

複数の全方位カメラによる人物動線計測システム

Customer Trajectory Detection System Using Multiple Omnidirectional Cameras

窪田 進

■ KUBOTA Susumu

丸山 昌之

■ MARUYAMA Masayuki

伊久美 智則

■ IKUMI Tomonori

高島 政実

■ TAKAHATA Masami

店舗内で顧客の行動を計測したいという要望が増加している。しかし、多数の人物が存在する環境下では人物相互の遮へいが生じ、1台のカメラで人物を安定して追跡することは難しい。

東芝と東芝テック(株)は、複数の全方位カメラを用いた人物動線計測システムを開発した。すべての領域を3方向から観測することで遮へいの問題を軽減し、背景差分、視体積交差法、及びパターン認識などの画像処理技術を互いの短所を補完し合うよう組み合わせ、多数の人物が入退出する実環境下でも安定した検出・追跡精度を実現した。

Demand is increasing for a system that is capable of recognizing customer behavior in retail stores. However, customers are often occluded by other customers in crowded situations, which causes difficulties in detecting and tracking people in a sequence of images.

Toshiba and Toshiba TEC Corporation have developed a customer trajectory detection system using multiple omnidirectional cameras. A group of omnidirectional cameras are located so that each area is observed from three different viewpoints in order to alleviate the occlusion problem. The system uses background subtraction, voxel carving, and pattern recognition to achieve stable results.

1 まえがき

立入禁止エリアへの侵入者や異常行動者を検知することで監視業務を省力化したり、店舗内での顧客の行動状況を計測し、売上げや商品の陳列を決める棚割りのデータと連携してマーケティングに役だてたりするなど、日常の様々な分野で、人の動きを計測したいというニーズが高まってきている。人の移動する道筋を動線の形で抽出する動線計測技術として、RFID (Radio Frequency Identification) などの無線タグを用いたものが既に実用化されているが、計測の対象にタグを持ってもらう必要があり、適用範囲が限定されてしまう。

一方、無線ではなく画像を用いる方式は、計測対象に無線タグなどの特別な装備を持たせる必要がない、無線方式と比較して高い位置精度が実現できる、動線とともに映像も同時に記録されるので、何が起こっていたのかを後から映像で確認できるなどの利点があり、大きな期待が寄せられている。しかし、計測の信頼性の点でまだ実用レベルに至っていない。

東芝と東芝テック(株)は、POS (Point of Sales: 販売時点情報管理) システムから得られる売上データと詳細な動線データを連携させ、これまで売上データだけでは把握できなかった店舗内での顧客の行動を可視化し分析するために、複数の全方位カメラを用いた人物動線計測システムを開発した。

ここでは、この人物動線計測システムの概要と、そこで用いられている画像処理技術について述べる。

2 全方位画像による人物動線計測システムの概要

現在のシステムは、1台のパソコン(PC)に最大9台のカメラを接続し、640×480画素のカラー画像を15フレーム/sで録画できる。各カメラには魚眼レンズ又は双曲面ミラーが装着され、360°の全方位画像が得られる。計測領域内のすべての場所が3方向から観測できるように、視野を重複させながらカメラを分散配置する。カメラ間の距離は5m程度である。すべてのカメラの撮影タイミングは同期しており、撮影された画像はHDD(磁気ディスク装置)に記録される。

入店から退店までの顧客の動きを1本の動線として抽出する必要があるが、これを完全に自動で行うことは、残念ながら現在の技術水準では難しい。そこで、画像処理で抽出された動線データの誤りを人手で修正して、完全な1本の動線を作製することを目指している。現在、リアルタイムでの動線計測処理はまだ実現できておらず、記録映像を別途PC上で処理して動線データを抽出している。

3 システムに用いられている画像処理技術

画像処理は背景差分処理、視体積交差法、及びパターン認識の三つの部分から成る。まず、背景差分処理で画像上の変化領域を検出し、次に、各画像上の変化領域のシルエットを空間に投影し、その交差領域を求める視体積交差法で人物の候補領域を絞り込む。最後に、パターン認識で候補領域に人が存在するか否かを判定する。検出された人物は次フレーム以

