

平成25年度

筑波大学情報学群情報科学類

卒業研究論文

題目 大規模な移動データの集約的な可視化

主専攻 知能情報メディア主専攻

著者 兵吾勇貴

指導教員 三末和男 志築文太郎 高橋伸 田中二郎

## 要 旨

GPS 端末の普及を背景に、位置データが大量に収集できるようになった。そして、収集された位置データをもとにした人や物の移動分析が広く活用されている。これらの分析において、対象となるデータを理解するためには可視化が有効な手段である。しかし、大規模なデータを対象とした移動の可視化においては、データが大規模かつ個々の移動が複雑なため、単純な手法によって、情報を提示する事が難しかった。

本研究は、地図上に四分位線を重畳表示し、個々にはベクトルや軌跡で表される移動を集約的に表せる可視化手法を開発した。四分位線とはある地点からの移動距離分布の四分位を表現する図で、移動の方角・距離・移動の量・それらの時空間的变化を読み取ることができる。加えてこの手法では、分析したい移動を静止画で表す事ができる。時間幅のある変化も瞬時に読み取ることができ、連続的に分析条件を変えて結果を見るインタラクティブな操作も可能にする。

# 目次

<b>第1章 序論</b>	<b>1</b>
1.1 位置データの普及	1
1.2 移動データ	1
1.3 移動の分析	2
1.4 一般的な分析における可視化	3
1.5 大規模な移動データの可視化における問題	3
1.6 目的	3
<b>第2章 関連研究</b>	<b>4</b>
2.1 人の移動データ	4
2.2 2次元ベクトル空間の可視化	4
2.3 地理空間上の移動の可視化	5
2.4 大規模データの可視化	5
<b>第3章 課題の整理</b>	<b>6</b>
3.1 分析に対する要求	6
3.1.1 個々の移動に対する要求	6
3.1.2 複数の移動に対する要求	6
3.2 可視化における問題	7
<b>第4章 提案手法</b>	<b>9</b>
4.1 概要	9
4.1.1 大規模データの集約的表現	9
4.1.2 多地点の移動を表すための small multiples 技法	9
4.2 前処理	10
4.3 アメーバ表現	10
4.3.1 表現の検討	11
i 移動を表すための情報	11
ii 表現が持つ視覚変数	11
iii 対応付けのパターン	13
iv 移動量の表現の検討	13
v 統括	15

4.4	アメーバコロニー表現 . . . . .	18
<b>第 5 章</b>	<b>可視化例</b>	<b>19</b>
5.1	利用したデータ . . . . .	19
5.2	アメーバ表現の様子 . . . . .	19
5.3	アメーバコロニー表現の様子 . . . . .	20
<b>第 6 章</b>	<b>結論</b>	<b>23</b>
	謝辞	24
	参考文献	25

## 目次

3.1	72,000 人の移動を線によって表した様子 . . . . .	8
4.1	アメーバ表現の概観 . . . . .	10
4.3.2	移動量を四分位線に割り当てた様子 . . . . .	16
4.3.3	移動量を方位ごとの面に割り当てた様子 . . . . .	17
4.4.1	アメーバコロニー表現の概観 . . . . .	18
5.2.1	東京都市圏 7 万 2 千人分の人の移動データの 4 地点におけるアメーバ表現 . . . . .	21
5.3.1	東京都市圏 7 万 2 千人分の人の移動データに対しアメーバコロニー表現を用いた様子 . . . . .	22

# 表目次

1.1	移動データ	2
4.1	アメーバ表現が含む図形とその視覚変数	12
4.2	表現する値と視覚変数との対応（原案）	13
4.3	移動量の表現案	14

# 第1章 序論

人々の生活や社会活動の中で、人や物の ID・時刻・位置がひも付けされた位置データが大量に記録されるようになった。これらのデータをもとにした移動の分析が広く行われており、これらの分析からは、広告・防災・都市計画など広い分野に渡る知見を得ることができる。一般的な分析において、対象となるデータを理解するには可視化が有効な手段である。しかし大規模な移動データにおいて、移動の軌跡を線で表すなどの素朴な方法による可視化では情報の読み取りに問題が生じる。この研究では、大規模な移動データの可視化における問題を解決した新たな可視化手法の開発を目的とする。

## 1.1 位置データの普及

身近な GPS 端末の普及などにより、大量の位置データが得られるようになった。今日の携帯情報端末には GPS が基本的な機能として搭載されており、それらの GPS 機能は、移動のナビゲーションだけでなく、位置情報を付加した Twitter・Facebook などへの投稿、自分の居場所の情報を共有するサービスなど、位置情報の多様な利用を可能にした。また、携帯情報端末以外から収集されている位置データもある。例えば鉄道の運賃支払に使われる交通系 IC カードは、駅の改札を利用した時刻と利用者の情報が保存されている。さらに電子媒体以外にも位置データの収集が行われている。例えば、公共団体等が公共交通機関の利用実態を調査する目的などでアンケートを行い、おおよその人の移動の様子を調べることが行われている。

人の位置データに関するもの以外では、車や飛行機といった乗り物の位置情報も多く得られるようになった。例えば車の場合には、車載ナビゲーションシステムの位置情報が自動車メーカーによって収集されている<sup>1</sup>。また高速道路の ETC 利用履歴なども位置データとして捉えることができる。

## 1.2 移動データ

ここで位置データについて整理してみる。位置データとは、緯度・経度のデータと、その経緯度が何の位置を表しているのが識別する ID が記録されたデータと定義する。識別する ID には、SNS のアカウントや GPS チップのロットナンバーなど、一意に物体を識別できるもの

---

<sup>1</sup>東日本大震災では多くの道路が通行できなくなった。本田技研工業（株）は、自社製車のカーナビから収集している GPS 情報をもとに、震災後に通行実績のあった道路のマップを公開した。その他にも建設機器メーカーのコマツでは、製品に組み込まれた GPS から得たデータを製品のアフターサービス等に活用している

が利用できる。位置に関しては、データが得られる媒体によって、緯度経度の形式で表されている場合が考えられる。その場合には、住所を経緯度に変換する Geocoding 技術などによって経緯度に変換可能である。したがって、位置に関するデータはほぼ <経緯度, ID> の形式に整えることが可能である。

位置データが時間的に連続して記録されたものを移動データ（表 1.1）と呼ぶこととする。時刻と、その時刻における位置データの組によって 1 つのレコードを成し、そのレコードが複数記録されたものが移動データである。移動データには、複数の ID の位置データを含むものとする。仮に同じ ID に対する複数のレコードが存在する場合、それらのレコードに記録された位置を時系列に追っていくことで、その ID の物体が移動していく様子がわかる。その時、記録される時間間隔が短くなれば詳細に移動経路を把握することができる。

本研究では、こうした移動データのうち、特に大規模な移動データを対象とする。データの規模としては数万以上の物体を対象とする。また位置情報は緯度・経度で表されているものとする。

表 1.1: 移動データ

識別 ID	緯度, 経度	時刻
50	36.110915, 140.099898	201304051300
50	36.094774, 140.098367	201304051700
51	36.098233, 140.105161	201304051300
:	:	:

実世界中の位置を一意に決めるためには緯度・経度が利用される。緯度経度は、角度によって地球上の位置を示すものである。緯度は赤道を 0 度とし、南北へ 90 度までの角度を割り当てている。同じ角度の数値でも、赤道より北の緯度は北緯、南は南緯で区別される。経度はイギリスの旧グリニッジ天文台を通過して北極と南極を結ぶ線（本初子午線）を 0 度として、東西へ 180 度までの角度を割り当てる。経度は東側を東経、西側を西経と呼び、区別される。

