分散協調視覚による顔画像の効果的な登録

加藤 丈和　　向川 康博　　尺長 健

岡山大学工学部情報工学科
〒 700-8530 岡山市津島中 3-1-1
E-mail: kato@chino.it.okayama-u.ac.jp

Effective Face Registration by Cooperative Distributed Vision

Takekazu KATO　　Yasuhiro MUKAIGAWA　　Takeshi SHAKUNAGA

Department of Information Technology , Faculty of Engineering,
Okayama University
Tsushima-naka 3-1-1 , Okayama , 700-8530 , Japan

あらまし: 本稿では，安定な顔認識のために必要な登録画像集合を，効率良く獲得する方法について述べる．実環境において安定な顔画像認識を実現するためには，照明条件や顔の向きなど，多様な条件のもとで撮影された顔画像が登録されていなければならない．対象人物を特定できない場合，すべての人物をあらかじめ登録しておくことができないため，登録されていない人物が新たに観測された時に，その人物を追跡しながら登録する必要がある．したがって，自由に移動する対象人物の顔画像を短時間で効率良く登録することが重要となる．我々は，多数の分散カメラを用いて複数人物の登録を行い，カメラと登録する対象人物との関係を動的に変化させることで効果的な登録を実現する分散協調登録を提案する．8 台のカメラを用いた実験系を構築し，提案手法の有効性を示す．

キーワード: 分散協調視覚，人物追跡，顔画像認識，顔画像登録

1　はじめに

人物監視は，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イスを実現するためにも重要な技術である．建物全体のような広い空間を観測範囲として人物の監視を行う場合，1 台のカメラで全体を観測することは難しい．したがって，多数のカメラを用いて観測し，異なる地点で観測された人物を同定することが必要であり，多様な環境下での安定な人物同定技術が重要となる．顔は実環境において対象人物に意識させることなく自然に撮影可能であることから，本研究では顔画像認識を用いた人物の同定を行う．

従来から，顔画像認識のために様々な手法が提案されている [2],[4],[5],[7]．これらの手法では，対象人物について様々な見え方の顔画像があらかじめ辞書に登録されていれば，ある程度の認識が可能である．しかし，対象人物を特定できない場合，すべての人物についてあらかじめ十分な登録をしておくことはできないため，未登録の人物が観測されたとき，その人物を追跡しながら新たに登録する必要がある．したがって，自由に移動する対象人物を検出・追跡しながら，認識のために必要な顔画像集合を効率良く獲得することが重要となる．従来の顔画像認識の研究では，あらかじめ限定された環境下で十分な登録画像集合が獲得できていることを仮定しており，実環境での登録については十分な検討がなされていない．そこで，本研究では，分散協調視覚 [3] の枠組に基づいて，人物の検出・追跡と顔画像の登録を統合的に扱う分散協調登録システムを提案する．本システムでは，複数の分散カメラの協調動作に基づいて，各カメラが登録する対象人物を動的に変化させることで，効果的な登録を実現する．

2　分散協調登録システムの概要

2.1　エージェントとエージェンシ

分散協調登録システムはパン・チルト・ズームを制御可能なアクティブカメラ，画像処理機能と通信機能を持つプロセッサ，および，ネットワークによって構成される．図 1 に示すように，アクティブカメラとプロセッサの組を構成単位として扱い，エージェントと呼ぶ．各エージェントは自律的に動作可能であり，また，通信によって協調動作を実現できる．

同じ目的を持つエージェントの集合をエージェンシと呼ぶ．逆に，エージェンシを構成するエージェントをメンバエージェントと呼ぶ．図 2 に示すように，分散協調登録システムには 2 種類のエージェンシが存在する．一定の観測範囲を持ち，新たに登場する人物を検出するエージェンシを検出エージェンシ，そのメンバエージェントを検出エージェントと呼ぶ．一方，ある特定の人物を追跡し，顔画像を登録するエージェンシを追跡エージェンシ，そのメンバエージェントを追跡エージェントと呼ぶ．

