The University of Osaka
Institutional Knowledge Archive

Title	分散視覚システムによる人物の追跡
Author(s)	中澤, 篤志
Citation	大阪大学, 2001, 博士論文
Version Type	VoR
URL	https://doi.org/10.11501/3184204
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

工学 8122

分散視覚システムによる人物の追跡

中澤 篤志

大阪大学大学院基礎工学研究科

博士学位論文

２００１年１月

分散視覚システムによる人物の追跡

博士学位論文

中澤　篤志

２００１年１月

要　約

本論文では，従来の視覚システムでは不可能だった広域環境のリアルタイム観測を可能にする「分散視覚システム」を提案する．本システムは，視覚センサを持ち自律して処理を行う計算機（観測ステーション）をネットワーク上に複数配置し，それらが協調・情報統合・競合解消などを行うことで，複数の視覚センサの観測領域に渡る広域シーンの認識を行うことができる．画像処理は各々のステーション上で行われるため，システム全体で処理が分散され，シーン全体を実時間で観測することが可能になる．

このシステムを用いた具体的なタスクとして，広域環境中の複数人物の実時間追跡・位置検出をテーマとした．その実現のために，以下の技術を開発した．

第１は，観測ステーション上で動作するリアルタイム人物追跡・位置検出手法である．この手法は，人物を３Ｄの楕円体で近似するモデルベース追跡アルゴリズムであり，環境の奥行きをシミュレートすることで，環境中の構造物等による人物像のオクルージョンに対処することができる．複数人物を対象にしても高速に動作し，人物位置を環境中の座標系で得ることができる．

第２は，広域観測を行うための観測ステーション間の協調手法である．本システムで用いる観測ステーションは，広域を部分領域に分割し分担観測する．これらの視野間を移動する人物を連続的に追跡しつづけるためには，観測ステーション間の情報伝達によるタスク制御が必須である．提案手法では，観測ステーションが互いの観測領域を把握し，観測視野が隣接するステーション間でメッセージを送りあうことで，人物の連続的追跡を可能にした．また，ステーションの障害に対するロバスト性，使用する通信帯域幅も考慮に入れ設計を行った．

第３は，複数ステーション間での複数対象の対応付け手法である．本システムでは観測ステーションが自律して観測する仕組みをとるため，ステーション間で同一対象のマッチングを行う必要がある．これを実現するために対象の位置情報など特定の特徴量を用いることも考えられるが，対象が近接した場合などには困難になる．他の特定の特徴量を用いたとしても状況によって同様の問題が生じるものと考えられる．本システムで提案するアルゴリズムでは，各々の観測ステーション上で観測対象と１対１に対応し観測を行う「観測エージェント」を複数個並列動作させる．観測エージェントは自らが獲得可能な全ての特徴情報を，他の観測エージェントの情報と比較することで対象の対応付けを実現する．この構成により，特定の特徴量に依存しない対象の対応付けが可能になった．

以上の手法を提案し，実験により有効性を確認することで，本システムが広域環境を実時間観測する視覚システムとして有効に働くことを示した．また，分散視覚処理というアプローチが広域・実時間観測を実現する手法として有効であることが示された．

目　次

1章 序論 1

1.1 本研究の背景 1

1.2 分散視覚研究の動向 3

1.3 分散システムとしての論点 5

1.4 エージェント研究としての論点 6

1.5 解決すべき技術課題と論文の構成 10

2章 楕円体モデルによるリアルタイム人物追跡・位置検出 11

2.1 人物の動作追跡に関する研究（関連研究） 11

2.2 処理の概要 15

2.2.1 モデル表現 15

2.2.2 人物位置管理マップ 17

2.2.3 カメラパラメータ 18

2.3 シミュレーション差分画像の生成 18

2.3.1 視野領域内人物の選択 18

2.3.2 楕円パラメータの計算 18

2.3.3 奥行きマップによるオクルージョンへの対処 20

2.4 比較処理 21

2.4.1 マッチングによる人物位置の決定 21

2.4.2 3値画像によるマッチング 21

2.5 拡張 22

2.5.1 複数人追跡への拡張 22

2.5.2 追跡モードと捕捉モード 23

2.6 実験と検討 24

2.6.1 実験環境 24

2.6.2 カメラのキャリブレーション 24

2.6.3 計測精度評価 25

2.6.4 複数人を対象とした追跡実験 27

2.6.5 検討 29

2.7 本章のまとめ 30

3 章 複数の観測ステーションによる広域観測 31
3.1 複数ステーションにおける人物の連続追跡アルゴリズム 32
3.2 システムの制御方式 33
3.3 システムの構成 35
3.3.1 観測ステーションと観測エージェント 35
3.3.2 システムパラメータ管理エージェント（System Parameter Management Agent:SPMA） 40
3.4 実験 41
3.4.1 実験環境 41
3.4.2 位置軌跡の評価実験（実験１） 41
3.4.3 エージェントの動作状態変化の観察（実験２） 43
3.4.4 一部のエージェントの追加・動作停止に対する実験（実験３） 43
3.5 検討 46
3.6 本章のまとめ 48

4 章 人物像の色分布を用いた人物識別 49
4.1 人物の誤対応問題 49
4.2 従来研究と本システムで要求される条件 50
4.3 人物像の色分布を用いた識別 51
4.4 処理の流れ 53
4.4.1 学習モード 54
4.4.2 識別関数モード 56
4.4.3 識別モード 58
4.5 実験 59
4.5.1 実験環境 59
4.5.2 ２人物の交差に対する識別 59
4.5.3 ３人物の交差に対する識別 61
4.6 検討 62
4.7 本章のまとめ 62

5 章 複数対象の同時観測 63
5.1 自律観測ステーション群による複数対象観測 63
5.2 提案するシステム構成とその特徴 67
5.3 アルゴリズムの概要 69

5.3.1 システムの構成 69
5.3.2 エージェントのグルーピング 70
5.3.3 確信度の更新 70
5.3.4 メディエータ 71
5.4 システムの実装 72
5.4.1 画像処理モジュール 73
5.4.2 通信モジュール 73
5.4.3 観測エージェント 73
5.4.4 メディエータ(mediator) 75
5.5 実験および結果 77
5.5.1 実験環境 77
5.5.2 追跡実験 77
5.6 検討 81
5.7 本章のまとめ 82

6章 結論 83
6.1 本論文のまとめ 83
6.2 分散視覚研究の今後 85

参考文献 87

1章　序論

本論文のテーマである分散視覚システムの研究とは，視覚機能を持つ複数の知的エージェントにより，より複雑かつ広域の実環境を観測しようとする試みである．この分野の研究は近年盛んに行われているが，その背景としてはＨＣＩやセキュリティ，福祉用途など，分散視覚の持つ実応用面における長所とともに，コンピュータビジョン（ＣＶ），人工知能（エージェント），通信技術など，学術的面から見ても新しいパラダイムである事が挙げられる．第１章では，本論文のテーマである分散視覚システム研究の背景を述べ，次に本論文の各学術分野の中における位置付けを述べる．

1.1 本研究の背景

近年のプロセッサの処理能力やネットワーク技術の向上はめざましいものがあり，前者は三年間で16倍，後者は64倍の成長率であるといわれている．この事実はすでに我々の日常生活に大きな変化をもたらしている．処理能力の向上によって，従来のコンピュータシステムが文字や数値などの「抽象的な記号情報」を主に取り扱ったのに対し，動画像・音声・３次元情報などの「より人間に近い実世界の情報」を取り扱えるようになった．また，ネットワーク技術の発達によって，インターネットのような大規模分散システムを有効に働かせることができるようになっただけでなく，近年では動画・音声をリアルタイムで伝送するストリーミングメディア，Voice over IP（VoIP）なども可能になってきている．

このようにコンピュータの能力を向上させるインフラストラクチャが整備される一方で，人間がそれらを「よりよく使う」ことを可能にするためのユーザインタフェース技術が求められている．近年実現しつつある音声による文字入力・コマンド入力インタフェースや，カメラから入力された動画像を認識することで実現されるユーザ認証，セキュリティ，ジェスチャインタフェースなどはその方向性である．しかしこれらのインタフェースは，現状では人間が機械に対して明示的に行動を示す必要がある．例えば，音声入力インタフェースでは人間がマイクに向かって音声を入力しなければならず，ジェスチャ認識やユーザ認証においては人間がカメラに向かって認識を促さなければならない．今後将来的にコンピュータシステムがより人間の生活に密着した存在となるためには，このような明示的なアクションを用いたユーザインタラクション技術ではなく，人が生活する実環境を把握する技術が必要になるものと思われる．

視覚システムによる実環境の理解

視覚システムによる実環境の理解は，適用可能なタスクの広さなどからこのような目的を実

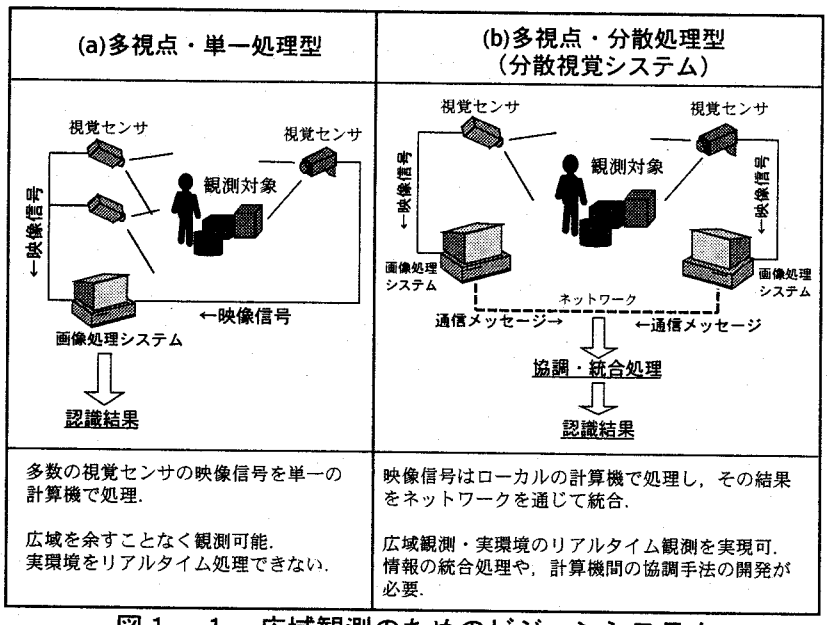


図１－１　広域観測のためのビジョンシステム

現できる可能性の高い技術だといえる．しかし１台の視覚システムでは観測できる範囲が限られるという問題がある．つまり人間は，視覚システムが観測する領域内に明示的に進入しシステムに対して認識を促さなければならない．

この問題を解決するため，環境中に複数の視覚センサ（カメラ）を設置し認識範囲を広げるという試みがある[Maeda92](図１－１（ａ)). これにより広範囲の領域を観測することが可能になった．このシステムでは，複数の視覚センサからのすべての映像信号を単一の計算機に入力し，順次処理を行うという構成をとっていた．単一の計算機で多数のカメラからの画像を処理するには多くの時間がかかり，リアルタイムでの広域観測は困難である．逆にこの構成によってリアルタイムの認識を行おうとすると，各カメラから得られた画像に対する認識処理を単純化する必要があるため，様々な状況が想定される実環境を対象とすることができない．

これに対し，視覚センサと画像処理装置を１対１に結合し，得られた画像をローカルの計算機で処理，その結果を通信プロトコルを用いて伝送し統合することで実環境の認識を行うという解決策が考えられる（図１－１（ｂ)). 前述のシステムよりも多くの台数の計算機が必要となるが，画像処理が分散化されるため高速処理が可能である．また，汎用ネットワークを用いて計算機を結合するためシステム構成の自由度も高い．これが本研究で提案している「分散視覚システム」である．分散視覚システムでは，視覚システムが結合された画像処理計算機を「観測ステーション」と呼ぶ．観測ステーションは互いの観測情報をネットワークを通じて交換・統合し環境の認識を行う．このような構成をとることにより，広域・実環境の観測をリアルタイムで実行することが可能になる．しかしこれを実現するには，各ステーション間での分散処理の度合いや交換するメッセージの内容，集められた情報の統合方法に関する検討を行う必要がある．

分散視覚の可能性

分散視覚システムによってリアルタイムでの広域環境観測が可能になれば，より人間に近い実世界の状況を把握するコンピュータシステムを実現できる．具体的な用途としては，以下のようなタスクが考えられる．

- 特別の機器を装着する必要のない人; 窿Ｉブジェクトのポジショニンキ
- 使用環境を選ばないジェスチャインタフェースの実現
- セキュリティ・交通管制システム等の防犯，安全監視システム[Tan97]
- 徘徊者定位等，病院や福祉介護施設等への利用[Tan97][Shimizu96]

また，分散視覚システムがネットワークによって結合されていることを考えると，この通信路を利用して特定／不特定多数のユーザに映像情報を提供する新たなネットワークメディアとしての可能性も開けてくる．すなわち，WWWに代表されるような分散データベースシステムと同様に,分散視覚システムによって入力されるシーンの映像を高度に処理することによって，ネットワークを通じて遠隔地に処理結果を提供することが考えられる．これにより，遠隔地のモニタリングや，遠隔地で生じたイベントをリアルタイムで通知するシステムなどへの応用も可能である．

本研究では，このような用途に利用可能な基礎的・汎用的技術として「広域環境内を移動する人物の追跡・位置検出」を実現することを目的とする．システムは全て同等の視覚センサと計算能力，通信機能を持つ観測ステーションによって構成され，広域環境を分割して分担観測を行う．人物が複数観測ステーションの視野間を歩きまわる状況において，システムはリアルタイムに人物の現在位置を提供したり，環境マップ上に通過軌跡を描き出すことができる．この技術が可能になれば，広域環境中の人物位置を提供するポジショニングシステムやＨＣＩシステム，セキュリティや安全監視，福祉用途などへの応用が期待できる．また追跡対象を，移動ロボットや車両等のオブジェクトに拡張することで，ＦＡ（Factory Automation）やＩＴＳ（Intelligent Transportation System）などへ適用することも可能なシステムを目指す．

1.2 分散視覚研究の動向

分散視覚システムは，従来の視覚認識システムでは不可能な広域環境の観測や，実環境のような複雑なシーンのリアルタイムでの観測を可能にする．このような目的のために視覚処理を分散させ，リアルタイム環境認識を行った研究として，交差点を通過する車両を対象にした西尾らのシステム[Nishio92]等をあげることができる．

「分散協調視覚(Cooperative Distributed Vision)」という用語を用い，分散視覚研究をプロジェクトとして推進したのが松山ら[Matsuyama97]である．ここではその意義として，従来の視覚・人工知能研究の「定性的な記号を元に知覚・推論・行動する」システムから，「知覚，行動を会

した外的世界との相互作用を基本とした知能システム」への移行を目指している．すなわち，従来のＣＶが対象とした静的環境（「積み木の世界」）から，時々刻々と変化する動的環境の理解を目指したものであるといえる．また，これを実現するためには様々な情報の「統合」を行うべきであり，ビジョンシステムにおいても抽出された情報統合の方式や，統合のための新たなアルゴリズムの開発が不可欠であるとする．

また，米国ではVSAM(Video Surveillance and Monitoring) Projectが提唱され，いくつかの研究がなされた[Strat97]．このプロジェクトも，複数のビデオセンサをネットワークで結合し，広域の観察タスクを実現するものであり，その中では，

- Cooperative Sensor Arrays
- Air/Ground Sensors
- Site Monitoring
- Vehicle Tracking, Activity Recognition
- Zone Instruction

などが技術的テーマである．ここでは，イメージメディアとしての応用研究が多く扱われている．例えば，ビル等に設置された複数のカメラからその視野内を通過するオブジェクトの追跡を行い，得られる通過軌跡のクラスタリング結果を提供したり，シーン内の動作イベントに対する検索を可能にするシステムを構築している．

これらのいずれの研究やプロジェクトを見ても，分散視覚に要求する方向性は似通っている．すなわち，

（１）　静的環境の理解から動的環境の理解へ

（２）　機械による環境理解から人間にとって有効な情報を提供するシステムへ

（３）　限定されたシーンの理解から，広域の実シーンの理解へ

という方向性である．（１）は分散視覚の処理体系の必然性（リアルタイムでの環境認識）からきたものであり，（２）はイメージメディアへの方向付けである．すなわち，システム開発の目的として「計算機による環境の理解」を目指すのみではなく，計算機によって獲得された環境の情報をいかに人間に伝えるかを表すものである．また（３）は，従来のＣＶが扱っていた「積み木の世界」からの脱却を強く目指したものであるといえる．すなわち，様々な条件が仮定され，「箱庭のように」限定された範囲の環境を対象とするのではなく，人が生活する広範囲の実環境を対象とすることで，視覚機能によって実世界と情報メディアの間を円滑につなぐことを目指すという分散視覚特有の方向性であるといえる．

1.3 分散システムとしての論点

本研究を分散システム研究の一つとしてとらえると，各観測ステーションの動作をどのように制御・管理するかという問題が生じる．すなわち，システム全体を鳥瞰的に見渡し各エージェントに指示を行う存在を設ける「集中制御構造」をとるか，エージェントが自律・分散して行動する「完全分散型」をとるかの違いがあげられる（図１－２）．この点に関しても，ロボット工学における議論を踏まえて検討する必要がある．まず，[Arai92]では複数移動ロボット系の計画問題として，この２方式の違いを以下の表のようにまとめている．

表１－１　マルチエージェントシステムにおける
集中型と分散型の違い（[Arai92]より）

	最適性の考慮	ロボットの台数増加時の計算量	フォールトトレラント性
集中型	○（可能）	×（計算量↑）	×（機能：集中）
分散型	×（不可能）	○（計算量→）	○（機能：分散）

また[Anzai92][Anzai94]でも同様に，集中制御と分散制御に関して以下のような特性をのべている．

集中制御：情報をある実体に集中するため通信の幅を広く取ることが必要であり，さらに，全ての情報が伝達されてから決定を行うため，遅延が大きいという欠点がある．しかしプラニングを要するような複雑な作業は構成しやすいという利点がある．

分散制御：各ロボットが自己の周囲の状況を把握し，周囲のロボットと通信するだけであるから，集中制御よりも通信量が少なくなる．このことは，処理速度の向上につながる．しかし分散制御では複雑な制御がやりにくいという欠点がある．

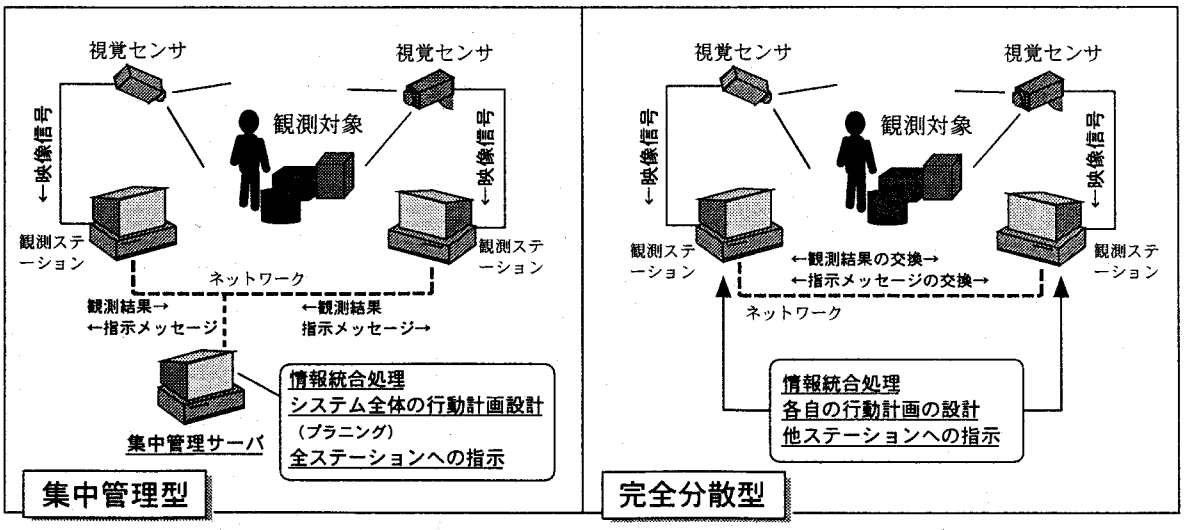


図１－２　集中管理制御と完全分散制御

両者の違いにはシステム設置の自由度という観点もある．一般的に，集中制御型では集中管理サーバが全観測ステーションの状況を把握する必要があるため，固定されたシステム設置しか扱えないという欠点がある．このような観点から本システムが対象とする「複数の視覚システムによる広域観測」というタスクについて考えると，その処理速度や耐障害性，システム設置の柔軟性の点から，基本的には分散型を取るべきである．しかし，タスク遂行のためには集中制御的な要素も取り入れておく必要がある．例えば，複数の観測ステーションが連携して同一の対象を観測する場合などでは，互いの観察視野を把握し協調動作を行う必要がある．これを実現するためには，一度全てのステーションの情報を獲得する必要があるため，集中的な情報管理を行うことが必要である．そこで本システムでは「異なった視野を観測する分散システムの制御」という点に注目し，観測ステーションの観測視野に基づいた制御方式を提案している．この詳細については，複数ステーションを用いた広域観測を取り扱うシステムの構築時（第３章）において再び議論することとする．

1.4 エージェント研究としての論点

分散視覚システムの各計算機で得られた映像情報は各々で処理され，計算機間で画像をそのままやり取りして処理することはない．すなわち，各計算機の視覚機能は独立しており，各々がなんらかの自律性をもって行動しているといえる．このような構成をとることから，分散視覚システムはロボティックス分野に見られる「マルチエージェントシステム」の一種であると考えることができる．ここではまず最初に，本論文の核となる概念の一つである「エージェント」について整理しておく．

エージェント(agent)とエージェンシー(agency)

人間の知性の働きのモデルとして「エージェント」という考え方を提案したのはMinskyである[Minsky96]．その著書の中で彼は「複雑な動作をしているように見える『知性』とは，より単純なプロセスの組み合わせ（社会）によって構成されている」という見方を提唱した．ここでの各プロセスが「エージェント」と呼ばれている．彼のエージェントおよびエージェンシーの概念の一例を図１－３に示す．

ここでは，「積み木によって塔を作る」という「知性」に対する例が示されている．この知性は「始める」「加える」「終わる」という単純なプロセスに分解される．そして，またこれらの各々のプロセスは，より小さなプロセス（エージェント）に分解される（例えば，「加える」は「見つける」「手に入れる」「置く」などの小プロセスに分解される）．すなわち，ある定義された「はたらき」は，より細かな，具体的なプロセスに分解することができる．逆に，あるレベ

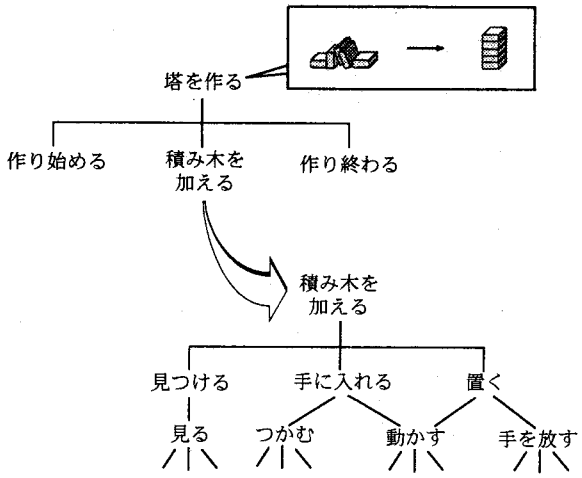


図１－３　エージェントによる人間の知能の表現（[Minsky96]より）

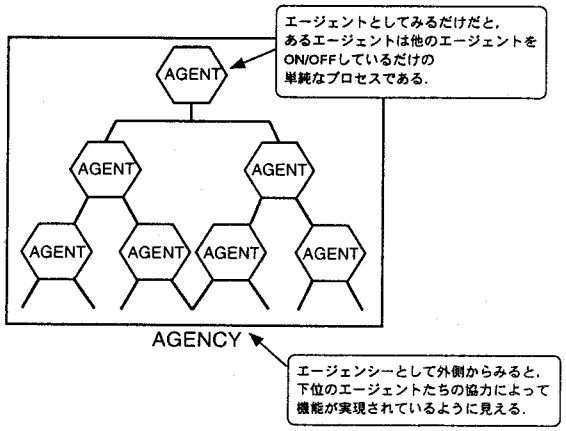


図１－４　エージェントとエージェンシー（[Minsky96]より）

ルのエージェントが互いに連携し，より高度な機能体となったものを「エージェンシー」と呼ぶ（しかし，このエージェンシーはまた上位のレベルからみると「エージェント」である.）

ここで重要な点は，ある「働き」はエージェントであるとともにエージェンシーであるという考え方である．すなわち，ある存在の「働き」がそのまま「定義されたもの」であると考えればその存在は「エージェント」であるとみなすことができるが，この「働き」をより詳しく考えてみると，その存在はより細分化されたエージェントの組み合わせからなる「エージェンシー」とみなすことができる（図１－４）．すなわち，ある機能体がエージェントであるかエージェンシーであるかは，「行為」をどのレベルの機能単位まで分割するかに依存している．

マルチエージェント

一方「マルチエージェント」とは，単にエージェントが複数存在している状態を指しているものと考えられる．すなわち，マルチエージェント状態のエージェント同士は，互いに階層的な関係を持つことでエージェンシーを構築することもあったり，まったく関係なく動作している状態であったりする．エージェント，エージェンシー，マルチエージェント各用語の定義関係を図１－５に示す．

ロボティックスの分野では，自律して動作するロボット・プログラムを複数用いることでマルチエージェントシステムを構築し，それらの協調動作によって高度なタスクを実行させる研究が盛んに行われている[Ishioka94][Asada97]. [Ishioka94]では，複数の移動ロボットの協調によって環境内の地図を生成するシステムを開発した．このシステムの移動ロボットはソナーセンサを持ち，周辺の環境情報を取得する．最終的にそれらの局所情報を統合して，環境マップを生成する. [Asada97]では，単純な画像処理能力を持つロボットが協調してサッカーを行うタスクを与えている．その中では，各ロボットが学習によって互いに協調の仕方を学習している過程を

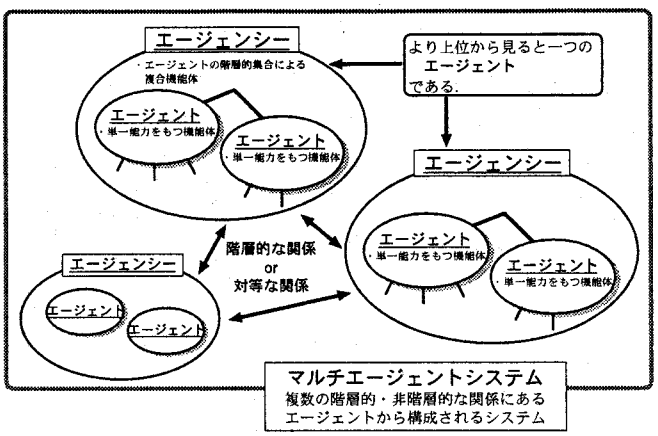


図１－５　エージェント，エージェンシー，マルチエージェントの関係

観察している．これらいずれの場合にも，各移動ロボットはエージェントでもあり，エージェンシーでもある．また，環境地図の生成やサッカーというタスク全体を考えてみると，それ一個が「エージェンシー」を構築しているという考え方をすることもできる．

このような点から分散視覚システムを考えると，視覚機能を持った計算機はエージェントであり，それらがネットワークで結合された全体のシステムがマルチエージェントシステムであるという考え方が可能である（図１－６（１））．また別の視点から見れば，各計算機自身も複数の認識アルゴリズムから形成されるマルチエージェントであると考えることもできる（図１－６（２））．

マルチエージェントとコミュニケーション

ここで重要な点は，１つのステーション上で動作するエージェント間や，異なったステーションで動作するエージェント間でやり取りされる情報の量と質である．

マルチエージェントを用いて環境認識を実行した例として，ヒトの視覚機能を模倣した認識システム[Yanai98][Ohta91]が挙げられる．これらのシステムでは，単一の画像処理システム中に複数の画像認識エージェントが並列動作し，それらのコミュニケーションによって処理が進められる．ここで対象とするシーンは，単一視点から得られた画像が用いられる．エージェント間のコミュニケーションは同一システム内で行われるため，交換される情報の量や交換方法に対する制約は少ない．

一方，分散視覚システム上のマルチエージェントは，複数のステーションによる複数視点の画像を認識に用いることができるという意味で，前者よりも優れた性能を発揮する可能性がある．しかし，エージェント間ではその情報を通信路を介して交換し統合しなければならない．ここでは，

１．各エージェントで行われる画像認識処理のレベル

２．複数エージェント間での情報統合

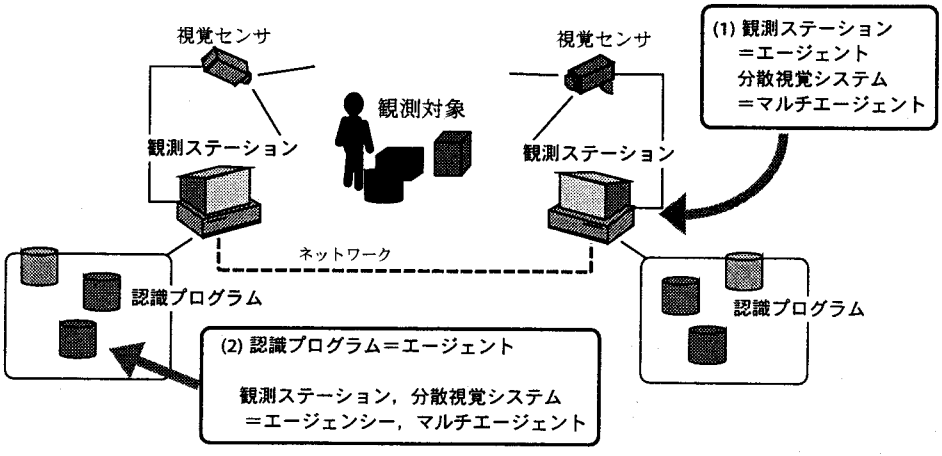


図１－６　分散視覚システムにおけるエージェント・エージェンシー
マルチエージェントの関係

表１－２　分散視覚システムのエージェントでの画像処理と
通信コスト，統合手法の関係

エージェントでの画像処理レベル	エージェント間の通信量情報統合に必要なコスト	その他の特徴
低レベル（画像に近い）	通信する情報量：大 統合時の計算量：大	統合アルゴリズムに柔軟性（多様性）がある
↕	↕	↕
高レベル（抽象的）	通信する情報量：小 統合時の計算量：小	各々のステーションでの処理に対する信頼性が重要

３．エージェント間で通信される情報量

４．通信にかかる時間

等が考慮すべき項目となる（表１－２）．１～３に関しては一般に，各エージェントで行う認識処理のレベルが低く，より画像の情報に近い情報をやり取りするほど統合アルゴリズムの柔軟性（多様性）は増すが，統合にかかるコストや通信量・通信時間は多くなる．逆に，各エージェントで行う認識処理のレベルが高く，より抽象的な情報をやり取りすれば，統合にかかるコストや通信量・通信時間は低く押されられるが，システム全体の環境認識の信頼性は各エージェントの認識処理に多く依存するため，その信頼性を把握する手段を用意しておくことが望ましい．一方４は，システムのリアルタイム性に関わってくる．あるエージェントで観測された結果は処理され他のエージェントに送られるが，その間にはそれぞれのエージェントに応じたの時間を要する．複数のエージェントから送られた情報の統合時には，これらの間のタイムラグを補正する処理が必要であるが，一般にその値が小さいほど正しい統合結果を得やすい．

このように，エージェントやエージェンシーをどのような単位として考え，どのような情報を交換するかは本研究における重要なテーマである．特に複数の観測対象を同時に取り扱い，そ

の間の対象対応付け手法を目指した５章において，エージェントとエージェンシーの枠組みをもう一度とらえ直すこととしたい．

1.5 解決すべき技術課題と論文の構成

本システムの目標とする「システムの観測視野内に存在する複数人物の連続的追跡・位置検出」を実現するためには，以下の技術要素を解決する必要がある．

- 観測ステーション上で行われる人物検出・追跡・位置計測（２章）
- 観測ステーション間の協調手法（３章）
- 同一画像内での対象（人物）の識別手法（４章）
- 複数の観測ステーションによる複数人物追跡手法（５章）

まず２章では，観測ステーション上で入力された画像を処理し，人物を検出・追跡・位置検出する手法について述べる．本手法は３Ｄモデルに基づいており，人物位置を環境座標系で得ることができる．３章では異なる視野を持つ観測ステーションが互いの観測視野を考慮して協調を行う手法を述べる．これにより，分散視覚システムにより広域を観測することが可能になったが，複数人物の同時観測にはいくつかの問題があった．これに対応するため４章では，人物の全身の色分布を用いて複数人物を判別する方法について述べる．この手法は計算コストやスピード・追跡ノイズなどの条件として実用には厳しい点があったが，そのアイデアや知見は最終的なシステムの構築時には重要であった．５章では最終的な目標となる複数人物の広域追跡システムについて述べる．最終章である６章では，本研究をまとめたうえ，分散視覚システムにおいて今後解決すべき問題や将来展望を述べる．

2章 楕円体モデルによるリアルタイム人物追跡・位置検出

本章では，観測ステーション上で実行される画像処理による人物追跡・位置検出アルゴリズムについて述べる．画像から人物動作の計測を行う研究は近年多く行われているが，取得しようとする情報やアプリケーション，手法によって様々なカテゴリに分類できる．そこで，まず本システムの最終的な目的の中で人物追跡処理に要求される項目を挙げ，これらの点から適切な手法についての検討を行う．すなわち，以下のような要件である．

