

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

市村 直幸

電子技術総合研究所 知能情報部
〒 305-8568 つくば市梅園 1-1-4
ichimura@etl.go.jp

あらまし 均一な運動をする物体の形状が4次元の線形空間で表現されることに基づく動きの分割方法を提案する。動きの分割は、複数物体の形状空間を部分空間の集合へ分解することにより行う。その分解には、形状空間への直交射影行列の要素の判別分析を用いる。1次元のデータの判別分析しか必要としないため、アルゴリズムは簡潔である。判別分析の結果に基づき動きの分割に有用な情報をもつデータのみを選択するため、雑音や外れ値を含むデータに対しても提案方法は適用できる。実画像を用いたオフラインおよびオンラインの実験により、提案方法の有効性について検討した。

キーワード 複数運動、セグメンテーション、形状空間、判別基準

Motion Segmentation Using Orthogonal Projection Matrix of Shape Space and Discriminant Criterion

Naoyuki ICHIMURA

Machine Understanding Division
Electrotechnical Laboratory
1-1-4, Umezono, Tsukuba, Ibaraki, 305-8568 Japan
ichimura@etl.go.jp

Abstract A novel algorithm for motion segmentation is proposed. The algorithm uses the fact that shape of an object with homogeneous motion is represented as 4 dimensional linear space. Thus motion segmentation is done as the decomposition of shape space of multiple objects into a set of 4 dimensional subspace. The decomposition is realized using the discriminant analysis of orthogonal projection matrix of shape space. Since only discriminant analysis of 1D data is needed, this analysis is quite simple. The algorithm based on the analysis is robust for data with noise and outliers, because the analysis can extract useful information for motion segmentation while rejecting useless one. The implementation results show that the proposed method is robust and efficient enough to do online task for real scenes.

key words multiple motions, segmentation, shape space, discriminant criterion

1 はじめに

動きの分割は、時系列画像を用いた物体の形状復元や運動解析において必要とされる。今まで提案されている動きの分割アルゴリズムは、大きく三つのカテゴリに分類することができる。一つ目は、サンプリングと検証に基づく方法である [1] [2]。いくつかの特徴対応データをランダムにサンプリングし、それらが単一物体のエピソード拘束を満たすかどうかを検証する。二つ目は、ベイズ推定など確率分布に基づく方法である [3]-[5]。これら二つのカテゴリのアルゴリズムは、「鶏と卵」の問題を含んでいる。何故なら、前者は同一の物体に属する特徴対応データを取り出しやすくするため、ランダムサンプリングに空間的な近接性を取り入れているし、また、後者は動き情報の適切な初期値を必要とし、これらのことは、分割に用いる情報を得るために、事前に「各グループに関する情報」を必要とすることを意味しているからである。

三つめのカテゴリは、鶏と卵の問題を回避するため、複数運動に関する情報を同時に求める方法を用いる [6]-[8]。その中で、因子分解法に基づく方法がいくつか検討されている [7][8]。これらの方法では、基本的に、物体の形状を表す線形空間（以後、形状空間）への直交射影行列が分割に用いられる。直交射影行列は、特徴対応データを要素とする計測行列の特異値分解の結果得られる形状空間の基底より計算できる。Costeiraらは [7]、直交射影行列を形状相互作用行列と呼び、その行列の要素の和から成る評価関数に基づく最適化を分割に用いた。また、Gear[8]は、形状空間に基づいて構成される二部グラフ (bipartite graph) を用い最適化を行った。これらの最適化は、分割に伴う組合せ問題を解くために使用された。しかし、これらの方法には、雑音や外れ値 (outlier) を含む特徴対応データに対するロバスト性の欠如という欠点がある [9]。また、最適化のための計算量も、オンライン処理では問題となる。

著者は組合せ問題を解くため、分割に含まれる組合せ数を特徴選択によって減少させる方法を以前提案したが [10]、この方法もロバスト性には問題があった。そのため、本論文では、ロバスト性を意識した特徴選択に基づく方法を提案する [11]。提案方法では、判別基準に基づく特徴選択を導入する。特徴選択は、画像中の一つの特徴に対応する直交射影行列の一つの行の選択により行われる。分割に最も有用

