

## 視体積の重なり状態の評価とゴースト物体特有な性質の評価を用いた 人物位置推定の高精度化

### Accuracy Improvement of Human Position Estimation using Evaluation of View Volume Intersection Status and Ghost Specific Properties

池田 浩雄<sup>†</sup>  
Hiroo Ikeda

大網 亮磨<sup>†</sup>  
Ryoma Oami

#### 1. はじめに

近年、監視カメラの普及と共にそれを用いた高度な映像監視やマーケティングへの期待が高まっている。特に、映像中から人/物の3次元位置を高精度に推定する技術は注目されつつあり、推定された位置から得られる動線を用いて、不審者、人/物の特定な動きを検知、理解することが求められている。

人物位置推定とは、画像情報から3次元空間上の人物位置を推定することである。従来法として、ステレオカメラによる人物位置推定法[1]がある。これは、3次元計測された特徴点群をクラスタリングすることで、人物を検出し位置を推定する手法である。しかし、画像間では対応点を取る必要があることから、人物をほぼ一方向から見ている。その為、棚や柱などの固定障害物や物体同士の重なりによるオクルージョンに弱い。

そこで、複数のカメラによる視体積情報を用いた人物位置推定を行う。視体積交差法(視体積のANDをベースとした物体復元)[2]を用いると、オクルージョンによる人物の未検出を改善できるが、一部でも物体領域(視体積)に欠損が生じると、ANDによる復元の為、物体復元が難しくなり未検出が起こりやすい問題がある。また、視体積は物体を含む広範囲に投影され広がる為、カメラと物体との位置関係で物体とは異なる空間に虚像物体(ゴースト物体)を形成し誤検出が起こる問題もある。

本稿では、物体を復元する際、視体積の重なり状態を評価し、さらに位置確定の際、ゴースト物体特有な性質を評価することで、人物の位置を精度よく推定する手法を提案する。

#### 2. 人物位置推定手法

##### 2.1. 概要

本手法の概要を図1で説明する。まず、複数のカメラ画像から物体領域を抽出し、抽出された物体領域を3次元空間上に逆投影する。逆投影した視体積データを用い、視体積の重なり状態を評価して3次元空間上に物体を物体尤度として復元する。

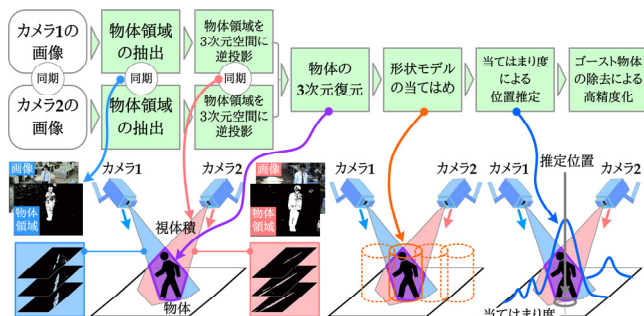


図1. 本手法の概要

復元データ(物体尤度)に検出対象の形状モデル(人物形状モデル)をスライディングさせながら当てはめ、当てはまり度(物体位置の尤度)を算出し、その度合が高い位置を人物位置として

推定する。さらに、ゴースト物体特有な性質を評価してゴースト物体を除去し高精度化する。当てはまり度は、形状モデル内の物体尤度の積分値として求める。視体積の重なり状態の評価による3次元復元については2.2節で、ゴースト物体特有な性質の評価による高精度化については2.3節で述べる。

##### 2.2. 視体積の重なり状態の評価による3次元復元

本手法は、視体積の重なり状態を2つの指標で評価し物体を物体尤度として復元する。1つ目の指標は、3次元位置 $x$ における各カメラからの視体積数 $L(x)$ を各カメラからの視界数 $M(x)$ で正規化した値 $r_1(x)$ で、式(1)にて求める(図2(a))。この指標は、交差する視体積の多さを表しており、物体が存在する可能性の高さを示す。

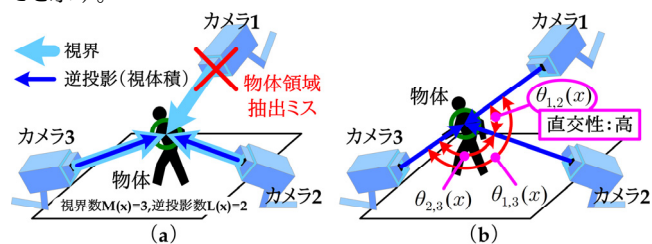


図2. 視体積の重なり状態の評価方法

$$r_1(x) = \frac{L(x)}{M(x)} \quad (1)$$

各空間の視界数はカメラ画角、固定障害物情報及びカメラパラメータを用いて事前に算出できる。視界数で正規化することで、画角や固定障害物等によるオクルージョンによる影響を指標に反映でき、ANDベースの視体積交差を用いた手法に比べ、物体領域の欠損による未検出に対してロバストとなる。

