Chin. Geogra. Sci. 2010 20(1) 018–022
DOI: 10.1007/s11769-010-0018-z

Fundamental Theories of Spatial Similarity Relations in Multi-scale Map Spaces

YAN Haowen[1, 2]
(1. School of Mathematics, Physics and Software Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China;
2. School of Geographic Science and Planning, Sun Yat-Sen University, Guangzhou 510275, China)

Abstract: Similarity relation is one of the spatial relations in the community of geographic information science and cartography. It is widely used in the retrieval of spatial databases, the recognition of spatial objects from images, and the description of spatial features on maps. However, little achievements have been made for it by far. In this paper, spatial similarity relation was put forward with the introduction of automated map generalization in the construction of multi-scale map databases; then the definition of spatial similarity relations was presented based on set theory, the concept of spatial similarity degree was given, and the characteristics of spatial similarity were discussed in detail, including reflexivity, symmetry, non-transitivity, self-similarity in multi-scale spaces, and scale-dependence. Finally a classification system for spatial similarity relations in multi-scale map spaces was addressed. This research may be useful to automated map generalization, spatial similarity retrieval and spatial reasoning.

Keywords: similarity relation; spatial relation; multi-scale map spaces

1 Introduction

Multi-scale map databases are the fundamental contents in the national spatial data infrastructure (NSDI), which provides geographical spatial positioning bases for the society in the fields of politics, economy, military, environment protection, traffic and transportation, telecommunication, *etc*. Nowadays, map databases of NSDIs in many nations are constructed using the method of "one database, multiple versions", i.e., maps at each scale are digitized and saved in separated map databases. The disadvantages of this method are obvious: 1) repeated construction of databases leads to the redundant data; 2) the consistency of the databases can not be ensured; 3) the renewal of the databases is not easy; 4) the transmission of the data through the Internet becomes difficult; and 5) the acquirement of an arbitrary scale map data is difficult (Wang, 1993). The ideal method for the construction of multi-scale map databases is "one database, one version", i.e., only one larger scale map database is constructed for one region and the other smaller scale map databases are generated from the larger one by means of automated map generalization. Nevertheless, employing "one database, one version" method in the construction of NSDI is still a dream to cartographers, for many theoretical and technical problems in automated map generalization are not solved yet, due to the complexity of representation of geographic information on multi-scale maps and the uncertainty of traditional manual map generalization procedures, which leads to the difficulty of the computer-acceptable formal description of map generalization process. Automated quality assessment of generalized maps is one of the problems that have puzzled researchers and cartographers and cumbered the realization of automated map generalization for years.

The quality of generalized maps was often assessed empirically by means of "comparing the similarity degree between the original map and the resultant map" by cartographers and map inspectors in the time of manual map generalization. As different cartographers produce different maps using the same original map, the cartographers and inspectors need to determine whether the resultant map coincides with the one in their brains, i.e., they need to evaluate the similarity degree between the resultant map and an "imaginary" map. On the other hand, different maps can be also generated by different algorithms using the same map data in automated map

Received date: 2009-05-20; accepted date: 2009-10-09
Corresponding author: YAN Haowen. E-mail: yhw0118@sina.com

generalization. Hence, a question arises, no matter how the maps are generalized manually or by computers: "which one is the best?" This problem has not been solved yet, therefore the procedure of map quality assessment in automated map generalization systems still relies on human beings manual work (Ruas, 2001), which hampers the automation of map producing process and goes against the improvement of map data quality and the acceleration of map producing period.

Automated map generalization process by computers, in essence, is the analog of human manual work. Hence, the objective of the quality assessment of digital maps by computers is to implement the automation of "comparing the original maps and the resultant maps and get their similarity degrees". Solving this problem depends on the definition, description, model and computation of spatial similarity relations in multi-scale map spaces. In other words, to automate the quality assessment process of map generalization, the theories of spatial similarity relations in multi-scale map spaces must be solved systematically.

The significance of spatial similarity relations is far from only supporting multi-scale map databases, it is also useful to the theories of spatial relations (Tobler, 1970; Miller, 2004; Goodchild, 2006) , spatial retrieval and analysis (Guo, 1997) , and spatial reasoning.

2 Definitions and Characteristics of Spatial Similarity Relations

Spatial similarity relation, as a subset of spatial relations (including topological, distance, direction, similarity and correlation relations), is the basic theory of Geo-Sciences (Goodchild, 2006; Egenhofer and Mark, 1995; Li, 1997), but the research on it has only attracted the attention in the community of cartography and geographic information sciences in recent ten years. The achievements have been made on similarity expressions of linear and areal objects before and after simplification (Ramer, 1972; Imai and Iri, 1988), description and calculation methods of spatial similarity relations (Ding, 2004), directional similarity of spatial areal objects in raster spaces (Guo and Ding, 2004), representation and calculation of topological similarity of multi-scale spatial objects (Lu and Wu, 2006), *etc*. Though so much work has been done, a definition for spatial similarity relation has not been presented by far, which is obviously not favorable to the establishment of theory and further research of spatial similarity relations. For this reason, a clear definition of spatial similarity relations is necessary indeed.

