Discussion Paper Series
No. 96r1

Department of Urban Engineering
University of Tokyo

Analysis of the relation among spatial tessellations

Yukio Sadahiro

Department of Urban Engineering, University of Tokyo

Analysis of the relation among spatial tessellations

Yukio Sadahiro
Department of Urban Engineering, University of Tokyo

Abstract

Spatial tessellation is one of the most important spatial structures in geography. It represents census tracts, postal zones, school districts, soil patterns, and so forth. Spatial tessellations in the same region are often closely related with each other. School districts and postal zones are sometimes based on administrative units, while market areas of drug stores and those of supermarkets affect with each other. This paper proposes a method for analyzing the relation among spatial tessellations. The relation between two tessellations is described by four topological relations, whose closeness is evaluated quantitatively by three distance measures. The relation among more than two tessellations is represented by a Hasse diagram and visualized by tree representations. The method is applied to the analysis of two sets of spatial tessellations in Japan: candidate plans for the new administrative system called 0reveals the properties of the method and quantitative measures used in analysis.

1. Introduction

Spatial tessellation is one of the most important spatial structures in geography. Some are defined for administrative purposes such as census tracts, postal zones, electoral and school districts. Others are based on natural phenomena that include land cover, vegetation and soil patterns. Market areas of retail stores, say, drug stores and gas stations, are often approximated by a set of polygons such as Voronoi diagrams.

Spatial tessellations in the same region are often closely related with each other (Okabe et al. 2000, Sadahiro 2002). School districts and postal zones are often based on administrative units. TAZ (Transportation Analysis Zones) are determined by census tracts and administrative units. A close relation exists between administrative units and land use pattern. Market areas of different categories of shops affect with each other because of consumers' propensity for one stop multipurpose shopping. To understand a spatial tessellation and its underlying structure, it is necessary to analyze not only the tessellation itself but also its relation with others representing different spatial phenomena.

Analysis of the relation among spatial tessellations starts with visual comparison of tessellations. Visual analysis is a useful means to obtain research hypothesis about spatial phenomena (Nielson *et al.* 1997, Kraak 2003, Slocum *et al.* 2004, Ware 2004, Wright 2006). Visual comparison of choropleth maps helps us to detect the similarities and differences among tessellations.

Visual analysis is often followed by quantitative analysis. Basic statistics are available such as the average area and perimeter of regions, their variance and standard deviation, spatial mean of their gravity centers, and so forth. Since they effectively summarize the geometrical properties of tessellations as numerical measures, they permit us to compare and analyze numerous tessellations simultaneously.

One disadvantage of the above statistics is that they do not detect differences in spatial arrangement of regions. They are calculated for each tessellation separately and compared among tessellations. Consequently, basic geometrical transformations such as translation, rotation and reflection are not reflected in statistics values.

Difference in spatial arrangement is not negligible in spatial analysis. To resolve the problem, this paper proposes a new method for analyzing spatial tessellations. It focuses on spatial tessellations of different attribute types defined in the same region. Tessellations defined by the same or comparable variables such as time-series vegetation maps and land cover maps obtained from different data sources are out of scope of this paper (discussions can be found in Congalton and Mead 1983, Rosenfield and Fitzpatrick-Lins 1986, Monserud and Leemans 1992, Pontius 2000, 2002, and Fritz and See 2005). The emphasis is on spatial rather than attribute relation among tessellations.

Section 2 describes a method for analyzing the relation among spatial tessellations. Topological relations between tessellations are defined, which is followed by numerical evaluation and visual representation of the relations. To discuss the benefits and limitations of the method, Section 3 applies it to the analysis of two sets of tessellations: candidate plans for the new administrative system in Japan and areas covered by branches of private companies. Section 4 summarizes the conclusions with a discussion.

2. Method

Suppose a set of spatial tessellations $\Omega=\{\Omega_1, \Omega_2, ..., \Omega_n\}$ defined in region S. Tessellation Ω_i consists of m_i regions $\{T_{i1}, T_{i2}, \ldots, T_{im_i}\}$. The relation between a pair of tessellations in Ω is discussed first, which is followed by analysis of the relation among more than two tessellations.

2.1 Topological relation between a pair of spatial tessellations

Topological relation between Ω_i and Ω_j is classified into one of the four categories: 1) identical, 2) hierarchical, 3) compatible, and 4) incompatible.

Two tessellations Ω_i and Ω_j are *identical* if any region in Ω_i perfectly fits a region in Ω_j and vice versa. This relation is more restrictive than "congruent" used in geometry, because the former does not allow congruent transformation such as translation and rotation.

If two tessellations are not identical, weaker relations are considered. Two tessellations Ω_i and Ω_j are *hierarchical* if any region in Ω_i is consistently and fully covered by a region in Ω_j or vice versa. Tessellations Ω_1, Ω_2 and Ω_3 in Figure 1 are hierarchical with each other. This relation is often found in administrative and management systems. If any region in Ω_i is fully covered by a region in Ω_j , Ω_i is a *lower level tessellation* of Ω_j. Tessellation Ω_i is generated by further dividing regions in Ω_j into smaller regions. If Ω_i is obtained by merging regions in Ω_j, Ω_i is a *higher level tessellation* of Ω_j. In Figure 1, Ω_3 is a lower level tessellation of Ω_2, while Ω_1 is a higher level tessellation of Ω_2.

Hierarchical relation holds when either "cover" or "be covered" relation consistently exists between tessellations. On the other hand, if the two relations exist simultaneously between tessellations, it is called *compatible* relation. An example is the relation between Ω_2 and Ω_4 in Figure 1. The largest polygon in Ω_2 fully covers five squares in Ω_4 while the rest four squares in Ω_2 are completely covered by the lower right polygon in Ω_4.

The compatible relation can also be regarded as local hierarchy. The largest polygon in Ω_2 is an upper level tessellation of the five squares in Ω_4, while the rest four squares in Ω_2 compose a lower level tessellation of the largest polygon in Ω_4. Though hierarchical relation is not consistent over the whole region, it is locally consistent between tessellations.

If tessellations Ω_i and Ω_j are neither identical, hierarchical, nor compatible, they are called *incompatible*. In Figure 1, tessellation Ω_6 is incompatible with Ω_4 and Ω_5. This relation appears if regions in two tessellations partially overlap with each other.

2.2 Overlay operations for a pair of tessellations

Two overlay operations are introduced to evaluate the topological relation between tessellations. *Merging overlay* of Ω_i and Ω_j is to make an overlay of Ω_i and Ω_j with keeping only the boundaries shared by both tessellations (Figure 2). It generates a higher level tessellation of Ω_i and Ω_j, which is called *ancestor* of Ω_i and Ω_j denoted by $A(\Omega_i, \Omega_j)$. *Dividing overlay*, on the other hand, keeps all the boundary lines in both Ω_i and Ω_j to yield their *descendant* $D(\Omega_i, \Omega_j)$. It divides regions into smaller ones as shown in Figure 2. If Ω_i and Ω_j are hierarchical, $A(\Omega_i, \Omega_j)$ and $D(\Omega_i, \Omega_j)$ are identical to the higher and lower level ones of Ω_i and Ω_j, respectively.

2.3 Computational evaluation of topological relation between a pair of tessellations

A computational procedure is proposed in the following that evaluates the topological relation between tessellations. To this end, some quantitative measures are defined first as mathematical preparation.

Region T_{ij} (j=1, ..., m_i) of tessellation Ω_i is represented by a binary function:

$$\rho\left(\mathbf{x};T_{ij}\right)=\begin{cases}1 & \text{if } \mathbf{x}\in T_{ij}\\ 0 & \text{otherwise}\end{cases}. \tag{1}$$

Two points $\mathbf{x}_1$ and $\mathbf{x}_2$ are contained in the same region in Ω_i if

$$\begin{aligned}\sigma\left(\mathbf{x}_1,\mathbf{x}_2;\Omega_i\right)&=\begin{cases}1 & \text{if } \sum_j \rho\left(\mathbf{x}_1;T_{ij}\right)\rho\left(\mathbf{x}_2;T_{ij}\right)=1\\ 0 & \text{otherwise}\end{cases}.\\ &=\sum_j \rho\left(\mathbf{x}_1;T_{ij}\right)\rho\left(\mathbf{x}_2;T_{ij}\right)\end{aligned} \tag{2}$$

The area of region T_{ij} is given by

$$a\left(T_{ij}\right)=\int_{\mathbf{x}\in S}\rho\left(\mathbf{x},T_{ij}\right)\mathrm{d}\mathbf{x}. \tag{3}$$

Topological relation between tessellations is evaluated by examining every pair of points in S whether contained in the same or different regions in two tessellations. The relation between $\mathbf{x}_1$ and $\mathbf{x}_2$ denoted by $R(\mathbf{x}_1, \mathbf{x}_2)$ is classified into one of the four categories:

R_{11}: $\mathbf{x}_1$ and $\mathbf{x}_2$ are contained in the same region in both Ω_i and Ω_j.
R_{10}: $\mathbf{x}_1$ and $\mathbf{x}_2$ are contained in the same region in Ω_i but in different regions in Ω_j.
R_{01}: $\mathbf{x}_1$ and $\mathbf{x}_2$ are contained in different regions in Ω_i but in the same region in Ω_j.
R_{00}: $\mathbf{x}_1$ and $\mathbf{x}_2$ are contained in different regions in both Ω_i and Ω_j.

