

実環境下における自律的な顔画像登録とその応用

加藤 丈和[†] 蔵田 武志[†] 坂上 勝彦[†] 向川 康博[‡] 尺長 健[‡]

[†]産業技術総合研究所 知能システム研究部門
〒 305-8568 茨城県 つくば市 梅園 1-1-1 中央第 2

[‡]岡山大学 情報工学科
〒 700-8530 岡山県 岡山市 津島中 1-1-1

E-mail : t.kato@aist.go.jp

あ ら ま し

本稿では、ネットワークで結ばれたユビキタスカメラやウェアラブルカメラによって、自律的に人物を観測し、顔画像を登録する技術とその応用について述べる。まず、ウェアラブルビジョンシステムにおいて、自律的な情報収集を行う着用型アシスタントである VizWear-Active のコンセプトを紹介し、VizWear-Active による顔画像追跡手法について述べる。また、環境のいたるところに配置したカメラ (ユビキタスカメラ) をネットワークでつないで協調動作させることにより、効果的な顔画像登録を行う分散協調登録システムの概要と、その主な協調メカニズムである動的役割分担と動的エージェンシー構成について紹介する。

キーワード ウェアラブルビジョン, 分散協調視覚, 顔画像登録, 顔画像認識

Autonomous Face Registration and Its Applications in Real Environments

Takekazu Kato[†], Takeshi Kurata[†], Katsuhiko Sakaue[†],
Yasuhiro Mukaigawa[‡], and Takeshi Shakunaga[‡]

[†]Intelligent Systems Institute,
National Institute of Advanced Industrial Science and Technology (AIST),
1-1-1 Umezono, Tsukuba, Ibaraki, 305-8568, Japan

[‡]Department of Information Technology, Okayama University
1-1-1 Tsushima-naka, Okayama, 700-8530, Japan

E-mail : t.kato@aist.go.jp

Abstract

In this paper, we discuss techniques and applications of autonomous face recognition using wearable and ubiquitous cameras connected via networks. We introduce a concept of VizWear-Active and methods of real-time face tracking by VizWear-Active. It can actively and robustly understand the wearer and his or her environment by controlling a wearable active camera according to image processing and motion sensors. We also introduce a cooperative distributed registration system which efficiently register face images by cooperations of the ubiquitous cameras. Two principal cooperation mechanisms of the cooperative distributed registration system, a dynamic role assignment and a dynamic agency configuration, are presented.

Keywords wearable vision, cooperative distributed vision, face registration, face recognition

1 まえがき

近年、空港やビル内、道路など様々な場所にカメラが設置されるようになり、また、携帯電話や PDA などの携帯機器に接続され、着用者が常に持ち歩くことのできるカメラなども発売されている。さらに、無線、有線のネットワークインフラの整備によって、それらのカメラから得られる情報を互いに結びつけることも可能となりつつある。このような各種機材やインフラの整備によって、実環境下でいつでもどこでも視覚情報を獲得することが可能となり、このような情報を有効利用するための研究がさかんに行われている。

カメラを用いた人物観測技術は、セキュリティシステムだけでなく、人間と環境との自然なインタフェースを実現するためにも重要な技術である。これらのシステムの実現のために重要なことは、人物に関する有用な情報を収集、整理し、観測対象人物の状況を把握することである。顔画像は、個人ごとの固有情報であり、自然に観測可能であるという点で優れた人物識別情報であり、またその表情など人物の状況を把握するためにも有用な情報である。

本稿では、人間がカメラや計算機を身につけることで、常に着用者を経験を共にするウェアラブルビジョンシステムと、環境のいたるところにカメラを設置し広範囲かつ多角的な観測を実現する分散協調視覚システム [5] の 2 つの観点から、人物の顔画像を獲得するための技術を紹介し、それらの技術の応用について議論する。

2 ウェアラブルビジョンシステムによる顔画像登録

ウェアラブルシステムは常に着用者やその周囲の状況を把握し、状況に応じて適切な情報を着用者に提示することで、適応的に着用者を支援する潜在能力を持つ。一方、コンピュータビジョン技術は状況把握のために有用な手段であり、ウェアラブルシステムへの応用が期待されている [1, 4]。

本節では、ウェアラブルシステムにおいて、自律的な状況把握と情報収集によって、適応的に着用者を支援する携帯型ペットロボットである *VizWear-Active* のコンセプトと、*VizWear-Active* による顔追跡手法について紹介する。

