Visualization Management of the Spatial Data Containing Uncertainty

Stanislav Belyakov  
Southern Federal University  
Taganrog, Russia  
beliacov@yandex.ru

Marina Belyakova  
Southern Federal University  
Taganrog, Russia  
mibeliacova@yandex.ru

Marina Savelyeva  
Southern Federal University  
Taganrog, Russia  
msavelyeva@sfedu.ru

Sergey Zubkov  
Southern Federal University  
Taganrog, Russia  
zubkov@intech.ru

Abstract—The paper is devoted to the development of a method of visualization management of the spatial data on electronic maps. It is assumed that spatial data are obtained from sources that didn’t allow to define fully and consistently all characteristics of objects or events. There are defects of display in the visual analysis of the maps containing cartographic objects with uncertainty. The method of visualization management is proposed, which allows to control the utility of visual representation of the map with defects. The special presentation of the analysis workspace is introduced, the structure of expert rules is described, allowing to realize the selection of information from the geographic information system database to reduce redundancy and ensure the integrity of the picture. The form of the utility function is defined. The restriction of perception is introduced, which is determined by the complexity of the analysis workspace. The complexity measure is the number of cartographic objects. The algorithm of visualization management is given.

Keywords—visualization management, spatial data, uncertainty, electronic maps.

I. INTRODUCTION

The visualization of spatial data by geoinformation systems (GIS) plays a big role in their practical use. Visual analysis of maps, schemes and plans allows users to solve especially difficult and informal tasks, stimulating from thinking by cartographic images on the screen. Dialogue process of studying by a user-analyst of a specific visual object is realized in a complex with software tools of spatial, statistical, topological and other special types of analysis, i.e. cartographically presented workspace of a common GIS map. Dialogue study of the workspace is directly or indirectly related to decision-making. The analyst extracts the necessary data and knowledge for evaluation and selection of alternatives, having a set of alternatives to the solution of the task. The quality of the extracted data markedly affects the value of the obtained solution. The important role at data retrieval is played by the organization of the visualization process. GISs provide a set of software picture management tools, which include zooming, panning, changing views, 3D-fly-by, controlling views, map overlay, the ability to change the cartographic representation of the same data. It can be argued that the analyst uses the listed tools to improve the perception on the screen of complex cartographic expressions. The perception isn’t meant a simple distinction of graphic objects by the human eye. The term "perception" is contacted with a higher level of representation of the cartographic expression in the analyst's consciousness. This view has the form of a holistic image, which includes a set of spatial entities and their relations. The image isn’t formed instantly, but during of visual analysis of the workspace. The user's dialogue with GIS, therefore, is an instrument of providing a level of perception sufficient for the formation of the image of the workspace of the map.

Modern means of communication and collection of spatial data provide new opportunities of the analysis of situations using GIS. This refers to spatial data obtained by indirect measurement. These include photo and video materials processed by the systems of recognition, texts generated in news streams and network communities of social networks. The spatial coordinates of objects and events, in this case, are calculated based on a number of assumptions and estimates. The peculiarity of such spatial data is incompleteness, inaccuracy, uncertainty and inconsistency. Their mapping is problematic because of the high cost and long deadlines for cartographic work. At the same time, the value of this type of spatial data can be so high that even with an incorrect display on the map, it will improve the quality of the developed solutions. For example, the display of the approximate river flood zone due to heavy rains, which was observed a few days ago, but not indicated at the moment on the map, can significantly correct the transportation route for large-capacity cargo.

