

複数の能動カメラによる移動体追尾

高見 将司[†] 宇都宮 将陽[†] 室井 基継[†] 全 炳東[‡]

† 千葉大学大学院自然科学研究科 〒263-8522 千葉市稻毛区弥生町 1-33

‡ 千葉大学総合メディア基盤センター 同上

E-mail: [†]{s.takami, p6um5504, m.muroi}@graduate.chiba-u.jp, [‡]zen.h@faculty.chiba-u.jp

あらまし パン・チルトなどが制御可能な複数の能動カメラを連携させ、広い範囲にわたって対象を追尾するための手法を提案する。各カメラはあらかじめ定義した状態遷移モデルにより自律的に動作する。各カメラ（ユニット）の追尾状況はサーバに集められ、相互に参照可能である。追尾対象の検出には、画素の色やエッジ強度などを特徴量とした AdaBoost 識別器とパーティクルフィルタを組み合わせた手法を用いる。識別器の学習は、追尾開始時に画像上で指定された検出対象によって行うが、対象のビューの変化に追従するため信頼度の高いパーティクルを用いた識別器の更新も行う。検出された対象が画像中央にくるよう、各カメラのパン・チルト・ズーム比率を制御する。一方、カメラの配置（相対的な位置関係）を事前に得ることにより、他のカメラが追尾している対象を発見・追尾できる。従って隠蔽などにより対象を見失った場合でも、他のカメラが追尾を継続していれば、再び対象を発見・追尾することも可能である。3台の能動カメラを用い、歩行者を対象とした実験を屋外でを行い、提案手法の性能を確認した。

キーワード 能動(PTZ)カメラ、自動追尾、ブースティング、パーティクルフィルタ

Object Tracking with Multi Active Cameras

Shoji Takami[†] Masaaki Utsunomiya[†] Mototsugu Muroi[†] and Heitoh Zen[‡]

† Graduate School of Science and Technology, Chiba University 1-33 Yayoi-cho, Inage-ku, Chiba,
263-8522 Japan

‡ Institute of Media and Info. Tech., Chiba University 1-33 Yayoi-cho, Inage-ku, Chiba, 263-8522 Japan

E-mail: [†]{s.takami, p6um5504, m.muroi}@graduate.chiba-u.jp, [‡]zen.h@faculty.chiba-u.jp

Abstract In this paper, a novel cooperative tracking technique that enables multiple active (PTZ) cameras to track a target autonomously in wide area. To track the target on an image plane, we combine the AdaBoost and the particle filter. After the initial learning, the classifiers are updated with the selected particles to follow the view-change of a target. Since every camera's status is gathered on the server in real-time, each camera should find the target again referring these status data, even if it loses the target. The outdoor experiments results with three cameras configuration is shown.

Keyword active (PTZ) camera, tracking, boosting, particle filter

1.はじめに

監視カメラを利用した物理セキュリティの確保は、社会的要請の高い課題であり、画像処理やコンピュータビジョン(CV)技術を駆使した実応用としてさまざまな研究開発が行われている。

監視カメラの多くは視点や視野が固定されているが、パン、チルト、ズームを遠隔で操作できる能動的カメラ(PTZカメラ)も増えつつある。こののようなカメラは、多くの場合、人が

操作することを前提に設置されている。一方、監視カメラに写った人物などを画像処理やCV技術で自動的に追尾する場合、カメラが固定なのか操作可能なのかによって、技術的には大きな違いが生じる。また監視すべき領域が広く、1台のカメラでは不足する場合には、複数のカメラを設置する必要があるが、固定カメラとPTZカメラでは、複数のカメラの設置方法や、協調／連携のさせかたも大きく異なる。

この研究では、多くのカメラを用いることで対

象の安定した追尾を行うために複数の PTZ カメラを協調動作させることを目指している。想定している環境（条件）はつぎのとおりである。

- ・ 任意のカメラで対象人物を指定されたら PTZ を制御しながら自動追尾
- ・ 他の能動カメラは対象の情報を基に追尾に加わる。
- ・ 追尾に失敗しても他のカメラの情報より追尾に復帰

すべてのカメラは、待機、追尾、探索の 3 つの状態で定義された状態遷移モデルに基づき動作を制御する。また対象人物の検出には色情報とエッジ情報を入力とした AdaBoost 識別器とパーティクルフィルタを用いる。