In Figure 3, for instance, $R(\mathbf{x}_1, \mathbf{x}_2)=R_{11}$ because $\mathbf{x}_1$ and $\mathbf{x}_2$ are contained in the same region in both Ω_1 and Ω_2. Similarly, the relation between other points is described as $R(\mathbf{x}_1, \mathbf{x}_3)=R(\mathbf{x}_2, \mathbf{x}_3)=R_{10}$, $R(\mathbf{x}_1, \mathbf{x}_4)=R(\mathbf{x}_2, \mathbf{x}_4)=R_{01}$, and $R(\mathbf{x}_3, \mathbf{x}_4)=R_{00}$.

Granularity of tessellation Ω_i is the ratio of point pairs contained in the same region in Ω_i to all the pairs in S. It is mathematically defined as

$$\begin{aligned} G(\Omega_i) &= \frac{\int_{\mathbf{x}_2 \in S} \int_{\mathbf{x}_1 \in S} \sigma(\mathbf{x}_1, \mathbf{x}_2; \Omega_i) \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}_2}{\{a(S)\}^2} \\ &= \frac{\sum_j \{a(T_{ij})\}^2}{\{a(S)\}^2} \end{aligned} \tag{4}$$

($0<G(\Omega_i)\leq 1$). It indicates the degree of fragmentation in tessellation Ω_i. A large value implies a tessellation consisting of a few number of large regions.

The ratio of point pairs in R_{kl} is denoted as $r_{kl}(\Omega_i, \Omega_j)$. The ratio $r_{11}(\Omega_i, \Omega_j)$ is given by

$$r_{11}(\Omega_i, \Omega_j) = \frac{1}{\{a(S)\}^2} \int_{\mathbf{x}_1 \in S} \int_{\mathbf{x}_2 \in S} \sigma(\mathbf{x}_1, \mathbf{x}_2; \Omega_i) \sigma(\mathbf{x}_1, \mathbf{x}_2; \Omega_j) \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}_2 . \tag{5}$$

Similarly,

$$\begin{aligned} r_{10}(\Omega_i, \Omega_j) &= \frac{1}{\{a(S)\}^2} \int_{\mathbf{x}_1 \in S} \int_{\mathbf{x}_2 \in S} \sigma(\mathbf{x}_1, \mathbf{x}_2; \Omega_i)\left(1-\sigma(\mathbf{x}_1, \mathbf{x}_2; \Omega_j)\right) \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}_2 \\ &= \frac{1}{\{a(S)\}^2} \int_{\mathbf{x}_1 \in S} \int_{\mathbf{x}_2 \in S} \sigma(\mathbf{x}_1, \mathbf{x}_2; \Omega_i) \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}_2 - r_{11}(\Omega_i, \Omega_j) \\ &= G(\Omega_i) - r_{11}(\Omega_i, \Omega_j) \end{aligned} , \tag{6}$$

$$\begin{aligned} r_{01}(\Omega_i, \Omega_j) &= \frac{1}{\{a(S)\}^2} \int_{\mathbf{x}_1 \in S} \int_{\mathbf{x}_2 \in S} \left(1-\sigma(\mathbf{x}_1, \mathbf{x}_2; \Omega_i)\right) \sigma(\mathbf{x}_1, \mathbf{x}_2; \Omega_j) \mathrm{d}\mathbf{x}_1 \mathrm{d}\mathbf{x}_2 \\ &= G(\Omega_j) - r_{11}(\Omega_i, \Omega_j) \end{aligned} , \tag{7}$$

and

$$\begin{aligned} r_{00}\left(\Omega_i,\Omega_j\right) &= \frac{1}{\left\{a\left(S\right)\right\}^2}\int_{\mathbf{x}_1\in S}\int_{\mathbf{x}_2\in S}\left(1-\sigma\left(\mathbf{x}_1,\mathbf{x}_2;\Omega_i\right)\right)\left(1-\sigma\left(\mathbf{x}_1,\mathbf{x}_2;\Omega_j\right)\right)\mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_2 \\ &= 1+r_{11}\left(\Omega_i,\Omega_j\right)-G\left(\Omega_i\right)-G\left(\Omega_j\right) \end{aligned} \quad (8)$$

The relationship among the above ratios is summarized in Table 1.

These ratios allow us to identify identical and hierarchical relations between Ω_i and Ω_j. Tessellations Ω_i and Ω_j are identical if $r_{10}(\Omega_i, \Omega_j)=r_{01}(\Omega_i, \Omega_j)=0$. If either $r_{01}(\Omega_i, \Omega_j)=0$ or $r_{10}(\Omega_i, \Omega_j)=0$, Ω_i and Ω_j are hierarchical. In the former case Ω_i is a higher level tessellation of Ω_j, while Ω_i is a lower level tessellation of Ω_j in the latter case.

Compatible and incompatible relations are identified as follows. Tessellations Ω_i and Ω_j are incompatible if two regions in different tessellations partially overlap with each other. Examples include the upper left rectangle in Ω_1 and the middle rectangle in Ω_2 in Figure 3. Partial overlap is characterized by point pairs such as $\mathbf{x}_3$ and $\mathbf{x}_4$ that are contained in different regions in both Ω_i and Ω_j but in the same region in their ancestor $A(\Omega_i, \Omega_j)$. Consequently, Ω_i and Ω_j are compatible if any region in $A(\Omega_i, \Omega_j)$ does not have a point pair contained in different regions in both Ω_i and Ω_j.

The above condition is mathematically represented as follows. The ratio of point pairs contained in the same region in $A(\Omega_i, \Omega_j)$ is $G(A(\Omega_i, \Omega_j))$. The ratio of point pairs in different regions in both Ω_i and Ω_j is $r_{00}(\Omega_i, \Omega_j)$. Since point pairs in the same region in either Ω_i or Ω_j are also contained in the same region in $A(\Omega_i, \Omega_j)$, Ω_i and Ω_j are compatible if

$$\begin{aligned} G\left(A\left(\Omega_i,\Omega_j\right)\right) &= 1-r_{00}\left(\Omega_i,\Omega_j\right) \\ &= r_{11}\left(\Omega_i,\Omega_j\right)+r_{10}\left(\Omega_i,\Omega_j\right)+r_{01}\left(\Omega_i,\Omega_j\right) \end{aligned} \quad (9)$$

2.4 Measurement of the closeness of relation between spatial tessellations

To evaluate the closeness of relation between tessellations, this subsection proposes three distances for a pair of tessellations: 1) granularity distance, 2) hierarchy distance, and 3) compatibility distance.

Granularity distance evaluates the difference in granularity between tessellations. Granularity distance between Ω_i and Ω_j is given by

$$d_G\left(\Omega_i,\Omega_j\right)=\left|G\left(\Omega_i\right)-G\left(\Omega_j\right)\right|. \quad (10)$$

Hierarchy distance indicates the degree of separation from perfect hierarchical relation. Hierarchy distance between Ω_i and Ω_j is defined as

$$\begin{aligned} d_H\left(\Omega_i,\Omega_j\right) &= \min\left(r_{10}\left(\Omega_i,\Omega_j\right), r_{01}\left(\Omega_i,\Omega_j\right)\right) \\ &= \min\left(G\left(\Omega_i\right)-r_{11}\left(\Omega_i,\Omega_j\right), G\left(\Omega_j\right)-r_{11}\left(\Omega_i,\Omega_j\right)\right). \\ &= \min\left(G\left(\Omega_i\right), G\left(\Omega_j\right)\right)-r_{11}\left(\Omega_i,\Omega_j\right) \end{aligned} \tag{11}$$

The distance is zero if two tessellations are completely hierarchical. The distance increases as the relation collapses.