2.1 Definitions of spatial similarity relations

In essence, similarity means 1-1 correspondence of the properties of things (Zhou, 1993; Liang, 1999). In the light of the previous achievements (Li, 2000), the definition of spatial similarity relations may be discussed from the viewpoint of set theory.

Definition 1: Suppose that A_1 and A_2 are two objects in the geographic space. Their property sets are C_1 and C_2, and $C_1 \neq \Phi$ and $C_2 \neq \Phi$. If $C_1 \cap C_2 = C_\cap \neq \Phi$, $C_\cap$ is called the spatial similarity relations of object A_1 and object A_2.

Definition 2: Spatial similarity degree is a value between [0, 1]. It is used for evaluating the similarity relations of spatial objects.

Spatial similarity degree is fuzzy and uneasy to calculate precisely (Fan, 1992).

According to the above two definitions, the following conclusions can be made:

(1) The lager $C_\cap$, the larger the similarity degree of the two objects.

(2) If $C_\cap = \Phi$, the two objects has no similarity property, therefore their spatial similarity degree is 0.

(3) If $C_1 = C_2 = C_\cap$, the property sets of the two objects are wholly same, therefore their spatial similarity degree is 1.

The definition of spatial similarity relations among k ($k \geq 2$) spatial objects may be given, if the Definition 1 is extended.

Definition 3: Suppose that $A_1, A_2, \ldots, A_k$ are k objects in the geographic space. Their property sets are $C_1, C_2, \ldots, C_k$, and $C_i \neq \Phi$ (i=1, 2, …, k). If $C_1 \cap C_2 \ldots \cap C_k = C_\cap \neq \Phi$, $C_\cap$ is called the spatial similarity relations of objects $A_1, A_2, \ldots, A_k$.

Definition 1 and Definition 3 are for spatial similarity relations of objects in a single scale space. An extension of them may induce the definition of spatial similarity relations of an individual object in multi-scale map spaces.

Definition 4: Suppose that A is an object in the geographic space. It is symbolized as $A_1, A_2, \ldots, A_k$ separately on the maps at scales $S_1, S_2, \ldots, S_k$. The property sets of A_i (i=1, 2, …, k) are $C_1, C_2, \ldots, C_k$, and $C_i \neq \Phi$ (i=

1, 2, …, k). If $C_1 \cap C_2 \ldots \cap C_k = C_\cap \neq \Phi$, $C_\cap$ is called the spatial similarity relations of the multiple representations of object A in multi-scale map spaces.

2.2 Characteristics of spatial similarity relations

Spatial similarity relations own at least the following five characteristics. They will be discussed here using the definitions given in section 2.1.

(1) Reflexivity. Any object has similarity relations with itself.

(2) Symmetry. If object A has similarity relations with object B, object B must have similarity relations with object A.

(3) Non-transitivity. The conclusion that object A has similarity relations with object C can not be made, even if object A has similarity relations with object B, and object B has similarity relations with object C.

An example of non-transitivity of similarity relations is given in Fig. 1. In Fig. 1, if take {shape, land type} as the properties for detecting similarity relations, the property set of (a), (b), and (c) are C_a={rectangle, settlement}, C_b={rectangle, vegetable land}, and C_c={irregular polygon, vegetable land}. $C_a \cap C_b$={rectangle} denotes that the objects in (a) and (b) have similarity relations; $C_b \cap C_c$={vegetable land} denotes that the objects in (b) and (c) have similarity relations; but the conclusion that the objects in (a) and (c) have similarity relations can not be made, for $C_a \cap C_c = \Phi$.

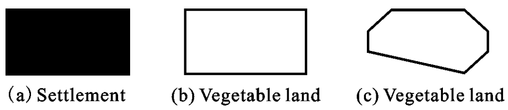

Fig. 1 An example of non-transitivity of similarity relations

(4) Self-similarity at multi-scales. An object can be expressed as different patterns and symbols on maps at different scales, but they have spatial similarity relations.

(5) Scale-dependence of self-similarity degree at multi-scales. The value of spatial similarity degree of an object on maps at different scales is scale-dependent. The greater of the scale span between two maps, the less the similarity degree is.