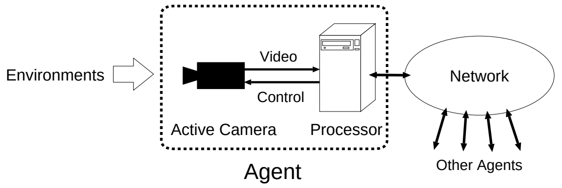


図 1: エージェントの構成

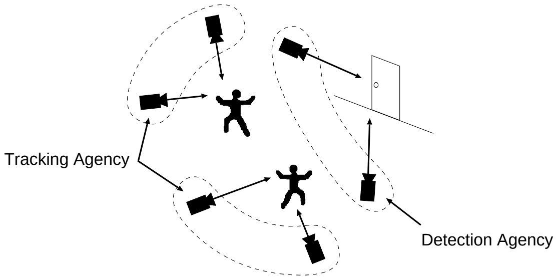


図 2: 追跡エージェンシと検出エージェンシ

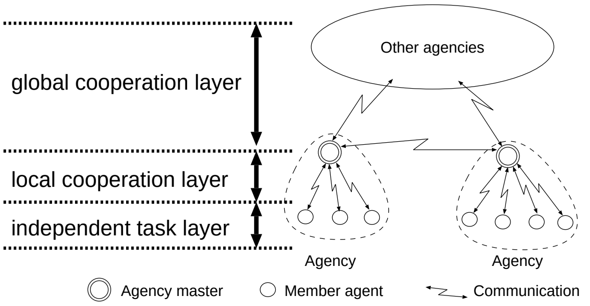


図 3: 3 層構造の協調動作

各エージェンシはエージェンシマスタと呼ぶ特別なエージェントを1つ持つ．エージェンシマスタはエージェンシの代表として働き，メンバエージェントの管理，および，他のエージェンシとの通信を行う．

初期状態においては，検出エージェンシのみが存在し，全てのエージェントは検出エージェンシに属する．検出エージェントが新たな人物を検出したとき，そのエージェントをエージェンシマスタとする追跡エージェンシを新たに生成し，人物の追跡を開始する．また，全てのメンバエージェントが対象人物を追跡できなくなると追跡エージェンシは消滅する．

2.2　3 層構造の協調動作

分散協調システムでは，限られた範囲，対象のみを観測する多数の分散エージェントが協調動作することによって，シーン全体の観測を実現する．安定な人物追跡のためには，実時間性が要求され，エージェント間での密な協調動作が必要になる．一方，効果的な登録のためには，全ての対象人物の登録状態や位置，姿勢などを考慮する必要があり，システム全体での協調動作が重要になる．しかし，システム全体で密に協調動作する場合，通信量が膨大になり，効果的な協調を実現できなくなる危険性がある．そこで，図3に示すような3層構造を導入し，各エージェントの協調する相手を目的に応じて適切な範囲に制限する．

独立処理層 (Independent task layer)

各エージェントは自律的に動作し，エージェント間の協調動作を必要とすることなく，人物の検出・追跡を行う．独立処理層におけるエージェント動作の詳細を3章で述べる．

局所協調層 (Local cooperation layer)

エージェンシ内のメンバエージェント間で局所的に協調動作し，エージェンシの目的を実現する．検出エージェンシでは，エージェンシマスタがメンバエージェントに観測範囲を割り当てる．追跡エージェンシでは，追跡対象の位置情報をメンバエージェント間で共有し，安定な追跡を実現する．局所協調層における協調動作の詳細を4章で述べる．

大域協調層 (Global cooperation layer)

エージェンシ間の大域的な協調動作により，エージェンシの構成を動的に変え，システム全体の動作を調整する．エージェンシ間の協調動作に必要な通信は，エージェンシマスタが代表して行うことで，システム全体の通信量を軽減する．大域協調層における協調動作の詳細を5章で述べる．

