

LI\_002

# パーティクルフィルタによる単眼動画像からのサッカーボール3次元軌道推定 3D Soccer-ball Trajectory from Monocular Video Based on Particle Filtering

三須 俊彦<sup>†</sup> 高橋 正樹<sup>†</sup> 藤井 真人<sup>†</sup> 八木 伸行<sup>†</sup>  
Toshihiko Misu Masaki Takahashi Mahito Fujii Nobuyuki Yagi

## 1. まえがき

近い将来に運用開始されるサーバ型放送では、番組を意味的なまとまりに区分し、その内容を記述したメタデータの付与 [1] が必須となる。多様なプレイ (イベント) が発生するスポーツ番組用メタデータは、サーバ型放送のほか、映像データベースにおける映像検索、ダイジェスト番組制作など幅広い応用も期待される。

シナリオのないスポーツ映像に対しメタデータを全て手動付与するには莫大な時間、労力、コストを要することから、その自動化を目指す研究が盛んになっており [2][3]、我々も映像から選手のフォーメーションを解析することでコーナーキックなどのイベントを検出する手法を開発した [4]。一方、サッカーでは、ボール位置も重要な役割を担っていることから、ボール位置に基づくイベント検出手法 [5][6] も提案されている。しかし、解像度の低さ、複雑なダイナミクス、頻繁なオクルージョン、照明変動などボール検出自動化には課題が多い。このため、3次元運動をモデル化し、ボール像位置に関する最小自乗フィッティングを画像上で行う手法 [7] などが提案されている。

本稿では、ボールの放物運動、反発、転がり抵抗といった複雑なダイナミクスと、微小なボール像の曖昧な観測とをモデル化し、パーティクルフィルタ [8] による状態推定を行うことで、単眼の広角画像からボールの3次元軌道推定が可能となったので報告する。特に、今回はボール像の重心 [9] に加えて、面積に関する観測を導入したほか、粒子の初期化においてボール初速の確率モデルを導入することで3次元推定精度の向上を図った。

図1に提案するボール追跡手法の構成を示す。本手法は、ボール等の動物体を背景から抽出するシルエット抽出部、ボールらしい色・大きさ・縦横比のシルエットを選定するボール候補選定部、ボール候補の画像座標と面積に基づきボールの3次元位置・速度・加速度の確率密度分布を推定するパーティクルフィルタ部により構成される。以降、各構成要素の動作やモデル化について述べ、実画像を用いた実験結果を示す。

## 2. シルエット画像生成

芝生や観客席からボール像を分離するため、(1) 背景色情報に基づく手法と (2) フレーム間差分に基づく手法の2手法を実装し、その両者の切り替えもしくは論理演算によりシルエット画像を生成する。前者は芝生上のボールを抽出するのに有効であるのに対し、後者はボールが高速移動する場合に有効である。

### 2.1 背景色情報に基づくシルエット抽出

RGB色空間を  $256 \times 256 \times 256$  に均等に分割したルックアップテーブル  $T$  を用い、時刻  $t$  に

<sup>†</sup>日本放送協会, NHK

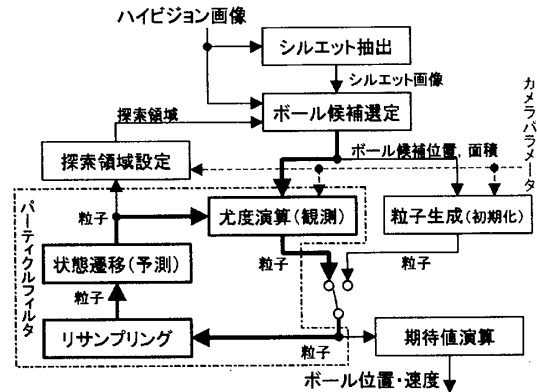


図1: 全体構成

おける入力画像の各画素  $X = [X, Y]^T$  の色  $[R(t, X), G(t, X), B(t, X)]^T$  が背景色であるか否かを判定することでシルエット  $S_C(t, X)$  を抽出する。

$$S_C(t, X) = T(R(t, X), G(t, X), B(t, X))$$

ここに、ボールを含む前景領域を  $S_C(t, X) = 1$ 、背景領域を  $S_C(t, X) = 0$  とする。なお、ルックアップテーブル  $T$  は、予め入力画像上の芝生領域の一部をマウスで指定することで作成した。