実際の緯度経度のデータにおいて緯度経度は、度・分・秒で表され、分と秒は 60 進数で表されている。秒をさらに細かに表す場合には、小数点以下を 10 進数にて表す。また度は [°], 分は ['], 秒は ["] の記号を用いることもある。(例えば 36°5'41.19"N は北緯 36 度 5 分 41.19 秒を表す)。また、近年では分・秒を用いず 10 進数にて 1 度未満の数値を表すことも行われる。(例えば N36.094774° は北緯 36 度 5 分 41.19 秒を表す)。

### 1.3 移動の分析

大量の移動データが日々生み出される中で、それらのデータから人や物の移動を分析することがなされるようになった。人や物の移動の分析は、防災・都市計画・広告など幅広い分野に有益な情報をもたらしている。具体例として静岡県熱海市では、携帯電話端末の GPS デー



タから観光客の移動の動線を調べ、周遊観光を後押しするための計画を立案している。他にも道路やバス運行計画などの都市計画や、感染症の拡大予測モデルの作成など様々なことに移動データを役立てることができる。

## 1.4 一般的な分析における可視化

一般に、データ分析の主な目的には予測と理解が挙げられる。予測とは与えられたデータから、今後起こりうる事象やその確率等を推し量ることである。理解とは与えられたデータの持つ傾向や性質を人間が把握することである。気象予報を例にとると、全国で観測された気象のデータをもとに、明日以降の天気や降水確率を求めることが予測にあたる。同じ気象の観測データから、現在までの大気の移動や降雨があった地域などの情報を知ることが理解にあたる。

データの理解では、人間がそのデータの様子を知る必要があり、データを人間にとってわかりやすい形で提示する必要がある。可視化はデータを人間にわかりやすく提示するための手段であり、理解を手助けする有効な手段である。

## 1.5 大規模な移動データの可視化における問題

データの理解には可視化が有効な手段ではあるが、大規模な移動データの場合には、視覚的混雑度が高くなり情報を読み取りにくくなる問題がある。例えば、都市圏全体の移動の様子を知るために移動の経路を地図上の線によって表現する可視化手法を用いるとする。すると人や物の往来が多いところでは線同士が重なりあい、視覚的混雑度が高まってしまう。この表現においては、距離や方向が多様に異なっている移動を読み取ることはできない。

## 1.6 目的

この研究では、大規模データの可視化における問題を解決した可視化手法を開発することを目的とする。人や物の数が大規模であったり、移動の範囲が広範囲に渡る情報を簡潔に表すことのできる可視化手法を開発する。

## 第2章 関連研究

関連研究について述べる。移動データの収集に関して、人の移動を調べるアンケートの結果をもとに、時間的・空間的に位置情報を内挿する技術がある。この技術によって内挿された移動データに対して提案手法の可視化を試みた。また、既に行われている移動の可視化について、2次元ベクトル空間の可視化と地理空間上の移動の可視化の2つについて調査を行った。2次元ベクトル空間の可視化は、洋上の風や流体の可視化などに用いられており、地理空間上の移動の可視化においても、2次元ベクトル空間の可視化手法を応用したものがある。最後に、大規模データの可視化に関する研究についても調査を行った。調査した論文中には大規模な移動データに対しても応用できる概念が述べられている。

### 2.1 人の移動データ

人の移動データは様々な媒体によって収集されている。Sekimotoらは、パーソントリップ調査と呼ばれる移動のアンケート調査から得られる断片的な位置データから、統計的な処理によって移動経路を補間する手法を開発した [1]。この手法は他の移動データにも応用することが可能である。

### 2.2 2次元ベクトル空間の可視化

大規模な移動（流れ）データの可視化に関して、もっとも簡潔な表現としてメッシュ状に区切られた領域に矢印を描くことが考えられる。Fieldらは矢印が連続的に連なっていると流れの認識が容易となることを示した [2]。更にTurkらは、矢印ではなくStreamlineと呼ぶ線を用いて効果的に流れの場を表している [3]。このStreamlineは、定義されたエネルギー関数の値を減らしていくように反復的に計算を行い、その収束結果から得られる。また、Laidlawらは2次元ベクトル空間の流れを表す6つの手法について、ユーザスタディを行い、いくつかの知見を得ている [4]。これらの手法は、平面上の1点において1つの方向への移動のみを表すものである。そのため個々に様々な方向・距離で移動する人や物の移動の情報を表すことが難しい。

## 2.3 地理空間上の移動の可視化

地理空間上における人の移動について、Tobler はコンピュータを用いた大規模なデータの可視化について考察を行っている [5]. 論文では、地点間を移動した量が記録されている”from-to” テーブルのデータに対して、離散的な表現と連続的な表現によって表せることを示した. 連続的な表現としては、2次元ベクトル場の可視化に用いられるような **Streamline** の手法を、離散的な表現には矢印を用いた手法を提案している. この論文が対象とするデータは、2地点間を移動する人数が記録されたデータであり、任意の地点における移動の方向や距離などを表すものではない.

Andrienko は移動の分析フローに基づいた可視化を行った [6]. 分析において、データが持つセマンティックな情報を分析を行う人に対して提示するため、データ以外の情報源へのリンクの提示を行っている. また移動が日や週などの一定周期で起きていることに着目し、大規模データから目的データを絞り込むための時間フィルタを組み込むなど種々の工夫が凝らされている. 本研究では、大規模データに対して時間フィルタを利用する点を参考とした.

## 2.4 大規模データの可視化

移動データに限らない大規模データの可視化について、Shneiderman は **atomic**, **aggregated**, **density plots** の3つの方法についてまとめている [7]. **Atomic** な可視化とは、1つのデータレコードを1つの図形に割り当てて描画するものである. 簡単な散布図が **Atomic** な可視化の例にあげられる. この方法では、大規模データの場合には視覚的混雑度が高まる影響や画面の解像度を越えた情報を提示できないという問題が発生する. それに対して、**Aggregated** な可視化では、1つの図形によって複数のデータレコードを割り当てるものである. 本研究では、大規模な移動データに対して **Aggregated** な可視化を行い、視覚的混雑度を抑えている.

## 第3章 課題の整理

移動データの分析に対する要求は様々に存在する。まずはじめに、それらの要求を、要求が対象としている移動の規模によって整理を行った。次に、実際に大規模な移動データを素朴な方法によって可視化を行い、その結果をもとに、可視化における課題をまとめた。

### 3.1 分析に対する要求

移動データの理解には、読み取りの対象となる移動の規模に応じた要求が存在していると考えられる。それは、個々の移動の様子を詳しく知りたいという要求と複数の移動の様子をまとめて知りたいという要求である。それぞれについて、3.1.1節、3.1.2節で述べる。

これらの分析要求を満たす可視化が求められるが、単純な可視化手法では要求を満たすことができない。このことを3.2節にて詳しく述べる。

#### 3.1.1 個々の移動に対する要求

1つの物体の移動を表すための基本的な情報には、移動の出発地・到着地・経路・移動時間などが考えられる。これら表現することがはじめに求められる。この基本的な情報をもとにして、出発・到着地点間の距離や方角といった2次的な情報も合わせて表現されるべきである。また、分析においてはなぜそのような移動が起きているのか、原因を推測するためには周辺の地理的な情報も必要である。例えば、道路や鉄道などの交通網・河川や丘などの地形・住宅地やオフィス街といった土地利用の状況などである。

