Similarity Calculation Under Multiple Expression of Polygon Entity

Qin ZONG[a,1], Qiu ZUO[a], Kuan TIAN[a] and Yanxia LIU[b]
[a] *Chongqing Jianzhu College, China*
[b] *Wuhan Geomatics Institute, China*

Abstract. Aiming to understand the morphological differences under the multiple expressions of spatial entities, after optimizing the shape segmentation ratio, the shape inner distance segmentation ratio matrix and the shape outer distance segmentation ratio matrix are established. The shape similarity model under the multiple expressions of a polygon entity is then established to provide quantitative indicators for the query, matching and change detection of spatial entities.

Keywords. Polygon entity, multiple expression, similarity keyword.
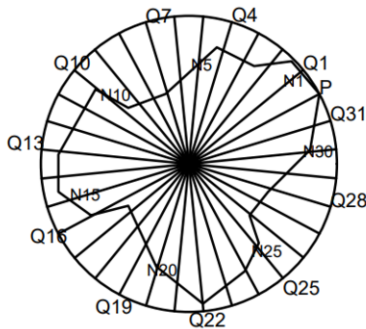
1. Introduction

According to past research, the process by which an entity with the same name changes from one form to another in a multi-scale vector map database is called a fuzzy similarity transformation. The relationship before and after the change is called the fuzzy similarity relationship [1-2], which is the fixed relationship between the entity under the scale transformation. The calculation model of the quantitative relationship is very important for the construction, query, matching and fusion in the multi-scale vector map database. Previously, scholars have established similarity relationship calculation models of point entities, line entities and polygon entities in different dimensions [3-12], which also have a certain optimization space. Considering that the point entity and line entity can be upgraded, this paper mainly studies the similarity calculation under the multiple expressions of a polygon entity.

2. Shape Descriptor Extraction

Firstly, the construction process of the referenced geometric feature segmentation ratio vector [13] is optimized, and the shape's inner distance segmentation ratio matrix and outer distance segmentation ratio matrix are constructed. These are used to establish the morphological similarity model under the multiple expressions of the polygon entity.

In order to establish the calculation model of each parameter, the referenced method [13] is improved to extract the shape descriptor of the polygon entity. As shown in figure 1, a feature line is made every 360 ° / M. The features of the polygon entity are further refined after being segmented by the feature line.

Figure 1. Characteristic line division.

The flow algorithm shown in figure 2 describing the mining of the shape description subset of the entity.

Figure 2. Shape descriptor mining process

Definition 1: The scale is S, S is related to 360 ° / M where M represents how many feature lines are divided. 360 ° / M represents the angle of each interval to make a feature line, and then S represents the scale of the polygon entity segmentation, which is increased from 1 to M according to step 1.

Definition 2: The shape inner distance refers to the distance from the centroid to the feature point, which are ON_1, ON_2... ON_M, etc., expressed as $d_{s,i}^{(I)}$ ($i = 1$、 2、 3、 ... 、 M).

Definition 3: The shape outer distance is the distance from the feature point to the corresponding intersection Q_i of the feature line and R, which are Q_1, N_1, Q_2, N_2,..., $Q_M N_M$, etc., expressed as $d_{s,i}^{(O)}$ ($i = 1$、 2、 3、 ... 、 M).

Two sequences are constructed: the shape's inner distance segmentation ratio sequence $\{d_{s,1}^I/D_{max}, d_{s,2}^I/D_{max}, d_{s,3}^I/D_{max}, ..., d_{s,M}^I/D_{max}\}$ and the shape's outer distance segmentation ratio sequence $\{d_{s,1}^O/D_{max}, d_{s,2}^O/D_{max}, d_{s,3}^O/D_{max}, ..., d_{s,M}^O/D_{max}\}$. If S is allowed to present multiple dimensional changes according to definition 1, two M-order matrices can be obtained:

(1) Shape's spacing partition ratio matrix

$$
\text{IDMATRIX} =
\begin{bmatrix}
\dfrac{d_{1,1}^I}{D_{max}} & \dfrac{d_{1,2}^I}{D_{max}} & \cdots & \dfrac{d_{1,M}^I}{D_{max}} \\
\dfrac{d_{2,1}^I}{D_{max}} & \dfrac{d_{2,2}^I}{D_{max}} & \cdots & \dfrac{d_{2,M}^I}{D_{max}} \\
\vdots & \vdots & \cdots & \vdots \\
\dfrac{d_{M,1}^I}{D_{max}} & \dfrac{d_{M,2}^I}{D_{max}} & \cdots & \dfrac{d_{M,M}^I}{D_{max}}
\end{bmatrix}
\tag{1}
$$

