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

亀山 直也　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：{kame,kato,mukaigaw,shaku}@chino.it.okayama-u.ac.jp

あらまし
我々は，複数カメラの協調動作によって，人物検出・追跡と顔画像の登録・認識を統合して扱うことのできる分散協調登録システムを提案してきた．本稿では，分散協調登録システムにおいて，カメラの役割変更によって安定な人物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る．各カメラが果たすべき2つの役割として，広角で人物追跡を行う追跡モードと，高解像度の顔画像を得る登録モードを考える．各カメラのモードを動的に切り替えることで，安定な人物追跡と効果的な顔画像登録を実現する．実際に12台のカメラを用いた実験系を構築し，モードを固定した場合と動的な役割変更を行った場合の比較実験を行い，提案手法の有効性を確認した．

キーワード　分散協調登録，人物追跡，顔画像登録，顔画像認識，役割変更

Stable Person Tracking and Effective Face Registration by Dynamic Role Assignment

Naoya KAMEYAMA　Takekazu KATO　Yasuhiro MUKAIGAWA　Takeshi SHAKUNAGA
Department of Information Technology, Okayama University

Tsushima-naka 3-1-1, Okayama,700-8530, Japan
E-mail : {kame,kato,mukaigaw,shaku}@chino.it.okayama-u.ac.jp

Abstract
We have been proposed a cooperative distributed registration system which integrates person tracking, registration and recognition. This paper discusses a method for realizing both stable person tracking and effective face registration by dynamic role assignment to each camera. Each camera has two possible roles in the system. The first role is tracking, in which the camera takes images with a short focal length for stable person tracking. The other is registration, in which the camera takes images with a long focal length to register fine face images. Both the stable person tracking and the effective face registration are realized by switching these roles. An experiment system is constructed using 12 cameras, and the effectiveness of our method is confirmed in comparison with the case of fixed roles.

Keywords　cooperative distributed registration, person tracking, face registration, face recognition, role assignment

1　はじめに

カメラを用いた人物監視は，セキュリティシステムなどに応用可能な技術である．人物監視では，観測範囲内に登場する人物を検出・追跡するだけではなく，人物の詳細な情報を獲得し，個人同定などの技術と組み合わせることが重要となる．本研究では，人物情報として個人同定に有効な顔画像の獲得を目的とする．

広域の人物監視を行う場合，1 台のカメラで全体を監視することは難しいため，多数のカメラを用いる必要がある．我々は分散協調視覚 [1] の枠組に基づいて，複数台のカメラの協調動作により，人物追跡と顔画像登録を行う分散協調登録システム [2] を提案している．しかし，このシステムでは，各カメラで人物追跡と顔画像登録を同時に行っているため，安定に人物追跡を行いつつ，高解像度の顔画像を獲得することが困難であった．そこで本研究では，各カメラに対して追跡と登録の 2 つの役割を与え，これらを動的に切り替えることで，安定な人物追跡と効果的な顔画像登録を両立させる．

2　分散協調登録システム [2]

2.1　システムの目的

分散協調登録システム [2] は，人物の検出・追跡と顔画像の登録・認識を統合して扱うシステムである．観測範囲に現れた人物を検出し，複数のカメラによって追跡する．このとき同時に，対象人物の顔画像の登録，及び認識を行う．

2.2　システムの構成要素

本システムは複数のエージェントからなる．図 1 に示すように，各エージェントはパン・チルト・ズーム制御可能なカメラと画像処理を行う計算機の組合せで構成される．基本的にそれぞれのエージェントは独立に動作するが，エージェント同士がネットワークによって通信することで協調動作することもできる．

2.3　協調追跡

2.3.1　各エージェントの独立動作

各エージェントはそれぞれ独立に人物追跡を行う．人物追跡手法としては，すでに様々な方法が提案されているが，本システムでは顔画像の撮影を目的とし，人物が後ろを向いていても検出・追跡が可能であること，高速な検出処理が必要であることを考慮し，楕円フィッティング [3][4] によって人物頭部の検出・追跡を行う．フィッティングのための評価基準として，人物頭部の色情報と，楕円の円周上の画素の輝度勾配を用いる．また，人物の顔部分を表す肌色領域の重心位置と頭髪部分を表す黒色領域の重心位置から顔の向きを推定する．