2.1 *VizWear-Active*

現在、多くのウェアラブルシステムでは、頭部など体の一部に固定されたカメラを使用している。頭部に着用したカメラでは、視線方向が着用者の頭部の動きに追従するため、着用者が注目する対

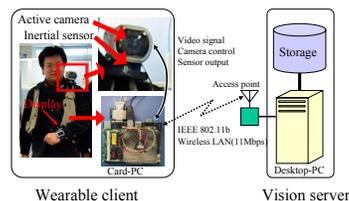


図 1: *VizWear-Active* の試作システム

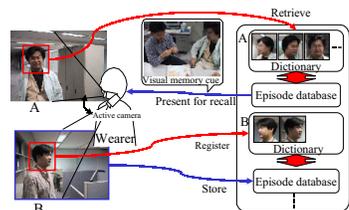


図 2: 記憶補助アプリケーション

象を自然に観測することが可能である。しかし、このような固定型のカメラでは、着用者の動きによって映像にぶれが生じたり、着用者の視線方向にない情報を見逃してしまうなどの欠点がある。特に人物を対象とする場合、着用者自身も対象人物も移動するため、対象人物が常にカメラの視野にいるとは限らない。そこで我々は、ウェアラブルアクティブカメラを備え、画像認識やモーションセンサの情報に基いて自律動作することで、着用者自身の視線方向とは独立に、能動的な情報収集を行うことで、注目対象を安定に観測する *VizWear-Active* を開発している [3, 8]。

図 1 に *VizWear-Active* の試作システムを示す。本システムは、着用者が身につけるウェアラブルクライアント、デスクトップ PC からなるビジョンサーバ、及びそれらを接続する無線 LAN から構成されている。ウェアラブルクライアントは、ウェアラブル PC、アクティブカメラ、モーションセンサから構成される。ウェアラブル PC は軽量かつ小型であり、本システムではベストの中に組み込まれている。アクティブカメラは肩に固定され、ウェアラブル PC からのパン・チルト制御が可能である。また、モーションセンサはアクティブカメラの土台に固定され着用者の動作にともなうカメラの動きを計測することができる。

多くのコンピュータビジョンアルゴリズムは、ウェアラブル PC 単体で処理するには計算コストが高い。また、個人認識処理や画像データの記録などを行うためには、大規模なデータベースが必要であり、これらをすべてウェアラブルクライアントに持たせるのは現実的ではない。本システムでは、このような処理を無線 LAN 経由でサーバに送り、協調的に処理を行うことで、実世界で動作するア

アプリケーションを構築できる。

VizWear-Active のアプリケーションとして、我々は特に顔画像認識技術に基づく記憶補助システムを目指している [3]。図 2 に示すように、システムは着用者が出逢った人物に関するエピソード (時刻、場所、ビデオログなど) を蓄積するエピソードデータベースと、顔画像からエピソードデータベースを検索するための顔辞書を持っている。まず、着用者が他の人物に出逢ったとき、*VizWear-Active* は顔画像を検出、追跡を行いながら、既に登録されている顔辞書を使って個人識別を行う。追跡されている顔画像が顔辞書に登録されていれば、関連づけられたエピソードを着用者に提示する。また、顔画像が未登録であれば、顔辞書を新規に構築するとともに、そのときの状況をエピソードデータベースに蓄積する。

このようなアプリケーションを実現するためには、実環境において安定かつ実時間で顔画像の検出、追跡が行えることが重要である。本節では特に *VizWear-Active* による顔画像検出、追跡技術について述べる。

2.2 *VizWear-Active* による顔追跡

ウェアラブルシステムでは着用者とともにカメラが移動するため、環境変化に対してロバストな追跡法が必要である。そこで我々は、逐次モンテカルロ追跡法 [2] をベースとし、*VizWear-Active* システムに適した顔追跡法を考案した [8]。

2.2.1 逐次モンテカルロ追跡法

逐次モンテカルロ追跡法は、ランダムサンプリングに基づく追跡手法である。人物頭部形状を楕円によって近似し、各時刻の追跡結果を N 個の楕円のサンプルと各サンプルの人物頭部らしさ表す評価値 (尤度) として表現する。ある時刻における楕円のサンプル集合は、直前のフレームのサンプル集合から、各サンプルの評価値を重みとして用いてランダムに選択し、あらかじめ与えた推移確率に従って生成する。

