Exploratory analysis of spatial relations among polygons

Y. Sadahiro

Department of Urban Engineering, University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
Email: sada@ua.t.u-tokyo.ac.jp

1. Introduction

Spatial entities represented as polygons are often in close relations. Land use and land cover are affected by soil and vegetation distributions. Ethnicity distribution is based on other cultural distributions such as race, religion, language, and occupation. Christaller's hierarchical system is a theoretical model of the relations among administrative units and market areas.

The relations between two polygons have been discussed intensively in computer science (Egenhofer and Franzosa 1991; Cohn et al. 1997; Egenhofer and Shariff 1998; Rodriguez and Egenhofer 2004; Schneider and Behr 2006). Many analytical methods are available to describe and analyze the relation between two polygons. Concerning the relations among more than two polygons, on the other hand, analytical methods are quite limited: visual analysis and spatial regression modeling (Anselin 1988, Fotheringham et al. 2002). The former is easy to implement but rather subjective and does not work on numerous polygons. The latter is sophisticated and objective but requires statistical knowledge and research hypothesis usually found in exploratory analysis.

To fill the gap between the two methods, this paper proposes a new exploratory method of analyzing the spatial relations among polygons. The objective of analysis is to understand the global relation among polygons and to detect useful local patterns in the relations among polygons. A focus is on the hierarchical relation among polygons, since spatial hierarchy is an essential concept in geography.

2. Graph Representations of the Relations among Polygons

2.1 Similarity between a pair of polygons

Suppose a set of polygons $\Omega=\{P_1, P_2, ..., P_M\}$ in region S. The relations between a pair of polygons can be described at various levels of details by existing methods mentioned above. Since this paper focuses on the relations among more than two polygons, we start with a basic set of topological relations defined between a pair of polygons. Two polygons P_i and P_j are *hierarchical* if one fully contains the other. The larger polygon is a *higher-level polygon* of the smaller one while the smaller one is a *lower-level polygon* of the larger one.

The similarity between two polygons is evaluated by two measures. *Size distance* between P_i and P_j is given by

$$D_S\left(P_i, P_j\right) = \left|A\left(P_i\right) - A\left(P_j\right)\right|, \tag{1}$$

where $A(P_i)$ is the area of P_i. *Hierarchy distance* is the area of the relative complement of the larger one in the smaller one:

$$D_H\left(P_i, P_j\right) = \min\left(A\left(P_j \setminus P_i\right), A\left(P_i \setminus P_j\right)\right)$$
$$= A\left(P_i \cup P_j\right) - \max\left(A\left(P_i\right), A\left(P_j\right)\right) \cdot \quad (2)$$

The latter evaluates the departure from complete hierarchy. The distance is zero if two polygons are completely hierarchical. It increases as the overlap of polygons decreases.

2.2 Hasse Diagram

Given a set of polygons Ω, we make the intersection of all the polygons with keeping the boundary lines. This yields fragmented polygons whose power set is denoted by Λ. The set Λ and Boolean operations {∩, ∪} form a lattice. The lattice can be visualized as a Hasse diagram, where nodes and links represent polygons and topological relation between them, respectively.

By adding intermediate polygons generated by polygon overlay, the Hasse diagram embeds all the original polygons as nodes in a single graph (Figure 1). It visualizes all the relations among polygons even if they are not hierarchical. Hierarchical relations are represented by either a single link or a set of links in the same direction.

