

統計GISの発展に向けて ～ カルトグラムの可能性を探る



日本写真測量学会関西支部
第41回空間情報話題交換会

平成19年10月26日

東京大学 大学院工学系研究科
社会基盤学専攻
清水英範

発表内容に関する主な参考文献

清水英範、井上亮
時間地図作成問題の汎用解法
土木学会論文集, No.765, pp.105-114, 2004.

井上亮、清水英範
連続エリアカルトグラム作成の新手法
土木学会論文集, No.779, pp.147-156, 2005.



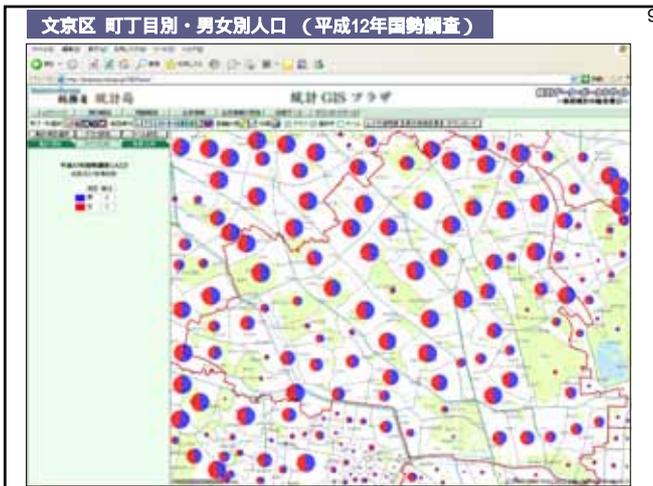
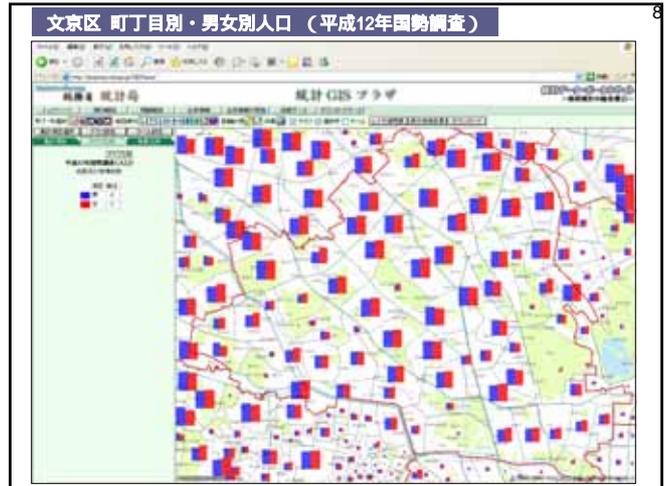
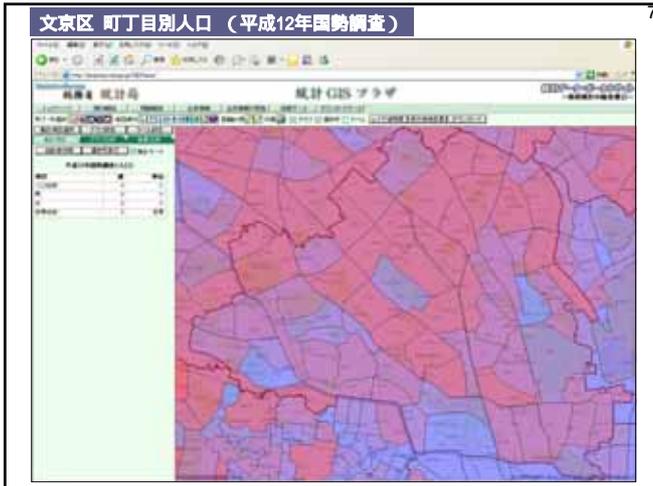
統計GIS

GISの簡単な応用として
従来から広く普及している



近年ではWebGISを利用した
情報提供も進展している





一般市民や研究者

多様な統計データを容易に入手
GISの多様な機能を通して表現

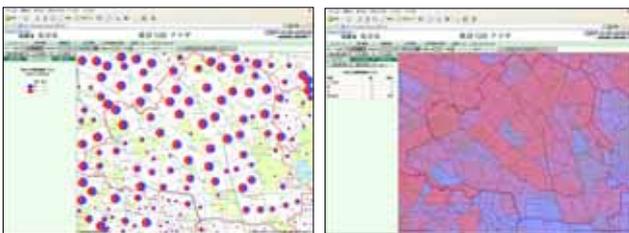


国土や地域の実態を容易に
把握・分析することが可能

GISに実装されている表現手法

グラフ表示、コロプレスマップ

紙地図以来の伝統的な表現



統計データの提供と活用に関わる近年の革新的な動向



- ・世論の高まり ・行政の前進
- ・情報通信技術の進展

GISの表現手法は
何も進化していない

統計GISの発展と活用に向けて

多くの国民が**統計データ**に関心をもち、統計データを通して地域の自然、社会経済状況を探ることに**興味を見出せる**ような環境の構築

・GISも貢献するべき

・GISの表現手法も進化するべき

2. カルトグラム(cartogram)とは

地点間(例:都市間)や地域(例:都道府県)に与えられた**統計データの大きさ**を**地点間の距離**や**地域の面積**で表現するように、地理的地図を変形させたもの



