

部分空間分離法による特徴点追跡のアウトライア除去

菅谷 保之 金谷 健一

岡山大学工学部情報工学科

{sugaya,kanatani}@suri.it.okayama-u.ac.jp

ビデオ画像上の特徴点の軌跡から誤った追跡結果を除去する手法について述べる。3次元形状復元や移動物体の分離のためにビデオ画像上で特徴点を追跡するアルゴリズムが知られているが、追跡結果がすべて正しいとは限らない。本稿では、正しく追跡した特徴点の軌跡はある部分空間に含まれるという拘束条件を用いてアウトライアを除去する。これはRANSACを用いて特徴点の軌跡に部分空間をロバストに当てはめ、しきい値処理によって残差の大きいデータを除去するものである。実画像を用いた実験により、静止シーンだけではなく、シーン中に動物体が存在する場合にも提案手法が有効であることを示す。そして実際に運動物体を移動する背景から分離し、提案手法によって分離の正解率が向上することを確認する。

キーワード: 特徴点追跡, アウトライア除去, 部分空間分離法, ロバスト推定, RANSAC

Outlier removal for motion tracking by subspace separation

Yasuyuki Sugaya and Kenichi Kanatani

Department of Information Technology, Okayama University, Okayama 700-8530 Japan

{sugaya,kanatani}@suri.it.okayama-u.ac.jp

We present a robust method for outlier removal in motion tracking. Many tracking techniques have been proposed for 3-D reconstruction and motion segmentation, but the resulting trajectories are not necessarily correct. In this paper, we propose an algorithm for detecting and removing outliers based on the knowledge that correct trajectories are constrained to be in a subspace of their domain. This algorithm robustly fits the subspace to the detected trajectories using RANSAC and removes those that have large residuals. We show real video stream examples to demonstrate that our method is effective not merely for static scenes but also when there exist objects moving independently of the scene. Actually separating moving objects from the background using real video images, we confirm that the separation accuracy indeed improves by our proposed technique.

Key words: feature tracking, outlier removal, subspace separation, robust estimation, RANSAC

1. まえがき

ビデオ画像からシーン中を移動する物体が存在するかどうかを判定し、存在すればそれを背景から分離することは多くのビデオ画像処理の基礎である。カメラが静止している場合にはフレーム間差分に基づく各種の手法が開発されているが、首振り監視カメラや移動する偵察車両からの撮影のようにカメラ自身が移動する場合は、移動する背景から独立に運動する物体を検出する必要があり、処理が極めて困難となる。

これに対してオプティカルフローや色やテクスチャなどに基づく多くの手法が研究されているが、Costeira・Kanade [1] は動画から3次元復元するTomasi-Kanadeの「因子分解法」[22]をヒントにして、特徴点の軌跡の分類をその画像座標の履歴から計算されるある行列の要素の非零判定に帰着させた。

その後、これをもとに種々の応用や改良が研究された。例えばGear [3]はCosteira・Kanade [1]とは異なる形の行列の標準形を用いてグラフのマッチングに帰着させた。市村・富田 [6]は行列のQR分解を用い、市村 [5]は大津の判別規準 [20]を適用した。Inoue・Urahama [10]はファジクラスタリングを用いた。

一方、黒澤・金谷 [15, 16]はCosteira-Kanadeの方法がTomasi-Kanadeの因子分解法とは無関係に、「アフィンカメラのもとでは剛体運動する物体上の特徴点の画像上の運動履歴がある4次元部分空間に拘束される」という事実のみからの数学的に帰結されることを示した。そして幾何学的AICによるモデル選択 [13]と最小メジアン法 [21]によるロバスト推定を組み合わせた「部分空間分離法」[15]、「アフィン空間分離法」[16]と呼ぶ効果的な分離手法を示した。牧ら [18, 17]はこの考え方を数学的に同じ構造を持つ運動物体の照度変化の解析に適用している。

しかし黒澤・金谷 [15, 16] の方法は, (i) 独立に運動する物体の個数は既知, (ii) 画像上の特徴点の追跡は正しい, という二つの仮定に基づいている. (i) の物体個数の推定は困難な問題であり, これまでの研究 [1, 3, 19] では発見的なしきい値処理しかされていない. これに対して金谷ら [14] は幾何学的 AIC および幾何学的 MDL による推定を提案している

一方 (ii) の特徴点の追跡は Kanade-Lucas-Tomasi の方法 [23] がよく知られているが, 特徴点が多いと誤った追跡が含まれる可能性がある. これを防ぐために市村ら [7, 8, 9] は非線形非ガウス型状態空間に基づく特徴点軌跡のフィルタリングや特徴点追跡の改良を試みている. 一方, Huynh ら [4] は静止したシーンの移動カメラによる動画像系列中の特徴点履歴に対して 4 次元部分空間をロバストに当てはめることによって誤まった追跡を排除する方法を試みている.

