

非同期多視点画像を用いた物体追跡

内海 章 鉄谷 信二

ATR メディア情報科学研究所

〒 619-0288 京都府相楽郡精華町光台 2-2-2

{utsumi, tetsutani}@atr.co.jp

あらまし: 非同期多視点画像を利用した物体追跡の一例として人の動きを追跡するシステムについて述べる。本システムは非同期に動作して各視点の画像を処理する複数の観測ノードと対象物体の追跡を行う追跡ノード、発見を行う発見ノードからなる。本システムでは、多視点で非同期に得られる観測情報をカルマンフィルタにより統合し効率的に追跡処理を行う。追跡モデルと画像特徴の対応づけは追跡ノードから送られる情報に従い各観測ノードで行われる。対応のとれた画像特徴は追跡ノードに送られ追跡モデルが更新される。対応のとれない特徴はシーン内に新たに現れた対象物体の発見を行うため発見ノードに送られる。本手法により、視点数の増加による処理効率の低下、観測間の冗長性の増大といった問題の発生を抑制しながら、大規模な追跡システムの構築が可能になる。

Object Tracking Based On Non-Synchronous Multiple Camera Images

Akira Utsumi Nobuji Tetsutani

ATR Media Information Science Laboratories

2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288 JAPAN

{utsumi, tetsutani}@atr.co.jp

Abstract: We describe a human tracking method using non-synchronous multiple camera observations. Our experimental system consists of three types of processes, discovering node for detecting newly appeared person, tracking node for tracking each target person, and observation node for processing one viewpoint (camera) images. In our method, each camera observes a scene fully independently and multiple observations are integrated with a Kalman-filter based algorithm for an efficient human tracking. Matching between tracking models and observed image features are performed in each observation node based on the information sent from the tracking node. Image features that match with tracking models are sent to the tracking node to update tracking model. Unmatched features are sent to the discovering node for new person detection. Due to the non-synchronous property, this system is highly scalable for increasing the detection area and number of observing nodes.

1 はじめに

我々は、計算機との非接触インタフェースを目指し、人の動きを画像処理によって検出する手法の検討を進めている。検出対象となる情報としては、追跡に必要な位置・移動方向の検出、着座動作を含むいくつかの動作の検出に加えて、各人物の同定に必要な顔画像、背丈、服の色等の抽出などが考えられる。これらの情報を検出することにより、仮想環境でのインタラクション、ユーザの位置に応じた情報提供システム、さらには監視システムといった幅広い応用が可能になると考えられる。

画像を用いた人物追跡については従来より多くの提案がある [1, 2, 3, 4]。これには画像を用いることで非装着、非接触の検出が可能になり、ユーザへの負担が軽減する、応用範囲が広がる等の理由が考えられる。これら手法の多くは単眼または2眼画像を利用するものであるが、限られた視点数ではオクルージョンに対応できない、検出エリアが狭い、十分な精度が得られないといった問題があった。これに対し多視点画像を利用した人物追跡システムはオクルージョンの発生を低減し、より安定した検出が可能になると期待され、いくつかの提案が行われている [5, 6, 7, 8, 9, 10]。

しかし、このようなシステムで広範囲の人間の動きを追跡するには検出エリアに合わせて多くのカメラが必要になり、そのことにより生じる問題も多い。例えば、多くのビジョンシステムは3次元計測のために各カメラが同時刻に観測を行うことを前提としており、そのための同期機構の導入によりシステムが複雑化する。また、各視点で得られた画像を処理する際には、全体の処理速度は最も処理の遅いプロセスにより制限されることになる。これにより、システムの処理効率が低下する。さらに、複数の観測を同時に行うことは観測間の冗長性を増大させる。また、多数のカメラ全てをあらかじめキャリブレーションしておくことは、視点数(カメラ数)が増えるにしたがって困難になる。これら同期を前提としたシステムを持つ問題は、利用する視点数が増加するに伴って、より顕著になると考えられる。

我々は、多視点画像による追跡システムにおいては、これら規模の拡大によって生じる問題が本質的であると考え、規模の拡大を容易にする手法

としてカメラ間の同期を前提としない追跡手法を提案してきた [11, 12]。そこでは、非同期に独立に動作する多数の観測ノードで得られる観測情報をカルマンフィルタを用いて統合することにより、カメラ間の同期を必要とせず、拡張が容易な追跡システムを実現している。これにより、観測間の冗長性を低減し視点数の増加による処理効率の低下も抑制できる。

