

オブジェクト抽出技術のスポーツ番組への応用

高橋 正樹[†] 三須 俊彦[†] 合志 清一[†]

[†] NHK 放送技術研究所 〒157-8510 東京都世田谷区砧 1-10-11

E-mail: [†] takahashi.m-iu@nhk.or.jp

あらまし 放送コンテンツの中でもスポーツ中継は視聴率が高い。スポーツ中継は完成の域に達しているように見えるが、新たな映像表現を用いればより多くの視聴者へさらにわかりやすい映像を供給できる可能性がある。映像内の個々の物体（オブジェクト）の抽出・追跡技術を利用することにより、オブジェクトの特徴量を利用した様々な映像表現の生成が可能となる。本稿では検出処理と幾何変換・作画処理を組み合わせた、スポーツシーンにおける新たな映像表現生成手法について述べる。また高速に移動するオブジェクトに対応したオブジェクト抽出・追跡法について述べる。さらに、上記手法を利用したスポーツコーダシステムを構築し、これを実験した結果仕様通りの映像が生成されることが確認できたのでこれを報告する。

キーワード オブジェクト, 抽出, Kalman フィルタ

An Application of the Object extraction technology for Sportscast

Masaki TAKAHASHI[†] Toshihiko MISU[†] and Seiichi GOHSHI[†]

[†] NHK Science and Technical Research Laboratories 1-10-11

Kinuta, Setagaya-ku, Tokyo, 157-8510 Japan

E-mail: [†] takahashi.m-iu@nhk.or.jp

Abstract A Computer-generated picture, when superimposed on a real image, would be powerful aids for TV viewers to enjoy TV programs. Using object extraction and tracking technologies, we have developed several new visual effects on sportscast. A sportscast is one of the most popular TV programs. The paper describes a technique of the extracting and tracking fast moving objects. It also describes a technique for combining processors for motion tracking and real time CG rendering. We report the CG effects on golf scenes generated by our proposed system.

Keyword Object, extraction, Kalman filter

1. はじめに

スポーツ中継では多種多様な映像制作が行われている。中でも実映像へ情報をスーパーインポーズした映像は、臨場感を損なわずに番組を分かりやすくする有効な映像表現である。

これまで選手名や経過時間等をスーパーインポーズする手法や、映像内の物体の動きに連動して仮想の図形や文字を作画する手法等が実用化されている。また映像内の動オブジェクト抽出に関して多数の研究が行われている[1],[2]。

しかし、ボールのように映像内に占める面積が小さく、高速で移動するオブジェクトを抽出・追跡することは一般に困難であり、ボールの特徴量を用いた映像制作はあまりされていない。本手法ではボールの様々な特性を利用した抽出処理を施すことでボールの抽出・追跡を実現した。また飛距離等の特徴量を自動的

に算出した。

近年、画面内のオブジェクトに張り付くスーパー画像等が放送されているが、これらの多くは高価な機器を用いたオフライン処理で制作されている。安価な装置を用いたリアルタイムで使用可能なスーパー映像作画システムが求められている。一方、近年のパーソナルコンピュータ（PC）の能力の向上は目覚しく、従来は不可能と考えられていた処理も可能となってきた。