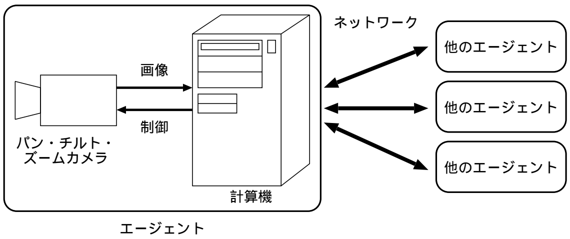

図 1: エージェントの構成

2.3.2　エージェントの協調動作

各エージェントが追跡している対象人物の位置の情報をもとに，三角測量の原理で対象人物の 3 次元位置を推定する．図 2 に示すように，対象人物は画像上での人物位置とカメラの投影中心を結ぶ直線上に存在する．複数のエージェントが対象人物を追跡している場合，各エージェントに対応する直線の交点として対象人物の 3 次元位置を推定することができる．しかし，実際には各エージェントの追跡結果の誤差や，人物を捉えた時刻のずれのため，直線のすべてが一点で交差する保証はない．そこで，図 2 のように，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．

エージェント i のカメラの投影中心を $\boldsymbol{X}_i$，追跡している対象人物とカメラの投影中心を結ぶ直線 l_i の方向ベクトルを $\boldsymbol{V}_i(||\boldsymbol{V}_i|| = 1)$ とするとき，対象人物の 3 次元位置 $\boldsymbol{X}$ と直線 l_i の距離 d_i は次式のように求められる．

$$\begin{aligned} {d_i}^2 &= |\boldsymbol{X} - \boldsymbol{X}_i|^2 - \left(\boldsymbol{V}_i^T(\boldsymbol{X} - \boldsymbol{X}_i)\right)^2 \\ &= (\boldsymbol{X} - \boldsymbol{X}_i)^T(\boldsymbol{I} - \boldsymbol{V}_i{\boldsymbol{V}_i}^T)(\boldsymbol{X} - \boldsymbol{X}_i) \end{aligned} \qquad (1)$$

また，すべてのエージェントに対応する直線との距離の二乗和 ϵ は次式で表される．

$$\epsilon = \sum_i {d_i}^2 \qquad (2)$$

このとき，ϵ を $\boldsymbol{X}$ で偏微分した結果を $\boldsymbol{0}$ とすることで $\boldsymbol{X}$ に関する線形方程式を導出でき，ϵ を最小とする $\boldsymbol{X}$ を求めることができる．

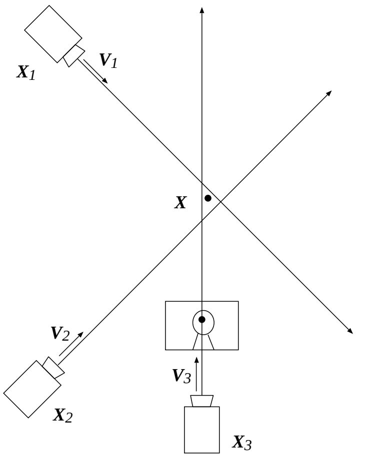

図 2: 対象人物の 3 次元位置

$$\frac{d\epsilon}{d\boldsymbol{X}} = 2\sum_i (\boldsymbol{I} - \boldsymbol{V}_i\boldsymbol{V}_i{}^T)(\boldsymbol{X} - \boldsymbol{X}_i) = \boldsymbol{0} \tag{3}$$

$$\boldsymbol{X} = \left(\sum_i (\boldsymbol{I} - \boldsymbol{V}_i\boldsymbol{V}_i{}^T)\right)^{-1} \sum_i (\boldsymbol{I} - \boldsymbol{V}_i\boldsymbol{V}_i{}^T)\boldsymbol{X}_i \tag{4}$$

このようにして求めた対象人物の 3 次元位置 $\boldsymbol{X}$ を用いて，各エージェントの追跡結果の検証，及び追跡に失敗したエージェントの再検出を行う．