本稿では Huynh ら [4] の方法を拡張し, シーン中に運動物体が存在する場合にも有効な方法を示す. これは特徴点の運動履歴に運動物体の数に応じた部分空間をロバストに当てはめることによってアウトライアを除去するものである.

以下, 本手法の基本原理を述べ, 一連の処理を実ビデオ画像を用いて実験し, その有効性を確認する.

2. 特徴点追跡と部分空間

N 個の特徴点 $\{p_\alpha\}$ を M 枚の画像に渡って追跡し, 第 κ 画像における p_α の画像座標を $(x_{\kappa\alpha}, y_{\kappa\alpha}), \kappa = 1, \dots, M, \alpha = 1, \dots, N$ とする. そして α 番目の特徴点の軌跡を次の $2M$ 次元ベクトルで表す.

$$p_\alpha = (x_{1\alpha} \ y_{1\alpha} \ x_{2\alpha} \ y_{2\alpha} \ \cdots \ x_{M\alpha} \ y_{M\alpha})^\top \quad (1)$$

物体に任意の物体座標系を固定し, 特徴点 p_α のその物体座標系に関する座標を $(a_\alpha, b_\alpha, c_\alpha)$ とする. 時刻 κ での物体座標系の原点の位置と各座標の基底ベクトルをそれぞれ $t_\kappa, \{i_\kappa, j_\kappa, k_\kappa\}$ とすると, 特徴点 p_α の時刻 κ における位置 $r_{\kappa\alpha}$ は次式で表せる.

$$r_{\kappa\alpha} = t_\kappa + a_\alpha i_\kappa + b_\alpha j_\kappa + c_\alpha k_\kappa \quad (2)$$

平行投影や弱透視投影, 疑似透視投影を一般化したアフィンカメラ [11] のもとでは, 点 $r_{\kappa\alpha}$ は画像上の次の位置に投影される.

$$\begin{pmatrix} x_{\kappa\alpha} \\ y_{\kappa\alpha} \end{pmatrix} = A_\kappa r_{\kappa\alpha} + b_\kappa \quad (3)$$

ただし A_κ, b_κ はそれぞれ時刻 κ でのカメラの位置や内部パラメータによって定まる 2×3 行列および 2 次

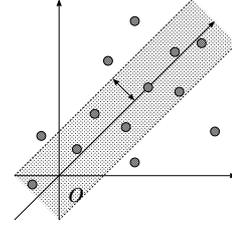


図 1: 部分空間の当てはめによるアウトライアの除去

元ベクトルである. 式 (2) を代入すると, これは次式のように書ける.

$$\begin{pmatrix} x_{\kappa\alpha} \\ y_{\kappa\alpha} \end{pmatrix} = \tilde{m}_{0\kappa} + a_\alpha \tilde{m}_{1\kappa} + b_\alpha \tilde{m}_{2\kappa} + c_\alpha \tilde{m}_{3\kappa} \quad (4)$$

ただし $\tilde{m}_{0\kappa}, \tilde{m}_{1\kappa}, \tilde{m}_{2\kappa}, \tilde{m}_{3\kappa}$ は時刻 κ でのカメラの位置や内部パラメータで決まる 2 次元ベクトルである. これを $\kappa = 1, \dots, M$ に渡って式 (1) のように縦に並べると, ベクトル p_α は次のように書ける.

$$p_\alpha = m_0 + a_\alpha m_1 + b_\alpha m_2 + c_\alpha m_3 \quad (5)$$

ただし $m_i, i = 0, 1, 2, 3$ は $\tilde{m}_{i\kappa}$ を $\kappa = 1, \dots, M$ に渡って縦に並べた $2M$ 次元ベクトルである.

これから, 同一物体に属する各特徴点 p_α の軌跡が $2M$ 次元空間の 1 点で表され, 点 p_α がベクトル m_0, m_1, m_2, m_3 の張る 4 次元部分空間に含まれることがわかる. これを「部分空間拘束条件」と呼ぶ [15].

同様に, シーン中に m 個の物体が独立に運動していると, すべての特徴点の履歴 $\{p_\alpha\}$ は一般に $4m$ 次元部分空間に含まれる.

3. 部分空間の当てはめ

前述の理論が成立するのは特徴点の追跡がすべて正しい場合であり, 誤った特徴点の軌跡は正しい部分空間には含まれない. Huynh ら [4] はこの知識に基づいて, 移動するカメラで撮影した静止シーンのビデオ画像上で追跡した特徴点の軌跡に 4 次元部分空間をロバストに当てはめることによって誤まった追跡を除去している. しかし前述のように, シーン中に運動物体が存在する場合にも部分空間拘束条件が成立する. そこで本稿では, RANSAC [2, 12] を用いたロバスト推定によって特徴点の軌跡ベクトルに $4m$ 次元部分空間を当てはめ, 当てはめた部分空間との距離がしきい値よりも大きい軌跡をアウトライアとして除去する (図 1).

