

## 背景・照明・対象サイズの変化に頑健なパーティクルフィルタ Robust Particle Filter for Changes of Background, Illumination and Target Size

河中 治樹† 岩堀 祐之‡ 高井 敏裕††  
Haruki Kawanaka Yuji Iwahori Toshihiro Takai

### 1. まえがき

コンピュータビジョンにおける動画画像からの移動物体追跡に関する手法の中で近年パーティクルフィルタ[1]と呼ばれる手法が注目されている。この手法は複雑背景の下で観測値が非ガウス型になる状況においても頑健な追跡を実現する。そのためパーティクルフィルタを用いた動画画像からの人物追跡に関する研究や人物の顔部追跡に関する研究などが多くなされている。

ところが従来のパーティクルフィルタでは照明の変化、さらには対象の移動による背景領域および対象のサイズの変化を考慮していないため、映像中でこれらの変化が起きるとその追跡精度が低下するという問題点がある。

そこで本論文では、パーティクルフィルタを用いた動画画像からの移動物体追跡の精度向上のため、照明変化や背景変化を考慮した参照パターンの生成方法および対象のサイズ推定を利用した仮説群生成の手法を提案する。そして、動画画像を用いた実験を行い、提案手法の有効性を確認する。

### 2. パーティクルフィルタによる追跡

時刻  $t$  において  $N$  個の重みつき仮説群  $s_i^{(t)}$  ( $i = 1, \dots, N$ ) はそれぞれ状態空間に追跡対象の状態量  $x_i$  を持ち、その確率は各仮説の重み  $\pi_i^{(t)}$  に比例する。ある時刻  $t$  における観測結果を  $z_t$  とし、時刻  $t$  までに得られる観測結果を  $Z_t = \{z_1, \dots, z_t\}$  とする。時刻  $t$  までの観測結果が得られた時の状態  $x_t$  の事後確率分布  $p(x_t|z_t)$  をベイズの法則に従って式(1)のように推定する。

$$p(x_t|z_t) = \alpha p(z_t|x_t)p(x_t|Z_{t-1}) \quad (1)$$

ただし、 $\alpha$  は正規化のための定数である。時刻  $t-1$  において観測  $Z_{t-1}$  が得られたときの状態量  $x_{t-1}$  の分布  $p(x_{t-1}|z_{t-1})$  が  $N$  個の重みつき仮説群  $s_i^{(t-1)}$  ( $i = 1, \dots, N$ ) で表されている時、各仮説の重み  $\pi_i^{(t-1)}$  の比に従って  $N$  個の仮説群  $s_i^{(t-1)}$  を選択する。その各仮説を事前に与えられた状態空間モデル  $p(x_t|x_{t-1})$  によって推移させることで確率分布  $p(x_t|Z_{t-1})$  を近似する。

本研究では追跡対象の画像上での座標  $(x, y)$  およびそれを中心とした追跡対象を内包する矩形の幅  $w$  と高さ  $h$  を追跡対象の状態変数とする。状態空間モデルは式(2)によって遷移させ、時刻  $t$  における  $N$  個の仮説群  $s_i^{(t)}$  を生成する。

$$x_i^{(t)} = \{x_{i-1}^{(t)} + \omega_x, y_{i-1}^{(t)} + \omega_y, w_{i-1}^{(t)}, h_{i-1}^{(t)}\} \quad (2)$$

ただし、 $\omega_x$  および  $\omega_y$  はガウシアンノイズであり、 $w_{i-1}^{(t)}$  および  $h_{i-1}^{(t)}$  は時刻  $t-1$  における確率分布から推測した幅および高さである。 $w_{i-1}^{(t)}$  および  $h_{i-1}^{(t)}$  については4節で詳しく述べる。

