

修 士 論 文

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

Tracking People using Adaptive Environmental Attributes
based on Long-Term Observation of Human Activities



東京大学大学院
情報理工学系研究科
電子情報学専攻

56426

杉 村 大 輔

指導教官

佐藤 洋一 助教授

平成 19 年 2 月

概要

視野を共有する複数のカメラを用いて, 三次元空間における人物頭部の実時間追跡を行う. 追跡に用いる指標として, 観測画像から得られる画像特徴量以外に, 環境属性情報を用いる. 環境属性情報とは, 対象シーン内の人物の三次元的な存在可能性を意味し, 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. 本研究では, 後者を獲得し, そしてそれを追跡の枠組みに統合する. 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. そして, ICONDENSATION の考えに基づき, 環境属性情報を Importance Function として追跡の枠組みに統合する. これにより人物追跡と対象シーンに特化した環境属性の獲得の両者を同時に実現することを可能とする. 最後に, 実環境における実験により, 本手法の有効性を確認する.

目次

第1章	序論	1
1.1	はじめに	1
1.2	本論文の構成	3
第2章	関連研究	4
2.1	パーティクルフィルタによる追跡	4
2.1.1	時系列フィルタリング	4
2.1.2	パーティクルフィルタ	5
2.1.3	パーティクルフィルタを利用した人物追跡における問題点	6
2.2	観測モデルの高精度化	9
2.2.1	複数情報の統合	9
2.2.2	複数情報の階層的評価	10
2.2.3	顔識別器の利用	12
2.2.4	環境属性情報の利用	12
2.3	サンプリングの効率化	15
2.3.1	ICONDENSATION	15
2.3.2	追跡器と検出器の併用	16
2.4	行動履歴の利用による環境情報の獲得	18
2.5	本研究の位置付け	20
第3章	行動履歴の反映による適応的環境属性を伴う三次元人物追跡	22
3.1	環境属性のオンライン学習と人物頭部追跡の同時実行	22
3.2	人物頭部の三次元追跡	24
3.2.1	人物頭部モデル	24
3.2.2	仮説のサンプリング	24
3.2.3	仮説の画像平面への投影	26
3.2.4	観測による人物頭部らしさの評価	26

3.2.5	人物頭部位置推定	29
3.2.6	複数人追跡への拡張	30
3.2.7	サーバ・クライアントモデルの構築	32
3.3	適応的環境属性の獲得	34
3.3.1	環境属性情報の定義	34
3.3.2	適応的環境属性の確率表現	35
3.3.3	動線の反映による適応的環境属性の獲得	36
3.3.4	人物の行動状態推定	38
第4章	実験	42
4.1	実験環境	42
4.2	適応的環境属性の獲得実験	43
4.2.1	環境属性の初期化	43
4.2.2	適応的環境属性の獲得	45
4.3	人物追跡実験	48
4.3.1	環境属性の効果	48
4.3.2	適応的環境属性を伴う人物追跡実験	49
第5章	結論	53
5.1	まとめ	53
5.1.1	環境属性を組み入れた人物頭部三次元追跡	53
5.1.2	適応的環境属性の獲得	53
5.2	今後の展望	54
5.2.1	環境属性の表現の再考案	54
5.2.2	観測モデルの高精度化	55
5.2.3	行動状態推定の詳細化	55
	謝辞	56
	参考文献	57
	発表文献	61

目次

2.1	確率密度関数の離散表現	6
2.2	パーティクルフィルタの処理手順	7
2.3	観測画像と画像特徴量	11
2.4	カスケード型顔識別器による人物頭部らしさの評価分布 [15,16]	13
2.5	エッジ情報による人物頭部らしさの評価分布 [15,16]	13
2.6	レンジセンサにより得られた室内形状 [30]	14
2.7	物理的環境属性情報 [30]	14
2.8	追跡器と検出器の混合分布 [22]	18
2.9	Activity Map [5]	21
3.1	システム全体図	23
3.2	人物頭部の画像平面への投影	27
3.3	評価する楕円の輪郭位置と法線ベクトル	29
3.4	サーバ・クライアントモデル	33
3.5	人物存在可能性の偏り	35
3.6	人物の滞留による環境属性の集中	39
3.7	人物の行動状態遷移図	40
4.1	観測視野	43
4.2	対象空間の物理的情報	44
4.3	環境属性の初期状態	44
4.4	適応的環境属性の獲得に用いた観測映像	46
4.5	適応的環境属性（動線のみ利用）	47
4.6	適応的環境属性（行動状態推定を伴う）	47
4.7	行動状態推定器による滞留行動状態の推定結果	48
4.8	環境属性の有無に伴う追跡結果比較	50
4.9	行動状態推定の有無に伴う追跡結果比較	52

表目次

4.1 適応的環境属性の獲得実験に用いたパラメータ	45
-------------------------------------	----

第1章 序論

1.1 はじめに

近年の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応用技術が注目されている. なかでも人物追跡技術は, セキュリティやマーケティングなど様々な分野から大きな期待が寄せられている. 例えばセキュリティ分野では, 犯罪を未然に防ぐために, 対象シーン内において不審な行動をとる人物を自動的に検出する技術が必要とされている. 不審者の自動検出, すなわちコンピュータが不審者の判別を行うためには, 観測視野内の人物の正確な動線が必要となる. その人物の動線を獲得するために, カメラから観測される映像から人物を自動的に追跡する技術が用いられる. このように, 人物追跡技術は様々なアプリケーションにおける必須技術であり, その用途は幅広い.

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, なかでも時系列フィルタの一つである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 ([10–13, 15, 18–22, 24, 26, 29–32, 34, 35] など). パーティクルフィルタとは, 追跡対象を, 状態量と重みを持つ多数の仮説群で離散的に近似した確率密度分布として表現する. そして, 事前の状態と予め定めた状態遷移モデルによる“予測”, 観測されるカメラ画像を用いて追跡対象らしさを評価する“観測”の処理を, 逐次的に繰り返すことで追跡を実現するアルゴリズムである. 確率密度分布を離散的に近似することで, 任意の形状の分布を表現することが可能になるため, 観測される分布が非ガウス型になるような複雑な場合においても頑健な追跡を実現できる.

しかしながら, 実環境における人物追跡には様々な問題が考えられる. 例を挙げると, 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がある. そのため, 観測画像から正確に人物らしさを評価することは難しい.

これらの問題に対処するために, 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 [15, 18–21, 24, 26, 29, 30, 35] や, 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 [12, 13, 22, 32, 34] 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, 観測画像から得られる人物の色や輪郭などを用いて人物らしさを評価する比較的単純な評価手法 [18, 21, 24, 26, 29, 30, 35] や, 視野を共有する複数台のカメラでシーンを観測し, それぞれのカメラから得られる観測による評価を統合することで人物らしさを評価する方法 [15, 19, 20, 30] などがある. また近年では, カスケード型 Adaboost による顔識別器 [33] のスコアを評価に利用した, より高精度な評価手法が提案されている [15, 16].

その一方で, 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, パーティクルフィルタの枠組みに Importance Sampling と呼ばれるサンプリング手法を組み入れた, ICONDENSATION と呼ばれる手法がよく知られている [12]. 追跡の手がかりに利用していない, 観測される他の情報源 (Importance Function と呼ばれる) から仮説を発生させることで, パーティクルフィルタによる追跡が不安定になる場合においても, 修正, 再初期化が働くため, 安定な追跡が可能になる. これ以外にも, 複数対象追跡を目的とした柔軟な初期化, 修正のために, Adaboost 検出器 [33] を追跡と併用する手法も提案されている [22].

このように, 画像から得られる特徴量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が, これらの手法が必ずしも十分であるとはいえない. そこで, より安定した追跡を実現するために, 画像以外の異なるモダリティの情報を組み入れる手法が考えられる.

鈴木らは, 画像から得られる指標に加え, 対象空間における人物の存在可能性を意味する環境属性という指標を, 仮説の評価に統合することで, 追跡の安定化を図る手法を提案している [30]. 彼らは, 環境属性を対象空間内の机や椅子などの障害物からの距離や配置に依存した人物存在可能性として定義している.

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, 対象視野内の人物の行動履歴に基づく環境属性が考えられる. 店舗内などの屋内環境を想定した場合, 人物は立ち読みをするために特定位置で滞留したり, 何らかの要因により複数ある通路のうちある特定の通路を通行するなど, 環境に依存した特定の行動をとることが多い. これは, 環境内における人物存在可能性には, 物理的な情報のみでは表現できない行動履歴に伴う偏りがあ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た環境情報を獲得する例では, Demirdjian らによる Activity Maps がある [5]. 蓄積した追跡データから, 類似する行動特徴をクラスタリングすることで, Activity Zone と呼ばれる人物が同様の行動をとりやすい領域に対象空間を分割する. これを更に起き得る行動の頻度について選定することにより, 人物がある特定の行動をとりやすいと意味付けされた領域図を意味する, Activity Map を得ることができるとしている. 彼らはその Context-aware Computing への応用を提案しているが, 追跡の安定化への応用は検討されていない.

このような様々な知見に基づき, 本研究では, 行動履歴を反映した適応的環境属性を獲得し, 対象空間内の人物三次元追跡の安定化に利用する手法を新たに提案する. 具体的には次の通りである. 観測視野内の複数人物追跡を, Mixture Particle Filter [32] の考えに基づき実現する. 個々の追跡には ICONDENSATION [12] の考え方を利用し, 環境属性を Importance Function として定義する. これにより, 追跡が不安定となる場合においても, 環境属性からのサンプリングにより三次元的に追跡が修正, 再初期化されるため, 安定な追跡を実現することができる. それと同時に, 毎フレーム推定される人物の状態を, オンライン EM アルゴリズム [25] により逐次的に環境属性に反映させる. これにより, 対象シーンに特化した適応的環境属性を事前知識無しに獲得することができる. このように, これらの二処理が相互に作用し合うことで本手法は実現される.

1.2 本論文の構成

第2章では, 関連研究についていくつか紹介することで, 本研究の位置付けを確認する. 第3章では, 本研究内容について詳細に述べる. 第4章では, 実環境における実験の結果を示し, 考察する. そして第5章にて本研究のまとめと今後の展望について述べる.

第2章 関連研究

本章では、まず時系列フィルタの一つであるパーティクルフィルタ [11] について説明する。そして、パーティクルフィルタの枠組みにおいて、頑健な人物追跡を行うために高度化すべき点を言及する。その上で、2.2, 2.3 節においてそれらの点について関連研究を分類分けし、説明する。また 2.4 節では、人物の行動の履歴を対象空間の環境に反映する関連研究について説明する。そして本章の最終節において、これらの関連する研究例と本研究を対比することにより、本研究の位置付けを確認する。

2.1 パーティクルフィルタによる追跡

2.1.1 時系列フィルタリング

フレーム間の輝度変化等による特徴抽出の誤りや、画像中のノイズにより不完全な観測データしか得られない状況で対象を追跡する場合、時系列フィルタリングが有効であることが知られている。代表的なものにカルマンフィルタ、パーティクルフィルタがある。

時刻 t における追跡対象の状態量を \mathbf{X}_t 、また、観測値を Z_t 、観測の履歴を $\mathcal{Z}_t = \{Z_1, Z_2, \dots, Z_t\}$ とする。一般に、時系列フィルタによる状態推定は、事後分布 $p(\mathbf{X}_t | \mathcal{Z}_t)$ を推定する問題として定式化され、次のベイズの法則により推定することができる。

$$p(\mathbf{X}_t | \mathcal{Z}_t) \propto p(Z_t | \mathbf{X}_t) p(\mathbf{X}_t | \mathcal{Z}_{t-1}). \quad (2.1.1)$$

ここで $p(\mathbf{X}_t | \mathcal{Z}_{t-1})$ は時刻 t における事前分布を表す。また、 $p(Z_t | \mathbf{X}_t)$ は尤度を表し、観測画像から推定される。

推定対象がマルコフ過程に従うと仮定したとき、事前分布 $p(\mathbf{X}_t | \mathcal{Z}_{t-1})$ は Chapman-Kolmogorov 方程式、

$$p(\mathbf{X}_t | \mathcal{Z}_{t-1}) = \int p(\mathbf{X}_t | \mathbf{X}_{t-1}) p(\mathbf{X}_{t-1} | \mathcal{Z}_{t-1}) d\mathbf{X}_{t-1}, \quad (2.1.2)$$

により計算される。これにより逐次的な状態量推定が可能となり、対象の追跡が実現される。

2.1.2 パーティクルフィルタ

パーティクルフィルタは, Fig.2.1 に示すように, N 個の離散的な仮説 $s_t^{(n)}$ とそれに対応する重み $\pi_t^{(n)}$ を用いて, 確率密度分布 $p(\mathbf{X}_t | \mathcal{Z}_t)$ を多数の仮説から形成される仮説群 $\{s_t^{(n)}; \pi_t^{(n)}\}$ ($n = 1, 2, \dots, N$) として表現する. これにより, 非ガウス型となるような任意の確率密度分布を近似することが可能となり, ノイズや環境変動に対して頑健な追跡を実現できる.

パーティクルフィルタの具体的な処理手順は以下の通りである.

(i) 仮説の選択

時刻 $t - 1$ における事後分布 $p(\mathbf{X}_{t-1} | \mathcal{Z}_{t-1})$ を離散的に近似した N 個の仮説 $\{s_{t-1}^{(1)}, \dots, s_{t-1}^{(N)}\}$ に対応する重み $\{\pi_{t-1}^{(1)}, \dots, \pi_{t-1}^{(N)}\}$ の比に従い, 仮説群 $\{s_{t-1}^{(1)}, \dots, s_{t-1}^{(N)}\}$ を選択する.

(ii) 状態遷移モデルに基づく仮説の伝播

選択された仮説群 $\{s_{t-1}^{(1)}, \dots, s_{t-1}^{(N)}\}$ を, 予め定義した状態遷移モデル $p(\mathbf{X}_t | \mathbf{X}_{t-1} = s_{t-1}^{(n)})$ ($n = 1, 2, \dots, N$) によりそれぞれ伝播させ, 時刻 t における事前分布 $p(\mathbf{X}_t | \mathcal{Z}_{t-1})$ に従う仮説群 $\{s_t^{(1)}, \dots, s_t^{(N)}\}$ を生成する. (一般に, 状態遷移モデルはあらかじめ定めてある運動モデルによる“移動”と, 移動させた仮説にランダム雑音を加える“拡散”により構成される.)

(iii) 観測による重み $\pi_t^{(n)}$ の決定

仮説 $s_t^{(n)}$ の重み $\pi_t^{(n)}$ を, 時刻 t の観測量から得られる観測モデル $p(\mathbf{Z}_t | \mathbf{X}_t = s_t^{(n)})$ を用いて評価することにより決定する.

これらのステップを毎時繰り返すことにより, 対象の追跡を実現する. 以上の処理手順を Fig.2.2 に示す.

しかしながら, パーティクルフィルタを実環境下での人物追跡に利用する場合, 様々な問題が考えられる.