統計データが示す**地域の実態**や**地域間格差**を印象的に表現する

カルトグラムの種類

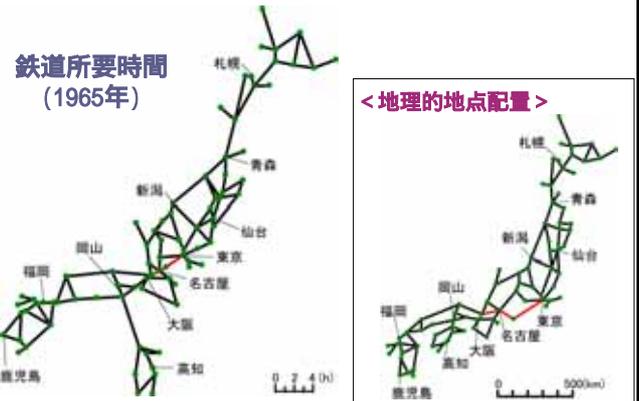
距離(distance)カルトグラム

地点間の統計データの大きさを地図上の**地点間の距離**で表現

面積(area)カルトグラム

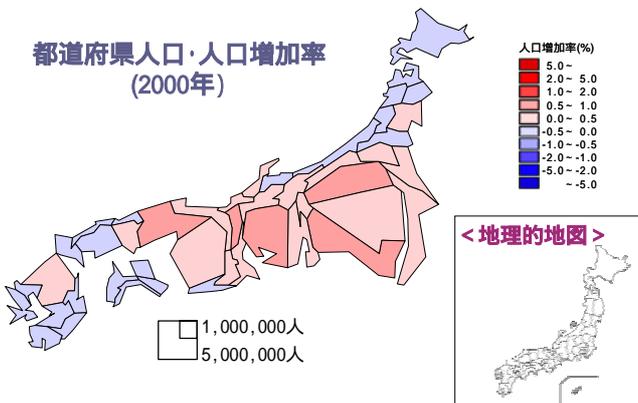
地域の統計データの大きさを地図上の**地域の面積**で表現

距離カルトグラムの例



面積カルトグラムの例

都道府県人口・人口増加率 (2000年)



これまでの研究・開発

- ・計量地理学、ORの分野で研究の蓄積
- ・実用性の高い、確立した手法はない

潜在的な魅力と可能性

- ・地図に画くのではなく、地図を変形させるという、既成概念を超えた魅力
- ・統計GISを前提とすれば、追加的なデータは不要という、高い実行可能性

手法次第では、**広く活用の可能性**

3. カルトグラム作成の課題

距離カルトグラム

鉄道所要時間 (1965年)

面積カルトグラム

都道府県人口・人口増加率 (2000年)

カルトグラムの作成は、統計データの大きさを地図上の地点間の距離や地域の面積で再現
カルトグラム作成は最適化問題

カルトグラム作成の最適化問題

距離カルトグラム

地点間に与えられた距離を再現する形状は無数に存在

面積カルトグラム

地域に与えられた面積を再現する形状は無数に存在

カルトグラム作成問題は不良設定問題
何らかの方法で正則化が必要

正則化の視点

カルトグラムは、地理的地図との比較対照を通して、統計データの特徴を解釈する

地理的地図

距離カルトグラム

統計データの大きさの再現とは関係のない無意味な地図変形を排除する必要がある

カルトグラム作成の操作性

統計GISの機能としての高い操作性

↓

ユーザーの試行錯誤的な作業を
(変数の初期値設定・正則化パラメータ設定など)
極力排除あるいは簡便化

↓

手法の全体像 概念的にも、
数学的にも可能な限り単純かつ明快に

論点

- ・不良設定問題の解消
- ・無意味な地図変形の抑制
- ・単純かつ明快な解法の構築

距離カルトグラム

鉄道所要時間 (1965年)

面積カルトグラム

都道府県人口・人口増加率 (2000年)

4. 本研究の基本的方法論

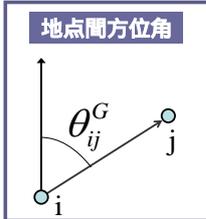
1. 作成手法の全体像を可能な限り単純かつ明快に問題を最小二乗法で記述

$$\sum (\text{統計データ} - \text{地点間距離})^2$$

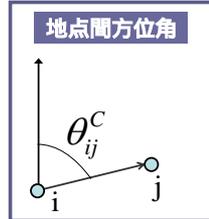
$$\sum (\text{統計データ} - \text{地域面積})^2$$

2. 不良設定問題の解消と、無意味な地図変形の抑制のために 方位角拘束による正則化

地理的地図



カルトグラム



≈

基本的方法論

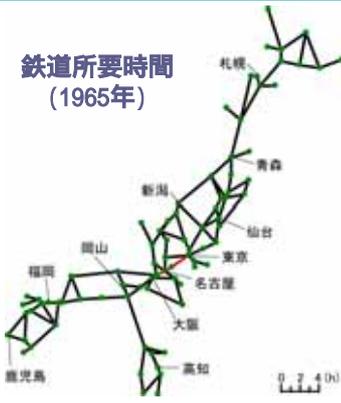
最小二乗法による問題記述
方位角条件による正則化



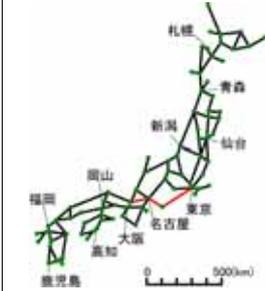
距離・面積カルトグラムの双方について、統計GISの機能として十分な実用性を有する手法を開発する。

