ISSN 8756-6990, Optoelectronics, Instrumentation and Data Processing, 2009, Vol. 45, No. 3, pp. 255–261.

Method of Directed Enumeration of Alternatives in the Problem of Automatic Recognition of Half-Tone Images

A. V. Savchenko

State University Higher School of Economy, ul. Bol'shaya Pecherskaya 25/12, Nizhnii Novgorod, 603155 Russia
E-mail: savchenko@tecomgroup.ru

Received January 21, 2009

Abstract—The problem of automatic recognition of half-tone images based on the minimum information discrimination principle is formulated and solved. A method of directed enumeration of the set of alternatives in the Kullback–Leibler information metric is proposed as opposed to the method of complete enumeration of competing hypotheses. A program and results of an experimental study of the method are presented. It is shown that the proposed algorithm is characterized by increased accuracy and reliability of automatic image recognition.

DOI: 10.3103/S8756699009030066

Key words: automatic image recognition, image recognition, learning recognition, minimum information discrimination criterion.

INTRODUCTION

The minimum information discrimination (MID) principle is known [1] to be an effective tool for solving various problems of automatic pattern recognition (APR). Meanwhile, its capabilities have not been fully exploited. In particular, almost no studies have addressed the advantages of the MID principle over traditional methods and approaches in problems of automatic image recognition (AIR), especially, for half-tone images as one of the most complex cases in the theory and practice of AIR [2, 3]. The present paper seeks to fill this gap. It proposes a method of directed enumeration (MDE) of a set of alternatives taking into account the metric properties of the decision statistic of MID, as opposed to the traditional method of complete enumeration of competing hypotheses.

MINIMUM INFORMATION DISCRIMINATION CRITERION

Let a set of $R > 1$ half-tone images $X_r = \|x^r_{uv}\|$ ($u = \overline{1, H}$ and $v = \overline{1, W}$) be specified. Here H and W are the image height and width, respectively, $x^r_{uv} \in \{1, 2, \ldots, x_{\max}\}$ is the intensity of an image point with coordinates (u, v); r is the reference number ($r = \overline{1, R}$), and $x_{\max}$ is the maximum intensity. It is assumed that the references X_r define some classes of images, for example, as a method of protection against noise. Furthermore, the objects belonging to each class have some features in common or similar characteristics. The common thing that unites objects in a class is called a pattern. It is required to assign a new input image $X = \|x_{uv}\|$ to one of the R classes. This is a typical example of diagnostics problem (learning image recognition) for objects of non-numerical nature (ONN) [4].

Recall that the term ONN refers to elements of a mathematical space that is not linear (vector). The procedures for constructing decision rules for the problem in question are generally divided into deterministic and statistical approaches. The deterministic approach is currently the most widely used. This approach seeks to determine a certain distance (measure of similarity) between any pairs of objects in the space of ONN. For AIR, one often uses the criterion based on the standard metric l_1:

$$\rho_1(X/X_r) = \frac{1}{WH} \sum_{u=1}^{H} \sum_{v=1}^{W} \left| x_{uv} - x^r_{uv} \right| \to \min\,. \tag{1}$$

However, this approach does not always provide satisfactory results. This is due, first, to the well-known [2] variability of visual patterns, and second, to the presence of noise in the input image X, such as unknown intensity of light sources or simply random distortions of some points of the image.

In the deterministic approach, the indicated problems are usually solved by adding new images to the set of reference images (SRI), which, in turn, leads to a sharp increase in its size. The above-mentioned difficulties are overcome by invoking a statistical approach [2, 3]. In this approach, it is assumed that a reference image X_r defines the spectral power density (SPD) of a certain (hypothetical) two-dimensional (spatial) random signal. This interpretation seems justified considering the unquestionable similarity between the rigorous specification of the SPD [2] as a function of the frequency distribution of the signal power and the definition of a half-tone image as a function of the black color intensity distribution in a plane, i.e., in spatial frequency. Thus, it is required to verify R hypotheses on the SPD X_r ($r = \overline{1, R}$) of the input image signal X. Here the general problem of counter hypotheses on the class of the examined hypothetical distributions arises. However, it can be solved by the simple logical reasoning: for a set of all conceivable distributions there is a unique, Gaussian, law, for which helpful information lies in the SPD shape of the analyzed signal. The above logic, together with the MID principle [1] and its well-known transformations in the frequency region [5], leads to the optimal decision rule