本報告では、位置・速度以外も対象として非同期多視点画像による追跡アルゴリズムを示し、観測・発見・追跡の3種類の処理ノードからなる人物追跡システムについて述べる。さらに位置追跡・動作追跡についての追跡例を挙げ本手法の有効性を示す。

2 追跡アルゴリズム

2.1 運動モデル

ここでは、追跡の対象として平面上の2次元運動および対象物体に関する付加的な情報(人の身長や服の色など)を考える。時刻 t における物体 h_j の位置を $X_{x_j, t_n} (= (X_{j, t_n}, Y_{j, t_n}))$ とする。ここで、 X_{h_j}, Y_{h_j} はそれぞれ世界座標 $X-Y$ 上での物体位置を示す。時刻 t における位置・速度状態は、4次元ベクトル $[X'_{x_j} \quad \dot{X}'_{x_j}]'$ で表される。

時刻 t_n における状態 $[X'_{x_j, t_n} \quad \dot{X}'_{x_j, t_n}]'$ が既知のとき、時刻 t_{n+1} における状態 $[X'_{x_j, t_{n+1}} \quad \dot{X}'_{x_j, t_{n+1}}]'$ は次の線形変換で表される。

$$\begin{bmatrix} X_{x_j, t_{n+1}} \\ \dot{X}_{x_j, t_{n+1}} \end{bmatrix} = F_x \begin{bmatrix} X_{x_j, t_n} \\ \dot{X}_{x_j, t_n} \end{bmatrix} + \Delta t \cdot w_{x_j, t_n}. \quad (1)$$

ここで、 $w_{\Delta t}$ は時間 $\Delta t (= t_{n+1} - t_n)$ の間に系に加わる外乱ベクトルで平均 0 、共分散が R_x とする。

また、 F_x は時間 Δt の間の位置・速度の推移を表す以下の行列である。なお、本システムでは、運動モデルとして等速直線運動を仮定している。

$$F_x = \begin{bmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}. \quad (2)$$

同様に、位置・速度以外の付加的な情報に関する状態ベクトルを X_{s_j} とし、時刻 t_n から時刻 t_{n+1}

への状態変化を次式で表す。

$$\mathbf{X}_{s_j, t_{n+1}} = \mathbf{X}_{s_j, t_n} + \Delta t \cdot \mathbf{w}_{s_j, t_n}, \quad (3)$$

\mathbf{w}_{s_j, t_n} は時間 $\Delta t (= t_{n+1} - t_n)$ の間に加わる外乱ベクトルである (平均 $\mathbf{0}$ 、共分散 \mathbf{R}_s)。

式 (1)、(3) は次式のようにまとめられる。

$$\mathbf{X}_{j, t_{n+1}} = \mathbf{F} \mathbf{X}_{j, t_n} + \Delta t \cdot \mathbf{W}_{j, t_n}. \quad (4)$$

ここで、 $\mathbf{F} = \begin{bmatrix} \mathbf{F}_X & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_S \end{bmatrix}$ 、 $\mathbf{X}_{j, t_n} = \begin{bmatrix} \mathbf{X}_{x_j, t_n} \\ \dot{\mathbf{X}}_{x_j, t_n} \\ \mathbf{X}_{s_j, t_n} \end{bmatrix}$ 、 $\mathbf{W}_{j, t_n} = \begin{bmatrix} \mathbf{w}_{x_j, t_n} \\ \mathbf{w}_{s_j, t_n} \end{bmatrix}$ 。 \mathbf{I}_S は \mathbf{X}_s と同じ次元数を持つ単位行列とする。 \mathbf{W}_{j, t_n} の共分散行列は以下のように定義される。

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_x & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_s \end{bmatrix}. \quad (5)$$