各サンプルは、楕円輪郭上に設定した E 個の観測点を通る法線上のエッジによって評価される。各法線上に 1 次元のエッジ抽出を行い、観測点とエッジ点との距離を求める。そして距離が近いほど評価値が大きく、距離が遠くなるにつれて評価値が減衰するように評価値を定義する。

また、色情報を併用することで、追跡の安定化、高速化をはかる。各時刻で生成されたサンプル集合に対して、色ヒストグラムの評価し、あらかじめ学習した標準の色ヒストグラムと大きく異なるサンプルを削除しておく。こうすることによって、余分なエッジが大きく存在するような複雑な背景に対し

て安定になり、またエッジの評価に対して色の評価の計算コストが低いことから高速化を実現できる。

このように、逐次モンテカルロ追跡法では、サンプル集合とそれらの評価値によって表現される尤度分布として各時刻の追跡結果を得ることができる。これによって、ノイズや一時的な隠れなどに対してロバストな追跡が実現可能である。なお、ある時刻において 1 つの解が必要な場合には、各サンプルをそれぞれの評価値で重み付けして平均することによって推定値を求める。

2.2.2 分散逐次モンテカルロ追跡法

VizWear-Active システムにおいて、アクティブカメラの制御を伴う処理を行うためには、遅延の少ない高速なフィードバックが不可欠である。ここでは、このようなカメラ制御を伴う動作を反射運動と呼ぶ。一方で、追跡対象人物から顔画像を抽出し認識、登録処理を行うためには、高精度な結果が必要である。

逐次モンテカルロ追跡法によって、安定かつ高精度な追跡を実現するためには、ランダムサンプリングの回数を増す必要があるが、計算コストが高くなるためウェアラブルクライアントのみで実時間処理を行うのは容易ではない。また、入力画像をそのままビジョンサーバへ送り、サーバ側で追跡処理を行う場合には、無線 LAN による遅延が生じるため、反射運動に必要な反応速度を得ることが困難となる。そこで、ウェアラブルクライアントとビジョンサーバに分散してサンプリング、評価を行うことで高速かつ高精度な追跡を実現する。

具体的には、ウェアラブルクライアントでは追跡を継続するために必要最小限のサンプリングのみを行い、同時にビジョンサーバによって追加的にサンプリングを行うことで高精度な追跡結果を得る。ウェアラブルクライアントは、各時刻において推定される対象人物位置が画像の中心になるようにカメラを制御することで、カメラを対象人物に追跡させながら追跡する。このとき、遅延の少ないフィードバックを実現するためにクライアントの結果のみを用いてカメラ制御を行う。一方、ビジョンサーバでは、大量かつ密度の高いサンプリングを行い精度の高い追跡結果を得る。また、追跡結果として得られる楕円内部からさらに詳細な解析を行うことによって、登録、認識のための顔領域を獲得する。

2.2.3 カメラ動作の補正

カメラ制御や着用者自身の動作によりカメラの姿勢が変化すると、追跡対象のみかけの変化が大きくなり、追跡を行うことが困難となる。そこで、カメラの制御パラメータとモーションセンサによ

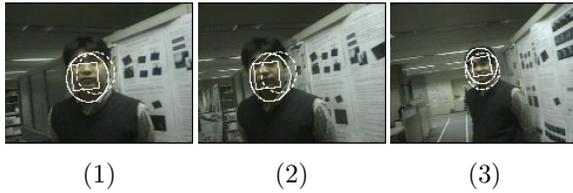


図 3: 人物追跡結果

て得られるカメラの角度変化に基づいて追跡結果を補正することによってこの問題を解決する。

時刻 t におけるカメラパラメータを $\theta_t^c, \phi_t^c, \psi_t^c$, モーションセンサから得られるカメラの姿勢を $\theta_t^m, \phi_t^m, \psi_t^m$ とする。このとき時刻 $t-1$ から時刻 t までのカメラの角度変化 $\dot{\theta}_t, \dot{\phi}_t, \dot{\psi}_t$ は次式によって与えられる。

$$\begin{aligned}\dot{\theta}_t &= (\theta_t^c - \theta_{t-1}^c) + (\theta_t^m - \theta_{t-1}^m) \\ \dot{\phi}_t &= (\phi_t^c - \phi_{t-1}^c) + (\phi_t^m - \phi_{t-1}^m) \\ \dot{\psi}_t &= \psi_t^m - \psi_{t-1}^m\end{aligned}\quad (1)$$

