

CSIS Discussion Paper #22

二つの領域分割図の適合度評価と統合化の手続き
**A procedure for evaluating the differences between two
spatial tessellations and unifying these tessellations**

増山 篤・岡部 篤行・貞広 幸雄・島村 秀樹・三好 輝生

Atsushi Masuyama, Atsuyuki Okabe, Yukio Sadahiro,

Hideki Shimamura and Teruki Miyoshi

The Department of Urban Engineering, School of Engineering,
University of Tokyo 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656
e-mail: masu@ua.t.u-tokyo.ac.jp

二つの領域分割図の適合度評価と統合化の手続き

増山 篤・岡部 篤行・貞広 幸雄・島村 秀樹・三好 輝生

A procedure for evaluating the differences between two spatial tessellations and unifying these tessellations

Atsushi MASUYAMA, Atsuyuki OKABE, Yukio SADAHIRO,
Hideki SHIMAMURA and Teruki MIYOSHI

Abstract: This paper shows a procedure for evaluating the differences between two spatial tessellations which are made by different mapmakers, and unifying these tessellations into one tessellation. First, we show a pre-processing method. Second, we show a method for detecting topological differences and adjusting these differences. Third, we define the 'skeleton' of a boundary to make two corresponding boundaries into a uniform coarseness and develop a method for analyzing whether or not skeletons of these boundaries are qualitatively similar. For qualitatively similar boundaries, we develop a method for measuring quantitative differences in terms of parallel motion, rotation and expansion. Last, we propose a method for unifying two boundaries into one boundary.

Keywords: 領域分割図 (Spatial Tessellation), 境界線 (boundary), 骨格 (skeleton)

1. はじめに

空間データには、いくつかのタイプがある。中でも、一つの地域をいくつかの小領域に分割した領域分割図は、代表的な空間データタイプの一つである。例えば、行政界、土地利用図、植生図など、枚挙に暇がない。そして、我々は、しばしば、二つの領域分割図を組み合わせる分析を行う。ところが、そのとき、同じ地域に関して同一となるべき領域分割図間に不一致が生じていることがある。それは、往々にして、それらが複数の主体によって作成されているという事情に起因する。例えば、新宿区の町丁目界という領域分割図は、国土地理院でも統計局でも作成されている。しかしながら、作成した主体が違うので、それらにおける領域分割等が同一である保証はない。実際は、むしろ、違う場合が多い。したがって、複数の主体によって作成された領域分割図を同時に用いて

分析を行うと、領域分割図間の不一致の程度によっては、深刻な分析上の問題を引き起こす可能性がある。例えば、人口密度はポリゴンの面積に直接依存する量であるから、二つの領域分割図で面積に不一致が生じている限りにおいては、この密度を用いた分析は信頼できない。

この論文では、国土地理院による 1/2500 の新宿区の町丁目界（以下、地理院地図と呼ぶ）と総務庁統計局（以下、統計局地図と呼ぶ）による新宿区の町丁目界を例にとり、以上のような不一致を見出し、その不一致の程度を評価し、二つの領域分割図を一つに統合するための一連の手続きを提案する。まず、第一に、これらのデータの処理について述べる。第二に、これらの町丁目界において、位相構造の違いを見つけ、それを修正する方法を提案する。第三に、町丁目界の骨格というものを定義し、二つの町丁目界の粗さを揃える方法を開発する。次に、それらの骨格が質的に一致するかどうかを調べる手法を開発する。第四に、もし骨格が質的に一致するならば、それらの平行移動、回転、拡大・縮小などの空間的関係を分析する方法を開発する。最後に、違いが存在す