時刻 t_n における状態 \mathbf{X}_{j, t_n} の予測値が平均 $\bar{\mathbf{X}}_{j, t_n}$ 、共分散 \mathbf{P}_{j, t_n} のガウス分布で表されるものとする、この分布に基づく時刻 t_{n+1} における状態の予測分布の平均 $\bar{\mathbf{X}}_{j, t_{n+1}}$ 、共分散 $\mathbf{M}_{j, t_{n+1}}$ は次式で与えられる。

$$\hat{\mathbf{X}}_{j, t_{n+1}} = \mathbf{F} \bar{\mathbf{X}}_{j, t_n}, \quad (6)$$

$$\mathbf{M}_{j, t_{n+1}} = \mathbf{F} \mathbf{P}_{j, t_n} \mathbf{F}^T + \Delta t \cdot \mathbf{R}_{j, t_n}. \quad (7)$$

2.2 観測モデル

次に、ここで時刻 t_n にカメラ (観測ノード) i により 1 回の観測が行われたとする。この観測により対象物体の位置は画像平面上に投影された特徴点として得られる (図 1)。カメラの位置・姿勢および内部パラメータは既知であるから、これにより対象物体がカメラ位置と画像平面上の特徴点を結んで決定できる直線 (視線) 上に位置することがわかる (図 1)。ここでは、この視線と世界座標の Y 軸のなす角度を φ_{ij, t_n} とする。

$$\mathbf{z}_{x_j, t_n} = \mathbf{H}_x \mathbf{R}_{\varphi_{ij, t_n}}^{-1} \mathbf{X}_{x_j} + \mathbf{v}_{x_j, t_n}. \quad (8)$$

ここで、 $\mathbf{H}_x = [1 \ 0]$ 。 \mathbf{v}_{x_j, t_n} は観測誤差で平均 $\mathbf{0}$ 、共分散行列 \mathbf{Q}_{x_j, t_n} とする。観測誤差はカメラから

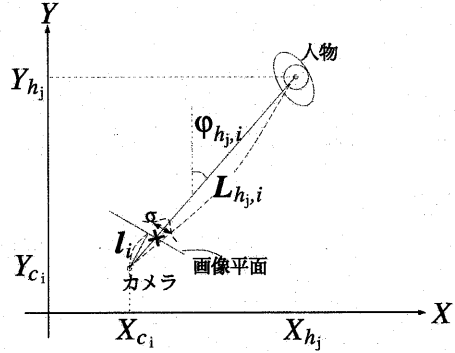


図 1: 観測モデル

の距離が大きくなるほど増加すると考え、次のように表す。

$$\mathbf{Q}_{x_j, t_n} = \frac{\bar{L}_{ij, t_n}}{l_i} \mathbf{Q}_0 \approx \frac{L_{ij, t_n}}{l_i} \mathbf{Q}_0. \quad (9)$$

L_{ij, t_n} はカメラ i -対象物体 j 間の距離、 l_i はカメラ i の焦点距離、 \mathbf{Q}_0 はカメラの観測誤差を示す定数である。なお L_{ij, t_n} は未知のため、 X_{h_j, t_n} の予測状態 \bar{X}_{h_j, t_n} により計算した \bar{L}_{ij, t_n} で近似する。

(8) 式の \mathbf{z}_{x_j, t_n} は、実際には観測された特徴点から得られた回転行列 $\mathbf{R}_{\varphi_{ij, t_n}}$ と世界座標上でのカメラ位置 \mathbf{C}_i から次式により求めることができる。

$$\mathbf{z}_{x_j, t_n} = \mathbf{H}_x \mathbf{R}_{\varphi_{ij, t_n}}^{-1} \mathbf{C}_i. \quad (10)$$

同様に、位置・速度以外の状態についての観測は以下のように表す。

$$\mathbf{z}_{s_j, t_n} = \mathbf{H}_s \mathbf{X}_{s_j, t_n} + \mathbf{v}_{s_j, t_n}. \quad (11)$$

ここで、 \mathbf{v}_{s_j, t_n} は観測誤差を示す (共分散行列 \mathbf{Q}_s)。 (10)、(11) 式は次のようにまとめられる。

$$\mathbf{Z}_{j, t_{n+1}} = \mathbf{H} \mathbf{X}_{j, t_{n+1}} + \mathbf{V}_{t_{n+1}}. \quad (12)$$

