

# 可視化技術をもちいた交通系ICカードデータの分析\*

## Visualization of Passenger Flow with Smart Card System \*

日下部貴彦\*\*・中島良樹\*\*・朝倉康夫\*\*\*

By Takahiko KUSAKABE \*\*・Yoshiki NAKAJIMA\*\*・Yasuo ASAKURA\*\*\*

### 1. はじめに

鉄道をはじめとした公共交通でICカードによる運賃決済の普及が進んでいる。これら交通系ICカードの決済による乗車履歴データ(以下ICデータと呼ぶ)には、①改札通過時刻を1分単位と詳細に取得できる、②長期にわたる定点観測データが収集できる、入出場記録が完結されなければ継続利用できないため、③データの欠損がないなどの特徴がある。

ICデータを長期に渡る継続的な観測データとして捉えた場合、様々な交通現象が観測されていることが期待でき、利用実態の把握やマーケティングなどに役立つ情報も多く含まれると考えられる。しかし、一般的に長期間のデータは膨大な量となることから、観測期間中に起こったすべての現象を予見して分析を行うことは難しいといえる。このようなことから、ICデータの分析では、大量に取得されたデータから未知の特徴や規則性をとらえる「データマイニング」によるアプローチの研究が積極的に試されるべきといえる。

本研究ではデータマイニングの一手法である可視化技術に着目する。この手法はデータを一定の法則で視覚化した画像から人間の視覚を用いて特徴的な交通現象の発見を行うものである。これまで交通分野では、高速道路に設置されている車両検知器による観測データを可視化する技術が日下部ら<sup>1)</sup>によって研究されている。車両検知器による観測データは、ICデータなどの自動改札機によるデータと同様に定点で長期間に渡って観測されたデータである。このことから、ICデータにおいても車両検知器での方法と同様の可視化技術を用いたデータマイニングによる分析が有効であると考えられる。可視化技術をICデータに応用することで、乗客流の長期的な変動の傾向や特徴的な現象を直感的に把握できることが期待される。

本研究の目的は、ICデータによる改札通過記録の可

\*キーワード: 可視化, ICカード, 旅客流動

\*\*学生員, 工修, 神戸大学大学院工学研究科

(神戸市灘区六甲台町1-1,

TEL078-803-6360, FAX078-803-6360)

\*\*\*正会員, 工博, 神戸大学大学院工学研究科

視化システムを構築することにある。第2章では、ICデータの可視化技術を構築する。第3章で可視化技術を用いた分析例を示し、鉄道の旅客分析における可視化技術の有用性について考察する。

### 2. 交通系ICカードの可視化技術

可視化技術を用いたデータマイニング手法での視覚化に求められる要件は、

- 適切な配色を用いた情報の正確な表示
- 高速かつシームレスな表示

である。一つ目の要件は、観測値の微小な変化をできるだけ正確に捉えるために必要なものである。可視化では観測値等を色によって視覚化し視認することによって現象を捉えるため、表示するデバイスと人間の視覚の双方の特性を考慮した配色が求められる。二つ目の要件は、分析者が大量のデータから様々な条件でデータを表示し、ブラウジングしながらデータから知識を精錬する際に求められるものである。可視化に時間がかかると様々な条件で何度もブラウジングすることが不可能になるためである。本章では、これらの要件を満たす可視化手法を構築する。第1節では、本研究で対象としているデータについて述べる。第2節で可視化の方法について述べ、第3節で構築した可視化システムについて説明する。

#### (1) 対象のデータ

本研究で想定しているICデータの形式を図1に示す。このデータ形式では、レコード毎に一回の鉄道利用が記録されているものである。本研究の方法で必須の項目は「使用年月日、入場駅、入場時刻、出場駅、出場時刻」である。

ICデータ					
使用年月日	入場駅	入場時刻	出場駅	出場時刻	カードID
2007年10月13日	A駅	7:10	C駅	7:23	A25687DK
2007年10月13日	A駅	7:11	C駅	7:23	B68677DS
2007年10月13日	A駅	7:11	B駅	7:18	B67732RR
⋮					
2007年10月13日	B駅	17:57	A駅	18:09	B67732RR
2007年10月13日	C駅	18:00	A駅	18:17	B89751RR
⋮					