3 エージェントのモード変更

3.1 安定な追跡と効果的な登録

人物の追跡と登録を実現するためには，対象人物を見失わないための安定な人物追跡と，顔画像の効果的な登録が必要となる．本システムでは，役割の動的な変更を実現するため，エージェントに 2 つのモードを考え，これを各カメラで切り替えることによって，安定な追跡と効果的な登録を実現する．

3.2 追跡モードと登録モード

パン・チルト・ズームカメラによって安定な人物追跡を行うためには，人物の急な移動にも対応できるように，ズームを広角にして撮影する必要がある．一方，高精度な認識のためには高解像度の顔画像を必要とするため，ズームを望遠にして撮影した方がよい．このように，撮影時のカメラのズームには，追跡に適した値と登録に適した値があり，これらを一台のカメラで両立させることは難しい．そこで，エージェントに追跡モードと登録モードの 2 つのモードを与え，それぞれのモードに適したズーム値で撮影を行う．

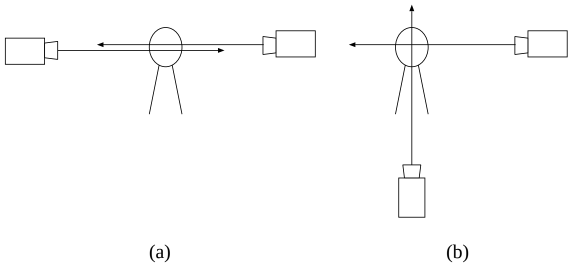

図 3: 人物とエージェントの位置関係

エージェントが追跡モードのとき，ズームを広角にして人物の検出・追跡を行う．このとき，画像上から顔の検出を容易に行うために，顔が一定のサイズとなるように，対象人物とカメラの距離に応じて，ズーム値を適切に変化させる．一方，エージェントが登録モードのとき，ズームを望遠にして撮影し，顔画像の登録を行う．

3.3 追跡評価値と登録評価値

各エージェントを追跡モードと登録モードに分けることで，個々のエージェントは安定な人物追跡と効果的な顔画像登録のいずれか一方を行う．ここで，系全体として，安定な人物追跡と効果的な顔画像登録を両立するためには，各エージェントと対象人物の位置関係や，それぞれのモードの組合せを考慮する必要がある．そこで，系全体として，追跡がどの程度安定しているか，及び登録がどの程度効果的に行えているかを評価するための基準を定義する．

まず，安定な人物追跡について考える．協調追跡では，推定した人物の 3 次元位置を用いて追跡を行うため，人物の 3 次元位置を安定に推定することが重要である．人物の 3 次元位置は，2.3.2 節で述べたように，エージェントと人物を結ぶ直線の交点を求めることにより推定している．このため，図 3(a) のようにエージェントと人物が一直線上に並んだ場合，推定される 3 次元位置は不安定になる．逆に図 3(b) のように垂直に近くなると，推定される 3 次元位置は安定になる．また，人物を捉えたエージェントの台数が多いほど追跡は安定する．このように，追跡の安定性は人物とエージェントとの位置関係と，追跡を行っているエージェントの台数によって変化する．

そこで，各エージェントと人物を結ぶ直線の内積を用いて，追跡の評価値 E_t を式 (5) のように定義

する．この評価値は値が大きい方が追跡が安定であることを示す．

$$E_t = \sum_j \sum_k (1 - |\boldsymbol{V}_j \cdot \boldsymbol{V}_k|) \tag{5}$$

次に，効果的な顔画像登録について考える．必要とされる顔画像集合は，顔画像の認識方法によって異なるが，一般的には顔の向きや照明変動など様々な条件の顔画像を揃えることができれば，どのような状況下でも対応可能な顔画像認識を行うことができる．ここでは簡単化のために，登録された顔画像の枚数によって登録画像集合を評価する．登録画像集合の評価値 E_r を次式のように定義する．

$$E_r = <\text{登録画像枚数}> \tag{6}$$

3.4 モードの切り替え

エージェントの役割を文献 [5] のように固定にすると，エージェントと人物の位置関係によっては安定な追跡や効果的な登録を行うことができない．これは，人物の移動にともなって，エージェントと人物の位置関係が変化していくためである．そこで，本研究ではシーンの状況に応じて各カメラのモードによって役割を動的に切り替えることで，安定な人物追跡と効果的な顔画像登録を両立させる．