降, Kalman Filter^(注1)によって次々に位置と速度を推定しながらトラッキングされる。

3.1 背景差分処理による移動物体検出

背景差分処理とは、あらかじめ背景となる画像を持っており、入力画像をこの背景画像と比較し、差の大きい領域を移動物体として抽出する技術である。簡単な差分としきい値処理で移動物体が検出できるので、処理が高速であるという利点がある。ただし、長時間観測を行う場合、照明条件の変化などで背景も変化するので、実際に背景差分処理を用いる場合は、背景画像を適宜更新する必要がある。また、背景を適宜更新すると、立読みやレジ待ちで長時間静止している人物が背景に溶け込んでしまうという問題があり、これに対しても対策する必要がある。

東芝と東芝テック(株)は、背景と差分処理のしきい値の更新に Adaptive Median Filterを用い、人手によるパラメータ調整の不要な背景差分処理を実現した。静止している人物が背景に溶け込む問題は、背景差分処理単体で解決しようとする非常に難しい。しかし、このシステムでは、後段の処理で人物と判定された領域を背景更新処理にフィードバックし、検出された人物領域では背景を更新しないということで比較的容易に解決できる。

3.2 視体積交差法による位置計測

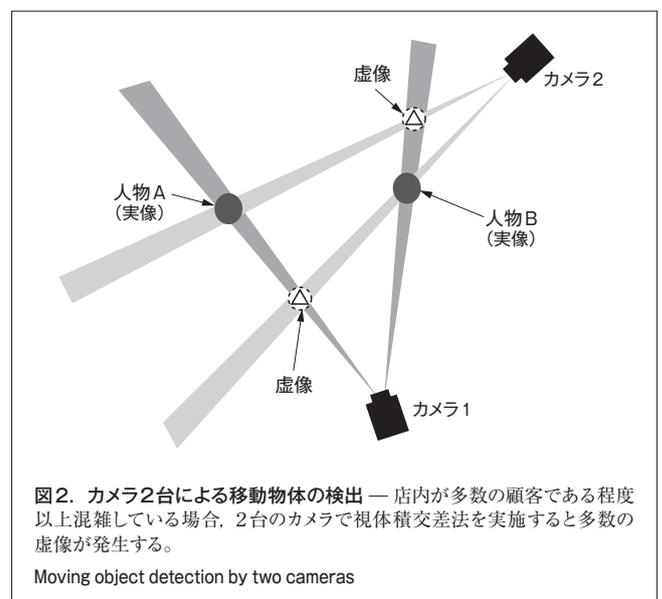
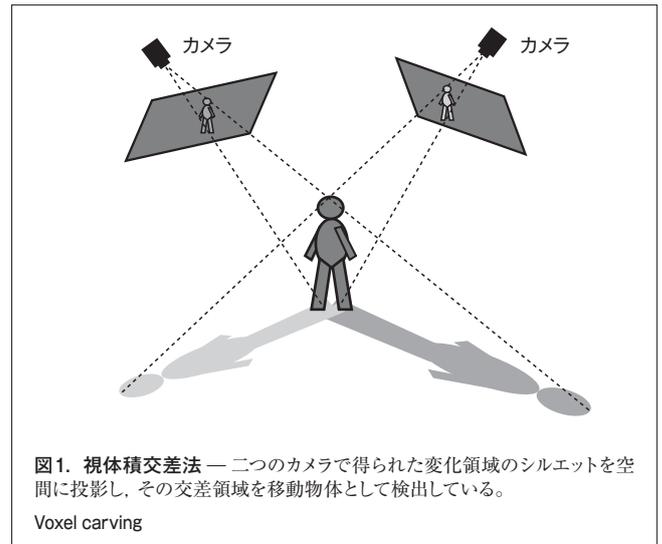
画像を用いた位置計測技術として、2台のカメラを用いるステレオ視がよく知られている。一般的なステレオ視は、2枚の画像の間の対応関係をブロックマッチング^(注2)などで求め(対応点探索)、得られた対応関係とカメラの位置から三角測量で対象の位置を求める技術である。しかし、対応点探索を精度よく行うには、画像間で対象の見え方が極端に変化しないようにするためカメラ間隔をあまり大きくとれず、また同様の理由から、ひずみの大きな広角レンズを用いることもできないため、計測範囲がどうしても狭くなってしまふ。このため、コンビニエンスストア程度の規模の店舗でも、全域をカバーするためには非常に多くのカメラが必要になってあまり現実的ではない。