5. 距離カルトグラムの作成手法

鉄道所要時間
(1965年)

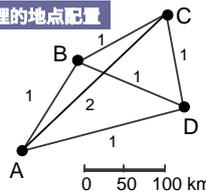


< 地理的地点配置 >



従来の作成手法(1)

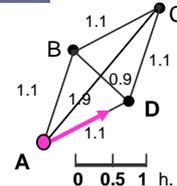
地理的地点配置



全ての地点間に
距離指標データあり
完全ネット型

MDS: 多次元尺度構成法

時間地図



非線形最小二乗問題

$$\min \sum_{i < j} \left(t_{ij} - \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \right)^2$$

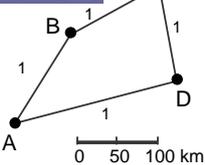
距離指標データ

カルトグラム上の距離

但し、データの表現精度が低くなる可能性あり

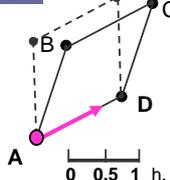
従来の作成手法(2)

地理的地点配置



一部の地点間に
距離指標データあり
部分ネット型

時間地図

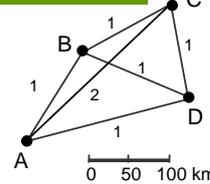


1点の座標と1リンクの方位角を
固定しても地点配置を一意に
定めることは不可能

完全な不良設定問題

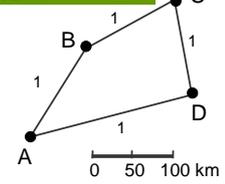
時間地図作成問題の汎用解法

完全ネット型



全地点間に距離指標データ

部分ネット型



一部地点間に距離指標データ

完全ネット型は部分ネット型の一形態

部分ネット型時間地図の解法は汎用解法

汎用解法の定式化

$$\min \left[\sum_{ij \in L} (t_{ij} - d_{ij})^2 + \mu \sum_{ij \in L} (\theta_{ij}^C - \theta_{ij}^G)^2 \right]$$

μ : 正則化パラメータ

地理的地図

地理的地図上辺 ij の方位角

カルトグラム

カルトグラム上辺 ij の方位角

地理的地図

カルトグラム

- 不良設定の解消
- 無意味な変形なし

解法

$$\min \sum_{ij \in L} \left[(t_{ij} - d_{ij})^2 + \mu (\theta_{ij}^C - \theta_{ij}^G)^2 \right]$$

↓

- 座標の近似値 (x'_i, y'_i) を用いて線形化
- 正則化項の重み $\mu = 1$ と設定

$$\min \left[\sum_{ij \in L} \left(t_{ij} \frac{x'_{ij}}{d'_{ij}} - (x_j - x_i) \right)^2 + \sum_{ij \in L} \left(t_{ij} \frac{y'_{ij}}{d'_{ij}} - (y_j - y_i) \right)^2 \right]$$

- 2つの独立な線形最小二乗問題 $x'_{ij} = x'_j - x'_i$
- 説明変数行列は不変 $y'_{ij} = y'_j - y'_i$

↓

計算時間は約 $\frac{1}{(\text{リンク数})}$ に 高速に計算可能

提案手法の特徴

- 入力データは統計データと地理的地図データのみ
- (統計GISに入力済み)
- 線形最小二乗問題の繰り返しで計算可能
- 計算時間 短(数秒程度)

$t_{ij}, (x_i^0, y_i^0) / (x_j^0, y_j^0)$: 地点 i の地理座標

$\theta_{ij}^0 \leftarrow (x_i^0, y_i^0)$

$\theta_{ij} \leftarrow \theta_{ij}^0$

$\min \sum_{ij \in L} (t_{ij} \sin \theta_{ij} - (x_j - x_i))^2$

$\min \sum_{ij \in L} (t_{ij} \cos \theta_{ij} - (y_j - y_i))^2$

$(x_i, y_i), (x_j, y_j) \Rightarrow \theta_{ij}$

$\theta_{ij} \leftarrow \theta_{ij}$

YES EPS: 収束条件

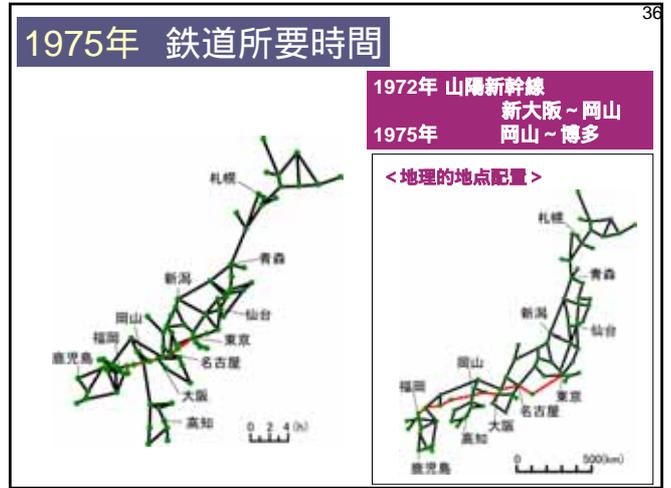
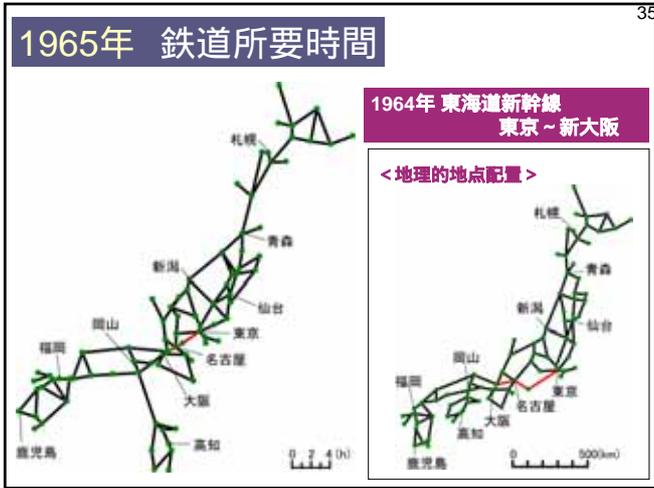
NO $|\theta_{ij} - \theta_{ij}^0| < EPS$

