Physics – Uspekhi **47** (10) 1017 – 1032 (2004)

METHODOLOGICAL NOTES PACS numbers: 02.10. – v, 42.30.Tz

Physical interpretation of the elements of image algebra

V A Utrobin

DOI: 10.1070/PU2004v047n10ABEH001638

Contents

Abstract. The physical and mathematical aspects of image algebra transformations, their elements and groups are reviewed, allowing some processes involved in innate visual perception mechanisms to be justified.

1. Introduction

As observers, we draw the main body of information on the environment through the system of visual perception. which is not considered to be a subject of physical studies because the optical properties of the eye were already determined by Helmholtz and the further processing of visual information is a subject of 'nonphysical' research. Such a formal approach does not provide insight into the general problem of mapping the outer physical world with its laws onto the 'inner world' of the perception system, without which we (observers) not only fail to understand what we see but would also be unable to walk or orient ourselves in the environment. Therefore, the system of visual perception, understood on the whole, should necessarily be adequate to the environment, that is, obey the laws of physics of space and time with its conservation (symmetry) laws. Furthermore, visual perception as a whole should be sufficient for judging adequately what we see and surviving in the environment. This means that the system of visual perception (among the other perception systems) should not be less sophisticated than the environment in which we are immersed, suggesting two conditions that are imposed on the system of visual perception and determining its necessary and sufficient properties.

V A Utrobin Nizhny Novgorod State Technical University
ul. Minina 24, 603600 Nizhniĭ Novgorod, Russian Federation
Tel. (7-8312) 36-82 28
Fax (7-8312) 36-23 11
E-mail: utrobin-va@yandex.ru

Received 11 April 2003, revised 15 February 2004
Uspekhi Fizicheskikh Nauk **174** (10) 1089 – 1104 (2004)
Translated by M V Tsaplina; edited by A M Semikhatov

First, the image is a subject of research corresponding to physical objects with their objective laws and space – time properties.

Second, the system of visual perception, as an identification system, should not yield to the 'physical complexity' of the object under study.

The innateness of visual perception mechanisms (requiring no training), being a product of nature, is responsible for the generality of the laws of transformation in the visual perception system [1]. Therefore, the known physical laws of nature are laws obeyed by the organization (from the standpoint of information transformations) of the system of visual perception. Knowing it, one can expect to obtain new regularities and information technologies of image processing. Such technologies are needed because the problems to be solved become increasingly complicated and require analysis of not so much one-dimensional signals as various scalar and vector fields. Such fields can readily be represented by images in a given two-dimensional domain of definition (the field of vision). However, the majority of image processing problems are ill-posed. For example, the identification of an object boundary using spatial filtration is problematic because this operation amplifies any high-frequency noise and thus obstructs correct decision making concerning the boundary.

For linear problems, regularization methods exist (e.g., regularization according to Tikhonov [2]), while no general approaches to the solution of nonlinear problems are known [3]. However, the method of image shape analysis using the variable resolution technique [4] is well known. A striking historical example is the Saturn enigma, which was solved by Huygens with the use of a telescope that had sufficient magnification and a higher resolution than Galileo's telescope. Galileo took the rings of Saturn for two lateral appendages 'bearing' Saturn, and when the rings were seen as a thin line, he failed to notice them and never mentioned them again [5]. This method in physics is a description of media and processes as a rough system with a consequent more detailed analysis.

We propose a new image processing technique that excludes regularization and applies a 'deep' integral transformation (Q-transformation), mapping the image as a whole

into an absolutely smooth manifold with a subsequent 'multidimensional' spatial differentiation that reveals differential structures of image description and their connections on the manifold. The details of the image are specified (if necessary) on a variable-resolution pyramid [6]. Figure 1 gives a schematic drawing of the procedure using the example of the analysis of Saturn rings.

We consider here the physical and mathematical operations applied within this method.

For the majority of problems solved in the course of image analysis, one cannot use the classification models applied in the pattern recognition theory and in machine perception at all because they do not allow taking the *a priori* structure of the image as an object of study into account. It is the structure (shape) that allows image description. For instance, in an analysis of a photograph with tracks of particles in a bubble chamber, not only a classification but also a description of the picture is needed [7]. Such a description should contain information on separate parts of the picture and relations between them. Similar procedures should be applied in the analysis of symbols, patterns, and forms for their identification and possible classification in the sense of ordering within a class. In perception theory, such descriptions are called a composition analysis of a scene (picture). The analysis of a visible scene, which is necessary for automatic robot-type machines with their technical sight systems, is also of importance in solving problems, e.g., involving the analysis of surfaces in various industries and in research work, in processing aerial photography data of the earth and other planet surfaces.

This problem is also topical for biomedical examination of medications and chemicals using a microscope or other technical means (X-ray, tomograph, and other devices). However, this problem appear to be practically unsolvable because the computational and time resources are rather limited. To describe structures, attempts were made to borrow concepts from the theory of formal languages with a construction of linguistic models [8]. The problem of model description turned out to be much wider than it had seemed before. It is, in a sense, analogous to the problem existing in the physics of nonlinear media, where the list of open systems that are capable of self-organization under certain conditions is increasingly extensive (Taylor vortices, Benard and Marangoni cells in liquid media, *T*-layers, *E* and *H*-fibers in low-temperature plasma, etc.) [9, 10]. The model description of such systems necessarily requires elaboration of relatively simple methods of form analysis, a set of simple forms that allow the description of complex structures being of importance. An example of such an approach is the use of eigenfunctions of a nonlinear medium, whose total set is guaranteed by localization of an initially continuous process (e.g., combustion) in a finite domain of definition. The problem lies in constructing multidimensional eigenfunctions of a nonlinear medium for controlling processes in such media (especially in real time) and for creating laws of unification of these media into complex structures [10].

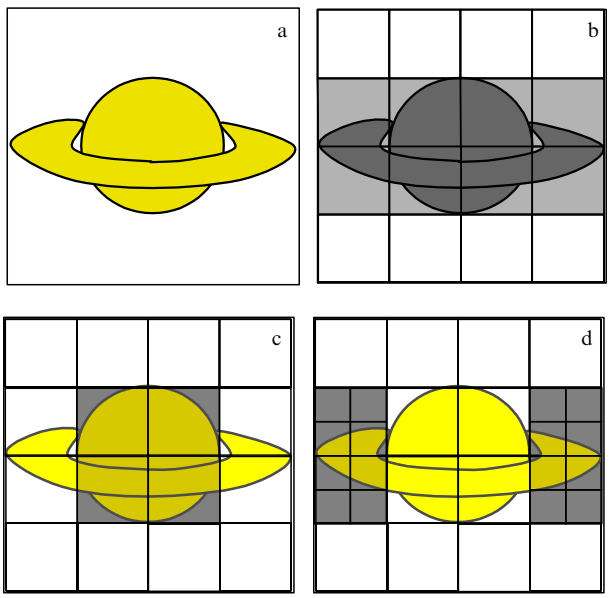

Figure 1. Example of the analysis of Saturn's shape: (a) Saturn and its rings in the field of vision; (b) the domain of definition of the object; (c) the subdomain of planet segmentation; (d) subdomains of specification of the ring geometry with high resolution on the planigon.

The proposed method of forming structural descriptions on the basis of the image algebra involving a complete set of simple forms (operators), which play the role of innate etalons, and the rules for their unification allow a reliable (in the sense of system roughness) and relatively accurate (at a given level of resolution) solution to the above-mentioned and similar problems. Such a 'roughly accurate' representation is a reflection of a natural desire to simplify the image, reducing it to a relatively small, but maximally informative number of its parts with their consequent unification into a meaningful formation [7].

The methodical results in this paper also include the proof of the completeness of image algebra elements from the standpoint of symmetry (Lie groups and conservation laws), substantiation of the variational principles of the analysis of visually observed media with consequent decision-making in the etalon space, and approaches to the reconstruction of the image shape from a single two-dimensional image (such a perception for the human system of visual perception is called a mono – stereo perception). The approaches to this problem (namely, the representation of continuous functions of three variables by superposition of continuous functions of two variables) were theoretically considered by A N Kolmogorov and V I Arnold as far back as the 1950s [11].

2. Statement of the problem

The active perception theory with the image algebra as the mathematical basis is devoted to the problems of *a priori* uncertainty of the image as an object of study [6, 12 – 16]. We single out the points of the technique and algebra under investigation that are necessary for our further presentation. To begin, we define the subject of our study.

Let an observer be immersed in the environment and not distort this environment. Then, the image M is a function of the observer (appearing on the retina of the eyeball) satisfying the following natural restrictions. The image is a set M, each element of which, at a fixed instant of time t, is a nonnegative real function of real arguments (observability property),

$$M_t = \begin{cases} \mu(x,y) & \text{if} \quad (x,y) \in \overline{G}\,, \\ 0 & \text{if} \quad (x,y) \notin \overline{G}\,, \end{cases}$$

defined on a finite set of points of a closed two-dimensional domain $\overline{G}$ of the Euclidean space, summable and square-integrable on this set (measurability property), and having the properties of ordering, structuredness, and discreteness in space and time.

Therefore, because the domain of definition is restricted, the image is a two-dimensional signal which at a fixed instant of time is represented by the function $\mu(x, y)$ or by a set of points (pixels), each point being lexicographically ordered and 'labeled' by the value of this function.

According to the adopted definition, we have two concepts of image:

A spatial, time-independent ('spacelike', static) image is a function $\mu(x, y)$ defined at the instant $t \in T_N \subset T$, where T_N is the observation interval from the set T;

A space – time ('timelike', dynamic) image is a function

$$M_{\mathrm{d}} = f(t, M_t)\,, \qquad 0 \leqslant t \leqslant T_N\,.$$

The properties of measurability and observability of M_t (henceforth simply M) imply that the set M is finite and has the cardinality

$$k_{\max} = P^{N \times M}\,,$$

where $N \times M$ is the size of the domain G (the spatial discreteness property is used) and P is the number of brightness gradations.

Because the object under study is finite and belongs to the set R^+ of positive real numbers, its point set converges in the mean square, that is, both the entire image and any of its subdomains of definition $G_i \subseteq G$ allow the projection transformation (Fig. 2)

$$m(G_i) = \iint_{(x,y)\in G_i} \mu(x, y)\,\mathrm{d}x\,\mathrm{d}y\,. \tag{1}$$

Transformation (1) is specified by the Q-transformation in view of the following properties:

1. It can be applied to any object of study of any dimensionality and in any frequency band satisfying relatively weak restrictions imposed by the definition on the image.

2. It realizes a mapping into a real, absolutely smooth space (a C^∞-manifold) and allows two model representations (function (1) is an averaged function):

a) because of harmonicity in $\overline{G}$ (or $\overline{G}_i \subseteq \overline{G}$), we have

$$\varphi(A) = m(B) \quad \forall A \in \overline{G}\,;$$

b) because of harmonicity in G (or $G_i \subseteq G$), we have

$$\varphi(A) = m(B) \lg r\,,$$

where $\varphi(A)$ is the potential created at an arbitrary point A by the 'charge' (or 'mass') $m(B)$ defined by transformation (1) and r is the distance between the points A and B.

3. It is fundamental (because the Cauchy sequence existing in M is fundamental), sufficient, and realizable. Because (1) involves integration ($\int$), there exists the inverse transformation, that of exterior differentiation (d). Together, they form a composition (which we call the U-transformation $U = \mathrm{d} \circ \int$, which removes the uncertainty of the object under study, i.e., of the image M in the field of vision G.