#### 3.1.2 複数の移動に対する要求

複数の移動では、物体それぞれによって移動の方向や距離が多様である。またそうした移動が広い範囲で起こっている。分析要求として、こうした多様で広範囲の地点で起こっている移動の様子を簡潔に読み取れるように示すことが求められる。一般に多様な値をとる情報の様子を知るためには統計量が用いられる。統計量によってその集団の特徴を知ることができるからである。例えば、小学校のクラス全体の学力を知りたいと思った時、テストの点数の平均や分散といった統計的な数値を知ることでおおよそのクラスの学力が推し量れる。移動の分析においても同様である。移動の特徴を表す統計量としては以下の物が考えられる。

- 似ている移動をしている移動量（人数・物量）

- ー 移動距離の分布（最大値や平均値，分散値など）
- ー 移動方向の分布（最頻値など）

これらを読み取ることによって移動の傾向を把握することができる。また広範囲の地点で起こっている移動を知るためには、各地点での移動の特徴が理解できるとともに、複数の地点が示す移動の傾向や特徴の情報が求められる。例えば、人の移動の場合、朝の時間帯に都心へ向かう移動が多い地点は郊外の住宅地であるということがわかる。こうした複数地点の移動の傾向や特徴を知るためには、地点ごとの移動が比較可能である必要がある。各地点の特徴が示されていても、比較が難しいと広域に渡る様子は理解できない。なぜならば、特徴的な移動を行っている地点を見つけ出すには、まず周辺の地点でどのような移動が起きているのかを知らなければならないからである。例えば、ある地点で北に向かう大きな移動があったとしても、周辺地点でもそのような移動が多ければ、北に向かう移動が特徴的な移動とはいえないからである。

これらに加えて、時間の経過を伴う移動を、アニメーションではなく静止画によって即座に読み取れると分析効率がよい。アニメーションでは、長時間に渡る移動を表すために、アニメーションの再生時間が長くなってしまふことによる時間的な効率が悪い。さらに、アニメーションによる表現の場合、場所や時間帯の異なる2つの移動の様子を比較する際に、2つの移動を目で追いかける必要があり、視覚的な認識に困難をきたす。

## 3.2 可視化における問題

移動の経路、出発地、到着地を表す方法として、地図上に個々の移動経路を線で描くことが考えられる。更に単純に、東京都市圏7万2千人の移動データを基に線で移動の様子を示したものが図3.1である。これは朝6時の位置と、3時間後の9時頃の位置を直線で結んだものである。こうしたシンプルな表現では、多くの人が行き来する場所では線が幾重にも重なろう。そのため視覚的混雑度が増す問題が生じる。図3.1では、人の往来の多い都心部において線を読み取ることができない。この問題を解決するために線を半透明にする方法も考えられるが、逆に個々の移動が見え難くなる。このように、単純な方法では個々の移動の詳細な様子を表しつつ、移動が多い地点における移動を表すことは難しい。

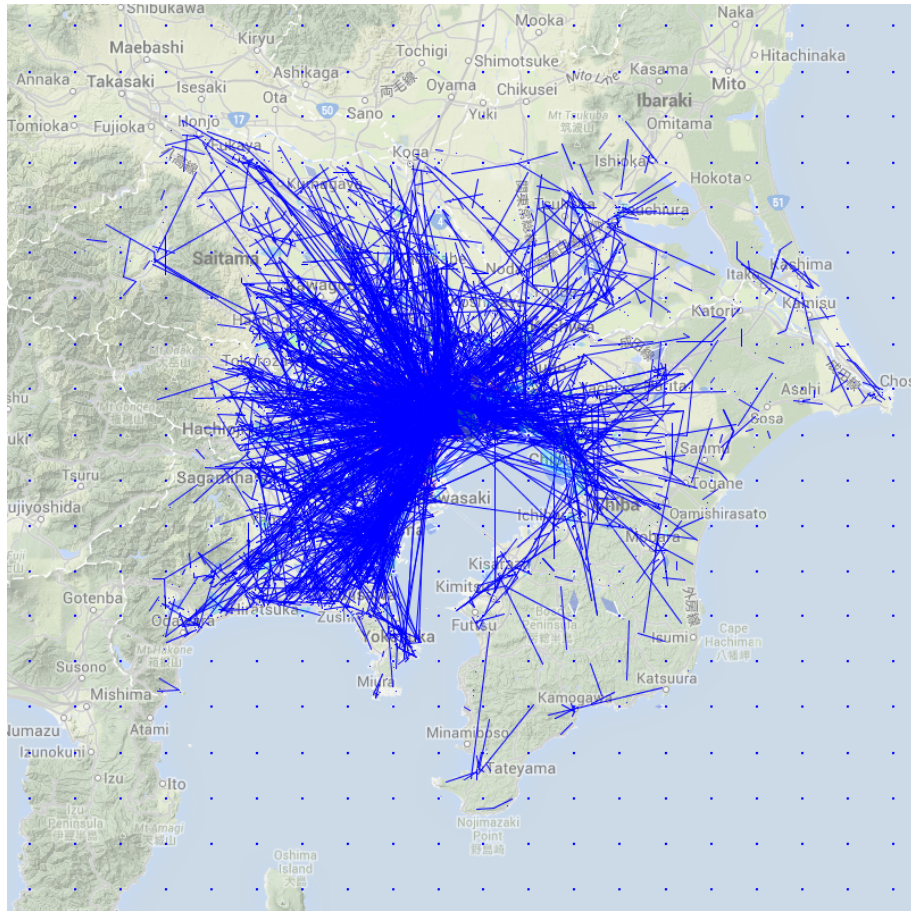


図 3.1: 72,000 人の移動を線によって表した様子

## 第4章 提案手法

移動データを表現するために、地理空間上のある1点の移動の様子を表す「アメーバ表現」と、広い範囲の移動を表す「アメーバコロニー表現」を開発した。どちらの表現も移動データを集約的に表すことによって、大規模なデータを簡潔に表現している。

### 4.1 概要

大規模な移動データの分析要求を満たすため、2つの解決法を用いることとした。1つ目は複数の移動を集約的に表現する方法である。これによって大規模データにおいても視覚的混雑度を抑えた簡潔な表現を実現できる。2つ目は複数の図を小さく並べる方法である。これは、広範囲に渡る移動データの様子を表すためのアイデアである。

#### 4.1.1 大規模データの集約的表現

大規模な移動データを可視化するために、集約的な表現を用いることとした。これは複数の移動の1つ1つを描画するのではなく、意味のある塊として描画する方法である。提案手法では、地理空間上の任意の1点から出発する複数の移動の集団を対象に、その集団の移動の様子を集約的に表現した。この方法では、個々の移動の経路や出発地・到着地を表現することはできないが、大規模な移動データを簡潔に表現することができる。

#### 4.1.2 多地点の移動を表すための small multiples 技法

対象データでは移動は地理空間上の広い範囲に及んでいると考えられるため、任意の1点のみの移動を表すだけでは不十分である。広い範囲の移動を表すために、small multiples の技法を用いることとした [8]。small multiples とは、規則的に小さな図形をならべる技法で、その小さな図形の情報をひと目で比較することができる。例えば、散布図を規則的にならべる散布図行列も small multiples の技法を用いていると言える。この技法を用いてアメーバ表現を多地点に並べることによって、地理空間上の広い範囲の移動の様子を提示し、比較可能にする。

## 4.2 前処理

移動データに対して、前もって統計量を求める処理を行う。前述の通り、本手法では地理空間上のある1点から出発する複数の移動を集約的に表現する。移動の様子を表す情報には、出発地・到着地・経路・移動時間などが考えられるが、本手法では、出発地と到着地の2地点に着目した。まずはじめに地理空間上のある1点の近傍から出発する移動の集団を抜き出す。この集団を、出発地点から到着地点へ向かう方位によって8つの小さな集団に分ける。方位には、東西南北と各方位の中間（北東、南東、南西、北西）の8つを選んだ。そして、各方位の中で移動の様子を表すために、移動距離の分布の統計量を求める。求める統計量は、8つの方位で分けた集団それぞれで、出発地から到着地までの距離を昇順に並べた時の四分位の値（第1四分位、中央値、第3四分位、最大値）である。最小値は全く移動を行わなかった場合の0mである。また各方位ごとに移動量（人数・物量）が異なっているので、それらも合わせて求める。