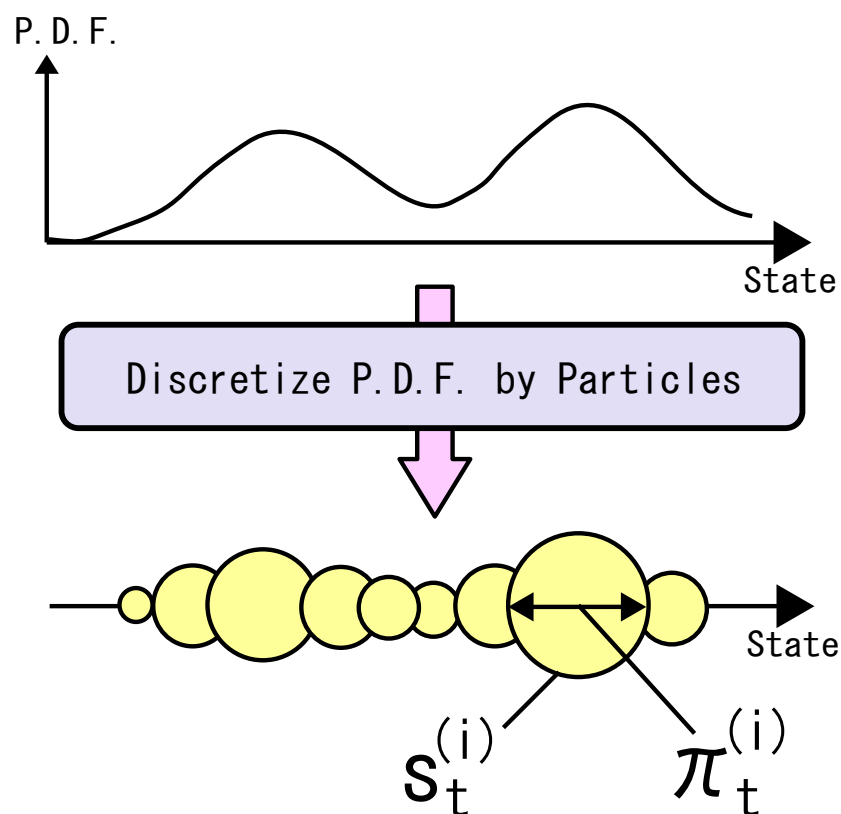


Fig. 2.1: 確率密度関数の離散表現

2.1.3 パーティクルフィルタを利用した人物追跡における問題点

実環境下での人物追跡において、以下のような問題により追跡対象の確率密度分布が正しく推定されず、複雑な分布形状を持つことがある。

- 見えの変化

観測視野内で行動する人物は、観測しているカメラに対して向きが不定であることから、常にその見えが変化する。そのため、様々な見えの変化に対応した観測モデルが必要となる。

- 遮蔽

屋内環境を想定した場合、観測視野内には机や椅子、本棚などの障害物が配置

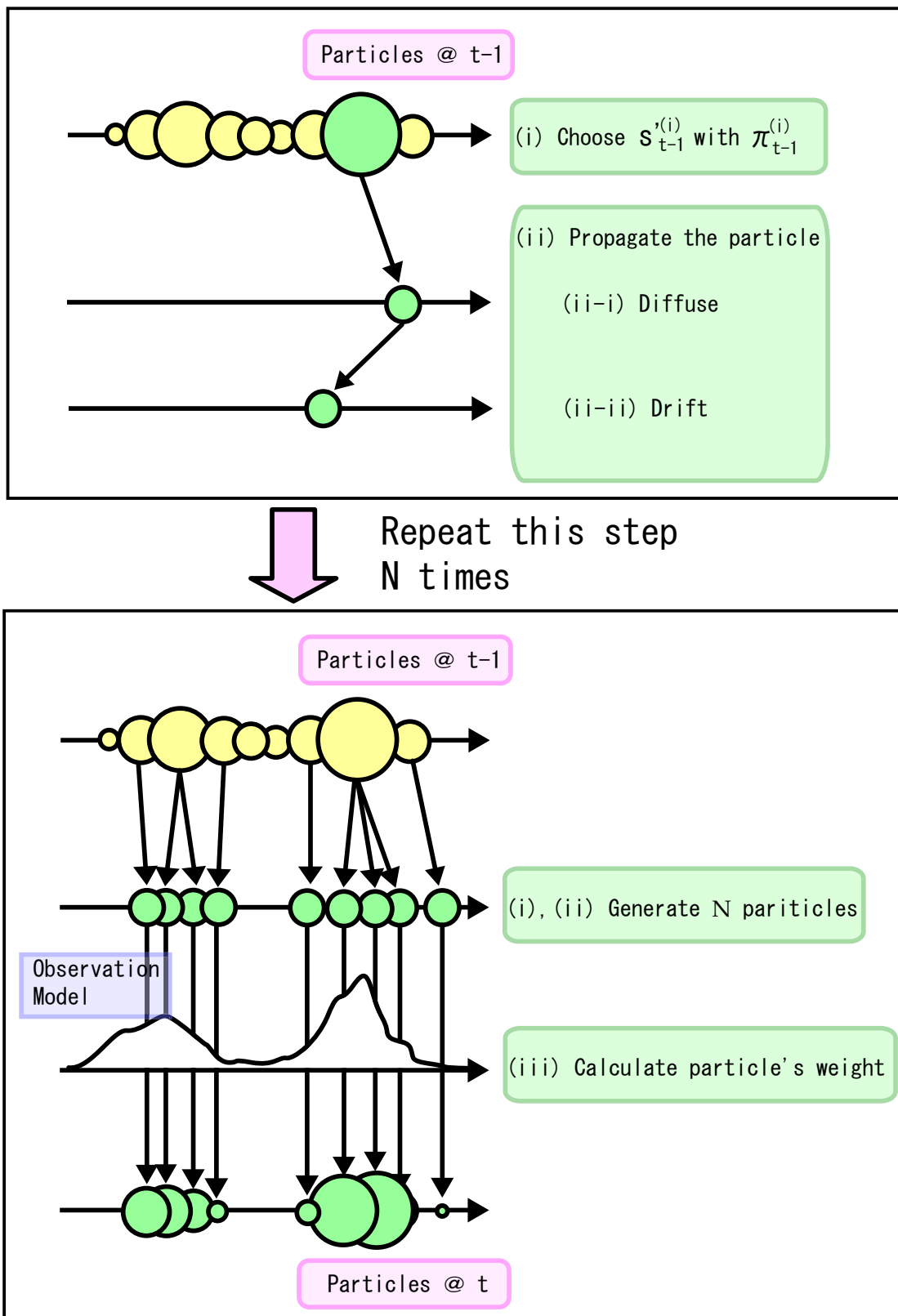


Fig. 2.2: パーティクルフィルタの処理手順

されていると考えられる。その配置とカメラの位置関係によって、人物全体、もしくは一部が遮蔽されることがある。また、観測視野内に複数の人物がいる場合、追跡対象とは別の人物によって遮蔽されることも考えられる。これにより、追跡対象の正しい確率密度分布が得られない可能性がある。

- 照明条件の変化

屋外環境において、朝、昼、夜といった時間帯や天候などに伴い、日照状況の変化が起きる。また、屋内では、ライトのスイッチのオン・オフにより急激に観測視野内の照明条件は変化する。このような照明条件の変化が起きると、観測される画像の特徴もまた大きく変化するため、正しい分布が得られないことがある。

- 低解像度

広い店舗内を考えた場合、監視カメラは天井に設置され、広範囲の視野を持つことが多い。このような場合、観測視野内の人物が非常に低解像度で映ることが考えられ、そのような画像から評価に必要な特徴量を取り出すことは難しい。

- 複雑背景

店舗内を対象シーンと考えると、そこには様々な商品が陳列されているため、観測される画像は複雑なテクスチャを持つことが多い。そのため、人物と背景領域との切り分けが難しくなり、正しく分布を推定することを困難にする。

そのため、このような状況下においてパーティクルフィルタを用いた人物追跡を行う場合、確率密度分布の近似精度を維持するために多数の仮説を必要とする。しかしながら、それに伴い多数の仮説を処理することになるため、計算コストの問題により実時間処理を維持することが難しくなる。逆に実時間処理を維持するために仮説の数を制限すると、状態空間内の仮説の密度が不足するため、追跡精度の低下を引き起こすことになる。

このような問題に対処するために、主に二つのアプローチがとられている。一つは、正確な確率密度分布を推定するために観測モデルを高精度化するアプローチ、二つ目は仮説数を制限した場合においても追跡精度の低下を抑えることができるように、サンプリングを効率化するアプローチである。次節より、それぞれについて関連する研究についてまとめる。

2.2 観測モデルの高精度化

人物追跡に関する研究の多くは観測モデルを高精度化するものが多く、様々な手法が提案されている。以下、主な手法をまとめる。

2.2.1 複数情報の統合

一つの特徴量しか利用しない観測モデルでは、前述した問題から確率密度分布が正しく得られず、追跡精度は低下してしまう。そこで、精度向上のために、人物らしさの評価に複数の情報を用いる手法が提案されている。

複数の画像特徴量の利用

観測画像から得られる特徴量のうち人物追跡によく用いられるものは、色分布情報、エッジ情報、背景差分による前景領域などが挙げられる。これらの特徴量は、上述した問題点の一部について頑健であるが、全ての問題に対処できるわけではない。例えば、エッジ情報を利用した評価では、人物頭部を追跡対象とした場合、画像に映る頭部のアピランスモデルを楕円とし、エッジと楕円の法線ベクトルの内積をとることにより評価する手法がよくとられる [19]。しかしながら、床や壁といった背景の色によって人物頭部の輪郭が明確でない場合や、画像中で十分な解像度を持った人物頭部領域が得られない場合に楕円モデルを適用することが難しいといった問題がある。

このように、実環境において上述したような単一の特徴量を手がかりに追跡をする場合、頑健な追跡を実現することは難しい。

そこで、画像から得られる複数の手がかりを統合することでより高精度な評価を実現する手法が提案されている [18, 21, 26, 29, 30, 35]。利用する特徴量それぞれに対して評価器を定義し、それぞれ得られる評価値を統合することにより、最終的なその時刻の人物らしさの評価値とする、すなわち、

$$\pi_t^{(i)} = \prod_j \pi_{j,t}^{(i)}. \quad (2.2.1)$$

ここで $\pi_t^{(i)}$ は時刻 t における i 番目の仮説の重みを表し、 $\pi_{j,t}^{(i)}$ は i 番目の仮説の j 番目の評価器による重みの評価値を意味する。これにより、より安定な追跡が実現される。

複数台のカメラの利用

単一視点のカメラによる追跡では、遮蔽の問題に対処できない。また、人物を三次元空間中で追跡する場合、人物らしさを二次元画像を用いて評価するため、奥行き方向に曖昧さが残るといった問題がある。これらの問題に対処すべく、視野を共有する複数視点のカメラ群を用いて追跡する手法が提案されている [15, 19, 20, 30, 34]。具体的には、仮説を各カメラ画像平面に投影し、各カメラで観測される画像を用いて評価する。そしてそれらを統合することで最終的な重みの評価とする、すなわち、

$$\pi_t^{(i)} = \prod_{cam=1}^M \pi_{cam,t}^{(i)}. \quad (2.2.2)$$

M は用いるカメラの台数、 $\pi_{cam,t}^{(i)}$ は cam 番目のカメラでの時刻 t , i 番目の仮説の重みの評価値を意味する。

複数台のカメラを用いたシステムでは、計算量の面から1台のPCでの処理に限界がある。そこで文献 [19, 20, 30] では、統合処理のためのサーバPCと、カメラ1台に対して1台のクライアントPCを割り当てた分散システムを構築している。

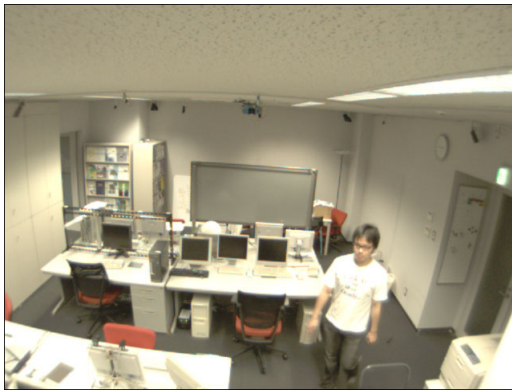
サーバは全てのクライアントに予測された仮説群を一斉送信する。仮説群を受信した各クライアントPCは、それぞれ画像平面への投影、人物らしさの評価を行い、その評価値をサーバに送り返す。全てのクライアントからの評価値を受信した後、サーバは式 (2.2.2) に従い評価値を統合し、最終的な仮説の重みとする。

これにより、計算コストのかかる人物らしさの評価の処理を分散させることができるため、高速な処理が可能となる。

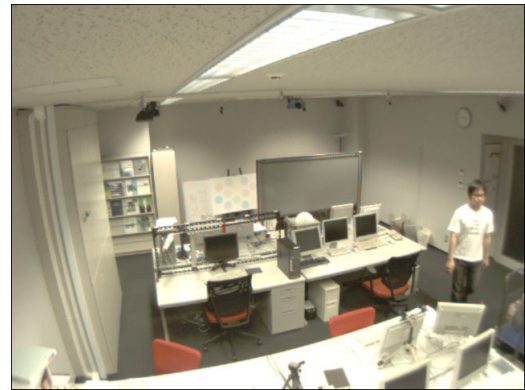
2.2.2 複数情報の階層的評価

複数情報を統合することで高精度化を図っている手法の多くでは、全ての仮説を複数の評価器で評価するため、用いる仮説の数や評価のアルゴリズム次第で計算コストが大きくなることがある。そのため、リアルタイム性を保つために、一つ一つの評価器は比較的高速かつ単純な評価手法が用いられることが多く、必ずしも頑健な追跡を実現できるとは限らない。

これに対し、Yangらは、高速かつ頑健な追跡を実現するために、複数の評価器を階層的に用いて評価する手法を提案している [35]。Coarse-to-Fine カスケードの考え方により、各指標の評価器をカスケード型に接続して段階的に評価する。明らかに評価



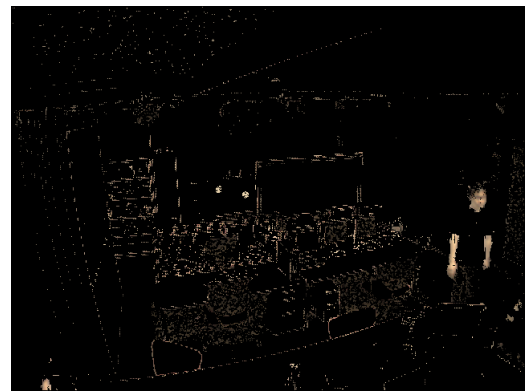
(i) 観測画像 (Camera 1)



(ii) 観測画像 (Camera2)



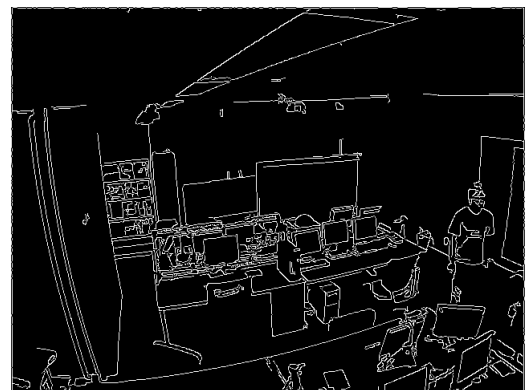
(iii) 肌色情報 (Camera 1)



(iv) 肌色情報 (Camera2)



(v) エッジ情報 (Camera 1)



(vi) エッジ情報 (Camera2)

Fig. 2.3: 観測画像と画像特徴量

値が小さくなるような仮説は、初期の単純かつ高速な評価器により棄却され、尤もらしい仮説は後段のより高精度な評価器によって評価される。このような段階的処理により、後段の高精度な評価をすべき仮説の数を少なくすることができるため、結果的に高速な処理が実現される。

2.2.3 顔識別器の利用

2.2.1 節で述べた比較的単純な画像特徴を利用した追跡手法は、実環境において必ずしも十分ではないことがある。そこで Kobayashi らは、高速かつ高精度なカスケード型 Adaboost を用いた顔識別器 [33] に着目し、仮説の人物頭部らしさの評価にこの識別器を利用した手法を提案している [15, 16]。人物頭部のアピアランスを楕円でモデル化し、観測される画像中の対象領域に顔識別器を適用する。そして、この識別器のカスケードを通過した段数を人物らしさ評価値とする。これにより、追跡における様々な問題に対処できる頑健な追跡が実現される。また、人物頭部の向きの変動による見えの変化に対応するために、人物頭部の向き毎に学習を行った顔識別器をそれぞれ用意し、毎時推定される顔の向きに伴い、識別器を選択的に利用している。これにより、視野を共有した複数のカメラにより様々な方向で観察される人物頭部に対し、人物の向きに適した顔識別器が選択されるため、安定した評価を行うことができる。

2.2.4 環境属性情報の利用

追跡対象の確率密度分布が正確に得られない場合、パーティクルフィルタの性質上、サンプリングした仮説が壁の中や床の下など、人物が存在しえない場所に発生することがある。このような領域に発生する仮説は、人物追跡において意味のない情報に他ならず、出来る限りこのような位置に仮説が発生することを防ぐ必要がある。

そこで鈴木らは、画像情報以外に、レンジセンサによって得られた室内の三次元形状から空間的な人物の存在可能性を定義し、それを追跡に利用することで追跡の安定化を図っている [30]。彼らは、椅子、棚、机などの障害物の位置関係から対象空間をいくつかの領域に分割し、それぞれの領域に人物の存在可能性に基づく重みを設定している (Fig.2.7)。この環境による重みを、観測画像から得られる手がかりに統合することで、仮説の無駄打ちを抑制した安定な追跡を実現する。これは、正確な確率密

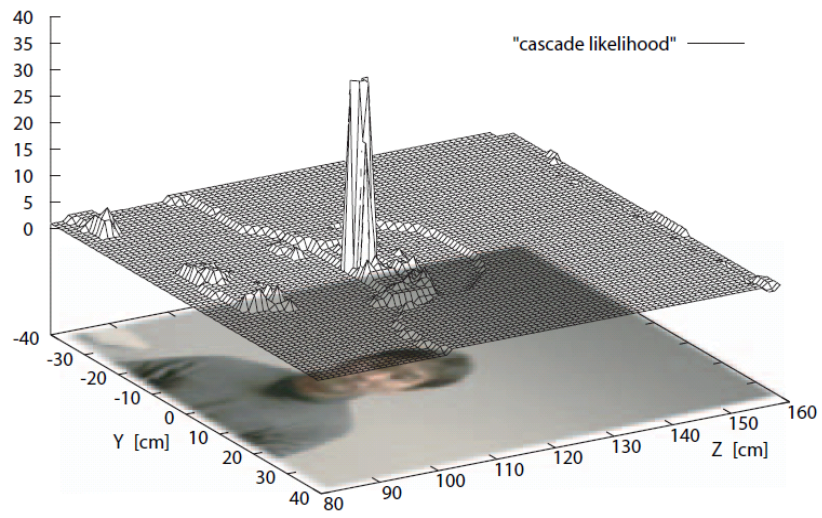


Fig. 2.4: カスケード型顔識別器による人物頭部らしさの評価分布 [15, 16]

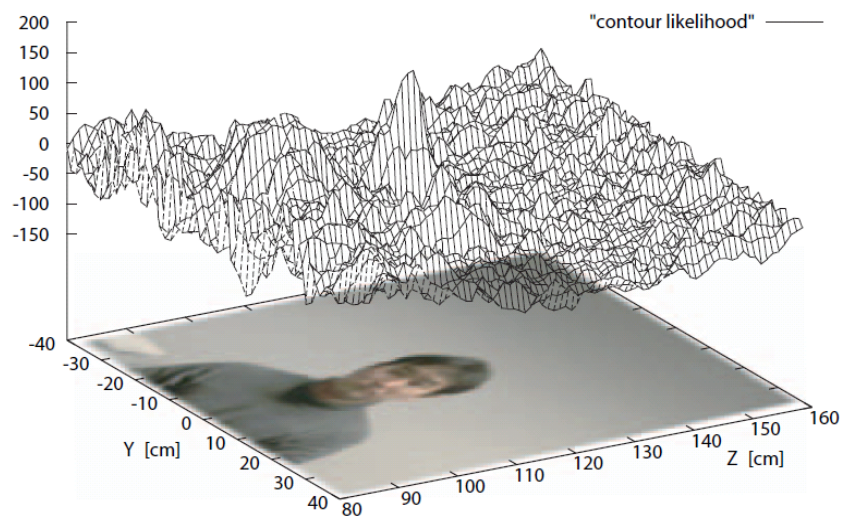


Fig. 2.5: エッジ情報による人物頭部らしさの評価分布 [15, 16]