Because the set M is finite, its representation in the space $N \times N \times P$ (we henceforth assume the domain G to have a square shape for convenience) defines this space by the general aggregate of simple events. Therefore, any image from the ('timelike') set M is equiprobable, the same as any point in the image M is equiprobable, and the relation of these points (or images) is an equivalence relation. Hence, any filter F, as a covering of M (i.e., a filter realizing the operation d after the Q-transformation), is a mask constructed by divisions (dichotomies) of the domain G into smoothly glued subdomains G_i.

From the standpoint of Riemannian geometry, such a filter (and a finite set of such filters, because G is finite) belongs to the tangent space A^n at a point A_0 (Fig. 3). In the ε-neighborhood of the tangency point, each vector from A^n corresponds to a differential operator

$$\nabla_i = \frac{\partial}{\partial x^i}$$

in the set of directions x^i. Given this, the vector V is uniquely defined on the difference

$$\Delta m(G) = m(G_j) - m(G_i)\,, \tag{2}$$

where $G = G_i \cup G_j$ is a dichotomy of the domain G into two nonintersecting subdomains for which the $m(G_i)$ values are determined from (1) (see Fig. 2).

Relation (2) makes it possible to reveal the position of the mathematical point on the straight-line segment of the direction x^k common for the subdomains (because the tangent space A^n is Euclidean, it follows that $x^k = x_k$ and because it is applicable to objects from E^3, it follows that

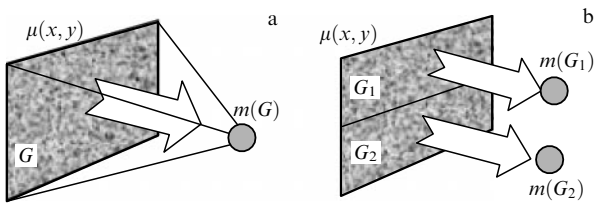

Figure 2. Realization of a G-transformation over the entire image area (a) and over two of its subdomains (b).

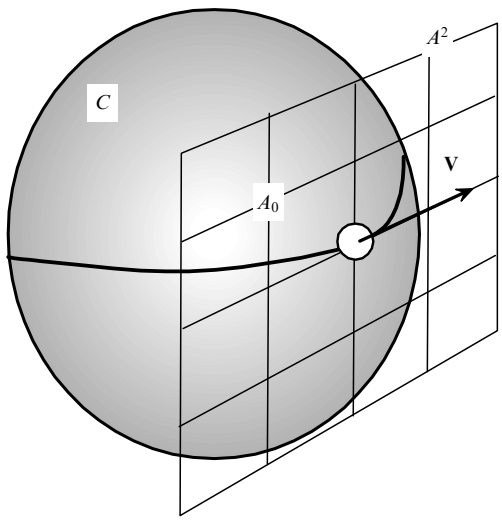

Figure 3. Planigon A^2 relative to the manifold C.

$n = 2$). To reveal all the binary relations between the pair of points in the domain of definition $G \subset E^2$, it is necessary to use (2) to investigate 16 versions of the dichotomy of G, including the zeroth version, i.e., the unit element. As a result, we have

— a canonical basis of the vector space representing M:

$$\mathbf{e}_1 = (1, 0, \ldots, 0)^{\mathrm{T}}, \quad \mathbf{e}_2 = (0, 1, \ldots, 0)^{\mathrm{T}},$$

$$\ldots, \mathbf{e}_{15} = (0, 0, \ldots, 1)^{\mathrm{T}},$$

where $()^{\mathrm{T}}$ is the transposition; and

— a set of filters $\{F_i\}$ realizing $\partial/\partial x^i$ in the 15 directions x^i upon obtaining a 15-dimensional vector $\boldsymbol{\mu} = \mu_i \mathbf{e}_i$, defined constructively on the two-dimensional Walsh function of the Harmuth system,[1] and ordered on the two-dimensional grid $V(x, y)$.

First, being binary, the Walsh functions simplify the analysis of relations on the set $\{m(G_i)\}$ ordered on the matrix $[m_{ij}]_{4\times 4}$ because the simplest operations, i.e., addition (1) and subtraction (2), are used. Therefore, the U-transformation has the minimum possible computational complexity as distinct from the standard transformations, which require a convolution and an arithmetical multiplication at the level of weight coefficients. (The scheme of information transformations according to (1) and (2) is presented in Fig. 5 and can be used for any subdomain of the image. For example, in Fig. 6, the number of levels of the variable resolution pyramid is determined by the required precision of the solution, and the version of choosing a necessary subregion is determined by the perception strategy; see Fig. 1.)

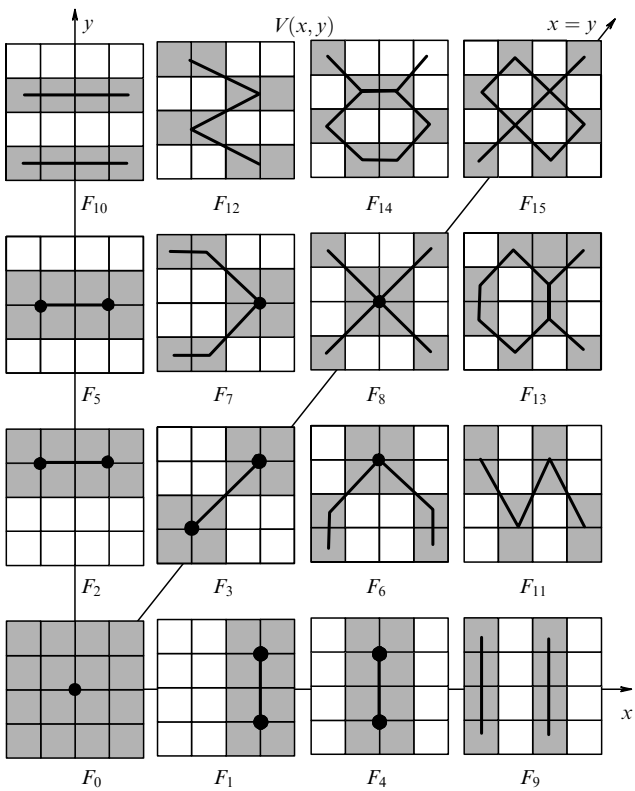

Figure 4. Filters $\{F_i\}$ and the equivalent operators $\{V_i\}$ on the grid $V(x, y)$.

[1] In Figure 4, the dark region corresponds to the weight factor $+1$ and the light region corresponds to the factor -1.

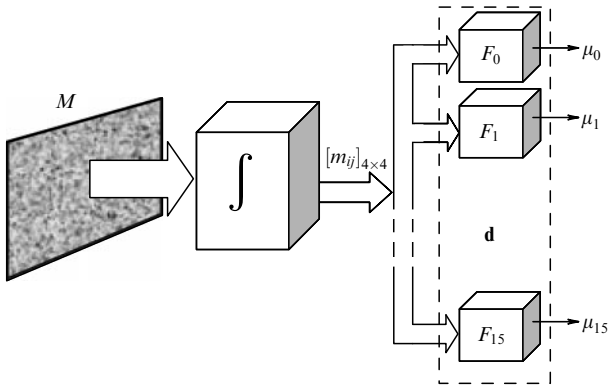

Figure 5. Scheme of information transformations.

Second, transformations (1) and (2) are operations with integers and hence the problem of error accumulation due to rounding off disappears.

Third, transformation (1) maps an image as a function into a real (rather than imaginary) space, where all the other operations of the analysis are realized.

Fourth, beginning our analysis from the upper levels of the pyramid (see Fig. 6) (the resolution of the lower levels is low), we take only the low-frequency image components into account, which are objects (e.g., buildings). In this case, the details are small and the disturbances seem to vanish.

The filters (their numeration is conditional) are constructed on a square cellular space (see Fig. 4), which we define by a planigon centered at the point A_0. If we take a rectilinear Cartesian coordinate system (x, y) in the domain of definition G, we can show that the set of filters corresponds to the set of transformations that are coefficients in the Taylor series (the functions of sensitivity of the series in the solution

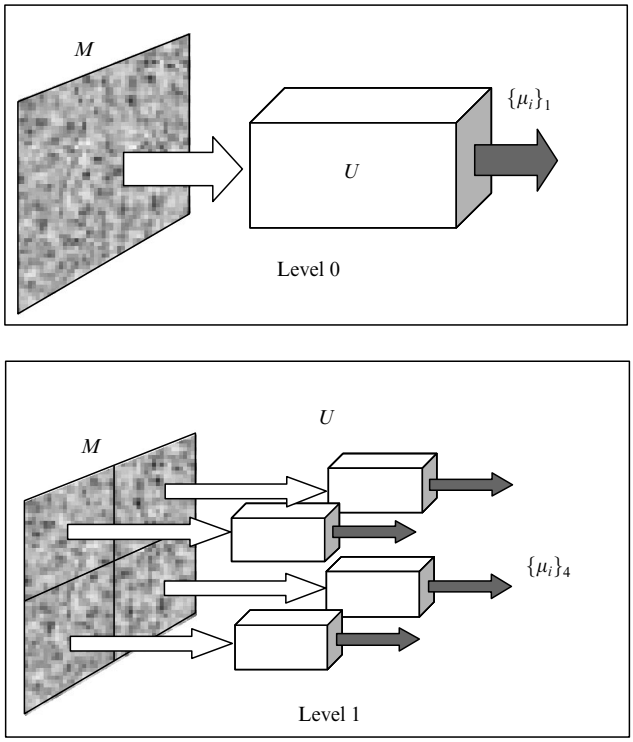

Figure 6. Variable-resolution pyramid.

of the active identification problem):

$$(F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8, F_9, F_{10}, F_{11}, F_{12}, F_{13}, F_{14}, F_{15}) \sim \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial^2}{\partial x\,\partial y}, \frac{\partial^2}{\partial x^2}, \frac{\partial^2}{\partial y^2}, \frac{\partial^3}{\partial x^2\,\partial y}, \frac{\partial^3}{\partial x\,\partial y^2}, \frac{\partial^4}{\partial x^2\,\partial y^2}, \frac{\partial^3}{\partial x^3}, \frac{\partial^3}{\partial y^3}, \frac{\partial^4}{\partial x^3\,\partial y}, \frac{\partial^4}{\partial x\,\partial y^3}, \frac{\partial^5}{\partial x^3 \partial y^2}, \frac{\partial^5}{\partial x^2\,\partial y^3}, \frac{\partial^6}{\partial x^3\,\partial y^3} \right). \quad (3)$$

If we redefine the filter homogeneity subregions, i.e., preserve their construction but assume the weights $+1 \to 1$, $-1 \to 0$, we obtain a set of binary operators $\{V_i\}$ for which the set-theoretic operations of unification (summation) and intersection (multiplication) are allowed. Thus, we arrive at the image algebra $A_V = \langle \{V_i\}\colon +, \times \rangle$ (Fig. 7).

If the set $\{F_i\}$ is applicable at the stage of decomposition of the image M and its analysis, then $\{V_i\}$ in the algebra A_V is applicable at the stage of synthesis and pattern imaging. One can show that the following algebra groups exist in the algebra A_V:

— complete (algebraic) groups P_{ni} formed on triples of operators (V_i, V_j, V_k) for which the relations $V_i + V_j + V_k \equiv e_1$ (the unit) hold; $V_i V_j V_k$ is the image (on the multiplication operation) on the planigon and the description of the group P_{ni} (Fig. 8);

— closed (algebraic) groups P_{si} generated by quadruples of operators (V_i, V_j, V_p, V_m), where

$$(V_i, V_j, V_k) \in P_{ni}\,, \quad (V_p, V_m, V_k) \in P_{nj}\,,$$

with the description $P_{si} = V_i V_j + V_p \overline{V}_m$ (where the required number of operator inversions is odd) and the unit $V_i + V_j + V_p + \overline{V}_m \equiv e_1$ (Fig. 9).

