The Joint Use of Image Equivalence and Image Invariance in Image Recognition[1]

I. B. Gurevich and I. A. Jernova

Scientific Council on Cybernetics, Russian Academy of Sciences, ul. Vavilova 40, Moscow, GSP-1, 119991 Russia

e-mail: igourevi@ccas.ru; jer_irina@mail.ru

Abstract—The work is devoted to the investigation of the key concepts of mathematical theory of pattern recognition: image equivalence and invariance. Different ways to determine equivalence on the image set are considered and relations between image equivalence and invariance are analyzed. It is proved that the image recognition task in standard formulation can be reduced to the task that has a correct algorithm in the framework of AEC algebraic closure (with certain restrictions on image transformations).

1. INTRODUCTION

Over several years, research into the development of the mathematical apparatus for analysis and estimation of information represented in the form of images has been carried out at the Scientific Council on Cybernetics of the Russian Academy of Sciences [9, 10, 21]. This work presents new results concerning the establishment of the conditions for existence of the class of effective algorithms that includes the algorithm for correct solution of an image recognition task. The proposed method for verifying the condition feasibility is based on the new definition of image equivalence introduced with reference to the special formulation of the image recognition task. It is shown that the efficient class of algorithms based on estimate calculations (AEC) [2] contains such an algorithm in its algebraic closure. The existence theorem is the main result.

The selection of the algorithm that correctly classifies images on the basis of their descriptions is a topical problem in image recognition. The approach to image recognition developed by the authors is a specialization of the algebraic approach to recognition and classification problems introduced by Yu.I. Zhuravlev [2, 20]. It is based on the following idea. No accurate mathematical models have been developed for poorly formalized fields, such as geology, biology, medicine, and sociology. However, decent methods based on heuristic considerations have a great practical effect in many cases. Therefore, it is sufficient to construct a family of such heuristic algorithms for solving corresponding tasks and then to construct an algebraic closure of this family. The existence theorem is proved; it asserts that any task from the set of tasks related to the investigation of poorly formalized situations turns out to be solvable in the closure [2].

Image recognition is one of the classical examples of the tasks with incompletely formalized and partially contradictory information. This gives reason to believe that the algebraic approach applied to image recognition can yield convincing results and, hence, the most perspective direction for developing the desired mathematical apparatus for the analysis and estimation of the information represented by images is the "algebraization" in this field.

Note that the idea of developing a unified algebraic theory covering different approaches and operations used in image and signal processing has a certain history that starts with the works of von Neumann and was continued by S. Unger, U. Grenander [8], M. Duff, Yu.I. Zhuravlev [2, 20], G. Matheron, G. Ritter [16], J. Serra [17], and others. The results presented here are related to the descriptive approach to image analysis and recognition [9, 10], a line of research in the field of image algebras different from those mentioned above, and they are completely original.

Unfortunately, the algebraic apparatus developed by Zhuravlev failed when applied directly to an image recognition task. This is mainly caused by the complexity of the object of recognition, the image, and, as a consequence, by the substantial differences between the image recognition task and the classical one:

• The standard object of classical recognition theory is usually described by the feature set; there is no natural way for image description without losing essential information about the image; the known ways of image description are either too complicated and time-consuming (e.g., image representation as a raster) or semantically primitive (feature set);

[1] This work was partially supported by the Russian Foundation for Basic Research, project nos. 02-01-00182, 03-07-06056; by the federal target oriented program "Research and Development in Priority Directions of Science and Technology" in 2002–2006, project no. 37.011.11.0016; and by the program for fundamental Research of the Department of Mathematical Sciences of the Russian Academy of Sciences "Algebraic and Combinatorial Methods for Mathematical Cybernetics," project no. 10002-251/OMN-2/024-025/120503-063.

Received July 25, 2003

• Several images that differ in brightness, contrast, scale, or viewpoint can correspond to the same object (scene); in the context of the recognition task, this means that different images of the same object should be equally classified by the recognition algorithm.

This explains the interest that has arisen in the study and use of image equivalence in recognition tasks [11, 13].

2. DIFFERENT DEFINITIONS OF IMAGE EQUIVALENCE

One of the basic ideas of the proposed approach to image recognition is the idea that a certain image is not the only visual representation of the object but one of many possible. This means that one object can correspond to several images that differ in scale, observation angle, illumination, etc.

Thus, image recognition tasks in the framework of the proposed approach consist in the following: (a) the image is a partial representation of a certain entity (object, scene) that should be described through these partial representations only; (b) on the image set, the equivalence relation can be defined as correspondence of the images to the same entity; (c) images of one class should have equal vectors of belonging; and (d) transition from images to features is not predetermined, so features can be varied in applied tasks.

There are several ways to determine the equivalence relation on the image set. Note that the equivalence relation should be reflexive, symmetric, and transitive at the same time it splits the entire set of images into nonoverlapping classes: image equivalence classes.

Hereafter, the image will be regarded as a function that satisfies the following definition.

Definition 2.1. *Image,* or *image function,* is a real-valued function $f(x, y)$ such that $f(x, y) \neq 0$ on the bounded support and $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)dxdy > 0$.

2.1 Equivalence Based on Groups of Transformations

Let us consider a definition of equivalent images based on the hypothesis that images of the same object (scene), differing in scale, illumination, and observation angle, are equivalent in a sense (in the common or everyday meaning of the term). Let us fix a certain set of transformations that describe changes in scale, illumination, or observation angle. At first, the case where the transformations form a group is considered.