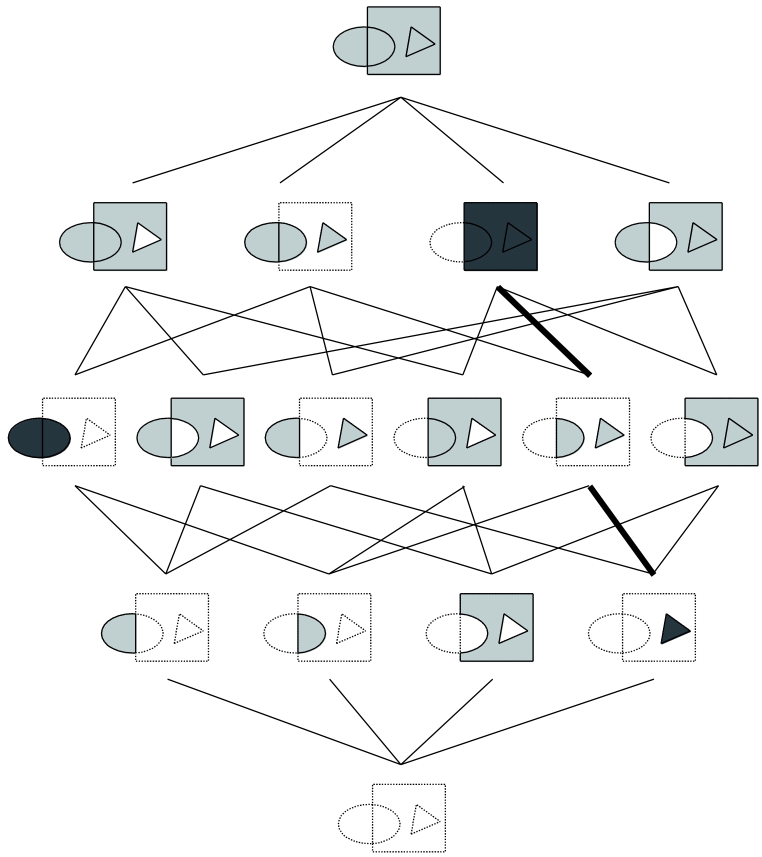

Figure 1. Hasse diagram generated from three polygons (oval, triangle, and rectangle).

2.3 Quasi-Hasse Diagram

The Hasse diagram often becomes too complicated for visual analysis. This paper proposes another graph representation to complement the Hasse diagram.

The distance between polygons P_i and P_j is measured by

$$d\left(P_i,P_j\right)=\frac{D_S\left(P_i,P_j\right)+D_H\left(P_i,P_j\right)}{A\left(P_i\cap P_j\right)}$$
$$=\frac{\left|A\left(P_i\right)-A\left(P_j\right)\right|+\min\left|A\left(P_i\setminus P_j\right),A\left(P_j\setminus P_i\right)\right|}{A\left(P_i\cap P_j\right)}. \tag{3}$$

Given Ω we choose a pair of polygons of the smallest distance and replace them by their intersection. We repeat this until only one polygon remains in Ω. *Intersection tree* is a tree representation of this process as shown in Figure 2.