新しいサンプル  $s_i^{(t)}$  の重み  $\pi_i^{(t)}$  は画像の色情報に基づく尤度から推定する。重みの和は1になるように正規化する。その結果、時刻  $t$  における  $p(x_t|z_t)$  の近似表現を得ることができる。追跡対象の最適状態量推定には仮説群の期待値を用いる。尤度関数  $L(z_t|x_t)$  を利用して、

$$\pi_i^{(t)} = L(z_t|x_i) \quad (3)$$

により重みを計算し、Sequential Importance Sampling[2]によりリサンプリングを行う。

次に本研究における尤度関数について述べる。表色系はHSVを使い、現在の画像における追跡対象を内包する矩形領域のHSの正規化色ヒストグラム  $H_u$  と予め得ておいた参照用ヒストグラム  $H_{ref}$  との類似度  $S$  を次式のSwainのHistogram Intersection[3]を用いて計算する。

$$S = \sum_{i=1}^T \min(H_u, H_{ref}) \quad (4)$$

ただし、 $T$  はHSの各階調を表す。この類似度  $S$  を用いて尤度関数を次式のように定める。

$$L(z_t|x_i) = e^{kS^2} \quad (5)$$

ただし、 $k$  は定数である。以上より、追跡対象の存在可能性が高い仮説に大きな重みを与えることができる。また時刻  $t$  での推定結果  $x_t^e$  を以下の式によって求めることにする。

$$x_t^e = \{x_t^e, y_t^e, w_t^e, h_t^e\} = \sum_{i=1}^N \pi_i^{(t)} x_i^{(t)} \quad (6)$$

### 3. 参照パターンの生成

#### 3.1 一様な照明変化に頑健なヒストグラムの作成

文献[4]では照明環境の違いに頑健な色ヒストグラムとして表色系にHSVを使い、二次元のHSヒストグラムと一次元のVヒストグラムを合わせて色ヒストグラムを作成している。文献[4]では、ヒストグラム作成の際に二つの閾値を設けているが、これらは特定の動画に対して経験的に定められている。また照明環境の変化の度合に応じて閾値を動的に定めるべきだと思われるが、手動で与えた閾値のまま固定で追跡を行っている。

図1は一様な照明変化が起きる映像において、追跡対象のHSヒストグラムを文献[4]と同様の大きさで作成したものである。この結果から、画像全体に対して一様に照明環境の変化が発生する場合、HSヒストグラムの形に大きな変化は発生しないことがわかる。そこで、本論文ではHとSの情報のみの二次元のHSヒストグラムをパーティクルフィルタでの追跡の際の特徴量として用いる。そうすることで、動的な閾値を設定することなく、照明変化に頑健な追跡を行うことができる。

#### 3.2 背景を考慮した参照用ヒストグラムの作成

各追跡対象の参照用色ヒストグラムの作成方法について

† 愛知県立大学 情報科学部

‡ 中部大学 工学部

†† 名古屋工業大学大学院 工学研究科

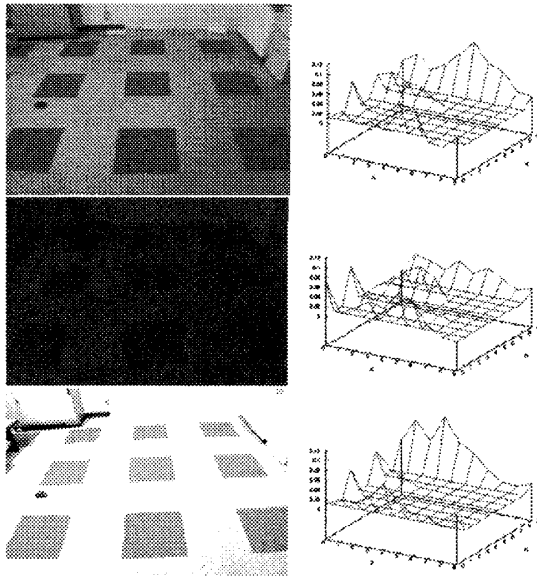


図1 一様な照明変化が起きた場合のHSヒストグラム

述べる。尤度関数のための参照用色ヒストグラムの作成には背景差分を用いる。その結果から各追跡対象を内包する矩形の位置と大きさが決定できる。その矩形内の色情報をそのまま用いて参照用色ヒストグラムとすると追跡対象を内包する矩形には背景の色情報も含まれているため、初期フレームと異なる背景の位置に追跡対象が移動した時に参照用色ヒストグラムとの尤度が低くなる。

そこで背景差分の後に追跡対象の色ヒストグラム  $H_{target}$  とは別に背景全体の色ヒストグラム  $H_{back}$  を作成しておき、これらを合成した色ヒストグラムを追跡対象の参照用色ヒストグラム  $H_{ref}$  とする。これにより初期フレームでの背景の参照用色ヒストグラムへの影響を軽減することができる。

