{w+N} \) we get a desirable affine invariant

\[ \left( I \right)^{w+N} = \left( \mu_{w+N} \right)^I \]

(it w is odd and \( J < 0 \) there is an additional factor \(-1\)).

We illustrate the general formula (14) on two simple invariants. First, let \( N = 2 \) and \( n_{12} = 2. \) Then

\[ I(f) = \int_{-\infty}^{\infty} (x_1 y_2 - x_2 y_1)^2 f(x_1, y_1) f(x_2, y_2) dx_1 dy_1 dx_2 dy_2 \]

\[ = 2(m_{20} m_{02} - m_{11}^2). \]

Similarly, for \( N = 3 \) and \( n_{12} = 2, n_{13} = 2, n_{23} = 0 \) we get

\[ I(f) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 y_2 - x_2 y_1)^2 (x_1 y_3 - x_3 y_1)^2 f(x_1, y_1) f(x_2, y_2) f(x_3, y_3) dx_1 dy_1 dx_2 dy_2 dx_3 dy_3 \]

\[ = m_{20}^2 m_{04} + 4 m_{20} m_{11} m_{13} + 4 m_{20} m_{02} m_{22} + 4 m_{11}^2 m_{22} - 4 m_{11} m_{02} m_{31} + m_{02}^2 m_{40}. \]

(16)

The above idea has an analogy in graph theory. Each invariant generated by formula (14) can be represented by a graph, where each point \((x_i, y_i)\) corresponds to one node and each cross-product \( T_{kj} \) corresponds to one edge of the graph. If \( n_{kj} > 1, \) the respective term \( T_{kj}^{n_{kj}} \) corresponds to \( n_{kj} \) edges connecting \( k \)-th and \( j \)-th nodes. Thus, the number of nodes equals the degree of the invariant and the total number of the graph edges equals the weight \( w \) of the invariant. From the graph one can also learn about the orders of the moments the invariant is composed of and about its structure. The number of edges originating from each node equals the order of the moments involved. Each invariant of the form (14) is in fact a sum where each term is a product of certain number of moments. This number is constant for all terms of one invariant and is equal to the total number of the graph nodes. Particularly, for the invariants (15) and (16) the corresponding graphs are shown in Fig. 1.

![Fig. 1. The graphs corresponding to the invariants (15) (left) and (16) (right)](image)

Now one can see that the problem of derivation of the AMI’s up to the given weight \( w \) is equivalent to generating all graphs with at least two nodes and at most \( w \) edges. This is a combinatorial task with exponential complexity but formally easy to implement. Unfortunately, most resulting graphs are useless because they generate invariants, which are dependent. Identifying and discarding them is very important but very complicated task.

There might be various kinds of dependencies in the set of all AMI’s (i.e. in the set of all graphs). The invariant which equals to linear combinations of other invariants or of products of other invariants is called reducible invariant. Other invariants than reducible are called irreducible invariants. Unfortunately, "irreducible" does not mean "independent" – there may be higher-order polynomial dependencies among irreducible invariants. Current methods [51] perfectly eliminate reducible invariants but identification of dependencies among irreducible invariants has not been resolved yet.

For illustration, let us consider AMI’s up to the weight 10. Using the graph method we got, after discarding isomorphic (15) graphs, 1519 AMI’s in explicit forms. Then we applied the
algorithms eliminating reducible invariants, which led to 362 irreducible invariants.

VI. INVARIANTS TO CONVOLUTION

Two previous sections were devoted to the invariants with respect to transformation of spatial coordinates only. Now let us consider an imaging system with ideal geometry, i.e., $\tau(x, y) = (x, y)$, but suffering from non-ideal optical/radiometrical properties. Assuming the system is shift invariant, degradation operator $D$ has a form of

$$g(x, y) = (f * h)(x, y),$$

where $g(x, y)$ is the point-spread function (PSF) of the system. This is a simple but realistic model of degradations introduced by out-of-focused camera ($h(x, y)$ has then a cylindrical shape), by camera and/or scene motion ($h(x, y)$ has a form of rectangular pulse), and by photographing through turbulent medium ($h(x, y)$ is then a Gaussian), to name a few. However, in real applications the PSF has more complex form because it may be a composition of several degradation factors. Neither the shape nor the parameters of the PSF use to be known. This high-level uncertainty prevents us from solving eq. (17) as an inverse problem. Although such attempts were published (see [52] or [53] for a basic survey), they did not yield satisfactory results.

