

地図を用いた人物検出システムの連携法に関する研究

森敦[†] 和田俊和[†]

本報告では、一台の広角カメラで撮影した画像上の人物位置を検出することができるシステムを複数台連携させ、広域の人物検出・追跡システムを構築する方法について述べる。単一画像からの人物検出法は、学習もしくはあらかじめ設計された人物モデルと撮影画像との照合計算によって画像上の人物位置を検出するものが多い。このような人物検出システムの主な用途は、仮想ゲートを通る人数のカウントであり、歪んだ画像上の人物位置を求めることが基本機能である。このようなシステムを複数連携させるためには、システム間で同一人物の判定をしなければならない。本報告では、各システムの画像上で検出された人物位置をシーンの地図（見取り図）上の位置に写像し、複数のシステムで検出された人物の近接性をこの地図上で評価し、同一人物の判定をする方法を提案する。このための写像は非線型写像になり、本研究では高速性を考慮して線型回帰木を用いる。しかし、人物モデルのマッチングは通常頭部に着目して行われるため、身長が判らない限り、画像上の人物の足下位置は不正確になる。この問題に対して、本手法では同一人物の判定と同時に地図上の人物位置のずれから、個々の人物の身長を推定するというソリューションを与える。推定された身長は、人物属性として使えるだけでなく、シーン中の人物位置の補正に用いることができる。

Map based Cooperation for Multiple Pedestrian Detection Systems

Atsushi MORI[†] Toshikazu WADA[†]

This report presents a cooperation method for multiple pedestrian detection systems, each of which has a single camera with wide view angle, for extending their observation area. Most single-viewpoint human detection algorithms perform matching between pedestrian models and captured images. Since the major task of the target system is people counting, the output of the system is rather simple. We assume that only the pedestrian positions on the image plane are available. For the cooperation of such systems, we propose map-based cooperation scheme, where the detected positions by multiple systems are transformed by nonlinear regression onto a single floor map, and the locations detected by different systems are merged based on the proximity. In our research, we employ linear regression tree for the nonlinear regression due to its short computation time. However, single camera people detection system cannot produce accurate floor position without the height information of each pedestrian. This is because the half upper body of the human model has strong matching weight, and hence, the head position is well estimated on the image plane but the foot position is not. Our method computes the correspondences of the locations of the same pedestrian, and the pedestrian's height is estimated simultaneously based on the discrepancy of their locations on the map.

1. はじめに

画像を用いた人物検出システムは、店舗の客数カウントなどの用途で広く用いられている。特に、デパートやチェーン店などでは来客パターンの分析の需要が高く、積極的に用いられる傾向がある。しかし、カメラで観測できる範囲には限界があり、カメラ一台のシステムで広い店舗内全ての客の移動を把握することには限界がある。より広角のカメラで一つの店舗をカバーする視野を持たせるような商品開発も行われているが、より広い範囲を観測するためには複数台のカメラを協調動作させる必要が生じる。

本報告では、地図を用いて人物検出システムを複数連携させ、広域の人物追跡システムを構築し、同一人物の対応付け問題について取り上げる。この追跡システムは、単眼の検出システムを用いるため安価に構築できる。また、この人物検出システムはカメラ画像に人物モデルを当てはめることでその位置を検出するものであり、人物の身長によって検出位置にずれが生じるという性質を持つ。

本手法では、複数のカメラで検出された人物位置を、広域人物追跡を行う領域である単一の地図座標系に変換する。

ここで、カメラ画像にはレンズの歪みがあるため、カメラ座標から地図座標への変換は非線型写像になる。人物追跡システムには高速性が求められるため、線型回帰木を用いてこの変換を行う。地図座標への変換はレンズ歪みを補正することに相当し、検出位置間の距離計算がしやすくなるので人物追跡に有利である。

本手法の人物追跡は、直前のフレームと比較して同一人物に同一の ID を与えることで実現される。この計算を同じカメラで検出された人物ごとに行い、カメラ間で同一人物に同一の ID を対応付ければ連携が可能になる。ここで、カメラ間での ID 対応付けの際、カメラの共通視野に相当する地図上の領域に限定して同一人物の判定を行うことで、連携に不必要な計算を省くことができる。

カメラ座標を地図座標に変換すると、前述の検出位置のずれは、地図上におけるカメラ位置と実際の人物位置を結ぶ直線上に発生する。そのため、このずれを考慮してカメラ間で ID を対応付ける。この対応付けができると、検出位置のずれを計算できるため、副次的にはあるが人物の身長を推定することができる。人物の身長が判れば検出位置のずれを補正することができるため、安定したカメラ間連携が可能になる。

