

三角網分割による連続エリアカルトグラムの作成手法*

New Method of Continuous Area Cartogram Using Triangulated Irregular Network*

井上 亮**・清水英範***・Mongkol TAWECHAITOSAPOL****

By Ryo INOUE**・Eihan SHIMIZU***・Mongkol TAWECHAITOSAPOL****

1. はじめに

近年、国や地方自治体などが収集する多種多様な統計データが、無料あるいは安価に提供されるようになってきている。このような膨大なデータから、有用な情報を抽出することは困難だが、データの視覚化はその一助となり得る手法である。

従来、計量地理学では、統計データを視覚化する手法が研究されてきており、その一手法として、カルトグラムと呼ばれる、各地域の統計データに基づき地図を歪め、地域の特徴を表現する主題図が提案されている。連続エリアカルトグラムとは、地域の統計データの大小を、カルトグラム上に表示された地域の面積の大小を用いて表現する手法である。この連続エリアカルトグラムでは、地域を色分けすることにより、同時に2つの異なる情報を視覚的に表現することが可能である。

さて、この連続エリアカルトグラム作成手法は様々提案されてきている¹⁾⁻³⁾。その中で、最良手法であるKocmoud and House³⁾による手法では、データの表現精度・視認性に優れたカルトグラム作成が可能である。しかし、そのアルゴリズムは極めて複雑で、また、多くの計算時間を要するため、視覚化手法として実用的とは言えない。

連続エリアカルトグラム作成の問題点は、対象とする地域の形状が多種多様なため、全てに適用可能な手法の構築が難しいことにある。そこで本研究では、地域を三角網分割することを前提とし、三角形の面積をデータに合わせるよう変形する、簡潔なアルゴリズムを構築する。

*キーワード：カルトグラム，視覚化

**学生員，修(工)，

東京大学大学院 工学系研究科 社会基盤工学専攻
(東京都文京区7-3-1, TEL:03-5841-6129,

E-mail: inoue@planner.t.u-tokyo.ac.jp)

***正員，工博，

東京大学大学院 工学系研究科 社会基盤工学専攻

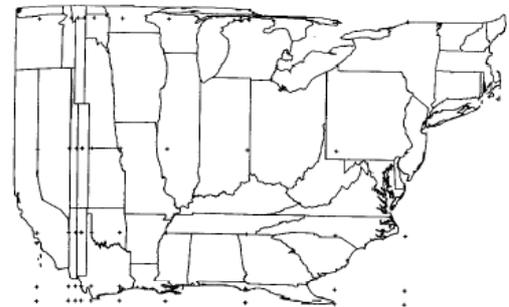
****非会員，修(工)，Department of Highways，

Ministry of Transport, Thailand

2. 既存の連続エリアカルトグラム作成手法

ここでは、1980年のアメリカ合衆国州人口を例に、既存の連続エリアカルトグラム作成手法について記す。

Tobler¹⁾では、経線・緯線の間隔を変化させて地域の面積を変化させているが、データの表現精度が



(a) Tobler¹⁾



(b) Gusein-Zade and Tikunov²⁾



(c) Kocmoud and House³⁾

図1 既存手法による1980年アメリカ合衆国州人口連続エリアカルトグラム

(Kocmoud and House³⁾より引用)

悪く、表現誤差の総和は全体の面積の 60%にも相当する(図 1-(a))。また、Gusein-Zade and Tikunov²⁾では、直線の境界を曲線に変換するため、地域の面積比較が行いにくく、視覚的に優れた作図法であるとは言えない(図 1-(b))。

既存の最良手法は、Kocmoud and House³⁾が提案した手法である。この手法による結果(図 1-(c))は高いデータ表現精度(誤差 4.4%)を示している。また、図 1-(b)に見られるような境界の直線部分を曲線へ変換することもなく、地域間の面積の比較が容易で、データの視認性が高い。このアルゴリズムは、データへの面積の適合、地域形状の保存、地域の隣接関係保持等の条件を満たすために、多くの複雑な制約条件を設定して解いている。しかし、目的関数が明確に示されておらず、数学的に明快な解法とはいえない。また、図 1-(c)を作成するために、CPU300Mhz の計算機で 6 時間を要するなど、長時間の計算を要する手法であり、視覚化ツールとしては実用性に乏しい。