る場合に、これら町丁目界を統合・修正する方法を示す。

二つの領域分割図の境界線の適合度評価法に関する既存研究としては、Goodchild and Hunter (1997)、Skidmore and Turner (1992)、Tveite and Langaas (1999) が挙げられる。しかしながら、これらの研究では、領域分割図の適合度評価に止まっており、また、先に述べたような、二つの境界線の質的違い、空間的關係などは扱われていない。また、本来同一であるべき、二つの空間データが一致しない状況下において、それらの違いを評価し、修正・統合する一連の手続きを示した研究としては Walter and Fritsch (1999)がある。具体的には、ドイツにおける二つの道路ネットワークデータを題材として、情報理論に基づいた適合度評価法などが取り上げられているが、領域分割図を取り扱ったものではない。

2. 非空間データの整合性のチェック

二つの領域分割図の適合度評価、統合を行うためには、それらを比較可能なものとするために、その前段階としてデータに対して処理を行う必要がある。具体的には、ファイル形式の統合、不要な線の除去、対応関係にあるポリゴンに対して共通の名称を振ることである。この論文では、こうした処理のことを前処理と呼ぶことにするが、ここでは、いくつかの前処理手続き手法を提案していこう。

Step 1: 地図ファイル形式の統一化

新宿区の町丁目界の場合、統計局地図のファイル形式は、シェイプファイルであるが、地理院地



図1 図郭線の除去

図の場合、独自のファイル形式のデータである。そこで、地理院地図のファイル形式を同一のファイル形式とする。

Step 2: 地理院地図の図郭線除去手続き

地理院地図の場合、元々のデータにおいては、図 1(a)にあるように図郭線が存在している。そこで、図 1(b)にあるように図郭線を除去する。

Step 3: 名称の対応チェックと修正

例えば、新宿区には、下落合 1 丁目、西早稲田 2 丁目、高田馬場 3 丁目など様々な町丁目が存在する。ところが、実際には、こうした町丁目のリスト間において、対応関係にエラーがあることが多い。そもそも町丁目数が一致しないということがありうるし、また町丁目数が一致したとしても、町丁目リストの内容が全く同一であるとは限らない。例えば、この研究で用いたデータの場合、町丁目数が一致しないということはなかった。しかしながら、例えば、下落合 1 丁目に関して言えば、統計局地図の場合、「下落合 1 丁目」、地理院地図の場合「下落合一丁目」と入力されている。この場合、「丁目」の前の漢数字を算用数字に置き換えるだけでよいので、この場合は、システムティックに内容を修正することができる。

3. 領域間の接続関係の違いの検出、統合・修正

Step 4: 位相（隣接）関係の表示

ここで言う位相関係とは、新宿区の町丁目界データに関して言えば、町丁目の隣接関係のことである。この関係を視覚的に表現するには、各町丁目を一つのノードと表し、二つの町丁目が隣接し

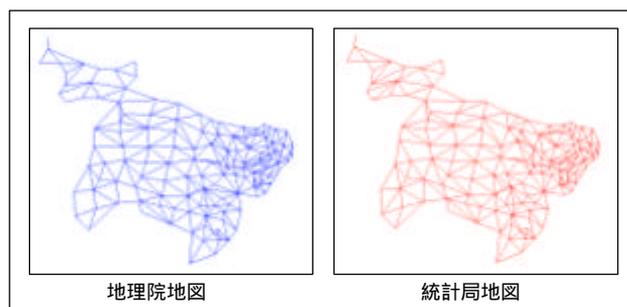


図2 町丁目の隣接関係を表すグラフ

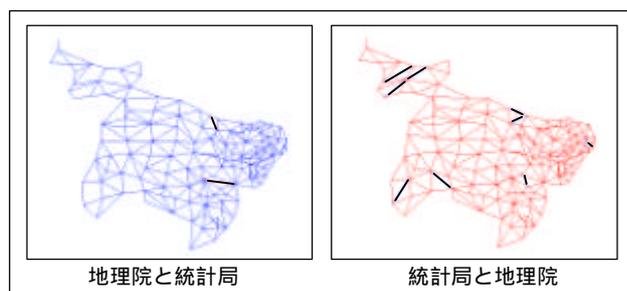


図3 位相関係の相違

ている場合には、それぞれを表すノードを結びリンクを張ってグラフを生成し、これを表示してやればよい。図2は、地理院地図、統計局地図について、このようにして生成したグラフを視覚的に表現したものである。