The sets $\{P_{ni}\}$ and $\{P_{si}\}$ are finite and have the respective cardinal numbers 35 and 105 (the 36th element of the set of complete groups is the operator V_0).

We consider an example of the analysis of a symbol, for instance, the letter A located in the field of vision. Suppose we have several versions of the letter image, including a 'constellation' of points (Fig. 10a). Let the m_{ij} values be binary and let them correspond to Fig. 10b for convenience. Then the values of the decomposition vector components have the form in Fig. 10c. When mapped onto the operator grid (Fig. 10d), they give graphs of the complete and closed groups, where the size of the circles corresponds to the value of the decomposition vector component, and the half-tint corresponds to the operator inversion. These groups provide a description of the symbol transform or, more precisely, its image $O(M) = P_n + P_s$. The operator $\overline{V}_{14}$, which has the maximum 'share holding' through the value of its component, is the established standard of the symbol in the above ways of writing it.

The aim of this paper is to reveal the possibilities of the elements $\{F_i\}$, $\{V_i\}$ and the groups $\{P_{ni}\}$, $\{P_{si}\}$ of the image algebra from the standpoint of the well-elaborated physical

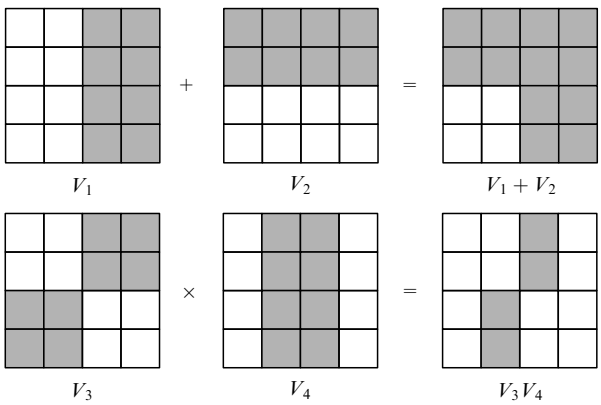

Figure 7. Examples of operator interactions.

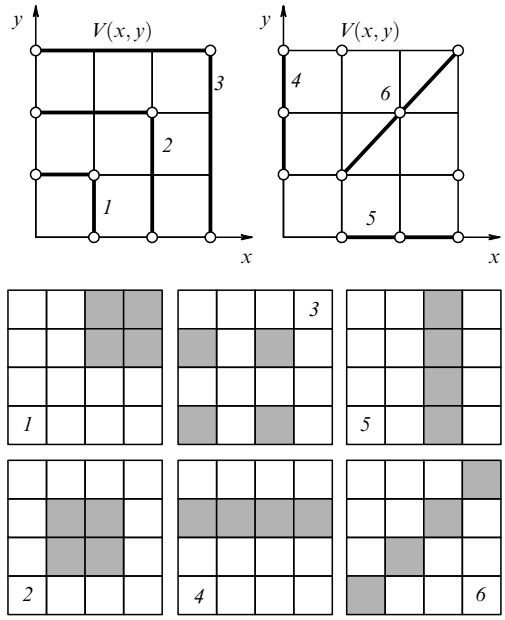

Figure 8. Examples of complete groups represented by graphs on the operator grid, and one of their images.

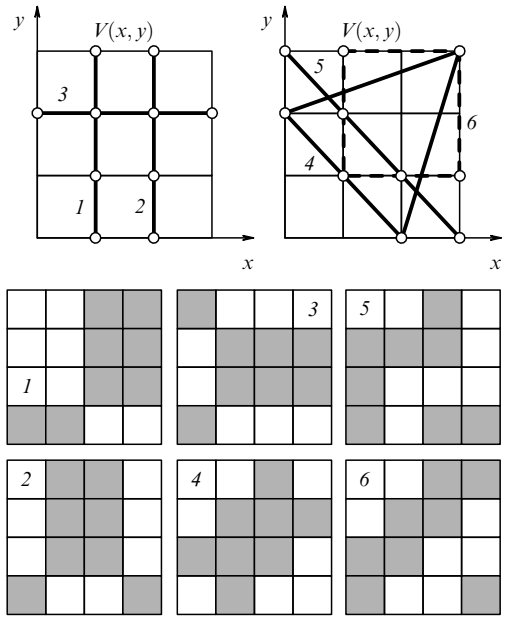

Figure 9. Examples of closed groups represented by graphs on the operator grid, and one of their images.

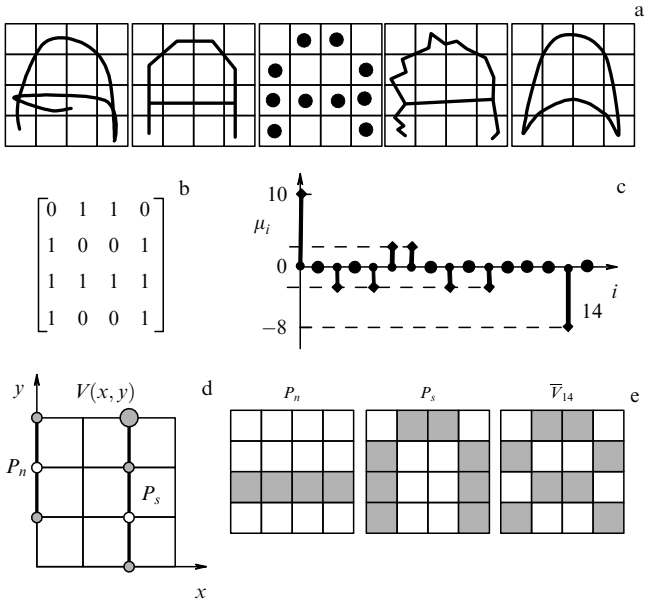

Figure 10. Example of formation of symbol description.

and mathematical apparatus for solving problems of structure analysis.

3. Field representations on a planigon

Let $\mu(i, j) \subset R^+$ be the value of a bounded function $\mu(A)$ at the point A of a finite domain of definition G in E^2. Then, the domain function

$$F(G) = \sum_G \mu(A)$$

exists, with the domain function density at a point given by

$$\varphi(A) = \lim_{G \to A} \frac{F(G)}{\omega(G)} ,$$

where $\omega(G)$ is the area of G.

Let the domain G be divided into two nonintersecting subdomains: $G = G_1 \cup G_2$. Then, in view of the additivity of $\{\mu(A)\}$, the additivity property

$$F(G) = F(G_1) + F(G_2)$$

holds for the domain function. Therefore, both the entire domain G and its nonintersecting subdomains G_i can be assigned domain functions. Because the relations between the function and the density are independent of the dimension of the domain, they hold for any spatial representation E^n. The existence of density in any $G_i \subseteq G$ makes it possible to define the domain function by an integral characteristic (the measure) of this domain,

$$F(G) = \int_G \varphi(A)\, \mathrm{d}\omega ,$$

which allows the interpretation of the domain 'mass' (for the image, it is the visual mass [6]). This implies that after transformation (1) and the representation on the mass matrix $[m_{ij}]_{4\times4}$, the pre-image becomes the set of domain functions lexicographically ordered in the domain of definition. Such domains (spots) are objects in the field of vision, maximally stable against disturbances (provided that the disturbances are smaller in size than the object in the domain of definition, otherwise they themselves become objects) and containing no elements of minor importance. The regularization problem is thus solved. The distinguished spots with 'labels' given by the visual mass values are structure elements of the images; the relation among them is revealed as follows.

Let each division be a revelation of the relation between domain functions that have the physical meaning of mass. Then these masses can be represented by their centers A_0, and therefore their relation is determined by the Euclidean line l of the direction $x^i = x_i$ that links these centers and passes through $A_0 \in G$. We can therefore use the Newton – Leibniz function to obtain

$$\int_{A_0 \in G_1}^{A_0 \in G_2} \frac{\partial \varphi(A)}{\partial x^i}\, \mathrm{d}x_i = F(G_2) - F(G_1) = \int_l \frac{\partial \varphi}{\partial l}\, \mathrm{d}s ,$$

where $\mathrm{d}s$ is an element of the line l. Because the basis vector $\mathbf{e}_i$ coincides with the direction of the unit vector $\mathbf{t}$ tangent to the Euclidean line l, the relation

$$\frac{\partial \varphi}{\partial x^i}\, \mathbf{e}_i = \mathrm{grad}_{x^i} \varphi \tag{4}$$

holds at the planigon center $A_0 \in G$, where φ is a density and a smooth function on $\{x^i\}$ and $\mathbf{e}_i$ is the basis vector of the direction x^i.

The set of divisions of the domain G corresponds to the set $\{F_i\}$, where any filter realizes a spatial differentiation operation. For this reason, representation (4) holds for any F_i. If the function $\varphi(x_1, \ldots, x_{15})$ is defined on the C^∞ manifold in the ε-neighborhood of the point A_0 (which is the center of the coordinate system and of the planigon), the set $\{F_i\}$ on transformations (4) is the first total differential $\mathrm{d}\varphi = (\partial\varphi/\partial x^i)\, \mathrm{d}x_i$ at the point A_0 or the 15-dimensional gradient vector at this point,

$$\nabla\varphi = \mathrm{grad}\, \varphi = \frac{\partial \varphi}{\partial x^i}\, \mathbf{e}_i ,$$

where $\{\mathbf{e}_i\}$ is a basis of the space E^{15} and ∇ is a linear operator.

Therefore, using the set $\{F_i\}$, we reveal gradient components on the matrix of visual masses obtained in transformation (1) and reconstruct the domain function by its density. As an example, we consider the reconstruction of an arbitrary digitized signal (Fig. 11). Such a signal can be regarded as the profile of a surface (a cut, a layer along the coordinate $z = \mathrm{const}$ for imaging, where z is the axis of gray-level imaging). In accordance with the technology presented, we divide the one-dimensional domain G of signal observation into four intervals (Fig. 11 does not keep to scale). For each interval, we find the visual mass m_i from (1) and the relation between the elements of the set $\{m_i\}$ from (2). As a result, we obtain versions of the 'binary' description of the signal (relations between two elements only are considered). If this signal is a cut of a massif, we have a description of such a massif as a whole (i.e., at a given resolution level). In the set $\{m_i\}$, let the conditions $\mu_1 > 0$, $\mu_4 > 0$, and $\mu_9 > 0$ be satisfied; then the massif is described by the triple of operators V_1, V_4, and V_9, which form a complete group with the description $V_1 V_4 V_9$ segmenting the region of extremum. If we use the averaged function model, the reconstructed signal itself (at the given level of resolution) is represented by

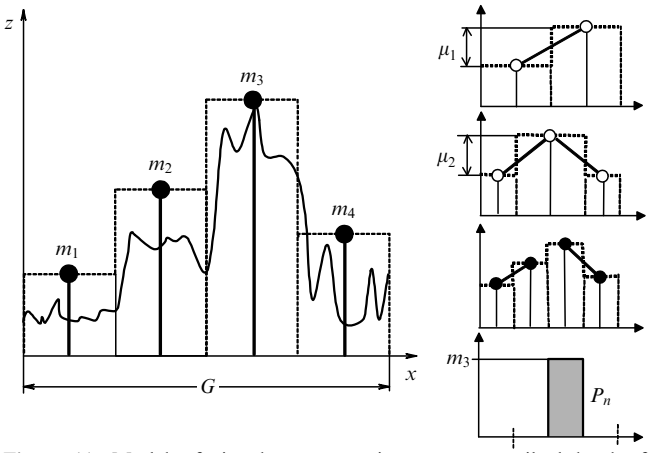

Figure 11. Model of signal reconstruction at a prescribed level of resolution.

four half-tint steps of height m_i ($i = 1, 2, 3, 4$); with allowance for the gradient representation, it is an approximation by first-order lines. For fast recognition (e.g., of relief features) such a rough description is often sufficient. We note that the height m_i can be scaled or assigned the value of m_3 if the height of the vertex is known *a priori*. It is of importance that the relation between the structure elements m_i is preserved (this is exactly how we see spatial relations between objects).

