

A Seamless Constraint Model of Multi-Scale Representation of Geographical Information

Di Chen^{1,3}, Han Yue^{1*} and Xinyan Zhu^{1,2}

¹*State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, China*

²*Collaborative Innovation Center of Geospatial Technology, Wuhan University, Wuhan, China*

³*China Data Center, University of Michigan, Packard Rd. 330, Ann Arbor, Michigan, USA*

**Corresponding author, Email: 1083378776@qq.com*

Abstract

At present, it has become a hot issue to provide a multi-representation mechanism and build multi-scale spatial databases in the field of GIS. Meanwhile, an inevitable problem is how to evaluate and preserve the consistency of multi-scale spatial data. In this paper, we introduce the concepts of homonymous points and homonymous entities based on innate characteristics of spatial data and put forward a seamless constraint model of multi-scale representation of geographic information that involves time, geometry, spatial relations and semantic characteristics. According to this model, a consistency assessment method has been implemented by computing similarities between vector data at three different scales. The main contributions of this paper are the proposal of a consistency constraint model and assessment system of multi-scale spatial data.

Keywords: *constraint model, multi-scale representation, spatial database, seamless representation, homonymous entities*

1. Introduction

With the development of technology and application of geographic information, requirements for high-quality geospatial representations pose new challenges for Geographic Information System (GIS), which can express geographical information from multi-view, multi-scale and multi-level perspectives. At present, it has become a hot issue to provide a multi-representation mechanism and build multi-scale spatial databases in the field of GIS, which involves a set of complex methods and technologies about spatial data consistency [1]. One important problem is how to evaluate and preserve the consistency of multi-scale spatial data representation. Most research so far has focused on similarity measurement between identical or similar scales, but there is still a lack of research on large scale differences. The inconsistency of spatial data usually results from location deviation and physical or logical inconsistency in spatial databases [2]. Therefore, the spatial similarity is mainly evaluated from aspects of geometrical shape [3]-[4] and spatial relationship [5][7], and is lack of characteristics such as time, semantic and global structure.

In this paper, we define the concepts of homonymous points and homonymous entities through deeply analyzing seamless expression characteristics of spatial data. A seamless constraint model of multiple representation of geographic information has been put forward. We use this model in spatial consistency assessment that computes similarity between vector data at three different scales. The contribution of our work is the proposal

of a consistency constraint model of seamless expression and establishment of a consistency assessment system of multi-scale spatial data.

2. Seamless Constraint Model

2.1. Space-time Homonymous Points

It is possible for a certain space-time point to have many different expressions in the digital world. An identical house in real world shown as the upper figure in Fig.1 could be expressed in two different ways, which corresponds to the lower two pictures, i.e. SD1 and SD2.

This kind of points on maps are known as homonymous points because they indicate a same position in real world. We use SD_{narrow} to denote the spatial data:

$$SD_{narrow} = \{X, Y, Z, T, A\} \quad (1)$$

Where X, Y, Z and T are 3-D coordinates and time, and A is attribute value respectively.