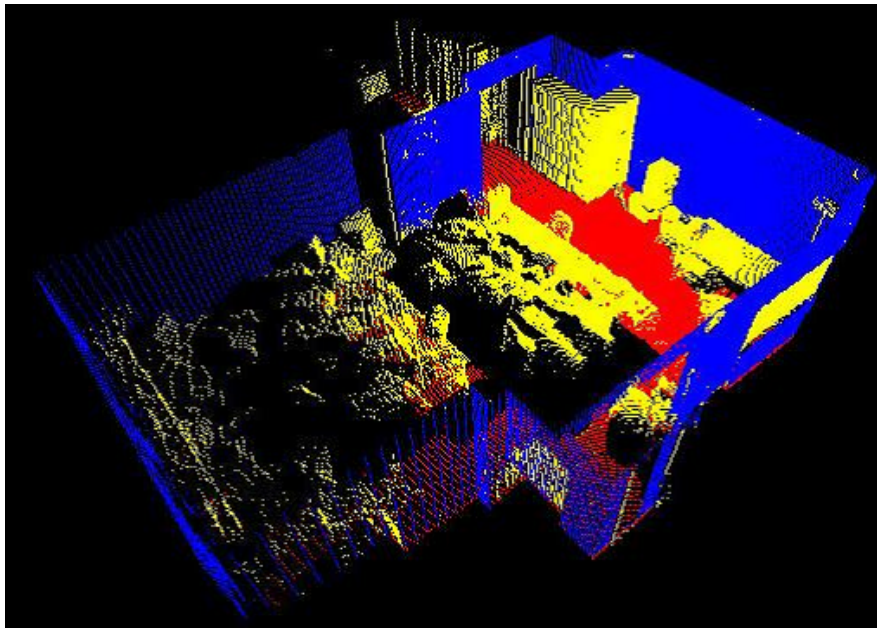


Fig. 2.6: レンジセンサにより得られた室内形状 [30]

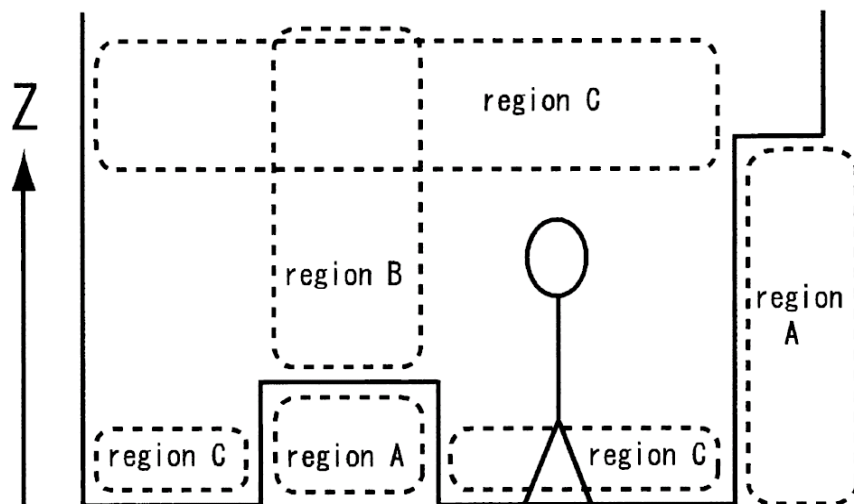


Fig. 2.7: 物理的環境属性情報 [30]

度分布を推定すること, それと後述する効率的にサンプリングすることの両者を高度化する手法であるといえる.

2.3 サンプリングの効率化

対象追跡において効率的な仮説のサンプリングをするために, 様々なサンプリング手法をパーティクルフィルタに組み入れた手法がいくつか提案されている. 本節では, Importance Sampling と呼ばれるサンプリング手法をパーティクルフィルタの枠組みに組み入れた ICONDENSATION [12] と, 検出器と追跡器を明示的に区別し, 主に追跡の初期化, 再初期化に重点を置いた手法 [13, 22, 34] について述べる.

2.3.1 ICONDENSATION

パーティクルフィルタによる追跡の際, 2.1.3 節で挙げた問題に起因して, 追跡が不安定, もしくは失敗することが考えられる. ICONDENSATION は, そのようなときに, Importance Sampling によって不安定な追跡を修正, もしくは再初期化することで安定な追跡を実現する手法である.

Importance Sampling とは, 事前分布 $p(\mathbf{X}_t | \mathcal{Z}_{t-1})$ が複雑な形状を持つ場合や, 正確に得られていない場合に, 事後分布 $p(\mathbf{X}_t | \mathcal{Z}_t)$ について重要な情報を含んでいると考えられる別の情報源 (Importance Function と呼ばれる) から仮説をサンプリングすることで, サンプリングを効率化する手法である. これをパーティクルフィルタの枠組みに適用することで, 追跡性能を高めることができる.

具体的には, ICONDENSATION は以下の三つの機能を持つ. そのために, 事前分布 $p(\mathbf{X}_t | \mathcal{Z}_{t-1})$ と Importance Function 両方の分布から, それぞれある割合で仮説をサンプリングする.

(a) 追跡の初期化

追跡失敗の際の再初期化, もしくは新たに視野内に入ってきた人物の追跡の初期化のために, Importance Function から仮説をサンプリングする. サンプリングされた仮説は, パーティクルフィルタの観測モデルにより人物らしさを評価される.

(b) 追跡の修正

追跡を, Importance Function から仮説をサンプリングすることにより修正する. 具体的には, Importance Function $g_t(\mathbf{X}_t)$ により, 観測モデルによる評価は次のように修正される.

$$\pi_t^{(i)} \sim \frac{p(\mathbf{X}_t = \mathbf{s}_t^{(i)} | \mathcal{Z}_{t-1})}{g_t(\mathbf{s}_t^{(i)})} p(\mathbf{Z}_t | \mathbf{X}_t = \mathbf{s}_t^{(i)}). \quad (2.3.1)$$

ここで, 右辺の $p(\mathbf{X}_t = \mathbf{s}_t^{(i)} | \mathcal{Z}_{t-1})/g_t(\mathbf{s}_t^{(i)})$ は, Importance Sampling による追跡の修正を意味する.

(c) パーティクルフィルタによる追跡

従来のパーティクルフィルタによる追跡を行う. 事前分布から仮説をサンプリングし, 観測モデルによって評価される処理を逐次的に行う.

Importance Function には任意の形状の確率密度分布を定義できるが, 仮説をサンプリングしやすい分布でなければならない. 主な理由は, リアルタイム性を保つためである. 過去の事後分布を元に事前分布を予測し, 仮説をサンプリングする時系列フィルタリングとは異なり, Importance Sampling は観測された情報から仮説をサンプリングする. そのため, 仮説のサンプリング次第ではリアルタイム性を保てなくなることがある. よって多くの手法では, 正規分布, もしくは混合正規分布が用いられることが多い.

2.3.2 追跡器と検出器の併用

ICONDENSATION では, Importance Function として定義した確率分布と, パーティクルフィルタの状態遷移モデルにより構築される事前分布の適合度により, 追跡を修正することが主な寄与であった. それに対し本節では, パーティクルフィルタによる追跡器と, 主に追跡の初期化を目的とした対象の検出, 再検出を行うための検出器を併用した手法について述べる. ICONDENSATION の枠組みにおいても同様の機能を実現できるが, 本節で述べるいくつかの手法は, 追跡の安定化のために陽に用いてはいないという点で区別できる. また, このように追跡器と検出器を区別する利点は, 複数対象追跡への拡張が容易, かつその方法が理論的に整理されている [32] 点である.

頭部検出器の利用

Jin らは、パーティクルフィルタによる追跡と同時に、背景差分、エッジ検出を組み合わせた独自の人物頭部検出器を用い、視野内に観測される人物頭部の検出処理を行っている [13]. そしてパーティクルフィルタの事前分布からのサンプリングとこの検出位置からのサンプリングをある割合で併用することで、検出結果を追跡に反映している。また、頭部検出結果に伴い、事前分布と頭部検出器からのサンプリングの割合を動的に変更している。これにより、観測シーンに応じた適応的な仮説のサンプリングを行うことができる。

複数台のカメラによる観測の利用

Wang らは、複数台のカメラを用い、各カメラの観測を人物検出器として利用している [34]. つまり、検出器による仮説のサンプリングは各カメラの観測にのみ依存し、いずれかのカメラにより観測されると、そのカメラの観測から仮説がサンプリングされる。また、仮説の重みの評価は、複数カメラによる評価の積をとるのではなく、各カメラの観測モデルによる評価値のうち最大となるものが用いられる。これにより、少なくとも一台のカメラで人物を観測することができれば追跡を実現できるため、遮蔽の問題に対して頑健な手法である。

Boosted Particle Filter

Okuma らは、検出器にカスケード型 Adaboost 検出器 [33] を利用し、追跡の初期化、再初期化のための、高速かつ高精度な人物検出を実現している [22]. 彼らは複数対象追跡を目的としており、複数対象追跡において有効な手法である Mixture Particle Filter [32] の枠組みにこの検出器を組み入れている。

彼らの実験では、アイスホッケーのゲーム中の複数の選手を追跡することを目的としているが、リンク脇の観客が検出器により過検出されてしまうことがあった。そこで、追跡対象の動きを考慮することにより過検出を減らすことができるとし、Mixture Particle Filter の状態遷移モデルと検出器による分布の混合分布を用いる、すなわち、

$$q_B^*(\mathbf{X}_t | \mathbf{X}_{t-1}, \mathcal{Z}_t) = \alpha q_{ada}(\mathbf{X}_t | \mathbf{X}_{t-1}, \mathbf{Z}_t) + (1 - \alpha)p(\mathbf{X}_t | \mathbf{X}_{t-1}). \quad (2.3.2)$$

ここで $q_{ada}(\mathbf{X}_t | \mathbf{X}_{t-1}, \mathbf{Z}_t)$ は時刻 t の観測に基づく Adaboost 検出器による分布で、ここでは正規分布としている。右辺第二項の分布、 $p(\mathbf{X}_t | \mathbf{X}_{t-1})$ は時刻 $t - 1$ の状態に基

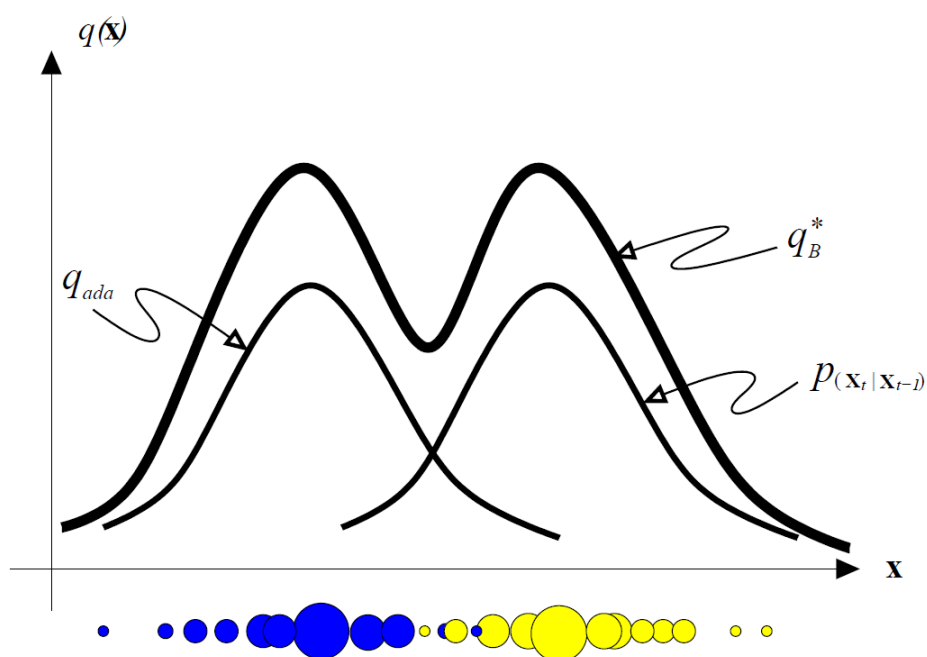


Fig. 2.8: 追跡器と検出器の混合分布 [22]

づく状態遷移モデルである。また、 α は Mixture Particle Filter と検出器の混合度を表すパラメータである。すなわち、もし $\alpha = 0$ ならば、それは Mixture Particle Filter と等価となることを意味する。Fig.2.8 にその概念図を示す。

もし予測される Mixture Particle Filter のうちのいずれかの要素の分布が Adaboost 検出器による分布と重なっているならば、混合分布 q_B^* からサンプリングする。重なっていないならば、 $\alpha = 0$ とし、Mixture Particle Filter の状態遷移モデルに従う。これにより、検出器による過検出を抑えることができ、リンク内の不定数の選手の頑健な追跡を可能とする。

2.4 行動履歴の利用による環境情報の獲得

文献 [30] では、机や椅子、棚の配置などの対象空間の物理的な形状から人物の存在可能性に基づく重みとして環境属性情報を定義し、追跡の安定化に利用している。その一方で、対象空間における人物の行動履歴を反映することで得られる、シーンに特化した適応的な環境属性情報が考えられる。

人物の行動履歴を反映させることにより環境情報を適応化する研究事例は、追跡の分野ではあまり多くないが、環境に意味付けをする Location-aware Computing への応用を目的とした研究について、関連するものがいくつかある。そこで本節では、カメラ画像を用いた人物追跡を利用している Activity Maps [5] について述べる。

Activity Maps

対象空間内の位置情報は多くのアプリケーションにおいて重要な情報である。例えば、机やディスプレイなどの位置情報とユーザの位置情報をシステムが理解していれば、ユーザの位置に伴い自動的にディスプレイをオンにしたり、自動的に照明を点灯する、といった処理を可能にする。このように、位置情報はユーザの環境をより高度化するために必要な情報である。

しかしながら、位置情報単体ではコンテキストとして不十分であり、それが何なのか、どのような処理がされるべきなのか、といったことを全て決定することはできない問題がある。また、その一方で、環境内におけるユーザの行動は位置情報と密接な関係があると知られている。彼らはそれに着目し、Activity Map と呼ばれる、対象シーンにおいて人物がとりやすい特定の行動で意味付けされた位置情報を用いることで、コンテキストを自動的に獲得できるとしている。

Activity Map の構築方法は次の通りである。

- Step 0*: 予め長時間対象シーンを観測し、人物の行動履歴(追跡データ)を獲得しておく。このようにして蓄積された追跡データから、人物の頭部位置(床からの高さ)、速度を計算し、これらを観測による特徴量 $f_i(x, y) = (h, v, v_h)$ とする。ここで v_h は平均速度を意味する。
- Step 1*: $f_i(x, y)$ の集合を、パラメータ空間において K-means クラスタリングをすることで N 個のクラスに分類し、それぞれのクラスにおいて平均特徴量 F_k を計算する。この平均特徴量に伴い、それぞれのクラスに特定の行動(歩いている、作業している、休んでいる、など)をラベル付けする。
- Step 2*: 同様の行動を持つ(同じクラスに属する)位置 (x, y) をグルーピングすることで、連結領域 R_k を推定する。このうち、ある程度大きな連結領域を Activity Zone, Z_k とする。この Activity Zone のうち、特徴量 $f_i(x, y)$ の数が少ない、すなわち頻繁にはとられない行動によりグルーピングされた Activity Zone を除外する。この操作の後に残った Activity Zone を Activity Map と定義する。

このような処理により Activity Maps は得られる. Fig.2.9 に一例を示す. Fig.2.9 の例では, あるオフィスで二人の人物が行動するシーンを考えており, 対象空間は5つの Activity Zone に分割されている. それぞれ番号が振られている領域には, 各々頻繁に観測される人物の行動のラベルが割り当てられている. 例えば, Zone1 (通路領域) では“ 歩く ”というラベル, Zone2, 3 (机がある作業領域) では“ 作業をする ”というラベルがつけられており, 対象空間の物理的な情報とも符合していることがわかる.

主な利点としては, 人物の位置推定にカメラ画像を用いた人物追跡器を利用しているため, 位置計測に特殊な装置を身に着ける必要がない. また, 行動の履歴から行動特徴を抽出, グループングしてコンテキストを獲得するため, プリンタやディスプレイなどの特定の装置の位置をシステムに教える必要がない. そして, 環境内のレイアウトや人数など, 対象とするシーンに特化したコンテキストを得ることができる.

2.5 本研究の位置付け

2.1 節では, 人物追跡において非常に効果的な手法である, パーティクルフィルタについて説明し, 主に, “ 観測モデルの高精度化 ”と“ サンプルングの効率化 ”の二点が, 頑健な人物追跡を実現するための重要な要素であることを述べた. そして2.2, 2.3 節では, 関連研究をそれぞれの要素に基づき分類し, 説明した. いずれにおいても, 多くの研究は観測される画像から得られる指標を元にしており, 他のセンサ情報や三次元情報を併用することで追跡を高度化する手法はまだ少ない.