Step 5: 位相（隣接）関係の違いを表示

以上のようにして生成された2つのグラフが同相であれば、二つの領域分割図間において、町丁目界の接続関係に違いはない。しかし、往々にしてこれらのグラフは同相ではない。しかも、全体的にグラフの位相関係が一致しないということではなく、部分的に位相関係が一致しないということが多い。実際、この新宿区町丁目界データに関して、二つのデータ間で位相関係が異なる箇所を太線で表示すると図3のようになる。

Step 6: 位相（隣接）関係の統合化

さて、このような位相関係の不一致が生じたとき、それをいかに修正すればよいか、ということを考えてみよう。

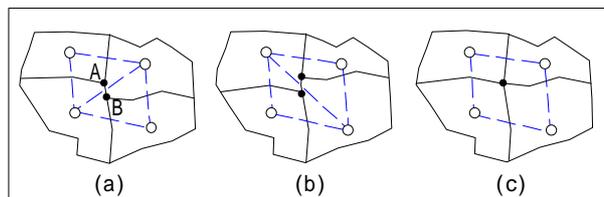


図4 機械的に修正できる位相関係

まず、このような位相関係の不一致がもっとも生じやすいケースとしては、図4のように町丁目界が「ほぼ十字」に交わっていると、「ほぼ

十字」の交点付近における境界線の交わり方が微妙に変化するだけで、図4(a)~(c)のように、位相関係が大きく変化する場合があります。このような場合、ある閾値を設定し、ポリゴンの境界線同士が交わる二点間の距離が、その閾値以下であった場合、それらを一つにまとめるような操作を行えばよい。例えば、図4(a)における二点A,B間の距離がある程度以下ならば、図4(c)のような位相関係に機械的に修正することができる。しかし、A,B間の距離がその閾値を超えた場合は、機械的な修正は不可能である。

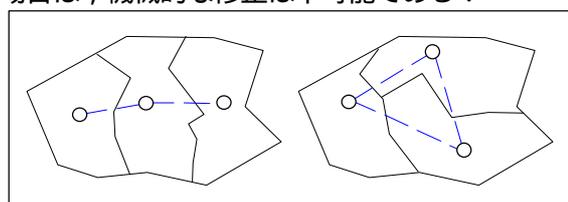


図5 致命的な位相関係の違い

さらに、図5のような場合もあり得る。この場合、機械的に位相関係を修正することはほとんど不可能である。おそらく、データの入力段階において致命的なミスが生じたものと考えられる。こうした場合は、人間の手によってデータを入力し直す必要がある。

この研究でも、上のような方法で、地理院地図と統計局地図の間で位相的に違う部分については、修正を行い、統合化する。図6は、その結果である。

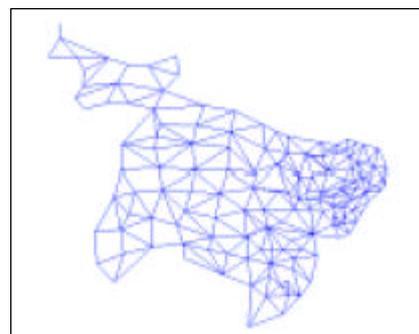


図6 修正された位相関係

なお、領域分割図における位相的問題に関しては、岡部・貞広(1999)において、より詳しく議論されている。

4. 二つの境界線の質的同一性の分析

町丁目の位相関係が一致すれば、町丁目の境界線同士に対応付けを行うことができる。しかしながら、これらはあくまで位相的に一致しているに過ぎず、これらの間には何らかのずれが生じていることが多い。例えば、境界のデジタル化精度が異なることは十分あり得る。また、入力精度が同じであったとしても、それらを重ね合わせるとき、完全に一致する保証はどこにもない。

この論文では、まず境界線の骨格を定義し、次に対応関係にある境界線の粗さを揃えるとは何かということについて述べる。次に、粗さが揃えられた二つの境界線の骨格が質的に同一かどうか調べる方法について述べる。最後に、もし、それらの骨格が質的に一致するならば、それらの空間的關係を調べる方法について述べることにする。