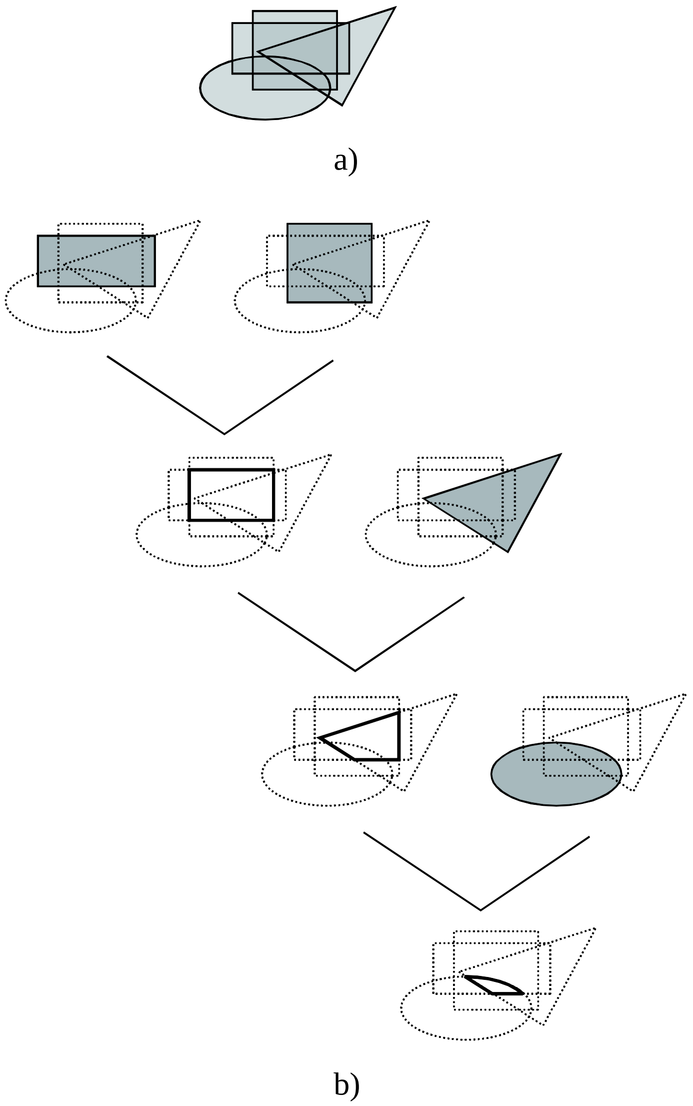

Figure 2. Construction of intersection tree from four polygons. (a) Original polygons, (b) intersection tree generated from the polygons. The two rectangles are chosen first to be replaced by their intersection. It is then overlaid on the triangle and ellipse in turn to generate a small polygon at the bottom of the tree.

Using union overlay instead of intersection overlay, we obtain a *union tree*. We combine it with the intersection tree to generate a *quasi-Hasse diagram* (*QH diagram*). Similar to the Hasse diagram, QH diagram shows the hierarchical relation by a chain of links. Polygons connected directly by a single link are completely hierarchical.

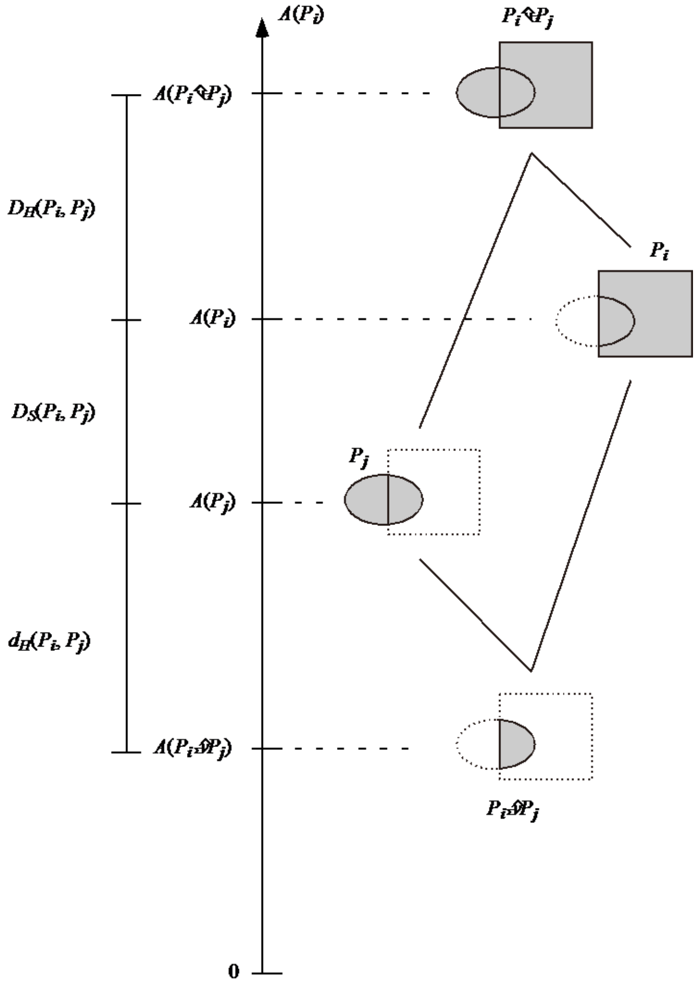

Figure 3. A QH diagram generated from two polygons. The vertical axis indicates the area of polygons.