It should be noted that the functional ability of filters (more precisely, of transformations realized by them) in finding variations in the field of vision is very high, which can be confirmed by a simple example. Let the image in the field of vision be uniform (e.g., during night-time observations). Let one pixel of the image 'disappear' (e.g., the light of a cigarette appeared somewhere). Then the filters 'work' to establish the fact that there have been changes in the image and localize it.

The existence of a 15-dimensional gradient vector uniquely defines the vector field $\mathbf{R} = \mathbf{R}(A_0)$ by the smooth field of the potential $\varphi(A_0)$ for which (according to the Ostrogradskii formula), in a closed domain G with the boundary $\gamma(G)$, the function of the domain

$$F(G) = \int_{\gamma(G)} \mathbf{Rn}\,\mathrm{d}\sigma$$

can be set in correspondence with its density, i.e.,

$$\operatorname{div}\mathbf{R} = \operatorname{divgrad}\varphi = \frac{\partial^2\varphi}{\partial x_i^2}\,,$$

where $\mathbf{n}$ is the unit vector of the outer normal to $\gamma(G)$ and $\mathrm{d}\sigma$ is a surface element. The existence of the divergence implies the presence of a vector field flux from sources (drains) existing in G. Any operator is a step function varying from 0 to 1, i.e., the domain of definition for a particular operator narrows to a subdomain for which the coefficient 1 is specified. This means that the entire mass (or the entire charge) is localized in this subdomain with a corresponding density.

Thus, having identical constructive organization, the filters and operators are different in realizable transformations. If the filters from $\{F_i\}$ work to find the gradient in accordance with the maximum-sensitivity direction from $\{x^i\}$, the operators from $\{V_i\}$, which help in estimating the source (drain) density, are endowed with the property of invariance and conservation of changes in a signal (within the limits of the covering). Therefore, $\{F_i\}$ is a set of sensitivity elements for control in finding $\mathbf{R} = \operatorname{grad}\varphi$, and the elements from $\{V_i\}$ in the composition of the functional

$$\int_G \operatorname{div}\mathbf{R}\,\mathrm{d}\omega = F(G)$$

are invariant under rotation and parallel translation in the limits of their coverings.

In the estimation of an object of imaging, the invariance of operators under certain transformations is determined by their measures (quality functions). Such measures are etalons (innate etalons in the visual perception system). Therefore, the decision making is realized in the etalon space by the following rule: the object belongs to the etalon if a minimum of $(\mu_{i\,\mathrm{et}} - \mu_{i\,\mathrm{im}})$ with respect to i exists between the etalon and the object; the rule is analogous for groups and descriptions in the algebra A_V. Moreover, one can show (see [6]) that if $\{V_i\}$ are model representations of simple neurons existent in the cortex of the lens, then $\{P_{ni}\}$ are complex and $\{P_{si}\}$ are supercomplex neurons.

We now consider examples. Suppose one points a video camera at an object, e.g., a post. When the camera turns automatically in the horizontal plane, the appearance (or disappearance) of the object in the field of vision is fixed by actuation of the filter F_1 (or $\overline{F}_1$), and the fact that the object is at the camera's crosshairs — by actuation of the filter F_4 ($\mu_4 = \mu_0$ if no obstructive factors exist in the field of vision). If two cameras are pointed at an object, the algorithms of their pointing are identical, and because the basic operations are particularly simple, the computational complexity of the algorithm is minimal compared to the known methods. Because the minimum-resolution level is used in solving the problem of control (but with the level that is sufficient to distinguish a necessary object), it follows that possible disturbances do not affect the result of control. This is how one solves the problem of pointing two cameras at an object in solving the problem of stereoscopic perception.

Let a square be observed in the field of vision (Fig. 12). It is described by the image of the complete group

$$P_n = \overline{V}_4 + \overline{V}_5 + \overline{V}_8$$

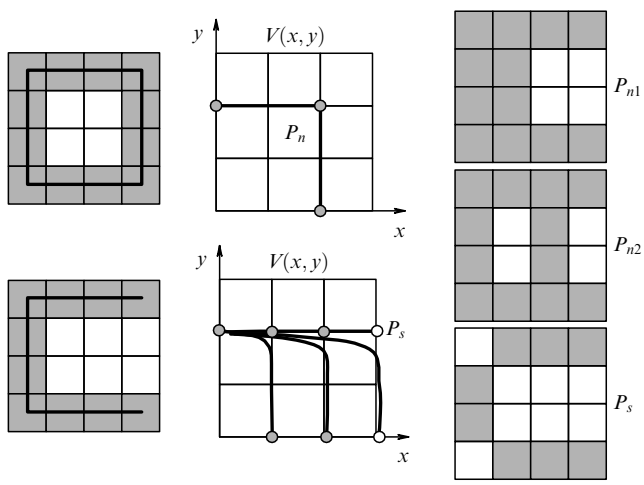

Figure 12. Example of the perception of shape with its possible deviations.

with the visual mass $\mu_0 = 12$ (arbitrary binary units). Let the square 'disappear' at the next instant of time; more precisely, let one of its sides be erased or screened. In this case, two cells in the planigon become white. As a result, the total visual mass of the image becomes equal to 10, and two new groups, P_{n1} and P_{n2}, and one closed group

$$P_s = \overline{V}_5 \overline{V}_7 + \overline{V}_8 V_{11}$$

appear in the description instead of the group P_n. Because the observer knew *a priori* that there was a square in the field of vision, he/she easily makes a decision about partial object screening. If the observer knows *a priori* that the field of vision involves a certain geometrical object (shape) with two elements screened, then from the observation of an object with mass 10, he/she assumes (foretells) three possible versions corresponding to the images of three complete groups P_n, P_{n1}, and P_{n2}. Furthermore, the image of the closed group P_s unites all these foretold shapes.

4. Functional spaces in image algebra

Let a planigon be a plane tangent to a smooth manifold (see Fig. 3), with a Cartesian coordinate system defined on the plane with respect to the absolute (external) coordinate system of the observer. Such a representation naturally defines the planigon by a constrained system. Because the set $\{m_{ij}\}$ is a result of mapping (1) of an actually observed process from E^3 into a manifold and the filters $\{F_i\}$ reveal the mass distribution with respect to the conditions

$$\frac{\partial}{\partial x^i} = \{0; > 0; < 0\}\,,$$

it follows that:

- the manifold represented on the planigon is a continuous medium;
- each F_i corresponds to the state of such a medium relative to the initial state corresponding to the 'unit' filter F_0; and
- each state of the medium F_i is related to another state F_j (we mean the images of the filters reflecting the state) through parallel transport and shear deformation of the masses of one planigon subdomain into another.

Therefore, each pair (F_i, F_j) of states of the planigon medium, related as $\Delta t = t_2 - t_1$, corresponds to its own version of the motion of the continuous medium. With respect to the external coordinate system, we have: 1) the movements from the point A_0 are possible movements in the planigon plane; 2) the movements from an arbitrary point A to B are kinematically possible movements; and 3) the absolute motion is revealed through the deviation from the kinematically possible motion represented on the set $\{V_i\}$ in the image algebra on the planigon (such an approach to the analysis of motion is a variational solution of the mechanical problem).

Thus, the analysis of a process on the planigon at a fixed time moment and within time intervals related via Δt is a construction of possible movements on images of the operators $\{V_i\}$ in the algebra A_V.

For instance, if transitions of the type

$$V_3 \to \overline{V}_3 \to V_3 \to \overline{V}_3 \to \ldots$$

are observed, we have rotation in the field of vision about the coordinate axis z orthogonal to the planigon plane and joining two coordinate systems. If it is known *a priori* that the observed object is a body (envelope), then, for example, the transitions

$$V_1 \to \overline{V}_1 \to V_1 \to \overline{V}_1 \to \ldots$$

represent rotation about the y axis of the planigon (this is also the case with the other operator images).

On the other hand, for the image of the operator V_1 regarded as a rigid body, rotation about the x axis (the sensitivity image of the x-filter) determines it by an object invariant under rotation; this holds for all operators because the respective filters are antisymmetric, that is, the condition

$$\frac{\partial \varphi(x)}{\partial x} = 0\,,$$

where

$$\varphi(x) \equiv \varphi(x^1, \ldots, x^{15})\,,$$

implies that the function $\varphi(x)$ is invariant with respect to the variable $x^i = x_i\,(E^{15})$.

The general solution of this equation is any arbitrary function independent of x^i. Therefore, according to this condition, the invariance is a symmetry with respect to all the coordinate axes from $\{x^i\}$ that are orthogonal to x^i (for example, in Fig. 10, equality of the component μ_i to zero is a manifestation of the symmetry in the direction x^i).

The opposite condition

$$\frac{\partial \varphi(x)}{\partial x} \neq 0$$

expresses the existence of antisymmetry in the relevant function $\varphi(x)$ along the x^i axis and, therefore, the difference of potentials along this axis, i.e., the existence of a 'current'. As a result, from the standpoint of the existence of circulation, we have a relation for the vector field $\mathbf{R}$ at an arbitrary point A of the domain G,

$$\int_G \operatorname{rot} \mathbf{R}\, \mathrm{d}\omega = -\mathbf{I}(G)\,,$$

which holds for any filter.

Let the image of the operator V_i be an ordered cellular space consisting of points, 'black atoms' on the white planigon. Such atoms are related among themselves and make up a 'construction' of this operator. Consequently, each operator has a corresponding geometric description, which is called a skeleton. This is a graphic description of images on a planigon, which is also applicable for images of complete and closed groups (for example, the line in Figs 4 and 12 is the image skeleton).

Let each atom be a point with mass. Then, a pair of neighboring atoms is connected by a line satisfying the Taylor series expansion for the equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y)$$

in the neighborhood of the point A,

$$y(x) \approx y(A) + y^{(1)}(A)\, x\,,$$

where

$$y(A) = y_0\,.$$

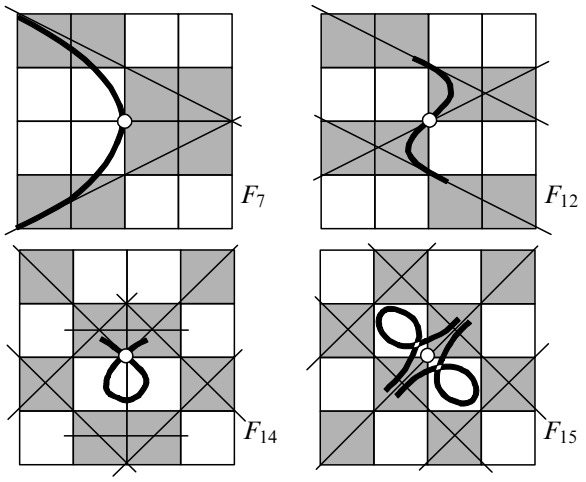

Figure 13. Integral curves on filter images.