なお、以下では、新宿区における全ての境界線の中から、図 7 にある特定の一組の境界線（以下、例題境界線と呼ぶことにする）を一例として取り上げ、これらの方法について説明していこう。

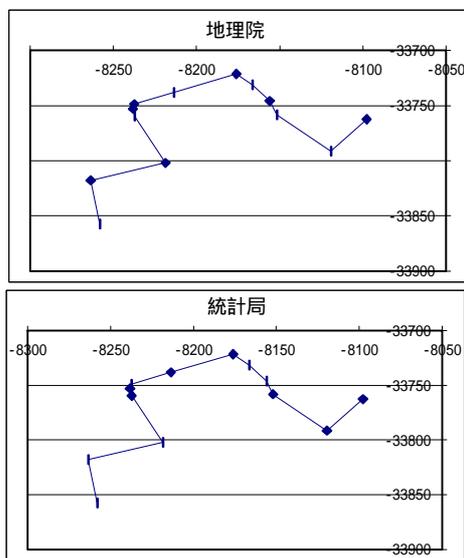


図 7 例題境界線

Step 7: 境界線の粗さを揃える

まず、境界線を長さの等しい小線分から成る線分で近似する。例えば、図 8 のような近似を行うということであるこのとき、近似線分において小線分の長さを十分小さくすることで、十分良

い精度での近似が可能である。

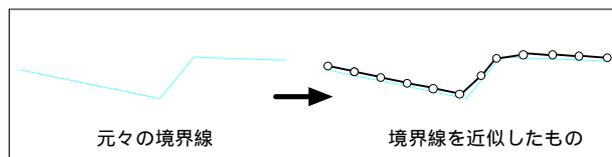


図 8 境界線の近似法

図 9 は、このような方法で、例題境界線について近似線分を求めたものである。

ここで、例えば東西軸を実数軸、南北軸を虚数軸とみることで、先の小線分からなる近似線分は複素数列とみなすことができる。具体的には、この近似線分が複素平面上の $(n+1)$ 個の点 $z(0), z(1), \dots, z(n)$ から成るものとする。なお、2点 $z(j), z(j+1)$ (ただし、 $j = 0, 1, \dots, n-1$) の距離は一定間隔 d とする。また、

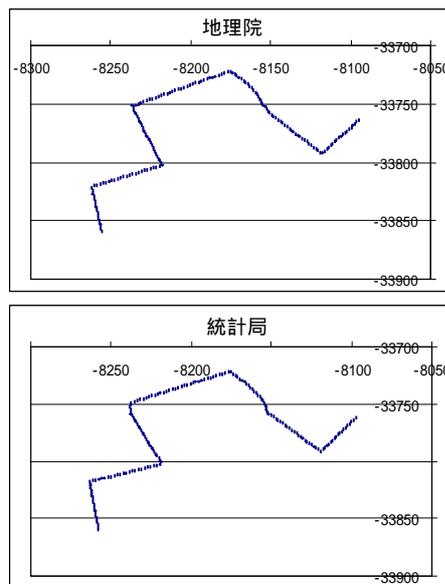


図 9 近似された例題境界線

$$w(j) = \frac{z(j+1) - z(j)}{d} \quad (1)$$

とおく。

さて、このようにして近似線分を複素数列と見なせば、それに対して複素数列の離散フーリエ変換（上坂, 1983）を行うことができる。複素数列の離散フーリエ変換とは、先の近似線分を三角関数の和として表現することであり、数式を用い

て表現すると次のようになる。

まず、

$$j_k(j) = \exp\left(-2\pi i \frac{jk}{n}\right), \quad k = 0, 1, \dots, n-1 \quad (2)$$