1．リアルタイムでの動作

2．複数人物の追跡

3．実環境で生じるノイズに対してもロバストであること

4．検出結果が複数の視覚システム間の統合に適していること（統合容易性）

1～3に関しては本システムの目的のために必要とされる点ではあるが，通常の画像処理研究としても重要な実現課題であるといえる．4は分散視覚システムにおいて結果を利用することを見据えた上での要件である．具体的には，検出結果が画像中の領域のみでなく，複数の観測ステーションに共通する座標系であることなどが望ましい．この点が実現できれば，複数の視覚システム間で観測結果を交換した場合でも容易に互いの情報を参照・統合することが可能となる．

2.1 人物の動作追跡に関する研究（関連研究）

視覚システムによって人物の動作追跡を実現する研究はこれまでにも数多く提案されている．古くはJohanssonらにおける，MLD(Moving Light Displays)と呼ばれるマーカを用いて人物動作を計測した研究[Johansson70]が挙げられるが，近年ではマーカ等に頼らない非拘束・非装着の手法に関する研究がほとんどとなっている．

非拘束・非装着を前提とした研究の中でも，応用分野や追跡手法によっていくつかに分類する事ができる[Gavrla99]．応用の点から見ると以下のような分類が考えられる．

1．ＶＲやテレコンファレンス，ＨＣＩ等

人物の姿勢をリアルタイムで取得する必要があるもの．得られた情報は，ＣＧで描かれたアヴァターの制御やジェスチャ認識に用いられる[Wren97][Takahashi2000]．

2．ＣＧのモデリングやスポーツ動作の解析

リアルタイムである必要はないが詳細な人物姿勢・位置を計測する必要があるもの[Sato97]．

3．セキュリティ等の監視用途

シーンからリアルタイムで人物を発見する必要があるもの．人物姿勢等は必要でないが，様々な環境下でロバストに働く必要がある[Okada97][Segen96]．

また，手法の点から考えると以下のように分類できる．

１．２Ｄ画像から形状モデルを用いずに実現する

人物の形状モデル等は使わず，画像から得られた特徴量のみから人物を抽出する手法である．背景差分画像から人物領域の特徴点を抽出し，フレームごとに対応付けを行うことによってトラッキングを実現するもの[Segen96]，オプティカルフローによる画像領域のクラスタリングを用いるもの[Okada97]，両眼視差画像と運動視差画像から生成されるヒストグラム上でクラスタリングを行うもの[Nordlund97]，B-spline関数で表される輪郭線によって歩行する人物を追跡するもの[Baumberg94]，色(R,G,B)と空間位置(x,y)による画像領域のクラスタリングによるもの[Heisele97]などがある．長所としては，人物の初期的位置などを与えなくとも動作する点などが挙げられる．逆に短所としては，画像特徴の対応づけやクラスタリングはノイズの多いシーンや複数人物の同時追跡に対して処理が困難になる点や，オプティカルフローや視差画像の計算をリアルタイムで実行するためには高速なハードウェアが要求されるといった点があげられる．また，人物の検出結果は画像中の座標系であるものがほとんどである．

２．２Ｄ画像と２Ｄの形状モデルを用いて実現する

画像内によって人物領域がどのように見えるかという先見知識を２Ｄの形状モデルで保持し，モデルと入力画像を比較することによって追跡を行う手法である．人物領域は背景差分などによって得られ，人物の形状モデルには２次元の「棒状モデル(stick-figure)[Leung95]」（図２－１）などが用いられる．また追跡においては，フレーム間における微小移動の仮定がされているケースが多い．この研究に分類されるものとして，図２－１のような棒状モデルをシルエット画像とフィッティングするもの[Leung95]，人物モデルを２次元のガウシアン画像領域とし画像との

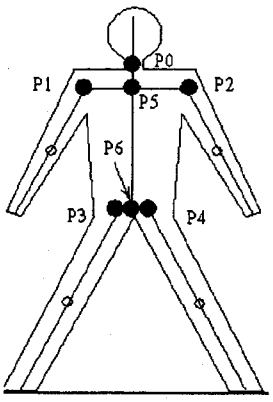


図２－１　２Ｄ人物モデルの例
([Leung95]より)

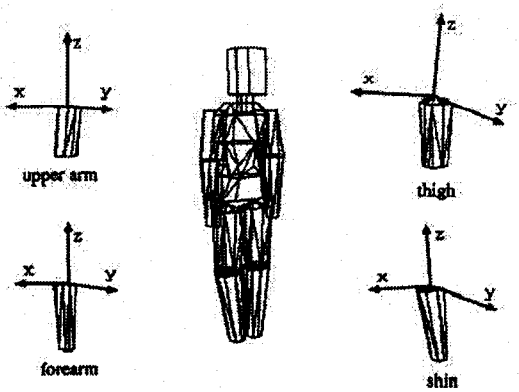


図２－２　３Ｄ人物モデルの例
([Sato96]より)

類似度の比較を行う方法[Wren97]などが挙げられる．この手法は人物の見かけの形状や大きさの変化が少ない状況においては，比較的安定して高速に動作する．また，モデルを用いることでノイズ等にも強いと考えられる．一方，モデルに対して初期位置を与えなければならない，観測方向やオクルージョンなどによって人物の形状・大きさが変化する状況ではうまく働かないという問題点が指摘されている．人物の検出結果は１と同様に画像内での座標系であるものが多い．

３．２Ｄ画像と３Ｄの形状モデルを用いて実現する

人物に対する何らかの形状モデルを仮定する点は２と同様であるが，３Ｄ形状モデルが仮定されており，モデルを投影した画像と入力画像を比較することで追跡を実現する方法である．人物モデルとしては，骨格に肉付けされたポリゴンモデルや一般化円筒による関節モデルなどが用いられる（図２－２）．この実現方法における基本的な処理の流れを図２－３に示す．

この研究の例としては，面の組み合わせによって表現された人物モデルを用いるもの[Badler80]，人物領域の輪郭と一般化円筒で表現された人物モデルを比較することで人物姿勢を求めるもの[Rohr94]，ポリゴンで表現された人物モデルの動きと入力画像の輝度から時空間勾配法によって推定した動きとのマッチングによる方法[Sato97]などがある．この手法の長所は，３Ｄ人物モデルを用いることで人物の位置・動きなどが環境座標系で得られる点や，人物モデルを画像に投影する段階で適切な投影を行えば人物の形状の変化や大きさ変化に対応できることが挙げられる．一方短所としては，正確な人物モデルを用いるほど関節角度変化の組み合わせが多くなり，次フレームの変化予測の探索にかかる時間が多くなる事や，モデルから得られた画像と入力画像とのマッチングに失敗するとその後の復帰に難しいこと，２と同様にモデルに対して初期位置．初期関節角度を与えなければならない等が挙げられる．

これらの人物追跡手法の応用分野とその長所・短所を表２－１にまとめておく．

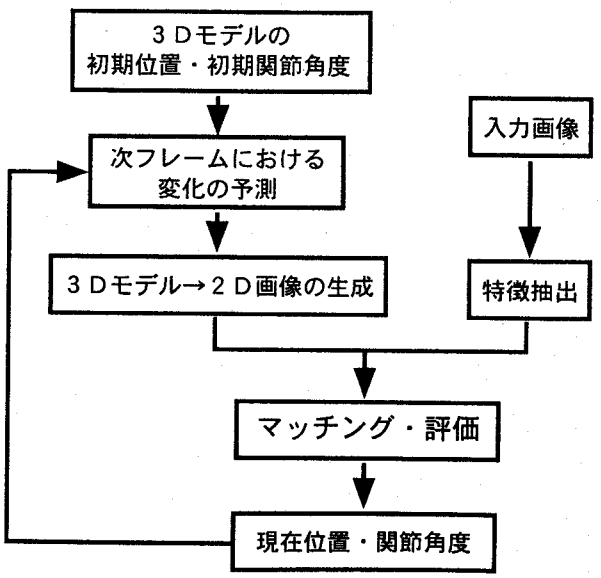


図２－３　２Ｄ画像と３Ｄ形状モデルを用いた追跡処理の流れ

これらの各手法の特性と冒頭で述べた本システムにおける人物追跡手法の要件から考えると，リアルタイム処理や複数人物追跡を行うための簡易性，ノイズに対するロバスト性の点から２のアプローチが考えられるが，人物の見かけの大きさ変化への対応や検出結果の統合容易性の点から見て，３の手法の要素も取り入れるべきであると考える．そこで，基本的には３のアプローチをとりながらも人物モデルとして多関節物体のような複雑なモデルを用いることは避け，使用するCADモデルを簡素化することで計算量を低減させる．具体的には，規定身長・規定幅を持った楕円体を用い，床面上を移動することを仮定する（図２－４）．これにより，３の手法では問題となる関節変化の組み合わせ等が生じず，環境座標系での位置変化のみで追跡可能になる．結果的に２とほぼ同じ計算量で処理が実行でき，検出位置も環境座標系で得ることができる．また複数人物の追跡に対しては，複数人物のCADモデルを同時に用いることによって可能になるものと考えられる．

モデルに基づいた手法で問題となるオクルージョンに対しては，以下のように対処する．まず，セルフオクルージョンに関しては本モデルの性質上起こり得ない．環境中に存在するオブジェクトによるオクルージョンに対しては，あらかじめ環境中のオブジェクトの位置・形状を記述しておき，楕円体人物モデルとオブジェクトの奥行きを比較することでオクルージョンのシミュレートをし対処する（図２－５）．

図２－４　入力画像（上），差分画像と楕円体人物モデル（下）

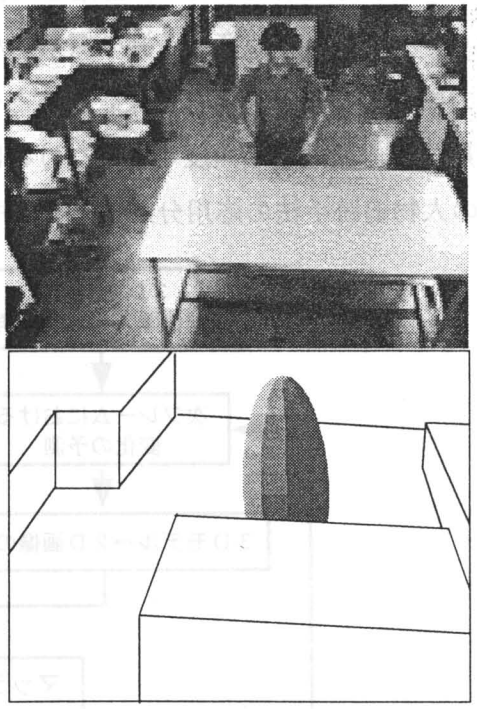


図２－５　オクルージョンが生じた入力画像（上），環境中の構造物記述による人物モデルのオクルージョンのシミュレーション（下）

表２－１　各追跡手法の特性とその応用分野

手法	1. モデルを用いない	2. ２Ｄの形状モデル	3. ３Ｄの形状モデル
長所	初期位置が不必要 計算量が比較的小	安定・高速な動作 ノイズに強い	検出結果が環境座標系で得られる 人物の見かけの形状・大きさ変化に対応可
短所	複雑なシーン・複数人物の追跡時に計算量大	人物の初期位置が必要 見かけの形状変化・大きさ変化に対応しづらい	初期位置・関節角度が必要，関節数に応じて計算量が増大，セルフオクルージョン
計測結果	画像座標系	画像座標系	環境座標系
応用分野	監視 通行人流調査等	監視・通行人流調査 動作解析，VR, HCI	モデリング，動作解析，VR, HCI

2.2 処理の概要

図２－６に提案する手法の概要を示す．まず，画像領域からの人物領域抽出には背景差分を用いる．カメラから入力された画像に対し，事前に取得しておいた背景画像と比較を行い，その結果を固定しきい値により２値化する．これにより生成された画像を実差分画像と呼ぶ．実差分画像には背景画像に対し閾値以上の濃度差のある領域が抽出される．理想的にはこの領域が人物領域となるが，実際には影や他の移動物などの領域の混入や人物と背景の濃度値の差が少ない部分における領域の欠損がノイズとして生じることになる．

人物の位置情報は，世界座標系で表現される人物位置管理マップに登録されている．その情報と事前に準備した環境モデル・カメラパラメータを用いて，実際のカメラから観測され背景差分・２値化処理が施された実差分画像，つまり人物領域のみが抽出された画像のシミュレーションを行うことができる．これにより生成された画像をシミュレーション差分画像と呼ぶ．そこで，人物位置管理マップから読み出した人物位置を環境座標系上で様々な方向に微小移動させ，複数枚のシミュレーション差分画像を生成する．これら複数のシミュレーション差分画像と実差分画像を比較し，実差分画像に最も類似したシミュレーション差分画像における人物位置を現在の人物位置として，人物位置管理マップの情報を更新する．この処理を繰り返す．

2.2.1 モデル表現

（１）環境モデル

先に述べたように，構造物による人物モデルのオクルージョンに対するシミュレーションや，追跡における移動可能領域を評価するために，環境座標系で記述される環境モデルを用意する．

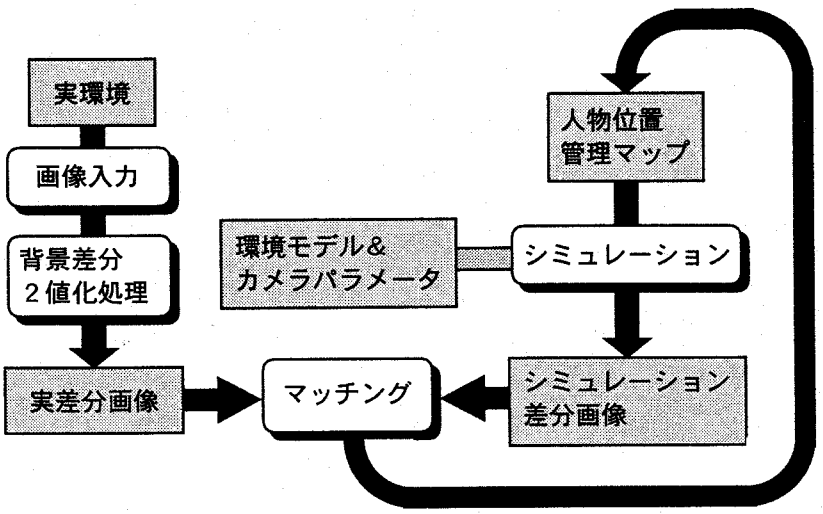


図２－６ 提案手法の概要

環境座標系は，(X,Y,Z)の３次元直交座標系とし，Ｚ軸を鉛直方向上向き，Z=0平面を床面としている．環境モデルは以下の２つの形式でシステム内に保持され，追跡に用いられる．

（ａ）人物移動可能領域の記述（図２－７）

壁や棚，パーティション壁などにより，人物移動は制限される．そこで，人物の移動可能領域を明示することで，現実には移動不可能な位置に対するシミュレーションを防止する．記述には図２－７のような形式を用い，人物の移動可能領域と不可能領域を環境座標系の矩形で記述することで表現する．

（ｂ）構造物の記述

環境中の構造物による人物のオクルージョンに対処するため，環境中のオブジェクト（机，パーティションなど）を３次元モデルとして記述しておき，人物モデルの奥行きと比較する（図２－８上）．これにより，生成されるシミュレーション差分画像はオクルージョンが考慮されたものとなる．ただし，この画像には人物領域のみを描画すればよく，構造物を描画する必要はない．つまり，人物領域を隠蔽しているかどうかという判定にのみ利用されるので，構造物の色やテクスチャー情報も必要なく，３次元形状情報のみが表現されていればよい．そこで，物体表面形状を表現するサーフェイスモデルでこれら構造物の記述を行う．また形状の記述精度や記述すべき物体の範囲に関しては，人物自体も１つの楕円体という近似を行っていることを考えると，物体の細かな凹凸を表現する必要はないと考え，１辺が数十ｃｍ程度の長方形パッチで構造物の概形を表現するに留めている．この処理の詳細は2.3.3節において詳説する．

```
# map description for tracking system.
# mode: sx ex sy ey height [mm]: type
#  object type 0:space,1:wall,2:object
1: -4500  5500 -6000 6000 3000: 1
0: -4200  2000 -1000 4600    0: 0
0:     0  2000  4600 5900    0: 0
2:   350  2000   500 3500  740: 2
2:  2000  3300 -1500  500    0: 3
2:  -200 -1500   100 1000  800: 2
2: -2500 -2000  1500 4500  650: 2
```

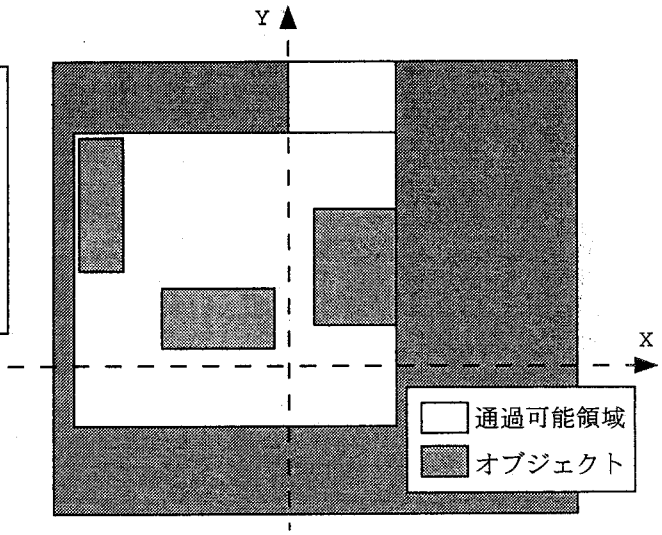


図２－７　人物の移動可能領域の記述例（左）と生成された移動可能領域（右）

（２）人物モデル（楕円体モデル）

前述したように，人物モデルとしては規定身長，規定幅を持つ楕円体を用いる．これにより2.1節で述べたようにセルフオクルージョンの回避や低い計算コストなどを実現でき，２Ｄの追跡アルゴリズムと同様の高速性・安定性を実現できる．また複数の人物モデルを同時生成しても比較的単純なシミュレーション画像が得られるため，複数人物追跡時の安定性にも寄与している．

2.2.2 人物位置管理マップ

追跡中のすべての人物の位置情報を記述する．現在は世界座標系のZ=0平面を床面としているので，世界座標系の残りの２成分であるX,Y座標値を記述している．またシミュレーション差分画像の生成で必要とされる楕円体モデルの高さと幅情報も各人物毎に与える．この値は他のセンサから獲得するか事前知識として準備しておく必要があるが，現在は成人を仮定して，高さ165[cm]，幅40[cm]の固定値を用いている．

2.2.3 カメラパラメータ

カメラは環境内の高い位置から人物が視野内に入るように固定されている．レンズに非線形ひずみが無く理想的な透視変換が仮定できれば，カメラの画像座標系(xc,yc)と環境座標系(X,Y,Z)の間は，媒介変数 h を用いた以下の式で表される．

$$\begin{bmatrix} hx_c \\ hy_c \\ h \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad \text{（式１）}$$

この 3x4 の行列をカメラパラメータと呼び，カメラの内部パラメータ（焦点距離，撮像素子のサイズ，撮像素子のひずみ等）およびカメラの外部パラメータ（環境座標系に対する位置，姿勢）が合成されてできる行列である．カメラの設置時には，この行列の 11 個のパラメータを求めることが必要となる．これには，画像中の６点以上の特徴点の世界座標系の３次元座標値と，それに対応する画像平面内での２次元座標値を最小二乗法を用いることによって導出できる[Inokuchi90]．

2.3 シミュレーション差分画像の生成

2.3.1 視野領域内人物の選択

本手法は分散視覚システムでの利用を考えているため，人物位置管理マップには，カメラに写らない他の観測ステーションの観測視野中に存在する人物も登録されていると仮定する．そこでまずカメラに写る人物を選択する．これは実際に画像内に描画される人物を事前に選択することで処理コストを削減するためである．

人物位置管理マップに登録されている人物 P の位置を(Xp,Yp)，その人物の高さを Hp とすると，（式１）に人物の最高点(Xp,Yp,Hp)と最低点(Xp,Yp,0)を代入することで，それらの画像平面内ので座標値を求めることができる．この座標値のどちらかが画像サイズ内に含まれる場合にその人物を描画するものとする．

2.3.2 楕円パラメータの計算

シミュレーション差分画像は人物領域が画素値１，背景が画素値０で塗りつぶされた２値画像である．人物を３次元空間内の楕円体で近似しているので，その楕円体がカメラに写った場合の見え方をシミュレーションすることになる．今，人物の世界座標系での位置を(Xp,Yp)，その人物の幅と高さを 2Ra，2Rb とすると人物を表現する楕円体の方程式は次のようになる．

$$\frac{(X-X_p)^2}{R_a{}^2}+\frac{(Y-Y_p)^2}{R_a{}^2}+\frac{(Z-R_b)^2}{R_b{}^2}=1 \qquad \text{（式２）}$$

また，カメラパラメータの基本式（式１）に対して形式的に４行目を追加し正方行列とすることで，その逆行列を求め，次の式を得ることができる．

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} hx_c \\ hy_c \\ h \\ 1 \end{bmatrix}$$
$$\cong \begin{bmatrix} D_{11} & D_{12} & D_{13} & D_{14} \\ D_{21} & D_{22} & D_{23} & D_{24} \\ D_{31} & D_{32} & D_{33} & D_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} hx_c \\ hy_c \\ h \\ 1 \end{bmatrix} \qquad \text{（式３）}$$

この式のX,Y,Zを（式２）に代入することで，以下に示すhに関する２次方程式を得る．

$$\begin{aligned}
&(a_1x_c{}^2+a_2y_c{}^2+a_3x_cy_c+a_4x_c+a_5y_c+a_6)h^2 \\
&+2(b_1x_c+b_2y_c+b_3)h+c_1=0 \\
&a_1=R_b{}^2(D_{11}{}^2+D_{21}{}^2)+R_a{}^2D_{31}{}^2 \\
&a_2=R_b{}^2(D_{12}{}^2+D_{22}{}^2)+R_a{}^2D_{32}{}^2 \\
&a_3=2R_b{}^2(D_{11}D_{12}+D_{21}D_{22})+2R_a{}^2D_{31}D_{32} \\
&a_4=2R_b{}^2(D_{11}D_{13}+D_{21}D_{23})+2R_a{}^2D_{31}D_{33} \\
&a_5=2R_b{}^2(D_{12}D_{13}+D_{22}D_{23})+2R_a{}^2D_{32}D_{33} \\
&a_6=R_b{}^2(D_{13}{}^2+D_{23}{}^2)+R_a{}^2D_{33}{}^2 \qquad \text{（式４）} \\
&b_1=R_b{}^2D_{11}(D_{14}-X_p)+R_b{}^2D_{21}(D_{24}-Y_p)+R_a{}^2D_{31}(D_{34}-R_b) \\
&b_2=R_b{}^2D_{12}(D_{14}-X_p)+R_b{}^2D_{22}(D_{24}-Y_p)+R_a{}^2D_{32}(D_{34}-R_b) \\
&b_3=R_b{}^2D_{13}(D_{14}-X_p)+R_b{}^2D_{23}(D_{24}-Y_p)+R_a{}^2D_{33}(D_{34}-R_b) \\
&c_1=R_b{}^2(D_{14}-X_p)^2+R_b{}^2(D_{24}-Y_p)^2+R_a{}^2(D_{34}-R_b)^2-R_a{}^2R_b{}^2
\end{aligned}$$

この解が重解の場合，カメラ画像平面内の座標値(xc,yc)を通過する視線方程式が（式２）の楕円体と接することを意味する．そこで判別式（式５）＝０とする方程式をつくることで，その式のxc,ycは，楕円体がカメラに写ったときの輪郭線を表現する．

$$\frac{1}{4}D=\left(b_1^2-a_1c_1\right)x_c^2+\left(b_2^2-a_2c_1\right)y_c^2+\left(2b_1b_2-a_3c_1\right)x_cy_c$$
$$+\left(2b_1b_3-a_4c_1\right)x_c+\left(2b_2b_3-a_5c_1\right)y_c+\left(b_3^2-a_6c_1\right)=0 \qquad \text{（式５）}$$

この式は最終的には，次の形式の楕円方程式にまとめることができる．

$$\mathrm{K}_1x_c^2+\mathrm{K}_2y_c^2+\mathrm{K}_3x_cy_c+\mathrm{K}_4x_c+\mathrm{K}_5y_c+\mathrm{K}_6=0 \qquad \text{（式６）}$$

2.3.3 奥行きマップによるオクルージョンへの対処

（式６）で表される楕円方程式は画像平面内での人物領域の輪郭線を表現しているので，この領域内部を画素値１で塗りつぶすことでシミュレーション差分画像が生成できる．しかし，環境中の構造物などによるオクルージョンを考慮する必要がある．これにはコンピュータグラフィックスの隠面消去法として一般的なＺバッファ法を利用する．

2.2.1節(b)で示した構造物の記述とカメラパラメータをもとに画像内の各画素毎に構造物までの距離を計算し，奥行きマップ，いわゆるＺバッファを事前に生成しておく（図２－８）．この距離値と描画される楕円体モデルの距離値を各画素毎に比較し，楕円体モデルの方がカメラに近い画素のみ描画を行う．

Ｚバッファの作成は事前に行うことができるためにその計算速度は問題にはならないが，塗りつぶされる楕円内の各画素に対して距離値を求めるための計算コストが問題となる．そこで，楕円内全領域の奥行き値を楕円体中心の奥行き値で簡略化させることで計算コストを削減した．また，Ｚバッファの生成も含めた距離計算には，３次元距離ではなく，Z=0 平面に射影した２次元平面内での距離値を利用することにした．

```
# depth map description for tracking system.
# x1, y1, z1, x2, y2, z2, x3, y3, z3, x4, y4, z4

-9999 -9999 0,  9999 -9999 0,  9999  9999     0, -9999  9999     0
-4500 -2000 0, 2000 -2000 0,  2000 -2000  3000, -4500 -2000  3000
-4200 -2000 0, -4200  4600 0, -4200  4600  3000, -4200 -2000  3000
-4500  4600 0,     0  4600 0,     0  4600  3000, -4500  4600  3000
    0  4600 0,     0  6000 0,     0  6000  3000,     0  4600  3000
 2400 -2000 0,  2400  6000 0,  2400  6000  3000,  2400 -2000  3000
```

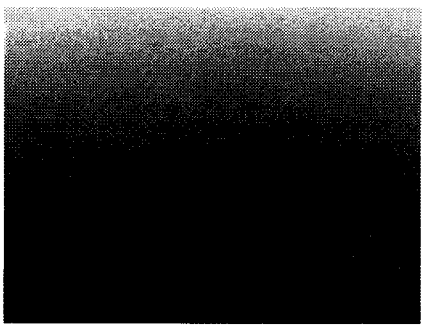


図２－８　環境中オブジェクトの記述と奥行きマップ（Ｚバッファ）
画像の輝度は値の大小（奥行き）を表す

2.4 比較処理

2.4.1 マッチングによる人物位置の決定

人物位置管理マップに記述されている人物位置を微少移動させ，2.3節で述べた方法により複数枚のシミュレーション差分画像を生成し，実画像との類似度から現時点での人物位置を決定する．カメラ視線方向を床面に投影し，その方向とそれに垂直な方向の２軸を考え，その各軸方向に，-50[mm], 0[mm], +50[mm] の３段階に人物位置を変化させ，合計９枚のシミュレーション差分画像を生成する．１回の画像入力に対して，この９枚のシミュレーション差分画像とのマッチングによる位置更新を３回繰り返し，最後に求められた人物位置を人物位置管理マップに記録する．つまり１回の画像入力で，最大150[mm]の移動に追従でき，10frame/sec の処理速度が達成できれば，1500mm/sec の移動速度まで対応可能になる．

2.4.2 ３値画像によるマッチング

マッチングでは，以下の式で評価を行う．

$$E = \frac{\sum_{x,y}\{\mathrm{sim}(x,y)\cdot \mathrm{real}(x,y)\}}{\sum_{x,y}|\mathrm{sim}(x,y)|} \qquad (式７)$$

sim(x,y), real(x,y)はそれぞれシミュレーション差分画像，実差分画像の座標値(x,y)での画素値である．両画像とも画像サイズは同一であり，基本的には0,1を画素値とする画像である．しかし，図２－９に示すようにシミュレーション画像が人物領域に完全に包含されるような場合，評価値は最大となりマッチングを誤る．また，影などの影響により画像に多くのノイズが発生した場合も，ノイズ領域を包含する楕円領域では評価値は高くなり，マッチングを誤る可能性が

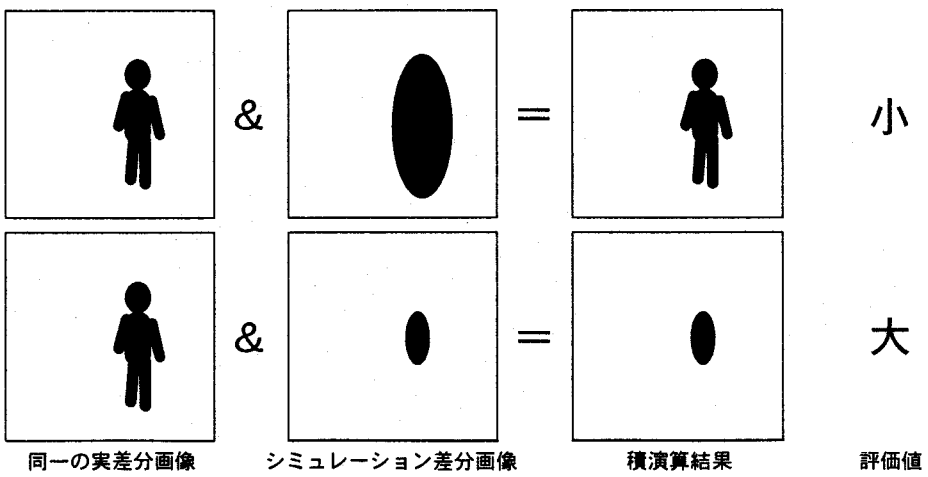


図２－９　３値画像による類似度計算

高くなる．このような影響を抑制するために輪郭のマッチング度合いを評価値に反映させることとした．

つまり，シミュレーション差分画像，実差分画像のそれぞれにおいて，1の値の領域に対して膨張処理を施し，膨張領域には-1の値を割り当て３値画像として考える．これらの画像で（式７）の計算を行うと，共に-1という値を持つ画素では評価値を高めるが，1と-1とくいちがう画素では評価値を下げる効果を生み，輪郭のマッチング度合いを評価値に反映させることができる．2.6節の実験においては膨張幅を５画素としている．

また，画素値のいずれかが0の部分は評価式分子の計算に不要なので，シミュレーション差分画像の描画領域情報から評価値計算領域を限定でき，ここでも計算コストを削減できる．

2.5 拡張

以上に述べたことが基本的な処理方法であるが，ここでは複数人の追跡のための処理，および，画像視野外から進入してくる人物の追跡に関する処理について説明する．

2.5.1 複数人追跡への拡張

複数人の追跡において問題となるのは，人物同士が見かけ上で重なる場合である．本手法においても，各人物毎に独立に処理を行ってはこの問題に適切に対処できない．そこで，すべての人物に対する処理を一括して行う．図２－10にその手順を示す．１人に対して９通りの位置

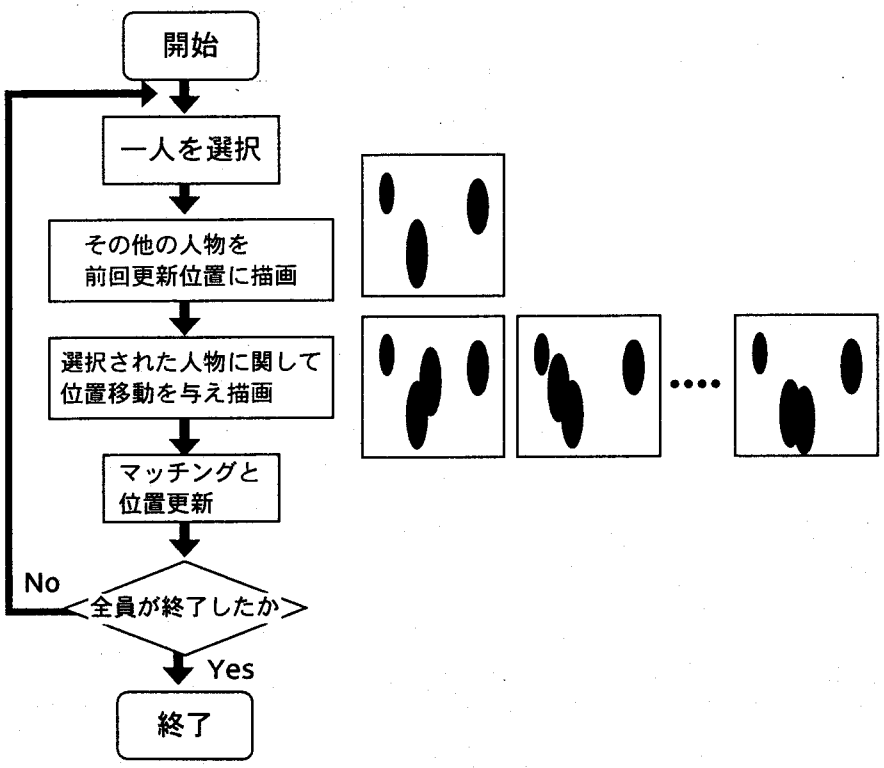


図２－10　複数人物追跡の処理手順

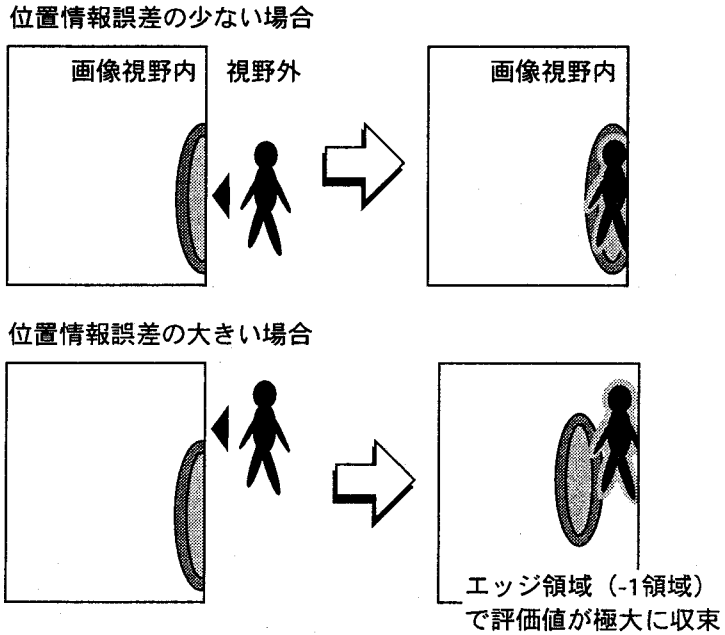


図２－11　類似度計算の問題点

変化候補を与えるとすると，N人だと9^N通りの組み合わせとなる．しかし，計算コストの点で問題が生じるので，一人を変化させる時には他の人物の位置は固定しマッチングを行い，その人物についてのみ位置を更新する．すべての人物について，順次この処理を繰り返す．この方法だと，N人だと9Nの組み合わせとなる．

