[13] [Brusilovsky, 2004] Brusilovsky, P. A component-based distributed architecture for adaptive Web-based education. In: U. Hoppe, F. Vardejo and J. Kay (eds.) Proceedings of International Conference on Artificial Intelligence in Education: Shaping the Future of Learning through Intelligent Technologies (AI-ED'2004), Sydney, Australia, July 20-24, 2004. Amsterdam: OIS Press, 2004, 386-388.

[14] [Kernighan, Ritchie, 1988] Kernighan, B., Ritchie, D. C Programming Language (2nd Edition). Prentice Hall PTR; 2 edition: 1988.

[15] [Sosnovsky et al., 2004] Sosnovsky S., Brusilovsky P., Yudelson M. Supporting Adaptive Hypermedia Authors with Automated Content Indexing. In Lora Aroyo and Carlo Tasso (eds.) Proceedings of AH2004 Workshops: Part II /Second Workshop on Authoring Adaptive and Adaptable Educational Hypermedia/, Eindhoven, The Netherlands, 2004, 380-389.

[16] [Gaines, Shaw, 1995] Gaines, B. R. and Shaw, M. L. G. Concept maps as hypermedia components. International Journal of HumanComputer Studies, 43(3), 1995, pp. 323--361.

[17] [Brusilovsky, 2001] Brusilovsky, P. Adaptive hypermedia. User Modeling and User Adapted Interaction, Ten Year Anniversary Issue (Alfred Kobsa, ed.) 11 (1/2), 2001, 87-110.

Author's Information

Sergey Sosnovsky – University of Pittsburgh; 135, North Bellefield Street; Pittsburgh, PA, 15217, USA phone: +1 (412) 624-5513; e-mail: sas15@pitt.edu

Tatiana Gavrilova – Saint-Petersburg State Polytechnic University; Intelligent Computer Technologies Dept.; Politechnicheskaya, 29; Saint Petersburg - 195251, Russia; phone: + 7 (812) 550-40-73; e-mail: gavr@csa.ru

DEVELOPMENT OF PROCEDURES OF RECOGNITION OF OBJECTS WITH USAGE MULTISENSOR ONTOLOGY CONTROLLED INSTRUMENTAL COMPLEX

Alexander Palagin, Victor Peretyatko

Abstract: the ontological approach to structuring knowledge and the description of data domain of knowledge is considered. It is described tool ontology-controlled complex for research and developments of sensor systems. Some approaches to solution most frequently meeting tasks are considered for creation of the recognition procedures.

Keywords: the tool complex, methods of recognition, ontology.

ACM Classification Keywords: I.4.8 Scene Analysis – Object recognition; I.2.9 Robotics – Sensors

Introduction

One of the ambitious purposes of the world civilization is construction of the knowledge-oriented society. In computer science, a main priority direction thus is intellectualization of computer resources and technologies, in particular creation knowledge-oriented ontology controlled intelligence systems for various assignments. Information technologies on their basis are composing components of all high technologies. Except for usage in spheres of socioeconomic activity (the most difficult) spheres of research and development activity which result are objects of new knowledge, engineering, high technologies are rather important. The majority of these applications of intelligent systems is related to the problem solving, recognition (identifications), the diversification of their settings and implementation which are extremely various. The present material is devoted to usage of the

instrumental ontology-controlled complex for development of sensor systems, and in particular new (and refinement of already known) methods of recognition.

Productivity

Productivity is the most important parameter for the tasks related to recognition of signals data acquisition from external sources and their processing. The logical approach to the execution of these conditions is creation of tool complexes. Instrumental complex (IC) should unite in itself the block of interaction with an environment, the block of digital signal processing, and the block of interaction with the user. (Fig. 1)