3. 連続エリアカルトグラム作成手法の提案

連続エリアカルトグラム作成問題を整理し、その解法の提案を行う。

(1) 連続エリアカルトグラム作成手法の整理

連続エリアカルトグラム作成問題の目的関数は、面積の大小でデータの大小を表現するように、地域の形状・大きさを合わせることである。一方、視認性の高い連続エリアカルトグラム作成のためには、地域の形状変化を抑え、また、地域の隣接関係を保ち、通常的地図との比較を容易にする必要がある。

これらの目的関数・制約条件を踏まえた解法の構築を目指す。しかし、連続エリアカルトグラム作成では、対象とする地域の形状は多種多様なため、全ての形状に適用可能な手法の構築は難しい。そこで本研究では、地域を三角網分割した上で、三角網をデータに合わせて変形する手法を提案する。

三角網分割を用いる最大の利点は、三角形は最も単純な形状であるため、面積の計算が簡単な点である。このため、三角形の面積をデータに合わせるという目的関数を簡潔に記述することができる。また、制約条件である地域の形状・隣接関係の維持に関しても、三角網を構成する各三角形の形状変化・反転を抑えることで解決が可能である。

そこで、以後、本研究では、地域が三角網に分割されていることを前提とし、三角形の面積を与えられたデータに合わせる手法の構築を行う。

(2) 三角網を用いた作成手法の提案

三角網を用いた連続エリアカルトグラム作成の目的関数は、三角形 i の面積 A_i をデータ D_i に合わせることである。 T を三角網に含まれる三角形の集合とすると、式(1)と書ける。

$$\min \sum_{i \in T} (D_i - A_i)^2 \quad (1)$$

ここで、三角形 i の頂点 l, m, n の座標 $(x_l, y_l), (x_m, y_m), (x_n, y_n)$ を用いて表すと、式(2)と表せる。

$$A_i = \frac{1}{2} |(x_m - x_l)(y_n - y_l) - (x_n - x_l)(y_m - y_l)| \quad (2)$$

ここで、式(2)の絶対値符号を消去するため両辺を二乗する。式(1)を書き直し、

$$\begin{aligned} & \min \sum_{i \in T} (D_i^2 - A_i^2)^2 \\ & = \min \sum_{i \in T} \left[D_i^2 - \frac{1}{4} \{(x_m - x_l)(y_n - y_l) - (x_n - x_l)(y_m - y_l)\}^2 \right]^2 \end{aligned} \quad (3)$$

と目的関数を改める。式(3)より、連続エリアカルトグラム作成問題は、与えられたデータをもとに三角網の頂点座標を定める、非線形最小二乗問題であるといえる。なお、表記上の都合により、

$$x_{lm} = x_m - x_l \quad (4)$$

$$y_{lm} = y_m - y_l \quad (5)$$

とする。式(4)(5)を用いて式(3)は式(6)と書き直せる。

$$\min \sum_{i \in T} \left[D_i^2 - \frac{1}{4} (x_{lm} y_{ln} - x_{ln} y_{lm})^2 \right]^2 \quad (6)$$

ここで、式(6)を線形化して解くこととする。まず、変数 x_i, y_i に対して近似値 x'_i, y'_i を与える。

$$x_i = x'_i + \Delta x_i \quad (7)$$

$$y_i = y'_i + \Delta y_i \quad (8)$$

式(6)を線形化すると、式(9)と書ける。

$$\begin{aligned} & \min \sum_{i \in T} \left\{ D_i^2 + 3A_i^2 \right. \\ & \left. + A_i (y'_{mn} x_l - y'_{ln} x_m + y'_{lm} x_n - x'_{mn} y_l + x'_{ln} y_m - x'_{lm} y_n) \right\}^2 \end{aligned} \quad (9)$$

但し、 $x'_{lm} = x'_m - x'_l, y'_{lm} = y'_m - y'_l : x_{lm}, y_{lm}$ の近似値 $A'_i = \frac{1}{2} (x'_{lm} y'_{ln} - x'_{ln} y'_{lm})$: 三角形 i の面積の近似値(符号を含む)と表記する。