As seen in Figure 3, similarity measures defined in Section 2.1 can be easily calculated in QH diagram. If two polygons are not completely but almost hierarchical, the shorter of two links connecting the polygons with their intersection or union becomes negligibly short so that the polygons look like connected directly by a single link. This helps us to obtain a general view of hierarchical relations among polygons.

Intersection and union trees can be also used for the classification of polygons. Subtrees naturally define groups of similar polygons that share a common property.

3. Empirical Application

This section applies the proposed method to analysis of the result of experiment in environmental psychology. The objective of experiment is to analyze the mental image of Aoyama area in Tokyo, Japan. 53 undergraduate and 17 graduate students were handed base maps of Aoyama area, and asked to draw a polygon indicating the extent of Aoyama area that they imagined in their mind.

A wide variety exists in the mental image of Aoyama area (Figure 4). We applied the method proposed in the previous section to understand the overall relation among mental images and to classify them into groups of similar properties. The results are shown in Figures 5 and 6.

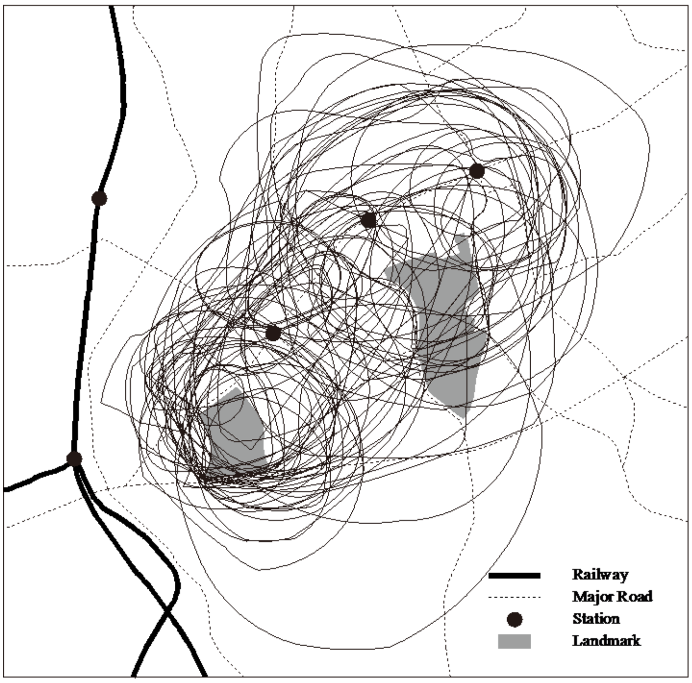

Figure 4. Mental images of Aoyama area drawn by 70 students.

