

## カメラセンサネットワークにおけるモバイルエージェントを用いた適応的手法選択型人物追跡システムの設計と実装

田中孝浩<sup>†</sup> 小川剛史<sup>††</sup> 沼田哲史<sup>†</sup> 板生知子<sup>†††</sup> 塚本昌彦<sup>†</sup>  
西尾章治郎<sup>†</sup>

<sup>†</sup> 大阪大学大学院情報科学研究科 〒565-0871 大阪府吹田市山田丘 1-5  
<sup>††</sup> 大阪大学サイバーメディアセンター 〒560-0043 大阪府豊中市待兼山町 1-32  
<sup>†††</sup> NTT 未来ねっと研究所 〒180-0012 東京都武蔵野市緑町 3-9-11  
E-mail: †{tanaka,numata,tuka,nishio}@ist.osaka-u.ac.jp, ††ogawa@ime.cmc.osaka-u.ac.jp,  
†††tomoko@ma.onlab.ntt.co.jp

**あらまし** 本稿では、広域にわたって分散配置された多数のカメラやセンサから得られる情報を利用した人物追跡システムの構築について述べる。本システムでは、モバイルエージェントがカメラセンサネットワーク内を移動する人物に追従し、背景差分法や形状認識法といった従来手法を適応的に組み合わせて人物の認識を行う。手法の組合せを動的に変更することで、周囲の環境や人物の外見的な変化に対してロバストな追跡を実現した。さらに、実装したシステムの試験運用による性能評価について述べる。

**キーワード** 人物認識, 従来手法の動的な組合せ, モバイルエージェント

## Design and Implementation of a Human Tracking System Using Mobile Agents in Camera and Sensor Networks

Takahiro TANAKA<sup>†</sup>, Takefumi OGAWA<sup>††</sup>, Satoshi NUMATA<sup>†</sup>, Tomoko ITAO<sup>†††</sup>, Masahiko  
TSUKAMOTO<sup>†</sup>, and Shojiro NISHIO<sup>†</sup>

<sup>†</sup> Grad. Sch. of Info. Science and Tech., Osaka Univ. 1-5, Yamadaoka, Suita, Osaka, 565-0871 Japan  
<sup>††</sup> Cybermedia Center, Osaka Univ. 1-32, Machikaneyamacho, Toyonaka, Osaka, 560-0043 Japan  
<sup>†††</sup> NTT Network Innovation Laboratories 3-9-11, Midoricho, Musashino, Tokyo, 180-0012 Japan  
E-mail: †{tanaka,numata,tuka,nishio}@ist.osaka-u.ac.jp, ††ogawa@ime.cmc.osaka-u.ac.jp,  
†††tomoko@ma.onlab.ntt.co.jp

**Abstract** In this paper, we explain the design and implementation of a human tracking system using data obtained by a number of cameras and sensors. In this environment, we use a mobile agent technology to track people. Mobile agents identify each person by combination and adaptation of conventional techniques. According to the situation, the mobile agents dynamically change the combination to achieve robust tracking. Finally, we show the experimental results of our system in a real environment which show the availability of the system.

**Key words** human identify, dynamic combination of conventional techniques, mobile agent

### 1. はじめに

カメラを用いて特定の対象を追跡するシステムに関して多くの研究が行われている。対象追跡システムは、例

えば、建物内や車内に設置した監視カメラの映像から不審な動作をする人物を自動検出する監視システムや、店舗内における客の動線を把握し、客の移動傾向を反映し

た位置へ商品を配置する販売戦略など、さまざまな用途への応用が考えられている。そのほか、人や車輛の通行量調査や室内における人物の位置管理などへの利用も考えられ、カメラを用いた対象追跡システムの必要性は今後ますます高まるものと考えられる。さらに、カメラ画像だけでなく、各種のセンサから得られるデータを統合して利用できれば、対象追跡システムのより多くの環境への適用が可能となる。