な情報をもつ特徴を選択するため、直交射影行列の各行の要素に対して判別分析を行う。選択された特徴の判別分析の結果に基づいて、均一な動きをもつと判断される特徴のグループを一つ抽出する。他のグループを抽出するため、残った特徴に対し同様の処理を再帰的に適用する。特徴選択により分割に有用な情報を持たない特徴が自動的に除去されるため、提案方法は雑音や外れ値に対するロバスト性を有する。また、直交射影行列の1つの行しか用いないため、分割に伴う組合せ問題も生じない。さらに、特徴のグループを再帰的に抽出するため、物体数に関する先見的知識を必要としない。

2節では、 N 個の物体がシーンに含まれる場合の計測行列の導出と、その計測行列の性質を示す。3節では、形状空間の直交射影行列の定義と性質を示す。4節では、判別基準を用いた特徴選択に基づく動きの分割方法を提案する。5節では、オフラインおよびオンラインの実験結果を示す。6節で、本論文をまとめる。

2 複数物体から得られる計測行列

本節では、アフィン投影の下で複数物体から得られる計測行列を導出する。また、導出の結果から得られる計測行列の性質を示す。以下において、 N は物体数、 P^i は i 番目の物体の特徴数、 $P = \sum_{i=1}^N P^i$ 、 F はフレーム数である。

2.1 計測行列の導出

世界座標系での i 番目の物体上の特徴の座標を行列 D_s^i ($i = 1, \dots, N$) で表す。

$$D_s^i = \begin{pmatrix} x_1^i & x_2^i & \dots & x_{P^i}^i \\ y_1^i & y_2^i & \dots & y_{P^i}^i \\ z_1^i & z_2^i & \dots & z_{P^i}^i \\ 1 & 1 & \dots & 1 \end{pmatrix} \quad (1)$$

世界座標系での i 番目の物体の j 番目 ($j = 1, \dots, F$) のフレームにおける運動パラメータを行列 M_j^i で表す。行列 M_j^i は、回転行列 R_j^i と並行移動ベクトル t_j^i より構成される。

$$M_j^i = \begin{pmatrix} R_j^i & t_j^i \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2)$$

i 番目の物体の全フレームを通じての運動パラメータを行列 M^i で表す。

$$M^i = \left(M_1^i, M_2^i, \dots, M_F^i \right)^t \quad (3)$$

ここで、行列 D_s と M_w を

$$D_s = \text{diag} \left(D_s^1, D_s^2, \dots, D_s^N \right) \quad (4)$$

$$M_w = \left(M^1, M^2, \dots, M^N \right) \quad (5)$$

とすると、各特徴の各フレームにおける位置は、次の行列 W_w によって表現される。

$$W_w = M_w D_s \quad (6)$$

画像上での特徴の位置は、この行列の要素を投影した結果得られる。投影モデルとして、次式のアフィンカメラモデルを用いる。

$$\lambda \tilde{m} = P_a' \tilde{X} \quad (7)$$

$$P_a' = \begin{pmatrix} p_{11}' & p_{12}' & p_{13}' & p_{14}' \\ p_{21}' & p_{22}' & p_{23}' & p_{24}' \\ 0 & 0 & 0 & p_{34}' \end{pmatrix} \quad (8)$$

ここで、 $\tilde{m}$ と $\tilde{X}$ は同次座標で表した特徴の画像上および3次元空間内での座標である。 $\tilde{m}$ と $\tilde{X}$ が無限遠点をとらないとすると、式(7)、(8)は次式で表現できる。

$$m = P_a (X^t, 1)^t \quad (9)$$

$$P_a = \begin{pmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \end{pmatrix} \quad (10)$$