## 4.3 アメーバ表現

地理空間上の任意の地点の移動の様子を表すため、地図上に重畳表示する視覚的表現を開発した。この図形を、その形状から「アメーバ表現」と名付けた。

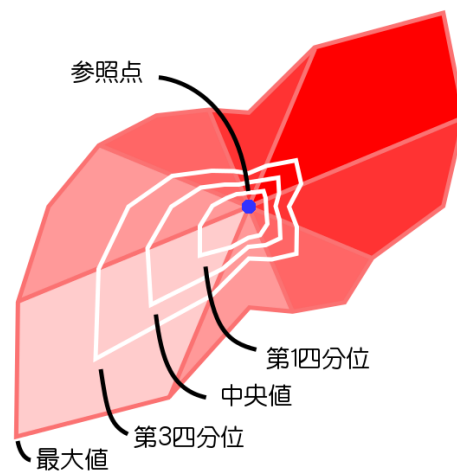


図 4.1: アメーバ表現の概観

アメーバ表現の外観を図 4.1 に示す。アメーバ表現は図形の中央に位置する参照点、それを囲む4本の環状の線と、線と面から構成される。線と面は参照点を中心に放射状に8つの方位で分かれており、線の参照点からの距離によって、それぞれの方位における四分位を表現している。そして面の色の不透明度によって方位ごとの移動量を表す。最も外側の線が各方位の移動の最大距離を表しており、距離はこのアメーバ表現が重畳表示されている地図の



縮尺に合わせられている。最大距離の線から参照点に向かって順に第3四分位，中央値，第1四分位を表す。これらの四分位を表す線を，四分位線と呼ぶこととした。各方位の四分位線は，隣接する方位同士で接続され，輪を描いている。線が接続する位置は，隣り合う四分位の算術平均から決定される。これは四分位線の輪ができるだけ滑らかにつながるようにするためである。また，各方位の移動量に応じて面を塗りつぶす色の不透明度を変化させる。移動量が大きいほど不透明度が高く，小さいほど不透明度は低い。

これらの表現により，参照点付近からの移動の方向・距離・人数を読み取ることができる。図4.1の例では，右上の方向へ向かって移動量が大きな移動があることがわかる。左下の方向へは移動量が小さい。四分位線を読むとさらに状況を細かく読み取ることができ，右上，左下ともに最大移動距離が大きな移動がある。しかしながら，右上の方向は他の四分位が参照点に近いので，ほとんどの移動は参照点近傍にとどまっている事がわかる。この様にして，アメーバ表現は多様で大規模な移動を，集約的な表現を用いることによって簡潔に表現している。

この表現は，分析したい時間帯を指定した上で利用される。指定する値は，移動の開始時刻と終了時刻（もしくは移動時間）である。これらを指定した後，地図上の任意の地点にアメーバ表現を表示することで，その地点の指定された時間帯の移動を表示する。

### 4.3.1 表現の検討

図4.1のようにアメーバ表現を設計し，その表現が適切であるかどうかの検討を行った。具体的には移動を表すための情報とアメーバ表現が持つ視覚変数の組合せを様々に変えた場合について，読み取りに誤解を生じることがないかどうかを考察した。

視覚変数とは，点や線などの図形を視覚的に修飾する属性のことである。例えば点の位置は1つの視覚変数である。Bertinは，自身の著書において，点・線・面の図形が持つ視覚変数を変えることによって，その図形が表現する情報を変えることができると述べた [9]。Bertinは視覚変数として，XY平面上の位置・大きさ・濃淡・きめ・色・方向・形を挙げた。

#### i 移動を表すための情報

移動の方向・距離・移動量を表現するために，アメーバ表現では各方位ごとに移動距離分布の四分位と数量を示している。また参照点の位置によって移動の出発地点を示している。

**出発地点** 地理空間上の移動の出発地点。1つのアメーバ表現に対し1つ。

**四分位** 各方位の移動距離分布の四分位にあたる距離。各方位に4つ。8方位分。

**移動量** 各方位への移動量（人数・物量）。8方位分。

#### ii 表現が持つ視覚変数

アメーバ表現が含んでいる図形と，その視覚変数を表4.1に列挙する。

表 4.1: アメーバ表現が含む図形とその視覚変数

図形	可変な視覚変数	変更	説明
参照点を表す丸い点	位置	不可	丸を置く位置
	大きさ	可能	丸の大きさ
	濃淡	可能	塗りつぶし色の濃さ
	きめ	可能	点をパターンで塗りつぶした時のきめの細かさ
	色	可能	色相, 明度, 彩度, 不透明度
	方向	可能	点が指し示す方向
	形	可能	点の形 (四角や三角等)
四分位を表す線	位置	不可	線が通る位置
	大きさ	可能	線の太さ
	濃淡	可能	塗りつぶし色の濃さ
	きめ	不可	線をパターンで塗りつぶした時のきめの細かさ
	色	可能	色相, 明度, 彩度, 不透明度
	方向	不可	パターンが指し示す方向
	形	不可	線の形 (四角や三角等)
各方位の面	位置	不可	面を置く位置
	大きさ	不可	面の大きさ
	濃淡	可能	塗りつぶし色の濃さ
	きめ	可能	面をパターンで塗りつぶした時のきめの細かさ
	色	可能	色相, 明度, 彩度, 不透明度
	方向	可能	パターンが指し示す方向
	形	不可	点の形 (四角や三角等)

表 4.1 に列挙した視覚変数の中で、変化させることのできない視覚変数が存在する。以下に挙げるものは、アメーバ表現上の制約によって変化させることができないものである。

参照点を表す丸い点 位置  
 四分位を表す線 位置  
 各方位の面 位置・大きさ・形

参照点を表す丸い点については、背景に表示する地図の場所を示す働きがあるため、その位置をずらすことはできない。また四分位を表す線についても、参照点からの距離・方向によって場所が一意に定まるので、変更することはできない。四分位を表す線の位置を変化で

きないので、それらの間に位置する面もその位置・大きさ・形を変えることができない。

その他に、四分位を表す線の視覚変数のうち、きめ・方向・形は変化させるべきではない。これは、線をパターンによって塗りつぶすことが困難だからである。四分位線は、移動量が極端に少ない場合や移動距離分布の分散が小さい時に重なる場合がある。また地図上に描かれている四分位を表す線は、線が占める領域によって、意図せず地理空間上の場所を指し示すように読み取られる可能性がある。こうした問題からアメーバ表現では、視認性を損なわない範囲で線を細く描画している。細い線では線上にパターンを描画できないため、きめ・方向・形の視覚変数を変化できない。

### iii 対応付けのパターン

図 4.1 に示しているアメーバ表現では、次の表 4.2 の通りに値と視覚変数を対応付けている。

表 4.2: 表現する値と視覚変数との対応（原案）

表現する値	図形	視覚変数	説明
出発地点	参照点	位置	地図上の移動の出発地点に参照点を置く
四分位	四分位線	位置	各方位の四分位を通るように線を引く
移動量	面	色（不透明度）	移動量が大きいほど不透明度を高くする

表 4.2 の対応付けについて、出発地点と四分位の表現はどちらもアメーバ表現の背景に表示される地図との対応付けしている。そのため直感的に値を読み取ることができる。そのため読み取りに誤解を生じることは少ないと考えられる。移動量の表現については、その他の視覚変数を用いても表現が可能である。よって移動量の表現について、アメーバ表現が持つ視覚変数との様々な対応付けを検討することとした。