2.5.2 追跡モードと捕捉モード

複数システムや他のセンサ情報との統合を考えると，カメラ視野外から視野内に人物が出現する場合や視野内から視野外に消滅する場合に対応する必要がある．本手法においては，視野外から視野内に出現する人物の追跡を開始する場合に，観測データの誤差に起因する問題が生じる．これは，2.4.2節で示した３値画像のマッチングに関連し，図２－11に示すように，-1という画素値の効果により実画像上の人物領域にシミュレーション画像の人物領域が収束しない状況が生じるからである．すなわち互いの輪郭領域（-1の画素を持つ領域）が重なった状況では，その位置の評価値が周りの位置の評価値に比べて局所的に高くなり，正しい解の探索が行えなくなる．そこで，視野外から視野内に出現する人物の初期補足の段階では，初期捕捉モードとして－１領域を設定しないマッチングを行う．そして（式７）の評価値が閾値を越えた時点で通常の追跡モードに復帰するようにした．

2.6 実験と検討

2.6.1 実験環境

提案手法の人物静止状態での位置検出精度の評価，および人物移動時の通過軌跡の評価実験を行った．カメラ配置を図２－12に示す．水平床面上の適当な場所に世界座標原点（Z軸が鉛直方向上向き）を置き，その座標系内において，人物を斜め上方から見下ろす配置でカメラ１，真上から見下ろす配置でカメラ２を設置した．この２種類の配置のカメラを用いてそれぞれ実験を行い，カメラ配置による影響も考察する．カメラからのビデオ信号は，パーソナルコンピュータ(Intel PentiumII-333MHz)で縦横各120x160画素の大きさで取り込まれ処理される．

2.6.2 カメラのキャリブレーション

縦横各300mm，高さ1100mmの基準直方体（キャリブレーションボックス，図２－13）を作成し，これを用いてカメラのキャリブレーションを行った．基準直方体を世界座標系での定点に置き，画像中の基準直方体のマーカ位置をマウスで指示すると共に，対応する世界座標系での座標値を入力する．基準直方体を移動させながら，この操作を繰り返し，30点のデータを用いて（式１）のカメラパラメータを最小２乗近似により求めた．得られたカメラパラメータより計算された各カメラの位置は，それぞれ(-1.00m, -1.74m, 2.64m), (-1.07mm, 1.32m, 3.09m)であった．また，カメラ２では広角レンズ使用による樽型レンズ歪みが観測されたために，その影響も調べた．キャリブレーションに用いたデータの画像中の座標値に対応する世界座標系での床

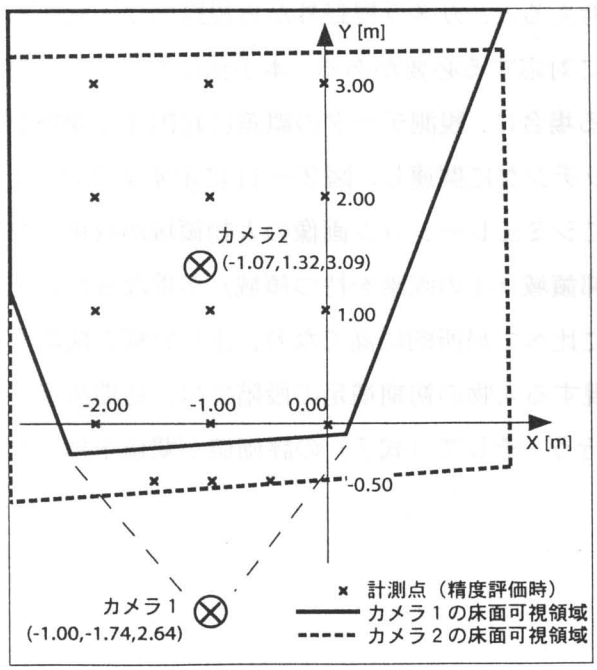


図２－12　２台のカメラ配置と観測視野

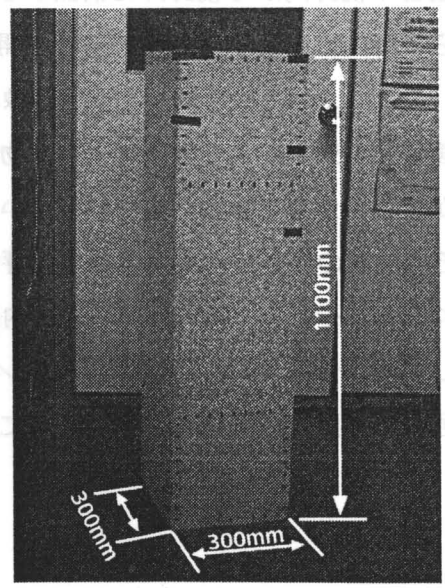
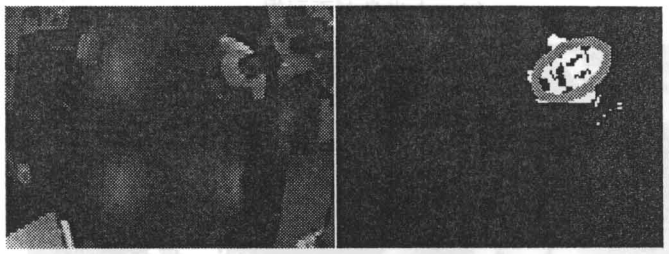


図２－13　キャリブレーションボックス

(a) カメラ１



(b) カメラ２

図２－14　２つのカメラ画像（左は入力濃淡画像，右は背景差分画像の画素値１の領域と選択されたシミュレーション画像の輪郭領域を重ね合わせたもの）

面位置をカメラパラメータを用いて逆算し，真値と比較したところ，画像中央部で0.02[m]程度，画像周辺部で0.10[m]程度の誤差が確認できた．つまり，カメラ２による計測では常にこの程度の誤差が生じると考えられる．

2.6.3 計測精度評価

人物静止状態での位置検出精度を評価するため，人物が定点に立ち止まっているときの計測誤差を評価した．図２－12に示すように計測点をとり，カメラ１では(x,y) = (0m,-0.5m) ～ (-2m,3m)の範囲内にある計15点，カメラ２では(x,y) = (0m,0m) ～ (-2m,3m)の範囲内にある計１２点で計測を行った．カメラ１，カメラ２による画像を図２－14に示す．背景差分処理により多少の領域欠損を伴って人物領域の抽出は行われている．また，影や映り込み部分も同時に抽出されていることが確認できる．また，背景の単調さ・複雑さによる人物領域抽出性能の変化はなかったが，単調な背景の上に同一の濃淡レベルの人物像が写る場合には大きな領域欠損が生じる場合もあった．

カメラ２では，人物を上方向から垂直に見下ろす視線となるのでほとんどの場合画面中には人物の全身が写るが，カメラ１では人物をななめ上方から見下ろす画像が得られるため，画像境界によって人物の上半身が写らない状態（図２－15ａ）や下半身が写らない状態（図２－15ｂ）が生じる．そこでカメラ１では，

(a) 人物の全身が写った場合(計測範囲(0m,0m)～(-2m,2m):計９点)

(b) 人物の下半身のみが写った場合（計測範囲(0m,3m)～(-2m,3m):計３点）

(c) 人物の上半身のみが写った場合（計測範囲(0.5m,-0.5m)～(-1.5m,-0.5m):計３点）

(a)　上半身が不可視

(b)　下半身が不可視

図２－15　画像境界で部分的なオクルージョンが生じた画像

表２－２　計測誤差

	計測点数	誤差平均[m]	Ｘ軸方向[m]	Ｙ軸方向[m]
カメラ１(a)	９点	0.11	0.05	0.09
(b)	３点	0.12	0.04	0.11
(c)	３点	0.06	0.05	0.03
カメラ２	１２点	0.21	0.11	0.17

の３タイプに分割して評価した．なお被験者は身長が166[cm]，胴体横幅最大値は42[cm]であった．楕円体モデルには高さ165[cm]，幅40[cm]という値を用いている．表２－２に計測誤差を示す．計測値と真値の距離の平均を誤差平均とした．また，ＸＹ各軸方向に関する誤差平均も示す．

計測値の量子化誤差は，楕円体モデルの生成間隔（0.005m）と入力画像１フレームに対する繰り返し処理回数（３回）から0.015[m]と算出できる．これを考慮すると計測結果は全体的には良好であると思われる．カメラ２の結果の誤差が大きいのは，レンズ歪みの影響の他に，ある計測点における人物の影による比較的大きな誤差が原因となった．これは真上から見下ろすというカメラ配置固有の問題ではなく，斜めからの場合でも照明との位置関係により起こりうる問題である．この点を考慮すると両カメラにおける計測精度の比較において画像のレンズ歪みによる違い以外の差はみられなかった．また，カメラ１の視線は世界座標系においてほぼＹ軸に沿ったものなので，Ｘ軸が画像面内方向，Ｙ軸が奥行き方向に対応する．表中カメラ１(a)(b)の結果から奥行き方向の誤差がわずかに大きなことが確認できるが，個別のデータからも奥行き方向に対する誤差が大きいことが確認できた．これは斜めから見下ろすカメラ配置において，奥行き方向の人物移動に対する画像変化が面内方向に比べて小さいことに起因する．

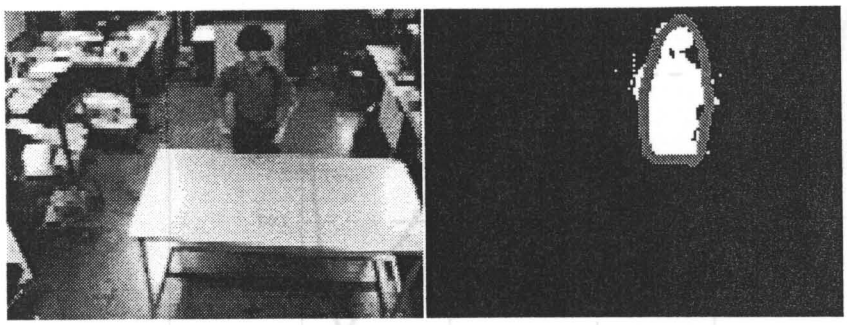


図２－16　構造物によるオクルージョンが生じた状態

表２－３　オクルージョンによる計測誤差への影響

	計測点数	誤差平均[m]	X軸方向[m]	Y軸方向[m]
カメラ１(a)	９点	0.18	0.06	0.15
(b)	３点	1.31	0.18	1.29

次にカメラ１において障害物によるオクルージョンが生じた場合の計測精度を評価する．計測範囲は(0m,0m)～(-2m,3m)の12点とし，すべての計測点において人物の足下が隠れるように，高さ0.73mの机を移動させながら計測した（図２－16）．なお，この机の配置はその都度環境モデルを更新した．計測結果を表２－３に示す．足下は隠れるが上半身が写っている(a)の場合では，Ｙ軸方向の誤差が多少増加したのがわかる．(b)の場合では，足元と頭部の両方が隠れてしまい胴体の一部が写る画像が得られる．この場合，人物のＹ軸方向の変化に対しシミュレーション差分画像はほとんど変化がなく，表からもＹ軸方向に大きな誤差が生じているのがわかる．頭部と足元が同時に隠れる状況に対しては，本手法では計測不能領域として取り扱わなければならない．

2.6.4 複数人を対象とした追跡実験

複数人物を対象にした追跡実験として，2.6.3と同様の２台のカメラによって，２人の人物が同時にカメラ視野内を移動する際の人物の通過軌跡を評価した．人物の歩行経路を図２－17に示す．人物１の被験者の身長は173[cm]，胴体横幅最大値は48[cm]で，人物２は2.6.3節の被験者と同一人物である．楕円体モデルの高さ・幅値は各160[cm]・40[cm]と変更していない．図中の破線で示した位置では，カメラ１の画像中で人物が重なって観測される（図２－18）．カメラ２では人物は重なって観測される箇所はない．計測された人物の歩行軌跡を図２－19に示す．全体的評価として歩行追跡は両方のカメラで可能だった．カメラ１においては人物が重なって観測される部分においても良好な結果が得られた．カメラ２の結果には計測誤差が目立つ．これはカメラ２において，画像中の人物領域が比較的小さいにもかかわらず，床に写る人物の影が大きく写ったことが影響した．

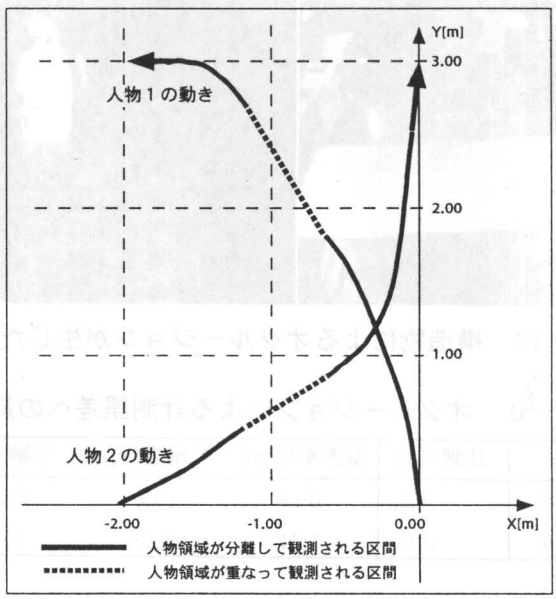


図２－17　２人物の歩行経路

図２－18　２人物が重なって観測される状態

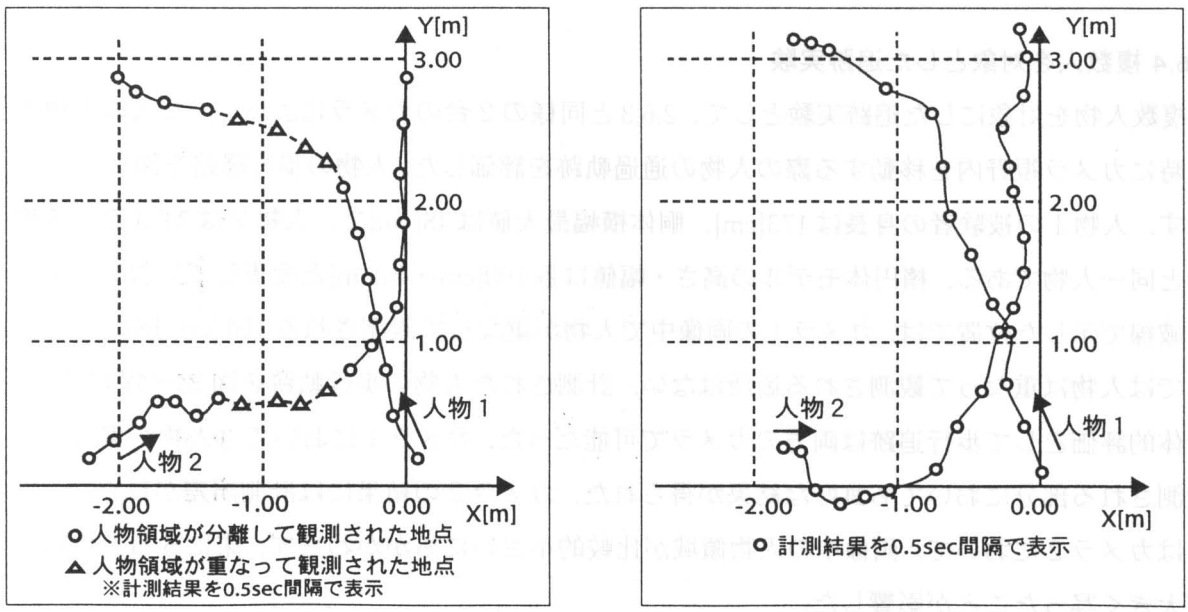


図２－19　歩行人物の追跡結果

2.6.5 検討

本実験は特別な環境を設置すること無く通常の屋内環境で行われており，影や映り込みが生じる環境にもかかわらず安定した追跡結果が得られた．処理速度は，１人の追跡の場合20～30 frame/sec，２人の場合5.6～10 frame/sec，３人の場合2.7～5.5 frame/secであった．１人の場合は小走りする人物の追跡も可能だったが，３人の場合ではゆっくりとした歩行でない場合には追跡に失敗する場合もみられた．本手法では，ノイズなどの影響による処理速度の変化はないが，シミュレーション画像の生成枚数を起因とする，追跡人数にほぼ比例した処理時間を必要とする．このため，処理速度の観点からは，現状の処理装置では３人の追跡が限界となっている．しかし，視野の重なる複数のシステムに処理を分担させることにより，この問題も解決可能と思われる．

複数人物が画像上で重なりあう状況に関しても，本手法では良好な追跡結果を得ることができた．しかし，人物同士が見かけ上ではなく実際に非常に接近して交差するような場合には，追跡人物が入れ替わるという問題が生じる．現在は２値の背景差分画像のみを情報として用いているが，このような人物の入れ替わりが予想される場合には，色情報などを用いた人物識別を検証処理として追加し，人物の入れ替わりを修正する方法が考えられる．また，人物の密度が高い環境においては，真上から見下ろすカメラ配置を採用することで，人物の見かけ上の重なりを低減することも可能である．

計測精度に関しては，本システムへの応用に対して十分な精度が得られたと考えられる．計測精度に影響する主な要因は，

１）人物モデルと実際の人物の大きさの差異

２）影・写り込み領域の発生・人物領域の欠損

３）画像内の人物領域の大きさ

４）人物の大きな姿勢変化

の４点であった．１に関してはモデルの高さ情報の値が実際の人物の身長と異なる場合に定常的な誤差が観測されたが，5%程度の差であれば，その影響により追跡不能に陥ることはなかった．２に関しては，影や写り込みが人物領域以上に大きな領域として観測されるような場合や人物像の50%以上の領域が抽出できなかった場合に，大きな誤差が発生し，追跡不能に陥ることもあった．影・写り込みをシミュレーションし画像に反映させる方法も考えられるが，処理時間の増加が予想される．領域欠損に関してはカラー画像による背景差分処理の利用も考えられる．さらに，実応用においては，影や写り込みの発生を押さえるような環境に対する工夫や，それらの影響を受けにくいカメラ配置の検討が必要になると思われる．また，照明環境の変化や人物以外の物体移動に対処するためには，背景画像の更新処理も導入しなければならない．３に関しては，今回の実験において人物全身が写る場合の人物領域の画像面積に対する割

合は3～8%程度であったが，その範囲における計測精度は2.6.3節で示したとおりである．しかし，廊下などに俯角を浅くカメラを設置した場合，遠くの人物まで観測することは可能である．この場合は人物移動に対する画像変化の減少に加え，画像内の人物領域が小さくなることによるノイズ領域による影響が大きくなる．画像面積に対する人物領域の割合が1%以下の画素数の場合では，事実上追跡は不可能であった．４に関しては，座り込むといった姿勢変化によって大きな位置誤差が生じたが，追跡の安定性には大きな影響を与えなかった．

本研究の最終目標である複数台のシステムでの統合を考えた場合，本手法では人物管理マップ上に人物位置を世界座標系で表現しているので，これを複数システムで共有することにより統合は容易に行えると考えられる．しかし，複数のカメラに同一人物が写るような場合の効果的な処理の分担方法などを検討する必要がある．また，現在は水平な床面を仮定しているが，追跡処理の中からその制約を取り除くことは容易であり，環境モデルの表現法を工夫することで，階段やスロープといった環境に対しても本手法は適応可能である．

2.7 本章のまとめ

本章では，分散視覚システムによる人物追跡を実現するため，各観測ステーションで実行する画像処理アルゴリズムについて述べた．本手法は，人物移動に対するカメラによる見え方を複数シミュレーションし，それら画像と入力画像のマッチングを行うことを特徴としている．人物を１つの楕円体で表現することにより，計算コストを低減し，リアルタイム処理が可能になり，複数人物の追跡も可能である．また，モデル生成を環境座標系で行うことによって，得られる結果も環境座標での人物位置となる．これにより，複数のステーション間での統合が容易な検出結果が得られることになる．

3章 複数の観測ステーションによる広域観測

本章では2章で実現された人物追跡機構を元に，複数の観測ステーション間で連続的に人物を追跡する手法に対して述べる．本手法を開発するにあたっては，対象領域を屋内の実環境と想定し，以下の点が実現されるべきであるとしている．

1．観測ステーションの観測視野の違いや隠蔽などによって，観測不可能な領域が存在する状況でも使用可能であること．
2．複数の観測ステーションから得られた観測情報の統合．
3．一部の観測ステーションの障害がシステム全体に影響を及ぼさなくすること．また新たに起動したステーションが柔軟にシステムに組み入れられること．

1の条件は，本システムが屋内実環境での利用を目的としているため，観測ステーションが異なった領域を観測することで広域をカバーすることが必要となる事や，観測範囲中には通路や扉などによる隠蔽，各ステーションの観測領域間で不可視な領域が存在することが考えられるために必要となる．また2の実現によって，複数の観測ステーションによる対象の同時観測時に，単独のカメラで行った場合より良い結果を得ることが期待できる．更に3に対処することで，分散システムにおいて想定される「一部のステーションの動作停止」や「新たなステーションの追加」といったイベントに柔軟に対応できるようになる．

本システムのように「視覚機能を持つ自律観測プログラム（エージェント）の協調」によってタスクを実現する試みは，分散協調視覚プロジェクト[Matsuyama97]やVSAM Project [Haritaoglu99]，西尾らのシステム[Nishio92]でも行われている．前者のプロジェクトで行われている研究[Matsuyama98]では，観測対象は複数のエージェントが同時観測可能な領域に存在することを想定しており，本研究で目的とする「不可視領域を含む実環境」のような状況での動作は対象としていない．また[Nishio92]は複数の視覚システムを用いて交差点での車両追跡を目指しているが，[Matsuyama98]と同様，常にいずれかのカメラで対象が観測可能であることを想定しているため同様の問題が生じる．

本章では，まず本システムで用いる複数ステーション間の人物追跡するアルゴリズムについて述べ，次にそのアルゴリズムを効率的に実現するためのシステム構成に関して検討を加える．その後，実システムの構成および実験結果について述べる．

3.1 複数ステーションにおける人物の連続追跡アルゴリズム

まず，冒頭で述べた環境下で対象を連続して観測するための基本的アルゴリズムを述べる．本システムにおける観測ステーションが持つカメラは環境に対して固定されているため，あらかじめすべての観測ステーションの観測可能領域を計算しておき，人物の現在位置に応じて追跡を行うステーションを切り替えるというタスク制御を行うことが考えられる．具体的には，各ステーションにおいて以下の３つのモードを切り替えることによって実現する．

1．人物がいずれかの観測ステーションの視野内に存在する場合は，その領域を観測するステーションで追跡・位置検出を行う．

2．人物が複数の観測ステーションによって同時に観測できる領域（共同観測領域）に進入した場合（図３－１），互いに検出した位置を交換しながら追跡を行う．複数視点から

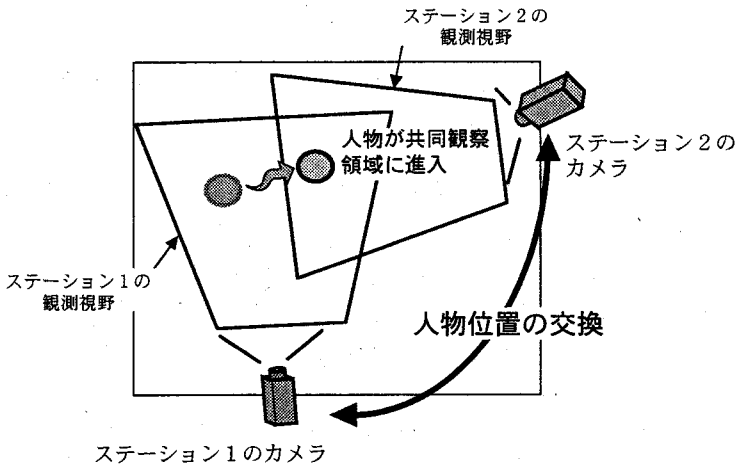


図３－１　人物が複数の観測ステーションの視野に存在する場合の協調アルゴリズム

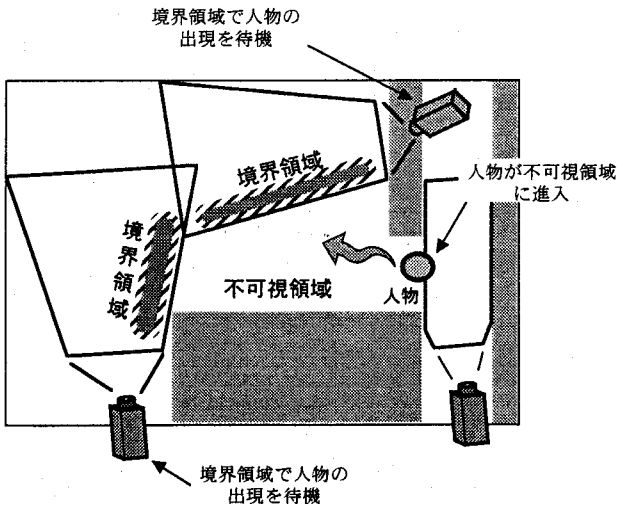


図３－２　人物が不可視領域に親友する場合の協調アルゴリズム

同時に対象人物の位置計測を行うことになるため，計測結果を統合することでより正確な人物位置を推定することができる．また，一方のステーションが追跡に失敗した場合でも，他方のステーションの計測位置情報を用いて捕捉動作を行うことで，人物を発見し追跡を再開することが可能になる．

３．観測ステーションの観測領域の間に「不可視領域」が存在する場合（図３－２）には，その領域に接する観測領域を持つステーションが，観測領域の境界で捕捉動作を行い人物の出現を待機する．捕捉動作を行ういずれかのステーションが人物の出現を検出すると，その情報を他のステーションに伝えることで捕捉動作を終了する．

１の人物追跡・位置検出は前章で提案した手法によって可能となっている．一方2および3の動作を実現するには，１．画像から人物を捕捉する，２．互いの検出誤差特性を考慮して推定位置を求める，３．他のエージェントの観測領域と自分の観測領域との関係から進入予測領域を求め，人物の出現を待機する，４．他のエージェントへ人物の進入予測位置を指示する，などの処理が必要となる．１に関しては前章で述べた捕捉モード（2.5.2節）によって実現することができるが，２～４の処理には，複数のエージェントがコミュニケーションを行う「協調処理」が必須である．

3.2 システムの制御方式

ここで，このアルゴリズムを実現するためのシステム全体の制御方式に関して考察を行う．1.3節において，本システムのタスクとそのためにとりうるべきシステム構成について簡単に触れた．すなわち，本システムが対象とする「複数の視覚システムによる広域観測」というタスクについて考えると，その処理速度や耐障害性，システム設置の自由度の点から分散型を取るべきであるが，タスク遂行のためにはステーション同士の観測視野の関係などを考慮する必要があり，集中制御型の要素も取り入れる必要がある．この点を前述したアルゴリズムに対して整理して考えると，以下のことに気づく．

- ステーション同士の観測視野の関係を求めるには，動作しているすべてのステーションの観測視野情報が必要である．

- 一度観測視野の関係が求まれば，人物の位置情報等の交換は観測視野が交差するステーション間（図３－１）あるいは不可視領域を挟んで隣接するステーション間（図３－２）のみで行えばよい．

以上の点から，まずステーションの起動時に集中的に情報を管理する存在から全ステーショ

ンの情報を取得し，それらの観測視野の関係を導き出す．その後のシステム動作では，観測ステーションの観測視野の関係に基づく部分的なコミュニケーションのみでよい．また冒頭でも述べたように本システムが分散システムであることを考えると，システムの動作中に一部のステーションが停止したり，逆に新たなステーションが追加される事にも対処する必要がある．これらの点を考慮し，システムの制御方式として以下のような構成を考えた（図３－３）．

1．システムの拡張性や一部ステーションの動作停止や起動イベントに対する柔軟性をはかるため，全観測ステーションに指示を与えるような存在はおかず，ステーションはすべて同等の立場の存在にある．

2．ただし，システム全体の情報を把握し，観測ステーションが要求したときにそれを配布できるような存在（システム情報管理エージェント）を設ける．このエージェントをシステム内に複数置くことを許すことで，耐障害性を持たせる．

3．観測ステーションは起動時にのみ，他の全ステーションに自分の視野情報を伝えるメッセージをブロードキャストすることで，自分の存在を伝える．

4．ステーションの障害は，システム動作中の観測ステーション間のメッセージ交換によって検知する．つまり，観測ステーション間のメッセージ交換においては，受信ステーションは送信ステーションに対して確認メッセージ(acknowledge)を返すという取り決めを作っておき，これが返されない場合はそのステーションが停止していると判断することにする．隣接ステーションが停止したと判断したステーションは，その情報を全ステーションへブロードキャストすることで障害を伝える．

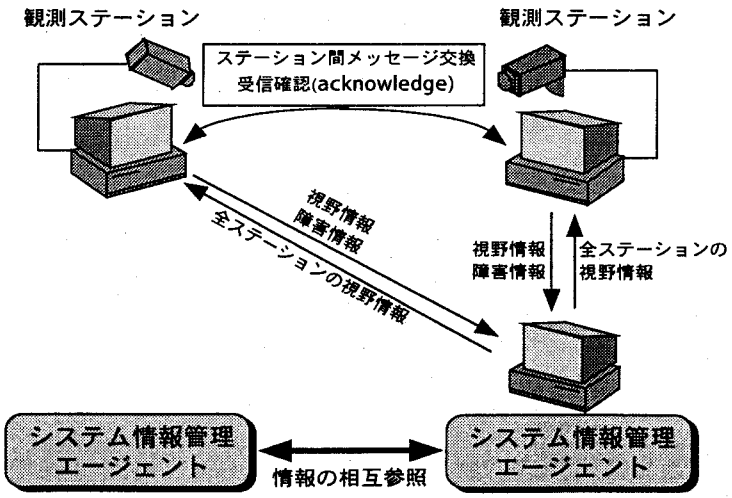


図３－３　システムの制御方式の概要

5．システム情報管理エージェントがステーションの起動メッセージや障害メッセージを受け取った場合，自己内に保持する全ステーションの動作状況データベースを更新する．また，システム内に複数のシステム情報管理エージェントが存在する場合は，一定時間ごとに通信し，相互の情報を参照し合うことで誤りを防ぐ．

システム情報管理エージェントは，インターネットの分散管理システムにおけるデータベースプロトコルであるＮＩＳ(Network Information Service)と同様の働きを果たす．ＮＩＳでは，ＮＩＳサーバと呼ばれるデータベースサーバが，クライアントの要求に応じてシステム共通の情報を返す．ＮＩＳサーバにはマスターサーバとスレイブサーバがあり，スレイブサーバはマスターサーバと一定時間ごとに通信することで，互いの情報を更新する．本システムでは，システム情報管理エージェントがＮＩＳサーバに対応し，観測ステーションはＮＩＳクライアントに対応するが，本システムにおいて特有な点は，提供する情報があらかじめサーバに用意されているのではなく，観測ステーションがブロードキャストする情報に応じて適宜更新されていくという点である．

以上のような構成をとることによって，システム動作中における一部ステーションの起動・停止に対処できるようになる．また，観測ステーション間の通信は基本的に，地理的に隣接するステーションのみで行われるため，全ステーション間で通信を行ったり，あるいは集中サーバから通信を行う場合に比べ使用する通信帯域幅を削減することができる．

3.3 システムの構成

以上の機構を実装した分散視覚システムを構築した．本システムは複数台の観測ステーションと１台以上のシステムパラメータ管理エージェントから構成される．また観測ステーション上では各々１つのプログラム（観測エージェント）が動作している．

3.3.1 観測ステーションと観測エージェント

観測ステーションはカメラが接続された画像処理ワークステーションで，この上では観測領域内の人物追跡や位置検出，隣接する観測ステーションのタスク制御等を行うプログラム（観測エージェント）が動作している．観測エージェントのソフトウェア構成を図３－４に示す．追跡状態変数はそのカメラで観測している人物数だけ確保され，各人物の位置や追跡状態（追跡・捕捉・境界捕捉）が記述されている．システムパラメータには，動作しているすべての観測ステーションのカメラパラメータと画像サイズが保持される．追跡ルールは，動作中の全観測ステーションの観測視野がマッピングされた環境マップで，ここから他の観測ステーションの観測領域，視野が重なる他のステーション，視野が隣接する他のステーションとそれに対応する境界領域の情報を得ることができる．

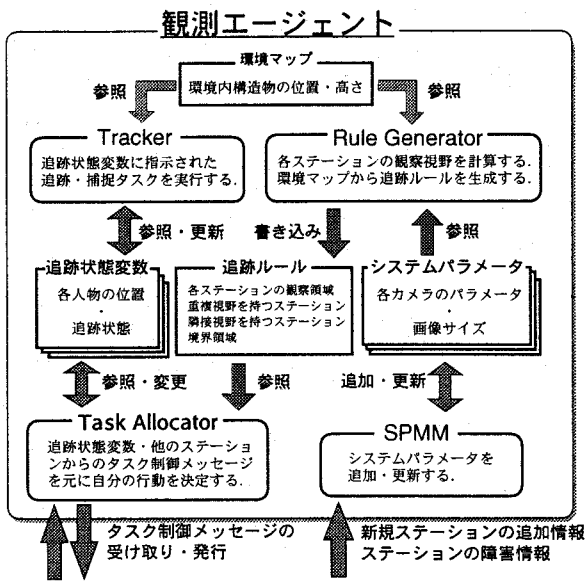


図３－４　観測エージェントのソフトウェア構成

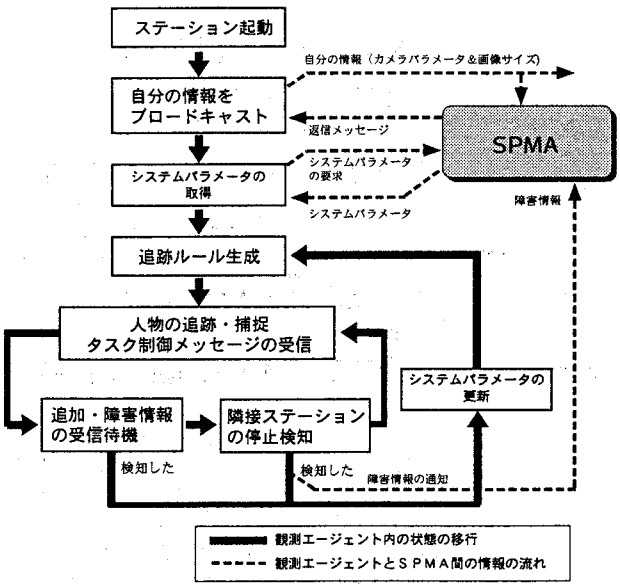
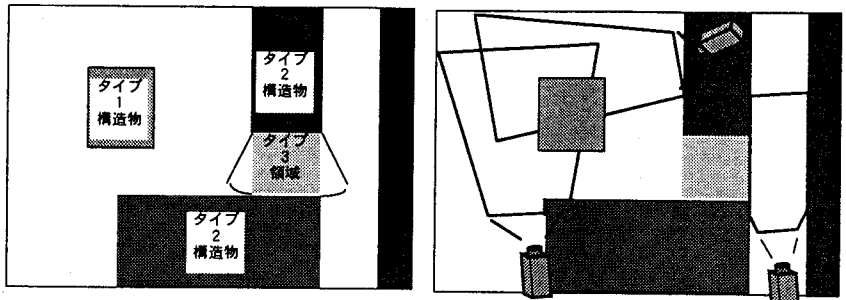


図３－５　観測エージェントの動作の流れ

図３－５に観測エージェントの動作の流れを示す．観測エージェントは起動すると自分の情報を他のエージェントにブロードキャストするとともに，システムパラメータ管理エージェントからシステムパラメータを取得する．次に得られたシステムパラメータから追跡ルールを生成し，それを元に追跡・捕捉・メッセージの待機などの動作を行う．他のエージェントから新規エージェントの追加や既存のエージェントの障害情報を受け取った場合は，システムパラメータを更新し追跡ルールを再構成する．以下では観測エージェントの各動作モジュールについて述べる．



a) 構造物の配置　　b) カメラ視野のマッピング

タイプ２）通過不能だが背後にあるものを隠蔽しないオブジェクト.