このシステムは約5m間隔で配置された全方位カメラを用いるため、対象の見え方の変化が大きく、対応点探索を行うことは難しい。そこで、対応点探索の不要な計測技術である視体積交差法を用いて、位置計測を行うことにした。視体積交差法とは、図1のように各カメラで得られた変化領域のシルエットを空間に投影し、その交差領域を移動物体として検出する技術である。

視体積交差法は、移動物体のシルエットさえ得られていれ

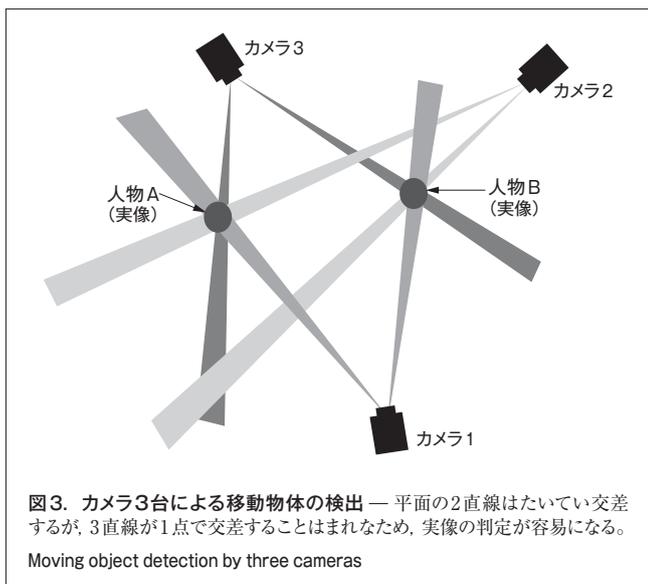
(注1) 1960年にR.E.Kalman氏によって提案された確率過程に基づいたフィルタリングの一つで、ノイズを含む観測データから計測対象の位置や速度の推定を行う手法。

(注2) 画像間の類似性を評価するために、比較する画像から領域を切り出し、その領域に対する輝度差の総和、輝度差の2乗和、及び正規化相互相関を用い、画像間の変位を画素単位で求める技術。



ば単純な投影処理で位置の計測が行える。しかし、あらかじめ画像間の細かい対応関係を求めないため、例えば図2のようにカメラ1から見た人物Aのシルエットとカメラ2から見た人物Bのシルエットが交差して虚像が生じることがある。カメラが2台の場合、視野内にN人の人がいると最大でN×Nの像が得られるが、そのうちN×(N-1)個は虚像である。つまり、2台のカメラで視体積交差法を行うと、店内がある程度以上混雑した時点で多数の虚像が発生する。視体積交差法ではその画像が虚像か実像かを判定することができないので、その判定は後段のパターン認識処理で行うことになるが、虚像が実像に比べてあまりに多いとパターン認識処理の負荷が大きくなり、実用的な性能を達成することが困難になる。

平面上の2直線はたいてい交差するが、3直線が1点で交差することはまれなので、図3のように3台目のカメラを追加する



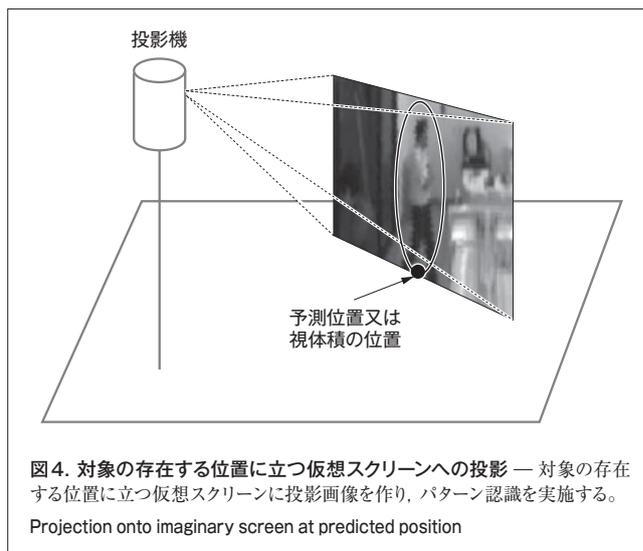
ことで、このパターン認識処理の負荷は大幅に軽減される。このシステムでは、計測エリアのすべての位置が3台のカメラでカバーされるようにカメラを分散配置し、すべての位置で3方向からシルエットを投影して視体積を求める。これで混雑時でも虚像が大量に発生することはなくなる。