4. アウトライア除去の手法

アウトライアの除去の手順は次のように行う. 以下, $n = 2M, d = 4m$ と置く.

1. 特徴点の運動履歴のベクトル $\{p_\alpha\}$, $\alpha = 1, \dots, N$ からランダムに d 個のベクトルを取り出し, q_1, q_2, \dots, q_d とする.
2. 取り出したベクトルの $n \times n$ モーメント行列

$$M_d = \sum_{i=1}^d q_i q_i^\top \quad (6)$$

を計算する.

3. 行列 M_d の大きい方から d 個の固有値 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_d$ と対応する単位固有ベクトルの正規直交系を $\{u_1, u_2, \dots, u_d\}$ を計算する.
4. $n \times n$ 射影行列 P_{n-d} を次のように計算する.

$$P_{n-d} = I - \sum_{i=1}^d u_i u_i^\top \quad (7)$$

5. 全データ $\{p_\alpha\}$ のうち

$$\|P_{n-d} p_\alpha\|^2 < (n-d)\sigma^2 \quad (8)$$

となるものの個数を S とする. ただし $\|P_{n-d} p_\alpha\|$ はベクトル p_α と当てはめた d 次元部分空間との距離である. また σ は特徴点の各座標の検出誤差の標準偏差である.

6. 以上の処理を反復し, S を最大とする射影行列 P_{n-d} を求める¹.
7. 次式を満たすベクトル p_α をアウトライアとして除去する.

$$\|P_{n-d} p_\alpha\|^2 \geq \sigma^2 \chi_{n-d;99}^2 \quad (9)$$

ここに $\chi_{r;a}^2$ は自由度 r の χ^2 分布の $a\%$ 点である. 特徴点の各座標に独立に期待値 0, 標準偏差 σ の正規分布に従う誤差が入るとすれば, n 次元空間の点 p_α の誤差の当てはめた d 次元空間に平行な成分 (d 次元) とそれに直交する成分 ($n-d$ 次元) とは互いに独立である. 点 p_α がインライアであり, 当てはめた d 次元空間が正しいとすると, それに直交する成分のノルムの二乗 (これを「残差」と呼ぶ) を σ^2 で割ったものは自由度 $n-d$ の χ^2 分布に従う. したがって残差の期待値は $(n-d)\sigma^2$ である. 上の手順ではそれ以下の点の個数が最大になるように d 次元空間を推定し, 有意水準 1% でインライアと見なされないものを除去している.

5. 実験

静止シーンと運動物体が存在するシーンについて, Kanade-Lucas-Tomasi の方法 [23] で追跡した特徴点データを用いて誤追跡点の除去実験を行った.

¹実験では 200 回連続して更新がないことを収束条件とした.

図 2(a), (b), (c) はカメラを移動しながら静止シーンを 100 フレームに渡って撮影した画像 (320×240 画素) の初期フレーム, 第 50 フレーム, および最終フレームであり, 追跡した 126 個の特徴点を \square 印で示す.

図 2(d) は $\sigma = 0.5$ としてアウトライアを除去した特徴点の軌跡であり, 図 2(e) は除去されたアウトライアの軌跡である. 図 2(f) は各特徴点の残差を表すグラフであり, 横線は式 (9) から得られるしきい値である. 図 2(g) に除去された特徴点を初期フレームに重ねて表示する.

アウトライアとされた軌跡が本当に誤っているかどうかを全フレームに渡って目視によってチェックした. 図 2(f) 中の \times 印は明らかに誤りであることがわかる軌跡であり, 残差が突出して大きい. しかし正しい軌跡でアウトライアとされたものもある. それらを注意深く調べると, 微妙に位置のずれた点を追跡した場合や, 追跡にふらつきのある点であった. このため正しい軌跡と比べて残差が若干大きくなり, アウトライアとみなされたものと思われる.

このように, 明らかな誤追跡点以外の点もアウトライアとして除去されているが, 曖昧な追跡は以降の 3 次元復元や運動物体の分離に悪影響を及ぼすので, ある程度信頼できない点は除去しても問題ないと思われる. 図 3 は同様に静止シーンを対象とした実験結果の例である.

図 4 は運動物体が存在する場合である. 図 4(a), (b), (c) は運動物体を注視するようにカメラを移動しながら 100 フレームに渡って撮影した画像 (320×240 画素) の初期, 第 50, 最終フレームであり, \square 印は追跡した 107 個の特徴点である.