我々は複数台のPCを用いた分散処理により、高速移動オブジェクトの抽出・追跡、及び映像内の物体の動きに連動した図形・文字のリアルタイム作画を実現した。

また開発したシステムを利用し、ゴルフ中継のシミュレート映像をもとに実験を行った結果、有効性が確認されたので報告する。

2. 提案システム

開発したシステムは、以下の機能を有する。

1. 高速移動物体の抽出と追跡
2. 放送カメラ映像内の物体の動きに連動した図形・文字のスーパーインポーズ

以下ではこれらの詳細について述べる。

2.1 システム構成

図1に提案手法に基づくシステムの概要を示す。

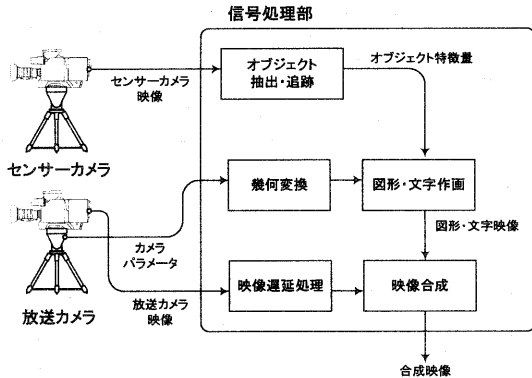


図1 提案手法の概要

抽出すべき映像上の対象物をオブジェクトと呼ぶ。図1に示すシステムはセンサカメラ、放送映像カメラ、信号処理部よりなる。センサカメラにて全景映像の撮影を行い、オブジェクトの抽出・追跡処理を行うための映像データを作成する。放送カメラにて放送用映像を撮影する。また放送カメラのパン・チルト角、ズーム量（カメラパラメータ）を計測する。信号処理部では、センサカメラ映像から抽出したオブジェクトの特徴量と、放送カメラ映像内の物体の動きに連動した図形・文字映像を作画する。作画した図形・文字映像と放送カメラ映像の同期をとり、合成して出力する。

2.2 信号処理手順

信号処理部の各処理の概要を示す。

- (1) オブジェクト抽出・追跡部では、センサカメラ映像を分析し、オブジェクトの抽出・追跡処理を行う。またオブジェクトの位置座標や面積等の特徴量を生成する。
- (2) 幾何変換部では、放送カメラ映像内の物体の動きに連動した、図形の幾何変換処理を行う。
- (3) 図形・文字作画部では、オブジェクト特徴量と幾何変換した図形のスーパー映像を作画する。
- (4) 上記(1)～(3)の処理により、図形・文字映像は放

送カメラ映像に対して遅延時間が生じる。映像遅延処理部では、放送カメラ映像に遅延時間を設け、2つの映像の同期をとる。

(5) 映像合成部では、放送カメラ映像と図形・文字映像を合成し、合成映像を出力する。

オブジェクト抽出・追跡手法の詳細については3で述べる。

2.3 システムの特長

本システムは以下の特長を有する。

(1) 映像の新規性

高速移動物体の特徴量、及び放送カメラ映像内の物体の動きに連動した、図形・文字スーパー映像の作画が可能。

(2) 低コスト性

画像・カメラパラメータ分析、作画処理等の複数の処理をPCでまかなうことが可能なため、必要となる機材が少ない。

(3) 高速性

分散処理、及び探索範囲を限定した高速オブジェクト抽出手法によりリアルタイム作画が可能。

(4) 操作性

本手法ではオブジェクトの自動抽出・追跡が可能となるため、ポインティングデバイスなどによる手動操作が不要。

3. オブジェクト抽出・追跡手法

オブジェクト抽出・追跡手法は、センサカメラの映像を分析し、ボールを抽出・追跡する。フレーム毎にボールの位置座標、面積等の特徴量を生成する。オブジェクト抽出・追跡手法のブロック図を図2に示す。以下では、“オブジェクト候補の抽出”、“オブジェクト選定”、“位置予測・探索範囲決定”の3ステップを順に説明する。

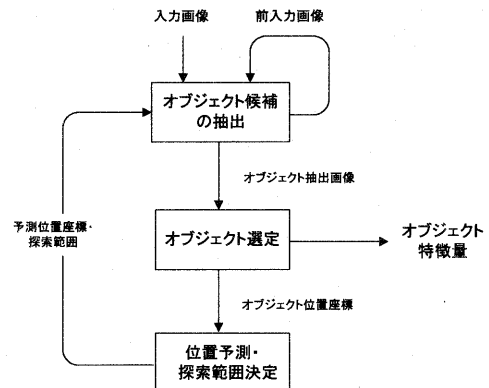


図2 オブジェクト抽出・追跡手法

3.1. オブジェクト候補の抽出

オブジェクト候補の抽出ステップでは、センサーカメラからの現フレーム画像及び前フレーム画像を分析し、背景部分とオブジェクト部分を分離した、オブジェクト抽出画像を作成する。図3にオブジェクト候補の抽出フローを示す。