### iv 移動量の表現の検討

移動量を表すために可能な視覚変数を表 4.3 に列挙した。これらを実際にアメーバ表現に対して適用した。図 4.3.2 に、線に移動量を割り当てたものを、図 4.3.3 に、面に移動量を割り当てたものをそれぞれ示す。線と面の視覚変数を変化させたそれぞれのアメーバ表現は、すべて同じデータを用いた。そのデータにおいて、各方位の人数は、上半分の方位では左から右に向かって人数を線形に変化させた。下半分については、左から右に向かって 2 倍となっていく等比数列状に変化させた。左と右の方位は上半分と下半分で共通となるようにしている。また四分位は参照点から外側に向かって 2 倍となっていく等比数列状に変化させた。この同じデータを用いた線と視覚変数を変化させる L1~L4、面の視覚変数を変化させる S1~S3 のアメーバ表現についてそれぞれに対して検討を行った。

表 4.3: 移動量の表現案

No.	図形	視覚変数	表現の仕方
L1	四分位線	大きさ	太さによって量表現
L2	四分位線	濃淡	不透明度によって量表現
L3	四分位線	きめ	線を点線とたどきの点の数
L4	四分位線	色	色相の連続的な変化によって量表現
S1	面	濃淡	不透明度によって量表現
S2	面	きめ	縞模様をほどこし縞の密度によって量表現
S2	面	色	色相の連続的な変化によって量表現

#### 四分位線の視覚変数を変化させる表現の検討

線に移動量を割り当てたものについては、各方位の4つの線に対して視覚変数を変化させた。線は同じ方位でつながっていないため、線に移動量を割り当てた表現は各方位を1つのまとまりとして認識しにくくなっている。線に移動量を割り当てた表現 L1 から L4 までを順に検討していく。

##### L1: 太さを変化させた表現 図 4.2(a)

線の太さを 0~10[px] までの 11 段階に変化させている。これ以上の段階で移動量を区別しようとした場合、近接する四分位同士で線が重なる問題が生じてしまう。

##### L2: 濃淡を変化させた表現 図 4.2(b)

赤色 (#FF0000) のアルファ値を 16~255 までの 240 段階に変化させている。この表現は濃淡の具合を読み取るためには線が細い。見やすくするために線を太くすると、近接する線が重なる問題が生じてしまう。

##### L3: きめを変化させた表現 図 4.2(c)

点の個数を移動量に応じ 4 段階に分けている。参照点に近い四分位線において点同士が近づいて重なってしまうため、段階数を 4 段階以上にするのが難しい。また隣り合う方位で点が連続しているように見えるため、区別がしにくくなっている。

##### L4: 色を変化させた表現 図 4.2(d)

色相を変化させることによって色を表している。色は量が多いところを赤、少ないところを青としている。この表現ではそれぞれの方位を区別することが比較的容易ではある。しかしながら、様々な色の大小関係がわかりにくい。

#### 各方位の面の視覚変数を変化させる表現

各方位の面の視覚変数 3 つを、方位への移動量をもとに変化させた。実際の個々の移動において、到着地点はアメーバ表現の面が占める地図上の領域の中に収まっている。面の視覚変数を用いる方法では、到着地点が面の示す地図上の領域に収まっていること直

接的に表現できている。そのため面の視覚変数に対して何らかの変化を持たせることは読み取りに誤解を生じにくいと考えられる。また、線の視覚変数を変化させる方法では、方位の区別が全般的に難しい点があったが、面の視覚変数を変化させる場合には、方位の区別は容易である。

#### **S1: 濃淡を変化させた表現 図 4.3(a)**

赤色 (#FF0000) のアルファ値を 16~255 までの 240 段階に変化させている。これはアメーバ表現の概観 (図 4.1) に用いている表現と同様である。

#### **S2: きめを変化させた表現 図 4.3(b)**

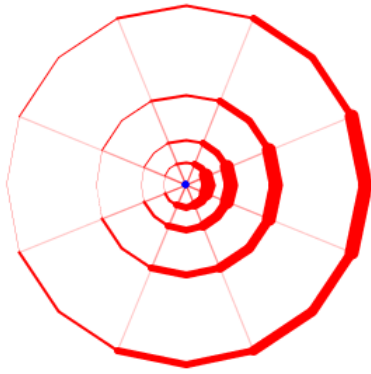
面に白黒の縞模様を描画し、その縞の太さを変化させた。縞模様の方向については、縞の方向の変化が情報を表しうるのですべて同じ縞の方向とした。量の読み取りは、線の幅を測ることによって厳密に求めることができる長所がある。しかしながら、視覚的混雑度が高く、背景に表示する地図の読み取りに影響をおよぼす可能性がある。

#### **S3: 色を変化させた表現 図 4.3(c)**

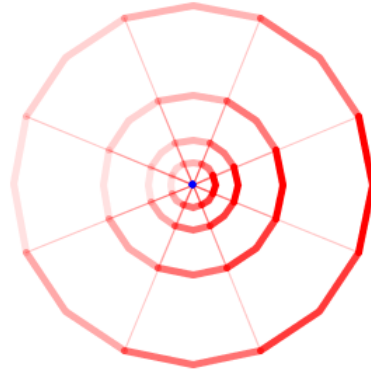
色相を変化させることによって色を表している。色は量が多いところを赤、少ないところを青としている。この表現は、L4 の表現と同じく、それぞれの方位を区別することが比較的容易ではある。しかしながら、様々な色の大小関係がわかりにくい。

## **v 統括**

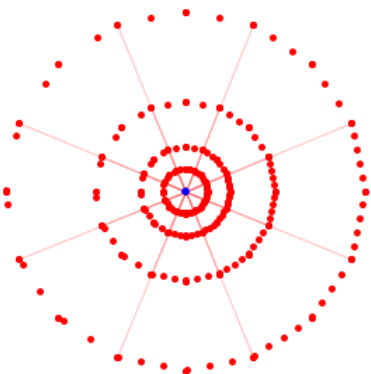
アメーバ表現が持つ図形の視覚変数と、表している値の対応付けについて検討を行った。アメーバ表現が表現中に持つ図形には、図形中央の点 (参照点)、点を環状に囲う線 (四分位線)、線で囲まれた面がある。そして、それらの図形に対して、移動の出発地点、到着地点の方位ごとの移動距離分布から求まる四分位、方位ごとの移動量の 3 つの変数を対応付けて表している。これらの変数のうち、出発地点と移動距離分布の四分位は、それぞれ図形中央の点、点を環状に囲う線によって表現している。この 2 つの変数は、アメーバ表現の背景に表示される地図と関連付けて表現されており、読み取りに誤解を生じにくいと判断できた。残る変数である方位ごとの移動量については、アメーバ表現がもつ図形のいくつかの視覚変数で表すことが可能であった。移動量の表現に用いることができる線と面の視覚変数を列挙し、実際に描画して検討した。線の視覚変数による移動量の表現では、全般的に方位の区別がつきにくいものとなった。これに対し、面の視覚変数を用いる場合には、方位の区別がつきやすい。さらに、個々の移動の到着地点が面の示す地図上の領域に収まっていること直接的に表現できていると考えられた。移動量を表現可能な面の視覚変数には、濃淡・きめ・色を挙げた。きめを移動量の表現に用いた場合には、背景の地図の読み取りに影響を及ぼしうる懸念があり、また色の色相を用いた場合には、量の大小関係がわかりにくいという問題があった。よって、移動量の表現には、面の濃淡を用いる方法が最良と考えられる。