The examples for characteristics (4) and (5) are shown in Fig. 2, Fig. 3 and Fig. 4. Both a linear object and an areal object can be represented using different symbols on maps at different scales (Fig. 2 and Fig. 3). They have spatial similarity relations, and their similarity degree changes with the change of map scale. The point feature can be represented using different point clusters on maps at different scales (Fig. 4). They also have spatial similarity relations, taking patterns, topological relations, and semantics as their property sets, and their spatial similarity degree changes with the change of map scale.

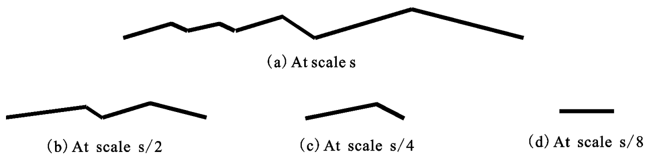

Fig. 2 Representations of a linear object on maps at different scales

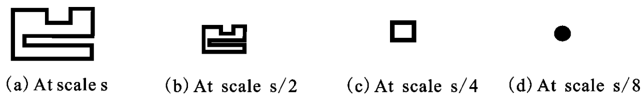

Fig. 3 Representations of an areal object on maps at different scales

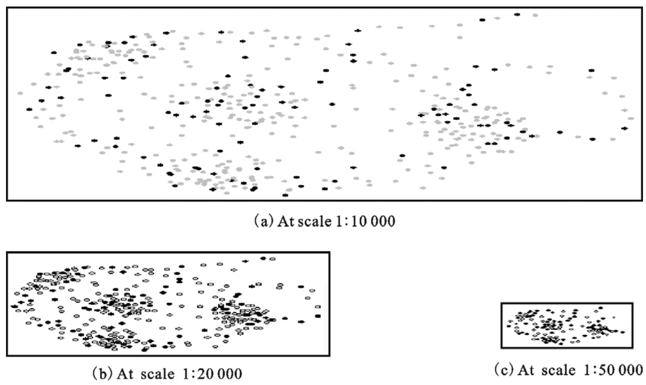

Fig. 4 Representations of point feature on maps at different scales

3 Classification of Spatial Similarity Relations

Generally, two rules must be obeyed in all classifications, i.e., completeness and exclusiveness. Completeness means that the union of all subsets of the sub-categories equals to the whole set, while exclusiveness means that the intersection of every two subsets is empty. To meet the demands of the two rules, appropriate criteria must be specified for the purpose of classification. Different criteria make different classification systems.

In the light of the definitions of spatial similarity relations in section 2.1, it sounds natural to classify spatial similarity relations by the scales of researched objects: Whether the objects are at same scale or different scales. The former is called horizontal similarity relations and the latter is called perpendicular similarity relations (Fig. 5). A typical application of horizontal similarity relations is spatial similarity retrieval. This paper focuses on perpendicular similarity relations.

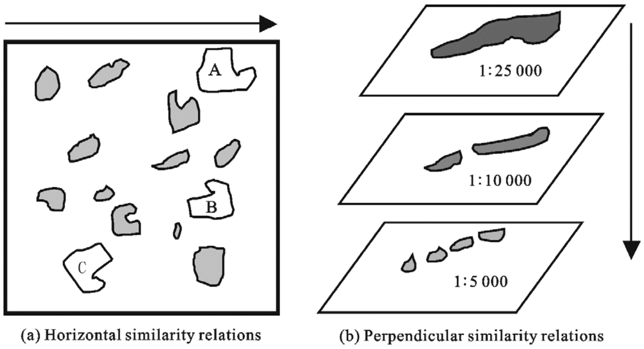

Fig. 5 A scale-based classification system for spatial similarity relations

As far as perpendicular similarity relation is concerned, the geometric graphics and attributes should be selected as classification criteria, for both of them are most important in the representation of spatial objects. Further, for a single object, its geometric properties include dimension, size, shape, length, area, etc. and for group objects, their geometric properties include topology, distance, direction and correlation. As to attribute properties, semantics and time (i.e. temporal aspects of spatial data) should be taken into consideration. Therefore, a framework of the classification system for spatial similarity relations in multi-scale map spaces may be given according to this thought (Fig. 6).

Based on the classification system shown in Fig. 6, spatial similarity relations may be discussed at the level of properties. For example, Fig. 1a and Fig. 1b have similarity relations in dimension, size, shape, and area, while Fig. 1b and Fig. 1c have similarity relations in semantics. The objects at four scales in Fig. 2 are similar in dimension, shape, and semantics, so are the objects in Fig. 3. As to the three point clusters at different scales, they are similar in topology, distance, direction, and semantics (Fig. 4).

