Akademie věd
České republiky

Teze disertace
k získání vědeckého titulu "doktor věd"
ve skupině věd: fyzikálně-matematické vědy

2D and 3D Image Analysis by Moments
Analýza obrazu pomocí momentů ve 2D a ve 3D

název disertace

Komise pro obhajoby doktorských disertací v oboru informatika a kybernetika

Jméno uchazeče: Ing. Tomáš Suk, CSc.

Pracoviště uchazeče: Ústav teorie informace a automatizace AV ČR, v.v.i.

Místo a datum: Praha 13.9.2016

Contents

1 Introduction **6**
- 1.1 Image analysis by computers 6
- 1.2 Invariants 9
- 1.3 Moments 10

2 2D Moment Invariants to Translation, Rotation, and Scaling **11**
- 2.1 Invariants to translation and scaling 11
- 2.2 Invariants to rotation 11
- 2.3 Moment invariants of vector fields 12
- 2.4 Rotation invariants of symmetric objects 13

3 3D Moment Invariants to Translation, Rotation, and Scaling **13**
- 3.1 Tensor method 13
- 3.2 The number of the independent invariants 14
- 3.3 Rotation invariants from 3D complex moments 15
- 3.4 Invariants of symmetric objects 17

4 Affine moment invariants **17**
- 4.1 Affine moment invariants generated by graphs 17
- 4.2 Derivation of the affine moment invariants from the Cayley-Aronhold equation 18

5 Orthogonal Moments **19**
- 5.1 2D moments orthogonal on a square 19
- 5.2 Orthogonal moments of a discrete variable 20
- 5.3 2D moments orthogonal on a disk 20
- 5.4 Rotation invariants from OG moments 21
- 5.5 3D orthogonal moments 21

6 Algorithms for Moment Computation **21**
- 6.1 Computing binary OG moments by means of decomposition methods 22
- 6.2 Geometric moments of graylevel images 22

7 Applications **24**
- 7.1 Image understanding 24
- 7.2 Image registration 25
- 7.3 Other applications 25

8 Contribution of the Thesis **28**
- 8.1 Outlook to the future 29

Resumé

The book [1] submitted as the dissertation thesis deals systematically with moments and moment invariants of 2D and 3D images and with their use in object description, recognition, and in other applications. It introduces basic terms and concepts of object recognition such as feature spaces and related norms, equivalences, and space partitions. Invariant based approaches are introduced together with normalization methods. Eight basic categories of invariants are reviewed together with illustrative examples of their use. It also recalls main supervised and unsupervised classifiers as well as related topics of classifier fusion and reduction of the feature space dimensionality.

First, 2D moment invariants with respect to the simplest spatial transformations – translation, rotation, and scaling are derived. A general method for constructing invariants of arbitrary orders by means of complex moments is presented. We prove the existence of a relatively small basis of invariants that is complete and independent. We also show an alternative approach – constructing invariants via normalization. We discuss the difficulties with the recognition of symmetric objects poses and present moment invariants suitable for such cases.

3D moment invariants with respect to translation, rotation, and scaling are introduced. We present the derivation of the invariants by means of three approaches – the tensor method, the expansion into spherical harmonics, and the object normalization. Similarly as in 2D, the symmetry issues are also discussed there.

Four main approaches showing how to derive moment invariants to the affine transformation of spatial coordinates in 2D are presented – the graph method, the method of normalized moments, the transvectants, and the solution of the Cayley-Aronhold equation. Relationships between the invariants produced by different methods are mentioned, and the dependency among the invariants is studied. We describe a technique used for elimination of reducible and dependent invariants. Numerical experiments illustrating the performance of the affine moment invariants are carried out, and a generalization to color images, vector fields, and 3D images is proposed.

A completely different kind of moment invariants is invariants to image blurring. We introduce the theory of projection operators, which allows us to derive invariants with respect to image blur regardless of the particular convolution kernel, provided it has a certain type of symmetry. We also derive so-called combined invariants, which are invariant to composite geometric and blur degradations. Knowing these features, we can recognize objects in the degraded scene without any restoration/deblurring.

Various types of orthogonal moments are presented. The moments orthogonal on a rectangle/cube as well as the moments orthogonal on a unit disk/sphere are described. We review Legendre, Chebyshev, Gegenbauer, Jacobi, Laguerre,

Gaussian-Hermite, Krawtchouk, dual Hahn, Racah, Zernike, pseudo-Zernike, and Fourier-Mellin polynomials and moments. The use of the moments orthogonal on a disk in the capacity of rotation invariants is discussed. Image reconstruction of an image from its moments is discussed. We explain, why orthogonal moments are more suitable for reconstruction than geometric ones, and a comparison of reconstructing power of different orthogonal moments is presented.

Computational issues are studied. Since the computing complexity of all moment invariants is determined by the computing complexity of the moments, efficient algorithms for moment calculations are of prime importance. There are basically two major groups of methods for moment computation. The first one consists of methods that attempt to decompose the object into non-overlapping regions of a simple shape, the moments of which can be computed very fast. The moment of the object is then calculated as a sum of the moments of all regions. The other group is based on Green's theorem, which evaluates the integral over the object by means of a less-dimensional integration over the object boundary. We present efficient algorithms for binary and graylevel objects and for geometric as well as for selected orthogonal moments.

Various applications of moments and moment invariants in image analysis are mentioned. We demonstrate the use of the moments in image registration, object recognition, medical imaging, content-based image retrieval, focus/defocus measurement, forensic applications, robot navigation, digital watermarking, and others.

1 Introduction

This thesis is brief guide to the book [1] submitted as the dissertation thesis. First, basic terms as "image analysis", "invariants", and "moments" are introduced.

1.1 Image analysis by computers

Image analysis in a broad sense is a many-step process the input of which is an image while the output is a final (usually symbolic) piece of information or a decision. Typical examples are localization of human faces in the scene and recognition (against a list or a database of persons) who is who, finding and recognizing road signs in the visual field of the driver, and identification of suspect tissues in a CT or MRI image. A general image analysis flowchart is shown in Fig. 1 and an example what the respective steps look like in the car licence plate recognition is shown in Fig. 2.

The first three steps of image analysis – image acquisition, preprocessing, and object segmentation/detection – are comprehensively covered in classical image processing textbooks [2–6], in recent specialized monographs [7–9] and in thousands of research papers. This work is devoted to one particular family of features which are based on image moments.

In *image acquisition*, the main theoretical questions are how to choose the sampling scheme, the sampling frequency and the number of the quantization levels such that the artifacts caused by aliasing, moire, and quantization noise do not degrade the image much while keeping the image size reasonably low (there are of course also many technical questions about the appropriate choice of the camera and the spectral band, the objective, the memory card, the transmission line, the storage format, and so forth, which we do not discuss here).

Since real imaging systems as well as imaging conditions are usually imperfect, the acquired image represents only a degraded version of the original scene. Various kinds of degradations (geometric as well as graylevel/color) are introduced into the image during the acquisition process by such factors as imaging geometry, lens aberration, wrong focus, motion of the scene, systematic and random sensor errors, noise, etc. (see Fig. 3 for the general scheme and an illustrative example). Removal or at least suppression of these degradations is a subject of *image preprocessing*. Historically, image preprocessing was one of the first topics systematically studied in digital image processing (already in the very first monograph [10] there was a chapter devoted to this topic) because even simple preprocessing methods were able to enhance the visual quality of the images and were feasible on old computers. The first two steps, image acquisition and preprocessing, are in the literature often categorized into *low-level processing*. The characteristic feature of the low-level methods is that

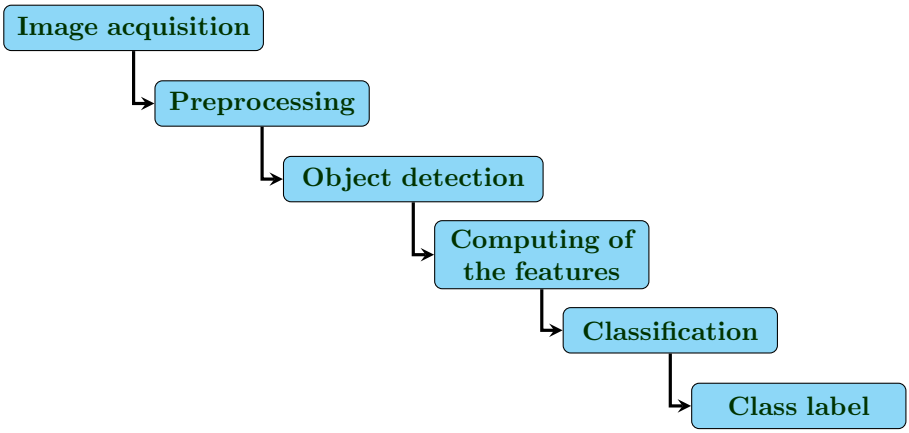


Figure 1: General image analysis flowchart.