これらの角度変化を用いて楕円の中心位置、及び傾きを補正する。他のパラメータ、つまりスケール及び偏平率に関しては、カメラの角度変化に対する変化が微量であるのでここでは無視する。時刻 t におけるあるサンプルの中心位置を (x_t^h, y_t^h) 、傾きを θ_t^h とすると、カメラの角度変化を考慮して補正されたサンプル位置 $(x_t'^h, y_t'^h)$ と傾き $\theta_t'^h$ は次式によって求められる。

$$\begin{aligned}x_t'^h &= f \tan(\arctan(\frac{x_t^h}{f}) + \dot{\theta}_t) \\ y_t'^h &= f \tan(\arctan(\frac{y_t^h}{f}) + \dot{\theta}_t) \\ \theta_t'^h &= \theta_t^h + \dot{\psi}_t\end{aligned}\quad (2)$$

ただし、 f は焦点距離とする。この補正されたパラメータを用いて次の時刻のサンプルを生成することによって、カメラ動作に対して安定な追跡を実現する。

2.3 実験結果

本節で述べた追跡手法を *VizWear-Active* の試作システムに実装し、実験を行った結果を示す。図 3 に追跡結果を示す。破線の楕円がウェアラブルクライアントによる追跡結果、実線の楕円がビジョンサーバで追加的なサンプリングを行った結果を示す。なお実験ではウェアラブルクライアント上で 200 回、ビジョンサーバ上で 1000 回のサンプリングを行った。

ウェアラブルクライアントによる 1 フレームの処理時間は約 0.15 秒であり、着用者が対象人物とすれ違うように歩きながら処理を行っても見失うことなく追跡を行うことができた。また、ウェアラ

ブルクライアントの結果にはスケールや位置のずれが生じているが、ビジョンサーバによって修正されていることがわかる。また図中の四角は、楕円から顔画像を抽出した結果を表わしている。多少の位置ずれがあるものの比較的安定に顔画像の追跡が行われていることがわかる。

これらの結果より、実環境下において、実時間で比較的安定に顔画像の追跡が実現できていることがわかる。しかし、現在の精度ではそのまま顔認識等に用いることは困難であり、また動きが激しいときなどに失敗することもあるため、今後、顔部品抽出などのさらに高精度な後処理を行い、これらの誤差や失敗を検出、修正する必要がある。

2.4 考察

本節では、ウェアラブルアクティブカメラを備えた着用型アシスタントである *VizWear-Active* のコンセプトと試作システムを紹介した。また、逐次モンテカルロ追跡法に基づき、*VizWear-Active* に適した追跡手法により、実時間で顔画像追跡を行う手法を示した。

ウェアラブルシステムの限られたリソースで安定かつ実時間の追跡を実現するために、クライアント/サーバで分散してランダムサンプリングを行う分散逐次モンテカルロ追跡法を提案した。また、着用者の動きやカメラ制御によって追跡が不安定になる問題を解決するために、モーションセンサの情報やカメラパラメータに応じて追跡処理を補正する方法を示した。実際に追跡アルゴリズムを *VizWear-Active* の試作システムに実装して実験を行い、本手法の有効性を示した。

本稿では顔画像の追跡に焦点を当ててきたが、本手法を用いた記憶補助アプリケーションの実現のためには、得られた顔画像の誤差や失敗の修正を含む登録処理、及びそれを用いた認識処理について検討する必要がある。また、現在の試作システムで用いているアクティブカメラは、常に着用するには大きすぎる上に、パン・チルト範囲が狭く、ウェアラブルシステムに適していない。現在、小型で可動範囲の広いウェアラブルアクティブカメラを用いた次期試作システムを開発中である。

3 分散協調登録システムによる顔画像登録

分散協調登録システム [7] は、分散協調視覚の枠組 [5] に基づいて、環境のいたるところに設置された多数のエージェント (カメラ) が協調動作することによって、それぞれのエージェントの役割や観測対象を動的に切り換え、観測範囲内に登場する人