以下、2 章では関連研究、3 章では地図を用いたカメラ

[†] 和歌山大学
Wakayama University

間連携と人物の身長推定の方法について述べる．4章では身長推定について実験を行い，5章にてまとめる．

2. 関連研究

かつて分散協調視覚システム[1]や，米国のVSAM[2]プロジェクトなどで，複数のカメラを分散配置したシステムで現実の空間で起きる様々な変化をモニタリングする研究がなされてきた．これらの研究を通じて，個々の人物検出・追跡の研究は進んだが，複数カメラを用いた広域監視システムは未だに実用化されていない．

この主な理由としては，実用環境では振動や風などで，カメラのキャリブレーションに変化が起こるため長期間の安定運用が行えないこと，より少ないカメラ台数で広い範囲のモニタリングを行うというニーズに十分応えてこなかったこと，などが挙げられる．

さて，複数カメラの協調動作の内容は，①共通視野内での処理（主に対象IDの対応付けと三次元復元）と，②単一視野内での処理（単一画像上での検出と追跡）の二つに分けられる．特別な例としては③共通視野を持たないシステム間で追跡対象を対応づける機能を持つものもある．①のみの機能を持つものは，撮影範囲に対してカメラ台数が増えるため，コストが高くなる．②のみであると連携ができないため，③も必要になる．つまり，①+②か，②+③の組み合わせでしか協調は行えない．①+②の手法にはParticle Filterを用いて人物を追跡する手法[3]が挙げられる．これは，画像から主に人間の直立方向を検出する手法である．同じくParticle Filter (CONDENSATION)を用いて，尤度の統合法をペアワイズな積の和にすることにより，オクルージョンに対して頑健な追跡を実現した手法[4]などもある．②+③の例を挙げると，共通視野の有無が混在する手法[5]などがある．③の例としては，交差点の地図情報と対象の属性と観測時刻を用いて，ストリームデータとして表現される対象を対応付ける手法[6]がある．

本研究では①+②の機能を持ったシステムを構築する．このような連携手法を考える上での問題点として，1)カメラキャリブレーションの困難さ[7]，2)単眼カメラシステムでの人物位置の不正確さがある．前者は，カメラ間の距離が近い場合はカメラの位置関係の推定誤差が大きく，カメラ間の距離が遠い場合は共通視野が少なくなることに起因している．後者は，単眼カメラの場合，人物モデルを利用するため，後述の人物の立ち位置の推定精度が低くなることを指している．本報告では，主に後者の問題について検討する．

3. 地図上での人物対応付けと身長の推定

本手法では，単眼カメラによる，モデル当てはめ型の人物検出システムを利用する．以降ではこの人物検出システムをターゲットシステムと呼ぶ．ターゲットシステムからは，

カメラ画像上での人物検出位置を得ることができる．本手法では，複数のシステムで検出した人物位置を，同一の地図座標にマッピングする．そして，これらの地図座標を用いて人物を追跡する．また，カメラ間で連携して人物追跡することにより，人物追跡を広域なものに拡張する．また，人物検出位置の誤差から人物の身長を推定する．

3.1 単眼人物検出システムの検出位置決定の不確定性

本節では，ターゲットシステムと，人物の身長によって検出位置がずれる性質について説明する．

ターゲットシステムでは，まず，入力として天井に取り付けたカメラで撮影した画像を与え，カメラ画像と人物モデルを比較する．そして，人物モデルが当てはまった場合，そのモデルの足元座標を出力する．

ターゲットシステムの人物検出アルゴリズムでは，人物の頭部を基準に同じ大きさの人物モデルを当てはめる．図1に示すように，どのような身長の人物であっても，モデル身長的人物として検出されるため，カメラ位置Cとモデル身長的人物の頭部 H_2 を結ぶ直線上に頭部 H_1 がある人物はすべて，同じ位置Pに足元があると検出される．つまり，検出した人物の身長がモデル身長と異なる場合，その検出位置が実際の位置からずれるという性質がある．このずれは地図において，カメラ位置と実際の人物位置を結ぶ直線上に起こる．なお，身長がモデルよりも低い人物は実際の位置よりもカメラに近い位置に検出され，その逆も同様である．また，検出位置Pから実際の位置Oのずれ量 Δd は図2から次式のように計算できる．