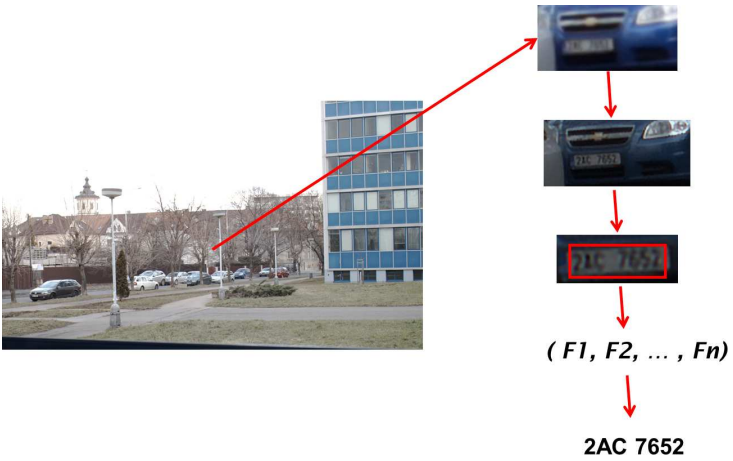


Figure 2: An example of the car licence plate recognition.

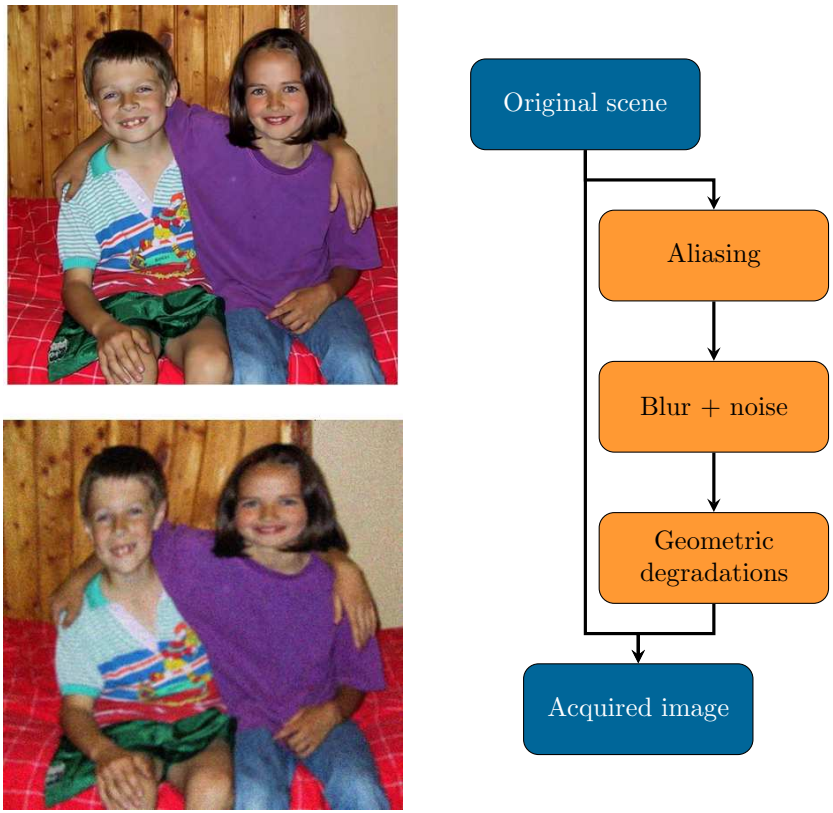


Figure 3: Image acquisition process with degradations.

both their input and output are digital images. On the contrary, in *high-level processing*, the input is a digital image (often an output of some preprocessing) while the output is a symbolic (i.e. high-level) information, such as the coordinates of the detected objects, the list of boundary pixels, etc.

Object detection is a typical example of high-level processing. The goal is to localize the objects of interest in the image and separate (segment) them from the other objects and from the background. Hundreds of segmentation methods have been described in the literature. Some of them are universal, but most of them were designed for specific families of objects such as characters, logos, cars, faces, human silhouettes, roads, etc. A good basic survey of object detection and segmentation methods can be found in [6] and in the references therein.

Feature definition and computing is probably the most challenging part of image analysis. The features should provide an accurate and unambiguous quantitative description of the objects. The feature values are elements of the *feature space* which should be for the sake of efficient computation of low dimensionality. The design of the features is highly dependent on the type of objects, on the conditions under which the images have been acquired, on the type and the quality of preprocessing, and on the application area. There is no unique "optimal" solution.

Classification/recognition of the object is the last stage of the image analysis pipeline. It is entirely performed in the feature space. Each unknown object, now being represented by a point in the feature space, is classified as an element of a certain class. The classes can be specified in advance by their representative samples, which create a *training set*, in such a case we speak about *supervised classification*. Alternatively, if no training set is available, the classes are formed from the unknown objects based on their distribution in the feature space. This case is called *unsupervised classification* or *clustering*, and in visual object recognition this approach is rare. Unlike the feature design, the classification algorithms are application independent – they consider neither the nature of the original data nor the physical meaning of the features. Classification methods are comprehensively reviewed in the famous Duda-Hart-Stork book [11] and in the most recent monograph [12]. The use of the classification methods is not restricted to image analysis. We can find numerous applications in artificial intelligence, decision making, social sciences, statistical data analysis, and in many other areas beyond the scope of this book.

1.2 Invariants

The differences among objects in one class can be described by some operator $\mathcal{D}(f)$ of the image f. The invariant is a functional $I(f)$, which depends on the class of the object in f, but does not depend on particular appearance of f. It means $I(f) = I(\mathcal{D}(f))$ for any f and any instance of $\mathcal{D}$.

The operator $\mathcal{D}(f)$ can be geometric transformation. Simple examples of the invariants under a similarity transformation are then angles and ratios, the object size is invariant to rotation, and the property of two lines "to be parallel" is invariant under affine transformation.

The invariant should also be discriminative, i.e. it should have distinct values on objects that belong to distinct classes. The trivial feature $I(f) = 0$ is perfectly invariant, but completely useless.

Other important properties of the invariants are independency and completeness. The set of invariants should not include any invariant dependent on others that does not bring any new information about the object. The system of the invariants should be so extendable that if we theoretically know all their values, we would be able to reconstruct the object (up to the operator $\mathcal{D}(f)$).

1.3 Moments

The moment $M_{\mathbf{p}}^{(f)}$ of image f is defined as

$$M_{\mathbf{p}}^{(f)} = \int_{\Omega} \pi_{\mathbf{p}}(\mathbf{x}) f(\mathbf{x}) \mathrm{d}\mathbf{x}, \tag{1}$$

where $\{\pi_{\mathbf{p}}(\mathbf{x})\}$ is a d-variable polynomial basis of the space of image functions defined on Ω and $\mathbf{p} = (p_1, \cdots, p_d)$ is a multi-index of non-negative integers, which show the highest power of the respective variables in $\pi_{\mathbf{p}}(\mathbf{x})$. The number $|\mathbf{p}| = \sum_{k=1}^{d} p_k$ is called the order of the moment. Depending on the polynomial basis, we recognize various systems of moments. The standard power basis $\pi_{\mathbf{p}}(\mathbf{x}) = \mathbf{x}^{\mathbf{P}}$ leads to geometric moments. Other important types of the moments are complex moments and various orthogonal moments.