Specifying these straight lines as tangent to the directional field in F_i, we obtain that the atom distribution over each F_i corresponds to the integral curve of the solution of a corresponding differential equation (Fig. 13).

Indeed, the transformation realized by a filter is applied to a smooth manifold that can be defined on the set of integral curves defined in the ε-neighborhood of the tangency point A_0 to the planigon on which the coordinate system (x, y) centered at A_0 is given. The result of the filter operation is an expansion vector for which the conditions

$$\mu_0 = \mu_i, \quad \mu_j = 0 \quad \forall j \neq i$$

mean that the observed function belongs to the subdomain of the covering of the filter F_i (the same is true for the mirror filter to F_i if $\mu_0 = \mu_i < 0$).

Because $\{F_i\}$ is the set of functions (3) in the Taylor series, the left-hand side of the equation $\mathrm{d}y/\mathrm{d}x = f(x, y)$ is opposed to the corresponding function from this series. This means that the equation establishes a dependence between coordinates of the point (within the covering) and the angular coefficient of the tangent to the graph of the solution at this point. Hence, any filter defines a directional field and therefore makes it possible to find integral curves (or a family of integral straight lines on the basis of which the envelope is defined) whose tangents at each point (on each atom) coincide with the field direction. We note that the obtained solution of the equation $\mathrm{d}y/\mathrm{d}x = f(x, y)$ (as in the well-known existence and uniqueness theorem) is only sufficient because the specificity of the point A_0 (the center of the planigon) is ignored, while, e.g., for the filter F_3 it can be the envelope for a sheaf of integral lines, or an inflection point for a cubic parabola, or can cover a pair of equilateral and mirror hyperbolas for which A_0 is the intersection point of asymptotes (coordinate axes x, y). Therefore, the obtained solution on the set $\{F_i\}$ is a hypothesis that needs to be confirmed by way of analysis at the next levels of expansion with a higher resolution.

Because $\{V_i\}$ and the algebra on this set are defined on the planigon that can be represented on a connected system, the planigon on this set is a functional space and each operator V_i within the limits of its image is a family of variationally close functions in a function space (e.g., the versions of images of a symbol in Fig. 10 covered by the same image allow a definition in such a family). That is why the sets $\{V_i\}$, $\{P_{ni}\}$, and $\{P_{si}\}$ defined on the planigon viewed as a functional space 'solve' variational problems of the form

$$\int_a^b F(x, y, y^{(1)})\,\mathrm{d}x$$

in this space within the limits of subdomains of coverings over the entire covering of the image of these sets.

We call this problem an inner variational problem because its 'solution' (solution is understood here as approximation to the integral lines on images of operators and operator groups) consists in finding the functional dependence $f(x, y)$ within the covering by the image on the planigon (we recall that to improve the accuracy of the solution it is necessary to move to the next level of expansion within a subdomain of the covering, applying the same technique of analysis). The exterior variational problem is an analysis of the image representation as a whole on the planigon.

Let the image of V_i be the characteristic of a given state of an object in the field of vision. The transition from V_i to V_j then determines a certain deformation tensor through comparison of two states of the object, and their product V_iV_j determines the complete group V_iV_j of such a transition. The same holds for the transitions

$$P_{ni} \to P_{nj}\,, \quad P_{si} \to P_{sj}$$

within the limits of their sets. Such transitions in time are deformations of the observed object in the planigon field.

Let the object be free from deformations; then the transitions in time

$$V_i \to \overline{V}_i \to V_i\,, \quad P_{si} \to \overline{P}_{si} \to P_{si}$$

are movements. If $\{V_i\}$, being an equivalent of $\{F_i\}$, is a basis for E^{15}, then $\{P_{ni}\}$ is the basis manifold for the description of motion on the planigon. Indeed, for $\{V_i\}$ and $\{P_{si}\}$, the image on the planigon occupies half of it and the motion is the 'rocking' of the image from direct to inverse and back. The image (on the intersection operation) of any group P_{ni} is then a compact (four planigon cells connected by the group image,[2] not necessarily glued topologically) whose motion on the planigon is equivalent to the motion of a rigid body represented by its center of mass. To this end, it suffices to consider any complete group on four out of eight of its images: for example, for the group (V_1, V_2, V_3), we have (Fig. 14)

$$V_1V_2V_3, \quad V_1\overline{V}_2\overline{V}_3, \quad \overline{V}_1V_2\overline{V}_3, \quad \overline{V}_1\overline{V}_2V_3\,.$$

Let the trajectory of the center-of-mass motion be a continuous curve, for example, a circle represented on the planigon as on a picture plane. Suppose the moving object has a size of the order of a quarter of the planigon (i.e., conditionally occupies a quadrant; otherwise it is necessary to move to a different level of resolution to either side). Let the object shape at the considered level of resolution be close to a square (if this is not so, it is necessary to either change the planigon shape by adapting it to the object or complement it

[2] We refer to such connectedness as p-connectedness, i.e., connectedness within the image of the complete group, as distinct from the well-known connectedness in chain coding.

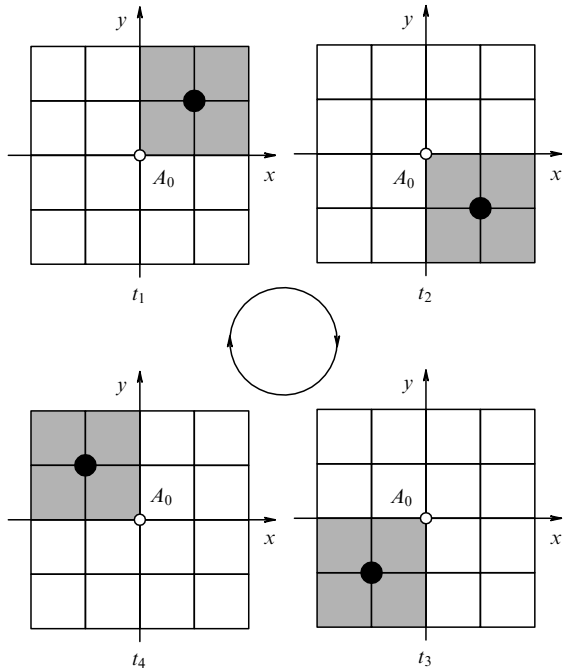

Figure 14. Motion of the compact of the complete group (V_1, V_2, V_3), represented by the center of mass (point) relative to the planigon center A_0.

to a square). Then, on a continuous path and on the set $\{t_i\}$ within the interval T_N, we obtain a sequence of 'pictures' of the type presented in Fig. 14, but additionally involving complete groups (V_1, V_5, V_7) and (V_2, V_4, V_6) when the center of mass of the object ends up on the x and y coordinates, respectively. And this is so if the velocity v of the center-of-mass motion within the interval Δt between neighboring observations satisfies the restriction for the pathlength

$$\Delta l = v\Delta t > \frac{1}{2}\,S\,,$$

where S is the spatial size of the planigon cell.

If the interval Δt is defined such that $\Delta l > (2/3)\,S$, the group (V_1, V_2, V_3) is sufficient for viewing the object.

Thus, $\{P_{ni}\}$ is a necessary and sufficient set describing the motions on the planigon of the center of mass of the object relative to A_0 (at an *a priori* given resolution level and at prescribed instants of time t_i on the observation interval T_N).

A striking example of problems of observation of a given object is the bionic 'frog and fly' problem (which has not yet been solved using standard approaches): a frog watches a fluttering fly and successfully catches it. The frog does not react to other objects that do not correspond to the fly in size or are at rest, even if the fly is right under its very nose. Moreover, using the proposed method, one can predict the object path within the family of variationally close functions.

Suppose a motion on a planigon is observed whose trajectory does not go beyond the region of the covering of V_i, P_{ni}, P_{si} and, therefore, allows a description within the limits of families of variationally close functions (an inner variational problem). Because the trajectory of the motion and current streamlines coincide (irrespective of the stationarity of the process) if and only if the velocities of motion at a given point in space vary in time only in magnitude and not in direction, it follows that the homogeneity region of operator and group images is the invariance region under possible variations and the region of the current streamlines. Hence, every image in the algebra A_V is invariant under a possible nonstationary process represented in the initial space and each image corresponds to the current streamlines. Because every image in the algebra is represented by a homogeneity subregion only and 'nothing exists' outside it, such an image is equivalent to a rigid body. In the translational motion of the body (i.e., such that any segment of a straight line taken on the body moves parallel to itself), irrespective of velocity variations (with its direction preserved), the current streamlines are trajectories. The situation is the same in the rotation of a body around a fixed axis or in arbitrary helical motions in space. Therefore, for a rigid body as a whole, its path is represented by current streamlines. From this, if streamlines are observed, a rigid body has the same path.

5. Spatial properties of algebra elements

We suppose that streamlines exist on images of algebra elements in their domain of definition. Because a planigon is a two-coordinate region, the current streamlines can always be put into correspondence with a line l that is not a streamline, but is a generatrix in the space E^3 in which the planigon is embedded. Then l and the current streamlines constitute a vector surface.

Let the equation of the generatrix l in E^3 have the form

$$\frac{x-\alpha}{a} = \frac{y-\beta}{b} = \frac{z}{1}\,,$$

where a, b, and 1 are direction coordinates and α, β, and 0 are coordinates of the track of the generatrix on the xy plane (in the general case). Then, for an arbitrary linear directrix transported parallel to itself along the generatrix, we obtain a surface equation of the form

$$a\left(\frac{\partial z}{\partial x}\right) + b\left(\frac{\partial z}{\partial v}\right) = 1\,.$$

If the generatrix l is closed, we have a vector (or vortex) tube. For example, for the image of the operator V_4, the current streamlines are the lines of the family $y = y(x, C)$, where C is a constant. The generatrix to them is not necessarily closed in the xz plane in view of cylindrical symmetry of the image (e.g., a circle). As a result, we have a cylindrical surface (not necessarily circular) up to its position on the depth axis z (and the sign of the curvature) with respect to the planigon as a picture plane (Fig. 15). A similar result can be obtained from the equation

$$\frac{\partial^i \varphi(x,y)}{\partial x^k \partial y^m} = \text{const}\,,$$

whose left-hand side corresponds to transformations (3) in the Cartesian coordinate system of the planigon. The solution of the equation is the corresponding surface $z = \varphi(x, y)$.

