COMPUTATIONAL PROBLEMS OF ELECTRICAL ENGINEERING

Vol. 1, No. 1, 2011

FACTORIZATION AND ALMOST FACTORIZATION OF NORMED SEPARABLE SPACES IN SIGNAL PROCESSING TASKS

Yuriy Rashkevych, Dmytro Peleshko

Lviv Polytechnic National University, Automated Systems Control Department
rashkev@polynet.lviv.ua

Abstract. The notion of the metrical space factorization in the tasks of processing a separate space as well as n- image sets based upon the almost factorisation of the images of normed separable topological spaces is exemplified. It is provided by formulating the notion of the almost factorspace as the generalization of the topological space factorization. The almost equivalency class, which gave the possibility of demonstrating the practical realisation of theoretical results is considered. The notions of the almost equivalency and the solution of the task of the superposition of the sets of equitype images for continuous and discrete cases are exemplified.

Keywords: factorization, metrical space, almost equivalency class, almost factorspace, topological space factorization

Introduction

The intensive technological development of the computer hardware and software caused the increase and low price of computing power resulting in large accessible computing resources. They began the expansion in versatile applied spheres of human activities giving some impulse for the active development of already existed scientific areas. The most significant and dynamically developing sphere is the artificial intelligence, i.e. signal, language and image processing. The main criterion of this innovative processing is its intellectuality aiming at the solution of many practical tasks in robotics, control systems, decision making processes, etc. The intellect support problem raises the need of the active development of existing methods of previous processing as well as of new ones for the further use of their results in the applied tasks of the analysis and synthesis of the artificial intelligence systems.

Classic methods of image processing developed by R.Honsales, H.Endryus, M.Tekalp, U.Prett, L.Shapiro, Ye.Putiatyn, V.Hrytsyk, Sh.Peleg, R.Vorobel are elaborated enough and provide effective results in solving the applied image processing tasks. The development of these methods has reached its peak because they were directed towards processing a separate signal.

The development of a new method for building effective computing tools demands the use of new approaches to seeking and finding additional information and knowledge. V. Kozhemiako, R. Tkachenko, E. Bodianskyi, R. Duda, P. Khart and others have contributed to this development. Thus the crucial task is considered to be the development of the effective means of signal set processing, which demands developing new models of their representation as well as methods and algorithms of their processing.

1. Factorisation of topological signal spaces

1.1. THE CASE OF A SEPARATE SIGNAL

Let n-dimensional signal be given

$$P\colon \mathbf{R}^n \to \mathbf{R}^1, \tag{1}$$

which in the discrete representation looks as follows

$$P\colon \mathbf{X}^{n,\infty} \to \mathbf{R}^1 \tag{2}$$

where $\mathbf{R}^n$ – the continuous space of real numbers, which in a discrete case was presented as an unlimited set $\mathbf{X}^{n,\infty}$.

Let the given topology $\Im$ [1, 5] be in the space $\mathbf{R}^n$ n-dimensional open areas $p_m \in \mathbf{R}^n$. E.g., in the case of language signals (one-dimensional signals) p_m – are fragments, and in the case of images (two-dimensional signals) p_m are fragments. The set $\Xi^\infty = \{p_m\}_{m\in[1,\infty]}$ is the topological space covering $(P,\Im)$ [1, 5], i.e. $\mathbf{R}^n \subset \bigcup\limits_{m\in[1,\infty]} p_m$.

For the class K(P) [16] of the signal P every element p_m of the covering Ξ^∞ corresponds to some integral characteristic $\Lambda_m^{\mathrm{K}} \in \mathbf{R}^1$ of the signal representation model [16]

$$p_m \to \Lambda_m^{\mathrm{K}}. \tag{3}$$

The representation (3) is exceptionally injected and gives the possibility for the covering Ξ^∞ to get individual space (as a rule a normed one)

$$\Delta^{\mathrm{K}} = \left\{\Lambda_m^{\mathrm{K}}\right\}_{m\in[1,\infty]}, \tag{4}$$

which elements Λ_m^{K} belong to the space $\mathbf{R}^1$ in accordance with [16]. In general case the characteristics Λ_m^{K} can be multidimensional, but it is compulsory for them to belong to some normed space.

The metrics $d_{\Delta^K}\left(\Lambda_m^K,\Lambda_n^K\right)$ of the individual space Δ^K makes it possible to induce the metrics of the space Ξ^∞

$$\forall p_m, p_n \in \Xi^\infty \quad \exists \Lambda_m^K, \Lambda_n^K \in \Delta^K, \; p_m \to \Lambda_m^K, p_n \to \Lambda_n^K : \\ d_{\Xi^\infty}\left(p_m, p_n\right) = d_{\Delta^K}\left(\Lambda_m^K, \Lambda_n^K\right). \quad (5)$$

Based upon the metrics (5) of the space Ξ^∞, the binary ratio of the equivalency ~ [8] may be introduced as

$$\forall p_m, p_n \in \Xi^\infty, \\ p_m \neq p_n : \; p_m \overset{\Delta^K}{\sim} p_n \leftrightarrow d_{\Xi^\infty}\left(p_m, p_n\right) = d_{\Delta^K}\left(\Lambda_m^K, \Lambda_n^K\right) = 0. \quad (6)$$

which determines the equivalency class $[p_m]$ of the element p_m in the space Ξ^∞ as follows

$$\left[p_m\right] = \{p_k \mid p_k \in \Xi^\infty\}_{k=\overline{1;\infty};k\neq m}. \quad (7)$$

In compliance with [7] the set of all equivalency classes forms a factor space $\Xi^\infty/\sim$ of the space Ξ^∞ by the space Δ^K

$$\Xi^\infty/\sim = \left\{\left[p_m\right] \mid p_m \in \Xi^\infty, m = 1..\infty\right\}. \quad (8)$$

The equivalency ratio (6) determines metrics (5) as semimetrics.

In practical signal processing tasks the determination area is considered to be closed and limited in the space $\mathbf{R}^n$, i.e. a compact, what gives the possibility by [1] out of any convergent covering Ξ^∞ to separate finite subcovering Ξ^N of the dimension N, which is a covering P and belongs to the topology $\Im$. In case of the finiteness of the space Ξ^N its factor space $\Xi^M/\sim$ by the space Δ^K is finite as well.