図 4(d) に各特徴点の残差, および $\sigma = 0.5$ としたしきい値を示す. この場合は明らかな誤追跡点のみが除去されている. 図 4(e) にその除去された特徴点を初期フレームに重ねて表示する.

図 4(f) はアウトライアが運動物体の分離に与える影響を調べるために, アウトライアを除去しないデータに黒澤・金谷 [15] の部分空間分離法を適用した結果である². \square 印は背景とされた点, \times 印は運動物体とされた点を示す.

誤って追跡した特徴点が存在する場合にも正しく分離されるべきものは比較的良好に分離が行われている. しかし誤って追跡した特徴点の分離が誤っている. それらのほとんどは移動物体の隠れ境界上の点であり, 画像上のみかけの点を追跡していることがわかる. 一方, 図 4(g) はアウトライアを除去した

²以下に使用したプログラムを公開している.
<http://www.suri.it.okayama-u.ac.jp/>

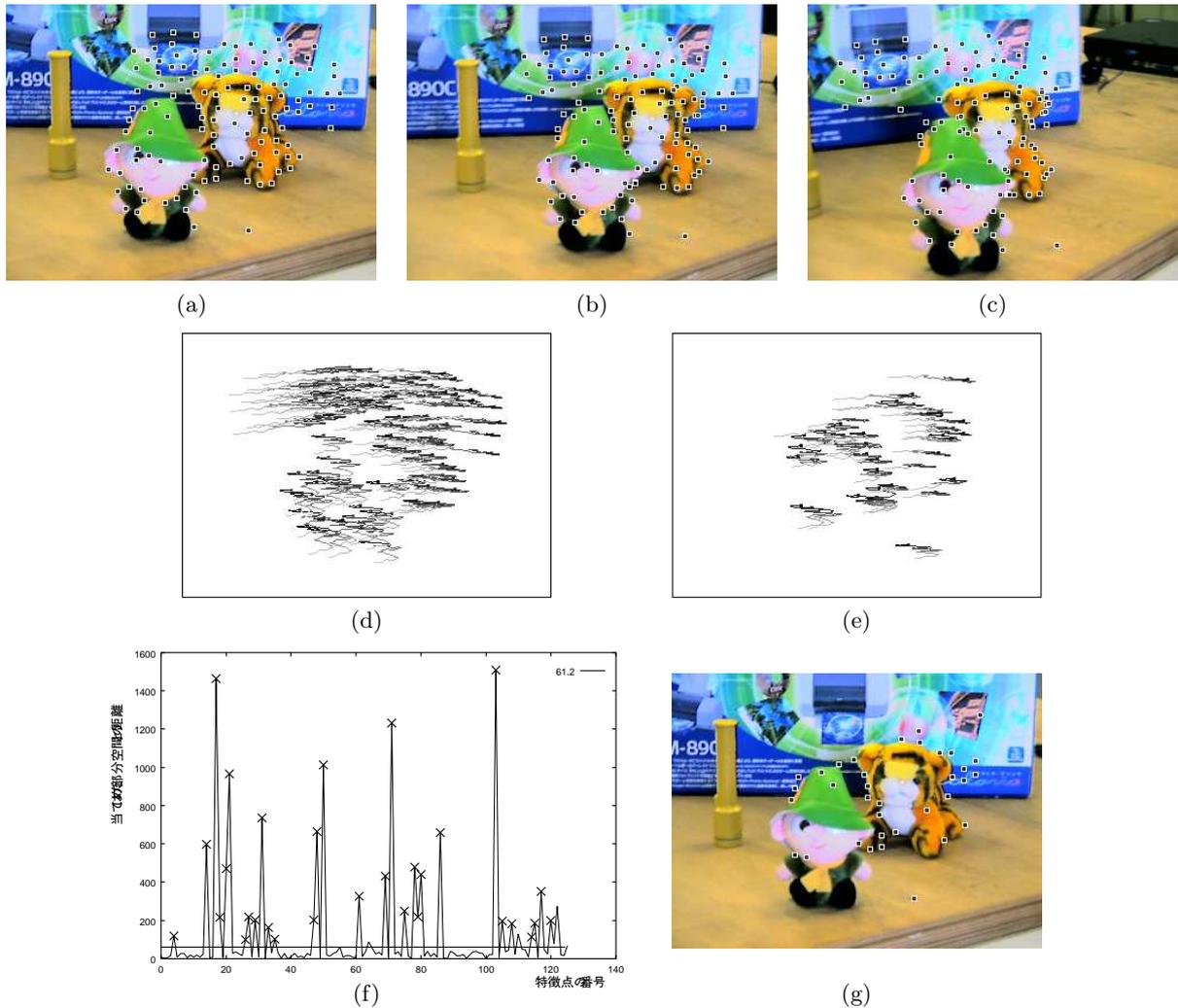


図 2: (a), (b), (c) 追跡した全特徴点 (第 1, 50, 100 フレーム), (d) インライアの軌跡, (e) アウトライアの軌跡, (f) 各特徴点の残差, (g) アウトライアとされた特徴点