ここで、 m と X は、その同次座標での表現が $\tilde{m}$ および $\tilde{X}$ となるユークリッド座標である。

式(6)の行列の各座標に対し、各フレーム毎の投影行列 P_a^i ($i = 1, \dots, F$) を用いて投影を行うと、次式を得る。

$$\begin{aligned} W_s &= A_p M_w D_s \\ &= M D_s \end{aligned} \quad (11)$$

$$M = A_p M_w \quad (12)$$

$$A_p = \text{diag} \left(P_a^1, P_a^2, \dots, P_a^F \right) \quad (13)$$

W_s は、運動情報をもつ行列 M と形状情報をもつ行列 D_s の積に分解される。

行列 W_s には、最初から P^1 個の列は物体1の特徴対応データ、次の P^2 個の列は物体2の特徴対応データというように、セグメンテーション情報が入っている。しかし、通常このセグメンテーション情報は得られない。つまり、画像上での特徴の走査順番などの計測状況に依存して、 W_s の列の入れ替えが生じる。その影響を入れ替え行列 P_{col} で表し、入れ替えを行った特徴対応データを要素とする行列を W で表す。

$$W = W_s P_{col} \quad (14)$$

この行列 W を、計測行列と呼ぶ。

2.2 計測行列の性質

式(11)より、次の計測行列の性質が得られる。

(i) 計測行列は、運動および形状情報をもつ二つの行列に分解される。

(ii) 計測行列のランクは、 N 個の物体の場合 $4N$ となる。

これらは、計測行列を用いた動きの分割のための基礎となる性質である。

3 形状空間の直交射影行列

2.2で示した計測行列の性質を、動きの分割に利用するために、まず計測行列の分解を考える。計測行列は、特異値分解により、次のように分解できる。

$$W = U_r \Sigma_r V_r^t \quad (15)$$

ここで、 U_r および V_r は、計測行列の列空間および行空間の直交基底であり、 $r = 1, \dots, \min(2F, P)$ である。対角行列 Σ_r は特異値より構成される。

2.2の第一の性質より、計測行列の行空間、つまり、式(15)の行列 V_r の直交基底により張られる空間は、物体の形状を表現している。よって、行空間を形状空間と呼ぶ。

形状空間の次元は $4N$ である。また、形状空間は、1つの物体に対応する4次元の部分空間の集合に分解できる。このことは、2.2の第二の性質より明らかである。よって、形状空間の分解を通じて、動きの分割を行うことができる。

次式の形状空間への直交射影行列が、形状空間の分解に利用される。

$$X_{P \times P} = V_r V_r^t = (x_1, \dots, x_P)^t \quad (16)$$

この行列の大きさは $P \times P$ であり、その行および列共に、画像中で抽出された P 個の特徴に対応する。特徴対応データに雑音や外れ値が含まれない場合には、 X の要素 x_{ij} は次の性質をもつ [7] [12]。

$$x_{ij} \begin{cases} \neq 0, & i \text{ 番目の行と } j \text{ 番目の列に対応する} \\ & \text{特徴が同一物体に属する場合} \\ = 0, & i \text{ 番目の行と } j \text{ 番目の列に対応する} \\ & \text{特徴が異なる物体に属する場合} \end{cases} \quad (17)$$

雑音や外れ値がなければ、この性質から容易に均一な運動をもつ特徴のグループを見い出せる。しかし、雑音や外れ値を含む実データでは、この性質が完全

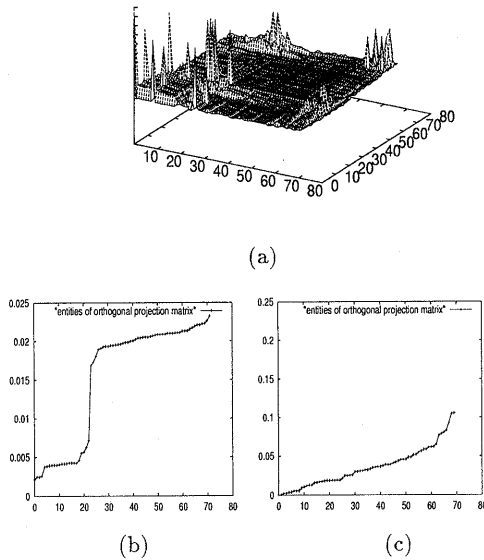


図 1: 分割に有用な情報をもつ特徴とそうでない特徴の例. (a) 実画像から得た直交射影行列. (b) 有用な情報をもつ特徴に対応する行のソートされた要素の変化. (c) 有用な情報をもたない特徴に対応する行のソートされた要素の変化.