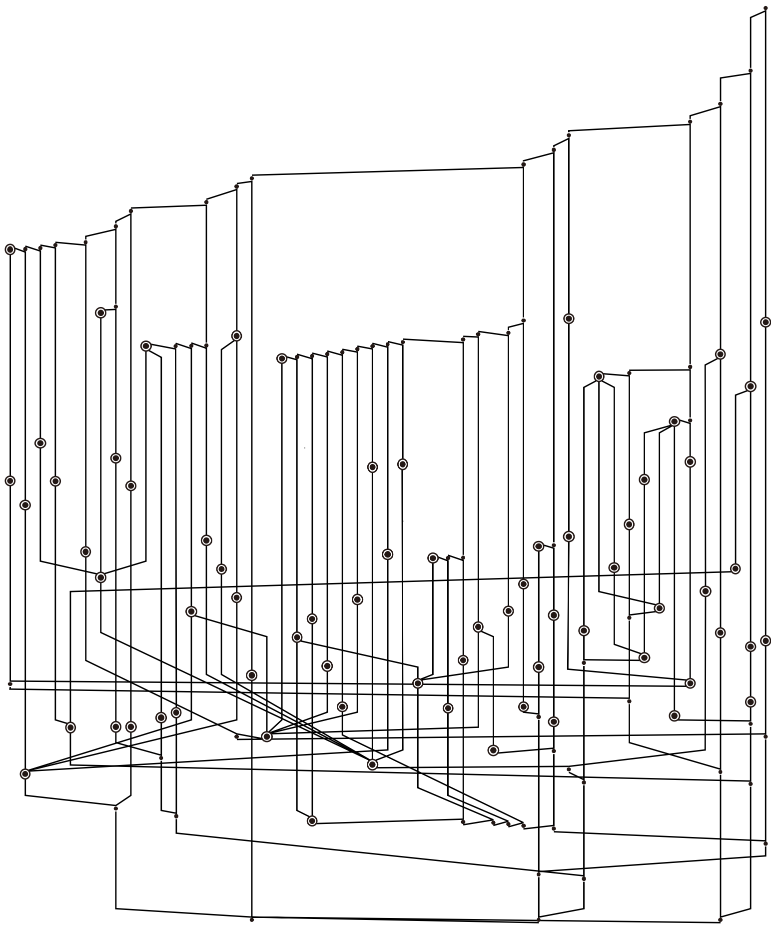

Figure 5. QH diagram representing the relations among mental images of Aoyama area.
Circles indicate the original images.

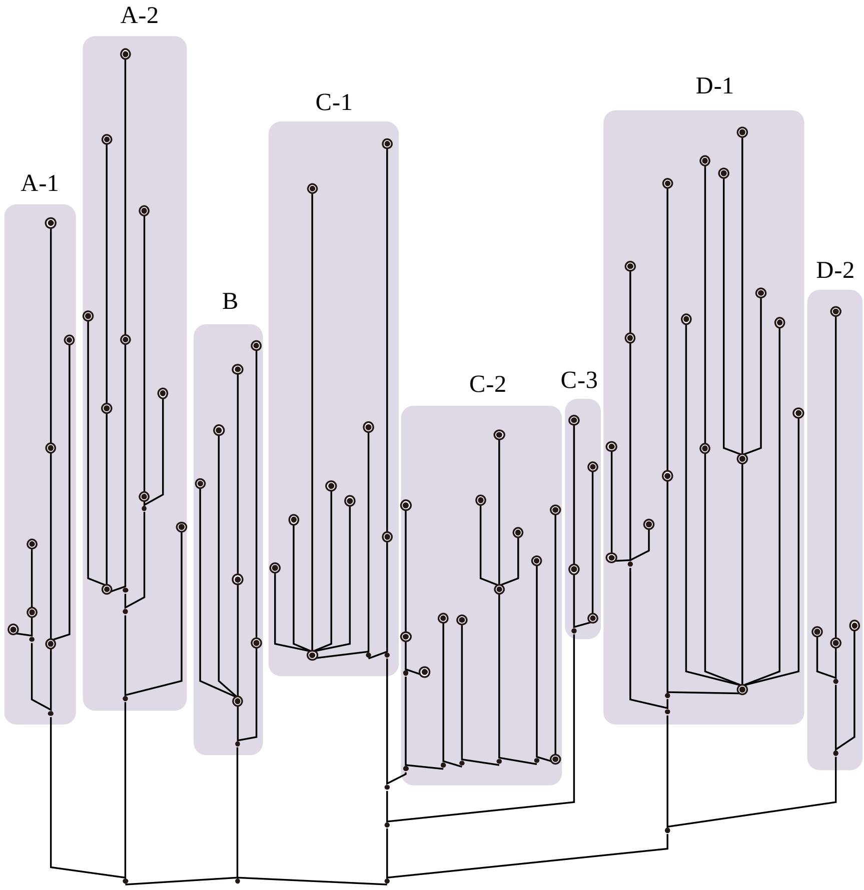

Figure 6. Classification of mental image by intersection tree. Gray shades indicate groups of similar images.

Classification of mental images by intersection tree suggests four groups (A, B, C, and D), each of which is further classified into smaller groups.