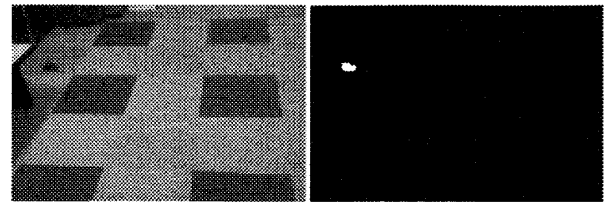
背景差分で検出された追跡対象の検出画素数を  $P_{detect}$ 、検出領域を内包する矩形の総画素数を  $P_{square}$  とすると、追跡対象を内包する矩形内における追跡対象と背景の割合は  $P_{detect} : P_{square} - P_{detect}$  となる。追跡対象を内包する矩形に対して追跡対象が占める割合が姿勢によって大きく変化することはないという仮定の元で、式(7)により参照用色ヒストグラム  $H_{ref}$  を作成する。

$$H_{ref} = \frac{P_{detect}}{P_{square}} H_{target} + \frac{P_{detect} - P_{square}}{P_{square}} H_{back} \quad (7)$$

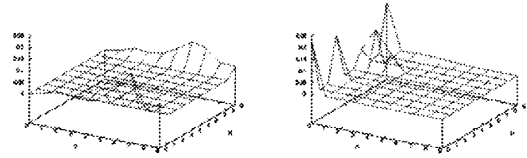
提案手法で生成した参照用色ヒストグラムの様子を図2に示す。図2(b)は(a)の入力画像に対して背景差分を行った結果、(c)は追跡対象のみの色情報を用いて作成された色ヒストグラム  $H_{target}$ 、(d)は背景の色情報のみを用いて作成された色ヒストグラム  $H_{back}$ 、(e)は追跡対象を内包する矩形内の色情報のみを用いた色ヒストグラム、(f)は(c)と(d)を式(7)で合成した色ヒストグラム  $H_{ref}$  である。(f)は対象が移動して背景が変わった場合にも尤度の低下があまりないヒストグラムになっている。

#### 4. 追跡対象の尤度を評価する矩形の大きさ推定

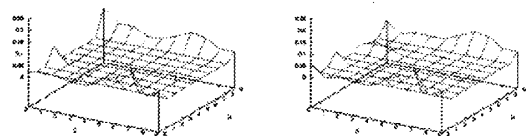
カメラから得られる画像上の追跡対象の位置  $(x, y)$  を追跡対象が存在する三次元平面の座標系へ射影変換することで、カメラから追跡対象までの距離を求められる。これに加え



(a)入力画像 (b)背景差分検出結果



(c)  $H_{target}$  (d)  $H_{back}$



(e) 追跡対象を内包する矩形内 (f)  $H_{ref}$

図2 参照用色ヒストグラムの例



図3 矩形の推定が正しく行われない例

てあらかじめ追跡対象の大きさに関する知識を与えることで追跡の精度向上と効率化をはかる。

従来は式(1)の4つの状態変数を用いられることが多く、これらの値にガウシアンノイズを加えることで確率ダイナミクス  $p(x_i|x_{i-1})$  を実現している。しかし  $w$  と  $h$  に関しては、追跡に用いている特徴量によってはガウシアンノイズだけでは不十分である。例えば図3にある3つの矩形は全て矩形内の黒色の頻度が同じであるため尤度が等しくなる。矩形内の色ヒストグラムの類似度を特徴量として用いることは、すなわち矩形内に占める追跡対象の面積の割合を調べることと同義であるため、単純にガウシアンノイズによる探索を行っただけでは追跡対象を内包する矩形とならない可能性が高い。

そこで本研究では事前に追跡対象の大きさに関する知識を与えておくことで、単純なガウシアンノイズによる推定よりも精度の高い  $w$  と  $h$  を推定する方法を提案する。

#### 4.1 弱透視投影での追跡対象の大きさ推定

カメラ中心座標系で  $(X, Y, Z)$  と表される点が画像の  $(x, y)$  に投影されるとする。図4に示すようなピンホールカメラを仮定した場合の弱透視投影を考えると式(8)が成立する。

$$x = f \frac{X}{Z}, \quad y = f \frac{Y}{Z} \quad (8)$$

ただし、 $Z$  は物体を画像と平行な面に投影した面までの  $Z$  軸方向の距離、 $f$  は焦点距離である。これにより、平行投影面上での物体の面積を  $S$  とすると、画像上での物体の面積  $S'$  は