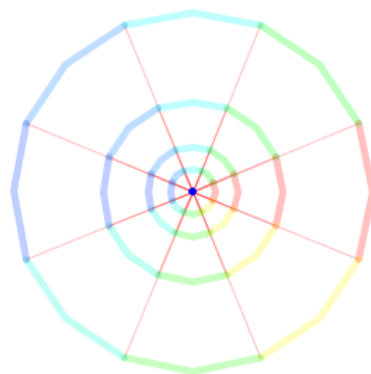
(a) L1: 線の大きさに割り当て



(b) L2: 線の濃淡に割り当て

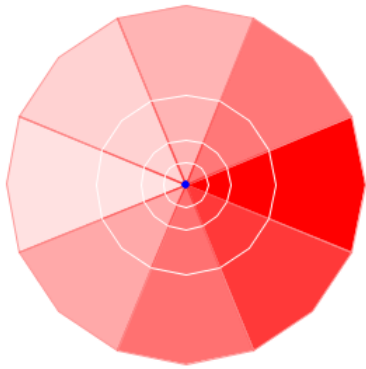


(c) L3: 線のきめに割り当て

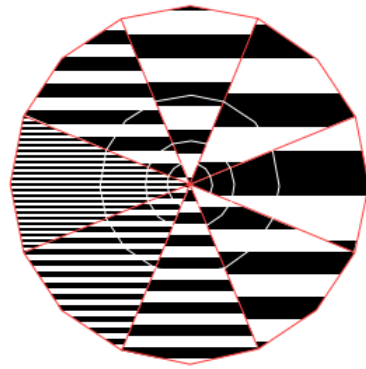


(d) L4: 線の色に割り当て

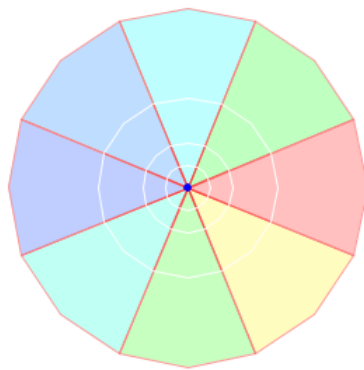
図 4.3.2: 移動量を四分位線に割り当てた様子



(a) S1: 面の濃淡に割り当て



(b) S2: 面のきめに割り当て



(c) S3: 面の色に割り当て

図 4.3.3: 移動量を方位ごとの面に割り当てた様子

## 4.4 アメーバコロニー表現

アメーバ表現では、参照点が地図に指し示す位置の近傍を出発地点とする移動を表現している。それに対して、アメーバ表現1つでは読み取ることのできない広範囲な地点の移動を表現する表現を開発した。これは、アメーバ表現を **small multiples** 技法を用いており、多地点の移動の様子を比較できるものである。アメーバ表現の集団からなる、多地点の移動を表すこの表現を「アメーバコロニー表現」と名づけた（図4.4.1）。

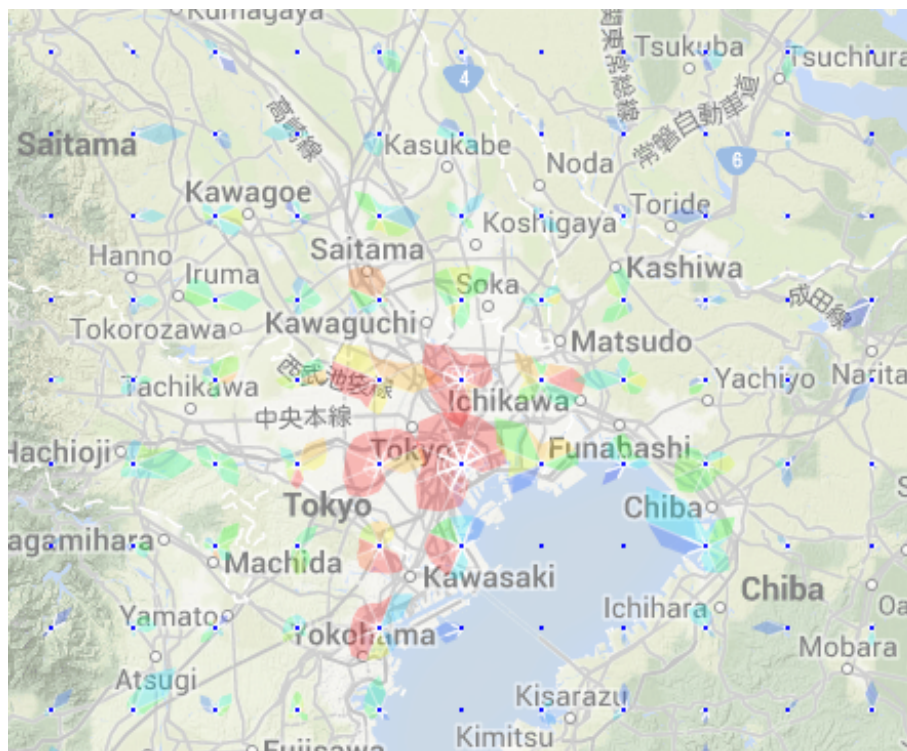


図 4.4.1: アメーバコロニー表現の概観

アメーバコロニー表現は、アメーバ表現を小さくメッシュ状に並べている。これは、地図の縮尺に応じて四分位を描いているアメーバ表現を、単純にメッシュ状に並べると、それらの四分位線同士が重なりあい、情報を読み取れないという問題が生じたからである。

アメーバコロニー表現は、多地点の移動の傾向や特徴的な地点を表現できる。それは、アメーバ表現では移動量を表すために面の濃淡を用いていたが、アメーバコロニー表現内の個々のアメーバ表現は移動量の表現に色相の変化を用いているからである。色相の変化は不透明度の変化に比べて区別がしやすいため、同じ傾向の移動をわかりやすくしたり、特徴的な移動を際立たせることができる。また、アメーバコロニー表現では、同じアメーバ表現を並べているため、各地点の移動の様子を比較可能である。



## 第5章 可視化例

実際に東京都市圏7万2千人の移動データをもとに可視化を行った。

### 5.1 利用したデータ

可視化には1998年10月1日の東京都市圏の7万2千人分の移動データを用いた。このデータでは1分ごとの位置データが24時間分、csv形式にて記録されている。各時刻の位置が1つのファイルに記録され、そのファイル中の1行(レコード)にID、時刻、緯度、経度の情報を持つ。

### 5.2 アメーバ表現の様子

東京都市圏の4つの地点(新宿駅、品川駅、横浜駅、東京スカイツリー)の近傍3kmの地点に午前6時にいた人たちの3時間での移動の様子を調べた。各地点におけるアメーバ表現の様子を図5.2.1に示す。

図5.2.1の4つの図の四分位線を読みとると、都庁や企業のビルが多く立ち並ぶオフィス街<sup>1</sup>である新宿駅、品川駅では、どの方位も距離の大きな移動を含んでいることが読み取れる。それに対して、都心から少し離れた横浜駅や東京スカイツリー周辺では大きな距離の移動は方位によって偏りがあることがわかる。また、移動量をあらゆる面の不透明度を読み取ると、新宿駅、品川駅、横浜駅の3つの地点では南から西にかけての方位で人数が多いことがわかる。東京スカイツリー付近では他の3つの地点とは人数の偏りが異なっており、北よりと東へ向かう人数が多い。

個々の地点の移動について見てみる。新宿駅付近の移動を示す図5.1(a)では、南西と西の方位の面の不透明度がもっとも高いことから、それらの方位へ向かう人が最も多いことがわかる。ついで北よりの方位と南への移動が多い。最も少ないのは南東の方角である。移動距離分布を読みとると、南東を除くすべての方位で、四分位線にあまり偏りが無いことがわかる。もっとも参照点に近い線である第1四分位の線が、東にむかってやや伸びており、東の方向に限ってはやや新宿から離れる移動が多い。また同じく第1四分位の北の方向は、参照点に近く、北の方向は新宿駅付近にとどまっている人が他の方位より多い。新宿駅から最も