$$\Delta d = d \frac{x - T}{H - T} \quad (1)$$

ここで， H はカメラの高さ， T はモデル身長， x は人物の身長， d はカメラ位置Cと人物位置Oの地図における距離を示している．

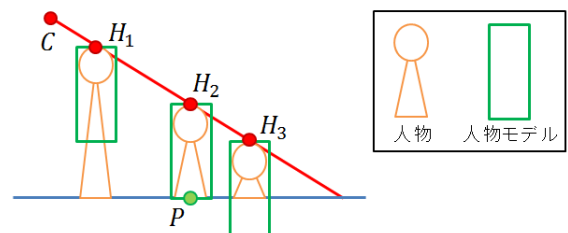


図1 同じ位置に検出される異なる身長の人物

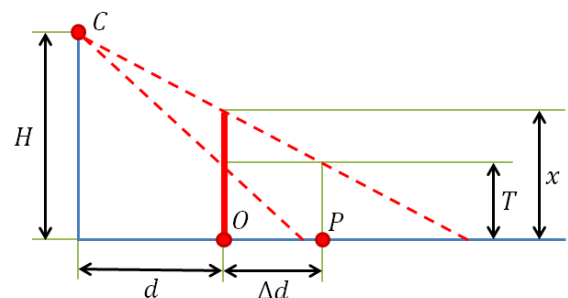


図2 身長による検出位置のずれ

3.2 非線型回帰によるキャリブレーション

本節では、非線型回帰をキャリブレーションに利用し、カメラの座標から地図の座標にマッピングする方法について説明する。

本手法では、カメラで検出した人物の座標を地図の座標に対応付ける。地図への対応付けは、各カメラのレンズ歪みを補正することになるので、地図上で人物座標を比較することでカメラ間連携ができるようになる。

一般に、カメラ画像から地図への変換を行う場合、カメラ画像と地図では座標系が異なるため、両者に共通するワールド座標系を考える。そして、ワールド座標系とカメラ座標系、また、ワールド座標系と地図の座標系の間でキャリブレーションすることで、ワールド座標系からカメラ座標系・地図の座標系への変換が可能になる。このようなカメラキャリブレーションは、レンズ歪みを補正するための内部パラメータのキャリブレーションと、カメラのワールド座標系における位置姿勢を定める外部パラメータのキャリブレーションに大別できる。

本手法では、歪み補正などのキャリブレーションをすべて合わせたものを、カメラ座標から地図座標への非線型写像で表すことができると考える。また、高速性が求められるため、以下では、線型回帰木を用いてカメラ座標系から地図の座標系へ直接変換する方法を示す。

3.2.1 PaLM-tree

本手法では PaLM-tree[8]を非線型写像の学習に用いる。これは、非線型写像を複数の線型写像で近似する線型回帰木である。

まず、 N 次元の入力空間と M 次元の出力空間を考える。そして、学習データとして入出力空間の間で複数の対応点が与えられている。そして、PaLM-tree は以下のように Split and Merge の要領で構築される。

始めに、線型回帰である Principal Component Regression (PCR) によって入力空間から出力空間への線型写像を求めることで、非線型写像を線型写像で近似する。このとき、近似誤差が許容誤差 E 以上であれば入力空間を 2 つの超直方体領域に分割する。そして、近似誤差が許容誤差 E 以下になるまで領域を分割していく。この分割のようすは回帰木で表現される。近似誤差が E 以下のとき、回帰木の終端ノードにその回帰係数を保存しておく。

次に、汎化性能を向上させるために、先ほど分割した領域を併合する。まず、入力空間の隣接する領域について、この 2 領域を併合した場合の、出力空間への線型写像を求める。そして、近似誤差が E 以下であればこれを併合し、対応する回帰木の終端ノードに保存されていた回帰係数を共有する。この併合を繰り返すことによって最終的に構築される回帰木が PaLM-tree である。

本手法では、カメラ画像 ($N = 2$) から地図 ($M = 2$) への写像を学習する。

3.2.2 マッピング

カメラ画像から地図への点のマッピングについて、あらかじめ、カメラ画像と地図の間で複数点を手動で対応付け、その対応点を PaLM-tree で学習しておく。そして、人物が検出された際、学習した PaLM-tree を用いて地図上の点を求めることでマッピングが行われる。以降では、カメラで検出された点を地図上にマッピングしたものを、人物の検出位置と呼ぶ。したがって、この検出位置には回帰木の精度に由来する誤差が含まれる。

3.3 人物検出・追跡システムの連携

本節では、地図を用いて複数カメラを連携する方法について説明する。

まず、検出した人物に ID を割り当て、カメラごとに人物を追跡する。次に、カメラ間で同一であると見なした人物同士が同一の ID になるよう ID を対応付ける。このとき、単一の地図上で同一人物かどうかを判定する。このように、地図を用いて複数カメラを協調動作させることで広域な人物追跡を実現する。

3.3.1 人物追跡

人物追跡はカメラごとに独立して行われる。あるカメラで現時刻に検出された人物について、1 時刻前に同一人物が検出されている場合、その ID を与えることで人物追跡が可能になる。また、1 時刻前に同一人物が検出されていない場合は、新たな ID を与えることで次の時刻から追跡される。ここでは、現時刻と 1 時刻前に検出された人物の間で、距離が最も近い人物から順に同一人物と見なす。

3.3.2 地図を用いた連携

地図を用いて、カメラ間で人物追跡を連携する。この連携は、複数カメラで撮影された同一人物について、ID が同一になるように ID を対応付けることで実現される。

あらかじめ、連携させるカメラ間で、それらの撮影範囲の一部を重ね合わせておく。そして、地図上において、重ね合わされた領域に該当する箇所に、矩形領域を手動で設定しておく。この領域を仮想ゲートと呼び、カメラ間連携の計算を行うための領域とする。仮想ゲートを設けることで、連携に不必要な計算を省くことができる。