そこで本研究では、広域に分散配置された多数のカメラやセンサから得られる情報を利用した人物追跡システムの構築を目的とする。多岐にわたる用途に応用可能な対象追跡システムを構築するためには、次のような問題点や要求事項が挙げられる。

まず、追跡対象がカメラやセンサといった分散リソース間を移動したり、複数の追跡対象が重なったりしても、追跡中の個々の対象同士を識別しながら追跡を続行する必要がある。本システムでは、モバイルエージェントによって人物認識および追跡を行う。モバイルエージェントを用いることで、分散リソース間を移動する個々の人物に追従できるほか、追跡システムにおける人物の消失や新たな人物の出現に応じたモバイルエージェントの消滅・生成といった動作の実現が可能となり、システムにおける対象人物の管理が容易となる。

次に、分散リソースによる広域での人物追跡を行う場合、得られる情報量が膨大になるため、これらの情報を統合し人物位置把握や人物認識を行うためには多くの計算資源が必要となる。そこで、豊富な計算資源の提供が可能な Grid 環境を適用することにより、直接利用する計算機が非力であっても人物追跡システムの構築が実現できる。

最後に、分散リソースから得られる多様な情報を統一して利用できるフレームワークが必要である。加えて、情報の有用性の低下などによる状況の変化への動的な対応が求められる。本稿では分散リソースから得られる情報の動的な統一利用を行うフレームワークを提案する。提案するフレームワークでは、背景差分法や形状認識法といった従来手法を複数組み合わせることで人物認識を伴った追跡を実現する。人物追跡中は、用いる手法の組合せを動的に切り替えることにより、周囲の環境やシステムの状況の変化に影響されにくい人物認識を行う。こうした手法の組合せの切り替えや、得られた情報の統合による人物認識はモバイルエージェントにより行う。

2. 節では、カメラによる追跡対象システムにおいて用いられる従来手法に関する関連研究を挙げ、その問題点を明らかにする。3. 節では、構築するシステムの概要を

示し、システムで用いるアルゴリズムについて 4. 節で詳しく説明する。実装したシステムの実環境下での運用による評価について 5. 節で述べ、最後に 6. 節で本稿をまとめる。

## 2. 関連研究

カメラによる対象追跡システムを実現するために、カメラから得られる画像から追跡の対象となる物体を抽出する手法に関する研究が行われている。フレーム間差分法 [1] では、1 フレームまたは数フレーム前の画像と現在の画像との差分をとることで動きのある物体を抽出し、背景差分法 [3][4] では、あらかじめ追跡対象のいない背景画像を取得しておき、現在のフレーム画像との差分をとることで物体を抽出する。頭部抽出法 [7] では、画像内で円形に見える部分を人間の頭部と認識して追跡する。また、抽出した個々の物体を識別する手法としては、さまざまな角度から撮影した人物の顔画像をあらかじめデータベースに登録しておき、得られたカメラ画像とのマッチングをとる手法 [5] や、人物の肌や髪の毛、衣服の色の組合せによって識別を行う手法 [2] などがある。

しかし、こうした手法にはいくつかの問題点がある。フレーム間差分法は、追跡対象が静止している場合に、背景差分法は、日照条件や照明などにより画像全体の輝度や色相が変化した場合に手法の有効性が失われる。頭部抽出法は、ボールなど、球形の物体を人間の頭部と誤認識する可能性がある。顔画像による識別法は、データベースとのマッチングのため、カメラで人物の顔を適切な角度から適切なサイズで観察しなければならない。そこで、顔画像による識別に関する研究では、パン・チルト・ズーム制御が可能なアクティブカメラが比較的多く用いられるが、人物ごとにカメラの制御が必要となるため、1 台のカメラで複数人物の観察を行っている場合、その同時追跡は困難である。