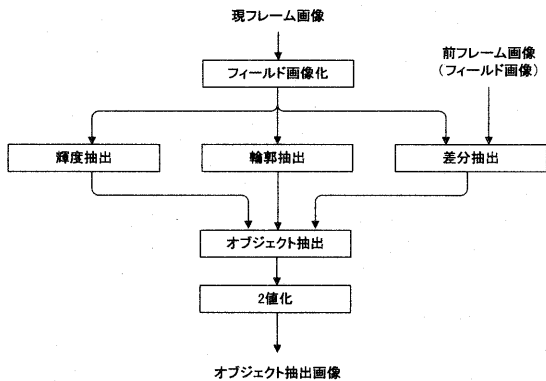


図3 オブジェクト候補の抽出フロー

始めに、センサーカメラからの720*480解像度フレーム画像を720*240解像度フィールド画像へ変換する。次にボールの予測位置・探索範囲に基づき、限定した画像範囲に対して輝度抽出、輪郭抽出、差分抽出処理を行う。ボールの位置予測・探索範囲決定手法に関しては3.3にて述べる。入力画像の特性に合わせて輝度画像、輪郭画像、差分画像のいずれか、もしくは複数の画像の組み合わせを利用する。その後、閾値により2値化処理を施す。

図4にフィールド画像化した原画像と、全画面を探索範囲とした輪郭抽出によるオブジェクト抽出画像を示す。ここではボール以外のオブジェクトも抽出されている。この中からボールを選定する手法については次節にて述べる。

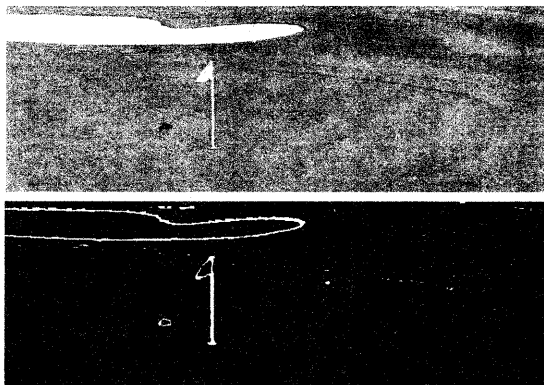


図4 オブジェクト抽出画像の例

3.2. オブジェクト選定

オブジェクト選定ステップでは、オブジェクト抽出画像に含まれる複数のオブジェクトの中からボールを選定し、位置座標や面積等の特徴量を出力する。図5にオブジェクト選定のフローを示す。

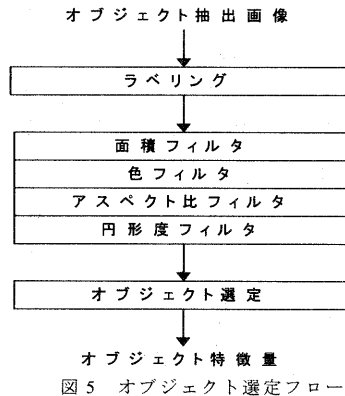


図5 オブジェクト選定フロー

画像に含まれる全てのオブジェクトにラベルを付与し、抽出すべきボールの候補とする。面積、色、アスペクト比、円形度フィルタを適用し、候補となるオブジェクトを絞る。各フィルタについて述べる。

面積フィルタは、面積の閾値処理により閾値外の面積を持つオブジェクトを候補から外す。閾値は前フレームでの抽出オブジェクト面積に応じて動的に値を定める。色フィルタは、オブジェクトの色の平均値、またはベースとなる背景色との明るさや色合いの変化度を求め、閾値外のオブジェクトを候補から外す。アスペクト比フィルタは、オブジェクトの縦方向の最大長と横方向の最大長の商に対して閾値処理を施し、閾値外のオブジェクトを候補から外す。円形度フィルタは、特定範囲外の円形度を持つオブジェクトを候補から外す。円形度は以下の式で表される[3]。

$$e = \frac{4\pi S}{l^2}$$

ここでSをオブジェクトの面積、lを周囲長、πを円周率、eを求める円形度とする。

全てのフィルタに適合したオブジェクトのうち、前回選定したオブジェクトの位置から最も近距離にあるものをボールとして選定する。選定したボールの位置座標、面積、飛距離等の特徴量として出力する。オクルージョン等の影響により抽出に失敗した場合、前回のオブジェクト特徴量を引き続き出力する。

