Application of morphological analysis method to identification of gray-scale images

V.T. Kalaida

Tomsk State University

Received July 17, 2003

The paper proposes a modified method of morphological analysis for identification of gray-scale images and a hardware/software pattern recognition system based on the proposed algorithm.

Image processing and identification are traditionally based on the principles and theory of linear systems and on the Fourier transform (or other similar transformations).[1,2] In many applications such an approach turned out rather efficient, but for images it does not permit numerical description of their shape and geometric structure. An alternative approach in analysis of images is mathematical morphology. It allows quantitative description of peculiarities of the geometric structure.

Morphological analysis of image shape based on the set theory, integral geometry, analysis of convex functions, stereology, and geometric probability theory was developed by J. Serra and Yu.P. Pyt'ev in the 1960s.[3,4] Active application of morphological analysis methods is caused by continuous development and enhancement of computer architectures used for morphological transformations of a signal. However, in spite of quite efficient mathematical structure of morphological analysis, these methods do not gain the common recognition yet.

The main idea of morphological analysis can be represented as follows.

Let R and Z be sets of real and integer numbers, and E is a d-dimensional continuous space R^d (d = 1, 2, 3, ...) or a discrete space Z^d. Then the d-dimensional signal can be presented as a function of the area R^d (continuous) or Z^d (discrete), whose range is either R at continuous amplitude variation or Z at quantized amplitude variation.

Binary signals can be presented through sets. Binary images often result from threshold selection of gray-scale images, and threshold selection is often used to represent gray-scale images through binary signals, that is, through sets. Serra[3] uses representation of a real d-dimensional function $f(x)$, where x means a d-dimensional vector, through an ensemble of its d-dimensional threshold sets defined as

$$T_a(f) = \{x : f(x) \geq a\}, \quad -\infty < a < \infty,$$

where the amplitude a fully covers R or Z depending on whether the signal f has a continuous or quantized range. Threshold sets have two important properties. They are linearly ordered, because $a < b \Rightarrow T_a(f) \supseteq T_b(f)$, and allow unambiguous reconstruction of the signal because

$$f(x) = \max\{a : x \in T_a(f)\}, \ \forall x .$$

Consequently, $f(x)$ can be reconstructed from the ensemble $f_a(x)$, since

$$f(x) = \max\{a : f_a = 1\}, \ \forall x .$$

Signal transformations in mathematical morphology are nonlinear operators that locally modify geometrical characteristics of multidimensional signals. Let $X \subseteq E$ be a set representation of the binary input signal and $B \subseteq E$ be a small-size compact set of a simple shape (for example, d-dimensional sphere). The set B is called a structuring element. Let $X \pm b = \{x \pm b : x \in X\}$ expresses the vector transfer of X onto $\pm b \in E$. The main morphological operators for sets are dilation $\oplus$ and erosion of X with the help of B; they are defined as follows:

$$X \oplus B = \bigcup_{b \in B} X + b = \{x + b : \ x \in X \ \text{ and } \ b \in B\},$$

$$X \quad B = \bigcap_{b \in B} X - b = \{z : \ (B + z) \subseteq X\}.$$

Based on these definitions, we can show that the yield of the dilation operator is a set of transferred points, so that the transfer of the reflected structuring element $\breve{B} = \{-b : b \in B\}$ forms a non-empty intersection with the input set, that is, $X \oplus B = \{z : (\breve{B} + z) \bigcap X \neq \varnothing\}$. Similarly, the yield of the erosion operator is a set of transferred points, so that the transferred structuring element is contained in the input set.

Other operators are defined as combinations of erosion and dilation. For example, two extra operators: opening $\circ$ and closing $\bullet X$ with the help of B are defined as

$$X \circ B = (X \quad B) \oplus B, \ X \bullet B = (X \oplus B) \quad B.$$

The described set of the operators can be extended to many-level (that is, non-binary) signals represented by really-significant functions. Serra used

0235-6880/03/09 791-04 $02.00

representation of the d-dimensional function $f(x)$ by the set of its threshold sets. In this case, the operation of dilation of all the threshold sets of the function f with the same compact set B gives the sets $T_a(f) \oplus B$, which are threshold sets of the new function $f \oplus B$ referred to as the dilation of the function f with B. This new function can be calculated either as $(f \oplus B)(x) = \max\{a : x \in T_a(f) \oplus B\}$ or from the direct equivalent formula

$$(f \oplus B)(x) = \max_{y \in B}\{f(x - y)\},$$

$$(f \ominus B)(x) = \min_{y \in B}\{f(x + y)\}.$$

Opening ($\circ$) and closing ($\bullet$) of the function f with B are determined as $f \circ B = (f \ominus B) \oplus B$ and $f \bullet B = (f \oplus B) \ominus B$.

Erosion of the function f with a small convex set B decreases the number of peaks and increases minima of the function. Dilation of the function f with B increases valleys and elongates maximums of the function. Opening with B smoothes the plot of the function f from below by cutting peaks, while closing smoothes it from above by filling valleys of the function f.