データを用いた分離結果であり、正しく分離されていることがわかる。

図 4(h) にそれぞれの分離の正解率と分離度を示す。ただし分離度とは物体点で背景点とされたものの個数、および背景点で物体点とされたものの個数である。図 4(i), (j), (k) にそれぞれ判定されたアウトライア、背景点、および物体点の軌跡を示す。

図 5 は別のシーンの実験結果である。この場合も正しい軌跡でも除去されたものがあるが、アウトライア除去によって正解率が向上していることがわかる。

この例で特徴的なのは、図 5(d) に見られるように、インライアの残差が大きいものと小さいものの 2 種類にはっきり別れていることである。これは明確な特徴点は常に正しく追跡され、不明確な特徴点は数画素ずれたまま、あるいは変動しながら追跡されるためであると思われる。特にこの画像例では輝度値の範囲が狭く、濃淡変化が乏しいためそれが顕著に

なったものと思われる。

このような場合は、特徴点位置の誤差が特徴点ごと、フレームごとに独立であるという仮定が成立していない。ロバスト推定法としてよく知られている最小メジアン法 [21] はその仮定に基づき、インライアの残差から誤差の大きさを推定してアウトライアを除去している。したがってこの例のような場合には最小メジアン法は適用できない。そこで本稿では RANSAC[2, 12] を採用し、残差の不均一性を考慮して定数 σ を定めている。

6. まとめ

本稿では、3 次元形状復元や移動物体の分離を目的とし、ビデオ画像上の特徴点の軌跡から誤った追跡結果を除去する手法を提案した。これは部分空間拘束条件に基づき、特徴点の画像上の運動履歴に部分空間をロバスタに当てはめることによってアウトライアの除去を行うものである。