## 3. 人物追跡システムの概要

本研究では、人物追跡および人物認識を実現するための、複数の従来手法の動的な組合せを行うアーキテクチャを考案し、人物追跡システムを設計した。今回、システムの構築にあたっては対象物体に含まれる色情報を用いて人物認識を行うが、設計したシステムでは、その他にも形状認識やパターンマッチングといったさまざまな手法を柔軟に利用できるフレームワークとなっている。以後、人物認識や人物追跡に用いる手法をフィルタと呼ぶ。

本システムでは、1 台の PC とそれに接続された 1 つのカメラからなるシステムをカメラシステムと呼ぶ。分散配置された複数のカメラシステム同士は、ネットワークによって互いに接続されている。人物認識および人物追

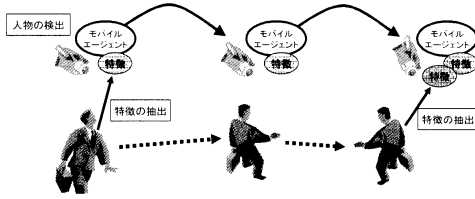


図1 システムの概略

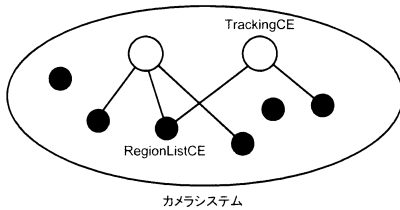


図2 モバイルエージェント

跡を行うモバイルエージェントは、カメラシステム上において動作する。

システムの概略を図1に示す。カメラによって人物が検出されると、そのカメラシステム上にモバイルエージェントが生成される。モバイルエージェントは、髪の毛や衣服の色、背の高さなど、追跡対象となる人物から抽出された特徴に基づき、その人物の追跡に最適な手法の組合せを決定する。人物が、あるカメラの観察可能領域から離脱し、他のカメラの観察可能領域に侵入した場合、モバイルエージェントは人物の移動に追従してカメラシステム間を移動し、フィルタを用いた追跡を続行する。追跡中に、人物の新たな特徴を抽出したり、現在使用しているフィルタの有用性が低下した場合は、フィルタの組合せを最適化し、よりその人物の追跡に適したフィルタに更新する。モバイルエージェントは、このフィルタの組合せによって追跡中の人物の識別を行う。

次に、モバイルエージェントの概略を図2に示す。カメラシステム上には、RegionListCE(CyberEntity)、TrackingCEと呼ぶモバイルエージェントが存在する。1つのRegionListCEは1つのフィルタを管理し、フィルタによってカメラ画像から抽出された結果を矩形情報として保持する。TrackingCEは、カメラシステム上に存在するRegionListCEの中から必要に応じて適切なものを選択して自分の子とし、子のもつ抽出結果情報を統合して人物認識を行う。そのため、基本的にTrackingCEと人物とは1対1に対応付けられる。

システムの動作の概要を図3に示す。カメラによって人物が検出されると、そのカメラシステム上にモバイルエージェントが生成される。人物追跡を行う際は、抽出された特徴に基づいてその人物の追跡に適したフィルタ

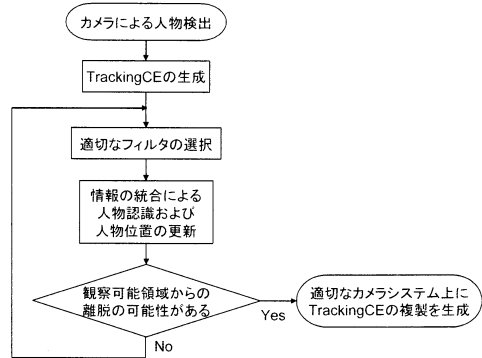


図3 システムの動作

を選択し、選択した各フィルタから得られた情報を統合することによって人物認識および人物位置の更新を行う。人物がカメラの観察可能領域から離脱し、他のカメラの観察可能領域に移動しそうな場合は、次に人物が現れると予想されるカメラシステム上にTrackingCEのコピーを生成し、人物が移動した後でも同様の追跡を続行する。