Another way of obtaining features is integral transformation (e.g. Fourier transformation, Laplace transformation, etc.). Its definition is formally similar to (1), but the indices (parameters) p_k are real values and $\{\pi_{\mathbf{p}}(\mathbf{x})\}$ are some general functions.

Between the moments and the integral transformations, there is a terminological gap. The term for the case, when p_k are integers and $\{\pi_{\mathbf{p}}(\mathbf{x})\}$ are not polynomials, is missing. Therefore some authors extend the definition to various "moments in wider sense". The basis functions then include various scalar factors and weighting functions in the integrand or even completely non-polynomial functions.

2 2D Moment Invariants to Translation, Rotation, and Scaling

The invariance with respect to translation, rotation, and scaling (TRS – sometimes called similarity transformation) is widely required in almost all practical applications, because the object should be correctly recognized regardless of its particular position and orientation in the scene and of the object-to-camera distance. We can use both geometric moments and complex moments for the construction of TRS invariants. The 2D geometric moments are defined

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)\mathrm{d}x\mathrm{d}y. \tag{2}$$

2.1 Invariants to translation and scaling

Invariance to translation can be achieved by shifting the coordinate origin into the object centroid. Such moments are called central

$$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-x_c)^p (y-y_c)^q f(x,y)\mathrm{d}x\mathrm{d}y, \tag{3}$$

where $x_c = m_{10}/m_{00}$, $y_c = m_{01}/m_{00}$ are the coordinates of the centroid.

Scaling invariance is obtained by a proper normalization of each moment, we normalize most often by a proper power of μ_{00}

$$\nu_{pq} = \frac{\mu_{pq}}{\mu_{00}^{w}}, \qquad \text{where} \qquad w = \frac{p+q}{2} + 1. \tag{4}$$

2.2 Invariants to rotation

The complex moment c_{pq} is obtained, when we choose the polynomial basis of complex monomials

$$c_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+iy)^p (x-iy)^q f(x,y)\mathrm{d}x\mathrm{d}y = \int_{0}^{\infty}\int_{0}^{2\pi} r^{p+q+1} e^{i(p-q)\theta} f(r,\theta)\mathrm{d}\theta\mathrm{d}r. \tag{5}$$

The rotation invariants can be constructed from them as an arbitrary product, where the sum of first indices equals the sum of the second indices. The complete and independent set (basis) of 2D rotation invariants can be constructed as

$$\Phi(p,q) = c_{pq} c_{q_0 p_0}^{p-q}, \tag{6}$$

where $p \geq q$, $p + q \leq r$, 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 admissible images. Usually, we use $p_0 = 2$ and $q_0 = 1$. An example of the set composed from the moments of the 2nd and 3rd orders is

$$\Phi(1,1) = c_{11},\ \Phi(2,1) = c_{21}c_{12},\ \Phi(2,0) = c_{20}c_{12}^2,\ \Phi(3,0) = c_{30}c_{12}^3. \qquad (7)$$

It contains six real-valued invariants, $\Phi(2,0)$ and $\Phi(3,0)$ are complex. There are other possibilities, how to achieve the rotation invariance, e.g. tensor method from geometric moments or normalization.

2.3 Moment invariants of vector fields

Some images have more than one value in each pixel, for example color images, multispectral images, or vector fields. The vector fields are often used for description of various flows, Fig. 4 is an example of wind. The vector fields are specific by connection of the rotation in coordinates with the rotation of the vectors (so called *total rotation*), we must then modify the construction of the invariants. The complex moments are computed from the x and y components

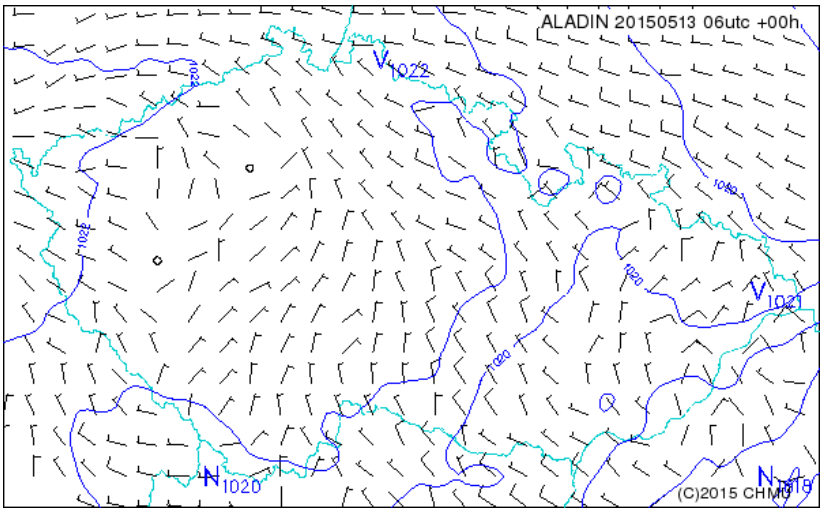


Figure 4: The wind velocity forecast for the Czech Republic (courtesy of the Czech Hydrometeorological Institute, the numerical model Aladin). The longer dash is constant part of the wind, the shorter dash expresses perpendicular squalls. The figure actually represents two vector fields.

of the vector field $\mathbf{f}(\mathbf{x}) = (f_1(x,y),\ f_2(x,y))^T$ as

$$c_{pq}^{(\mathbf{f})} = c_{pq}^{(f_1)} + i c_{pq}^{(f_2)}.$$

The sum of the second indices in the product must then be higher than the sum of the first indices by the number of the moments, e.g. c_{12}, $c_{00}c_{02}$, $c_{10}c_{02}^2$ or $c_{20}c_{02}^3$. The moment c_{00} is not invariant to the total rotation of the vector field, it cannot be used as a sole normalization factor. Instead, we have to use some rotation invariant for normalization, e.g. $(c_{00}c_{02})^{(p+q+3)/8}$.

2.4 Rotation invariants of symmetric objects

In many applied tasks, we want to classify shapes having some kind of symmetry. Different classes may or may not have the same symmetry, which makes the task even more difficult. Moment-based classifiers suffer from the fact that some moments of symmetric shapes are zero and corresponding invariants do not provide any discrimination power. That is why it is necessary to design special invariants for each type of symmetry [13].

If $f(x, y)$ has an N–fold rotation symmetry (N finite), we obtain non-zero c_{pq} only when $(p - q)/N$ is an integer. If $f(x, y)$ has circular symmetry, then only c_{pp} are non-zero. We should only use the invariants that bring some information; the invariants that are zero for all admissible objects should be omitted.

3 3D Moment Invariants to Translation, Rotation, and Scaling

Sensors of 3D images can be divided into two groups. Medical imaging devices (MRI, CT) provide full 3D graylevel images of the object represented by a 3D array of voxels. The rangefinders measure only the distance of each point on the object surface from the sensor and cannot see inside the object. The result is 3D binary image often represented as a triangulated surface.

While translation and scaling invariance of 3D moments can be achieved by similar way as in 2D, i.e. by normalization to scaling by zeroth-order moment and by normalization to translation by first-order moments, achievement of rotation invariance is more difficult. We can use either tensor method or complex moments.

3.1 Tensor method

The tensor method [14] is based on moment tensor

$$M^{i_1 i_2 \cdots i_k} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^{i_1} x^{i_2} \cdots x^{i_k} f(x^1, x^2, x^3) \mathrm{d}x^1 \mathrm{d}x^2 \mathrm{d}x^3, \qquad (8)$$

where $x^1 = x$, $x^2 = y$ and $x^3 = z$. If p indices equal 1, q indices equal 2 and r indices equal 3, then $M^{i_1 i_2 \cdots i_k} = m_{pqr}$. The definition can be easily modified to the different number of dimensions.