(2) Shape's outer distance partition ratio matrix

$$
\text{ODMATRIX} =
\begin{bmatrix}
\dfrac{d_{1,1}^O}{D_{max}} & \dfrac{d_{1,2}^O}{D_{max}} & \cdots & \dfrac{d_{1,M}^O}{D_{max}} \\
\dfrac{d_{2,1}^O}{D_{max}} & \dfrac{d_{2,2}^O}{D_{max}} & \cdots & \dfrac{d_{2,M}^O}{D_{max}} \\
\vdots & \vdots & \cdots & \vdots \\
\dfrac{d_{M,1}^O}{D_{max}} & \dfrac{d_{M,2}^O}{D_{max}} & \cdots & \dfrac{d_{M,M}^O}{D_{max}}
\end{bmatrix}
\tag{2}
$$

The shape's inner distance segmentation ratio matrix reflects the size and shape characteristics of multiple expressions of the polygon entity. The shape's outer distance segmentation ratio matrix reflects the approach degree to regular circles under the multiple expressions of the polygon entity, and needs to meet the invariance under fuzzy similarity transformation. This means that the shape's outer distance segmentation ratio matrix and shape inner distance segmentation ratio matrix will not change as the translation of the spatial entities rotate and their scale changes. Firstly, when the spatial entity is translated, it is obvious that it will not affect the calculation results of the matrix, thus has translation invariance under fuzzy similarity transformation. When the spatial entity is scaled, the shape's inner distance partition ratio matrix and shape's outer distance partition ratio matrix are reduced by $d_{s,i}^{(I)}$, $d_{s,i}^{(O)}$ and D_{max}, which has the same scaling coefficient whilst eliminating has scaling invariance. When the spatial entity rotates, the position of the starting vector is different, but the two sequences choose the maximum modulus as the starting vector. These are still arranged counterclockwise according to the starting vector, which does

not affect the size of the matrix. Therefore, they have scaling invariance under fuzzy similarity transformation.

3. Calculate Similarity

We can conclude from the analysis in the previous section that the shape's inner distance segmentation ratio matrix and shape's outer distance segmentation ratio matrix do not change with the translation, rotation and scaling of the spatial entity. Therefore, we can conclude that the two matrices can be directly compared to measure the difference of the polygon entity. If the spectral characteristic sets of face entity A and B are SHAPEDES^A and SHAPEDES^B respectively, the Euclidean distance between them can be calculated, then the second norm of matrix difference, and finally the difference degree of the polygon entities A and B can be obtained, i.e.

$$\text{Diversity}(A,B)_{areaentity} = (\|IDMATRX^A - IDMATRX^B\| + \|ODMATRX^A - ODMATRX^B\|)/(2sqrt((T-1)M)) \tag{3}$$

Where, $\|\cdot\|$ represents the second norm of the calculation matrix, $sqrt$（·）represents the calculation square, $Deg(A,B)_{areaentity} \in [0,1]$, therefore, the shape similarity between A and B is

$$Shape(A,B)_{areaentity} = 1 - \text{Diversity}(A,B)_{areaentity} \tag{4}$$

When it is necessary to establish a similarity measurement model with other similarity influencing factors such as location, direction and size, it is possible to allocate models' respective weights and calculate the weighted average to establish a comprehensive similarity calculation model. This can be described as:

$$\begin{bmatrix} Shape(A,B)_{entity} & Location(A,B)_{areaentity} \\ Direction(A,B)_{areatntity} & Size(A,B)_{areaentity} \end{bmatrix}^T$$

respectively representing the similarities between A and B in shape, position, direction and size. $weight_i (i = 1,2,3,4)$ is the weight of each measuring element. The comprehensive similarity calculated using the weighted average method is:

$$Sim(A,B) = (Shape(A,B)_{entity}\, weight_1 + Location(A,B)_{areaentity}\, weight_2 + Direction(A,B)_{areatntity} weight_3 + Size(A,B)_{areaentity} weight_{41})/\sum_{i=1}^{4} weight_i \tag{5}$$