図1 ICデータのデータ形式の例

## (2) 可視化の方法

本研究では、乗客流の長期間の変動を捉えるための可視化の方法として、日下部ら<sup>12)</sup>の研究で用いられているTraffic Contour Map (TCM)を用いる。TCMは、定点で継続的に観測された時系列的な交通データを観測値点毎にコンピュータのディスプレイなどの二次元上に表現するための可視化方法である。本研究では、「Day-to-day TCM」と「Time-value TCM」を用いる。はじめに、Day-to-day TCMとTime-value TCMについて説明する。次にTCMで用いる色について述べ、最後にTCM上に表示する観測値について述べる。

### a) Day-to-day TCM

Day-to-day TCMは、図2に示すように縦軸に日付、横軸に時刻をとり、色で改札通過人数などの観測値を示すものである。TCMを用いることによって、横軸方向に1日の中の時刻による変動を、縦軸方向に日毎の変動を表現することができる。

### b) Time-value TCM

Time-value TCMは、図3のように縦軸に観測値、横軸に時刻をとる。縦軸に対応する観測値が観測期間中に観測された頻度を時刻毎に色を用いて表示する。このTCMでは時刻別の観測値の分布を表現することができる。

### c) 色

TCMでは、観測値等を色で表現することから、より正

確に観測値の変動を捉えるためには、人間の視覚に対応した連続的な色の割り当てが必要である。このため本研究では、日下部らの研究で提案<sup>1)</sup>されている方法である、CIE 1976 (L\*,a\*,b\*) 均等知覚色空間<sup>3)</sup>を用いて色の見え方の変化である色差と観測値の変化量が比例するように色を割り当てる方法を用いる。

表1 TCMで表示する観測値

表示方法	対象のトリップ	基準となる地点	基準となる時刻	集計値
1	特定の駅で入場したもの	入場駅	入場時刻	人数
2	特定の駅で出場したもの	出場駅	出場時刻	人数
3	特定の駅で入場し、特定の方面の駅で出場したもの	入場駅	入場時刻	人数
4	特定の駅で出場し、特定の方面の駅で出場したもの	出場駅	出場時刻	人数
5	特定の駅間ODを利用したもの	入場駅	入場時刻	人数
6				平均旅行時間
7		出場駅	出場時刻	人数
8				平均旅行時間

### d) TCMで表示する観測値

TCMは、それぞれの地点で時刻毎の一つの観測値を視覚化するものである。したがって、駅間トリップ毎に観測されたICデータでは、何らかの方法で集計した値を用いることになる。ICデータの各レコードには、入場駅、入場時刻、出場駅、出場時刻が記録されていることから、可視化のために集計する際に基準となる時刻と地点に複数の組み合わせが考えられる。これらのうち最も単純な組み合わせは、それぞれの駅での時刻別の入場者数又は