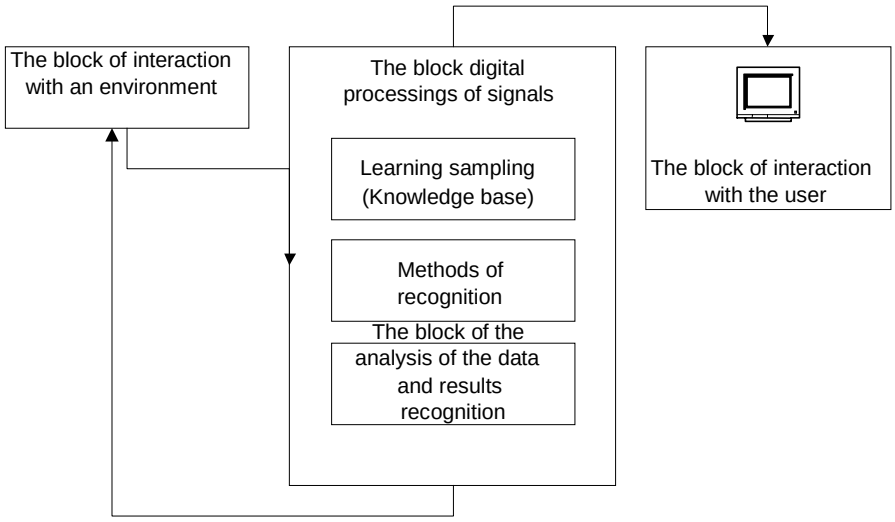


Fig. 1 Block-scheme of the tool complex.

The block of digital signal processing (DSP) contains the ontological component. Recently ontology's designing and knowledge processing become the object of steadfast notice of contributors of the various domains mainly ones operating in the knowledge engineering area. Among others, it is possible to mark such important directions of their interests:

- Knowledge management. To this directions such sections may be relevant as (intelligent) search, an automatic information accumulation from various sources (channels of news operating RSS), extract of knowledge from texts (Text mining) or sets of others - unstructured documents (text, databases, HTML, XML, etc.). The result of such analysis should become the generated document, which briefly formulates the major positions of the document, or groups of documents.
- Attempt to create advanced (system like ontology for WEB) [1]. Two key standards, which subsequently were used as a basis of the project named Semantic Web are completed. It is possible, that we are on a threshold of significant events' variations comparable with those, which have brought the Internet, World Wide Web and HTML. The given development - attempt to correct an ancient disadvantage of the Internet - its weak structure ability. New standards are Resource Definition Framework (RDF) and Web Ontology Language (OWL). They are the part of Semantic Web project and the main idea consists in making the information transmitted on the Network more clear, having provided possibility of identification, sorting and processing. Till now Web has been mainly oriented to operation of the person, but Web of the following generation, by opinion of developers of the project will be designed on the computer processing of the formation [1]. As a basis of the future WEB is assumed to be not only to search and read, but also to understand a contents of the Internet - information, and to reach it not through the creation of programs of the artificial intelligence,

simulating activity of the person, but through usage of resources of expression of semantics of the data and their links [2].

- Ontologies designing on the basis of available program systems which are reduced to filling special forms by the description of this or that data domain. Such developments are carried on more often in the research (educational) projects, for example [7].

Ontology

Ontology (O) as the formal description of the defined data domain can be represented as:

$$O = \{X, R, F\},$$

Where:

O - a finite set concepts (terms, quantum's of knowledge);

R - a finite set of the relations between concepts;

F - set of functions of interpretations of concepts and/or relations.

The set X is frequently bunched as subsets $X = \{X1, X2, \ldots Xq\}$, $q=1,2, \ldots, Q$ on subsumption to tags, making tree-like network structure. In a case $R \subset \varnothing$, $F \subset \varnothing$, ontology $O = \{X\}$ represents the usual dictionary - glossary.

Relation R - are served for association concepts in orderly structure (semantic model of a field of knowledge). Link is always unidirectional - one of the concepts is a grandparent, and another - the descendant. Link can define the ratio between quantity of grandparents and descendants 1:1, 1:N, N:1, N:M. Link can be hierarchical i.e. if A it is coupled with B, and B - with C and so on.

For effective scientific operation in this or that field of knowledge - it is necessary to fulfill the following standard procedures: to structure knowledge which concern both to a researched theme, and to adjacent areas, to study existing methods of researches, putting forward hypotheses and trying to check them on concrete examples to develop new methods.