またこれらのオブジェクト情報のほかに,

タイプ３）通過可能だが隠蔽されている領域.

も設定しておく．これは扉によって囲まれた領域などをあらわす（図３－６ａ）．一方システムパラメータには全てのステーションのカメラパラメータと画像サイズが記述されているので，ここから各ステーションの観測領域を計算する．次に，各ステーションの観測領域をビットマップ表現された環境マップ上に投影する（図３－６ｂ）．この際，「ステーションのレンズ主点との間にタイプ１のオブジェクトが存在する領域」は，そのステーションから不可視とする．この情報を用いて，環境マップを３種に分割する（図３－６ｃ）．

タイプＡ）単独のステーションによる観測可能領域

タイプＢ）複数のステーションによる観測可能領域

タイプＣ）不可視領域

このように観測視野がマッピングされ分割された環境マップを追跡ルールとして記録する．これを用いることで「観測領域が重なるステーション」,「不可視領域を挟んで隣接しているステーション」,「タイプＣと接しているタイプＡ領域の境界領域」を得ることができる（図３－６ｄ）．この追跡ルールによって，ステーションの行動決定時に必要となる他のステーションとの関係（観測領域の重なりや隣接）や不可視領域との境界，人物の存在領域検出などの情報獲得が可能になる．

（ｄ）タスク指示モジュール (Task Allocator)

観測エージェントの追跡状態変数と他のエージェントからのメッセージを元に追跡・捕捉タスクを決定し，他の観測エージェントにタスク制御メッセージを送る．その構造は図３－７に示すように，追跡状態変数と他のエージェントからのタスク制御メッセージを遷移条件として内部状態を変化させる状態遷移器である．以下ではタスク指示モジュールがとる状態について説明する．

（１）IDLING 状態

他のエージェントから「捕捉」あるいは「境界捕捉」のタスク制御メッセージを受け取るとACQUISITION 状態あるいは ACQUISITION ON BORDER 状態に移行する．

（２）ACQUISITION 状態

指定された座標において人物の捕捉動作を行う．捕捉に成功すれば，タスク制御メッセージを割り当てたエージェントに成功メッセージを送り TRACKING 状態に移る．

（３）ACQUISITION ON BORDER 状態

不可視領域との境界部分で捕捉動作を行う．複数の不可視領域と接している場合は，タスク制御メッセージの送信元エージェントと接する不可視領域との境界を選択する．人物が出現し

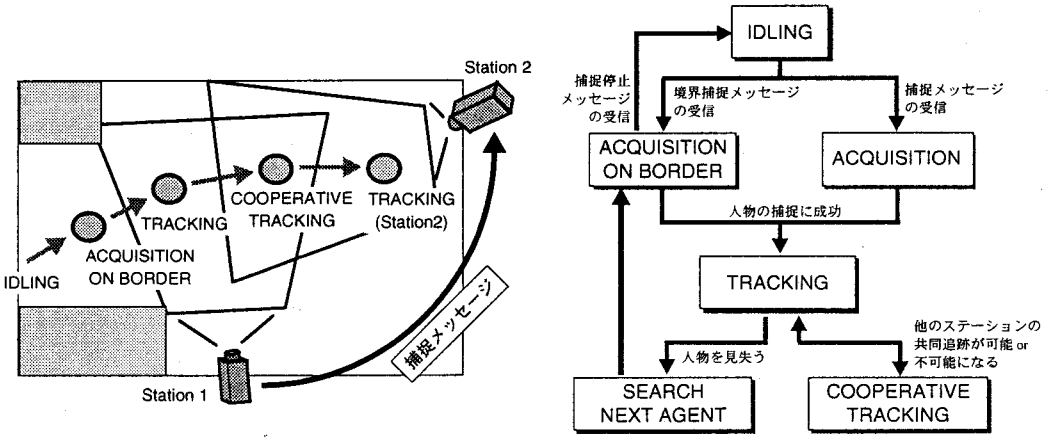


図３－７　人物の移動（左）とタスク指示モジュールの状態遷移（右）

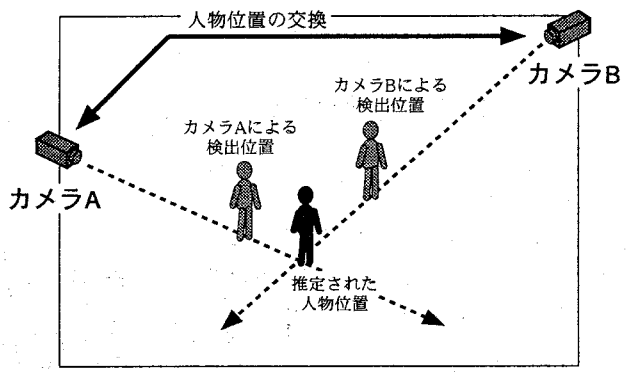


図３－８　複数視点からの観測結果の統合

捕捉に成功した場合は，不可視領域に接したすべてのエージェントに捕捉停止メッセージを送りTRACKING状態に移る．他のエージェントから捕捉停止メッセージを受け取った場合はIDLING状態に戻る．

（４）TRACKING状態

単一の観測エージェントで追跡を行う．検出された人物位置が他の観測エージェントとの共同観測領域に入れば，そのエージェントに捕捉動作のタスク制御メッセージを送る．そのエージェントが人物の捕捉に成功し，成功メッセージを返した場合にはCOOPERATIVE TRACKING状態に移る．また，人物の追跡に失敗した場合はSEARCH NEXT AGENT状態に移行する．

（５）COOPERATIVE TRACKING状態

複数の観測ステーションで同時に人物を追跡する．ステーションは互いの検出結果を交換し，検出結果を統合して人物位置の推定を行う．ここでは，各観測ステーションのカメラ位置（レンズ主点）と，その観測ステーションで得られた人物位置を結ぶ直線の交差点を推定人物位置とする（図３－８）．これはTrackerで行われる人物追跡手法ではカメラの視線方向（奥行き方向）への検出誤差が大きいため，この影響を受けにくくするためである．相手のステーション

から一定時間検出結果を受け取らなかった場合，そのステーションは追跡不能になったと判断し TRACKING 状態に戻る．

（６）SEARCH NEXT AGENT 状態

人物が不可視領域に進入した場合の処理を行う．観測エージェントは追跡ルールを用いて人物が移動したと考えられる不可視領域を求め，その領域に隣接するステーションに境界捕捉メッセージを送る．

3.3.2 システムパラメータ管理エージェント (System Parameter Management Agent:SPMA)

ＳＰＭＡはシステムパラメータを保持・更新し，観測エージェントからの要求に応じてシステムパラメータを与える．ネットワーク上に複数のＳＰＭＡが存在する場合，各ＳＰＭＡが持つシステムパラメータは同一になるように保たれる．ＳＰＭＡの構成は図３－９に示すように，システムパラメータ，情報管理モジュール，情報通信モジュール，情報同期モジュールからなる．情報管理モジュールは観測エージェントにあるものと同じで，システムパラメータを追加・更新する役割を果たす．情報通信モジュールは，観測エージェントからの要求に応じて自らが保持するシステムパラメータを送信する．情報同期モジュールは，他のシステムパラメータ管理エージェントと一定時間ごとに通信し，互いの情報が同じになるように保つ．

ＳＰＭＡは，１つのシステムに対して複数動作させることができる．これは耐障害性の点で有利なだけでなく，通信の障害などで「一部のＳＰＭＡには登録されているが，他のＳＰＭＡには登録されていない観測ステーションが存在する」という状況が生じても互いの情報を参照しあうことで防ぐことができる．

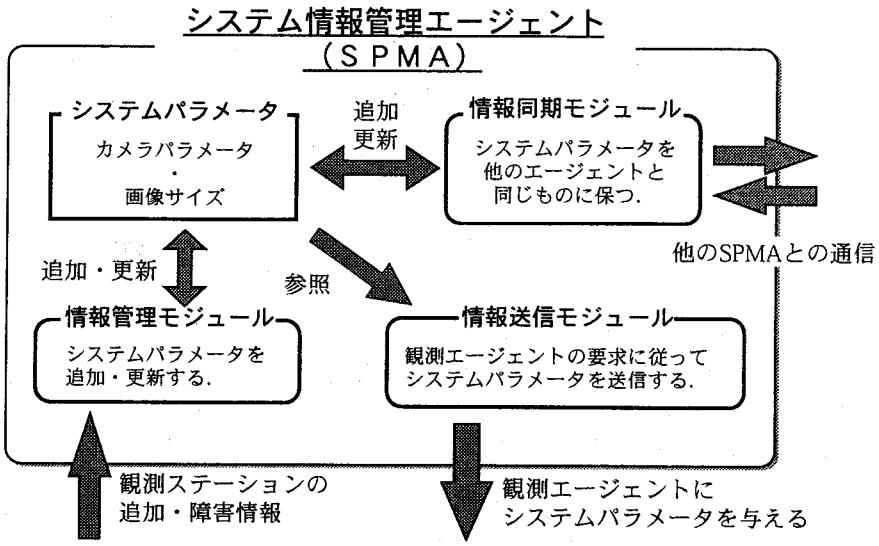


図３－９　システム情報管理エージェント（ＳＰＭＡ）の構成

3.4 実験

本手法の有効性を確認するため，複数の観測ステーションを用いて人物の追跡実験を行った．

3.4.1 実験環境

実験環境およびカメラ配置，その観測視野を図３－10に示す．水平床面上の適当な場所に世界座標系原点（Ｚ軸が鉛直方向上向き）を置き，人物を斜め上方から見下ろす配置で４台のカメラを設置した．部屋内はカメラ１，２の２台のカメラで観測され，通路も同様にカメラ３，４によって観測される．部屋内を観測する２台のカメラの観測領域は重複しているが，通路を観測するカメラの観測領域は離れている．また部屋と通路の間には２つの扉で挟まれた不可視領域が存在している．カメラ１～４にはそれぞれ，SGI-Octane(MIPS R10000-175MHz), O2(R10000 175 MHz), Indigo2 (R4400-250MHz) ,Indy (R4400-133MHz)が接続され，映像は縦横各120x160 Pixelの大きさで取り込まれ処理される．また，システム管理エージェントは他のワークステーション(Indy R4400-100MHz)上で動作させておく．

3.4.2 位置軌跡の評価実験（実験１）

部屋内を観測する２台の観測ステーション１，２によって２人の人物を同時に追跡し，検出軌跡を評価する．２人物の通過軌跡を図３－11に示す．人物通過範囲は両カメラが共に観測で

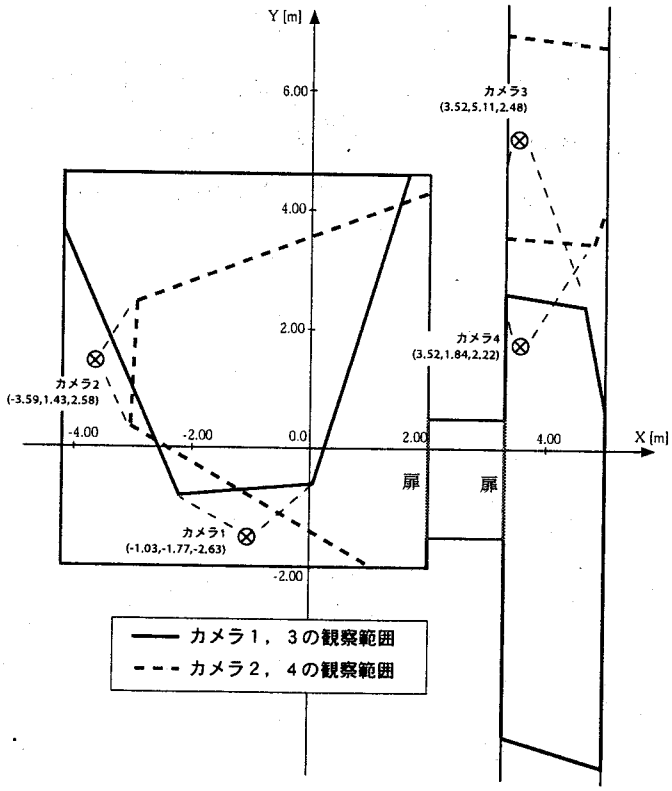


図３－10　４台のカメラ配置

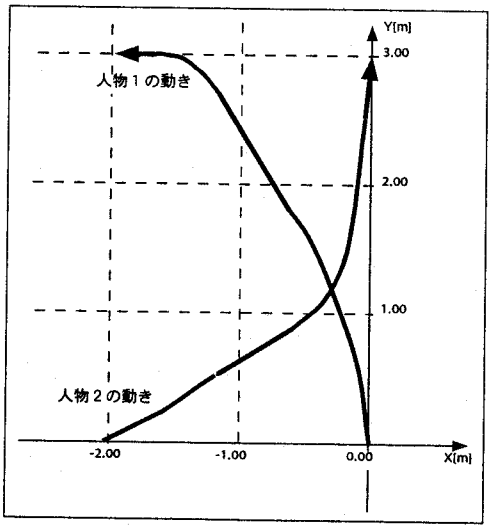
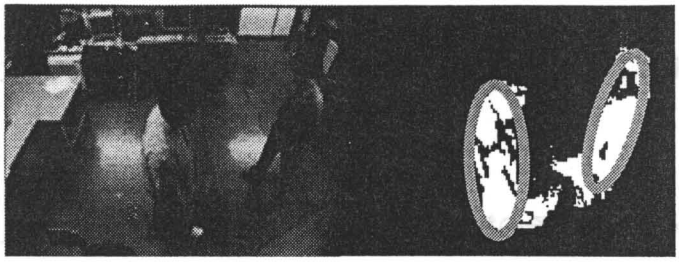


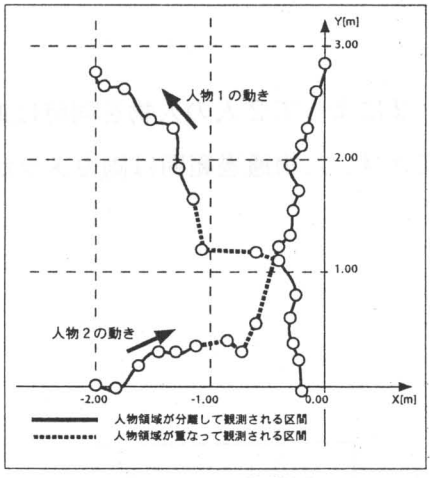
図３－11　２人物の通過軌跡（実験１）



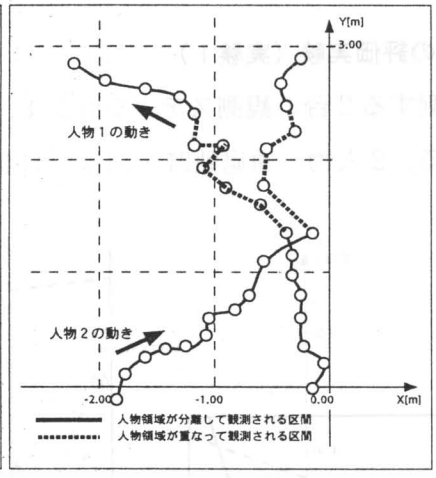
(a) 人物領域が分離している状態

(b) 人物領域が重なっている状態

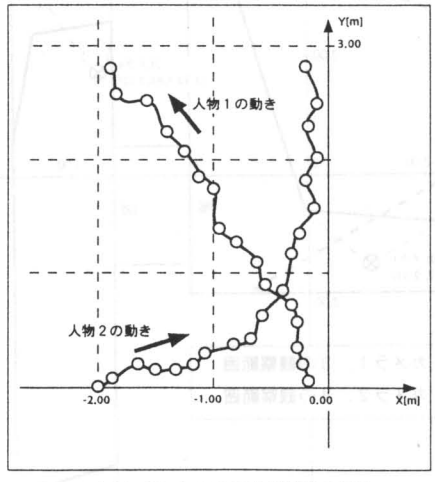
図３－12　実験中の画像（実験１）



(a)カメラ１による計測結果



(b)カメラ２による計測結果



(c)カメラ１＆２の計測結果の統合

図３－13　実験結果（実験１）

きる領域にあり，追跡中のすべてのシーンにおいて２台のカメラが追跡しつづけることが可能だった．実験はカメラ１，２の単独追跡およびカメラ１，２の共同追跡の３種で行った．実験においては，あらかじめカメラ１，２で撮影される画像をビデオディスクに保存しておき，それらをリアルタイムに再生することで同じ入力画像を用いた３種の計測を行うことができる．

実験中の映像を図３－12に，各実験で計測された通過軌跡を図３－13に示す．観測ステーション１，２の単独追跡では，画面上の人物領域が重なる状態（図３－12（ａ））において計測誤差が大きくなっていることが確認できる．一方，観測ステーション１と２による同時追跡を行い検出結果を統合した結果では，同様の状況でも安定した通過軌跡が得られている．これは，一方のステーション上の画像で人物領域が重なっても，他方のカメラで分離して観測されているため，両者の検出結果を統合することでより安定した位置検出結果が可能になったものと考えられる．

3.4.3 エージェントの動作状態変化の観察（実験２）

４台の観測ステーションが観測する環境中を２人物が歩き回るタスクを行い，その計測結果および観測エージェントの動作状態を観察する．人物の通過軌跡を図３－14に示す．人物１，２はそれぞれ通路の両端からカメラ３，４の観測領域へ進入し，順次扉を通って部屋内に入り歩きまわる．この状況に対して生成された追跡ルールを図３－15に示す．カメラ１とカメラ３，カメラ３とカメラ４の間に不可視領域が存在し，境界領域が設定されていることが確認できる．実験に際しては，カメラ３，４に対して初期的に通路の両端で境界捕捉動作を行うように設定しておく．４台の観測ステーションから得られた人物位置の検出結果を図３－16に，人物１に対する観測ステーションの状態変化を図３－17に示す．これらからわかるように，通路での不可視領域や部屋と通路間の不可視領域の境界部分で捕捉領域が設定されており，妥当な追跡ルールが生成されている．また実際の実験でも，扉部分での捕捉動作によって，不可視領域から出現する人物を捕捉・追跡することが確認できた．

3.4.4 一部のエージェントの追加・動作停止に対する実験（実験３）

ステーションの追加や動作停止といったイベントを意図的に発生させ，システムの耐障害性を評価する．実験環境やカメラ配置，人物の移動などは前節と同様である．まず実験開始時には観測ステーション１，３，４を動作させておき，人物１が部屋内に進入するのを待つ．この際３台のステーションが生成した追跡ルールは図３－18のようになった．人物が部屋内に進入しステーション１が追跡を開始すると，ステーション２を起動させ人物２の入室を監視する．このときすでに起動していたステーション１，３，４はステーション２の起動を検知し，追跡ルールの再構築を行った．この際生成された追跡ルールは前節の実験と同様（図３－15）であった．

最後にステーション２を動作停止させ，ステーション１単独で追跡を行わせる．この際生成された追跡ルールは再び図３－18のようになった．４台のステーションの状態変化を図３－19に示す．実験中には，ステーション２の起動および動作停止というイベントが生じたが，それに応じて各ステーションが追跡ルールを再構築し，適切なプランを実行しているのが確認できた．特に通路と部屋間の不可視領域から部屋内に人物が出現する際，ステーション２が動作している際にはステーション２が扉付近で捕捉動作を行ったが，ステーション２が動作していない場

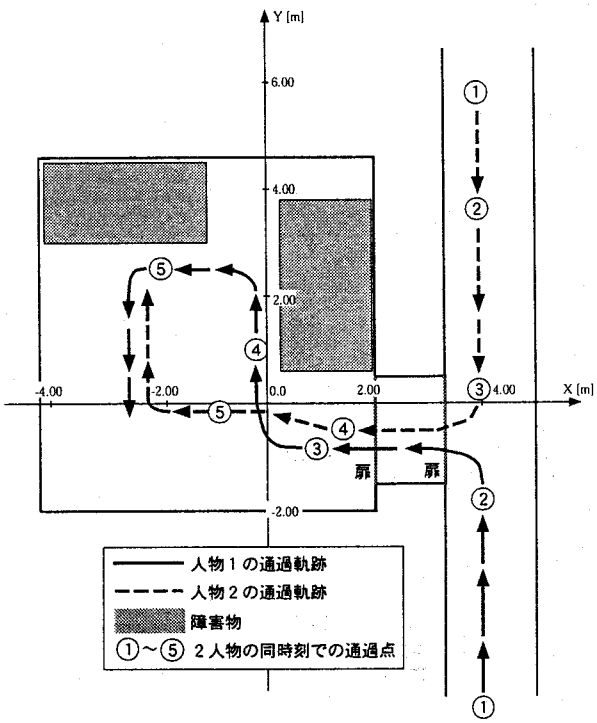


図３－14　人物の移動軌跡（実験２）

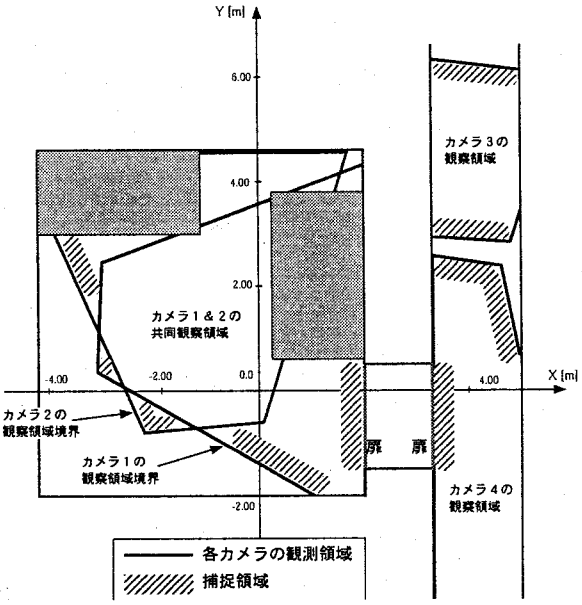


図３－15　生成された追跡ルール

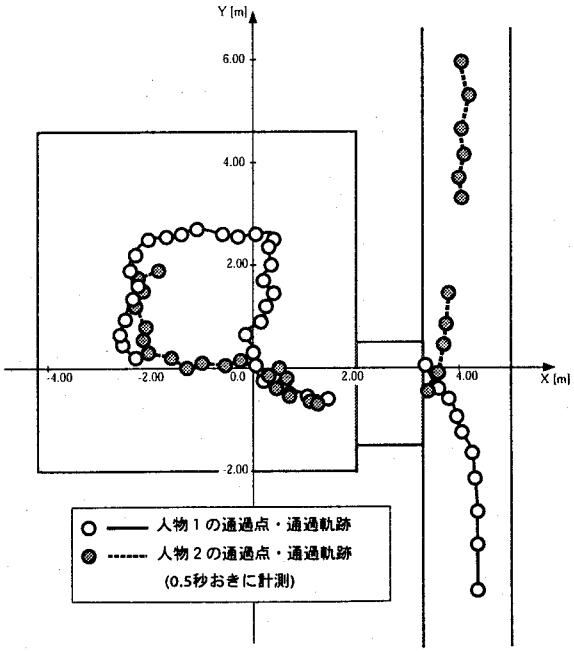


図３－16　人物の移動軌跡の検出結果（実験２）

合にはステーション１が捕捉動作を行う，といった役割分担が動的に変更されることが確認できた．

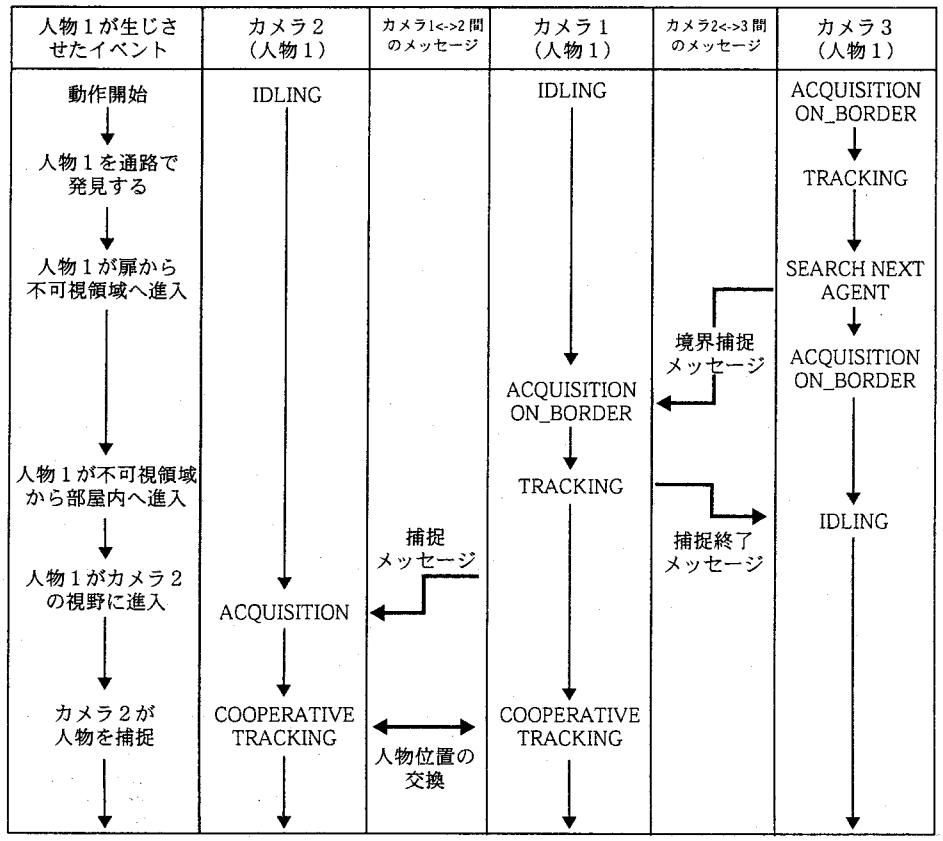


図３－17　観測ステーションの状態変化（実験２）

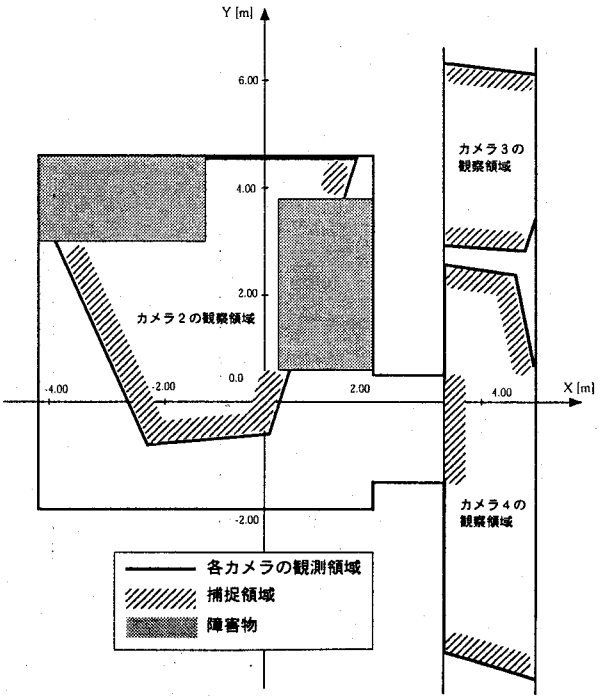


図３－18　初期的に３台のステーションで生成される追跡ルール（実験３）

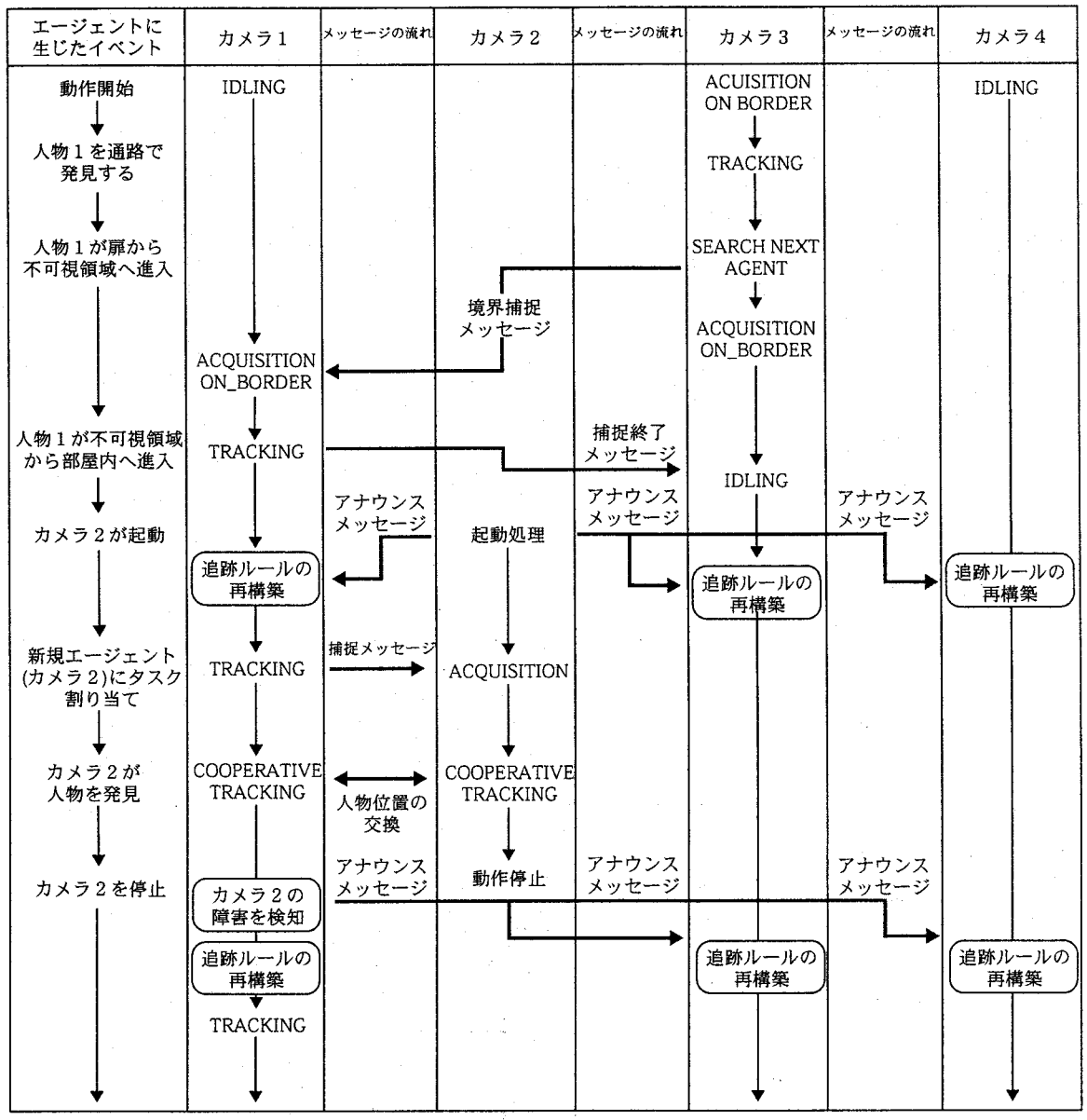


図３－19　ステーションに障害が生じた場合の状態変化（実験３）

3.5 検討

（１）システムの性能

想定された環境において，本手法が有効に機能することが確認できた．２人物の歩行軌跡を評価する実験では，人物を１台の観測ステーションで追跡・位置検出した結果と，２台の観測ステーションの検出結果を統合して位置推定した結果を比較した．１台の観測ステーションのみでは，画面上での人物領域が交差したときの計測誤差が大きくなったが，２台の観測ステーションを用いると安定した歩行軌跡が得られることが確認できた．第２の実験では不可視領域を挟んだ広域環境で，複数のステーションが協調して人物追跡を行うことが確認できた．ステーションの追加や障害に対する第３の実験では，小規模なシステムではあるがステーションの追

