

行動履歴を反映させた適応的環境属性を伴う三次元人物追跡

杉村 大輔[†] 小林 貴訓[†] 佐藤 洋一[†] 杉本 晃宏[§]

[†] 東京大学生産技術研究所

〒 153-8505 東京都目黒区駒場 4-6-1

E-mail : {sugimura, yosinori, ysato}@iis.u-tokyo.ac.jp

[§] 国立情報学研究所

〒 101-8430 東京都千代田区一ツ橋 2-1-2

E-mail : sugimoto@nii.ac.jp

あらまし パーティクルフィルタを用いた人物追跡技術は、これまでに様々な手法が提案され、その有用性が報告されている。一方で、追跡の安定化に有効な、画像以外の手がかりとして、環境属性情報がある。環境属性情報とは、対象シーン内の人物の存在可能性を意味し、これを追跡に利用することで人物が存在しにくい箇所への仮説の生成を抑制することができる。したがって環境属性情報は、効率的な仮説の生成に役立つと考えられる。この指標には、障害物の配置などの物理的な制約に依るもの、人物の行動履歴に依るものがある。本研究で提案する手法では、特に後者に焦点を当て、環境属性情報の獲得と追跡の枠組みへの統合を行う。具体的には、環境属性情報を混合正規分布で表現し、オンライン EM アルゴリズムにより人物行動履歴を逐次的に学習することで獲得する。さらに、ICONDENSATION の考えに基づき、追跡の枠組みに統合する。実環境における実験により、本手法の有効性を確認した。

キーワード：人物追跡，ICONDENSATION，オンライン EM アルゴリズム

People Tracking with Adaptive Environmental Attributes using the History of Human Activity

Daisuke SUGIMURA[†] Yoshinori KOBAYASHI[†] Yoichi SATO[†] Akihiro SUGIMOTO[§]

[†]Institute of Industrial Science, The University of Tokyo
4-6-1 Komaba, Meguro-ku,
Tokyo 153-8505, JAPAN

E-mail : {sugimura, yosinori, ysato}@iis.u-tokyo.ac.jp

[§]National Institute of Informatics
2-1-2 Hitotsubashi, Chiyoda-ku,
Tokyo 101-8430, JAPAN

E-mail : sugimoto@nii.ac.jp

Abstract Various tracking techniques based on particle filters have been proposed. To enhance the robustness of tracking, it is very significant to consider environmental attributes which represent an existing probability of people in a scene. They can be used for an effective hypothesis generation by considering the area where people are likely to exist. The environmental attributes can be considered in two aspects: the one is based on the physical configuration of objects in a scene and the other is based on the history of people activities. In this paper, we establish the history based environmental attributes that are updated by people tracking result everyframe using the online EM algorithm. Furthermore, we incorporate them into our tracking algorithm by using the ICONDENSATION framework. Our experimental results demonstrate the effectiveness of our method.

Keywords : people tracking, ICONDENSATION, online EM algorithm

1 はじめに

コンピュータの高速化やカメラの小型化、低価格化に伴い、カメラ映像を用いた様々な応用技術が注目されている。なかでも人物追跡技術は、防災・防犯を目的とするセキュリティや、効果的な店内レイアウトの設計に向けたマーケティングなどへの応用が期待されている。

カメラ映像を用いた人物追跡技術は、これまでに多くの手法が提案されているが、なかでもパーティクルフィルタの有用性が数多く報告されている ([2, 4, 5, 6, 8, 9, 10, 11] など)。パーティクルフィルタによる人物追跡では、追跡対象を状態量と尤度を持つ多数の仮説群により

離散的な確率密度として表現し、仮説の人物らしさをカメラ画像を用いて評価する。しかし、実環境での人物追跡では、カメラに対して人物の向きが不定であることによる見えの変化、人物が障害物や他の人物により隠れてしまう遮蔽、屋外の日照状況に代表される照明条件の変化、複雑背景などの問題により、画像から、いかなる状況でも正確に人物らしさを評価することは難しい。

この問題に対処し、安定な追跡を実現するために、複数の指標を用いて仮説の評価を行うことで、追跡の失敗を可能な限り防ぐアプローチや [2, 5, 6, 8, 9, 10, 11], 追跡が失敗した際に修正, 再初期化を速やかに行うことで, 正しい追跡を再開するアプローチ [4] がとられてきた。

複数の指標を用いて仮説の評価を行うことで, 追跡の失敗を可能な限り防ぐアプローチでは, 人物の色や輪郭などの複数の指標から人物らしさを評価する方法 [2, 5, 8, 9, 11] や, 複数の視野を共有するカメラによって人物を観察し, それぞれのカメラから得られる指標を統合することで人物らしさを評価する方法 [6, 9, 10] などがある。