本研究の先行研究にあたる [30] は, 環境属性という対象空間の三次元的な幾何情報に基づいた人物存在可能性による重みを追跡に応用しており, それは追跡の安定化に有効であるといえる. しかしながら, そのような物理的な情報の利用には, いくつか問題点がある. 一つは, 対象空間の物理的な形状, すなわち机や棚などの障害物の形状が既知であること, 二つに, 障害物の移動などによる環境変動に弱いことが挙げられる.

その一方で, 行動履歴を反映することで獲得できる対象シーンに特化した環境属性が考えられる. 2.4 節で述べたように, 環境情報は対象空間内の人物の行動に強く関係する. そのため, シーン内の人物の行動履歴を環境に反映することは, その対象シーンに特化した環境属性情報の獲得に非常に有効な方法であるといえる. これにより獲得される環境情報は, 事前知識無しに容易に獲得することができるため, これを利用した様々な応用が考えられる.

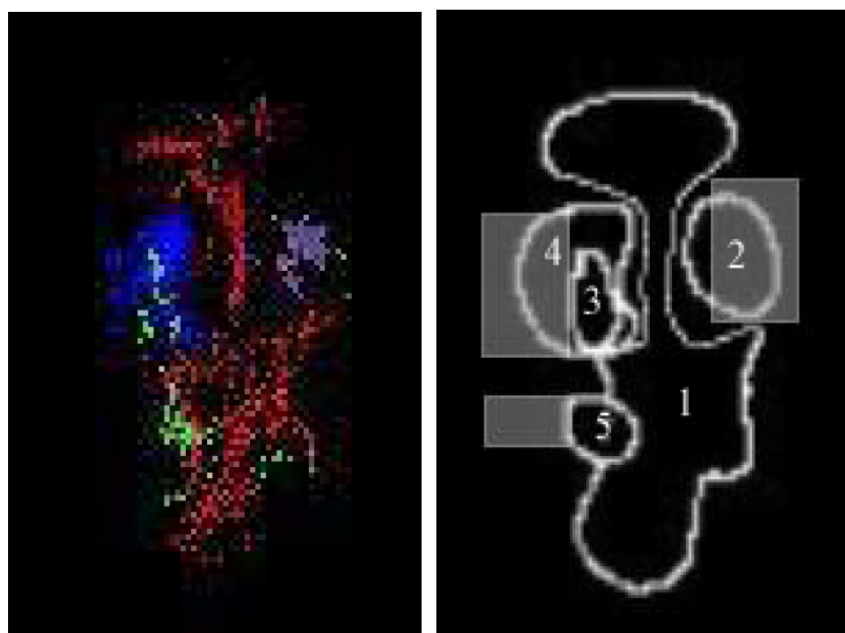


Fig. 2.9: Activity Map [5]

このような、追跡の安定化のための三次元情報の利用、人物の行動履歴の利用による適応的な環境情報の獲得、という二つの知見に基づき、本研究では、行動履歴を用いた環境属性のオンライン学習と人物追跡の両者を同時に実行することで、適応的な環境属性の獲得と追跡の安定化の両方を実現する手法を提案する。

第3章 行動履歴の反映による適応的環境属性を伴う三次元人物追跡

本章では, 人物の行動履歴の反映による環境属性のオンライン学習と人物頭部追跡が相互に作用することで, 対象シーンに特化した適応的環境属性の獲得, そしてそれに伴う追跡の安定化を実現する本手法について詳しく説明する. まず次節 3.1 節では本提案手法の全体像について説明すると共に, その新規点, 利点などについて述べる. 3.2, 3.3 節では, 3.1 節で述べる本システムの具体的な実現方法についてそれぞれ詳細に述べる.

3.1 環境属性のオンライン学習と人物頭部追跡の同時実行

Fig.3.1 に本手法で提案するシステムの全体図を示す. この図にあるように, 本手法は大きく次の二つの部分に分けることができる.

1. 人物頭部の三次元追跡

対象視野内において, 複数の人物の頭部を Mixture Particle Filter [32] の考え方に基づいて三次元的に追跡する. また個々の対象追跡には ICONDENSATION [12] の考え方を利用し, 適応的環境属性情報を Importance Function として定義する. これにより安定な人物追跡を実現する.

2. 適応的環境属性の獲得

対象空間内における人物の存在確率を表す環境属性を獲得する. 空間内を移動する人物を追跡することにより推定される動線より, その人物の行動状態を推定する. そして, 推定された行動状態に伴い, 人物頭部位置をオンライン EM アルゴリズム [25] により逐次的に環境属性へ反映させる. 長時間の観測により, 対象とするシーンに特化した適応的環境属性を得ることができる.

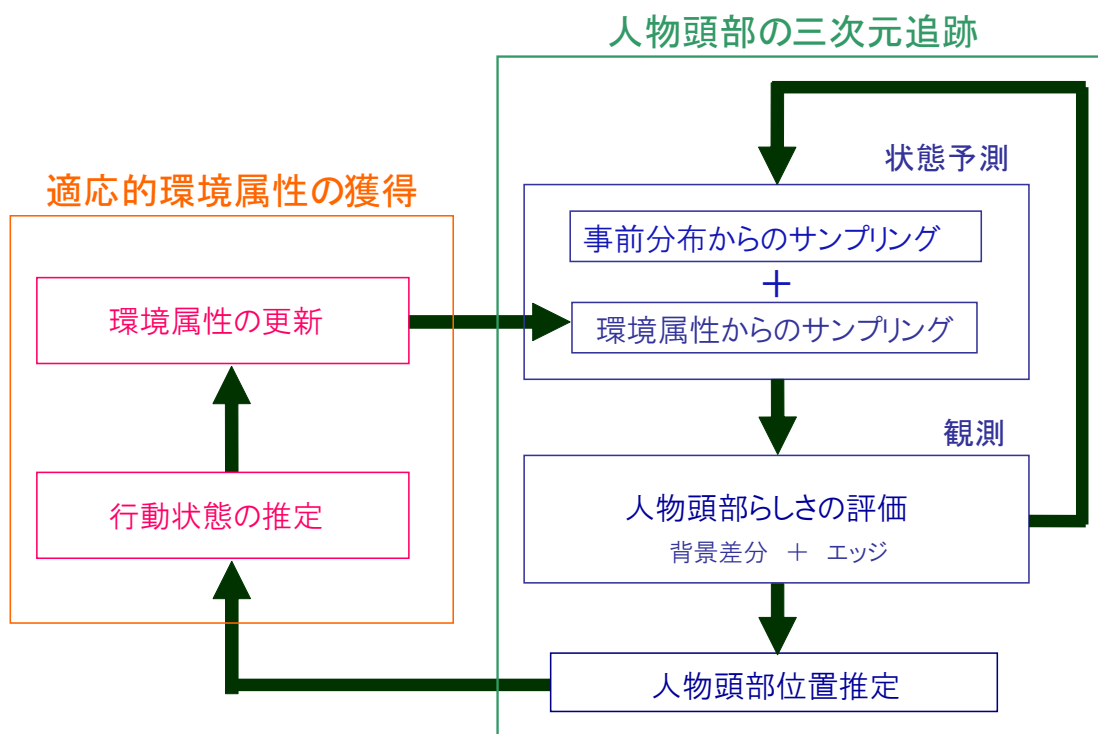


Fig. 3.1: システム全体図

このような二つの機構の相互作用により本手法は実現される。また、本手法の大きな特徴は次の3点である。

- 三次元情報の利用による追跡の安定化
- 事前情報無しに環境情報の獲得が可能
- 人物追跡の安定化と適応的環境属性の獲得の同時実現

次節以降で、その詳しい実現方法について述べる。

3.2 人物頭部の三次元追跡

視野を共有する複数台のカメラを用いて、三次元位置を状態量に持つ人物頭部追跡を行う。また、Mixture Particle Filter [32] の考え方にに基づき、複数人物追跡を対象とする。個々の人物頭部を追跡するアルゴリズムは、2.3.1 節で述べた ICONDENSATION アルゴリズム [12] を利用し、Importance Function として環境属性を用いる。人物頭部らしさの評価手法は、2.2.1 節で述べた複数の評価器による評価を統合する手法をとることとし、背景差分による評価とエッジ情報による評価を利用する。そして複数台のカメラ各々から得られる評価を統合することにより、その時刻における最終的な評価とする。このような手法により、全ての仮説を評価することで事後分布を推定する。そしてこの推定される事後分布における状態量の期待値を人物頭部位置推定とする。また、計算コストの削減と分散環境に対応するために、1 台のサーバ用 PC と複数台のクライアント PC によるサーバ・クライアントモデルを構築する [19, 20, 30]。

3.2.1 人物頭部モデル

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 $O-XYZ$ を定義する。また、追跡対象である人物頭部を、頭部形状は不変であることから、アスペクト比を一定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。すなわち、パーティクルフィルタにおける n 番目の仮説は、 $s_t^{(n)} = [x_t^{(n)}, y_t^{(n)}, z_t^{(n)}]^T$ という三次元ベクトルの状態量を持つ楕円体 $\Gamma_t^{(n)}$ として表現される。このような状態量とそれに対応する重みを持つ仮説群 $\{s_t^{(n)}; \pi_t^{(n)}\}$ を用い、状態推定、評価を逐次的に行うことで人物追跡を実現する。

3.2.2 仮説のサンプリング

ICONDENSATION の考え方にに基づき、追跡対象の事前分布 $p(\mathbf{X}_t | \mathcal{Z}_{t-1})$ からのサンプリングと、Importance Function, すなわち環境属性 $g_t(\mathbf{X}_t)$ からのサンプリングを、ある割合で用いる。また、環境属性からのサンプリングは、“追跡の修正”と“追跡の初期化”の二用途を考える。

時刻 t , n 番目の仮説 $s_t^{(n)}$ をサンプリングするために、乱数 $\alpha \in [0, 1)$ を発生させる。この値と予め定めた割合, q, r に伴いサンプリング方法が選択される。そして、以下のそれぞれの場合の手順に従い、時刻 t , n 番目の仮説 $s_t^{(n)}$ がサンプリングされ

る. この処理を全ての仮説に対して行うことにより, 時刻 t の状態を推定する仮説群 $\{s_t^{(1)}, \dots, s_t^{(N)}\}$ が得られる. また, Importance Function である環境属性 $g_t(\mathbf{X}_t)$ の具体的な形については, 3.2 節で詳しく述べる.

(a) $\alpha < q$ の場合

“ 追跡の初期化 ” 用として, 仮説 $s_t^{(n)}$ を環境属性 $g_t(\mathbf{X}_t)$ からサンプリングする.

(b) $q \leq \alpha < q + r$ の場合

“ 追跡修正 ” 用として, 仮説 $s_t^{(n)}$ を環境属性 $g_t(\mathbf{X}_t)$ からサンプリングする.

(c) $q + r \leq \alpha$ の場合

パーティクルフィルタの予測ステップに従う. 以下の手順により, 事前分布から仮説がサンプリングされる.

(1. 仮説の選択)

時刻 $t - 1$ の事後分布 $p(\mathbf{X}_{t-1} | \mathcal{Z}_{t-1})$, すなわち仮説群 $\{s_{t-1}^{(n)}; \pi_{t-1}^{(n)}\}$ の重みの比に従い, 仮説群 $\{s_{t-1}'^{(1)}, \dots, s_{t-1}'^{(N)}\}$ をサンプリングする.

(2. 仮説の移動・拡散)

予め定義した状態遷移モデル $p(\mathbf{X}_t | \mathbf{X}_{t-1})$ に従い, 仮説群 $\{s_{t-1}'^{(1)}, \dots, s_{t-1}'^{(N)}\}$ を伝播させる. 本手法では, 次の 1 次 AR モデルを用いる.

$$\mathbf{s}_t'^{(n)} = \mathbf{s}_{t-1}'^{(n)} + \boldsymbol{\omega}. \quad (3.2.1)$$

ここで, $\boldsymbol{\omega}$ は, 平均 0, 分散 σ^2 を持つガウス雑音である. また, 式 (3.2.1) 右辺第一項が“ 移動 ” にあたり, 第二項が“ 拡散 ” にあたる.

これにより, 事前分布,

$$p(\mathbf{X}_t | \mathcal{Z}_{t-1}) = \sum_{n=1}^N \pi_{t-1}^{(n)} p(\mathbf{X}_t | \mathbf{X}_{t-1} = \mathbf{s}_{t-1}'^{(n)}), \quad (3.2.2)$$

が構築される. そして, この事前分布より, 仮説 $s_t^{(n)}$ をサンプリングする.

このようにして得られた仮説群 $\{s_t^{(1)}, \dots, s_t^{(N)}\}$ は、各カメラの観測画像により人物頭部らしさが評価される。推定される仮説群は三次元世界座標系にあるため、仮説はカメラ画像平面にどのように投影されるのか、すなわち世界座標系と各カメラの画像座標系との対応関係を予め知っておく必要がある。

3.2.3 仮説の画像平面への投影

時刻 t における n 番目の仮説を表す楕円体 $\Gamma_t^{(n)}$ を、 i 番目のカメラ画像平面へ投影したときに得られる楕円領域を $\Omega_{i,t}^{(n)}$ とする。世界座標系と i 番目のカメラ画像座標系との対応関係を、投影関数 $\mathcal{F}_i(\cdot)$ で表現したとき、仮説の投影は次式のように表現される。また、投影された人物頭部の中心位置を $p_{i,t}^{(n)}$ とする。

$$\Omega_{i,t}^{(n)} = \mathcal{F}_i(\Gamma_t^{(n)}) . \quad (3.2.3)$$

画像平面に投影された人物頭部の大きさは、次のようにして求められる。

三次元空間中の人物頭部の大きさを、簡単のために、個人差による変動を考えず、ある決まった値を用いる。これにより、楕円体の中心座標が決まることで頭部の端点の位置も決定される。そして、これら得られた端点の座標を同様に画像平面へ投影することで、画像平面に投影された人物頭部の横幅 $w_{i,t}^{(n)}$ と縦幅 $h_{i,t}^{(n)}$ を得ることができる。

3.2.4 観測による人物頭部らしさの評価

画像平面に投影された仮説の人物頭部らしさを観測画像を用いて評価する。本手法では背景差分をすることで得られる前景領域と、Sobel Filter をかけることにより得られるエッジ画像の両者を用いて評価する。そして、複数台のカメラによる評価を統合することで、仮説の最終的な評価値とする。

背景差分による評価

i 番目のカメラの、背景差分により獲得した前景画像から、仮説 $s_t^{(n)}$ の背景差分による重み $\pi_{i,t}^{fore,(n)}$ を求める。

得られる i 番目のカメラの前景画像において、楕円 $\Omega_{i,t}^{(n)}$ 内部の全画素数を $|\Omega_{i,t}^{(n)}|$ 、楕円 $\Omega_{i,t}^{(n)}$ 内部の前景領域に含まれる画素数を $f_{i,t}^{(n)}$ としたとき、重み $\pi_{i,t}^{fore,(n)}$ は次式で与

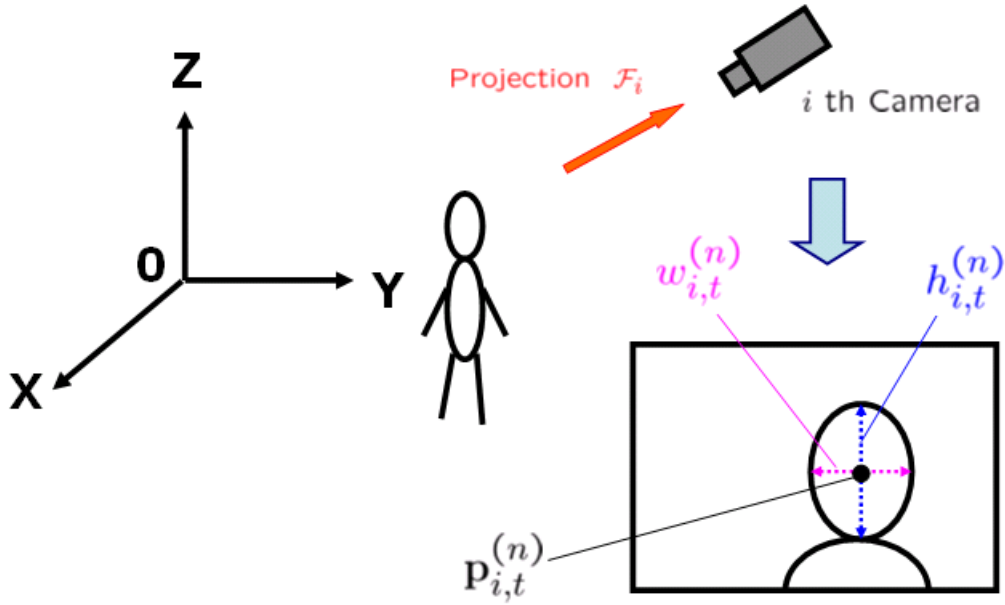


Fig. 3.2: 人物頭部の画像平面への投影

えられる。また、計算高速化のために、前景領域の画素数の計算には Integral Image [33] を用いる。

$$\pi_{i,t}^{fore,(n)} = \frac{f_{i,t}^{(n)}}{|\Omega_{i,t}^{(n)}|}. \quad (3.2.4)$$

これにより、前景領域である人物の領域に発生した仮説の重みが大きくなる。

エッジ情報による評価

i 番目のカメラから得られる観測画像 $I_{i,t}(x, y)$ に Sobel Filter をかけることにより、水平方向のエッジ画像 $D_{x_{i,t}}(x, y) \equiv \partial I_{i,t}(x, y) / \partial x$ と、垂直方向のエッジ画像 $D_{y_{i,t}}(x, y) \equiv \partial I_{i,t}(x, y) / \partial y$ を獲得する。これらの情報を用いて、仮説 $s_t^{(n)}$ のエッジ情報による重み $\pi_{i,t}^{edge,(n)}$ を求める。本手法では、楕円の輪郭位置の、エッジ勾配 $\mathbf{D} \equiv (D_{x_{i,t}}(x, y), D_{y_{i,t}}(x, y))^T$ と法線ベクトル $\mathbf{n} \equiv (n_x(x, y), n_y(x, y))^T$ の内積を評価する手法を用いる [19, 20].

i 番目のカメラの画像平面に投影された n 番目の仮説の位置 $\mathbf{p}_{i,t}^{(n)}$ と、人物頭部の横

幅 $w_{i,t}^{(n)}$, 縦幅 $h_{i,t}^{(n)}$ から, 楕円 $\Omega_{i,t}^{(n)}$ の輪郭位置は次式で与えられる.

$$x_{c_{i,t}}^{(n)}(\theta) = \frac{w_{i,t}^{(n)}}{2} \cos \theta, \quad (3.2.5)$$

$$y_{c_{i,t}}^{(n)}(\theta) = \frac{h_{i,t}^{(n)}}{2} \sin \theta. \quad (3.2.6)$$

ここで θ は画像座標系 (u, v) における u 方向を 0 [deg.] とし, 半時計回りの方向に増加する値をとる角度パラメータを意味する. また, 輪郭位置を $\mathbf{c}(\theta) \equiv (x_{c_{i,t}}^{(n)}(\theta), y_{c_{i,t}}^{(n)}(\theta))^T$ とする.

