

## ISZOTにおける追跡動線クラスタリングを用いたショッパー行動解析 Clustering of intra-camera Trajectories for Behavior Analysis of Shoppers

○越中谷 俊樹† 奈良 博之† 金子 俊一† 李 媛‡ 三好 雅則‡ 藤吉 弘亘\* 渡辺 琴美\*\*

† 北海道大学 ‡ 株式会社日立製作所 日立研究所

\* 中部大学 \*\* 生活協同組合 コープさっぽろ

E-mail: etchuya@ssc.ssi.ist.hokudai.ac.jp

### 1 序論

スーパーやコンビニ、家電量販店などでは、売場に来店する客が何を購入したか、どのような行動をとったかという情報は顧客分析や売場分析、キャンペーンなどを行う上で非常に重要なことである。これらの分析をどれほど正確に行うかで、その後の売りに大きく影響する。

これまでは、POS(Point Of Sale system)によって得た購入商品の情報から分析を行っていた。しかし、POSでは購入した商品名や時間はわかるが、どのような経緯、順序で購入したかは不明であった。迷った結果購入しなかった商品、何も買わずに出て行った人の行動などはわからなかったため、マーケティングへの応用も限られた範囲でしか行うことができなかった。これに対して、RFIDなどの発信機を付けたカートを利用して店内行動を取得して解析を行う研究 [1][2] が行われている。しかし、現場で利用する場合、導入のために多くのコストがかかってしまうという問題がある。

一方、社会的な安全、安心の意識の増加や防犯意識の向上、カメラの価格低下により、現在では多くの店で監視カメラが設置されている。しかし、取得された画像は実際にはほとんど使用されておらず、犯罪が起きた場合の事後確認のために使用されたりと、撮影した大部分の情報はそのまま捨てられてしまっているというのが現状である。行動解析を行う際も、これまでは撮影された映像を人がすべてチェックしたり、店舗で客の後ろを追跡して解析していた。しかし、人が手作業で行うには解析対象が多すぎて作業が困難ある。

この問題に対し、行動目的が類似しているショッパー(顧客)同士をまとめることが出来れば、解析対象は格段に減り、これまで困難であったショッパーの店内での行動に注目した行動解析を容易に行うことが出来るようになる。販売店側の視点から見ると、類似するショッパーとは行動目的が同じ人であるといえる。行動目的が類似しているショッパーは、同じような経路を移動することが考えられる。このことから、経路が同じショッパーを集めることが出来れば、行動目的が類似するショッパー



図 1: ショッパーの動線

同士をまとめることが出来る。そこで本論文では、監視カメラの映像からショッパーの動線を取得し、類似するショッパーをクラスタリングする手法を提案する。

商店に来店する顧客のことを“ショッパー”と呼び、ショッパーの移動した経路を“動線”と呼ぶ。画像処理を用いてショッパーの動線を取得し、大量の動線を用いてショッパーの行動解析を行う本システムを ISZOP(Image based sensing system for analyzing Shopper's Zone Path) と呼ぶ。ISZOPには画像からの人物検出、追跡、類似度算出、クラスタリングの4つのパート(図2)がある。本書では類似度算出、クラスタリングについて詳しく述べる。

### 2 商品を基準とした解析空間“Zone”

ショッパーの行動目的は商品を購入することであり、効果的に解析するには商品とのインタラクションを考える必要がある。そのため、商品に対してショッパーが関与できる領域を定め、これを用いてショッパーの行動を解析するのが最も効果的である。そこで我々は商品に対し、ショッパーが関与できる領域を“Zone”と呼び、これを用いて動線を解析する。本章では Zone の定義と Zone を用いた動線の記述方法を図3を用いて説明する。

