

# 視野を共有しないカメラ群を用いた人物行動軌跡の推定

小林 貴訓

## 1 はじめに

安全・安心な社会環境の実現のため、監視カメラ活用への期待は大きい。しかし一方で、人々の広大な活動空間や、プライバシーの問題などにより、人物の行動すべてをカメラで観察することは困難であることが多い。このような状況で、人物の行動を計測するためには、環境に疎に分散配置された視野を共有しないカメラ群からの情報を統合し、人物の行動を計測する技術が重要となる。

そこで、本研究では、屋内における人物の行動計測を目的とし、環境の知識をシステムに与えることで、人物の行動軌跡をカメラで観察されない箇所においても良好に推定する手法を提案する。

## 2 提案手法

Rahimi ら [1] と同様の考えに基づいて、ある時刻  $t$  における人物頭部の状態量  $\mathbf{x}_t$  を式 (1) のように定義する。

$$\mathbf{x}_t = [u_t \ \dot{u}_t \ v_t \ \dot{v}_t \ w_t \ \dot{w}_t]^\top \quad (1)$$

ここで、 $u_t, v_t, w_t$  は世界座標系における人物頭部の位置、 $\dot{u}_t, \dot{v}_t, \dot{w}_t$  はそれぞれの速度成分を表す。

また、観測開始から時刻  $T$  まで ( $t = 1 \dots T$ ) の軌跡  $\mathbf{X}$  を  $\mathbf{x}_t$  の状態集合として表現すると、軌跡  $\mathbf{X}$  は式 (2) のような  $1 \times 6T$  の行列として表現される。

$$\mathbf{X} = [\mathbf{x}_1^\top \ \mathbf{x}_2^\top \ \dots \ \mathbf{x}_T^\top]^\top \quad (2)$$

この軌跡  $\mathbf{X}$  は時間経過に伴う位置の履歴を含んでいるため、人物頭部の移動軌跡の推定は、この軌跡  $\mathbf{X}$  を推定する問題に帰着される。

本手法では、最大事後確率推定の枠組みに基づき、ある観測  $\mathbf{Y}$  が得られたとき、以下のように、事後確率が最大となる軌跡  $\mathbf{X}$  を選ぶことで、尤もらしい軌跡を推定する。

$$\begin{aligned} \arg \max_{\mathbf{X}} p(\mathbf{X}|\mathbf{Y}) \\ \approx \arg \max_{\mathbf{X}} p_m(\mathbf{X}) p_e(\mathbf{X}) p(\mathbf{Y}|\mathbf{X}) \end{aligned} \quad (3)$$

$p_m(\mathbf{X})$ ,  $p_e(\mathbf{X})$ ,  $p(\mathbf{Y}|\mathbf{X})$  は運動モデル、環境モデル、観測モデルを表しており、それぞれ以下のようにモデル化する

### 2.1 運動モデル

人物頭部の動きは急激に変化しないと仮定し、運動モデルを式 (4) のように準備する。

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{v}_t \quad (4)$$

ここで  $\mathbf{A}$  は状態遷移行列であり、 $\mathbf{v}_t$  は平均  $\mathbf{0}$  共分散行列  $\Sigma_v$  を持つノイズである。これより、軌跡  $\mathbf{X}$  の確率を式 (5) のように表現できる。

$$p_m(\mathbf{X}) = \prod_{t=1}^T N(\mathbf{x}_t | \mathbf{A}\mathbf{x}_{t-1}, \Sigma_v) \quad (5)$$

ここで  $N(\mathbf{x}_t | \mathbf{A}\mathbf{x}_{t-1}, \Sigma_v)$  は、 $\mathbf{x}_t$  の分布が平均ベクトル  $\mathbf{A}\mathbf{x}_{t-1}$  共分散行列  $\Sigma_v$  の多変量正規分布に従うことを表す。

### 2.2 環境モデル

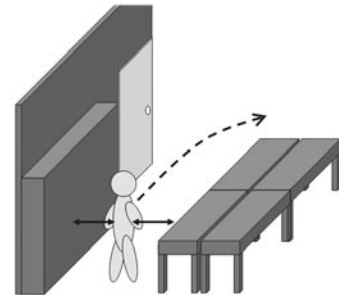


図 1: 環境モデル

通常、人物は壁を通過することはなく、机や棚などの障害物を越えて移動することもほとんどないという知見に基づき、障害物との距離に基づいて人物の存在確率を与える関数を環境モデルとして準備する。障害物との距離を  $d_h(\mathbf{x}_t)$  として、存在確率を  $e_h(d_h(\mathbf{x}_t))$  で表す。

また、人物頭部の高さについても、平均的な身長からの乖離に基づいて存在確率を与える関数を準備する。先と同様に平均的な身長からの乖離を  $d_v(\mathbf{x}_t)$  として、存在確率を  $e_v(d_v(\mathbf{x}_t))$  で表す。

これらより、軌跡  $\mathbf{X}$  に対する環境モデルを式 (6) のように与える。

$$p_e(\mathbf{X}) = \prod_{t=1}^T \exp\{e_h(d_h(\mathbf{x}_t)) + e_v(d_v(\mathbf{x}_t))\} \quad (6)$$

### 2.3 観測モデル

人物頭部の三次元位置の観測を、各カメラ画像への射影とノイズの和で記述し、観測モデルを式 (7) のように準備する。

$$\mathbf{y}_t^i = \pi^i(\mathbf{x}_t) + \omega_t \quad (7)$$

ここで  $\pi^i(\mathbf{x}_t)$  は、カメラの画像座標への射影を行う関数であり、 $\omega_t$  は平均ベクトル  $\mathbf{0}$ 、共分散行列  $\Sigma_\omega$  を持つ