登録モードのエージェントを多くすると，短時間で登録の評価値を大きくすることができる．しかしこの場合，追跡モードのエージェントが少なくなるため，3 次元位置の推定結果が不安定となり，系全体での追跡に失敗する場合がある．つまり，安定な追跡と効果的な登録を同時に実現するためには，追跡の評価値を一定以上に保ったまま，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．

安定な追跡を保ったまま効果的な登録を行うために，エージェントのモードが切り替わった後の追跡の評価値と，エージェントの登録に対する貢献度を求める．エージェント i のモードが追跡から登録に切り替わった後の追跡の評価値 $E'_t(i)$ を式 (7) のように定義する．この値が閾値 τ 以上の時，エージェント i を登録モードに切り替え可能であると判断する．

$$E'_t(i) = \sum_{j \neq i} \sum_{k \neq i} (1 - |\boldsymbol{V}_j \cdot \boldsymbol{V}_k|) \tag{7}$$

また，あるエージェント i が必要な顔画像を撮影できるかどうかを，対象人物とエージェントの位置関係によって判断し，エージェント i の登録に対する貢献度 $C_r(i)$ を次式のように定義する．

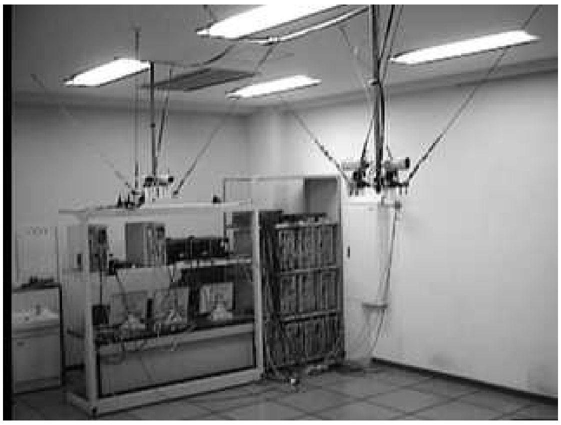
図 4: 実験環境

$$C_r(i) = \begin{cases} 1 & \text{顔画像撮影可能} \\ 0 & \text{それ以外} \end{cases} \tag{8}$$

初期状態では，すべてのエージェントは追跡モードであり，人物の検出及び追跡を行う．人物追跡が始まると，追跡モードの各エージェントについて $E'_t(i)$ と $C_r(i)$ を求め，$E'_t(i) \geq \tau$ かつ $C_r(i) = 1$ となるエージェントが存在するならば登録モードに切り替える．また，登録モードのエージェントについて，$E'_t(i) < \tau$，または $C_r(i) = 0$ となるエージェントが存在するならば追跡モードに切り替える．

4 実験

提案手法の有効性を確認するために，12 個のエージェントを用いて実験を行った．各エージェントをパン・チルト・ズーム制御可能なカメラ (SONY 製 EVI-G20) と PC(PentiumIII 750MHz Dual) で構成し，各カメラは図 4 のように室内に固定した．また，各エージェント間は 100Base-TX のネットワークによって接続した．各カメラの位置はあらかじめキャリブレーションによって求めている．図 5 に各カメラの位置及び撮影可能な角度を示す．実験では，対象人物が部屋の中を一周する間の追跡及び登録を行った．なお，登録対象とした顔画像は，顔の正面を中心として左右 75 度の範囲とし，モードの切り替えに用いる閾値 τ は 1.6 とした．

図 6 に実験中のエージェントと人物の位置関係の一例を示す．図中の丸印に囲まれた A は推定された人物の位置を表す．白丸は追跡モードのエージェントを，黒丸は登録モードのエージェントをそれぞれ表す．また，エージェントから人物の方向に