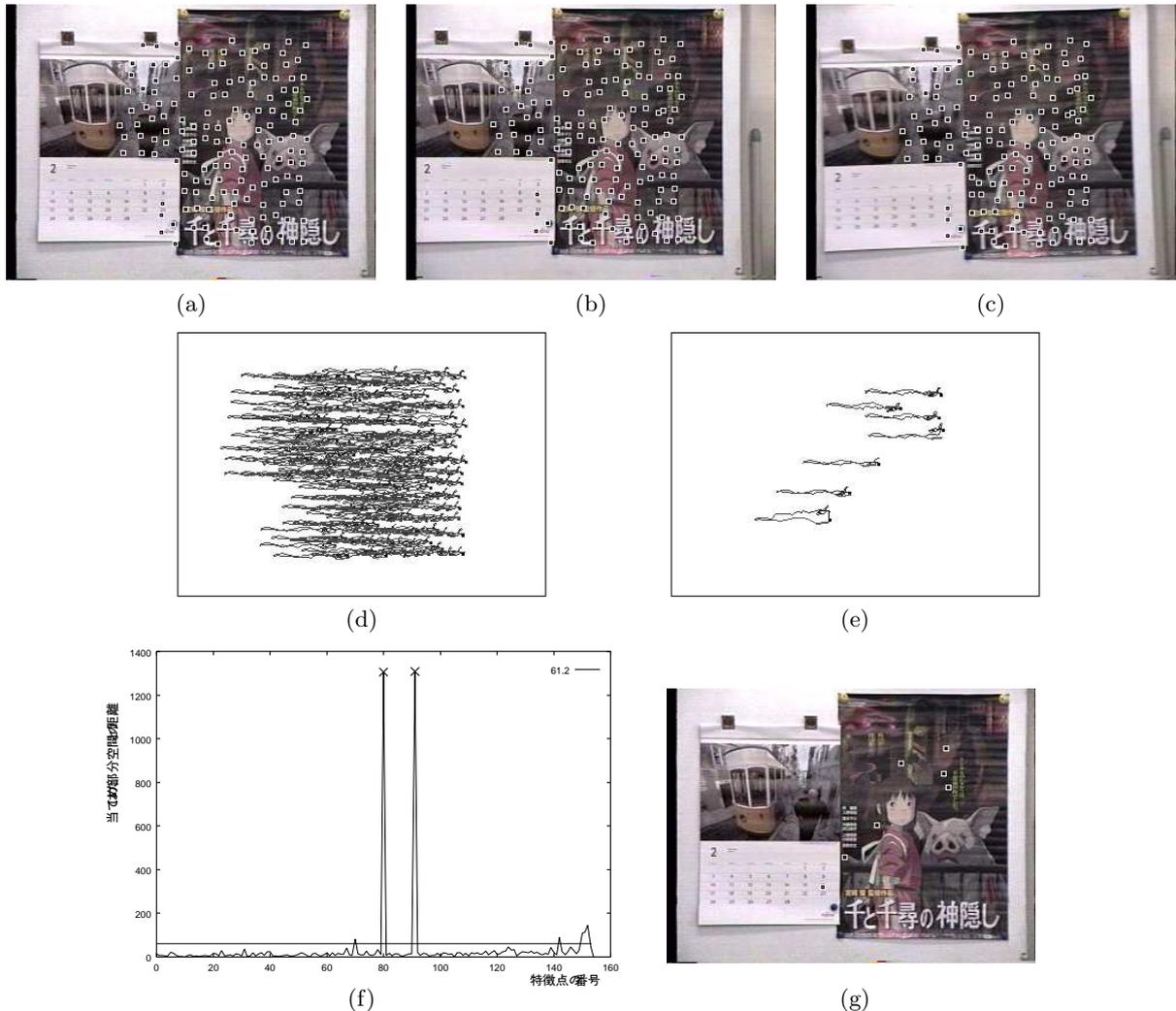


図 3: (a), (b), (c) 追跡した全特徴点 (第 1, 50, 100 フレーム), (d) インライアの軌跡, (e) アウトライアの軌跡, (f) 各特徴点の残差, (g) アウトライアとされた特徴点

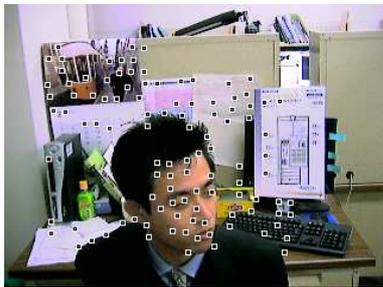
実験結果から、静止シーンだけではなくシーン中に運動物体が存在する場合にも提案手法が有効であることがわかった。そして実際に運動物体を移動する背景から分離し、提案手法によって分離の正解率が向上することを確認した。

市村ら [7, 9] のフィルタリングは正しい軌跡が“自然な動きをする”という確率モデルに基づくのに対して、提案手法は正しい軌跡が“剛体運動と解釈できる”という幾何学モデルに基づいている。両者は相補的な性格を持っているので、併用することによってさらに性能が向上すると期待される。

謝辞: 有益なコメントを頂いたオーストラリアの Murdoch 大学の Du Huynh 氏および産総研の市村直幸氏に感謝します。本研究の一部は文部科学省科学研究費基盤研究 C (2) (No. 13680432), テレコム先端技術研究支援センター, 柏森情報科学振興財団の助成による。

参考文献

- [1] J. P. Costeira and T. Kanade, “A multibody factorization method for independently moving objects,” *Int. J. Computer Vision*, **29**-3, pp. 159–179, Sept. 1998.
- [2] M. A. Fischer and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Comm. ACM*, **24**-6, pp. 381–395, Jun. 1981.
- [3] C. W. Gear, “Multibody grouping from motion images,” *Int. J. Comput. Vision*, **29**-2, pp. 133–150, Aug./Sept. 1998.
- [4] D. Q. Huynh and A. Heyden, “Outlier detection in video sequences under affine projection,” *Proc. of CVPR*, pp. 695–701, Dec. 2001.
- [5] 市村直幸, “形状空間への直交射影行列と判別基準を用いた複数運動の分割,” 情報処理学会研究報告, 2000-CVIM-120-3, pp. 17–24, Jan. 2000.
- [6] 市村直幸, 富田文明, “形状行列からの特徴選択に基づく動きの分割,” 電子情報通信学会論文誌 D-II, J81-



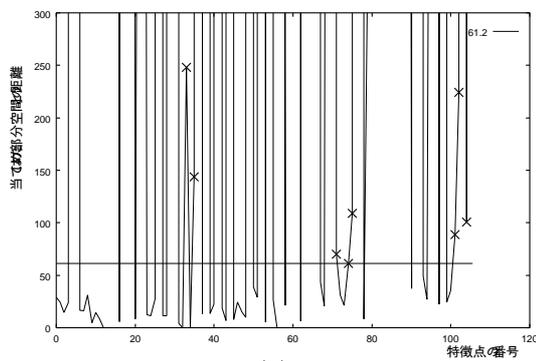
(a)



(b)



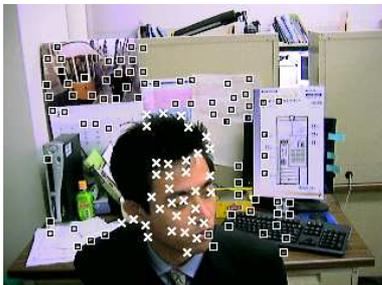
(c)