ここで、 $\mathbf{Z}_{j, t_{n+1}} = \begin{bmatrix} \mathbf{z}_{x_j, t_{n+1}} \\ \mathbf{z}_{s_j, t_{n+1}} \end{bmatrix}$ 、 $\mathbf{H} = \begin{bmatrix} \mathbf{H}_x \mathbf{R}_{\varphi_{ij, t_n}}^{-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H}_s \end{bmatrix}$ 、 $\mathbf{V}_{j, t_{n+1}} = \begin{bmatrix} \mathbf{v}_{x_j, t_{n+1}} \\ \mathbf{v}_{s_j, t_{n+1}} \end{bmatrix}$ (共分散行列は、 $\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_x & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_s \end{bmatrix}$)。

2.3 状態の更新

観測モデル (式 (12)) によりカルマンフィルタを構成し、時刻 t_{n+1} における対象物体 h_j の状態予

測分布の平均 $\bar{X}_{j,t_{n+1}}$ 、共分散 $P_{t_{n+1}}$ を得る [13]。

$$\bar{X}_{j,t_{n+1}} = \hat{X}_{j,t_{n+1}} + K_{j,t_{n+1}} (Z_{j,t_{n+1}} - H \hat{X}_{j,t_{n+1}}) \quad (13)$$

ここで、

$$K_{j,t_{n+1}} = P_{j,t_{n+1}} H' Q_{j,t_{n+1}}^{-1} \quad (14)$$

$$P_{t_{n+1}} = (M_{j,t_{n+1}}^{-1} + H' Q_{j,t_{n+1}}^{-1} H)^{-1} \quad (15)$$

上式により、一回の観測についての状態の更新を行う。

3 人物追跡システムの構成

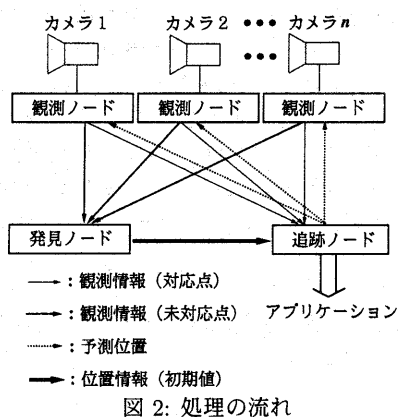


図 2: 処理の流れ

本システムは、図 2 に示すように観測・発見・追跡の 3 種類の処理ノードからなる。観測ノードではカメラ毎に独立して得られた入力画像に対し特徴抽出処理を行う。各観測ノードは非同期に動作し、他の観測ノードに左右されることなく自身の処理能力に応じた処理を行うことができる。現在の実装では、各観測ノードの平均的な処理速度は約 5~6Hz である。発見ノードでは観測ノードから送られてきた未対応点情報を用いて、シーン中に新たに現れた人物の検出を行う。追跡ノードでは、発見ノード・各観測ノードから送られる観測情報をそれぞれ初期値・入力値として各人物情報を更新する。次節以降では、各ノードの処理について詳しく述べる。

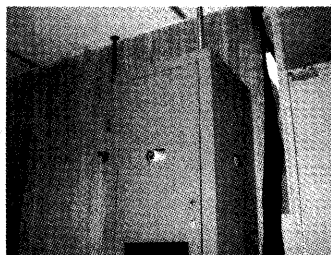


図 3: 観測ノード

3.1 観測ノード

3.1.1 特徴抽出

観測ノード (図 3) では、まずカメラから得られる入力画像を人物領域と背景領域に分割し [14]、得られた人物領域に対して距離変換を行う (図 4)。この距離変換された画像における領域内の各画素は、最近傍の領域境界までの距離を画素値として持つ。領域内で画素値が最大となる点をその領域の特徴点として選択する。図 5 に特徴点の検出例を示す。右図×印が検出された特徴点である。さらに頭頂点・服の色等を検出する。

