A Method of Measuring Shape Similarity between multi-scale objects

Dongliang Peng, Min Deng

Department of Geo-informatics, Central South University, Changsha, China

Email: pengdlzn@163.com; dengmin028@yahoo.com

1. Introduction

Similarity measure is a key issue in evaluation of map generalization, object matching and object recognition. The measures of similarity include shape similarity, location similarity and semantic content similarity (Frank & Ester, 2006). Among these similarity measures, the shape similarity measure is very important because of the easy collecting of the necessary parameters and the well matching of human intuition (Zhang *et al*., 2002). As a matter of fact, most existing shape similarity measures (e.g. Arkin *et al*., 1991; Basri *et al*., 1998; Veltkamp & Latecki, 2006) are developed based on the concept of dissimilarity by measuring the distance or cost between objects. In practice, it is not as easy as a real similarity measure to distinguish how similar two objects are.

Among the shape dissimilarity measures, the turning function based method proposed by Latecki & Lakämper (2000) is an impactful one. In this method, the most possible correspondence of the maximal convex/concave arcs is established first, and then the integral azimuth of every edge is computed. Finally, the integral of turning angles (the difference of integral azimuths) between corresponding edges is defined as the shape dissimilarity. As for two vectors ($\vec{a}, \vec{b}$), there is always an angle θ between them in [0, π], therefore an effective way is proposed by Charikar (2002) to translate the turning angle idea to similarity measure, that is, defining the similarity of two vectors ($\vec{a}, \vec{b}$) as 1- $\theta(\vec{a}, \vec{b})/\pi$.

Note that corresponding objects in different scales maps are similar in shape. The corresponding vectors can be well established based on the similar structures. Moreover, except the angle itself, the similarity is also influenced by the extension of the angle, which can be revealed by using the lengths of the vectors as weights. Through these analyses, a new method of measuring shape similarity of corresponding multi-scale objects is proposed in this study.

2. The new method of measuring shape similarity

2.1 The establishment of corresponding vectors

To establish the corresponding vectors, the classic Binary Line Generalization trees (BLG-trees in abbreviation, van Oosterom & van den Bos, 1989) of corresponding objects are built, which are constructed upon the Douglas-Peucker algorithm (Douglas & Peucker, 1973), and can represent objects in a hierarchy with increasing accuracy in the lower levels of the trees. Particularly, the line segments stored in the nodes of the BLG-tree are used for similarity measuring, which is different from the original BLG-tree for line generalization. Therefore, the line segments will be stored in the nodes, instead of the baseline of the line segment. As shown in figure 1, the line segment $\widehat{p_0 p_2}$ instead of the straight line $\overline{p_0 p_2}$ is stored in the corresponding nodes.

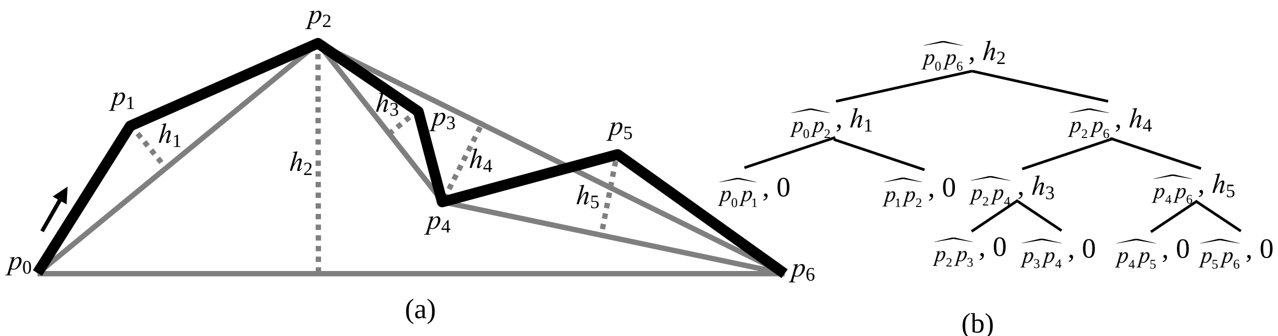


Figure 1. The structural representation of a linear object.
(a) The linear object. (b) The BLG-tree of the linear object.

After building the BLG-trees for the corresponding objects, the matching of their nodes will occur. Here the ratio of the lengths of the baselines is used as the indicator for the matching, which is expressed as:

$$Ratio_{BL}\left(L_1, L_2\right) = \frac{Length(BaseLine(L_1))}{Length(BaseLine(L_2))} \tag{1}$$

Where, the baselines for line segments $L_1$and L_2 are the straight lines connected by the starts and the ends of the two line segments.