まず、ショッパーの足元位置  $\mathbf{p}$  と動線  $\mathbf{P}$  を以下のように定義する。

$$\mathbf{p}_i = [ X_i \ Y_i \ t_i ]^T \quad (1)$$

$$\mathbf{P} = \{ \mathbf{p}_i \}_{i=1,2,\dots,I} \quad (2)$$

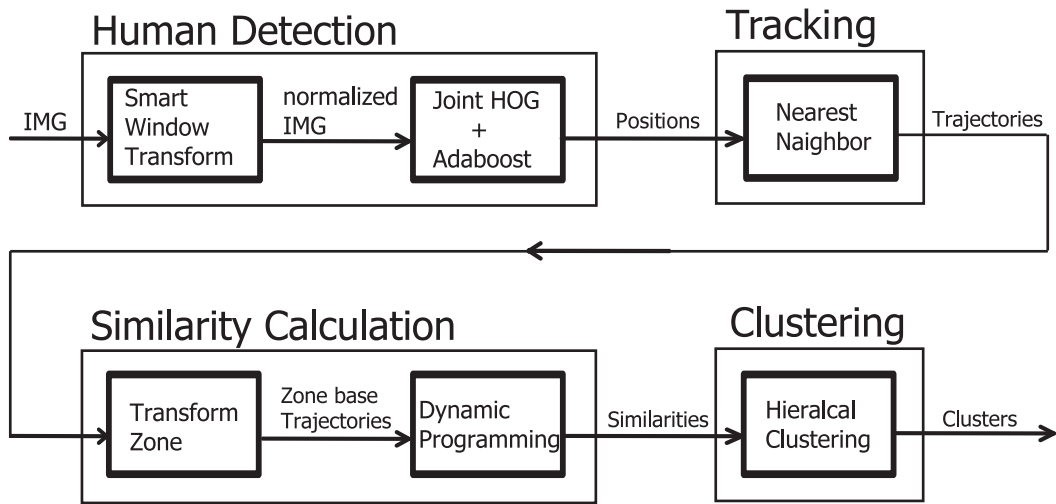


図 2: Zone ベース時系列データクラスタリング法の処理手順

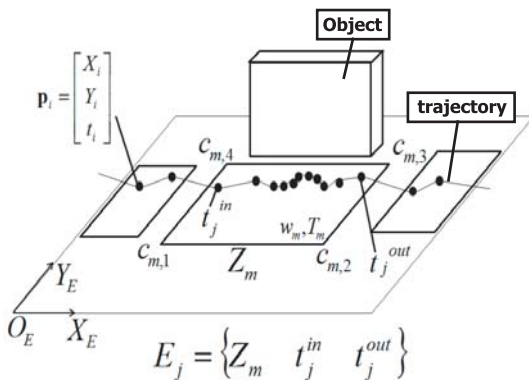


図 3: zone の定義図

そこで我々は実空間上に商品を基準にした領域を Zone  $Z$  と呼び、以下のように定義する。

$$Z = \{c_1, c_2, c_3, c_4, w, T\} \quad (3)$$

$$c_n = [X_n \ Y_n]^T \quad (4)$$

ここで,  $c_n$  は Zone の各頂点,  $w$  は重み,  $T$  は後に述べる滞留を判別するためのしきい値である. ショッパーが移動した Zone  $Z_m$ , Zone に入った時間  $t^{in}$ , 出た時間  $t^{out}$  の 3 つをまとめてエンタリー  $E$  と呼ぶ. また, エンタリーの時系列で表した動線をエンタリーベース動線  $P^E$  と呼び、以下のように定義する。

$$E = \{Z_m, t^{in}, t^{out} \mid 1 \leq m \leq M\} \quad (5)$$

$$P^E = \{E_j\}_{j=1,2,\dots,J} \quad (6)$$

ショッパーは目的の商品や興味のある商品の Zone で購入, 検討のために立ち止まるため, 同じ場所に立ち止まったショッパーは同じ商品に興味を持っているといえる. 本論文では立ち止まることを滞留と呼ぶ. このため, ショッパーの行動解析において立ち止まることは大変重要な要素であると考えられる. ショッパーが Zone 内にいる時間を滞留時間  $t^{out} - t^{in}$  とし, 滞留時間のしきい値  $T$  を満たすエンタリーをストップエンタリー