Generally, we can distinct two types of tensors, affine and Cartesian. The affine tensors describe behavior under affine transformation. They have two types of indices: covariant and contravariant. They are suitable for derivation of affine invariants. The Cartesian tensors have one type of indices and they are suitable for derivation of rotation invariants (generally invariants to orthogonal transformation). The advantage of the tensor method is easy use in any number of dimensions.

The moment tensor can be used both as affine (then it has all indices contravariant) and as Cartesian. Then each total contraction of a Cartesian tensor or of a product of Cartesian tensors is rotation invariant. We can compute the total contraction of the moment tensor product, when each index is included just twice in it.

A simple example of such total contraction is M_{ii}. The corresponding invariant is

$$\Phi_1^{3D} = (\mu_{200} + \mu_{020} + \mu_{002})/\mu_{000}^{5/3}.$$

Each invariant has its 2D counterpart. Here it is

$$\Phi_1^{2D} = (\mu_{20} + \mu_{02})/\mu_{00}^2.$$

Another example is $M_{ij}M_{ij}$, which yields

$$\Phi_2^{3D} = (\mu_{200}^2 + \mu_{020}^2 + \mu_{002}^2 + 2\mu_{110}^2 + 2\mu_{101}^2 + 2\mu_{011}^2)/\mu_{000}^{10/3}.$$

Its 2D counterpart is

$$\Phi_2^{2D} = (\mu_{20}^2 + \mu_{02}^2 + 2\mu_{11}^2)/\mu_{00}^4.$$

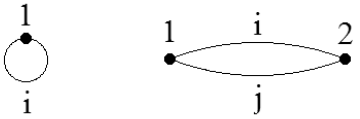
Each invariant generated by the tensor method can be represented by a connected graph, where each moment tensor corresponds to a node and each index corresponds to an edge of the graph. The graphs can include self-loops, see Fig. 5a.

3.2 The number of the independent invariants

An intuitive rule suggests that the number n_i of independent invariants created from n_m independent measurements (i.e. moments in our case) is

$$n_i = n_m - n_p, \tag{9}$$

where n_p is the number of independent constraints that must be satisfied. In most cases, we estimate n_p as the number of transformation parameters. In the



(a) (b)

Figure 5: The generating graphs of (a) both Φ_1^{2D} and Φ_1^{3D} and (b) both Φ_2^{2D} and Φ_2^{3D}.

case of 2D TRS, $n_p = 4$ (scaling, two translations and rotation). In the case of 3D, $n_p = 7$ (scaling, three translations and three rotations). The number of moments of orders from 0 to r equals

$$n_m = \binom{r+d}{d}, \tag{10}$$

where d is the number of dimensions of the space, i.e. $d = 2$ in 2D and $d = 3$ in 3D. Ideally, we should only use just n_i independent invariants, but sometimes, it is difficult to prove their independence.

3.3 Rotation invariants from 3D complex moments

3D complex moments provide an alternative to the tensor method [15]. The definition (5) of 2D complex moments in polar coordinates contains the harmonic angular function $e^{i(p-q)\theta}$. In 3D, analogous functions are spherical harmonics. They are defined as

$$Y_\ell^m(\theta,\varphi) = \sqrt{\frac{(2\ell+1)}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}} P_\ell^m(\cos\theta) e^{im\varphi}, \tag{11}$$

where the degree $\ell = 0, 1, 2, \ldots$, the order $m = -\ell, -\ell+1, \ldots, \ell$, and $P_\ell^m(x)$ is an associated Legendre function. The first few spherical harmonics can be seen in Fig. 6. The 3D complex moments are defined as

$$c_{s\ell}^m = \int_0^{2\pi}\int_0^{\pi}\int_0^{\infty} \varrho^{s+2} Y_\ell^m(\theta,\varphi) f(\varrho,\theta,\varphi) \sin\theta \mathrm{d}\varrho\mathrm{d}\theta\mathrm{d}\varphi, \tag{12}$$

where s is the order of the moment, ℓ is latitudinal repetition and m is longitudinal repetition.

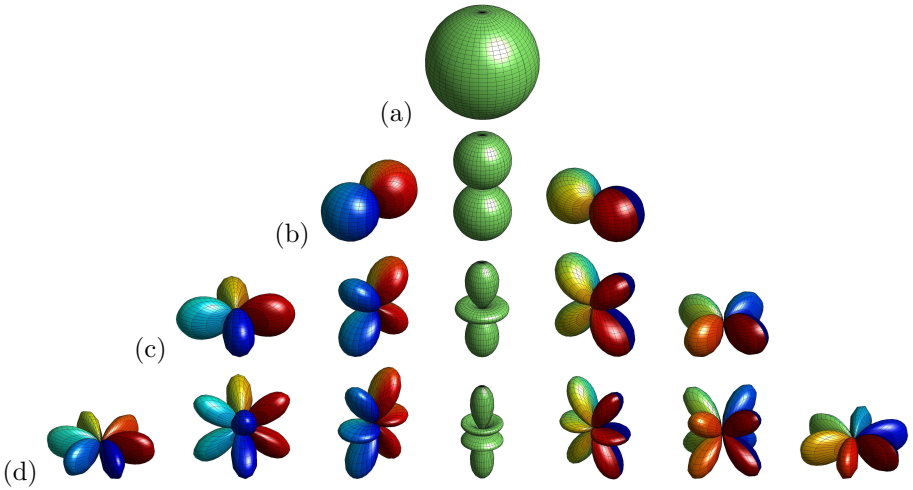


Figure 6: The spherical harmonics. (a) $\ell = 0$, $m = 0$, (b) $\ell = 1$, $m = -1, 0, 1$, (c) $\ell = 2$, $m = -2, -1, 0, 1, 2$, (d) $\ell = 3$, $m = -3, -2, -1, 0, 1, 2, 3$. Imaginary parts are displayed for $m < 0$ and real parts for $m \geq 0$.

The moment c^0_{s0} itself is an invariant. For other invariants we have to construct composite complex moment forms

$$c_s(\ell, \ell')^k_j = \sum_{m=\max(-\ell, k-\ell')}^{\min(\ell, k+\ell')} \langle \ell, \ell', m, k-m | j, k \rangle \, c^m_{s\ell} \, c^{k-m}_{s\ell'}, \tag{13}$$

where $\langle \ell, \ell', m, k-m | j, k \rangle$ are Clebsch-Gordan coefficients. The form $c_s(\ell, \ell)^0_0$ is then 3D rotation invariant. We can further combine the composite complex moment forms and the complex moments

$$c_s(\ell, \ell')_j c_{s'} = \frac{1}{\sqrt{2j+1}} \sum_{k=-j}^{j} (-1)^{j-k} c_s(\ell, \ell')^k_j c^{-k}_{s'j} \tag{14}$$

or two composite complex moment forms

$$c_s(\ell, \ell')_j c_{s'}(\ell'', \ell''')_j = \frac{1}{\sqrt{2j+1}} \sum_{k=-j}^{j} (-1)^{j-k} c_s(\ell, \ell')^k_j c_{s'}(\ell'', \ell''')^{-k}_j. \tag{15}$$

An example of the feature set for the orders two and three is c^0_{20}, $c_2(2,2)^0_0$, $c_2(2,2)_2 c_2$, $c_3(3,3)^0_0$, $c_3(1,1)^0_0$, $c_3(3,3)_2 c_2$, $c_3(3,1)_2 c_2$, $c_3(1,1)_2 c_2$, $c_3(3,3)_2 c_2(2,2)_2$, $c^2_3(3,3)_2$, $c_3(3,3)_2 c_3(3,1)_2$, $c_3(3,3)_2 c_3(1,1)_2$, and $c_3(3,1)_2 c_3(1,1)_2$.

3.4 Invariants of symmetric objects

Similarly as in 2D, the 3D objects having some kind of symmetry often cause difficulties, because some of their moments are zero. While in 2D, we have 2 infinite sequences (rotational and dihedral) and the circular symmetry, in 3D, there are 7 infinite sequences (only some of them have special names as prismatic or anti-prismatic), 7 separate groups of symmetries of polyhedra (e.g. tetrahedral or octahedral) and 3 infinite symmetries (conic, cylindrical and spherical). A detailed comprehensive study of this problem was published in [16].