### 2.2 フレーム間差分に基づくシルエット抽出

「前後の両時点よりも現時点の輝度が十分高い画素」または「前後の両時点よりも現時点の輝度が十分低い画素」を検出することで移動物体を検出することができる。本手法では、時刻  $t - \alpha$ ,  $t$ ,  $t + \beta$  における入力画像 (芝生上でのボール弁別に有効な赤成分を使用) に注目し、以下の基準により動物体領域  $S_M(t, X)$  を検出する。

$$S_M(t, X) = \begin{cases} 1 & (\text{条件\#1または\#2が成立時}) \\ 0 & (\text{上記以外}) \end{cases}$$

$$\#1: R(t, X) > \max\{R(t - \alpha, X), R(t + \beta, X)\} + \rho_0$$

$$\#2: R(t, X) < \min\{R(t - \alpha, X), R(t + \beta, X)\} - \rho_1$$

ここに、 $\alpha > 0$ ,  $\beta > 0$  を撮像時間間隔とし、 $\rho_0 > 0$ ,  $\rho_1 > 0$  を輝度変化の閾値とする。

### 2.3 両手法の組み合わせ

本システムではボール候補選定のためのシルエット画像  $S(t, X)$  を生成するにあたり、撮影位置などの環境に応じて以下の4手法の中から適当なものを選択できるようにした。

$$S(t, X) = S_C(t, X) \quad (1)$$

$$S(t, X) = S_M(t, X) \quad (2)$$

表 1: 物理的なパラメータ

物理パラメータ	文字	実験に用いた値
ボール半径	$B$	0.11 m
重力加速度	$g$	9.8 m/s <sup>2</sup>
動摩擦係数	$\mu$	3.0/g
摩擦を受ける高度上限	$H$	0.30 m
反発係数	$e$	0.35
撮像パラメータ	文字	実験に用いた値
画像解像度		1920 × 540 画素
焦点距離	$f$	15 mm
シャッター時間	$T$	1/59.94 s
撮像素子画素サイズ	$s_x \times s_y$	5.0 × 10.0 μm <sup>2</sup>
リミッタ $\phi_l$	文字	実験に用いた値
ボール位置の制限範囲 $D$	$p_{x0} : p_{x1}$ $p_{y0} : p_{y1}$ $p_{z0} : p_{z1}$	-52.5 : 52.5 m -34.0 : 34.0 m 0.0 : 30.0 m
ボール速度の制限範囲	$v_{x0} : v_{x1}$ $v_{y0} : v_{y1}$ $v_{z0} : v_{z1}$	-50 : 50 m/s -50 : 50 m/s -50 : 50 m/s
ボール加速度の制限範囲	$a_{x0} : a_{x1}$ $a_{y0} : a_{y1}$ $a_{z0} : a_{z1}$	-1.0 : 1.0 m/s <sup>2</sup> -1.0 : 1.0 m/s <sup>2</sup> -10 : -9.6 m/s <sup>2</sup>

$$S(t, X) = \min\{S_C(t, X), S_M(t, X)\} \quad (3)$$

$$S(t, X) = \begin{cases} S_C(t, X) & (\text{ボール像推定位置が場内}) \\ S_M(t, X) & (\text{上記以外}) \end{cases} \quad (4)$$

(1)は高視点でボールの背景が常に芝生である場合, (2)はボールの動きが大きい場合に有効である. (3)および(4)は,  $S_C$  と  $S_M$  の長所を補完しあう手法である.

### 3. ボール候補の選定

シルエット画像  $S(t, X)$  の各単連結領域に対し, その面積  $A$ , 縦横比  $P$ , および該領域内における入力画像の平均色  $[\bar{R}, \bar{G}, \bar{B}]^T$  を求め, その各値が上下限2つの閾値内にあるもののみをボール候補として残すことで絞込みを行う. ボール候補が得られた場合には閾値範囲を狭め, 候補が得られなかった場合には範囲を拡げる動的閾値制御を行うことで, 照明変動などへの適応性を保ちつつ偽オブジェクトの混入を防いでいる.