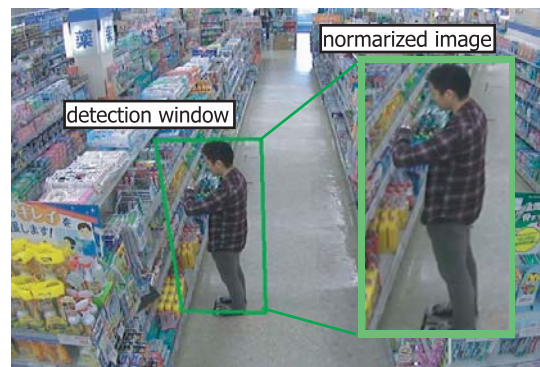


図 4: SWT を用いた検出ウィンドウ作成と正規化画像

$E^S$  と呼ぶ. また, ストップエンタリーの時系列で表した動線をストップエンタリーベース動線  $P^{SE}$  と呼び、以下のように定義する。

$$E^S = \{E_j \mid t_j^{out} - t_j^{in} \geq T_j\} \quad (7)$$

$$P^{SE} = \{E_1, \{E_k^S\}_{k=1,2,\dots,K}, E_J\} \quad (8)$$

### 3 Zone ベース動線クラスタリング法

本章ではカメラ映像から人物検出技術を用いて動線を取得する手法と, エンタリーベース動線の類似度を算出する手法, およびクラスタリング手法について述べる. 処理の流れを図 2 に示す. 監視カメラの映像から人物検出を行い, 取得した足元位置をトラッキングして図 1 のような動線を作成する. 作成した動線を, Zone を用いて変換してエンタリーベース動線とし, それぞれの類似度を算出する. 算出した類似度を基に階層的クラスタリングによりクラスタリングする.

#### 3.1 SWT(Smart Window Transform) を用いた人物検出

カメラ画像から人物検出を行う手法は様々提案されている [3]. 従来の人物検出手法は, 画像上で検出ウィンドウをラスタスキャンし, そのウィンドウ内部の特徴量から人かどうかを判断していた. これらの人物検出手法は画面内で映り方があまり変化しない場合には



図 5: 手動で検出したショッパーの動線



図 7: Zone を用いて動線を圧縮した Zone ベース動線

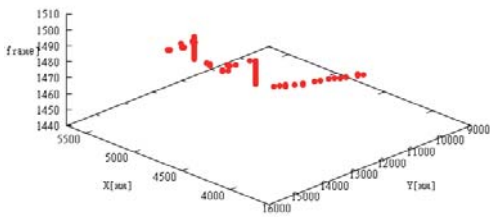


図 6: 時空間動線表記によるショッパー動線

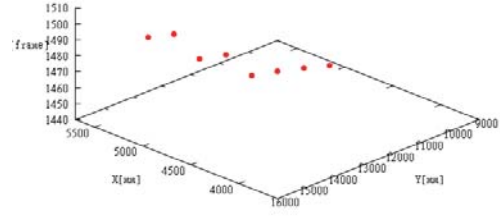


図 8: 時空間動線表記によるショッパー動線

有効であるが、監視映像内では、カメラの設置条件により歪みや傾きが発生する。また、対象までの距離が近い場合が多いため、画面の手前と奥で映り方に大きな差が生じてしまう。このため、ラスタスキャンを用いた従来の検出ウィンドウの作成方法では検出精度が低くなるという問題がある。

そこで我々は検出ウィンドウの作成に、カメラの設置条件や画面内での映り方の変動にロバストな手法として李らが提案する SWT(Smart Window Transform)[4]を用いる。SWTでは実空間上に定義した円筒に、円筒の中心を通りカメラに正対する切断面を定義する。カメラパラメータを用いてこの切断面を画面上に変換することで、カメラの設置条件や映り方の違いにロバストな検出ウィンドウを作成することが出来る。図4はSWTで作成した検出ウィンドウと正規化した画像である。

人物検出には藤吉らが提案する照明変動にロバストな Joint-HOG 特徴量と Adaboost[5]を用いる。Joint-HOG 特徴は2つ HOG 特徴の共起に注目した特徴量で、人のように対称的な形状の物体に対して高い検出精度が得られる。

SWTは実空間上で検出ウィンドウを作成するため、人を検出した際に、足元位置も容易に求めることが出来る。人物検出を行うと人の周辺に複数の足元位置候補が検出されるので、MeanShift[6]により、一人の人の一つの足元位置を定義する。