Usually, the contributor fulfils all this, drawing up all logical chains as speak, in mind. In this case, designed in process of research operations ontology can help him in solution of a lot of the important tasks essentially. It is correctly and full constructed model of the field of knowledge can become the power factor for research and development designing operations. Even preliminary constructed variants of ontology give more complete "imbedding" in a theme researches (general domain). The description of objects and links connecting them will allow presenting more precisely processes, which occur in this or that system (a fragment of the system).

Ontology in the application to methods of recognition fulfils such functions:

- Formalistic - structuring of knowledges;
- Information- retrieval - carries out relevant navigation;
- Creative- generation of applications;
- Transforming- perfecting of the base methods of recognition;
- Extension- support of extended model computer ontology.

Computer Ontology

In [7] computer ontology is described which have been created on Lotus Notes platform (though at the given stage editors of ontology are developed, such as OntoEdit, OILed, Protege which give a graphic interface and create an output file according to standards of representation for Semantic Web). The program fulfilled on Lotus-Domino platform, is non-relational database which has the defined advantages: the convenient system of replication, the power 128 bit system of encoding, a built-in system of navigation, broadcasting of contents of base - programs in Web, etc.

The system is suitable for filling by the information from arbitrary fields of knowledges. In process of creation, the area of methods of recognition, which can be used, for construction of sensor systems has been selected.

The program will consist of the following blocks:

Categories - the description of categories which concern to a given theme,
Applications – examples of usages (demonstrations) of operations of different kind,
Navigation - relevant navigation by category (both through headers, and internal content),
Glossary – description of general scientific terms of ontology,
The library - storage of necessary files,
Diary - clone of an organizer. Allows to plan the operations and to bring various arbitrary records,
Help for the user.