$$\rho_{KL}(X/X_r) = \frac{1}{WH} \sum_{u=1}^{H} \sum_{v=1}^{W} \left(\frac{x_{uv}}{x_{uv}^r} - \ln \frac{x_{uv}}{x_{uv}^r} \right) - 1 \to \min. \tag{2}$$

Here the statistic $\rho_{KL}(X/X_r)$ defines the Kullback–Leibler information discrimination (ID) [6] between the observed image signal X and its rth reference from the SRI $\{X_r\}$.

Thus, the AIR procedure in this case involves a multichannel processing scheme in which the number of channels R is given by the number of reference images. Decision making is based on the statistic minimum criterion from expression (1) for traditional AIR methods [2] or from expression (2) when using the MID principle.

METRIC PROPERTIES OF DECISION STATISTIC OF MINIMUM INFORMATION DISCRIMINATION

We consider the theoretically and practically most important case $R \gg 1$, where the AIR problem is solved with a SRI containing hundreds and thousands of images. For the specified conditions, practical implementation of the optimal decision rule (2) by the R-channel processing scheme encounters the obvious problem of its computational complexity and even feasibility, especially considering the labor-consuming procedure of image equalization in many parameters: size, color, project view, etc. The present work seeks to develop methods other than complete enumeration of SRI to solve the above problem.

We first note the metric properties of the MID decision statistic $\rho_{KL}(X/X_r) \geqslant 0$, which is equal to zero only in the ideal case of coincident input and reference signals. Therefore, we first transform the MID criterion (2) to a simplified (for practical implementation) form [7]:

$$W_\nu(X): \quad \rho_{KL}(X/X_\nu) < \rho_0 = \text{const}. \tag{3}$$

Here ρ_0 is the threshold for the admissible ID on the set of similarly-named images due to their known variability. The value of this threshold is easy to find experimentally. In fact, expression (3) defines the stopping condition for the enumeration procedure using the MID criterion (2).

Thus, in decision-making based on the MID principle (3), instead of looking through all references, one needs to calculate the value of ID only until it becomes smaller than a certain threshold level. It is easy to see that this circumstance in itself should reduce the amount of enumeration by 50% on the average. In other words, use of the stopping rule (3) halves the amount of computation and thus considerably relieves the problem of real-time implementation of AIR. This is the main advantage of the MID over all best-known statistical analogs based on the classical (Bayes) criteria: minimum average risk, maximum a posteriori probabilities, and others [2]. Meanwhile, a gain in computational complexity and productivity is far from being the only advantage of the MID principle in the APR problem.

Indeed, the general formulation (2) allows one to consider it as the optimization problem and apply optimal solution search algorithms with the specified stopping condition (3). In this problem, it is required to find an image X_ν from the set of reference images $\{X_r\}$ that would minimize the MID statistic. In this case, the method involving complete enumeration of SRI is one of the many known optimization methods.

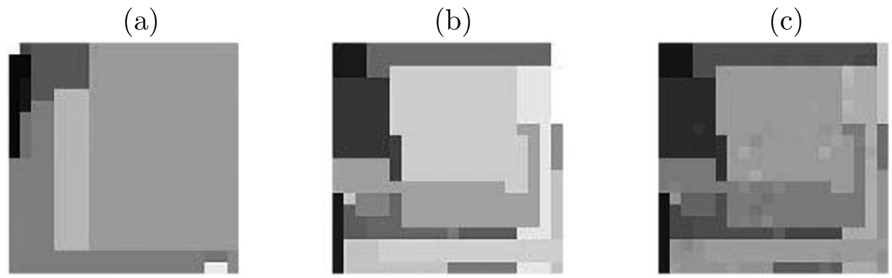

Fig. 1.