次に、カメラ間で検出された人物が同一人物であるかを判定する。3.1 節で述べたように、人物の検出位置は実際の位置よりもずれることがある。そのため、この検出位置のずれを考慮し、実際の位置に近い人物同士を同一人物と見なす。ここでは、仮想ゲートから出た人物に対して、その人物と他の人物の実際の位置が、仮想ゲート内にいた間どの程度近かったかを計算することによって判定する。

以下では、すべての人物が同じ仮想ゲート内に検出されているとして、実際の人物位置の近さを計算する方法について述べる。ここでは検出位置を中心とする Gaussian を考え、その等確率楕円の重なりから人物間のスコア (2 人物がどれだけ同一人物かを表す指標) を求める。

(1) 等確率楕円と Gaussian の生成

カメラ位置 c と検出位置 p を結ぶ直線に沿う等確率楕円を生成する. この等確率楕円は, 図 3 のように標準偏差 σ_1, σ_2 と単位ベクトル φ_1, φ_2 によって表される. これらのパラメータを求めることで等確率楕円を生成することができる. 標準身長に対する身長の変動幅が高低ともに等しい場合, それを床面に射影するとカメラに近い方の変化幅は小さく, 遠い方は大きい. 等確率楕円はこの近似になる.

σ_1, φ_1 は, 身長による検出位置のずれが起こり得る範囲を表している. ここで, 人間の身長の分布が平均 μ_T , 標準偏差 σ_T の Gaussian になると仮定する. そして, 等確率楕円と身長の Gaussian の図 4 のような関係になると考える. この概念図において, $p(x)$ は, 位置 p にいる身長 x の人物が検出される位置を示している. つまり, σ_1 は身長 $\mu_T + n\sigma_T$ の人物に関する検出位置のずれ量に等しいと考える. ここで, n は等確率楕円の長さに関するパラメータである. 以上の考えに基づくと, σ_1 および φ_1 は式(1)を用いて次のように表すことができる.

$$\sigma_1 = \|p - c\| \frac{(\mu_T + n\sigma_T) - T}{H - T} \quad (2)$$

$$\varphi_1 = \frac{p - c}{\|p - c\|} \quad (3)$$

σ_2, φ_2 は, φ_1 方向以外の検出位置のずれの範囲を表している. これにはマッピングの精度が影響し, 単純には計算できないため, σ_2 は σ_1 よりも小さい任意の定数とする. また, φ_2 は $\varphi_1^T \varphi_2 = 0$ を満たし, $\|\varphi_2\| = 1$ である.

以上のパラメータを用いて, この等確率楕円を持つ 2 次元 Gaussian $g(x)$ は次式のように計算することができる.

$$g(x) = \frac{1}{2\pi\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right) \quad (4)$$

$$\Sigma = \sigma_1^2 \varphi_1 \varphi_1^T + \sigma_2^2 \varphi_2 \varphi_2^T \quad (5)$$

ここで, Σ は分散共分散行列であり, $\mu = p$ である.

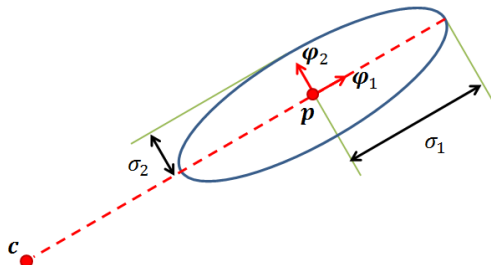


図 3 検出位置を中心とする等確率楕円

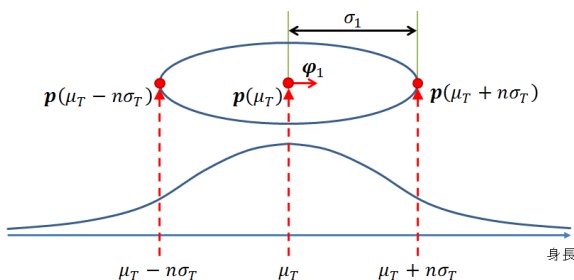


図 4 等確率楕円 (上) と身長の Gaussian (下) の関係

(2) 人物 ID の対応付けとスコア計算

等確率楕円の重なりから 2 人物間のスコアを計算し, 人物 ID の対応付けを行う.