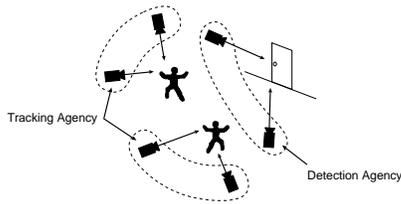


図 4: 分散協調登録システム

物を効率良く観測するシステムである。本節では、分散協調登録システムの概要を紹介し、その主な協調メカニズムである動的役割分担と動的エージェント構成について述べる。

3.1 分散協調登録システム

図 4 に示すように、分散協調登録システムは観測エージェント（以下エージェント）とネットワークによって構成される。エージェントはパン・チルト・ズームを制御可能なアクティブカメラと、顔画像検出、追跡機能、及び通信機能を持つプロセッサによって構成される。各エージェントは自律動作が可能であり、また通信によって互いに協調動作することができる。同じ目的を持つエージェントの集合をエージェントと呼び、エージェントを構成するエージェントをメンバエージェントと呼ぶ。エージェントには一定の観測範囲を持つ検出エージェントと、特定の人物を追跡する追跡エージェントが存在する。

分散協調登録システムの目的は、顔画像認識技術に基づく広域人物行動監視である。そのためには、観測範囲内に登場する人物を検出、追跡しながら、その顔画像を効率良く獲得することが重要である。顔画像は対象人物の姿勢や撮影条件によってその見え方が変化するため、顔画像認識などに有用な顔画像集合を獲得するためには、その時々を対象人物の状況に応じて効果的な観測を行うことが必要である。そのために本システムでは、各エージェントが互いに協調することによって、対象人物の状況を把握しながら、それぞれの役割や観測対象を動的に切り換えることで、効果的な観測を実現する。

3.2 動的役割分担

観測範囲内に登場する人物を追跡しながら、有用な顔画像の登録を行うためには、対象人物を見失わないための安定な追跡と、効果的な顔画像の登録を両立する必要がある。

安定な追跡を実現するためには、人物の急激な移動にも対応できるように広角のカメラによって観測することが望ましい。一方、高精度な認識のた

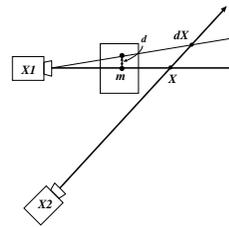


図 5: 3次元位置推定の誤差

めには、高解像度の顔画像を必要とするため、人物をアップで観測することが望ましい。しかし、広角で観測することとアップで観測することを一台のカメラで同時に実現することは困難である。そこで、各エージェントに追跡の役割と登録の役割の2つの役割を与え、それらを状況に応じて動的に変化させながら、それぞれの役割に従ってカメラのズームを変化させることで、システム全体として安定な追跡と効果的な登録を実現する方法を提案した [9]。

追跡を行うためのエージェントを追跡モードのエージェントと呼び、登録を行うためのエージェントを登録モードのエージェントと呼ぶ。追跡モードでは、ズームを広角にして観測を行うことで、対象人物を安定に追跡すること優先し、また登録モードではズームを望遠にして対象人物をできるだけアップで観測することを優先する。

3.2.1 追跡の安定性

複数エージェントが同一対象を観測している場合、対象人物の位置情報をエージェント間で交換し、その3次元位置を推定しながら追跡を行うことで、各エージェントの失敗を補いながら追跡することができる [7]。このとき、3次元位置を安定に推定することができれば、全体の追跡が安定になる。そこで、複数エージェントによる3次元位置推定の安定性を調べることによって、全体の追跡の安定性を評価する。

3次元位置推定の安定性は、各エージェントの追跡の安定性と、エージェントと追跡対象の位置関係によって定まる。例えば、2個のエージェントによって3次元位置を推定する場合、エージェントと対象人物が1直線に並ぶ場合、3次元位置推定は不安定になるが、垂直に並ぶ場合、3次元位置推定は安定でとなる。そこで図 5 に示すように、まず各エージェントの追跡の安定性を画像上での誤差として表現し、その誤差によって生じる3次元位置の誤差を全体の安定性の評価値とする。

3.2.2 登録の有効性

登録の有効性とは、登録された顔画像集合が認識のためにどの程度有効であるかを評価するもの

である．どのような顔画像集合が有効であるかは認識手法によって異なるが，一般に顔の向きや照明条件など様々な顔画像を揃えることができれば，安定な認識を実現することができる．ここでは簡単に，顔認識に用いるために必要な条件を備えた顔画像の枚数によって，登録の有効性を評価する．顔画像が満たすべき条件とは，1) 顔の向きが正面向きから一体範囲内であることと，2) 画像中の顔のスケールが一定値以上であることの2点とする．この1),2)の条件を満たす顔画像の枚数を登録有効性の評価値とする．