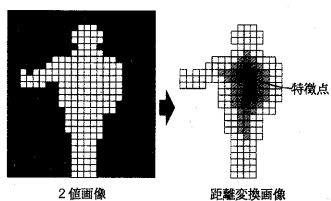


図 4: 距離変換による特徴点検出



図 5: 特徴点検出の例

3.1.2 特徴点の対応づけ

続いて、抽出された特徴点と既に発見されている追跡目標(モデル)間の対応づけの方法について述べる。

例として、図6のようなシーンを考える。時刻 t_a にカメラ l 、続いて時刻 t_b にカメラ m で観測が行われたものとする。ただし $t_a < t_b$ である。このうち2人の人物 h_0, h_1 は時刻 t_a で既にシーン中に存在しており、 h_2 は時刻 t_b で初めて現れたと考える。図中×印は各画像において検出された特徴点を示している。

h_0, h_1 については、追跡ノードへの問い合わせにより時刻 t_a までの観測結果に基づく時刻 t_b の予測状態を得る。時刻 t_n における人物 h_j の状態は2.1節で述べたように平均 $\hat{X}_{j,t_{n+1}}$ 、共分散 $M_{j,t_{n+1}}$ のガウス分布で表される。従って、時刻 t_n におけるカメラ i による観測値の予測分布は、平均 $H\hat{X}_{j,t_{n+1}}$ 、共分散 $HM_{j,t_{n+1}}H'$ のガウス分布となり、これは画像 i 上での人物の存在確率を示す。

$$P_i(Z_i) = \frac{1}{|HM_{j,t_{n+1}}H'|} \exp\left(-\frac{1}{2}(Z_i - H\hat{X}_{j,t_{n+1}}) \cdot (HM_{j,t_{n+1}}H')^{-1}(Z_i - H\hat{X}_{j,t_{n+1}})\right). \quad (16)$$

上記確率を最大にする特徴点を観測時間での人物 h_j に対応する観測値とし、 h_j のラベルをつける。対応づけられた特徴点の位置と観測時刻は、観測情報として追跡プロセスに送信される。ただし、複数の人物に対応づけられた特徴点は、観測時にオクルージョンが発生していると判断し、送信されない。

これらの処理の後、対応づいていない特徴点が存在する場合、これらは未知の人物に属すると考えられる(図中の h_2 がこれにあたる)。未対応点の観測情報(位置・時刻)は発見ノードに送られる。

3.2 追跡ノード

追跡ノードでは、観測ノードから送られてきた観測情報(位置・時刻)に基づき目標人物の位置・速度を更新する。

追跡ノードでは、各観測ノードから観測情報が送られる毎に独立に(6),(13)式による更新処理を行う。特定の時刻の状態予測分布は、観測ノード

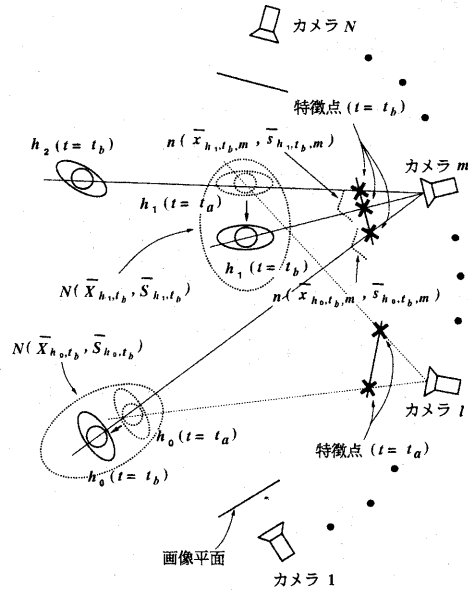


図6: 位置検出

の要求に応じて(6)式により計算され、観測ノードにおいて前節で述べた通り特徴点の対応づけに利用される(3.1.2節)。検出範囲外へ移動した人物モデルは削除し、その人物の追跡を中止する。

発見ノードで新しい人物が発見されると、追跡ノードに新たな人物モデルが生成される。人物状態の初期値は、発見ノードから送信される新規モデル情報によって決定される。

3.3 発見ノード

発見ノードでは観測ノードで未対応となった特徴点をもとに、新たにシーン内に現れた人物を検出し、対応するモデルを追跡ノードに追加する。

本システムでは観測情報を非同期に獲得するため、観測の同時性を仮定した通常のステレオ法をそのまま適用することはできない。そこで前述の等速運動の仮定から次のような時系列情報による対応(発見)手法を用いる。

まず、観測ノードから送られてきた過去 M 回の観測情報(それぞれ1点以上の未対応点を含む)のうち、異なる4時刻の観測情報について未対応点を各一点ずつ選ぶ。それぞれの未対応点の観測時刻・観測したカメラの状態ベクトルを $t_a, t_b, t_c, t_d \cdot C_a, C_b, C_c, C_d$ とする($t_a \leq t_b \leq t_c \leq t_d$)。各観