(x_i, y_i)

距離カルトグラムの作成例

- 鉄道時間地図とその変遷 -

1965, 1975, 1985, 1995

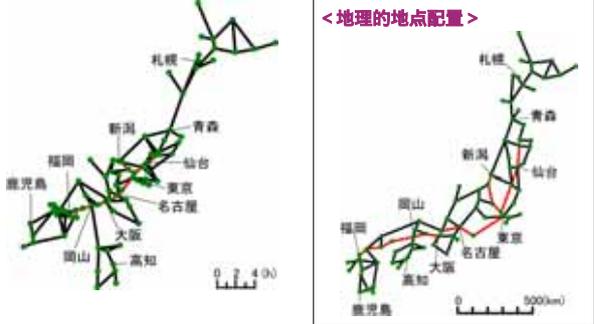


1985年 鉄道所要時間

37

1982年 東北新幹線 大宮～盛岡
上越新幹線 大宮～新潟
1985年 上野～大宮

<地理的地点配置>

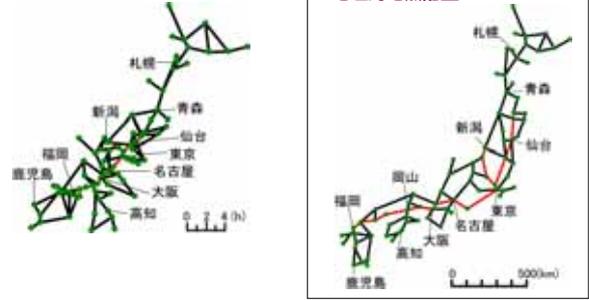


1995年 鉄道所要時間

38

1988年 青函トンネル
瀬戸大橋

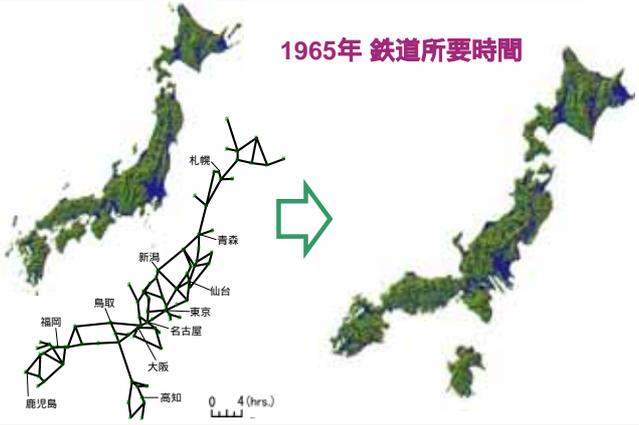
<地理的地点配置>



距離カルトグラムへの地理情報の内挿

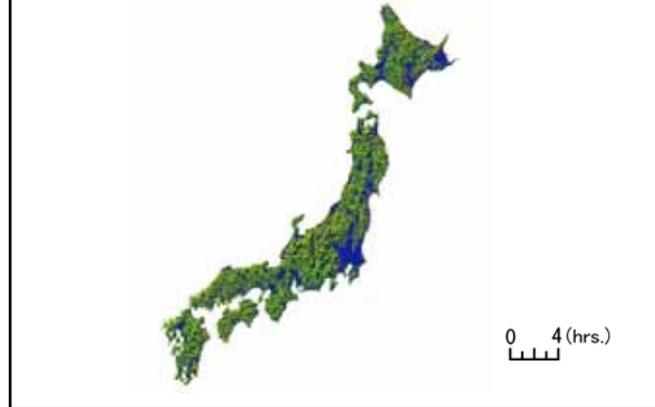
39

1965年 鉄道所要時間



地図から時間地図への変化(1965)

40



時間地図の変化(1965～1995)