これに対し, なんらかの理由により追跡が不安定となる, もしくは失敗した際に, 速やかに修正, 再初期化を行うことで正しい追跡を再開するアプローチでは, ICONDENSATION [4] がよく知られている。ICONDENSATION では, 画像から得られる指標のうち, 追跡に用いるものとは異なる指標を Importance Sampling の枠組みで利用する。これにより, 追跡に用いる特徴量が不安定となる場合でも, Importance Sampling により修正, 再初期化が働くため, 安定な追跡を行うことができる。

このように, 画像から得られる指標を用いて安定な追跡を実現する試みは一定の成果をあげているものの, 画像からの指標のみでは, 十分ではない場合も多い。鈴木ら [9] は画像から得られる指標に加えて, 環境属性を仮説の評価に統合することで, 追跡の失敗を防ぎ, 追跡の安定化を図る手法を提案している。そこでは, 対象空間内の机や椅子などの障害物からの距離や, その配置に依存した人物の存在可能性を環境属性として用いている。

しかし, 環境属性はこのような物理的なものだけではなく, 人物の存在確率ともいえる人物の行動履歴に基づく環境属性が考えられる。例えば, 人間は机や椅子の付近に滞留したり, 特定の通路を何度も通ることが多い。これは, 対象空間内の物理的な形状が同様であっても,

存在確率は人物の行動履歴に大きく依存することを示唆している。このことから, 行動履歴に基づいた人物の存在確率 (以後, 適応的環境属性と呼ぶ) を定義し, 利用することは, 追跡の安定化に有効であると考えられる。

人物の行動履歴に基づいて環境属性を獲得する例では, Demirdjian らによる Activity Maps がある [3]。彼らは, 蓄積した追跡データから, 類似する軌跡をクラスタリングすることで, 人間の行動のラベルを割り当てた activity zone と呼ばれるいくつかの領域に対象空間を分割している。これにより, 彼らは, 特定の位置において人物がどのような行動をとりやすいか, という地図 (Activity Map) を得ることができるとし, Context-aware Computing への応用を提案している。しかし, 追跡の安定化への応用は検討されていない。

そこで本研究では, 適応的環境属性の獲得とその追跡の枠組みへの統合を新たに提案する。適応的環境属性は混合正規分布で表現し, オンライン EM アルゴリズム [7] により人物行動履歴を逐次的に学習することで獲得する。また, ICONDENSATION の考えに基づいて適応的環境属性を追跡の枠組みに統合することで, 追跡が不安定となる場合でも, 速やかな修正, 再初期化を可能とする。

2 人物追跡における環境属性情報

人物追跡における環境属性情報は, 次の二つに分類することができる。一つは, 障害物の配置などに依存するものであり, “物理的環境属性” と呼ぶ。鈴木らはこれを仮説の尤度評価に統合することで追跡の安定化を図っている [9]。二つ目は対象シーンにおける人物の行動履歴に依存するものであり, 本稿ではこれを “適応的環境属性” と呼ぶ。本研究ではこれを獲得し, 追跡に統合することを目指す。以下, それぞれについて詳しく述べる。

2.1 物理的環境属性

壁の中, 床の下などの人物が存在しえない領域や, 机の下や棚の上など人物が存在しづらい領域は, どのような空間においても必ず存在する。パーティクルフィルタの枠組みでは, そのような領域にも仮説が発生することが起こり得て, これはサンプルの無駄打ちに他ならない。こういった領域に仮説が発生することを抑制するため

に、障害物の配置などの知識から人物の存在確率を定義する。これを物理的環境属性と呼ぶことにする。

2.2 適応的環境属性

環境属性は物理的なものだけでなく、対象空間内の人物の行動履歴に基づくものが考えられる。例を挙げて説明する。人間は机や椅子などの作業場所に滞留したり、複数ある通路のうち特定の通路を何度も通ることが多い。これは、つまり、対象空間内の人物の存在確率は空間内の人物の行動履歴に大きく依存していると思なすことができる。このことから、人物の行動履歴を長時間観測することにより、対象とするシーン（環境）に特化した人物存在確率を定義することができる。本研究ではこれを適応的環境属性と定義する。

3 提案手法

Fig.3.1 に提案手法の流れを示す。人物追跡は、人物の移動モデルにより伝播される事前確率と環境属性情報から仮説をサンプリングし、仮説の人物頭部らしさを評価する処理を逐次的に行うことで実現する（Fig.3.1 Tracking 部）。そのため、追跡開始前に、追跡の初期化、初期環境属性情報の構築が必要となる（Fig.3.1 Initialize 部）。そして追跡により各フレームで推定される人物頭部の三次元位置を用い、これをオンライン学習することにより、環境属性を更新する（Fig.3.1 Update 部）。このように、追跡と環境属性の学習が相互に作用しながら逐次的に処理が進む。その結果、長時間の観測により対象とするシーン（環境）に特化した適応的環境属性を獲得することができる。