### 4. 状態空間

本システムではボールの3次元かつ複雑な運動を推定することが目的であるため, サッカー場に固定した座標系(以下, ワールド座標系)における3次元位置, 速度および加速度からなる9次元の空間を状態空間にとった.

また, パーティクルフィルタにおける粒子の総数は  $N$  個とし, 時刻  $t$  における第  $n$  番目の粒子の状態ベクトルを  $x^{(n)}(t)$  とおく. 以降, 誤解のない範囲で,  $(t)$  または上付きの  $(n)$  を省略することがある.

$$x^{(n)} = [p_x^{(n)}, p_y^{(n)}, p_z^{(n)}; v_x^{(n)}, v_y^{(n)}, v_z^{(n)}; a_x^{(n)}, a_y^{(n)}, a_z^{(n)}]^T \\ \triangleq [(p^{(n)})^T; (v^{(n)})^T; (a^{(n)})^T]^T$$

ここに,  $p_i^{(n)}$ ,  $v_i^{(n)}$  および  $a_i^{(n)}$  を, それぞれ位置, 速度および加速度成分とする. 以降, 各種のパラメータは, 特に断りのない限り表1に示す文字で表記する.

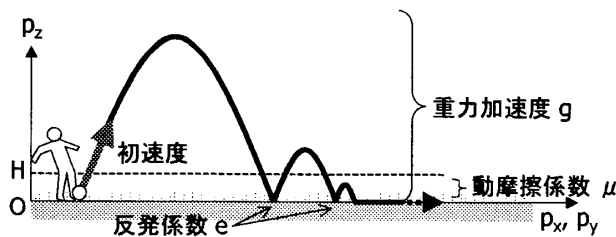


図 2: ボールのダイナミクス

### 4.1 状態遷移

サッカーボールの典型的な運動として, (1) 放物運動  $\phi_p$ , (2) バウンド  $\phi_r$ , (3) 転がり  $\phi_f$ , (4) キック・ヘディング, (5) その他の外乱(空気抵抗, 風, ピッチの凹凸など)が挙げられる. 今回は, 図2に示すように(1)~(3)を運動方程式でモデル化し, (4)および(5)はノイズ  $\nu$  により吸収することとした. また, 状態ベクトルの各成分の取り得る値にリミッタ  $\phi_l$  を設けることで, ボール位置が場外に出るといった追跡対象外の状態や, 速度 50 m/s (180 km/h) 超で飛行するなど物理的に通常あり得ない状態の粒子が生ずることを防いだ. 以下のとおり, 1 時点前(時刻  $t - \alpha$ ) から現時点  $t$  における粒子の状態遷移を定式化する. なお, 下式において  $\circ$  は合成関数を表す.

$$x(t) = (\phi_l \circ \phi_p \circ \phi_r \circ \phi_f)(x(t - \alpha)) + \nu(t - \alpha)$$

$$\phi_l(x) = \begin{bmatrix} \max\{p_{x0}, \min\{p_{x1}, p_x\}\} \\ \max\{p_{y0}, \min\{p_{y1}, p_y\}\} \\ \vdots \\ \max\{a_{z0}, \min\{a_{z1}, a_z\}\} \end{bmatrix}$$

$$\phi_p(x) = \begin{bmatrix} I & \alpha I & (\alpha^2/2)I \\ O & I & \alpha I \\ O & O & I \end{bmatrix} x$$

$$\phi_r(x) = \begin{cases} x & (p_z \geq B) \\ [p_x, p_y, 2B - p_z, v_x, v_y, -v_z, a_x, a_y, a_z]^T & (\text{上記以外}) \end{cases}$$

$$\phi_f(x) = \begin{cases} x & (p_z > H) \\ x - \left[ 0, \dots, 0, \frac{-[p_x, p_y]}{\sqrt{p_x^2 + p_y^2}} \mu g, 0 \right]^T & (\text{上記以外}) \end{cases}$$

ここに,  $I$  および  $O$  は, それぞれ 3 行 3 列の単位行列および零行列である. また,  $\nu$  はガウスノイズとした.