カメラ C で検出された人物 (ID: x) が仮想ゲート V から出たとき, 以下の条件を満足する, C 以外のカメラ C' で検出された人物の ID: $y(x)$ を求め, これらに対応付ける.

$$y(x) = \operatorname{argmax}_{y \in S} \left(\sum_t \operatorname{score}(G(x, t), G(y, t)) \right) \quad (6)$$

ただし, S は, 人物 (ID: x) が仮想ゲート V の内部に検出されたすべての時刻について, C 以外のカメラで仮想ゲート V の内部に検出されたすべての人物の ID の集合を示す. また, $G(id, t)$ は, 時刻 t における人物 (ID: id) に関する確率分布を示す. $\operatorname{score}(G(a, t), G(b, t))$ は異なるカメラで検出された 2 人物 (ID: a, b) 間のスコアを示し, 次式のように計算される.

$$\operatorname{score}(G(a, t), G(b, t)) = \sum_{x \in D} s(x; a, b, t) \quad (7)$$

$$s(x; a, b, t) = g(x; a, t) g(x; b, t) \quad (8)$$

特に式(8)は, ある 1 点 x における時刻 t での 2 人物 (ID: a, b) 間のスコアを示している. ここで, $g(x; id, t)$ は, 時刻 t における人物 (ID: id) の位置 p_{id} を平均ベクトルとする Gaussian であり, 式(4)から求められる. また, D はスコアを計算する領域であり, 等確率楕円の重なりに相当する. ここでは, 図 5 のように等確率楕円を囲う矩形領域を設定し, その矩形領域が重なる領域を D とする. ただし, 図 5 において, c_i, c_j は異なるカメラの位置を示している ($i \neq j$). なお, 実際には高速化のために, 式(7)では領域 D のすべての点を用いずにランダムに選んだ N_R 個の点で計算する.

以上をまとめると, 式(6)は, 各時刻について人物 (ID: x) とその他の人物とのスコアを計算し, 最終的に人物 (ID: x) が仮想ゲートの内部にいたすべての時刻を通じて, スコアの和が最も高かった人物を示していることになる.

3.4 身長推定

本節では, 検出位置のずれから人物の身長を推定する方法について述べる. 身長を推定することができれば, 人物の身長に基づいて検出位置のずれを補正することができるようになり, 後の ID 対応付けの精度を向上させることが期待できる.

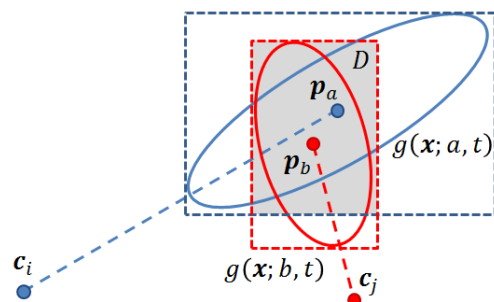


図 5 楕円を囲う矩形領域とスコアを計算する領域

実際の人物位置が判れば、式(1)における $\Delta d, d$ を求めることができ、身長を逆算することが可能である。ここで、検出位置はカメラ位置と実際の人物位置を結ぶ直線上にずれるため、図 6(a)の場合は2直線の交点を求めることで実際の人物位置を推定できる。図 6において、 c_i はカメラ位置 (i はカメラの番号)、 p_i はカメラ i で検出された人物位置、 e は推定位置である交点を示す。しかしながら、一般には図 6(b)のような場合が考えられるので、交点だけで単純に推定することはできない。

そこで、前節の計算を利用して実際の人物位置を推定する。図 5 の領域 D において最大のスコアを取った点は、同一人物の判定に最も寄与する点であると考えられるので、この点を人物位置として推定する。したがって、推定した人物位置 e は式(8)を用いて次式で表すことができる。

$$e = \operatorname{argmax}_{x \in D} (s(x)) \quad (9)$$

異なるカメラ c_1, c_2 で検出された同一人物 (Gaussian はそれぞれ $g_1(x), g_2(x)$ とする) に関する図である図 7 について、式(1)より人物 i の身長 x_i は次のように求めることができる。

$$x_i = T + \frac{\Delta d_i}{\|p_i - c_i\|} (H - T) \quad (10)$$

$$\Delta d_i = \begin{cases} \|p_i - e\| & \text{if } e, c_i \in R \text{ or } e, c_i \notin R \\ -\|p_i - e\| & \text{else} \end{cases} \quad (11)$$

ここで、 R は、 p_i を通り $p_i - e$ と直交する直線 l によって分割される領域である。 p_i に関する領域 R を図 8 に示す。

式(10)に示すように、1つの推定位置に対して2つの推定身長が求まる。また、2人物が仮想ゲートにいる間、各時刻で身長が推定される。以下では、ある同一人物に対して複数求まる身長を1つにまとめるための、2種類の方法について述べる。