本手法まででオブジェクト抽出は可能であるが、動画面上での追跡精度及び処理時間を向上させるため、探索範囲を限定する手法を用いた。

3.3. 位置予測・探索範囲決定

位置予測・探索範囲決定ステップでは、高精度且つ高速な追跡処理を実現するため、次フレームでのボール位置座標予測と探索範囲の限定を行う。

本手法では Kalman フィルタを用いている。推定すべき次フレームでの画像座標(x,y)を含む \mathbf{x} を推定する状態 (パラメータ) とする。

$$\mathbf{x} = [x \ y \ v_x \ v_y]^T$$

ここで v_x, v_y はそれぞれ x, y 方向の速度を表す。

観測の感度行列 H , 観測共分散行列 R , プロセス共分散行列 Q を以下のようにおく。

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad R = \begin{bmatrix} R_x & 0 \\ 0 & R_y \end{bmatrix}$$

$$Q = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & Q_x & 0 \\ 0 & 0 & 0 & Q_y \end{bmatrix}$$

R_x, R_y, Q_x, Q_y はそれぞれの行列の初期値とする。

現フレーム (時刻 t) にてボールの抽出に成功した際は、ボールの位置座標を観測値 $\mathbf{y} (x, y)$ として状態更新式

$$\mathbf{x}_t = \mathbf{x}_{t-1} + K(\mathbf{y} - H\mathbf{x}_{t-1})$$

$$P_t = (I - KH)P_{t-1}$$

$$K = P_{t-1}H^T(H P_{t-1}H^T + R)^{-1}$$

に当てはめ、Kalman ゲイン K , 状態 \mathbf{x} , 状態共分散 P を更新する。

またフレーム毎に状態予測式

$$\mathbf{x}_t = F\mathbf{x}_{t-1}$$

$$P_t = F P_{t-1} F^T + Q$$

に当てはめ、次フレームにおける \mathbf{x} と P を決定する。

F は状態遷移行列である。本手法ではボールの等速直線運動を想定して追跡を行ったため、以下のように定めた。

$$F = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

また追跡を始める際に \mathbf{x} の初期座標と初期探索範囲を与える必要がある。初期探索範囲を全画面とし、最初に抽出したボールの座標を \mathbf{x} の初期値とする。もしくは初期探索範囲を小さめに取り、ポインティング

デバイスで初期座標を指定することが可能である。

Kalman フィルタの更新の際、観測値 \mathbf{y} と予測値 \mathbf{x} の誤差の大小に応じて状態共分散 P の制御を行う。これにより、オクルージョン等の影響でオブジェクトを見失ったときには状態予測による伝搬、オブジェクトを抽出した際には観測重視という適応的な追跡が可能となる。また抽出に成功したときには共分散を小さく、失敗したときには共分散を大きくとる。この共分散の変化をオブジェクト探索範囲に対応させることで時々刻々と最適な探索範囲へと変化させることが可能となる。

4. 座標系の変換

放送カメラの映像から物体を抽出してその位置座標を示し、映像に連動した図形画像を作画するためには汎用的な手法を用いた方が計算も容易である。また、他の目的への応用も実現しやすい。このため、ワールド座標を導入した。ワールド座標とは 3 次元空間に設置した一般的なデカルト座標である。

4.1. ワールド座標系への変換

本システムを放送現場で使用する際には、カメラやオブジェクトの相対的な位置を実測し、測定値をワールド座標系に変換する。実測はカメラを基準点として行うので (以下カメラ座標系と呼ぶ)、カメラ座標系からワールド座標系の変換が必要となる。

図 6 にカメラ座標系とワールド座標系の関係を示す。

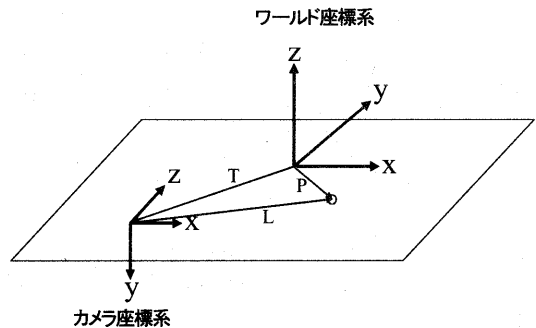


図 6 カメラ座標系とワールド座標系

ここでパン、チルト、ロールをそれぞれ α, δ, ϕ とし、基準姿勢時のカメラ座標系からワールド座標系への姿勢変換行列を S とおく。また (W) 及び (C) でワールド座標系とカメラ座標系を区別する。カメラ座標系からワールド座標系への姿勢変換行列 $R_c^{(W)}$ は次式で表される。