に成立することは稀である。そのような実データに対するロバスト性の欠如が、従来の方法の問題点であった [7]-[10]。

4 特徴選択に基づく動きの分割

本節では、ロバスト性をもった動きの分割のために、特徴選択に基づく方法を提案する。

4.1 特徴選択の概要

提案方法では、分割のために最も有用な情報を有する一つの特徴のみを選択する。この特徴選択により、雑音や外れ値の影響が軽減されると共に、分割に伴う組合せ問題も生じない。

特徴選択の有用性を、例を用いて説明する (図 1)。二つのグラフ、図 1 (b) と (c) は、実際の時系列画像から得られた直交射影行列 (図 1 (a)) の、2つの行のソートされた要素の変化を示している。図 1 (b) の要素は、図 1 (c) の要素より分割に有用な情報を有している。何故なら、図 1 (b) の要素は容易に 2つのグループに分けられるからである。この例のように、全特徴の中の一部のみが、分割に有用な情報を有している。そのため、提案方法は、分割に有用

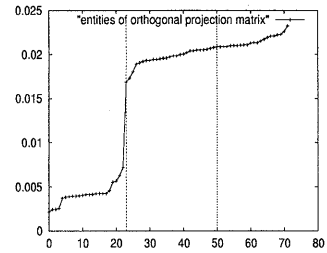


図 2: 判別基準の計算の例.

な情報をもつ特徴に対応する、図 1 (b) のような直交射影行列の行を選択する。

4.2 判別基準を用いた特徴選択

特徴選択の問題は、どのようにして分割に最も有用な情報をもつ一つの行を選択するかである。提案方法では、判別基準 [13] を選択に用いる。式 (15) の r が与えられた下で、直交射影行列 X を計算する。そして、 X の各行 x_k ($k = 1, \dots, P$) の要素をソートする。次式の判別基準により、行 x_k の要素を二つに分割する。

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2} \quad (18)$$

$$\sigma_B^2 = N^1 N^2 (\bar{\varepsilon}_1 - \bar{\varepsilon}_2)^2 \quad (19)$$

$$\sigma_W^2 = N^1 \sigma_1^2 + N^2 \sigma_2^2 \quad (20)$$

ここで、 $\bar{\varepsilon}_1$, $\bar{\varepsilon}_2$, σ_1^2 , σ_2^2 , N^1 , N^2 は各グループの要素の平均、分散、要素数であり、 σ_B^2 と σ_W^2 は二つのグループ間およびグループ内の分散である。

λ を最大化するグループ分けを見い出すため、行 $x_k = \{x_{k1}, x_{k2}, \dots, x_{kP}\}$ に対して、各要素 x_{kl} ($l = 1, \dots, P$) をしきい値として用いる。例えば、図 2 に示す一つの行の要素を、横軸の 50 にある線で二つのグループに分けたとする。この場合、左側のグループ内の分散が大きくなるため σ_W^2 も大きくなり、判別基準は大きな値とならない。一方、横軸の 23 にある線で二つのグループに分けたとすると、この場合の判別基準は 50 にある線を用いた場合に比べ大きくなる。何故なら、グループ間の分散が大きくなると共に、グループ内の分散が小さくなるためである。この例にあるように、要素に対する判別分析により、自動的に二つのグループ間の適切な境界を見い出せ、しきい値を必要としない。

λ_k を、 k 番目の行の最大の判別基準とする。図 2 の行では、23 の線に対応する判別基準となる。分割

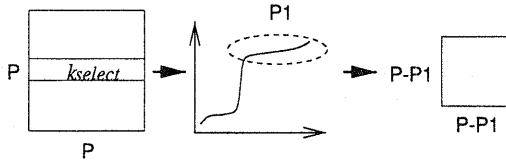


図 3: 分割のための再帰的処理.

