

# 移動対象の協調的追跡のための観測可能領域モデル生成・更新法

浮田宗伯 松山隆司

京都大学大学院情報学研究科

souhaku@vision.kuee.kyoto-u.ac.jp tm@i.kyoto-u.ac.jp

## Incremental Observable-Area Modeling for Cooperative Tracking

Norimichi Ukita Takashi Matsuyama

Graduate School of Informatics, Kyoto University

あらまし: 本稿では, 複数の視覚を備えた自律システム (シーンを観測するためのアクティブカメラと, ネットワーク接続された計算機から構成される Active Vision Agent, 略して AVA) によって実時間対象追跡を行う際に, シーンの観測可能領域モデルを用いることによって, 各 AVA の協調性を向上させる手法を提案する. 複数の自律システムが協調動作を行うためには, タスクが何であったとしても, 相互に相手の能力・現在の状況に関する知識を持ち合わせる必要がある. 特に移動物体の追跡を行う際には, AVA は協調相手である他の AVA がシーン中のどの領域を観測できるかを把握し, 相互の状況を考慮した上で各 AVA 間で役割分担を行うべきである. このような協調動作を実現するために, 本稿で提案する追跡システムは, 追跡を実行する過程で獲得される各 AVA に関する観測可能領域情報を集めて, シーンの観測可能領域モデルを生成・更新する. そして, この観測可能領域モデルを用いることにより, シーンの構造と対象の動きの両方を考慮にいたした協調追跡を実現する.

## 1 はじめに

本稿では, 複数の能動視覚エージェント (Active Vision Agent, 略して AVA) [1] によって構成される実時間協調視覚システムについて述べる. AVA とは, シーンを観測するためのアクティブカメラと, 相互にネットワーク接続された計算機から構成される自律システムの論理モデルである. 複数自律視覚システムの協調動作は, 1) 広範囲観測可能, 2) 多角的情報の統合による頑健性, 3) システム構成の柔軟性, 4) 並列処理による高速性, 5) 故障に対する補償, など多くの利点を持ち合わせている. それゆえ, 文献 [2] に見られるような関連研究が, 近年数多く報告されている.

こうした分散協調視覚システムの利用により, 実世界を対象とした様々な視覚システムの実現が可能である. 中でも, 移動対象の追跡システムは, 1) 入退室管理などの広域監視システム, 2) ITS (Intelligent Transport System), 3) 対話型遠隔会議・講義システムなどの応用システム実現のための重要な基盤技術の一つである.

本稿では, 多様な実世界の状況下における移動対象の実時間追跡を研究の目的として, システム全体として効率的な対象追跡を実現するために必要な機能について述べる. 即ち, 分散協調視覚システムに基づいた対象追跡システムにおいては, 各エージェント間の協調性の優劣がシステムの性能に大きな影響を与えるた

め, 実世界で有効に機能するシステムの実現のためには, この協調性の向上が必要である. そこで本稿では, 移動対象の実時間追跡を行う際に, 各 AVA の能力 (即ち, シーン中の観測可能領域) に関する知識を AVA 間で共有することによって, エージェント間の協調性を向上させる手法を提案する. また, 提案手法による複数 AVA の協調の有効性を, 実験によって示す.

## 2 協調追跡

### 2.1 協調追跡時の AVA の動作記述

従来の我々の追跡システム [3] では, 相互にネットワーク接続された複数の AVA が一つのグループを形成して, 障害物に妨げられることなく, ある一つの対象物体を追跡し続けることが可能であった. このシステムは, 以下に示す特徴を持っている.

- 各 AVA は, 視点固定型パン・チルト・ズームカメラを備えている.
- 全カメラに関して外部パラメータ, 即ちカメラ位置の 3 次元座標はキャリブレーション済である.
- 各 AVA はユニークな ID を持っている.