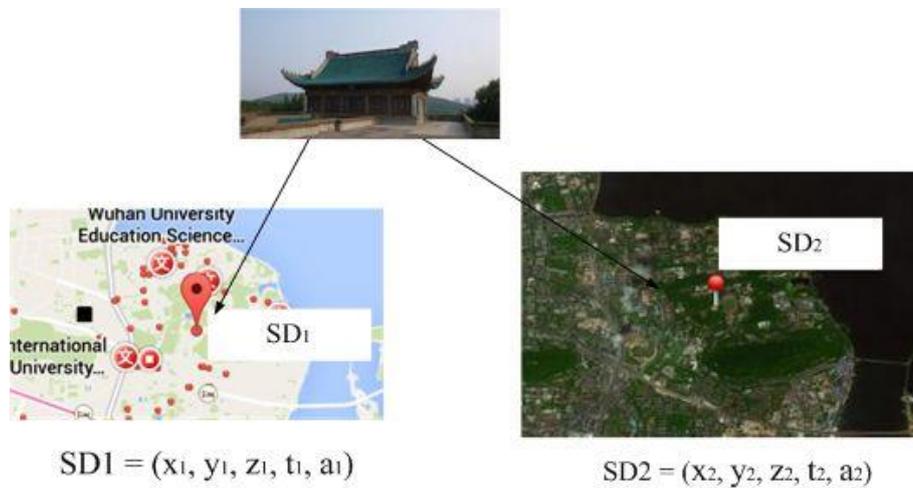


Figure 1. Homonymy Points

So it leads to a seamless concept in a narrow sense. Given two representations of spatial data, $SD1_{narrow}$ and $SD2_{narrow}$, the differences between them can be defined as follows.

$$\begin{aligned} \Delta x_i &= \bar{x}_{1i} - \bar{x}_{2i}; \\ \Delta y_i &= \bar{y}_{1i} - \bar{y}_{2i}; \\ \Delta z_i &= \bar{z}_{1i} - \bar{z}_{2i}; \\ \Delta t_i &= \bar{t}_{1i} - \bar{t}_{2i}; \\ \Delta a_i &= \bar{a}_{1i} - \bar{a}_{2i}; \end{aligned} \quad (2)$$

Where \bar{x}_{1i} , \bar{y}_{1i} , \bar{z}_{1i} , \bar{t}_{1i} and \bar{a}_{1i} indicate to transform these spatial expressions into the same space-time reference system. So the difference of spatial expression is actually the distance in the 4D space between homonymous points. If the SD1 and SD2 satisfied the constraints ($\Delta x_i < \varepsilon \cap \Delta y_i < \varepsilon \cap \Delta z_i < \varepsilon \cap \Delta t_i < \varepsilon$, where ε is the tolerance), we can deduce that they are narrowly defined seamless expressions. Therefore, the inconsistency of spatial expressions generally refers to the geometrical inconsistency in a narrow sense.

2.2. Homonymous Entities

The characteristics of spatial data are not only reflected in geometrics but also in geographic semantics and spatial relations between objects, including topological, directional, metrics relationships. So the spatial expression of geographic entities can be expressed as formula (3).

$$S = \{O, S_{top}, S_{dir}, S_{mea}\} \quad (3)$$

Where O is the set of objects in the real world, and S_{top} , S_{dir} , S_{mea} is topological, directional and matric relationships, respectively. It is worth noting that an identical geographic entity is expressed as different geometric forms on different maps. The different representations of the same entity may have different scales, different dimensions and even different sizes. Therefore, the concept of homonymous entities can be raised, similar to homonymous points. Given $o_1, o_2 \dots o_n$ are different representations in spatial database, if they represent an identical entity e in the real word, they are considered as homonymous entities. The spatial database needs not only to record the geometric shapes of geographical entities but also their attributes, semantics and scales. It is also necessary to store the detail of data processing in order to provide a more comprehensive geographical expression, which will benefit to find out the cause of data inconsistency. Broadly defined, the spatial data representation model can be formalized with a 5-tuple (4).

$$SD_{broad} = \{S, T, A, SC, SM, U\} \quad (4)$$

Where S and T are spatial expression and time; A denotes attributes; SC and SM are spatial scale and spatial semantics, U is data processing. The common data processing includes generalization, data transformation, symbolization etc.

2.3. The Broadly Seamless Expression

The broadly seamless expression can be described according to the definition of narrowly seamless expression. Given two representations of spatial data $SD1_{broad} \{S_1, A_1, T_1, SC_1, SM_1, U_1\}$, $SD2_{broad} \{S_2, A_2, T_2, SC_2, SM_2, U_2\}$, they are of the broadly seamless expressions when some consistency conditions are satisfied, including time coherence, same spatial scale, content similarity, compatible spatial dimension, spatial relationship similarity and semantic consistency, etc. It is really accessible for the concept of seamless expression of multi-scale representation. It does not request the representations of spatial data belong to the same scale ($SC_1 \neq SC_2$) but the other conditions are still valid ($S_1 \circ S_2 \cap A_1 \circ A_2 \cap T_1 \circ T_2 \cap SM_1 \circ SM_2 \cap U_1 \circ U_2$). These constraints will be discussed in next section, and the seamless constraint model of multi-scale representation is put forward here first as Fig.2.