に最も有用な情報をもつ特徴は、次式で選択される。

$$k_{select} = \arg \max_k \lambda_k \quad (21)$$

特徴選択の後に、 k_{select} 番目の行に対する判別分析の結果から、1つのグループを抽出できる (図3の左と真中の図)。次に、残った特徴に対応する要素を元の直交射影行列から取り出し、新しい行列を作る (図3の右の図)。そして、他のグループを取り出すために、その行列の各行に対して判別基準を計算する。この処理を、選択された特徴に対する判別基準が前もって与えられたしきい値 $\lambda_{threshold}$ より小さくなるまで、再帰的に繰り返す。

$$\lambda_{k_{select}} < \lambda_{threshold} \quad (22)$$

上式は、分割に有用な情報が直交射影行列に含まれていないことを意味している。

式 (15) のパラメータ r は、計測行列のランクが特異値より推定できる場合には、その値とする。しかし、実データを用いた場合、ランクの推定は困難である。本論文では、パラメータ r を分割の観点から決定した。上で説明した再帰的処理を、ある範囲の r 、例えば [3:12]、に対して適用する。そして、抽出された g 個のグループに対応する $\lambda_{k_{select}}$ の和 $\lambda(r)$ を最大化する分割結果を用いた。

$$\lambda(r) = \sum_{i=1}^g \lambda_{k_{select}}(i, r) \quad (23)$$

この式は、要素が最大の分離度をもつ直交射影行列に対応するランクを用いることを示している。

4.3 提案方法の特徴

提案方法の特徴を以下にまとめる。

- (i) ロバスト性を有する: 特徴選択によって有用な情報を持たないデータが除去される。
- (ii) 分割に伴う組合せ問題が生じない: 一つの行には P 個以下の組合せしかないので、判別基準の枚挙的な計算によってグループを抽出できる。

(iii) 簡潔な数値計算: 特異値分解および判別基準の計算よりアルゴリズムは構成される。

(iv) 物体数に関する先見的知識を必要としない: グループは再帰的処理によって逐次抽出される。

5 実験結果

オフラインおよびオンライン実験の結果を示す。オフライン実験の目的は、外れ値除去の効果を確認することである。オンライン実験では、実シーンにおけるロバスト性と計算量について検討する。すべての実験において、式 (22) のしきい値は 5.0、式 (15) の r は [3:12] の範囲で変化させた。画像からの特徴点抽出には、Chabat らの角検出器 [14] を用いた。

5.1 オフライン実験

車の時系列画像 (図4) では、10 フレームを用いた。カメラが車を追うように動いていたため、背景は静止していない。車と背景に対応する二つのグループが抽出され (図5 (a), (b))、残りは誤対応によって生じた不均一な動きをもつ特徴点に対応している (図5 (c))。このように、提案方法は、主要な動きを分割すると共に、外れ値を除去している。

人の歩行画像 (図6) でも、10 フレームを用いた。カメラは固定されていた。人と背景に対応する二つのグループが抽出された (図7 (a), (b))。誤対応と、背景と人の特徴点の動きに伴う重なりによって、不均一な動きが生じていた (図7 (c))。提案方法は、このような小さな動きの違いを検出できる。よって、外れ値を含んだデータから、均一で主要な動きを抽出できている。

5.2 オンライン実験

Pentium II 450 MHz および Pentium III 500 MHz をもつ2台のパーソナルコンピュータ (以下、PC1 および PC2) による並列処理を用いた。特徴点の追跡は PC1 で行い、マッチングには、正規化相関値を計算するハードウェアをもった画像ボード (日立 IP5010) を使用した。動きの分割は PC2 で 30 フレーム毎に行った。分割に用いる計測行列を PC1 より PC2 に送り、計算終了後、分割結果を PC1 に戻した。その結果に従い、各特徴点にグループ番号を割り当てた。二つの PC 間の通信には、PVM ライブラリ [15] を使用した。この並列処理によって、追跡を中断することなく、動きの分割を行うことができた。

図8および図9に結果を示す。これらの図の中の

正方形は、一つのグループと見做され、かつ、30フレームの間に10ピクセル以上の動きをもつ特徴点の重心位置を示している。

人の時系列画像では(図8)、背景の白い壁等で多くの誤対応が生じているにも関わらず、頭の動きが800フレーム以上に渡って分割、かつ、追跡された(図8(a)-(f))。