Let F be a space of image functions and $G = \{g_1, \ldots, g_n\}$ be a group of transformations.

Definition 2.1.1. Two images $f_1(x, y)$ and $f_2(x, y)$ are *equivalent* if $\exists g \in G$: $f_1(x, y) = g(f_2(x, y))$.

Let us introduce the notation $f_1 R f_2$; that is, f_1 and f_2 are equivalent; i.e., they are related by the equivalence relation R.

The equivalence relation should be reflexive, symmetric, and transitive. Let us check these properties for the definition proposed.

(1) Reflexivity.

$f = e(f)$ (by virtue of the existence of a unit element in a group).

$\forall f \Rightarrow f R f$ is fulfilled.

(2) Symmetry.

$f_1 = g(f_2) \Rightarrow f_2 = g^{-1}(f_1)$ (since each nonzero element of a group has an inverse element).

$\forall f_1, f_2 \; f_1 R f_2 \Rightarrow f_2 R f_1$ is fulfilled.

(3) Transitivity.

$f_1 = g_1(f_2)$, $f_2 = g_2(f_3) \Rightarrow f_1 = g_1(g_2(f_3)) = g_3(f_3)$ (by virtue of closure).

$\forall f_1, f_2, f_3 \; f_1 R f_2, f_2 R f_3 \Rightarrow f_1 R f_3$ is fulfilled.

All three properties are fulfilled; therefore, R is the equivalence relation.

The definition proposed is a constructive definition in the sense that it suggests a way to construct classes of equivalent images. Let us consider simple examples of equivalent classes.

Example 2.1.1.

Let a finite group of transformations of order 3 be given; it is described by the following equations: $G = \{g, g^{-1}, e\}$, $g \cdot g = g^{-1}$, $g^{-1} \cdot g^{-1} = g$, and $g \cdot g^{-1} = g^{-1} \cdot g = e$. Let F be the space of image functions. Each element of the group of transformations operates in the following way [4]:

g: $f_1(x, y) \in F \longrightarrow f_2(x, y) \in F$, g is a bijection.

The effect of the element of the group of transformations on the image function can be interpreted in the following way. Let us consider $f(x, y)$ in polar coordinates $f(\rho, \theta)$, by replacing the variables $x = \rho\cos\theta$, $y = \rho\sin\theta$. Let g be a rotation trough $2\pi/3$ angle, g^{-1} a rotation trough $-2\pi/3$, and e a rotation trough 0. Here, the equations $g \cdot g = g^{-1}$ and $g^{-1} \cdot g^{-1} = g$ are fulfilled. Then,

$$g(f(\rho, \theta)) = f\left(\rho, \theta + \frac{2\pi}{3}\right). \qquad (2.1.1)$$

Image equivalence classes are formed according to the following rule: two images are equivalent if one can be derived from another by rotating it trough $2\pi/3$, $-2\pi/3$, or 0.

Example 2.1.2.

Let $G = \{g, g^{-1}, e\}$ be a group of transformations described by the equations $g \cdot g = g^{-1}$, $g^{-1} \cdot g^{-1} = g$, and $g \cdot g^{-1} = g^{-1} \cdot g = e$. Let $F = \{f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8\}$ be a space of image functions.

The mapping g is defined as follows

$$g : \begin{pmatrix} f_1 & f_2 & f_3 & f_4 & f_5 & f_6 & f_7 & f_8 \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ f_2 & f_3 & f_1 & f_5 & f_6 & f_4 & f_7 & f_8 \end{pmatrix}. \qquad (2.1.2)$$

Let us recall that equivalence relation R was introduced as follows: $f_i R f_j$, if $\exists g^0 \in G$: $f_i = g^0(f_j)$, $i, j \in \{1, 2, \ldots, 8\}$.

The following relations are fulfilled in this example:

$$f_1 R f_1,\ f_1 R f_2,\ f_1 R f_3,\ f_2 R f_2,\ f_3 R f_3,\ f_2 R f_3, \qquad (2.1.3)$$

$$f_4 R f_4,\ f_4 R f_5,\ f_4 R f_6,\ f_5 R f_5,\ f_6 R f_6,\ f_5 R f_6, \qquad (2.1.4)$$

$$f_7 R f_7, \qquad (2.1.5)$$

$$f_8 R f_8. \qquad (2.1.6)$$

The following image equivalence classes are formed:

$$K_1 = \{f_1, f_2, f_3\}, \quad K_2 = \{f_4, f_5, f_6\}, \quad K_3 = \{f_7\}, \quad K_4 = \{f_8\}. \qquad (2.1.7)$$

The notion of image equivalence introduced above is convenient, and it has indubitable advantages in respect to applicability to pattern recognition theory. Let us formulate the major advantages of this definition.

(1) The physical interpretation of this definition conforms to our view of image equivalence, according to which images are considered to be equivalent if one differs from another due to a certain transformation, e.g., rotation, translation, scaling, illumination change, etc.

(2) The introduced definition allows one to reduce an image recognition task in terms of equivalence classes to the image recognition task in standard formulation. This reduction is carried out by substituting each equivalence class by a single image, a class representative, with certain restrictions on the type of transformations. As will be shown further, the theorem on the correctness of an algebraic closure of AEC algorithms can be proved for the reduced image recognition task [13].

(3) It is essential that, in the case where transformations form a group, the mathematical methods of image invariant construction are developed, which allows images to be compactly described with no extra calculating effort [15]. Particularly, the parametric image models can be constructed on the basis of invariant features. In this case, equivalent images have the same descriptions according to the equivalence definition given above.