とおく(ここで、 i は虚数単位である)。このとき、

$$c(k) = \frac{1}{n} \sum_{j=0}^{n-1} w(j) j_k(j) \quad (3)$$

とすると、この式で求められる c が w の離散フーリエ変換と呼ばれる。

次に、

$$\tilde{c}(k) = \begin{cases} c(k) & k = 0, 1, \dots, \frac{n}{2} \\ c(k-n) & k = \frac{n}{2} + 1, \dots, n-1 \end{cases} \quad (4)$$

と変数変換すると、 $\tilde{c}(k)$ において、 k が0に近いほど波長が長い成分となっている。以下、そのような $\tilde{c}(k)$ のことを低周波成分と呼ぶことにする。逆に、 k の絶対値が大きいほど、波長が短い成分となっており、そのような成分のことを高周波成分と呼ぶことにする。

さて、ここで、最低周波数成分から数えて N 個の $\tilde{c}(k)$ を用いて線分を構築することを考える。具体的には、まず、

$$w_N(j) = \sum_{|k| \leq N} \tilde{c}(k) j^*(j) \quad (5)$$

なる $w_N(j)$ を求める(ここで、 $j^*(j)$ は $j(j)$ の複素共役)。次に、これを用いて

$$z_N(j) = z(0) + d \sum_{r=0}^{j-1} w_N(r) \quad (6)$$

なる式によって、点列 $z_N(0), z_N(1), \dots, z_N(n)$ からなる線分を構築する。先に述べた $\tilde{c}(k)$ の性質より、 N を大きくするほど、フーリエ変換以前の近似線分の微細な特徴まで復元される。逆に、 N を小さくするほど、近似線分のマクロ的な特徴が抽出される。

この論文では、上のようにして構築された線分のことをレベル N の骨格と定義する。また、対応関係にある二つの境界線に関し、共通のレベル N で骨格を抽出することを“粗さを揃える”ということにする。

なお、図 10 は、 $N = 72$ として例題境界線の粗さを揃えた結果である。

さて、ここで、二つの境界線があったとき、 N をパラメトリックに変化させながら、粗さを揃えたとしよう。このとき、二つの骨格が質的に一致するように見えることもあれば、そうでない場合もある。

では、次に、あるレベル N において骨格が質的に一致するとはどういうことかということを考えてみよう。

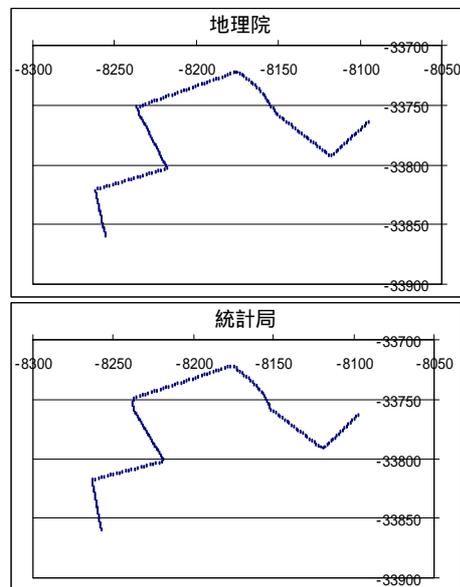


図 10 例題境界線の骨格 ($N=72$)

