SEGMENTATION OF MULTIDIMENSIONAL IMAGES BY CLUSTERING ALGORITHM BASED ON A BILATERAL CRITERION FOR HOMOGENEITY

Yu.V. Gridnev and K.T. Protasov

Institute of Atmospheric Optics,
Siberian Branch of the Russian Academy of Sciences, Tomsk
Received December 29, 1994

An algorithm for automatic image classification based on a bilateral criterion for homogeneity of multidimensional samples with the use of the Kolmogorov–Smirnov distance has been proposed to solve problems of thematic interpretation of aerospace video data obtained from observations of the Earth's underlying surface and cloudiness. Applicability to a set of adjusted images in multispectral channels of visible, IR, and microwave ranges of electromagnetic spectrum is a salient feature of the constructed algorithm.

Aerospace images of the Earth's underlying surface and cloudiness obtained by simultaneous recording of data in multispectral channels of visible, IR, and microwave ranges of electromagnetic radiation are the main source of routine information when solving the problems of exploration of natural resources and climatic–ecological monitoring. Subsequent processing of multidimensional observations contains a stage of thematic interpretation consisting in recognition of texture–homogeneous zones of images by a segmentation algorithm. To construct a segmentation procedure, it is natural to use the algorithms for automatic classification. In so doing, we must, first, to formulate the texture parameters of image fragments; second, to introduce a measure of adequacy of the textures being analyzed.[1,4]

A peculiarity of the algorithm for automatic classification considered further consists in analysis of a set of multidimensional images adjusted in scales. In spite of the ease of recognition of the fragments of image with the texture of the same type by an interpreter, up to now there has been no universal formal definition of the texture. The known special definitions reflect only heuristic approaches to the formation of the texture–parameter space and methods of its automatic classification.[2,3] The texture is manifested as a spatial image characteristic, and the corresponding functional of the probability density (for continuous observation) is exhaustive statistic description of the texture. It is impossible to assign this functional *a priori*. Therefore, the use of more simple models to describe the statistical properties of digitized radio–brightness fields, for example, with the help of brightness distribution histograms, matrices of interelement connectivity, and regression and autocorrelation functions is justified.[2,3,4]

First a problem of contouring of images is considered to eliminate these fragments of images as interfering at the stage of the texture analysis. The images under analysis are assumed to be digitized, adjusted in scales, and normalized on brightness, so that the digital representation has the form of a three–dimensional spatial matrix of the numbers $\{z^{ijk}\}$ of order $M\times N\times K$, where z^{ijk} is the digitized brightness magnitude for a point (pixel) with the coordinates (i, j) of the plane of observed image of format $M\times N$, and k is the serial number of spectral channel in the set of images of visible, IR, and microwave ranges, $k = 1, \ldots, K$. A set of the elements $\{z^{ij}\}$ of any image from the set $1, \ldots, K$ with the coordinates (i, j) belonging to a square of $(2l + 1)\times(2l + 1)$ pixels will be referred to as the image fragment with the central element $(i = 0,\ j = 0)$ and the local coordinate system $i = -l, \ldots, -1, 0, +1, \ldots, +l$; $j = -l, \ldots, -1, 0, +1, \ldots, +l$ introduced on it, where l is the window size parameter. Correspondingly, the fragment on the square of window $(2l + 1)\times(2l + 1)$ of multicomponent image is a set of the elements $\{z^{ijk}\}$, $(i, j, k) \in (2l + 1)\times(2l + 1)\times K$.

To describe the brightness $\{z^{ij}\}$ within the fragment $(2l + 1)\times(2l + 1)$, a simple model introduced in Ref. 5 is used. In this case the image local characteristics are described in terms of the density sections or facets.