Compatibility distance is a measure of separation from complete compatibility. Tessellations Ω_i and Ω_j are compatible if and only if Equation (9) holds. Consequently, compatibility between Ω_i and Ω_j can be measured by the difference between the left and right sides of the equation:

$$\begin{aligned} d_C\left(\Omega_i,\Omega_j\right) &= G\left(A\left(\Omega_i,\Omega_j\right)\right)-\left(1-r_{00}\left(\Omega_i,\Omega_j\right)\right) \\ &= G\left(A\left(\Omega_i,\Omega_j\right)\right)-r_{11}\left(\Omega_i,\Omega_j\right)-G\left(\Omega_i\right)+r_{11}\left(\Omega_i,\Omega_j\right)-G\left(\Omega_j\right)+r_{11}\left(\Omega_i,\Omega_j\right) \\ &= G\left(A\left(\Omega_i,\Omega_j\right)\right)+r_{11}\left(\Omega_i,\Omega_j\right)-G\left(\Omega_i\right)-G\left(\Omega_j\right) \end{aligned} \tag{12}$$

Substituting Equation (11) into (12), we have

$$d_C\left(\Omega_i,\Omega_j\right) = G\left(A\left(\Omega_i,\Omega_j\right)\right)+\min\left(G\left(\Omega_i\right),G\left(\Omega_j\right)\right)-d_H\left(\Omega_i,\Omega_j\right)-G\left(\Omega_i\right)-G\left(\Omega_j\right) \tag{13}$$

The distance is zero if Ω_i and Ω_j are completely compatible. It increases as the relation diminishes.

2.5 Graph Representation of relation among spatial tessellations

Having discussed the relation between a pair of spatial tessellations, this subsection extends it to the case of more than two spatial tessellations.

Suppose an overlay of all the tessellations in Ω with keeping all the boundary lines. This yields a tessellation $\Lambda(\Omega)$, which is the finest and lowest level tessellation in $\{\Lambda(\Omega)\}\cup\Omega$.

Let $\Psi=\{\Psi_1, \Psi_2, ..., \Psi_N\}$ be the set of all the partitions of $\Lambda(\Omega)$. In each partition, adjacent regions are merged into one region. Merging and dividing overlay operations are denoted by binary operations $\wedge$ and $\vee$, respectively. The algebraic structure $(\Psi, \wedge, \vee)$ is a lattice because of its following properties (Anderson 2002, Davey and Priestly 2002).

1) Commutativity: $\Psi_i \wedge \Psi_j = \Psi_j \wedge \Psi_i$, $\Psi_i \vee \Psi_j = \Psi_j \vee \Psi_i$

2) Associativity: $\Psi_i \wedge (\Psi_j \wedge \Psi_k) = (\Psi_i \wedge \Psi_j) \wedge \Psi_k$, $\Psi_i \vee (\Psi_j \vee \Psi_k) = (\Psi_i \vee \Psi_j) \vee \Psi_k$

3) Absorption: $\Psi_i \wedge (\Psi_i \vee \Psi_j) = \Psi_i$, $\Psi_i \vee (\Psi_i \wedge \Psi_j) = \Psi_i$

The least element of Ψ is $\Lambda(\Omega)$ while the greatest element is the tessellation consisting of one region covering the whole area of S. It also contains all the original tessellations in Ω.

Since a lattice is a partially ordered set, a binary relation representing the hierarchical relation between tessellations can be defined. If Ψ_i is the same or a higher level tessellation of Ψ_j, it is represented as

$$\Psi_j \leq \Psi_i . \tag{14}$$

If $\Psi_j \leq \Psi_i$, Ψ_i is called a successor of Ψ_j, while Ψ_j is a predecessor of Ψ_j. Specifically, if $\Psi_j < \Psi_i$ and there is no Ψ_k such that $\Psi_j < \Psi_k < \Psi_i$, Ψ_i is called an immediate successor of Ψ_j while Ψ_j is an immediate successor of Ψ_i.

We should note that the above terms used in discrete mathematics do not correspond directly to those defined in this paper. For instance, hierarchical relation is in a sense described in the opposite direction. An ancestor of a certain tessellation, which is a higher level tessellation, is called a successor in mathematics. A descendant, a lower level tessellation, is a predecessor. In addition, a descendant is always a successor, but the converse does not always hold. A successor of Ψ_i is also its descendant only if there exists a tessellation in Ψ with which Ψ_i yields the successor by merging overlay.

A partially ordered is often visualized as a Hasse diagram (Davey and Priestley 2002, Pemmaraju and Skiena 2003). It is a simplest representation of a partially ordered set that conveys the least information necessary to specify the structure of a partially ordered set. Nodes represent tessellations, while links indicate immediate successors and predecessors of tessellations (Figure 4). Immediate successors of Ψ_4, for instance, are Ψ_9, Ψ_{11}, and Ψ_{12} which are connected upward directly with Ψ_4.

Tracing links upward, we can find successors of a tessellation, which are higher level tessellations. Downward tracing tells us lower level tessellations. Tessellations reachable by unidirectional tracing are hierarchical, while non-hierarchical tessellations are not reachable by unidirectional tracing.

In a Hasse diagram, the ancestor of a pair of tessellations is their successor of the lowest level. In Figure 4, for instance, tessellations Ψ_4 and Ψ_5 have two common successors Ψ_9 and Ψ_{15}, and the lower successor Ψ_9 is their ancestor obtained by merging overlay. Tessellations Ψ_4 and Ψ_7 share only one successor Ψ_{15}, and thus it is their ancestor. On the other hand, the descendant of a pair of tessellations is their predecessor of the highest level. The predecessors of Ψ_{11} and Ψ_{14} are Ψ_1 and Ψ_7, the latter of which is the descendant of Ψ_{11} and Ψ_{14}.

Merging and dividing overlays that generate an ancestor and a descendant from two tessellations are not always represented as a pair of links connecting directly the original tessellations and their ancestor or descendant. Ancestor and descendant of Ψ_4 and Ψ_7 are Ψ_{15} and Ψ_1, where overlay operations are both represented as a pair of two links. In contrast, ancestor and descendant of Ψ_4 and Ψ_8 are Ψ_1 and Ψ_{15}, each of which is connected with their original tessellations by indirect links.

2.6 Tree representations of the relation among spatial tessellations

Hasse diagram is a basic representation of the relation among elements in a partially ordered set. However, if a set consists of numerous elements, Hasse diagram inevitably becomes so complicated that is difficult to understand and interpret the whole structure of relation. In addition, Hasse diagram visualizes only complete hierarchical relation between tessellations. It is too restrictive because tessellations in almost hierarchical relation are often found in the real world. They are not completely hierarchical but it is practically important to find such imperfect relations as well as perfect ones. To resolve the above problems, this subsection proposes tree representations of the relation among spatial tessellations.

Ancestor tree of the set of tessellations $\Omega=\{\Omega_1, \Omega_2, ..., \Omega_n\}$ is constructed as follows. First, from all the $n(n-1)/2$ pairs of tessellations in Ω, a pair of tessellations in the closest relation is chosen. The closeness of relation between Ω_i and Ω_j is evaluated by a distance measure $d(\Omega_i, \Omega_j)$, whose definition is based on granularity, hierarchy, compatibility distances and so forth (details will be discussed in Section 3). If Ω_i and Ω_j are chosen, merging overlay generates their ancestor $A(\Omega_i, \Omega_j)$ to replace them in Ω. We repeat this process until Ω consists of only one tessellation.

The above process is represented as a tree graph what we call an ancestor tree (Figure 5). Nodes indicate original tessellations in Ω and their ancestors, while links represent merging overlay operation. The vertical axis indicates the granularity of tessellations. An ancestor tree grows upward from original tessellations.

A *descendant tree* is constructed by using dividing overlay instead of merging overlay in the above process. It is generated by connecting pairs of tessellations with two links to generate their descendant in order of distance measure $d(\Omega_i, \Omega_j)$. A descendant tree extends downward from original tessellations.

In ancestor and descendant trees, the distances defined by Equations (10)-(13) are represented as differences in granularity (Figure 6). Suppose two tessellations Ω_1 and Ω_2, where $G(\Omega_1) \leq G(\Omega_2)$. Granularity distance $d_G(\Omega_1, \Omega_2)$ is the difference in granularity between Ω_1 and Ω_2. Hierarchy distance $d_H(\Omega_1, \Omega_2)$, on the other hand, is the difference in granularity between Ω_1 and $D(\Omega_1, \Omega_2)$. Compatibility distance $d_C(\Omega_1, \Omega_2)$ is given by subtracting $d_H(\Omega_1, \Omega_2)$ from the difference in granularity between $A(\Omega_1, \Omega_2)$ and Ω_2.