以上のようなアーキテクチャにより、汎用性が高く、周囲の環境やシステムの状態の変化に影響されにくいシステムの構築が可能になると考えられる。さらに、IDカードなどによって付加的に個人情報が得られる環境下であれば、TrackingCEによって各個人に合わせたサービスを提供するシステムも実現できる。

#### 4. アルゴリズム

本節では、動的なフィルタの組合せによる人物認識アルゴリズムや、モバイルエージェントの具体的な動作など、設計したシステムで用いているアルゴリズムについて詳しく述べる。

##### 4.1 特徴の抽出とTrackingCEの生成

カメラシステムによって人物が検出された場合、まずその人物の特徴を抽出し、追跡に適したフィルタの組合せを決定する必要がある。本研究では、髪の毛や衣服など、追跡対象となる人物に含まれる色の組合せによって人物認識を行う。以下では、カメラから得られるフレーム画像から色情報を抽出し、その情報に基づいてTrackingCEを生成するアルゴリズムについて説明する。

本システムでは、背景差分法などの従来手法において一般的に用いられるRGB値ではなく、色相値によって色を表現する。色相を用いることで、日照条件や照明条件などによる画像全体の輝度や彩度の変化に影響されずに、人物に含まれる厳密な色味の抽出が可能となる。色相は、 $R = 0^\circ, G = 120^\circ, B = 240^\circ$ である色相環において、 $0^\circ$

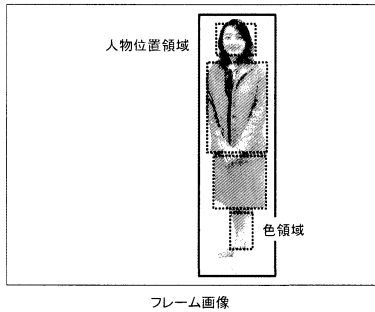


図4 人物初期位置の設定とフィルタリスト情報の抽出

からの角度で表す。

まず色相環をいくつかの小領域に等分割する。各 RegionListCE が管理するフィルタは、色相環のある小領域内に含まれる色相値をもつピクセルをフレーム画像から抽出し、近隣のピクセル同士を統合して、フレーム画像から注目する色が存在する領域を求める。RegionListCE は、得られた領域のバウンディングボックスを矩形情報としてリストに保持する。

TrackingCE は、追跡対象となる人物の位置と、その人物の追跡に必要な一連の RegionListCE (以後、フィルタリストと呼ぶ) の情報をもつ。TrackingCE をカメラシステム上に生成する際は、人物の初期位置とフィルタリストの情報をあらかじめ与えておく。TrackingCE は、フィルタリスト情報によって自身の子にすべき RegionListCE を決定する。人物の初期位置の設定とフィルタリスト情報の抽出の概略を図4に示す。カメラによって人物が初めて検出された場合は、まずそのときのフレーム画像を取得し、現在の実装では人手によってフレーム画像における人物の位置を矩形として設定する。次に、設定した人物位置領域 (実線) において、顔や衣服など、1つの色として比較的まとまりのある領域を色領域 (点線) として設定する。カメラシステムは、それぞれの色領域内のピクセルが最も多く含まれる小領域を色相環から見つけることで、フィルタリスト情報を得る。さらに、TrackingCE は、自身のフィルタリストの RegionListCE が管理するフィルタが人物位置領域に対してどの部分の領域を抽出するものであるかを、TrackingCE の生成時にカメラシステムから与えられたパラメータによって把握する。以後、フィルタが人物位置領域において色領域を抽出すべき領域を担当領域と呼ぶ。

#### 4.2 人物認識

人物が振り返ったり、上着を脱いだり追跡対象や環境の特徴が変わった場合でも、人物の追跡および認識を継続して行うためには適切なフィルタを選択する必要がある。図5に適切なフィルタを選択する処理を示す。TrackingCE