2つ目の指標は、3次元位置 $x$ において、交差する視体積の全てのペアの集合 $S$ を抽出し、ペアのカメラ $c_1, c_2 \in S$ の光軸角度 $\theta_{c_1, c_2}(x)$ の中で最も直交に近い角度を、直交する場合1になるよう正規化した値 $r_2(x)$ で、式(2)にて求める(図2(b))。この指標は、位置特定における物体復元の有効性を示す。

$$r_2(x) = 1 - \min_{c_1, c_2 \in S} |\cos \theta_{c_1, c_2}(x)| \quad (2)$$

$$P(x) = P_1(r_1(x))P_2(r_2(x))$$

$$P_1(r_1(x)) = \begin{cases} r_1(x) & : r_1(x) > 0.5 \\ 0.25r_1(x) & : otherwise \end{cases}, \quad P_2(r_2(x)) = \begin{cases} 1 & : r_2(x) > th2 \\ 0 & : otherwise \end{cases} \quad (3)$$

この指標が大きい(直交に近い)程、物体以外の領域が復元される割合が低くなり、物体位置の特定に有効になる。逆に、直交から離れると、交差する角度が浅くなり、物体以外の領域も含まれる割合が高くなる。結果、位置推定において多くの誤検出を招く。特に、式(1)のような交差する視体積の多さを評価する場合、その影響は強くなるので、この指標による評価が重要である。

<sup>†</sup> 日本電気(株) 情報・メディアプロセッシング研究所

本手法は、上記で説明した2つの指標を評価し、式(3)のように3次元空間上の物体を物体尤度  $P(x)$  として復元する。

### 2.3. ゴースト物体特有な性質の評価による高精度化

2.2節までの手法では、当てはまり度の評価によって、ある程度ゴースト物体を除去可能だが、ゴースト物体による誤検出を完全になくすことは困難である。そこで、図3のように、実物体の追跡と、ゴースト物体らしさの評価による手法を追加して、ゴースト物体に対処する。

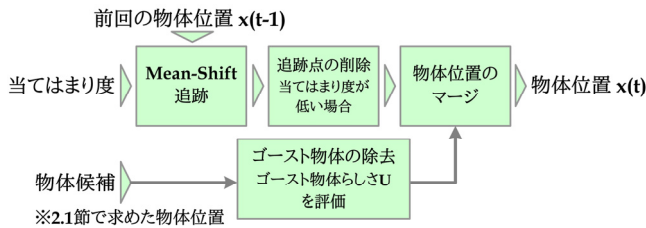


図3. ゴースト物体特有な性質の評価方法

まず、2.1節で求めた当てはまり度(物体位置の尤度)を重み  $\omega(x)$  とし、式(4)を用いた Mean-Shift 追跡を行う。ゴースト物体は3次元空間上に突然現れるという性質を持つことから、追跡によって、ゴースト物体を除去する。

$$\Delta x = \frac{\sum_{i=0}^N K(|x_i - x_0|, \sigma) \omega(x_i) (x_i - x_0)}{\sum_{i=0}^N K(|x_i - x_0|, \sigma) \omega(x_i)} \quad (4)$$

$\Delta x$  は移動量、 $K$  はカーネル関数(Gaussian)、 $x_0$  は初期位置、 $x_i$  は初期位置における周辺位置、 $N$  は周辺位置の数である。

さらに、2.1節で求めた物体位置を物体候補とし、それら物体候補に対してゴースト物体特有な性質から得られるゴースト物体らしさを評価する。結果、ゴースト物体らしさが低い物体候補を選択し、追跡結果とマーージすることで最終的な物体位置を求める。これにより、出入口の配置等により3次元空間上に突然現れる人物にも対応しながら、ゴースト物体による誤検出を低減できる。

ゴースト物体らしさの評価について説明する。ゴースト物体は、必ず実物体の視体積から構成されている。その為、ゴースト物体は、全ての視体積に対して物体候補が含まれやすい性質がある(図4)。この性質をゴースト物体らしさとして評価し、ゴースト物体を除去する。