各 AVA は, 視点固定型パン・チルト・ズームカメラの特性を利用することによって, 予め撮影された画像数枚から任意のパン・チルト・ズームの組み合わせで撮影される背景画像を生成可能である. それゆえ, AVA

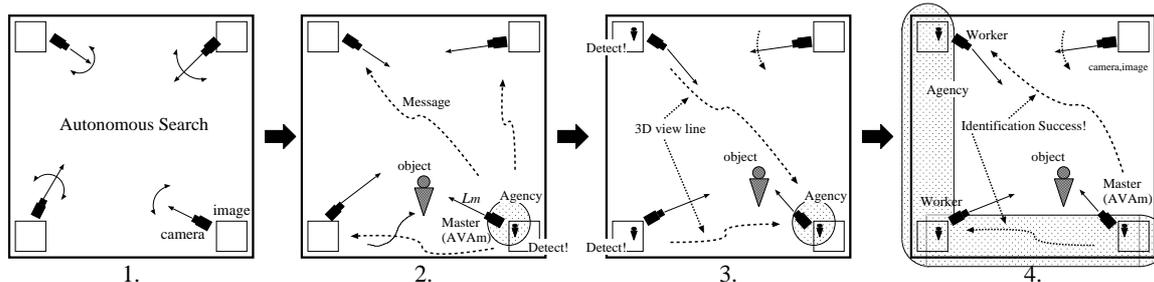


図 1: 複数 AVA による対象の協調追跡システム

は適応的にズームを変更させながら広範囲を観測する過程で、背景差分によって異常物体を検出できる。その後、検出された対象領域が画像中の中心付近、且つ適当な大きさで観測されるように、カメラのパン・チルト・ズームパラメータを制御することによって、一台のカメラによる移動対象の単独追跡が可能となる。

また、全カメラの外部パラメータが既知なので、観測画像上で検出された対象領域から、その対象の 3 次元情報 (位置・方向など) を復元することも可能である。

以上の仮定の下で、以下のようにシステムの動作 (以降、協調追跡基本方式と呼ぶ。) を設計した。

1. 全 AVA は、自律対象探索を行う (図 1 中の 1.)。
2. ある AVA (以降、 $AVA_m$ ) が異常物体を検出すると、これを追跡対象とみなす。そして、他の AVA の視線をこの追跡対象へ向けるために、この対象が存在する方向を他の AVA に対してブロードキャストする。この対象方向は、カメラの投影中心と画像中の検出領域の重心によって決定される 3 次元直線  $L_m$  によって表される。この瞬間、ある特定の対象を追跡対象と見なす AVA のグループ (エージェント) が生成される (図 1 中の 2.)。
3. ブロードキャストされた視線方向を受信した AVA は、その視線方向に沿って対象を探索する。探索の結果対象を検出した AVA (以降、 $AVA_w$ ) は、検出された対象方向を表す 3 次元直線  $L_w$  を  $AVA_m$  に対して返信する (図 1 中の 3.)。
4.  $AVA_m$  は、返信されたメッセージを受信後、 $L_m$  と  $L_w$  の間の距離を求める。そして、その距離が閾値以下であれば、 $AVA_w$  で検出された対象領域と  $AVA_m$  自身の検出領域が同じ対象を観測した領域であると判断する。こうして対象同定された全 3 次元直線からの距離の平均が最も短い点 (最近点) を求め、その座標を対象の 3 次元位置とみなす。
5.  $AVA_m$  は、同じ対象方向を注視できるように、復元した対象の 3 次元位置を  $AVA_w$  に知らせる。これを、 $AVA_m$  による視線誘導と呼ぶ。対象同定が成功した  $AVA_w$  はエージェントに組み込まれ、