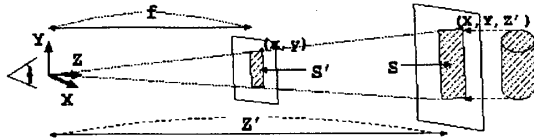


図4 弱透視投影を仮定した大きさ推定

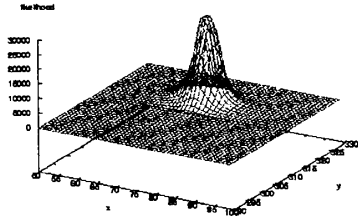


図5 理想的な環境における尤度分布

$$S^{(i)} = f^2 \frac{S + \omega_S}{Z^2} \quad (9)$$

として求められる。ただし、 $\omega_S$ は平均が0、標準偏差が $r_S$ のガウシアンノイズである( $r_S$ は0.0~1.0の定数)。カメラ中心から $Z$ 離れた追跡対象の画像上での大きさ(画素数)の推定値として $S'$ を用いる。 $S$ の値はパーティクルフィルタの初期状態設定時に既知の $Z$ および $S'$ から予め求めておく。式(9)で $S$ にガウシアンノイズを加えるのは追跡対象の姿勢変化によって画像上での追跡対象の画素数が変化することに対応するためである。 $r_S$ は各フレーム間の追跡対象の姿勢変化によって追跡対象の見えの大きさが変化する割合であり、追跡対象の大きさに関する知識から事前に与えておく。これにより追跡対象の姿勢変化に起因する画像上での見え(大きさ)の変化にも対応することができる。

#### 4.2 尤度分布を用いた追跡対象の縦横比推定

追跡対象をより正確に捉えているパーティクルの確率ほど高くなるため、理想的な環境においては確率分布が図5に示すように追跡対象を中心として綺麗な山形となるはずである。よって、分布の極大点から一定以上の確率である範囲に追跡対象が存在していると考えるのは妥当である。

そこで時刻 $t$ の尤度分布から追跡対象を内包する矩形の縦横比を推定し、その結果を時刻 $t+1$ の状態変数に用いることで、より正確に追跡対象を捕らえ矩形の形で追跡を行うことができる。図5のように、最も高い尤度となる座標は追跡対象の中心座標であると考えられるので、以下のようにして矩形の縦横比を推定する。

まず全ての仮説に対して時刻 $t$ の追跡対象位置の推定結果 $(x_t^e, y_t^e)$ からのずれ $d^{(i)}$ を式(10)により求める。

$$d^{(i)} = \sqrt{(x_t^e - x_t^{(i)})^2 + (y_t^e - y_t^{(i)})^2} \quad (10)$$

そして $d^{(i)}$ が小さくてかつその尤度が閾値 $l$ より大きい仮説の $x_t^{(i)}$ と $y_t^{(i)}$ からそれぞれの最大値と最小値 $x_{\max}, x_{\min}, y_{\max}, y_{\min}$ を求め、時刻 $t$ の追跡対象領域を表す矩形の推定縦横比 $R_t$ を次式により求める。

$$R_t = \frac{y_{\max} - y_{\min}}{x_{\max} - x_{\min}} \quad (11)$$

表1 一様な照明変化が生じる映像での実験における状態変数の理論値と推定値の平均誤差の比較

	$x$	$y$	$w$	$h$
提案手法	4.79	8.96	15.25	12.35
文献[4]	39.88	14.74	29.98	16.20

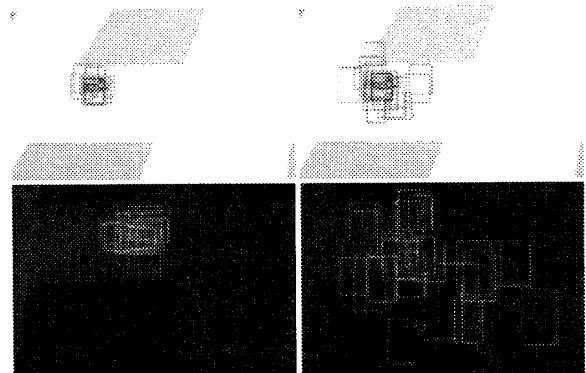


図6 一様な照明変化が生じる映像での追跡結果の比較(左:提案手法, 右:文献[4]の手法)