この輪郭位置において, 正規化されたエッジ勾配と法線ベクトルの内積の和を計算する. ただし, 人物の顔の部分はエッジが出にくいいため, 全輪郭位置については計算せず, 角度 $-30 \leq \theta \leq 210$ の範囲, つまり頭部の位置のみを考慮する (Fig.3.3). これより, 仮説 $s_t^{(n)}$ のエッジ情報による重み $\pi_{i,t}^{edge,(n)}$ は次式で与えられる.

$$\pi_{i,t}^{edge,(n)} = \sum_{-30 \leq \theta \leq 210} \frac{\mathbf{D}|_{\mathbf{c}(\theta)} \cdot \mathbf{n}|_{\mathbf{c}(\theta)}}{\|\mathbf{D}|_{\mathbf{c}(\theta)}\| \|\mathbf{n}|_{\mathbf{c}(\theta)}\|}. \quad (3.2.7)$$

これは, エッジ勾配の方向と法線ベクトルの方向が同一方向を示すときに最大となる. つまり, 楕円 $\Omega_{i,t}^{(n)}$ と観測されたエッジ画像の類似度を意味する.

評価の統合

背景差分による評価 $\pi_{i,t}^{fore,(n)}$ と, エッジ情報による評価 $\pi_{i,t}^{edge,(n)}$ を統合することにより, i 番目のカメラによる n 番目の仮説の人物らしさの重み $\pi_{i,t}^{(n)}$ とする, すなわち,

$$\pi_{i,t}^{(n)} = \pi_{i,t}^{fore,(n)} \pi_{i,t}^{edge,(n)}. \quad (3.2.8)$$

これにより, i 番目のカメラによる観測において, 前景領域周辺, かつ投影された人物頭部を表す楕円 $\Omega_{i,t}^{(n)}$ と類似する領域に発生した仮説の重みが大きくなる.

複数台カメラによる評価の統合

各カメラ毎に得られる人物頭部らしさを表す重み $\pi_{i,t}^{(n)}$ を統合することにより, 時刻 t の n 番目の仮説の最終的な人物頭部らしさの重み $\pi_t^{(n)}$ とする. 視野を共有するカメラの台数を M としたとき, 次式で表現される.

$$\pi_t^{(n)} = \prod_{i=1}^M \pi_{i,t}^{(n)}. \quad (3.2.9)$$

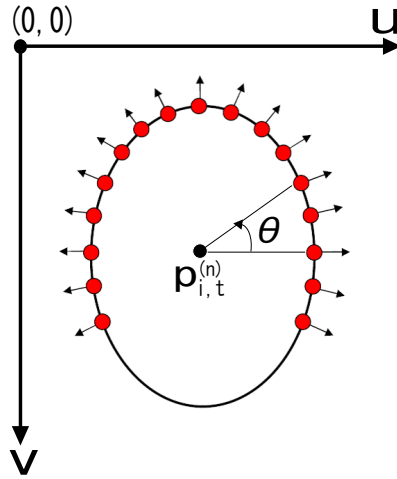


Fig. 3.3: 評価する楕円の輪郭位置と法線ベクトル

ただし、“追跡修正”のために環境属性 $g_t(\mathbf{X}_t)$ からサンプリングされた仮説の場合 (3.1.2節, (b)のサンプリング), 式 (3.2.9) は, 次のように修正される.

$$\pi_t^{(n)} = \frac{p(\mathbf{X}_t = \mathbf{s}_t^{(n)} | \mathcal{Z}_{t-1})}{g_t(\mathbf{s}_t^{(n)})} \prod_{i=1}^M \pi_{i,t}^{(n)}. \quad (3.2.10)$$

3.2.5 人物頭部位置推定

全ての仮説の評価を終えた後, 全仮説の重みを正規化する, すなわち,

$$\pi_t^{(n)} = \frac{\pi_t^{(n)}}{\sum_{n'=1}^N \pi_t^{(n')}}. \quad (3.2.11)$$

正規化された仮説の重みにより, 時刻 t の事後分布 $p(\mathbf{X}_t | \mathcal{Z}_t)$ は次のように与えられる.

$$p(\mathbf{X}_t | \mathcal{Z}_t) \approx \sum_{n=1}^N \pi_t^{(n)} \delta_{\mathbf{X}_t}(\mathbf{X}_t). \quad (3.2.12)$$

本手法では, 人物頭部推定位置はこの事後分布における期待値 \mathbf{s}_t^{EAP} とする, すなわち,

$$\mathbf{s}_t^{EAP} = \sum_{n=1}^N \pi_t^{(n)} \mathbf{s}_t^{(n)}. \quad (3.2.13)$$

このような手法で人物頭部位置を逐次推定することにより、追跡を実現する。また、この推定される頭部位置 s_t^{EAP} は、適応的環境属性の獲得に利用される。詳しくは3.3節にて述べる。

3.2.6 複数人追跡への拡張

複数対象追跡についてパーティクルフィルタを利用する場合、一つの重要な問題について考慮しなければならない。パーティクルフィルタは確率密度分布を多数の仮説により離散的に近似するため、複数対象追跡を想定する場合、追跡対象同士の遮蔽や視野内への出現、退去などにより追跡対象の分布推定が曖昧になってしまう問題がある。この問題に対処し、頑健な追跡を実現するために、近年様々な手法が提案されている [10, 20, 22, 31]。

Vermaak らは M 人の追跡対象全体の確率密度分布 $p(\mathbf{X}_t | \mathcal{Z}_t)$ を、互いに独立な追跡対象それぞれの確率密度分布 $p_m(\mathbf{X}_t | \mathcal{Z}_t)$ の重み付き混合分布、

$$p(\mathbf{X}_t | \mathcal{Z}_t) = \sum_{m=1}^M \Pi_{m,t} p_m(\mathbf{X}_t | \mathcal{Z}_t), \quad (3.2.14)$$

で表現することにより、視野内の追跡対象の管理を容易にできるとしている [32]。個々の追跡対象それぞれにパーティクルフィルタによる追跡器を割り当て独立に動作させ、追跡器の相互関係に基づき追跡対象のマージや削減、追加などが行われる。また、各々の追跡器は一つの追跡対象用であるため、追跡アルゴリズム自体に大きな変更を必要としない。そのため、複数追跡への拡張は非常に容易である。

そこで、本手法もまた文献 [32] の考え方にに基づき複数人追跡の枠組みへ拡張する。追跡対象の追加、削除、結合などの追跡対象の管理の方法については以下で詳しく述べる。

追跡対象の追加（検出）

環境属性を人物頭部検出器として利用する。検出方法は追跡のそれと同様で、環境属性からサンプリングした仮説 $s_{t,det}^{(n)}$ ($n = 1, \dots, N$) を各カメラ画像平面へ投影し、追跡と同じ観測モデルにより人物頭部らしさを評価する。そして以下の二つの条件を同時に満たすとき、その仮説の位置に新たに人物が現れたと判定する。

1. 検出用に環境属性からサンプリングされた仮説 $s_{t,det}^{(n)}$ の評価値 $\pi_{t,det}^{(n)}$ が、閾値 T_{det} 以上である
2. 時刻 t の時点で追跡中の他の人物頭部とのユークリッド距離が T_{id} 以上である

条件1は人物頭部らしさが小さい仮説を除外するためであり、条件2は現在追跡している人物を新たな人物として誤検出することを防ぐためである。

追跡対象の削除

人物頭部を追跡しているとき、その密度分布 $p_m(\mathbf{X}_t|Z_t)$ は追跡対象が存在する付近でピークを持つと考えられる。逆にピークを持たない分布であるとき、それは追跡に失敗している、もしくは追跡対象が視野外に出て行ったと考えられる。このことから、分布の形状に伴い人物を追跡できているか判断することができる。

ところで、パーティクルフィルタの利用における大きな問題として、仮説の縮退現象がある。これは、全ての仮説のうち、1～数個程度の仮説以外は無視できるほど小さい重みを持つようになる現象のことで、これにより追跡の失敗、計算資源の非効率利用などの問題が引き起こされる。それを回避するために、パーティクルフィルタの縮退度を判定する指標、

$$N_{eff}(t) \equiv \frac{1}{\sum_{n=1}^N (\pi_t^{(n)})^2}, \quad (3.2.15)$$

が一般的に利用されている [1, 9]。これは、時刻 t 現在の追跡において意味のある仮説の数を表しており、これが用いる仮説の数より十分小さい、すなわち $N_{eff} \ll N$ であるとき縮退を起こしていると判定することができる。また逆に $N_{eff} = N$ となるとき、つまり全ての仮説に差がないことを意味するため、このとき均等でなだらかな分布であると判断することができる。

本手法ではこの指標に着目する。 $N_{eff} \ll N$ となるときは縮退を起こしていると考えられるが、ある一定数 T_{del} 以上であり、かつ $N_{eff} < N$ となるとき、それは人物を追跡していることによる分布の偏りであるとみなすことができる。これより、縮退度を判定する指標 N_{eff} を用いて次のようにして判断することにする。

- $T_{del} < N_{eff}(t) < N$: 「追跡中」と判定 追跡続行
- $N_{eff}(t) = N$: 「追跡失敗 or 観測視野からの退去」と判定 追跡対象削除

追跡対象の結合

複数の追跡対象が互いに近い領域に存在する場合, 周りの対象により部分的, もしくは全体が遮蔽されることがあり, これにより追跡対象の混同, 追跡の失敗といった問題を引き起こすことがある. 特に問題となるのは, 追跡対象を混同した際に, 一人の対象に二人分の追跡器が働くことである. この状況は明らかに計算資源の無駄であり, なおかつこの状態から追跡を是正することが難しい. そのため, このような問題に対処すべく, 追跡器が混同しないように様々な手法が提案されている [20, 31].

これに対して本手法では, 追跡器の他に検出器を動かしているため, 追跡に失敗しても直ちに検出器による初期化が働く. そのため, 高度な処理は必要無く, 一人の人物について二つの追跡器が常に働き続ける状況のみを回避すればよい.

そこで, 複数の追跡対象がある距離の閾値 T_{merge} より近づいた場合, その段階で追跡対象を結合し, あたかも一人を追跡しているものと判断する. そして対象が離れたとき, 検出器により追跡の初期化が働き, それぞれ追跡を再開できる. このような単純な処理により, 一人の対象を複数の追跡器が追跡し続けることを回避できる.

3.2.7 サーバ・クライアントモデルの構築

文献 [19, 20, 30] のように, サーバ用 PC 1 台とカメラ 1 台に対して 1 台のクライアント PC を割り当てたサーバ・クライアントモデルを構築する. Fig.3.4 にモデル図を示す.

具体的な追跡の処理手順は以下のようなになる.

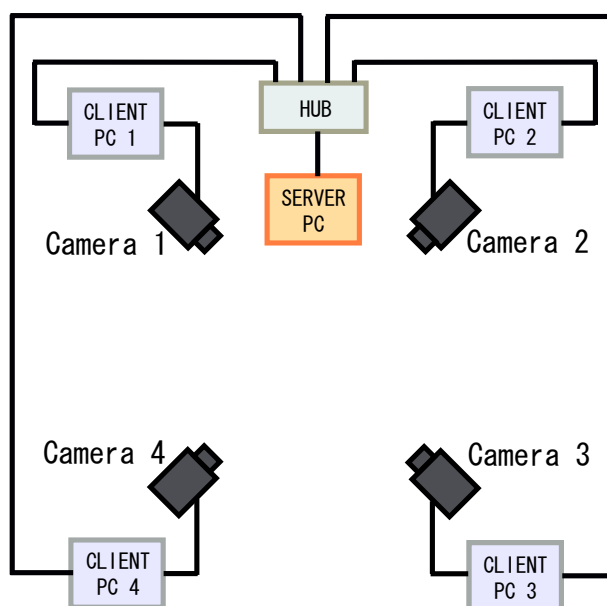


Fig. 3.4: サーバ・クライアントモデル

- (1) サーバ : 仮説群を全てのクライアントに送信
- (2) クライアント : 仮説群受信
- (3) クライアント : 観測画像に基づく人物頭部らしさの評価
- (4) クライアント : 評価値をサーバに送信
- (5) サーバ : 各クライアントの評価値を統合し, 最終的な評価値を得る
- (6) サーバ : 仮説サンプリング. その後 (1) へ戻る

観測画像を用いた人物頭部らしさの評価をそれぞれのクライアントPCで行うため, 高速な処理が可能となる.

3.3 適応的環境属性の獲得

本節では、人物の行動履歴を反映することにより得られる適応的環境属性の獲得方法について詳細に述べる。前節で述べたように、追跡の安定化のために環境属性を Importance Function $g_t(X_t)$ として用いる。そのため、環境属性を、仮説のサンプリングが容易な確率分布として表現する必要がある。そこで本手法では、複雑環境下における分布の多峰性に対応し、かつ行動履歴の反映を容易にするパラメトリックなモデルである混合正規分布を用いて表現する。このような仮定の下、人物の行動履歴を環境属性に反映する。具体的には、追跡により推定される人物の動線と、その動線により推定される行動状態に伴い、オンライン EM アルゴリズム [25] を環境属性に適用することにより実現する。そして長時間の観測により、対象シーンに特化した適応的環境属性を獲得することができる。

3.3.1 環境属性情報の定義

第2章でも述べたように、環境属性情報は、“物理的環境属性”と“適応的環境属性”の大きく二つに分類することができる。

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 [17, 30]。これを物理的環境属性と呼ぶ。

適応的環境属性

環境属性は物理的なものだけではなく、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見なすことができる。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

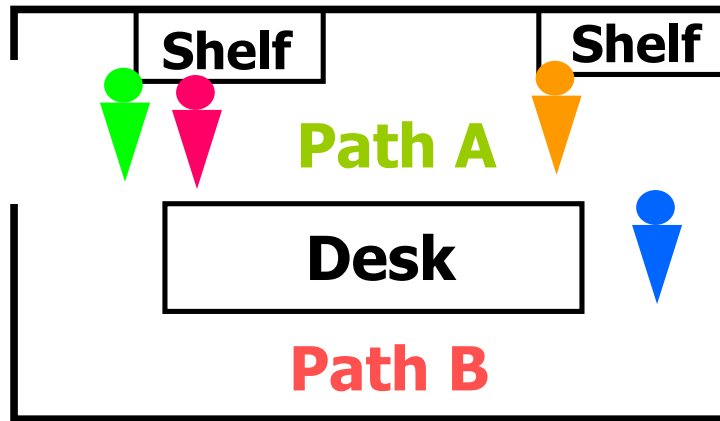


Fig. 3.5: 人物存在可能性の偏り：コンビニなどの店舗内を対象空間とした場合、立ち読みをするために本棚周辺において数人の人が滞留することがある。このような人物の行動に伴い人物の存在可能性は大きく偏ると見なすことができるが、文献 [17, 30] で提案されている物理的環境属性ではこの偏りを表現できない。

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。これを適応的環境属性と定義する。

3.3.2 適応的環境属性の確率表現

環境属性を Importance Function $g_t(\mathbf{X}_t)$ として用いるために、仮説のサンプリングが容易な確率分布として表現する必要がある。複雑環境、かつ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ため、多峰性の分布を表現可能であることが求められる。また、行動の履歴を反映を容易にすることができるパラメトリックなモデルが望ましい。そこで本手法では、混合正規分布で環境属性を表現する。すなわち、

$$g_t(\mathbf{X}_t) = \sum_{i=1}^K \omega_{i,t} \mathcal{N}(\mathbf{X}_t | \boldsymbol{\mu}_{i,t}, \boldsymbol{\Sigma}_{i,t}). \quad (3.3.1)$$

ここで $\mathcal{N}(\cdot)$ は平均 $\boldsymbol{\mu}_{i,t}$ 、共分散 $\boldsymbol{\Sigma}_{i,t}$ 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。