The main factors preventing the use of a more effective optimization method in the problem considered is that, first, the problem refers to the area of discrete mathematics and, second, in this problem, it is required to find the global minimum of the decision statistic (2). Apparently, the most appropriate method of finding the global extremum for the specified conditions is the random-walk method. Unfortunately, the well-known classical optimization techniques (for example, the genetic algorithm) ignore information on images and discriminations between them. Moreover, most of the discrete optimization algorithms do not have a precise criterion for stopping enumeration [3]. In this case, the MID principle is again helpful. Based on expression (3), it is possible to formulate the required algorithm stopping criterion. This ensures that if a solution of the problem exists (i.e., if the input image belongs to one of the classes given by $\{X_r\}$), it will be found. A natural development of this idea is the method of directed enumeration of SRI proposed below, which fully exploits the metric properties of the MID decision statistic (3).

DIRECTED ENUMERATION METHOD

Following the general computation scheme (2), (3), we reduce the problem of AIR of X to a check of the first N variants $X_1, \ldots, X_N$ from the specified R-set of alternatives $\{X_r\}$ subject to the condition $N \ll R$. If, at least, one of them, namely X_ν ($\nu \leqslant N$), meets the stopping requirement (3), the enumeration of the optimal solution by the MID criterion (2) will end with it. However, it can generally be assumed that none of the first N alternatives passes the check (3) in the first step. Then, it is possible to check the second group from N reference images within the set $\{X_r\}$, then the third group, etc., until condition (3) is satisfied. There is also another, more rational, method to solve the problem in question.

Let us arrange the images of the first control sample $X_1, \ldots, X_N$ in decreasing order of their IDs $\rho_{KL}(X/X_n), n = \overline{1, N}$. As a result, we have an ordered (ranged) sequence of reference images of the form

$$\{X_{i_j}\} = \{X_{i_1}, X_{i_2}, \ldots, X_{i_N}\}, \qquad i_j \leqslant N.$$

The corresponding sequence $\{\rho_j\}$ of their IDs $\rho_j = \rho_{KL}(X/X_{i_j})$ ($i_j \leqslant N$) is a monotonically decreasing dependence. Its form, in particular, the rate of decrease, depends on both the pattern of the control sample of references $X_1, \ldots, X_N$ and its information theoretical characteristics, in particular, the value of the mutual information discriminations (MIDs) $\rho_{(j+1)/j} = \rho_{KL}(X_{i_{j+1}}/X_{i_j})$ between pairs of neighboring elements in the ordered control sample $\{X_{i_j}\}$.

Let us illustrate this dependence by using an experimental SRI specially generated for this purpose. During its generation, we obtained a large-size ($R = 5000$) set of reference images $\{X_r\}$ of size 20×20 pixels. Each image (Figs. 1a and 1b) consists of several rectangles with intensity randomly chosen from the range of [1, 256]. Intensity equal to 1 corresponded to absolutely black color, and intensity equal to 256 corresponded to white color. This provided stringent conditions for the subsequent automatic recognition of half-tone images X.

In the first step of the computations, the control sample pattern $X_1, \ldots, X_N$ was chosen at random and its size was set equal to $N = 100$. The ID value was calculated using expression (2). From the computation results, an ordered control sample $\{X_{i_j}\}$ (Figs. 2a and 2b) was obtained. Its fragment containing the last 45 elements $X_{i_{56}}, \ldots, X_{i_{100}}$ is depicted in Fig. 2a as a curve of their IDs $\{\rho_j\}$ relative to a certain input image $X \in \{X_r\}$. For comparison, Fig. 2b shows a curve of the corresponding MIDs $\{\rho_{(j+1)/j}\}$ (curve *1*). It is evident from the figure that, for the ordered control sample elements $\{X_{i_j}\}$, the last dependence, as well as the sequence of IDs $\{\rho_j\}$, has the nature of damped oscillations.