ただし、パーティクルフィルタは確率分布を離散サンプリングしているため、使用する粒子数によっては図5のような綺麗な分布が得られない。そのため $R_t$ が $R_{t-1}$ から急激に変化した場合にはその推定縦横比は誤っていると見て、1フレーム前の $R_{t-1}$ を $R_t$ として用いることにする。

式(9)および(11)を用いて以下のように式(1)の $w_{t-1}^{(i)}$ および $h_{t-1}^{(i)}$ を計算する。

$$w_{t-1}^{(i)} = \sqrt{\frac{S^{(i)}}{R_{t-1}}} \quad h_{t-1}^{(i)} = \sqrt{\frac{S^{(i)}}{w_{t-1}^{(i)}}} \quad (12)$$

#### 5. 実験と考察

本実験に用いた映像は720×480ピクセル、24bitカラーであり、計算機はCPUがAthlon XP 2200+, メモリ512MBである。粒子数を40個、式(6)のパラメータ $k$ は経験的に20と定めた。

まず3.1節で提案した、一様な照明変化に頑健なヒストグラムによる追跡の精度評価を行った。実験に用いた映像では模様や障害物のある床面を模型自動車が多様な方向に移動する。撮影の最中に16段階の明るさ調整を行った。明るさの変化は画像内で一様に発生している。この16種類の映像に対して提案手法を用いた場合と文献[4]の方法を用いた場合の追跡結果の比較を行った。提案手法が二次元のHSヒストグラムであるのに対して、文献[4]のヒストグラムは二次元のHSヒストグラムと一次元のVヒストグラムの組み合わせである。ヒストグラム作成の際の閾値も文献[4]と同じにした。図6に各手法での追跡の様子を示す。図中における薄色の矩形は時刻 $t$ での各粒子によるサンプリングされた位置、濃色の矩形は推定結果 $x_t^e$ である。またシーン内での各照明変化前後の計32フレームにおいて、目視で確率変数の理論値を測定して推定値との誤差の平均を求めた。表1に比較結果を示す。

図6より照明が暗くなると文献[4]の手法では追跡精度が悪くなるだけでなく、粒子群が拡散してしまい追跡不可能

表2 参照用ヒストグラムの背景合成の有無による状態変数の理論値と推定値の平均誤差の比較

	$x$	$y$	$w$	$h$
提案手法	5.74	3.29	14.24	5.19
追跡対象の色情報のみ	5.55	3.01	19.41	16.32

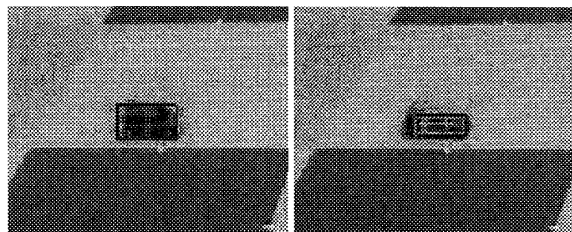


図7 参照用ヒストグラムの背景合成の有無による追跡結果の比較(左:提案手法, 右:追跡対象の色情報のみ)

になった。一方、提案手法は照明の変化があっても対象を捕らえ続けることができた。その際の各状態変数の誤差も文献[4]の手法での結果に比べて小さい。この結果から提案手法では照明変化に頑健な追跡を行えることが確認できる。

次に3.2節の式(7)で作成した参照用色ヒストグラムを用いた場合と背景差分で検出された追跡対象のみの色情報から作成した参照用色ヒストグラムを用いた場合の追跡結果の比較を行った。ヒストグラムは二次元のHSヒストグラムで $T=10$ である。図7に各手法での追跡の様子を示す。図中における薄色の矩形は時刻 $t$ での各粒子によるサンプリングされた位置、濃色の矩形は推定結果 $x_t$ である。またシーン内の90フレームにおいて目視で状態変数の理論値を測定し、推定値との誤差の平均を求めた。表2に比較結果を示す。