In either ancestor or descendant tree, tessellations in complete hierarchical relation are connected by either a single link or a set of unidirectional links. A multilevel hierarchy is represented as a unidirectional chain of more than two tessellations of different granularity such as that indicated by the dotted line in Figure 7. Chains of tessellations that share a common ancestor or descendant form a cluster of tessellations.

If two tessellations are not completely but almost hierarchical, the shorter link between the tessellations and their ancestor/descendant becomes so short that the tessellations look connected directly by a single link. This allows us to find almost hierarchical relation between tessellations in ancestor and descendant trees.

Two compatible tessellations form a parallelogram with their ancestor and descendant. It is confirmed Figure 6 because $d_C(\Omega_1, \Omega_2)=0$ when Ω_1 and Ω_2 are compatible.

Ancestor and descendant trees are not subsets but deduced representations of Hasse diagram. Both trees consist of at most $n(n-1)/2$ nodes and $n(n-1)/2-1$ links while a Hasse diagram has a far more dense structure. Every link in the trees corresponds to a single or a chain of links in Hasse diagram. A link in ancestor and descendant trees can be deduced by tracing links in Hasse diagram upward or downward. Ancestor and descendant trees are even simpler representation than Hasse diagram suitable for visual analysis of the relation among numerous tessellations.

Ancestor and descendant trees may look quite similar to dendrograms used in cluster analysis. They are all tree representation of similarity relation among tessellations. However, ancestor and descendant trees are different from dendrograms in the following aspects. First, original tessellations are not aligned at the same level of the trees except when they all have the same granularity. Second, original tessellations can be located in the middle of trees if they are completely hierarchical. Third, while dendrograms grow from the bottom to the top, ancestor and descendants trees are generated from any pair of tessellations if they are in the closest relation. Fourth, the link length indicates the similarity in granularity, not the overall closeness of relation between tessellations. It is a measure of the closeness of relation only if the relation is evaluated only in terms of granularity by neglecting hierarchical and compatible relations.

3. Application

Using the method proposed above, this section analyzes two sets of tessellations both of which cover the whole area of Japan. One is candidate plans for a new administrative system called Doshusei system, a set of administrative units now under consideration. The other is a set of areas covered by branches of private companies. They consist of 33 and 34 tessellations, respectively. All the tessellations are based on prefecture units, each of which is contained in one region without division in every tessellation. Figure 8 shows an example of a candidate plan for Doshusei system.

The purpose of analysis is to classify tessellations into several groups. In classification, hierarchical relation should be considered explicitly as well as granularity, because administration and management often adopt multilevel hierarchical system. Tessellations in close hierarchical relation should be detected. Distance measure $d(\Omega_i, \Omega_j)$ is defined to meet this objective in four different ways.

A natural choice is hierarchy distance $d_H(\Omega_i, \Omega_j)$ that evaluates the separation from perfect hierarchy. In ancestor and descendant trees it is reflected in the length of the shorter link from Ω_i and Ω_j to their ancestor or descendant (recall Figure 6). Consequently, to evaluate the separation from perfect hierarchy, method 1 defines $d(\Omega_i, \Omega_j)$ as

$$d\left(\Omega_i,\Omega_j\right)=d_H\left(\Omega_i,\Omega_j\right)+d_C\left(\Omega_i,\Omega_j\right) \tag{15}$$

and

$$d\left(\Omega_i,\Omega_j\right)=d_H\left(\Omega_i,\Omega_j\right) \tag{16}$$

for ancestor and descendant trees, respectively.

In method 2, on the other hand, emphasis is shifted to the relation in granularity. Distance measure defined by the length of the longer link from Ω_i and Ω_j to their ancestor or descendant:

$$d\left(\Omega_i,\Omega_j\right)=d_H\left(\Omega_i,\Omega_j\right)+d_C\left(\Omega_i,\Omega_j\right)+d_G\left(\Omega_i,\Omega_j\right). \tag{17}$$

and

$$d\left(\Omega_i,\Omega_j\right)=d_H\left(\Omega_i,\Omega_j\right)+d_G\left(\Omega_i,\Omega_j\right),\tag{18}$$

for ancestor and descendant trees, respectively.

As mentioned earlier, ancestor and descendant trees in nature grow from any pair of tessellations. Tessellations of high granularity may be connected at an early stage by long links directly with those of low granularity. It is not desirable because such links conceal multilevel hierarchical relation.

To avoid this, method 3 considers the granularity of ancestor and descendant of tessellations. Distance measure is given by

$$d\left(\Omega_i,\Omega_j\right)=G\left(A\left(\Omega_i,\Omega_j\right)\right).\tag{19}$$

and

$$d\left(\Omega_i,\Omega_j\right)=G\left(D\left(\Omega_i,\Omega_j\right)\right),\tag{20}$$

for ancestor and descendant trees, respectively. In an ancestor tree, pairs of tessellations whose ancestors are low in granularity are chosen earlier so that the tree grows upward gradually from original tessellations. Method 3 emphasizes the relation in granularity while it totally neglects the hierarchical relation among tessellations.

Method 4 extends method 3 by taking hierarchical relation into account explicitly. Distance measure is defined as

$$d\left(\Omega_i,\Omega_j\right)=G\left(A\left(\Omega_i,\Omega_j\right)\right)+d_H\left(\Omega_i,\Omega_j\right)+d_C\left(\Omega_i,\Omega_j\right).\tag{21}$$

and

$$d\left(\Omega_i,\Omega_j\right)=G\left(D\left(\Omega_i,\Omega_j\right)\right)+d_H\left(\Omega_i,\Omega_j\right),\tag{22}$$

for ancestor and descendant trees, respectively. As seen in its definition, method 4 is based on both methods 1 and 3; it considers the closeness in both hierarchical and granularity relations. Tessellations of low granularity in close hierarchical relation are connected earlier in tree construction.

Application result is discussed in the following. Properties of ancestor and descendant trees generated are summarized in Table 2 and Table 3. The focus is on the link length because it greatly affects the effectiveness of analysis. As seen in the tables, ancestor trees consist of longer links than descendant trees. This is because merging overlay is more sensitive to a difference in location of boundary lines. As seen in Figure 9, it often generates an ancestor tessellation quite different in granularity from its original ones.

Since long links conceal multilevel hierarchical relation among tessellations, discussion focuses only on descendant trees hereafter. Figure 10 and Figure 11 show the descendant trees for Doshusei systems and company branches, respectively. They reveal that the form of descendant trees heavily depends on the distance measure used in tree construction rather than the properties of tessellations. The form is quite different even within the same set of tessellations while the same measure generates similar trees from different tessellation sets.

In either set of tessellations, the trees generated by method 1 look quite different from those by others. The comb-like trees consisting of long links give us little information about multilevel hierarchical relation among tessellations. Methods 3 and 4 generate shorter links than methods 1 and 2 as intended. Emphasis on relation in granularity reduces the link length. It is also confirmed by comparing the results from methods 1 and 2, the latter of which considers the difference in granularity explicitly.

Methods 3 and 4 are different in that the former tends to connect tessellations similar in granularity, while the latter often chooses tessellations of different granularity. Method 4 seems more successful to reveal the multilevel hierarchical relation represented by chains of tessellations, and consequently, is the most effective among the four methods to understand the relation among tessellations.

Figure 12 shows the details of descendant trees generated by method 4. In the figure we naturally find clusters of tessellations such as those shown by dotted lines. Though they are not defined based on an objective theory, they clearly reveal multilevel hierarchical relations in each cluster as expected.

Let us focus on the cluster on the right end of descendant tree of Doshusei system in Figure 12. It consists of eight tessellations shown in Figure 13. They can be further classified by chains of tessellation into three subgroups $\{\Omega_{17}, \Omega_{19}\}$, $\{\Omega_{21}, \Omega_{26}, \Omega_{12}\}$, and $\{\Omega_{22}, \Omega_{1}, \Omega_{8}\}$, each of which has its own multilevel hierarchical relation. In $\{\Omega_{21}, \Omega_{26}, \Omega_{12}\}$, for instance, Ω_{12} is a lower level tessellation of Ω_{21} and Ω_{26}. Tessellations Ω_{21} and Ω_{26} consist of 5 and 6 regions respectively, and Ω_{12} can be almost obtained by further dividing the regions in Ω_{21} and Ω_{26}. The subgroup $\{\Omega_{22}, \Omega_{1}, \Omega_{8}\}$, on the other hand, shows a three-level hierarchical relation. Tessellation Ω_{1} is a lower level tessellation of Ω_{22}, and Ω_8 is a lower level tessellation of Ω_1. What distinguishes the three subgroups is the spatial configuration of tessellations in the central area of Japan. Tessellations share a similar configuration within each subgroup while they are different among different subgroups.