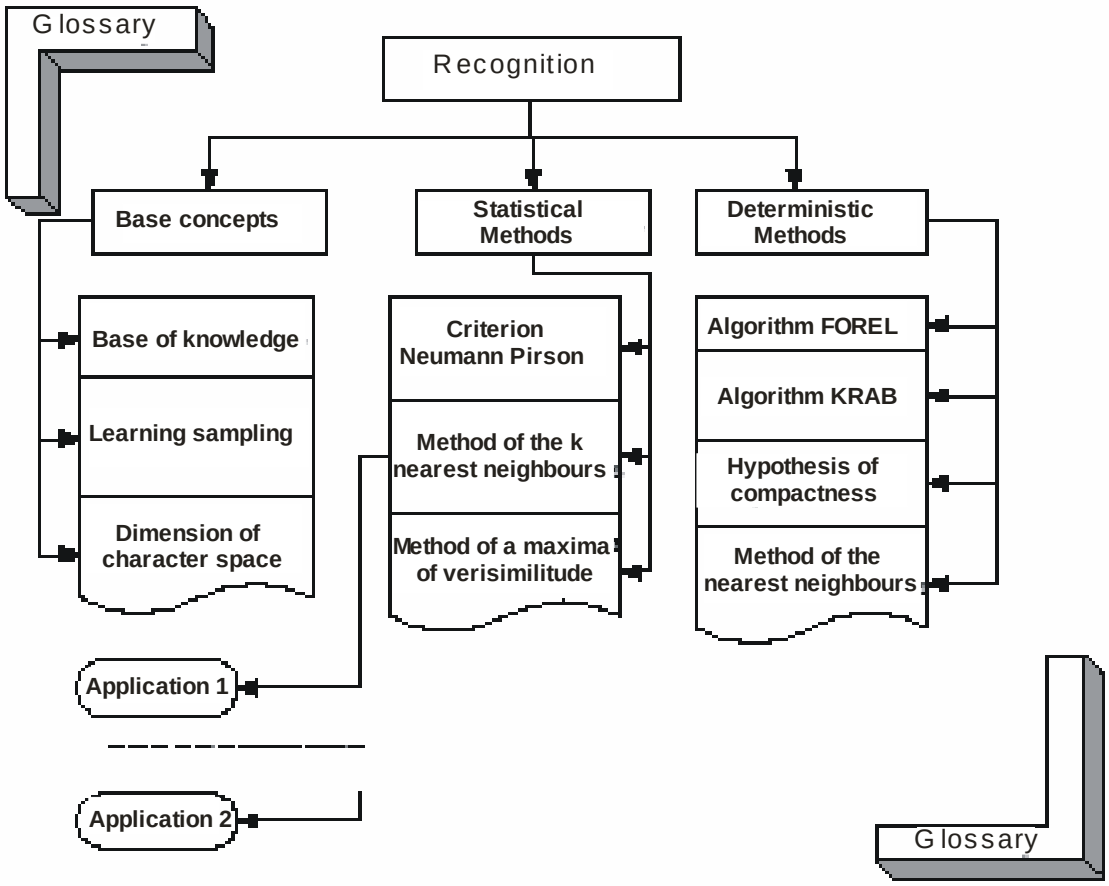


Fig. 2 Scheme ontology

Ontology represents an outline (see fig. 2). The theme ontology "Recognition" has three main categories: base concepts, statistical methods and deterministic methods. Inside these subsections also are concepts - quantum's of knowledge. It is necessary to mark that each of concepts can be referred to more than one subsection. For example, some methods of recognition are simultaneously statistical and deterministic ones. It is possible to connect the derived object of the application in which practical application of some concept is described in parent concepts to each quantum of knowledge. Besides, from anyone concept the reference to arbitrary others concepts, units of the application or a glossary can be created.

For each of concepts, (for each quantum of knowledge) it is underlined:

- Theme;
- The type of the message (definition, an explanation, an example, etc.)
- Category (in a case with methods of recognition - the class of a methods);
- Keywords;
- Weight coefficient of concept (on a five-point scale);
- The language of the message (Russian, English);
- Source.

More detailed description of ontological components of the complex is resulted in [7].