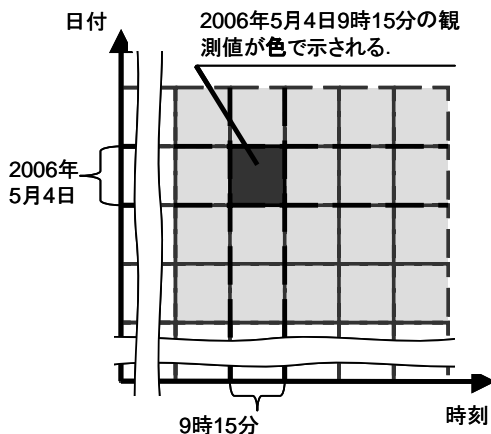


図2 Day-to-day TCM

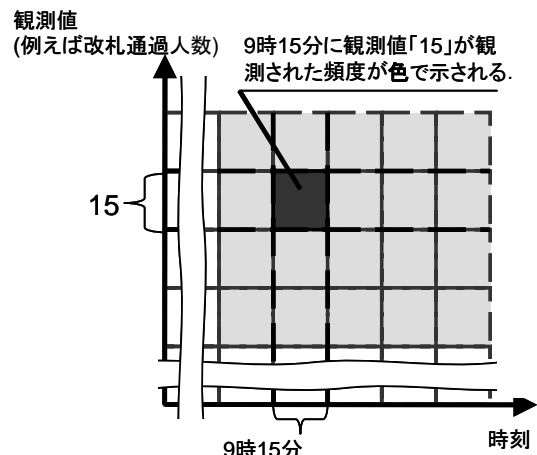


図3 Time-value TCM

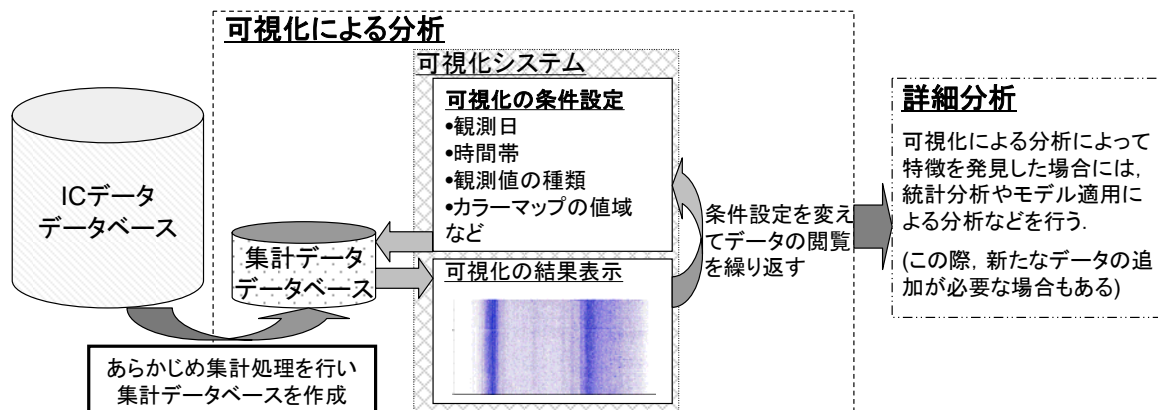


図4 可視化システムと可視化による分析手順の概要

出場者数を集計したものである。これらの他にも、上り下りの方向別や、特定の駅間OD毎に時刻別の入場者、出場者数に集計したものや、これらの入場者及び出場者に対応する平均旅行時間を用いることも考えられる。ICデータを用いたTCMでの表示が想定される主な観測値を