4. Conclusion

This paper has proposed a new method for analyzing spatial tessellations. It considers the relation among tessellations in terms of granularity, hierarchy, and compatibility. Three distance measures are introduced to evaluate the relation between tessellations. The relation is represented by a Hasse diagram and visualized by ancestor and descendant trees. The method was applied to the analysis of Doshusei systems and branch areas of private companies. It revealed the properties of the method and distance measures proposed in this paper.

We finally discuss some limitations and extensions of the paper for future research. First, this paper focuses on three aspects of relation among tessellations. Though they are essential elements in the concept of relation, other aspects should also be discussed further. The similarity between tessellations, for instance, can be evaluated by using basic statistics of regions such as the mean and variance of area, perimeter, and shape indices. Second, this paper evaluates the relation between tessellations equivalently for every pair of points over the whole region. However, as seen in the discussion on spatial autocorrelation, relation between close points is often considered more important than that between distant points. One possible extension in this direction is to introduce a distance decay function in evaluating the relation between tessellations. Third, the method should be applied to a wider variety of tessellations. Specifically, it would be useful to apply the method to tessellations representing natural phenomena, because their properties are quite different from those of artificial tessellation such as administrative units. Unlike artificial tessellations, tessellations defined from natural phenomena do not usually share common boundaries, and they are often unstable and indeterminate. Such application will allow us a deeper understanding of the benefits and limitations of the method.

Acknowledgement

This paper is an extension of the master's thesis of T. Sasaya completed under the author's supervision (Sasaya 2008). Discussions with him were always stimulating and fruitful. The author would also like to thank A. Okabe, Y. Asami, and T. Ishikawa for their valuable comments.

References

Anderson, I. (2002) *Combinatorics of Finite Sets.* Dover.

Congalton, R. G. and Mead, R. A. (1983). A Quantitative Method to Test for Consistency and Correctness in Photointerpretation, *Photogrammetric Engineering and Remote Sensing* 49, 69-74.

Davey, B. A. and Priestley, H. A. (2002) *Introduction to Lattice and Order*. Cambridge University Press.

Fritz, S. and See, L. (2005). Comparison of Land Cover Maps Using Fuzzy Agreement, *International Journal of Geographical Information Science*, 19(7), 787-807.

Kraak, M.-J. (2003) *Cartography: Visualization of Geospatial Data*. Pearson Education.

Monserud, R. A. and Leemans, R. (1992) Comparing global vegetatic maps with Kappa statistic, *Ecological Modelling*, **62**, 275-293.

Nielson, G. M., Hagen, H., Muller, H. (1997) *Scientific Visualization: Overviews, Methodologies, and Techniques*. IEEE.

Okabe, A., Boots, B., Sugihara, K., and Chiu, S.-N. (2000): *Spatial Tessellations - Concepts and Applications of Voronoi Diagrams* . John Wiley & Sons.

Pemmaraju, S. and Skiena, S. (2003): *Computational Discrete Mathematics: Combinatorics and Graph Theory with Mathematica* . Cambridge University Press.

Pontius Jr., R. G. (2000). Quantification Error versus Location Error in Comparison of Categorical Maps, *Photogrammetric Engineering & Remote Sensing*, 66, 1011-1016.

Pontius Jr, R. G. (2002). Statistical methods to partition effects of quantity and location during comparison of categorical maps at multiple resolutions, *Photogrammetric Engineering & Remote Sensing*, 68(10), 1041-1049.

Rosenfield, G. H. and Fitzpatrick-Lins, K. (1986). A Coefficient of Agreement as a Measure of Thematic Classification Accuracy. *Photogrammetric Engineering and Remote Sensing*, 52(2), 223-227.

Sadahiro, Y. (2002): An Exploratory Method for Analyzing a Spatial Tessellation in Relation to a Set of Other Spatial Tessellations, *Environment and Planning A*, 34(6), 1037-1058.

Sasaya, T. (2008): *Methods for Evaluating the Similarity between Tessellations*. Master's thesis, Department of Urban Engineering, University of Tokyo.

Slocum, T. A., McMaster, R. B., Kessler, F. C. and Howard, H. H. (2004) *Thematic Cartography and Geographic Visualization*. Prentice-Hall.

Ware (2004) *Information Visualization: Perception for Design*. Morgan Kaufmann.

Wright, H. (2006) *Introduction to Scientific Visualization*. Springer.

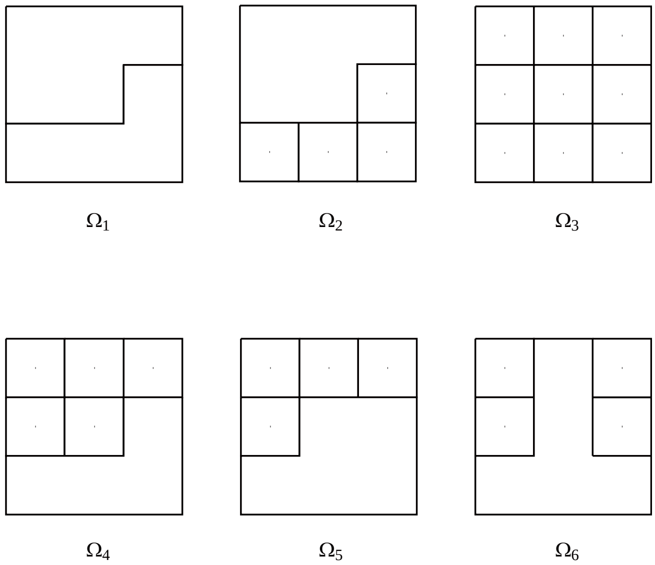

Figure 1 Hierarchical and compatible relations between tessellations. Tessellations Ω_1, Ω_2, and Ω_3 are hierarchical, while tessellations Ω_2 and Ω_4 are compatible.

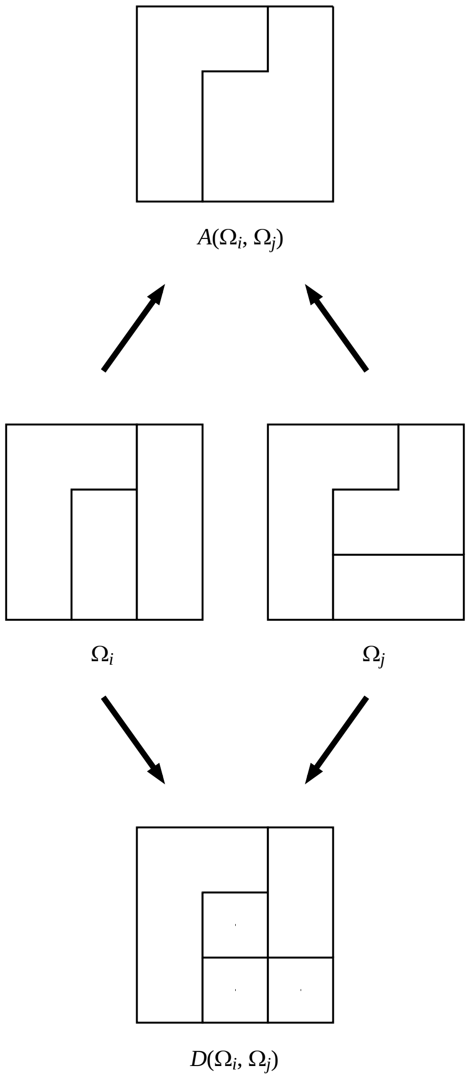

Figure 2 Merging and dividing overlay performed on tessellations Ω_i and Ω_j. They yields their ancestor and descendant tessellations denoted by $A(\Omega_i, \Omega_j)$ and $D(\Omega_i, \Omega_j)$, respectively.

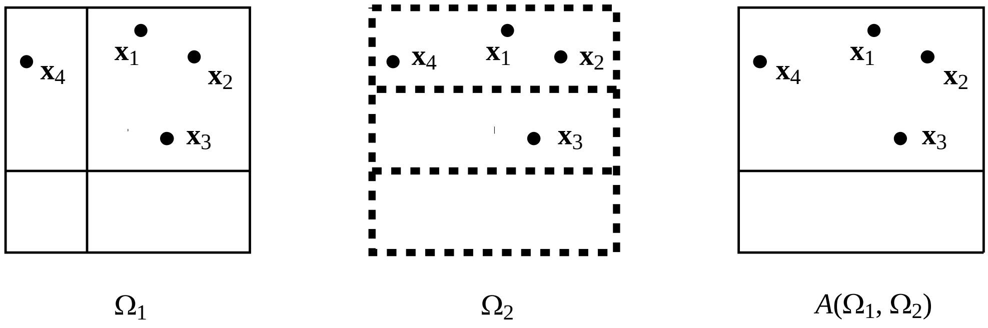

Figure 3 Two spatial tessellations Ω_1 and Ω_2, and their ancestor $A(\Omega_1, \Omega_2)$ obtained by merging overlay.