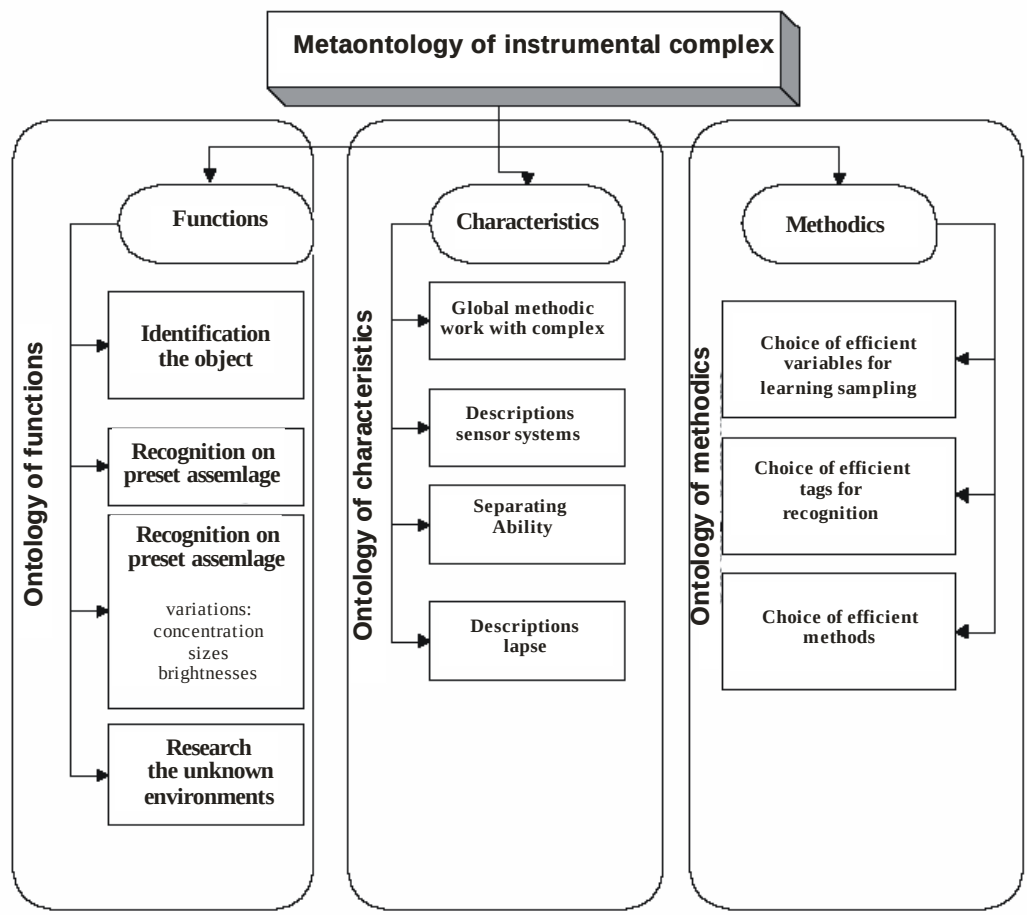


Fig. 3 Functional diagram instrumental complex

The modern level of development of the methods of a multivariate statistical analysis allows to carry out classification of objects on a wide and objective basis, in view of all essential structural - typological tags and characters of object allocation in the preset system of tags.

At logical level IC should contain such descriptions:

- Common procedures of operation with the complex for various conditions;
- Sets of characteristics of the block of interaction with an environment (a set of sensor controls, procedures of operation with them, errors, separating ability for various sorts of defined objects);
- A method (or several methods) of recognition;
- Procedures of choice of effective variables, for samplings;
- Procedures of choice (automatic or hand-held) of effective methods for operation with each concrete object or group of objects (fig. 3)

