

H-034

画像処理とベイジアンネットワークを用いた店舗内における顧客の全身像からの購買行動認識の基礎的検討

Fundamental Study of Recognizing Customers' Entire Bodies' Behaviors in Shops Using Image Processing and Bayesian Network

Jing Yan[†] 堀 隆之[‡] 大谷 淳[†]Jing Yan[†] Takayuki Hori[‡] Jun Ohya[†]

1. はじめに

従来のPOSシステムが収集できない顧客が買い物する過程での行動を認識する研究の一環として、本研究では店舗内にある監視カメラのビデオ画像から、顧客の購買行動の中で店舗のビジネスにとって重要度の高い「迷い」の有無と、「購入」の有無を認識することを目的とする。具体的には、画像処理とベイジアンネットワークを用いて認識する方法の検討を行う。

従来の筆者らの研究[2]では、人物の追跡や行動特徴の抽出は手動で行われていた。これに対し、本稿では、画像処理を用いて自動的にこれらの追跡や行動特徴を抽出する方法を検討する。さらに、ベイジアンネットワークに、抽出された特徴等を入力することにより、行動の認識率を評価する。

2. 研究手法

本稿では、「顧客の追跡」、「行動特徴量の抽出」、「迷いと購買の有無の認識」を検討する。

「顧客の追跡」では、Deformable Part Model[1]を用いて、顧客の追跡処理を行う方法を検討する。「行動特徴量の抽出」の段階では、購買領域の座標系を構築し、顧客の移動軌跡から行動特徴量を抽出する方法を提案する。

2.1 顧客の追跡

顧客が買い物する時の移動情報を捉えるために、顧客ごとにトラッキングする必要がある。

まず、ビデオ画像から5フレーム/秒の頻度で画像を抽出し、Felzenszwalbらが提案したDeformable Part Model (DPM)[1]により人体の検出処理を行う。DPMとは、物体の全体像とパーツの組み合わせでモデル化し(DPMで学習

した人体のモデルを図1に示す)、全体とパーツの形状およびパーツの位置ずれ・変形に対するコストで評価を行い、線形SVMにより識別処理をする。ビデオ画像から抽出したあるフレームの人体検出結果を図2に示す。

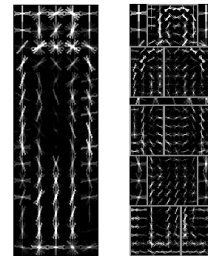


図 1 DPM の人体モデル (特徴量: HOG 特徴量)



図 2 顧客の全身像と各パーツの検出結果

次に、検出した人物の全身像(図2の赤い Bounding box で示す)から color histogram を求め、前のフレームで検出した人体の color histogram とマッチングし、一致する場合、同じ人物だと判定し、移動軌跡を更新する。追跡結果を図3に示す。顧客の存在する位置を把握するため、全身像の中心座標ではなく、両足の中心座標を検出し、顧客の軌跡を求める。

2.2 追跡軌跡を用いた行動特徴量の抽出

図4(a)の赤い線が示す領域をドリンクコーナーの購買領域と定義し、この領域に入る顧客を実験対象とする。この購買領域の頂点座標を求めるには、床に貼ってあるタイルを基準に、エッジ検出を行い、タイルのエッジから購買領域の4頂点の座標を算出する(図4(b))。そして、前述の

[†] 早稲田大学大学院国際情報通信研究科
Graduate School of Global Information and Telecommunication
Studies, Waseda University

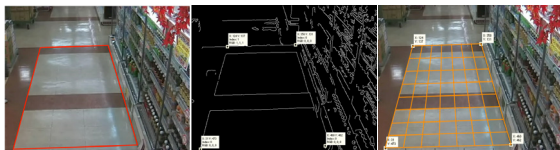
[‡] 総合警備保障株式会社セキュリティ科学研究所
SOHGO SECURITY SERVICES CO.,LTD.