2. 能動カメラを用いた追尾

2.1. 対象の画像上ででの追跡

画像処理による追跡で重要な要素は、対象を特定する色や輪郭などの画像特徴と、それらより対象の位置や大きさを決定する推定法である。

画像特徴は色ヒストグラム[3]やエッジなど数多く存在するが、追跡に最適な画像特徴量を決定するのは困難である。この問題を解決するため、[4]は移動体追跡を対象と背景画素の識別問題とした。色とエッジヒストグラム[5]を特徴ベクトルとした対象と背景画素を識別する AdaBoost 識別器[2]により、対象の見え方や周囲の環境変化に対応する安定した追跡を実現した。しかし、対象の位置推定に Mean Shift 法を用いているため、対象の移動量が大きい場合や対象の隠蔽が生じると追跡が不安定になる。

そこで本研究では、AdaBoost 識別器と対象の位置、大きさの推定手法であるパーティクルフィルタを用いた追跡手法を提案する。パーティクルフィルタとは複数の仮説から対象の確率分布を推定することで対象の隠蔽に頑強である。色とエッジ情報を特徴ベクトルとし対象と背景画素を識別する AdaBoost 識別器を逐次学習し、対象の見え方や環境変化に対応する。また、AdaBoost 識別器の応答値をパーティクルフィルタの観測値に適用することで対象の隠蔽や複雑背景に対して安定した対象の追跡を行う。

2.2. 複数のカメラによる追尾

単体のカメラを用いて追尾を行う場合、対象に長時間の隠蔽が生じた場合や対象がカメラの視野範囲から離脱した場合などは追尾が失敗する可能性が高い。しかし、その対象を同時に別の視点から追尾を行っていれば対象の追尾は失敗せず追尾を持続することが可能となる。

複数の固定・能動カメラによる追尾の研究は数多く行われている[6][7]。しかし、固定カメラでは視野が限られているため、対象追尾可能な範囲が狭くなってしまう[6]。複数の能動カメラによる対象の追尾の研究の多くは、広域での多数対象の追尾を目指しており单一対象の追尾には適していない[7]。

そこで、本研究では複数の能動カメラが協調して同一の対象の追尾を行うことで追尾の安定性を向上させる。

以下、3 章で手法の概要と処理の流れ、4 章で対象の画像での検出、5 章でカメラ間の協調について詳しく述べる。

3. 状態遷移に基づく追尾

ここでは全体的な処理の流れをカメラの状態遷移に基づいて述べる。本研究ではカメラとカメラの接続された PC のことを合わせてユニットと呼ぶ。ユニット間はネットワークで接続され、サーバを介して追尾中の対象の情報のやり取りを行う。

提案手法を実装するため、本研究では各々のユニットに図 1 のように待機、追尾、探索という 3 つの状態を設定する。各ユニットはイベントに対してその 3 つの状態を遷移しながら自律的に動作する。

ユニットの初期状態は待機であり、画像で対象を指定すると追尾状態に移行する。

- 以下で各遷移とその条件について述べる。
- ① 待機→追尾：対象の指定を受ける。サーバに対象情報を送信し、追尾へ移行する。
 - ② 追尾→追尾：画像から対象を検出し、カメラの制御を行う。また、対象情報をサーバに送信する。各ユニットの対象情報が参照されると、比較をして追尾の成功判定を行う。
 - ③ 待機→探索：他ユニットが追尾を開始すると、その対象情報を基に探索へ移行する。対象情報よりカメラを制御し画像か

ら対象を探索する。

- ④ 探索→追尾：探索にて対象の検出に成功。
- ⑤ 追尾→探索：対象情報を比較して追尾失敗と判定される。探索状態へ移行する。
- ⑥ 追尾→待機・⑦探索→待機：追尾終了の指示を受ける。ユニットは動作を終了し待機状態へ移行する。

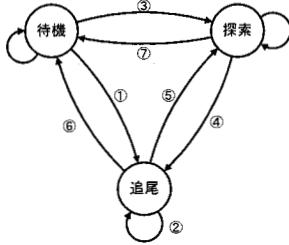


図 1 カメラの状態遷移図

4. AdaBoost 識別器とパーティクルフィルタによる追跡

4.1. AdaBoost 識別器の学習

AdaBoost は、識別性能の低い弱識別器を組み合わせ、識別性能の高い識別器を構築する教師つき機械学習アルゴリズムである[2]。以下に識別器 H の式を示す。

$$H(\mathbf{F}) = \sum_{m=1}^M \alpha_m h_m(\mathbf{F}) \quad (1)$$

ただし、 h_m は 2 クラスの弱識別器、 α_m は定数、 M は弱識別器の数、 \mathbf{F} は特徴ベクトルを示す。

AdaBoost は、学習サンプルの重みを更新しながら弱識別器を逐次学習していく。このとき識別率が高い弱識別器の重み α_m を大きくする。また誤識別した学習サンプルの重みを大きくすることで前の弱識別器が誤識別したサンプルを重点的に学習し、識別器 H の識別率を上げる。