Let us consider some additional aspects of image equivalence regarding groups of transformations.

Only transformations that form a group were used in this definition. The question arises as to how important this requirement is. In other words, what will happen to the equivalence definition if the requirement to be a group on the transformation set is weakened? For instance, let us remove the axiom of inverse element existence from group axioms. Let us consider a semigroup $G = \{g_1, \ldots, g_n\}$. Suppose that two images $f_1(x, y)$ and $f_2(x, y)$ are connected with equivalence relation R, if $\exists g \in G$: $f_1(x, y) = g(f_2(x, y))$. Let us check whether this relation is reflexive, symmetric, and transitive, i.e., whether this relation is an equivalence relation.

(1) Reflexivity.

$f = e(f)$ (by virtue of the existence of a unit element).

$\forall f \Rightarrow f R f$ is fulfilled.

(2) Symmetry.

$f_1 = g_1(f_2) \Rightarrow f_2 = g_2(f_1)$?

$\forall f_1, f_2\ f_1 R f_2 \Rightarrow f_2 R f_1$ is not fulfilled.

(3) Transitivity.

$f_1 = g_1(f_2)$, $f_2 = g_2(f_3) \Rightarrow f_1 = g_1(g_2(f_3)) = g_3(f_3)$ (by virtue of a closure).

$\forall f_1, f_2, f_3\ f_1 R f_2, f_2 R f_3 \Rightarrow f_1 R f_3$ is fulfilled.

The reflexivity and transitivity properties are fulfilled, while the symmetry property fails without the axiom of existence of inverse elements. Consequently, this relation is not an equivalence relation. Further weakening of group requirements leads to the failure of transitivity and reflexivity property fulfillment.

Since the practical importance of the introduced equivalence definition consists in its application for proving the theorem about the correctness of algebraic closure of AEC in an image recognition task, the group requirement can be considered necessary for the transformation set.

Also note that the question arises as to how to interpret group application for image functions with the use of an arbitrary group G of transformations of order n.

2.2. Equivalence Directed at the Recognition Task

Image equivalence definitions directed at the special setup of an image recognition task can be considered. Let a certain set of allowable images described by n-dimensional feature vectors be given. The set of allowable images is covered by a finite number of subsets called classes; let l classes $K_1, \ldots, K_l$ be given. Let a recognition algorithm A be given which constructs an l-dimensional information vector on the basis of an n-dimensional description vector. Let us recall that a vector of the object's belonging to a class is called an information vector; here, information vector element values (0, 1, Δ) are interpreted as follows: "object does not belong to the class," "object belongs to the class," and "the algorithm fails to determine whether the object belongs to the class or not," respectively [2]. We consider that each recognition algorithm $A \in \{A\}$ can be represented as a consequent implementation of algorithms B and C, where B is a recognition operator that transforms training information and a description of the allowable object into a numerical vector, called the estimation vector, and C is a decision rule that transforms an arbitrary numerical vector into an information vector.

The performance of a recognition algorithm ($A = BC$) can be sketched in the following way.

$$\begin{array}{c} \text{Feature description of an image } \alpha = \\ = (\alpha_1, \alpha_2, \ldots, \alpha_n). \\ \Big\downarrow \textit{Recognition operator B.} \\ \text{Vector of estimations for classes } \beta = \\ = (\beta_1, \beta_2, \ldots, \beta_l). \\ \Big\downarrow \textit{Decision rule C.} \\ \text{Information vector } \gamma = (\gamma_1, \gamma_2, \ldots, \gamma_l). \end{array}$$

Thus, during the solution of an image recognition task, an object of recognition, i.e., an image, is described with the help of three different vectors: the n-dimensional feature vector, the l-dimensional vector of estimations for classes, and the l-dimensional information vector. This permits consideration of three levels of equivalence.

Definition 2.2.1. Images are called *equivalent* with respect to the recognition algorithm A if their information vectors constructed by the algorithm A coincide.

Definition 2.2.2. Images are called *equivalent* with respect to the recognition operator B if their vectors of estimations for classes constructed by the algorithm B coincide.

Definition 2.2.3. Images are called *equivalent* if their feature vectors coincide.

It is easy to verify that the introduced equivalence definitions are correct, since they represent an equality relation which is a special case of the equivalence relation. Moreover, image equivalence in terms of Definition 2.2.3 implies image equivalence in terms of Definition 2.2.2, which, in turn, leads to image equivalence in terms of Definition 2.2.1.

Let us consider the equivalence definition based on the group of transformations introduced in the previous paragraph. If an image is described by the vector of features invariant relative to the group transformations, images equivalent in terms of Definition 2.1.1 will have the same feature descriptions. In this case, image equivalence based on group transformations implies image equivalence in terms of Definitions 2.2.1, 2.2.2, and 2.2.3, i.e., equivalence directed at the recognition task.

2.3. Equivalence with Respect to a Metric

In some cases, when images in the recognition task are not strictly equivalent in terms of the definitions introduced above, it makes sense to consider image proximity with respect to a metric specified in a space of formal image descriptions or in a space of features representing images. As this takes place, ε-equivalence of images can be considered for a certain *a priori* given parameter ε. We recall that we regard an image as a real-valued function $f(x, y)$ meeting the conditions of Definition 2.1. Let $D(f(x, y))$ be a certain formal image description (vector of features, image defined at each point of a raster, analytical representation of image function, etc.) and $L(D(f(x, y)))$ be a metric space of formal image descriptions with metric ρ.