3　独立処理層における検出・追跡

3.1　検出エージェントによる人物顔の検出

顔領域検出には，部分空間法 [6] と顔構造解析法 [5] を併用する．部分空間法によって顔領域候補をいくつか検出し，各候補領域に対して顔構造解析法を適用する．顔構造解析法は，顔部品 (目，鼻，口) の位置関係を検証することで顔領域を検出する手法で，高精度な検出を実現できる．

3.2　追跡エージェントによる人物顔の追跡

カメラのパン・チルト制御を伴う追跡を安定に行うためには，顔領域が常に画像の中心になるようにカメラを制御することが望ましい．そのため，高速に処理を行うことが重要となる．一方，顔画像は登録・認識に用いられるため，高精度な顔領域の位置決めが必要である．本研究では，以下に述べる2段階の処理によって，高速かつ高精度な顔領域の追跡を実現する．

第 1 段階　直前フレームで得られた顔領域を用い，テンプレートマッチングにより顔領域の粗い位置決めを行い，得られた顔領域が画像の中心になるようにパン・チルトを制御する．この方法では高速な探索が可能であるが，詳細な顔の位置，および，向きを求めることは難しい．また，テンプレートを更新するたびに位置ずれが蓄積するという問題がある．これらの誤差，位置ずれは次の段階で修正する．

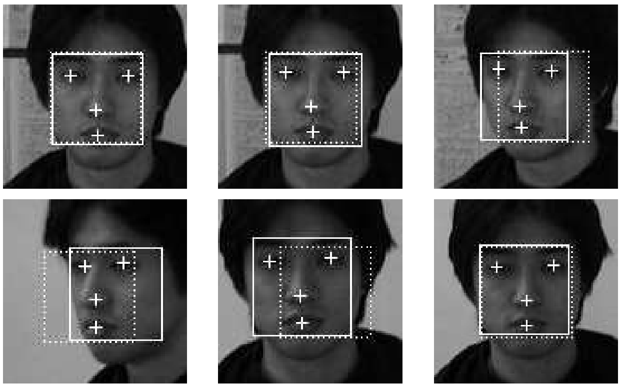


図 4: 独立処理層における追跡結果

第 2 段階 前段階で得られた顔領域に対して顔構造解析法を適用し高精度な位置決めを行う．顔構造解析法によって得られた顔部品の位置に基づいて，顔領域の位置，方向，スケールを修正する．この処理で得られた顔領域を登録・認識に用いる．また，前段階のテンプレートを更新し，誤差，位置ずれを修正する．

3.3 独立処理層における検出・追跡実験

パン・チルト・ズームカメラ (SONY 製 EVI-G20) と PC(PentiumIII-600) によりエージェントを構成し，独立処理層のおける人物顔の追跡実験を行った結果を図 4 に示す．図中の破線の枠は第 1 段階で得られた顔領域を示し，実線の枠は第 2 段階で修正された後の顔領域を示す．また，4 つの ‘+’ は 顔構造解析法によって抽出された顔部品の位置を示す．第 1 段階の処理はビデオレート (33ms/frame) 以内で行うことができ，実時間の追跡が実現された．また，第 1 段階の追跡の誤差が第 2 段階の詳細な抽出の結果によって修正され，高精度な顔領域の位置決めが実現されていることが分かる．

4 局所協調層における協調動作

4.1 検出エージェンシによる観測範囲の分割

検出エージェンシは人物が新たに登場する場所を観測範囲として持つ．エージェンシマスタは観測範囲を各メンバエージェントに分担して割当て，メンバエージェントは割当てられた観測範囲内で人物を検出する．観測範囲を割当てる方法は，シーン中で観測する必要のある範囲，およびメンバエージェントの数と位置によって異なる．本研究では，簡単化のために人物が新たに登場する場所を一ヶ所に限定し，1 つのメンバエージェントによって検出する．

4.2 追跡エージェンシによる協調追跡

各エージェントによる追跡では，対象人物が障害物によって遮蔽される場合や，カメラと対象人物の距離が遠い場合に，追跡が不安定になる．そこで，メンバエージェント間で対象人物位置の整合性をとることによって，安定な追跡を実現する．