4 Affine moment invariants

The affine moment invariants are invariant with respect to affine transformation of the spatial coordinates

$$\mathbf{x}' = \mathbf{A}\mathbf{x} + \mathbf{b}. \tag{16}$$

Most of the traditional photographs are central projections of 3D world onto a plane, which is modeled by a perspective projection. The perspective projection is hard to work with and the affine transformation often serves as a reasonably accurate and simple approximation.

4.1 Affine moment invariants generated by graphs

The graph method [17] can be used for generation of the affine moment invariants. It is based on integration of "cross-products"

$$C_{12} = x_1 y_2 - x_2 y_1$$

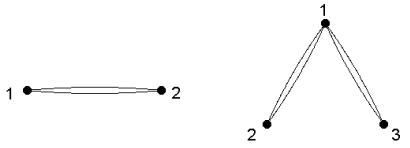
of pairs of points (x_1, y_1), (x_2, y_2) over an image. These integrals can be expressed in terms of moments and they yield affine invariants. Having $r > 1$ points $(x_1, y_1), \cdots, (x_r, y_r)$, we define functional I depending on r and on non-negative integers n_{kj} as

$$I(f) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{k,j=1}^{r} C_{kj}^{n_{kj}} \cdot \prod_{i=1}^{r} f(x_i, y_i) \mathrm{d}x_i \mathrm{d}y_i. \tag{17}$$

After scaling normalization, $I(f)$ is the affine invariant.

Two examples: for $r = 2$ and $n_{12} = 2$

$$\begin{aligned} I(f) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1 y_2 - x_2 y_1)^2 f(x_1, y_1) f(x_2, y_2) \mathrm{d}x_1 \mathrm{d}y_1 \mathrm{d}x_2 \mathrm{d}y_2 = \\ &= 2(m_{20} m_{02} - m_{11}^2). \end{aligned} \tag{18}$$



(a) (b)

Figure 7: The graphs corresponding to the invariants (a) (18) and (b) (19).

The corresponding affine invariant is $I_1 = (\mu_{20}\mu_{02} - \mu_{11}^2)/\mu_{00}^4$. Second example: for $r = 3$ and $n_{12} = 2, n_{13} = 2, n_{23} = 0$ we obtain

$$I(f) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_1y_2 - x_2y_1)^2(x_1y_3 - x_3y_1)^2 \\ f(x_1, y_1)f(x_2, y_2)f(x_3, y_3)\mathrm{d}x_1\mathrm{d}y_1\mathrm{d}x_2\mathrm{d}y_2\mathrm{d}x_3\mathrm{d}y_3 = \\ = m_{20}^2 m_{04} - 4m_{20}m_{11}m_{13} + 2m_{20}m_{02}m_{22} + 4m_{11}^2 m_{22} - 4m_{11}m_{02}m_{31} + \\ + m_{02}^2 m_{40}. \tag{19}$$

The normalizing factor is in this case μ_{00}^7.

The affine moment invariant generated by the formula (17) can be represented by a graph, where each point (x_k, y_k) corresponds to a node and each cross-product C_{kj} corresponds to an edge of the graph. The graphs corresponding to our examples are in Fig. 7. Unlike the 3D rotation invariants, this graph representation cannot contain self-loops.

4.2 Derivation of the affine moment invariants from the Cayley-Aronhold equation

There is another method for derivation of the affine moment invariants [18]. It relies on a decomposition of the affine transformation into elementary transformations: horizontal and vertical translation, scaling, stretching, horizontal and vertical skewing and possible mirror reflection. Each of these transformations imposes one constraint on the invariants. The invariance to translation and scaling can be achieved traditionally by centroid and m_{00}. The invariance to stretching is provided by the constraint that the sum of the first indices in each term equals the sum of the second indices.

If a function I should be invariant to the horizontal skew, then its derivative with respect to the skew parameter t must be zero. From this constraint, we

can derive the Cayley-Aronhold differential equation

$$\sum_p \sum_q p\mu_{p-1,q+1} \frac{\partial I}{\partial \mu_{pq}} = 0. \tag{20}$$

We can obtain the affine moment invariant as a solution of this equation.

5 Orthogonal Moments

Neither the geometric nor the complex moments are orthogonal, what leads to high correlation between the moments of adjacent orders. The most significant digits of the geometric and complex moments are highly correlated with the moments of lower orders, while their information contribution is concentrated in the less significant digits. If their computer representation is not sufficiently precise, these digits are forgotten and their contribution is lost. The solution in some applications can be use of orthogonal (OG) moments.

The OG moments use an orthogonal polynomials as the basis functions. Polynomial basis $\{\pi_{\mathbf{p}}(\mathbf{x})\}$ is orthogonal (possibly with weight function $w(\mathbf{x})$) if

$$\int_\Omega w(\mathbf{x})\pi_{\mathbf{p}}(\mathbf{x})\pi_{\mathbf{q}}(\mathbf{x})\mathrm{d}\mathbf{x} = \begin{cases} 0 & \text{for } \mathbf{p} \neq \mathbf{q} \\ n_g(\mathbf{p}) & \text{for } \mathbf{p} = \mathbf{q}. \end{cases} \tag{21}$$

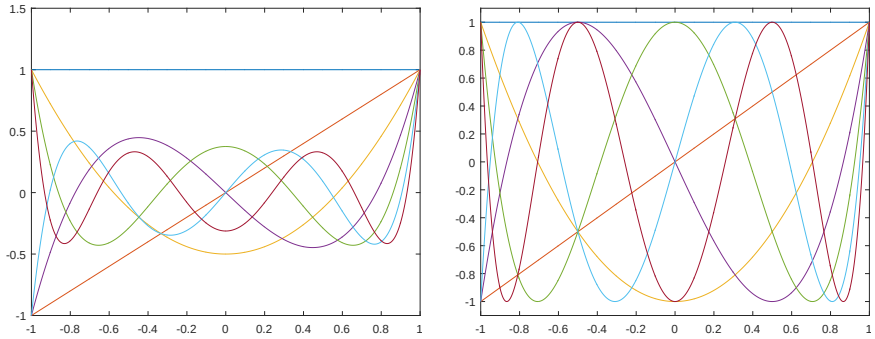
If $n_g(\mathbf{p}) = 1$ for any $\mathbf{p}$, the polynomial basis is called orthonormal. The set $\Omega \subset \mathbb{R}^d$ is called the area of orthogonality. We can divide OG moments into two basic groups. The polynomials orthogonal on a square/cube and the polynomials orthogonal on a disk/sphere.

5.1 2D moments orthogonal on a square

These moments are based on products of 1D polynomials in x and y

$$M_{pq} = n_p n_q \iint_\Omega \pi_p(x)\pi_q(y) f(x,y)\mathrm{d}x\mathrm{d}y. \tag{22}$$

There is a group of polynomials orthogonal on a finite interval (usually $\langle -1, 1\rangle$): Legendre polynomials (see Fig. 8a) with unit weight function $w(x) = 1$, Chebyshev polynomials of first (they have unit amplitudes, see Fig. 8b) and second kinds with weight functions $1/\sqrt{1-x^2}$, respectively $\sqrt{1-x^2}$. Their generalization is Gegenbauer polynomials with one parameter.The Jacobi polynomials are further generalization with two parameters. Laguerre polynomials are specific by the area of orthogonality $\langle 0, \infty\rangle$. Hermite polynomials orthogonal on $\langle -\infty, \infty\rangle$ are more important in image processing.



(a) (b)

Figure 8: The graphs of the OG polynomials up to the 6th degree: (a) Legendre polynomials, (b) Chebyshev polynomials of the first kind.