Definition 2.3.1. Two images $f_1(x, y)$ and $f_2(x, y)$ are ε-*equivalent* with respect to metric ρ if $\rho(f_1(x, y), f_2(x, y)) < \varepsilon$, where ε is an *a priori* given constant.

3. INVARIANT-BASED CONSTRUCTION OF EQUIVALENCE CLASSES

The construction of equivalence classes becomes a topical problem since the notion of equivalence is one of the crucial notions in the proposed approach to image recognition. Let us consider here one of the ways to decompose an image set into equivalent classes based on the construction of image invariants.

Let the space F of image functions $f(x, y)$ be given. Let F be a ring. Let us define a homomorphism $\varphi : F \longrightarrow F'$. The kernel of homomorphism is $\mathrm{Ker}\,\varphi = \{f(x, y)|\varphi(f(x, y)) = 0\}$. Let us construct a factor set F from the kernel of homomorphism $F\backslash\mathrm{Ker}\,\varphi$. The set (ring) of image functions is split into nonoverlapping cosets from the kernel of homomorphism. These cosets are considered to be image equivalence classes.

$$\{f(x, y)|\varphi(f(x, y)) = a\} = \mathrm{Class}(a). \qquad (3.1)$$

(Each coset is defined as follows: Class $= b\mathrm{Ker}\,\varphi = \{bh|h \in \mathrm{Ker}\,\varphi\}$.)

The main idea is to use invariants as the basis for constructing the kernel of the homomorphism. Let us recall one of the definitions of an invariant.

Definition 3.1. [3] An *invariant* is a mapping φ of the considered aggregate of mathematical objects M, supplied with a fixed equivalence relation ρ, into another aggregate of mathematical objects N, which is constant on equivalence classes M on ρ (more precisely, the invariant of equivalence relation ρ at M).

The definition of invariant implies that the mapping is constant on equivalence classes of the set. Hence homomorphism φ is invariant. In other words, the invariant φ must be constructed for the set F of functions $f(x, y)$ and, therewith, the set φ is the kernel of a certain homomorphism. In order to construct image equivalence classes, the kernel of homomorphism should be constructed and, then, a factor-ring from $\mathrm{Ker}\,\varphi$.

4. INVARIANTS IN IMAGE RECOGNITION

It is obvious from the above reasoning that invariance is a very important notion in image recognition theory. Note that image invariance is often employed for practical purposes. In many tasks of image analysis and processing, it is convenient to describe images by sets of invariant characteristics: they are more resistant to noise as compared to other features and, as a rule, provide high recognition quality. The proposed approach treats invariants not only as effective numerical features but also as an algebraic instrument for constructing equivalence classes in image recognition tasks. In order to cover these two positions of invariant application in image recognition, we recall some knowledge of invariant theory (Section 4.1) and briefly discuss the invariant features that will be most often used in image analysis (Section 4.2).

4.1. Basic Facts from the Theory of Invariants

The theory of invariants [1] in its classical definition (the algebraic theory of invariants) is an algebraic theory concerned with algebraic expressions (polynomials, rational functions and their aggregates) changing in a certain way during nondegenerate linear replacements of the variables. Here, generally speaking, not the entire linear group is considered (i.e., not the complete set of nondegenerate linear replacements of the variables) but its certain subgroup (e.g., orthogonal, symplectic, etc.)

One of the first objects invariant theory was concerned with were so-called *invariants of the form*. The form of order r of n variables with undefined coefficients is treated as

$$f(x_1, \ldots, x_n) = \sum_{i_1 + \ldots + i_n = r} a_{i_1 \ldots i_n} x_1^{i_1} \ldots x_n^{i_n}; \qquad (4.1.1)$$

after linear replacement of variables $x_i \longrightarrow \sum_{j=1}^{n} \alpha_{ij} x_j$, $1 \le i \le n$, where α_{ij} are real or complex numbers, it converts into the following form:

$$f'(x_1, \ldots, x_n) = \sum_{i_1 + \ldots + i_n = r} a'_{i_1 \ldots i_n} x_1^{i_1} \ldots x_n^{i_n}, \qquad (4.1.2)$$

so that the linear replacement of variables mentioned above determines a certain transformation of form coefficients $a_{i_1 \ldots i_n} \longrightarrow a'_{i_1 \ldots i_n}$.

Definition 4.1.1. [3] Polynomial $\varphi(\ldots, a_{i_1 \ldots i_n}, \ldots)$ of form $f(x_1, \ldots, x_n)$ coefficients is called *(relative) invariant of the form* if the following equation is fulfilled for all nondegenerate linear replacements of variables:

$$\varphi(\ldots, a'_{i_1 \ldots i_n}, \ldots) = |\alpha_{ij}|^q \varphi(\ldots, a_{i_1 \ldots i_n}, \ldots), \qquad (4.1.3)$$

where $|\alpha_{ij}|$ is the linear transformation determinant and q is a constant (weight). If $q = 0$, then invariant is called *absolute*. Thus, the simplest examples of an invariant are the discriminant $D = b^2 - ac$ of a binary quadratic ($n = r = 2$) form $f(x, y) = ax^2 + 2bxy + cy^2$ and the discriminant $\Delta = 3b^2c^2 + 6abcd - 4b^3d - 4ac^3 - a^2d^2$ of a ternary ($n = 2$, $r = 3$) form $f(x, y) = ax^3 + 3bx^2y + 3cxy^2 + dy^3$.