Step 8: 境界線の特徴点を抽出

今、仮に、図 11 のように、境界線の骨格があったとしよう。このとき、この骨格上における白

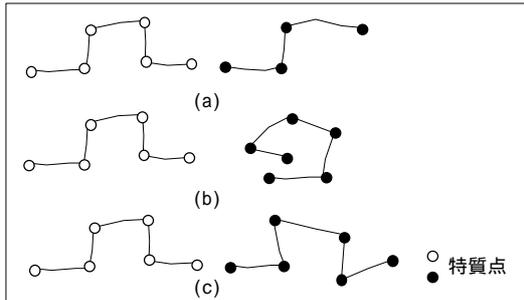


図 11 質的に一致する骨格と一致しない骨格

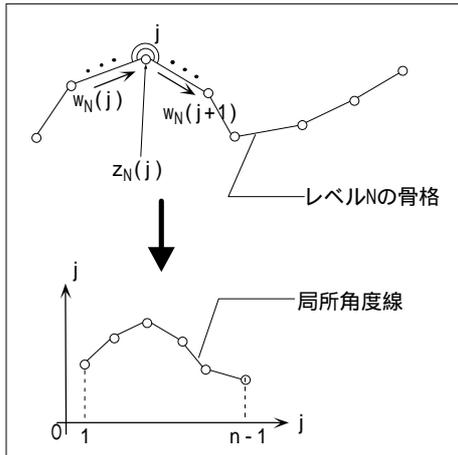


図 12 局所角度線の求め方

丸および黒丸のような点は特徴的な点だと言える。では、これらの点がどのような性質を持った点であるか、ということを考えてみると、こうした点においては、この骨格が 180 度から大きくかけ離れた角度で折れ曲がっていることがわかる。以下、この研究では、このような点のことを特質点と呼ぶことにする。

今、二つの骨格それぞれについて、特質点が求められたとしよう。このとき、図 11(a)のように、二つの骨格間で特質点の数が異なるとき、これらが質的に一致しているとは言い難い。また、図 11(b)のような場合、特質点の数は一致しているものの質的に一致しているとは考えにくい。なぜならば、特質点における左右への折れ曲がり方の順序が違うためである。図 11(c)のような場合であれば、これらは質的に一致しているとごく自然に考えられる。そこで、この論文では、左右への折れ曲がり方も考慮して、特質点のシーケンスが一致するとき、“二つの骨格は質的に一致する”とすることにする。

では、骨格が質的に一致するかどうか調べる

にはどうすればよいか考えてみよう。まず、レベル N の骨格に関し、骨格の始点から終点に沿って隣り合う小線分同士がなす角度の列を求める。このためには、複素平面上で、 $w_N(j+1)$ と $w_N(j)$ がなす角度を q_j としたとき、角度の列 $\{q_1, q_2, \dots, q_{n-1}\}$ を求める。この論文では、この角度の列を局所角度線と言うことにする(図 12)。次に、局所角度線において、180 度から大きく外れ、顕著なピークまたはボトムとなっている点を求める。これら顕著なピークまたはボトムが特質点に対応しているから、顕著なピーク・ボトムのシーケンスを求め、それらが一致しているならば、二つの骨格は質的に一致していることになる。

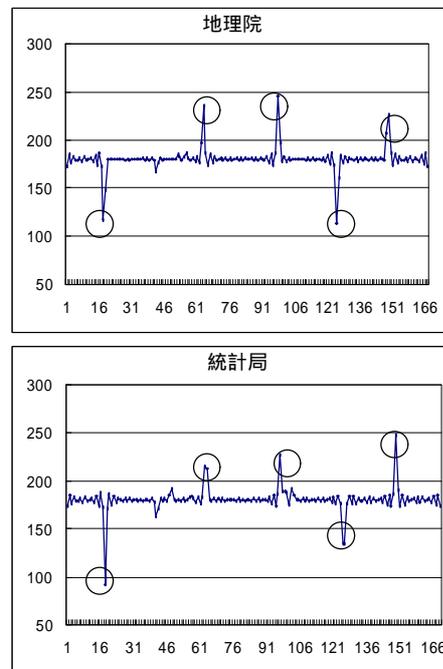


図 13 例題境界線の局所角度線

図 13 は、 $N = 72$ として、例題境界線に関して求めた局所角度線である。ここで、記号を用いて、顕著なピークを P 、顕著なボトムを B と表すことにすると、地理院地図、統計局地図それぞれの局所角度線がともに (B, P, P, B, P) というシーケンスを持っていることがわかる。したがって、例題境界線は、 $N = 72$ で、骨格が一致することが分かる。