The factorization according to formulae (3)-(7) is exemplified for a continual case (1), but can be applied for a discrete case. It is taken that the issue of discretization $\mathbf{R}^n$ was considered previously. Then, in practice *the task of factorization of the signal* [7] is solved via semimetrics formation (in accordance with the characteristics of the representation model) for the finite covering Ξ^N of the limited space $\mathbf{X}^n$ and formation by this metrics the equivalency class $[p_m]$ for every element p_m of the covering Ξ^N, which form the factor space $\Xi^N/\sim$ of the space $\mathbf{X}^n$. The formulation of the factorization task is alike in the case of the continual and unlimited area of the signal determination area P.

1.2. THE CASE OF THE SIGNAL SPACE

In the case of the signal space existence

$$\mathbf{P} = \{P_z \mid P_z : \mathbf{R}^n \to \mathbf{R}^1\}_{z\in[1,\infty]}. \quad (9)$$

Then alike the case of the covering Ξ^∞, when the topology $\Im$ of the open sets $p_{z,m}$ is introduced, for every image P_z of the set $\mathbf{P}$ the set of coverings $\bar{\boldsymbol{\Xi}}^\infty$ should be considered

$$\bar{\boldsymbol{\Xi}}^\infty = \left\{\Xi_z^\infty\right\}_{z\in[1,\infty]} = \left\{\left\{p_{z,m}\right\}_{m\in[1,\infty]}\right\}_{z\in[1,\infty]}. \quad (10)$$

Let's remark that in case of the set $\mathbf{P}$ an element $p_{z,m}$ of the topology $\Im$ can be the whole signal P_z.

It means that there are equivalency classes in other spaces of the set $\bar{\boldsymbol{\Xi}}^\infty$ (9) for every element $p_{z,m}$ except of the equivalency classes $\left[p_{z,m}\right] = \left\{p_{z,k} \mid p_{z,k} \in \Xi_z^\infty\right\}_{k=\overline{1;\infty}, k\neq m}$ in the space Ξ_z^∞ of the signal P_z, calculated by (7), what gives possibility to consider a total equivalency class in the space $\bar{\boldsymbol{\Xi}}^\infty$

$$\left[p_{z,m}\right]/\sim = \bigcup_{z=1}^{\infty}\left[p_{z,m}\right] \quad (11)$$

In compliance with (8) and (11), in the case of the signal space (9) the factor space $\bar{\boldsymbol{\Xi}}^\infty/\sim$ of the signal space is considered as the equivalency classes set (11)

$$\bar{\boldsymbol{\Xi}}^\infty/\sim = \left\{\left[p_{z,m}\right]/\sim\right\}_{\substack{z\in[1,\infty]\\ m\in[1,\infty]}}. \quad (12)$$

Practical tasks are characterized with the finiteness of the set $\mathbf{P}$, i.e. dim $\mathbf{P} < \infty$, as well as with closed and limited determination areas $\mathbf{X}_z^n$ for every signal P_z. The use of computer calculating environments additionally demands the discretization of these areas resulted in the fact that for every signal P_z the discrete space $\mathbf{X}_z^n$ is obtained.

Alike the case of the separate signal, for $\mathbf{X}_z^n$ the finite subcovering $\bar{\boldsymbol{\Xi}}^N = \left\{\Xi_z^{N_z}\right\}$ can be separated, where N_z – the dimension of the finite covering of the signal P_z. Then *the signal space factorization task* $\mathbf{P}$ [7] lies in building the factor space $\bar{\boldsymbol{\Xi}}^N/\sim$ via introducing into the spaces $\Xi_z^{N_z}$ by the representation of the model parameter of the signal P_z of the semimetrics (6) and building total equivalency classes $\left[p_{z,m}\right]/\sim$ for every element $p_{z,m}$ of the topology $\Im$.

Let's remark that in the most practical tasks for all signals P_z of the set $\mathbf{P}$ the identical representation model is used, and inside this model the identical parameter is applied. Secondly, for all signals P_z the same determination area is considered, i.e. $\forall z \in [1,\infty]: \mathbf{X}_z^n = \mathbf{X}^n$ is true.

2. Almost factorization of topological signal space

2.1. THE CASE OF A SEPARATE SIGNAL

Let the space Ξ^∞ of the signal P be determined as the equivalency ratio (6), i.e. the semimetrics (5) is given. According to the semimetrics *the almost* *equivalency class* $[p_m|\varepsilon]$ is a class determined almost everywhere within the space Ξ^∞, i.e. within the set $[p_m]_0 \subset [p_m|\varepsilon]$ the following condition is true

$$\varepsilon > 0, \forall p_n \in [p_m]_0 :\ d_{\Xi^\infty}(P_m, P_n) \le \varepsilon. \quad (13)$$

The condition (13) is the condition of the almost factorization, and the variable ε of the formula (13) is denominated as the length (parameter) of the almost factorization.

The almost equivalency class or ε-equivalency $[p_m|\varepsilon]$ is the union of two classes: the equivalency class (7) and the null measure set $[p_m]_0$

$$[p_m|\varepsilon] = [p_m] \cup [p_m]_0. \quad (14)$$

So far the $[p_m]_0$ set measure is equal to 0 (i.e., $\mu([p_m]_0) = 0$), then taking into account the measure additiveness characteristic, the classes $[p_m|\varepsilon]$ and $[p_m]$ become equal by the space measure Ξ^∞

$$\mu([p_m|\varepsilon]) = \mu([p_m]). \quad (15)$$

The almost factor space or the ε-factor space $\Xi^\infty/\overset{\varepsilon}{\sim}$ is defined as such a factor space of the topological space Ξ^∞, where at least one of the space classes is the almost equivalency class.

Statement 1. The measure of almost factor space $\Xi^\infty/\overset{\varepsilon}{\sim}$ is equal to the factor space measure $\Xi^\infty/\sim$.

◁ to proof the statement before let's use the mathematical method of induction:

1st step. Let's consider the case of the space $\Xi^1/\overset{\varepsilon}{\sim}$, i.e. both the set $\Xi^1/\overset{\varepsilon}{\sim}$ and the dimension $\dim\left(\Xi^\infty/\overset{\varepsilon}{\sim}\right) = 1$ comprise the same class $[p_m|\varepsilon]$. Then

$$\mu\left(\Xi^\infty/\overset{\varepsilon}{\sim}\right) = \mu([p_m|\varepsilon]) = \mu([p_m]) = \mu(\Xi^\infty/\sim) \quad (16)$$

Thus, in the case of the space $\Xi^1/\overset{\varepsilon}{\sim}$ the former statement is true.