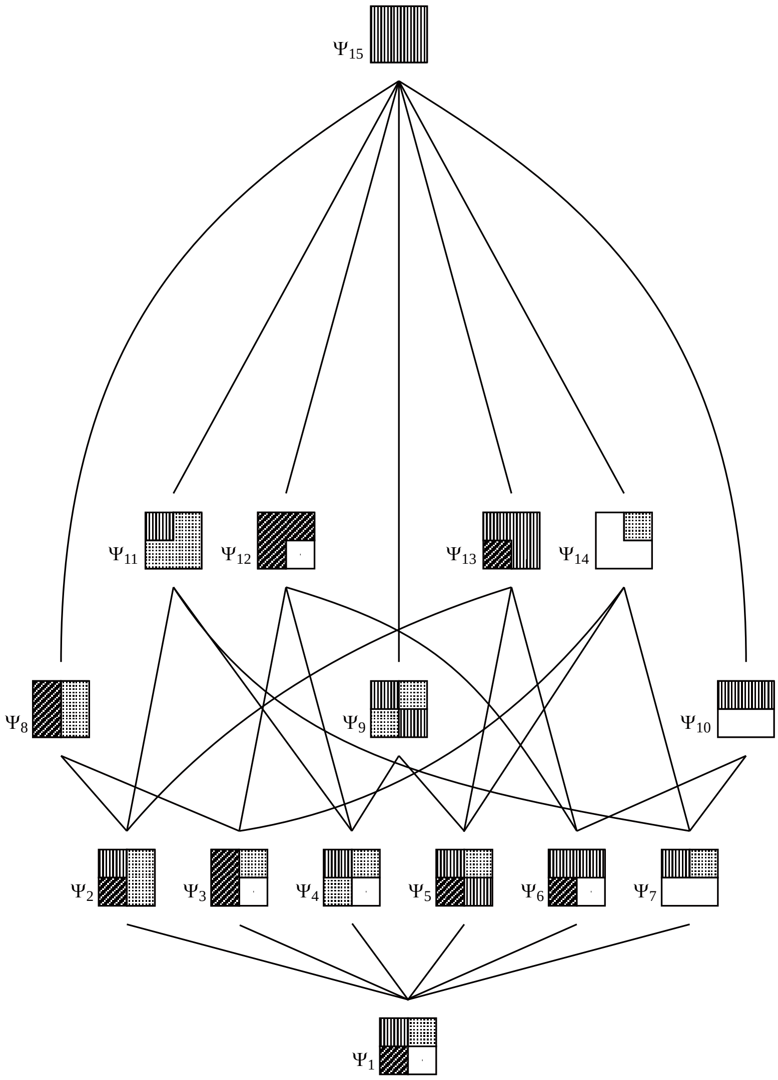

Figure 4 A Hasse diagram where the finest partition consists of four regions. Nodes and links represent tessellations and immediate relations, respectively.

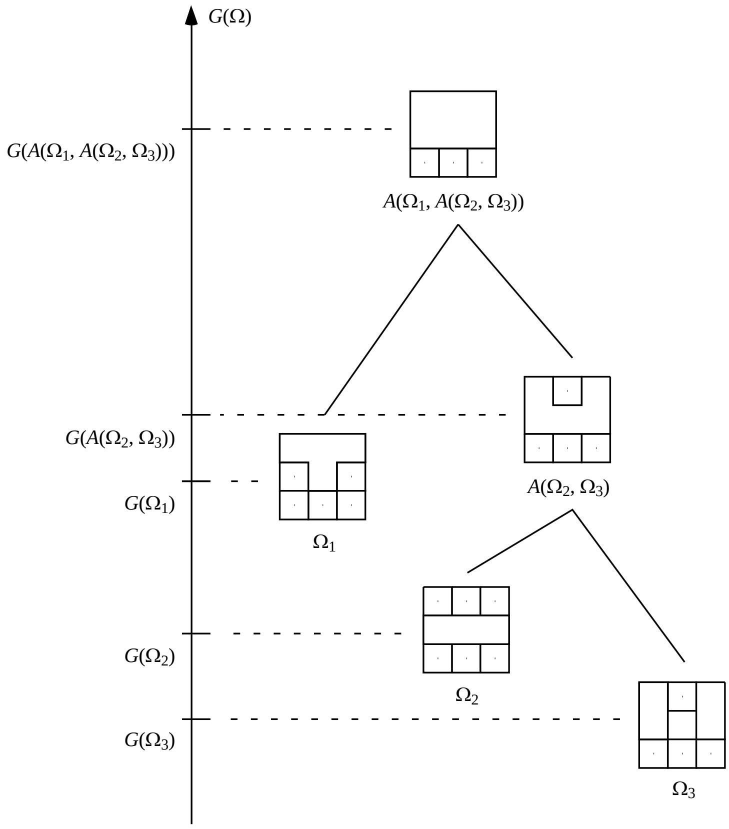

Figure 5 An ancestor tree constructed from three tessellations Ω_1, Ω_2 and Ω_3. Tessellations Ω_2 and Ω_3 are chosen first to generate $A(\Omega_2, \Omega_3)$, and then tessellations Ω_1 and $A(\Omega_2, \Omega_3)$ yield their ancestor $A(\Omega_1, A(\Omega_2, \Omega_3))$.

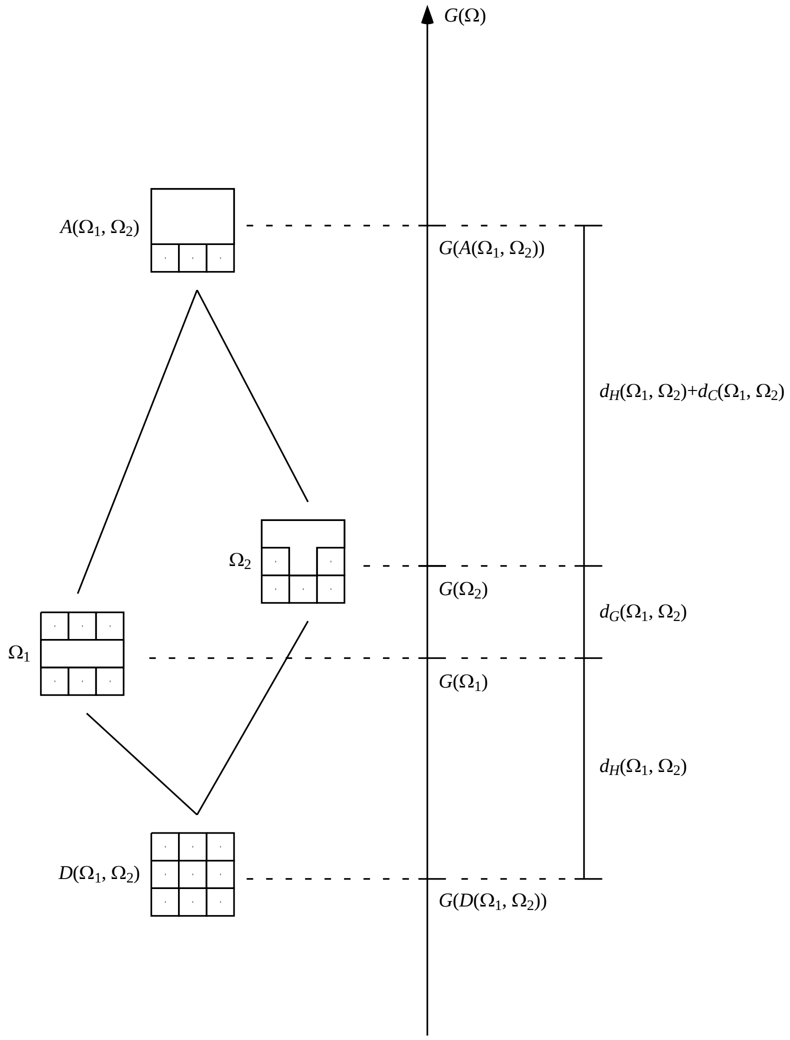

Figure 6 Relationship among distance measures and the granularity of tessellations in ancestor and descendant trees.