このようなパラメトリックな確率モデルとして表現した環境属性に行動履歴を反映するということは、行動履歴を用いてモデルが持つパラメータを更新することと

同義である。すなわち、これは、追跡により推定される動線の情報から混合正規分布のパラメータを推定する問題に帰着される。

3.3.3 動線の反映による適応的環境属性の獲得

適応的環境属性獲得のために、環境属性に行動履歴、すなわち時刻 t の事後分布により推定される人物頭部の三次元位置 s_t^{EAP} を環境属性に反映させる。具体的には、時刻 $t-1$ において各正規分布が持つパラメータ、 $\omega_{i,t-1}$, $\mu_{i,t-1}$, $\Sigma_{i,t-1}$ ($i = 1, \dots, K$) と推定される人物頭部の三次元位置 s_t^{EAP} を用いて、EM アルゴリズムをオンライン利用に拡張したオンライン EM アルゴリズム [25] を環境属性に適用することで実現する。

EM アルゴリズムとは、最尤推定手法の一つであり、観測できない、もしくは未観測の事象が存在する場合においても用いることができる汎用的な最尤推定手法である [2, 6, 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つのステップを解が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。

$$\begin{aligned} Q(\theta | \theta^{(t)}) &= E \left[\ln p(\mathcal{X}, \mathbf{Y} | \theta) | \mathcal{X}, \theta^{(t)} \right] \\ &= \sum_j \ln p(\mathcal{X}, \mathbf{y}_j | \theta) p(\mathbf{y}_j | \mathcal{X}, \theta^{(t)}) . \end{aligned} \quad (3.3.2)$$

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、

$$\theta^{(t+1)} = \underset{\theta}{\operatorname{argmax}} Q(\theta | \theta^{(t)}) . \quad (3.3.3)$$

本手法では EM アルゴリズムを環境属性、すなわち混合正規分布のパラメータ推定に適用する。このとき、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.

$$\boldsymbol{\mu}_i^{(t)} = \frac{\langle \mathbf{x} \rangle_i(m)}{\langle 1 \rangle_i(m)}, \quad (3.3.4)$$

$$\Sigma_i^{(t)} = \frac{\langle \mathbf{x} \mathbf{x}^T \rangle_i(m)}{\langle 1 \rangle_i(m)} - \boldsymbol{\mu}_i^{(t)} \left(\boldsymbol{\mu}_i^{(t)} \right)^T, \quad (3.3.5)$$

$$\omega_i^{(t)} = \langle 1 \rangle_i(m). \quad (3.3.6)$$

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分布 $p(i | \mathbf{x}_l, \theta^{(t)})$ の重み付き平均,

$$\langle f(\mathbf{x}) \rangle_i(m) \equiv \frac{1}{m} \sum_{l=1}^m f(\mathbf{x}_l) p(i | \mathbf{x}_l, \theta^{(t)}), \quad (3.3.7)$$

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

$$\begin{aligned} p(i | \mathbf{x}_l, \theta^{(t)}) &= \frac{p(i | \theta^{(t)}) p(\mathbf{x}_l | i, \theta^{(t)})}{p(\mathbf{x}_l | \theta^{(t)})} \\ &= \frac{\omega_i^{(t)} \mathcal{N}(\mathbf{x}_l | \boldsymbol{\mu}_i^{(t)}, \Sigma_i^{(t)})}{\sum_{j=1}^K \omega_j^{(t)} \mathcal{N}(\mathbf{x}_l | \boldsymbol{\mu}_j^{(t)}, \Sigma_j^{(t)})}. \end{aligned} \quad (3.3.8)$$

このような EM アルゴリズムは, 観測データが全て得られた後にバッチ的に処理される. そのため, データのサンプル数が増えるほど計算コストもまた増大するため, オンライン処理を想定する本手法には不向きである.

一方で, EM アルゴリズムをオンライン処理のような観測データが逐次的に与えられる場合においても適用できるように拡張した, オンライン EM アルゴリズムが提案されている [25]. 全観測データを用いてその都度パラメータ推定するのではなく, 時刻 t において得られた観測データと時刻 $t-1$ で推定されたパラメータを用いて時刻 t におけるパラメータを推定する. このとき, 重み付き平均, 式 (3.3.7) を次式で置き換える.

$$\langle \langle f(\mathbf{x}) \rangle \rangle_i(t) = (1 - \alpha) \langle \langle f(\mathbf{x}) \rangle \rangle_i(t-1) + \alpha f(\mathbf{x}) p(i | \mathbf{x}(t), \theta^{(t-1)}). \quad (3.3.9)$$

ここで α は学習率を意味する. これより, 式 (3.3.4) ~ (3.3.6) はそれぞれ次のように置き換えられる.

$$\boldsymbol{\mu}_i^{(t)} = \frac{\langle \langle \mathbf{x} \rangle \rangle_i(t)}{\langle \langle 1 \rangle \rangle_i(t)}, \quad (3.3.10)$$

$$\Sigma_i^{(t)} = \frac{\langle \langle \mathbf{x} \mathbf{x}^T \rangle \rangle_i(t)}{\langle \langle 1 \rangle \rangle_i(t)} - \boldsymbol{\mu}_i^{(t)} \left(\boldsymbol{\mu}_i^{(t)} \right)^T, \quad (3.3.11)$$

$$\omega_i^{(t)} = \langle \langle 1 \rangle \rangle_i(t). \quad (3.3.12)$$

このように、過去の推定結果を学習率 α で伝播させることにより、高速にパラメータ推定をすることを可能にする。本手法では逐次的に得られる人物頭部の三次元位置 s_t^{EAP} を新しい学習サンプルとしてこのアルゴリズムを利用することにより、行動履歴の環境属性への反映を実現する。

3.3.4 人物の行動状態推定

毎フレーム推定される人物頭部位置を環境属性に反映させることで、物理的環境属性では表現できない人物存在確率の偏りを表現できることを前節で述べた。ところが、このような推定される頭部位置をただ単純に反映させる方法の場合、環境属性の本来の役割である追跡の安定化の機能を失うことがある。それを引き起こす最も単純な人物の行動として、滞留状態が考えられる。

人物が滞留している間、追跡している人物頭部の位置はほとんど変化しない。その一方で、推定される頭部位置を環境属性に反映させるため、長時間の滞留が起きた場合この周辺位置のみ人物存在確率が高くなる。これにより、他の領域の人物存在確率は小さくなるため、そのような領域において人物追跡が不安定になったり、新たに視野内に入ってくる人物の検出失敗などの様々な問題を引き起こす可能性がある (Fig.3.6)。これは、追跡対象である人物がとる行動を考慮していないために起こる問題であるといえる。

このような問題へ対処するために、追跡対象である人物の行動を推定し、その状態に応じて環境属性への反映方法を変化させる方法をとる。本手法では、追跡において重要であると考えられる基本的な4つの行動状態を定義する。また、人物はそれらの状態のうちいずれかに必ず属すると仮定し、Fig.3.7のような状態遷移モデルにより人物の行動状態を記述できるものとする。

- 視野外：観測視野内に追跡対象である人物が存在しない
- 出現：人物が観測視野内に現れる
- 活動：視野内において任意の動き（歩く、作業をする、etc.）をとる
- 滞留：立ち読みやデスクワークなどにより、ほぼ今いる位置から動かない

このような仮定の下、毎フレーム推定される人物頭部位置の情報を用いてその人物の行動状態を推定する。そしてその状態に伴い、オンラインEMアルゴリズムの学

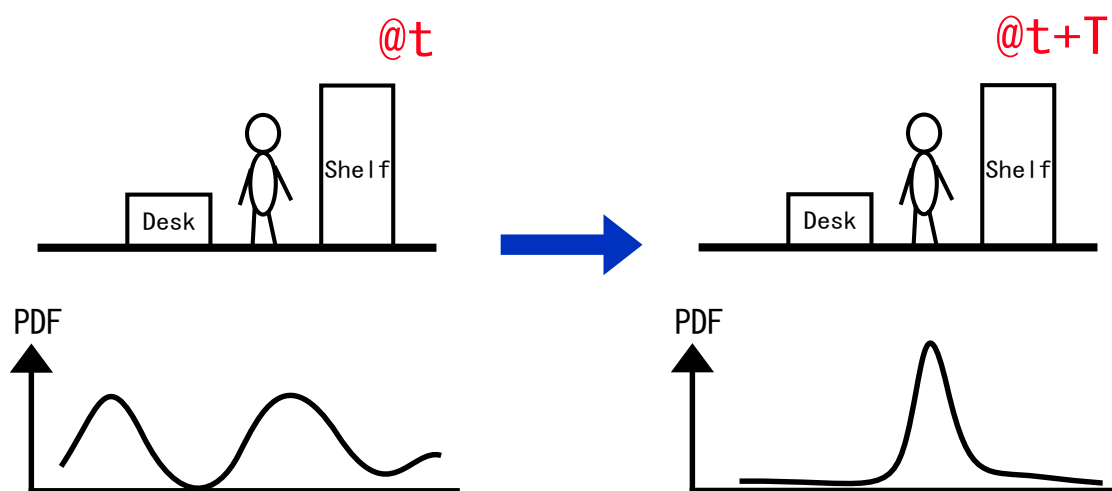


Fig. 3.6: 人物の滞留による環境属性の集中：人物頭部位置を行動状態を考慮せずに環境属性へ反映させ続けると、長時間の滞留に伴い環境属性が一箇所に鋭いピークを持つ分布形状になる。これにより他の領域において追跡の初期化、再初期化の失敗を引き起こす。

習率を変更する。これにより、滞留による環境属性の確率密度値の集中の問題を回避できる。以下、その状態遷移の判定方法について詳しく述べる。

“視野外”/“出現”/“活動”の行動状態判定

これらの行動状態判定は、追跡が成功しているかの判定方法、すなわち、3.2.6節の追跡対象の削除の判定法と同義であるといえる。よって、3.2.6節と同様にパーティクルフィルタの縮退度を表す指標 N_{eff} を用いて、次の条件により人物が視野内に存在するかどうかの判定、すなわち“視野外” “出現”、“活動” “視野外”の人物行動状態遷移を判定する。

- $N_{eff}(t) = N$ ：“視野外”の状態へ遷移
- $N_{eff}(t-1) = N$ かつ $T_{del} < N_{eff}(t) < N$ ：“視野外”から“出現”へ状態遷移
- $T_{del} < N_{eff}(t-1) < N$ かつ $T_s < N_{eff}(t) < N$ ：“活動”の状態へ遷移

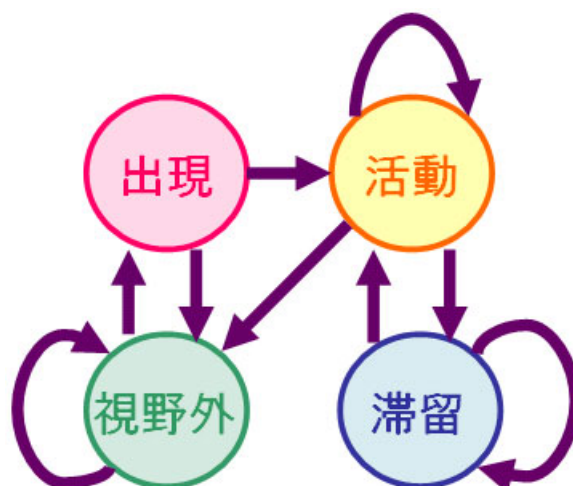


Fig. 3.7: 人物の行動状態遷移図

“活動”/“滞留”行動状態判定

“滞留”状態とは、対象視野内において同じ位置に長時間留まることである。このことから、“滞留”状態であると判断するための条件を次のように表現する。

1. (空間的条件)

「同じ位置からほとんど動かない」という空間的な条件を記述する。

“活動”状態に遷移したときの初期位置を s_0 として格納する。この位置と毎フレーム推定される頭部位置 s_t^{EAP} とのユークリッド距離が、ある距離の閾値 T_{dist} より小さいとき空間的に滞留状態であるとみなす。すなわち、

$$\|s_t^{EAP} - s_0\|^2 < T_{dist}. \quad (3.3.13)$$

2. (時空間的条件)

1の空間的条件に、「長時間留まる」という時間的な条件を加える。

空間的条件、式(3.3.13)を、時間判定のための閾値 T_{stay} 以上満たすとき、その人物は“滞留”状態にあると判定する。

これにより対象視野内における人物の行動状態は判定される。以下に“行動” “滞留”状態遷移判定についてまとめる。

- “活動”状態かつ2の条件を満たしている：“活動”状態から“滞留”状態へ遷移
- “活動”状態かつ1,もしくは2の条件を満たさない：“活動”状態へ遷移
- “滞留”状態かつ1の条件を満たしている：“滞留”状態へ遷移
- “滞留”状態かつ1の条件を満たさない：“活動”状態へ遷移

第4章 実験

実環境において、本提案手法を用いて次の2つの実験を行った。

(I) 適応的環境属性の獲得実験

対象空間を長時間観測することにより、対象シーンに特化した適応的環境属性の構築を行った。追跡対象の行動状態を推定することによる効果を比較、確認するために、行動状態を考慮しない場合の適応的環境属性の獲得実験もまた行った。

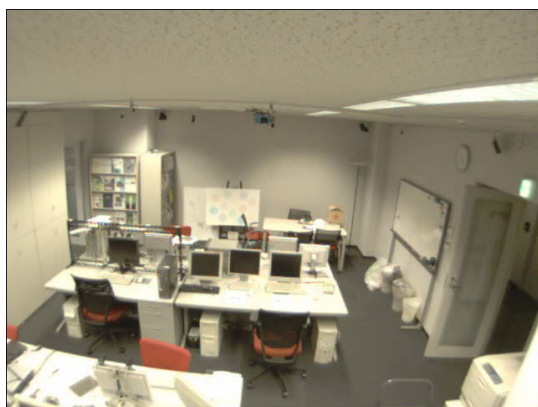
(II) 人物追跡実験

実験 (I) にて得られた適応的環境属性を用いて複数人の人物追跡実験を行った。適応的環境属性の効果を確認するために、様々な状況や方法により得られる環境属性を用いて実験を行った。

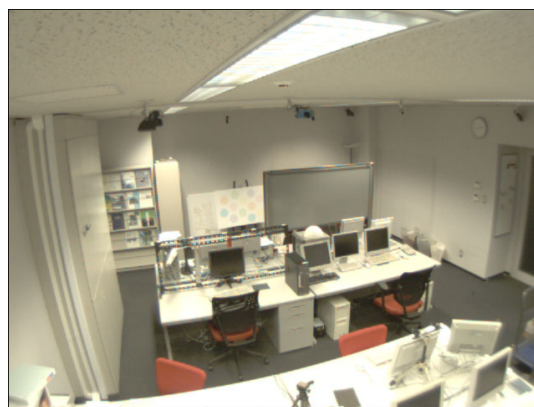
本章では、これらの実験による詳細な結果を示すと共に、得られた結果について詳しく考察する。

4.1 実験環境

室内天井に校正済みの IEEE1394 カメラ (Point Grey Research 社製 Flea) を 2 台設置する。解像度 640 x 480 ピクセル、毎秒 60 フレームの速度で得られるカメラ画像を用いる。Fig.4.1 に観測されるカメラ画像を示す。このようなカメラ 1 台毎に 1 台の汎用 PC (CPU: Intel Pentium4 2.8GHz, RAM: 2GByte, OS: WindowsXP) を割り当て、それぞれクライアント PC とする。またそれらの情報を統合するサーバの役割を果たす汎用 PC (CPU: Intel Pentium4 3.2GHz, RAM: 1GByte, OS: WindowsXP) を用意し、これらを通信速度 1Gbps の Gigabit Ethernet で接続してサーバ・クライアントモデルを構築した。またデータ通信には DirectX9.0c に含まれる API の一つである



(i) 観測画像 (カメラ1)



(ii) 観測画像 (カメラ2)

Fig. 4.1: 観測視野

DirectPlay [23] を利用し, 通信プロトコルは TCP/IP プロトコルを用いた. これにより Winsock [36] を利用するよりもはるかに容易にネットワークプログラムを構築できる.

4.2 適応的環境属性の獲得実験

対象空間を長時間観測することにより, 対象シーンに特化した適応的環境属性の構築を行った. 本節ではその結果と考察について述べる.

まず行動履歴を環境属性に反映するための環境属性の初期状態の設定について簡単に説明する. 本実験では, 簡単のために対象空間内の大まかな物理的な情報を利用して初期状態を設定した. そしてこの得られた初期状態を元にし, 適応的環境属性の獲得実験を行った. 具体的には, 追跡対象の行動状態推定の効果を比較, 確認するために, 行動状態を考慮する場合としない場合の実験を行った. それぞれの実験により得られた結果を比較し, 考察する.