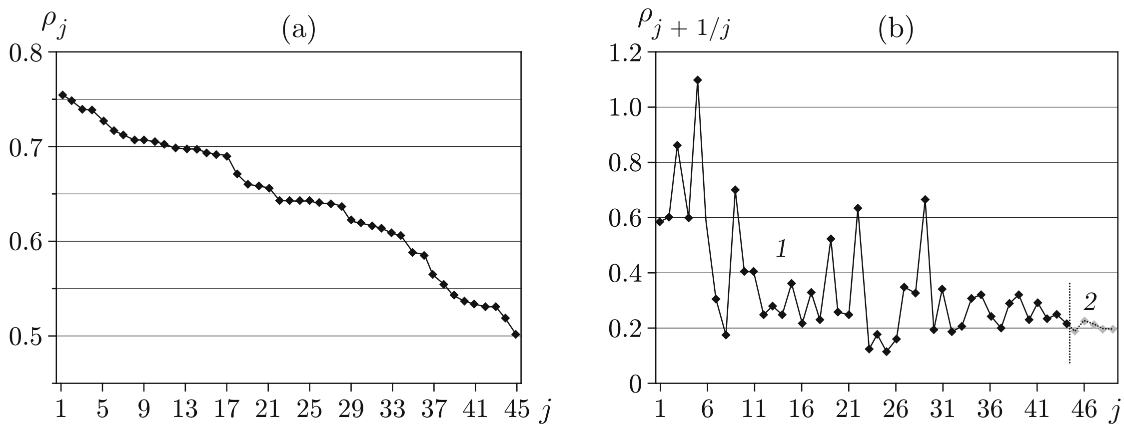

Fig. 2.

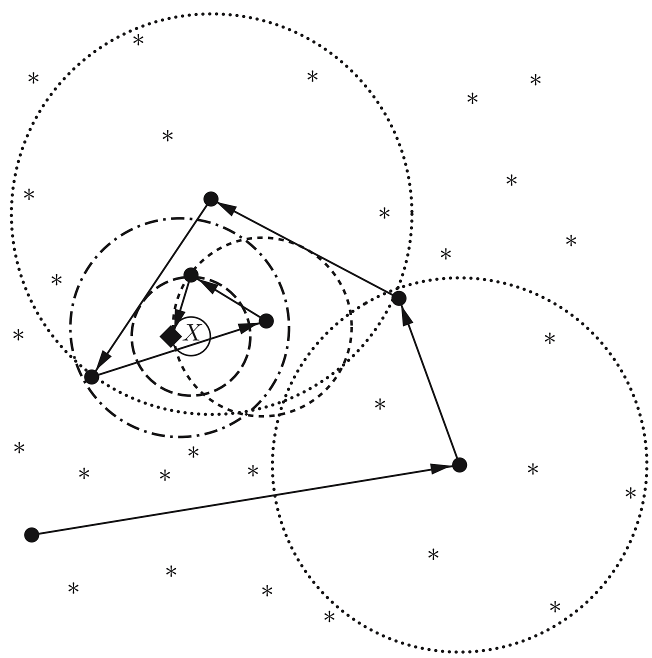

Fig. 3.

This observation is strengthened by the results of extrapolation of the sequence of MIDs in Fig. 2b (curve *2*) five steps beyond the boundaries of the control sample using the Pth-order linear prediction formula:

$$\widehat{\rho}_{(N+1)/N} = \sum_{i=1}^{P} a_i \rho_{(N-i+1)/(N-i)}. \tag{4}$$

Here $\{a_i\}$ is the vector of the autoregression (AR) coefficients. The Berg–Levinson recursive computational procedure was used for the AR analysis [8], and the autoregressive order was $P = 10$. This implies the main idea of the MDE: the last element X_{i_N} of the ordered control sample $\{X_{i_j}\}$, as the best approximation of the required image X, is used as the reference point for enumeration of the most suitable candidates for the next control sample. In this case, the data of extrapolations of the MIDs (4) can serve as a reference point for determining the maximum possible differences (in the information theoretical sense) between the reference images from the future control sample with respect to the reference point X_{i_N}. This is illustrated by a diagram of the MDE enumeration procedure (Fig. 3). Here asterisks denote all available reference images, X is the input image, and the rhombus denotes the reference the closest to X. It determines the required optimal solution of the problem. The enumeration trajectory is shown by a broken directed line. The bold points on it denote the sequence of the images the closest to the optimum X_{i_N} after several successive computation steps. Circles denote the boundaries of the corresponding control sample point $X_1, \ldots, X_N$. Their radii are determined according to expression (4). It is evident that the enumeration trajectory has the shape of a twisted helix, in accordance with the conclusions on Figs. 2a and 2b.

SYNTHESIS OF ALGORITHM

Following the definition of ID (2), we generate an $(R \times R)$ matrix $\mathrm{P} = \|\rho_{ij}\|$ of ID values $\rho_{ij} = \rho_{KL}(X_i/X_j)$, $i, j \leqslant R$. This computationally complex operation needs to be made only once: in the preliminary computation step and for each concrete SRI. After that, within the available R-set of references $\{X_r\}$, we specify an arbitrary first control sample $X_1, \ldots, X_N$ of a certain fixed size N, use it to obtain a number of data $\{X_{i_j}\}$ ranged by the MID criterion (2), and ultimately find the first local optimum X_{i_N}. The first computation step ends with this. In the second step, for the distinguished reference image X_{i_N} from the matrix P, we find the set of $M < R$ images $X^{(M)} = \{X_{i_{N+1}}, \ldots, X_{i_{N+M}}\}$ $(i_j \leqslant R)$ that are separated from the image X_{i_N} by the distance (2) not exceeding the threshold value $\widehat{\rho}_{(N+1)/N}$:

$$\left(\forall\, X_i \notin X^{(M)}\right)\left(\forall\, X_j \in X^{(M)}\right)\ \Delta\rho(X_i) \geq \Delta\rho(X_j). \tag{5}$$

Here

$$\Delta\rho(X) = \left|\rho_{KL}(X/X_{i_N}) - \widehat{\rho}_{(N+1)/N}\right|$$

is the deviation of the MID $\widehat{\rho}_{(N+1)/N}$ extrapolated by formula (4) relative to the ID between the pair of images X and X_{i_N}. In Fig. 3, each such set is bounded by a corresponding circle with center at the point X_{i_N}. To this set we add one more $(M+1)$th element $X_{i_{N+M+1}}$ that did not fall in the control sample in the previous computation step. This brings some randomness to the search procedure as a method of attaining a global optimum in a finite number of steps (computation steps). As a result, for the analysis we obtain the second control sample of reference images

$$\{X_{i_N}, \ldots, X_{i_{N+M+1}}\}, \qquad i_j \leqslant R.$$

Next, all computations of the first step are repeated cyclically until, in some Kth step, an element X_{i_N} satisfies the stopping condition (3):

$$\rho_{KL}\left(X/X_{i_N}\right) < \rho_0. \tag{6}$$

At this moment, the input image in Fig. 3 is within the set of the control points of the last computation step. In this case, a decision is made in favor of the closest pattern X^* or, at worst, after enumeration of all alternatives from the set $\{X_r\}$ but in the absence of a solution from (6), the conclusion is drawn that the input image X cannot be assigned to any class from the SRI and that it is necessary to switch to the decision feedback mode. Generally, there may be a considerable gain in the total number $N + (M+1)K \leqslant R$ of checks carried out according to (6) compared to the size of the SRI used. This is the effect of the directed enumeration.

Thus, the system of expressions (2)–(6) defines the proposed MDE in the AIR problem.