II. REVIEW OF METHODS OF VISUALIZATION MANAGEMENT OF SPATIAL DATA

Cartographic visualization is the basis of visualization methods of spatial data [1,2]. The greatest attention in the in researches of this direction is given to the design of the
workspace of analysis. The task consists in selection of the initial components for designing: maps, schemes, databases and multimedia documents, as well as previously constructed thematic maps. The search and selection of initial cartographic materials should be completed by constructing a thematic map that provides the actual spatial data of the required accuracy for the decision of the application task. Further manipulation of the picture of the workspace in works on cartographic visualization isn’t investigated, since the functionality of the GIS toolkit is accepted as sufficient. Such an emphasis in the visualization management is reasonable at the stage of preparation of the workspace. However, the dialogue control becomes more important as increase in complexity of the map i.e. the number of cartographic objects, classes, layers, views, links, etc. To form a single cartographic image, the analyst should do a lot of routine work on picture management. It negatively affects the formation of the image. In addition, the originally created workspace is often modified during the analysis. Therefore, the method of the dialog control becomes important at the stage of creation of the workspace.

The problems of mapping of objects and events in real time are investigated in neocartography [3]. The problems are related to the fact that modern means of communication give the chance to obtain cartographic data from nonprofessional people, to link professionally created maps with information objects of any nature, to use camera records instead of maps, which has generated new technologies for visualization of maps. These technologies are also explored within the direction called cybercartography [4,5]. Researches concentrate mainly on filtration of spatial data and models for representing inconsistent data. The methods of using maps containing inaccurate and not completely defined objects remain low-investigated.

Researches of the user’s interface with GIS are carried for many years [5]. Object of research is the perception of cartographic pictures in the sense indicated above. However, if in earlier works the problems of visualization with strict limitations on the speed and volume of memory were considered, modern researches changed their context. The problems of information flow management of on the network [6] and knowledge extraction from visualized data [7] come to the fore. The analysis of the works of this direction shows the need to study the mechanism of management of cartographic information flow, and the use for this purpose of knowledge.

The problem of limited resources for the visualization of spatial data has acquired a certain specificity relating to the use of network cartographic services by mobile devices. In the works of this direction, the problems of optimizing the client interaction of geo-service for solving certain classes of applied problems [8]. The dynamic aspect of interaction is researched using special data models, the implementation of which on the server side will allow the required quality of interaction. In this case, the limitation of perception is considered from the standpoint of psychophysical perception of information from the screen of the mobile device.

Analysis of publications on research related to scientific visualization [10,11] shows the importance of the problem of a holistic perception of the visualized data. The proposed solutions affect different aspects, beginning with the possibility of differently to see the same data, ending with the introduction of a special set of metaphors for displaying data and knowledge. It is noteworthy that in [10] such metaphors became characteristic for GIS objects and relations of urban development of the real world.

In the works of the authors [12,13], an approach to visualization is developed, based on a special representation of the structure of the workspace of analysis and the use of knowledge to construct informative cartographic pictures. The proposed model can be modified for the case of using uncertain elements. This will allow to implement the process of maximizing the utility of visualization of the workspace from a unified position.

III. PRINCIPLES OF MAP VISUALIZATION MANAGEMENT WITH UNCERTAIN ELEMENTS

Consider the features of the process of using cartographic pictures in the analyst's dialogue with GIS, the influence on the dialogue of objects containing uncertainty, and formulate the principles of visualization management.

Objects containing uncertainty, we mean cartographic objects whose form and position aren’t determined by the values of their properties. For brevity, we will further call to such objects as u-objects. Any u-object is an instance of one of the GIS object classes. We will assume that each class includes the Uncertain property of type Boolean, which takes the value True if at least one property of the instance has inaccurate, uncertain, incorrect, insufficiently reliable, etc. value [14]. True value is established by GIS procedures that construct instances based on incomplete data. The property of uncertainty is considered by the visualization system, which displays similar objects specific way to particular GIS. The form, color, texture or special mnemonic designation is used for this.

The typical scenario for solving the applied problem using GIS cartographic materials includes the following steps:

1) the user-analyst builds the workspace of the map for the study. To do this, through a set of dialog menus, he forms a set of requests to the GIS server for selection of fragments of thematic maps, layers, classes and instances of cartographic objects, raster camera records. Thanks to these actions, the analyst forms a mental image necessary for generation of alternatives and choice of the solution to the problem;

2) the generated workspace is processed to visual analysis. From the psychological point of view, [15] is an examination of an object that complements and modifies its existing mental image. Visual analysis includes the invocation of functions of spatial, topological, statistical and other types of analysis;
3) if the semantic content of the workspace isn’t enough, the analyst modifies it, adding new and deleting the studied components. Then go to step 2. Otherwise, the formation of the workspace is completed by decision-making for the application task.