Maps of mental images are shown in Figure 7. Images in Figure 7a are clustered around Aoyama 1-Chome Station. Those in group A-1 cover the northeast region of the station while group A-2 spreads to the opposite side of the station with partially covering Aoyama Cemetery. Group C in Figure 7c is a typical example of multi-level hierarchical structure. All the images in this group share the same landmark (Aoyama Gakuin University), and they are classified into three groups each of which has a common larger area containing the university and other different landmarks. Group C-1 covers a larger area from the university to the north and east ends of the study area. Images in C-2 are concentrated around the north of the university with covering Omote Sando Station. Images in C-3 are located to the south of those in C-2. Groups D-1 and D-2 spread from Omote Sando and Aoyama 1-Chome Stations to the outer boundary of the study area. Images in D-2 are smaller than those in D-1.

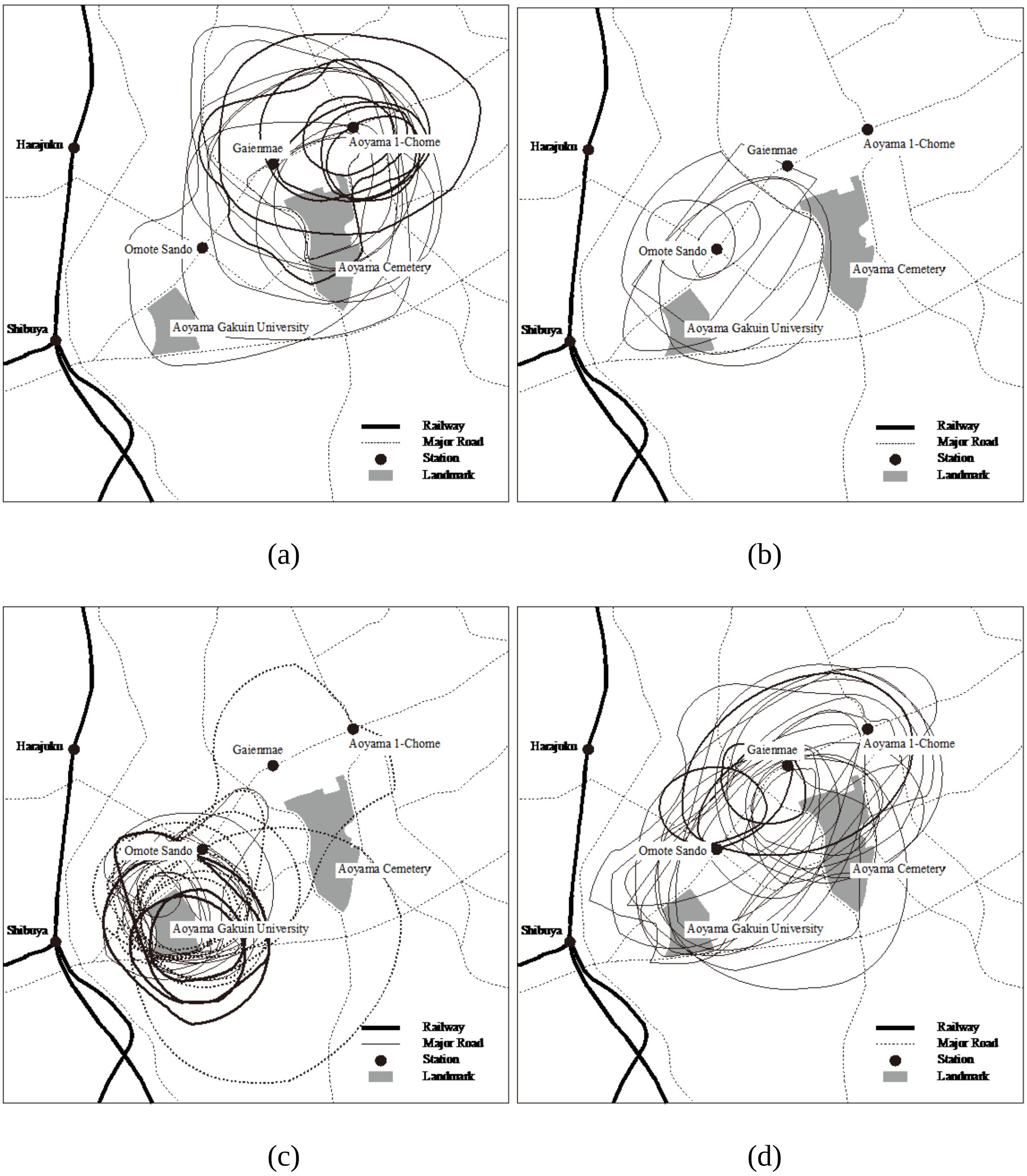

Figure 7. Mental images classified by intersection tree. (a) Group A (bold lines: A-1, thin lines: A-2), (b) group B, (c) group C (bold dotted lines: C-1, thin lines: C-2, bold lines: C-3), (d) group D (thin lines: D-1, bold lines: D-2).

The above result can be partially explained by the attributes of students. Group C mainly consists of mental images of undergraduate students while groups A and D are those of graduate students. This is probably because young students are only familiar with Aoyama Gakuin University and its neighborhood while senior students have experience of visiting high-end boutiques and fancy restaurants around Omote Sando and Aoyama 1-Chome Stations.

Frequency of visiting this area also seems to affect the mental image. In general, the mental image of Aoyama area expands with frequency. Frequent visitors in groups A and D drew large circles. On the other hand, undergraduate students who rarely visit this area drew very small circles around landmarks whose name contains “Aoyama” such as Aoyama 1-Chome Station and Aoyama Gakuin University.

4. Conclusion

This paper has proposed a new method of analyzing the relations among polygons. The QH diagram permits us to grasp the whole structure of the relations among polygons, to detect local spatial patterns in the relations, and to classify the polygons into groups of similar properties. The application of the method revealed the properties of the diagram as well as provided useful findings in environmental psychology.

