

LOCAL ANALYSIS OF IMAGES BY MEANS OF ANISOTROPIC SPATIAL FILTERS

1966

SovietRxiv

View the original and related papers at <https://sovietrxiv.org/items/ru-196601.76297>

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.

Abstract

Full Text

UDC 612.825.8.001.57 : 51 : 681.142

CYBERNETICS AND CONTROL THEORY

V. P. ROMANOV

LOCAL ANALYSIS OF IMAGES BY MEANS OF ANISOTROPIC SPATIAL FILTERS

(Presented by Academician A. A. Dorodnitsyn on 12 XI 1965)

1. To ensure high reliability in the automatic reading and recognition of images of complex form, the usual geometric representations, according to which the main task of recognition is the construction of a multidimensional feature space and the drawing in this space of boundaries separating classes of patterns ⁽¹⁾, are already becoming insufficient. In this connection, a more general conception has recently emerged, in which patterns are regarded as constructions in a certain formal system ^(2,3). Here the principal task becomes the discovery of the set of rules (grammar) of the formal system by which the patterns are constructed, as well as the analysis of an unknown image in order to compose its description in terms of this grammar.

In the article ⁽⁴⁾, methods are described for extracting rectilinear and curvilinear elements of the form of an image by means of two-dimensional spatial filters. In the present work a new method is described for analyzing the form of images, using certain ideas of theoretical mechanics. The analysis of an image begins with the study of its local structure and of the spatial orientation of small regions. The decision to assign a given element one or another index is made with the use of ordinary geometric recognition methods. The field of amplitudes (brightnesses) and the field of directions (phases) are then used in a system for recognizing letters of the alphabet and digits.

2. A two-dimensional image is completely determined by specifying a function $g(x, y)$, usually associated with the degree of blackening at the point (x, y) , $0 \leq g(x, y) \leq 1$. Let us isolate a small region $\Omega(x', y')$ in the neighborhood of the current point (x', y') . To find the directionality of this region we shall use ideas from theoretical mechanics. Let the function $g_{\Omega}(x', y') = g(x, y)\theta_{\Omega}(x', y')$, where $\theta_{\Omega}(x', y')$ is the characteristic function of the domain $\Omega(x', y')$, specify the distribution of masses of a thin plate in the neighborhood of the point (x', y') . The anisotropy of this distribution is characterized by the ellipse of inertia with center at the point (x', y') :

$$I_{xx}x^2 + I_{yy}y^2 - 2I_{xy}xy = 1.$$

It is natural to assign to the isolated region the direction of the minor semiaxis of the ellipse of inertia, i.e., the direction that corresponds to the maximum moment of inertia. Bearing in mind the subsequent use of a discrete algorithm, we shall confine ourselves to finding the maximum moment by considering the following four directions:

$$I_{yy} (\varphi = 0), \quad I_{xx} (\varphi = \pi/2), \quad I_{\eta\eta} (\varphi = \pi/4), \quad I_{\xi\xi} (\varphi = -\pi/4).$$

Let us form the differences

$$B_1 = I_{yy} - I_{xx}, \quad B_2 = I_{\eta\eta} - I_{\xi\xi}.$$

In order to determine the direction of the minor semiaxis of the inertia ellipse, it is sufficient to compare B_1 and B_2 in absolute value and to take into account the sign of the one whose absolute value is larger. The decision rule assigning to each image element one of four directions in the discrete model described here has the form

$$\alpha(\rho, \varphi) = \begin{cases} 0 \pm k\pi & \text{if } |B_1| - |B_2| \geq 0, \quad B_1 \geq 0, \\ \pi/2 \pm k\pi & \text{if } |B_1| - |B_2| \geq 0, \quad B_1 < 0, \\ \pi/4 \pm k\pi & \text{if } |B_1| - |B_2| < 0, \quad B_2 \geq 0, \\ -\pi/4 \pm k\pi & \text{if } |B_1| - |B_2| \leq 0, \quad B_2 < 0. \end{cases} \quad (1)$$