2冊の本の時系列画像では(図9)、2冊の本の動きが分離(図9(a)-(c))および統合(図9(d)-(f))されているにもかかわらず、動きが分割された。この結果は、提案方法が物体数に関する先見的知識を必要としないことを示している。このシーンでも、誤対応や背景と動いている領域の重なりによる外れ値が数多く含まれているが、2300フレーム以上に渡って主要な動きが分割、かつ、追跡された。

動きの分割の計算時間および特徴点の追跡のフレームレートは、約100個の特徴点に対し、300[ms]および10[frame/s]程度であった。これらの結果より、動きの分割の計算量が、通常のPCでオンライン処理可能な程度になっていると言える。

6 まとめ

形状空間の直交射影行列に基づく動きの分割方法を提案した。提案方法は、判別基準に基づく特徴選択によるロバスト性と計算の容易さを有している。オフラインおよびオンライン実験により、実シーンでの有効性を確認した。

提案方法は、式(11)の計測行列に基づいているため、アフィンカメラモデルによる投影の近似を必要とする。近似のない透視投影の下でのアルゴリズムの開発が、今後の課題の一つである。

謝辞

本研究の機会を与えて下さいました電総研知能情報部、大津展之部長に感謝致します。

参考文献

- [1] E. Nishimura, G. Xu and S. Tsuji: "Motion segmentation and correspondence using epipolar constraint," Proc. ACCV, pp.199-204, 1993
- [2] P. H. S. Torr: "Geometric motion segmentation and model selection," Phil. Trans. R. Soc. Lond. A, Vol.356, pp.1321-1340, 1998
- [3] D. W. Murray and B. F. Buxton: "Scene segmentation from visual motion using global optimization," IEEE Trans. Pattern Anal. & Mach. Intell., Vol.9, No.2, pp.220-228, 1987
- [4] J. Konrad and E. Dubois: "Bayesian estimation of motion vector fields," IEEE Trans. Pattern Anal. & Mach. Intell., Vol.14, No.9, pp.910-927, 1992
- [5] Y. Weiss and E. H. Adelson: "A unified mixture framework for motion segmentation: Incorporating spatial coherence and estimating the number of models," Proc. CVPR, pp.321-326, 1996
- [6] M. Shizawa: "Transparent 3D motions and structures from point correspondences in two frames: A quasi-optimal, closed-form, linear algorithm and degeneracy analysis," Proc. First Asian Conf. on Comp. Vis., pp.329-334, 1993
- [7] J. P. Costeira and T. Kanade: "A multi-body factorization method for independently moving objects," Internat. J. Comp. Vis., 29, 3, pp.159-179, 1998
- [8] C. W. Gear: "Multibody grouping from motion images," Internat. J. Comp. Vis., 29, 2, pp.133-150, 1998
- [9] T. Kanade and D. D. Morris: "Factorization methods for structure from motion," Phil. Trans. R. Soc. Lond. A, 356, pp.1153-1173, 1998
- [10] 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," 信学論D-II, Vol.J81-D-II, No.12, pp.2757-2766, 1998
- [11] N. Ichimura: "Motion segmentation based on factorization method and discriminant criterion," Proc. 7th Int. Conf. on Comp. Vis., Vol.I, pp.600-605, 1999
- [12] 金谷健一: "因子分解を用いない因子分解法: 複数物体分離," 信学技報, PRMU98-117, 1998
- [13] N. Otsu: "A threshold selection method from gray-level histograms," IEEE Trans. Sys., Man, and Cybern., Vol.SMC-9, No.1, pp.62-66, 1979
- [14] F. Chabat, G.Z. Yang and D.M. Hansell: "A corner orientation detector," Im. and Vis. Comp., 17, pp.761-769, 1999
- [15] http://www.epm.ornl.gov/pvm/pvm_home.html