加や障害を自動的に検知し追跡プランを更新する仕組みによって，人物を連続的に追跡し続けることが確認できた．処理スピードに関しては，人物追跡処理がもっとも大きな計算処理を必要とし，協調処理部分はそれほど大きな計算量は必要としない．現在の人物追跡手法では，１台のステーションで１人を追跡した場合 30 frame/sec，２人の場合でも 15 frame/sec ほどの処理速度で行われており，協調処理を加えても実時間処理が可能であった．また今回の実験では，機材や開発環境の関係からワークステーションを用いたが，汎用ＰＣを用いても同等以上の性能が得られると考えられる．

（２）エージェント間の複数人物の誤対応

実験１において２人の人物が画面上で交差した場合に「人物モデルと追跡人物の入れ換わり」という現象が生じ，２台の観測ステーション間で２人物の対応付けを誤ってしまうことがあった．これにより統合結果自体も誤りが大きくなることが観測された．これは本システムにおける追跡手法に依存する問題ではあるが，今後カメラ間の人物識別機構を導入することで解決していくことが考えられる．

（３）観測の非同期性によって生じる問題

観測ステーションでの観測は非同期で行われているため，ステーション間の観測時刻の相違が観測結果の統合時に問題となる場合がある．

観測時刻の相違は１）画像取得・処理タイミングの相違，２）通信による遅延，の２種類の要因が考えられる．このうち１）に関しては，本システムがLAN上に構築されているということもあり無視できるが，２）に関してはステーション間に最大で追跡１フレームの処理時間分の時刻差が生じることになる．画像処理モジュールの性能は１フレーム間に最大150mm移動する場合まで追跡可能なため，これによって生じるステーション間の計測値の差異は最大150mmとなる．この差異が統合結果に及ぼす影響は，カメラ配置や人物の移動方向に左右されるが，状況によっては１台の計測よりも誤差が大きくなる可能性は否定できない．この問題に対しては，各ステーションから送られる情報に観測時刻や人物の移動速度・移動方向の情報を付け加えることで，人物位置を予測し統合を行うことなどの解決方法が考えられる．

（４）追跡ルールの妥当性

本システムでは，環境をステーションの観測領域に応じて分割し，人物は隣接した分割領域間を移動すると仮定されている．しかし，分割された各領域が画像処理モジュールで１フレーム間に計測できる人物移動距離(150mm)よりも小さい場合，人物がその領域を飛び越して，離れた領域に移動してしまうことが考えられる．この場合先に述べた仮定が成り立たないことになり，システムが正常動作できない．つまり，システム構築時のカメラ配置の段階から，分割された各領域の大きさが十分大きくなるように注意する必要がある．

また，今回の実験システムではステーションの台数が4台と少なかったためルール生成に要する計算量は少なかったが，より多くのステーションを用いた大規模システムではルール生成に要する計算量も多くなり，実時間処理が困難になる可能性がある．この問題に対しては「明らかに観測領域が離れてしまうステーション」の情報は用いないなどの対処方法が考えられる．

3.6 本章のまとめ

本章では，分散観測システムによって広域環境を観測する手法を開発した．システムは，自律的に動作する複数の観測ステーションによって構成され，広域を通過する人物を連続的に追跡し，その軌跡を得ることが可能である．

観測ステーション間の協調手法としては，環境地図とカメラ視野から追跡プランを生成し，地形的に隣接するステーションの通信によってタスク制御を行う手法を提案した．自律的に動作するステーションの組み合わせによって動作するため，一部のステーションの障害や新規ステーションの追加に柔軟に対処されることが確認された．また，通信する範囲は地形的に隣接するステーション間に絞られるため，使用する通信帯域も少なくすることができる．一方で，動作開始時に必要となるシステム全体の情報獲得を実現するために，システム情報を管理するデータベースを用いて観測ステーションに配信する機構を用いた．結果として，集中制御型の情報共有機構を有した自律型分散観測システムの構築を行うことができたと考える．

問題点としては，複数の対象の同時追跡というタスクに対して，ステーション間での対象の誤対応が生じるケースが見られた．この問題に対しては，次章での人物識別手法や，次々章で述べるステーションのソフトウェア構成の改良によって解決を試みる．

４章 人物像の色分布を用いた人物識別

前章までにのべた手法で，単一人物の複数カメラ間にわたる連続追跡が可能であることを示した．また，各ステーションでは複数の人物を同時追跡することが可能なため，機構的には複数の人物の同時追跡が可能となっている．しかし前章のまとめでも触れたように，実環境の使用においては信頼性の上でいくつかの問題点がある．主なものとして挙げられるのが，複数の観測ステーションで同時に複数人物を追跡した場合，その間で同一人物と対応づけられた人物モデルが実際には異なる人物であるという「人物の誤対応」問題である．本章では，この問題を解決するための「人物像の色分布を用いた人物識別手法」について述べる．以下ではまずこの手法の適用条件と役割について述べ，概要の説明および実験結果を示す．

4.1 人物の誤対応問題

前章までに提案したシステムで，観測環境中の全人物の位置を知ることが可能である．しかし実験を行う中で，以下のような状況で人物の誤対応問題が生じることが観測された．

（１）単一カメラ上での人物領域交差（図４－１）

各エージェントで行われる楕円体モデルに基づいた追跡手法は，複数人物の同時追跡が可能である．これは２章で述べたように，複数の楕円体人物モデルが同時に投影されたテンプレート画像を用意し，差分画像とマッチングすることで実現している．しかし複数の人物が非常に接近し互いの人物領域が重なった場合「モデルと人物が入れ替わる」現象が確認される．これにより観測ステーション間において人物の誤対応が生じることになる．

（２）観測ステーション間の不可視領域（図４－２）

３章の複数ステーション間の協調手法では，観測ステーションの観測領域間に不可視領域が存在することを想定している．複数の人物が不可視領域に進入すると，その近接領域を観測するエージェントが境界部分で「捕捉動作」を行うが，複数の人物が不可視領域に進入すると，出現人物がどの人物であるかを特定することは不可能であり，１と同様の問題が生じることになる．

これらの問題を解決するには各ステーションで行われる人物追跡手法を拡張し，ステーション内およびステーション間での人物識別機構を備えることが考えられる．ここでは，以下の２種類の条件下での人物識別機構が必要になる．

第１は単一カメラの視野中での人物の学習および識別である．これは（１）の問題を解決するためのもので，同一画像中で複数の人物が交差する問題を扱う．つまり，単一ステーション

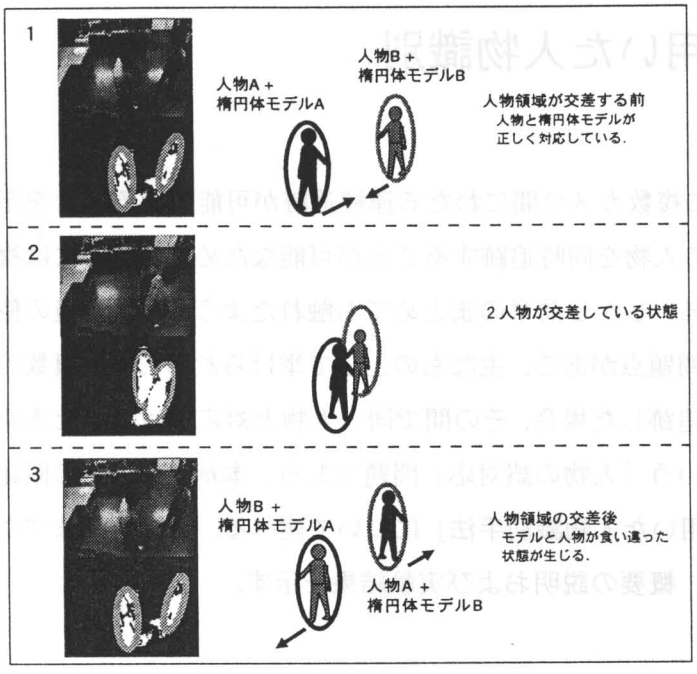


図４－１　人物の画像領域の交差によって生じる人物と人物モデルの誤対応

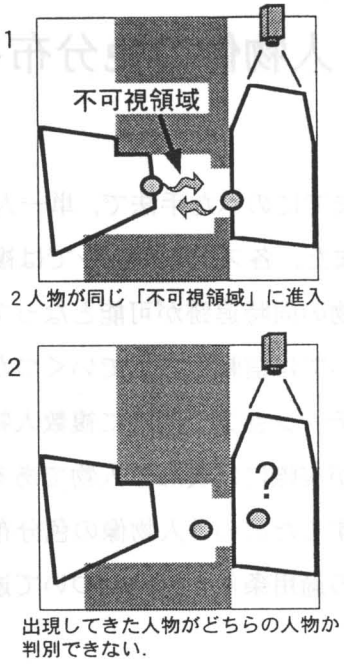


図４－２　不可視領域による人物と人物モデルの誤対応

の取得画像内で人物領域が交差する前後の人物の対応付けを行う．すべての人物情報の学習・識別が同一のカメラ内で行われるため，カメラの特性や環境照明などの撮影条件が同等であると考えられる．

第２は複数カメラ間での人物の学習および識別問題である．これは（２）の問題を解決するためのものである．この場合，学習を行うカメラと識別を行うカメラが異なる可能性があるだけでなく，学習に使われたカメラが人物に応じて異なるケースや，識別の行われるカメラが人物に応じて異なるといったケースがありうる．また複数のステーション間で協調して識別を行う手法を開発する必要がある．

本章では基本的に（１）の問題を取り扱うものとするが，人物像の色を用いるという点では次章で取り扱う（２）の解決手法の一部ともなっている．

4.2 従来研究と本システムで要求される条件

人物識別の研究は従来から多くなされているが，多くはセキュリティ用途などを目的としたもので，顔画像照合による開錠システムなどあらかじめ登録された人物中に対象人物が存在するかを判断するタスクを目的としたものが多い．そのため画像の撮影条件を統制できる，人物画像の取得・学習がすでに行われている，認識時間に比較的余裕がある等の条件が与えられている[Shioya92]．

一方，動画像からリアルタイムで人物の追跡・識別を行った例として，パン・チルト・ズー

ム可能なカメラを用い，肌色情報によって顔領域をトラッキングしながら顔画像を学習し，Eigen-Face法[Turk91]に基づいた個人識別を行った[Shakunaga98]などがある．ただし識別を行う条件としては，あらかじめ拘束された環境条件下で対象人物の顔画像の学習が行われていることを前提としている．また[Heisele98]では，歩行する人物像の足の動きをＴＤＮＮ（Time Delay Neural Networks）を用いて学習・個人識別する手法を提案している．これも動画像からの対象認識を目指したものであるが，学習があらかじめ行われていることや，学習自体がリアルタイムで行われないなどの制約がある．また，動画像（顔画像，歩行画像特徴）と音声により個人識別を行った研究[Maeda96]もあるが，複数の特徴量を用いるために新たなセンサーを設置しなければならない．これらいずれの手法においても，顔画像の取得や歩行画像の取得のためにカメラ設置に対する制約を受ける．

本章で行おうとしている人物識別は，一つのカメラ視野内で数人の画像領域が交差した場合での，交差前後の人物対応付けを目的とする．この場合，対象となる人物（クラス）数は少ないため識別問題としては単純であるが，人物とカメラの位置関係は拘束できず，顔のように人物の特定部位の画像を用いた識別を行うことは困難である．また想定する状況としては，観測ステーションの視野に次々と進入する不特定多数の人物を対象とすることが考えられるため，人物に対する事前の学習データを使用することはできない．すなわち，追跡を行いながらリアルタイムで学習・識別を行える程度の計算負荷の少ない手法を開発する必要がある．

4.3 人物像の色分布を用いた識別

人間は，多人数が混在する複雑なシーン下で複数の人物がすれ違った場合でも，対象となる人物を追跡し注視し続ける事ができる．このような事が可能なのは，人が観測対象となる人物の特徴を把握し，それをもとに識別をおこなっているからだと考えられる．ここで用いられる特徴量には，身長や体型，人物の動き，動作のくせ等様々なものが考えられるが，本システムでは画像処理によって最も取得しやすい特徴量のひとつである「対象人物の着衣の色分布」を用いて識別を行うことを試みる．

同様の提案は[Konparu98]でもみられる．ここでは全方位視覚センサを用いて全方位の画像を取得し，特定の人物を発見・注視するタスクを行っている．まず初期学習として，背景差分により人物領域を取り出しそのＲＧＢヒストグラムを取得する．次に，あらかじめ取得していた背景画像と現在の取得画像をＲＧＢ空間中で比較する．背景と異なるＲＧＢ値をとる領域を人物領域とし，領域中のＲＧＢヒストグラムと学習した人物領域ヒストグラムを比較することで，人物の識別を行う．この手法では，全方位画像から得られた人物領域から，ＲＧＢヒストグラムを用いて人物識別する方法が提案され，実現されている．一方で，この手法では人物像の色情報を利用しているが，その空間的な組み合わせに関する情報を有効に利用しているとはいえ

ない．例えばこの方法では「上半身：青，下半身：赤」という特性をもつ人物と，「上半身：赤，下半身：青」という特性をもつ人物の違いを見分けることは困難であると考えられる．また，初期学習に用いる人物像は１つの方向から見た像のみであるため，着衣の色分布が見る方向により変化するものであった場合に識別できないものと考えられる．

本手法で目指しているのは，人物の着衣色のさまざまな組み合わせを学習し，識別に用いる事である．たとえば複数の人物が上半身に同色の着衣をしている場合でも，他の部分（下半身）の色を評価すれば識別は可能である．また，同一の人物の全身像を複数フレームにわたって長時間蓄積し学習することで，多方向から見た人物像の色分布を取得し識別に用いることができる．つまり，着衣色を用いるとともにその空間的な分布も識別に用いることを目指している．

この識別手法は，２章で述べた楕円体人物モデルによる追跡手法と組み合わせることで実現できる．２章の手法では楕円体が人物の立像と対応すると想定し，差分画像から人物領域の検出を行っている．この人物領域を床面に水平な平面で分割する処理を行えば，分割された各領域は人物の上半身・下半身等の部位に相当するものと考えられる（図４－３）．ここで，分割された各領域から色情報を抽出して人物像の着衣の色情報とし，追跡する全てのフレームにわたり記録する．

学習段階で問題となるのがそのデータ量である．動画像からの画像特徴量をすべて学習しようとするとそのデータ量は膨大となり，記録・処理に必要とされるメモリや計算量が多大になる．そこで，人物領域から得る色情報は分割された各領域における代表色のみとする．また，記録段階においては「新たに観測された色を，特徴ベクトルの新たな次元として動的追加する」方法によって記録容量の削減をはかる．これにより，特徴・識別空間の次元は「対象とする人物像中で出現する代表色の数」程度に押さえられる．

本手法で行われる学習および識別は，特徴ベクトル（人物像の色分布）とそのクラス（人物）が既知である「教師付き学習[Russell95]」にあたる．そこで，教師付きデータによる識別方法である線形判別分析[Nagao82][Kitagawa66]とベイズ識別法[Nagao82]の組み合わせによる手法

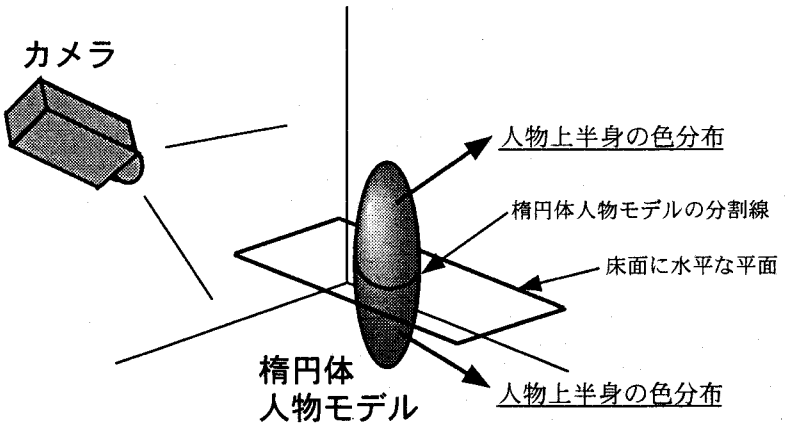


図４－３　楕円体モデルの分割と色分布の抽出

[Kanehara95]を用いて対象の識別を行う．ベイズ識別法という確率的な手法を用いるのは，影や反射等によって追跡が失敗した場合，学習データである人物領域中に，影領域や背景領域などのノイズが含まれる可能性があるからである．このようなノイズは一時的なものであるため，仮にノイズを含むデータが学習に使われたとしても，長期的な観測と確率的識別手法によって影響が避けられるものと考えられる．

4.4 処理の流れ

図４－４に処理の概要を示す．各人物に対しては，学習，識別空間生成，識別の３モードが用意される．人物領域が分離しておりモデルと人物の対応付けが正しく行われている間は「学習モード」となり，人物領域の代表色の色分布を学習する（図４－４（１））．人物領域の交差がおこると「識別空間生成モード」に移り，学習結果を線形判別分析することで判別関数を得る（図４－４（２））．再び人物領域が分離すると「識別モード」となり，判別関数を用いて分離した各々の人物の識別を行う（図４－４（３））．以下では，各モードに対して説明する．

4.4.1 学習モード

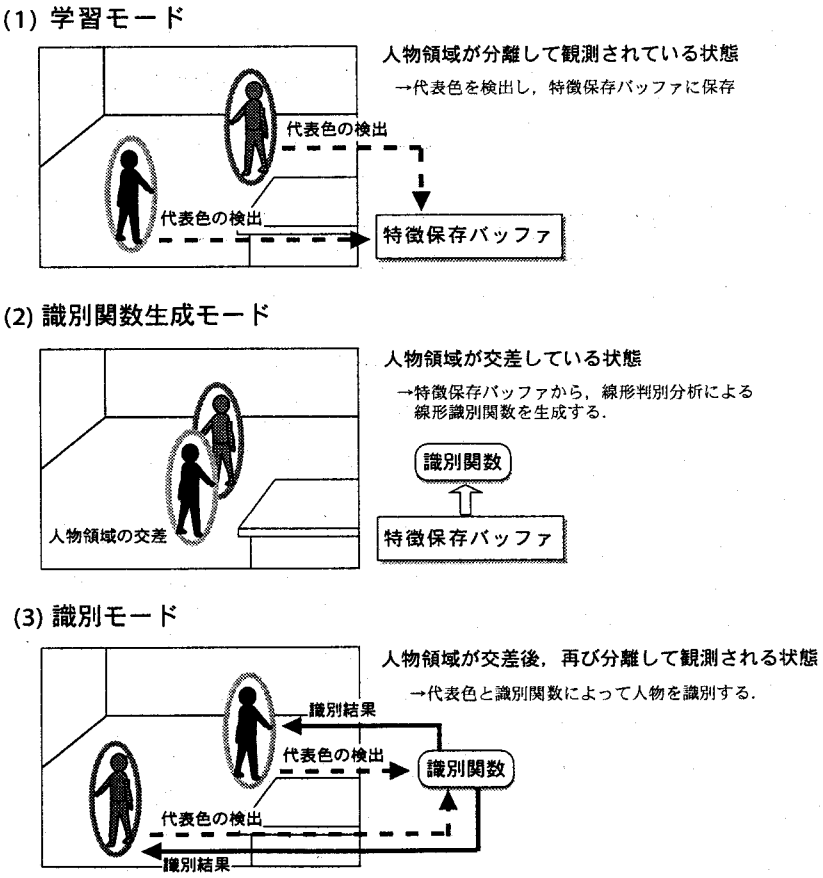


図４－４　人物識別処理の流れ

人物領域を分割し色分布を学習する．この処理は，

（１）人物領域の分割

（２）分割領域ごとのＲＧＢ代表色の取得

（３）学習空間の生成および更新

（４）学習空間上での色分布情報の保存

の４ステップから構成される．

（１）人物領域の分割

画像中での人物領域を，床面からの高さによっていくつかの部分に分割する．２章で述べた楕円体モデルによる追跡・位置検出によって，画像上の人物領域，人物の環境座標系における位置を得ることができる．また楕円体人物モデルの環境座標系における幅および身長は固定値である．この楕円体モデルを楕円体を床面と水平ないくつかの平面（Z＝k）で分割する．

世界座標系(X,Y,Z)における楕円体モデルの式は以下のようになる．ただし，人物追跡によって検出された人物位置を(X_p,Y_p)，人物の幅を$2R_a$，身長を$2R_b$としている．

$$\frac{(X-X_p)^2}{R_a^{\ 2}}+\frac{(Y-Y_p)^2}{R_a^{\ 2}}+\frac{(Z-R_b)^2}{R_b^{\ 2}}=1 \qquad \text{（式４－１）}$$

この式と床面に水平な平面Ｚ＝ｋとのなす図形は，以下の楕円の方程式になる．

$$(X-X_0)^2+(Y-Y_0)^2=R_w^{\ 2}R_h^{\ 2}\{1-\frac{(k-R_h)^2}{R_h^{\ 2}}\} \qquad \text{（式４－２）}$$

また，画像座標系(x_c,y_c)と環境座標系の変換をあらわす透視変換の式は前節の式１のように表せるが，これを正方行列化し逆行列をとることで，以下の式を得る．

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}=\begin{bmatrix} C_{11} & C_{12} & C_{13} & C_{14} \\ C_{21} & C_{22} & C_{23} & C_{24} \\ C_{31} & C_{32} & C_{33} & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}^{-1}\begin{bmatrix} hx_c \\ hy_c \\ h \\ 1 \end{bmatrix}$$
$$\cong\begin{bmatrix} C'_{11} & C'_{12} & C'_{13} & C'_{14} \\ C'_{21} & C'_{22} & C'_{23} & C'_{24} \\ C'_{31} & C'_{32} & C'_{33} & C'_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} hx_c \\ hy_c \\ h \\ 1 \end{bmatrix} \qquad \text{（式４－３）}$$

これを式４－２に代入すると以下の式を得る．

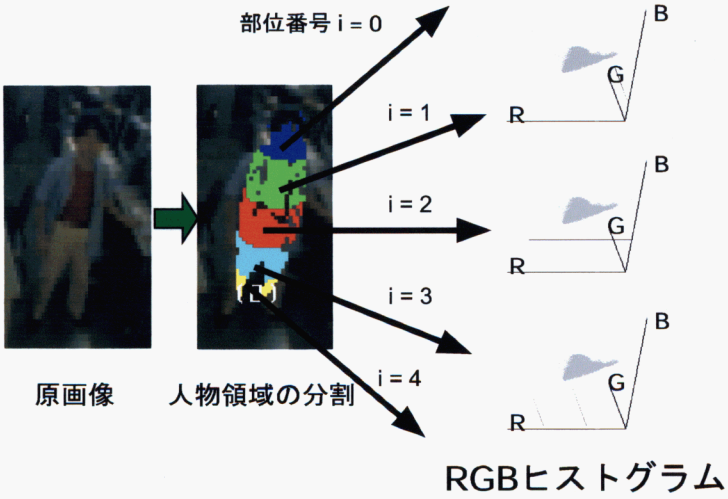


図４－５　人物領域の分割とＲＧＢヒストグラムの作成

$$\{h(C_{11}'x_c + C_{12}'y_c + C_{13}') + C_{14}' - X_0\}^2$$
$$+\{h(C_{21}'x_c + C_{12}'x_c + C_{13}') + C_{24}' - Y_0\}^2 = R_w^2 R_h^2 \{1 - \frac{(k - R_h)^2}{R_h^2}\} \qquad \text{（式４－４）}$$
$$h = \frac{k - C_{34}'}{C_{31}'x_c + C_{32}'y_c + C_{33}'}$$

この式は，式４－２で表された３次元中の楕円を式４－３で表される変換によって２次元画像に投影したものである．３次元中の楕円を２次元に投影したものなので，投影された像も楕円を描く．ここで，本システムでのカメラは上から見下ろす方向で固定されていることから，画像に投影された楕円のうち画像中の下側に見える（床面に近い）値を求める分割線とする．

この分割を，楕円体モデルを生成した人物の身長の 1/5, 2/5, 3/5, 4/5 の４つの平面で行い，楕円体を５つの領域に分割する（図４－５）．以下分割された各領域を，頭部から数えて領域番号 i = 0, 1, 2, 3, 4 として表すこととする．

（２）分割領域毎のＲＧＢ色分布の獲得

分割された各領域中の有効ピクセルのＲＧＢヒストグラムを得る（図４－５）．ここでの「有効ピクセル」とは，追跡処理での背景差分で得られたピクセルを指す．入力画像では各ピクセルのＲＧＢ値が各０～２５５までの値で表現されるが，これを各軸ごとに１７分割した座標軸上でヒストグラムを求める．すなわちこのヒストグラムは(R,G,B)=(0,0,0)～(16,16,16)の座標値をとる．次に，ヒストグラムの各値を分割領域中の有効ピクセル数で割ることで正規化する．

こうして得られたヒストグラムを「正規化ヒストグラム」と呼ぶ．このうち頭部（領域番号 i = 0）および足部（領域番号 i = 4）は，トラッキング時のずれや影などのノイズを多く含んでいるので使用しないこととし，人物の胴体部分（i = 1, 2, 3）を以降の処理に利用する．

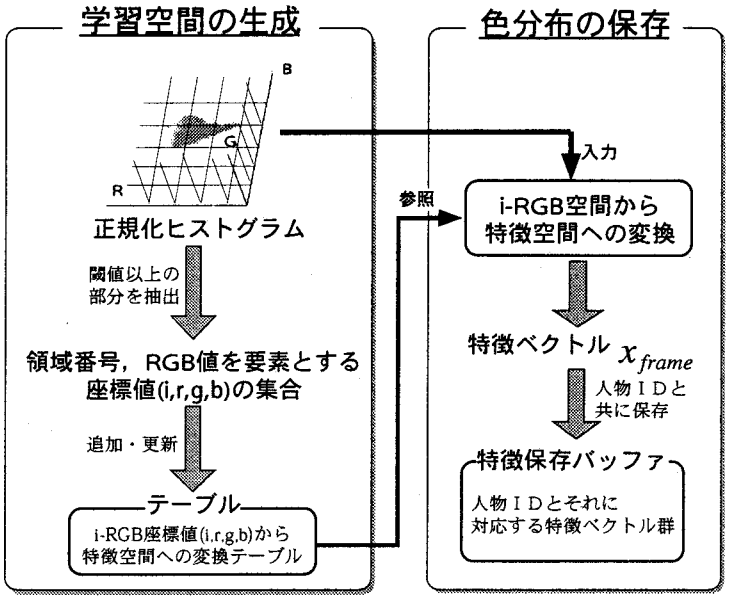


図４－６　学習空間の生成と特徴ベクトルの保存

（３）学習空間の動的生成

以上の処理で各分割領域における色分布のＲＧＢヒストグラムを求める事ができるが，一般に人物像の色分布はＲＧＢ空間中の一部分にしか分布しないため，このヒストグラムをそのまま保存するのは無駄が多い．そこで，フレーム毎に順次得られる正規化ヒストグラムから動的に学習空間を生成することで，効率の良い代表色保存方法を考えた．その流れを図４－６に示す．観測ステーション内の人物識別部には人物の特徴量を学習するための「特徴保存バッファ」と共に，領域番号とＲＧＢ値からなる座標値(i,r,g,b)から特徴ベクトルの次元番号（要素番号）へ変換する対応表を記録するテーブルが保持されている．このテーブルは，全フレーム・全追跡中人物を通して共通のものが使われる．

あるフレームの画像から正規化ヒストグラムが得られると，その中から閾値以上の値を持つ座標を選び出す．次に，この座標値(i,r,g,b)に対応する番号をテーブルから探し出す．テーブルにこの座標値に対応する変換が登録されていなかった場合，新たに追加する処理を行う．この処理を正規化ヒストグラム中でしきい値以上の値をもつ全座標について行い，そのフレームに対応する特徴ベクトル x_{frame} を得る．

（４）色分布の投影と保存

得られた x_{frame} を特徴保存バッファに保存する．この際，特徴ベクトルが得られた人物のＩＤも共に保存する．このＩＤがこの特徴ベクトルの属するクラスを表す．

4.4.2識別関数生成モード

複数人物が画面上で交差している間は，学習モードで蓄積された全フレームの特徴ベクトル

から識別関数を生成する処理を行う．これは以下の２つの処理で構成される．

1．線形判別分析を用い，特徴ベクトルの分布がクラス（人物ＩＤ）間で統計的に最も大きく分離する直交変換を求める．

2．各クラスに属する特徴ベクトルの分布が識別空間中で正規分布に従うことを仮定し，各クラスの平均および分散を求める．

１は線形判別を行うための基本的な処理であるが，２は後に述べる識別モードにおいて確率的な判別手法（ベイズ識別法）を用いるために行う．

（１）線形判別分析[Aoki99][Nagao83][Kitagawa66]

特徴保存バッファには，人物毎の色分布を表す特徴ベクトル集合 $\vec{x}_{in_i}$ が保存されている．ただし，ｉは人物ＩＤ，n_i はクラスｉに属する特徴ベクトルの数（学習サンプル数）である．線形判別分析はこれらのデータから，クラス（人物ＩＤ）内特徴ベクトルの分散が小さく，クラス間の分散が大きくなるような変換行列を求める方法である．２次元平面上で模式的に示すと図４－７のようになり，ここで最も高い判別が行える変換ｆを求めることを目的とする．

変換ｆは特徴ベクトル $\vec{X}=(x_1,x_2,\cdots,x_n)$ から全特徴ベクトルの平均 $\overline{X}=(\bar{x}_1,\bar{x}_2,\cdots,\bar{x}_n)$ を引いたベクトル $\vec{X}-\overline{X}$ をn次元ベクトル $\vec{a}$ により射影したものと考えることができる．この射影ベクトルは，以下の処理で求めることができる[Aoki99]．

1．各クラスに属する特徴ベクトルから，クラス内共分散 Σ_w，クラス間共分散 Σ_B を求める．これは以下の式で表される．

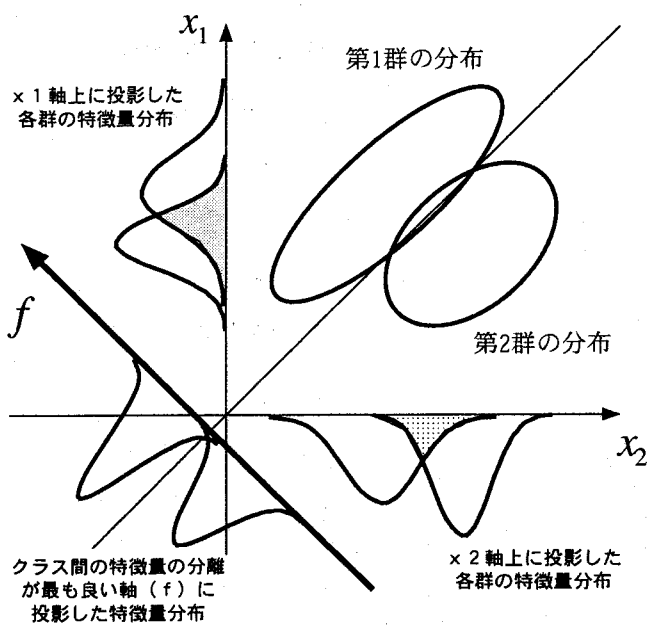


図４－７　線形判別分析の概念図

$$\Sigma_B = \sum_{k=1}^{K} \omega_k (\overline{X}_k - \overline{X})(\overline{X}_k - \overline{X})^T$$

$$\Sigma_w = \sum_{k=1}^{K} \omega_k \Sigma_k$$

$$\Sigma_k = \frac{1}{N_k} \sum_{i=1}^{N_k} (\vec{X}_{ik} - \overline{X}_k)(\vec{X}_{ik} - \overline{X}_k)^T$$

$$\omega_k = \frac{N_k}{N}$$

ただし，$\overline{X}_k$はｋ群に属する特徴ベクトルの平均であり，N_kおよびNはｋ群に属する特徴ベクトルの数および全特徴ベクトルの数を表す．

２．以下の固有値問題を満たす行列Ａを求める．

$$\Sigma_B A = \Sigma_w AL$$

$$A^T \Sigma_w A = I$$

３．行列Ａのうち，最大の固有値をとる固有ベクトルが求めるベクトル$\vec{a} = (a_1, a_2, \cdots, a_n)$となる．求める識別関数ｆとは，識別対象となる特徴ベクトルとベクトル$\vec{a}$との内積で表される．