In general, the dialogue can be described as follows. There is the GIS database with a lot of cartographic objects \( \Omega = \{o_1,o_2,\ldots,o_k\} \). The analyst forms set of queries \( Q_i(X_S,X_T,X_C),i=0,N \), in the query language for the GIS database, where \( X_S \) is set of spatial query parameters, \( X_T \) is set of temporal query parameters, \( X_C \) is set of semantic query parameters. Query responses are redundant, which forces the analyst to apply knowledge about the importance of instances, object classes, relations, and to perform auxiliary actions to reduce redundancy and preserve the integrity of the workspace. This solves the problem

\[
\begin{align*}
I(m_W) & \rightarrow \text{max}, \\
C(m_W) & \leq C^*, \\
m_W : Q(X_S,X_T,X_C) = true, i = 1,N.
\end{align*}
\]

Here \( I(m_W) \) is utility function, \( C(m_W) \) is function of laborious analysis of picture \( m_W \). The function \( C(m_W) \) determines the number of auxiliary operations necessary to eliminate redundancy in the workspace. \( C^* \) is the restriction reflecting the influence of secondary actions on the quality of the mental image. The this number is higher, the quality of the image is lower.

Analysis of the model (1) allows us to draw the following conclusions.

Using u-object in the process of solving the problem (1) increases the number of auxiliary operations, because u-objects break the integrity of the workspace. Since the function type \( C^* \) isn’t known, the task is solved by applying the analyst's knowledge of complete cartographic expression.

The restriction of the number of auxiliary operations \( C^* \) is subjective and is determined by the psycho-physiological features of the analyst's knowledge of cartographic expression. The presence in the expression of u-object doesn’t change \( C^* \), since any of them is an element of the cartographic expression.

The utility function \( I(m_W) \) is determined by the perception of the picture of the workspace, depending on its complexity. As the complexity of the picture increases, its utility decreases. At the same time, the utility of simple pictures is small due to the lack of meaningful content in them. Therefore, the utility function must have a maximum at some level of complexity of the workspace.

The presence or absence of u-object in the workspace can’t be considered as a factor that changes the complexity of the workspace. The main role is played by the holistic perception of the picture, which doesn’t depend on the concrete type of objects. As a result, it is impossible to talk about increasing the utility of the picture. The effect of emergence of u-object is shown in direct impact on a mental image of the situation.

The modern development of psychophysiology doesn’t allow us to explicitly link the parameters of the mental image with the solution of the problem (1), but it can be argued that the technical support of the dialogue to reduce the number of auxiliary operations \( C(m_W) \) increases the quality of the mental image. Consequently, the probability of the effective analyst’s decisions-making increases.

The obvious problem of reducing \( C(m_W) \) is the subjectivity of perception and thinking of the analyst. Nevertheless, it can be assumed that there is a "reasonable" way of selection of information useful for analysis, which can be reproduced programmatically in the GIS environment. This assumption is proved by the following:

- there are classes of applied problems that are solved by specialists of a certain profile with the known set of competencies. The thematic content of the spatial data necessary for them has sufficient certainty to not consider this content unique for each user individually. Professional groups of analysts can describe the properties of useful information for decision-making;
- specialists have knowledge about "reasonable" selection of information from the GIS database for their activities. Such knowledge can be extracted and reused.

Thus, the principles of visualization management, which implements GIS, can be formulated as follows:

1. The purpose of visualization management is to achieve the subjective utility of the picture for the professional group of analysts, who solves similar problems.
2. Knowledge should be used to form complete non-redundant cartographic expressions.
3. The basis of the semantic content of the workspace is the sequence of analyst requests to the GIS database. Many replies are redundant and must be processed to increase the level of the semantic content of the workspace.
4. The utility function should depend on the complexity of the workspace, but not on its content.
5. U-objects in the workspace directly affect the mental image in the analyst's mind, change the structure of the workspace, but don’t affect its utility.