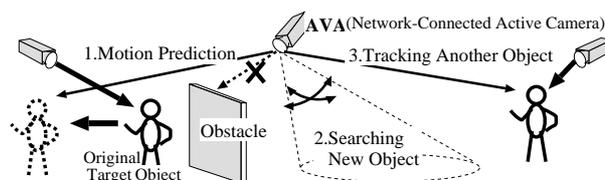


図 2: AVA の動的な役割変化

以降  $AVA_m$  からのメッセージに従って動作し、この対象の協調追跡に従事する (図 1 中の 4.)。

ある一つのエージェントにおいて、 $AVA_m$  をマスターと呼び、その他の AVA をワーカと呼ぶ (図 1 中の 4.)。マスターの持つ権限は、対象追跡の安定性の維持のために必要に応じてワーカに対して委譲される。マスター権限は、対象検出結果の信頼性<sup>1</sup>が高いワーカに対して委譲される。このワーカが新たなマスターに遷移すると同時に、先のマスターはワーカへと遷移する。

以上の動作記述に従うことにより、シーンの 3 次元情報 (障害物の位置・姿勢など) が未知であっても、全 AVA が単一対象を注視し続けることが可能である。

## 2.2 マスタによる視線誘導の問題

協調追跡基本方式では、全 AVA は、マスター AVA からの強制的な視線制御に従って単一対象を追跡し続けることが可能である。しかし、シーン全体の効率的な監視を実現するためには、障害物の遮られて対象を観測できない AVA は、自身の役割を変更するべきである。このような AVA に対する有効な役割分担として、以下の例が考えられる (図 2)。

1. 次に追跡対象が現れる位置を予測し、それが自身の観測可能領域内であればその方向へ視線を向けて待ち伏せる (図 2 中の 1.)。
2. 他どの AVA も観測していない領域を、新たな追跡対象を求めて探索する (図 2 中の 2.)。
3. 他エージェント<sup>2</sup>の追跡対象を観測可能な場合、そのエージェントに加わる (図 2 中の 3.)。

<sup>1</sup> 1) 画像中の検出領域の大きさ、2) 画像中の検出領域の重心と画像中心間の距離、3) シーン中での対象物体とカメラの間の 3 次元距離、によって決定される。

<sup>2</sup> 複数対象を追跡するシステムでは、それぞれの追跡対象に対して一つのエージェントが形成される。

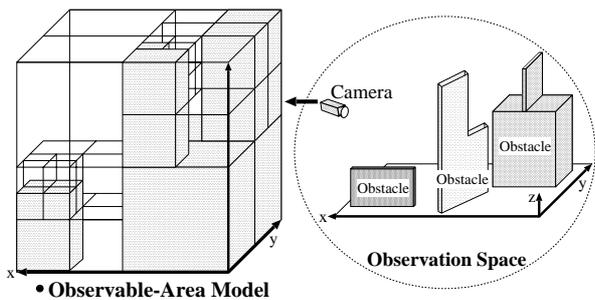


図 3: 観測可能領域モデルのデータ構造

そこで本稿では、この視線誘導問題を解決し、例に挙げた各機能を実現するために、各 AVA におけるシーン中の可視・不可視領域をそれぞれ特定し、視線誘導の際にこの可視・不可視領域情報を利用するように協調追跡基本方式を拡張したシステムを提案する。

提案システムでは、マスタは可視・不可視領域情報と現在の対象の位置を比較し、各 AVA が対象を観測できるか否かを調べた結果に基づいて、各 AVA に対して適切な役割分担を決定する。また、マスタが全 AVA に関する情報を効率的に参照できるように、全ての可視・不可視領域情報を一括してシーンのモデルとして管理する。この協調追跡のためのシーンモデルを、観測可能領域モデルと呼ぶ。

### 3 観測可能領域モデル

#### 3.1 データ構造

観測可能領域モデルのデータ構造には Octree 表現 [4] を採用した。Octree 表現では、空間全体を一つの立方体で表現している。この立方体を再帰的にオクタントと呼ばれる八分空間に分割し、各オクタントにラベルを与えることにより、空間中の各領域が持つ情報を表現する。一つの立方体の分割によって生じる各オクタントのラベルが一つでも異なる場合に、その立方体は 8 分割される。観測可能領域モデルのデータ構造に Octree 表現を用いることにより、以下の利点が得られる。