Considering that errors always exist, $Ratio_{BL}(L_1, L_2)$ doesn't necessarily equal to 1. To allow these errors, a range U is considered so that line segments $L_1$and L_2 is determined as correspondence if there satisfies

$$Ratio_{BL}\left(I_1, I_2\right) \in U \tag{2}$$

Simply, the range U is defined as follows:

$$U = \left[1 - T_L, 1/(1 - T_L)\right], \quad T_L \in \left[0, 1\right] \tag{3}$$

Where, T_L is a parameter to control the range. As the lengths of corresponding baselines should be very close, a high value 0.98 is simply set to T_L.

The matching of the nodes is a recursive process. The array *CorrespondingNodes* is defined to record the corresponding nodes in the matching process.

- For the current corresponding nodes, if they both have child nodes and their left- and right-child nodes correspond to each other, that is, equation (2) is satisfied, in this case the matching is successful. Further the two pairs of corresponding nodes will be regarded as current corresponding nodes for a deeper matching process.
- Otherwise, the current corresponding nodes are recorded in *CorrespondingNodes*.

To the corresponding nodes recorded in *CorrespondingNodes*, the starts and ends of their stored line segments can be used as breaks to split the two objects. Therefore, a set of corresponding line segments are obtained. Finally, corresponding vectors are established by using a linear interpolation algorithm (Nöllenburg *et al*., 2008) for every pair of corresponding line segments. To be clear, Figure 2 shows an example of the establishment of corresponding vectors.

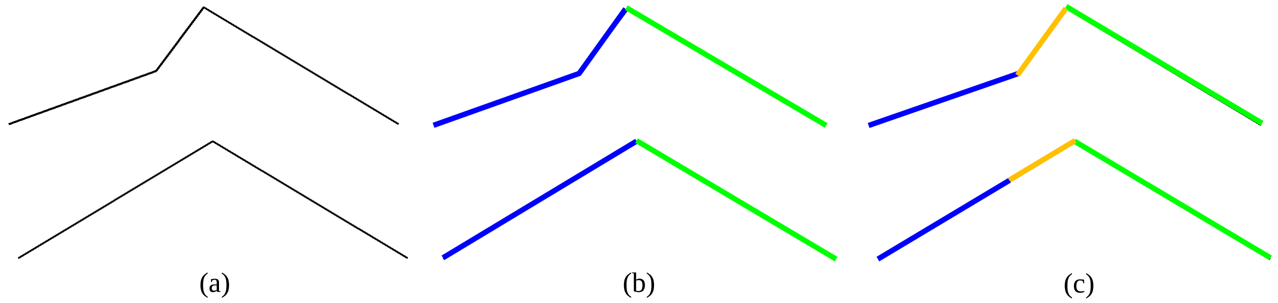


Figure 2. The establishment of corresponding vectors. (a) The original linear features. (b) The corresponding line segments obtained based on BLG-trees. (c) The corresponding vectors established based on the linear interpolation algorithm.

2.2 The definition of shape similarity measure

It can be deduced that the angle between two vectors is in [0, π], and a smaller angle indicates that these two vectors are more consistent in direction. In particular, 0 means same directions and π means counter directions. Furthermore, the lengths of the two vectors are used to define the weight to reveal the extension of the angle. As a result, the shape similarity measure is defined as:

$$sim(A,B) = 1 - \frac{\sum_i (|\vec{a}_i| + |\vec{b}_i|)\theta(\vec{a}_i, \vec{b}_i)}{\pi(|A| + |B|)} \tag{4}$$

Where, A and B are the two objects, and $|A|$, $|B|$ are respectively their lengths; $\vec{a}_i$ and $\vec{b}_i$ are a pair of corresponding vectors, and $\theta(\vec{a}_i, \vec{b}_i)$ is the angle between them.

3. Case study

3.1 Similarity measuring of a set of basic objects

First, a set of basic objects, including circle, octagon, hexagon, pentagon, quadrangle and triangle, are selected for shape similarity measuring. The computational results of the shape similarity between the set of basic objects are shown in figure 3. Based on the circle row, the similarities decrease according to the reducing of points, which satisfy human recognition well. Among these similarities, the one between quadrangle and triangle is 0.806 which is the minimum.