3.1 人物頭部追跡

3.1.1 ICONDENSATION

人物頭部追跡の枠組みは ICONDENSATION[4] を用いる。ICONDENSATION とは、パーティクルフィルタの枠組みに Importance Sampling を組み入れることで、追跡対象の確率分布が複雑になる場合に起こり得るサンプルの無駄打ちを抑制し、より安定な追跡を行う手法である。具体的には、(i) 追跡初期化のための Importance function からのサンプリング、(ii) 追跡修正のための Importance function からのサンプリング、(iii) 事前確

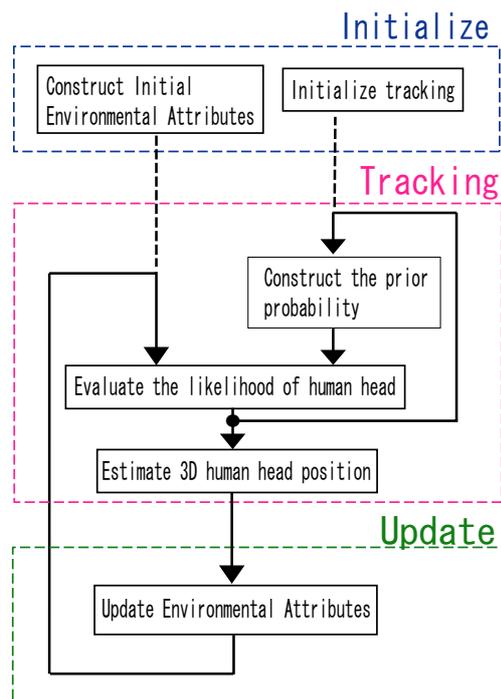


Fig. 3.1: System flow

率からのサンプリング（パーティクルフィルタ）の三つの仮説の発生手法を一定の割合で用いる。

3.1.2 追跡の枠組み

室内空間において、床面を XY 平面に一致させ、高さ方向に Z 軸をとる三次元世界座標系 XYZ を考える。追跡対象は一人の人物頭部とし、天井に固定された複数台のカメラを用いて三次元空間内で追跡する。

人物頭部を形状を不変とした楕円体でモデル化し、その中心座標を (x, y, z) とする。また時刻 t における n 番目の仮説が持つ状態量を $\mathbf{s}_t^{(n)} = [x_t^{(n)}, y_t^{(n)}, z_t^{(n)}]^T$ とする。時刻 t の仮説の状態量をとる確率変数を \mathbf{X}_t 、観測の状態量をとる確率変数を \mathbf{Z}_t とし、時刻 t までの観測列を $\mathcal{Z}_t = \{\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_t\}$ とする。

3.1.3 人物頭部らしさの評価

人物頭部らしさの評価は、画像平面に投影された仮説 $\mathbf{s}_t^{(n)}$ を、観測される画像の特徴量から評価する。本稿では、背景差分により抽出される前景領域と、エッジ勾配を用いて評価する。

i 番目のカメラの背景差分による評価値を $\pi_{i,t}^{fore,(n)}$ 、 i 番目のカメラのエッジ勾配の利用による評価値を

$\pi_{i,t}^{edge,(n)}$ とすると、各カメラで評価される人物頭部らしさの重みは次のように表現される。

$$\pi_{i,t}^{(n)} = \pi_{i,t}^{fore,(n)} \cdot \pi_{i,t}^{edge,(n)}. \quad (3.1)$$

ただし、ICONDENSATION のサンプリング手法 (ii) より発生した仮説の重みは、次のように修正される。

$$\pi_{i,t}^{(n)} = \frac{p(\mathbf{X}_t = \mathbf{s}_t^{(n)} | \mathcal{Z}_{t-1})}{g(\mathbf{s}_t^{(n)})} \pi_{i,t}^{fore,(n)} \cdot \pi_{i,t}^{edge,(n)}. \quad (3.2)$$

ここで $g(\mathbf{X}_t)$ は Importance function であり、本提案手法では適応的環境属性を用いる。適応的環境属性については 3.2 節で詳しく説明する。