IV. UTILITY FUNCTION OF THE MAP WORKSPACE

The utility of the cartographic representation of the workspace is shown that the analyst can perceive and
comprehend it without losing concentration because of the need to perform auxiliary visualization operations. The utility function depends on the complexity of the workspace and a number of factors that reflect the physiology of perception, training, the experience of the analyst, and some others. We will assume that the sole argument of the utility function is the complexity of the workspace, since accounting for any other factors will unduly complicate the task. Complexity measure is number $|m_F|$ of cartographical objects in workspace. Note that this measure is local for the current session of the analyst's work with the GIS. It means that it isn’t intended for use in different sessions by the same user or different users.

The assumed measure allows to consider integrally the subjectivity of the utility function through the limiting for perception number of cartographic objects $m^*$ that the user sets as GIS configuration parameter. In essence, $m^*$ is the average number of cartographical objects, the excess of which leads to a decrease in the level of perception of the picture in general. For values $|m_F| < m^*$, the perception level doesn’t decrease, however, pictures with a small number of elements aren’t of interest for visual analysis. It is unlikely that a small number of cartographic elements received at the request displays a situation of the real world somehow deeply and meaningfully. Therefore, we can say that in the formal sense the point $m^*$ corresponds to the maximum of the utility function.

Thus, the utility function can be represented in the form

$$ I(m_w^*) = \begin{cases} 1 - (m^* - |m_F|)/m^*, & |m_F| \leq m^*, \\ 1 - (|m_F| - m^*)/m_{\text{max}}, & |m_F| > m_{\text{max}}. \end{cases} \tag{2} $$

Here $m_{\text{max}}$ is the maximum possible number of picture elements determined by the properties of the display device.

It can be seen from (2) that the maximum utility value $I(m_w^*) \leq 1$ is reached at the point $m^*$, which is the unique maximum point. Deviation from $m^*$ in the greater or lesser side leads to a decrease in utility and $I(m_w^*) \to 0$.

V. STRUCTURE OF THE WORKSPACE WITH DEFECTS

Considering the concept of u-objects, introduced above, we represent the structure of the workspace in the form

$$ m_w = B \cup E \cup B^+ \cup B^-, \tag{3} $$

where set $B$ is the skeleton of workspace containing the objects received in response to queries $Q_i(X_S, X_T, X_C), i = 0, N$; set $E$ is the environment of the skeleton from objects that complement the semantic content of the skeleton for giving meaning to the picture; set $B^+$ is u-objects that don’t generate picture defects; set $B^-$ is which have become the reason of defects.

The selection of the set $E$ reflects the peculiarity of cartographic visualization, which consists in the fact that the GIS users’ requests don’t describe in detail the boundary and elements of space, time and semantics, which form the holistic view of the real world. The user is forced to create missing "details" by means help of additional requests. In fact, we consider such requests as auxiliary, and the reduction of their number plays a major role in the quality of visual analysis. To minimize the number of auxiliary requests, it is suggested to enter the GIS procedure for creating the environment in any request. The procedure should be based on knowledge of non-redundant and holistic cartographic expressions.

The environment $E$ is formed by applying expert rules $K(B, \Omega)$ constructing the response image to the skeleton of the query $B$:

$$ \omega_i \in E \Rightarrow \exists k_m(X_S, X_T, X_C) \subseteq K(B, \Omega): k_m = \text{true}, m \in [1, |K(B, \Omega)|], i \in [1, |m_w|]. \tag{4} $$

Thus, the utility function can be represented in the form

$$ \omega_i \in E \Rightarrow \exists k_m(X_S, X_T, X_C) \subseteq K(B, \Omega): k_m = \text{true}, m \in [1, |K(B, \Omega)|], i \in [1, |m_w|]. $$. 