- 通常、可視・不可視領域はシーン中で集中しているので、Octree 表現によってデータの総量を減少させることが可能である。
- Octree 内の任意の場所の立方体をオクタントに分割することによって単位領域のサイズ変更が容易になり、観測可能領域モデルの解像度を局部的に設定することができる。

図 3 は、あるシーン・カメラ配置における観測可能領域モデルの例である。白色、灰色の領域が、それぞれ可視領域、不可視領域に対応している。

Octree 内の各立方体には、各 AVA 毎の可視・不可視ラベル (以降、単にラベルと呼ぶ) が代入される。代入されるラベルには、以下の三種類がある。

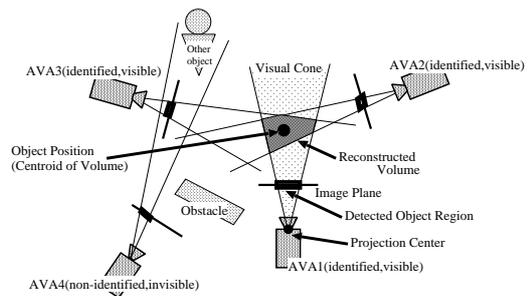


図 4: 視体積交差法による対象の体積・位置の復元と可視・不可視領域情報の生成

UNDEFINE 観測可能であるか否か不明な領域。

VISIBLE 観測可能領域。

INVISIBLE 観測不可能領域。

#### 3.2 可視・不可視領域情報の生成法

協調追跡基本方式では、マスタは検出対象方向を示す 3 次元直線間の近接関係に基づいて対象同定と対象の 3 次元位置復元を行っていた。しかし、本稿で提案する手法では、各 AVA が検出した対象領域とそのカメラの投影中心によって決定される視体積 (Visual Cone) の交わりを求め、交わりが存在するならばそれらの検出領域は同一対象の検出結果であると判断する。また、こうして求められた視体積の交わりは対象の体積を表しており (視体積交差法による 3 次元体積復元)、その体積の重心を対象の 3 次元位置と見なす (図 4)。

以上のように、協調追跡を実行中は、視体積交差法によって対象の体積が復元されている。もし、ある AVA の観測画像中の検出領域が実際に体積復元に関与しているのであれば、復元された対象の体積領域は、この AVA にとって可視であることが分かる。逆に、対象方向を観測しているにも関わらず、この体積復元に関与していない AVA にとって、この領域は不可視であるということも分かる (図 4)。

よって、システムは追跡を行う過程でマスタによって行われる対象の体積復元時に、観測可能領域モデルの生成に必要な可視・不可視領域情報を獲得できる。

#### 3.3 可視・不可視領域情報を用いた観測可能領域モデルの生成・更新

マスタは、可視・不可視領域情報を獲得すると、観測可能領域モデルを更新するために、新しい可視・不可視領域情報中の全ラベルを現在の観測可能領域モデル中の各ラベルと比較する。比較の結果、新しく生成された情報と現在のモデルの間で同じ領域中のラベルが異なっていれば、その領域に対応する観測可能領域モデル中の立方体は 8 分割される。この分割は、以下の条件が共に満たされる限り実行され続ける。

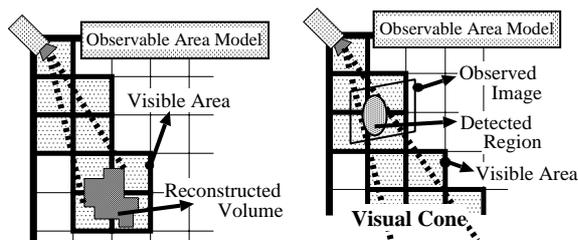


図 5: 可視領域の伝搬 (左: 対象の体積の復元成功時, 右: 対象の体積の復元失敗時)