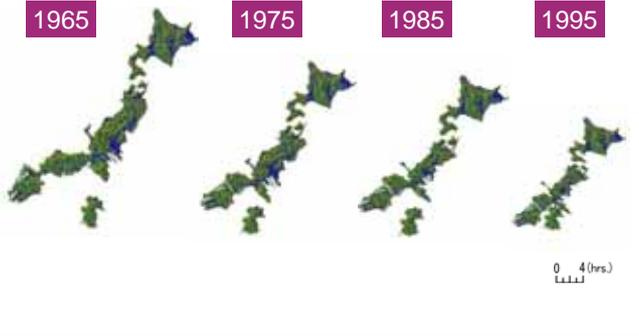
41

1965

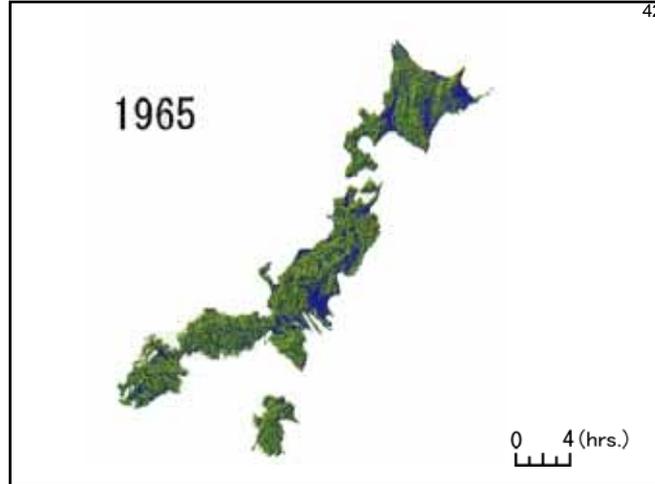
1975

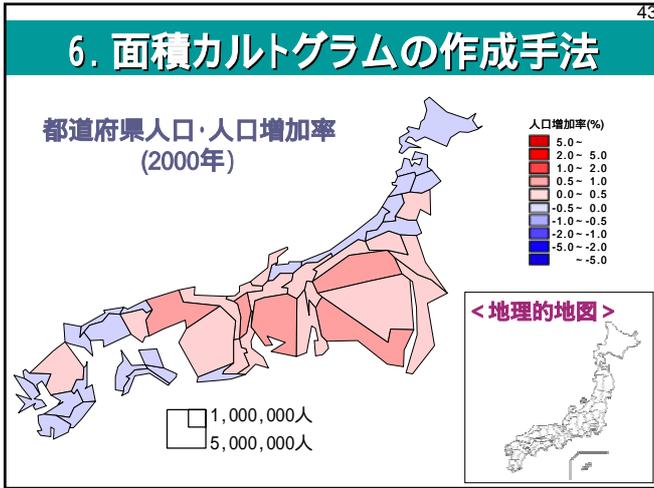
1985

1995



1965





従来の作成手法

1980年 アメリカ合衆国 州人口

- ・Pseudo-Cartogram Method (Tobler, 1986)
データ表現精度 低
- ・Line Integral Method (Gusein-Zade&Tikunov, 1993)
直線 曲線 分かりにくい図
- ・Constraint-Based Method (Kocmoud&House, 1998)
解法が不明快・複雑なパラメータ設定・長時間の計算

従来の作成手法

1980年 アメリカ合衆国 州人口

CartoDraw
(Keim, North & Panse, 2004)

短時間で計算可能
実用的な解法

問題点

- ・再現精度が低い
- ・初期値依存性が高い

過小表示 過大表示

面積カルトグラム作成手法の定式化

$$\min \sum_{i \in R} (D_i - A_i)^2$$

R: 地域の集合
D_i: 地域 i の統計データ
A_i: 地域 i の面積

地域を三角網に分割
(Delaunay三角形分割)

三角形分割を用いた作成手法の定式化

$$\min \left[\sum_{t_{ijk} \in T} (D_{ijk} - A_{ijk})^2 + \mu \sum_{e_{mn} \in E} (\theta_{mn}^C - \theta_{mn}^G)^2 \right]$$

↓

t_{ijk}: 頂点 i, j, k を結ぶ三角形
e_{mn}: 頂点 m, n を結ぶ辺
E: 三角網を構成する辺の集合
μ: 正則化パラメータ

頂点座標(x, y)を用いて表現

地理的地図
辺 mn の方位角

カルトグラム
辺 mn の方位角

$$\min \sum_{t_{ijk} \in T} \left\{ 1 - \frac{1}{4D_{ijk}^2} \left((x_j - x_i)(y_k - y_i) - (x_k - x_i)(y_j - y_i) \right)^2 \right\}^2$$

$$+ \mu \sum_{e_{mn} \in E} \left\{ \frac{(x_n - x_m)(y_n - y_m) - (x_n - x_m^G)(y_n - y_m^G)}{d_{mn} d_{mn}^G} \right\}^2$$

d_{mn}: 辺mnの距離

提案手法の特徴

- ・入力データは統計データと地理的地図データのみ
(統計GISに入力済み)
- ・初期値設定が必要な変数は正則化パラメータ μ のみ
- ・線形最小二乗問題の繰り返しで計算可能
- ・計算時間 短(1分程度)

地域形状 D_{ri}, x_i⁰, y_i⁰ D_{ri}: 地域統計データ
x_i⁰, y_i⁰: 頂点地理座標
x_i⁰, y_i⁰: 頂点地理座標 (総面積) = (総データ) に拡大・縮小

α = ∑ D_{ri} / ∑ A_{ri}
x_i⁰, y_i⁰ ← √α x_i⁰, √α y_i⁰

各地域の三角網分割
D_{ri} ← D_{ri} を按分
A_{ri} ← x_i⁰, y_i⁰

min ∑ [1 + (M_{ijk}² / D_{ijk}²) + (A_{ijk}² / D_{ijk}²) (x_n - x_m - x_n^G - x_m^G) (y_n - y_m - y_n^G - y_m^G) / (d_{mn}² d_{mn}^{G2})]