表2から追跡対象の色情報のみでは $w$ および $h$ ともに誤差が大きいことがわかる。これは参照用色ヒストグラムに背景の色情報が含まれていないために、追跡対象からはみ出さずに背景の色情報を持たない仮説の尤度が高くなってしまったためである。一方、提案した手法では追跡対象の座標推定精度を保ったまま、 $w$ と $h$ の推定誤差が減っている。このことから追跡対象のみの色情報だけではなく、背景の色情報を合成することで追跡精度が向上することが確認できる。

最後に4節で提案した、知識による推定効率化に関する検証として追跡対象の大きさ推定を行った場合と行わなかった場合の追跡結果の比較を行った。実験結果を図8および表3に示す。式(9)の $\omega_s$ の標準偏差 $r_s$ は0.2とした。追跡対象を内包する矩形の推定縦横比の閾値 $l$ は経験的に25とした。またシーン内で追跡対象が射影変換の特徴点上を通過するフレームにおいて目視で確率変数の理論値を測定し、誤差の平均を求めた。

図8より、推定を行わなかった場合の結果では、追跡対象の一部だけに枠がかかっていることがわかる。一方、推定を行うことで対象を内包する矩形で追跡を行うことができている。また、 $w$ と $h$ の推定誤差だけでなく、追跡対象の座標 $(x, y)$ の推定値の誤差も小さくなっている。これは矩形の大きさを精度良く推定できたことにより、色情報を的確に捕らえられるようになったためと思われる。

図7のような様な照明変化が発生する条件下での実験

表3 追跡対象の大きさ推定の有無による状態変数の理論値と推定値の平均誤差の比較

	$x$	$y$	$w$	$h$
大きさ推定する場合	3.23	2.57	7.23	4.10
大きさ推定しない場合	9.30	5.93	15.90	11.67

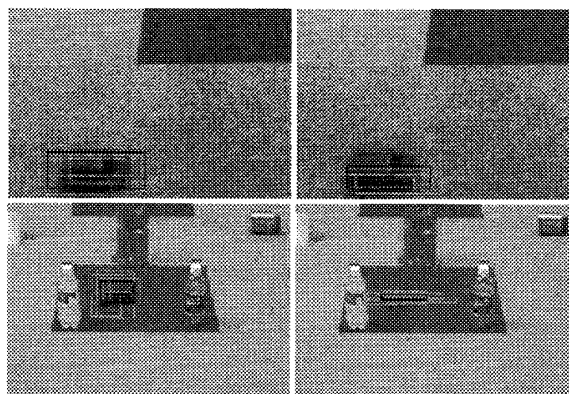


図8 追跡対象の大きさ推定の有無による結果の比較(左:推定をする場合, 右:推定をしない場合)

では矩形の大きさを推定しない場合、色情報の信頼度が落ちて追跡対象を正確に捕らえることが困難になる。しかしながら、矩形の大きさ推定を行うことで照明変化による色情報の信頼度低下があっても精度良く矩形の推定が可能となり、追跡精度が大幅に向上した。

## 6. おわりに

本論文では、パーティクルフィルタによる移動物体追跡を行う際の精度向上のための照明変化や背景変化を考慮した参照パターン生成手法および対象のサイズ推定を利用した仮説群生成の手法を新たに提案した。提案手法を評価した結果、既存手法より精度が向上することを確認できた。

追跡対象と背景対象が似た色の環境になると、尤度分布が一つの山型にならず、縦横比の推定に失敗することがある。このように尤度分布の極大点が多数ある場合へ対応し、追跡対象と似た色情報をもつ背景環境での追跡精度を向上させることは今後の課題である。

謝辞 本研究は科学研究費補助金(基盤研究(C)16500108)及び中部大学研究費のサポートによる。ここに感謝申し上げます。

## 参考文献

- [1] M. Isard and A. Blake: "Condensation- conditional density propagation for visual tracking", Int. J. Computer Vision, Vol. 29, pp. 5-28, 1998.
- [2] A. Doucet, S. Godsill and C. Andrieu: "On sequential Monte Carlo sampling methods for Bayesian filtering", Statistics and Computing, Vol. 10, pp. 197-208, 2000.
- [3] M. J. Swain and D. H. Ballard: "Color indexing", Int. J. Computer Vision, Vol. 7, pp. 11-32, 1991.
- [4] P. Perez, C. Hue, J. Vermaak and M. Cange: "Color-based probabilistic tracking", Proc. European Conf. on Computer Vision, Vol. 1, pp. 661-675, 2002.