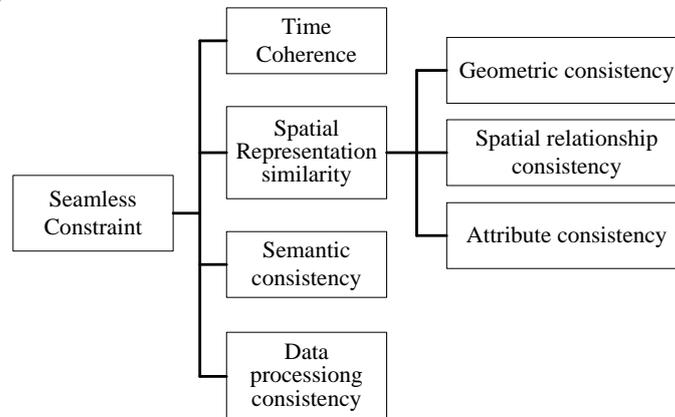


Figure 2. Seamless Constraint Model of Multi-scale Representation

3. Seamless Constraints of Multi-Scale Representation

Spatial data are commonly collected by different users at different time and contain different levels of details [8]. The inconsistency of geometric shape, spatial relationship and geographical semantics is ubiquitous in multi-scale spatial databases due to data acquisition approaches, map generalization algorithms, etc. A more important thing is to recognize the constraints of consistency rather than how to resolve the problem of inconsistency. The seamless issues can be divided into logical inconsistency and physical inconsistency [9]. In this section, we will discuss the constraint model of seamless multi-scale representation from logical perspective.

3.1 Time Coherence

The spatial representations would always be time coherent until the geographical entity would have changed at the time t_2 . In other word, time coherence mainly depends on whether the entity is unchanged in terms of location and attribute. In Figure 3, SD_1 and SD_3 are time coherent in $[t_1, t_2]$ until t_i , when a new entity e_i is added to the spatial scene.

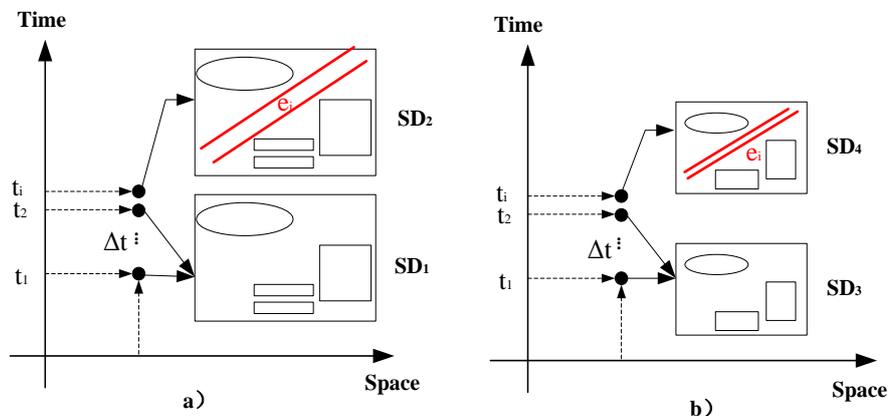


Figure 3. Time Coherence: a) Large-scale, b) Small-scale

3.2 Spatial Representation Similarity

Geographical entities are generally presented by geometric shapes and object attributes. So the content similarity of multi-scale representation refers to geometric and attribute consistency.

3.2.1 Geometric Consistency: As the spatial scale or resolution changes, the geographical entities will be expressed by different geometric forms. For such maps, we can state in general that at finer resolution the numbers of areas and volumes will increase, their boundaries will be more detailed and homogeneous [10]. In the narrow sense, the homonymy entities at different resolutions are inconsistent or incompatible in terms of shape, structure and dimension. However, they are geometric consistent from the broad sense as long as the geometric morphing can satisfy the constraints of map generalization at three levels.