3.3 パターン認識による人物の頭部検出

視体積交差法では、原理上実像か虚像かを判定できないので、ある程度虚像が発生することは避けられない。また、実像であっても自動ドアや台車の上の荷物など、検出対象でない物体であることもあるので、いずれにせよ何らかの手段で検出対象かどうかを判定する必要がある。今検出対象が人であるため、画像から人かどうかを判定するが、店舗内では棚などによる遮へいで全身が見えないことが多い。そこで比較的遮へいされにくい頭部に着目し、パターン認識による頭部検出を行うことにした。

パターン認識で頭部検出を行う場合、あらかじめ頭部の様々なパターンを用いて学習を行い、識別器のパラメータを求めておく必要がある。視体積交差法又は前のフレームまでのトラッキング結果に基づく予測で、対象の位置はおおむねわかっているので、まず図4のように、対象の存在する位置に立つ仮想スクリーンへの投影画像を作る。この投影画像上では対象のスケールがほぼ一定で、変形もあまり大きくないので、パターン認識は比較的簡単である。

残念ながら、現在の技術水準では100%の精度で人を認識できず、ある程度の見落としや誤検出は避けられない。したがって、認識誤りが生じることを前提としたうえで信頼性の高い判定処理を行う枠組みを作る必要がある。このシステムでは1か所を3台のカメラで観測するので、図5のような3方向から得られる三つの投影画像に対して認識処理を行うことができ、判定の信頼性を向上できる。



視体積交差法で検出される移動物体の候補位置において3方向から見た投影画像を作り、それぞれの画像上でパターン認識により頭部を検出する。

パターン認識処理で頭部が検出されると、カメラの光学中心と検出された頭部を結ぶ直線、つまりカメラから頭部を見る視線が得られる。複数のカメラで同一の頭部が検出された場合、各カメラからの視線は空間中の頭部位置において交わる。逆に各カメラで検出された頭部がそれぞれ異なる場合は、3次元空間上で2直線が交差することはまれなので、視線は交差しない。これを拘束条件として用いることで、異なる方向から見た投影画像上で検出された頭部の対応関係と、3次元空間上での位置を求めることができる。空間上で対応付かない頭部の組は一方又は両方が誤検出されたものと判断し、二つ以上の方向から見た頭部が空間上で対応付けられた場合に初めて人物と判定されるものとする。三つの投影画像のうちいずれか二つで正しく頭部が検出されればよいので、各投影画像上の多少の検出漏れは許容され、誤検出の多くは前述の拘束条件で排除される。

拘束条件を満たす頭部の組が見つかった後、その3次元位置を初期位置として次フレーム以降で追跡処理を行う。

3.4 Kalman Filterを用いた追跡処理

前のフレームで検出された人物と現在のフレームで検出された人物の対応付けを行い、Kalman Filterで平面上の位置と速度を次々に推定し、人物を追跡する。パターン認識による頭部位置の計測精度は、カメラとの位置関係にもよるがおおむね10 cm未満なので、頭部の検出ができていれば前後のフレーム間での対応付けは容易であり、混雑時にも別人と取り違えることはほとんどない。ただし、毎フレーム確実に頭部が検出されるとは限らないので、その場合は、予測位置の近傍で検出される視体積の位置を観測位置とする。いったん人物として検出されれば、現在のフレームで改めて視体積が実像か虚像かを判定する必要はない。パターン認識による頭部検出に続けて失敗しても、視体積をトラッキングすることで、とぎれない動線が得られる。