<sup>1</sup>東京23区のうち、千代田区、中央区、港区は都心3区と呼ばれ、東京都の中でも際立って官公庁や企業のビルが立ち並ぶ。この定義は東京都や政府による統計調査などで用いられている。さらに都心3区に隣接する新宿区、渋谷区を合わせて都心5区と呼ぶ場合もある

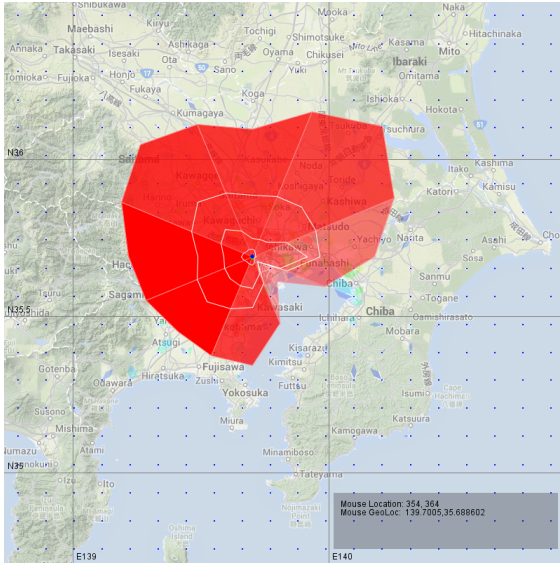
遠くへの移動は、茨城県土浦市付近まで移動している。次に、品川駅付近の移動を図 5.1(b)より読み取る。特徴的な点について述べると、南西と南の方位への移動人数が多くなっていることが、面の不透明度から読み取れる。また、東の方位は面の不透明度より人数が少ないことがわかるものの、四分位線より東へ向かう人のほとんどが遠くへ向かっている事がわかる。続いて都心から少し離れた横浜駅の移動について見てみる。横浜駅付近の移動を表した図 5.1(c)からは、面の不透明度より、南・南東・東と北への移動が多いことがわかる。もっとも特徴的な点は四分位線の偏りであり、もっとも大きな距離の移動は北東方向ではあるものの、最大値以外の四分位を読みとると、南西方向への移動距離が大きいことがわかる。最後に、東京スカイツリー付近の移動を読み取る。図 5.1(d)が示す特徴的な点は、北よりの方位と東への方位への移動人数が、他の方位に比べて大きい点である。面の不透明度を細かく読みとると、最も人数が大きいのは東の方位である。そして、移動人数が多い方位では、移動距離も他の方位に比べて長くなっていることが四分位線から読み取れる。

### 5.3 アメーバコロニー表現の様子

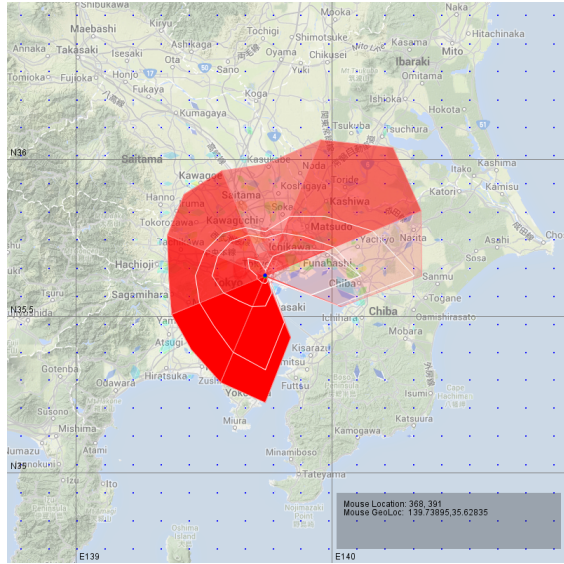
対象データに対して、アメーバコロニー表現を描いたものを図 5.3.1 に示す。

図 5.3.1 に示されたアメーバコロニー表現からは、面の色が赤色で示されている人数の多い移動が都心付近で起きていることがわかる。この移動は横浜やさいたまのそれぞれ一部の方位にも見られる。また、面の色が青色で示されている人数の少ない移動が公開で見られ、それらは都心へ向かう方位で距離が伸びていることがわかる。

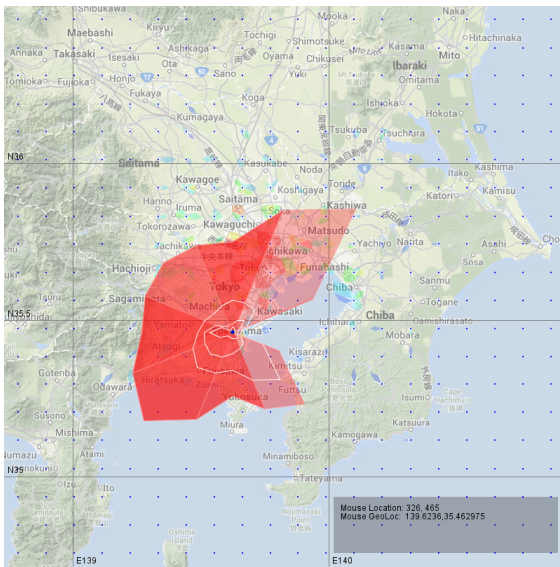
図 2 にアメーバ表現を、図 3 にアメーバコロニー表現を表す。どちらも朝 6時から 9時までの 3 時間の移動を表現している。図 3 では人数の表現に色相を用いた。赤は人数が多いことを、青は少ないことを表す。図 2 からは、新宿駅付近を起点とする北東方向への移動は最大距離が大きいものの、四分位線の多くが起点近傍に密集していることから、ほとんどの人の移動距離が短いことがわかる。図 3 からは、都心付近では人数も多く、四分位線も各方位に均一に広がっていることから、各方位に同じ程度の多くの人移動していることがわかる。