各々のカメラについて評価した後、それぞれの重み $\pi_{i,t}^{(n)}$ を統合することで、その時刻における人物らしさの評価値とする。統合された重み $\pi_t^{(n)}$ は、各カメラによる重みの積、すなわち、

$$\pi_t^{(n)} = \prod_i \pi_{i,t}^{(n)} \quad (3.3)$$

とする。

3.2 適応的環境属性情報の構築

3.2.1 環境属性情報の確率表現

環境属性を追跡の枠組みに統合する、すなわち Importance function として用いるために、確率モデルを定義する。広い空間を対象とする場合、人物の存在しやすい領域は複数存在すると考えられる。そこで本研究では、混合正規分布で環境属性を表現する。すなわち、

$$g(\mathbf{X}_t) = \sum_{i=1}^K \omega_{i,t} \mathcal{N}(\mathbf{X}_t | \boldsymbol{\mu}_{i,t}, \Sigma_{i,t}). \quad (3.4)$$

ここで $\mathcal{N}(\cdot)$ は平均 $\boldsymbol{\mu}_{i,t}$ 、共分散 $\Sigma_{i,t}$ を持つ正規分布を表し、 $\omega_{i,t}$ は重み、また K は正規分布の数を表す。

3.2.2 初期環境属性情報の構築

環境属性を更新するために、追跡結果を反映させる前の初期状態、すなわち、混合正規分布の初期パラメータ $\omega_{i,0}, \boldsymbol{\mu}_{i,0}, \Sigma_{i,0}$ ($i = 1, \dots, K$) の推定が必要である。本提案手法では、パラメータ推定手法に EM アルゴリズムを用いる。EM アルゴリズムとは、観測できない、もしくは未観測の事象が存在する場合にも使うことができ

る汎用的な最尤推定手法であり、様々な分野に応用されている [1, 7]。

推定すべきパラメータを θ 、観測されるデータ集合を $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$ とし、観測できない値をとる確率変数を \mathbf{Y} とする。このとき以下の 2 つのステップを収束するまで繰り返し計算する。

E ステップ :

次の対数尤度の条件付期待値を計算する。

$$\begin{aligned} Q(\theta | \theta^{(t)}) &= E \left[\ln p(\mathcal{X}, \mathbf{Y} | \theta) | \mathcal{X}, \theta^{(t)} \right] \\ &= \sum_j \ln p(\mathcal{X}, \mathbf{y}_j | \theta) p(\mathbf{y}_j | \mathcal{X}, \theta^{(t)}). \end{aligned} \quad (3.5)$$

M ステップ :

$Q(\theta | \theta^{(t)})$ を θ について最大化し、その θ を $\theta^{(t+1)}$ とする、すなわち、

$$\theta^{(t+1)} = \operatorname{argmax}_{\theta} Q(\theta | \theta^{(t)}). \quad (3.6)$$

これを混合正規分布のパラメータ推定に適用したとき、以下の計算を全ての i についてパラメータが収束するまで行う。そして得られた $\omega_i^{(t)}, \boldsymbol{\mu}_i^{(t)}, \Sigma_i^{(t)}$ ($i = 1, \dots, K$) を用いて環境属性、(3.4) の初期状態を求める。

$$\boldsymbol{\mu}_i^{(t)} = \frac{\langle \mathbf{x} \rangle_i(m)}{\langle 1 \rangle(m)}, \quad (3.7)$$

$$\Sigma_i^{(t)} = \frac{\langle \mathbf{x}\mathbf{x}^T \rangle_i(m)}{\langle 1 \rangle_i(m)} - \boldsymbol{\mu}_i^{(t)} \left(\boldsymbol{\mu}_i^{(t)} \right)^T, \quad (3.8)$$

$$\omega_i^{(t)} = \langle 1 \rangle_i(m). \quad (3.9)$$

ここで $\langle \cdot \rangle_i(m)$ は i 番目の正規分布が選択される事後確率 $p(i | \mathbf{x}_l, \theta^{(t)})$ の重み付き平均、

$$\langle f(\mathbf{x}) \rangle_i(m) \equiv \frac{1}{m} \sum_{l=1}^m f(\mathbf{x}_l) p(i | \mathbf{x}_l, \theta^{(t)}) \quad (3.10)$$

を意味する。また、事後確率 $p(i | \mathbf{x}_l, \theta^{(t)})$ は、ベイズの定理により次のように表される。

$$\begin{aligned} p(i | \mathbf{x}_l, \theta^{(t)}) &= \frac{p(i | \theta^{(t)}) p(\mathbf{x}_l | i, \theta^{(t)})}{p(\mathbf{x}_l | \theta^{(t)})} \\ &= \frac{\omega_i^{(t)} \mathcal{N}(\mathbf{x}_l | \boldsymbol{\mu}_i^{(t)}, \Sigma_i^{(t)})}{\sum_{j=1}^K \omega_j^{(t)} \mathcal{N}(\mathbf{x}_l | \boldsymbol{\mu}_j^{(t)}, \Sigma_j^{(t)})}. \end{aligned} \quad (3.11)$$