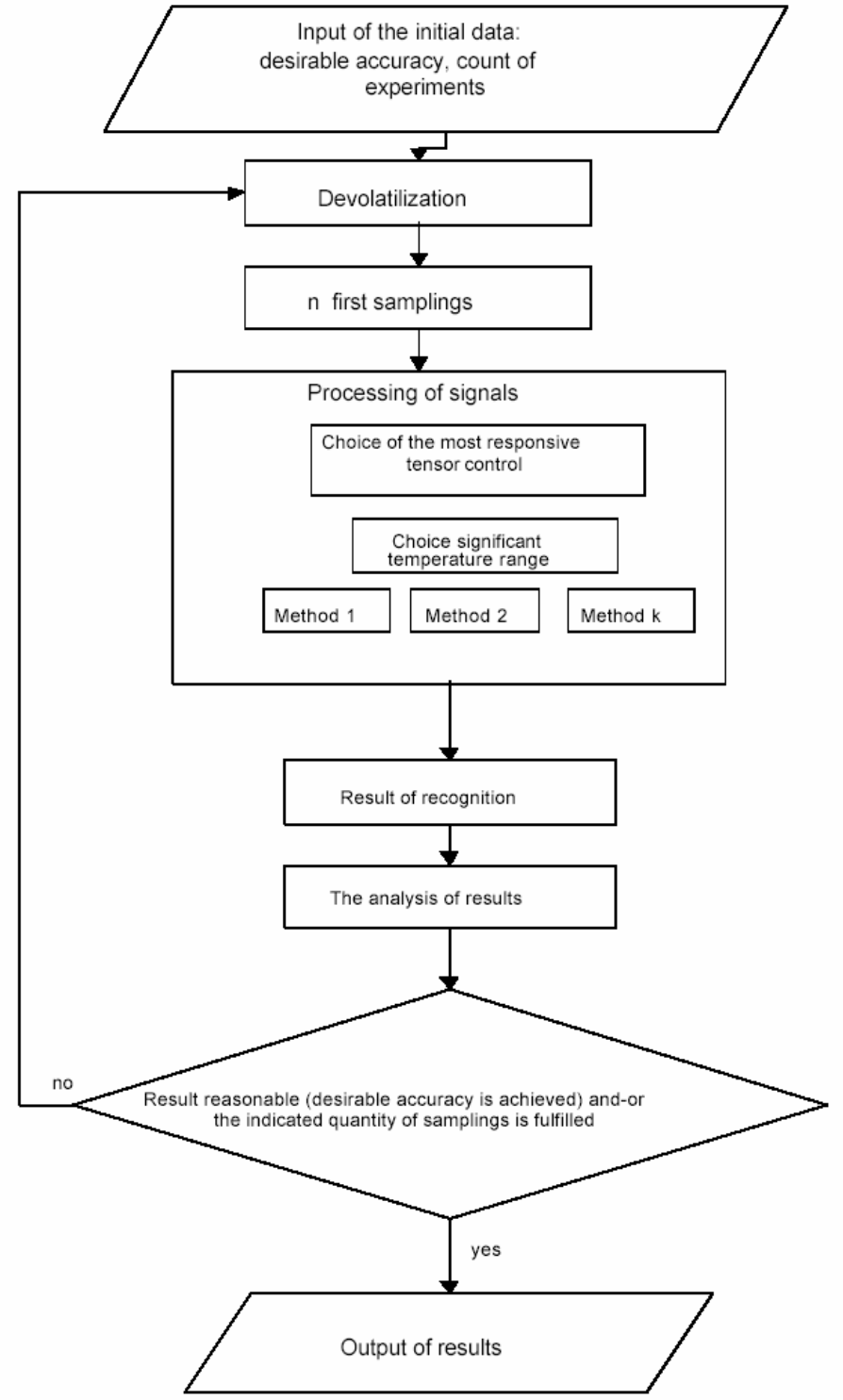


Fig. 4 Procedure of sampling

This tool complex is intended, in particular, for research of the gaseous environments, and coupled to a set of the sensor controls included in the projected instrument (system). Sampling procedure consists in removal of metrics from sensor controls at heating. The scheme of the procedure of removal of the data with the subsequent processing is represented on fig. 4.

Classification of Multivariate Objects

There are many methods of classification of multivariate objects with the help of a computer. Methods of the first group are connected with task of "recognition", identifications of objects and have received the name of methods of pattern recognition. The sense of recognition consists in that any object showed to the computer with the least probability of an error has been referred to one of beforehand-generated classes. Here to the computer all over again show a taught sequence of objects (about each it is known to what class or "image" ' it belongs), and then, "having trained", the computer should recognize to what classes from the investigated collection the proposed object is belonging.

More common approach to the classification includes not only reference of objects to one of classes, but also simultaneous creation of "images", the number of which can be beforehand unknown. At absence of a taught sequence such classification is made on the basis of tendency to collect in one group somewhat similar objects moreover so that objects from different groups (classes) were whenever not similar. Such methods have received the name of methods of automatic classification.

Now tens and hundreds of the various algorithms realizing multivariate classification automatically are developed. They are based on various hypotheses about character of allocation of objects in multivariate space of tags, on various mathematical procedures. Browses of these methods widely represented in the literature.

Various requirements are characteristic for various types and procedures of recognition of the objects to the measured data. In general, *choice given* (variables for recognition) is the most challenge in all chain of the operations coupled to recognition. Thus, it is possible to speak about two types of choice:

- About choice of the measured data: the most sensitive for these objects a range, periodicity, etc., that is those aspects which are defined *by a procedure of samplings* (and, accordingly, can be modified at this stage);
- About choice of variables – tags for *recognition.*

The Requirement of Independence

The requirement of independence of tags is typical of the majority of methods of recognition. The requirement is logically proved: if the data are dependent, the information contained in one tag, already is presented at the learning sampling, and in other measured variables the method of recognition can "tangle" its repetitions only. For example, for method Bayes (posterior probability) this requirement is extremely strict and mandatory. For other methods, this rule can be neglected. To such ones the methods based on clusterizations, "cognizance" objects (a method to the nearest units, a method of measurement standards) are concerned.

However, the independence of tags can be frequently guaranteed by the very sensor subsystem. For support of independence of tags pertinently takes advantage the check of correlation tags.