### 5. 観測と尤度

移動するボールをカメラで観測した場合, ボール像の重心の画像座標(視線)に基づき, ボールの存在領域を限定することができる. また, ボール像のシルエット面積には, カメラからボールまでの奥行きや動きばけの情報が含まれている. そこで, ボール候補として選択された  $M$  個のシルエットに関して, 各画像重心  $\{[\bar{X}^{[m]}, \bar{Y}^{[m]}]^T\}_{m=1}^M$  と面積  $\{A^{[m]}\}_{m=1}^M$  からなる  $3M$  次元の観測ベクトル  $y$  を下記のとおり定義する.

$$y = [\bar{X}^{[1]}, \bar{Y}^{[1]}, \dots, \bar{X}^{[M]}, \bar{Y}^{[M]}; A^{[1]}, \dots, A^{[M]}]^T \\ \triangleq [y_G^T; y_A^T]^T$$

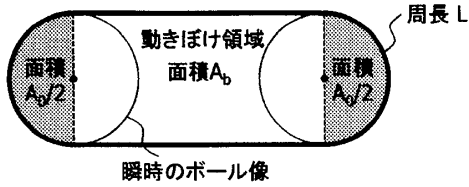


図 3: 動きぼけを伴ったボールのシルエット

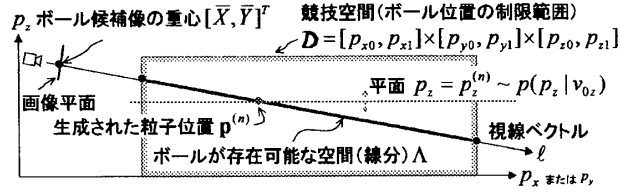


図 4: 粒子の初期化

また、尤度  $p(\mathbf{y} | \mathbf{x})$  は、重心を観測した際の尤度  $p(\mathbf{y}_G | \mathbf{x})$  と、その後で面積を観測した際の尤度  $p(\mathbf{y}_A | \mathbf{x}, \mathbf{y}_G)$  との積により定義する。

$$p(\mathbf{y} | \mathbf{x}) = p(\mathbf{y}_G | \mathbf{x}) \cdot p(\mathbf{y}_A | \mathbf{x}, \mathbf{y}_G)$$

以下において、状態ベクトル  $\mathbf{x}$  (またはその位置成分  $\mathbf{p}$ ) を画像座標に写像するモデルとして、ピンホールモデルによる透視投影変換  $\mathbf{h}(\mathbf{x} | \theta)$  (または  $\mathbf{h}'(\mathbf{p} | \theta)$ ) を用いることとする。なお、 $\theta$  はカメラ位置  $\mathbf{q}$ , 姿勢, 焦点距離  $f$ , 光軸位置および撮像素子の画素サイズ  $s_x \times s_y$  からなるカメラパラメータであり、事前の画像計測により求めた値を用いた。

### 5.1 シルエット重心の観測

粒子を投影した画像座標の近傍にボール候補のシルエットが存在すれば、その粒子の状態は尤もらしいといえる。そこで、まず、各粒子に関し、状態ベクトル  $\mathbf{x}$  の像  $\mathbf{h}(\mathbf{x} | \theta)$  を求める。続いて、観測されたボール像の重心群  $\{[\bar{X}^{[m]}, \bar{Y}^{[m]}]^T\}_{m=1}^M$  の中から  $\mathbf{h}(\mathbf{x} | \theta)$  に最も近いものまでの距離  $d$  を求める。なお、距離の算出にあたっては撮像素子の画素の縦横比を考慮し、 $X$  (水平) 方向の画素数単位で  $d$  を求めた。

$$d = \min_{m=1, \dots, M} d^{[m]}, \quad \hat{m} = \arg \min_{m=1, \dots, M} d^{[m]}$$

$$d^{[m]} = \left\| \left\{ [\bar{X}^{[m]}, \bar{Y}^{[m]}]^T - \mathbf{h}(\mathbf{x} | \theta) \right\} \begin{bmatrix} 1 & 0 \\ 0 & s_y/s_x \end{bmatrix} \right\|$$