本提案手法では、EM アルゴリズムを用いる際に必要なサンプル集合 $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_l, \dots, \mathbf{x}_m\}$ は、対象空間内の障害物の配置などの物理的な情報を用いることで発生させる。具体的には、対象空間全体を S 、障

害物のある領域を B_i ($i = 1, 2, \dots, l$) とし, 障害物がある領域全体を $P = \bigcup_{i=1}^l B_i$ としたとき, 人物が存在し得る領域 $E = S \cap \overline{P}$ に, 一様乱数を発生させる. これをサンプル集合として用いる.

3.2.3 適応的環境属性の獲得

人物の行動履歴を反映することで適応的環境属性を獲得する. 具体的には, 人物頭部追跡を行い, 各フレームで推定される人物頭部の三次元位置情報を用いて環境属性のパラメータを逐次更新する. この処理を実現するために, 本研究ではオンライン学習に対応した, オンライン EM アルゴリズムを用いる [7].

オンライン EM アルゴリズムとは, EM アルゴリズムのように観測データが全て得られた後にバッチ的に学習する手法ではなく, データが逐次的に与えられるような場合に, その都度パラメータを更新するアルゴリズムである. 時刻 t におけるパラメータは, 時刻 t にて得られたサンプルと, 時刻 $t-1$ で推定されたパラメータを用いて推定される. このとき, 重み付き平均, 式 (3.10) は次式で置き換えられる.

$$\langle\langle f(\mathbf{x}) \rangle\rangle_i(t) = (1 - \alpha) \langle\langle f(\mathbf{x}) \rangle\rangle_i(t-1) + \alpha f(\mathbf{x}) p(i | \mathbf{x}(t), \theta^{(t-1)}). \quad (3.12)$$

ここで α は学習率を意味する. これより, (3.7) ~ (3.9) 式はそれぞれ次のように置き換えられる.

$$\mu_i^{(t)} = \frac{\langle\langle \mathbf{x} \rangle\rangle_i(t)}{\langle\langle 1 \rangle\rangle_i(t)}, \quad (3.13)$$

$$\Sigma_i^{(t)} = \frac{\langle\langle \mathbf{x}\mathbf{x}^T \rangle\rangle_i(t)}{\langle\langle 1 \rangle\rangle_i(t)} - \mu_i^{(t)} (\mu_i^{(t)})^T, \quad (3.14)$$

$$\omega_i^{(t)} = \langle\langle 1 \rangle\rangle_i(t). \quad (3.15)$$

本手法では, 毎フレーム推定される人物頭部の三次元位置 s_i^{EAP} をサンプルとし, オンライン EM アルゴリズムを時刻 $t-1$ の環境属性に適用する. この操作により, 人物の行動の履歴が逐次環境属性に反映され, 長時間の観測により適応的環境属性を獲得することができる.

4 実験

実環境において, 室内天井に校正済みの IEEE1394 カメラ (Point Grey Research 社製 Flea) を 2 台設置して実験を行った. 解像度 640x480 ピクセル, 毎秒 30 フ

レームの速度で得られるカメラ映像を, 1 台の汎用 PC (CPU: Intel Pentium4 3.2GHz, Memory: 1GByte, OS: WindowsXP) で処理した. Fig.4.1 に観測されるカメラ映像を示す. ICONDENSATION に用いた仮説数は 216 個とし, 3 種類のサンプリング手法は等しい割合で用いた. また状態遷移モデルはランダムウォークを用いた.

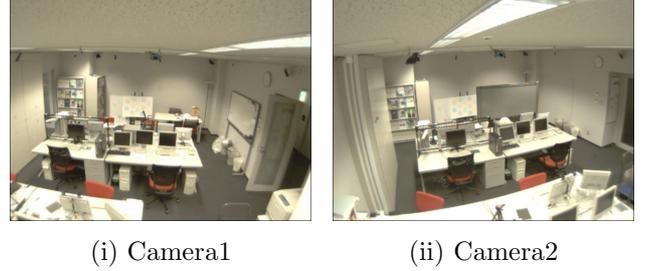


Fig. 4.1: Target space

4.1 適応的環境属性の獲得

4.1.1 初期環境属性の構築

Fig.4.2 に示すような対象空間の物理的な情報に基づき, サンプルを 10000 個発生させた. そして, そのサンプル集合を用いて EM アルゴリズムを適用することにより, 初期環境属性のパラメータを推定した. また正規分布の個数は $K = 7$ とした.