4. Conclusion and Prospect

One of the new "4E standardized products" in basic surveying and mapping is stepless map expression. This has been the subject of long-term research in many geographic information fields, such as automatic cartographic generalization, spatial database management, spatial data model, and geographic entity expression amongst others. Although the process is arduous, it has achieved fruitful results. The detailed theoretical research carries the undercurrent of upper technical problems. The similarity calculation under the multiple expressions of geographical entities has important bottom significance. Future research is required to develop further ideas and studies, in particular for the following aspects:

(1) Application of stepless feature expression. The definition of new basic surveying and mapping is that it is a basic project that takes a "geographical entity" as the perspective and object, carries out surveying and mapping according to "spatial

accuracy and entity granularity", constructs "basic geographical entity databases" as the goal, and assembles "4E standardized products". These products include combined aggregation entity sets, stepless map expressions, terrain level real time 3D maps and city level real time 3D maps, which can be used for surveying for public welfare and mapping behavior. There is a relationship between stepless map expression and multiple expressions in this definition, and the specific contribution of quantitative calculations to spatial entity similarity in stepless map expression.

(2) Similarity and complexity of spatial 3D objects. At present, data acquisition methods are intelligent and diversified. Low altitude digital aerial photogrammetry can obtain three-dimensional models of geographical entities much more rapidly. This could be applied to the similarity descriptors of spatial three-dimensional objects in the space-time dimension and the calculation models of their similarity.

References

[1] Zong Q, Jiang S H, Peng Q. Mapping of Fuzzy Similarity Relation in Multi-scale Vector Map Database. Science of Surveying and Mapping.2020Jan; 45(3):117-121.

[2] Zong Q, Jiang S H, Liu Y X. Fuzzy Similarity Transformation and Its Metric Model in Multiscale Vector Map. Science of Surveying and Mapping.2018Feb; 43(1):72-78.

[3] Xu F, Niu J Q, Lin H, et al. Establishment of the Similarity Metric Model of Multi-scale Spatial Object Using Isometry. Geomatics and Information Science of WUHAN University.2019Jan; 44(9):1399-1406.

[4] Cheng M M, Sun Q, Xu L, et al. Polygon contour similarity and complexity measurement and application in simplification. Acta Geodaetica et Cartographica Sinica.2019 Apr; 48(4):489-501.

[5] Cheng M M, Sun Q, Li S, et al. Generalized Hausdorff Distance of Multi-scale Point Group and Its Application in Similarity Measurement. Geomatics and Information Science of WUHAN University.2019Apr;44(6)：885-891.

[6] Ma J Z, Sun Q, Xiao Q, et al. Research on comprehensive similarity measurement of face and solid using multi-level chord length arch height complex function. Journal of Image and Graphics.2017 May; 22(4):0551-0562.

[7] Li Y, Gao D Q. Research on entity similarity calculation in knowledge map. Journal of Chinese Information Processing.2017Apr; 31(1):140-154.

[8] Cheng Z L, Lv M L, Wu L, et al. SpaceScene Similarity Metrics Based onFeature Matix and Associated Graph. Geomatics and Information Science of WUHAN University.2016Jul; 41(11):1-7.

[9] Cheng Z L, Qin M J, Wu L, et al. Establishment of the Comprehensive Shape Similarity Model for Complex Polygon Entity by Using Bending Mutilevel Chord Complex Function. Acta Geodaetica et Cartographical Sinica.2016, Feb; 45(2):224-232.

[10] Cheng Z L, Xu Y Y, Xie Z. Discussion on geometrical similarity measurement of vector planar elements. Science of Surveying and Mapping, 2016Jan ;(6):12-16.

[11] Cheng Z L, Gong X, Wu L, et al. A Quantitative Calculation Method of Composite Spatial Direction Similarity Concerning Scale Differences. Acta Geodaetica et Cartographica Sinica.2016Mar; 45(3):362-371.

[12] Xia Y, Zhu X Y. A geographic entity similarity query method based on shape features. Geography and Geo-Information Science.2015Jan; 31(1):6-11.

[13] Zou J, Chen Y G, Gong J Q, et al. A Efficient Matching Algorithm Based on Vector Graphics Using Multi-dimensional Object Segmentation Ratio. Geomatics and Information Science of WUHAN University .2020 ,Oct; 10(45):1626-1632