ここで、正規分布  $\mathcal{N}(\mathbf{0}, \text{diag}(\sigma_G^2, \sigma_G^2))$  に従う観測ノイズを仮定し、尤度  $p(\mathbf{y}_G | \mathbf{x})$  を次式により定める。

$$p(\mathbf{y}_G | \mathbf{x}) \propto \exp(-d^2 / (2\sigma_G^2))$$

### 5.2 シルエット面積の観測

前節で距離  $d^{[m]}$  が最小となったシルエット  $\hat{m}$  について、その面積の尤もらしさを評価する。

ボールのシルエットは、シャッター開放時間  $T$  のボール像の集合和として観測される。すなわち、図3に示すように、観測されるシルエットの面積  $A$  は、瞬時のボール像の面積  $A_0$  (次式第1項) に動きぼけ領域の面積  $A_b$  (次式第2項) を加えたものとなる。

$$A \simeq \pi B^2 f^2 / (z^2 s_x s_y) + 2Bf^2 T \|v_{\perp}\| / (z^2 s_x s_y)$$

$$z = \hat{\ell} \cdot (\mathbf{p} - \mathbf{q}), \quad v_{\perp} = \mathbf{v} - (\hat{\ell} \cdot \mathbf{v}) \hat{\ell}$$

ここに、ワールド座標系における光軸方向 (被写界を正) の単位ベクトルを  $\hat{\ell}$ , カメラの設置位置を  $\mathbf{q}$  とした。なお、 $z$  はカメラに対する粒子位置の奥行きを、 $v_{\perp}$  は粒子速度の視線直交成分をそれぞれ表している。

実際に抽出されるシルエット形状は、背景や閾値設定の違いなどによって、本来の姿よりも膨張・収縮して観測されることがある。観測面積の変動は、モルフォロジカルな膨張・収縮の量 (一定幅を仮定する) と周長  $L$  との積で近似できる。そこで面積の誤差を周長  $L$  で除したものが正規分布  $\mathcal{N}(c, \sigma_A^2)$  に従うと仮定し、尤度  $p(\mathbf{y}_A | \mathbf{x}, \mathbf{y}_G)$  を定義した。

$$p(\mathbf{y}_A | \mathbf{x}, \mathbf{y}_G) \propto \exp \left\{ \frac{-1}{2\sigma_A} \left( \frac{A - A^{[m]}}{L} - c \right)^2 \right\}$$

$$L \simeq 2 \| \mathbf{h}'(\mathbf{p} | \theta) - \mathbf{h}'(\mathbf{p} - \mathbf{v}T | \theta) \| + 2\pi \sqrt{(r_a^2 + r_b^2) / 2}$$

$$r_a \simeq \frac{Bf}{z \min\{s_x, s_y\}}, \quad r_b \simeq \frac{Bf}{z \max\{s_x, s_y\}}$$

なお、上の  $L$  の近似式の右辺1項は図3の直線部分の長さ、第2項は曲線部分の長さの近似式 (第2種完全楕円積分の近似式を使用) である。

## 6. 粒子の初期化

ボール候補シルエットの重心  $[\bar{X}, \bar{Y}]^T$  に基づき  $N_0$  個の粒子を生成する。以下に、粒子1個あたりの生成手順を示す。まず、図4に示すように投影中心から  $[\bar{X}, \bar{Y}]^T$  を通る視線ベクトル (半直線  $\ell$ ) を引く。ボールの存在位置は、半直線  $\ell$  と競技空間  $D$  の共通部分  $\Lambda = \ell \cap D$  内に拘束される。

次に、ある時点のボールの高度  $p_z$  と鉛直速度  $v_z$  は、以前にピッチ上 ( $p_z = 0$ ) にあったボールがある鉛直初速度  $v_{0z}$  で投げ上げられた結果であると考え、粒子  $n$  の鉛直初速度  $v_{0z}^{(n)}$  を  $v_{0z}^{(n)} \sim \mathcal{N}(0, \sigma_{v_{0z}}^2)$  なる正規分布からのサンプリングにより与える。