1. 新しい可視・不可視領域情報と観測可能領域モデル中のラベルが異っている。
2.  $\frac{distance}{focallength} < \frac{constant}{2^{depth}}$ .  $distance$  はカメラと対象間の距離,  $focallength$  はカメラの焦点距離,  $depth$  は Octree でのそのオクタントの深さである。

体積の復元精度は  $distance$  と  $focallength$  に依存するので, 上記の不等式によって Octree を分割する回数定義される。

分割終了後, 以下の優先順位にしたがって各立方体に対応した可視・不可視ラベルが代入される。

INVISIBLE > VISIBLE > UNDEFINE

例えば, 既に INVISIBLE であると判定されている領域には, VISIBLE も UNDEFINE も記録されない。

また分割・代入終了後に, ある立方体が分割されてきたオクタントの全ラベルが等しくなった場合, これらのオクタントは統合されて一つの立方体に戻る。

更に, 観測可能領域モデルの生成を促進させるために, 観測可能領域モデル内で可視領域を伝搬させる。可視領域の伝搬には次の二通りが考えられる (図 5)。

対象の体積の復元成功時: まず, 復元された対象の体積領域と対応する観測可能領域モデル中の各立方体は, 可視であると判定される。この場合, 体積領域とカメラ間の領域も同様に可視であることは自明である。よって, この領域に対応する観測可能領域モデル中の立方体も可視に更新される。

対象の体積の復元失敗時: 対象の体積が復元できないので, 視体積に含まれる領域全体を仮の可視領域として更新する<sup>3</sup>。この操作によって, その時点で未確定であったが, 実際には不可視である領域を誤って可視であると判定してしまう可能性がある。そこで, この誤りをその後の更新によって訂正可能にするために, この仮の可視領域は VISIBLE, INVISIBLE ラベルによって上書き可能である。

### 3.4 観測可能領域モデルの管理

観測可能領域モデルは, 各 AVA への視線誘導時に参照されるので, この参照を行う AVA はマスタに限定

<sup>3</sup> 但し, 既に不可視であると判定されている領域は除く。

される。しかも, モデルの生成に必要な情報はマスタによって生成される。しかし, 追跡の安定性維持を目的として, マスタとなる AVA はエージェント内で動的に入れ替わる。よって, 各エージェントのマスタが独自に観測可能領域モデルを持つようにシステムを設計すると, マスタが変わるたびに AVA 間で観測可能領域モデルを送受信する必要がある。また, エージェントが複数存在する場合には, 各マスタ間で頻繁に可視・不可視領域情報を送受信する必要もある。これは, システムの即応性の維持, ネットワーク負荷の抑止という点で望ましくない。

そこで, 提案システムでは, マスタが行う

- 観測可能領域モデルの更新と一括管理。
- 観測可能領域モデルを参照した, 各 AVA への動的役割分担の案の決定。

を専用のモジュールに委任する。このモジュールを観測可能領域モデルコントローラと呼ぶ。各マスタは, 新しい可視・不可視領域情報を獲得すると, この情報を観測可能領域モデルコントローラに送信する。

観測可能領域モデルコントローラの利用により, ネットワークの負荷だけでなく, 観測可能領域モデルの管理に関する AVA の負担が削減され, AVA は協調追跡基本方式による対象注視に専念できるため, システムの安定性や即応性の向上も実現できる。

## 4 観測可能領域モデルによる協調

本章では, 1) 観測可能領域モデルの更新と, 2) このモデルを用いた各 AVA への動的な役割分担, のために必要な観測可能領域モデルコントローラと AVA 間の通信 (図 6) のプロトコルについて述べる。

エージェントから観測可能領域モデルコントローラ

各フレームにおいて, マスタ AVA は以下のメッセージを観測可能領域モデルコントローラに対して送信する。

- VISIBLE/INVISIBLE MAP: マスタが生成した可視・不可視領域情報が含まれている。これを受信した観測可能領域モデルコントローラは観測可能領域モデルを更新する。
- OBJECT POSITION: マスタが計算した追跡対象の 3 次元位置が含まれている。観測可能領域モデルコントローラは受信した対象の 3 次元位置を記録し, 対象の移動軌跡を推定する。