連続エリアカルトグラム作成の際には、結果を図の座標系に固定するために、1頂点の座標を与え、残りの頂点の座標を求める問題となる。しかし、三角形の面積が与えられただけでは、残りの頂点の座標を一意に定めることはできない。そこで、式(7)を解くためには、何らかの正則化条件が必要である。

前述のように、連続エリアカルトグラム作成上の制約として、地域の形状・地域の隣接関係の維持が

必要である。そこで、これらの制約を見たす解を求めるため、三角形の形状変化・反転を抑制する項を、正則化条件として式(7)に導入することを考える。

三角網上で、各三角形の方向を保ち、反転を防ぐためには、三角網を構成する各辺の方向変化を小さくすることで対応が可能である。そこで、三角網を構成する辺の方向変化を抑える項を導入する。すなわち、正則化項

$$+\mu \sum_{l_{mn} \in L} \left(\frac{y'_{mn}}{d'_{mn}} x_{mn} - \frac{x'_{mn}}{d'_{mn}} y_{mn} \right)^2 \quad (8)$$

(但し、 l_{mn} : 頂点 m, n を結ぶ辺,

L : 三角網を構成する辺の集合, μ : 変数,

$d'_{mn} = \sqrt{x'^2_{mn} + y'^2_{mn}}$: 辺 mn の長さの近似値)

を式(7)に導入することにより式(9)が得られる。

$$\begin{aligned} \min & \left[\sum_{i \in T} \left\{ D_i^2 + 3A_i^2 \right. \right. \\ & \left. \left. + A_i \left(y'_{mn} x_l - y'_{ln} x_m + y'_{lm} x_n - x'_{mn} y_l + x'_{ln} y_m - x'_{lm} y_n \right) \right\}^2 \right. \\ & \left. + \mu \sum_{l_{mn} \in L} \left\{ \frac{y'_{mn}}{d'_{mn}} (x_n - x_m) - \frac{x'_{mn}}{d'_{mn}} (y_n - y_m) \right\}^2 \right] \quad (9) \end{aligned}$$

ここで、式(9)で解を得ることができることを確認する。未知変数の数は、一点を座標系に固定するため、(頂点数-1)×2 となる。一方、観測方程式数は、(三角形数)+(辺数)となる。三角網では常に、(未知変数数) ≤ (観測方程式数) が成り立つことから、式(9)から解を一意に求めることができる。

式(9)より、頂点座標の近似値を入力すると、三角形の面積 A をデータ D に合わせるように変形した、頂点座標を得ることが可能である。

ただし、連続エリアカルトグラム作成の制約条件、すなわち地域の形状および隣接関係の維持、を満たすためには、全ての三角形の反転を防がなければならない。正則化項に対する重み μ を小さく設定すると、三角網の辺の方向変化が大きくなり、三角形の反転が生じる可能性が大きくなる。そのため、反転を防ぐためには、 μ の値を大きく設定し、三角網を構成する各辺の方向変化に対する制約を大きくする必要があり。しかし、 μ 値の増加は、収束速度の低下をもたらす、収束計算回数・計算時間の増加を招く。三角形の反転を防ぎ、かつ、収束計算時間の短い最小の μ 値を収束計算毎に設定することが望ましいが、この値を求めることは不可能である。そこで、 μ の値を減少させながら収束計算を行い、三角形の反転が観測された場合には μ を増加させて再計算を行うことにより、三角形反転の防止、および、計算時間の短縮を図る。なお、三角形反転は、頂角の