まとめると表1のようになる。

(3) 可視化システム

図4は、可視化システムと可視化による分析手順の概要を示したものである。可視化によるデータマイニングでは、分析者は様々な条件でデータを表示し、何度も

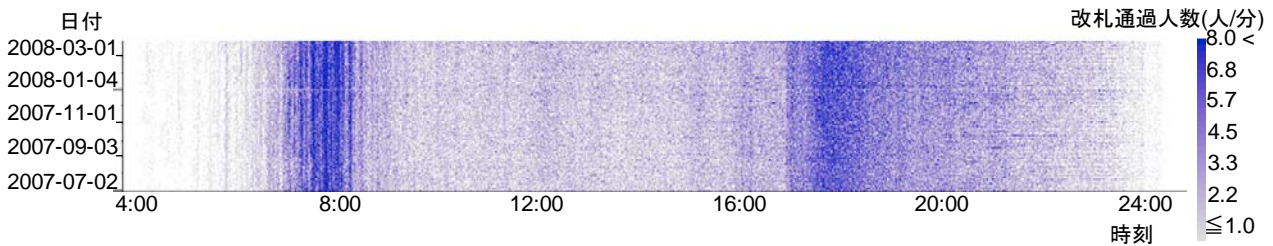


図5 A駅のDay-to-day TCM

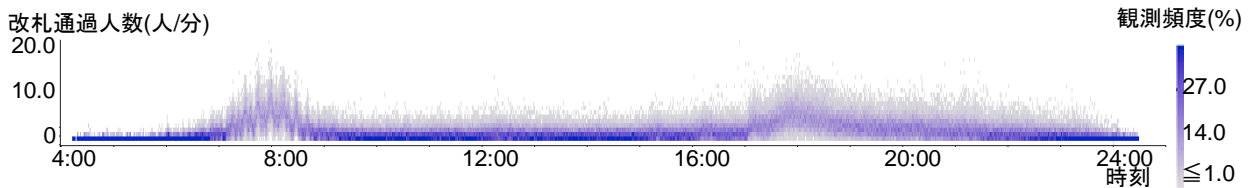


図6 A駅のTime-value TCM

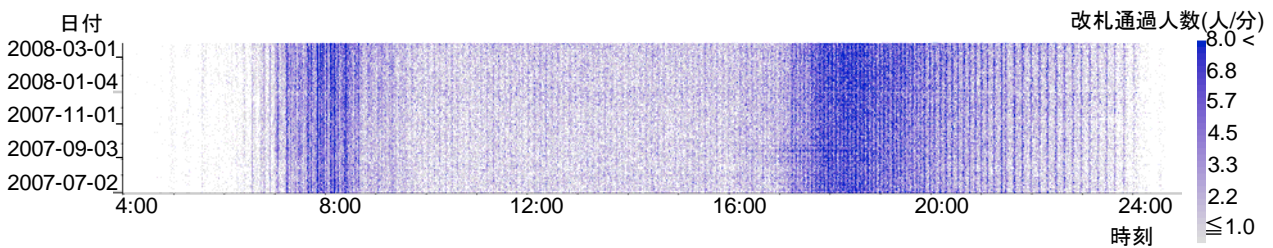


図7 B駅のDay-to-day TCM (他社線乗換駅)

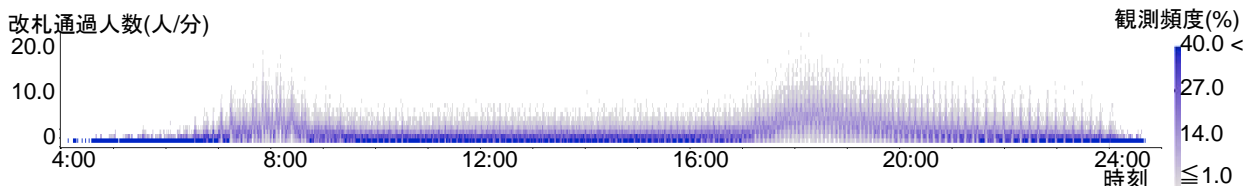


図8 B駅のTime-value TCM (他社線乗換駅)

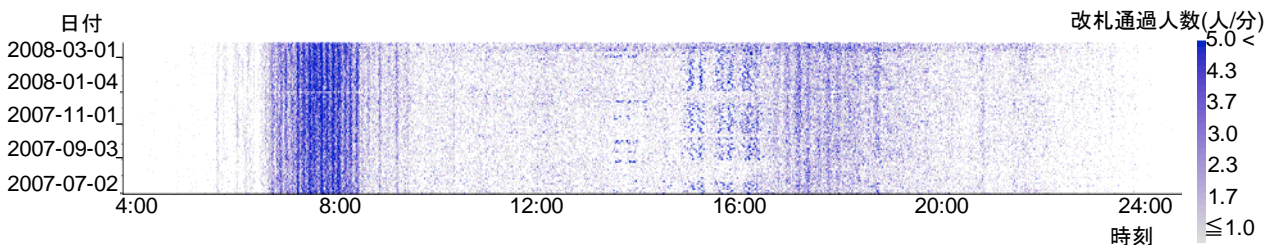


図9 C駅のDay-to-day TCM (学校近隣の駅)

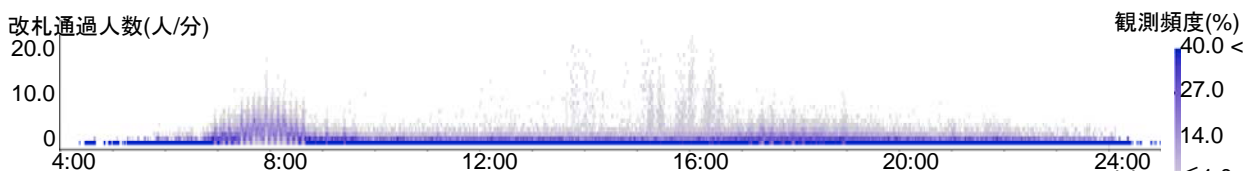


図10 C駅のTime-value TCM (学校近隣の駅)

ブラウジングすることが求められるため、高速な表示が求められる。対象としているICデータから表示の対象となるレコードを検索し、集計を行い、その結果を表示するという手順をすべて可視化の際に行ったうえで、高速な表示という条件を満たすことは、現在一般的に用いられているコンピュータでは困難である。このため本研究の可視化システムでは、あらかじめすべての地点、駅間ODについてICデータを表1に示した観測値に集計した集計データを作成し、可視化の際には集計データを用いて画像に変換する方法を採用し表示の高速化を図った。