鉛直初速度  $v_{0z}$  で投げ上げたボールが、高度  $p_z$  を通過するときの鉛直速度は  $v_z = \pm \sqrt{v_{0z}^2 - 2gp_z}$  である。また、このボールが達し得る最高高度は  $p_{z \max} = v_{0z}^2 / (2g)$  である。このボールが高度  $p_z$  にある確率密度  $p(p_z | v_{0z})$  は、鉛直速度の絶対値に反比例するので、

$$p(p_z | v_{0z}) = \begin{cases} g / (v_{0z} \sqrt{v_{0z}^2 - 2gp_z}) & (0 \leq p_z \leq p_{z \max}) \\ 0 & (\text{上記以外}) \end{cases}$$

となる。そこで、粒子  $n$  の高度  $p_z^{(n)}$  を、 $p_z^{(n)} \sim p(p_z | v_{0z}^{(n)})$  なるサンプリングにより生成する。

続いて、ボールが存在し得る空間  $\Lambda$  と平面  $p_z = p_z^{(n)}$  の交点を求め、その座標を粒子  $n$  の位置成分  $\mathbf{p}^{(n)}$  とする。速度成分は  $v_x^{(n)} \sim \mathcal{N}(0, \sigma_{v_x}^2)$ ,  $v_y^{(n)} \sim \mathcal{N}(0, \sigma_{v_y}^2)$ ,  $v_z^{(n)} = \pm \sqrt{(v_{0z}^{(n)})^2 - 2gp_z^{(n)}}$  (符号は 1/2 の確率で決定) とし、加速度成分は  $\mathbf{a}^{(n)} = [0, 0, -g]^T$  とする。

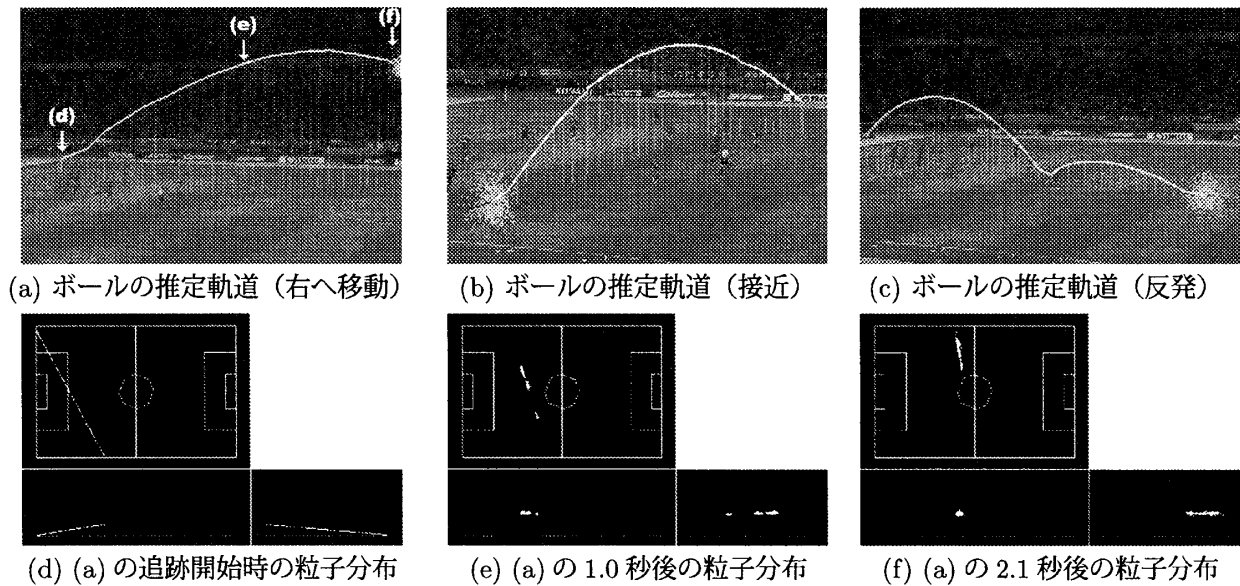


図 5: 実験結果

## 7. ボール推定位置, 速度の出力

パーティクルフィルタ部では, ボールの状態ベクトルの確率密度分布が得られる. 本手法では, この確率密度分布に基づき位置, 速度, 加速度の期待値を求めることで, 最終的に出力すべき位置等の推定値とした.