ただ、こうした顕著なピーク、ボトムのシーケンスが一致しなくても、骨格が一致する場合が

あり得る．例えば，片方の局所角度線が (B, P, P, B, P) というシーケンスを持ち，もう片方の局所角度線が，その逆 (P, B, P, P, B) というシーケンスを持っているときに，それらの骨格が一致するということがあり得る．しかしながら，そうした場合が生じるのは，二つの境界線間において，始点と終点の対応関係が入れ替わっている場合など，非常に極端な場合にのみ生じる．したがって，二つのデータ間で標準的なデータ構造が採用されているのならば，逆のシーケンスまで考慮する必要はない．

なお，図 14 は，例題境界線の局所角度線における顕著なピーク，ボトムに対応する特質点を抽出した結果である．

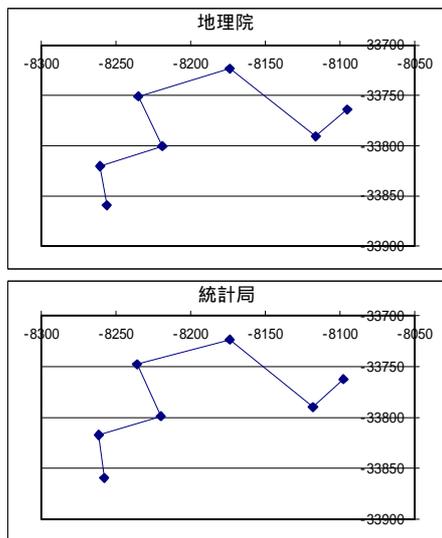


図 14 例題境界線の特質点

Step 9: 特質点の数で質的な相違を評価

例題境界線の場合， $N = 72$ で骨格が一致したが，どのレベル N においても骨格が一致しない場合もあり得る．そのような場合，特質点数の差を用いることで，二つの境界線の質的な差の評価を行うことが考えられよう．

4. 二つの境界線の空間的関係の分析

Step 10: 骨格が質的に一致する場合における計量的相違の評価

ある粗さで二つの境界線の骨格が質的に一致したとき，片方の特質点の組に関して平行移動操

作を行うことで，もう片方の特質点の組と重ね合わせることができるかもしれない．また，回転あるいは拡大・縮小といった操作によって，重ね合わせることができるかもしれない(図 15)．

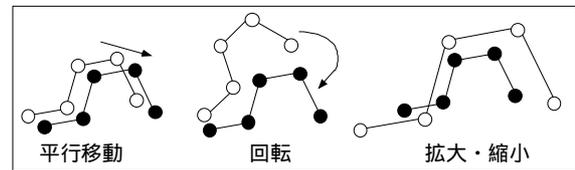


図 15 二つの特質点の組の空間的關係

二次元平面上における二つの点列の空間関係の相違を計量的に評価するための分析手法としては，複素数回帰分析がある (King, 1969)．この分析は，二つの点列が与えられたとき，平行移動，回転，拡大・縮小の三つのどれによって，どの程度まで重ね合わせられるかを分析する手法である．

具体的には，複素平面上における二つの点列 $\{W_1, W_2, \dots, W_n\}$, $\{Z_1, Z_2, \dots, Z_n\}$ が与えられているとする．このとき，以下のようなモデルによって Z_j によって， W_j の予測値 \hat{W}_j の予測値を求めるとい問題を考えてみる．

$$\hat{W}_j = A + BZ_j \tag{7}$$

(ただし， $A = a_1 + ia_2$, $B = b_1 + ib_2$)

このとき，以下の最適化問題を解けばよい．

$$\min_{a_1, a_2, b_1, b_2} \sum_{j=1}^n |\hat{W}_j - W_j|^2 \tag{8}$$

また，この最適化問題の解は以下のように一意に定まる (Dryden and Mardia, 1998)．

$$\begin{pmatrix} a_1 \\ a_2 \\ b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} n & 0 & \sum_{j=1}^n x_j & -\sum_{j=1}^n y_j \\ 0 & n & \sum_{j=1}^n y_j & \sum_{j=1}^n x_j \\ \sum_{j=1}^n x_j & \sum_{j=1}^n y_j & \sum_{j=1}^n (x_j^2 + y_j^2) & 0 \\ -\sum_{j=1}^n y_j & \sum_{j=1}^n x_j & 0 & \sum_{j=1}^n (x_j^2 + y_j^2) \end{pmatrix}^{-1} \begin{pmatrix} \sum_{j=1}^n u_j \\ \sum_{j=1}^n v_j \\ \sum_{j=1}^n (x_j u_j + y_j v_j) \\ \sum_{j=1}^n (x_j v_j - y_j u_j) \end{pmatrix} \tag{9}$$