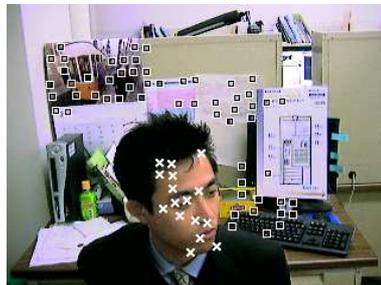
(d)



(e)



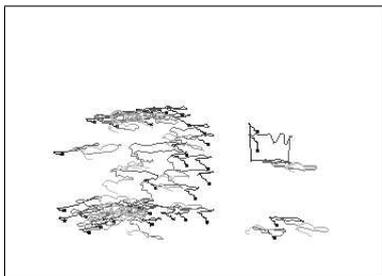
(f)



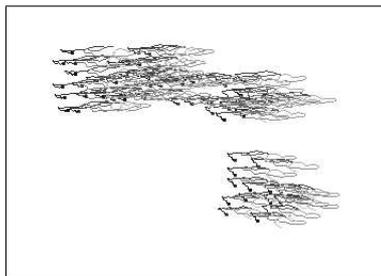
(g)

	正解率	分離度
アウトライアあり	94.4	6/40
アウトライアなし	100.0	0/16

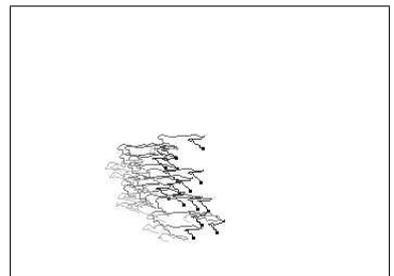
(h)



(i)

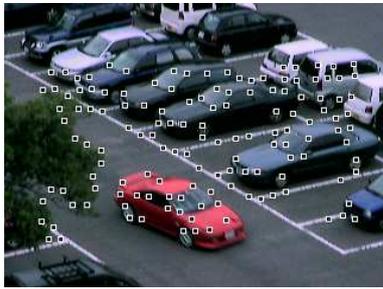


(j)

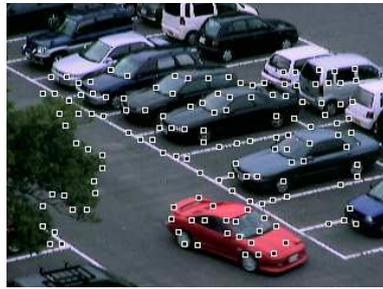


(k)

図 4: (a), (b), (c) 追跡した全特徴点 (第 1, 50, 100 フレーム), (d) 各特徴点の残差, (e) アウトライアとされた特徴点, (f) アウトライアを含んだデータからの運動物体の分離, (g) アウトライアを除去したデータからの運動物体の分離, (h) 分離の正解率と分離度, (i) アウトライアの軌跡, (j) 背景点の軌跡, (k) 物体点の軌跡



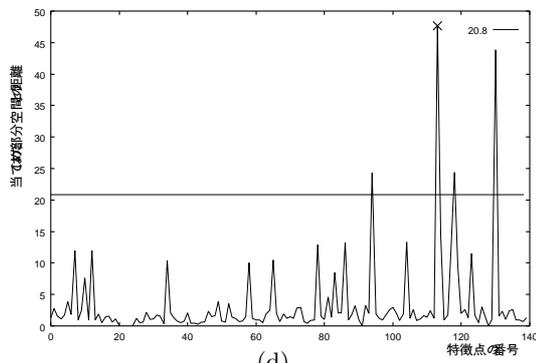
(a)



(b)



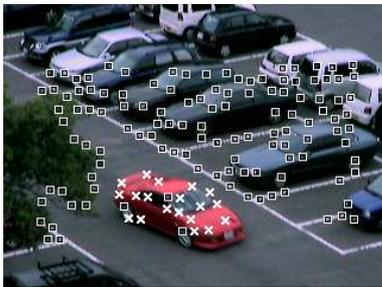
(c)



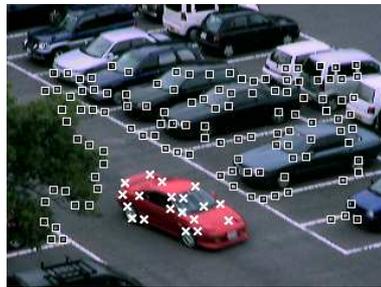
(d)



(e)



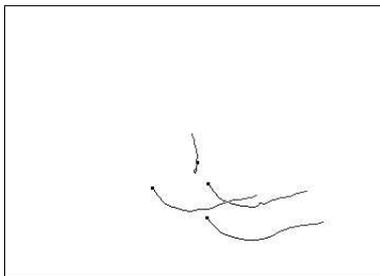
(f)