2nd step. Let's assume that in the case $\Xi^N/\overset{\varepsilon}{\sim}$, i.e. when $\dim\left(\Xi^\infty/\overset{\varepsilon}{\sim}\right) = N$, the previous statement and the following equality are true

$$\mu\left(\Xi^N/\overset{\varepsilon}{\sim}\right) = \sum_{m=1}^{N}\mu([p_m|\varepsilon]) = \sum_{m=1}^{N}\mu([p_m]) = \mu(\Xi^N/\sim) \quad (17).$$

3rd step. Let's check the truthfulness for the case of the space $\Xi^{N+1}/\overset{\varepsilon}{\sim}$. If $\dim\left(\Xi^\infty/\overset{\varepsilon}{\sim}\right) = N+1$ then by the definition of the almost factor space the following equality is true

$$\Xi^{N+1}/\overset{\varepsilon}{\sim} = [p_{N+1}|\varepsilon] \cup \bigcup_m^N [p_m|\varepsilon] \quad (18)$$

Taking into consideration the additiveness measure, the characteristic (15) and the assumption (17) we will obtain the following

$$\begin{aligned}&\mu\left(\Xi^{N+1}/\overset{\varepsilon}{\sim}\right) = \sum_{m=1}^{N}\mu([p_m|\varepsilon]) + \mu([p_{N+1}|\varepsilon]) = \\ &\sum_{m=1}^{N}\mu([p_m|\varepsilon]) + \mu([p_{N+1}]) = \\ &\mu(\Xi^N/\sim) + \mu(\Xi^1/\sim) = \mu(\Xi^{N+1}/\sim)\end{aligned}. \quad (19)$$

It means that in case of the space $\Xi^{N+1}/\overset{\varepsilon}{\sim}$ the measures of factor space and almost factor space are equal. Thus, the statement has been proved ▷

In the case of the closed and limited P signal determination area within the space $\mathbf{R}^n$ almost factorization is considered according to the finite subcovering Ξ^N (p.1.1). In the case the almost factor space $\Xi^N/\overset{\varepsilon}{\sim}$ by the space Δ^K is finite as well.

The almost factor space should be considered as a logical development of classical factor spaces. It is considered to be obvious that ε-factor space $\Xi^\infty/\overset{\varepsilon}{\sim}$ is the continuation (by the value *ε*) of the factor space $\Xi^\infty/\sim$ but only if both exist. Generally, the ε-factor space can exist without factorspace Ξ^∞, but in the case its measure is null according to *Statement 1*. The reverse situation is acceptable: Ξ^∞ – exists, but for the *ε* set there is no element determined for which it is possible to build the set $[p_m]_0$, and the almost factor space $\Xi^\infty/\overset{\varepsilon}{\sim}$ and the factor space $\Xi^\infty/\sim$ converge. Thus, their dimensions in both finite and non-finite cases are equal: $\dim\left(\Xi^\infty/\overset{\varepsilon}{\sim}\right) = \dim(\Xi^\infty/\sim)$. In general we have

$$\dim\left(\Xi^\infty/\overset{\varepsilon}{\sim}\right) = \dim(\Xi^\infty/\sim) + \dim\left(\Xi^\infty(\varepsilon)/\overset{\varepsilon}{\sim}\right), \quad (20)$$

where $\Xi^\infty(\varepsilon)/\overset{\varepsilon}{\sim}$ – the set of all almost equivalency classes.

In the discrete case (2) for the limited determination area $\mathbf{X}^n$ of the signal *P*, *the almost factorization task* [6, 7] is similar to the factorization task (p.1.1) and lies in building the almost factor space $\Xi^N/\overset{\varepsilon}{\sim}$ by the semimetrics (5), determined with the representation model parameter and built being based on the semimetrics factor space $\Xi^N/\sim$ and the class set of the almost equivalency $\Xi^\infty(\varepsilon)/\overset{\varepsilon}{\sim}$.

2.2. THE CASE OF SIGNAL SPACE

In the case of the signal space existence (9) and introduced $\bar{\Xi}^{\infty}$ covering set (10) for every element $p_{z,m}$, together with the equivalency classes $[p_{z,m}]$ it is necessary to consider the almost equivalency classes $[p_{z,m}|\varepsilon]$ similar to the classes $[p_m|\varepsilon]$. The existence of the class $[p_{z,m}|\varepsilon]$ for the signal P_z of the set **P** provides the total almost equivalency class existence

$$\left[p_{z,m} \mid \varepsilon\right] / \sim = \bigcup_{i=1}^{\infty}\left[p_{i,m} \mid \varepsilon\right]. \qquad (21)$$

Then, similar to the definition (12), the almost factor space $\bar{\Xi}^{\infty} / \overset{\varepsilon}{\sim}$ of the signal space **P** will be determined as following

$$\bar{\Xi}^{\infty} / \overset{\varepsilon}{\sim} = \bar{\Xi}^{\infty} / \sim \cup \left\{\left[p_{z,m} \mid \varepsilon\right] / \sim\right\}_{\substack{z\in[1,\infty] \\ m\in[1,\infty]}} \qquad (22)$$

The space $\bar{\Xi}^{\infty} / \overset{\varepsilon}{\sim}$ exists as well in the case of the finite set **P** and limited determination areas of the signals P_z. If the signals are discrete, the factorization space analogy $\bar{\Xi}^{N} = \left\{\Xi_z^{N_z}\right\}$ of the set **P** by the definition (22) is true. Then *the signal almost factorization task* **P** [6, 7] lies in building the almost factor space $\bar{\Xi}^{N} / \overset{\varepsilon}{\sim}$ via forming according to the chosen parameters $\bar{\Xi}^{\infty} / \sim$ of the factor space representation models (in accordance with the semimetrics (5)) and total almost equivalency classes $\left[p_{z,m} \mid \varepsilon\right] / \sim$ (under the almost equivalency condition (13)).