図 2 に対象領域 R_{target} と背景領域 R_{back} から、学習サンプルクラスラベル $I_i^{(t)}$ (1: 対象、-1: 背景) を以下のように取得する。

$$I_i^{(t)} = \begin{cases} +1 & \mathbf{x}_i^{(t)} \in R_{target} \\ -1 & \mathbf{x}_i^{(t)} \in R_{back} \end{cases} \quad (2)$$

ただし、 $\mathbf{x}_i^{(t)}$ は位置 (x, y) を表わす。学習サンプル $(\mathbf{F}(\mathbf{x}_i^{(t)}), I_i^{(t)})$ より初期の AdaBoost 識別器 H_1 を構築する。ここで \mathbf{F} は画像特徴ベクトルとする。

本研究では画像特徴ベクトルとして Luv 色空間の画素の平均値と分散値、Edge Orientation Histogram(EOH)[5]を用いる。Luv 画素の平均分

散値特徴は局所領域で Luv 画素の平均値と分散値を計算する。ここで L は輝度値、 uv は色相、彩度を表す。Luv 色空間は色と輝度値が分離でき、影の影響を軽減できるため Luv 色空間を用いる。また、EOH は局所領域のエッジ強度で重み付けしたエッジ方向ヒストグラムで、対象の形状などを表現できる。EOH の階級数 9 とし、画像特徴ベクトル \mathbf{F} は 15 次元となる。



図 2 初期対象領域と背景領域

4.2. パーティクルフィルタによる対象の位置、大きさ推定

ある時刻 t の状態 \mathbf{x} を持つ N 個の粒子群を $\{\mathbf{s}_i^{(t)}\}_{i=1,\dots,N}$ とする。状態 \mathbf{x} は位置 (x, y) とし、各粒子群の尤度を $\{\pi_i^{(t)}\}_{i=1,\dots,N}$ とする。パーティクルフィルタの選択、予測、観測の各ステップに従い、対象の確率分布を表現する粒子群を以下のように逐次更新する。

(1) 選択

前時刻の粒子群 $\{\mathbf{s}_{i-1}^{(t)}\}_{i=1,\dots,N}$ から尤度の大きさに応じて新たな粒子群 $\{\mathbf{s}_{i-1}^{(t')}\}_{i=1,\dots,N}$ を選択する。

(2) 予測

次式の運動モデルに従い時刻 $t-1$ の各粒子 $\mathbf{s}_{i-1}^{(t')}$ を遷移させ時刻 t の各粒子 $\mathbf{s}_i^{(t')}$ を予測する。

$$\mathbf{s}_i^{(t')} = \mathbf{s}_{i-1}^{(t')} + \mathbf{w}_{i-1} \quad (3)$$

ただし、 \mathbf{w}_{i-1} は、ガウスノイズである。

(3) 観測

各粒子の観測値（識別器の応答値）から各粒子の尤度 $\pi_i^{(t')}$ を以下のように計算する。

$$\pi_i^{(t')} \propto \frac{1}{1 + \exp\{-a H(\mathbf{F}(\mathbf{s}_i^{(t')}))\}} \quad (4)$$

ただし、 a は定数、 H は識別器の応答値である。

対象の重心位置 $\bar{\mathbf{x}}_t = (\bar{x}_t, \bar{y}_t)$ を各粒子群の期待値として次式より求める。

$$\bar{\mathbf{x}}_t = E[\mathbf{s}_i] = \sum_{i=1}^N \pi_i^{(t')} \mathbf{s}_i^{(t')} \quad (5)$$

また、対象の大きさは対象と分類された粒子の広がりで決まるとして、対象の重心位置 $\bar{\mathbf{x}}_t$ に対する各粒子の尤度を重みとした重み付き共分散行列 Σ_t' を用いたマハラノビス距離で表現する。