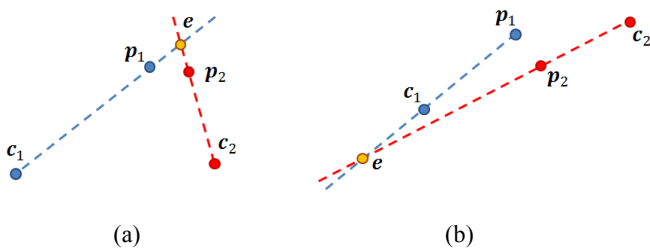


図 6 交点による人物位置推定

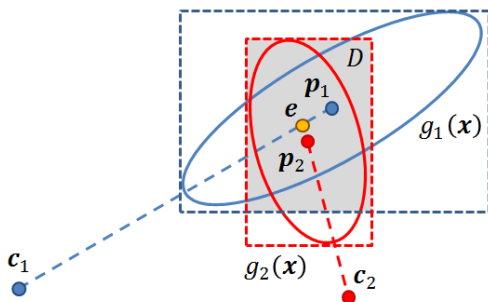


図 7 人物の推定位置

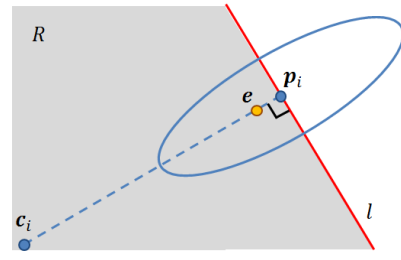


図 8 領域の判定 ($e, c_i \in R$ の場合)

(1) 平均による推定身長の統合

ある同一人物について、時刻 i における人物 j の検出位置を $p_j^{(i)}$ 、その推定位置を $e^{(i)}$ とする。例えば、図 7 では $j = 1, 2$ となる。また、検出位置 p について、式(10)で求めた身長を $h(p)$ とする。ここでは、これらの平均を取ることで同一人物の推定身長とする。したがって、ある同一人物に関する身長 $h(p_j^{(i)})$ の集合を S とすると、その身長 $h_m(S)$ は次式のように求められる。

$$h_m(S) = \frac{1}{|S|} \sum_{h \in S} h \quad (12)$$

(2) 重み付き平均による推定身長の統合

図 9 に示すように、推定位置 e が期待する位置 e' よりも外れる場合がある。このような場合、式(12)で計算すると統合結果に悪影響をおよぼす恐れがある。そこで、求めた身長 $h(p_j^{(i)})$ について、時刻ごとに重みを加えることで影響を減らすことを考える。

次式のように、式(8)を用いて、 $p_j^{(i)}$ に対応する推定位置 $e^{(i)}$ におけるスコアを重み $w(p_j^{(i)})$ とする。

$$w(p_j^{(i)}) = s(e^{(i)}) \quad (13)$$

ある同一人物に関する重み付き身長 $w(p_j^{(i)}) h(p_j^{(i)})$ の集合を S' とすると、その身長 $h_w(S')$ は次式のように求められる。

$$h_w(S') = \frac{1}{W} \sum_{h \in S'} h \quad (14)$$

ここで、 W は重みの合計を示す。

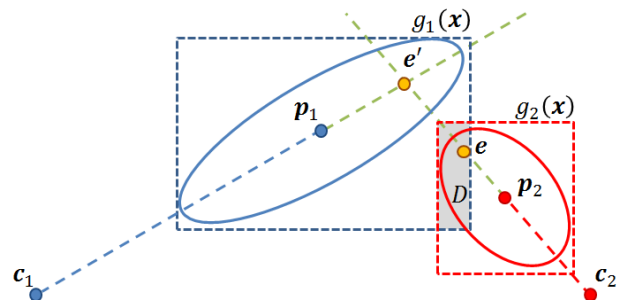


図 9 期待した推定位置が求まらない場合

4. 実験

身長推定の精度について確認するため、実験を行った。実験では身長の異なる3人物について、平均・重み付き平均の2通りの方法で身長を推定した。

4.1 実験環境

実験に用いたパラメータは次の通りである。

PaLM-tree 構築時の許容誤差	$E = 5.0[\text{cm}]$
身長 Gaussian の平均[9]	$\mu_T = 167.3[\text{cm}]$
身長 Gaussian の標準偏差[9]	$\sigma_T = 7.0[\text{cm}]$
カメラの高さ (実測値)	$H = 266.0[\text{cm}]$
モデルの身長	$T = 160.0[\text{cm}]$
等確率楕円の幅	$\sigma_2 = 5.0[\text{cm}]$
等確率楕円の長さのパラメータ	$n = 2$
スコア計算に用いる点の数	$N_R = 500[\text{個}]$

また、人物 1~3 の身長はそれぞれ、 $x_1 = 159.0, x_2 = 170.0, x_3 = 179.0$ である。人物検出システムには、技研システム株式会社の PALOSSIE AIO を使用した。このカメラが検出に用いる人物モデルの身長が不明であるため、ここでは仮に $T = 160.0$ と定めている。

実験では、図 10 に示す地図を用い、 c_i に位置する 3 台のカメラ C_i に対して 4 つの仮想ゲート V_j (矩形領域) を設定して連携させる。ここで、色で塗られた領域は壁や机など人物の侵入を想定しない領域を示している。また、格子は $50 \times 50[\text{pixel}]$ の大きさを示し、 $1[\text{pixel}]$ あたり $1[\text{cm}]$ の縮尺で地図が描かれている。