実空間において高さ方向を Z 軸、床平面を X, Y 軸とし、実空間で人が存在する X, Y 平面上の座標値を人物の足元位置とする。一人の人物の足元位置を時系

列に並べたものを時系列位置データと呼び、時系列位置データによって動線は定義される。

### 3.2 Zone ベース動線類似度算出法

Zone を用いてエントリーベース動線  $P^{SE}$  の類似度を算出する手法を“Zone ベース動線類似度算出法”と呼ぶ。エントリーベース動線はショッパーによってデータ数が異なるため、単純な1対1比較で計算することが出来ない。そこで、本論文では動的計画法(Dynamic Programming[7]以下, DP)を用いて時系列 Zone データの最適なマッチングを行い、類似度を算出する。DPは、音声認識の分野でも用いられている手法で、データ長の違う2つの一次元データに対して長さの違いを考慮しながら、データどうしの最適なマッチングを高速に導き出せる手法である。Zone ベース動線類似度算出法は以下の式である。各変数は次の通り。  $r(i, j)$  はデータの挿入の有無,  $d_{tra}$  は距離の差,  $d_{ins}$  はデータを挿入する際のコスト,  $\alpha, \beta$  はこれらの調整係数,  $d(i, j)$  は  $i$  と  $j$  の差にひとつ前の最適解を足したもの,  $s_{p,q}$  は動線  $P$  と  $Q$  の類似度,  $d_{p,q}$  は相違度,  $d_{max}$  は DP の算出結果の最大値である。

$$d_{tra}(i, j) = |E_i(Z_n(c_1)) - E_j(Z_m(c_1))|^\alpha \quad (9)$$

$$d_{ins}(i, j) = r(i, j) \cdot \beta \quad (10)$$

$$d(i, j) = d_{tra}(i, j) + d_{ins}(i, j) + \min d(i-1, j-1) \quad (11)$$

$$d_{p,q} = d(I, J) \quad (12)$$

$$s_{p,q} = \frac{d_{max} - d_{p,q}}{d_{max}} \cdot 100 \quad (13)$$



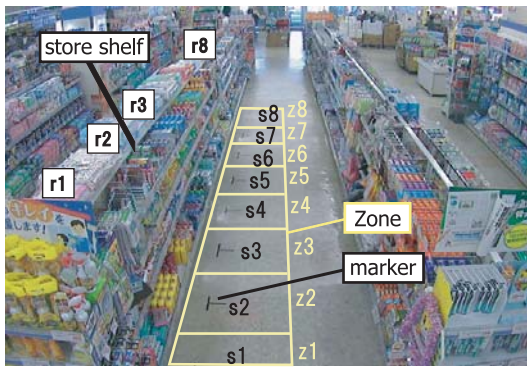


図 9: マーカーと Zone の設定図

### 3.3 階層的クラスタリング

本研究では Zone ベース動線類似度算出法で算出した動線の類似度を基に、階層的クラスタリングを行う。階層的クラスタリングは閾値の値によってクラスタの数を決められるという特徴がある。解析には階層的クラスタリングの最も基本的な手法である ward 法 [8] を用いた。ウォード法ではクラスタ  $C_i$  と  $C_j$  が融合したクラスタ  $C_{ij}$  とクラスタ  $C_k$  の距離  $d(C_{ij}, C_k)$  を以下の式を用いて求める。  $n$  はクラスタに含まれるデータ数である。

$$d(C_{ij}, C_k) = \frac{n_k + n_i}{n_k + n_i + n_j} d(C_i, C_k) + \frac{n_k + n_j}{n_k + n_i + n_j} d(C_j, C_k) - \frac{n_k}{n_k + n_i + n_j} d(C_i, C_j) \quad (14)$$

## 4 SWT による位置推定誤差の検証

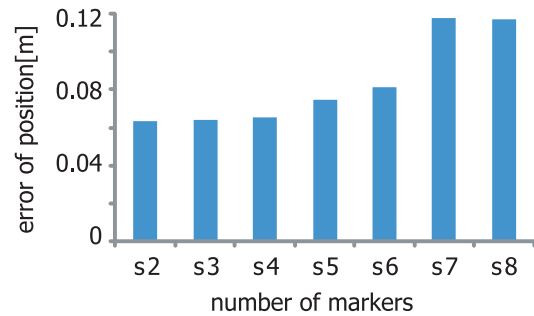
### 4.1 検証内容

実験は北海道を中心に 106 店舗を展開している生活協同組合コープさっぽろの協力で行った。実験店舗はコープさっぽろの余市店内のドラッグ売り場である。

SWT による人物位置検出では検出ウィンドウを作成するためにカメラパラメータを必要とする。カメラパラメータはキャリブレーションにより取得するが、このキャリブレーションによって誤差が発生する。店内に設置した 5 台のカメラのうちの 1 台のカメラのビュー内に設置した 8 つのマーカーの位置を、実際に計測した真値と SWT により計測した値を比較した。マーカーの設置条件は図 9 の通りである。