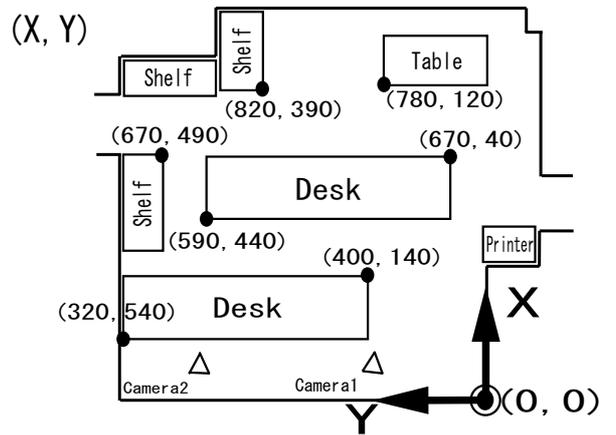


Fig. 4.2: Physical information of target space

推定された初期環境属性を Fig.4.3 に示す. Fig.4.2 のレイアウト図と比較して見ると, 人物が存在しやすいと考えられるカメラから見て手前側の通路 ($400 < X < 600, 100 < Y < 450$) の領域で存在確率が大きな値を示していることがわかる. また奥側の通路 ($650 < X <$

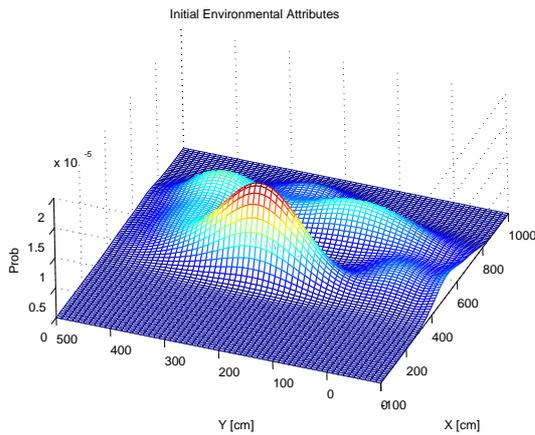


Fig. 4.3: Initial Environmental Attributes

800, $0 < Y < 500$) の領域にも手前の領域ほどではないが、存在確率の山が確認される。そしてその間にある机の領域 ($590 < X < 670, 40 < Y < 440$) で存在確率がこの両者の領域に比して十分小さくなっていることが見てとれる。これらの領域以外でも、人物が存在しにくい障害物のある領域では存在確率が低くなり、存在しやすい空いた領域では高くなっていることがわかる。

以上より、障害物のある領域には人物は存在しにくいという物理的な制約を記述できていると考えられるため、この分布は初期環境属性としては妥当な分布であるといえる。

4.1.2 適応的環境属性の獲得

対象空間を長時間観測することで、その間に追跡した人物の行動履歴を環境属性に反映し、適応的環境属性を得た。

Fig.4.4 に対象空間を 1 時間観測した結果を示す。Fig.4.3 と比較して Fig.4.4 で見られる大きな変化としては、まず第一に室内の出入口付近 ($400 < X < 670, -100 < Y < 0$) にピークが立っている。これは、対象空間内から外に出るにはこの出入口しかないため、部屋を出入りする人物は必ずこの付近を通過する。そのため、この付近の領域において三次元頭部位置が環境属性に頻繁に反映される。そしてその結果、ピークが立ったと考えられる。次に見られる顕著な変化としては、初期状態 (Fig.4.3) では、カメラから見て手前側の通路の領域 ($400 < X < 600, 100 < Y < 450$) の方が奥側の通路の領域 ($650 < X < 800, 0 < Y < 500$) よりも環境属性は高い値を示していたにも拘わらず、1

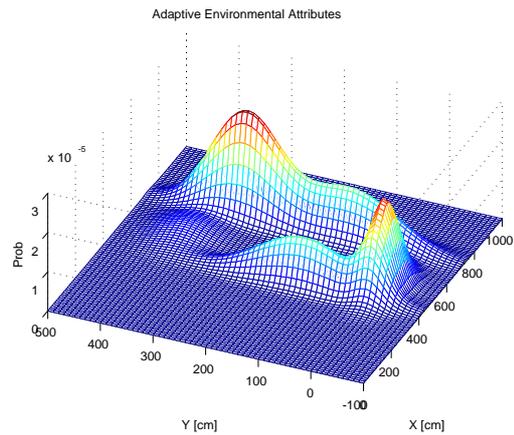


Fig. 4.4: Adaptive Environmental Attributes (After 1hour)