The geometric properties and attribute properties of the classification system in Fig. 6 do not crown all of

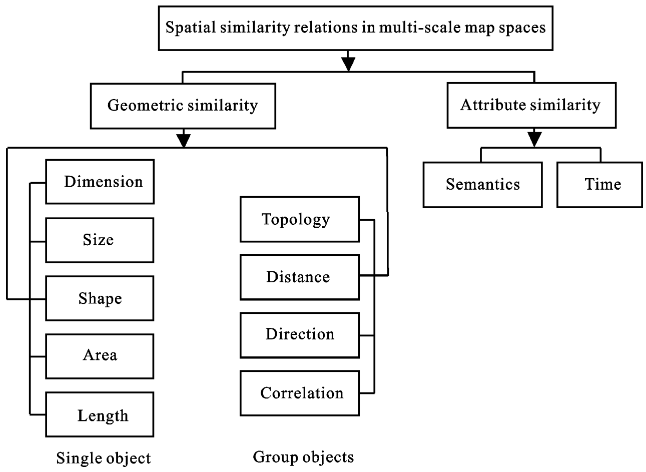

Fig. 6 A classification system for spatial similarity relations in multi-scale map spaces

them. It just provides an idea for the classification and presents a classification framework. The properties in the classification system may be added and deleted according to practical applications.

4 Conclusions

Spatial similarity relation in multi-scale map spaces is a fundamental issue in the community of cartography and geographic information sciences at the level of theory and applications, which is significant to the construction of multi-scale map databases, spatial retrieval and analysis, and spatial reasoning. This paper presented its definition by means of set theory, and discussed the characteristics of spatial similarity relations in detail, and addressed its classification system at the level of properties.

Indeed, the theory of spatial similarity relation in multi-scale map spaces is a new task arising in automated map generalization, and little achievement has been made. Some problems, such as the factors that affect the judgment of spatial similarity relations, the formal description models and calculation methods of spatial similarity relations in multi-scale map spaces, are complicated and difficult to solve. This paper only puts emphasis on some fundamental problems of it, and aims at providing a theoretical foundation to the further research on the models and calculation methods of spatial similarity relations in multi-scale map spaces.

References

Ding Hong, 2004. Research on the theory and calculation models of spatial similarity. Dissertation for Doctor′s degree. Wuhan: Wuhan University. (in Chinese)

Egenhofer M, Mark D, 1995. Naive geography, spatial information theory—A theoretical basis for GIS. In: Frank A, Kuhn W (eds.). *Proceeedings of International Conference COSIT '95*. Berlin: Springer-Verlag, 988: 1–15.

Fan Jian, 1992. Similarity relations: the foundation of analog methods. *Studies in Dialectics of Nature*, 8(8): 48–50. (in Chinese)

Goodchild M F, 2006. The law in geography. In: *Technical Report in GIS the GIS Institute*. Zurich: University of Zurich.

Guo Qingsheng, Ding Hong, 2004. Similarity for spatial directions between areal objects in raster data. *Journal of Geomatics and Information Science of Wuhan University*, 29(5): 447–450. (in Chinese)

Guo Renzhong, 1997. *Spatial Analysis*. Wuhan: Press of Wuhan Technical University of Surveying and Mapping. (in Chinese)

Imai H, Iri M, 1988. Polygonal approximation of a curve-formulations and algorithms. In: Tousaint G T (ed.). *Computational Morphology*. Amsterdam: Elsevier Science Publishers. 71–86.

Li Deren, 1997. Some thoughts about geographic information theory. *Journal of Wuhan Technical University of Surveying and Mapping*, 22(2): 3–5. (in Chinese)

Li Jiachun, 2000. From graphic similitude to object similitude. *Journal of Sichuan University of Science and Technology*, 19(2): 126–128. (in Chinese)

Liang Junxiong, 1999. Fundamental concepts of similarity relation and its weak equivalence properties. *Systems Engineering-theory & Practice*, 28(7): 106–111. (in Chinese)

Lu Xiuqing, Wu Fan, 2006. Representation of topological relationships for spatial objects in multi-scale representation. *Journal of Geomatics*, 31(2): 29–31. (in Chinese)

Miller H J, 2004. Tobler′s first law and spatial analysis. *Annals of the Association of American Geographers*, 19(2): 284–289.

Ramer U, 1972. An iterative procedure for the polygonal approximation of plane curves. *Computer Vision Graphics and Image Processing*, 1(3): 244–256.

Ruas A. 2001, Automating the generalization of geographical data. In: *Proceedings of the 20th International Cartographic Conference*. Beijing, China, 1943–1953.

Tobler W R, 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46(2): 234–240.

Wang Jiayao, 1993. *Principles of General Map Generalization*. Beijing: Surveying and Mapping Press. (in Chinese)

Zhou Meili, 1993. *Similarity Theories*. Beijing: China Science and Technology Publishing House. (in Chinese)