### 4.2 結果と考察

比較結果の図 10 をみると、誤差は最大でも 0.12m である。人の大きさは上から見ると、直径約 0.50m の円である。よって、誤差は最大でも人の大きさの約 24% 程度であるため、SWT の影響による位置推定誤差は無視できるほど小さいといえる。画面奥に行くほど誤差が大きくなるのは画像内での解像度が低いので、画像内で SWT の位置を合わせる精度が悪くなったためであると考えられる。



(s1 is too near to be seen in the view of camera)  
図 10: 各マーカーの位置ごとに測定した SWT の影響による位置推定誤差

## 5 Zone ベース動線クラスタリング法によるクラスタリング実験

### 5.1 実験内容

実店舗データを用いて Zone ベース動線類似度算出法によるクラスタリングを行った。実験は 4 章と同じ cam4 の映像を用いた。ショッパーに見立てた被験者に指定の滞留位置で棚に向かって 3 秒間立ち止ってもらった。

人物位置検出、動線作成は手動で行った。ただし、画面上の座標からカメラパラメータを使って実空間上の座標に変換するため、4 章で調べたキャリブレーションによる誤差が生じる。

本実験において Zone ベース動線類似度算出法は滞留位置および滞留回数が同じことに重点を置いて類似度を算出するようにパラメータを設定した。

### 5.2 結果と考察

Zone ベース時系列データクラスタリング法の結果を図 11 に示す。ショッパーのクラスタが統合される間隔から、Cluster 1,2,3 に分類した。類似度が低い (dissimilarity が高い)、最も大きなクラスタを作る Clustering 1 をみると、滞留回数が同じショッパー同士が同じクラスになっていることが分かる。また、類似度が高いクラスタリングの Clustering 3 をみると、同じ位置で滞留したショッパーが同じクラスになっている。Cluster 1 と 3 の間の Cluster 2 では滞留回数が 1 回、2 回の中でも滞留した位置の関係が近いショッパーが同じクラスになっている。

このことから、滞留位置が類似する動線をクラスタリング出来た。また、滞留位置は多少違うが滞留回数と同じクラスになるクラスタリングを一回の処理で階層的に行うことが出来た。今回使用した動線は手動で作成した動線ではあったが、キャリブレーションによる誤差を含むなかで、Zone ベース動線クラスタリング法により効果的にクラスタリング出来た。