We shall regard the analysis process as a transformation Φ that assigns to the original image a vector field on the plane. Let the direction of the vector at each point (x', y') coincide with the direction of the minor semiaxis of the inertia ellipse of the neighborhood $\Omega(x', y')$ of the point (x', y') , and let the modulus of the vector be equal to the mean value of the function $g_\Omega(x', y')$ in this neighborhood. In the case when the original image is distorted by noise, the error in determining the phase and modulus is smaller if a weight function is specified in the neighborhood $\Omega(x', y')$ and generalized moments and the weighted mean value of the function $g_\Omega(x', y')$ are used.

Experiments with weight functions of various types, for example triangular, exponential, and Gaussian, have shown that the smallest error occurs when a Gaussian weight function is used. Apparently this is connected with the spectral properties of the image and of the weight function. As is known, the product of the width of a Gaussian function by the "width" of its spectrum is the least possible, as a result of which it introduces a smaller error in determining the modulus and phase of the vector field. We therefore define the value of the

vector modulus at each point of the vector field $g(x, y)$ by means of the integral transform:

$$|g(\rho, \varphi)| = \frac{1}{2\pi\sigma^2} \int_0^\infty \int_0^{2\pi} r g(r, \psi) \exp \left\{ -\frac{1}{2\sigma^2} [\rho^2 - 2\rho r \cos(\varphi - \psi) + r^2] \right\} dr d\psi.$$

The direction of the vector at each point of the vector field is established by means of a decision procedure, which we shall now describe in more detail. First of all, let us write the explicit expression for the generalized moments I_{xx} and $I_{\xi\xi}$:

$$I_{xx} = \frac{1}{2\pi\sigma^2} \int_0^\infty \int_0^{2\pi} r^3 \sin^2 \psi \exp \left[-\frac{r^2}{2\sigma^2} \right] g(r, \psi) dr d\psi, \quad (2)$$

$$I_{\xi\xi} = \frac{1}{2\pi\sigma^2} \int_0^\infty \int_0^{2\pi} r^3 \sin^2 \left(\psi - \frac{\pi}{4} \right) \exp \left[-\frac{r^2}{2\sigma^2} \right] g(r, \psi) dr d\psi. \quad (3)$$

Replacing sin by cos in formulas (2) and (3), we obtain expressions for I_{yy} and $I_{\eta\eta}$. The differences B_1 and B_2 will then be written in the form

$$B_1 = \frac{1}{2\pi\sigma^2} \int_0^\infty \int_0^{2\pi} r^3 \cos 2\psi \exp \left[-\frac{r^2}{2\sigma^2} \right] g(r, \psi) dr d\psi,$$

$$B_2 = \frac{1}{2\pi\sigma^2} \int_0^\infty \int_0^{2\pi} r^3 \cos 2 \left(\psi - \frac{\pi}{4} \right) \exp \left[-\frac{r^2}{2\sigma^2} \right] g(r, \psi) dr d\psi.$$

Since B_1 and B_2 must be computed at each point of the image, we are dealing with two spatial anisotropic filters, one of which is rotated by the angle $\pi/4$ relative to the other. In combination with the decision rule (1), they constitute a means for analyzing directionality.

3. The experiments were carried out on images of letters of the Latin and Russian alphabets, numerals, certain geometric figures, and also on more complex images containing several geometric objects. The original images are represented by the values of the function

$$g(x_k, y_l) = \begin{cases} 1 \\ 0 \end{cases} \quad (k, l = 1, 2, \dots, 32).$$

At the initial stage of the analysis it is proposed to use the results of analysis in the system of a reading automaton developed in the laboratory of electrical modeling of the All-Union Institute of Scientific and Technical Information of

Direction field for two types of images: a –image of a rhombus, b –image of the letter O

Figure 1: Direction field for two types of images: a –image of a rhombus, b –image of the letter O

the Academy of Sciences of the USSR ⁽⁵⁾. For this purpose, a comparison of two recognition algorithms was carried out. The first consisted in comparing the amplitudes of the vector field of an unknown image with standards, without taking phase relations into account, and identifying by the minimum distance in the functional space