μ増加

三角形反転 あり

反転回数 < N_{max}

NO

YES

A_{ri} ← x_i⁰, y_i⁰
μ ← ∑ (D_{ri} - A_{ri}) / ∑ (D_{ri} - A_{ri})

NO

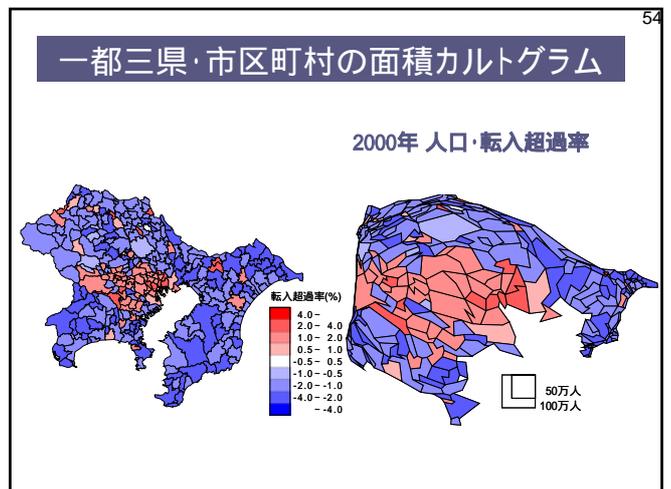
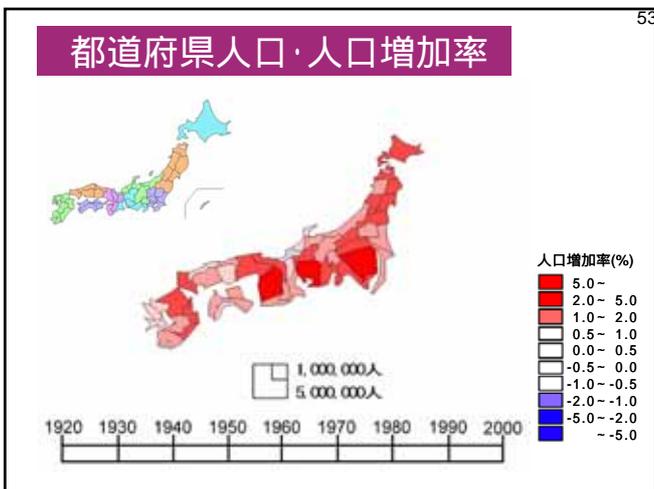
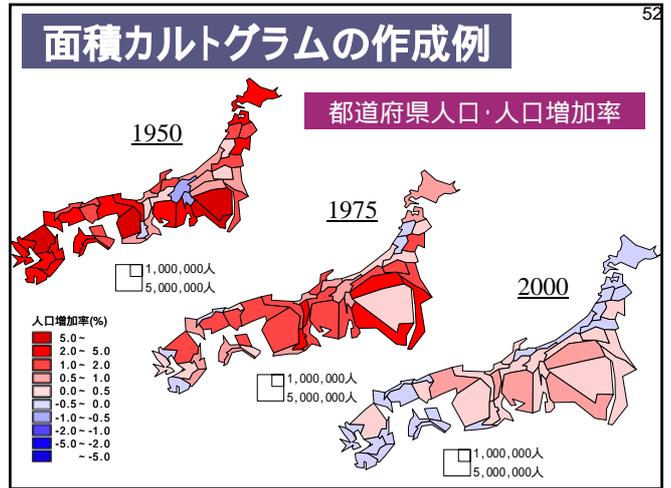
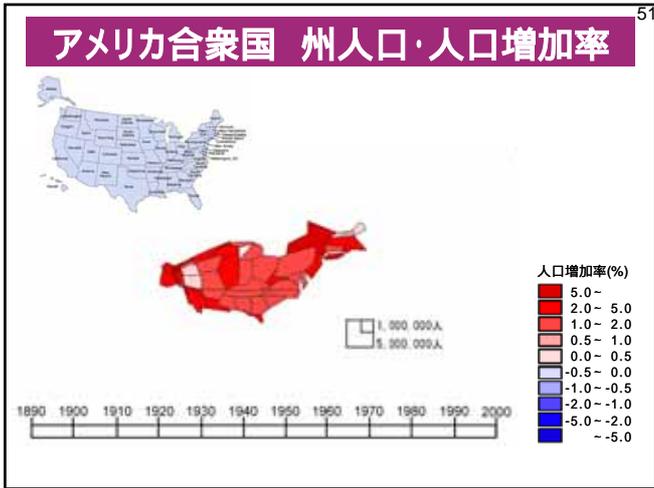
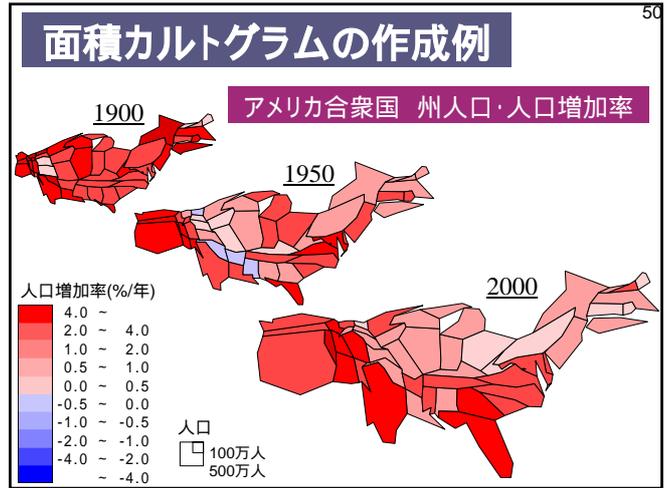
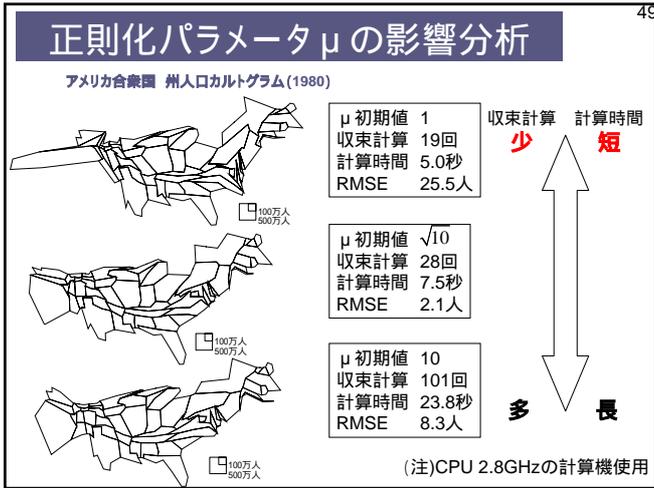
YES