3. Almost factorization in the signal representation models

3.1. THE MODEL OF STATISTICAL REPRESENTATION

According to the model the parameters of the image representation are [10, 16]: $M = M(P)$ – mathematical expectation, $D = D(P)$ – dispersion, $\sigma = \sigma(P)$ – the deviation of the random value P, resulting in the appropriate characteristic spaces Δ^M, Δ^D and Δ^σ : when the topology $\Im$ is set and the covering Ξ^{∞} is determined, by (3) any object p_m of this covering in every space Δ^M, Δ^D and Δ^σ can be represented as

$$\begin{aligned} &p_m \to M_m = M\left(p_m\right); \\ &p_m \to D_m = D\left(p_m\right); \quad p_m \to \sigma_m = \sigma\left(p_m\right) \end{aligned} \qquad (23)$$

where M_m, D_m, σ_m – appropriate mathematical expectation, dispersion, the deviation of the random P within the determination area of the element p_m.

The semimetrics introduction within the spaces Δ^M, Δ^D and Δ^σ gives the possibility to determine the equivalency ratio (6) in every appropriate class. The simplest variant of such a semimetrics by [6] is the certain Euclid's distance in accordance with the model representation parameters. For example, in the case of the characteristic space Δ^M such semimetrics looks like

$$\forall p_m, p_n \in \Xi^{\infty}: \quad d_{\Xi^{\infty}}\left(p_m, p_n\right) = \left|M_m - M_n\right|. \qquad (24)$$

In the discrete finite case

$$\forall p_m, p_n \in \Xi^{N}: \quad d_{\Xi^{N}}\left(p_m, p_n\right) = \\ = \left|\frac{1}{N_{p_m}} \sum_{c_i \in p_m} c_i - \frac{1}{N_{p_n}} \sum_{c_i \in p_n} c_i\right|. \qquad (25)$$

where c_i – elements of the fragments p_n and p_m of the signal P, and N_{p_m}, N_{p_n} – dimensions of their determination areas.

If the objects of the covering Ξ^N have the same dimension, i.e. $N_{p_m} = N_{p_n} = N_p$, the semimetrics determination becomes simpler

$$\forall p_m, p_n \in \Xi^{M}: \quad d_{\Xi^{M}}\left(p_m, p_n\right) = \frac{1}{N_p}\left|\sum_{i=1}^{N_p}\left(c_{i,m} - c_{i,n}\right)\right|. \qquad (26)$$

The second index below determines belonging of the element c_i to the appropriate signal fragment.

In general case for different objects of the covering Ξ^N different characteristic spaces can be chosen. Nevertheless, in the most of practical tasks it is enough to have one characteristic space. It is conditioned mainly by the necessity to minimize calculating spendings in practical signal processing tasks.

In the case of the absence of the characteristic space to formulate and determine the tasks of separable Ξ^N space factorization and almost factorization it is enough to set any semimetrics, based upon the mathematical statistics elements, e.g. the ratio

$$\forall p_m, p_n \in \Xi^{M}: \quad d_{\Xi^{M}}\left(p_m, p_n\right) = \left|\left|r\left(p_m, p_n\right)\right| - 1\right|, \qquad (27)$$

where $r(p_n, p_m)$ – correlation between the elements of the fragments p_n and p_m by [6] is the metrics and the semimetrics of the space Ξ^N. Then the task of factorization and almost factorization anticipates the use of the ratio of equivalency and almost equivalency, which are based on the semimetrics (27). The main disadvantage of the metrics (27) is its high calculating complexity.

If the signal set is given (9), then all exemplified theoretical statements, in particular the introduction of the metrics similar to (24), (25) by the parameters of the representation model or semimetrics (27), are true but only concerning the whole signal P_z.

3.2. ENERGY THEORY OF LINEAR MODELS OF STOCHASTIC SIGNALS

The base of the mathematical model of energy representation is an energy theory of the linear models of stochastic signals, which is suggested in [3, 12, 16]. In compliance with the theory the signal P is considered to be a random process by the abstract function C, indicating the random values within Hilbert space. Then every image corresponds to such energy parameters [3]:

- for impulse signals (the signal class **ε**, the characteristic space Δ^{ε}) – it is energy (the space $\mathbf{L}^p$ – integrated by Lebeh square function [3] is considered)

$$P \to e = \int_{\mathbf{R}^n} \left|C(x)\right|^p dx, \quad p > 1, \qquad (28)$$

where e – the energy of the signal P;

- for signals which don't fade out (the signal class **π**, the characteristic space Δ^{π}) – it is power (the space $\mathbf{B}^p$ – Hilbert space with the metrics of Bora-Bezykovych [3] is considered)

$$P \to \text{п} = \lim_{\theta\to\infty} \frac{1}{2\theta} \int_{-\theta}^{\theta} \left|C(x)\right|^p dx, \qquad (29)$$

where п – the power of the signal P.

Further, let's consider only the signal class **ε**. The samples for the **π** class signals will be analogical. If the topology $\Im$ is set and its covering Ξ^{∞} is determined, by (3) any object p_m of this covering of the classes Δ^{ε} is represented via the energy $e_m = e(p_m)$ calculated according to (28). Similar to (24) it is possible to calculate the equivalency ratio, formulate and solve the tasks of factorization and almost factorization.

In the finite case the representation model parameter will be finite energy. If the covering Ξ^{∞} is finite and disjunctive, then the representation (28) can be written as following

$$P \to \sum_{i=1}^{N} e_m,\ e_m = e(p_m),\ p_m \in \Xi^N, \qquad (30)$$

For the signal space (9), similar to the statistical model case all theoretical samples are possible as well. The other method suitable for setting the factorization and almost factorization tasks is to use the known seminorm of the space $\mathbf{L}^p$[3]

$$\left\|P_z\right\|_p = \left(\int_{\mathbf{R}^n} \left|C(x)\right|^p dx\right)^{\frac{1}{p}}, \qquad (31)$$

The seminorm (31) gives the possibility to set the equivalency ration which is almost everywhere similar to the signal similarity which precision is equal to null. Appropriately, the factor space in this case, which is determined in accordance with the equivalency ratio, is the almost factor space of some almost factorization parameter value. Let's note that the approach of the seminorm (31) can be used for the cases of both factorization and almost factorization of a separate signal of the topological space.