購買領域をさらに図4(c)のように小領域に区切り、いずれかの小領域に入った顧客の両足の中心座標を記録する。

筆者らは既に、顧客の買い物の行動から、4つの認識カテゴリー（表1に示す(A)～(D)）と因果関係を持つ20個の行動特徴量を提案し、手動で特徴量を抽出した[2]。本研究では、顧客の追跡結果と画像処理を用いて自動的に前述の20個の中での8個の特徴量（時間範囲、停止時間、停止時間割合、同じ棚での複数回停止、棚前の停止回数、棚に遠い時の停止回数、棚との最小距離、棚への接近量）を抽出する。例えば「停止時間」という特徴量を取得する場合、顧客の位置座標が同じ小領域（図4(c)）に5フレーム連続して留まると、停止していると判断し、停止時間を計測する。



図3 運動軌跡追跡結果



(a) 購買領域 (b) エッジ検出 (c) 区切り結果

図4 購買領域の確定

2.3 Bayesian Networkによる認識処理

本論文では、Bayonetというソフトウェアを利用し、顧客の購買行動の認識を行う。筆者らは既に、認識対象（迷う、買う）を、他の全てのノードの親ノードに設定し、186個の学習データ（ビデオ）を使い、正三角形モデルを構築した（図5）。本研究では、自動的に抽出した8個の特徴量と従来手動で抽出した12個の特徴量を組み合わせたデータを用いて正三角形モデルを構築し、4つのカテゴリーを認識する。

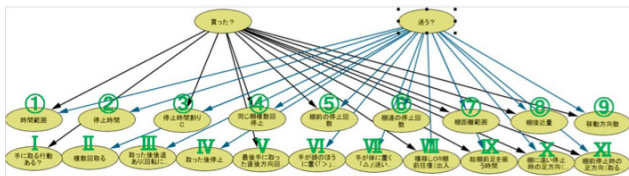


図5 学習データで構築された正三角形モデル

3. 実験評価

本稿では、36個のテストデータ（ビデオ）から抽出された特徴量をベイジアンネットワークに入力し、認識実験を行った。この36個のテストデータの中で、(A)が12個、(B)が7個、(C)が11個、(D)が6個である。認識結果を表1に示す。

本研究で利用した正三角形モデルを構築する時、「買った?」ノードと「迷う?」ノードと明らかに因果関係が強いノードだけではなく、それ以外のノードも因果関係があると判断され、より多くのノードの因果関係により、特に行動特徴がはっきりしていない(D)の認識率も83.33%に達し、全体的に良好な認識結果が得られた。表1のように、「迷った」、「買った」は高い認識率である。

表1 認識精度表

		迷わず購入した(A)	迷いながら購入した(B)	迷わず購入しなかった(C)	迷ったが購入しなかった(D)	迷った?	買った?
正三角形モデル	認識精度	100.00%	85.71%	90.91%	83.33%	94.44%	97.22%
	誤りの例	なし	A(1個)	D(1個)	B(1個)		

4. まとめ

本論文では Deformable Part Model を用いた人物の追跡法を提案し、画像処理技術により自動的に行動特徴量を抽出した。抽出された特徴量をベイジアンネットワークに入力することにより、行動パターンの認識処理を行い、提案手法の有効性を確認した。今後は「商品を手にする」等行動特徴量の抽出や、被験者の数を増やす必要がある。さらに、複数人物が現れた場合の追跡時のオクルージョンの問題を解決し[3]、信頼性の高い認識法の実現を目指すことも重要な課題である。

参考文献

- [1] P. Felzenszwalb, et al., "Object Detection with Discriminatively Trained Part Based Models," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 32, No. 9, Sep. 2010
- [2] 鄭曜, 大谷淳, "ベイジアンネットワークを用いた店舗内における消費者の行動認識に関する研究," FIT2012, 第3分冊, H-046, p221-222, Sep. 2012
- [3] Guang Shu, et al., "Part-based Multiple-Person Tracking with Partial Occlusion Handling," CVPR 2012, Providence, RI, June 16-21, 2012.