Formally, the vicinity of the point (x, y) is described by the following equation of plane for the continuous coordinates x, y, z of the tree–dimensional space

$$z = \alpha x + \beta y + \mu, \tag{1}$$

where z is the net brightness for the facet model; α, β, and μ are the parameters of the plane. With i and j denoting the discrete values of the coordinates x and y in Eq. (1), respectively, we adjust model (1) to actual observations z^{ij} of the image fragment $(2l + 1)\times(2l + 1)$ best of all in the sense of the least–squares technique, namely

$$I(\alpha, \beta, \mu) = \sum_{j=-l}^{+l} \sum_{i=-l}^{+l} [\alpha i + \beta j + \mu - z^{ji}]^2 = \min_{\{\alpha,\beta,\mu\}} .$$

The estimated unknown parameters minimizing $I(\alpha, \beta, \mu)$ have the following form:

$$\hat{\alpha} = 3 \sum_{i=-l}^{+l} i \sum_{j=-l}^{+l} z^{ij} \,/\, l(l + 1)\,(2l + 1)^2,$$

$$\hat{\beta} = 3 \sum_{j=-l}^{+l} j \sum_{i=-l}^{+l} z^{ij} \,/\, l(l + 1)\,(2l + 1)^2, \tag{2}$$

$$\hat{\mu} = \sum_{j=-l}^{+l} \sum_{i=-l}^{+l} z^{ij} \,/\, (2l + 1)^2.$$

Description of an image fragment with the help of facet model (1) is used to solve a problem of distinguishing the segments of image gradients. A value of the image gradient within the fragment is estimated by the spatial derivative determined as a ratio of the area dS of inclined plane (1) bounded by the square fragment $(2l + 1)\times(2l + 1)$ to the area of the base $d\Delta$ of this fragment. If Eq. (1) is represented in the form of the equation of plane in terms of the direction cosines, then

$$\frac{dS}{d\Delta} \cong \left| \left(\frac{3 \sum_{i=-l}^{+l} i \sum_{j=-l}^{+l} z^{ji}}{l(l+1)(2l+1)^2} \right)^2 + \left(\frac{3 \sum_{j=-l}^{+l} j \sum_{i=-l}^{+l} z^{ji}}{l(l+1)(2l+1)^2} \right)^2 + 1 \right|^{1/2} . \quad (3)$$

Estimated value of the gradient is attributed to the central point of fragment with the local pixel coordinates $i = 0$, $j = 0$. If now every image under analysis is differentiated in the same way and the values of the corresponding gradients are attributed to the central elements of scanning window with the parameter l, then we can transfer from initial images of the radio–brightness relief $\{z^{ijk}\}$ of aerospace photographs to the images with the underlined gradient having the values $\{w^{ijk}\}$ determined from formula (3).

Now a histogram of gradient distribution for a gradient image is constructed and two regions are selected on this histogram: the left no gradient region starting from zero and the right region of gradients.

Formalized statement of this problem leads to a task of decomposition of the joint gradient distribution $p(w)$, specified by the above–mentioned histogram, into two component distributions[1,4]:

$$p(w) = P\, f(w) + Q\, g(w), \quad (4)$$

where $g(w)$ is the function of conditional probability density for the hypothesis H_1 that the observation w falls in the gradient class; $f(w)$ is the density function for the alternative H_0 that the observation w falls in the not gradient class; P and Q are the parameters of mixture; $P + Q = 1$.

As parametric models of distributions comprising mixture (4), the Jonson S_B–distribution[6] is used having great approximating abilities in combination with few parameters among which are the form factors

$$g(w),\ f(w) = \frac{\lambda}{\sqrt{2\pi}\,(w - \varepsilon)\,(\lambda - w + \varepsilon)} \times$$

$$\times \exp\left\{ -\frac{1}{2} \left[\gamma + \eta \ln\left(\frac{w - \varepsilon}{\lambda - w + \varepsilon} \right) \right]^2 \right\}, \quad (5)$$

where $\varepsilon \le w \le \varepsilon + \lambda$, $\eta > 0$, $\lambda > 0$, $-\infty < \gamma < \infty$.