3.3. THE INFORMATION MODEL OF THE SIGNAL REPRESENTATION

The information model of the signal representation is based upon the information theory, which development is dynamic information theory fully set in [2, 9, 16]. If the signal P is interpreted as a continual one within the topological manifold $\mathbf{R}^n$ or as a random function or random process C, then by the dynamic information theory it can correspond with the following information parameters [2]:

1) informativeness $\bar{I}^{\delta}$ – the quantity of information, which is a constituent of the realization of the random process C in the case of its representation by the step function of the step value equal to δ_x (the characteristic space $\Delta^{\bar{I}^{\delta}}$)

$$P \to \bar{I}^{\delta} = \frac{1}{\delta_x} \int_{\mathbf{R}^n} \left|C'(x)\right| dx. \qquad (32)$$

2) δ-entropy H^{δ} – the average value of a random process change in quanta δ_x within the interval Δx – the measure of the random process change indetermination (the characteristic change $\Delta^{H^{\delta}}$)

$$P \to H^{\delta} = \frac{\Delta x}{\delta_x} \mathrm{M}\left[\frac{dC}{dx}\right]; \qquad (33)$$

3) set δ-entropy h^{δ} – the indetermination change, which is calculated by the distribution law and is within the range (0-1) (the characteristic space $\Delta^{h^{\delta}}$)

$$P \to h^{\delta} = \frac{\mathrm{M}\left[C'(x)\right]}{\max\limits_{x\in\mathbf{R}^n}\left[C'(x)\right]} \qquad (34)$$

If the topology $\Im$ is set and its covering Ξ^{∞} is determined in accordance with (5), (32)-(34) are metrics of the characteristic spaces $\Delta^{\bar{I}^{\delta}}$, $\Delta^{H^{\delta}}$ and $\Delta^{h^{\delta}}$, and the space Ξ^{∞} is following

$$\begin{aligned}&\Delta^{\bar{I}^{\delta}} : d_{\Xi^{\infty}}(p_m, p_n) = \left|\bar{I}_m^{\delta} - \bar{I}_m^{\delta}\right|;\ \Delta^{H^{\delta}} : d_{\Xi^{\infty}}(p_m, p_n) = \\ &= \left|H_m^{\delta} - H_m^{\delta}\right|; \\ &\Delta^{h^{\delta}} : d_{\Xi^{\infty}}(p_m, p_n) = \left|h_m^{\delta} - h_m^{\delta}\right|\end{aligned} \quad . \qquad (35)$$

The metrics provide setting the factorization and almost factorization tasks in the information representation model case.

In the finite discrete case (2) the representations (32)-(34) will look like the following

$$\begin{aligned}&P \to \bar{I}^{\delta} = \frac{1}{s}\sum_{k}^{N^{\delta}} I_{(k)}^{\delta} = \frac{1}{s}\sum_{k}^{N^{\delta}} \log_2\left(N_{(k)}^{\delta}+1\right); \\ &I_{(k)}^{\delta} = \log_2\left(N_{(k)}^{\delta}+1\right)\end{aligned} ; \qquad (36)$$

$$\begin{aligned}&P \to H^{\delta} = \sum_{k=1}^{N^{\delta}} p_{(k)} I_{(k)}^{\delta} = \sum_{k=1}^{N^{\delta}} p_{(k)} \log_2\left(N_{(k)}^{\delta}+1\right), \\ &p_{(k)} = \frac{N_{(k)}^{\delta}}{s}\end{aligned} ; \qquad (37)$$

$$P \to h^{\delta} = H^{\delta}\left(\log_2\left(\frac{\left|x_{\max}\right|}{\delta_x}+1\right)\right)^{-1}, \qquad (38)$$

where $I^{\delta}_{(k)}$ – the quantity of information, which is at the k level of rating (quantization); s – the signal dimension; $N^{\delta}+1$, $N^{\delta}_{(k)}$ – the quantity of quantization levels within the area of values of the signal P and the k random value.

In the case of the signal set (9) the parameters (32)-(34) or (36)-(38) refer to both the elements of every signal topology P_z and the signal P_z in general.

3.4. VECTOR REPRESENTATION MODEL

The vector representation model [6, 14, 16] of the signal P is based upon the hypothesis concerning the existence of some vector function **C** determined within the space $\mathbf{R}^3$. Thus, the first parameter of the model is the stream of the vector **C** through the hypersurface S

$$P \to \Phi = \int_{S \subset \mathbf{R}^n} \mathbf{C} \cdot dS , \qquad (39)$$

existing under the following conditions: the vector **C** is taken to be directed by the normal towards the surface S; the vector element dS of the surface exists almost everywhere within S and coincides by the normal with the direction **C**. In (5) the instance of the space $\mathbf{R}^3$ is considered because in the case of the space $\mathbf{R}^n$ the stream visualization is lost partially, nevertheless all the hypersurface characteristics coincide in accordance with the theorem of Ostrohradskyi-Hauss. That is why the surface $\mathbf{\Phi}$ creating the characteristic space Δ^{Φ} exists.

The 2nd parameter of the model is a color gradient which is alike the vector function of Euclid's space of the **C** vector coordinates within the surface S. The gradient diversity forms the vector (potential) field.

Let's consider different methods of semimetrics development to solve the tasks of factorization and almost factorization in the vector representation model.

In compliance with [6, 11, 14] if the topology $\Im$ is set and its covering Ξ^N is determined the expressions below

$$\langle p_m, p_n \rangle = \Phi_m \Phi_n ;$$

$$d_{\Xi^{\infty}}\left(p_m, p_n\right) = \left|\Phi_m - \Phi_n\right| , \qquad (40)$$

are considered to be a scalar product and the metrics of the Hilbert space $\left(\Omega_{\mathbf{C}}, \|\cdot\|_{\mathbf{C},\Phi}\right)$, which is separable as well, appropriately. $\Omega_{\mathbf{C}}$ – the normed space of the **C** vector function with the norm $\|\cdot\|_{\mathbf{C},\Phi}$, which is determined through the stream Φ as: $\|p_m\|_{\mathbf{C},\Phi} = \Phi_m$.

In the case of the finite space $\mathbf{X}^2$ the dimension $l \times h$ of [6] we have the following:

- the space $\left(\Omega_{\mathbf{C}}, \|\cdot\|_{\mathbf{C},\det}\right)$ with the determined norm $\|p_m\|_{\mathbf{C},\det} = \left|\det\left(C_m\right)\right|$;
- the space $\left(\Omega_{\mathbf{C}}, \|\cdot\|_{\mathbf{C},p}\right)$ with the determined norm $\|p_m\|_{\mathbf{C},p} = \sqrt[p]{\sum_{j=1}^{h}\sum_{i=1}^{l}\left(c_{m,(i,j)}\right)^p}$ (if $p = 2$ – Frobenius metrics);
- the space $\left(\Omega_{\mathbf{C}}, \|\cdot\|_{\mathbf{C},\mathrm{singl}}\right)$ with the multiplicative metrics $d_{\Xi^N}\left(p_m, p_n\right) = \max_i \sigma_i\left(C_m - C_n\right)$ (where σ_i – the metrics number).