In this section, we present functional invariant to convolution with arbitrary centrosymmetric PSF (in image analysis literature they are often called “blur invariants” because common PSF’s have a character of a low-pass filter). Blur invariants were firstly introduced by Flusser and Suk [33]. They have found successful applications in face recognition on out-of-focused photographs [34], in normalizing blurred images into the canonical forms [35], [36], in template-to-scene matching of satellite images [33], in blurred digit and character recognition [37], [13], in registration of images obtained by digital subtraction angiography [38] and in focus/defocus quantitative angiography [39].

The assumption of centrosymmetry is not a significant limitation of practical utilization of the method. Most real sensors and imaging systems, both optical and non-optical ones, have the PSF with certain degree of symmetry. In many cases they have even higher symmetry than the central one, such as axial or radial symmetry.

Principal theorem on convolution invariants is the following.

**Theorem 3**: Let functional $C: L_1(R^2) \times N_0 \times N_0 \rightarrow \mathbb{R}$ be defined as follows:

If $(p + q)$ is even then

$$C(p, q)^{(f)} = 0.$$

If $(p + q)$ is odd then

$$C(p, q)^{(f)} = \mu_{pq}^{(f)} - \frac{1}{\mu_{pq}} \sum_{n=0}^{p} \sum_{m=0}^{q} \binom{p}{n} \binom{q}{m} C(p-n, q-m)^{(f)} \cdot \mu_{nm}^{(f)}.$$

Then

$$C(p, q)^{(f * h)} = C(p, q)^{(f)}$$

for any image function $f$, any non-negative integers $p$ and $q$, and for any centrosymmetric PSF $h$.

Theorem 3 tells that blur invariants are recursively defined functionals consisting mainly from odd-order moments. Although they do not have straightforward “physical” interpretation, let us make a few notes to provide a better insight into their meaning. Any invariant (even different from those presented here) to convolution with a centrosymmetric PSF must give a constant response on centrosymmetric images. This is because any centrosymmetric image can be considered as a blurring PSF acting on delta-function. It can be proven that if $f$ is centrosymmetric then $C(p, q)^{(f)} = 0$ for any $p$ and $q$. The opposite implication is valid as well. Thus, what image properties are reflected by the $C(p, q)$’s? Let us consider a Fourier-based decomposition $f = f_s + f_a$, where $f_s$, $f_a$ are centrosymmetric and antisymmetric components of $f$, respectively. Function $f_a$ can be exactly recovered from odd-order moments of $f$ (while even-order moments of $f_a$ equal zero) and vice versa. A similar relation holds for the invariants $C(p, q)$. Thus, all $C(p, q)$’s reflect mainly properties of the antisymmetric component of the image, while all symmetric images are in their null-space.

VII. ALGORITHMS FOR MOMENT COMPUTATION

Since computing complexity of all moment invariants depends almost solely on the computing complexity of geometric moments themselves, we review efficient algorithms for moment calculation in a discrete space. Most of the methods are focused on binary images but there are also a few methods for graylevel images. Basically, moment computation algorithms can be categorized into two groups: decomposition methods and boundary-based methods. The former methods decompose the object into simple areas (squares, rectangles, rows, etc.) whose moments can be calculated easily in $O(1)$ time. The object moment is then given as a sum of moments of all regions. The latter methods calculate object moments just from the boundary, employing Green’s theorem or similar technique.

In the discrete case, the integral in the moment definition must be replaced by a summation. The most common way (but not the only one) how to do that is to employ the rectangular (i.e. zero-order) method of numeric integration. Then (2) turns to the well-known form

$$m_{pq} = \sum_{z=1}^{N} \sum_{y=1}^{N} x^p y^q f_{ij},$$

where $N$ is the size of the image and $f_{ij}$ are the grey levels of individual pixels.

VIII. CONCLUSION

This paper presented a review of moment-based invariant functionals, their history, basic principles, and methods how to construct them. We demonstrated that invariant functionals can be used in image analysis as features for description and recognition of objects in degraded images.

Invariant-based approach is a significant step towards robust and reliable object recognition methods. It has a deep practical impact because many pattern recognition problems would
not be solvable otherwise. In practice, image acquisition is always degraded by unrecoverable errors and the knowledge of invariants with respect to these errors is a crucial point. This observation should influence future research directions and should be also incorporated in the education.

REFERENCES


[9] J. Flusser and T. Suk, “A moment-based approach to registration of invariants with respect to these errors is a crucial point. In practice, image acquisition is always degraded by unrecoverable errors and the knowledge of invariants with respect to these errors is a crucial point. This observation should influence future research directions and should be also incorporated in the education.

REFERENCES


[9] J. Flusser and T. Suk, “A moment-based approach to registration of invariants with respect to these errors is a crucial point. In practice, image acquisition is always degraded by unrecoverable errors and the knowledge of invariants with respect to these errors is a crucial point. This observation should influence future research directions and should be also incorporated in the education.

REFERENCES