We solve the decomposition problem for distributions (4) with regard to models (5) numerically, choosing as a criterion for optimum a minimum of the following square criterion:

$$[\, \hat{p}(w) - Pf(w) - Qg(w)\,]^2 = \min_{\{\theta\}}, \quad (6)$$

where $\hat{p}(w)$ is the image gradient histogram; $\theta = \{\varepsilon_0, \lambda_0, \gamma_0, \eta_0; \varepsilon_1, \lambda_1, \gamma_1, \eta_1; P\}^{\mathrm{T}}$ is the set of unknown parameters; superscript T denotes transposition; ε_0, λ_0, γ_0, and η_0 are the parameters of the distribution $f(w)$ given by Eq. (5) for the alternative H_0; ε_1, λ_1, γ_1, and η_1 are the parameters of the distribution $g(w)$ given by Eq. (5) for the hypothesis H_1; P is the weight of the distribution $f(w)$ in Eq. (4); and, $Q = 1 - P$. The problem of minimization of multiextremum functional (6) with respect to the set of eight parameters can be solved numerically by standard methods of random search for the gradient descent points with subsequent gradient descent.

After estimating the parameters of the distributions $f(w)$ and $g(w)$, the decision rule in testing the hypotheses that contour (gradient) exists or not is written in the form of the Bayes decision rule

$$u(w^{ij}) = \arg\max \{\hat{p}\, \hat{f}(w^{ij}),\ \hat{Q}\, \hat{g}(w^{ij})\}, \quad (7)$$

where $u(w^{ij})$ is the solution; w^{ij} is the magnitude of the gradient image under analysis at the point (i, j). When we deal with a set of gradient images, the generalized gradient is determined by the vector $\frac{d\mathbf{S}}{d\Delta} = \left(\frac{dS_1}{d\Delta_1}, \ldots, \frac{dS_K}{d\Delta_K} \right)^{\mathrm{T}}$.

The components of this vector are the magnitudes of individual gradient images.

To solve the problem of retrieving the multidimensional probabilistic distributions of the components of the vector $\frac{d\mathbf{S}}{d\Delta}$ for model (4), it is natural to use the multidimensional analog of the Jonson S_B–distribution[6] that will substantially complicate a problem of evaluation of the unknown parameters. Simplification can be made at the cost of deterioration of the solution quality by introducing a concept of the gradient vector norm

$$\left\| \frac{d\mathbf{S}}{d\Delta} \right\| = \sqrt{ \left(\frac{dS_1}{d\Delta_1} \right)^2 + \ldots + \left(\frac{dS_K}{d\Delta_K} \right)^2 } .$$

This makes it possible to use completely the above–described variant with individual gradient image.

At last, let us transfer to consideration of the main problem of the texture analysis of multidimensional video data with excluded contours of large gradient.

As the texture manifests itself as a spatial image characteristic, initial images should be represented in the form of fragments, and the statistical characteristics of every fragment should be studied.

Let us use one of the criteria for testing hypotheses of homogeneity of two samples[7] using as a test the Kolmogorov–Smirnov distance between multidimensional empirical distributions retrieved on pairs of fragments being compared.

Let $\{z^{ijk}\}$ and $\{v^{ijk}\}$ be the sets of radio brightness recorded on two image fragments being analyzed for homogeneity with the continuous integral distribution functions $F(z)$ and $G(z)$, respectively. In this case, a measure of identity of the image fragment textures is a value of the Kolmogorov–Smirnov distance between the multidimensional distributions retrieved on the corresponding fragments[7]

$$D_{mn} = \sup_{\{z\}} | F_m(z) - G_n(z) |, \quad (8)$$

where $F_m(\cdot)$ and $G_n(\cdot)$ are the multidimensional empirical distribution functions of the following form:

$$F_m(z) = \frac{1}{m} \sum_{\{i,j\}} \prod_{k=1}^{K} C\,(z^k - z^{ijk}),$$

$$G_n(z) = \frac{1}{n} \sum_{\{i,j\}} \prod_{k=1}^{K} C\,(z^k - v^{ijk}), \quad (9)$$

m is the number of pixels in the fragment $\{z^{ijk}\}$; n is the number of pixels in the fragment $\{v^{ijk}\}$; moreover, these fragments have no elements belonging to contour if it is; $C(\cdot)$ is the comparison function, $C(t) = \{1,\ t \geq 0;\ 0,\ t < 0\}$.