In future research, further variation of the spatial relations between polygons such as those mentioned earlier should be considered. While the refinement of spatial relations permits us detailed description of the whole structure of polygon relations, its visual representation inevitably becomes much complicated. A more sophisticated method of visualization has to be developed. Vagueness and uncertainty in spatial representation and measurement should also be addressed in future (Winter 2000). The topological relation between polygons can be affected by only a slight shift of location. One option to treat this problem is to convert polygons into continuous surfaces by incorporating location uncertainty in the smoothing functions. Statistical methods, fuzzy theory and other approaches should be discussed. In addition, extension of the method to the spatiotemporal domain is important. Research questions include 1) how do we describe the relations among polygons of different times, and 2) how do we analyze the change of the relations among different polygons?

References

Anselin L, 1988, *Spatial econometrics: Methods and models.* Kluwer, Dordrecht.

Cohn AG, Bennett B, Gooday J and Gotts MM, 1997, Qualitative spatial representation and reasoning with the region connection calculus. *Geoinformatica,* 1: 275-316.

Egenhofer MJ and Franzosa RD, 1991, Point set topological spatial relations. *International Journal of Geographical Information Systems,* 5: 161-174.

Egenhofer MJ and Shariff ARBM, 1998, Metric details for natural-language spatial relations. *ACM Transactions on Information Systems,* 16: 295-321.

Fotheringham AS, Brunsdon C and Charlton M, 2002, *Geographically weighted regression: The analysis of spatially varying relationships*. Wiley, Chichester.

Rodriguez, MA and Egenhofer MJ, 2004, Comparing geospatial entity classes: an asymmetric and context-dependent similarity measure. *International Journal of Geographical Information Science,* 18: 229-256.

Schneider M and Behr T, 2006, Topological relationships between complex spatial objects. *ACM Transactions on Database Systems,* 31: 39-81.

Winter S, 2000, Uncertain topological relations between imprecise regions. *International Journal of Geographical Information Science,* 14: 411-430.