可視化による分析の際には、表示の対象とする駅又は駅間OD、対象とする観測日、時間帯、観測値の種類、カラーマップに割当てた観測値の値域を指定してTCMによって表示する。これらの項目の指定を変更するなどして、順次データの閲覧を行い特徴の探索を行う。

### 3. 可視化技術による分析例

本章では、鉄道会社A社のICカード乗車券システムにより取得されたICデータを用いて可視化システムを用いた分析例を示し、可視化技術による分析の適用可能性などを視点とした考察を行う。分析では、可視化システムを用いて各駅の入場者数を可視化し、乗客流の変動に関する特徴の発見を行う。

このとき、Day-to-day TCMとTime-value TCMを併用することで、データを三次元的に捉えることが可能となり、より直感的に状況把握が可能となる。

なお、今回の分析で使用したICデータの仕様は2章で示したものであり、対象期間は2007年7月2日～2008年3月31日の平日とし、平日に限定した理由は、平時の鉄道利用状況を、より確認しやすくするためである。

#### (1) 一般的な利用状況の駅

図5、6は中規模市街地にあるA駅の入場者数を可視化したものである。図によると8時および18時前後に利用者が集中する時間帯が確認できる。これは通退勤を示している。

ここで、通勤の8時前後では等間隔の縦縞が見えるが、この間隔はA駅の優等列車の列車間隔と同じであり、利用者がダイヤにあわせて駅に到着している状況を表している。一方、18時前後ではこの特徴はなく、利用者は退勤時にはダイヤを強く意識していないといえる。

#### (2) 他の交通機関の影響を受ける駅

図7、8は利用者数がA駅とほぼ同じのB駅の入場者数を可視化したものである。図によると通退勤の8時および18時前後に利用者が集中する特徴はA駅と同じであるが、全ての時間帯で細かな等間隔の縦縞が見える点が大きく異なっている。

この間隔はB駅のダイヤではなく、隣接する他社線の

駅のダイヤと整合しており、入場者の多くが隣接駅からの乗り継ぎ流入であることがわかった。

これは、駅の入場者数による比較ではA駅と同様と考えられるB駅が、可視化により利用状況が明確に異なることを発見できた例である。

#### (3) 沿線施設の影響を受ける駅

図9、10は高校や大学が付近にある中規模のC駅の入場者数を可視化したものである。A駅やB駅と異なり、18時前後に利用者が集中する特徴が弱く、15時から17時にかけて間隔の広い縦縞が見られる。

この間隔はC駅のダイヤではなく、学校の終了時限にあわせて利用者が駅にやってくる状況を示している。また、2007年8月、12月にはこの間隔の広い縦縞がない期間があり、夏期、冬期の長期休暇の状況も現れている。

これは、駅周辺の施設により、利用状況の時間変動や季節変動が異なることを発見できた例である。

### 4. おわりに

本研究では、長期間に渡って観測され、蓄積されているICデータの可視化手法を構築し、可視化技術を用いた分析事例を示した。

ICデータの可視化は、精緻なデータ特性の恩恵もあり、各駅の特徴が顕著に現れる結果となった。また、事前に集計データを準備することで、駅別、入出場別などの条件変更をしていく際も、即時に可視化でき全体像の把握や、条件別の比較が円滑に実施できることが確かめられた。

これを活用することで、鉄道沿線のイベントや集客施設、鉄道のダイヤ変更や新駅・新線の開業の影響や効果などを、可視化されたデータから直感的に把握することで、事業者のニーズにより直接的に対応できる分析が可能となる。さらに、駅や鉄道の利用時間帯や利用状況から読み取れる利用者特性などから、改善施策やマーケティングの検討への展開も見込める。

#### 参考文献

- 1) 日下部貴彦, 井料隆雅, 朝倉康夫: 車両検知器データを用いた交通流可視化技術の開発, 交通工学, 43(4), p.59-68, 2008.
- 2) Kusakabe, T., Iryo, T., and Asakura, Y.: Data Mining for Traffic Flow Analysis: Visualization Approach, Proceedings of International Workshop on Traffic data Collection & its Standardization, Barcelona, 2008.
- 3) 日本色彩学会: 「新編色彩科学ハンドブック」, 東京大学出版会, 1989.