5.2 Orthogonal moments of a discrete variable

A special group of polynomials fulfill the condition of discrete orthogonality. Given N points $x_1 < x_2 < \ldots < x_N$, the discrete orthogonality condition in 1D is

$$\sum_{\ell=1}^{N} w_\ell \, \pi_m(x_\ell) \, \pi_n(x_\ell) \; = 0 \text{ for } m \neq n. \tag{23}$$

The discrete moments are then defined as

$$M_p = \sum_{\ell=1}^{N} \pi_p(x_\ell) f(x_\ell).$$

Since the points x_ℓ may stand for sampling grid nodes, the discrete orthogonality is particularly suitable in digital signal/image analysis applications. The examples of the discrete OG moments used in image analysis are discrete Chebyshev moments, Krawtchouk moments, Hahn moments, dual Hahn moments and Racah moments.

5.3 2D moments orthogonal on a disk

These moments are constructed as products of a radial and angular factor

$$v_{pq} = n_{pq} \int_0^{2\pi} \int_0^1 R_{pq}(r) e^{-iq\theta} f(r,\theta) r \mathrm{d}r \mathrm{d}\theta \qquad p = 0, 1, 2, \ldots \; , q = -p, \ldots, p \; , \tag{24}$$

where n_{pq} is some normalizing factor, (r,θ) are polar coordinates, $R_{pq}(r)$ is a 1D radial polynomial and $e^{-iq\theta}$ is an angular part of the respective basis function. The area of orthogonality is usually a unit disk. To calculate the moments, image f must be appropriately scaled and shifted such that it is fully contained in Ω. The typical moments orthogonal on a disk are Zernike moments, Fourier-Mellin moments, Chebyshev-Fourier moments and Jacobi-Fourier moments.

5.4 Rotation invariants from OG moments

The moments orthogonal on a disk can be easily normalized to rotation. We took some moment of low order $v_{p_r q_r}$ (typically v_{31} for non-symmetric objects) and compute the normalizing phase from it

$$\theta = \frac{1}{q_r} \arctan\left(\frac{\mathcal{I}m(v_{p_r q_r})}{\mathcal{R}e(v_{p_r q_r})}\right). \tag{25}$$

The normalized moments are then rotation invariants

$$\hat{v}_{pq} = v_{pq} e^{-iq\theta}. \tag{26}$$

In the case of Hermite moments, we have another possibility. The 2D Hermite polynomials $H_{pq}(x,y) \equiv H_p(x)H_q(y)$ change under rotation by angle α in the same way as the monomials $x^p y^q$. It implies that if we substitute the geometric moments in a rotation invariant by the corresponding Hermite moments, we obtain also a rotation invariant.

The Hermite moments partially normalized by the Gaussian function are called Gaussian-Hermite (GH) moments. For the detailed discussion on high-order GH invariants, their numerical properties and the reconstruction power see [19, 20].

5.5 3D orthogonal moments

Similarly as in 2D, also in 3D OG polynomials and moments can be categorized according to their area of orthogonality Ω. We recognize polynomials orthogonal on a cube (see e.g. [21]), on a sphere, and on a cylinder. The polynomials OG on a cube are very easy to construct and calculate, while the polynomials OG on a sphere are generally more appropriate for construction of 3D rotation invariants. The polynomials OG on a cylinder can be used for construction of rotation invariants only when we know the axis of rotation.

6 Algorithms for Moment Computation

There is a number of algorithms that enable to accelerate or to improve accuracy of the moment computation for some types of images. Some approaches

are based on better numerical methods of integration. The methods for fast computation of the moments of binary images can be divided into two groups: boundary-based methods and decomposition methods.

The boundary-based methods employ the property that the boundary of a binary object contains a complete information about the object. Provided that the boundary consists of far fewer pixels than the whole object, an algorithm that calculates the moments from only the boundary pixels should be more efficient than the direct calculation by definition.

The decomposition methods attempt to decompose a binary object into a set of non-overlapping blocks. When the moment of each block can be calculated in $\mathcal{O}(1)$ time, then the overall complexity of $M_{pq}^{(\Omega)}$ is $\mathcal{O}(K)$, where K is the number of blocks. If it is much less than the number of pixels, the speed-up comparing to the definition may be significant.

A comparison of the efficiency of various decomposition methods can be found in [22]. The main decomposition methods are delta method that creates one-pixel thin rectangles, generalized delta method that unifies the adjacent thin rectangles into thicker ones, quadtree decomposition with regular partition into squares until they are uniform, distance transformation decomposition (the biggest square first) and graph-based decomposition that yields the optimal number of blocks, see Fig. 9.

6.1 Computing binary OG moments by means of decomposition methods

To calculate moments orthogonal on a rectangle we can, in principle, take any decomposition scheme used previously and simply replace the monomials $x^p y^q$ by respective OG polynomials. Particularly attractive are Legendre moments, because primitive functions of Legendre polynomials are again Legendre polynomials.

For Zernike moments (and for other moments orthogonal on a disk) the decomposition methods cannot be used in the Cartesian coordinates, but can be employed in polar coordinates, where "rectangular blocks" are actually angular-circular segments. Anyway, we can use the "polar delta method".

6.2 Geometric moments of graylevel images

When calculating moments of a graylevel image, we cannot hope for a significant speedup over the direct calculation without a loss of accuracy. Yet certain fast algorithms do still exist. However, they are meaningful for specific images only and/or do not yield exact results. There are two basic groups – graylevel slicing and graylevel approximation. The graylevel slicing is further divided into intensity slicing and bit slicing.

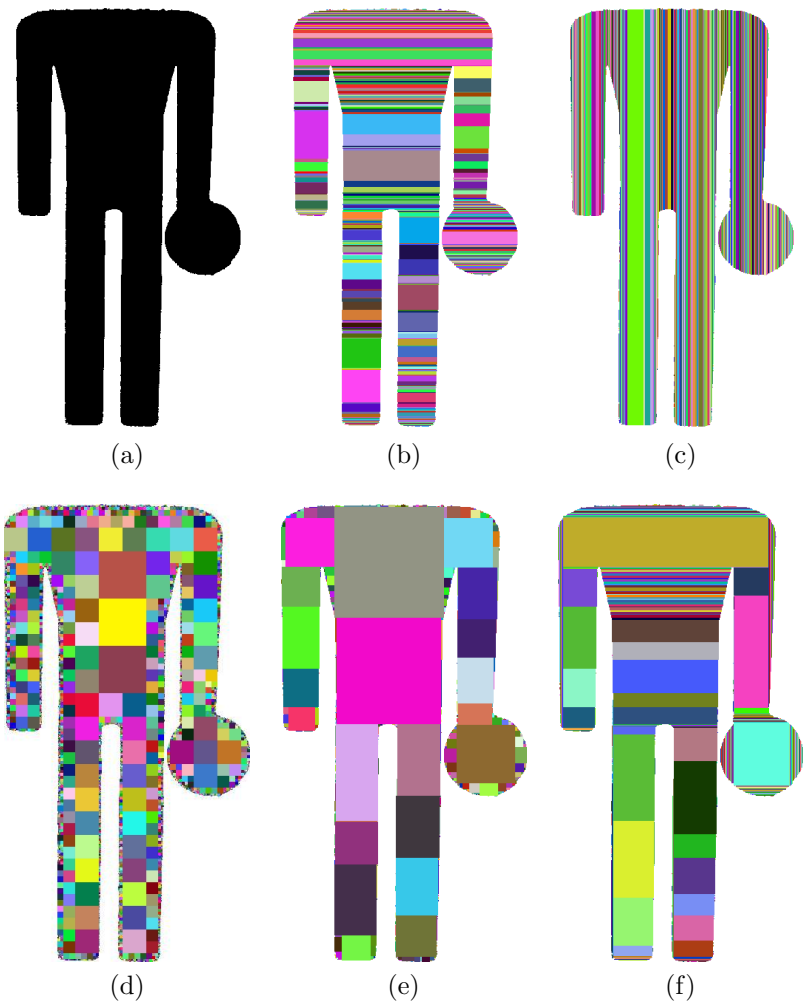


Figure 9: Decomposition methods: (a) the original image (black=1) with 124 212 pixels, (b) the generalized delta method applied row-wise (395 blocks, the basic delta method generated 1507 blocks) (c) the generalized delta method applied column-wise (330 blocks), (d) quadtree (3843 blocks), (e) distance transformation (421 blocks), (f) graph-based method (302 blocks).