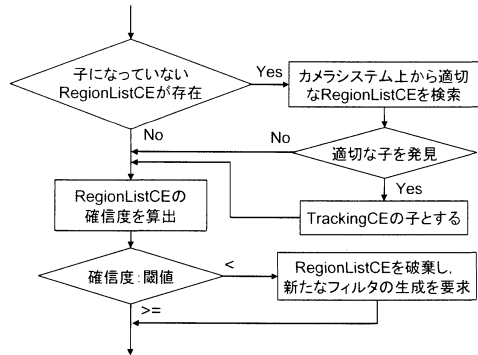


図5 フィルタ選択アルゴリズムのプロチャート

による人物追跡を行う際は、まず子にすべき RegionListCE のうち、まだ見つからないものをカメラシステム上から検索し、適切な RegionListCE が存在すれば子とする。続いて、既に子である RegionListCE の管理するフィルタから得られた情報に基づいて各 RegionListCE の確信度を算出する。確信度は、あるフィルタを用いて抽出された色領域がそのフィルタの担当領域を被覆するとき、被覆される面積が担当領域の面積に占める割合として定義する。フィルタの確信度が閾値を下回った場合は、そのフィルタを管理する RegionListCE を TrackingCE から切り離し、自身のフィルタリストを更新したうえで、人物の位置と切り離れたフィルタの担当領域をパラメータとして、カメラシステムに新たな RegionListCE の生成を要求する。

カメラシステム側では、フレーム画像において受け取ったパラメータに該当する領域内の色が最も多く含まれる色相環の小領域を探し、カメラシステム上に要求に応じた RegionListCE を生成する。TrackingCE は適切な RegionListCE が見つかった時点でそれを子とし、フィルタリストを更新する。

TrackingCE に破棄された RegionListCE は、自身を必要とする TrackingCE がカメラシステム上に存在した場合、再びその TrackingCE の子となる。一定時間が経過しても親となる TrackingCE が見つからない場合は、RegionListCE は自動的に消滅する。そのため、カメラシステム上に不要な RegionListCE が溢れることはない。

また、TrackingCE は、追跡中の人物が正しい追跡対象であることを示す指標として確信度パラメータをもつ。これは、自身の子である RegionListCE が管理するフィルタの確信度の平均値として定義される。

#### 4.3 人物位置の更新

TrackingCE は、フィルタリストの RegionListCE のうち、既に自身の子となっているものから得られる矩形情報を統合することにより、自身の保持する人物位置情報

の更新を行う。人物位置の更新手順を以下に示す。基本的には、RegionListCE が管理するフィルタによって人物位置領域およびその周辺から色領域を抽出し、得られた矩形を統合することによって人物位置を更新する。

(1) TrackingCE のもつ人物位置の情報を基に、人物位置領域を定数倍に拡大する。

(2) TrackingCE の生成時に与えられたパラメータにより、RegionListCE が管理するフィルタそれぞれに対して、拡大された人物位置領域における担当領域を算出する。

(3) RegionListCE のもつ抽出結果の矩形リストを辿り、担当領域と一定以上の重複部分がある矩形を探す。そして、拡大された人物位置領域を超えない範囲で、見つかった矩形を統合する。

(4) (3) でそれぞれの RegionListCE から得られた矩形を統合することにより得られた矩形と、各フィルタの担当領域を表すパラメータを用いて人物位置領域を算出し、新たな人物位置とする。

#### 4.4 カメラ間を移動した人物の追跡

本システムにおいて、それぞれのカメラシステムは自身の設置位置に関する情報を実空間における 3 次元座標として保持しており、カメラシステム同士はネットワークにより互いに接続されている。TrackingCE のもつ人物位置の情報によって、追跡中の人物がカメラの観察可能領域から離脱しそうだと判断された場合は、カメラシステム間で通信を行ってそれぞれのカメラシステムの設置位置を把握し、次にその人物が現れると予想されるカメラシステム上に TrackingCE の複製を生成する。