The classical setup of the main task of the theory of invariants is to calculate invariants and provide their complete description. Various formal processes (polarization, restitution, Kapelli identity law, Kelly Ω-process, etc.) were developed for this purpose. The development of the so-called symbolic method in the theory of invariants was the culminant point of this activity; it is a formal way to calculate all invariants of a degree no more than that given.

The global theory of semisimple groups and their representations developed by the 1930s promoted the following most common setup of the main task of the classical theory of invariants. Let an arbitrary group G and its finite-dimensional linear representation ρ in the linear space V on the field k be given. If $x_1, \ldots, x_n$ are the coordinates in V (in some basis), then each element $g \in G$ determines a linear replacement of variables $x_1, \ldots, x_n$. Performing this replacement of variables in the arbitrary polynomial $\varphi(x_1, \ldots, x_n)$, one obtains a new polynomial; hence g induces a certain transform (automorphism) of the ring of all polynomials $k[x_1, \ldots, x_n]$ on variables $x_1, \ldots, x_n$ on the field k. The polynomial $\varphi(x_1, \ldots, x_n)$, not changing under these transforms (i.e., when g runs over entire G), is called $\varphi(x_1, \ldots, x_n)$-*invariant of representation* ρ *for group* G.

From the very beginning, invariant theory revealed the circumstance that made it possible to survey the entire system of invariants as a whole: in all the considered cases, it was possible to separate a finite number of *basic invariants* $\varphi_1, \ldots, \varphi_m$, i.e., the invariants that made it possible to express each other invariant φ with the prescribed representation as their polynomial; i.e., $\varphi = F(\varphi_1, \ldots, \varphi_m)$. In other words, the algebra of invariants turned out to be finitely generated.

A new stage in invariant theory is connected to the extension of the number of tasks and geometric applications of the theory. Contrary to the classical theory of invariants, where, as a main object, the ring of polynomials of n variables over the field k with a group of automorphisms induced by the linear replacements of variables was considered, the contemporary theory of invariants considers arbitrary, finitely generated k-algebra R and an algebraic group G of its k-automorphisms. Instead of linear space V and representation ρ, an arbitrary affine algebraic manifold X and an algebraic group G of its algebraic transformations (automorphisms) are considered, so that R is a ring of regular functions on X and the effect of G on R is induced by the effect of G on X. Elements of R stationary with respect to G are invariants; their complete set constitutes k-algebra of R^G.

The definition of invariant connected with the group performance on the set is of interest for image recognition theory. Let Γ be a group acting in the set E. This means that a mapping $(\sigma, x) \longrightarrow \sigma \cdot x : \Gamma \times E \longrightarrow E$ is given with the following properties:

(1) $(\sigma \cdot \tau) \cdot x = \sigma \cdot (\tau \cdot x)$ for every $x \in E$, $\sigma, \tau \in \Gamma$;

(2) $\varepsilon \cdot x = x$ for every $x \in E$, where ε is a unit element of group Γ.

Then, the mapping $\mu_\sigma : E \longrightarrow E$ defined by the formula $x \longrightarrow \sigma \cdot x$ will be a bijection of the set E. By virtue of (1), mapping $\sigma \longrightarrow \mu_\sigma : \Gamma \longrightarrow \Gamma_E$ of the group Γ into the group of bijections of the set E is a homomorphism.

Definition 4.1.2. [1] Element $x \in E$ is called Γ-*invariant* or *invariant* if $\mu_\sigma(x) = x$ for all $\sigma \in \Gamma$.

The concept of invariants is one of the most important in mathematics, since the study of invariants is closely related to the classification tasks for objects of different kinds. In essence, the goal of each mathematical classification is to construct a certain complete system of invariants (the simplest, whenever possible), that is, a system separating any two nonequivalent objects from the set under consideration.

4.2. Image Description with the Help of Invariant Features

One of the most commonly used ways to describe images is their representation by feature sets. Invariants are often used as features for image description in practice; this is based on the following facts. The most effective methods among widely used image recognition algorithms are those invariant under certain transformations: in the simplest case, translation, rotation, and scaling, and in real application tasks, changes in intensity and contrast, blurring, viewpoint changes (for 3D scenes), etc., can be added to those mentioned above. The easiest way to achieve this is to use characteristics invariant to these transforms as algorithmic parameters. It is not always possible to ensure strict invariance of image features, since images are usually discrete. However, in some cases, it is possible to construct characteristics that are nonstrictly invariant and noise-immune. For the cases when transforms which necessitate the invariance form a group, the methods for constructing the invariants were developed. Moreover, an image can be restored with a certain accuracy from the fixed set of invariants of the same type. Computational experiments and the results of solved applied tasks show that invariants can be efficient for constructing an image description.

While constructing image descriptions, a question arises concerning the invariants to be chosen. The problem of invariant selection, as well as the number of invariants, is complicated by the fact that the known invariants are conceptually dissimilar in their properties, form, and geometric and physical interpretation. It should be taken into account that a formalized image description depends on the recognition task at hand.

It is necessary to systematize knowledge about invariants and their properties for effective construction of image representations. For this purpose, invariant classification was carried out in [12] by generalizing and normalizing knowledge about invariants and taking into account specific properties of different types of invariants.

The following basic principles are suggested for classification: (a) globality vs. locality, (b) transform under which the invariance is required, (c) noise sensitivity, (d) information redundancy, and (e) reconstruction ability of the invariant.