3.2.3 エージェントの役割分担

各エージェントの役割を固定した状態で追跡，及び登録を行うと，エージェントと対象人物の位置関係によっては効果的な登録が行えない状況となる可能性がある．そこで，各エージェントの役割を追跡の安定性と登録の有効性に従って動的に変化させることによって，安定な追跡と効果的な登録を両立する．

多くのエージェントを登録モードとして登録を行うと，短時間でたくさんの有効な顔画像を獲得することができる．しかしこの場合，追跡モードのエージェントが少なくなり，追跡が不安定となるため，追跡に失敗する危険がある．つまり，安定な追跡と効果的な登録を両立するためには，追跡の安定性を一定値以上に保ったまま，できるだけ多くのエージェントを登録モードにする必要がある．

本システムでは，初期状態においては全てのエージェントを追跡モードとする．そして，エージェントと対象人物の位置関係から，有効な顔画像を獲得可能なエージェントを選択し，その中から追跡の安定性の評価値に貢献していないエージェントを順に登録モードに変更して，追跡の安定性が閾値以上である間，登録モードのエージェントを増していく．また，追跡の安定性が閾値以下になった場合には，追跡の安定性を保つために必要なエージェントを登録モードのエージェントから選択し，順に追跡モードに変更していく．

3.3 動的エージェント構成

観測範囲内に1人の対象人物しかいない場合は，観測可能な全てのエージェントで同一人物を観測することにより，様々な方向の顔画像を獲得することが可能である．しかし，同時に複数の人物が観測範囲内に存在する場合，どのエージェントがどの人物を観測するかによって，登録される顔画像に偏りが生じる可能性がある．このように偏りが生じた場合，ある人物に関しては十分な登録画像が集り安定な認識が可能になるが，別の人物に関しては登録画像が不十分であり安定な認識ができない

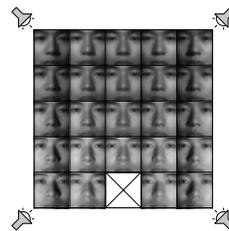


図 6: 標準照明パターン集合

いという状況になってしまう．そこで，各エージェントの観測対象をその時々登録状況に応じて切り換えることで，複数の人物について偏りなく，かつ効率良く顔画像を獲得する手法を提案した [7]．

3.3.1 エージェント移動プロトコル

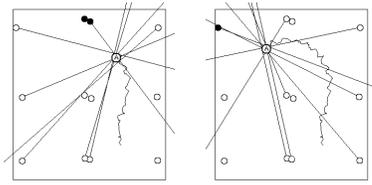
本システムでは，特定の人物を観測するエージェントの集合を追跡エージェントと呼び，1まとまりのグループとして扱う．複数の対象人物を観測する場合には，それぞれ対象人物に対応した追跡エージェントを生成することによって，各対象人物の観測を行う．つまり，どのエージェントがどのエージェントに属するかというエージェント構成によって，各エージェントの観測対象が決まる．

そこで，エージェント間の協調動作によって，各人物の登録状況を把握しながら，それに応じてあるエージェントに属するエージェントを別のエージェントに移動させることによって，各エージェントの観測対象を切り換え，効率の良い登録を実現する．そのためのエージェント間の協調メカニズムをエージェント移動プロトコルと呼ぶ．

エージェント移動プロトコルの具体的な手順について説明する．まず，各エージェントはそれぞれの観測対象人物に関する登録画像集合を，顔の向きによって分類しそれぞれの登録状況を常に評価しておく．ここで，ある向きに関しての登録画像が不足している場合，その向きの顔画像を観測可能なエージェントを他のエージェントに要求し，エージェントを移動させることで不足している方向の顔画像を獲得する．これらの処理を繰り返すことにより，各人物に関して偏りのない登録を実現する．

3.3.2 登録画像集合の評価

安定な認識を実現するためには，様々な見え方の顔画像が必要であるが，顔画像の見え方を変化させる大きな要因として，顔の向きと照明条件が挙げられる．顔の向きに関しては，登録画像集合をいくつかの向きに分類し，各向きごとに評価を行うことで，全ての向きに関する評価が高ければ，十分に様々な向きの顔画像が登録されていると判断することができる．また，照明条件に関しては，



(a) シーン 1 (b) シーン 2

図 7: エージェントのモードと人物位置



図 8: 登録モードで観測された顔画像

図 6 に示すようなあらかじめ容易した標準的な照明パターンを用いて評価する．これを標準照明パターン集合と呼び，多数の人物についていくつかの典型的な照明条件のもとで，顔画像を撮影し，各照明条件ごとに平均をとることで生成する．