Therefore, the image of each filter can be regarded as a result of the orthogonal projection of a certain surface, with its position over the planigon plane taken into account. For example, the image of the filter F_8, whose integral representations in the xy-plane are hyperbolas, can be put into correspondence with a hyperbolic paraboloid (see Fig. 15) and the image of F_6 (or F_7) with a conic section. It is of importance to know *a priori* the necessary initial and

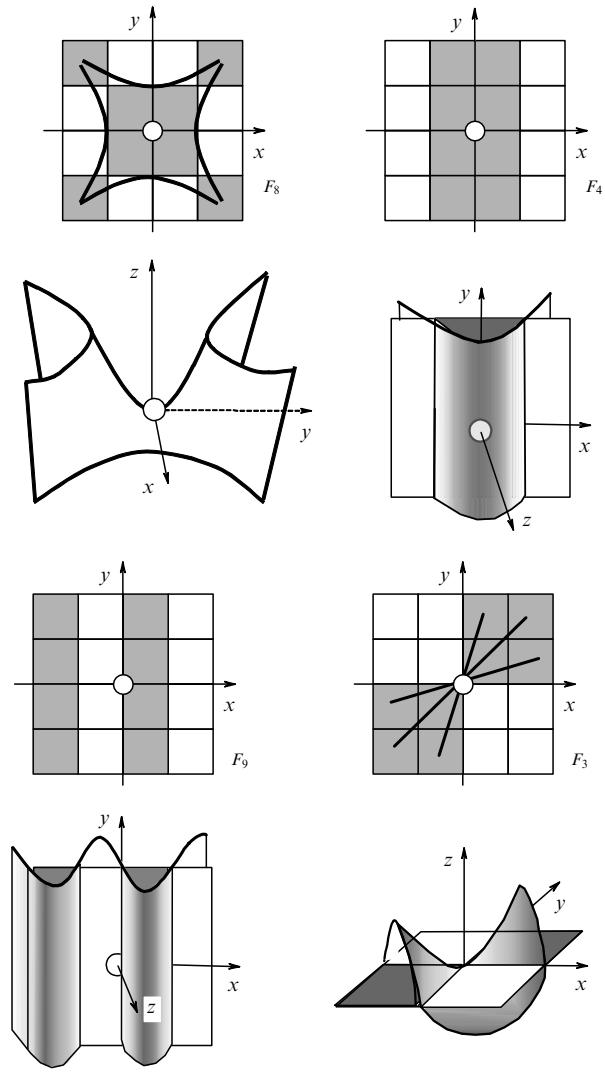

Figure 15. Example of spatial interpretation of images on the planigon.

boundary conditions that cannot readily be obtained from the image only (in order to check the correctness of a visible shape, one needs to touch it). Nevertheless, up to the sign of curvature and distance coordinate along the line of sight, the system of visual perception can rather rapidly reconstruct the shape of the object envelope. The algorithm is approximately the same as in the case of signal reconstruction (see Fig. 11).

We imagine an observable post (a tube, a tree trunk) and assume that the problem of its segmentation is solved. As a result, we have the image of the operator V_4 at a given level of resolution (in this case, the necessary condition for the existence of an extremum, namely, that the first derivatives are equal to zero, i.e., $\mu_1 = \mu_2 = 0$, holds for F_1 and F_2). If we now impose a gray matrix $[m_{ij}]_{4\times 4}$ (but with a higher resolution), we see a variation of half-tones (brightness) along the generatrix. Adopting *a priori* the condition 'the closer, the darker' (or vice versa, depending on the state of the environment — day, night, etc.), we can reconstruct the shape of the surface. An analogous result can be obtained, for instance, by applying a regularly repeating pattern (texture) to the post surface; then a more complicated (in realization) algorithm of analysis of the texture cell (texcel) size along the generatrix can be used (but up to the sign of the surface curvature).

Such an analysis is the solution of the variational problem of finding the extremum of a functional: if each operator image in E^3 corresponds to a vector surface $z = z(x, y)$, then the extremum of the functional

$$v\big(z(x,y)\big) = \iint_G F\left(x, y, z, \frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}\right) \mathrm{d}x\,\mathrm{d}y$$

exists, with the chosen boundary conditions of the covering V_i. In particular, if nonintersecting curves that belong to the family $y = y(x, C)$, where C is a constant, pass through the domain G of the covering V_i, then the family of curves in G is a proper field. For example, for V_4, the current streamlines parallel to the y axis constitute the proper field.

If the domain G of the covering V_i contains a point through which the curves of the family $y = y(x, C)$ pass and have no intersections, such a field is the central field of a sheaf of curves. For example, for V_3, we have the central field for which the point A_0 is the center of the sheaf of lines that do not intersect farther along (see Fig. 15).

The proper or central field of a family of extremals of the variational problem for elements of the image algebra is a field of extremals for which the angular coefficient of the tangent to a curve of the family of the proper field at the point (x, y) is the slope of the field at this point. The situation is similar for a surface in E^3.

By virtue of what has been said above, the plane problem consists in finding the shape of a curve from the family of variationally close functions, to be solved in the first approximation at the level of tangents. On the other hand, defining the extremals by equal-level lines (by mass values on the matrix $[m_{ij}]_{4\times 4}$), one can pass to a solution of the spatial problem of seeking an extremum for a two-dimensional surface in E^3 on the set $\{V_i\}$. To do so, it suffices to know *a priori* the type of problem — plane or spatial — to be solved. Regarding visual perception, the latter type of problem enjoys priority because a person normally looks into the distance (into an improper point of the Euclidean space). In this case, the image, which is a plane because a three-dimensional scene is imaged onto the two-dimensional surface of the retina, is reconstructed according to the laws of perspective with model representation in the form of three-dimensional objects and a subsequent specification of their shapes.

We consider a spatial reconstruction of images of complete and closed groups. If the set $\{F_i\}$ corresponds to a finite set of basis directions, then the set $\{V_i\}$ can be assigned the set of basis vectors $\{\mathbf{e}_i\}$ via equivalence of their elements. In the algebra A_V, these basis vectors are polyad products: $\mathbf{e}_i\mathbf{e}_i$ are dyads and $\mathbf{e}_i\mathbf{e}_j\mathbf{e}_k$ are triads. Therefore, to any operator from $\{V_i\}$ in E^3, we can assign the tensor surface

$$T_{11}(\mathrm{d}x^1)^2 + T_{22}(\mathrm{d}x^2)^2 = \mathrm{const}\,,$$

where degeneration in one of the components is permitted.

Indeed, because any operator, being a divergence component, is a scalar and enters the composition of a planigon (a tangent space) defined on semigeodesics, it follows that any V_i in an orthogonal coordinate system x^1, x^2, x^3 corresponds to the tensor surface

$$T_{11}(\mathrm{d}x^1)^2 + T_{22}(\mathrm{d}x^2)^2 + T_{33}(\mathrm{d}x^3)^2 = \mathrm{const}\,,$$

where T_{ij} are principal components along the principal axes of the tensor

$$T = T_{ij}\,\mathbf{e}^i\mathbf{e}^j = T^{ij}\mathbf{e}_i\mathbf{e}_j$$

in the coordinate system $\{x^i : i = 1, 2, 3\}$ with the origin at the point A_0 of the planigon for which

$$x^3 \equiv z\,, \qquad z(A_0) = 0\,.$$

In this case, any complete group P_{ni} from $\{P_{nj}\}$, generated on the triple (V_i, V_j, V_k), corresponds to a triad $\mathbf{e}_i\mathbf{e}_j\mathbf{e}_k$. In view of the completeness of the group $\{P_{ni}\}$ and the existence of the basis $\mathbf{e}_i$, $\mathbf{e}_j$, $\mathbf{e}_k$, such a group of operators is the linear vector space E_i^3. As a result, we have 35 three-dimensional spaces on the set of complete groups.

Because any two operators from a group are sufficient for the description of the group, each complete group is representable on a dyad from its basis $\mathbf{e}_i\mathbf{e}_j$, $\mathbf{e}_j\mathbf{e}_k$, or $\mathbf{e}_i\mathbf{e}_k$. In a curvilinear coordinate system $(\partial/\partial u, \partial/\partial v)$, where u and v are curvilinear coordinates of the group P_{ni}, such a basis determines a piece of surface (on the set of descriptions, we have a set of pieces)

$$\mathbf{r} = \mathbf{r}(u, v)$$

with an accuracy up to its position in the enclosing space, i.e., in the space of this group.

Any group P_{ni} belongs to its Euclidean space E_i^3 with the basis $\mathbf{e}_i$, $\mathbf{e}_j$, $\mathbf{e}_k$. The corresponding surface in E_i^3, a planigon to which is the tangent plane at the point (0, 0), is defined by quadratic forms, and because the space E_i^3 of the group P_{ni} is 'adapted to its basis', the quadratic form for it has the canonical form (see Fig. 15).

Next, any closed group P_{si} is a superposition of two complete groups

$$P_{ni} = (V_i, V_j, V_k)\,, \quad P_{nj} = (V_n, V_m, V_k)\,,$$

connected by the common operator V_k (i.e., by the common coordinate direction), for example,

$$P_{si} = V_i V_j + V_n \overline{V}_m\,.$$

Because each complete group corresponds to E_i^3, it follows that a closed group is a topological gluing of two such spaces along their common coordinate direction (with an accuracy up to the sign). Hence, any closed group allows a representation in the form

$$\mathbf{e}_i\mathbf{e}_i + \mathbf{e}_n\mathbf{e}_k$$

(up to the inversion of directions). If each complete group (within one of its descriptions) defines a piece of the surface, then two groups glued together define two coordinate-connected pieces of surfaces.

We suppose that the filter F_3 specifies a (positive) direction of x_1 in the invariance subdomain (i.e., in the subdomain of conservation of the weight factor $+1$; Fig. 16a). Then the filter F_8 with the transformations

$$\frac{\partial^4}{\partial^2 x\,\partial^2 y} \equiv \frac{\partial^2}{\partial x\,\partial y}\left(\frac{\partial^2}{\partial x\,\partial y}\right) \equiv \frac{\partial^2}{\partial x^2}\left(\frac{\partial^2}{\partial y^2}\right) \equiv \frac{\partial^2}{\partial y^2}\left(\frac{\partial^2}{\partial x^2}\right) \quad (5)$$

defines a new coordinate system (x_1, y_1) with the center A_0 on the planigon and $B_0 \equiv F_8$ on the grid $V(x, y)$ superposed on the planigon (Fig. 16b). Such a planigon is called a PV-planigon. The center-of-mass displacement to the point A_0 is expressed by two right-hand transformations in (5) (i.e.,

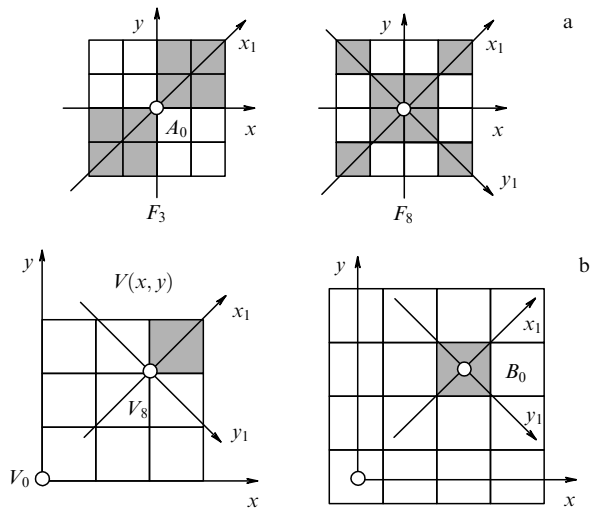

Figure 16. PV-planigon.

the filters F_4 and F_5), and the rotation inherent in F_3 by the composition of its transformations in (5). Therefore, the filter F_8 in its central part (in the closest neighborhood of the point A_0 of its image on the planigon; Fig. 16a) reflects the concentration of sources (masses) in this part through the action of F_4 and F_5, and the concentration of vortices with the centers spaced along the axes through the action of the composition of transformations of the filter F_3. If the direction of the axes is represented as shown in Fig. 16, the coordinate systems (x, y) and (x_1, y_1) with the common center A_0 cannot be superposed by rotation in the planigon plane.