One reason to use invariants as features for image description in recognition tasks is that invariants contain more information about images as compared to other features. Thus, an invariant set makes it possible to describe images and objects of a scene in a sufficiently compact manner. It should be mentioned, however, that different invariants convey information of different types. Invariants characterizing, for example, image texture or shapes of objects in the image can be distinguished in the set of known invariants. Therefore, the use of different types of invariants depends on the recognition task. For instance, invariant features based on moments of different types [5, 6, 14, 18] are often used in plane identification, recognition of ship images, symbol recognition, scene matching, and 3D object recognition from 2D projections. Some types of invariants are used for texture analysis and recognition. The following invariants can be assigned to the basic types of invariants used in image recognition: (a) texture invariants based on a cooccurrence matrix for a grayscale image [7]; (b) invariants regarding blurring, affine transformations, and changes in intensity and contrast based on moments of different types [5, 6, 14, 18]; (c) integral invariants based on integral transformations of Fourier- and Melline-type [19]; and (d) invariants of geometric objects in the image [15].

5. MATHEMATICAL SETUP OF THE RECOGNITION TASK

Let us return to the problem of selecting the algorithm that correctly solves the image recognition task. In order to prove the theorem on the existence of an algorithm that correctly solves the image recognition task, let us lean toward the similar Zhuravlev theorem on the correctness of algebraic closure of AEC for a pattern recognition task [2]. The results of the theorem cannot be employed immediately in the case when the image is an object for recognition. There are several reasons for this. Firstly, representing an image as a feature vector (as a standard object for recognition) usually results in the loss of a considerable amount of information about the image and, as a consequence, in incorrect classification. Secondly, the existence of equivalence classes substantially differ the image recognition task from the recognition task in the classical formulation.

A method based on image invariance and equivalence was proposed in order to eliminate these differences. Here, an image is described not by an arbitrary feature set but by a set of invariants. As was shown above, this representation is not only compact but also sufficiently informative and immune to image distortions. An image recognition task in terms of equivalent classes can be reduced to the classical setup with the use of the allowable transformation notion. To show this, a standard formulation of the recognition task, as well as different setups of an image recognition task, should be considered in detail.

5.1 Classical Setup of Pattern Recognition Task Z

Let us briefly recall a pattern recognition task in the standard setup formulated by Zhuravlev [2].

Let $Z(I_0, S_1, \ldots, S_q, P_1, \ldots, P_l)$ be a recognition task, where I_0 is allowable initial information, $S_1, \ldots, S_q$ is a set of allowable objects described by feature vectors, $K_1, \ldots, K_l$ is a set of classes, and $P_1, \ldots, P_l$ is a set of predicates on allowable objects, $P_i = P_i(S)$, $i = 1, 2, \ldots, l$. The task Z is to calculate predicate values $P_1, \ldots, P_l$.

Definition 5.1.1. [2] The algorithm is *correct* for the task Z if the following equation is fulfilled:

$$A(I, S_1, \ldots, S_q, P_1, \ldots, P_l) = \|\alpha_{ij}\|_{q \times l}, \qquad (5.1.1)$$

where $\alpha_{ij} = P_j(S_i)$.

5.2 Mathematical Setup of Image Recognition Task Z^1

The task setup consists in the following [13]. Let the set K of objects to be classified be given. It is known that the set K can be represented as a sum of subsets $K_1, \ldots, K_l$ called classes: $K = \bigcup_{j=1}^{l} K_j$. Let also some information $I_0(K_1, \ldots, K_l)$ about the classes be given.

According to the approach described above, an image is a visual representation of a certain essence S (e.g., object, scene) and each scene S has a corresponding set of images that differ in scale, viewpoint, illumination, etc. In other words, these images are equivalent (in terms of the definitions provided above). Let us consider the situation when each image is represented by a vector of invariant features.

The main task is to calculate predicate values $P_j(I)$: "$I \in K_j$"; $j = 1, 2, \ldots, l$ from information about classes $I_0(K_1, \ldots, K_l)$ and from image I description $D(I) = (a_1, a_2, \ldots, a_n)$.

Let A be a recognition algorithm transforming training information $I_0(K_1, \ldots, K_l) = I_0(l)$ and object description $D(I) = (a_1, a_2, \ldots, a_n)$ into an information vector $\{\alpha_j^A\}_{1 \times l}$ composed of elements 0, 1, Δ. The following equation is valid:

$$A(I_0(l), D(I)) = \{\alpha_j^A\}_{1 \times l}.$$

Standard interpretation is accepted for elements α_j^A:

$\alpha_j^A = 1$, the image I belongs to the class K_j;

$\alpha_j^A = 0$, the image I does not belong to the class K_j;

$\alpha_j^A = \Delta$, the algorithm A failed to identify whether the image I belongs to the class K_j; $j = 1, 2, \ldots, l$.