観測可能領域モデルコントローラは, 更新された観測可能領域モデルと対象の移動軌跡を参照して, 各 AVA に与える新たな役割分担の案を決定する。

観測可能領域モデルコントローラからエージェント

観測可能領域モデルコントローラが, ある AVA に

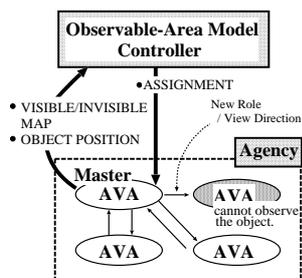


図 6: メッセージフロー

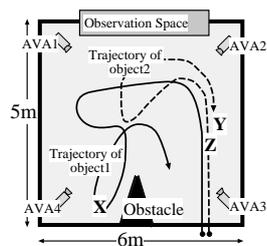


図 7: 実験環境

対して新たな役割の分担を決定した際には，次のメッセージがマスタに対して送信される．

- **ASSIGNMENT** : 1) 新しい役割が分担される AVA の ID, 2) 役割の内容, が含まれている．

マスタが ASSIGNMENT を受信し，そのメッセージが有効であると判断すれば，そのメッセージ内容に従って新たな視線誘導方向や役割分担を対象となるワークに対して送信する．このように，ワークに対する指示システムを同一エージェントのマスタに限定することにより，あるワークに対して同時に異なる命令を送ってしまう混乱を回避する．

## 5 実験結果

協調追跡に対する提案モデルの有効性を検証するための実験を行った．

実験は，図 7 に示される環境で行った．各 AVA は，PC(Pentium600MHz)と視点固定型カメラ (SONY EVI-G20) によって構成されている．加えて，全ての PC はネットワークによって接続されている．また，各カメラは床から約 2.8(m) の高さに設置されている．

上記の環境下で，まず対象 1 が観測空間に侵入して図 7 中の軌跡に沿って移動した後，地点 X において暫く停止する．対象 1 が地点 X にいる時には，AVA3 からは観測不可能であることに注意しておく．次に，対象 2 が観測空間に侵入して図 7 中の軌跡に沿って移動した後，地点 Y で停止する．この後，対象 1 が再び軌跡に沿って移動をはじめる．

AVA2 と AVA3 によって撮影された画像系列の一部を図 8 に示す．撮影画像のサイズは  $320 \times 240$  [pixel] である．画像 A1 から A14 と B1 から B14 は，それぞれ観測可能領域モデルを利用しない協調追跡 (協調追跡基本方式) の際に AVA2 と AVA3 によって撮影された画像である．一方，画像 a1 から a14 と b1 から b14 は，それぞれ観測可能領域モデルを利用した協調追跡 (提案方式) の際に AVA2 と AVA3 によって撮影された画像である．また，1) 画像中の黒線と白線で囲まれた領域は，それぞれ対象 1，対象 2 の検出結果，2) A1 と