Expert rules $K(B, \Omega)$ display knowledge of how the boundaries of the workspace $m_w$ for the given skeleton $B$ are constructed and how the received area is filled with database objects. Each rule $k_m(X_S, X_T, X_C)$ depends on the parameters of the cartographic objects, with a certain combination of values of which the program procedures for determining the boundaries of the workspace and filling it with objects are performed. The rules structure corresponds to one of two types:

- rules for defining the boundaries of the workspace

$\text{IF} \quad \text{Correspond}(\text{Parameters}, (B), \text{SetBenchmarkParameters}) \quad \text{THAT} \quad \text{SkeletonBoundary} = \text{ConstructBoundaryWay}(i).$

$\text{IF} \quad \text{Correspond}(\text{Parameters}, (B), \text{SetBenchmarkParameters}) \quad \text{THAT} \quad \text{SkeletonBoundary} = \text{ConstructBoundaryWay}(i).$

$\text{IF} \quad \text{Correspond}(\text{Parameters}, (B), \text{SetBenchmarkParameters}) \quad \text{THAT} \quad \text{SkeletonBoundary} = \text{ConstructBoundaryWay}(i).$

Here, the logical function $\text{Correspond}(X, Y)$ is true if the set of parameters of the existing skeleton $X$ in a given sense corresponds to a set of benchmark values of the parameters $Y$. The correspondence is determined by a fuzzy classification of parameter values. Each class is defined by a set of fuzzy restrictions on the values of the parameters. Function $\text{Parameters}, (B)$ forms a set of skeleton parameters for the i rule. Function $\text{ConstructBoundaryWay}(i)$ forms the skeleton boundaries, if its parameters correspond to the logic of the i rule. Variable $\text{SkeletonBoundary}$ stores the value of the computed boundaries or an undefined value null, if none of the rules in the knowledge base allowed to calculate boundaries;

- rules for selecting environment objects

$\text{IF} \quad \text{SkeletonBoundary} <> \text{null} \& \text{WayOfConstructBoundary} = i \quad \text{THAT} \quad E = \text{ConstructEnvironmentWay} (\text{SkeletonBoundary}, i).$

$\text{IF} \quad \text{SkeletonBoundary} <> \text{null} \& \text{WayOfConstructBoundary} = i \quad \text{THAT} \quad E = \text{ConstructEnvironmentWay} (\text{SkeletonBoundary}, i).$

The function $\text{ConstructEnvironmentWay} (...)$ builds a set $E$ for the given value of $\text{SkeletonBoundary}$ depending on the class which number is set by value $i$. 

43
Applying rules $K(B, \Omega)$ leads to a reduction in the redundancy of pictures, but doesn’t finally determine the utility of the picture for analysis. Here, the level of perception of the picture must be considered, which is expressed by the equality $|m_W| = \alpha$. Therefore, if $|m_W| > \alpha$, a part of the objects must be deleted. This part should be minimally significant for analysis. To implement this restriction, the expert rules $(X, X, X, K)$ are used to sort objects by preference, where $ID$ is identifier of a professional group of analysts. The set of rules is organized in such the way that

\[
\omega_a \neq \omega_b \Leftrightarrow \exists r_j \in K_s(X_S, X_T, X_C, ID) : r_j = true, \\
\neg \omega_a \neq \omega_b \Leftrightarrow \forall r_j \in K_s(X_S, X_T, X_C, ID) : r_j = false.
\]  

(5)

Thus, using rules, the binary function $S(\omega_a, \omega_b)$, used by any sorting algorithm:

\[
S(\omega_a, \omega_b) = \begin{cases} 
true, & \omega_a \neq \omega_b, \\
false, & \neg \omega_a \neq \omega_b.
\end{cases}
\]

(6)

The task of visualization management, thus, acquires the following form:

\[
I(m_W) \rightarrow max, \\
|m_W| = \alpha, \\
\omega_t \in E \Rightarrow \exists k_m(X_S, X_T, X_C) \in K(B, \Omega), \\
k_m = true, m \in [1, |K(B, \Omega)|], i \in [1, |m_W|], \\
\omega_t \in B : Q_j(X_S, X_T, X_C, X_E) = true, j = 1, N.
\]