4.2.1 環境属性の初期化

対象空間全体を S , 椅子や机などの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 $O \equiv \bigcup_{i=1}^L B_i$ とする. このとき, 人物が存在し得る領域

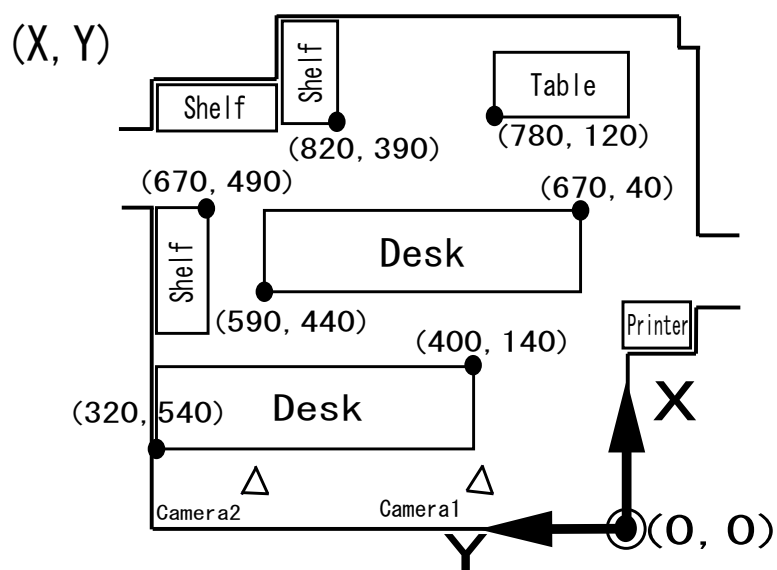


Fig. 4.2: 対象空間の物理的情報：対象空間における世界座標系 $O - XYZ$ を意味する。また、図中に記載されている数字は、世界座標系における各位置の (X, Y) 座標であり、単位は $[cm]$ である

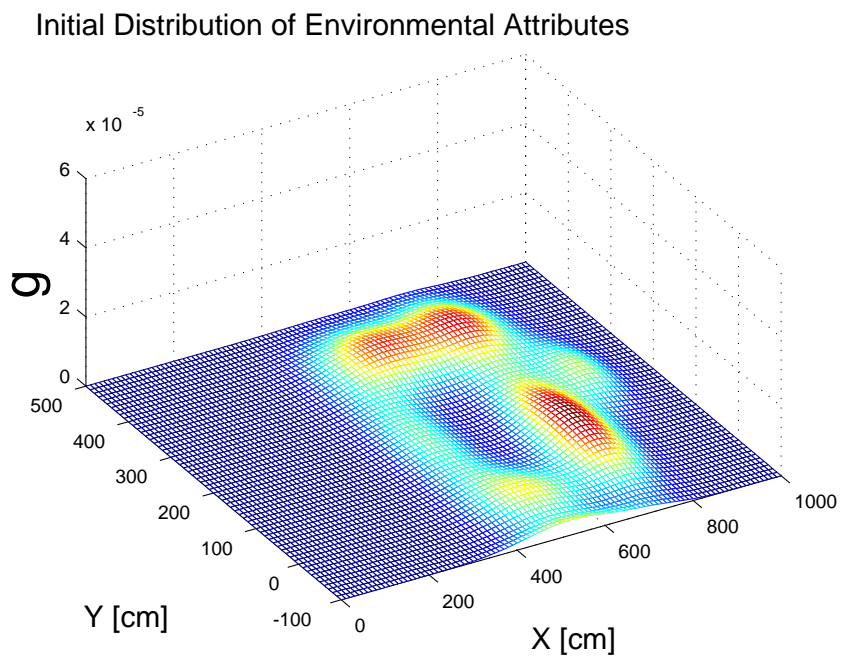


Fig. 4.3: 環境属性の初期状態

Table. 4.1: 適応的環境属性の獲得実験に用いたパラメータ

記号	値 [単位]	パラメータの詳細
K	7 [個]	環境属性の正規分布の個数
N	200 [個]	1 追跡対象あたりに用いる仮説数
T_{det}	$1/N$ [-]	新規追跡対象検出 (3.2.6 節) 用人物頭部らしさ重み閾値
T_{id}	150 [cm]	新規追跡対象検出 (3.2.6 節) 用誤検出防止用距離閾値
T_{del}	30 [particle]	追跡対象の削除判定用 (3.2.6 節) の有効仮説数閾値
T_{merge}	70 [cm]	追跡対象の結合判定用 (3.2.6 節) の距離閾値
T_{dist}	50 [cm]	活動 / 滞留 行動状態判定用 (3.3.4 節) の距離閾値
T_{stay}	200 [frame]	活動 / 滞留 行動状態判定用 (3.3.4 節) の時間閾値

$P \equiv S \cap \bar{O}$ に一様に乱数を発生させ、10000 個のデータ集合 $\{x_1, \dots, x_m\}$ を得た。そしてこれらを用いて EM アルゴリズムを実行することで、環境属性の初期パラメータ $\omega_{i,0}, \mu_{i,0}, \Sigma_{i,0}$ ($i = 1, \dots, K$) を獲得した。また正規分布の個数は $K = 7$ とした。推定された環境属性の初期状態を Fig.4.3 に示す。Fig.4.2 の対象空間のレイアウト図と比較してみると、通路などの人物の存在しやすい領域に峰があり、初期状態の分布として妥当であるといえる。

4.2.2 適応的環境属性の獲得

対象空間を長時間観測することにより、適応的環境属性を獲得する。本実験では、行動状態推定の有無による結果を比較するために、予め撮影した 15 分程度の観測映像 (Fig.4.4) についてそれぞれのアルゴリズムを適用した。また用いた種々のパラメータ値は Table.4.1 にまとめた。

それぞれの推定結果を Fig.4.5, 4.6 に示す。両結果共に環境属性の初期状態 (Fig.4.3) と比較して分布形状が大きく変化しており、追跡対象の行動履歴が反映されていることがわかる。

動線の反映のみにより得られた適応的環境属性は、ある領域 ($650 < X < 750, 0 <$



(a) 特定の通路の頻繁な通行

(b) 特定位置での長時間の滞留

Fig. 4.4: 適応的環境属性の獲得に用いた観測映像

$Y < 100$) に分布が集中しており, 他の領域ではその確率密度値が非常に小さな値を示していることがわかる. 用いた観測映像と比較すると, Fig.4.4 にあるように確率密度が集中している領域周辺において二人の人物が長時間雑談している様子が観測された. 追跡対象の行動履歴の反映が正しく行われていることが確認される一方で, これは 3.3.4 節で述べた人物の滞留行動が引き起こす環境属性の確率密度の集中であり, これにより他の人物が存在しやすい領域の情報が失われてしまっていることが考えられる.

行動状態推定を利用することにより得られた適応的環境属性 (Fig.4.6) では, 人物の滞留行動が観測された領域周辺にピークが観測される以外にも奥側の通路領域 ($650 < X < 750$, $0 < Y < 450$) においてなだらかな峰が観測される. これは Fig.4.4 に観測されたように滞留行動以外に奥側の通路の頻繁な通行が起きたことを意味する. 前者では長時間の滞留行動により他の情報が失われてしまっていたが, 後者では行動状態推定器により追跡対象の行動状態が推定され, その状態に伴い環境属性への行動の反映の強さが適応的に変更されたため, 滞留行動による確率密度の集中が抑えられたと考えられる. 確認のために行動状態推定器により追跡対象が滞留行動状態であると推定された対象空間中の三次元位置をプロットした図を Fig.4.7 に示す. この図からわかるように, 確かに滞留行動が観測された領域周辺に推定された位置が集中していることが見て取れる. これによりこの領域周辺への確率密度の集中が抑えられたといえる.

以上のことから, 行動状態を推定することにより獲得される適応的環境属性の方

Environmental Attributes WITHOUT estimating Human Activity

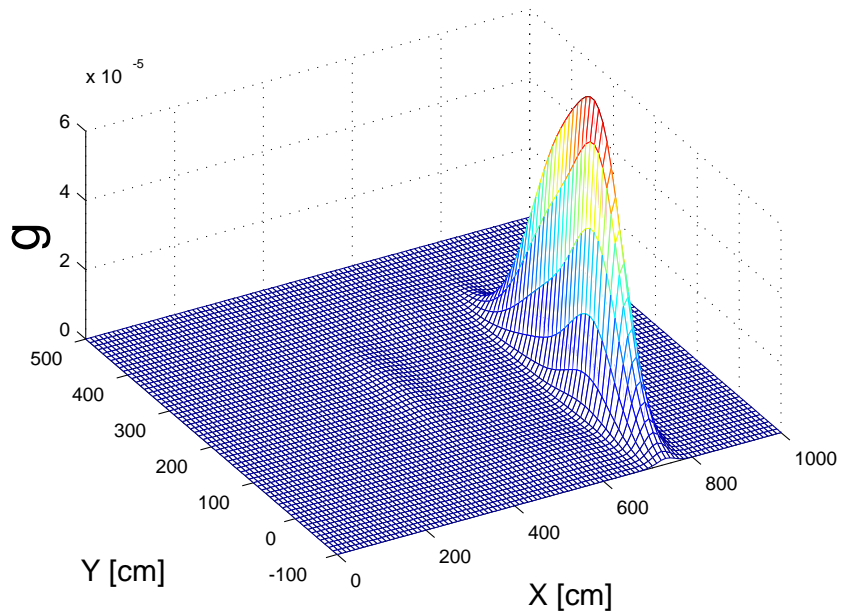


Fig. 4.5: 適応的環境属性（動線のみ利用）

Environmental Attributes with estimating Human Activity

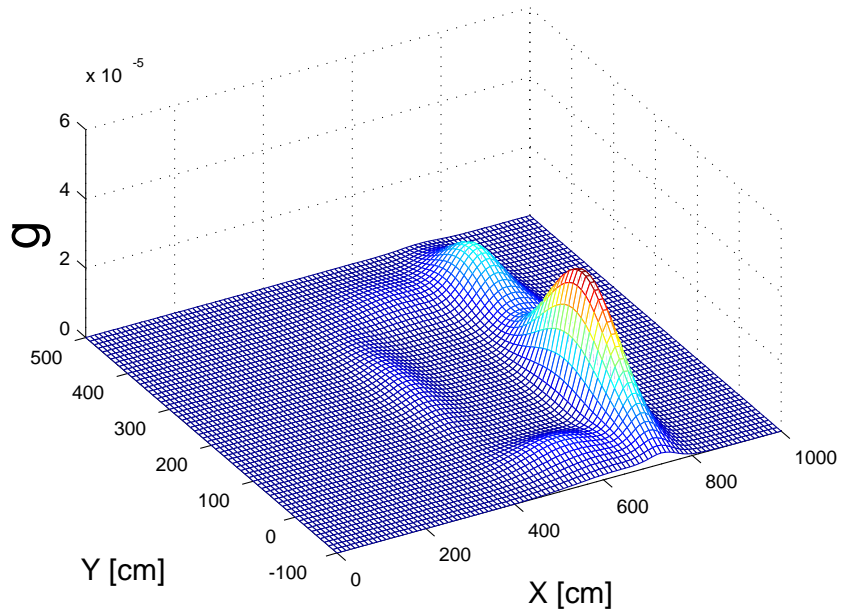


Fig. 4.6: 適応的環境属性（行動状態推定を伴う）

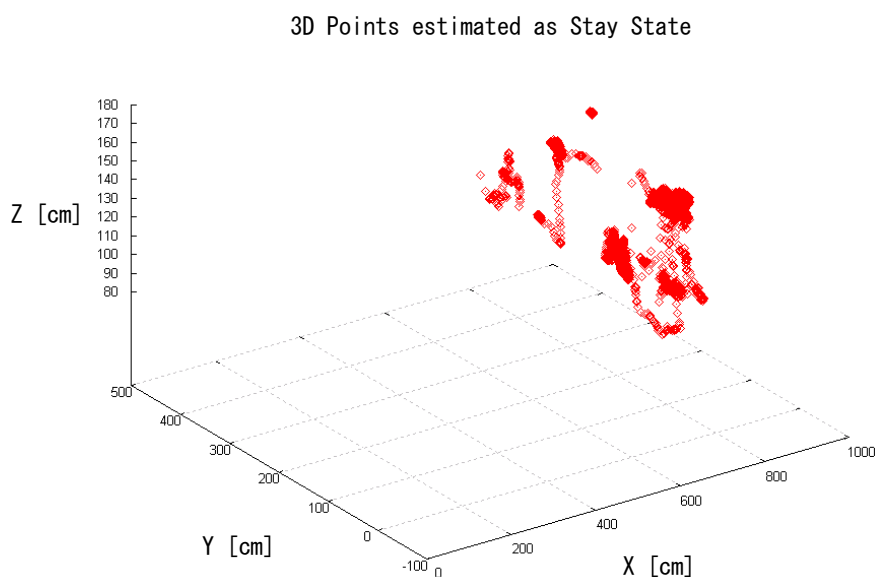


Fig. 4.7: 行動状態推定器による滞留行動状態の推定結果

が、人物の行動をより詳細に表現した環境情報を獲得できているといえる。また、これは追跡の安定化にも有効であると考えられる。詳しくは次節にて述べる。

4.3 人物追跡実験

4.2節の実験により得られた適応的環境属性を用いて人物追跡実験を行った。本節では、その結果を示すと共に考察する。

まず、環境属性の有無による人物追跡結果を示す。この結果より環境属性を追跡へ応用することの利点を示す。そして次に4.2節にて得られたそれぞれの適応的環境属性を用いた複数人の人物追跡実験の結果を示し、それらの結果を比較することで、追跡対象の行動状態推定の追跡精度への寄与について考察する。

4.3.1 環境属性の効果

人物追跡における環境属性の効果を確認する。具体的には、“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、ボードの裏側で留まる。その後、奥の部屋へ向かって歩くことで視野外から出る。”というようなビデオ映像を予め撮影し

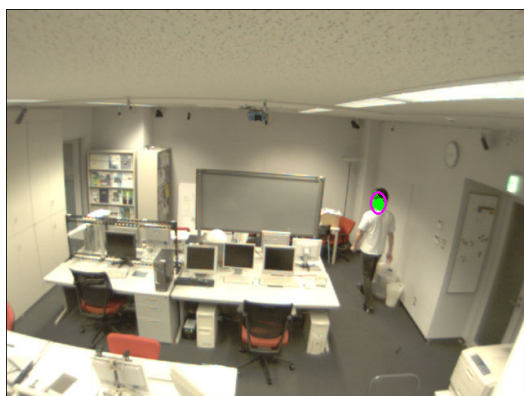
ておき、それに対し環境属性を伴う場合とそうでない場合それぞれの追跡アルゴリズムを適用した。

Fig.4.8 に両者それぞれの場合の追跡結果の一部を示す。(a), (c), (e) は環境属性を用いていない、基本的なパーティクルフィルタによる追跡結果、(b), (d), (f) は環境属性を導入したアルゴリズム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、両者共に人物頭部を追跡できている。しかし、ボードの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、環境属性を伴う追跡による結果では、正しい追跡を再開できていることがわかる((f))。これは、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、障害物や他の人物による遮蔽により追跡対象を見失った場合においても、容易に正しい追跡を再開できるため、環境属性を追跡に導入することは有効であるといえる。

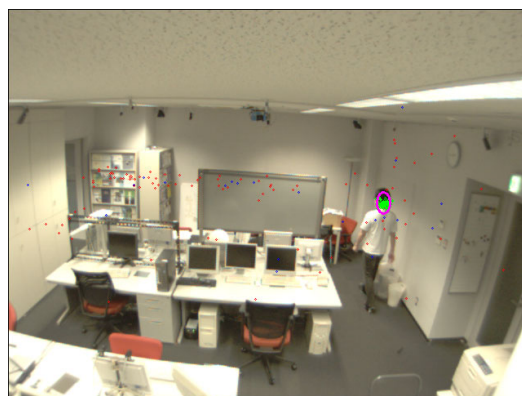
4.3.2 適応的環境属性を伴う人物追跡実験

4.2 節にて得られた、行動状態推定を伴う適応的環境属性とそうでないものそれぞれを用いた複数人の人物追跡実験を行った。両者の追跡結果を比較するために、予め 30fps で撮影しておいた観測映像についてそれぞれの追跡アルゴリズムを適用した。

Fig.4.9 にそれぞれの場合の追跡結果の一部を示す。視野外から人物が現れた場合(ii-a, b)、動線の反映のみにより得られた適応的環境属性を伴う追跡(ii-a)の方では新しい追跡対象の初期化が行われていないことに対し、行動状態推定を伴う追跡(ii-b)においては初期化が正しく行われていることがわかる。これは、前述したように、動線のみによる適応的環境属性は人物の滞留行動により、確率密度分布が滞留領域に集中してしまっていることに起因すると考えられ、この結果のように他の領域から人物が現れた場合に対処できない。その一方で、行動状態推定を伴う場合は、滞留の情報以外にも人物の存在しやすい領域の情報を保持しているため、(ii-b)のように高速かつ正しい追跡対象の初期化を可能とする。動線の反映のみの追跡において新たな追跡対象が初期化されたとき(ii-e)は、それは滞留行動により密度が集中している領域に人物がさしかかったときであり、これはあたかも特定の位置に初期化用の仮説を発生させている状態と同様で、柔軟な複数追跡を実現しているとはいえない。また行動状態推定を伴う追跡と比較すると結果的におよそ 150 フレームの差が確認された(ii-e, f)。以上の事柄から、行動状態推定を用いることにより得られた適応的環境属性は、追跡の安定化、特に追跡対象の初期化、再初期化に関して非常に有効であ