Two coordinate systems on the PV-planigon (and on the operator grid), $K_1 = (x, y)$ and $K_2 = (x_1, y_1)$, naturally distinguish five complete groups:

$$P_{nx} = (V_1, V_4, V_9), \quad P_{ny} = (V_2, V_5, V_{10})$$

are groups whose operators are ordered in the coordinate system K_1;

$$P_{nx1} = (V_3, V_8, V_{15}), \quad P_{ny1} = (V_8, V_{11}, V_{12})$$

are groups whose operators are ordered in the coordinate system K_2; and

$$P_{n0} = (V_4, V_8, V_5)$$

is a connection (image) group of the two coordinate systems.

Let the planigon be the tangent plane at the point A_0 to the surface from the smooth manifold $[z(A_0) = 0]$ and let two coordinate systems, connected by the z direction,

$$K_1 = (x, y)\,, \quad K_2 = (x_1, y_1)\,,$$

be given in this tangent plane. Then, any complete group $P_{ni} = (V_i, V_j, V_k)$ on its descriptions is a tensor surface representation in either of the coordinate systems $\{K_1, K_2\}$.

Indeed, any complete group in the algebra A_V on a set of three variables has eight images: four images on the $+$ operation and four on the $\times$ operation. Because each operator V_i corresponds to the basis vector in the direction $\mathbf{e}_i$ at the point A_0 and the set $\{V_i, V_j, V_k\} \in P_{ni}$ corresponds to the set of orthogonal frames $\{\mathbf{e}_i, \mathbf{e}_j, \mathbf{e}_k\}$, they form a family of

accompanying frames and make it possible to analyze the geometry of a complicated surface in E_i^3. Because any complete group is defined in a tangent plane, for the description of the geometry of the surface it suffices to use the representation

$$T_{11}(\mathrm{d}x^1)^2 + T_{22}(\mathrm{d}x^2)^2 = \mathrm{const}$$

on a pair of operators from the triple, where the first operator gives an estimate of one component and the second gives an estimate of the other. Because the whole set of operators is ordered on the pair of coordinate systems (K_1, K_2), it follows that:

• the unions $V_i + V_j$, $V_i + V_k$, and $V_j + V_k$ are representations on

$$T_{11}(\mathrm{d}x^1)^2 + T_{22}(\mathrm{d}x^2)^2 = \mathrm{const}\,,$$

where $\mathrm{d}x^i$ is the differential with respect to x^i from (K_1, K_2);

• any of the intersections V_iV_j, V_iV_k, and V_jV_k is the tensor representation

$$T = T^{nm}\mathbf{e}_n\mathbf{e}_m\,,$$

where n and m are chosen from (i, j, k), and a surface of the type

$$T_{11}(\mathrm{d}x^1)^2 + T_{22}(\mathrm{d}x^2)^2 = \mathrm{const}\,.$$

Such intersections define images of a complete group by 'new' tensor surfaces.

Thus, for any complete group, each description is a canonical representation on a pair of coordinate systems (K_1, K_2), where eigenvalues play the role of coefficients. Given this, each operator reveals its 'own' surface in $\mathbf{e}_i$; each complete group is a surface that is more complicated in its organization on operators and a closed group is a topological gluing of a pair of surfaces, which thus increases the dimension of the analyzed manifold.

The existence of the pair (K_1, K_2) makes it possible to extend the functionality of the algebra elements in what concerns the reconstruction of spatial relations between objects, including perspective relations.

Let a smooth manifold be one-dimensional and represented by a line. Then a two-dimensional planigon with images from $\{F_i\}$ is a natural 'object' represented as a fiber bundle for a one-dimensional manifold. First, a natural coordinate system (x, y) is allowed on a plane planigon. Second, for a line aligned, e.g., with x, the x axis is a coordinate for each fiber. For instance, for the filter F_5, this is a straight line (skeleton) passing along x through the point A_0. This straight line is the base of the fiber bundle and y is the coordinate inside the fiber, used in the construction of a tangent to the generic layer, translated parallel to the base from fiber to fiber (or vice versa for F_4). The vector $\mathbf{V}$ at an arbitrary point A of the base inside the planigon is then represented in the basis

$$\mathbf{V} = y\,\frac{\partial}{\partial x}\,,$$

where y is the component of $\mathbf{V}$.

The property of foliation on filter images is valid not only in the x, y directions of the Cartesian coordinate system centered at A_0 but also in the z direction of the absolute frame of reference of the observer. In this case, the planigon can be regarded as a picture plane. Such a foliation along the z coordinate is readily realized by the system of visual perception. If we focus our eye on an object, the space in the depth is separated by the picture plane into the front and rear parts. Realizing different focusing points, we obtain information on the depth in different fibers.

The property of space depth reconstruction is an innate property of the visual system. Because the set of transformations realized by the filters belongs to a five-dimensional space (for details, see below), the preferred complete groups

$$P_{nx},\ P_{ny},\ P_{nx1},\ P_{ny1},\ P_{n0}$$

allow a definition of the basis groups of the set $\{V_i\}$. In this case, taking the properties of foliation into account, we can represent a planigon as a 'window' through which we can see (Fig. 17a): the 'window-frame' is the region of peripheral sight and the 'window-pane' is the region of clear vision. With the existence of basis groups taken into account, we have the following interpretation of such a window on the PV-planigon: K_1 is the absolute reference frame and K_2 is the reference frame of the observed object.

To represent objects of a multidimensional, for example, four-dimensional space as regular bodies of a three-dimensional space, the improper space model [17] is used instead of the parallel projection. The model representation in Fig. 17a then specifies the point A_0 of the planigon by an improper point (a horizon point; this is so when we look into the distance: if we are looking at a close object, the 'frame' of the peripheral sight and the region of clear vision exchange places). The situation is analogous to the version in Fig. 17b. If this version corresponds to monocular sight (or the sight through a video camera), then by gluing together two planigons, we obtain the panorama viewed by two eyes with left and right coordinate systems. Figure 17c illustrates the

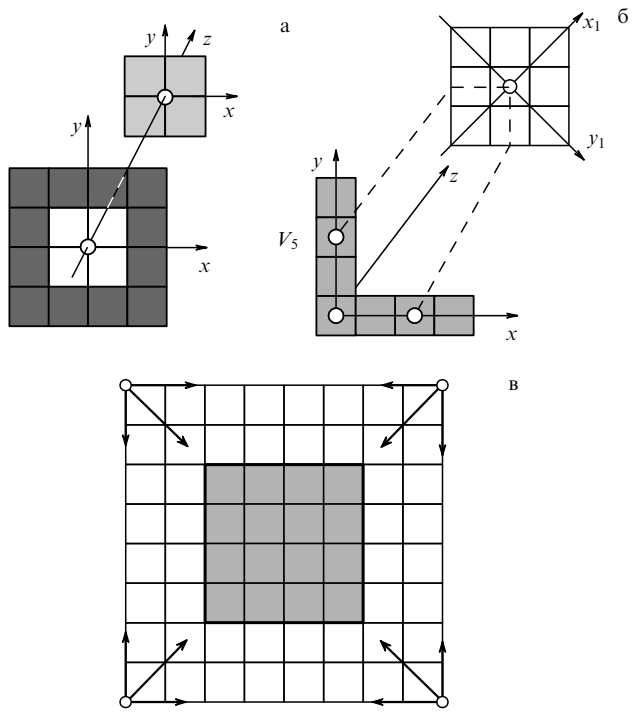

Figure 17. Spacing of planigon elements on the depth axis.

version of interaction of four coordinate systems (and the corresponding four planigons glued together) on a common planigon for a unit of four video cameras.

We now consider examples. A well-known model of a four-dimensional cube is a hypercube; in projection onto a plane from the side of a facet, it is represented by two cubes embedded in one another and connected by the vertices. On a planigon, such a projection corresponds to the image of the operator V_8, which is the center B_0 of the model in Fig. 16b if the square of the planigon is considered as the first square, the second (the internal one) as a square of the image of $V_4 V_5 V_8$, and the 'ears' of the operator as covering the vertex connection lines. If we define the planigon center as a horizon point, we obtain a model of a five-dimensional coordinate system, for which one coordinate is the depth direction (Fig. 18).

Let the image of the closed group

$$P_s = V_4 V_6 + V_8 \overline{V}_{14}$$

be observed on a planigon (see Fig. 9). Each closed group can be assigned a complete group with the operators endowed with trigger properties [6]. For the closed group under consideration, this is $P_{nx} = (V_2, V_5, V_{10})$. The trigger property consists in singling out pairs of complete groups that constitute the image of the closed group:

for the operator V_2, we have

$$V_2 V_4 V_6\,, \quad \overline{V}_2 V_8 \overline{V}_{14}\,;$$

for V_5,

$$V_5 V_4 V_8\,, \quad \overline{V}_5 V_6 \overline{V}_{14}\,;$$

and for V_{10},

$$V_{10} V_6 V_8\,, \quad \overline{V}_{10} V_4 \overline{V}_{14}\,.$$

If we superpose a gray (half-tone) matrix on the closed-group image and assume that dark objects (uniform within the limits of a compact) are closer to the observer in depth, the trigger operators give rise to a spatial separation of the compacts singled out on the closed-group image along the z axis (Fig. 19). The number of versions is equal to the number of trigger operators.

In the considered example, it is necessary *a priori* to make the following conventions. First, the objects are observed in the space E^3. Second, the observation is carried out with a limiting resolution, i.e., at a level of low resolving capacity, where the influence of disturbances is at a minimum. Third, it is necessary to decide what object on the planigon, light or dark, is to be considered the closest.

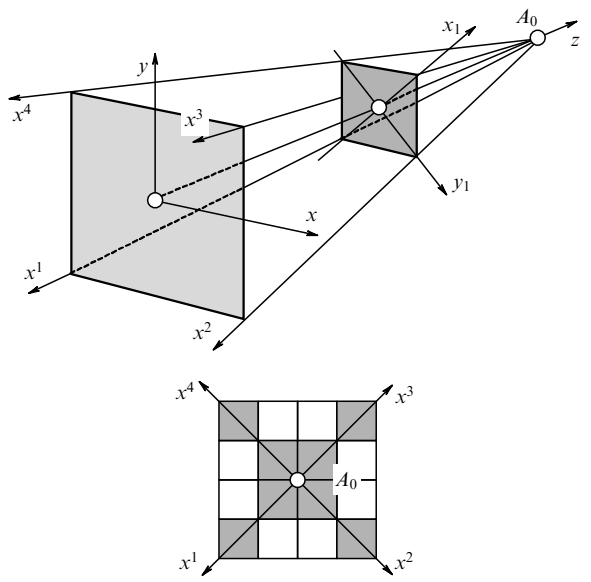

Figure 18. Hypercube projection onto the planigon and the 'perspective' cone with a horizon point z.