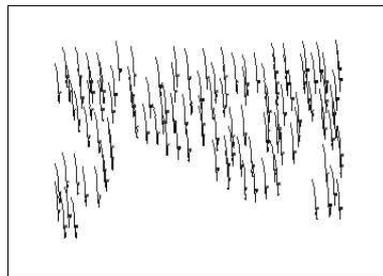
(g)

	正解率	分離度
アウトライアあり	97.9	3/22
アウトライアなし	100.0	0/19

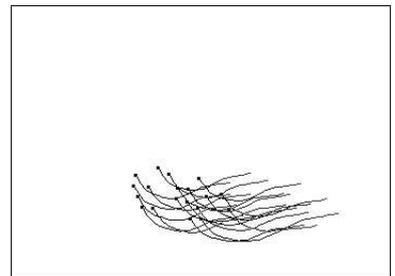
(h)



(i)



(j)



(k)

図 5: (a), (b), (c) 追跡した全特徴点 (第 1, 50, 100 フレーム), (d) 各特徴点の残差, (e) アウトライアとされた特徴点, (f) アウトライアを含んだデータからの運動物体の分離, (g) アウトライアを除去したデータからの運動物体の分離, (h) 分離の正解率と分離度, (i) アウトライアの軌跡, (j) 背景点の軌跡, (k) 物体点の軌跡

- D-II-12, pp. 2757–2766, Dec. 1998.
- [7] 市村直幸, “自己組織化型状態空間モデルを用いた運動軌跡のフィルタリング,” 情報処理学会研究報告, 2001-CVIM-128-2, pp. 9–16, July 2001.
- [8] 市村直幸, “フレーム毎の特徴点抽出に基づく特徴点の追跡,” 情報処理学会研究報告, 2001-CVIM-130-5, pp. 31–38, Nov. 2001.
- [9] 市村直幸, 生駒哲一, “非ガウス型状態空間状態モデルを用いた特徴点位置系列のフィルタリング,” 情報処理学会研究報告, 2000-CVIM-122-3, pp. 17–24, May 2000.
- [10] K. Inoue and K. Urahama, “Separation of multiple objects in motion images by clustering,” *Proc. 8th Int. Conf. Comput. Vision*, Vancouver, Canada, Vol. 1, pp. 219–224, July 2001.
- [11] 金出武雄, コンラッド・ポールマン, 森田俊彦, “因子分解法による物体形状とカメラ運動の復元,” 電子情報通信学会論文誌 D-II, **J74-D-II-8**, pp. 1497–1505, Aug. 1993.
- [12] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, U.K., 2000.
- [13] K. Kanatani, *Statistical Optimization for Geometric Computation: Theory and Practice*, Elsevier, Amsterdam, 1996.
- [14] 金谷健一, 黒澤典義, 松永力, “モデル選択によるランク推定と複数運動の分離,” 情報処理学会研究報告, 2001-CVIM-126-3, pp. 17–24, Jan. 2001.
- [15] 黒澤 典義, 金谷 健一, “部分空間分離法とモデル選択による運動物体の分離,” 情報処理学会研究報告, 2000-CVIM-124-4, pp. 25–32, Nov. 2000.
- [16] 黒澤 典義, 金谷 健一, “アフィン空間分離法による運動物体の分離,” 情報処理学会研究報告, 2001-CVIM-125-3, pp. 25–32, Mar. 2001.
- [17] 牧淳人, 服部寛, “輝度部分空間を用いた運動物体の分離,” 電子情報通信学会技術研究報告, PRMU2001-134, pp. 147–153, Nov. 2001.
- [18] 牧淳人, 渡邊睦, C. Wiles, “Geotensity 拘束による3次元形状獲得,” 電子情報通信学会論文誌 D-II, **J83-D-II-8**, pp. 1714–1752, Aug. 2000.
- [19] 長崎健, 川嶋稔夫, 青木由直, “因子分解法に基づく運動画像列解析による多関節物体の構造推定,” 電子情報通信学会論文誌 D-II, **J81-D-II-3**, pp. 483–492, March 1998.
- [20] 大津展之, “判別および最小2乗規準に基づく自動しきい値選定法,” 電子通信学会論文誌 D, **J63-D-4**, pp. 349–356, 1980.
- [21] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*, Wiley, New York, 1987.
- [22] C. Tomasi and T. Kanade, “Shape and motion from image streams under orthography—A factorization method,” *Int. J. Comput. Vision*, **9-2**, pp. 137–154, Nov. 1992.
- [23] C. Tomasi and T. Kanade, “Detection and Tracking of Point Features,” CMU Tech. Rep. CMU-CS-91-132, Apr. 1991; <http://vision.stanford.edu/~birch/klf/>.