（２）特徴分布の仮定

変換関数ｆを施した各クラスの特徴量の平均μ・分散σをもとめておく．各クラスの特徴量はこの空間において，この平均・分散をもつ正規分布に従うと仮定する．

4.4.3識別モード

識別モードでは，現在の人物領域から得られた特徴ベクトルを，先に求めた識別関数ｆによって変換し識別を行う．ここでは，確率を基にした識別手法であるベイズ識別法を用いる．

（１）ベイズ識別法[Nagao82]

ベイズ識別法とは，Ｋ種類のカテゴリからなる対象の学習サンプルの集合Ckが与えられたときに，未知の認識対象ｘが各カテゴリCkに属する条件付き確率を計算し，最大の確率値をもつカテゴリに分類する方法である．$p(C_k \mid x)$はベイズの定理から以下のように表される．

$$p(C_k \mid x) = \frac{p(C_k) p(x \mid C_k)}{p(x)} \qquad \text{（式４－５）}$$

ただし，p(x)は事象ｘの生起確率，p(Ck)はカテゴリCkに属する事象の生起確率である．p(x)は全てのｋについて共通であるため無視すると，識別は

$$\max p(C_k) p(x \mid C_k) \qquad \text{（式４－６）}$$

となるクラスＣｋを求めればよい．

（２）人物とモデルの入れ替わりに対する評価

今回の識別タスクでは画像領域の交差前・後の人物・人物モデルの対応付け確率は，全て等

しいものとする．すなわち，識別対象となるクラス数（交差が起こった人物数）をＮとすると $p(C_1)=p(C_2)=..=1/N$ の事前確率をもつ．

同様に，このタスクでは起こりうる状況がいくつかの有限の種類に絞られるため，起こりうる全ての組み合わせに対する尤度比を評価することで識別することが可能である．人物１，人物２の２人物が画面内が交差した後の状況を考えると，

Ａ．　楕円体人物モデル１，２と人物１，２が交差後も正しく対応付けられている．

Ｂ．　楕円体人物モデル１，２と人物１，２が交差後入れ替わっている．

という２種類の仮定に対してそれらの尤度を求め，それらの比から判断すればよいことになる．すなわち，

Ａ．人物１と楕円体モデル１，人物２と楕円体モデル２が対応している確率

$$p_A=\frac{1}{2}p(x_1\mid C_1)+\frac{1}{2}p(x_2\mid C_2)$$

Ｂ．人物１と楕円体モデル２，人物１と楕円体モデル１が対応している確率

$$p_B=\frac{1}{2}p(x_1\mid C_2)+\frac{1}{2}p(x_2\mid C_1)$$

を求め，これらの尤度比 p_A/p_B としきい値を比較することで，生起した事象を判断できる．ただし x_1, x_2 は現在の楕円体人物モデルから得られた観測結果であり，$p(x\mid C_k)$ は識別関数生成モードで得られたクラスｋの平均 μ_k，分散 σ_k を持つ正規分布（$\frac{1}{\sqrt{2\pi}\sigma_k}\exp\{-\frac{x-\mu_k}{2\sigma_k^2}\}$）を表している．

この処理によって，しきい値以上の尤度比が得られた場合，対象とする人物モデルに対して正しい人物ＩＤを割り振る．その後，再び学習モードに移行する．

4.5 実験

提案手法の有効性を確かめるために，２人物の交差，３人物の交差に対する実験を行った．

4.5.1 実験環境

２章での追跡実験と同様の屋内環境で実験を行った．処理にはSGI-Octane (MIPS-R10000-175MHz,Memory 256MBytes)を用い，CCDカラーカメラからの入力信号は160x120Pixel, 24Bit-Colorで取り込んで処理する．カメラは環境を斜め情報から見下ろす視線で設置され，カメラと人物との距離平均は約４～５ｍである．

4.5.2 ２人物の交差に対する識別

画像中に２人物が歩き回る状況で実験を行った．各人物には図４－９で示す経路に従い歩行

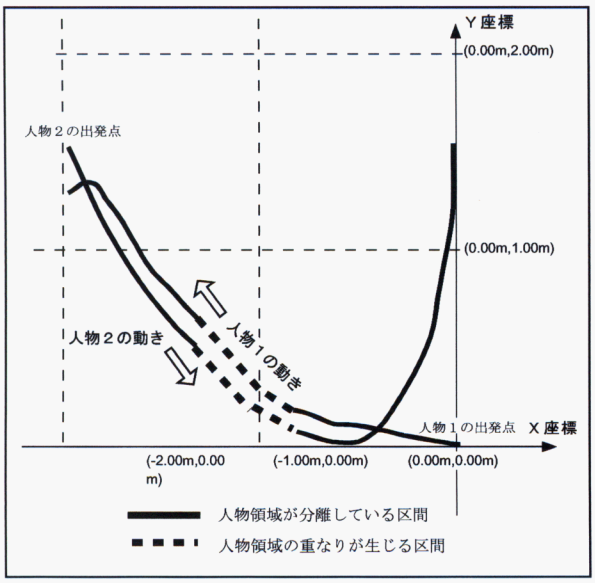


図４－９　２人物の通過軌跡

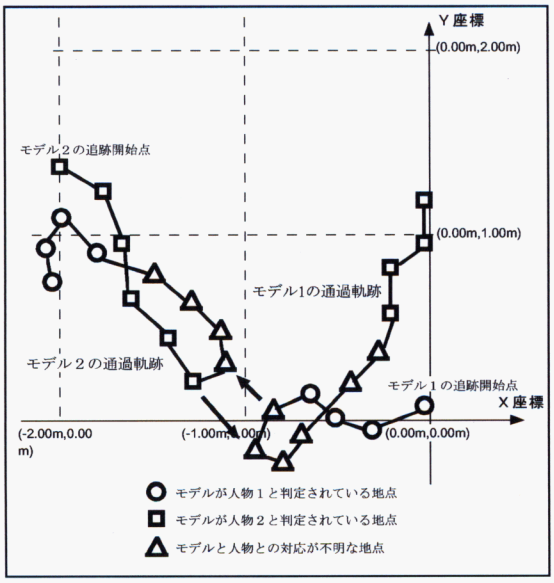


図４－10　通過軌跡検出結果と識別結果

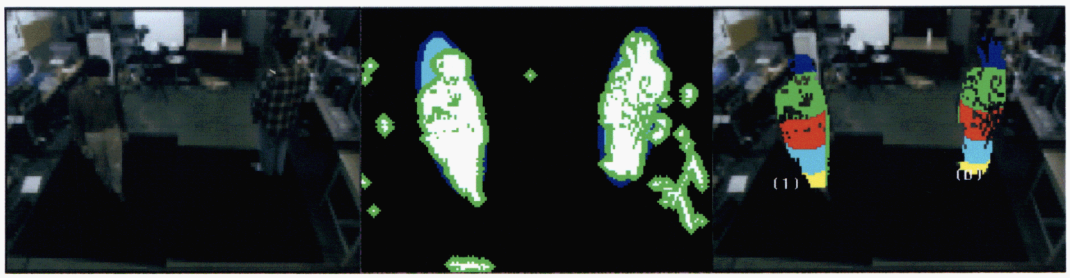


図４－11　２人物の識別実験
左：入力画像，中央：追跡結果画像，右：分割された人物領域

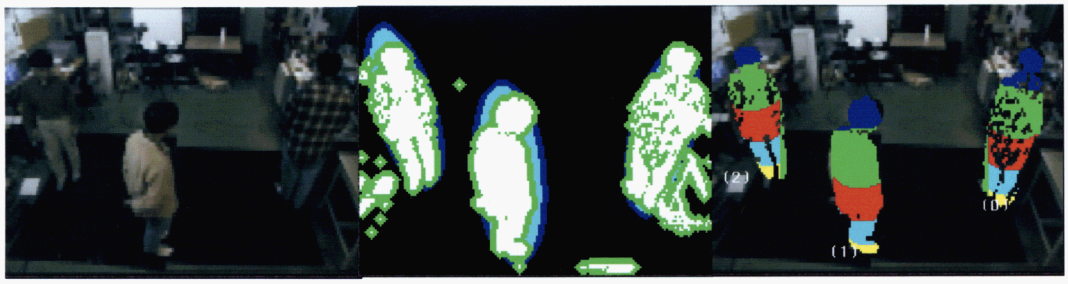


図４－12　３人物の識別実験
左：入力画像，中央：追跡結果画像，右：分割された人物領域

するよう指示した．

位置検出結果および実験中の画像を図４－10，11に示す．図４－10からわかるように，人物領域の重なりが生じている区間でモデルと人物との入れ替わりが生じ，検出結果が不自然になっている．しかし人物領域が分離した後，人物とモデルの正しい対応付けが行われており，追跡人物の識別が正しく行われていることがわかる．本実験では，この後再び人物領域の重なりが生じたが，この場合も人物の識別を正しく行うことが確認できた．また処理時間は平均7.3frame/secで行われており，処理の最適化を行えばリアルタイムでの識別も可能であると考えられる．いくつかのカメラ配置や被験者を用いて実験したが，ほぼすべてのケースにおいて正しく判別することが可能だった．しかし，人物がある程度高速に歩行したり影等のノイズが多く生じる環境では，追跡における人物位置検出や人物像のセグメンテーションがうまくいかず，影や背景などを人物領域として登録してしまう事が見られた．

4.5.3 3人物の交差に対する識別

次に，３人物が歩き回る状況で実験を行った．実験中の画像を図４－12に示す．３人に対する処理ではフレームレートが3 - 4 frame/secほどに低下するため，あらかじめ入力画像をビデオディスクに撮影したあと逐次入力しながら処理を行った．図に示した実験では，３人のうちの２人が交差する状況が３回，３人が同時に交差する状況が２回生じたが，全てのケースで正しい対応付けを行うことができた．しかし，いくつかのカメラ配置や服装の組み合わせを試したところ，正しく識別できないケースも見られた．とくに，影や床面の写り込みが生じるケースでは，追跡における人物領域のセグメンテーションがうまくいかず，写り込み領域を人物領域として学習する状況がみられ，これが最終的な誤識別の原因になったものと思われる．

２人，３人を対象にした両実験とも，いくつかの着衣の組み合わせで行った．下半身の着衣が同色（青），上半身の着衣と他の人物の下半身の着衣が同色というケースについても行ったが正しく判別する事が確認でき，組み合わせの違いにより判別を行うことが可能であると思われる．しかし問題点としては，影や反射などのノイズにより人物追跡結果段階で大きな誤差が生じた場合に，人物領域が大きく異なってしまい，背景や床面の影などを人物領域として学習してしまうことである．本手法では，識別において一般にノイズに強いとされているベイズ識別法を用いているが，カメラ配置や照明条件によっては影が常に生じるケースもあり，識別結果に影響を及ぼすケースも見られた．特に，人物の着衣色がこれらのノイズ領域と同系色であった場合に，この影響が顕著であった．

4.6 検討

実験結果から，前章までに用いてきた差分２値画像を用いた追跡処理で問題となった人物とモデルとの入れ替わり現象に対して，本手法が有効に働くことが確認された．処理速度としては，２人の識別に対してはほぼリアルタイムでの処理が可能であったが，３人以上を対象とした実験では，実時間処理は不可能だった．しかし処理の最適化や計算機性能の向上によって実時間処理を行うことも可能であると考えられる．母集団の数(識別人物の数)に関しては，本追跡システム自体が１つの観測ステーションで３人までを追跡すると仮定しており，本手法程度の識別率・処理速度であれば十分実用的であると考えている．しかしながら，本手法では追跡手法で得られた人物領域のセグメンテーションを元にしているため，影やノイズ等の追跡誤差が大きく影響してくる．特に３人の同時追跡や人物像が大きくなることで計算時間がかかる場合や，人物が比較的高速に移動する事で追跡処理が間に合わないなどのケースにおいて，この影響が顕著に見られた．また背景差分で得られた有効ピクセルのみを学習・識別に用いるため，人物領域と背景領域が同等の明るさである場合，人物領域であっても学習・識別に用いられないピクセルが生じ，正確な色情報を抽出できないという問題点も見られた．

4.7 本章のまとめ

本章では，複数対象を観測するタスクにおける対象の識別問題を，単一ステーションでの画像領域内で解決する手法について提案を行った．人物とカメラの位置関係は拘束できないという想定条件があるため，顔のように人物の特定部位の画像を用いた識別を行う事はせず，人物全身の色分布を用いて，追跡動作を行いながら高速に人物特徴の学習・識別を行う手法を提案した．人物全身の代表色とその組み合わせを識別できる手法をとることにより，同系色の着衣色の人物であっても，その組み合わせを評価することで識別することが可能であることが示された．一方で，追跡における人物領域の抽出エラーや，影や背景ノイズによって識別が困難になるケースも見られた．

本章では同一画像内での人物の対応付け問題を取り扱ったが，分散視覚システムにおいては「ステーション間の複数対象の対応付け」というより大きな問題が存在する．次章では本章で得られた知見を元に，この問題を中心に検討を行う．

5章　複数対象の同時観測

前章では複数対象を同時観測する場合において，単一ステーションの画像内での人物領域交差によって生じる人物の誤対応問題を取り扱った．しかしこの手法では，人物が観測ステーション間の不可視領域に進入した場合等に必要となる「複数のステーション間での対象の対応付け」には適用できないため，依然として複数ステーションによる複数対象の同時観測が安定して実現できるまでには至らなかった．本章では，複数ステーション間の対象の対応付け問題を解決するために，まず全体的なシステム構成について検討を行い，その上で複数のステーション間において対象を対応付ける手法を示す．

5.1 自律観測ステーション群による複数対象観測

前章では同一ステーションでの画像内で生じる人物の画像領域の交差問題を取り扱った．具体的には，観測ステーションの取得画像内で交差する２，３人物を，交差前・交差後で対応付けする手法を提案した．しかし，複数視覚ステーションが同時に観測する状況下では，この仕組みのみによって対象の対応付け問題が解決できるとはいえない．例えば，複数のステーション上で同時に人物領域の交差が生じた場合を考える（図５－１）．この場合，ステーションは交差前・後において人物像の対応付けを行うが，すべてのステーションで同一の解（対応付け）が得られるとは限らない．異なった解が得られた場合，どのステーションが正しい解を持っているかを判断するのは非常に難しい問題である．なぜなら，対象の対応付けの信頼性は各ステーションのパターン認識処理に依存するが，画像の認識処理自体の信頼性を決定することは一般的に困難な問題であるからである．このように，複数の自律観測ステーションを用いたシステ

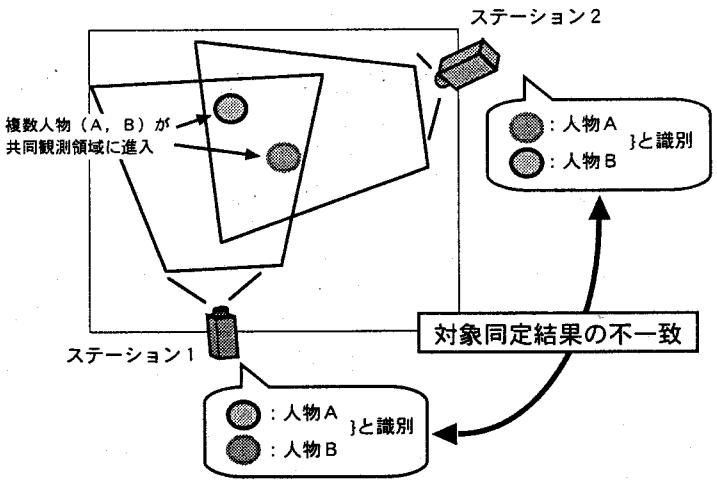


図５－１　複数ステーションの間で人物の識別結果が異なる場合

ムでは，複数対象の対応付け問題に関しても単一処理の観測システムとは異なった問題をもっている．これは，各々のステーションが持つ画像センサや画像の処理手法の向上のみによって解決される問題ではなく，システム構成や統合手法の改良も必要であると考えられる．

以上の点を踏まえて，分散された視覚システムによって複数対象の追跡・識別を行った関連研究（[Rao93],[Nishio92], [Ukita2000]）に対して，そのシステム構成と情報の統合手法の点から考察を行う．

（ａ）Raoらのシステム[Rao93]

このシステムではネットワークを通じて，複数の観測ステーションの観測情報を統合し，対象の認識を行うことができる．この構成を図５－２に示す．観測ステーションは同じ領域を観測していることが前提とされており，各ステーションは対象に対する各種の画像特徴量（移動速度，環境座標系でのサイズ，大きさ）を記録する．一方，識別対象となるオブジェクト（人物，車両，ロボット等）とその投影像の間には，対象に対する画像特徴の生起確率が与えられている．各ステーションでは，観測された対象の画像特徴量から識別対象である確率を推定し他のステーションに送信する．最終的には，全ステーションの識別結果を集め，ベイズ推定を行うことでシステム全体の対象識別結果を出力する．

このシステムにおいて複数対象を同時に識別する場合は，ステーション間での対象対応付けは完了している事を前提としている．ここでは，対象の環境座標系での位置情報のみが用いられている．このため，本システムでの対象問題である「複数ステーション間の対象の対応付け」を主に扱っているとはいえないが，特徴的な点は各ステーションで対象の識別を確定させるのではなく，各対象識別候補に対して重み付けされた確率値を出力し，全ステーションの情報を用いて対象の識別を確定させるという形式をとる点である．この方式であれば，冒頭で述べたステーション間の対象対応付け結果の競合という問題を回避することができる．

（ｂ）西尾・大田のシステム[Nisho92]

西尾らは交差点にカメラを設置し，通過する複数の車両を追跡・識別するシステムを構築した．システムの構成を図５－３に示す．観測ステーションのカメラは交差点の各道路に設置されている状況を想定しており，その観測領域は基本的に重複している場合を想定している．また，各ステーション上では視野内の動物体の追跡やその画像特徴の抽出を行うことが可能である．

このシステムで特徴的な点は，観測を行うリソース（ＳＡ:Sensor Agent）と，観測対象（車両）に対して１対１に割り当てられる観測プログラム（ＶＡ: Vehicle Agent）を明確に区別し，ＳＡはＶＡの要求にしたがってセンシングを行い，その結果を返すという機構を提案していることである．具体的には，以下のような機構で動作する．

1. ＳＡは車両を発見すると，これに対応するＶＡを生成する．
2. ＶＡは観測対象の位置や画像特徴を保持するとともに，全ＳＡに車両の予測移動位置を送信する．
3. ＶＡからの位置情報を受け取ったＳＡは，自分の観測視野内から受け取った予測位置に対応するオブジェクトを探索する．オブジェクトが観測された場合はその情報を返す．
4. ＳＡから探索成功の情報を受け取ったＶＡは，そのＳＡからの情報を元に，自己の情報を更新する．

つまりこのシステムでは，観測対象に対応するＶＡはシステム全体を通して唯一生成され，ＶＡが存在しないステーションはその情報の更新のために用いられている．このように，観測対象に対応する存在をシステム全体で１台に限ることで，ステーション間の対象の対応付け問題自体を回避している．

（ｃ）Ukita らのシステム[Ukita2000]

Ukita らのシステムでは，パン・チルト・ズーム可能なカメラを有した観測ステーションを用いる．これにより，観測ステーションはより広域の観測や，対象をズームすることによる詳細な情報の獲得が可能になっている．また各々の観測ステーションを，対象を観測するためのエージェントとみなしており，複数の観測ステーション（エージェント）が同一の対象を注視することが可能になっている（図５－４）．

同一対象を観測するステーション（エージェント）は「エージェンシー（Agency）」と呼ぶエージェントの集団にグループ化される．エージェンシー内の１台のステーションがマスタエージェント（master agent）となり，同じエージェンシー内の他のエージェント（worker agent）に指示を与えることで同時観察を実現する．一方，マスタからの要求を受けない場合には，自律的な観測によって対象の発見動作を試みる．

この一連の研究では，同一対象を注視する存在（エージェント）がエージェンシーというグループを作るというシステム構成を提案した．エージェンシー内のステーションでは，その情報を交換して，対象に対してより詳細な情報を獲得することができる．

これらのシステムを整理すると，個々のシステムが定義するエージェントおよびエージェンシーの概念に違いがあることがわかる．エージェントの考え方を導入した（ｂ）の構成では，「観測対象＝エージェント（ＶＡ）」という構成をとることによって，対象の同定問題を解決している．これは，多数の対象を同時に扱う仕組みとしてはシンプルであり優れているが，一部のＳＡが観測ミスを起こすと，その後の復帰が難しい可能性がある．すなわち，ＳＡは常にＶＡの観測要求に従っており自律した観測を行わないため，観測に対する突然のノイズなどに弱い可能性がある．

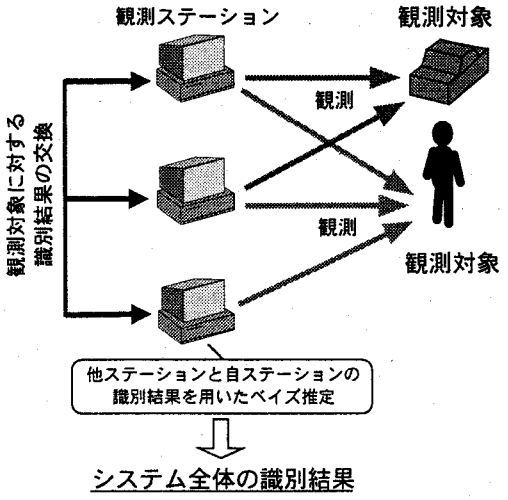


図５－２　Rao らのシステム（ａ）

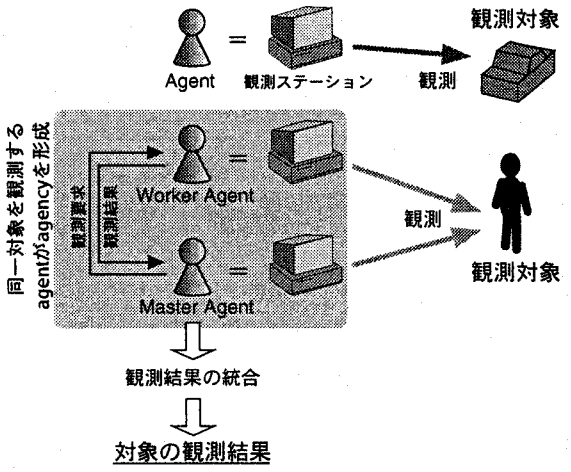


図５－４　Ukita らのシステム（ｃ）

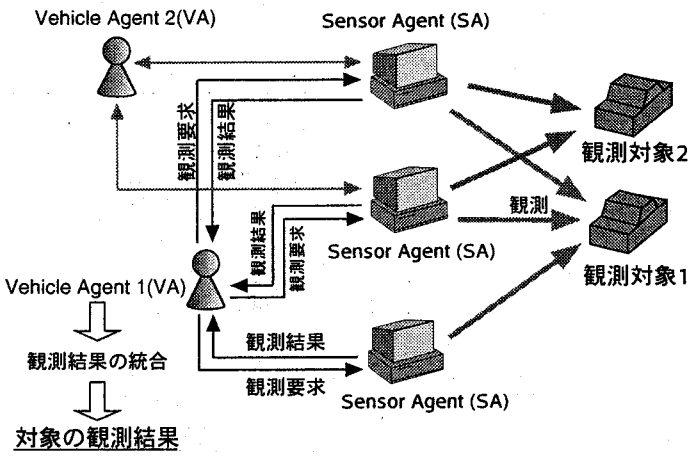


図５－３　Nishio らのシステム（ｂ）

一方（ｃ）のシステムは，個々に自律的観測機能を持つ観測ステーション（観測エージェント）を用い，観測対象とはエージェントのグループ（エージェンシー: agency）が１対１に対応する．すなわち，「観測対象＝エージェンシー」という関係によって対象同定を実現する．各々の観測エージェントは自律して観測を行う機能があるが，「観測ステーション＝１つの対象を注視するエージェント」という構成をとっているため，観測ステーション数以上の対象を同時観測することができない．

ただし，これらのシステム構成の違いは，それぞれのもつ目的が異なっているという側面もある．(ｂ)のシステムは車両追跡を目的としており，ある程度環境条件が拘束された範囲内で，観測視野内の全対象を追跡することを目的としているのに対し，(ｃ)のシステムでは特定対象のより詳細な観測を目的としている．つまり，同じ「複数対象の追跡」タスクを行うシステムであっても，その具体的な目的を設定することによってシステムの構成は変化する．

5.2 提案するシステム構成とその特徴

本システムでは複数対象の同時観測タスクの目標として，観測環境に存在する全人物の観測を目的とする．提案するシステムでの観測対象・観測エージェント・観測ステーションの関係を図５－５に示す．この構成の特徴として，以下の点が挙げられる．

（１）１台のステーションにより多数の対象の同時観測が可能

観測環境内に存在する全人物の観測を行うためには，前節（ｂ）の構成のように単一のステーション上で複数対象の同時観測が可能になる構成が必須である．これは，観測を行うプログラム（観測エージェント）を単一ステーション上で複数並列して実行することによって実現できる．

（２）複数の観測エージェントの自律動作によるロバスト観測

観測エージェントの対象の追跡・観測動作は自律的に行われ，集中指示サーバや他のステーション等の指示に従った構成はとらない．すなわち，（ｂ）あるいは（ｃ）のシステムで行っていた「ＶＡ(Vehicle Agent)」や「マスタエージェント」という概念は導入しない．しかし，複数の観測ステーション上で，同一対象を観測するエージェント同士は，（ｃ）のシステムと同様エージェントのグループ（エージェンシー:agency）を構築する．

本システムで提案する「エージェンシー」と，（ｂ）および（ｃ）のシステムで提案されている構成との違いを図５－６に示す．図５－６（左）のように，（ｂ）および（ｃ）において用いられるＶＡやマスタエージェントは，観測を行う存在（ＳＡ（Sensor Agent），ワーカエージェント）からの情報を収集・統合し，その結果から立てられた観測指示をＳＡ，ワーカエージェントにフィードバックする．一方で本システムでは，各観測エージェントから得られた情報は

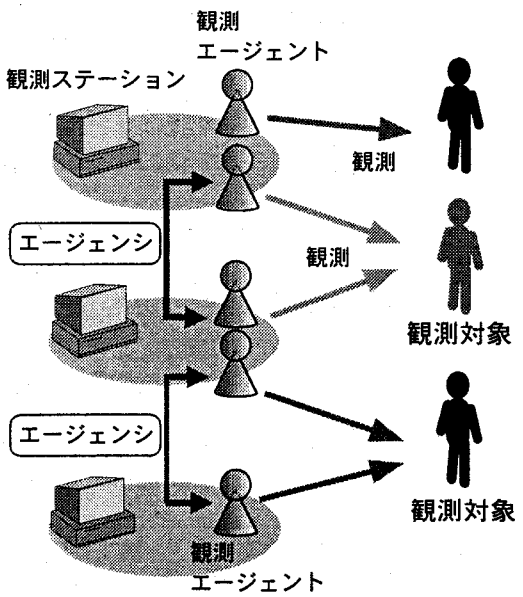


図５－５　提案するシステムでの観測対象・観測エージェント・観測ステーションの関係

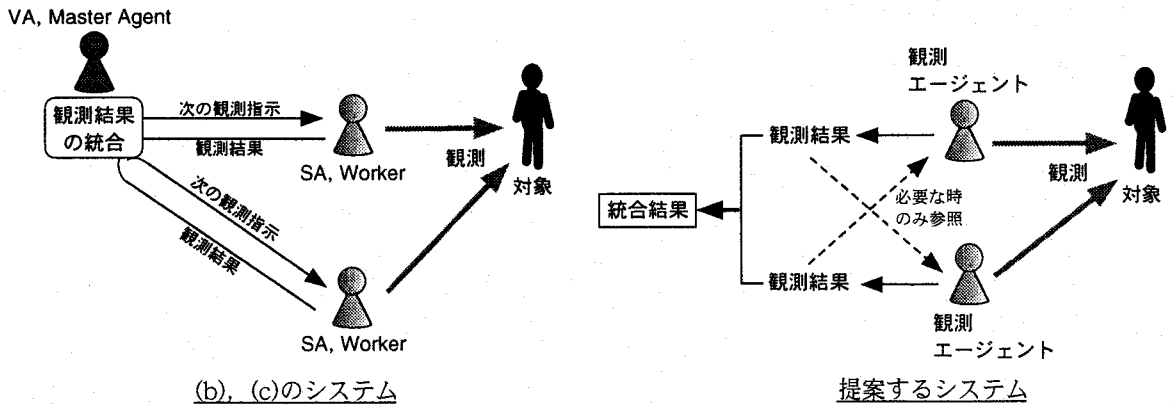


図５－６　各システムにおける情報の流れ

システム全体の統合結果を導き出すためにもちいられるが，そこから導き出された結果等を常に観測エージェントにフォードバックすることはない．観測エージェントは通常，自らの情報のみに基づいて追跡処理を行い，対象の観測が不可能になる等の状況下でのみ，同一エージェンシーに属する他のエージェントの観測結果を利用する．

これらの方針はシステムのロバスト性を考慮したものである．ここでいうロバスト性とは２つの意味を持つ．

第１にステーションのエージェントの障害に対するロバスト性が挙げられる．本システムでは各々のエージェントの観測は自律して行われ，他の存在からの指示を受けないため，一部のエージェントの動作停止等がシステム全体に影響を及ぼすことはない．

第２に観測に対するロバスト性が挙げられる．この性質は，複数のエージェントで同一対象を同時観測する場合，顕著に示される．ＶＡやマスタエージェントを用いた方式の場合，ＳＡやワーカエージェントで得られた観測結果はＶＡやマスタエージェントで統合され，その結果から次の観測指示が決定されＳＡ・ワーカエージェントに送られる．一方提案する構成では各エージェントの観測は自律しているので，それらの保持する観測結果は異なったものであり，次の観測指示も各エージェントで独自に決定される．この２者の違いは，システム内で有する対象に対する「情報の多様性」の有無であるといえる．前者では，対象に対する観測結果はシステム内（エージェンシー内）で唯一であるため，ノイズや観測ミス等でＶＡ・マスタエージェントの統合結果にエラーが生じると，全てのＳＡ・ワーカエージェントへの観測指示にエラーが含まれてしまい，以降の観測への影響が大きくなる．一方提案する構成では，全てのエージェントが独自の解を保持し次の観測指示も独自に決定するため，一部のエージェントに観測エラーが生じた場合でもエージェンシー全体に影響を及ぼすことはない．また観測に失敗したエージェントは，他のエージェントの観測結果を参照することによって，自らの観測を再開することも可能である．このように観測する存在（観測エージェント）に自律性をもたせ，システム全体で観測結果を多様化することで，観測のロバスト性を向上させることを目指している．

（３）エージェンシーによる対象の対応付け

本システムにおいて同一対象を観測するエージェントは，エージェントの集合（エージェンシー）を構築すると考える．つまり，本システムにおいてエージェンシーは対象と１対１に対応している．すなわち，複数のステーション間の対象の対応付け問題は，エージェンシーをどのように構成するかを考えれば良いことになる．本システムでは，観測エージェント同士が互いの観測結果を相互比較し，自律的にエージェンシーを構成する手法によってこの問題を解決した．

（４）多様な特徴量の導入が容易な構成

従来のシステムでは，環境座標系での対象の位置等を基準としてステーション間の対象の対応付けを行い，その上で他の特徴量を対象識別等に用いる．しかし，画像から獲得される特定の特徴量は状況や環境に応じて有利に働く場合，不利に働く場合が存在する．例えば，複数の人物が近接して歩行した場合などは，対象の位置情報では対応付けが困難である．すなわち，対象の対応付け手法を一般性のあるものし，様々な状況で適用可能にするには，観測可能なすべての特徴量を対応付けに用いることのできる仕組みをとることが望ましい．本システムでは，様々な特徴量を一括して比較し，その結果を「確信度」パラメータで表すことによってこれを実現した．確信度は（３）で述べた，エージェンシーを構成する基準となる．

5.3 アルゴリズムの概要

次に，提案するシステムの構成およびアルゴリズムの概要を述べる．

5.3.1 システムの構成

本システムのソフトウエア構成を図５－７に示す．観測ステーションはカメラが結合された計算機であり，環境中の複数の地点に配置される．対象の観測を行うプログラム（観測エージェ

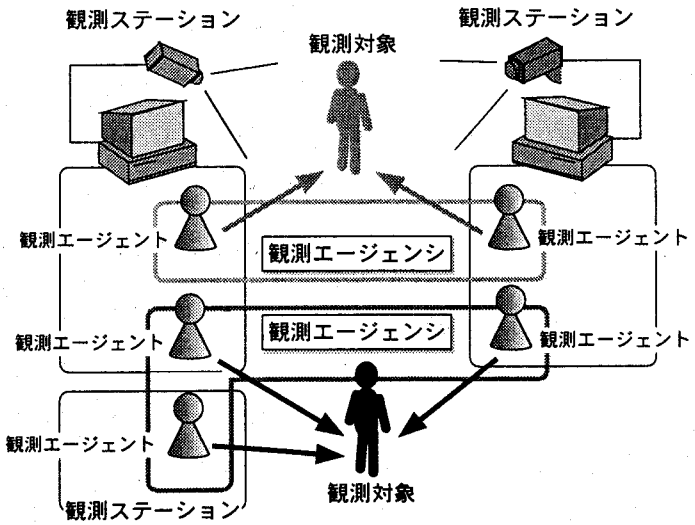


図５－７　システムのソフトウェア構成

ント)は，１台のステーション上で複数並列動作し，各々が自律して対象を観測する能力を持っている．それぞれの観測ステーションにおいて，観測エージェントは観測可能な対象に対して１対１に生成され，対象を観測しその画像上の特徴量を計測する．

ある対象が複数ステーション上のエージェントで同時観測される場合，同一対象を観測するエージェント同士は「エージェンシー(agency)」と呼ぶエージェントの集合にグルーピングされる．つまり本システムにおける観測ステーション間の対象の対応付けとは，このグルーピング処理を行うことと同義である．同様に，観測ステーション間の対象の対応付けエラーの解消は，“エージェントが現在所属するエージェンシーから離脱し他のエージェンシーへ参加する”といった「エージェンシーの再構成」を行うことで実現できる．