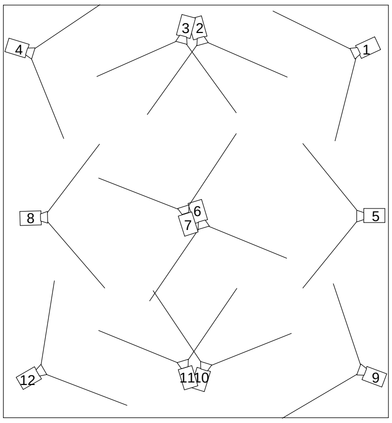

図5: カメラの配置と撮影範囲

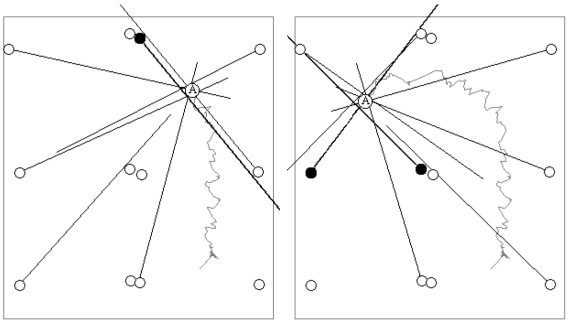

図6: カメラと人物の位置関係 (モード変更)

引かれた直線は，エージェントが人物を捉えたときのエージェントと人物を結ぶ直線を表している．人物の移動にともない，適切な顔画像の撮影ができるエージェントが登録モードになり，それ以外のエージェントが追跡モードになっていることが分かる．

図7に追跡モードのエージェントによる入力画像と追跡結果を示す．図中の楕円が，追跡している頭部領域を表している．また，図8に登録モードのエージェントによって登録された顔画像の一例を示す．追跡モードのときの入力画像に比べ，十分な解像度の顔画像が登録されていることが分かる．

図9にこのときの追跡評価値と登録評価値の時間的推移を示す．図中の実線が追跡評価値，破線が登録評価値を表している．この図より，ほぼすべての時刻において，安定な追跡を行いながら，顔画像の登録が行えていることが分かる．また，13秒の辺りで追跡評価値が一時的に下がっているが，その後追跡が安定になるようにエージェントのモード変更が行われ，追跡評価値が上昇している．

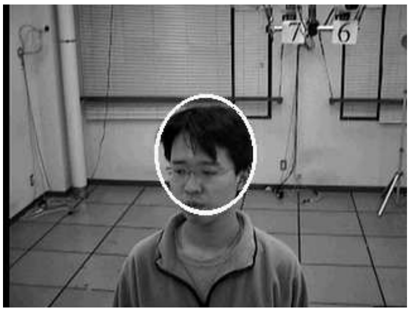

図7: 楕円フィッティングによる顔領域の検出

図8: 登録された顔画像

また，提案手法と比較するためにエージェントのモードを固定にして実験を行った．カメラの配置から，追跡・登録ともに疎になる空間がないように，エージェントのモードを決めた．図10にこのときのエージェントと対象人物の位置関係を示す．図6と同様に，図中の白丸が追跡モードのエージェント，黒丸が登録モードのエージェントである．対象人物はほぼ同じコースを通り部屋を一周している．このとき，人物がどの位置にいても，エージェントのモードが変わることはないため，登録に対して貢献していないエージェントが存在していることが分かる．

それぞれの手法によって得られた追跡評価値のグ

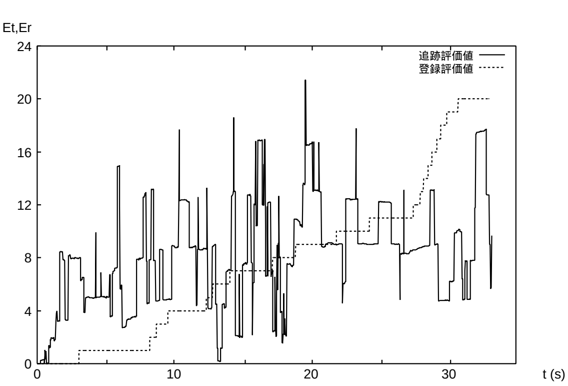

図9: 追跡評価値と登録評価値

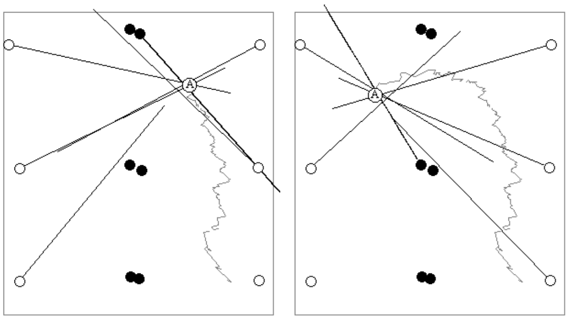

図 10: カメラと人物の位置関係 (モード固定)