In accordance with the denomination norms and metrics C_m stands for the metrics of the coordinates of the vector function C within the hypersurface p_m, and $c_{m\,(i,j)}$ being its elements.

If (9) is the image set, there is the 3rd parameter of the vector representation model, i.e. Euclid space $\{\Phi_z\}$:

$$\mathbf{P} \to \left\{\Phi_z\right\}_z \qquad (41)$$

Semimetrics is to set the tasks of factorization and almost factorization, which refer to the separate signal case and can be used for the image set without reserve.

4. Factorization and almost factorization on the example of the solution of the image superposition task

To realize the former theoretical studies in practice let's consider the superposition task in the set of one type digital images. In the case the images are considered as a discrete signal determined within two-dimensional limited area. Then the set (9) looks like

$$\mathbf{P} = \{P_z \mid P_z : \mathbf{X}^2 \to \mathbf{N}\}_{z=\overline{1,K}} . \qquad (42)$$

The superposition task lies in similarity detection between the set images or between their separate fragments and in further determination of superposition parameters, in particular horizontal and vertical superpositions, turning angle and scaling coefficients, i.e. the whole set of affine transformations is considered. To simplify practical realization let's set limits searching only for horizontal and vertical superpositions.

Similarity detection algorithms depend on the characteristics of the stochastic interconnection of the compared image fragments [4]. Traditionally to realize the procedure of detecting images superposition the correlation tether (correlation maximum) of the digital images is used [6].

The main disadvantage of the correlation tether method is the great combinatory complexity demanding significant calculating and time resources. Time spendings are the most undesirable, so far as they do not make possible to build image analysis systems in real time (or approximated to real time). All efforts to accelerate the correlation tether lay in unparalleling the calculating algorithm which is not effective in cases of large-dimensional sets.

Under the conditions of the task of superposition concerning the set (42) we have the following data: the fixed image $P_{\text{фікс}} = P_1$, the frame $\mathbf{X}^2_{\mathbf{fr}1,\text{зад}}\left(\Delta_{x\,1,\text{зад}},\Delta_{y\,1,\text{зад}},l_{\mathbf{fr}1},h_{\mathbf{fr}1}\right)$ (the set of pixels as a rectangular subarea of the determination area $\mathbf{X}^2$) and the corresponding fragment $p_{1,\text{зад}}$ (the set of integer values of every pixel color/intensity of the frame $\mathbf{X}^2_{\mathbf{fr}1,\text{зад}}$). Here we have the coordinates of the beginning (a top left angle) and the length and height $\Delta_{x\,1,\text{зад}},\Delta_{y\,1,\text{зад}},l_{\mathbf{fr}1},h_{\mathbf{fr}1}$ of the images. The superposition task lies in forming the set $\mathbf{P}'$ in the condition that

$$\mathbf{P}' = \mathbf{P}\setminus\left\{P_1\right\} = \left\{P_z\right\}_{z=\overline{2,K}} .$$

and all the images of the set $\mathbf{P}'$ are equal with regard to $P_{\text{фікс}}$ within a pixel. The topology $\Im$ of the set (42) is a set with the fragments of the frames [6, 11], which are equal to the frame $\mathbf{X}^2_{\mathbf{fr}1,\text{зад}}$ by their physical dimensions.

Under the set of initial conditions and with the use of semimetrics separated in p. 3 there are such superposition methods developed:

- methods based upon the mathematical expectation and dispersion (semimetrics (26)) [10];
- method based upon the finite energy/average power (semimetrics (26) but with the parameter (28)) [12];
- vector method based upon the color vector-function stream (semimetrics (40));
- method based upon signal/noise peak correlation (semimetrics $d\left(p_m,p_n\right)=\left|\mathrm{PSRN}\left(p_m\right)-\mathrm{PSRN}\left(p_n\right)\right|$) [13];
- method based upon informativeness / entropy / exemplified entropy (semimetrics (35)) [9];
- vector method based upon the Frobenius metrics (the space $\left(\Omega_{\mathbf{C}},\|\cdot\|_{\mathbf{C},2}\right)$) [6];
- vector method based upon the determinated norm (the space $\left(\Omega_{\mathbf{C}},\|\cdot\|_{\mathbf{C},\det}\right)$) [15];
- vector method based upon the singular metrics (the space $\left(\Omega_{\mathbf{C}},\|\cdot\|_{\mathbf{C},\text{singl}}\right)$) [6];
- the method based upon the peak signal / noise correlation as images similarity measure (semimetric

$$d_{\text{psnr}}\left(p_m,p_n\right)=\text{psnr}\left(p_m,p_n\right)=10\lg\left(\frac{h_{\mathbf{fr}1}l_{\mathbf{fr}1}\max\limits_{(i,j)}{}^2 c_{z(i,j)}}{d_{\mathbf{C},\text{Фрб}}\left(p_m,p_n\right)}\right),$$

here $d_{\mathbf{C},\text{Фрб}}\left(p_m,p_n\right)$ – Frobenius metrics) [13].

The general scheme of all the methods comprises the following stages:

1st stage. By the elements of the topology $\Im$ of the set $\mathbf{P}'$ the $\boldsymbol{\Xi}^N=\left\{\Xi_z^{N_z}\right\}$ determination of finite coverings.

2nd stage. Via solving the almost factorization task within the set $\mathbf{P}'$ determination of the almost factorization class $\left[p_{1,\text{зад}}\mid\varepsilon\right]/\sim=\left\{\left[p_{1,\text{зад}}\mid\varepsilon\right]_z/\sim\mid z=\overline{2,K}\right\}$ and corresponding frame for every class element. Here $\left[p_{1,\text{зад}}\mid\varepsilon\right]_z/\sim$ is a subclass of the class $\left[p_{1,\text{зад}}\mid\varepsilon\right]/\sim$, which comprises z image fragments.

3rd stage. Constricting of every subclass $\left[p_{1,\text{зад}}\mid\varepsilon\right]_z/\sim$ to the set $\{p_{z,\max}\}$ consisting one element of the frame $\mathbf{X}^2_{\mathbf{fr}z,\max}\left(\Delta_{x\,z,\max},\Delta_{y\,z,\max},l_{\mathbf{fr}1},h_{\mathbf{fr}1}\right)$. The element $p_{z,\max}$ is determined via solving the task of searching for the correlation maximum within the subclass $\left[p_{1,\text{зад}}\mid\varepsilon\right]_z/\sim$.