$$R_c^{(W)} = S R_y(\alpha) R_x(\delta) R_z(\phi)$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} \cos \alpha & 0 & \sin \alpha \\ 0 & 1 & 0 \\ -\sin \alpha & 0 & \cos \alpha \end{bmatrix} \\ \times \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \delta & -\sin \delta \\ 0 & \sin \delta & \cos \delta \end{bmatrix} \begin{bmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

カメラ座標系 (x_c, y_c, f) の位置にあるオブジェクトの世界座標系での位置 $P(x_w, y_w, z_w)$ を決定する。 f はカメラの焦点距離を表す。

カメラ座標系におけるオブジェクトの3次元座標は (ax_c, ay_c, af) と表せ (a は未知の係数)、以下の変換式が成立する。

$$P^{(w)} = R_C^{(w)} \begin{bmatrix} ax_c \\ ay_c \\ af \end{bmatrix} + T^{(w)}$$

ここで T は世界座標系の原点からカメラへ至るベクトルである。 a が未知であるため、 P を一意に定めることが不可能であるが、本手法ではオブジェクトが一つの平面内、すなわち $z_w=0$ とすることで P を決定している。

オブジェクトの世界座標系での位置座標が求まれば、オブジェクトから世界座標系の特定点までの距離を求めることが可能となる。

4.2. カメラ座標系への変換

カメラ映像内の物体の動きに連動して図形を作画する場合、世界座標系にある図形をカメラパラメータに応じてカメラ座標系へ変換する。

世界座標系からカメラ座標系への変換は、カメラ座標系から世界座標系への変換のほぼ逆演算により表現できる。

5. 実験

提案手法の有効性を確認するため、ゴルフ中継を模した実験を行った。ボール落下地点付近に設置したセンサーカメラの映像からボールを抽出・追跡し、飛距離を文字及び直線にて表示した。また放送カメラの動きに連動した距離線を表示した。

5.1. 概要

撮影にはセンサー・放送カメラ共に HD カメラを用いた。図7にカメラの設置状況を示す。

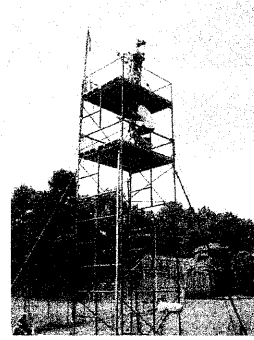


図7 カメラ設置状況

センサーカメラは地上から約 4.5m の高さに設置。ボールの落下地点を広くカバーできる画角で固定撮影した。オブジェクト抽出・追跡用 PC へ入力する際にはダウンコンバートして 720*480 の解像度に変換している。本実験においては入力画像のフレーム差分画像を抽出用画像としてボールの抽出・追跡を行った。抽出したボール位置座標を世界座標に変換し、世界座標系の特定点までの距離を計算して幾何変換及び図形・文字作画用 PC へ送信した。

放送カメラは地上から約 6m の高さにカメラパラメータ計測機能付雲台とレンズエンコーダを取り付けて設置。実際のゴルフ中継と同様のカメラワークを行った。パン・チルト角、ズーム量をフレーム毎に幾何変換及び図形・文字作画用 PC へ送信した。

幾何変換及び図形・文字作画用 PC ではボールの飛距離文字スーパーと距離線の図形スーパーを作画した。映像遅延装置によりスーパー映像と放送カメラ映像の同期をとり、映像合成装置にて合成、出力した。

5.2. 結果

本実験では以下の2シーンについての実験を行った。いずれも初期探索範囲は画面全体とし、自動でボールを検出した。Kalman フィルタの観測共分散行列、プロセス共分散行列の初期値は $R_x=1, R_y=1, Q_x=100, Q_y=100$ とした。

1. ティーショット

図8にティーショットにおけるボール抽出・追跡画面を示す。センサーカメラ映像上にボールの探索範囲を矩形で表示した。また過去のボールの位置座標を点で表示している。

図9は出力合成画像である。ティーショットのボールフォロー操作を行う放送カメラ映像に、グリーン上10ヤードごとに距離線をリアルタイム表示した。加えてボールの飛距離を文字・線にて表示した。