5.3.2エージェントのグルーピング

エージェントのグルーピング（エージェンシーの形成）を行う基準としては，各エージェントが持つ「確信度」パラメータが用いられる．これは各エージェントが与えられた対象を正しく観測している可能性を示す値であり，０～１の値を取る．エージェントは自身の観測結果に応じてこの値を逐次更新する．システム内では，同一の対象に関する確信度の高いエージェントがグルーピングされ，エージェンシーを形成する．確信度が低下し閾値以下になったエージェントは，エージェンシーから離脱し単独で観測を行う．

5.3.3確信度の更新

前述したようにエージェントは，自らの観測結果を元に確信度を逐次更新する．すなわち，与えられた対象を正しく観測している場合には確信度を上げ，異なった対象を観測している可能性がある場合には確信度を下げる処理を行う．エージェントがこの可能性を知ることのできる情報として以下の要素が挙げられる．

（１）画像上のイベント

追跡中に注視対象の画像領域が他の対象の画像領域と交わる等のイベントが生じた場合，対象の対応付けミスが生じる可能性が高い．この場合には確信度を下げる操作を行う．

（２）他のエージェントの観測結果との比較

同一対象を観測しているエージェント（同一エージェンシーに属するエージェント）の観測結果は，エージェントごとの計測誤差は含まれるものの，その特徴空間上の距離は近いものと考えられる．逆に，他のエージェントに比べ大幅に異なった観測結果が得られた場合，そのエージェントは異なった対象を観測している可能性が高い．ゆえに，エージェントの観測結果と，(同一エージェンシー内の)他のエージェントの観測結果との差が閾値以上である場合には確信度を下げ，閾値以下の場合には確信度を上げる操作を行う．

（３）過去の観測結果との比較

正しい対象を追跡している可能性が高い（確信度が高い）状態において，観測結果から得ら

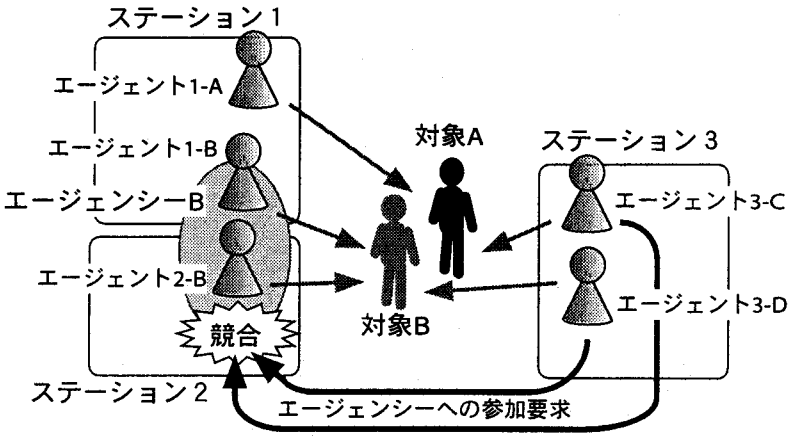


図５－８（ａ）グルーピング時のエージェント同士の競合

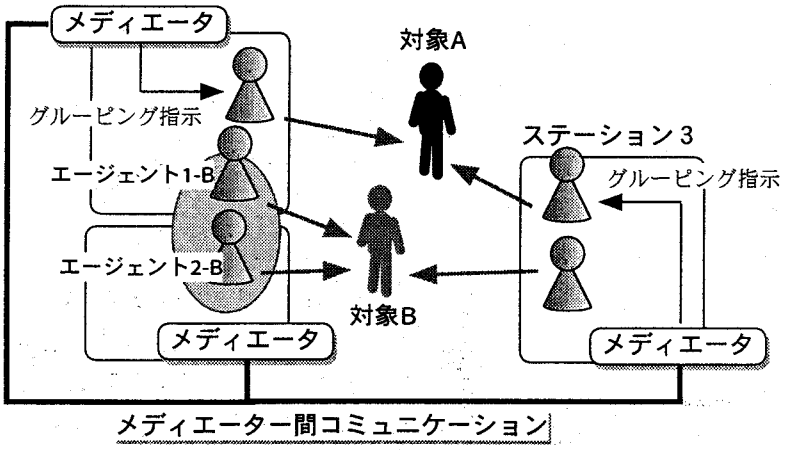


図５－８（ｂ）　メディエータによる競合の解消

れる対象の特徴量を蓄積しておき，確信度が低い状況では，過去に蓄積された特徴量と現在の観測結果を比較することで，正しい対象を観測しているかを判断し確信度を更新する．

5.3.4メディエータ

単独で観測するエージェントを既存エージェンシーに参加させる操作や，単独で観測するエージェント同士を組み合わせてエージェンシーを生成する操作などは，観測エージェント間のやり取りのみでは競合が生じる場合がある．

図５－８（ａ）の状況では，観測ステーション１上のエージェント（１－Ｂ）と，観測ステーション２上のエージェント（２－Ｂ）が，対象Ｂを観測するエージェンシーＢを作っている状況である．一方ステーション３から見ると，対象Ｂとそれに近接する対象Ａは近接しており，それぞれを区別することは困難である．この状態において，ステーション３上のエージェント３－Ｃ，３－Ｄは単独で対象を追跡する状況となっていた場合，エージェント３－Ｃあるいは３－Ｄが個々にグルーピング処理を実行しようとすると，双方がエージェンシーＢに参加要求を起こすため競合が生じる．

この問題を解決するため，各ステーション上に「メディエータ」というエージェントをおき，単独で観測するエージェント間の仲介処置を行う．すなわち，図５－８（ａ）のように対象の対応付けの判断が困難な状況ではグルーピング処理は行わず，図５－８（ｂ）のように対象が分離できる状況においてエージェンシーの再構築を観測エージェントに指示する．

5.4 システムの実装

以上のアルゴリズムを観測ステーションに実装した．その構成を図５－９に示す．システム内のモジュールやエージェントは，UNIXスレッドによって並列動作している．このうち，各モジュールはシステムが動作する間常に動作しているが，観測エージェントの個数はそのステーションで観測している対象数に応じて変化する．また，モジュール・エージェント間の情報のやり取りは３種の共有メモリによって行われる．

各々の観測エージェントには，システム内でユニークな値を持つエージェントＩＤがつけられている．また，各エージェントには観測対象ＩＤも与えられているが，これによりエージェントが現在観測している対象を識別する．観測対象ＩＤは０以上の値をもつが，エージェント

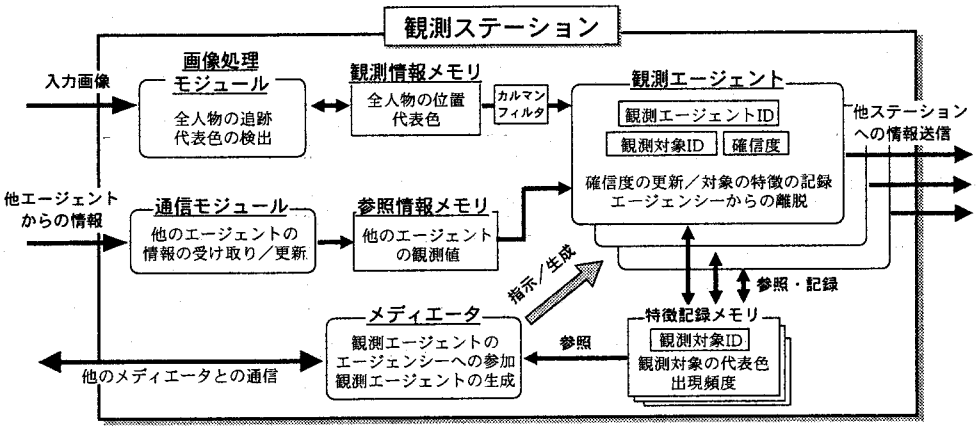


図５－９　観測ステーションの構成

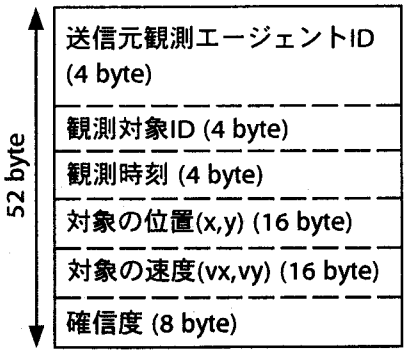


図５－10　エージェント間の通信に用いられるパケット

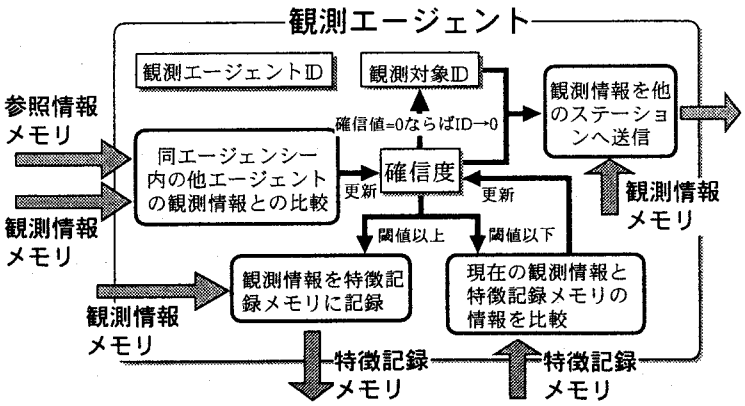


図５－11　観測エージェントの構成

が与えられた対象を見失い単独追跡を行っている場合にはこの値が０となる．同一の観測対象ＩＤ（＞０）を持つエージェント同士は，複数の観測ステーション間でエージェンシーを形成し，互いの情報を比較しあう．

5.4.1 画像処理モジュール

画像処理モジュールでは，２章で述べた手法によって人物の追跡や位置の計測を行う．２章での「人物位置管理マップ」は，本システムでの「観測情報メモリ」に相当するが，画像からの特徴量（人物の代表色や追跡時に生じた画像上のイベント等）を保存できるように拡張されている．このモジュールは以下の操作を，観測可能な全人物に対して逐次行うことで人物位置を更新する．

1．観測情報メモリから人物の現在情報を読み出す．
2．読み出した位置を初期位置として追跡処理を行う．
3．追跡結果を観測情報メモリに書き戻す．追跡処理中のエラーが発生や他の追跡対象と画像領域が重複した場合にはフラグを立てる．それ以外の場合は，追跡対象領域の３代表色（色相・彩度成分）を観測情報メモリに記録する．

5.4.2 通信モジュール

通信モジュールは，他のエージェントからネットワークを通じて送られる情報をすべて受け取り，参照情報メモリに格納する．通信モジュールが受信する通信パケットを図５－10に示す．通信パケットの先頭部分には送信元のエージェントＩＤおよび観測対象ＩＤが付けられている．通信モジュールはパケットを受けとると，この２種のＩＤを元に分類して参照情報メモリへ格納する．

5.4.3 観測エージェント

観測エージェントには観測エージェントＩＤ，観測対象ＩＤ，および確信度（初期値１．０）が与えられている．観測エージェントの構造を図５－11に示す．以下では，観測エージェントが行う動作について述べる．

（１）確信度の更新

5.4.3 節で述べた要素を基準に，確信度 c を更新する．これは以下のように定式化できる．

$$
\begin{aligned}
c_t = c_{t-1} &+ k_1 \cdot eval_event \\
&+ k_2 \cdot eval_other_agent \\
&+ k_3 \cdot eval_knowledge
\end{aligned}
\qquad \text{（式５－１）}
$$

各評価値は，以下のようにして求める．

（ａ）*eval_event*:画像上のイベントの有無による評価値

画面上で他の追跡対象と交差するイベントがあった場合には－１とし，それ以外の場合には０とする．すなわち，画像による人物追跡処理の中で「人物の対応付けが不安定になる状況」が生じた場合において確信度を下げる役割を果たす．

（ｂ）*eval_other_agent*:他のエージェントの観測結果との比較による評価値

参照情報メモリ内に，同じ観測対象ＩＤを持つエージェント（同じエージェンシーに属するエージェント）の情報が存在し，その確信度が自分の確信度よりも高い場合，その検出位置と自分の検出位置の差を計算し確信度に反映する．この評価値*eval_other_agent*は以下のように計算できる．

$$eval_other_agent = k21*(dist_diff - dist_thresh) + k22*(color_diff0 - color_thresh0)$$

ただし，　*dist_diff*:エージェント間の位置検出結果の差異

dist_thresh: *dist_diff*のしきい値

color_thresh0 : *color_diff0*のしきい値

（ｃ）*eval_knowledge*:過去の観測結果との比較による評価値

エージェントの観測対象ＩＤと同じＩＤを持つ特徴記録メモリを参照し，観測された代表色の過去の出現頻度を求める．これは以下の式で表される．

$$eval_knowledge = k31*(color_diff1 - color_thresh1)$$

ただし，　*color_diff1*: 過去に観測された代表色と現在の代表色の差

color_thresh1: *color_diff1*のしきい値

*color_diff1*を求めるためには様々な方法があるが，今回は代表色の出現頻度から(*color_diff1* - *color_thresh1*)の値を決定した．すなわち現在の３代表色のうち２色以上の出現頻度が閾値以上であれば＋１，１色であれば０，０色であれば－１を(*color_diff1* - *color_thresh1*)の値として与えることとした．

（２）しきい値の決定

確信度更新で用いられる各特徴量のしきい値は，各観測ステーションでの計測誤差に依存している．本システムでは位置の差および色（色相）の差を用いているが，前者のしきい値（*diff_thresh*）に関しては，各観測ステーションによる位置誤差範囲が床面上の各方向に対して±0.3m程度のため，対象人物のミスマッチングが生じている場合ステーション間の検出位置の差は計測誤差の２倍以上の差があるものと考え，その２倍程度（0.6m）を閾値とした．一方色相差のしきい値(*color_thresh0*)は複数のステーション間での色相計測誤差によって決定するが，本システムでは経験的に±20度を用いた．

（３）係数の決定

以上の各評価値の計算式より，確信度更新を表す式５－１は以下のように書き換えられる．

$$
\begin{aligned}
c_t = c_{t-1} \\
&+k_1 \cdot eval_event \\
&+k\{k_{21}(dist_diff - dist_thresh) + k_{22}(color_diff_0 - color_thresh_0)\} \\
&+k\{k_{31}(color_diff_1 - color_thresh_1)\}
\end{aligned}
$$

これらの係数に対しては，以下のように決定できる．まずk1およびkは，画像上のイベントに対するペナルティとエージェンシー形成のダイナミクスを制御するパラメータである．前者は画像のイベントがどの程度誤対応を生じさせる可能性があるかを示しているが，本システムの場合２～３人物が画面中で交差することが想定されるため，誤対応する確率から考えて0.4の値を与えている．また，後者はエージェンシーの形成および解消のスピードの早さを制御する．この値が大きい場合そのスピードは速くなり，小さいと遅くなるが，これは目的とするタスクや状況に依存するといえる．短いインターバルで人物領域が交差するイベントが生じる場合などでは，この値を大きくして対応付けを早めに解消・再決定することも考えられるが，逆に少しのノイズ等によってエージェンシーの対応付けが解消されてケースも想定できる．このパラメータは，実験において経験的に0.01～0.05までの値を与えることとする．

またk22とk23は，ステーション間で対象を対応づける際に用いる各特徴量の重要度を決定するパラメータである．本システムでは経験的に距離１ｍの誤差が色相で100度に対応するように設定している．この重要度に対しては，各特徴量が対象シーンに対してどの程度有効かに依存する．対象の識別に対して，「位置情報」を用いることが有効であればその重み付けを大きくすべきであり，また他の特徴量が有効であればその重みを大きくすることが有効であると考えられる．全体を通してこれらの係数は，

１．識別対象となる複数人物像の色分布

２．環境条件（環境光やカメラ特性がどの程度統一されているか？）

３．各特徴量の観測誤差はどの程度か？

４．対象シーンに対してどの特徴量が有効か？

によって判断することができる．

（４）観測対象に関する代表色出現頻度の更新

特徴記録メモリには，観測対象ＩＤに対応する対象の代表色出現頻度が記録されている．エージェントの確信度が高い状態において，観測された代表色を特徴記録メモリに記録し出現頻度を更新する．

（５）観測結果の通知

エージェントがいずれかのエージェンシーに属しており（観測対象ＩＤ≠０），かつ追跡に成功している場合，現在の観測結果および確信度を他のステーションに送信する．送信には図５

－８で示した通信パケットを用いるが，この中には観測が行われた時刻およびカルマンフィルタによって予測した人物移動速度も記述しておく．また送信先ステーションとしては，３章で示した方法を用いて「観測領域の重なるステーション」に限定し，必要とする通信帯域を小さくする．

（６）追跡失敗時の処理

画像処理モジュールが対象の観測に失敗した場合，同一エージェンシー内の他エージェントの観測結果を参照情報メモリから読み出し，対象の位置情報を観測情報メモリに書き出す．画像処理モジュールではこの値を元に観測をやり直す．この位置情報は，エージェンシー内で最も確信度の高いエージェントの情報を元にしており，その観測時刻と現在の時刻・対象の移動速度から予測した位置を用いる．一定時間この処理を繰り返しても発見できなかった場合，対象が観測視野外に出たと判断し観測エージェントを消滅させる．

（７）エージェンシーからの離脱

確信度が０となった場合，現在参加中のエージェンシーから離脱して単独で追跡を続ける．これは，自らの観測対象ＩＤを例外値の０に設定することで行われる．

5.4.4 メディエータ(mediator)

メディエータは観測ステーション内に１つずつ動作しており，新たな観測エージェントの生成や単独で観測中のエージェントからのエージェンシー形成など，観測エージェントを管理する役割を果たす．メディエータの役割は以下のようなものである．

（１）観測エージェントの生成

参照情報メモリ中に現在ステーション内で観測していない対象が記述され，それが視野内に存在すれば新たな観測エージェントを生成する．対象の追跡対象ＩＤおよび初期位置，初期確信度は参照情報メモリに従う．

（２）単独観測エージェントのエージェンシーへの組み入れ／複数の単独観測エージェントによるエージェンシーの生成

ステーション内に単独で観測しているエージェントが存在する場合，以下の操作を行う（図５－12）．

（ａ）エージェントの観測結果と，全対象の過去の特徴情報，他の観測エージェントで得られた観測結果を比較し，その結果をメディエータ内の「確信度リスト」に記録する．この比較処理は，エージェントがエージェンシーに属している場合の確信度更新アルゴリズムと同様である．

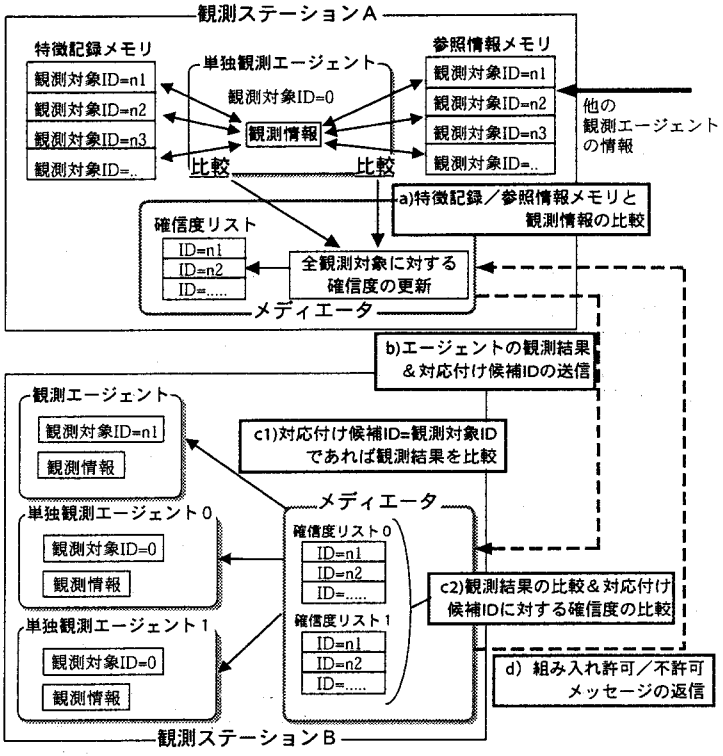


図５－12　メディエータの動作

（b）確信度リストのいずれかの値が閾値以上であり，そのＩＤと同一の観測対象ＩＤを持つエージェントがステーション内に存在しない場合，確信度リスト内のＩＤを対応付け候補ＩＤとする．次に，他ステーションのメディエータにエージェントの観測結果および対応付け候補ＩＤをメッセージとして送信する．

(c1) メッセージを受信したメディエータは，ステーション内に対応付け候補ＩＤと等しい観測対象ＩＤを持つエージェントが存在するかを確かめる．存在した場合，メッセージ内の観測結果とエージェントの観測結果の相関を評価する．

(c2) (c1)にあてはまるエージェントが存在しなかった場合，ステーション内の全単独追跡エージェントに対し以下の比較を行う．

i. 観測結果の差が小さいか？

ii. ステーション内の単独追跡エージェントに対応する確信度リストをチェックし，送信された対応付け候補ＩＤに対する確信度が最も高いか？

（d）(c1)および(c2)における比較で該当するエージェントが発見できた場合は，送信元エージェントに組み入れ許可メッセージを返し，発見できなかった場合には不許可メッセージを返す．

5.5 実験および結果

本手法の有効性を確認するため，３台の観測ステーションが観測するシーンの中で，２人物が歩き回るタスクを行った．

5.5.1 実験環境

観測ステーションは視点固定型のＣＣＤカラーカメラとＰＣ（Intel PentiumIII-500MHz dual）で構成され，カメラからのビデオ信号は240x180 pixelの大きさで取り込まれ，処理される．各カメラは屋内環境に斜め上方から見下ろす角度で設置され，図５－13に示す観測視野を持っている．カメラ１は部屋の入り口付近を観測し，入室する人物を捕捉する．このため，この観測エージェントには入り口付近から人物が出現するという情報をあらかじめ与えておき，エージェントは与えられた情報と自らが持つ環境マップに基づき，入り口部分で初期的に人物を捕捉する動作を行う．また，３台のカメラのキャリブレーションとして，環境中の１点に座標原点を設定し，環境中のマーカと入力画像中の点を手動で対応付けることでカメラパラメータを得ておく．また，各ステーションの内部時計はNTP[Mills88]を用いて補正しておく．

5.5.2 追跡実験

３台の観測ステーションの視野内を歩き回る２人物を追跡する実験を行った．人物の通過軌跡を図５－14，実験中の画像を図５－15に示す．各観測エージェントの人物の検出位置，確信度の変化およびエージェンシーの形成状況を図５－16～18に示す．なお，図中の人物通過点①～⑧は図５－14のものと対応している．実験を通して観測されたイベントを列挙すると以下のようになる．

１．出現を待機する観測ステーション１で入室する２人物が順次捕捉され，他のステーションにその位置情報が伝えられる（人物１①－②，人物２③－④）

２．ステーション２，３で人物が捕捉され，追跡が行われる．

３．地点⑤において２人物が交差する状況（図５－15（ｂ））が生じた．ステーション２では交差前後の人物とエージェントの対応関係は同一（同じエージェントが連続して同じ人物を追跡する）であったが，ステーション３では人物・エージェント間の対応関係が入れ替わる状況が観測された．

４．両ステーションは人物の観測を続け，地点⑧ではステーション２において人物領域が交差する状況が生じた．

確信度変化（図５－17）から見ると，画像中での対象の重なる状況（地点⑤）において，ステーション２，３上のエージェントの確信度に大きな影響が生じていることが確認できる．こ

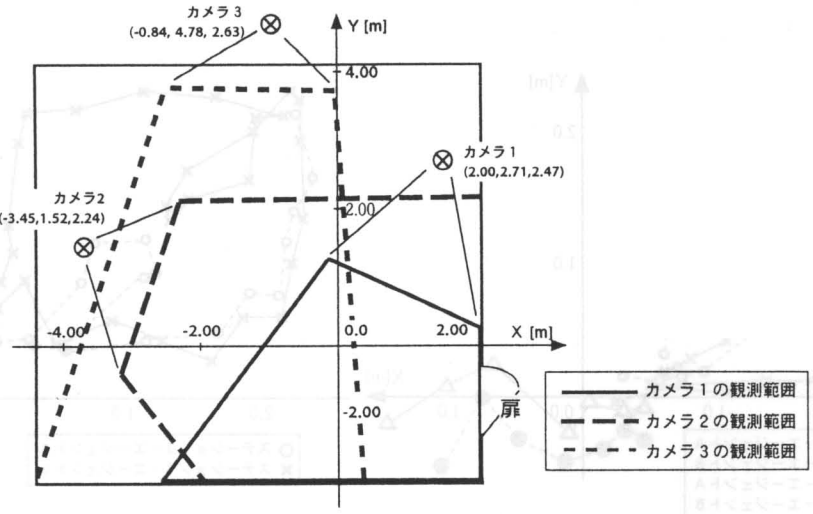


図５－13　３台のカメラ配置

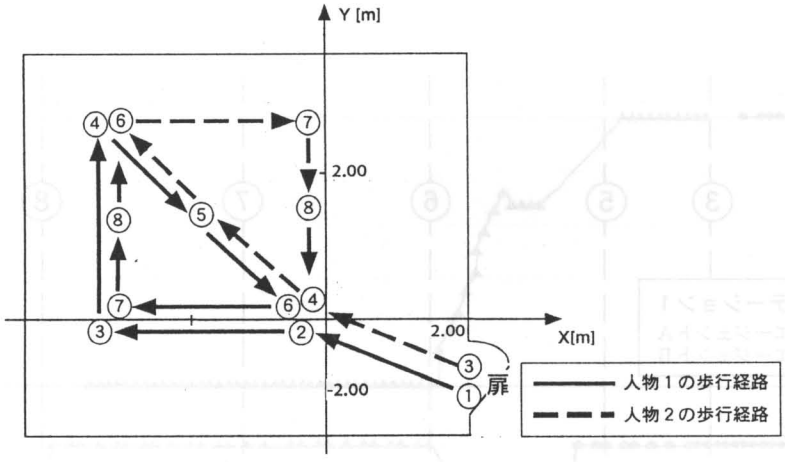
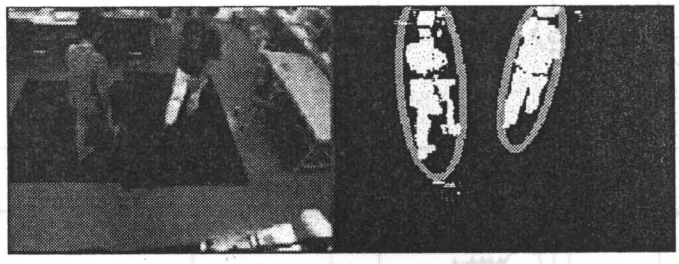
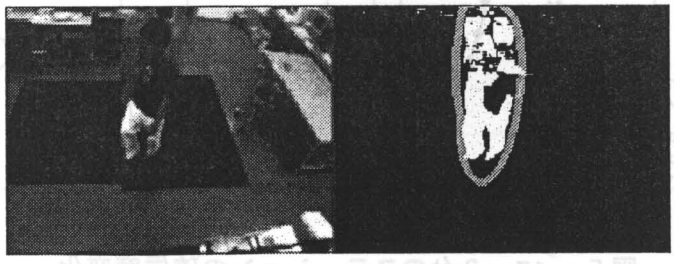


図５－14　人物の歩行軌跡



(a)人物領域が分離している状態



(b)人物領域が重なっている状態

図５－15　実験中の画像

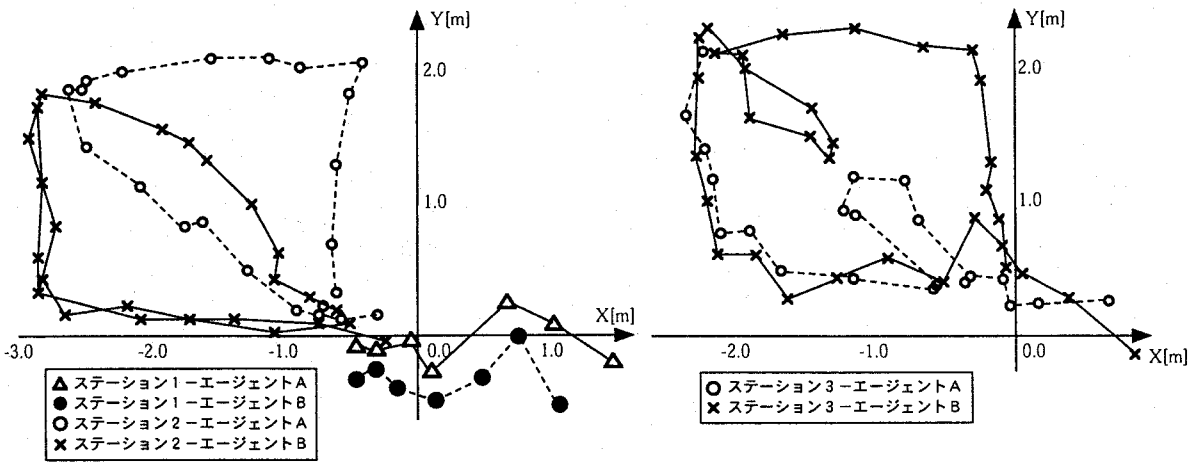


図５－16　３台のステーションによる人物軌跡の検出結果

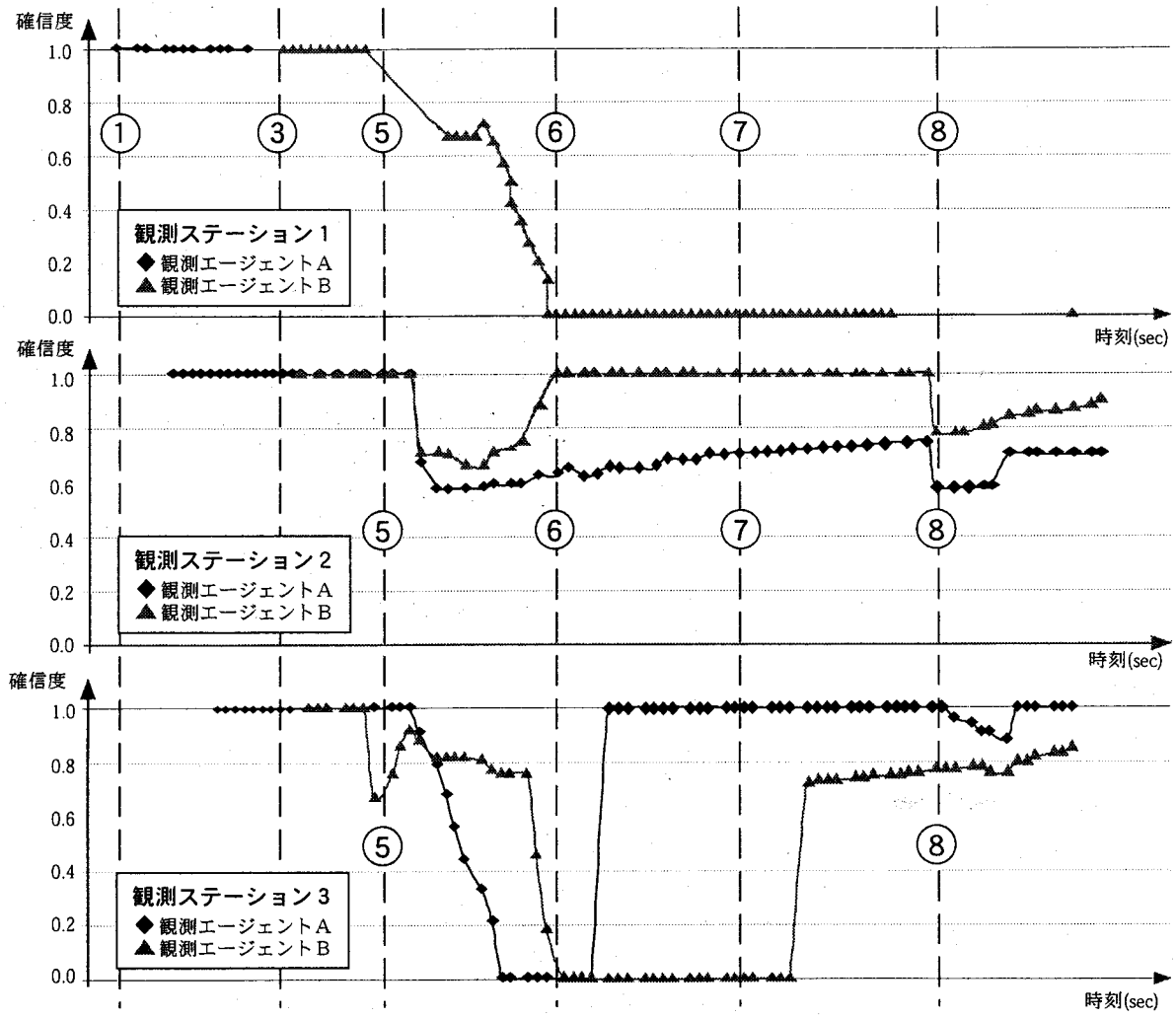


図５－17　３台のステーションの確信度変化

れは，画像中での対象領域が交差するイベントが生じたためであるが，その後ステーション２では確信度が上昇傾向にあるのに対し，ステーション３では逆に下がり，単独観測エージェントに移行している．これは前述したように，ステーション２では対象の交差前後において同一の人物・エージェントが対応付けられたのに対し，ステーション３では誤対応が生じたためである．このことは，各エージェントの人物位置検出結果（図５－16）からも確認できる．ステーション２では人物の通過軌跡を連続して検出できているのに対し，ステーション３では⑤の地点で人物の軌跡を誤って検出している．

しかしその後，ステーション３上のエージェントもメディエータによって正しい対象と対応付けられ，ステーション２のエージェントとエージェンシーを形成することが確認できる．その後，地点⑧においてもカメラ２において人物領域の交差イベントが生じ，確信度の低下が生じているが，最終的には正しい対応付けが得られた．