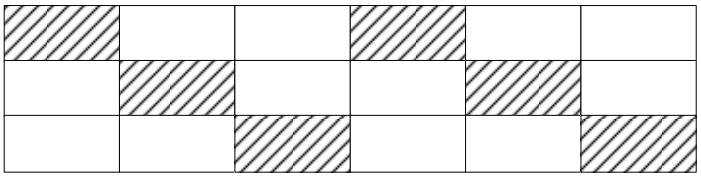
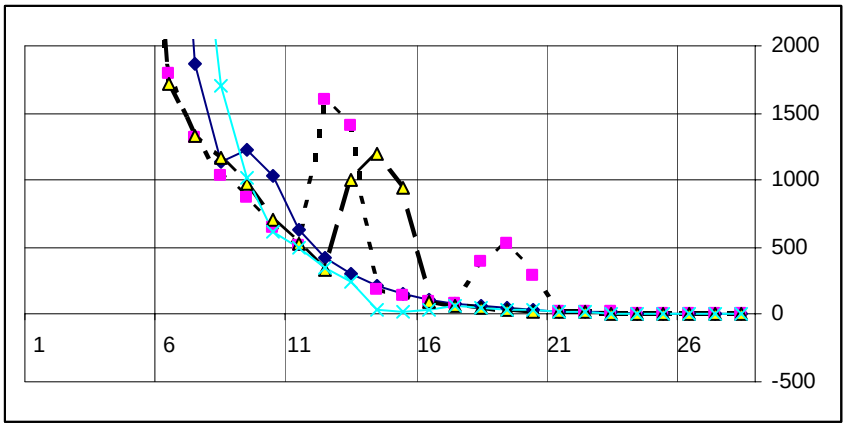


Fig. 5 Scheme of division of an interval of sampling on 6 parts, for definition of minimum correlation

For example, for obtaining an independent data set from 3 samplings the same process (object) it is possible to divide the data into intervals, and to check up on correlation each of intervals. Then the most correlated intervals are deleted from learning sampling (fig. 5.). Thus, it is possible to investigate, for definition of optimal quantity of parts. It is possible also, for the analysis of a digital data to make the various sorts of the intermediate

conversions, for more visual data representation, and, more convenient extract of the information from these data. For example, at presence of a plenty of the schedules constructed in one range, it is possible to construct a matrix in which to point that quantity of schedules which passes through each of coordinate squares. This information (fig. 6) can be used for development and modification of existing methods of recognition. In particular, it carries the information on potential possibilities of definition of measured objects on various intervals of a scale of argument.



4	1	0	0	0
4	2	0	0	0
4	4	1	0	0
0	4	4	4	4

Fig. 6 Schedules of the data, and a matrix of quantities

Conclusion

Considered tool ontology- controlled complex is the tool for creation and learning of procedures of identification of objects. Some approaches are resulted in creation of new procedures of the analysis and direct recognition.

Bibliography

1. [T. Berners-Lee, 1999]. T. Berners-Lee Weaving the Web, Harper, San Francisco, 1999 P. 2-13
2. [T. Berners-Lee, 2001]. T. Berners-Lee. -"Semantic Web Road map ", http://www.w3.org/DesignIssues/Semantic.html
3. [Stin Deker, 2001] Stin Deker, Sergey Melnik, Franc van Hermelen, etc. - " Semantic Web: roles XML and RDF ", "Open systems", #9 2001 of page 41-70;
4. [Fu K.S., 1977] Fu K.S. Structural methods in pattern recognition. - M.: the World, 1977, page 12-15
5. [Voloshin G.Ya., 2000] Voloshin G.Ya. Method of pattern recognition (the abstract of lectures) Page 9, 36-40, 42-43
6. [Gorelik A.L., 1977] Gorelik A.L., Skripkin V.A. Method of recognition. - M.: Vissh. shk., 1977, page 37-45
7. [Palagin A.V.,2005] Palagin A.V., Peretyatko V.J. Tool ontology controlled complex for research and developments of touch systems. UsiM.-N 2, 2005.

Authors' Information

Alexander Palagin - Deputy director, professor. Institute of Cybernetics, NASU, Kiev, Ukraine e-mail: palagin_a@ukr.net

Victor Peretyatko - postgraduate student of the Institute of Cybernetics, NASU, Timoshenko st., apt 3-a, 12, Kiev, Ukraine, e-mail: victor@iapm.edu.ua