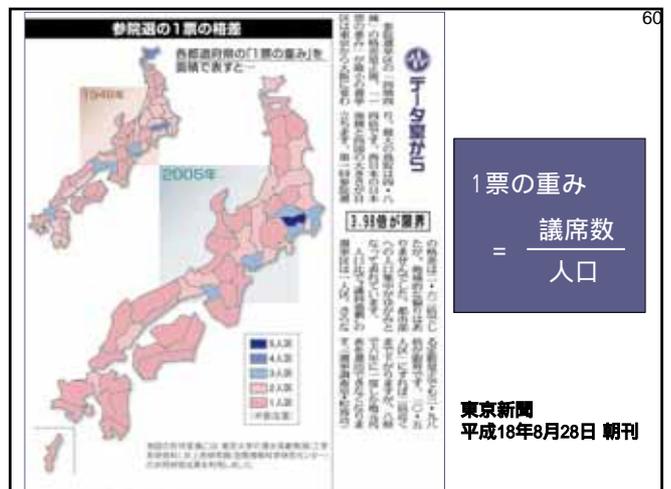
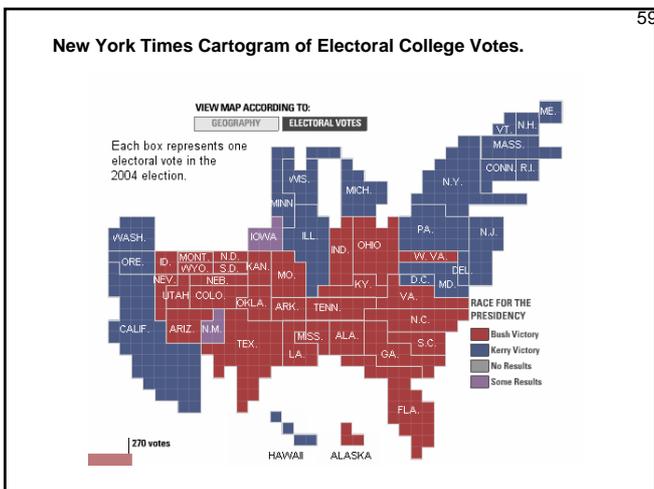
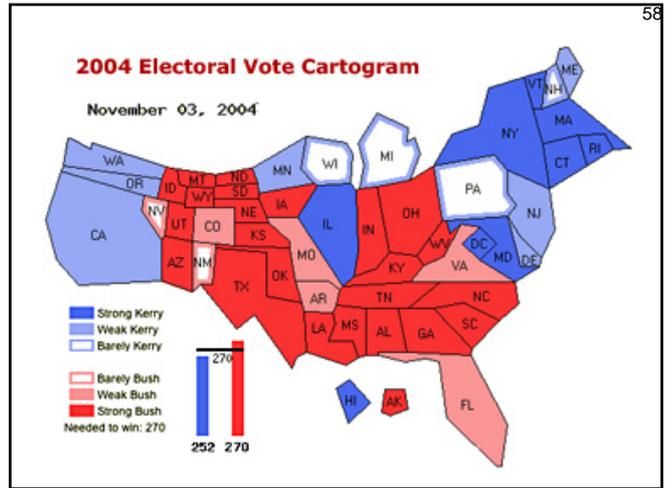
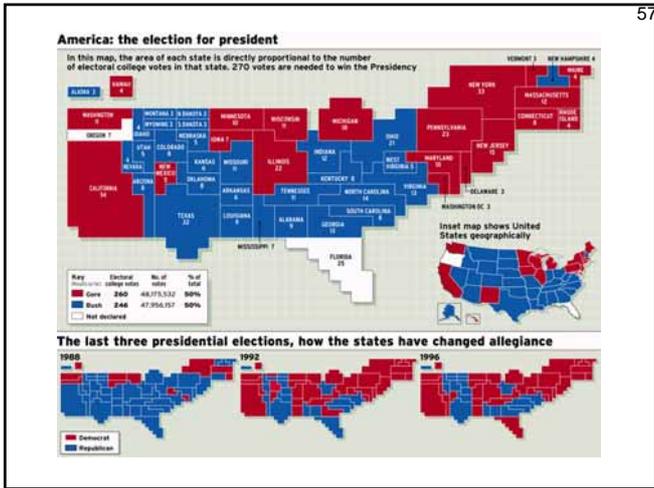
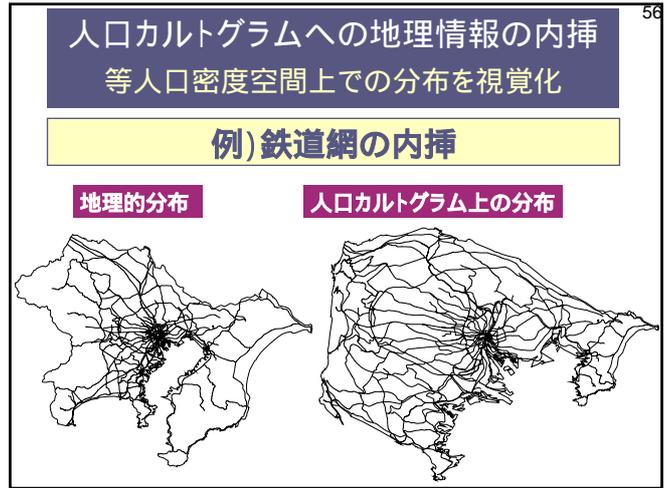
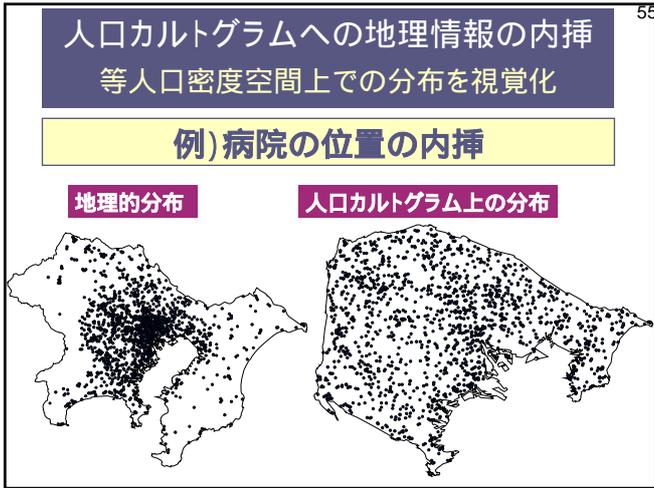
D_{ri} - A_{ri} < EPS