- (1) **Overall.** The overall consistency insists that the overall structures or patterns of multi-scale representations are similar. It can be evaluated by comparing the spatial distribution of key features in two spatial scenes.
- (2) **Local.** The local sense refers primarily to the unit that is consisted of a set of adjacent features. It is reflected by the spatial relations between two features. The consistency can be assessed through computing the similarity of spatial relations.
- (3) **Individual.** The individual consistency emphasizes to evaluate the similarity of individual geometric characteristics, such as shape, coordinate, dimension and

structure. It is necessary for a consistency evaluation to consider the variations that result from generalization operation.

Moreover, another important constraint is dimension compatibility. It will be guaranteed that the dimension of features shown on map is reasonable with the changes of scale or resolution.

3.2.2 Similarity of Spatial Relationships: The similarity of spatial relationships is also an important feature of multi-scale representation, it can be distinguished in detail as follows:

- (1) **Topological consistency.** When the scale variates, the topological relation will not only change in type (e.g. intersection, overlap), but also in structure (e.g. point/point, point/line, line/area). The topological consistency actually demands that the topological relationship of homonymous entities maintains consistent at different scales. And the concept distance and fuzzy calculation model presented in previous literatures are adopted to measure the similarity of topological relationship.
- (2) **Directional consistency.** Broadly, this consistency claims that the directional relations of homonymous entities remain constant when the scale changes. In [11] the direction concept distance is defined as the number of steps to transform one type of direction matrix to another. For example, the shape of entity B shows in two different forms under two spatial resolutions, thus the direction relation between A and B may change as shown in Fig.4. The direction relation changes from east to northeast.
- (3) **Metrical consistency.** Metric relationship is the strongest constraint in the three spatial relations, i.e. topology, direction, metric. A and B are two homonymous entities in Fig.5. The distances between them are d_1 and d_2 at different spatial scales. If the condition $|d_1 - d_2| < \Delta d$ meets, the representations at two scales are metrical consistent, where Δd is the threshold of metrical consistency and depends on the range of scale variation.

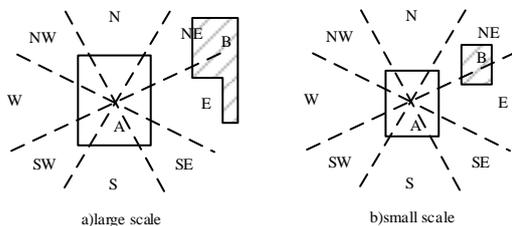


Figure 4. Directional Consistency

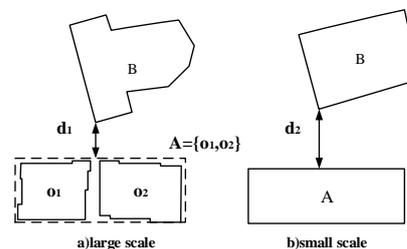


Figure 5. Metrics Consistency

3.3 Semantic and Attributive Consistency

Geographical semantic is also scale-dependent, and the semantic scale is closely related to time and space scales. Therefore, it is an inherent demand that scale variations of semantic keep pace with time and space. In general, the more subtle the spatial representation is, the higher the level of semantic concept is in geographic ontology. That is, the resolution or level of semantic changes with time and spatial scale. Here attributes denote the information that describes the geographic entities from quantitative perspective. They can be divided into spatial attributes and non-spatial attributes. In order to maintain the continuity of semantic, it is inevitable to update the attribute information to an appropriate semantic level. For example, the attribute value or field of spatial expression need to be aggregated or merged when the spatial resolution get rough.

3.4 Data Process Consistency

The contents above mainly discusses the consistency of multi-scale representation under static expressions. Besides innate inconsistency derived from spatial data itself, data processing is another main source of inconsistency, including data transformation,

generalization processing, spatial query and analysis. The data process consistency can be defined as follows: Given two seamless spatial representations, which represent an identical region at different scales ($SC1 \neq SC2$), if the consistency or similarity still maintains after undergoing the same process ($U1, U2 \dots$), they are of data process consistency.