sin の符号変化により判断することができる。

本提案手法による計算手順は図2の通りである。

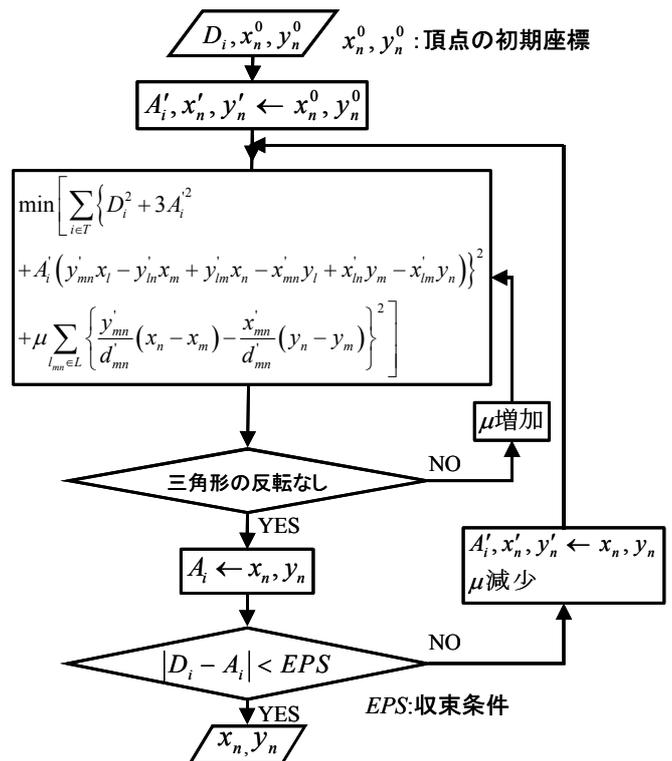


図2 本提案手法のフローチャート

4. 提案手法の適用

ここで、本提案手法を 1980 年のアメリカ合衆国州人口データに対して適用を行い、既存手法の結果(図1)との比較を行う。本提案手法適用に際しては、手作業にて作成した図3の三角網を用いている。頂点数は 409、三角形数は 698 である。

計算結果を図4に示す。図4作成には CPU1.7GHz の計算機で 12.1 分を要した。データ表現誤差の総和は、全体の面積の 3.9%にあたる。既存手法の Kocmoud and House³⁾では、誤差が 4.4%の作図結果(図1-(c))を得るために、CPU 300Mhz の計算機を用いて 6 時間を要していたことから、本提案手法は既存手法に比べて約 5 倍高速に計算できる

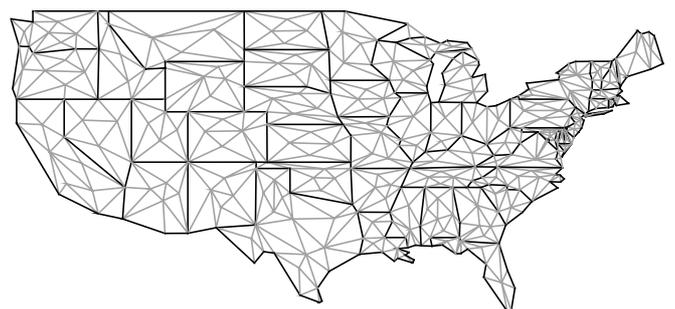


図3 使用した三角網

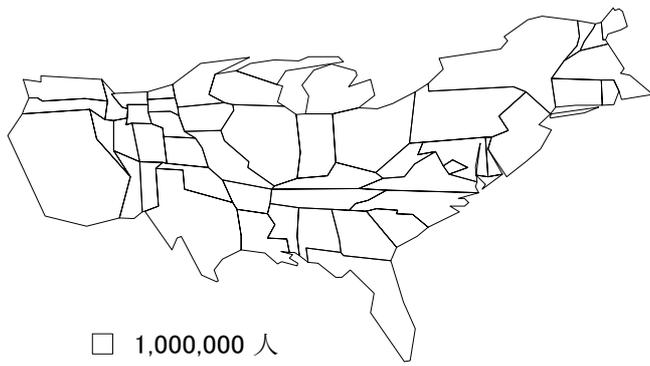


図4 提案手法による1980年アメリカ合衆国
州人口 連続エリアカルトグラム

ことが分かる。

また、日本の1925年から2000年まで5年毎の総務省統計局整備の都道府県人口データを用いて、提案手法により連続エリアカルトグラムを作成した。適用に当たっては、頂点数540、三角形数866の三角網を用いている。ここでは、1950・1975・2000年の都道府県人口連続エリアカルトグラムを人口増加率に応じて色分けした図を図5に記す。計算時間は、CPU1.7Ghzの計算機で、1950年が1.5分、1975・2000年が2.5分である。

これらの図より、戦後、三大都市圏に人口が集中していく様子を視覚的に表現することができた。また、これらの図を用いてアニメーション表示することにより、データの変化を効果的に視覚化することが可能である。

5. おわりに

本研究では三角網分割を利用した連続エリアカルトグラムの作成手法を提案した。本提案手法の特徴は、地域を三角網分割することを前提として、三角形の面積をデータに合わせる手法を構築している点である。このため、目的関数が簡潔かつ明快に記述でき、また、短時間で計算が可能である。

本手法をGISの拡張機能として実装すると、一般ユーザーが統計データを容易に視覚化することが可能となる。このツールの開発を今後の課題とする。

参考文献

- 1) Tobler, W. R. : pseudo-cartograms, The American Cartographer, Vol. 13, pp.43-50, 1986.
- 2) Gusein-Zade, S. M. and Tikunov, V. S : a new technique for constructing continuous cartograms, Cartography and Geographic Information Systems, Vol.20, pp.167-173, 1993.