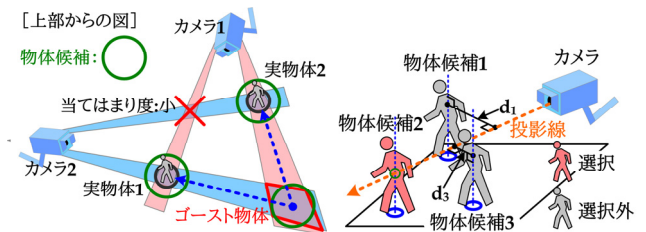


図4. ゴースト物体特有な性質 図5. ゴースト物体らしさの評価

まず、図5のように、物体候補を1つ選択し、物体候補内の3次元位置と対象カメラの位置を結ぶ投影線を視体積と仮定する。選択外の物体候補と投影線との距離  $d_i$  を計算し最小距離を求める。求めた距離に対して小さい程1、大きい程0に近づくように正規化し、対象カメラからの視体積に対して物体候補がどのくらい含まれているかという値を求める。この値は、選択した物体候補の高さ方向に対して複数求め、最終的には平均して用いる。求めた値を全てのカメラの視体積に対して取得し、その最小値を求

める。ゴースト物体ならば、全ての視体積に対して物体候補が含まれやすいので最小値が大きくなり、ゴースト物体らしさ  $U$  として評価できる。具体的には、式(5)のように計算する。

$$U = \min_{l \in S''} \left( \frac{\sum_{k=1}^R (f(\min_{j \in S'}(d_j)))_k}{R} \right)_l \quad (5)$$

$S'$  は選択外の物体候補の集合、 $S''$  は視界があるカメラの集合、 $R$  は選択した物体候補内の対象3D位置の総数、 $f$  は正規化関数である。

### 3. 実験及び考察

視体積交差法による従来法、視体積の重なり状態の評価を用いた本手法(2.2節)、さらに、ゴースト物体特有な性質の評価を追加した本手法(2.3節)を比較評価する。評価に用いるデータは、オフィスフロアで撮影された17シーンの動画(登場人物1~5名、カメラ5台)と、コンビニで撮影された10シーンの動画(登場人物5~10名、カメラ9台)である。また、データは、机、棚、柱などの固定障害物により、人物の全身が隠れたりする複雑な環境下である。各手法の性能は、リコール(再現率)とプレジジョン(適合率)のROC曲線を用いて評価する。正解の基準は、正解の人物位置に対して、最も近い推定位置を対応づけ、その距離が50cm以内の場合、正解としている。手法別のROC曲線と、推定位置の棒状アバタ表示及び地図表示の推定結果例を図6に示す。

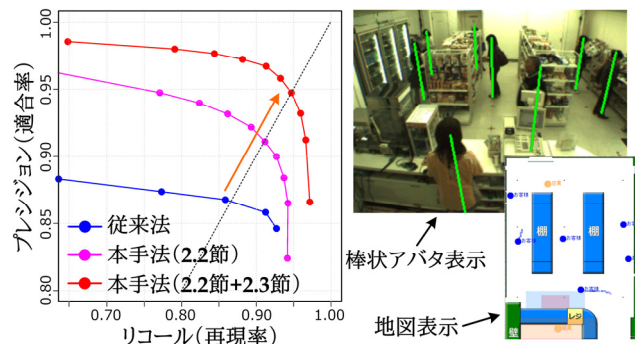


図6. 手法別のROC曲線と推定結果例

従来法と比べると、2つの評価を段階的に適用する毎にROC曲線が右上に上がり、各評価の効果と性能の改善が確認できた。さらに、本手法は、両レートが同値になる条件で、従来法86.7%を94.7%まで改善できた。これらにより、本手法の有効性を示した。

### 4. おわりに

本稿では、視体積の重なり状態とゴースト物体特有な性質とを評価することで、人物の位置を精度よく推定する手法を提案した。実験の結果、物体領域(視体積)の欠落による未検出と、ゴースト物体による誤検出が改善されることを確認し、本手法の有効性を示した。また、本手法は特殊なカメラ、特別なカメラ姿勢を必要とすることなく、人物位置推定が可能であることを示した。

### 参考文献

[1] 池谷ら, "ステレオビジョンセンサを用いた人物検知", 第10回画像センシングシンポジウム(SSII04)講演論文集, pp.231-234, 2004.  
 [2] A.Laurentini, "The Visual Hull Concept for Silhouette Based Image Understanding", IEEE, Transactions on Pattern Analysis and Machine Intelligence, Vol.16, No.2, pp.150-162, 1994