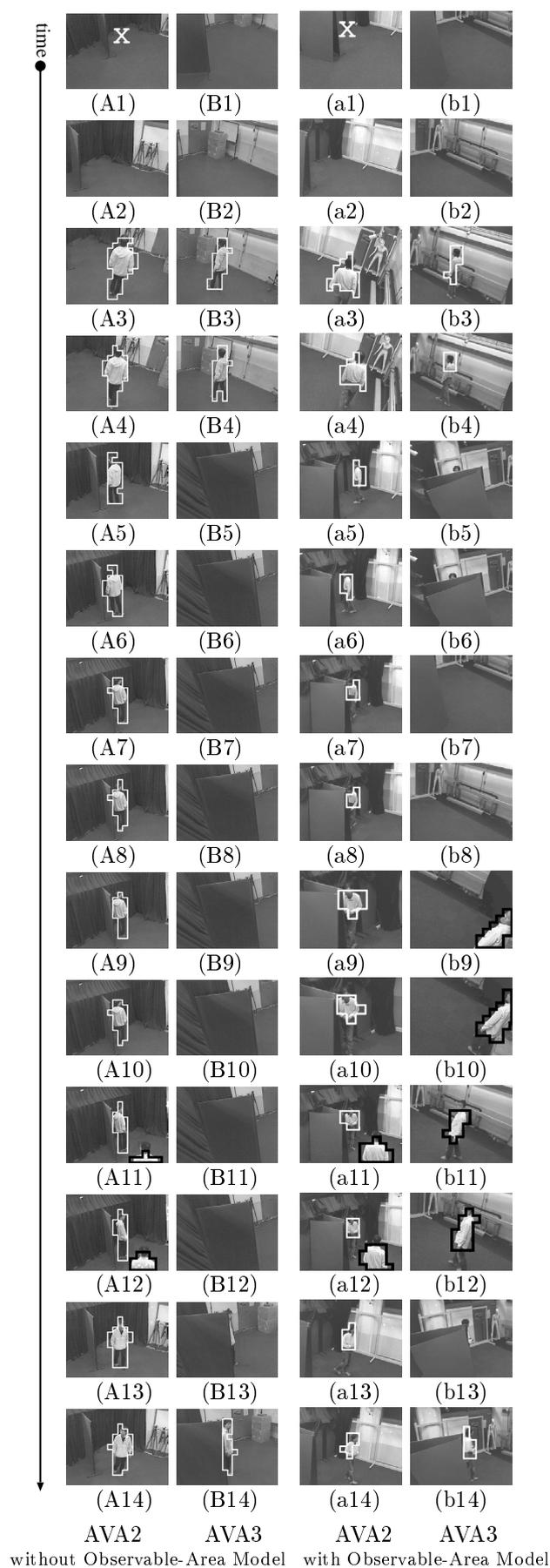


図 8: 観測画像の一部

a1 中の X は、撮影画像中の地点 X を示している。各 AVA は、平均 0.5 秒<sup>4</sup> 間隔で画像撮影を行う。

観測可能領域モデルを利用しない協調追跡では、探索 (A1,A2 と B1,B2) の後、対象 1 を検出、これを追跡対象とみなして、協調追跡を開始した (A3,A4 と B3,B4)。しかし、AVA3 は障害物に遮られて対象を観測できなくなった後も、マスタ AVA から送信される対象の存在する方向を注視し続けた (B5 ~ B13)。それゆえこの間、AVA3 は対象 1, 2 のどちらに関しても有益な情報を獲得することができなかった。一方、AVA2 は追跡対象とみなした対象 1 を追跡し続けていた (A5 ~ A14)。

観測可能領域モデルを利用した協調追跡では、探索 (a1,a2 と b1,b2) の後、観測可能領域モデルを利用しなかった場合と同様に、対象 1 を検出して協調追跡を開始した (a3,a4 と b3,b4)。しかし、観測可能領域モデルコントローラが「現在 AVA3 は対象 1 を観測不可能である」と判断すると (b5,b6)、コントローラはマスタに対してその判断結果を送信する。これを受信したマスタは、AVA3 に対して他対象の探索を新たな役割とする命令を送信した。AVA3 は、この命令に従い再度シーンの探索を開始して (b7,b8)、図 7 中の地点 Z において新たに対象 2 を検出した後、対象 2 を追跡対象として単独追跡を開始した (b9 ~ b12)。一方、AVA2 に対しては、観測可能領域モデルコントローラが保持している対象の移動軌跡を参照して、対象 1 が暫く同じ場所にとどまっていると判断された結果がマスタを経由して知らされる。このメッセージを受信した AVA2 は、高解像度の対象観測画像を得るようにズームパラメータを調整した (a9, a10)。また、対象 1 だけでなく対象 2 も観測画像中に映っている場合でも、エージェントとして追跡対象とみなしている対象 1 を注視し続けた (a11,a12)。この後、対象 1 が移動を再開し、AVA3 からも観測可能である領域に近づくと、マスタは AVA3 の視線方向を対象 1 の方向へ制御した (b13)。その結果 AVA3 は、対象 2 の単独追跡を中止して、エージェントに加わり対象 1 の追跡を再開した (b14)。