7 Applications

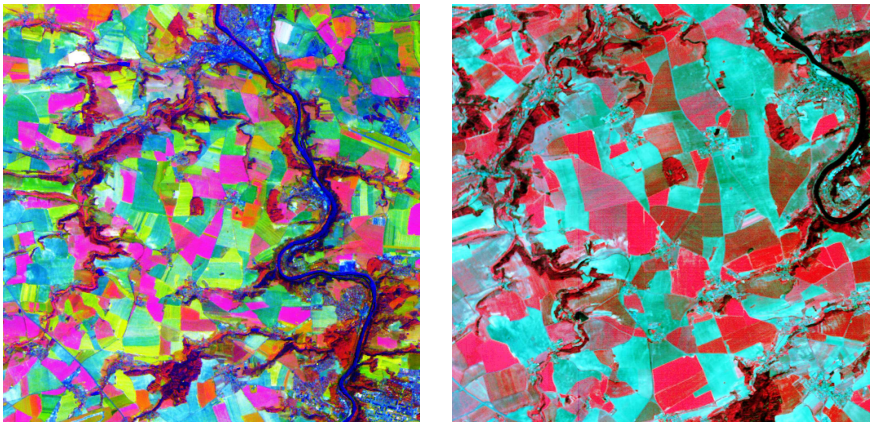
The moments have found numerous applications in practical image analysis and object recognition tasks. They often serve as "gold standard" methods.

7.1 Image understanding

One category of applications can be called image understanding or object recognition. Examples of such applications can be recognition of animals, fish, insects, faces (see [23]) and other human parts. Special subcategories are character recognition [24], logo recognition and recognition of vegetation, e.g. recognition of tree leaves (see [25], Fig. 10) or weed visual recognition in [26].

Figure 10: Image retrieval: Suk and Novotný [25] proposed the system for recognition of woody species in Central Europe, based on Chebyshev moments and Fourier descriptors.

Another such subcategory is traffic-related recognition that includes recognition of aircrafts, ships and vehicles, detection of an obstacle in traffic scene situations and traffic sign recognition. There are also various industrial recognition tasks.



(a) (b)

Figure 11: Satellite images: (a) Landsat image – synthesis of its three principal components, (b) SPOT image – synthesis of its three spectral bands.

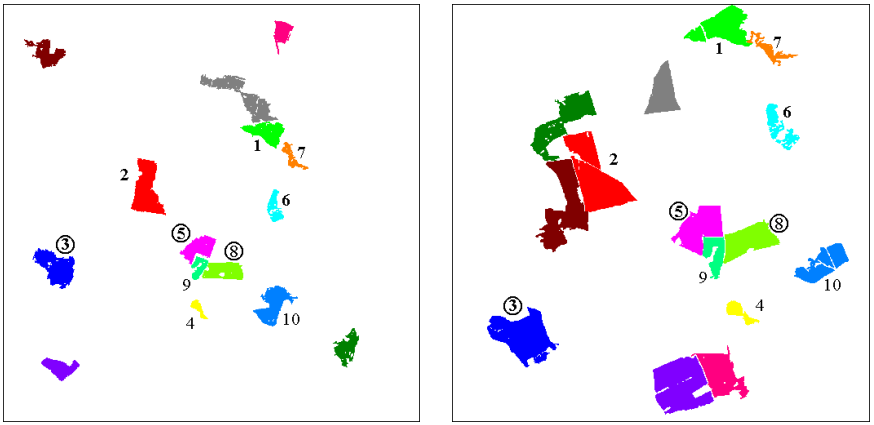
7.2 Image registration

The image registration is a process of overlaying two or more images of the same scene taken at different times, from different viewpoints, or by different sensors. It aims at the geometrical alignment of the reference and the sensed images. The image registration task appears for example in remote sensing, in medicine, and in computer vision. An example of such a task can be registration of satellite images by affine moment invariants of segmented closed boundary regions in [27], see Figs. 11–13.

7.3 Other applications

Other application areas cover robot & autonomous vehicle navigation and visual servoing, focus and image quality measure, content-based image retrieval, watermarking, medical imaging, forensic applications and astronomy. An example from astronomy is description of solar flares (Fig. 14), where 1D moments are used for description of the phenomenon in time (see [28], Fig. 15).

In the content-based image retrieval, the search engine looks for images, which are the most similar (in a pre-defined metric) to the given query image. In digital watermarking a chosen marker or a random signal is covertly inserted into the image to be protected. Forensic applications include e.g. person identification or photo forgery detection. The moments were also used for e.g. gas-liquid flow categorization, 3D object visualization or object tracking.



(a) (b)

Figure 12: segmented regions: (a) Landsat, (b) SPOT image. The regions with a counterpart in the other image are numbered, the matched ones have numbers in a circle.

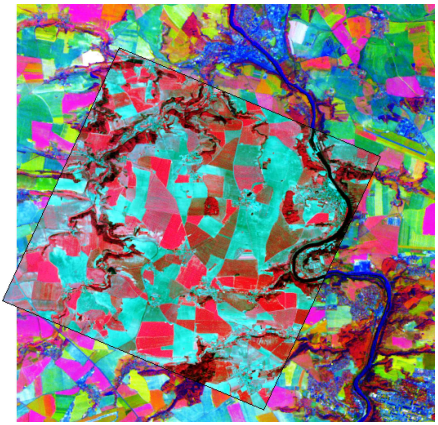


Figure 13: The superimposed Landsat and the registered SPOT images.

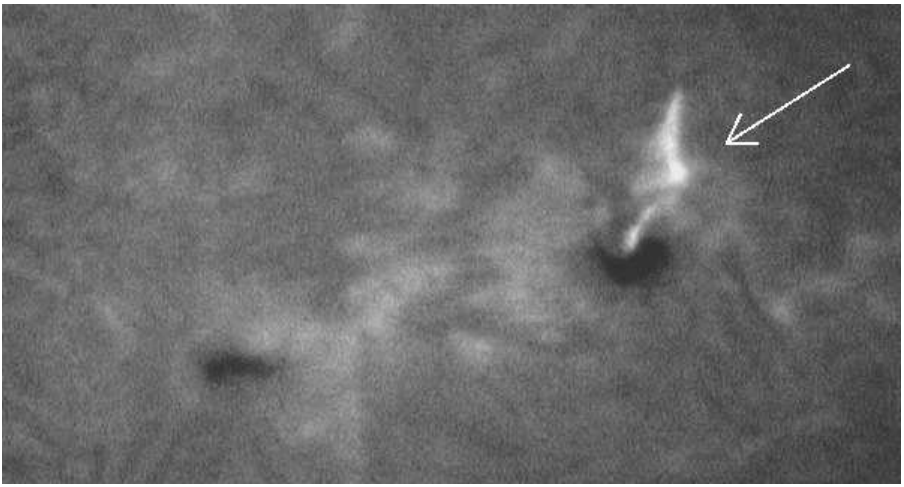
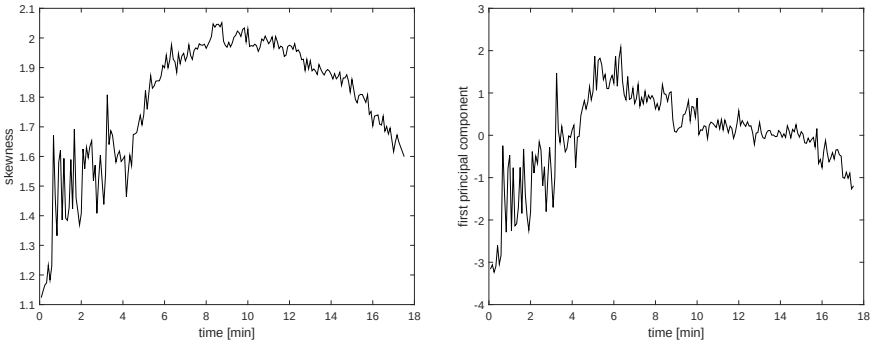


Figure 14: An example of the solar flare (the arrow-marked bright object). Dark areas correspond to sunspots.