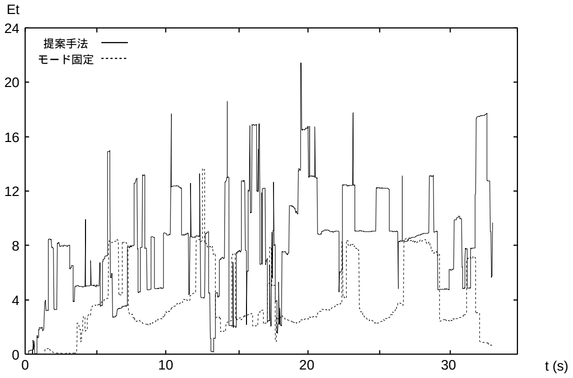

図 11: 追跡評価値の比較

ラフを図 11 に示す．提案手法の方が，全体的に高い値を示していることがわかる．また，同様に登録評価値のグラフを図 12 に示す．最終的な登録枚数は，提案手法が 20 枚，モード固定の手法は 17 枚であり，それほど差は表れていないが，各時刻における登録枚数の差は大きい．これは，モードを固定した場合は，対象人物と登録モードのエージェントとの位置関係によって顔画像の登録を行えない場合があるのに対し，提案手法では，適切なエージェントが登録モードになって撮影を行ったためである．

また，図 6 と図 10 を比較すると，人物を捉えている登録モードのエージェントの台数はほぼ同じであるが，人物を捉えている追跡エージェントの数が提案手法の方が多いことが分かる．これは，提案手法の方がより安定な人物追跡が可能であることを示している．

5 おわりに

本稿では，エージェントに追跡モードと登録モードの 2 つのモードを与え，そのモードを動的に切り替えることによって，安定な人物追跡と効果的な顔画像登録を同時に実現する方法を提案した．また，実際に 12 個のエージェントを用いてモードを固定にした場合との比較実験を行い，提案手法の有効性を確認した．今後の課題として，実際の認識方法に適した登録の評価値の定義，及び複数人物への対応などが挙げられる．

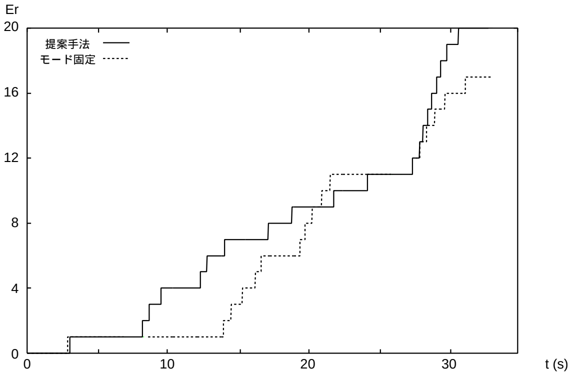

図 12: 登録評価値の比較

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] 松山隆司：“分散協調視覚”, 情処研報 CVIM 121-6,pp.41-48,2000.

[2] 加藤丈和，向川康博，尺長健：“分散協調視覚による顔画像の効果的な登録”, 画像の認識・理解シンポジウム (MIRU2000),vol.I,pp.445-450,2000.

[3] 永田亮一，川口剛：“膚の色情報と頭のだ円近似を用いる顔検出”, 画像の認識・理解シンポジウム (MIRU2000),vol.II,pp.121-126,2000.

[4] 谷内清剛，和田俊和，松山隆司：“視点固定型パン・チルト・ズームカメラを用いた適応的見え方モデルに基づく人物頭部の検出・追跡”, 画像の認識・理解シンポジウム (MIRU2000),vol.I,pp.9-14,2000.

[5] 伊藤渡，上田博唯：“広域監視用親カメラと追尾監視用子カメラによる協調監視システム”, 画像の認識・理解シンポジウム (MIRU2000),vol.II,pp.229-234,2000.