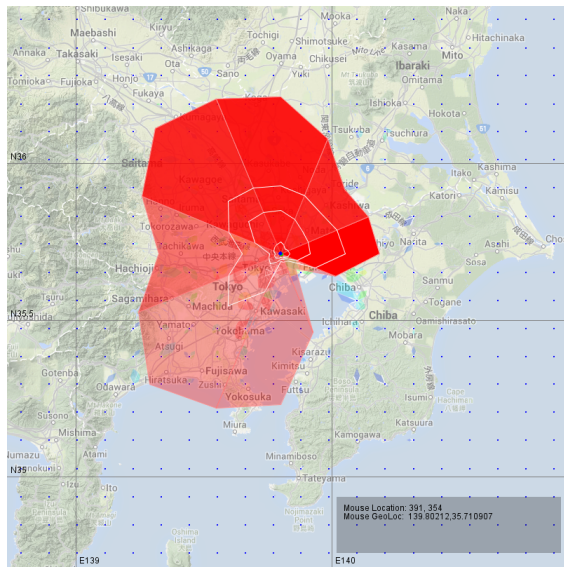
(a) 新宿駅付近



(b) 品川駅付近



(c) 横浜駅付近



(d) 東京スカイツリー付近

図 5.2.1: 東京都圏 7 万 2 千人分の人の移動データの 4 地点におけるアメーバ表現

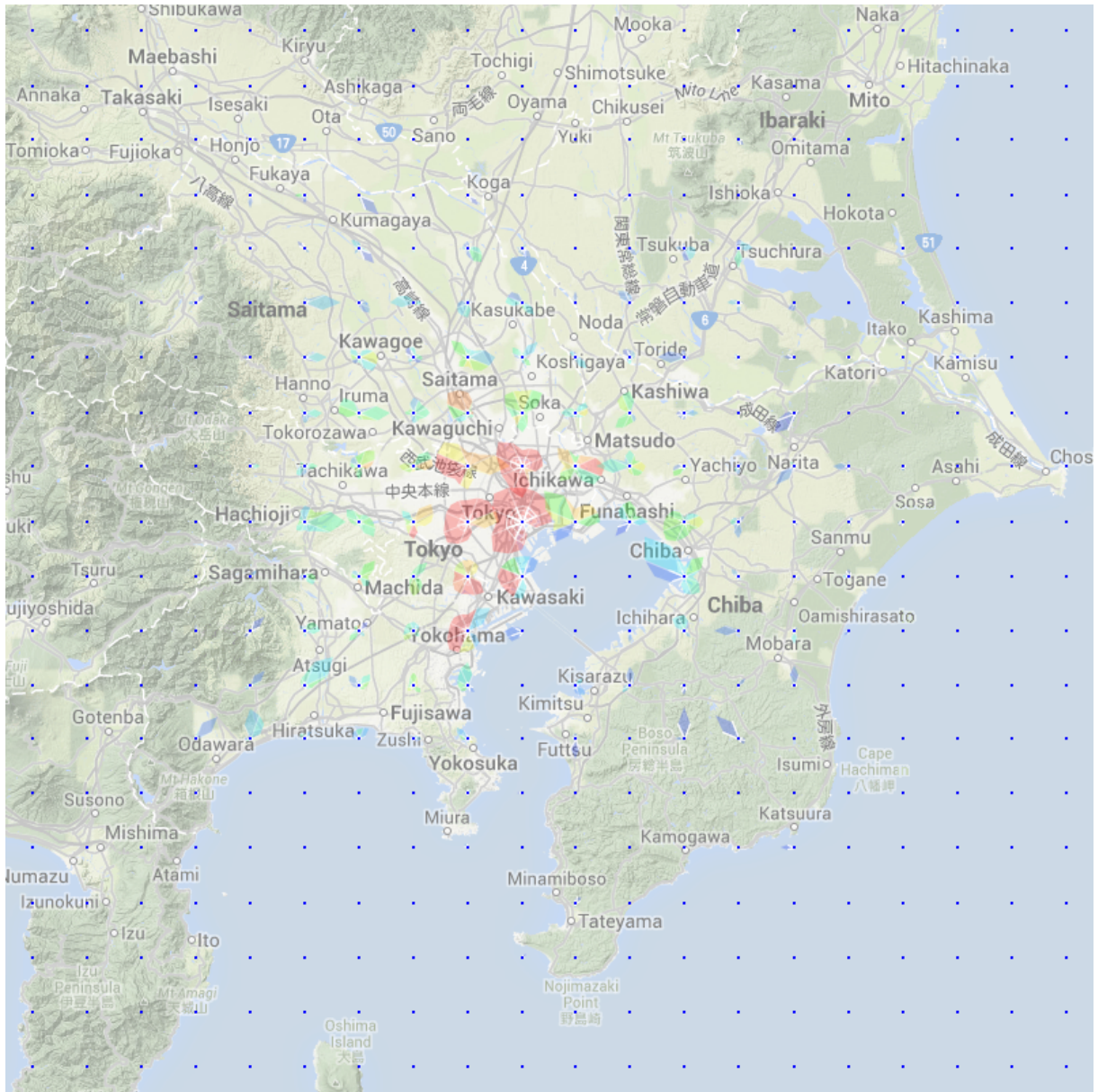


図 5.3.1: 東京都市圏 7 万 2 千人分の人の移動データに対しアメーバコロニー表現を用いた様子

## 第6章 結論

大規模な移動データに基づいて人や物の移動の様子を集約的に可視化する手法を開発した。対象とした大規模な移動データとは、地理空間上の複数の物体に対して、時刻と、その時刻における物体の位置が大量に記録されたデータである。このデータからは物体が地理空間上を移動していく様子を知ることができる。しかしながら、移動の様子を視覚的に理解するために、移動を単純に線によって描画した場合には、線同士の重なりによって情報を読み取ることが困難であった。

本研究では、そうした大規模な移動データを、集約的な表現を用いて可視化することによって、指定された点の近傍を始点とする1つ以上の物体の、移動方向、移動距離、移動量（人数や物量）を視覚的混雑度を抑えて表現する可視化手法を開発した。開発した手法では、移動する物体をその移動の方向によって8つの集団に集約し、それぞれの集団に対し、移動距離分布の四分位と移動量（移動の数量）を求めた。これらの方位ごとの四分位と移動量を表現することによって、複数の移動を集約的に表現している。開発した可視化手法を **small multiples** 手法によって複数個提示することにより、地理空間上の広い地点の様子を一度に表すことも可能にした。

## 謝辞

提案手法の可視化例を作成するにあたり，東京大学空間情報科学研究センター一人の移動プロジェクトのデータを利用いたしました．ここに感謝申し上げます．

本研究に際して，三末和男先生には多大なご指導を賜りました．お忙しい中にもかかわらず，細やかな助言・提案を与えて下さり深く感謝しております．また田中二郎先生，志築文太郎先生，高橋伸先生には，ゼミの場を通して有意義なご意見を賜り，研究をより一層深めることができました．ありがとうございました．

そして，研究室生活において何より身近な存在でありましたインタラクティブプログラミング研究室 NAIS チームのみなさまには，公私ともに大変お世話になりました．論文の添削や進捗の相談といった場面以外でも，様々な形で支えられ，卒業論文を仕上げる事が出来ました．厚くお礼申し上げます．

最後に，私をここまで育ててくださいました家族に心より感謝申し上げます．

## 参考文献

- [1] Yoshihide Sekimoto, Ryosuke Shibasaki, Hiroshi Kanasugi, Tomotaka Usui, and Yasunobu Shimazaki. Pflow: Reconstructing people flow recycling large-scale social survey data. *IEEE Pervasive Computing*, Vol. 10, No. 4, pp. 27–35, 2011.
- [2] David J Field, Anthony Hayes, and Robert F Hess. Contour integration by the human visual system: Evidence for a local “association field”. *Vision research*, Vol. 33, No. 2, pp. 173–193, 1993.
- [3] Greg Turk and David Banks. Image-guided streamline placement. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pp. 453–460. ACM, 1996.
- [4] David H Laidlaw, J Scott Davidson, Timothy S Miller, Marco da Silva, RM Kirby, William H Warren, and Michael Tarr. Quantitative comparative evaluation of 2d vector field visualization methods. In *Proceedings of the conference on Visualization'01*, pp. 143–150. IEEE Computer Society, 2001.
- [5] Waldo R Tobler. Experiments in migration mapping by computer. *The American Cartographer*, Vol. 14, No. 2, pp. 155–163, 1987.
- [6] Gennady Andrienko, Natalia Andrienko, and Stefan Wrobel. Visual analytics tools for analysis of movement data. *ACM SIGKDD Explorations Newsletter*, Vol. 9, No. 2, pp. 38–46, 2007.
- [7] Ben Shneiderman. Extreme visualization: squeezing a billion records into a million pixels. In *Proceedings of the 2008 ACM SIGMOD international conference on Management of data*, pp. 3–12. ACM, 2008.
- [8] Edward R. Tufte. *Envisioning Information*. Graphics Pr, 1990.
- [9] J. Bertin and W.J. Berg. *Semiology of Graphics: Diagrams, Networks, Maps*. Economic & Social Research Institute, 2010.