(a) (b)

Figure 15: Time curve of a solar flare – (a) skewness, (b) first principal component. Data from Ondřejov observatory, the scanning began on the 18th December 2003, at 12:25 (time 0 in the graphs).

8 Contribution of the Thesis

The theory of moments and moment invariants has been covered step by step in this book. Our aim was to address the main categories of 2D and 3D moments and moment invariants, the computational issues, and their applicability in various domains. In this way, we hoped to provide both material for theoretical study and references for algorithm design. We also provided readers with an introduction to a broader context – the shape feature design and the classification algorithms for object recognition.

The individual classes of moments were introduced, comprising geometric moments, complex moments, and a wide variety of orthogonal moments such as Legendre, Zernike, Chebyshev, and many others. They were presented together with the comments on their advantages and relevant weak points. Illustrative examples accompany the theoretical parts. The core section of the book is dedicated to the invariance of moments with respect to various geometric and intensity degradations. We consider these moment properties the most interesting ones from the image analysis point of view. In all cases, we presented a consistent theory allowing a derivation of invariants of any orders; in some cases we even demonstrated several alternative approaches. Certain properties of the recently proposed features such as affine invariants and especially the blur invariants have never been published prior to this book. To make the book comprehensive, widely known invariants to translation, rotation, and scaling are also described in detail.

The moments are valued for their representative object description. For time-optimal applications of moments, methods and algorithms used for their computation were reviewed to ensure the time efficiency of the solutions based on moments. The mentioned application areas make use of the moment ability to express the shape variations and at the same time to stay invariant under common image deformations.

The specific contribution of Tomáš Suk is mainly in

- affine invariants and their derivation by the graph method
- 3D rotation invariants derived by tensors and by complex moments
- the study of the influence of object symmetries on the vanishing invariants
- decomposition-based algorithms for moment computation
- stable algorithms for moment-based image reconstruction

In all these areas, T. Suk introduced novel methods and designed original algorithms.

8.1 Outlook to the future

Regarding open problems and desirable future research directions, we foresee several areas of rapid development in the near future.

We envisage breaking the limitations of fixed polynomial basis that have been mostly used so far. The future "moments" should not be restricted on the polynomials, and the basis functions should be dynamically adapted to the given task and to particular datasets. The functional bases should be designed in such a way that they optimize a criterion, which is important for the given problem, such as minimum information loss, maximum separability of the training sets, sparse object representation, and/or fast computability, to name a few examples.

Since the importance and the availability of 3D imaging technologies has been growing continually, the future development of moments should be more focused on the 3D domain, especially in the aspects where the 2D experience cannot be transferred into 3D in a straightforward manner.

The theory and applications of *vector field invariants* is a recently opened area with a big application potential, particularly in fluid dynamics. Detection of flow singularities such as vortices and sinks is of a great interest in many branches of engineering. These tasks are intrinsically three-dimensional, but 3D vector field invariants have not been developed at a sufficient level.

Finally, in moment research, just as in most areas of computer science, the development of theory and applications are intertwined issues. We can imagine the appearance of completely new classes of moment invariants w.r.t. the degradations, the models of which have not been described/utilized so far. Such degradations may appear in new, unexplored application areas and/or in connection with future advanced imaging devices and technologies. This may stimulate theoretical research, which would again, if successful, contribute to opening yet other application areas.

References

[1] J. Flusser, T. Suk, and B. Zitová, *2D and 3D Image Analysis by Moments*. Chichester, U.K.: Wiley, 2016.

[2] A. Rosenfeld and A. C. Kak, *Digital Picture Processing*. Academic Press, 2nd ed., 1982.

[3] H. C. Andrews and B. R. Hunt, *Digital Image Restoration*. Prentice Hall, 1977.

[4] W. K. Pratt, *Digital Image Processing*. Wiley Interscience, 4th ed., 2007.

[5] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Prentice Hall, 3rd ed., 2007.

[6] M. Šonka, V. Hlaváč, and R. Boyle, *Image Processing, Analysis and Machine Vision*. Thomson, 3rd ed., 2007.

[7] P. Campisi and K. Egiazarian, *Blind Image Deconvolution: Theory and Applications*. CRC Press, 2007.

[8] P. Milanfar, *Super-Resolution Imaging*. CRC Press, 2011.

[9] A. N. Rajagopalan and R. Chellappa, *Motion Deblurring: Algorithms and Systems*. Cambridge University Press, 2014.

[10] A. Rosenfeld, *Picture Processing by Computer*. Academic Press, 1969.

[11] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. Wiley Interscience, 2nd ed., 2001.

[12] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Academic Press, 4th ed., 2009.

[13] J. Flusser and T. Suk, "Rotation moment invariants for recognition of symmetric objects," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3784–3790, 2006.

[14] T. Suk and J. Flusser, "Tensor method for constructing 3D moment invariants," in *Computer Analysis of Images and Patterns CAIP'11*, vol. 6854–6855 of *Lecture Notes in Computer Science*, pp. 212–219, Springer, 2011.

[15] T. Suk, J. Flusser, and J. Boldyš, "3D rotation invariants by complex moments," *Pattern Recognition*, vol. 48, no. 11, pp. 3516–3526, 2015.

[16] T. Suk and J. Flusser, "Recognition of symmetric 3D bodies," *Symmetry*, vol. 6, no. 3, pp. 722–757, 2014.

[17] T. Suk and J. Flusser, "Affine moment invariants generated by graph method," *Pattern Recognition*, vol. 44, no. 9, pp. 2047–2056, 2011.

[18] J. Flusser and T. Suk, "Pattern recognition by affine moment invariants," *Pattern Recognition*, vol. 26, no. 1, pp. 167–174, 1993.

[19] B. Yang, J. Flusser, and T. Suk, "Design of high-order rotation invariants from Gaussian-Hermite moments," *Signal Processing*, vol. 113, no. 1, pp. 61–67, 2015.

[20] B. Yang, T. Suk, M. Dai, and J. Flusser, "2D and 3D image analysis by Gaussian-Hermite moments," in *Moments and Moment Invariants - Theory and Applications* (G. A. Papakostas, ed.), (Xanthi, Greece), pp. 143–173, Science Gate Publishing, 2014.

[21] B. Yang, J. Flusser, and T. Suk, "3D rotation invariants of Gaussian-Hermite moments," *Pattern Recognition Letters*, vol. 54, no. 1, pp. 18–26, 2015.

[22] T. Suk, C. Höschl IV, and J. Flusser, "Decomposition of binary images – a survey and comparison," *Pattern Recognition*, vol. 45, no. 12, pp. 4279–4291, 2012.

[23] J. Flusser, S. Farokhi, C. Höschl IV, T. Suk, B. Zitová, and M. Pedone, "Recognition of images degraded by Gaussian blur," *IEEE Transactions on Image Processing*, vol. 25, no. 2, pp. 790–806, 2016.

[24] J. Flusser and T. Suk, "Affine moment invariants: A new tool for character recognition," *Pattern Recognition Letters*, vol. 15, no. 4, pp. 433–436, 1994.

[25] P. Novotný and T. Suk, "Leaf recognition of woody species in Central Europe," *Biosystems Engineering*, vol. 115, no. 4, pp. 444–452, 2013.

[26] J. Flusser, T. Suk, and B. Zitová, "On the recognition of wood slices by means of blur invariants," *Sensors and Actuators A: Physical*, vol. 198, pp. 113–118, 2013.

[27] J. Flusser and T. Suk, "A moment-based approach to registration of images with affine geometric distortion," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 32, no. 2, pp. 382–387, 1994.

[28] S. Šimberová, M. Karlický, and T. Suk, "Statistical moments of active-region images during solar flares," *Solar Physics*, vol. 289, no. 1, pp. 193–209, 2014.