It is necessary to test the hypothesis H_0 that the integral functions of brightness distribution being retrieved on these fragments are the same in the statistical sense, and two samples from which they have been retrieved belong to the same statistical ensemble. This is indicative of statistical homogeneity of the pair of fragments being compared. The alternative H_1 states inequality of integral density functions and, hence, absence of the statistical equivalence of the textures of these fragments.

In other words, the hypothesis H_0: $F(z) = G(z)$ should be tested against the alternative H_1: $F(z) \neq G(z)$. But it should be borne in mind that for statistics (8) of bilateral test for homogeneity, the asymptotic distributions D_{mn} have been obtained only for the one–dimensional functions $F(\cdot)$ and $G(\cdot)$.

The way out in this situation can be found using an idea of decomposition of the mixed statistics D_{mn} into the component distributions subsequently used to construct the Bayes decision rule for testing the hypotheses H_0 and H_1 as in the case of model (4).

In this case, a process of clustering the texture–homogeneous fragments is constructed in the following manner. Initial multidimensional image is decomposed by the coordinate grid with a given step into individual multidimensional fragments with the magnitudes of radio brightness $\{z^{ijk}\}$, $i = 1, \ldots, l_x$; $j = 1, \ldots, l_y$; $k = 1, \ldots, K$. The size of fragments $l_x \times l_y$ is chosen for reasons that the texture features can manifest themselves completely within the fragment. In this case, the contradictory requirements must be taken into account, namely, choosing a fragment of larger dimensions, its texture characteristics can be described more completely thereby increasing the stability of distribution estimations, but the accuracy of presentation of homogeneity zone boundaries deteriorates in details; moreover, within larger fragment the texture internal homogeneity most likely will be disrupted.

Once the set of image fragments is fixed, we choose one of them as a reference and subsequently each remaining – as the second one. Then degree of homogeneity of the fragment pairs is evaluated in the sense of distance (8). This process yields the statistics $\{D_{mnj}\}$, $j = 1, \ldots, M$, where M is the number of fragments being compared with the reference one. According to this statistics, we construct the mixed distribution histogram describing the behavior of D_{mn} for the hypothesis H_0 and alternative H_1.

Using again model (4) and assigning the parametric approximations of unknown conditional density functions for the hypothesis H_0 and alternative H_1, these functions can be retrieved together with *a priori* cluster probabilities by solving a problem of identification of distribution mixture (6). After that the decision rule for the hypothesis H_0 and alternative H_1 has the standard Bayes form given by Eq. (7). Among the possible variants of the choice of the reference fragment, the variant is found by exhaustion for which the histogram of D_{mn} statistical distribution is concentrated near zero to the greatest extent that testifies the compactness of the cluster.

This variant with its reference fragment is chosen as a working one for clustering of fragments. The decision rule divides all set of fragments into the set comprising a class of texture–homogeneous fragments and the set of fragments out of this class; the latter ones are again subjected to the above–described procedure of choice of the reference fragment with further clustering.

After a certain number of iterations some fragments remain that can be characterized as heterogeneous ones or observation overshoots.

Let us considered an illustrative example of segmentation of a concrete satellite photograph, which was kindly provided at our disposal by V.I. Khamarin (Institute of Ecology of Natural Complexes of the Siberian Branch of the Russian Academy of Sciences). The photograph is a set of three digitized magnitudes of radio brightness recorded in the spectral ranges 0.5–0.6 μm (the first channel), 0.6–0.7 μm (the second channel), and 0.8–0.9 μm (the third channel).

FIG. 1. Initial image in the third channel.