4th stage. For every image P_z we search the horizontal $\Delta_{x,(z)}$ and the vertical $\Delta_{y,(z)}$ of the coordinate superpositions $\Delta_{x,(1,z)}$ and $\Delta_{y,(1,z)}$ for the image P_Z

$$\Delta_{x,(z)}=\Delta_{x,z,\max}-\Delta_{x,1,\text{зад}};\quad \Delta_{y,(z)}=\Delta_{y,z,\max}-\Delta_{y,1,\text{зад}} .$$

In Fig.1-3 the results of experiments at the different input parameters of image sets are exemplified. The influence of different characteristics upon the rapidity and quality of the set of equitype images the solution of the superposition task has been studied. The classic correlation tether method is chosen for comparison.

In Fig.1 the average values of the work time of the algorithms are exemplified with the following set characteristics: the set dimension – K = 99 images; grey gradation image; the dimension of every image – l = 34; h = 54 pixels; $P_{\text{фікс}} = P_0$. The parameters of the set frame are $\mathbf{X}^{2,+,d}_{\mathbf{fr}0,\text{зад}}:\Delta_{x,0,\text{зад}}=\Delta_{y,0,\text{зад}}=10;\quad l_{\mathbf{fr}0}=h_{\mathbf{fr}0}=10;$ $\varepsilon = 0.01$.

As you can see in Fig.1 the algorithms built on the base of the developed methods operate faster then the algorithms based on the well-known methods.

In Fig.2 the dependence of the operating time of the algorithms on the set dimension is illustrated.

According to these results we may state that in cases of small dimension sets it is recommended to use the superposition methods based upon the mathematical statistics as the fastest and having satisfactory superposition results. When the set dimensions are large, it’s better to use the methods of the image representation vector model, because the rapidity of the algorithms built upon the almost factorization of representation vector model Hilbert spaces is increasing more dynamically, if K is rising.

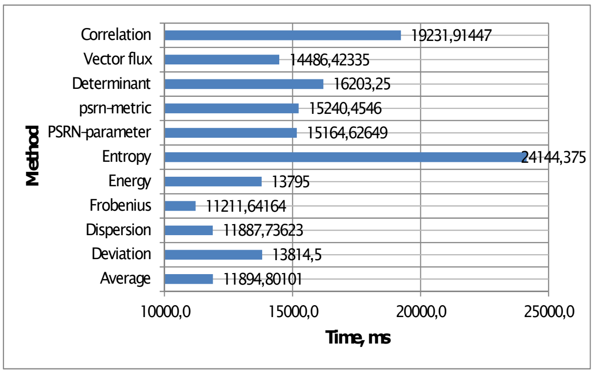


Fig. 1. Comparison of work time of image set superposition by different algorithms (K = 99 images)

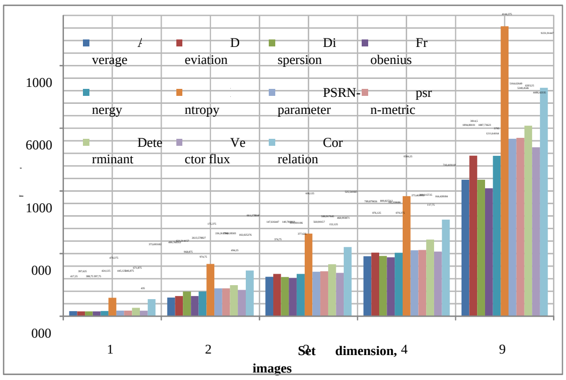


Fig. 2. Comparison of time results of image set superposition methods work at different set dimensions

During experiments the worst correlation values of the set fragment (standard) square and the general image square have been determined. The problem of the standard choice is thoroughly studied in [17] and is not considered in the paper. Moreover, the practical experiment results prove that the waste time dynamics is the lowest in the algorithms built upon the information model parameters. Thus if the standard dimensions are close to the critical correlation, it is necessary to use the methods based upon entropy, set entropy or informativeness. Furthermore, it is possible to influence upon the algorithm operating speed in general via the choice of quantization methods.

The results of the mistakes appeared during the developed superposition algorithms at different values of the almost factorization parameter are illustrated in Fig.3 The data obtained certify the dependence of the operation quality of the superposition algorithms on the parameter ε. Qualitatively the dependence is inversely proportional to the operation quality of the algorithms.

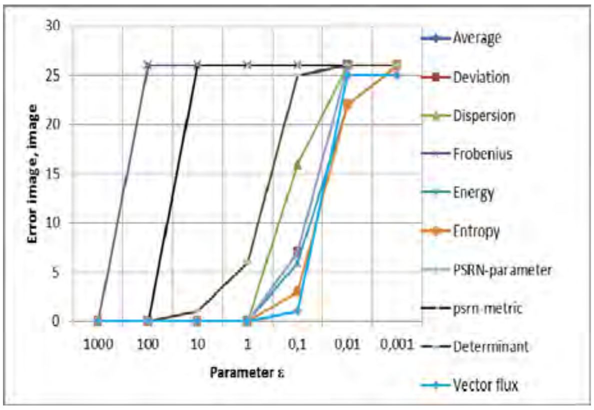


Fig.3. Quantity of mistakes at different values of almost factorization parameters of different superposition algorithms

On the other hand, more mistakes appear if ε values are small. Thus while choosing a method for practical experiments it is significant to take into consideration not only the dimensions of the set and fragment but the parameter ε.

The developed methods work effectively for both semitone and color images. To build the parametersof the base representation model in the superposition methods the integral color values can be used, since it accelerates the algorithm performance in general. All methods together with the image set superposition task give possibility to solve the classical task of searching object by the sample given. They can be partially or fully used while solving the applied tasks of segmentation, classification and recognition.

Conclusions

The formulated notion of almost factorization of the topological space signal is the extension of the metrics space notion, if the topology and the equivalency and almost equivalency ration is set. It gives the possibility to generalize separate classical tasks and create the single system approach in the signal processing area of the artificial intelligence systems.

The development of the almost factor space can be considered as an operation, which gives the possibility to formulate factor spaces within insignificant assumptions.