また、カメラ画像と地図の対応点を与え、PaLM-tree を構築しておく。ここでは格子に位置する点を対応点とすることで、カメラ C_1, C_2, C_3 に対してそれぞれ、102, 106, 102 組の対応点を手動で求めた。そして、これらの対応点を学習データとしてカメラごとに PaLM-tree を構築した。

4.2 身長推定の実験と考察

人物を 1 人ずつ歩かせて身長を推定した。その経路を図 11 に示す。人物は u_1 を始点に u_2 で折り返し u_1 に戻る。この往復を 3 回行うことで、のべ 12 個の仮想ゲートを出入りする。仮想ゲートを出る度に、平均・重み付き平均の 2 通りの方法で身長が推定される。なお、仮想ゲート V_3, V_4 は使用しなかった。

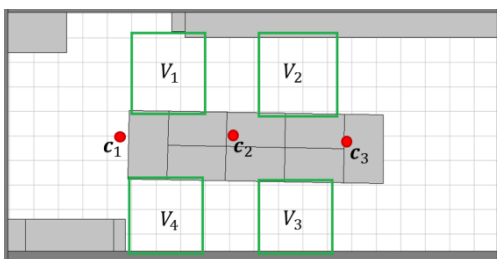


図 10 仮想ゲートとカメラ位置

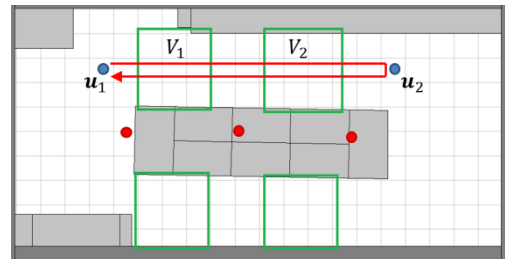


図 11 実験における人物の歩行経路

身長推定の結果を図 12~図 14 に示す。ここで、横軸は何番目の仮想ゲートを出たときの結果であるかを示しており、番号 k は V_1 を右に、 $k+1$ は V_2 を右に、 $k+2$ は V_2 を左に、 $k+3$ は V_1 を左に、仮想ゲートを出たときと対応している ($k = 1, 5, 9$)。

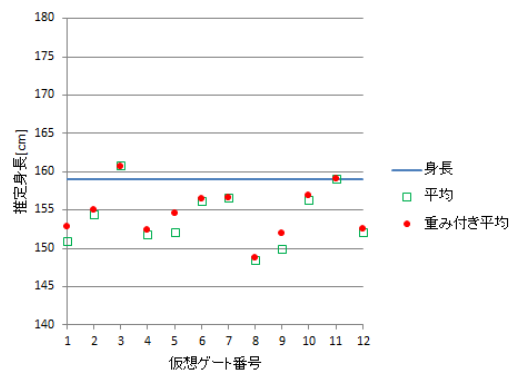


図 12 人物 1 (身長 159[cm]) の身長推定結果

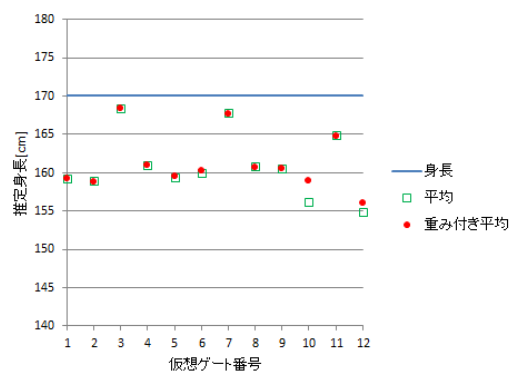


図 13 人物 2 (身長 170[cm]) の身長推定結果

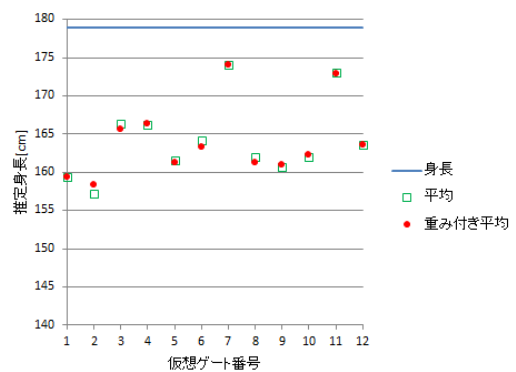


図 14 人物 3 (身長 179[cm]) の身長推定結果

以上の実験から、まず、実際の身長よりも身長が低く推定される傾向が見られる。特に平均を用いて計算した場合は、重み付き平均のときよりも低くなることもある。ただし、3人物の間で身長の低い人物は低く、身長の高い人物は高く推定されていて、身長の大小関係は全体的に維持されている。身長が低く推定される理由として、パラメータの一つであるモデル身長が適切でない可能性が考えられる。身長 170[cm]である人物 2 の場合、推定身長がモデル身長である 160[cm]の近くに分布している。身長がモデル身長と同じであると推定されるのは、人物の推定位置と検出位置に差がないとき、すなわち、身長による検出位置のずれが起こらない場合である。つまり、ずれが起こらないとき、人物の身長とモデル身長が等しいと考えることができるので、実際のモデル身長は 170[cm]程度ではないかと思われる。