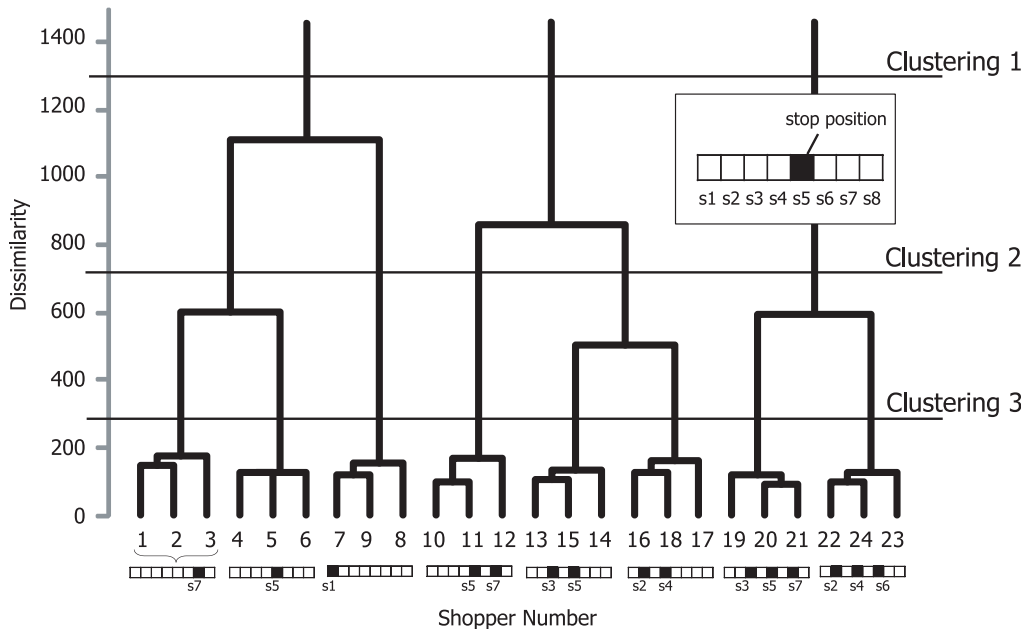


図 11: Zone ベース動線クラスタリング法による実店舗の動線 24 人のクラスタリング結果

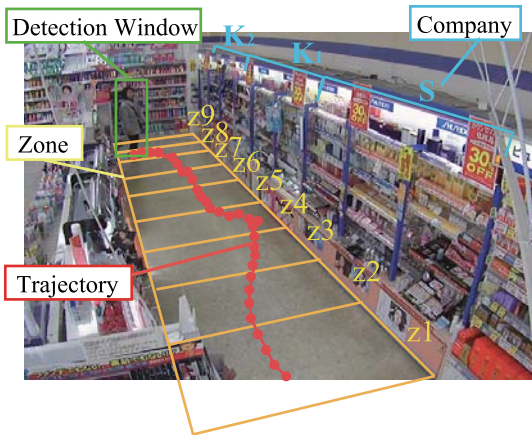


図 12: 化粧品コーナーのメーカーと Zone の設定

表 1: 滞留者数と購入数の関係

	9/8		10/27	
	販売数 (11h)	滞留者数 (2h)	販売数 (11h)	滞留者数 (5h)
A 社	7	18	29	24
B 社	5	27	21	30
C 社	2	33	23	33

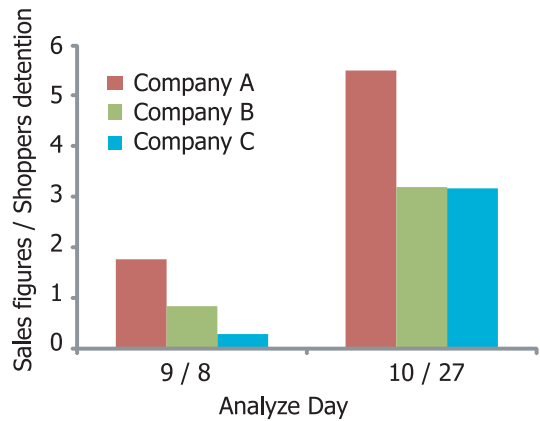


図 13: 滞留購入率

## 6 モデルパスクラスタリング

Zone ベース動線クラスタリング法を用いて滞留人数のカウントを行う。ZONE  $Z_n$  に滞留した人数をカウントするとき、まず  $Z_n$  に滞留した動線をモデルパス  $P_m^{SE}$  として与える。モデルパス  $P_m^{SE}$  を含めて階層的クラスタリングを行うと、モデルパス  $P_m^{SE}$  を含むクラスを作成することができる。モデルパス  $P_m^{SE}$  を含むクラスを  $P_m^{SE}$  のモデルクラス  $MC_m$  とする。モデルクラス  $MC_m$  に含まれるショッパーの人数を数えることでZONE  $Z_n$  に滞留した人数をカウントすることができる。モデルパスを用いた階層的クラスタリング法を“モデルパスクラスタリング”とよぶ。

## 7 滞留人数カウントとPOSデータを用いた解析

モデルパスクラスタリングを用いて滞留人数のカウントを行った。実験を行ったのはコープさっぽろ余市店の化粧品コーナーである。図 12 のように9つのZONEを設けた。各ZONEに対し、1つのモデルパスを与え

てクラスタリングを行った。与えたモデルパスは滞留人数をカウントするZONE  $Z_n$  に一度だけ滞留したものとした。化粧品コーナーには3つのメーカーの商品があり、それぞれのメーカーが3つずつZONEをもっている。滞留人数のカウント結果は各ZONEでの滞留人数をメーカーごとにまとめたものである。また、今回の実験ではコープさっぽろの協力で、実際の売上データであるPOSデータを提供していただいた。滞留人数のカウント結果は表1のようになった。滞留人数に占める購入者の割合を“滞留購入率”と呼び、以下のよう

に定義する.

$$\text{滞留購入率} = \frac{(\text{一時間あたり}) \text{購入数}}{(\text{一時間あたり}) \text{滞留者数}} \cdot 100$$

9月8日と10月27日のそれぞれでメーカー別の滞留購入率を算出した結果を図13に示す.