RESULTS OF EXPERIMENTAL STUDIES

After the generation of the experimental SRI presented above, the following experiment was performed 200 times. A new distorted image X was generated on the basis of a certain reference image X_r (each time chosen at random) by reducing the intensity of all of its points (blanking) and a subsequent change in the intensity of some points X_r. Typical diagrams of exactly this pair of images (X_r and X) are presented in Figs. 1b and 1c. In each case, the problem of AIR of X from the set of all its admissible alternatives $\{X_r\}$ was solved. This was first done using the metric l_1 from expression (1). The following MDE parameters were chosen: $N = 128$ and $M = 80$. The threshold $\rho_0 = 30$ was chosen experimentally. Here the exact solution $X^* = X_r$ was obtained in 92.5% of the cases. For each solution, it was required to check, on the average, 2550 images, or 51% of the size of the entire SRI. A histogram of the number of checks of images from the SRI (in percentage) is given in Fig.4a. In view of the computational complexity of the image dynamics equalization procedure (which is necessary in such cases), about a factor of two increase in the processing rate is at first sight a considerable achievement. However, this conclusion needs to be explained. First, the recognition error is large enough. In particular, it turned out that the input image X in Fig. 1c was closer [in metric (1)] to the image in Fig. 1a than to the initial image X_r (see Fig. 1b). In addition, for 37% of

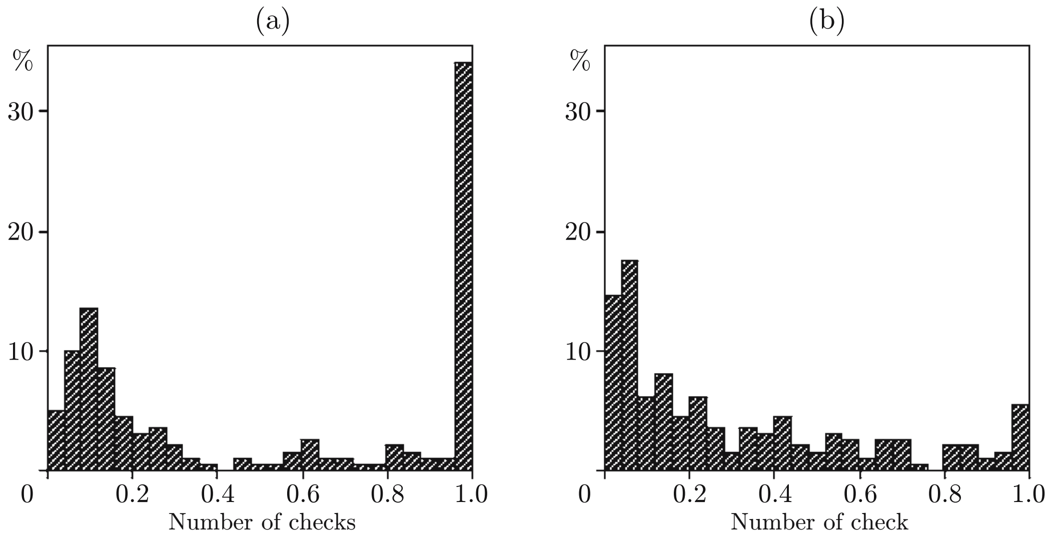

Fig. 4.

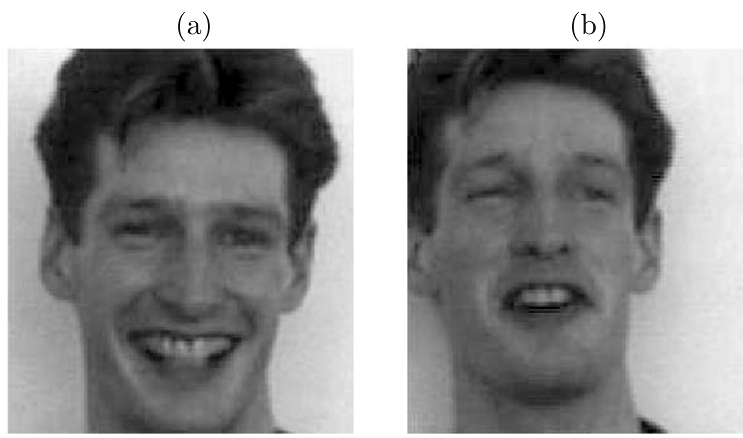

Fig. 5.

the input images X, the stopping condition (3) was not satisfied, and the algorithm therefore checked all R references.