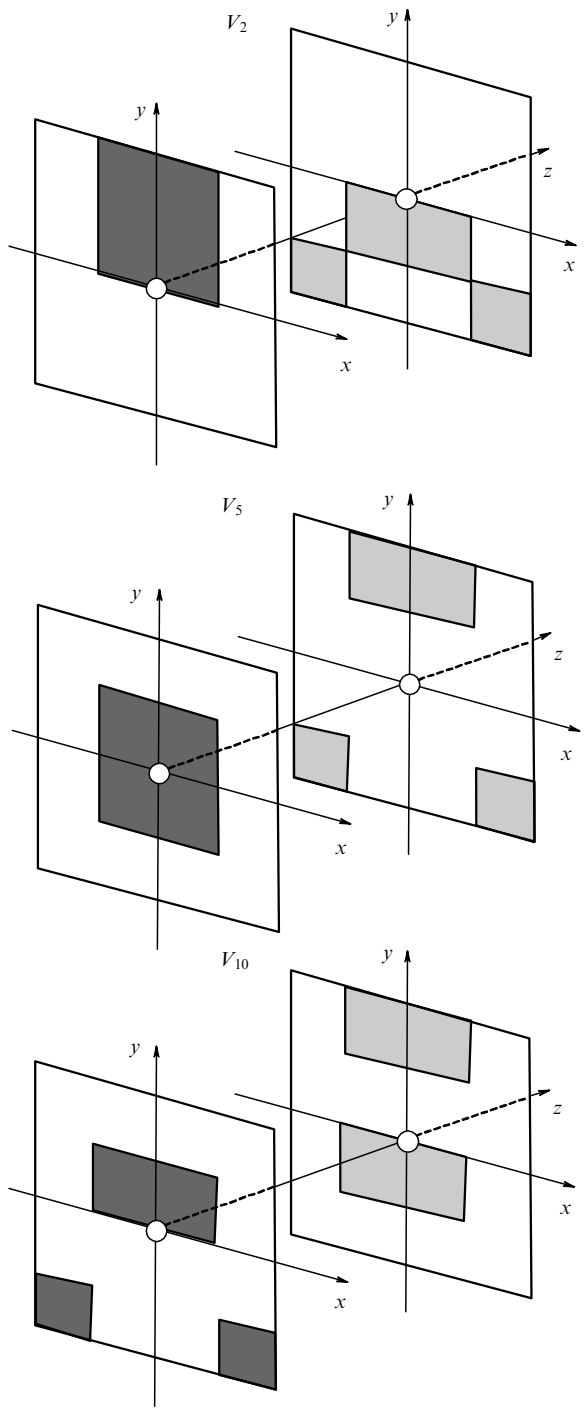

Figure 19. Example of a closed-group image foliation along the z axis.

In painting, perspective is given by a 'road stretching far into the distance' and converging towards the horizon (as in computer graphics). This is more obvious in the side view than in the view from above, that is, along the diagonal of the picture plane. Such a diagonal on the PV planigon is the direction $x = y$ on which the operators V_3, V_8, V_{15} of the complete group P_{nx1} are ordered. Therefore, a planigon in the model representation of Fig. 17 is a natural object on which perspective representations are allowed.

6. Symmetric properties of algebra elements

Let $\{\mathbf{e}_i\} \equiv \{\partial/\partial x^i\}$ be a coordinate basis of the gradient vector field defined on the set $\{F_i\}$. Because a base (line l) of a fiber bundle exists for any F_i, it follows that l is a geodesic on the manifold. Hence, defining the orthogonal set $\{\mathbf{e}_i\}$, the set $\{F_i\}$ defines an orthogonal set of geodesics with the basis $\{\mathbf{e}_i\}$ on the manifold in a natural way. This implies that the set $\{F_i\}$ with the coordinate basis $\{\mathbf{e}_i\} \equiv \{\partial/\partial x^i\}$ defines the Killing vector fields, a maximally symmetric manifold, and a basis of the Lie algebra of Killing vector fields in E^{15}.

Because the power of the set $\{F_i\}$ is equal to 15 (not counting the unit), we have 15 Killing vectors defining the five-dimensional manifold embedded in E^6. Therefore, the set $\{F_i\}$ with the basis $\{\partial/\partial x^i\}$ makes it possible to analyze the symmetry of such a manifold and determines 15 independent equations of motion on a set of 15 normal coordinates.

Thus, we have 15 conservation laws for the space E^5. The model of such a space (a hypercube) was presented in Fig. 18 (on the image of the operator V_8). It naturally determines a necessary coordinate system in the solution of problems of external space geometry reconstruction, the position of the observer in this space, and his/her orientation via observations 'through' the planigon (see Fig. 17).

We show the completeness of the system of transformations (3) from the standpoint of conservation laws [8]. Because a planigon is a connected finite system, it is conservative and the integral variational principles of mechanics with the corresponding conservation laws and equivalence between the total energy and the mass are valid for it. Because the process represented by the mass of domain (1) is analyzed in the planigon field, the 16th conservation law (of mass) on the filter and operator sets is expressed by the elements $F_0 \equiv V_0$.

Suppose a process represented by the function of the domain $F(G)$ is observed on a planigon. It can then be shown that the condition

$$\frac{\mathrm{d}m(G)}{\mathrm{d}t} = 0$$

is necessary and sufficient for conservation of the observed process in time. This condition implies the invariance (uniformity) of the observed image at steps Δt within the observation interval under all transformations (3). The involved convective derivative expresses structure conservation on the image and the local derivative expresses stationarity of the observed process.

Any filter realizes the transformation $\partial/\partial x^i$, and therefore, first, $\operatorname{grad}\varphi = 0$ is the condition of homogeneity of the process in the space E^{15}. Second, the condition $\operatorname{divgrad}\varphi \neq 0$ for any component (i.e., for any value μ_i that can be assigned V_i) is a consequence of the 'action of balanced forces' in x, y directions of the coordinate system defined on the planigon and a manifestation of the momentum conservation law under parallel transports within the limits of the covering V_i.

We note that the conditions

$$\operatorname{grad}\varphi = 0\,, \quad \operatorname{divgrad}\varphi \neq 0$$

make it possible to analyze texturized (regular) images in a natural way, for instance, crystalline structures, dissipative media, fingerprints, etc. The algorithm is relatively simple. If $\mu_0 \neq 0$ at a considered level of resolution and we have $\mu_i = 0$ for all i, then a level (of resolution) exists on which $\operatorname{divgrad}\varphi \neq 0$. Because transformation (1) is a mapping into an absolutely smooth manifold, the planigons of all levels are smoothly glued (a planigon with a low-resolution level is an atlas and subsequent smoothly glued planigons are maps).

The last (for mechanically closed systems) conservation law associated with spatial isotropy is the angular momentum conservation law. Because any operator V_i considered as a rigid body on two coordinate systems (x, y) and (x_1, y_1) connected in the z direction (along the normal to the planigon) has a rotation axis with respect to which this equivalent body is invariant under rotation, this law is satisfied.

Thus, the set $\{V_i\}$ (including V_0) on a planigon as a connected system reflects all three conservation laws. If V_0 is sufficient from the standpoint of the total energy (mass) conservation, the operators V_1 and V_2 are sufficient from the standpoint of the momentum conservation law for a deformation-free mechanical system (body) in x, y directions of the Cartesian coordinate system; the other operators allow one deformation or another, and thus also invariance under parallel transport. The angular momentum conservation law necessarily requires the introduction of a coordinate system and specifying the center of mass (from the standpoint of mechanics). Therefore, it is from the standpoint of this law that the differentiation of operators on different coordinate systems — (x, y) and (x_1, y_1) — is realized. As a result, we have

- $\{V_1, V_4, V_9, V_2, V_5, V_{10}\}$ is a subset defined in coordinates x, y and having cylindrical symmetry;
- $\{V_3, V_8, V_{15}\}$ is a subset defined in coordinates x_1, y_1 and having spherical symmetry;
- $\{V_6, V_7, V_{13}, V_{14}\}$ is a subset defined in coordinates x, y and having conic symmetry; and
- $\{V_{11}, V_{12}\}$ is a subset defined in coordinates x, y and having screw symmetry on a cylinder.

7. Conclusion

Possible physical interpretations of image algebra elements that extend the range of their comprehension and provide insight into the problem of 3-D (the shape of the envelope) reconstruction from a single two-dimensional image are considered. Approaches to the space geometry reconstruction from the standpoint of forward-looking properties of mapping onto a plane are presented. Ways of finding, observing, and describing an object and its motion with a possible prediction of its position at the next instant of time are demonstrated using variational principles.

The developed method has a rather low computational complexity and a high confidence in the sense of the presence of the object of observation within a given accuracy.

The work in this field of research is supported by the RFBR (projects 01-01-00452, 01-01-00459).

References

1. Vavilov S I *Glaz i Solntse* (The Eye and the Sun) (Moscow: Izd. AN SSSR, 1961)
2. Tikhonov A N, Arsenin V Ya *Metody Resheniya Nekorrektnykh Zadach* (Methods of Solution of Incorrect Problems) (Moscow: Radio i svyaz', 1979) [Translated into English: *Solutions of Ill-Posed Problems* (Washington: Winston, 1977)]
3. Hadamard J *Lectures on Cauchy's Problem in Linear Partial Differential Equations* (New York: Dover, 1952) [Translated into Russian (Moscow: Nauka, 1978)]
4. Chang S-K (Ed.) *Principles of Visual Programming Systems* (Englewood Cliffs, NJ: Prentice Hall, 1990) [Translated into Russian (Moscow: Mir, 1994)]
5. Kudryavtsev P S *Kurs Istorii Fiziki* (A Course in the History of Physics) (Moscow: Prosveshchenie, 1974)
6. Utrobin V A *Informatsionnye Modeli Sistemy Zritel'nogo Vospriyatiya dlya Zadach Komp'yuternoĭ Obrabotki Izobrazheniĭ* (Informational Models of the Visual Perception System for Computer Processing of Images) (Nizhniĭ Novgorod: Izd. NGTU, 2001)
7. Duda R O, Hart P E *Pattern Classification and Scene Analysis* (New York: Wiley, 1973) [Translated into Russian (Moscow: Mir, 1976)]
8. Fu K S *Syntactic Methods in Pattern Recognition* (New York: Academic Press, 1974) [Translated into Russian (Moscow: Mir, 1977)]
9. Bloembergen N, in *Nonlinear Spectroscopy* (Ed. N Bloembergen) (Amsterdam: Elsevier, 1975) [Translated into Russian (Moscow: Mir, 1979)]
10. Kurdyumov S P et al., in *Komp'yutery i Nelineĭnye Yavleniya: Informatika i Sovremennoe Estestvoznanie* (Computers and Nonlinear Phenomena: Information Science and Contemporary Natural Science) (Ed. A A Samarskiĭ) (Moscow: Nauka, 1988)
11. Kolmogorov A N *Teoriya Informatsii i Teoriya algoritmov* (Information Theory and the Theory of Algorithms) (Moscow: Nauka, 1987) [Translated into English (Dordrecht: Kluwer Acad. Publ., 1993)]
12. Kondrat'ev V V, Utrobin V A *Dokl. Ross. Akad. Nauk* **347** 316 (1996) [*Dokl. Math.* **53** 295 (1996)]
13. Kondrat'ev V V, Utrobin V A *Dokl. Ross. Akad. Nauk* **350** 315 (1996) [*Dokl. Math.* **54** 800 (1996)]
14. Kondrat'ev V V, Utrobin V A *Dokl. Ross. Akad. Nauk* **356** 178 (1997) [*Dokl. Math.* **56** 801 (1997)]
15. Hilbert D, Cohn-Vossen S *Anshauliche Geometric* (New York: Dover Publ., 1944) [Translated into English: *Geometry and the Imagination, transl. from the German* (New York: Chelsea, 1952); translated into Russian (Moscow: Nauka, 1981)]
16. Kondrat'ev V V, Utrobin V A *Pattern Recogn. Image Anal.* **7** 55 (1997)
17. Kondrat'ev V V, Utrobin V A *Pattern Recogn. Image Anal.* **7** 250 (1997)
18. Landau L D, Lifshitz E M *Mekhanika* (Mechanics) (Moscow: Nauka, 1965) [Translated into English (Oxford: Pergamon Press, 1976)]