現在の登録画像集合から部分空間を生成し，標準照明パターン集合の各画像との DFPS(distance from feature space)[6] を求める．もし，登録画像集合が様々な照明条件の顔画像を含んでいれば，すべての DFPS 小さくなり，照明条件に偏りがあれば足りない照明条件の画像との DFPS が大きくなる．そこで，これらの DFPS の最大値を登録画像集合の評価値として用いる．

3.4 実験結果

3.4.1 動的役割分担

12 個のエージェントを用いて 1 人の対象人物を動的役割分担を行いながら観測した結果を示す．実験ではエージェントが配置された部屋の内部を対象人物が一周する間の観測を行った．図 7 に観測途中の各エージェントのモードと観測対象の位置を示す．図中の黒丸は登録モードのエージェントを示し，白丸は追跡モードのエージェントを示している．また，エージェントから引かれた線は各エージェントの追跡結果をカメラからの方位として示している．A と書かれた丸は推定された対象人物位置を表し，対象人物位置まで続く折れ線は，対象人物の軌跡を表している．シーン 1 とシーン 2 を見比べると，エージェントのモードを変更しながら，顔を撮影可能なエージェントが登録モードとなっていることが分る．

また，図 8 に登録モードのエージェントによって観測された顔画像の例を示す．この結果から顔画像を十分な大きさと捉えて観測していることが分る．

3.4.2 動的エージェント構成

次に，実際に 2 人の対象人物が同時に室内を移動する場合に，動的エージェント構成によって各エージェントの対象人物を切り換えながら観測を行った結果を示す．この実験では全部で 8 台のカメラを用い，2 人の対象人物が登場する場合の観測を行った．

図 9 に，このときのカメラと対象人物の関係を示す．図中の T_A が対象人物 (Target-A, B)， E_1 が検出エージェント， E_2 と E_3 がそれぞれ Target-A, B に対する追跡エージェントを表す．また，追跡エージェントから対象人物に向けて引かれている直線は，各エージェントの追跡結果を表わしている．時刻 t_1 において，最初の対象人物 (Target-A) が，時刻 t_4 において 2 番目の対象人物 (Target-B) が，それぞれ検出されている．その後，時刻 t_6 から t_{11} の間，各エージェントは対象人物 (Target-A, B) を切り換えながら追跡を続けている．これらの結果から，対象人物が複数存在する場合は，顔画像の登録状況に応じて各カメラが対象人物を切り換えながら追跡できていることがわかる．

3.5 考察

本節では，分散協調登録システムのコンセプトを紹介し，本システムによって観測範囲内に登場する人物の顔画像を効率良く獲得する方法について述べた．まず，エージェントに追跡と登録の 2 つの役割を持たせ，それらを動的に切り換えながら，カメラのズームを適切に設定することで安定な追跡と効果的な顔画像登録を同時に実現する方法について述べた．また，観測範囲内に複数人物が登場する場合において，顔画像の登録状況に応じてエージェント構成を動的に変化させることで，効果的な登録を実現する方法を示した．

本システムの最終的な目的は，広域を対象とする人物監視システムであり，そのためには，本節で述べた顔登録技術にもとづいて，人物の登録と認識を統合したシステムとそれに基づく具体的なアプリケーションの構築が必要である．

4 むすび

本稿では，ウェアラブルビジョンシステムと分散協調視覚システムの 2 つのシステムにおいて，人物の顔画像を効果的に獲得する方法を紹介し，またそれらの技術の応用について議論した．人物情報としての顔画像は，個人識別や表情認識などの技術とともに用いることで，セキュリティやインタラクションなど様々な分野に利用可能な優れた情報であり，今後このような技術を用いたアプリ