4. Case Study

In order to illustrate our model, we collect multi-scale spatial scenes which represent the same region, Wuhan University (WU) as Figure 6. There may be various numbers or dimensions of features in these spatial scenes. For example, WU is represented by a point feature in Figure 6 (a), two areas in Figure 6 (b), and an area group in Figure 6(c). And the geometric type of Bayi road changes from double lines to a single line.

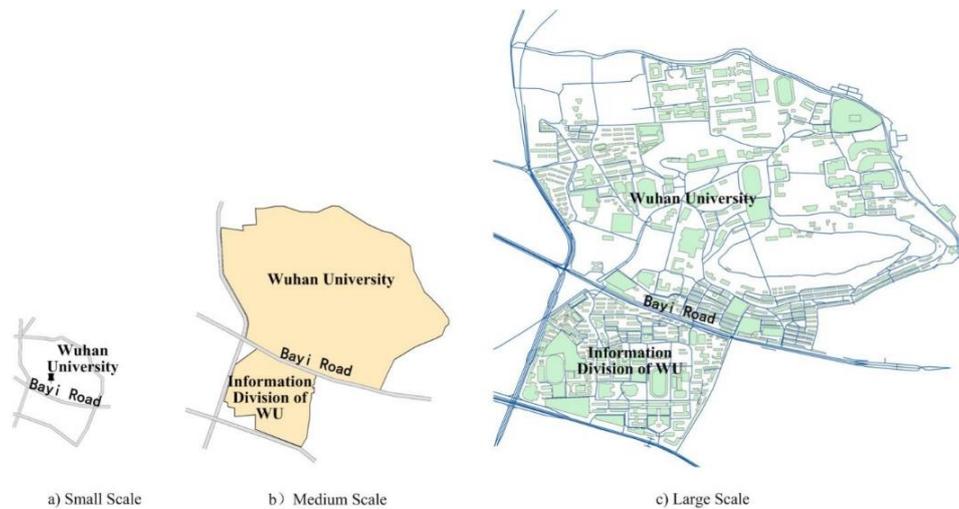


Figure 6. Homonymous Entities at Different Spatial Scales

Table 1. The Results of Multi-scale Representation Similarity

Element	Time	Spatial Representation				Semantic	Data Process	Consistency (Y/N)
		Geometric Shape (0.2)	Topology (0.3)	Direction (0.3)	Metric (0.2)			
Small/Medium	1		0.93			1	1	Y
		1	0.89	0.87	1			
Medium/Large	1		0.97			1	1	Y
		0.90	1	0.92	1			
Small/Large	1		0.86			1	1	Y
		1	0.89	0.86	1			

These representations are obviously inconsistent from narrow perspective, but seem not to hold from broader. Therefore, the seamless constraint model can be used to induct the consistency evaluation. It is noteworthy that the process directly or indirectly takes advantage of some existing similarity methods of multi-scale spatial scene or spatial relationship, such as concept distance of spatial relation[12], fuzzy description model for direction relation [13], Geo-ontology[14], etc. The similarity between these spatial representations can be calculated, and the result of assessment is a series of quantitative values. The similarities of elements are shown in Table 1. And the whole similarity can also be obtained after assigning the weights to elements. The similarities of Figure 6 (a) and Figure 6 (b) is 0.93; Figure 6 (b) and Figure 6 (c) is 0.97; Figure 6 (a) and Figure 6 (c) is 0.93. Thus they are high similar or consistent.

5. Conclusion

The seamless assessment of multi-scale spatial data is implemented from both logical and physical perspectives. This paper discusses the consistency of multi-scale representations from views of logical seamless. The concepts of space-time homonymous points and homonymous entities are proposed by analyzing the substantive characteristics of spatial data. And the seamless constraint model of multi-scale representation is established based on these concepts. Some various consistency constraints are described, including time coherence, spatial representation similarity, semantic consistency, data processing consistency, etc. We adopt spatial scenes at three different scales to illustrate how to use our model and prove its completeness and rationality.

Different from previous studies which mainly focused on local spatial representation similarity and are lack of concern for important aspects such as time, semantic, data processing, our research is more complete in theoretical framework. The model proposed in this paper can be applied into the process of building multi-scale spatial database, and facilitates applications such as spatial query, spatial reasoning, map quality assessment and so on.