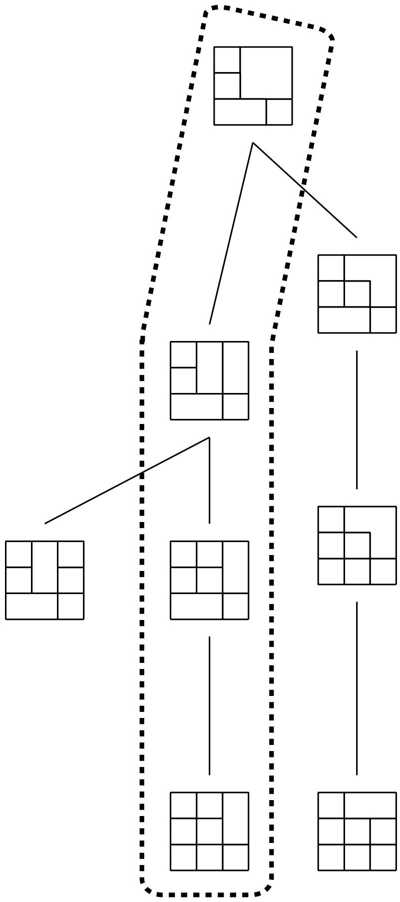

Figure 7 A cluster of tessellations that share a common ancestor. It consists of three chains of tessellations in multilevel hierarchical relation, one of which is indicated by the dotted line.

Figure 8 A candidate for Doshusei system. Thin and thick lines indicate the boundary of prefectures and new administrative units, respectively.

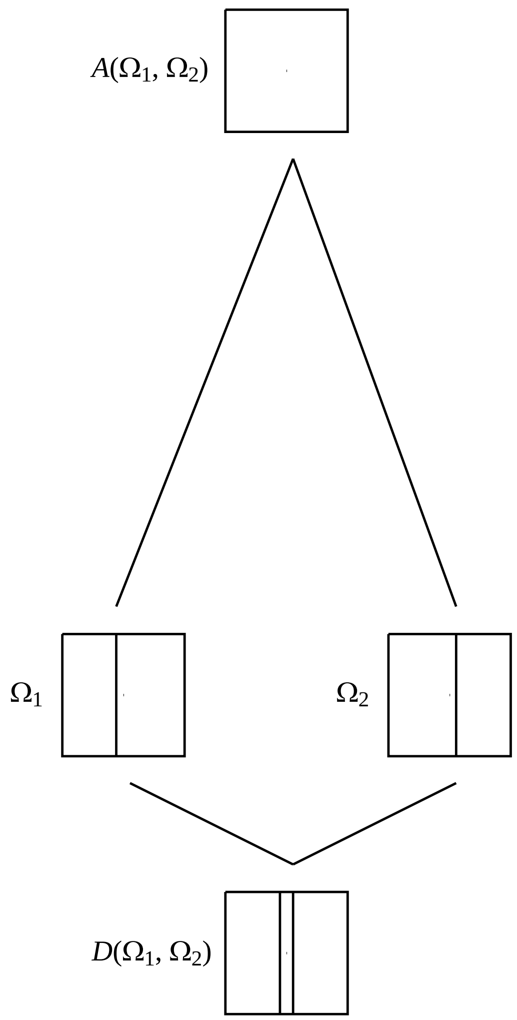

Figure 9 Merging and dividing overlay of two tessellations. A slight difference in boundary location greatly affects the result in the former while it is not critical in the latter.

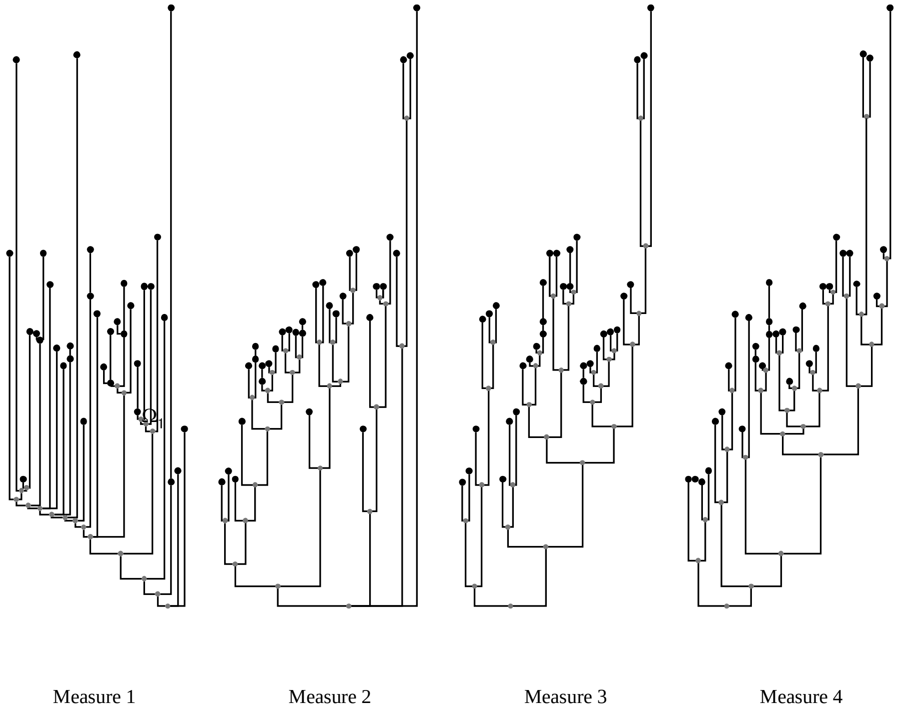

Figure 10 Descendant trees for Doshusei systems. Black and gray circles indicate original tessellations and their descendants, respectively.

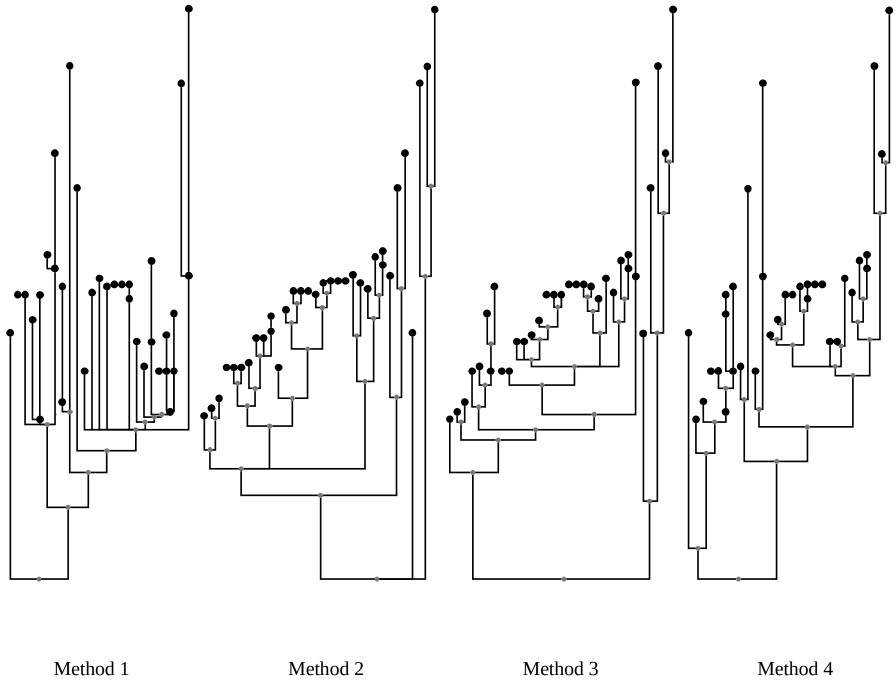

Figure 11 Descendant trees for company branches. Black and gray circles indicate original tessellations and their descendants, respectively.