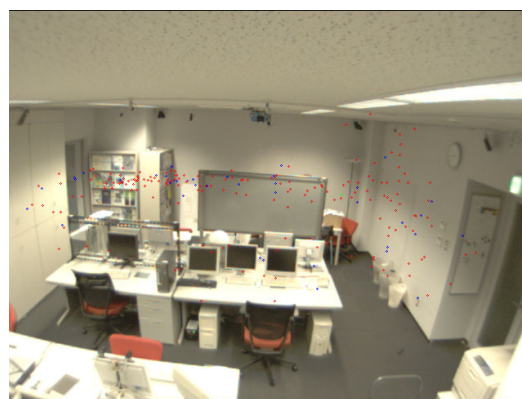
(i-a) # 300



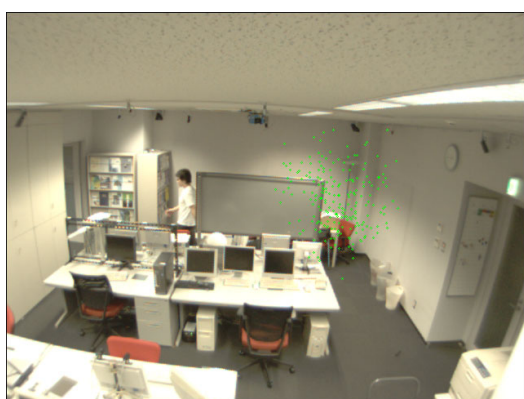
(i-b) # 300



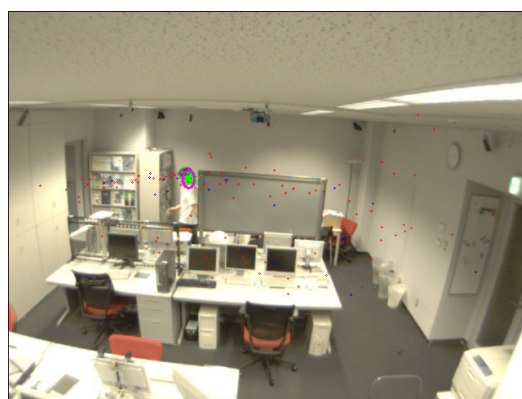
(i-c) # 400



(i-d) # 400



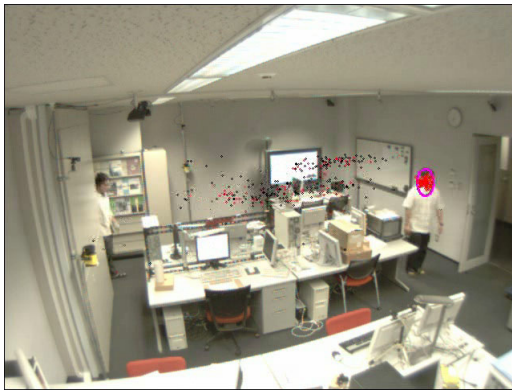
(i-e) # 700



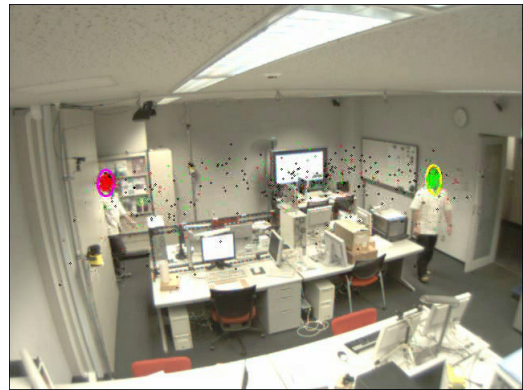
(i-f) # 700

Fig. 4.8: 環境属性の有無に伴う追跡結果比較：左列 ((i-a),(i-c),(i-e)) は環境属性無しの場合の追跡結果, 右列 ((i-b),(i-d),(i-f)) は環境属性を用いた場合の追跡結果を意味する.

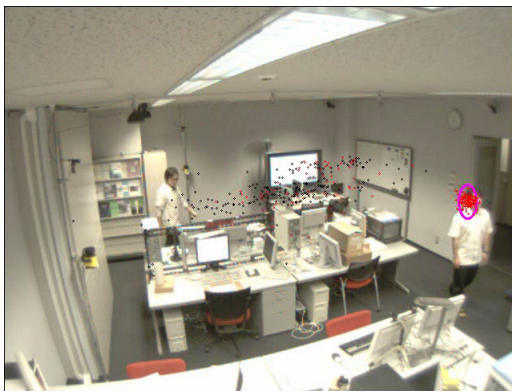
るといえる。



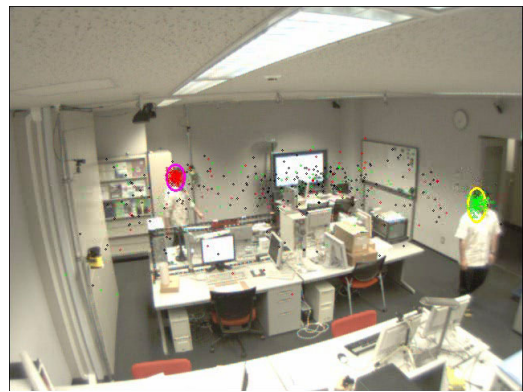
(ii-a) # 989



(ii-b) # 989



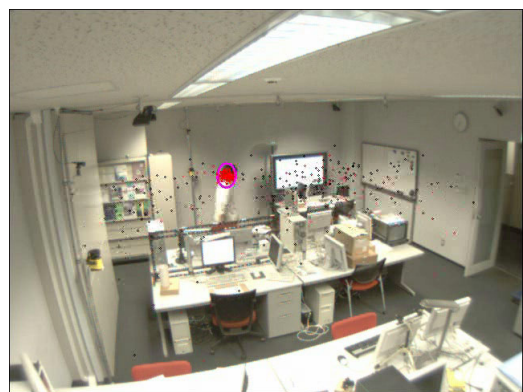
(ii-c) # 1041



(ii-d) # 1041



(ii-e) # 1141



(ii-f) # 1141

Fig. 4.9: 行動状態推定の有無に伴う追跡結果比較：左列 ((ii-a),(ii-c),(ii-e)) は動線の反映のみにより得られた適応的環境属性を伴う追跡結果, 右列 ((ii-b),(ii-d),(ii-f)) は行動状態推定を用いることにより得られた適応的環境属性を伴う追跡結果を意味する.

第5章 結論

5.1 まとめ

対象空間内において、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といったなんらかの行動の特性を持つことから、対象空間中の人物の存在可能性はこのような行動特性に大きく依存すると考えられる。本研究ではこの点に着目し、人物の行動履歴を逐次利用することにより、対象シーンに特化した人物の存在確率を獲得できるとした。これを適応的環境属性と定義し、その構築法、追跡の枠組みへの統合性について述べた。具体的には次の二つの機構が相互に作用し合うことにより、本システムは実現される。

5.1.1 環境属性を組み入れた人物頭部三次元追跡

ICONDENSATION の考え方にに基づき、環境属性をパーティクルフィルタの枠組みに導入した。環境属性を Importance Function として利用することで、追跡が不安定になる、もしくは失敗した際に、環境属性により追跡が三次元的に修正、初期化される。これにより正しい追跡への復帰を可能とし、結果的に追跡を安定化することができる。これは特に遮蔽の問題について効果があり、環境属性を伴わないパーティクルフィルタとの追跡結果を比較することによりその有効性を確認した。また、適応的環境属性による追跡の安定化も実環境の実験を通して確認することができた。

5.1.2 適応的環境属性の獲得

対象空間における人物の存在確率、すなわち環境属性と人物の行動は強い関連があることから、人物の行動履歴はシーンに特化した適応的な環境属性を獲得する上で重要な手がかりであるといえる。本研究では、毎時推定される人物頭部三次元位置を、オンライン EM アルゴリズムを用いて環境属性へ反映することにより、対象とす

るシーンに特化した適応的環境属性を獲得することができる考えた。しかしながら、この手法を用いた場合、人物の滞留行動により環境属性の確率密度値が一箇所に集中することがあり、その結果追跡の不安定化、失敗を引き起こすことがあった。

このような問題に対処するために、本研究ではこの枠組みに更に人物の行動状態推定を導入した。具体的には、得られる人物動線から追跡対象の行動状態を推定し、その状態に伴い環境属性への反映の重みを適宜変更することで、上述した問題に対処した。その効果を確認するために、長時間の観測映像を録画し、行動状態推定を導入した場合とそうでない場合に得られる適応的環境属性の比較、考察を行った。これにより前者は滞留行動による密度集中の問題に対応できることを確認した。このことから、適応的環境属性の獲得における人物の行動状態推定は、非常に重要かつ効果的であるといえる。

5.2 今後の展望

5.2.1 環境属性の表現の再考案

正規分布の個数設定の問題

現在の手法では、環境属性を混合正規分布で近似し、EM アルゴリズムを用いてオンライン学習を行っているが、正規分布の数は経験的に定めている。しかしながら、EM アルゴリズムによる混合正規分布のパラメータ推定において、正規分布の個数の決定は難しい問題であり、本研究においてもまた対処しなければならない問題である。例えば正規分布の個数が足りていない場合、追跡の初期化、再初期化が働かない可能性が考えられる。逆に個数が多すぎる場合、計算コストの非効率利用の問題が起きる。そのため、シーンに動的に対応できる適応的な分布数の推定手法について考慮する必要がある。また、EM アルゴリズムではないが、混合正規分布の個数を適応的に決定する手法 [27, 28] の利用も有効であると考えられる。

ノンパラメトリックな分布推定手法の利用

環境属性を混合正規分布で近似するのではなく、パルツェン窓 [4, 7, 8] などのノンパラメトリックな分布推定を利用する手法が考えられる。この手法の利点としては、混合正規分布の個数問題を考慮する必要がないこと、そしてより正確な密度推定を

行うことができることが挙げられる。しかしながら、全てのサンプルを用いて密度推定するため、サンプル数が増大した際にリアルタイム性を保つことが難しくなる問題が考えられる。また、その推定された分布から仮説を高速に発生させる手法が必要とされるため、詳細な調査、手法の考案が必要であると考えられる。

5.2.2 観測モデルの高精度化

追跡の安定化において、本研究で主に注目した点は効率的な仮説のサンプリングの実現であり、環境属性を Importance Function や検出器として利用する手法を提案した。そのため、観測モデルによる評価の高精度化については深く言及せず、単純な指標(背景差分, エッジ)の統合による評価手法を利用した。しかしながら、室内などの実環境における複雑背景下では明らかに精度不足であり、追跡において誤検出、誤追跡をすることが多々確認された。このような観測モデルの精度の低さによる問題を解決するために、より高精度な観測モデルである顔識別器を利用した手法 [15, 16] の利用、もしくは更なる高精度な観測モデルの提案が必要となる。

5.2.3 行動状態推定の詳細化

本研究では、人物の行動パターンを最も単純な4つの状態で表現しているが、これを更に詳細に分類し、より厳密に表現することで、更なる詳細な環境属性の獲得をすることができると考えられる。またそれに伴い文献 [5] などのように、Location-aware Computing への応用も考えられる。

謝辞

本研究を行うにあたり、日頃から熱心に御指導を下さいました佐藤洋一助教授に深く御礼申し上げます。

また、忙しい中、本研究に関する内容をはじめ、論文その他様々な面でご助言下さいました、国立情報学研究所の杉本晃弘教授に心より御礼申し上げます。

佐藤研究室の岡部孝弘助手には、普段より研究内容をはじめとした様々なことについて相談にのって下さり、大変感謝しております。

御自分の御研究でお忙しいにも拘わらず、常日頃から私の些細な疑問についても親身になって相談にのって下さった小林貴訓さんに深く感謝しております。

また、同じ研究室で研究生活を共に過ごすことのできた佐藤研究室の皆様に御礼を申し上げます。

最後に、何一つ不自由ない生活を援助して頂いた両親に深く感謝致します。ありがとうございました。

平成 19 年 2 月

杉村 大輔

参考文献

- [1] M. S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, " A Tutorial on Particle Filters for Online Nonlinear / NonGaussian Bayesian Tracking, " *IEEE Trans. Signal Processing*, vol.50, no.2, pp.174-188, 2002.
- [2] 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.
- [3] 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.
- [4] D. Comaniciu, " An Algorithm for Data-Driven Bandwidth Selection, " *Trans. on Pattern Analysis and Machine Intelligence*, vol.25, no.2, pp.281-288, 2003.
- [5] 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.
- [6] A. P. Dempster, N. M. Laird and D. B. Rubin, " Maximum-likelihood from incomplete data via the EM algorithm, " *Journal of Royal Statistical Society B*, Vol.39, pp.1-22, 1977.
- [7] R. O. Duda, P. E. Hart and D. G. Stork, " Pattern Classification, " John Wiley and Sons, Inc., 2001.
- [8] A. Elgammal, D. Harwood and L. Davis, " Non-parametric Model for Background Subtraction, " *Proc. European Conference on Computer Vision*, vol.2, pp.751-767, 2000.
- [9] F. van der Heijden, " Consistency Checks for Particle Filters, " *Trans. on Pattern Analysis and Machine Intelligence*, vol.28, no.1, pp.140-145, 2006.

- [10] M. Isard and J. MacCormick, " BraMBLe: A Bayesian Multiple-Blob Tracker, " *Proc. IEEE International Conference on Computer Vision*, vol.2, pp.34-41, 2001.
- [11] M. Isard and A. Blake, " CONDENSATION - Conditional Density Propagation for Visual Tracking, " *International Journal of Computer Vision*, vol.29, no.1, pp.5-29, 1998.
- [12] M. Isard and A. Blake, " ICONDENSATION: Unifying low-level and high-level tracking in a stochastic framework, " *Proc. European Conference on Computer Vision*, vol.1, pp.893-908, 1998.
- [13] 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.
- [14] 徐剛, 辻三郎, " 3次元ビジョン, " 共立出版, 1998.
- [15] Y. Kobayashi, D. Sugimura, Y. Sato, K. Hirasawa, N. Suzuki, H. Kage and A. Sugimoto, " 3D Head Tracking using the Particle Filter with Cascaded Classifiers, " *Proc. the British Machine Vision Conference*, pp.37-46, 2006.
- [16] 小林貴訓, 杉村大輔, 平澤宏祐, 鈴木直彦, 鹿毛裕史, 佐藤洋一, 杉本晃宏, " カスケード型識別器を用いたパーティクルフィルタによる人物三次元追跡, " 画像の認識・理解シンポジウム (MIRU2006), 2006.
- [17] 小林貴訓, 佐藤洋一, 杉本晃宏, " 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, " 情報処理学会研究報告, 2005-CVIM-150-21, pp.169-176, 2005.
- [18] G. Loy, L. Fletcher, N. Apostoloff and A. Zelinsky, " An Adaptive Fusion Architecture for Target Tracking, " *Proc. 5th IEEE International Conference on Automatic Face and Gesture Recognition*, pp.261-265, 2002.
- [19] 松元郁祐, 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.

- [20] 松元郁祐, 加藤丈和, 和田俊和, “複数カメラを用いた CONDENSATION によるオクルージョンにロバストな人物追跡,” 画像の認識・理解シンポジウム (MIRU2006), pp.501-506, 2006.
- [21] K. Nummiaro, E. Koller-Meier and L. Van Gool, “An Adaptive Color-Based Particle Filter,” *Image and Vision Computing*, vol.21, no.1, pp.99-110, 2003.
- [22] K. Okuma, A. Taleghani, N. Freitas, J. Little and D. Lowe, “A Boosted Particle Filter: Multitarget Detection and Tracking,” *Proc. European Conference on Computer Vision*, vol.3021, pp.28-39, 2004.
- [23] 大澤文孝, “書籍版 DirectX9 実践プログラミング,” 工学社, 第5章 DirectPlay, pp.261-311, 2005.
- [24] P. Perez, J. Vermaak and A. Blake, “Data Fusion for Visual Tracking with Particles,” *Proc. IEEE*, vol.92, no.3, pp.495-513, 2004.
- [25] M. Sato and S. Ishii, “On-line EM Algorithm for the Normalized Gaussian Network,” *Neural Computation*, 12(2), 2000.
- [26] J. Sherrah and S. Gong, “Fusion of perceptual cues for robust tracking of head pose and position,” *Pattern Recognition*, vol.32, no.8, 2001.
- [27] 島田敬士, 有田大作, 谷口倫一郎, “混合ガウス分布による動的背景モデルの分布数増減法,” 画像の認識・理解シンポジウム (MIRU2006), pp.746-751, 2006.
- [28] C. Stauffer and W. Grimson, “Learning Patterns of Activity using Real-Time Tracking,” *Trans. on Pattern Analysis and Machine Intelligence*, vol.22, no.8, pp.747-757, 2000.
- [29] 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.
- [30] 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.
- [31] D. Tweed and A. Calway, “Tracking Many Objects Using Subordinated CONDENSATION,” *Proc. the British Machine Vision Conference*, pp.283-292, 2002.

- [32] J. Vermaak, A. Doucet and P. Perez, " Maintaining Multi-Modality through Mixture Tracking, " *Proc. IEEE International Conference on Computer Vision*, vol.2, pp.1110-1116, 2003.
- [33] P. Viola and M. Jones, " Rapid Object Detection using a Boosted Cascade of Simple Features, " *Proc. the IEEE International Conference on Computer Vision and Pattern Recognition*, vol.1, pp.511-518, 2001.
- [34] 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.
- [35] 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.
- [36] 横田秀次郎, " 32ビット通信プログラム作成法, " CQ 出版社, 1997.

発表文献

1. 杉村大輔, 小林貴訓, 佐藤洋一, 杉本晃宏, ”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,” 情報処理学会コンピュータビジョンとイメージメディア研究会, 2006-CVIM-156-21, pp.171-178, 2006.
2. 小林貴訓, 杉村大輔, 平澤宏祐, 鈴木直彦, 鹿毛裕史, 佐藤洋一, 杉本晃宏, “カスケード型識別器を用いたパーティクルフィルタによる人物三次元追跡,” 画像の認識・理解シンポジウム (MIRU2006), 2006.
3. Y. Kobayashi, D. Sugimura, Y. Sato, K. Hirasawa, N. Suzuki, H. Kage and A. Sugimoto, “3D Head Tracking using the Particle Filter with Cascaded Classifiers,” *Proc. the British Machine Vision Conference*, pp.37-46, 2006.