時間観測した後の適応的環境属性は逆に奥側の通路の方が明らかに高い値を示している。この結果は、手前側の通路よりも奥側の通路の方が人物はよく通過することを意味し、実際の状況と正しく符号する。奥の棚付近 ($670 < X < 820, 200 < Y < 400$) でピークが現れた理由は、その辺りで作業をしていた人がいたためであり、また奥の部屋と行き来できる出入口付近 ($670 < X < 820, 390 < Y < 490$) にも近いためとも考えられる。

以上の考察により、1 時間観測後に得られた適応的環境属性は、対象空間内の人物の行動履歴を正しく反映しているといえる。

4.2 適応的環境属性を用いた人物追跡

4.2.1 環境属性情報の有効性

環境属性の効果を確認するために、追跡の比較実験を行った。比較のためのサンプルとして、部屋の出入口から入室し、そのまま奥にあるテーブルの方向へ歩き、そしてボードの裏側で滞留した後、奥の部屋へ向かう行動をとる映像を、予め撮影し、準備した。その映像に対して、パーティクルフィルタ、提案手法それぞれのアルゴリズムを適用することにより、追跡結果を比較した。

Fig.4.5 に両者の手法による追跡結果を示す。(a),(c),(e) はパーティクルフィルタによる追跡結果、(b).(d),(f) は提案手法による追跡結果である。人物がテーブルの方へ歩いている ((a),(b)) ときは、両者の手法共に人物頭部を追跡できている。しかし、ボードの

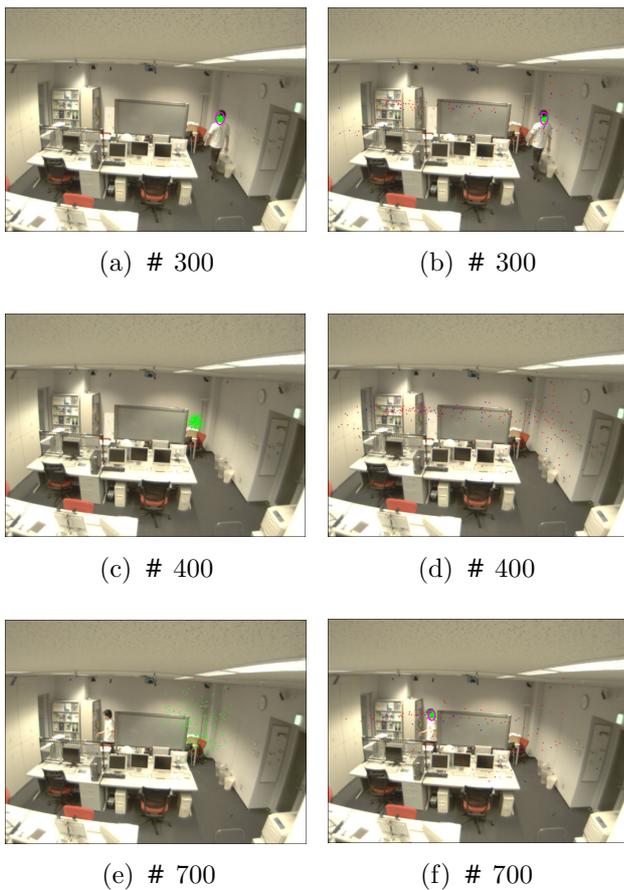


Fig. 4.5: Comparison of Tracking Result
(Partile Filter v.s. Proposed Method)
left column((a),(c),(e)): using particle filter
right column((b),(d),(f)): using proposed method

遮蔽により人物を見失った ((c),(d)) 後、パーティクルフィルタによる追跡では人物を見失った領域付近で探している ((e)) が、本提案手法では正しい追跡を再開できていることがわかる ((f))。これは、適応的環境属性からの仮説のサンプリングにより、三次元的に追跡の再初期化が働いたためである。このことから、本手法は一時的な遮蔽に対して頑健な手法であるといえる。

4.2.2 適応的環境属性の有効性

人物の行動履歴を反映することの有効性を確認するために、行動履歴を反映する前後の環境属性を伴う追跡の比較実験を行った。前節の実験方法と同様で、予め撮影しておいた映像（奥の出入口から、奥の通路を通り、外の出入口へ向かう行動）に対してそれぞれのアルゴリズムを適用することで、追跡結果を比較した。