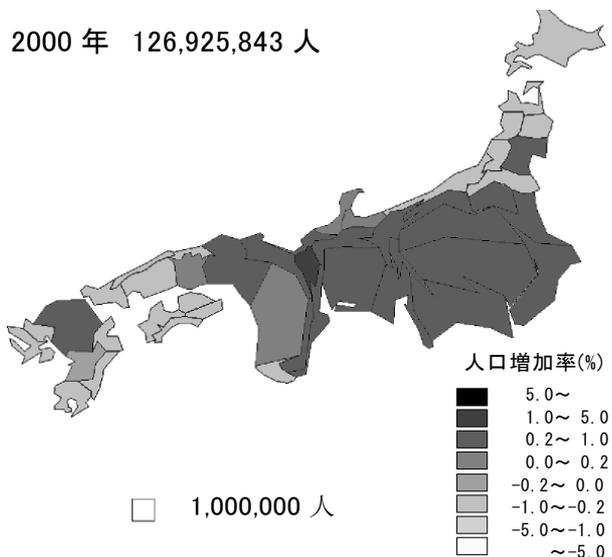
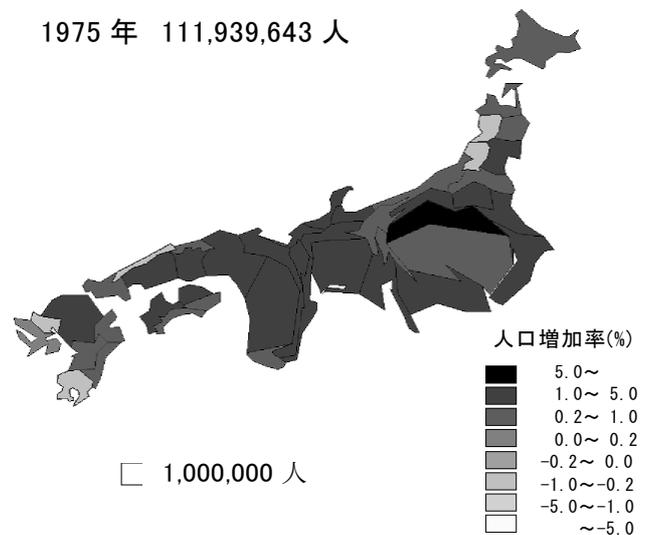
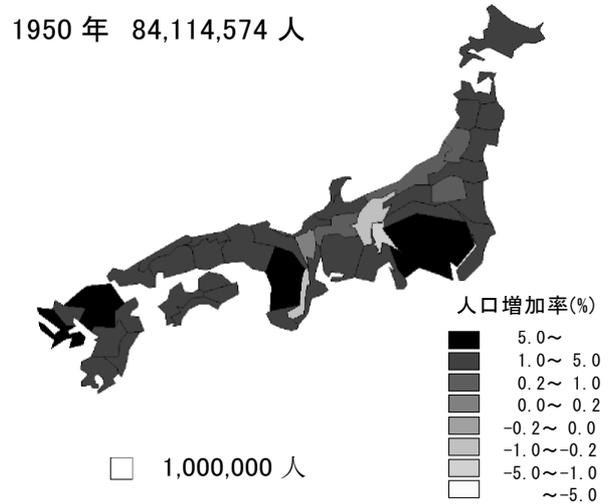


図5 提案手法による1950・1975・2000年
都道府県人口・人口増加率 連続エリアカルトグラム

- 3) Kocmoud, C. J. and House, D. H. : a constraint-based approach to constructing continuous cartograms, The Eighth International Symposiums on Spatial Data Handling Proceedings, pp.197-204, 1998.