(7)

GIS implements task (7) as information support of the solution of the problem (1). The generalized solution algorithm (7) has the form:

1) current analyst’s request is processed by the GIS server, and the selected objects are added to either the current skeleton set, to the skeleton u-object set, or to the set of skeleton defects;
2) environment for the current skeleton is formed;
3) objects in the workspace are ordered on the importance;
4) the subset of the least significant objects is removed from the workspace to provide the best level of perception, which corresponds to the maximum of the utility function;
5) workspace is visualized.

Analysis (7) allows us to note the following. Firstly, the structure of the workspace (3) makes it possible to process in different ways the constructed cartographic objects from a single position. The existence of defects will be displayed in a special way, but doesn’t require a radical restructuring of the visualization process in general.

Secondly, the representation (3) creates an additional opportunity to diversify the visual analysis by changing the "semantic angle" of the picture. The essence of such operation is as follows: sets from (3) can be combined, expert rules can be applied to them in different ways, which gives the specific visual image. We assume that

\[
m_W = B \cup E \cup B^+ \cup B^- = B^* \cup E^*, \\
B^* = B \cup B^+ \cup B^-, E^* = E.
\]

(8)

The inclusion of objects on request in skeleton, u-objects with defects and without them generates the most useful picture for generating the set of solution alternatives. Thanks to existence of u-objects, the variety of possible space-time and semantic situations increases.

If you create a visualization with the representation

\[
B^* = B \cup B^+, E^* = E,
\]

(9)

in the picture, plausible-looking u-objects will be displayed in the context of the solvable task. Such picture is useful in assessing the risk of decisions that must be implemented quickly and soon.

When visualization, when

\[
B^* = B^+ \cup B^-, E^* = E,
\]

(10)

the picture is constructed, containing u-objects in the context of the solvable task. If these are recent events, then this semantic angle is useful in the risk analysis of previously implemented projects in the workspace.

If representation is used

\[
B^* = B, E^* = B^\cup B^+ \cup E,
\]

(11)

then its visualization corresponds to the position of the analyst prone to caution in decision-making. All u-objects in the semantic sense only supplement the general picture of the situation.

VI. CONCLUSION

The main area of application of the proposed method of visualization management are GIS, used for decision-making. The main effect of using the method is improvement of the quality of the made decisions thanks to providing the
user-analyst with non-redundant and complete cartographic expressions. As shown by the experimental study of the software model of the system, the greatest interest to it is shown by analysts, who solves the tasks of operational planning. Visualization of the most recent information about objects and events gives the chance to reduce the risk of incorrect decision-making, even if the display isn’t absolutely correct and the parameters are partially uncertain.

The effectiveness of the application of the method also increases as the information base of the GIS becomes more complex. The majority of GIS are focused on the continuous accumulation of data [1] and the growth factor of the database begins to seriously affect the ability to analyze the accumulated information. It was determined experimentally that this effect is noticeable when the number of objects is more than several thousand, when the number of layers becomes more than 30. The boundary of perception of workspaces by different analysts was adjusted from 50 to 500 objects, that is several orders lower than complexity of the general GIS map.

The problem of using objects with uncertain parameters belongs to the fundamental problems of GIS [1]. The proposed method doesn’t solve the problem of cartographic display such objects, but allows to regulate the flow of cartographic information, enabling the analyst to subjectively assess an uncertain situation with the use of a special visualization tool.

The proposed method is based on knowledge, the quality of which is defining in assessment of the GIS behavior. This feature is typical for all systems that use knowledge. Therefore, further research should be focused on developing effective methods for extracting and using knowledge about the formation of non-redundant and complete workspaces of analysis.

ACKNOWLEDGMENT

This work has been supported by the Russian Foundation for Basic Research, Project № 17-01-00119 and № 17-01-00060.

REFERENCES