このようにして，先のモデルのパラメータが推

定されれば, $\{W_1, W_2, \dots, W_n\}$ は $\{Z_1, Z_2, \dots, Z_n\}$ を

- ・ $a_1 + ia_2$ 方向に平行移動
- ・ $\sqrt{b_1^2 + b_2^2}$ 倍に拡大
- ・ $0 \leq q < 2\pi$, $q = \cos^{-1}\left(b_1 / \sqrt{b_1^2 + b_2^2}\right)$,
 $q = \sin^{-1}\left(b_2 / \sqrt{b_1^2 + b_2^2}\right)$ を同時に満たす q
 だけ反時計回りに回転

させたものと解釈を与えることができる。

5. 町丁目界の統合・修正

Step 11: 境界線の統合

いずれかのレベル N において粗さを揃え, 二つの境界線の骨格が一致するのであれば, 図 16 のように, 対応関係にある特質点間の中点を通っていくような新しい境界線を作成し, 町丁目界の統合・修正を行なうという方法がある。こうした統合・修正はシステマティックに行うことができよう。しかしながら, いずれのレベルで粗さを揃えても, 骨格が一致しない場合については, 人間の手によって修正を行う。

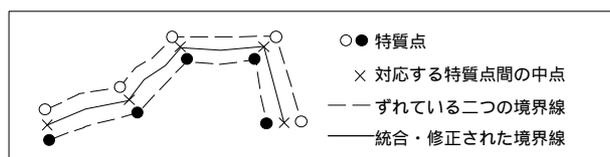


図 16 ずれている境界線の修正・統合

6. おわりに

この論文では, 国土地理院, 総務庁統計局によって作成された新宿区の町丁目データを題材として, 二つの領域分割図間に生じた領域同士の接続関係の相違, 境界線同士の質的相違, 空間的關係を分析する手法を提案した。また, 領域分割図間に生じた不一致を解消する手法も提案した。

この研究では, 手法ごとにプログラムを作成し, 分析, 統合・修正を行ってきたが, これらをまとめ, GIS 上において対話型システムとして実装

することが, 今後の課題として挙げられよう。

謝辞

この研究は東京大学空間情報科学研究センターとの共同研究の一部を筆者らの責任でまとめたものである。

参考文献

- Dryden, I.L. and Mardia, K.V. (1998) Statistical Shape Analysis, Chichester, Wiley
- Goodchild, M.F. and Hunter, G.J. (1997) A simple positional accuracy measure for linear features, International Journal of Geographical Information Science, 11, 299-306
- King, L.J. (1969) Statistical Analysis in Geography, New Jersey, Prentice-Hall
- 岡部篤行, 貞広幸雄 (1999) 二つの領域分割図の位相的適合状態を評価する方法と統合化する方法, Discussion Paper, 11, Center for Spatial Information Science, University of Tokyo
- Skidmore, A.K. and Turner, B.J. (1992) Map Accuracy Assessment Using Line Intersect Sampling, Photogrammetric Engineering and Remote Sensing, 58(10), 1453-1457
- Tveite, H. and Langaas, S. (1999) An accuracy assessment method for geographical line data sets based on buffering, International Journal of Geographical Information Science, 13(1), 27-47
- 上坂吉則 (1983) かたちのスペクトル分析 視覚からみたパターン認識, 数理科学, 246, 38-45
- Walter, V. and Fritsch, D. (1999) Matching spatial data sets: a statistical approach, International Journal of Geographical Information Science, 13(5), 445-473