以下に各粒子 $s_i^{(t)}$ の重心位置 \bar{x}_i に対するマハラノビス距離 $D_i^{(t)}$ を示す。

$$D_i^{(t)} = \left\{ (\mathbf{s}_i^{(t)} - \bar{\mathbf{x}}_i)^T \Sigma_i^{-1} (\mathbf{s}_i^{(t)} - \bar{\mathbf{x}}_i) \right\}^{\frac{1}{2}} \quad (6)$$

となる。ただし、 Σ_i^{-1} は Σ_i の逆行列である。マハラノビス距離が等しい点を結ぶと楕円を描き、楕円の大きさはサンプルの信頼区間となる。対象楕円は、粒子群 80% 含む $D \approx 1.40$ とし、背景楕円はすべての粒子群を含む $D=10$ とした。図 3 に対象楕円と背景楕円を示す。また、図中の白点は識別器の応答値が正（対象）の粒子、黒点は応答値が負（背景）の粒子を表す。



図 3 対象楕円と背景楕円

4.3. 対象の信頼度

全粒子の尤度の空間的な重み付け平均値を対象の信頼度として提案する。以下に対象の信頼度 c_i の定義式を示す。

$$c_i = \sum_{i=1}^N w_i^{(t)} \pi_i^{(t)} \quad (7)$$

ただし、 $w_i^{(t)}$ は各粒子の空間的重みである。図 4 に各粒子の空間的重みと尤度を示す。

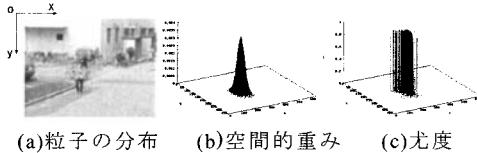


図 4 各粒子の空間的重みと尤度

4.4. AdaBoost 識別器の更新

追跡に用いた各粒子を利用し AdaBoost 識別器を更新する。図 3 の対象楕円内の領域を R_{target} 、対象楕円外でかつ背景楕円内の領域を R_{back} とする。以式に学習サンプルクラスラベル $l_i^{(t)}$ を示す。

$$l_i^{(t)} = \begin{cases} +1 & \mathbf{s}_i^{(t)} \in R_{target} \cap H(\mathbf{F}(\mathbf{s}_i^{(t)})) > 0 \\ -1 & \mathbf{s}_i^{(t)} \in R_{back} \end{cases} \quad (8)$$

このように外れ値を除いた学習サンプル $(\mathbf{F}(\mathbf{s}_i^{(t)}) = \mathbf{x}_i^{(t)}, l_i^{(t)})$ を取得する。さらに更新前の識別器で誤識別した学習サンプルを重点的に学習す

るような重みを与える。この学習サンプルより各時刻 t に K 個の弱識別器 h_i を追加学習する[3]。

過剰更新を防ぐため、初期の識別器 H_i と、更新による識別器 H_i を統合する。以下に H_i 、 H_i の式をそれぞれ示す。

$$H_i(\mathbf{F}) = \sum_{m=1}^M \alpha_{i,m} h_{i,m}(\mathbf{F}) \quad (9)$$

$$H_i(\mathbf{F}) = \begin{cases} \sum_{t=2}^T \left(\sum_{k=1}^K \alpha_{i,k} h_{i,k}(\mathbf{F}) \right) & (t \leq T) \\ \sum_{t=T}^i \left(\sum_{k=1}^K \alpha_{i,k} h_{i,k}(\mathbf{F}) \right) & (t > T) \end{cases} \quad (10)$$

ただし、 $h_{i,m}$ は $t=1$ の弱識別器、 T は更新識別器の履歴数、 $h_{i,k}$ は $t>1$ の弱識別器、 c_i は対象の信頼度である。識別器 H_i 、 H_i を以下のように統合し識別器 H を更新する。

$$H(\mathbf{F}) = \beta_i H_i(\mathbf{F}) + \beta_i H_i(\mathbf{F}) \quad (11)$$

ただし、 β_i 、 β_i はそれぞれ H_i 、 H_i の学習サンプルにおける識別率による重みである。

5. カメラ間の協調

ここでは、カメラ間の協調として対象の探索と追尾の成功判定について述べる。

5.1. カメラ座標系

本システムにおいて、問題を記述するために以下の座標系を定義する（図 5 参照）。

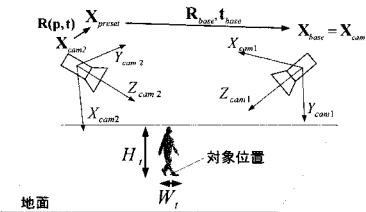


図 5 座標間の関係

カメラはカメラ座標系 \mathbf{X}_{cam} において、 Y_c 軸まわりの回転(パン)、 X_c 軸まわりの回転(チルト)を行うものとする。カメラのパン p ・チルト t による回転 $\mathbf{R}(p,t)$ からプリセット状態 ($p=0, t=0$) でのカメラ座標系 (プリセット座標) \mathbf{X}_{preset} への変換は次式となる。