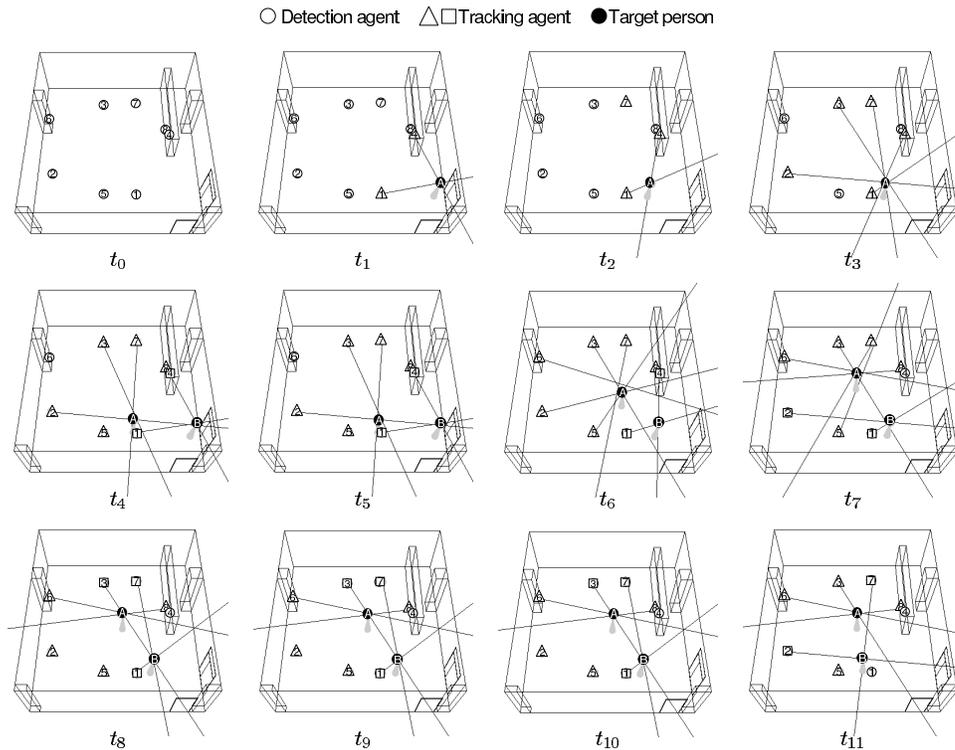


図 9: エージェント構成の変化

ケーションやサービスの提案を行っていくことが必要である。

ウェアラブルビジョンシステムの利点は、着用者と常に経験を共にすることであり、着用者自身の状況を把握し、適切な情報を着用者に提示することで、適応的に着用者の支援を行う潜在能力を持っていることである。一方、環境のいたるところにカメラを配置し、協調的に観測を行う分散協調視覚システムの利点は、観測対象を多角的かつ詳細に観測可能なことと、部屋の状況など環境に固有の情報を扱うことが可能であることである。これらの利点は相補的であり、それぞれの技術を融合することによって、さらに有用なアプリケーションやサービスを実現できる可能性を持つと考えられる。

謝辞

本稿の内容のうち、ウェアラブルシステムに関する研究は、文科省科学技術振興調整費の支援による。また、分散協調視覚システムに関する研究は、日本学術振興会未来開拓学術推進事業 (JSPS-RFTF 96P00501) の補助を受けて行った。記して謝意を表す。

参考文献

[1] <http://www.is.aist.go.jp/vizwear/>.
 [2] Micheal Isard and Andrew Blake. Condensation – conditional density propagation for visual track-

ing. *International Journal of Computer Vision*, Vol. 29, No. 1, pp. 5–28, 1998.

[3] Takekazu Kato, Takeshi Kurata, and Katsuhiko Sakaue. Face registration using wearable active vision systems for augmented memory. In *DICTA2002*, pp. –, Jan 2002. (to appear).
 [4] Takeshi Kurata, Takashi Okuma, Masakatsu Kourogi, Takekazu Kato, and Katsuhiko Sakaue. VizWear: Toward human-centered interaction through wearable vision and visualization. In *PCM2001*, pp. 40–47, 2001.
 [5] Takashi Matsuyama. Cooperative Distributed Vision – Dynamic integration of visual perception, action, and communication. In *Image Understanding Workshop*, pp. 365–384, Monterey, CA, November 1998.
 [6] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. *IEEE Trans. Pattern Anal. & Mach. Intell.*, Vol. vol.19, No. no.7, pp. 696–710, July 1997.
 [7] 加藤文和, 向川康博, 尺長健. 安定な顔認識のための分散協調登録. 電子通信学会論文誌 D-II, Vol. J84-D-II, No. 3, pp. 500–508, Mar 2001.
 [8] 加藤文和, 蔵田武志, 坂上勝彦. VizWear-Active - 記憶補助のための顔画像検出, 追跡, 登録 -. 画像情報システム研究会, 情報メディア学会, pp. 41–46, Dec 2001.
 [9] 亀山直也, 加藤文和, 向川康博, 尺長健. カメラの役割変更による安定な人物追跡と効果的な顔画像登録. 信学技報 PRMU2000-140, pp. 75–80, Dec 2000.