Another extension of the list of morphological operators for the functions was made by Sternberg,[8] who used the representation of the d-dimensional function $f(x)$ by a $(d + 1)$-dimensional set. The corresponding umbra operator is expressed as

$$U(f) = \{(x, a) : \ a \le f(x)\},$$

that is, the umbra is a set of points lying below the surface represented by $f(x)$. In general, the set expressing the umbra is extended up to $a = -\infty$. The function can be reconstructed from its umbra, because

$$f(x) = \max\{a : (x, a) \in U(f)\}, \ \forall x.$$

The operations of dilation or erosion of the umbra of f with the help of the umbra of g yield the umbrae of new functions: dilation or erosion of f with g. These new formulas can be derived from the direct formulas

$$(f \oplus g)(x) = \max_{y}\{f(y) + g(x - y)\},$$

$$(f \ominus g)(x) = \min_{y}\{f(y) - g(y - x)\},$$

where the ranges of x and y are defined as the intersection of the support of the function f and the support of the function g (shifted by x). The support of the function f is a set of x values, at which $f(x) \ne -\infty$. The function g is assumed to have a compact support and play the role of the structuring element. Opening and closing of f with g are expressed as

$$f \circ g = (f \ominus g) \oplus g \quad \text{and} \quad f \bullet g = (f \oplus g) \ominus g.$$

Morphological filters are widely used in image processing and analysis. If W is a small symmetric two-dimensional binary structuring element, then the difference of the sets $X \setminus (X \ominus W)$ gives the boundary of the binary image X, while the algebraic difference is

$$EG(f) = f - (f \ominus W). \qquad (1)$$

A similar operator improving the contour is the dilation gradient

$$DG(f) = (f \oplus W) - f. \qquad (2)$$

Combining operators (1) and (2), we can obtain new contour operators, which provide for more symmetric processing of an image and its background.

The stability of detecting contours by these morphological contour operators can be improved through prior smoothing of the input signal of the image f using linear smearing or self-tuning filters.

One of the fundamental problems in analysis of images is adequate mathematical description of images reproducing their content and meaning. In other words, this description should represent only significant (from the viewpoint of the problem to be solved) peculiarities of an image and be independent of insignificant details. In the morphological analysis, such insignificant characteristics are the conditions of recording the image of an object or scene (illumination, variation of optical properties of the object, variation of observation angles) and the parameters of the recording instrumentation.

An important class of problems preceding comparison of image shapes includes the problems of image recording and separation of fragments. The image shape is understood as a maximum invariant of image transformations at variation of the observation conditions, parameters of the recording instrumentation, and so on.[5]

Thus, the shape is not only determined by the object of scene under study, but it is closely connected with the model of recording of an image or scene. One of the ways to construct the shape is to specify the areas of constant brightness from the physical properties of the object, that is, from the arrangement of homogeneously luminous or reflecting sides or boundaries with respect to an observer. Assigning various brightness values to these areas, we obtain the image shape as a set of shapes. If such detailed information about the object is unavailable, we can construct the shape from one image, having known which transformations of the brightness of this image result from variation of the observation conditions. Any image, whose shape is not more complicated than f, can be obtained by fitting the brightness values C_i:

$$f(x, y) = \sum_{i=0}^{n} C_i \chi_i(x, y).$$

Here C_i is the brightness of the ith point; $\chi_i(x, y)$ is an indicator function, which is defined as follows:

$$\chi_i(x,y) = \begin{cases} 0 \notin A_i \\ 1 \in A_i \end{cases},$$

where A_i is the set of points of a given intensity.

Thus, construction of the image shape consists in determination of the indicator functions $\chi_i(x, y)$. In the simplest case, the range of variation of the image brightness f can be divided into n identical intervals.

Based on these suggestions, we can propose the following algorithm for identification of images.

Initially, the image is transformed into the gray scale with $n$bit gradation of brightness in the range $(\overline{0,n})$. Through successive application of morphological operations, an image fragment (subject image) is separated out and reduced to the "universal" size. For the selected class of images, the indicator functions are calculated from brightness values of one (reference) image through pixel-by-pixel analysis of the image. This operation yields the indicator functions of the image shape for the corresponding brightness values.

Then the indicator functions are used to compensate for the difference in brightness characteristics without distortion of the shape of the analyzed image ("candidate"). For this purpose, the correcting coefficient is calculated for each indicator function of the analyzed image, since comparison can be carried out only when the image recording conditions are practically identical.

This result can be obtained by applying the least-squares method (LSM). Toward this end, the brightness of the analyzed image is reduced to the intensity of the reference image so that the sum of the square deviations for each indicator function is minimum:

$$\sum_{i=0}^{m}(C_i - K_j C_i^*)^2 \to \min,\ j = (\overline{0,n}),$$

where C_i is the brightness of the reference image; C_i^* is the brightness of the ith point of the candidate image; K_j is the correcting coefficient; m is the number of points for each indicator function.

As a result, we have the reduced image without distortion of the shape, which can be compared with the basic image by finding the difference between pixels at every point i, k:

$$\left|C_{ik} - C_{ik}^{**}\right| = A_{ik},$$

where A_{ik} is the difference between images; C_{ik} is the brightness of the reference image; C_{ik}^{**} is the brightness of the reduced image.