ノイズである。これより、ある軌跡  $\mathbf{X}$  に対して観測  $\mathbf{Y}$  が得られる確率を式 (8) のように表現できる。

$$p(\mathbf{Y}|\mathbf{X}) = \prod_{(t,i) \in o} N(\mathbf{y}_t^i | \pi^i(\mathbf{x}_t), \Sigma_\omega) \quad (8)$$

ここで  $(t, i) \in o$  は時刻  $t$  において  $i$  番目のカメラで観測されることを表す。

## 2.4 各モデルによる評価の統合と最適化

これまでの各モデル  $p_m(\mathbf{X})$ ,  $p_e(\mathbf{X})$ ,  $p(\mathbf{Y}|\mathbf{X})$  を統合し、評価関数を書き下すと式 (9) のように表現できる。

$$\begin{aligned} \mathbf{X}^* = & \arg \min_{\mathbf{X}} \left( \mathbf{X}^\top \Lambda_{\mathbf{X}}^{-1} \mathbf{X} + \frac{1}{\sigma_\omega^2} \sum_{(t,i) \in o} (|\mathbf{y}_t^i - \pi^i(\mathbf{x}_t)|)^2 \right. \\ & \left. + \lambda \sum_t (e_h(d_h(\mathbf{x}_t)) + e_v(d_v(\mathbf{x}_t))) \right) \quad (9) \end{aligned}$$

ここで  $\lambda$  は環境モデルにより与えられる評価の調整に用いる変数である。また  $\Lambda_{\mathbf{X}}$  は以下で表される行列である。

$$\Lambda_{\mathbf{X}} = \begin{bmatrix} \mathbf{A}^\top \Sigma_v \mathbf{A} & -\mathbf{A}^\top \Sigma_v & 0 & \dots \\ -\Sigma_v \mathbf{A} & \Sigma_v + \mathbf{A}^\top \Sigma_v \mathbf{A} & -\mathbf{A}^\top \Sigma_v & \dots \\ 0 & -\Sigma_v \mathbf{A} & \Sigma_v + \mathbf{A}^\top \Sigma_v \mathbf{A} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

この評価関数を非線形計画法を用いて最適化することで、尤もらしい軌跡を推定する。本手法で用いる評価関数は凸関数ではないため、非線形計画法を適用した場合、局所的最適解に収束してしまう可能性が考えられるが、評価関数の最適化を始めるための初期軌跡  $\mathbf{X}_0$  を経験的に複数与え、それぞれに対して最適化を行い、得られた評価関数の値が最小となる軌跡を採用することで、この問題を回避する。

## 3 実験

これまでの提案手法に基づいて行った実験結果を図 2 に示す。図 2 の (A), (B), (C), (D) はカメラの位置を示す。得られたカメラ画像の例を図 3 に示す。

図 2 において、黒色の点線が本手法を用いて推定を行った結果である。灰色の実線は、人物が実験で移動したあらかじめ決められた軌跡あり、ground truth として表示している。また、環境モデルの導入による効果を明示するため、同じ観測データを用いて環境モデルなしで推定を行った結果を黒色の一点鎖線で表示している。

図 2 においてカメラの設置箇所付近で推定軌跡が ground truth に近づいている。これは評価関数の観測モデルにより、推定軌跡が拘束されているためと考えられる。一方、カメラで観測されない部分においては ground truth と差が大きくなっているが、軌跡は滑らかに接続されており、運動モデルによる拘束が有効に働いていると考えられる。また、環境モデルなしで推定した結果と比較して、机や壁などの障害物を避けた人物の行動軌跡を

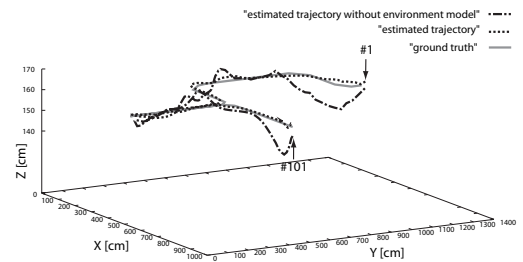
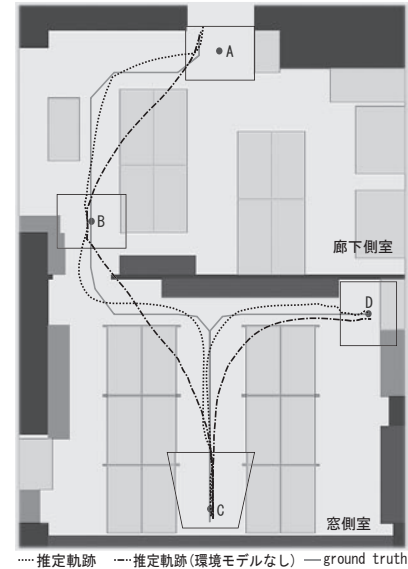


図 2: 室内レイアウト図と推定された軌跡

推定できていることから、環境モデルによる拘束が有効に働いていると考えられる。また、鉛直方向について乖離が少ないことから、環境モデルによる拘束が有効に働いていることが分かる。



図 3: カメラ画像の例

## 4 おわりに

本稿では、視野を共有しないカメラ群を用いて、人物の行動軌跡を、運動モデルと観測モデルに環境モデルを加えた評価関数の最適化を行うことで推定する手法を提案した。

今後は、視野外における人物の停留への対応や、観測に基づいた運動モデルの獲得、観測履歴からの学習に基づいた環境モデルの構築などの手法の検討を行う予定である。

## 参考文献

- [1] A. Rahimi, B. Dunagan, and T. Darrell, "Tracking People with a Sparse Network of Bearing Sensors," Proceedings of European Conference on Computer Vision (ECCV) 2004, pp.507-518, 2004.