測情報を入力として2.3節に述べたカルマンフィルタの更新処理を行う。ただし、初期分散 $P_{j,t_1}^{-1} = 0$ とする。

この操作によりこの4回の観測による予測軌跡が得られるので、次にその軌跡を用いて各観測時刻における人物状態を推定する((1)式、但し $\Delta t = t_i - t_d, i \in \{a, b, c, d\}$)。ここで観測値の集合を $\{Z_{t_a}, \dots, Z_{t_d}\}$ 、推定された人物状態の集合を $\{\hat{X}_{t_a}, \dots, \hat{X}_{t_d}\}$ とする。

観測情報の対応関係は、推定された人物状態の(12)式による投影と観測値間のマハラノビス距離の和として定義される次の誤差評価関数 f_{error} により評価することができる。

$$f_{error} = \sum_{i \in \{a, b, c, d\}} \left(Z_{t_i} - H \hat{X}_{t_i} \right)' Q^{-1} \left(Z_{t_i} - H \hat{X}_{t_i} \right) \quad (17)$$

発見ノードでは過去 M 個の観測間で得られる全ての組み合わせについて、 f_{error} を計算する。その値が一定の閾値以下である組み合わせは、新規人物に属する特徴点集合とし、最新観測時刻(ここでは t_d)における推定位置を初期発見位置として追跡ノードに送信する。

4 追跡結果

図8はひとりの人物についての追跡結果の例を示している。予め観測範囲内に人物の移動コースを設定し、その上を歩行する被験者の動きを4台のカメラにより追跡した。ここで破線は設定値、実線は推定結果を示している。ここにみられるように設定値と推定結果のずれはほとんどの個所で10[cm]以下であり、安定した追跡が行われているといえる。図9はこの追跡において、追跡モデルの更新に用いられた観測についてその時刻をカメラ毎に示す。人物の位置の変化に応じて追跡に使われるカメラが動的に変化している様子が分かる。本システムでは各観測ノードが非同期的に観測を行うため各カメラの観測時刻は一致しない。

図10は追跡モデルの更新に用いられた観測について、隣接する観測間の時間差の分布を示している。本実験における観測間隔の分布は約50[ms]付近にピークを持つ。前述の通り各観測ノードの処

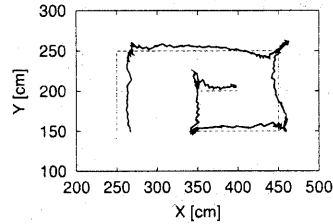


図8: 位置追跡結果(追跡人数: 1人)

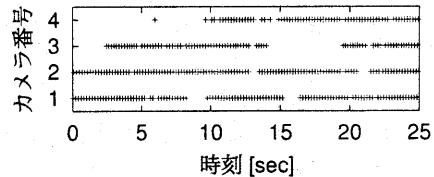


図9: 観測ノード毎の観測時刻(追跡人数: 1人)

理速度は約5~6Hzであるが、独立に動作する複数の観測ノードの情報を統合することにより、追跡モデルの更新が密に行われていることが分かる。

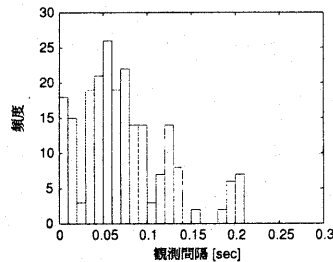


図10: 観測間隔の分布

続いて二人の人物がシーン内で行き交う状況を5台のカメラで追跡した例を示す(図11)。ここで実線は人物1、破線が人物2の追跡結果をそれぞれ示している。図12にカメラ毎にこの追跡において追跡に寄与した観測の時刻をカメラ毎に示す。図中×印は人物1、○印は人物2の追跡に寄与した観測を表している。この例では人物の近接時に複数の視点でオクルージョンが発生している。しかしながら、本手法では多視点情報を利用しているため、このような場合も残る視点の情報により追跡を継続することができる。図7に5台のカメラによる観測時の入力画像の例を示す。