Since real recognition tasks concern not objects but their images, we consider that the entire set of images is somehow divided into equivalence classes. Therewith, we assume that there is a correspondence between image equivalence classes and objects; however, we will drop objects in setting up a recognition task from here on. Taking into account the introduced notion of image equivalence, the image recognition task can be formulated as follows.

$$Z^1\left(\{I_i^{j_i}\}_{i=1,2,\ldots,q}^{j_i=1,2,\ldots,p_i}, \quad \{M_i\}_{i=1,2,\ldots,q}, \quad \{K_t\}_{t=1,2,\ldots,l}, \quad \{P_t^{ij_i}\}_{t=1,2,\ldots,l}^{i=1,2,\ldots,q;\ j_i=1,2,\ldots,p_i}\right)$$

is the image recognition task Z^1, where $\{I_i^{j_i}\}$ are images, $i = 1, 2, \ldots q$, j_i is an image number within the ith equivalence class, p_i is the number of images in the ith equivalence, $j_i = 1, 2, \ldots, p_i$; $M_i = \{I_i^1, I_i^2, \ldots, I_i^{pi}\}$, $i = 1, 2, \ldots, q$ are equivalence classes in the set $\{I_i^{j_i}\}$; $K_1, K_2, \ldots, K_l$ are the classes in the image recognition task; and $P_t^{ij_i}$: "$I_i^{j_i} \in K_t$", $t = 1, 2, \ldots, l$, $i = 1, 2, \ldots, q$, $j_i = 1, 2, \ldots, p_i$ are the predicates. Task Z^1 consists in calculating values of the predicate $P_t^{ij_i}$.

5.3 Mathematical Setup of Image Recognition Task Z^2

The difference between task Z^2 and task Z^1 is that each equivalence class is replaced by a single image, class representative, with the number n_i, $1 \le n_i \le p_i$, where i is the number of the equivalence class. The replacement is made by introducing the notion of allowable transformation.

Definition 5.3.1. An arbitrary transformation f: $\{I_i^j\} \longrightarrow \{I_i^j\}$ is called an *allowable transformation* if $f(I_i^j)$ and I_i^j belong to the same equivalence class for each I_i^j.

$$Z^2\left(\{I_i^{n_i}\}_{i=1,2,\ldots,q}^{1 \le n_i \le p_i}, \{K_t\}_{t=1,2,\ldots,l}, \{P_t^i\}_{t=1,2,\ldots,l}^{i=1,2,\ldots,q}\right)$$ is

an image recognition task Z^2, where $I_i^{n_i}$, $i = 1, 2, \ldots, q$, are images; $I_i^{n_i} \in M_i$; $K_1, K_2, \dots, K_l$ are the classes in the image recognition task; and P_t^i: "$I_i^{n_i} \in K_t$"; $t = 1, 2, \dots, l$, $i = 1, 2, \dots, q$ are the predicates. The task Z^2 consists in calculating values of the predicate P_t^i.

6. COMPLETENESS CONDITIONS FOR THE AEC CLASS IN IMAGE RECOGNITION TASK

The main result of the research concerns the computationally efficient class of recognition algorithms for estimate calculations [13]. These algorithms are based on formalization of the concept of precedents or partial precedents; i.e., the "proximity" between partial descriptions of objects already classified and the object to be classified is analyzed. Let standard descriptions for objects $\{\tilde{S}\}$, $\tilde{S} \in K_j$ and $\{S'\}$, $S' \notin K_j$ be given, as well as a method for finding the proximity of some parts of description $\{I(\tilde{S}), I(S')\}$ and corresponding parts of descriptions $j = 1, 2, \dots, l$, where S is the object of recognition. By calculating proximity estimations for parts of descriptions of $\{I(\tilde{S})\}$ and $\{I(S')\}$, $I(S)$ and $I(S')$, respectively, one can make a generalized estimation of the proximity between S and object sets $\{\tilde{S}\}$ a and $\{S'\}$ (in the simplest case, the generalized estimation equals the sum of proximity estimations between description parts). Then, the overall estimation of the object over the class is formed from the estimation set; this is the value of the function of an object's belonging to a class.

The following theorem on the existence of an algorithm for estimate calculations, solving the recognition task Z correctly, is proved for algebraic closure of AEC.

Theorem 6.1. [2] Let the natural assumptions about the difference in descriptions of classes and objects under recognition be true for feature vectors in the recognition task Z. Then, the algebraic closure of the AEC class is correct for the task Z.

It should be noted once again that the equivalence is not employed in the classical setup of a recognition task, which prevents the immediate application of the proven theorem on existence algorithm to correct image recognition.

The task Z^1 differs from Z in that it employs image equivalence classes in an explicit form. In order to reduce an image recognition task Z^1 to a standard recognition task Z, it is necessary to move from classification of object groups to classification of a single object. The task Z^2 is distinguished from Z^1 by the presence of allowable transformations that do not move an image out of the equivalence class; it is possible to operate with a single object, a representative of the equivalence class, under certain restrictions on allowable transformations.

The following theorem is the direct generalization of Theorem 6.1. for the task Z^2.

Theorem 6.2. Let allowable image transformations $\{f_1, f_2, \dots\}$ form a transitive group. Then, image recognition task Z^1 can be reduced to recognition task Z^2 and algebraic closure of the AEC class is correct for the task Z^2.

This theorem establishes the conditions of existence of the correct algorithm in the image recognition task and proves that such an algorithm can be found in algebraic closure of AEC.

CONCLUSIONS

The problem of selecting a correct algorithm for image recognition tasks was investigated in the paper. Notions of image equivalence and invariance related to this topic were examined. Different ways to define image equivalence were considered, namely, equivalence based on groups of transformations, equivalence directed at the recognition task, and equivalence with respect to a metric. Examples of equivalence classes were constructed for the case of equivalence definition based on groups of transformations. It was shown that equivalence is one if the key notions of image recognition theory.

The connection between image equivalence and invariance was studied. The notion of an invariant both in terms of its practical application in image analysis and processing and in terms of algebraic theory of invariants was analyzed. Different definitions of an invariant were given.