$$\mathbf{X}_{preset} = \mathbf{R}(p,t)\mathbf{X}_{cam} \quad (12)$$

各カメラのうち 1 つのカメラのプリセット座標を基準座標として扱う。カメラ間で対象の位置情報のやり取りは基準座標にて行う。プリセット座標から基準座標への変換をカメラ間の回転 \mathbf{R}_{base} と並進 \mathbf{t}_{base} を用いて次式で表す。

$$\mathbf{X}_{base} = \mathbf{R}_{base} \mathbf{X}_{present} + \mathbf{t}_{base} \quad (13)$$

対象は平面上に存在するものとし、また対象の高さを H_i 、幅を W_i と仮定する。

5.2. 対象の位置の算出

対象の位置の算出方法は、追尾状態のユニット数が単体か複数かによって分けられる。

(1) 単体のユニットが追尾

検出された対象の画像上での位置を (x, y) 、カメラの焦点距離を f とすると、対象のカメラからの観測方位は $\mathbf{X}_{direction} = (x/f, y/f, f)^T$ となる。この観測方位と地面の交点を対象の位置とする。しかし、対象の足元を正確に検出するとは限らない。そこで、図 6 のように対象の高さ平面との交点を加えた範囲を対象の位置とする。

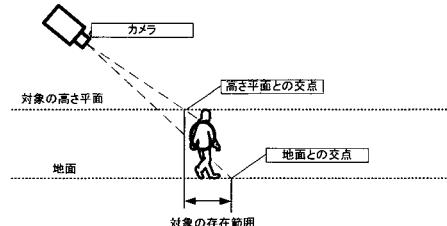


図 6 対象の位置の算出

(2) 複数のユニットが追尾

複数のユニットが追尾を行っている場合、各ユニットの対象の観測方位の最短点を対象の位置とする。予測される対象の位置との距離と対象の信頼度 c_i を重みとして最小二乗法を解くことで誤った追尾による誤差の影響を軽減する。

5.3. 対象の探索

対象の探索はカメラの制御と画像上での探索の 2 段階で行われる。まず、参照される対象の情報より対象が適切な大きさで中心に映るようパン・チルト・ズームを制御する。

対象の画像上での探索には追尾と同様の手法を用いる。AdaBoost 識別器には追尾を開始したユニットの初期識別器 H_i を用いる。パーティクルフィルタの粒子を分布する範囲は、対象の位置・高さ・幅を基に画像に投影される対象の領域に分布する。

5.4. 追尾の成功判定

追尾の成功判定は複数ユニットによる対象

の位置とユニット自身の対象の位置を比較することで行う。比較は対象の方向と距離の 2 つで行い、比較によって算出される対象の追尾度 $\alpha_{track}(t)$ によって追尾の成功判定を行う。

(1) 対象の方向での比較

複数ユニットによる対象を対象の位置・高さ・幅を基に画像に投影して比較する。

投影された対象の大きさを D 、ユニット自身の対象の位置と投影された対象の位置との距離を d とする。方向の追尾度 $\alpha_{direction}$ を次式で表す。

$$\alpha_{direction} = \begin{cases} k_{direction} \left(1 - \frac{d}{D} \right) & (\alpha_{direction} > -k_{direction}) \\ -k_{direction} & (\alpha_{direction} \leq -k_{direction}) \end{cases} \quad (14)$$

ここで $k_{direction}$ は係数である。

(2) 対象の距離での比較

複数ユニットによる対象の位置とユニット自身の対象の位置を比較する。

対象の位置から一定の範囲を求め、各ユニットの平均範囲 $R_{distance}$ を求める。複数ユニットによる対象位置とユニット自身の対象位置の距離を $r_{distance}$ とする。距離の追尾度 $\alpha_{distance}$ を次式で表す。

$$\alpha_{distance} = \begin{cases} k_{distance} \left(1 - \frac{r_{distance}}{R_{distance}} \right) & (\alpha_{distance} > -k_{distance}) \\ -k_{distance} & (\alpha_{distance} \leq -k_{distance}) \end{cases} \quad (15)$$