(a). 1 つのメンバエージェントのみ追跡に成功している場合，他のメンバエージェントは，対応するエピポーラ線の周辺領域で対象人物を検出する (図 5(a))．顔領域を見付けたエージェントは追跡を開始する．このとき，エージェンシの処理は (b) に移行する．一方，他のメンバエージェントが顔領域を見付けることができないときは，(a) の処理を繰り返す．

(b). 2 つ以上のメンバエージェントが追跡に成功している場合，三角測量の原理で対象人物の 3 次元位置を推定できる．図 5(b) に示すように，対象人物は各メンバエージェントの追跡結果に対応する 3 次元直線の交点に存在する．ただし，各メンバエージェントの追跡結果には誤差，および，各メンバエージェントが対象人物位置を獲得する時刻のずれのため，各直線が厳密に一点で交差する保証はない．そこで，各直線からの距離の二乗和が最小となる点を対象人物の 3 次元位置とする．

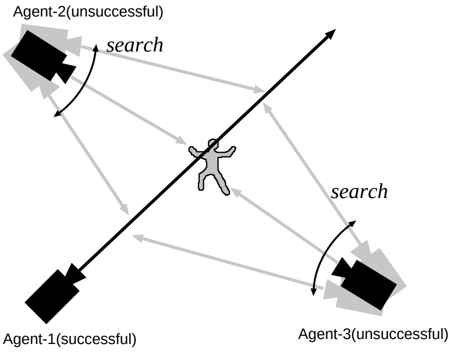
追跡に失敗しているメンバエージェントは，対象人物の 3 次元位置の周辺領域で顔領域の検出を行い，顔領域が見付かれば追跡を開始する．追跡に成功しているメンバエージェントは，追跡結果と対象人物の 3 次元位置を画像上に投影した点との距離が閾値以下の場合，追跡結果が正しいと判断し，追跡を継続する．距離が閾値を超えた場合は，追跡結果が誤っていると判断し，追跡失敗の場合と同様に顔領域の検出を行う．

このように，各エージェントの追跡結果を優先し，追跡に失敗した場合，もしくは追跡結果に整合性がとれない場合にのみ，3 次元位置を用いて追跡を修正することで，各エージェントの追跡誤差，および時刻のずれにロバストな追跡を実現できる．

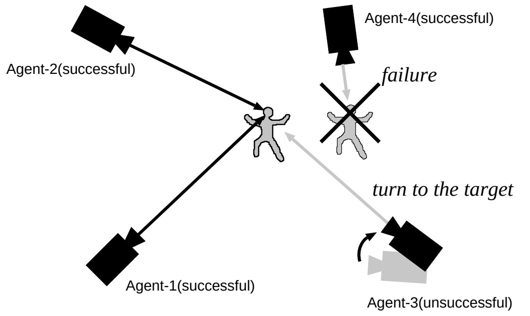
4.3 協調追跡実験

エージェンシ内の協調による追跡の実験結果を示す．4 つのエージェントにより実験系を構成し，各エージェントを 3.3 と同じ構成とする．各エージェント間は 100Base-TX のネットワークにより接続した．図 6 に各エージェントの入力画像系列の例を示す．図中の枠は各エージェントが追跡している顔領域を示す．

時刻 t_1 において Agent-1 が新たな人物を検出し，追跡を開始した．このとき，Agent-3 と 4 は，図 6 の破線で示すエピポーラ線上の探索を開始している．次に，時刻 t_2 において Agent-4 が顔領域を検出し，追跡を開始した．Agent-3 は t_3 において追跡に失敗しているが，時刻 t_7 において修正された．これは Agent-1 と 4 の追跡結果から求められた 3 次元位置に基づいて失敗を検出し，再探索を行ったためである．このように各エージェントの追跡の誤りを他のエージェントの結果を用いて修正し，全体として安定な追跡が実現さ