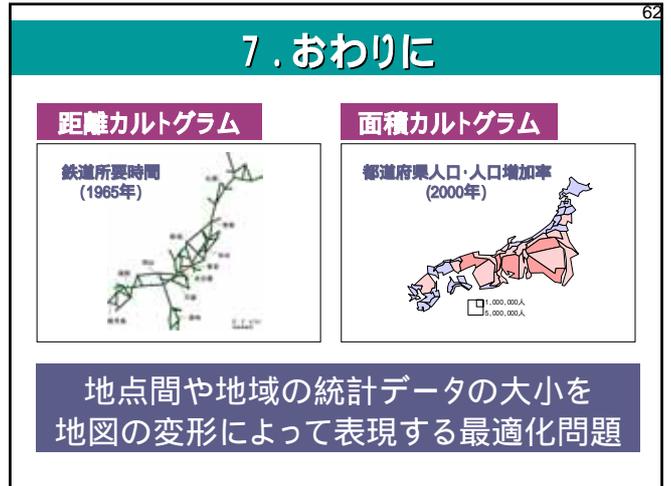
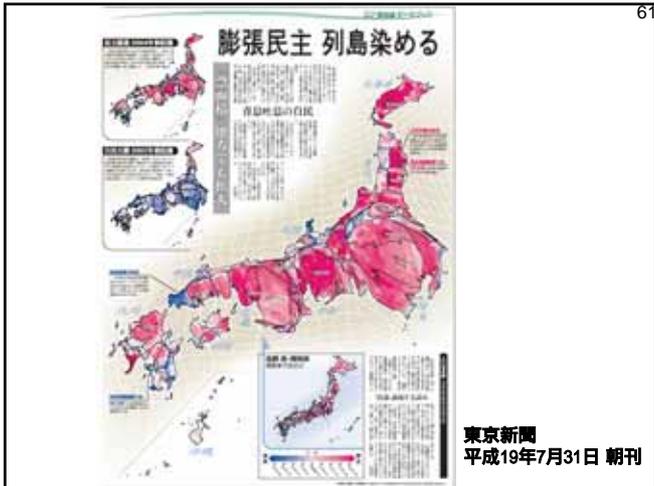
YES

EPS: 収束条件

x_i, y_i







論点

- 不良設定問題の解消
- 無意味な地図変形の抑制
- 単純かつ明快な解法の構築

距離カルトグラム

鉄道所要時間
(1965年)

面積カルトグラム

都道府県人口・人口増加率
(2000年)

論点

- 不良設定問題の解消
- 無意味な地図変形の抑制
- 単純かつ明快な解法の構築

論点が研究者間で十分に共有されず

➡ これまで

- 実用的な手法が殆ど無し
- レビュー研究、比較研究が殆ど無し

本研究で提案する基本的な方法論

- 全体像を単純かつ明快に
最小二乗法で記述
- 不良設定の解消、
無意味な地図変形の抑制
方位角拘束による正則化

➡ 距離・面積カルトグラムの双方について、
一応の実行可能性を有する手法を構築