最後に位置・速度以外の例として、身長と簡単な動作検出の結果を図14に示す。本システムでは、

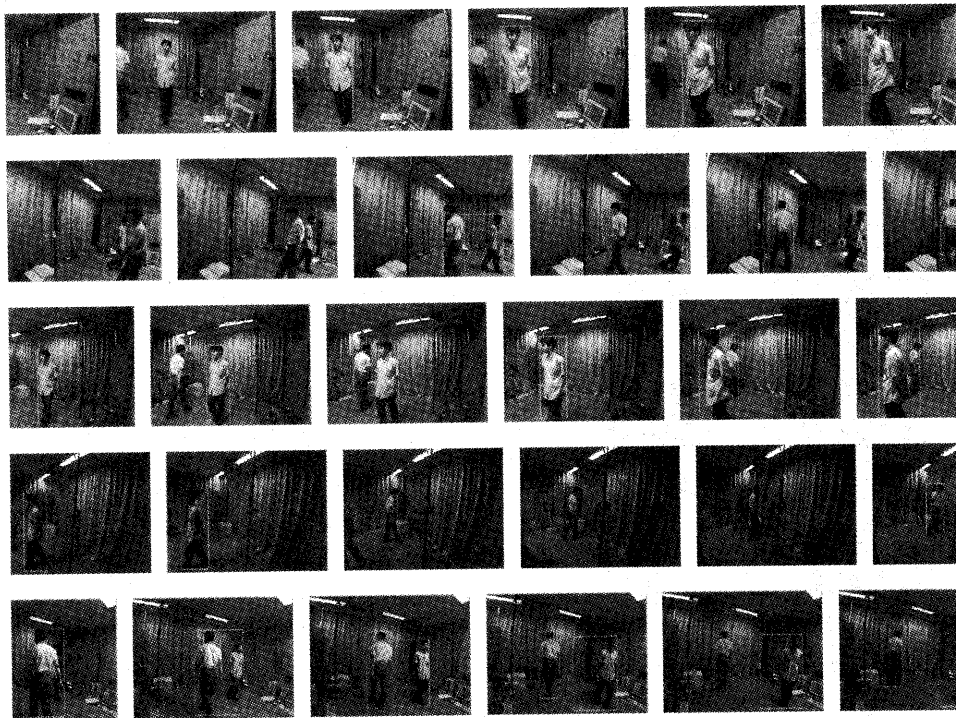


図 7: 各観測ノードでの観測画像の例

速度と身長の変化から Walking, Standing, Sitting の 3 状態 (図 13) を検出する。図 14 下には、状態の検出結果を示している。

以上の実験結果から、本手法により非同期に得られる観測情報を利用して安定した人物追跡が可能であることが示された。

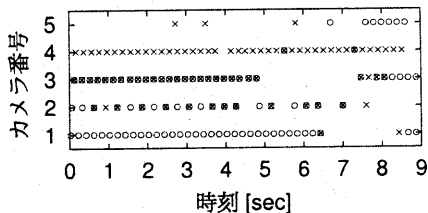


図 12: 観測ノード毎の観測時刻 (追跡人数: 2 人)

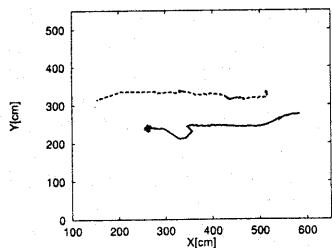


図 11: 位置追跡結果 (追跡人数: 2 人)

5 まとめ

非同期多視点画像を利用した追跡システムについて述べた。本システムでは、各カメラで非同期に観測される画像情報をカルマンフィルタにより統合することで追跡処理を行う。カメラ間の同期をとる必要がないためシステムの実装を簡素化でき、また観測を非同期に行うことから観測間の冗長性を低減し効率的な追跡処理が可能となる。