複製された TrackingCE は複製先のカメラシステム上で目的の人物の出現まで待機する。人物の検出に成功した TrackingCE は、自身以外の複製とオリジナルの TrackingCE に人物検出成功メッセージを送る。メッセージを受け取った TrackingCE は消滅し、メッセージを送った TrackingCE がその人物の追跡を続行する。

複製された TrackingCE は、自身のフィルタリストに基づき、複製先のカメラシステム上から自身の子となるべき RegionListCE を探す。適切な RegionListCE が見つからない場合は、4.2 節の場合と同様にカメラシステムに新たな RegionListCE の生成を要求する。

## 5. 実験

本研究で実装した人物追跡システムの動作を確認するため、実験を行った。カメラシステムは、ノート PC(CPU: Intel Pentium M 1.40GHz, メモリ: 512MB, OS: Red Hat Linux 9.0) と Web カメラ (Logitech 製 QV-700N) により構成した。各カメラシステム上では Ja-Net on GRID [6] が動作しており、人物追跡システムはその上

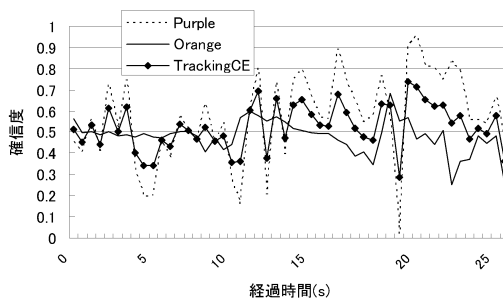


図 6 2 色の物体の追跡における確信度の変化

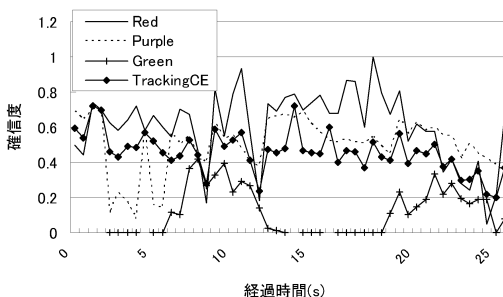


図 7 3 色の物体の追跡における確信度の変化

に実装している。動作実験では、単一の移動物体を追跡している間、確信度の変化に応じて適切にフィルタの組合せが変更されていることを確認した。本節では、実験結果を示し、考察を行ってシステムの問題点を明らかにする。

実験では、色相環を 36 の小領域に分割した。そして、複数の色をもつ物体の追跡を行い、それぞれの色に対応するフィルタの確信度の変化を時間とともに計測した。2 色 (紫, 橙) の物体の追跡結果を図 6 に、3 色 (赤, 紫, 緑) の物体の追跡結果を図 7 に示す。横軸は経過時間 (second)、縦軸は確信度である。確信度が大きいほど、追跡対象に含まれる色領域を正確に追跡できていることを表す。また、それぞれの図において TrackingCE で示されるグラフは各フィルタの確信度の平均値であり、この値が大きいほど、TrackingCE が追跡中の人物が正しい追跡対象である可能性が高いことを表す。閾値は 0.3 とし、確信度がこの値を下回った場合は、そのフィルタを管理する RegionListCE を子から切り離し、カメラシステムへの新たな RegionListCE の生成を要求する。

図 6 において、橙の領域を抽出するフィルタの確信度は平均的に高いが、紫の領域を抽出するフィルタの確信度は閾値 (= 0.3) を何度も下回っており、それに伴うフィルタの変更が行われている。図 7 において、紫と緑の領域を抽出するフィルタについても同様のことがいえる。

つまり、色を色相によって表現する場合、フィルタによる色領域の抽出に適さない色が存在すると考えられる。フィルタによる色領域の抽出が難しい場合には、他の手法やセンサから得られる情報を統合利用することにより、確信度の改善が期待できる。