The fragment under analysis displays the radio brightness of the forest zone of the Siberian taiga with cleared strips and felling areas and has a dimension of 640×488 pixels. Figure 1 displays initial image in the third channel. To distinguish contours in the video data, the window of facet model (1) was set with the parameter $l = 1$, and the image of the gradients determined by the norm of the spatial gradient was obtained. The histogram of gradient distribution shown in Fig. 2*a* provided initial information for solving a problem of decomposition of the mixed distribution into two component density functions (4), assigned in the form of the parametric Jonson S_B–distribution (5). The retrieved density functions with the weight coefficients are depicted in Fig. 2*a*. They determine Bayes decision rule (7) for recognition of contours in the gradient image. The distinguished contours are displayed by light tone in Fig. 3. Then the video data (a rectangle of 640×488 pixels) were decomposed, with the help of a coordinate grid, into individual fragments of 16×16 pixels and the obtained set of fragments was automatically classified, excepting those containing elements of contours.

Using three–dimensional Kolmogorov–Smirnov distance (8) on pairs of fragments being analyzed one of which was chosen as a reference fragment and each remaining was successively chosen as the second, the histograms of the distribution of the statistics $\{D_{mn}\}$ were obtained. One of the obtained mixed distribution of the statistic D_{mn} retrieved with the help of the nonparametric estimation of the unknown density function with the Epanechnikov kernel is displayed in Fig. 2*b*. The decision rule forms the next class of statistically homogeneous fragments from a part of the whole set of fragments connected with the left mode of the retrieved density function. Five classes were distinguished at the stage of clustering. They are shown by variation of gray tone in Fig. 3.

It should be noted in conclusion that in the above–described scheme of construction of the automatic classification procedure the other bilateral criteria can be used to test two samples for homogeneity, for example, ω^2–criterion.[7]

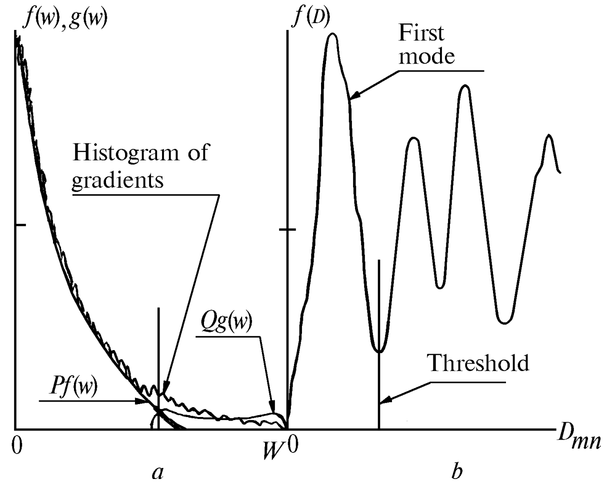

FIG. 2. Histogram of the image gradient distribution, its approximation (a), and nonparametric retrieving of the Kolmogorov–Smirnov distance (b).

FIG. 3. Image with distinguished contours and classes of homogeneous fragments.

REFERENCES

1. R. Duda and P. Hart, *Pattern Recognition and Analysis of Scenes* [Russian translation] (Mir, Moscow, 1976), 512 pp.
2. Yu.N. Khomyakov and V.A. Sushkin, Zarub. Radioelectron., No. 2, 33–46 (1986).
3. D.A. Denisov and V.A. Nizovkin, Zarub. Radioelectron., No. 10, 5–30 (1985).
4. A. Rosenfeld, *Recognition and Processing of Images on Computers* [Russian translation] (Mir, Moscow, 1972), 230 pp.
5. P.A. Chochia, in: *Coding and Processing of Images* (Nauka, Moscow, 1988), pp. 69–87.
6. K.T. Protasov, Atmos. Oceanic Opt. **7**, No. 6, 448–451 (1994).
7. F.P. Tarasenko, *Nonparametric Statistics* (Tomsk State University Publishing House, Tomsk, 1976), 292 pp.