他の傾向として、仮想ゲート番号 3,7,11 のときに、実際の身長に比較的近い値が求まることが分かる。特に人物 2 のときに顕著である。また、これらはすべて仮想ゲート V_2 を左に出たときに求められた身長である。これは単なる外れ値と見なすこともできるが、このような推定結果になる理由について今後検討していく。

5. おわりに

本報告では、非線型回帰計算でマッピングした座標を用いてカメラ間で人物追跡を連携する方法を示した。また、このカメラ間連携で考慮した人物位置のずれから、人物の身長を推定する方法を示した。実験では 2 通りの方法で身長を推定したが、精度に関してはまだ改善の余地があることがわかった。

今後は、図 6 で示した交点を用いて推定位置を求める方法を取り入れて、図 9 の場合でも人物の位置をより良く推定できるようにする。また、推定身長を検出位置の補正に用いて、ID 対応付けの精度向上を図る。

謝辞 本研究は技研トラステム株式会社からの支援を受けて実施された。

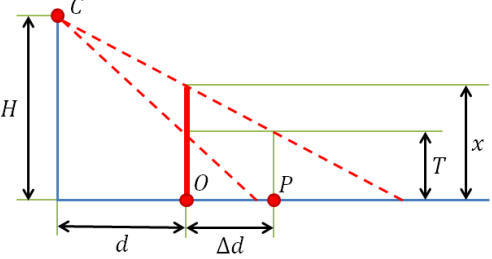
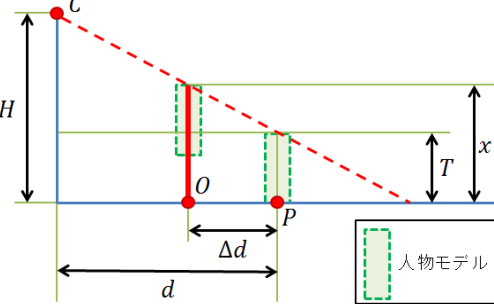
参考文献

- [1] 松山隆司, “分散協調視覚—研究成果と今後の展望—,” 情報処理学会研究報告, CVIM 121-6, pp.41-48, 2000.3.
- [2] R. T. Collins, A. J. Lipton, and T. Kanade, “Introduction to the Special Section on Video Surveillance,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, No.8, pp.745-746, 2000.8.
- [3] W. Du and J. Piater, “Multi-camera People Tracking by Collaborative Particle Filters and Principal Axis-Based Integration,” ACCV, 2007.
- [4] 松元郁佑, 加藤丈和, 和田俊和, “複数カメラを用いた CONDENSATION のための隠れに対して頑健な重み統合法,” 情報処理学会論文誌, CVIM, Vol.48, No.SIG9 (CVIM18), pp.100-114, 2007.6.

- [5] J. Black, T.J. Ellis, and D. Makris, “Wide area surveillance with a multi camera network,” Intelligent Distributed Surveillance Systems (IDSS-04), pp.21-25, 2004.
- [6] 和田俊和, 田村牧也, 松山隆司, “広域分散監視システムにおける分散協調型対象同定法,” 画像の認識・理解シンポジウム(MIRU'96)講演論文集, Vol. I, pp.103-108, 1996.
- [7] J. Verges-Llahi, T. Wada, “A General Algorithm to Recover External Camera Parameters from Pairwise Camera Calibrations,” ICIAR2008, pp.294-304, 2008.
- [8] 中村恭之, 加藤丈和, 和田俊和, “非線形写像学習のための PaLM-Tree の提案,” 日本ロボット学会誌, Vol. 23, No. 6, pp.732-742, 2005.
- [9] 厚生労働省, “平成 24 年国民健康・栄養調査報告,” 2012.3.

正誤表

「地図を用いた人物検出システムの連携法に関する研究」の内容に誤りがございましたので、訂正してお詫び申し上げます。

ページ	誤	正
2	<p>ここで、Hはカメラの高さ、Tはモデル身長、xは人物の身長、dはカメラ位置Cと人物位置Oの地図における距離を示している。</p>	<p>ここで、Hはカメラの高さ、Tはモデル身長、xは人物の身長、dはカメラ位置Cと検出位置Pの地図における距離を示している。</p>
2	 <p>図 2 身長による検出位置のずれ</p>	 <p>図 2 身長による検出位置のずれ</p>
5	<p>実際の人物位置が判れば、式(1)における$\Delta d, d$を求めることができ、身長を逆算することが可能である。</p>	<p>実際の人物位置が判れば、式(1)におけるΔdを求めることができ、身長を逆算することが可能である。</p>