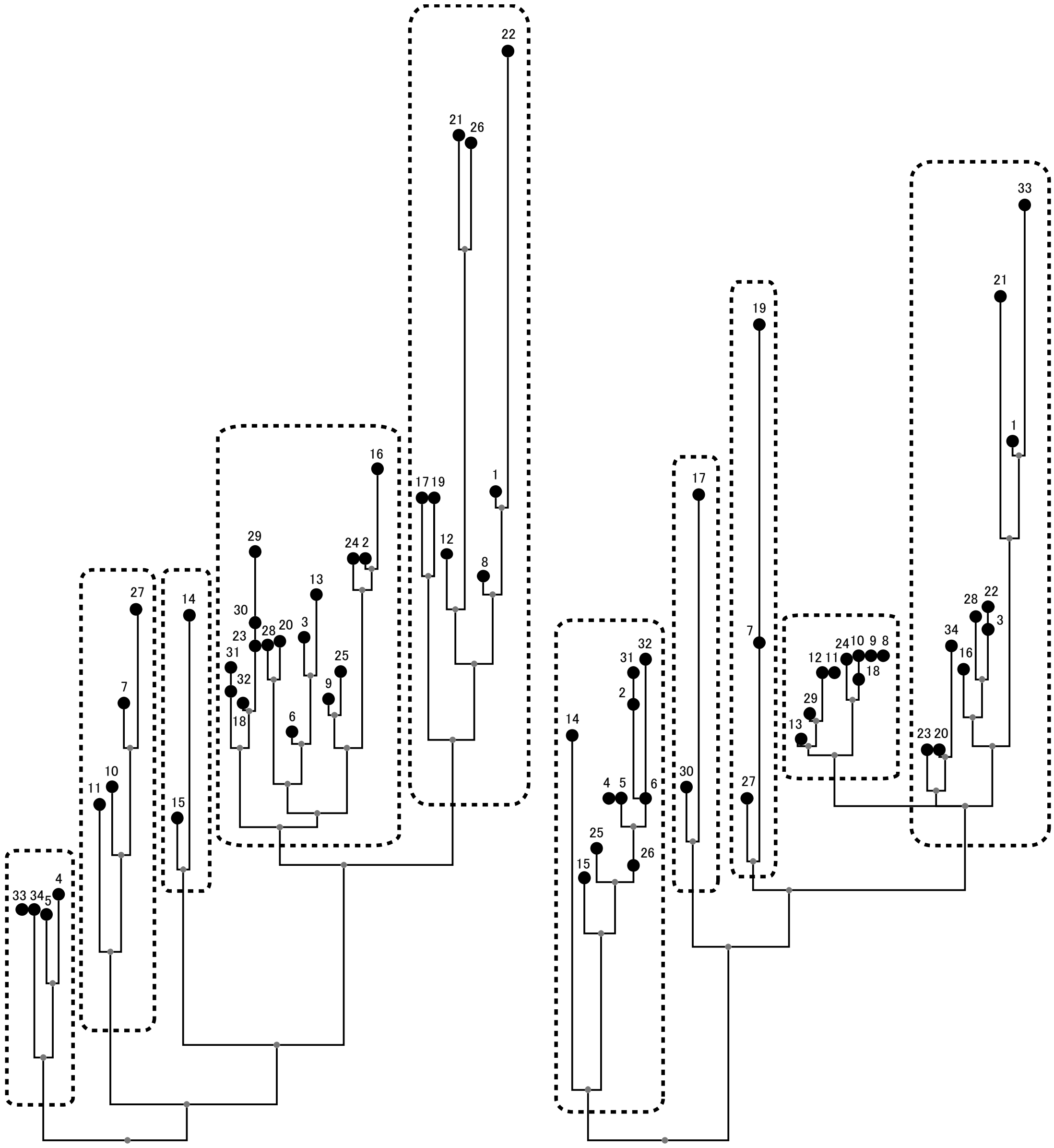

Figure 12 Descendant trees of Doshusei systems (left) and company branches (right). Black circles with identification numbers indicate original tessellations while small gray circles indicate their descendants. Tessellations surrounded by dotted lines form clusters, in each of which tessellations are in a close hierarchical relation.

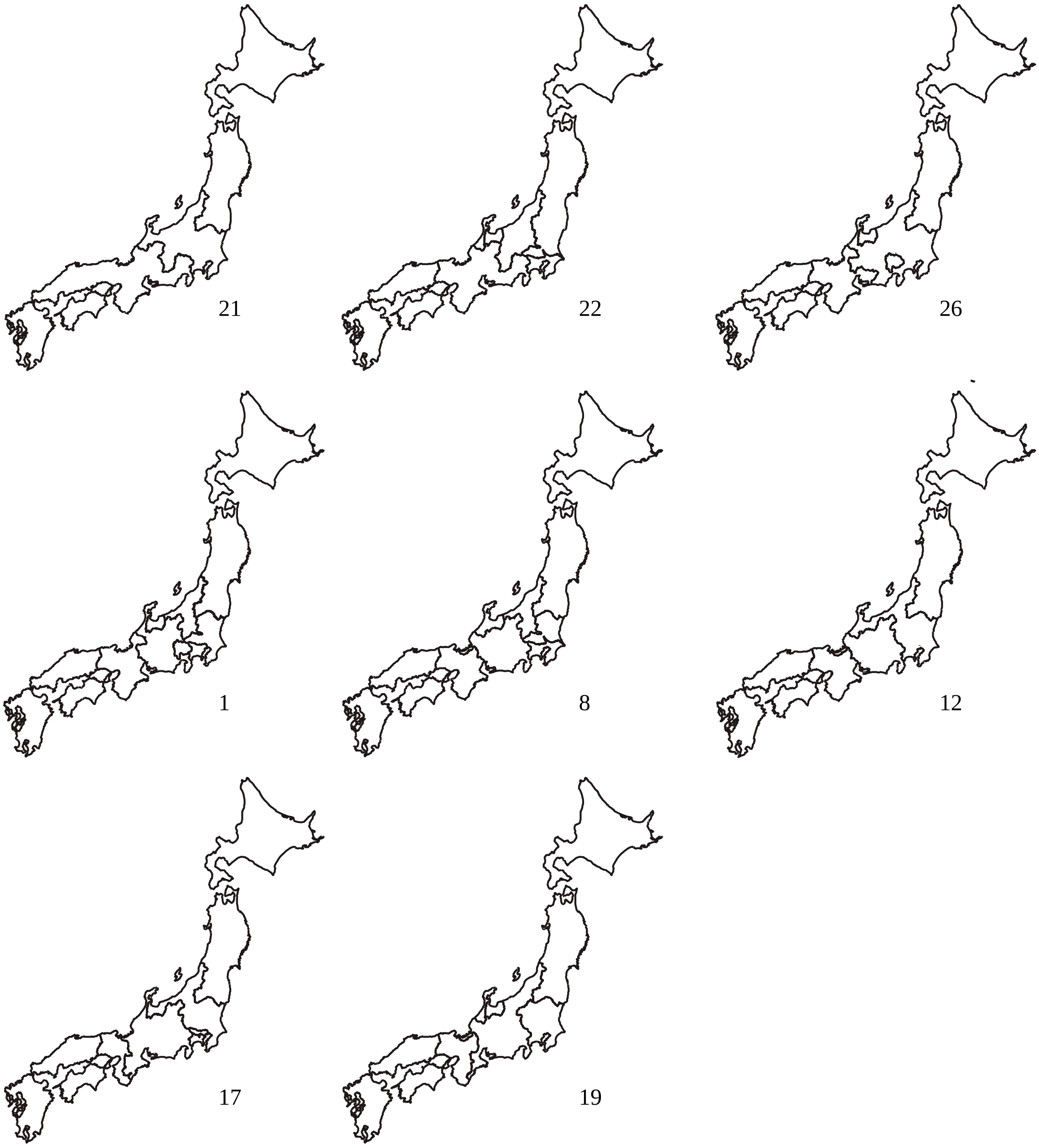

Figure 13 Candidate plans for Doshusei system.

Table 1 Relationship among the ratios of point pairs in S.

			Sum
	$r_{11}(\Omega_i, \Omega_j)$	$r_{10}(\Omega_i, \Omega_j)$	$G(\Omega_i)$
	$r_{01}(\Omega_i, \Omega_j)$	$r_{00}(\Omega_i, \Omega_j)$	$1-G(\Omega_i)$
Sum	$G(\Omega_j)$	$1-G(\Omega_j)$	1

Table 2 Total and average length of links in ancestor trees.

Tessellations	Methods	Total length	Average length
Doshusei system	1	6.4962	0.2030
	2	6.0598	0.1836
	3	5.9765	0.1828
	4	6.0210	0.1830
Company branches	1	5.5908	0.1694
	2	4.9588	0.1503
	3	4.8103	0.1458
	4	4.8231	0.1462

Table 3 Total and average length of links in descendant trees.

Tessellations	Methods	Total length	Average length
Doshusei system	1	2.0914	0.0654
	2	1.3889	0.0421
	3	1.1227	0.0351
	4	1.1942	0.0351
Company branches	1	1.5165	0.0460
	2	1.4830	0.0449
	3	1.0177	0.0308
	4	1.0638	0.0322