A similar problem using the same MDE was solved later in the Kullback–Leibler metric (2). A histogram of the number of checks carried out by algorithm (2)–(6) for this case is shown in Fig. 4b. The average number of checks was about 31% of the size of the SRI. With a probability of 80%, it does not exceed 2900 or 58% of the total number of reference images for the check. In this case, condition (3) was not satisfied for any reference from the SRI for only 2.5% of the initial images; therefore, all R alternatives were checked. In 99.5% of the cases, the exact solution $X^* = X_r$ was obtained. There is an unquestionable advantage in using the Kullback–Leibler discrimination in this problem and this is due to the fact that, in our example, the input images were artificially distorted (blanked).

Let us now consider the case where all images are equally illuminated. We use a real database of photographs of people.[1] Of 5000 photographs, 300 different people were selected as references $R = 1200$ of the most different images (Fig. 5a). In addition, 1000 more photographs of the same people were used in the test (Fig. 5b).

In this case, for the Kullback–Leibler metric (2) and the MDE (3)–(6) with the parameters N, M, and ρ_0 from the previous experiment, we obtain an average number of checks of 32% of the size of the SRI at an error probability of 4.5%. However, use of the metric l_1 from expression (1) and the MDE algorithm with a threshold $\rho_0 = 15$ gives much better results. The error probability reduces to 1.5%, and the average number of checks decreases to 23% of the size of the SRI.

[1] http://cswww.essex.ac.uk/mv/allfaces/index.html.

CONCLUSIONS

The problem of increasing the computation speed has attracted considerable interest of experts in both the theory and practice of AIR. When the power of the set of references is hundreds and thousands of units, most of the well-known algorithms operating by comparing an input image with each reference cannot be implemented in the real-time mode. Therefore, the problem of reducing the computational complexity for large SRIs has recently been given increased attention. For this purpose, this paper proposed a directed enumeration method based on an information theoretical approach which uses the metric properties of the MID decision statistic [1] and possesses wide functional capabilities and high operational properties. The point of fundamental importance in this enumeration method is the automatic stopping rule (3). In the most unprofitable version of its use, the method almost halves the amount of computation. The use of the proposed method in the formulation (2)–(6) reduces the computational complexity by 25–30%. The quality of the solution X^* reached by the MDE is comparable to that obtained by continuous enumeration of the SRI. Moreover, as shown in the experiments, the MDE algorithm can yield good results not only with the Kullback–Leibler metric (2) but also with the traditional metric l_1 (1).

REFERENCES

1. V. V. Savchenko and A. V. Savchenko, "Minimum Information Discrimination Principle in the Problem of Recognition of Discrete Objects," Izv. Vuzov Rossii, Ser. Radioelektronika, No. 3, 10–18 (2005).
2. *Discrete-Time Signal Processing*, Ed. by A. V. Oppenheim (Prentice-Hall, 1989).
3. Yu. E. Voskoboinikov and L. A. Litvinov, "Choosing the Stopping Iteration in Iterative Algorithms of Image and Signal Reconstruction ," Avtometriya **40** (4), 3–10 (2004) [Optoelectr., Instrum. Data Process. **40** (4), 3–9 (2004)].
4. A. I. Orlov, "Mathematical Methods for Material Study and Diagnostics," Zavod. Lab. **69** (3), 53–64 (2003).
5. V. V. Savchenko, "Discrimination of Random Signals in the Frequency Domain," Radiotekh. Elektron. **42** (4), 426–431 (1997).
6. S. Kullback, *Information Theory and Statistics* (Wiley, New York, 1959).
7. D. Yu. Akatiev and V. V. Savchenko, "Detection of the Stochastic Process Discord Based on the Principle of Minimum Informative Divergence," Avtometriya **41** (2), 68–74 (2005) [Optoelectr., Instrum. Data Process. **41** (2), 61–66 (2005)].
8. S. L. Marple, Jr., *Digital Spectral Analysis with Applications* (Englewood Cliffs, New York, 1987).