また本実験では、TrackingCE がもつ人物位置領域の情報を可視化することで、人物位置の更新が適切に行われていることを確認した。4.3 節で示したように、人物位置領域はフィルタの担当領域を示すパラメータを用いての算出するため、適切に色領域を抽出できるフィルタを管理する RegionListCE が TrackingCE の子として 1 つ以上存在すれば、人物位置領域におけるそれらのフィルタの担当領域を把握することで人物位置領域を求められる。そのため、適切な色領域の抽出に失敗したフィルタが存在しても、正確な人物位置を得ることができたと考えられる。

## 6. おわりに

本研究では、広域にわたって設置された多数のカメラやセンサなどの分散リソースから得られる情報を用いた人物認識を伴う人物追跡システムの構築を目的とし、分散リソースから得られる多様な情報を状況の変化に応じて動的に統一して利用するフレームワークを提案した。設計したシステムでは、形状認識やパターンマッチングといったさまざまな従来手法を柔軟に組み合わせて利用できる。そのため、周囲の環境や人物の外見的な変化の影響を受けにくくなり、安定した動作が可能となる。

本研究では、人物に含まれる色情報のみを用いたシステムの動作実験を行ったが、情報の信頼性の低下によって、情報の統合による適切な追跡が行われない場合があることが示された。本システムを実環境へ適用する際には、より豊富な計算資源を背景として、複数の従来手法やさまざまな種類のセンサからの情報を組み合わせて統合利用することにより、さらにロバストで精度の高い人物追跡システムが実現できると考えられる。

今後の課題としては、色情報だけでなく、背景差分法や形状認識法といった手法や、各種のセンサから得られる情報を統合利用した人物追跡によるシステムの評価や、それに伴うアルゴリズムの改良などが挙げられる。

## 謝 辞

本研究を進めていくにあたり、貴重な御助言と多大なる御協力を頂いたサイバーメディアセンターの中澤篤志先生に心より感謝の意を表する。

なお、この研究の一部は文部科学省 21 世紀 COE プログラム「ネットワーク共生環境を築く情報技術の創出」ならびに、文部科学省特定研究領域 (C)「Grid 技術を適応

した新しい研究手法とデータ管理技術の研究」(プロジェクト番号:13224059)、科学研究費補助金(基盤研究(B)(2))「大規模な仮想空間システムを構築する放送型サイバースペースに関する研究」(プロジェクト番号:15300033)によっている。ここに記して謝意を表す。

## 文 献

- [1] 馬場功淳, 大橋健, 乃万司, 松尾英明, 江島俊朗, “Head Finder: フレーム間差分をベースにした人物追跡.” 画像センシングシンポジウム 2000, pp.329-334 (2000).
- [2] T. Darrell, G. Gordon, M. Harville, J. Woodlil, “Integrated Person Tracking Using Stereo, Color, and Pattern Detection.” International Journal of Computer Vision 37(2), pp.175-185 (June 2000).
- [3] 早坂光晴, 富永英義, “動画像からの背景画像生成を用いた移動物体抽出方法に関する一検討.” 情報処理学会研究報告, オーディオビジュアル複合情報処理, IPSJ-AVM 29-1, pp.1-6 (June 2000).
- [4] 中澤篤志, 加藤博一, 井口征士, “分散カメラエージェントを用いた広域人物位置検出システム.” 情報処理学会論文誌, Vol. 41, No. 10, pp.2895-2906 (October 2000).
- [5] Chahab Naster, Matthias Mitschke, “Real-Time Face Recognition Using Feature Combination,” in Proc. of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition (FG 1998) (April 1998).
- [6] 沼田哲史, 板生知子, 小川剛史, 塚本昌彦, 西尾章治郎, “動的なグリッド環境における効率的でセキュアなリソース利用のためのモバイルエージェントシステム, Ja-Net on Grid.” 情報処理学会論文誌「データベース」(TOD 24) (2004).
- [7] Kiyotake Yachi, Toshikazu Wada, Takashi Matsuyama, “Human Head Tracking Using Adaptive Appearance Models with a Fixed-Viewpoint Pan-Tilt-Zoom Camera,” in Proc. of the 4th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2000) (March 2000).