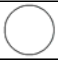
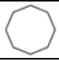
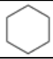
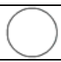
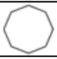
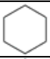
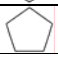
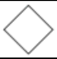
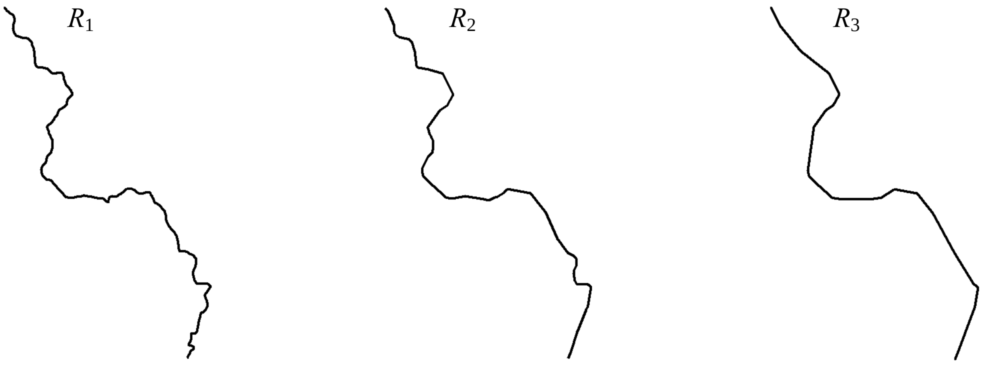
						
	1	0.938	0.917	0.900	0.875	0.833
	0.938	1	0.903	0.887	0.875	0.826
	0.917	0.903	1	0.878	0.861	0.833
	0.900	0.887	0.878	1	0.850	0.822
	0.875	0.875	0.861	0.850	1	0.806
	0.833	0.826	0.833	0.822	0.806	1

Figure 3. The shape similarities of some basic objects obtained by the method proposed in this paper

3.2 Similarity measuring of railway polylines

The railway polylines are from National Fundamental Geographic Information System of China, as shown in figure 4. Based on our method, $sim(R_1,R_2)$ is 0.923, $sim(R_1,R_3)$ is 0.886 and $sim(R_2,R_3)$ is 0.927. $sim(R_1,R_2)$ and $sim(R_2,R_3)$ are quite close because the corresponding scale differences are the same, which are both 5,000,000. Moreover, since the scale difference is larger, $sim(R_1,R_3)$ is smaller than $sim(R_1,R_2)$ and $sim(R_2,R_3)$. To sum up, the results are reasonable.



(a) Scaled at 1: 5,000,000 (b) Scaled at 1: 10,000,000 (c) Scaled at 1: 15,000,000

Figure 4. The railway polylines

4. Conclusions

A method of measuring shape similarity between multi-scale objects is proposed in this paper. First, corresponding vectors are established by BLG-tree based method. Second, shape similarity is computed based on the turning angles of corresponding vectors. The case study demonstrates that the proposed method is reasonable and satisfies human cognition. Furthermore, this method doesn't depend on scale and position of objects, but depends on objects' orientation.

References

Arkin, E. M., Chew, L. P., Huttenlocher, D. P., Kedem K. & Mitchell J. S. B., 1991. An Efficiently Computable Metric for Comparing Polygonal Shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(3), pp. 209-216.

Basri, R., Costa, L., Geiger, D. & Jacobs D., 1998. Determining the similarity of deformable shapes. *Vision Research*, 38, pp. 2365-2385.

Charikar, M. S., 2002. Similarity Estimation Techniques from Rounding Algorithms. In: *Proceedings of the 34th annual ACM symposium on Theory of computing*, New York, pp. 380-388.

Douglas, D. H., Peucker, T. K., 1973. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *The Canadian Cartographer*, 10(2), pp.112-122.

Frank, R., Ester, M., 2006. A quantitative similarity measure for maps. In: *Proceedings of the 12th International Symposium*, *Progress in Spatial Data Handling*, Springer, pp. 435-450.

Latecki, L. J. & Lakämper, R., 2000. Shape Similarity Measure Based on Correspondence of Visual Parts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10), pp. 1185-1190.

Nöllenburg M., Merrick D., Wolff A. & Benkert M., 2008. Morphing polylines: A step towards continuous generalization. *Computers, Environment and Urban Systems*, 32, pp. 248-260.

van Oosterom, P. & van den Bos, J., 1989. An Object-Oriented Approach to the Design of Geographic Information Systems. In: *the first symposium on Design and implementation of large spatial databases*, NewYork, pp. 255-269.

Veltkamp, R. C. & Latecki L. J., 2006. Properties and Performance of Shape Similarity measures. In: *Proceedings of IFCS 2003 Conference: Data Science and Classification*, pp. 1-9.

Zhang, Q., Li, D., Gong, J., 2002. Shape Similarity Measures of Linear Entities. *Geo-spatial Information Science*, 5(2), pp. 62-67.