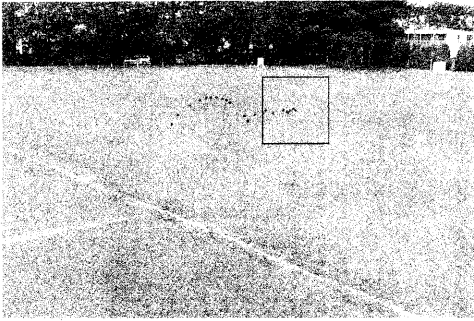


図8 ティーショットボール抽出・追跡画面



図9 ティーショット合成画像

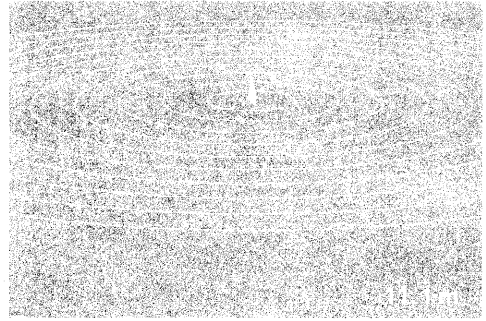


図11 アプローチショット合成画像

2. アプローチショット

図10にアプローチショットにおけるボール抽出・追跡画面を示す。ティーショットシーンと同様にセンサーカメラ映像上にボールの探索範囲を矩形で、過去のボールの位置座標を点で表示している。

図11は出力合成画像である。グリーンへのアプローチショットのボールフォロー操作を行う放送映像に、同心円状に距離線を表示した。加えてボールのカップまでの距離を文字にて表示した。

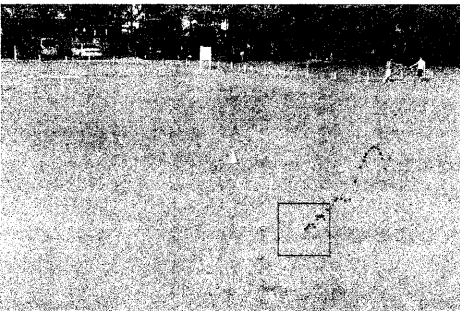


図10 アプローチショットボール抽出・追跡画面

ティーショット、アプローチショット共にボールの高速移動やバウンドに対しても見失うことなく、安定したボール抽出・追跡、及び距離線の描画を行えることを確認できた。また距離線もカメラの動きに応じて芝に張り付いたように描画することができた。

ボールの飛距離の測定誤差はセンサーカメラの画角とキャリブレーションの精度によるため評価することは難しいが、本実験においてはティーショットシーンで5m程度、アプローチショットで1m程度であった。

また3フレーム程度(約0.1秒)の処理速度で合成映像を作画することができた。

6. むすび

スポーツ映像制作における新たなスーパー映像生成手法を提案した。本手法により、映像からの高速移動オブジェクトの抽出・追跡。また、オブジェクトの特徴量と、カメラ画像内の物体の動きに連動したスーパー映像をリアルタイムで作画可能であることを確認できた。

今後、センサーカメラを複数台にすることによるオブジェクトの高さ方向の位置座標測定。また実用化を目指した操作性の改善などをはかる。さらに本手法を他のスポーツにも適合させ、様々なスポーツシーンにおける新たな映像表現を生み出し、実用化を進めていく。

文 献

- [1] 三須俊彦, 苗村昌秀, 境田慎一, 鄭 文涛, 金次保明, “複数情報の融合によるサッカー選手のロバストな追跡法,” 信学技報, IE2001-47, July 2001.
- [2] 鄭 文涛, 三須俊彦, 苗村昌秀, 和泉吉則, 金次保明, “複数カメラの協調動作による動オブジェクトの抽出” 信学論(D-II)vol.j86-D-II No.3 pp.397-408 2003.3
- [3] 齋藤 康之, 剣持 雪子, 小谷 一孔, “パラメトリックな眼鏡フレームモデルを用いた顔画像内の眼鏡フレーム領域の抽出と除去” 信学論 Vol.J82-D-II No.5 pp.880-890 1999.5