Based on the notion of image equivalence, the standard mathematical setup of the image recognition task was modified, and the image recognition task was formulated in terms of equivalence classes. It was proved that the image recognition task in the classical setup can be modified to a reduced task that has a correct algorithm in the framework of algebraic closure of AEC under certain restrictions on image transformations.

In our further investigations, we plan to study in more detail the image equivalence and relations between image equivalence and invariance. The obtained results will be applied to automated diagnostics of lymphatic system tumors from hematological specimens.

REFERENCES

1. Djodonne, G. Kerrol. Theory of invariants. Old and New. (1971) "Mercury-PRESS" Publishing House, Cherepovets, 2000 (in Russian). Translated from: Dieudonne J., La theorie des invariants au XIX siecle, Seminaire Bourbaki, 1970/71, Berlin, Springer, 1972.
2. Zhuravlev, Yu.I., An Algebraic Approach to Recognition or Classification Problems, *Pattern Recognit. Image Anal.* 1998, vol. 8, No. 1, pp. 59–100.

3. *Matematicheskaya Entsiklopediya* (Mathematical Encyclopedia), Vinogradov, I.M., Ed., vol. 2, Moscow: Sov. Entsiklopediya, 1979.
4. Shafarevich, I.P., *Osnovnye ponyatiya algebty* (Basic Notions of Algebra), Izhevsk: NITS Regular and Chaotic Dynamic, 2001.
5. Flusser, J. and Suk, T., Degraded Image Analysis: An Invariant Approach, *IEEE Trans on Pattern Analysis and Machine Intelligence*, 1998, vol. 20, no. 6, pp. 590–603.
6. Flusser, J. and Suk, T., Blur and Affine Moment Invariants, *Proc. 16th Int. Conf. on Pattern Recognition (ICPR2002)*, Quebec, 2002, vol. 4, pp.339–342.
7. Fu, K.-S. and Mui, J.K., Automated Classification of Nucleated Blood Cells Using a Binary Tree Classifier, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1980 vol. PAMI-2, no. 5, pp. 429–436.
8. Grenander, U., *General Pattern Theory: A Mathematical Study of Regular Structures*, Oxford: Clarendon, 1993.
9. Gurevich, I.B., The Descriptive Framework for an Image Recognition Problem, *Proc. 6th Scandinavian Conf. on Image Analysis*, Oulu, 1989, vol.1 pp. 220–227.
10. Gurevich, I.B., Descriptive Technique for Image Description, Representation, and Recognition, *Pattern Recognit. Image Anal.*, 1991, vol. 1, no. 1, pp. 50–53.
11. Gurevich, I.B. and Smetanin, Y.G., On the Equivalence of Images in Pattern Recognition Problems, *Proc. 12th Scandinavian Conf. on Image Analysis*, Bergen, 2001, pp. 679–685.
12. Gurevich, I.B., Jernova, I.A., and Smetanin, Y.G. A Method of Image Recognition Based on the Fusion of Reduced Invariant Representations: Mathematical Substantiation, *Proc. 16th Int. Conf. on Pattern Recognition (ICPR2002)*, Quebec, 2002, vol.3, pp. 391–394.
13. Gurevich, I.B. and Zhuravlev, Y.I. An Image Algebra Accepting Image Models and Image Transforms, *Proc. 7th Int. Workshop "Vision, Modeling, and Visualization 2002" (VMV2002)*, Erlangen, Greiner, G., Niemann, H., *et al.*, IOS Press, B.V. Amsterdam, Infix, Akademische Verlagsgesellschaft, Aka GMBH, Berlin, 2002, pp. 21–26.
14. Khotanzad, A. and Hong, Y.H. Invariant Image Recognition by Zernike Moments, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1990, vol. 12, no. 5, pp. 489–497.
15. Mundy, J.L. and Zisserman, A., *Towards a New Framework for Vision: Geometric Invariance in Computer Vision*, Mundy, J. and Zisserman, A., Eds., 1992, pp. 1–39.
16. Ritter, G.X. and Wilson, J.N. *Handbook of Computer Vision Algorithms in Image Algebra*, CRC Press Inc., 2001, 2nd ed.
17. Serra, J. *Image Analysis and Mathematical Morphology*, London: Academic, 1982.
18. The, C.-H. and Chin Roland, T., On Image Analysis by the Methods of Moments, IEEE Trans. on Pattern Analysis and Machine Intelligence, 1988, vol. 10, no. 4, pp. 496–512.
19. Wood, J., Invariant Pattern Recognition: A Review, *Pattern Recognition*, 1996, vol. 29, no. 1, pp. 1–17.
20. Zhuravlev, Yu.I., Algebraic Methods in Recognition and Classification Problems, *Pattern Recognit. Image Anal.*, 1991, vol. 1, no. 1, pp. 1–4.
21. Zhuravlev, Yu.I. and Gurevitch, I.B., Pattern Recognition and Image Recognition, *Pattern Recognit. Image Anal.*, 1991, vol. 1, no. 2, pp. 149–181.

Igor B. Gurevich. (See p. 568).

Irina A. Jernova. Born 1981. Graduated from the Faculty of Computational Mathematics and Cybernetics of Lomonosov State University, Moscow, in 2003. Since 2001, works at the Scientific Council on Cybernetics, Russian Academy of Sciences. First-year postgraduate student of the Scientific Council on Cybernetics. Author and coauthor of six papers. Scientific interests include image algebras and image processing, analysis, and recognition. Member of the Russian Association of IAPR.