獲得された観測可能領域モデルから AVA3 に関する可視・不可視情報だけを抽出した様子を図 9 に示す。左側、右側の図は、それぞれシーンを鉛直上向き、横方向から観測した様子である。図中の P, Q は、それぞれ可視領域、不可視領域を示している。また、R の領域が障害物領域に相当することも確認できる<sup>5</sup>。

以上のように、観測可能領域モデルを利用した場合としなかった場合の二つのケースを比較することで、提

<sup>4</sup> 実験に利用したシステムで、カメラパラメータの変更、画像撮影、画像処理、通信に必要な時間である。

<sup>5</sup> 各 AVA の観測可能領域情報を統合することにより、シーンの幾何情報を獲得可能である。

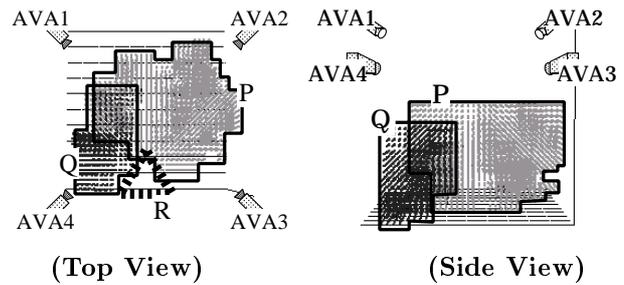


図 9: 獲得された観測可能領域モデル

案モデル有効性を確認できた。

## 6 まとめ

本稿では、シーンの観測可能領域モデルを用いた移動対象の協調的追跡を提案した。観測可能領域モデルを利用することによって、協調追跡を行っている各 AVA に適切な役割を割り当て、各 AVA の動作の協調性を向上させることが可能になる。本稿で提案したモデルは、視覚を持ち合わせたマルチエージェントシステムにあわせて設計されたものである。しかし、自律システムの協調動作のためには協調相手の能力・現在の状況に関する知識が必要であるという基本的な考えは、協調を伴う様々なタスクに応用可能である。

今後は、以下の点を中心に研究を進める予定である。

- 複数エージェントによる複数対象追跡問題への拡張を行う [5]。観測された対象の数にあわせてエージェント数を動的に増減させる必要がある。
- 複数エージェントが存在する場合、各エージェントが追跡対象を安定に追跡できるように所属する AVA を割り当てる必要がある。そこで、観測可能領域モデルを参照することにより、バランスのとれたエージェントの動的な再構成を行う。

本研究は、日本学術振興会未来開拓学術研究推進事業 (JSPS-RFTF 96P00501) の補助を受けて行った。

## 参考文献

- [1] T. Matsuyama, "Cooperative Distributed Vision - Dynamic Integration of Visual Perception, Action and Communication -", Proc. of Image Understanding Workshop, pp. 365-384, 1998.
- [2] T. Kanade, *et al.*, "Cooperative Multisensor Video Surveillance", Proc. of Image Understanding Workshop, pp. 3-10, 1997.
- [3] 松山隆司, 和田俊和, 丸山昌之, "能動視覚エージェントによる移動対象の協調的追跡", 画像の認識理解シンポジウム (MIRU'98), Vol.I, pp.365-370(1998).
- [4] N. Ahuja, J. Veenstra, "Generating Octrees from Object Silhouettes in Orthographic Views", IEEE Trans. on PAMI, Vol2, No.2, pp. 137-149, 1989.
- [5] 長尾卓, 浮田宗伯, 松山隆司, "能動視覚エージェント群による複数対象の協調的追跡", 情報処理学会研究会資料, CVIM121-8, pp.57-64(2000).