## 8. 実験

表 1 に示したパラメータを用い, ボール追跡を行った結果を図 5 に示す. パーティクルの総数は  $N = 1000$  個とした. フレーム間差分の閾値は  $\rho_0 = \rho_1 = 20$  とし, シルエット抽出は式 (4) の手法を採用した. なお, 追跡開始時のボールの探索範囲に限り手動で与えた.

(a) はボールの 3 次元推定軌道を原画像上にマッピングして表示したものである. 曲線 (黄色) が推定軌道を, 垂直の線 (灰色) が高さを, 点群 (水色) が粒子像をそれぞれ表している. 単眼画像であるにもかかわらず, 自然な 3 次元情報が得られている.

(d)~(f) は粒子分布を第三角法で示したものである. (d) はボールを最初に捕捉した際に, 第 6 節の手法で生成された粒子である. 粒子群は視線上に生成されている. 追跡開始間もない時点では, (e) に示すように粒子分布が多峰になっており, まだ解が絞られていないことを示している. 十分に時間が経つと (f) のように粒子が局在化し, 推定が収束していることがわかる.

(b) は接近するボールの推定軌道である. 視線方向に近い移動であるため単眼では推定が難しい例であるが, 定性的には正しい軌道が推定できている. (c) は跳ね返り (地面付近での脚との接触) のある場合の推定軌道である. 反発モデルとプロセスノイズが有効に働いた結果, ほぼ正しい軌道が得られている.

なお, 本追跡に要した処理時間は, Intel Xeon プロセッサ (3.2GHz) 2 個搭載のパーソナルコンピュータ使用で 1 フレームあたり約 0.13 秒であった. 図 5 はオフライン処理による追跡結果であるが, コマ落としの頻度を上げることで実時間追跡も可能であった (ただ

し, 精度や頑健性は若干劣る).

## 9. まとめ

サッカーボールの複雑なダイナミクスと広角画像の曖昧な観測をモデル化し, パーティクルフィルタによって状態推定を行うことにより, 単眼画像からの 3 次元軌道の推定が可能であることが示された. 精度や頑健性の定量評価, 多視点画像への応用, 物理パラメータやボール選定の閾値などの最適化/機械学習が今後の課題である.

## 参考文献

- [1] 馬場, 西本, 石川, 中村, 吉村, 栗岡: サーバー型放送におけるメタデータ利用技術の一検討, 信学技報 MoMuC2004-57, pp.11-16, 2003.
- [2] 新田, 馬場口, 北橋: 言語と画像の情報統合によるスポーツ映像へのアノテーション付け, MIRU2000, pp. I-319-I-324, 2000.
- [3] 瀧, 長谷川: チームスポーツにおける集団行動解析のための特徴量とその応用, 信学論 D-II, vol. J81-D-II, no. 8, pp. 1802-1811, 1998.
- [4] 三須, 高橋, 蓼沼, 八木: サッカー映像のフォーメーション解析に基づく実時間イベント検出, 情報科学技術レターズ (FIT2005), vol. 4, pp. 141-144, 2005.
- [5] 窪田, 有木, 塚田: 嗜好分類に基づく個人適応型サッカー映像の自動生成技術, 信学技報, PRMU2005-115, pp.7-12, 2005.
- [6] 橋本, 小沢: 選手とボールの動きに基づいた集団スポーツの試合解析~サッカーのオフサイド自動判定システム~, 2006 信学総大, D-12-40, p.172, 2006.
- [7] Y. Ohno, J. Miura, Y. Shirai: Tracking Players and Estimation of the 3D Position of a Ball in Soccer Games, ICPR 2000, pp. 1145-1148, 2000.
- [8] M. Isard, A. Blake: CONDENSATION — Conditional Density Propagation for Visual Tracking, Int. J. Computer Vision, vol. 29, no. 1, pp. 5-28, 1998.
- [9] 三須, 高橋, 藤井, 八木: 逐次モンテカルロ法による実時間サッカーボール追跡, 2006 信学総大, D-12-39, p.171, 2006.