ここで $k_{distance}$ は係数である。

方向と距離の追尾度を加えた対象の追尾度 $\alpha_{track}(t)$ は次式で算出される。

$$\alpha_{track}(t) = \alpha_{track}(t-1) + \alpha_{direction} + \alpha_{distance} \quad (16)$$

$\alpha_{track} \leq 0$ ならユニットは探索へと移行する。

6. 実験

提案手法を用いて実験を行った。今回の実験では、追尾中のユニットが追尾に失敗するが、その後追尾に復帰する。

実験でのカメラ間の位置・姿勢は弱キャリブレーション[8]を用いて算出を行った。

6.1. 実験環境

図 7 のようにカメラを設置し実験を行った。カメラ高は 2.3m、観測画像は 640×480pixel の画像である。対象人物は点 A から点 C まで移動し、途中点 B でユニット 2 に隠れが生じる。

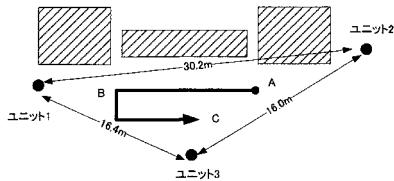


図7 実験環境

6.2. 実験結果

実験で対象を追尾する様子を図8に示す。まず、ユニット3にて手動で対象の指定が行われ追尾を開始する。ユニット1・2は対象情報を基に探索を行い、追尾状態へと移行する。130フレームにてユニット2に対象の隠蔽が生じて追尾に失敗する。ユニット2は追尾から探索へ移行する。175フレームで隠蔽から回復したユニット2は対象を検出し、追尾に復帰した。

実験より、複数カメラが協調することで長時間の隠蔽に対して安定性が増すことが確認された。ユニット単体での追尾と複数での追尾について以下にまとめると

(1) 単体ユニットでの追尾

パーティクルフィルタを用いて対象の検出を行うことにより一部分の隠蔽や短い間で隠蔽が発生した場合には対象を見失うことなく追尾を継続することができた。一方、長時間隠蔽が続いた場合、追尾失敗から復帰できなかった。

(2) 複数ユニットでの追尾

複数のユニットが互いに協調することで対象の追尾に失敗したユニットも自動的に追尾に復帰することが確認された。

7. おわりに

本論文では、複数の能動カメラの協調により追尾の安定性を向上する手法を提案し、実験によりその実現可能性を確認した。

しかし、追尾中のユニットの多くが誤った対象を追尾した場合、正しい追尾を行っているユニットが誤って追尾失敗と判定されたために追尾に失敗する。

今後、追尾の成功判定の改良を行う必要がある。追尾の成功判定に対象の幾何学以外の情報を加える事などを検討している。

謝辞 本研究を進めるにあたり、小糸工業株式会社の須藤聰一様には御助言と御協力いただき

きましたことを心から御礼申しあげます。



図8 追尾結果

文 献

- [1] M. Isard and A. Blake: "CONDENSATION-Conditional Density Propagation for Visual Tracking," *International Journal of Computer Vision*, vol.29, no.1, pp.5-28, 1998.
- [2] J. Friedman, T. Hastie, and R. Tibshirani, "Additive Logistic Regression: a Statistical View of Boosting," *technical report, Department of Statistics, Stanford University*, 1998.
- [3] M. Muroi and Heitoh Zen, "Human Tracking Based on Particle Filter in Outdoor Scene", *MVA*, pp.461-464, 2007.
- [4] S. Avidan, "Ensemble Tracking," *CVPR*, vol.2, pp.494-501, 2005.
- [5] K. Levi and Y. Weiss, "Learning Object Detection from a Small Number of Examples: The importance of good features," *CVPR*, vol.2, pp.53-60, 2004.
- [6] Wei Du, Justus Piater "Multi-camera People Tracking by Collaborative Particle Filters and Principal Axis-Based Integration" *ACCV 2007*, Part I, pp.365-374, 2007
- [7] 浮田 宗伯, "能動視覚エージェント群の密な情報交換による多数対象の実時間協調追跡", *電子情報通信学会論文誌 D-I Vol. J88-D-I No.9* pp.1438-1447, 2005
- [8] R. Hartley et al.: "Multiple View Geometry in Computer Vision," CAMBRIDGE UNIVERSITY PRESS, 2000.