Fig. 1. Direction field for two types of images: **a** –image of a rhombus, **b** –image of the letter O.

$$d_{xi}^1 = \iint (|\mathbf{g}_x(x, y)| - |\mathbf{g}_i(x, y)|)^2 dx dy, \quad (4)$$

$$i = 1, 2, \dots, N, \quad N \text{ is the number of standards.}$$

The method proposed in the present work consists in comparing the vector fields of the unknown image and the standards according to the formula

$$d_{xi}^2 = \iint |\mathbf{g}_x(x, y)|^2 dx dy + \iint |\mathbf{g}_i(x, y)|^2 dx dy - 2 \iint |\mathbf{g}_x(x, y)| |\mathbf{g}_i(x, y)| [\cos \alpha_x(x, y) - \cos \alpha_i(x, y)] dx dy,$$

$$i = 1, 2, \dots, N.$$

The system being modeled contains a subroutine that operates according to relation (4), and two subroutines of anisotropic filtering, operating in alternation. Both filters examine, along with each image point, 80 of its neighbors, forming a square of size 9×9 points. Different phase values are denoted by the digits 0, 1, 2, 3, depending on whether the point belongs to a direction forming with the abscissa axis the angle $\varphi = 0$, $\varphi = -\pi/4$, $\varphi = \pi/4$, or $\varphi = \pi/2$.

In Fig. 1 are shown phase fields, printed by the Ural-4 machine, for images of a rhombus and the letter O. The results of the program show high accuracy in analyzing the directions of segments regardless of stroke thickness. All points belonging to a rectilinear element receive one and the same phase value. Errors occur mainly at special G terminal and nodal points. This circumstance makes it possible to use the results of the analysis for constructing a graph-scheme of a character, independent of image scale and stroke thickness.

We give the minimum values of d_{ij}^1 and d_{ij}^2 , for combinations of the most similar characters of the Latin alphabet and digits, selected from the triangular matrix d_{ij} .

Letter pairs	O— C	8— S	8— B	H— R	B— H	P— F	M —N	C— G	H— S	G— O	8— H	B— E
d_{ij}^1	.0269	.0355	.0417	.0461	.0475	.0482	.0527	.0557	.0569	.0640	.0646	.0650
d_{ij}^2	.1339	.2131	.2642	.2030	.1918	.2155	.2263	.1710	.2540	.1945	.2764	.2792

The presented data can be compared directly, since the calculation in both cases was performed on the same scale. As can be seen, when criterion (4) is used (and it is precisely this criterion that is used in the majority of existing automatic reading devices), combinations of characters arise whose recognition reliability is low: O—C, P—F, 8—S, etc. The use of anisotropic filtering makes it possible to increase the difference between similar characters by means of a more complete account of their form. It is important to note that the difference between the most similar and the least similar characters decreases, owing to which recognition reliability increases.

The proposed method permits a comparatively simple technical implementation and may find application in such areas as the analysis of photographs of tracks of elementary particles in a bubble chamber for the automatic determination of the angles of their intersection and scattering. In processing aerial-photograph data by such methods, the directions of landmarks, for example roads, and the placement of objects relative to these landmarks can be determined.

All-Union Institute
of Scientific and Technical Information
Academy of Sciences of the USSR

Received
5 XI 1965

REFERENCES

1. G. S. **Sebastian**, *Decision-Making Processes in Pattern Recognition*, Kiev, 1965.
2. R. **Narasimhan**, *Information and Control*, **7**, No. 2, 151 (1964).
3. R. A. **Kirsh**, *IEEE Trans. on Electronic Computers*, EC-13, 4, 363 (1964).
4. V. P. **Romanov**, *Scientific and Technical Information*, No. 7, 24 (1964).
5. M. L. **Avrukh**, *Reading Devices*, Moscow, 1962, p. 116.

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.