これを見ると, いずれの日もメーカーAの滞留購入率が最も高いことがわかる. 滞留購入率が高いということは, 立ち止まった人数が同じでも, 実際に商品を購入するショッパーが他のメーカーよりも多いということである.

9月8日と10月27日を見ると, 明らかに10月27日の方が滞留購入率が高いことがわかった. 給料日が毎月25日であることを考えるとこの結果は妥当だと言える. よって, 10月27日の方がショッパーの購買意欲が高く, より商品を購入しやすいことがわかった.

ショッパーの店内の行動を取得することで, これまでわからなかった立ち止まった人数に占める購入数を算出することができた. これにより, 商品PRがうまくいっている商品と改善すべき商品を見つけることができた. また, ショッパーの購買意欲の高さを調べることもできた.

## 8 結論

本論文では画像センシング技術を用いたショッパーの行動解析システムとしてISZOPを提案した. その中で, ショッパーの動線をクラスタリングする手法としてZoneベース動線クラスタリング法を説明した. Zoneベース動線クラスタリング法の主な技術としてSWTを用いた人物検出, Zoneベース動線類似度算出法, モデルパスクラスタリングの3つの技術を記載した. 実験ではコープさっぽろ余市店に8つのマーカを設置し, 実測とSWTで取得した推定位置の差からSWTの位置推定誤差を求めた. また, 設置したカメラを用いて24人分のデモ動線を取得し, Zoneベース動線クラスタリング法によるクラスタリングを行い, 滞留回数, 滞留位置の違いによってクラスタリング出来ることを示した. POSデータとの連動によるショッパー解析を行うために, 化粧品コーナーにおいて, 実際のショッパーのデータを用いて実験を行った. 手動で取得した50人の動線と, 滞留人数をカウントするために作成した9つのモデルパスを用いてモデルパスクラスタリングを行った. この実験で, POSデータとクラスタリング結果から滞留購入率を算出し, 商品PRの改善, ショッパーの購買意欲の推定が行えることを示した.

## 9 謝辞

本論文において実験データの提供および, 顧客行動解析の貴重なご意見をいただいた生活協同組合コープさっぽろ 米田氏, 購買動向調査研究所 中野氏, ピップ株式会社 野川氏に感謝する.

## 参考文献

- [1] 服部可奈子, 豊嶋伊知郎, 板倉豊和, 吉田琢史, 折原良平. 時系列の位置データを用いた人間行動分析顧客行動分析, 工場ムダ作業, 不審者発見に対する適用例. 2007. The 21st Annual Conference of the Japanese Society for Artificial Intelligence, 2007.
- [2] 矢田勝俊. スーパーマーケットにおける顧客動線分析と文字列解析. Vol. 56, No. 2, 9 2008. 特集「データマイニングと統計数理」[研究ノート].
- [3] K. Levi and Y. Weiss. Learning object detection from a small number of examples: the importance of good features. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, Vol. 2, pp. II-53. IEEE, 2004.
- [4] 李媛, 伊藤誠也, 三好雅則, 金子俊一, 藤吉弘亘. Smart Window Transform を用いた人物検出. 精密工学会秋季大会学術講演会講演論文集, pp. 920-921, 2011.
- [5] T. Mitsui and H. Fujiyoshi. Object detection by joint features based on two-stage boosting. In *Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on*, pp. 1169-1176. IEEE, 2009.
- [6] Y. Cheng. Mean shift, mode seeking, and clustering. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, Vol. 17, No. 8, pp. 790-799, 1995.
- [7] Bertsekas. *Dynamic programming and optimal control*, Vol. 1. Athena Scientific Belmont, MA, 1995.
- [8] J.M. Lattin, J.D. Carroll, and P.E. Green. *Analyzing multivariate data*. Thomson Brooks/Cole Pacific Grove, CA, 2003.