(a)1 つのエージェントが追跡に成功している場合



(b)2 つ以上のエージェントが追跡に成功している場合

図 5: 3 次元位置を利用した協調追跡

図 6: 協調追跡結果

れていることがわかる．また，時刻 t_{11} において，対象人物が Agent-1 と 4 の観測範囲から外に出たため，これらのエージェントは追跡に失敗するが，その後も Agent-2 と 3 によって追跡は続行されている．

図 7 に，時刻 t_0 から t_{11} のエージェントと対象人物との関係を示す．図中の白い丸印はエージェントの位置を示し，黒い丸印は推定された対象人物の位置を示す．また，エージェントから対象人物に向って引かれた直線は，各エージェントの追跡結果に対応する 3 次元直線を示す．この結果より，推定された 3 次元位置には多少の誤差が含まれているが，全体としては安定な追跡が実現されていることがわかる．

5 大域協調層における協調動作

本システムによって登録された顔画像集合は，部分空間法による顔認識のための学習画像集合として用いることができる．部分空間法によって安定な顔認識を行うためには，顔の向きや照明条件など，多様な条件下で撮影された顔画像が偏りなく登録されている必要がある．複数の対象人物を多数のエージェントで登録する場合，各エージェントと対象人物との関係が固定されていると，得られる顔画像は対象の位置に依存し，多様な顔画像が得られない．更に対象人物があるエージェントの観測範囲の外に出ると登録を行うことができず，そのエージェントは無駄になる．そこで，各エージェントと登録する対象人物との関係，すなわちエージェンシの構成を，より効果的な登録が実現できるように動的に変更する．5.1 節では登録画像集合を評価するための基準を定義し，5.2 節ではエージェンシ構成を調整するためのプロトコルについて述べる．

5.1 登録画像の評価

登録画像集合が安定な認識を実現するために有効であるかを評価するための基準を定義する．安定な認識を実現するためには，顔の向きの変化や照明変動に対応する必要がある．

まず，顔の向きを幾つかに分け，各向きごとに部分空間を生成する．ここで，各向きの部分空間があらゆる照明条件の変化を含んでいる場合，部分空間と任意の照明条件の顔画像との距離がすべて小さくなり，どのような照明条件においても安定な顔認識が可能となる．一方，登録が不十分な場合，不足している照明条件の顔画像との距離が大きくなり，安定な認識を行うことができない．そこで，これらの距離の最大値を登録画像の評価値とし，これが小さいほど理想的な登録画像であると判断する．

実際には対象人物は不特定であるため，各人物に対してあらゆる照明条件の顔画像をあらかじめ準備しておくことはできない．そこで，平均画像を使った標準的な照明パターンを代わりに用いる．あらかじめ多数の人物の顔画像をいくつかの典型的な照明条件のもとで撮影しておき，各照明条件毎に平均画像を求める．この平均画像の集合を標準照明パターン集合と呼び，登録画像集合の評価に用いる．