Acknowledgements

Authors wish to thank Natural Science Foundation of China (No. 41271401).

References

- [1] Tinghua AI, Jianguo CHENG, 'Key Issues of Multi-Scale Representation of Spatial Data', Editorial Board of Geomatics and Information Science of Wuhan University (2005), Vol.30, No.5, pp.377-382.
- [2] Gong P, Mu L, 'Error Detection through Consistency Checking', Geographic Information Sciences (2000), Vol.6, No.2, pp.188-193.
- [3] AN Xiaoya, Qun SUN, Qiang XIAO. 'A Shape Multilevel Description Method and Application in Measuring Geometry Similarity of Multi-scale Spatial Data', Acta Geodaetica et Cartographica Sinica (2011), Vol.40, No.4, pp.495-501.
- [4] Tinghua AI, SHUAI Y, Jingzhong LI, 'A Spatial Query Based on Shape Similarity Cognition', Acta Geodaetica et Cartographica Sinica (2009), Vol.38, No.4, pp.356-362.
- [5] Bruns T, Egenhofer M, 'Similarity of Spatial Scenes', Seventh International Symposium on Spatial Data Handling (1996), Delft, The Netherlands, pp.31-42.
- [6] Argemiro J, Paiva C, Paiva A C, 'Topological Equivalence and Similarity in Multi-representation Geographic Databases' (1998).
- [7] Kang H K, Kim T W, Li K J, 'Topological Consistency for Collapse Operation in Multi-Scale Databases', Conceptual Modeling for Advanced Application Domains, Springer Berlin Heidelberg (2004), pp. 91-102.
- [8] Du S, Guo L, Wang Q, 'A Scale-Explicit Model for Checking Directional Consistency in Multi-Resolution Spatial Data', International Journal of Geographical Information Science (2010), Vol.24, No.3, pp.465-485.
- [9] Zhu X Y, Zhang J C, Li D R, 'Concepts, Implementation and Problems of the Seamless Spatial Database', Geomatics and Information Science of Wuhan University (2002), Vol.27, No.4, pp. 382-386.
- [10] Goodchild M F, 'Scale in GIS: An Overview', Geomorphology (2011), Vol.130, No.1, pp.5-9.
- [11] Goyal R K, 'Similarity Assessment for Cardinal Directions between Extended Spatial Objects' (2000), The University of Maine.
- [12] Egenhofer M J, Franzosa R D, 'On the Equivalence of Topological Relations', International Journal of Geographical Information Systems (1995), Vol.9, No.2, pp.133-152.
- [13] CHEN Di, ZHU Xinyan, ZHOU Chunhui, DAI You, 'A Fuzzy Description Method for Directional Relationship Based on Adaptive Sampling Granularity Model', Acta Geodaetica Et Cartographica Sinica (2013), Vol.42, No.3, pp.359-366.
- [14] Schwering A, 'Approaches to Semantic Similarity Measurement for Geo-Spatial Data: A Survey', Transactions in GIS (2008), Vol.12, No.1, pp.5-29.

Authors



Di Chen, he was born on 12.02.1985, Wuhan, Hubei, China. His current position, grades: PhD at Wuhan University. At present, he is a post-doctor at China Data Center of Michigan University. He obtained his PhD (2014) on cartography and geographical information engineering at Wuhan University. He is a specialist in the field of spatial database. His works are devoted to multi-representation of spatial database and spatial relationship.



Han Yue, he was born on 12.23.1990, Wuhan, Hubei, China. His current position, grades: Master at Wuhan University. He obtained his bachelor's degree(2013) on surveying and mapping engineering at Wuhan University. He is interested in Geographical Information Science and is now doing research on the multi-representation of spatial scenes.



Xinyan Zhu, he was born on 11.23.1963, Wuyuan, Jiangxi, China. He is currently a professor of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University. His PhD (photogrammetry and remote sensing) at Wuhan University. He is interested in Geographical Information System, Spatial database. He is an expert in the theory of Geographical Information System and spatial database. In last decade, he obtained significant results in the field of WebGIS, Spatial database and Location-based service.