4 実店舗における実証実験

これまでに、営業中のコンビニエンスストアにおいて3回の実証実験を行った。このシステムで検出された人物と動線の例を図6と図7に示す。

現時点での性能は、追跡処理に失敗して動線がとぎれる頻度が、通常時で動線1本につき約4分に1回、昼時の混雑時で2分に1回程度である。顧客の平均滞在時間は4分弱だったので、人手で修正作業を行う場合、1本ずつ目視で動線を確認しながら1、2回の編集作業が必要になる。この性能では、人手による修正を前提にすればなんとか実用に耐えられるが、全自動で動線データを生成するためにはまだまだもの足りない

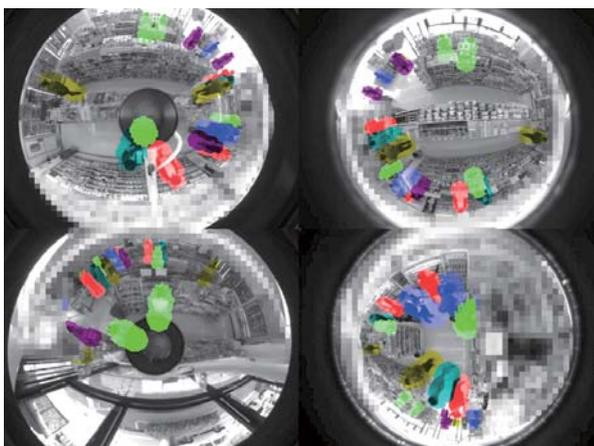


図6. 実店舗での人物の検出例 — 営業中のコンビニエンスストアで行った実証実験での人物の検出例である。検出された人物を色つきの円筒で表示している。

Customer detection in actual store

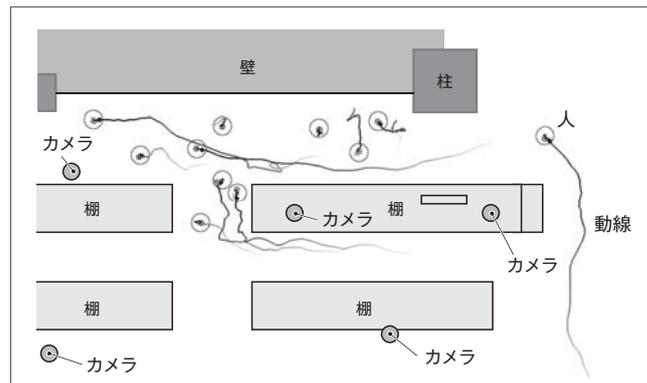


図7. 検出された人物と動線の例 — 追跡に失敗して動線がとぎれる場合もあるが、そのときは人手で修正する。

Example of detected customers and trajectories

レベルであり、今後いっそうの向上を図る必要がある。

5 あとがき

複数の全方位カメラを用いた人物動線計測システムと、そこで用いられている画像処理技術について述べた。これらの画像処理技術は、単独での精度はそれほど高くないが、互いの長所を組み合わせ、短所を補い合うことで、システム全体としての精度を向上できる。

現在は、コンビニエンスストア規模程度の屋内での計測を対象とした実証実験の段階にある。そこで得られたデータと知見を基に、今後は規模がより大きな店舗や屋外での実用化を目指し、性能の向上に取り組んでいく。



窪田 進 KUBOTA Susumu

研究開発センター マルチメディアラボラトリー研究主務。
画像処理及びパターン認識の研究・開発に従事。
Multimedia Lab.



丸山 昌之 MARUYAMA Masayuki

研究開発センター マルチメディアラボラトリー研究主務。
画像認識の研究・開発に従事。
Multimedia Lab.



伊久美 智則 IKUMI Tomonori

東芝テック(株) 技術本部 コア技術開発センター専門主査。
動線計測及び動線分析技術の研究・開発に従事。
Toshiba TEC Corp.



高島 政実 TAKAHATA Masami

東芝テック(株) 技術本部 コア技術開発センター。
動線計測及び動線分析技術の研究・開発に従事。
Toshiba TEC Corp.