The task of the one type images set superposition has been solved basing upon the almost factorization task setting in the case of finite discrete two-dimensional signals. The approach used can be considered as basic for many practical methods of signal and image processing.

Introducing other semimetricses as well as ratios of equivalency and almost equivalency, it is possible to formulate other tasks of factorization and almost factorization and develop new methods of signal processing .

References

1. Александров П. С. Введение в теорию множеств и общую топологию/ П.С. Александров. – М.: Наука, 1977. – 368с.

2. Боюн В.П. Динамическая теория информации. Основы и приложения/ В.П.Боюн. – Киев: ин-т кибернетики им. В.М.Глушкова НАН Украины, 2001. – 326с.

3. Драган Я.П. Енергетична теорія лінійних моделей стохастичних сигналів/ Я.П. Драган. – Львів: Центр стратегічних досліджень еко-біо-технічних сис-тем, 1997. – 333с.

4. Зорьян Л. Б. Алгоритм распознавания двоичных объектов при взаимно– коррелированных признаках/ Л. Б.Зорьян// Кибернетика. – 1976. – № 5. – С. 147 – 148.

5. Милнор Дж. Дифференциальная топология/ Дж. Милнор, А М. Уоллес. – М: Мир, 1972. – 279с.

6. Пелешко Д.Д. Суміщення наборів однотипних зображень/ Д. Д. Пелешко. – Львів: Національний університет "Львівська політехніка", 2010. – 140с.

7. Пелешко Д.Д. Задача факторизації просторів визначених на топологіях зображень/ Д.Д. Пелешко, Ю.М. Рашкевич// Бионика интеллекта: науч.-техн. журнал. – 2010. – № 1(72). – С. 98

8. Пелешко Д. Майже факторизація гільбертового простору на основі метрики Фробеніуса для вирішення деяких задач обробки зображень/ Д. Пелешко, Н. Кустра, З. Шпак// Журнал "Вісник Хмельницького національного університету". – 2010. – № 1(144). – С. 204 – 208.

9. Пелешко Д. Інформативність та ентропія динамічної теорії інформації в окремих задачах оброблення зображень у наборах/ Д. Пелешко// Наук. вісник НЛТУ України: Зб. наук.-техн. праць. – 2009. – Вип. 19.9. – С. 291 – 303.

10. Пелешко Д.Д. Використання математичного сподівання для швидкого центрування наборів зображень/ Д.Д. Пелешко// Зб. наук. праць інституту проблем моделювання в енергетиці ім. Г. Є. Пухова НАН України: "Моделювання та інформаційні тех-нології". – 2009. – Вип. 50. – С. 135 – 144.

11. Пелешко Д.Д. Топології зображень та наборів зображень/ Д. Пелешко// Наук. вісник НЛТУ України: Зб. наук.-техн. праць. – 2009. – Вип. 19.4. – С. 236 – 242.

12. Пелешко Д.Д. Використання характеристик енергетичної теорії моделей стохастичних сигналів для вирішення задачі суміщення однотипних зображень в наборі/ Д.Д. Пелешко, М.С. Пасєка// Наук. праці: наук.-методичний журнал ЧДУ ім. П. Могили. Серія "Комп'ютерні науки". – Миколаїв: ЧДУ ім. П. Могили. – 2009. – Вип. 104. – Т. 117. – С. 48 – 58.

13. Пелешко Д. Використання пікового співвідношення сиґнал/шум для вирішення задач майже факторизації топологічних просторів наборів зображень/ Д. Пелешко, Н. Кустра, З. Шпак// Наук. вісник Чернівецького університету: Зб. наук. праць. Комп'ютерні системи та компоненти. – 2009. – Вип. 479. – С. 42 – 48.

14. Адитивні групи функціоналів визначених на топологіях зображень: Матеріали 4-ої міжнародної конференції ["Комп'ютерні науки та інформаційні технології CSIT'2009"], (Львів, 15-17 жовтня вересня 2009)/ Національний університет "Львівська політехніка". – Львів: ПП "Вежа і Ко ", 2009. – С. 123-126.

15. Використання детермінатної метрики для майже факторизації гільбертового простору на топологіях зображень: Матеріали міжнародної наук. конференції ["Інформаційні – телекомунікаційні технології в сучасній освіті: досвід, проблеми, перспективи "], (Львів, 2009)/ – Львів, 2009. – Т. 2. – Част. 1. – С. 228 – 233.

16. Класифікація моделей представлення зображень та наборів зображень як стохастичних зображень та полів: Матеріали наук.-практ. конференції ["Інтелектуальні системи прийняття рішень та проблеми обчислювального інтелекту ISDMCI'2009"], (Євпаторія, 18-22 травня 2009)/ Херсонський національний технічний університет. – Херсон: ХНТУ, 2009. – Т. 2. – С. 401 – 405.

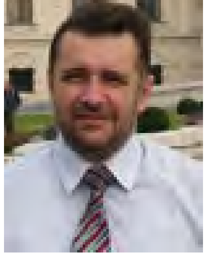
17. Шлезингер М.И. О построении эталонов для корреляционных читающих автоматов/ М.И. Шлезингер, Л.А. Святогор// III Всесоюзная конференция по информационно поисковым системам и автоматизированной обработке научн.-техн. информации. – М.:. – 1967. – Т. 3. – С. 129 – 139.

ФАКТОРИЗАЦІЯ ТА МАЙЖЕ ФАКТОРИЗАЦІЯ НОРМОВАНИХ СЕПАРАБЕЛЬНИХ ПРОСТОРІВ В ЗАДАЧАХ ОБРОБКИ СИГНАЛІВ

Розглянуто поняття факторизації метричного простору в задачах обробки як окремого так і простору n-вимірних сигналів для неперервного та дискретного випадків. Запропоновано поняття майже еквівалентності та класу майже еквівалентності, що дало можливість сформулювати поняття майже факторпростору як узагальнення факторизації топологічного простору. З метою демонстрації практичної реалізації теоретичних результатів приведено вирішення задачі суміщення наборів однотипних зображень на основі майже факторизації нормованих сепарабельних топологічних просторів зображень.

Yuriy Rashkevych

Dr.Sc., Prof., Vice-Rector for International Relations, Head of Automated Systems Control Department, Institute of Computer Science and Information Technologies, Lviv. Polytechnic National University



Dmytro Peleshko

Dr.Ph., Affiliate Professor Automated Systems Control Department, Institute of Computer Science and Information Technologies, Lviv Polytechnic National University.