Fig.4.6 に両者の手法による追跡結果を示す。(g),(i),(k) は初期の環境属性を伴う追跡結果, (h),(j),(l) は提案手法による追跡結果である。1180 フレーム目の画像 ((g),(h)) において、初期の環境属性を伴う追跡は初期化に失敗している ((g)) のに対し、適応的環境属性を伴う本提案手法では、正しく初期化が行われ、追跡を開始している ((h))。1200 フレーム目 ((i),(j)) では、ここで初期環境属性を伴う追跡の方は再初期化が働き、多少ずれてはいるものの、人物頭部の追跡を開始している ((i))。そして 1220 フレーム目 ((k),(l)) では、両者で正しい追跡が行えていることがわかる。この結果は、行動履歴を反映することで、速やかに人物の追跡を開始できることを示している。環境属性は、4.1.2 節で述べたように、人物の行動履歴を反映することにより、奥の部屋との出入口付近において人物存在確率が高くなっている。これは、この付近の領域で仮説が密にサンプリングされることを意味しており、より安定に頭部位置の検出が行われると考えられる。その結果、速やかに正しい位置での頭部検出が行われ、正しい追跡が開始される。

5 おわりに

本稿では、人物は特定の通路を何度も通ったり、特定の位置で滞留することが多いというような特性を、適応的環境属性という人物の行動履歴に依存する人物の存在確率として定義し、その構築方法、追跡への統合方法を提案した。追跡の枠組みには ICONDENSATION を利用し、適応的環境属性の構築にはオンライン EM アルゴリズムを適用することで実現した。

適応的環境属性を伴う追跡手法の安定性、特に、追跡対象を見失った際に、速やかに再初期化を行うことで追跡を再開できることを、実環境における精度比較の実験で確認した。

本稿では視野内に人物が一人であるという前提を用いているが、実環境においては視野内に同時に複数の人物が存在することは多々考えられる。また、同時に複数人の行動を環境属性情報反映させることは実際の環境により適応した属性情報を獲得することができると考えられる。そのため、今後は本手法を複数人に対応させることで、本手法の有用性をより実際の状況に近い形で再確認する予定ある。また、追跡の安定化以外に、適応的環境属性情報は駅構内などにおける混雑解消のための指標としての応用が考えられる。このような応用につ



(g) # 1180



(h) # 1180



(i) # 1200



(j) # 1200



(k) # 1220



(l) # 1220

Fig. 4.6: Comparison of Tracking Result
(Initial Env. v.s. Adaptive Env.)
left column((g),(i),(k)): using initial env.
right column((h),(j),(l)): using adaptive
env. (proposed method)

いてもまた今後検討する予定である。

参考文献

- [1] J. A. Bilmes, "A Gentle Tutorial of the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models," *Technical Report ICSI-TR-97-021*, 1997.
- [2] S. Birchfield, "Elliptical Head Tracking Using Intensity Gradients and Color Histograms," *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, pp.232-237, 1998.
- [3] D. Demirdjian, K. Tollmar, K. Koike, N. Checka and T. Darrell, "Activity maps for location-aware computing," *Proc. IEEE Workshop on Applications of Computer Vision*, 2002.
- [4] M. Isard and A. Blake, "CONDENSATION: Unifying low-level and high-level tracking in a stochastic framework," *Proc. European Conference on Computer Vision*, vol.1, pp.893-908, 1998.
- [5] Y. Jin and F. Mokhtarian, "Data Fusion for Robust Head Tracking by Particles," *Proc. IEEE International workshop on Visual Surveillance and Performance Evaluation and Tracking*, 2005.
- [6] 松本郁祐, 加藤丈和, 和田俊和, "Network Augmented Multisensor Association-CONDENSATION: CONDENSATION の自然な拡張による 3次元空間内での人物頭部の実時間追跡," *情報処理学会研究報告*, 2005-CVIM-150-21, pp.161-168, 2005.
- [7] M. Sato and S. Ishii, "On-line EM Algorithm for the Normalized Gaussian Network," *Neural Computation*, 12(2), 2000.
- [8] 杉本晃宏, 谷内清剛, 松山隆司, "確信度付き仮説群の相互作用に基づく複数対象追跡," *情報処理学会論文誌*, vol.43 no.SIGCVIM 4, pp.69-84, 2002.
- [9] 鈴木達也, 岩崎慎介, 小林貴訓, 佐藤洋一, 杉本晃宏, "環境モデルの導入による人物追跡の安定化," *電子情報通信学会論文誌 DII*, vol.J88-DII no.8, pp.1592-1600, 2005.
- [10] Y. Wang, J. Wu and A. Kassim, "Particle Filter for Visual Tracking Using Multiple Cameras," *Proc. IAPR Conference on Machine Vision Applications*, pp.298-301, 2005.
- [11] C. Yang, R. Duraiswami and L. Davis, "Fast Multiple Object Tracking via a Hierarchical Particle Filter," *Proc. IEEE International Conference on Computer Vision and Pattern Recognition*, vol.1, pp.212-219, 2005.