今後さらに、観測ノードにおける特徴抽出の高度化、カメラ校正の自動化による設置・維持コストの低減など、多視点システムの利用を容易にす

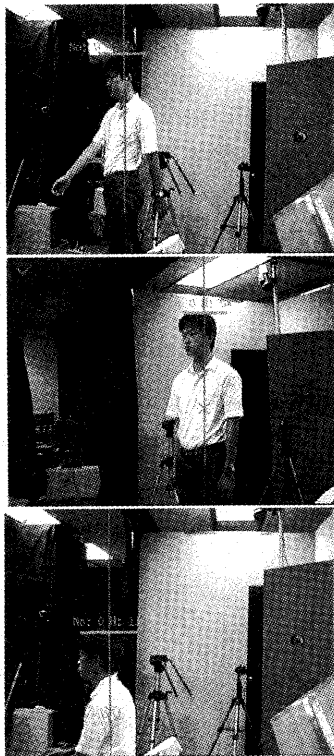


図 13: 人物状態 (上: Walking, 中: Standing, 下: Sitting)

る手法の開発を進めるとともに、個人識別やジェスチャ認識などと組み合わせて、非接触人物動作検出を用いたマンマシンインタフェースなど応用システムの構築を目指したい。

参考文献

[1] D. M. Gavrila and L. S. Davis. 3-d model-based tracking of humans in action: a multi-view approach. In *Proc. of Computer Vision and Pattern Recognition*, pp. 73–80, 1996.

[2] M. Patrick Johnson, P. Maes, and T. Darrell. Evolving visual routines. In *Proc. of Artificial Life IV*, pp. 198–209, 1994.

[3] Ali Azarbayejani and Alex Pentland. Real-time self-calibrating stereo person tracking using 3-d shape estimation from blob features. In *Proceedings of 13th International Conference on Pattern Recognition*, pp. 627–632, 1996.

[4] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland. Pfunder: Real-time tracking of the human body. In *SPIE proceeding vol. 2615*, pp. 89–98, 1996.

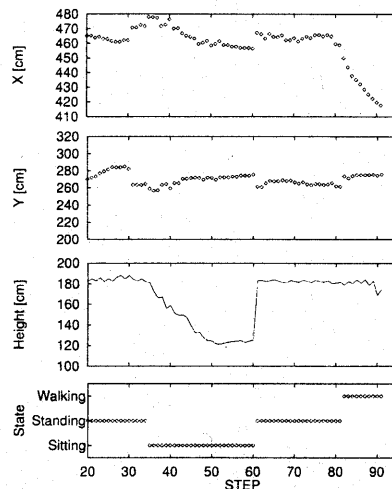


図 14: 状態検出の例

[5] Q. Cai, A. Mitiche, and J. K. Aggarwal. Tracking human motion in an indoor environment. In *Proceedings of 2nd International Conference on Image Processing*, pp. 215–218, 1995.

[6] Jakub Segen and Sarma Pingali. A camera-based system for tracking people in real time. In *Proceedings of 13th International Conference on Pattern Recognition*, pp. 63–67, 1996.

[7] 松山隆司, 和田俊和, 丸山昌之. 能動視覚エージェントによる移動対象の協調的追跡. 画像の認識・理解シンポジウム (MIRU'98) 講演論文集 I, pp. 365–370, 1998.

[8] Q. Cai and J. K. Aggarwal. Tracking human motion in structured environments using a distributed-camera system. *IEEE Pattern Analysis and Machine Intelligence*, Vol. 21, No. 11, pp. 1241–1247, 1999.

[9] 加藤博一, 中澤篤志, 井口征士. 楕円体モデルを用いたリアルタイム人物追跡. 情処学論, Vol. 40, No. 11, pp. 4087–4096, 1999.

[10] Daisaku Arita, Rin ichiro Taniguchi, Satoshi Yonemoto, and Yoshio Hamada. A real-time multi-view image processing system on pc cluster. In *Proceedings of Fourth Asian Conference on Computer Vision*, pp. 270–275, 2000.

[11] 森大樹, 内海章, 大谷淳, 谷内田正彦. 非同期多視点情報を用いた複数人物追跡手法の検討. 信学技報, PRMU98-178, pp. 15–22, 1999.

[12] 森大樹, 内海章, 大谷淳, 谷内田正彦, 中津良平. 非同期多視点画像による人物追跡システムの構築. 信学論 (D-II), Vol. J84-D-II, No. 1, pp. 102–110, 2001.

[13] 加藤寛一郎. 最適制御入門. 東京大学出版会, 1987.

[14] 内海章, 大谷淳, 中津良平. 画素値分布の推定による時系列画像からの動物体領域抽出. 信学論, Vol. J81-D-II, No. 8, pp. 1768–1775, 1998.