同様の実験を異なる人物により複数回行ったがおおむね良好な結果が得られた．特に，地点⑤での交差においてステーション２，３で同時に誤対応が生じるケースもあったが，この場合でも最終的には正しい対応付け・対象同定が行われることが確認できた．

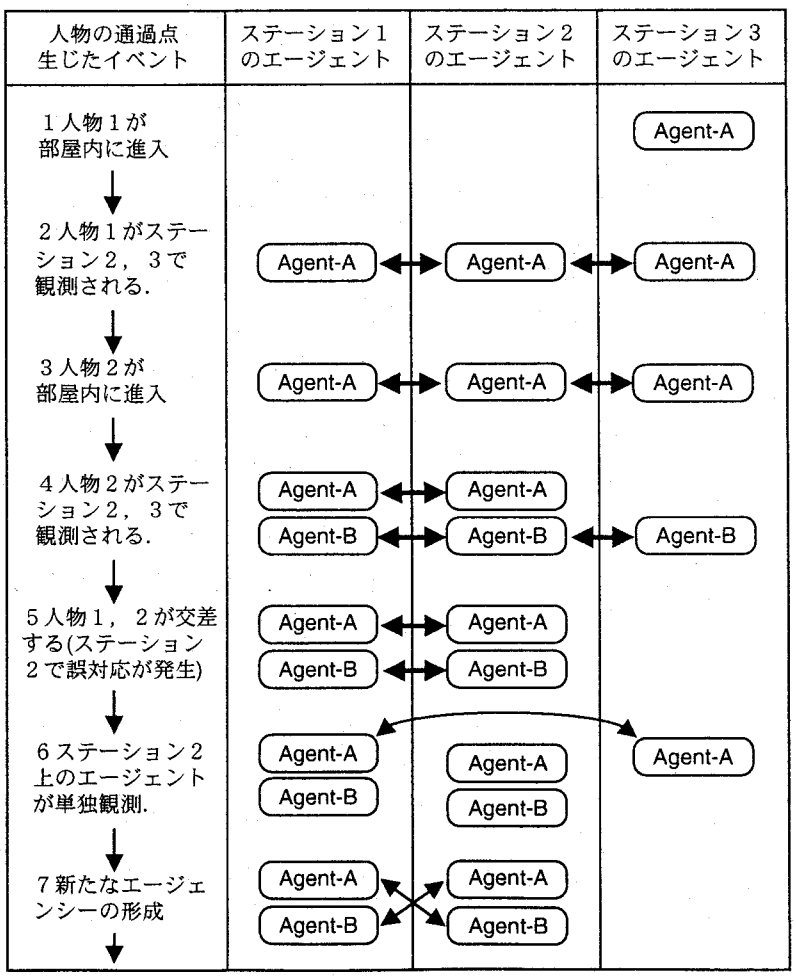


図５－18　エージェンシ形成の変化

5.6 検討

複数の実験を通じて，複数観測ステーション間での人物の対応付けを行う本システムの機構が有効に働くことを確認できた．実験中には，人物像が完全に重なるといった状況が生じ，観測エージェントが一時的に異なった人物を観測する場合もあったが，エージェンシーの解消，再構築によって正しい対応付けが行われることが確認された．処理速度としては，２人物の場合10～15frames/sec，３人物の場合７～12frames/secで可能であり，通常の速度で歩行する複数人物を追跡することが可能であった．

しかし人物の着衣の代表色が近い場合，エージェンシーの形成（対象の対応付け）は正しく行われるものの，過去に蓄積された特徴量との対応付けが成功せず，あやまった観測対象ＩＤを与えてしまうケースもあった．実験システムが用いる特徴量は代表色と位置情報のみである以上，取り扱いのできないタスクだと考えられるが，他の特徴量・比較手法を導入することで解決することができる．前述したように，本システムでは新たな識別特徴量を導入する事は容易だが，その際には以下の点を考慮しなければならない．

（１）特徴量が異なるステーション間で比較可能か？

（２）特徴量の比較時に用いる閾値と観測誤差の関係

（３）確信度更新時の係数

（４）確信度更新の各評価値がとりうる値に対する配慮

（１）は特徴量の性質に関する事項である．対象に関する知識が得られた場合，それが観測ステーションを通じて不変であるか，あるいは観測条件による影響が大きいために直接の比較が困難であるかを考慮する必要がある．今回の実験システムでは，人物の観測位置情報を前者（直接比較可能），人物の代表色出現頻度を後者（直接比較不可能）であると考え，前者に関してはエージェント間での比較を行うが，後者はエージェント内で用いるにとどめた．これは，今回の実験環境の照明環境，設置環境の条件によりステーション間の比較が困難であると判断したためだが，これらの条件が異なれば，代表色の比較も可能になると考えられる．

（２）は，エージェント間の特徴量の比較に用いる閾値をどの程度に設定すべきかという問題だが，これは5.4.3節で示したように，その特徴の観測誤差に依存するといえる．実験システムでの位置の比較を例に取ると，今回の位置検出アルゴリズムでは±0.3m程度の誤差が生じることが確認されているため，エージェント間の検出位置の差が0.6m以内であれば対象が一致していると判断している．この値は用いる特徴量や特徴抽出方法に依存する問題であり，システムの実装時に考慮すべきパラメータである．

（３）も5.4.3節で示した問題で，特徴量比較の結果が一致，不一致であった場合に，どの程度の値でエージェントの確信度を加減するかという問題である．これはエージェンシーの解消・再構築のダイナミクスに影響を与える．一般にこの値が大きいほどエージェンシーの解消・再

構築は頻繁になるが，逆に正しいエージェンシーを構築していても一時的な観測ノイズ等の影響によりエージェンシーが解消されてしまうといった欠点があるため，タスクに応じた値を設定することが重要である．

（４）は，確信度更新の各評価値がとりうる値に対する配慮である．今回の実験システムでは，各評価値のうち画像中のイベントによる評価値（*eval_event*）を{0, -1}の２値，エージェント同士の観測情報の比較による評価値（*eval_difference*）を連続値に，過去の観測結果との比較による評価値（*eval_knowledge*）を{-1, 0, 1}の３値とした．このように，各評価値がとることのできる値を異なる設定としたのは以下のような理由による．

（ａ）画像上のイベントが対象の誤対応を生じさせる影響に対して，連続値的な評価を与えるのが困難であったこと．

（ｂ）過去の観測結果と現在の観測結果の比較には様々な手法があるが，リアルタイムで処理するためには手法の簡略化が必要であったため，４章で用いたような代表色の出現確率を評価する手法を採用したこと．

すなわち，各評価値に対しては本来連続値を用いるべきではあるが，今回の実験システムでは本手法の有効性を評価するため近似的な確信度更新パラメータを用いたといえる．上記２点を解決することによって全てのパラメータに対して連続値を用いることができるため，性能の向上が得られるものと考えられる．

また今回用いた実験システムでは，全てのステーションを観測ステーションとして動作させたため，実験結果は全ステーションを個々にモニタリングする事によってのみ獲得可能であった．しかし実システムとしての応用を考えると，全ステーションの情報を集約，統合しユーザに提供する新たな存在（ステーション）を設けることが考えられる．これにより，追跡動作自体は本章および３章で示した完全分散型のアルゴリズムを用いながらも，ユーザはシステム全体で得られた観測結果をモニタリングすることが可能になる．

5.7 本章のまとめ

本章では，分散視覚システム上で複数対象を同時観測するための手法について述べた．まず冒頭では，複数視覚システムによって複数対象を同時追跡する関連研究について検討を行い，そのシステム構成上の利点・欠点・目的に対して検討を加えた．その上で，本システムが目的とする「観測環境中の全人物の追跡」タスクに対してふさわしい構成を考えた．

まず，観測環境中に多数の対象が進入することを考えると，単一の観測ステーションの視野内に複数の対象が進入する状況が考えられる．このために単一の対象を観測できる観測エージェントを，観測ステーション上で複数並列動作させる仕組みを導入した．また，同一の対象が複

数の観測ステーションの視野内に進入した場合には，各々の観測ステーションで独立した観測を行い，互いの情報は参照しあうのみにとどめた．このように，各ステーション上の観測エージェントが自律して観測することにより，システム全体での観測情報の多様性が生まれ，観測誤差等にロバストなシステムであると考えられる．

一方で，前々章から問題となってきた「観測ステーション間の対象の対応付け」問題に対しては，各ステーション上の観測エージェントによる観測結果を用い，それらをグルーピングすることによって解決する手法を提案した．これにより，同一の対象を複数視点で観測することが可能となり，より高精度な観測情報の獲得や，追跡のロバスト性の向上などが実現された．

以上のような工夫により本手法は，広域観測やロバストな観測を可能にする分散観測システムをより可用性の高いデバイスとして拡張できたと考えている．

6章　結論

6.1　本論文のまとめ

本論文では，複数の視覚センサの観測領域にまたがる広域シーンの認識を行う視覚機能をもつ計算機（観測ステーション）群をネットワーク上に複数配置し，それらが協調・情報統合・競合解消などを行うことで，「分散視覚システム」を提案し，システムの実装を行った．目的とするタスクとして，オクルージョン等を生じさせる障害物が存在する実広域環境内を歩きまわる人物の追跡・位置検出を行った．これは，従来の単眼の視覚システムでは認識範囲の限界や観測できない領域が生じ，多数の映像信号を集中し処理するシステムではリアルタイムの認識が不可能なタスクである．試作システムはこのような環境でも有効に動作することが確認され，本提案が有効なものであることが確認できたと思われる．このシステムを構築するために，３つの技術を開発した．

第一は，２章で示した画像処理によるリアルタイム人物追跡・位置検出手法である．この手法は，人物を３Ｄの楕円体で近似し，背景差分画像と比較することでリアルタイム人物追跡を実現する．通常用いられる多関節の３Ｄモデルを用いた人物トラッキングでは，関節角度の複雑さやセルフオクルージョンなどが生じ，実時間での動作や安定した追跡が困難な場合があるが，本手法では人物をシンプルな楕円体に近似することで実時間での複数人物追跡が可能になった．また，２Ｄベースの手法ではなく３Ｄモデルを用いているため，検出結果を環境座標系上で得られるのも大きな特徴である．同時に，環境中の既知のオブジェクトの３次元情報を利用し，オブジェクトが人物像に及ぼすオクルージョンにも対処可能にすることで，ロバストかつ正確な追跡が可能になった．実験結果から，本手法が分散視覚システムの応用に対して十分な高速性（10 ～ 15frame/sec）と精度をもたらすことが確認できた．

第二は，環境を複数の自律ステーションによって分散観測するための協調手法（３章）である．本システムで用いる観測ステーションは広域を部分毎に分割して分担観測する．これらの視野間を連続的に移動する人物を追跡しつづけるためには，観測ステーション間の情報伝達によるタスク制御が必須である．本研究では，システムの耐障害性や効率良い情報伝達の点を考慮し設計を行った．まず，観測ステーションの初期動作として，現在動作中の全てのステーションの情報を取得し互いの観測領域を把握する必要がある．このため，観測ステーションの他にシステム全体の情報を収集，配布する役割を果たす「システム情報管理エージェント（ＳＰＭ

A)」を設けた．動作を開始する全ての観測ステーションは，このエージェントに自分の情報を伝えるとともに，システム全体の情報を獲得し，自己の観測視野と他の観測視野の関係を与えられた環境マップ上に描くことで，互いの観測視野の関係を把握する．

追跡動作を開始すると，ステーション間では観測された人物位置に応じて適切なステーションを選択し，タスク制御メッセージを送りあう．人物はステーションの観測視野のうち，隣接した領域間しか移動しないため，タスク制御メッセージの通信は地理的に隣接した領域を観測するステーション間のみで十分である．このような構成をとることにより，集中制御サーバのような存在をおくことなく，システム全体の分散処理によって複数人物を追跡することが可能になった．実験結果からは，複数のステーションで同時に対象を観測する場合には，単独で観測する場合よりもより良い検出結果が得られることが確認できた．また，耐障害性の点に関しても，一部ステーションの動作停止イベントが生じてもシステム全体に影響を及ぼさないことが確認できた．

第三は，画像中の人物交差イベントや不可視領域の存在などにより，追跡中の人物の識別が不可能になる問題に対する解決である．これに対しては，問題を２種類に分類し対処した．第１のケースは，同一画像内での人物領域交差に対する問題である．画像内で複数人物の領域が交差すると，交差前の人物と交差後の人物を判別（対応付ける）ことは一般に困難な問題である．このため，追跡中の人物像からリアルタイムに特徴を抽出し，必要に応じて照合するアルゴリズムを実装した．具体的には，検出された人物領域を床面に水平な複数の領域に分割し，分割された各領域のＲＧＢ色分布を用いて判別を行う手法を提案した．動画像からの画像特徴の学習は一般に情報量が大きくなる事が問題となるが，本手法では学習空間を動的生成することで対処した．実験結果から，同系統の服装をした人物であっても，その空間的な分布を解釈することで識別が正しく行われることを確認した．一方で，床面の影や映りこみ，人物が高速に移動する場合の追跡誤差等，ノイズによる影響が大きいことも確認された．

この知見をもとに，第２のケースである複数ステーション間の対象対応付け問題に対する手法を検討した（５章）．ここでは，まず複数対象の同時追跡を可能にするためシステム構成に対して新たに検討を加えた．提案する構成では，各々の観測ステーションにおいて観測対象と１対１に対応し観測を行うプログラムである「観測エージェント」が並列に動作する．観測エージェントは与えられた対象の画像特徴を連続して取得・保存する．観測エージェントは観測ステーションをまたがって互いの観測結果を比較しあい，同様の結果を得たものが自律的にエージェントのグループ（エージェンシー）を築く．このエージェンシーがステーション間での複数対象対応付け結果となる．また，ステーション間の対象対応付けの誤りに関しては，エージェンシーを動的に形成・解消することを可能にすることで実現する．この手法では，従来の「観測ステーション＝エージェント」という考え方から，「観測ステーション上で１つの対象を観測

する存在＝エージェント」という考え方を導入した．これにより単一のステーションでも，複数の対象を観測できるように拡張することが可能になると共に，4章で取り扱った同一画像内での人物識別問題も解決することが可能になった．この対象対応付け手法は一時的な画像処理のみによって確定するのではなく長時間の観測に基づいて判断が行われるため，システムの動作中に一時的に対象の誤対応が生じた場合でも，長時間の観測によって最終的に正しい結果が得られることが実験によって確認された．

6.2 分散視覚研究の今後

本研究では，分散視覚システムの特徴として，

1．視覚機能を分散させることによる実時間観測

2．多視点観測による広域観測

3．複数の視覚センサの導入による計測性能の向上

4．分散システムによる動作の頑健性

の4点を想定しシステムを構築した．このような点は，分散視覚システムでなければ実現できない特徴であり，今後もこれらの特徴を生かしたシステムの研究が続けられるものと思われる．

多数物体の追跡・位置検出に関しては，本論文を含め西尾・大田のアルゴリズム[Nishio92]，松山らのアルゴリズム[Matsuyama98]などにより，基礎技術としてはほぼ完成されたものとなったと考えられる．今後は，これらの手法をタスクに応じて使い分けることで，実際のシステム応用へと発展させていく必要がある．そこでは，更なるシステムや観測の頑健性への配慮，距離センサや装着型センサなどの他センサと視覚センサとの統合，対象物体に関する特徴量のデータベース化とセンシングとの融合などが新たな課題になるものと思われる．

一方，分散視覚研究の中で今後新たに考えられる展開としては，イメージメディアとしての応用と，動的シーンの理解への方向性であると考える．

分散視覚を用いたイメージメディアとしての応用として有名なものが，視体積交差法を用いて自由視点からの映像生成を可能にする3次元形状復元システムである[Kitahara2000]．この手法自体は従来から提案されてきたものではあるが，実時間での形状復元は計算量の点から不可能だったため，他視点動画像を一度記録しておき，その後順次再生させながら処理することで目的となる形状データを得た．しかし，分散視覚システムを用いれば，処理の分散化を行うことによりリアルタイムで形状復元を行うことが可能になる．この取り組みは既にWxuらによってはじめられているが[Wxu2000]，実現のためには更に多くの技術的課題があるものと思われる．

本研究で取り組んだ環境中の物体の位置・移動軌跡の情報提供は分散視覚システムの応用の一例ではあるが，より高度な利用を行うためには，この情報から人間の必要とする「意味的な」情報を抽出することが必要になる．すなわち，従来のビジョン研究における重要なタスクの一つであった「静的シーンの理解」のように，分散視覚の導入による実時間性を生かした「動的シーンの理解」を行うことが考えられる．この取り組みもいくつかの例があげられるが，ビル等の高所から人物・車両などの移動軌跡を連続的に観測し，移動のパターンや「不審な動き」を抽出するもの[Stauffer99]，環境を移動する複数人物の軌跡から，通過する人物の行動パターンをいくつかの動作カテゴリに分類するもの[Hosie98]などが挙げられる．このようなタスクは，対象となる環境や目的に依存しやすい傾向があるが，よりロバストで般化された基礎技術が必要であるものと思われる．

分散視覚を取り巻く環境から見ると，計算機の処理能力，通信技術，両者に必要なコスト等，その実利用が可能になる条件が整いはじめたのはこの２，３年であり，基盤技術の開発や応用システムの開発が本格化するのもこれからであると思われる．通信の分野では，ファイルやハイパーテキストなどの静的な情報のやり取りから，ストリーミングメディア等の動的な情報の伝送への方向が急速に進んできており，その上で実行されるアプリケーションとしての分散視覚の応用も盛んになるものと思われる．

本研究は分散視覚システム応用における必要技術としての「広域環境における多物体追跡」に取り組んだものであり，この分野での基盤研究の一つとして社会に貢献することができれば幸いである．

参考文献

[Arai92] 新井民夫，太田順，“複数移動ロボット系の計画”，日本ロボット学会誌Vol. 10 No.4 pp.444-449，1992

[Anzai89] 安西祐一郎，“認識と学習”，岩波書店，1989

[Aoki99] Shigenobu Aoki，“統計学自習ノート”，http://aoki2.si.gunma-u.ac.jp/lecture/Discriminant/index.html，1999

[Asada97] Minoru Asada,Eiji Uchibe,Koh Hosoda,“Cooperative Behavior Acquisition for Mobile Robots in Dynamically Changing Real Worlds via Vision-Based Reinforcement Learning and Development”，Proceedings of First International Workshop on Cooperative Distribute

[Badler80] J.O’Rourke and N.Badler，“Model-based image analysis of human motion using constraint propagation”，IEEE Transactions on Pattern Analysis and Machine Intelligence No.2 Vol.6 pp.522-536, 1980

[Baumberg94] A. Baumberg, D.Hogg，“An efficient method for contour tracking using active shape models”，Proceedings of IEEE Workshop on Motion of Non-Rigid and Articulated Objects pp. 194-199, 1994

[Funakubo91] 舟久保登，“パターン認識”，共立出版，1991

[Gavrla99] D.M.Gavrila,“The Visual Analysis of Human Movement: A Survey”, Computer Vision and Image Understanding vol.73 no.1 pp.82-98, 1999

[Haritaoglu99] I.Haritaoglu, D.Harwood, and L.S.Davis,“Hydra: Multiple people detection and tracking using silhouettes.”, 2nd International Workshop on Visual Surveillance, pp. 6-113, 1999

[Heisele97] B.Heisele, U.Kressel, W.Ritter，“Tracking non-rigid, moving objects based on color cluster flow.”，Proceedings of IEEE Conference on Computer Vision and Pattern Recognition pp. 257-260, 1997

[Heisele98] B.Heisele, C.Woehler，“Motion-Based Recognition of Pedestrians”，Proceedings of ICPR98 pp. 1325-1329, 1998

[Hogg83] David Hogg，“Model-based vision: a program to see a walking person”，Image and Vision Computing Vol. 1 No. 1 pp. 5-20, 1983

[Hosie98] Robin Hosie, Svetha Venkatesh, Geoff West,“Classfying and Detecting Group Behavior from Visual Surveillance Data”，Proceedings of ICPR98 pp.602-604, 1998

[Inokuchi90] 井口征士，佐藤宏介，“三次元画像計測”，昭晃堂，1990

[Ishioka94] 石岡宏治,開一夫,安西祐一郎,“MARSHA:複数の自律移動ロボットの個体差を考慮した地図獲得システムの設計と実装”, 日本ロボット学会誌 Vol. 12 No.6 pp.846-856, 1994

[Johansson70] R.F. Rashid, “Towards a system for the interpretation of moving light display”, IEEE Transactions on Pattern Analysis and Machine Intelligence Vol 2 No 6 pp574-581, 1980

[Kanehara96] 金原玄, 加藤博一, 井口征士, “統計的手法を用いた犬と猫の顔画像の判別”, 第39回システム制御情報学会研究発表講演会, 1995

[Kawamura95] 河村邦明, 清水孝一, 山本克之: “痴呆性徘徊対処システムの開発ー移動体通信とGPSによる徘徊者定位の試みー”, 電子情報通信学会技術報告 MBE 94-96 pp.93-98, 1995

[Kitagawa66] 北川敏男編, “多変量解析論”, 共立出版, 1966

[Kitahara2000] 北原格, 大田友一, 金出武雄, “多視点画像の融合によるスポーツシーンの自由視点映像生成”, 画像の認識・理解シンポジウム(MIRU2000)講演論文集 I-3-8, 2000

[Konparu98] 金春利春、八木康史、谷内田正彦, “全方位視ロボットと両眼視ロボットの連携による人物の発見と注視”, 画像の認識・理解シンポジウム講演論文集 Vol.98 No.10, pp.7-12, 1998

[Leung95] M.Leung, Y.Yang, “First Sight: A human body outline labeling system.”, IEEE Transactions on Pattern Recognition and Artificial Intelligence No.17 Vol.4 pp.359-377, 1995

[Maeda92] Tomoharu Maeda, Hirokazu Kato and Seiji Inokuchi, " Image Fusion System for Object Tracking", Proc. of the Japan U.S.A. Symposium on Flexible Automation 365-368, 1992

[Maeda96] 前田茂則, 岡本真 明, 河原達也, 美濃導彦, 池田克夫, 堂下修司, “顔画像特徴, 歩行画像特徴および音声特徴の統合による個人識別”, 電子情報学会論文誌 D-II Vol.J79-D-II No.4 pp.600-607, 1996

[Matsuyama97] 松山隆司, 浅田 稔, 美濃導彦, 和田俊和, “分散協調視覚プロジェクトー分散協調視覚研究, システム開発の概要ー”, 情報処理学会研究報告 97-CVIM-103 pp.25-34, 1997

[Matsuyama98] 松山隆司, 和田俊和, 丸山昌之, “能動知覚エージェントによる移動物体の協調的追跡”, 画像の認識・理解シンポジウム(MIRU’98) pp.I-365-370, 1998

[Mills88] Mills D., “Network Time Protocol (Version 1) Specification and Implementation”, IETF RFC 1059, 1988

[Minsky86] Marvin Minsky, “The Society of Mind”, Simon & Schuster, 1986 (邦訳: 安西祐一

郎訳，“心の社会”，産業図書，1990）

[Nagao82] 長尾真編，“パターン認識と図形処理”，岩波書店，1982

[Nishio92] Shuichi Nishio, Yuichi Ohta, “Tracking of Vehicles at an Intersection by Integration of Multiple Image Sensors”, IAPR Workshop on Machine Vision Applications(MVA'92), pp.321-324, 1992

[Nordlund97] Peter Nordlund, Jan-Olof Eklundh, “Towards a Seeing Agent”, Proceedings of First International Workshop on Cooperative Distributed Vision pp. 93-120, 1997

[Ohta91] Y.Ohta, M. Watanabe, Y.Sumi, “Approaches to parallel computer vision”, IEICE Transactions E74(2) pp.417-426, 1991

[Okada97] 岡田 隆三，白井 良明，三浦 純，久野 義徳，“オプティカルフローと距離情報に基づく動物体追跡”，電子情報通信学会論文誌 D-II Vol. J80-D-II No. 6 pp. 1530-1538，1997

[Rao93] Bobby S. Rao and Hugh Durrant-Whyte, “A Decentralized Bayesian Algorithm for Identification of Tracked targets”, IEEE Trans. on Man and Cybernetics, Vol. 23, No. 6, pp. 1683-1698，1993

[Rohr94] K. Rohr, “Towards Model-Based Recognition of Human Movements in Image Sequences”, CVGIP: Image Understanding Vol. 59 No. 1 pp. 94-115, 1994

[Russell95] Stuart Russell, Peter Norvig, “Artificial Intelligence -A Modern Approach-”, Printice-Hall, 1995（邦訳：古川康一監訳，“エージェントアプローチ人工知能”，共立出版，1997

[Sato97] 佐藤明知，川田聡，大崎喜彦，山本正信，“多視点動画像からの人間動作の追跡と再構成”，電子情報通信学会論文誌 D-II Vol.J80-D-II No.6 pp.1581-1589，1997

[Segen96] Jakub Segen, Sarma (Gopal) Pingali, “A Camera-Based System for Tracking People in Real Time”, Proceedings of 13th Internatilnal Conference of Pattern Recognition (ICPR96) pp. 63-67, 1996

[Shakunaga98] Takeshi SHAKUNAGA, Yasuhiro MUKAIGAWA, Takekazu KATO, Shohei OKI, “Person Detection, Tracking, Registration and Recognition by Cooperative Distributed Vision”, Proceedings of Second International Workshop on Cooperative Distributed Vision pp. 169-186, 1998

[Shimizu96] 清水孝一，“徘徊者定位システムの開発ーバイオテレメトリによる在宅医療支援”，日本ME学会雑誌 Vol.10 No.5 pp3-10，1996

[Shioya92] 塩谷充，真田英彦，“個人識別技術の最近の研究動向”，電子情報通信学会技術研究報告 IE-92-49 pp.1-8, 1992

[Start97] Thomas M. Strat, Stephen J. Henessy, “Video Surveillance and Monitoring”, Proceedings of First International Workshop on Cooperative Distributed Vision, pp. 195-219, 1997

[Stauffer99] Chris Stauffer, "Automatic hierarchical classification using time-based co-occurrences.", Proceedings of CVPR99, 1999

[Takahashi2000] 高橋和彦，坂口竜己，大谷淳，“実時間非接触非装着型３次元人物姿勢推定法に関する一考察”，電子情報通信学会論文誌 D-II Vol.83 No.5 pp.1305-1314, 2000

[Tan97] 丹康雄，日比野靖，“位置情報ブローカの提案とそのアーキテクチャ”，情報処理学会研究報告 97-MBL-2, pp7-12，1997

[Turk91] Turk M., Pentland A.,"Eigenface for Recognition", Journal of Cognitive Neuroscience Vol.3 No.1 pp.71-86, 1991

[Ukita2000] Norimichi Ukita, Takashi Matsuyama, “Incremental Observable-Area Modeling for Cooperative Tracking”, Proceedings of 15th Internatilnal Conference on Pattern Recognition (ICPR2000) pp.192-196, 2000

[Wada98] 和田俊和，浮田宗伯，松山隆司，“視点型パンチルトズームカメラとその応用”，電子情報通信学会論文誌 D-II Vol.J81 No.6 pp.1182-1193, 1988

[Want92] R. Want, A. Hopper, V. Falcao, J. Gibbons, “The Active Badge Location System”, ACM Trans. on Information Systems Vol.10 No.1 pp.91-102, 1992

[Wren97] Christopher Richard Wren, Ali Azarbayejani, Trevor Darrell, Alex Paul Pentland, “Pfinder: Real-Time Tracking of the Human Body”, IEEE Transactions on Pattern Analysis and Machine Intelligence Vol.19 No. 7 pp.780-785, 1997

[Wxu2000] ウ小軍，東海彰吾，和田俊和，松山隆司，“PC クラスタを用いた身体動作の実時間３次元映像化”，情報処理学会研究報告 2000-CVIM-121 pp.65-72，2000

[Yanai948] Keiji Yanai, Kouichiro Deguchi, “An Architecture of Object Recognition System for Various Images Based on Multi-Agent”, Proceedings of ICPR’98 Vol I pp.278-281, 1998

謝辞

私が５年間に渡って続けた本研究の成果は，多くの方々のご協力・助言・激励の賜物であり，皆様に深く感謝の意を表します．

井口征士教授には，長期にわたる研究環境のサポート，コメント，激励をはじめ，さまざまな学会活動への参加や生活面におけるサポートを続けていただきました．本研究が充実したものとなり，また学会等においても幾分かの意義を認めていただけるまでに至ることができたのは，一重に先生のお蔭であり，深く感謝の意を表します．また，本論文に対して詳細な検討を加えていただきました．

谷内田正彦教授，西田正吾教授には，ビジョン，ヒューマンインタフェース各分野における学会活動，学内活動等の場において，長きにわたり様々なコメントをいただきました．また，本論文に対して各分野にわたり詳細な検討を加えていただきました．

佐藤宏介助教授には，研究室内発表等において本研究に対する貴重なコメントをいただきました．また学会発表時には，発表に対してさまざまなアドバイスをいただき，実のある学会活動を行うことができました．また，本論文に対して真摯なコメントをいただき，本論文がよりよいものとなるように力を尽くしていただきました．

現広島市立大学の加藤博一助教授には，本研究に対する多大なご協力をはじめ，さまざまなアドバイスや学会活動等，多くの恩恵を賜りました．また，研究者としての基本的な態度や考え方を示していただきました．先生への感謝は言葉を尽くしても表せるものではありませんが，ここに感謝の意を表します．

井口研究室助手の日浦慎作氏には，研究面における様々なアドバイスをはじめ，分散協調視覚プロジェクトに対するコミットメントの機会を与えてくださいました．また，本論文に関して詳細な検討を加えていただきました．

井口研究室助手の木村朝子氏には，長期にわたる様々なコメント，サポート，激励をいただきました．またヒューマンインタフェース分野での学会活動をはじめ，長期にわたる研究生活を，様々な面から潤いあるものにしていただきました．深く感謝いたします．

現奈良先端科学技術大学院大学の真鍋佳嗣助教授には，研究面でのサポート等様々な恩恵を賜りました．ありがとうございました．

分散協調視覚プロジェクトの皆様方には，学会活動やプロジェクトミーティングの場で多くのアドバイスをいただき，本研究をよりよいものとすることができました．お礼申し上げます．

井口研究室において長きにわたり共に研究生活を送っていただいたスタッフの方々，諸先輩方，同級生の皆様，後輩の皆様には深く感謝しております．本研究は，皆様方の研究への熱意に動かされて成り立っている部分が往々にしてあり，研究をよりよいものにすることができました．今後とも，良き伝統が続いていくことを願っております．

最後に，長期にわたる学生生活への理解・およびサポートを続けてくれた両親と妹に感謝いたします．

２００１年１月

中澤篤志

業績リスト

■ 学術論文

1. 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”，情報処理学会論文誌　第40巻 第11号 pp.4087-4096, 1999
2. 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　第41巻 第10号 pp.2895-2906, 2000
3. 中澤篤志，日浦慎作，加藤博一，井口征士，“分散視覚エージェントを用いた複数人物追跡システム”，情報処理学会論文誌（投稿中）

■ 国際会議

1. Atsushi Nakazawa, Hirokazu Kato, Seiji Inokuchi, “Human Tracking using Distributed Vision Systems.”, 14th International Conference on Pattern Recognition (ICPR '98), pp. 593-596, 1998
2. Atsushi Nakazawa, Hirokazu Kato, Shinsaku Hiura, Seiji Inokuchi, “Tracking Multiple People using Distributed Vision Systems”, The 8th IEEE International Conference on Computer Vision (Submitting).

■ 国内口頭発表

1. 中澤篤志，加藤博一，井口征士，“分散カメラシステムによる既知環境内での人物追跡”，第41回システム制御情報学会研究発表講演会，大阪，論文集 pp.275-276，1997
2. 中澤篤志，加藤博一，井口征士，“分散カメラシステムによる人物の追跡”，第42回システム制御情報学会研究発表会，大阪，論文集 pp.451-452，1998
3. 中澤篤志，加藤博一，井口征士，“分散カメラシステムによる人物の追跡”，画像の理解・認識シンポジウム（MIRU'98），岐阜，論文集 II pp.1-6，1998
4. 中澤篤志，加藤博一，井口征士，“分散カメラエージェントによる広域人物位置検出システム”，第43回システム制御情報学会研究発表講演会，京都，論文集 pp.487-488，1999
5. 中澤篤志，日浦慎作，加藤博一，井口征士，“分散観測エージェントによる複数人物の追跡”，電子情報通信学会基礎・境界講演会，広島，論文集 pp.1082，2000
6. 中澤篤志，日浦慎作，加藤博一，佐藤康介，井口征士，“分散観測エージェントによる複数人物追跡システム”，電気学会システム・制御研究会，大阪，2000
7. 中澤篤志，日浦慎作，加藤博一，佐藤康介，井口征士m“分散観測エージェントによる複数人物の追跡”，画像の認識・理解シンポジウム(MIRU2000)，講演論文集 I-15-20,

長野，2000

8．　木村朝子，中澤篤志，加藤博一，井口征士，“WWWを利用した家電製品の遠隔操作インタフェースの検討”，第41回システム制御情報学会研究発表講演会，大阪，論文集 pp.593-594，1997

9．　木村朝子，中澤篤志，加藤博一，井口征士，”WWWと赤外線リモコンによる家電製品の遠隔地制御”，第12回ヒューマン・インタフェース・シンポジウム，大阪，論文集 pp.199-204，1996

１０．　加藤博一，中澤篤志，井口征士，“イメージフュージョンによる人物追跡”，第３９回自動制御連合講演会，奈良，論文集 pp.425-426，1996

１１．　加藤博一，中澤篤志，井口征士，“イメージフュージョンによる移動体追跡とその応用”，第１３回センシングフォーラム予稿集，奈良，1996