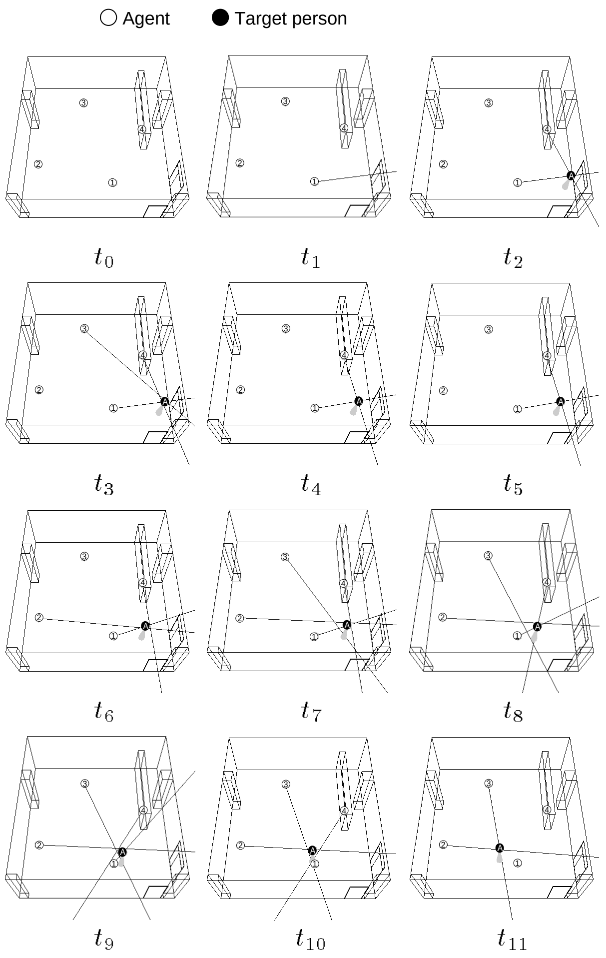


図 7: 協調追跡結果 (3 次元位置)

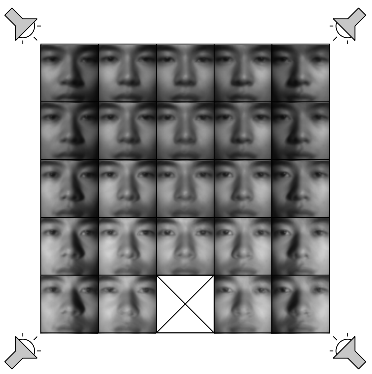


図 8: 標準照明パターン集合 (正面顔)

図8に評価に用いた標準照明パターン集合を示す．左右に5段階，高さを5段階に変化させ，カメラ自身の影となる照明位置を除く24通りの照明条件のもとで撮影を行った

5.2　エージェント移動プロトコル

エージェンシ間でエージェントを移動させることによって，エージェンシの構成を動的に変化させる．移動するエージェントは，エージェント移動プロトコルによって決定する．エージェント移動プロトコルは，以下に示すエージェンシ間のメッセージ交換によって定義する．

Request メッセージ:

各追跡エージェンシは対象人物の顔の向きごとの登録画像集合について，評価値を計算する．ある向きについての評価値が閾値以上である場合，エージェンシ (要求エージェンシと呼ぶ) は，エージェントを要求する Request メッセージを他のエージェンシ (被要求エージェンシと呼ぶ) に送信する．また，要求するエージェントの条件として，必要な顔の向き，登録画像集合の評価値，対象人物の3次元位置，顔の3次元方向をメッセージに付加する．

Ack メッセージ:

Request メッセージを受け取った被要求エージェンシは，メンバエージェントの中からメッセージに付加された条件に合うエージェントを候補エージェントとして選ぶ．ここで，複数の候補がある場合には，エージェンシ内での貢献度を比較し，最も貢献度の小さいエージェントを選択する．なお，貢献度は，そのエージェントが現在観測している対象人物の顔の方向の評価値とする．つまり，評価値が小さいほど，すでに十分な顔画像が得られている方向を観測しているので，エージェンシに対して貢献していないと判断する．ただし，エージェントが対象人物の追跡に失敗している場合は貢献度を0とみなし，これを選択する．

候補エージェントの貢献度が Request メッセージに付加された評価値より小さい場合，各被要求エージェンシは Ack メッセージを要求エージェンシに返す．このとき，候補エージェントの貢献度をメッセージに付加する．

Done，Cancel メッセージ:

要求エージェンシは，Ack メッセージを受けとると，候補エージェントの中から最も小さい貢献度が付加されているものを選択し，そのメッセージの送り主に Done メッセージを送り，実際にエージェントを移動させる．また，選択されなかったエージェンシには，Cancel メッセージを送り，要求を取り消す．

5.3　実験結果

分散協調登録の有効性を確認するため，複数の対象人物が登場する場合の追跡実験を行った．図9に実験環境を示す．図中の白い丸で示す部分は実験に用いたカメラのうちの1台であり，この実験では全部で8台