In the case of identical images the difference between them is zero, but when comparing real images it may be nonzero. For assessment of the residual image, the moment analysis is used.[6] This analysis is based on the fundamental theorem: an infinite set of image moments $\{m_{\alpha\beta}\}$ or $\{\mu_{\alpha\beta}\}$ is unambiguously determined by the function $f(x, y)$, and, *vice versa*, the function $f(x, y)$ is unambiguously determined by the set of image moments $\{m_{\alpha\beta}\}$ or $\{\mu_{\alpha\beta}\}$ (α, β) = 0, 1, …. (Ref. 7):

$$m_{\alpha\beta} = \iint_{\infty} f(x,y)x^{\alpha}y^{\beta}\mathrm{d}x\mathrm{d}y$$

is the intrinsic moment;

$$\mu_{\alpha\beta} = \iint_{\infty} f(x,y)(x - x_0)^{\alpha}(y - y_0)^{\beta}\mathrm{d}x\mathrm{d}y$$

is the central moment.

As a result of calculations, we obtain the set of moment characteristics. For the set of characteristics of all reference images, the reference hyperplane is constructed:

$$\sum_{i=1}^{n} a_i m_i + a_{n+1} = 0,$$

where m_i are moment characteristics; a_i are hyperplane coefficients;

$$m_k = \sum_{i=1(i\neq k)}^{n} b_i m_i + b_{n+1},$$

$$b_i = \frac{a_i}{a_k},\ \ i = 1,\ldots,n+1\,.$$

For calculation of the coefficients of the reference hyperplane for this class (subject) of images, require for the moment characteristics calculated through the hyperplane coefficients to achieve minimum deviation from the calculated moment characteristics for all images of this class:

$$L(\mathbf{b}) = \sum_{j=1}^{m}\left[m_k{}^j - \left(\sum_{i=1(i\neq k)}^{n} b_i m_i^j + b_{n+1}\right)\right]^2 \to \min,$$

where $\mathbf{b} = (b_1,\ldots,b_{n+1})$, $i \neq k$.

To assess these hyperplane coefficients, we construct the system of linear equations

$$\frac{\partial L(\mathbf{b})}{\partial b_m} = 0,\ \ m = 1,\ldots,n(m \neq k),$$

whose solution gives b_i for this class of images.

These calculations yield the database of the vectors of the reference hyperplane for different classes of images.

For the image coming from a video recorder (candidate), images from the database are successively selected. For two images the reduced difference and moment characteristics are calculated. The coefficients of the reference hyperplane from the database are used to assess the difference between the calculated characteristic and that obtained from the hyperplane coefficients:

$$m_{k_{\rm re}}^l = \sum_{i=1}^{n} b_i m_i^l + b_{n+1}(i \neq k), \quad l = 1,\ldots,N,$$

where $m_{k_{\rm re}}^l$ is the calculated moment characteristic of the lth reference image; N is the number of reference images in the database; l is the current reference image.

The reference image, whose difference with the current image is minimum, is fixed

$$\left| m_{k_{\rm re}} - m_{k_{\rm hyp}} \right| = \Delta,$$

where $m_{k_{\rm hyp}}$ is the moment characteristic calculated by the hyperplane coefficients.

The obtained value of Δ is compared with the acceptable deviation for the given class of images. If it is smaller than the acceptable deviation, then the candidate belongs to this class. Otherwise, the candidate does not fall in any class (*FOREIGN*). As an acceptable value, we can take the estimate 3σ, where σ is the root-mean-square (standard) deviation for this class of images.

The method proposed allows comparison of images and their classification by certain indices.

The main advantage of this method is image correction without loss of the information about the shape. This method can be employed in devices applicable in various technological fields, in medicine, science, etc.

References

1. E.O. Brigham, *The Fast Fourier Transform* (Englewood Cliffs, New York, 1974), 62 pp.
2. R.C. Gonzalez and P Wintz, *Digital Image Processing* (Addison-Wesley, Mass., 1987), 14 pp.
3. J. Serra, *Image Analysis and Mathematical Morphology* (Academic Press, London, 1982).
4. Yu.P. Pyt'ev, Dokl. Akad. Nauk SSSR **269**, No. 5, 1061–1064 (1983).
5. Yu.P. Pyt'ev, in: *Mathematical Methods for Investigation of Earth's Natural Resources from Space*, ed. by V.G. Zolotukhin (Nauka, Moscow, 1984), pp. 41–83.
6. V.T. Kalaida, V.V. Belov, V.A. Esipova, and V.M. Klimkin, Solanus, No. 13, 80 – 92 (1999).
7. B.V. Anisimov, V.D. Kurganov, and V.K. Zlobin, *Pattern Recognition and Digital Processing* (Vysshaya Shkola, Moscow, 1983), 295 pp.
8. P. Maragos and R.W. Shafer, Proc. IEEE **78**, No. 4, 690–710 (1990).