図 9: 実験環境

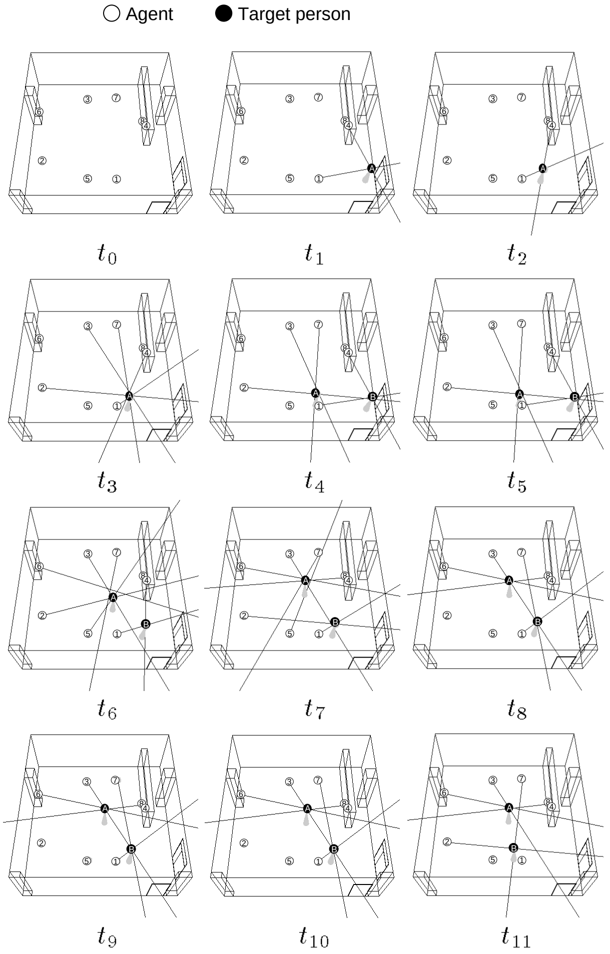


図 10: 大域協調層におけるエージェンシ構成の変更

のカメラを用いた．また，図中の白い四角で示す部分が入口であり，ここから 2 人の対象人物が入室し，それらの人物の追跡，登録を行う．

図 10 に実験結果を示す．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．

6 まとめ

本稿では，効果的な顔画像の登録のための分散協調登録について述べた．エージェンシ構成を動的に変化させることで効果的な登録を実現した．また，3 層構造の協調動作を導入することにより，協調に必要な通信量を軽減することができ，規模の増大にも対応可能であると考えられる．実際に 8 台のカメラを用いた実験系を構築し，提案するシステムの有効性を確認した．

今後の課題として，独立処理層における顔領域の検出・追跡の安定化，認識・登録を統合したシステムの構築とその評価法の確立があげられる．

謝辞

本研究は，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．

参考文献

[1] Haritaoglu, I., Cutler, R., Harwood, D. and Davis, L., “Backpack: Detection of People Carrying Objects Using Silhouettes,” Proc. ICCV’99, pp.102–107, 1999.

[2] Kirby, M. and Sirovich, L., “Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, no. 1, pp.103–108, 1990.

[3] Matsuyama, T., “Cooperative Distributed Vision – Dynamic Integration of Visual Perception, Action, and Communication –,” Proc. Image Understanding Workshop, pp. 365–384, 1998.

[4] Moghaddam, B. and Pentland, A., “Probabilistic Visual Learning for Object Representation,” IEEE Trans. Pattern Analysis and Machine Intelligence, vol.19, no.7, pp.696–710, 1997.

[5] Shakunaga, T., Ogawa, K. and Oki, S., “Integration of Eigentemplate and Structure Matching for Automatic Facial Feature Detection,” Proc. FG’98, pp.94–99, 1998.

[6] Sugiyama, Y. and Ariki, Y., “Facial Region Tracking and Recognition by Subspace Method”, Proc. VSSM’96, pp.225–230, 1996.

[7] Turk, M. and Pentland, A., “Eigenfaces for Recognition,” Journal of Cognitive Neuroscience, vol.3, no.1, pp.71–86, 1991.