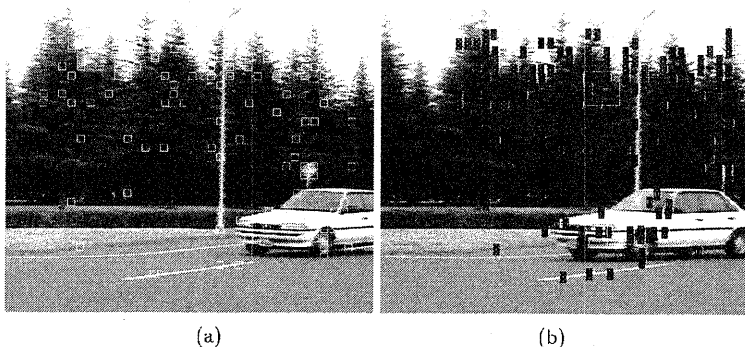


図 4: 車の時系列画像. (a) フレーム 1. (b) フレーム 10.

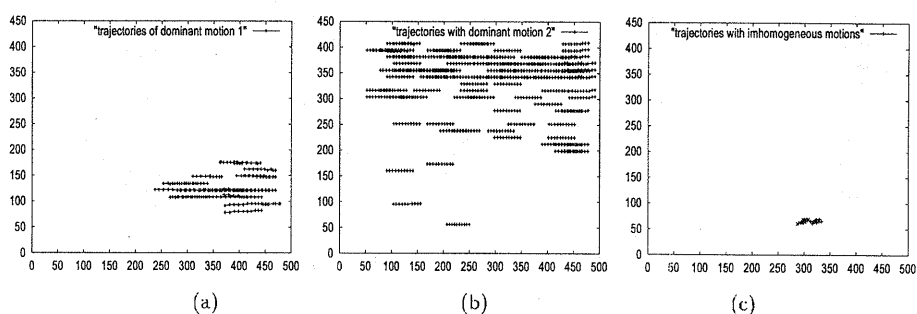


図 5: 車の時系列画像の分割結果. (a),(b) 主要な動きの特徴点の軌跡. (c) 外れ値と見做された特徴点の軌跡.

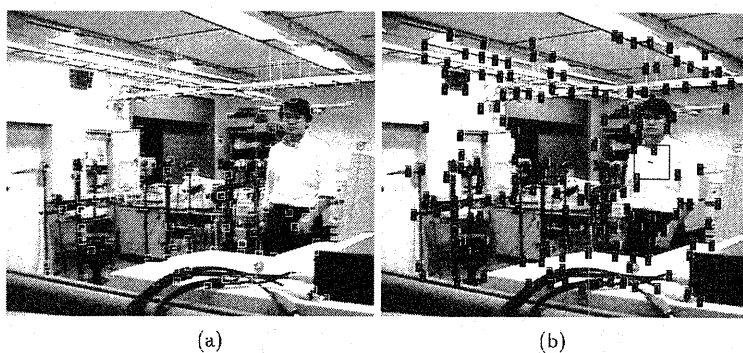


図 6: 人の歩行の時系列画像. (a) フレーム 1. (b) フレーム 10.

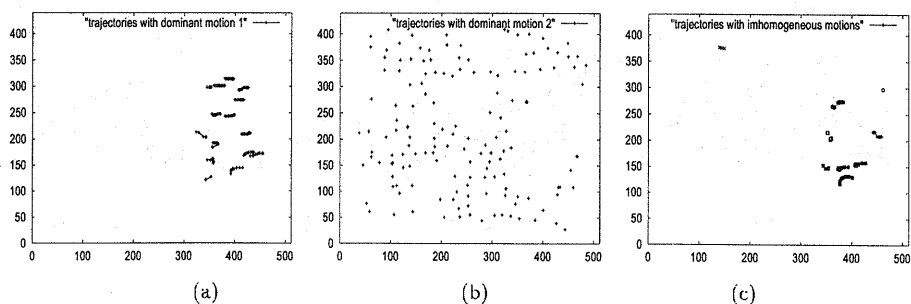


図 7: 人の歩行の時系列画像の分割結果. (a),(b) 主要な動きの特徴点の軌跡. (c) 外れ値と見做された特徴点の軌跡.

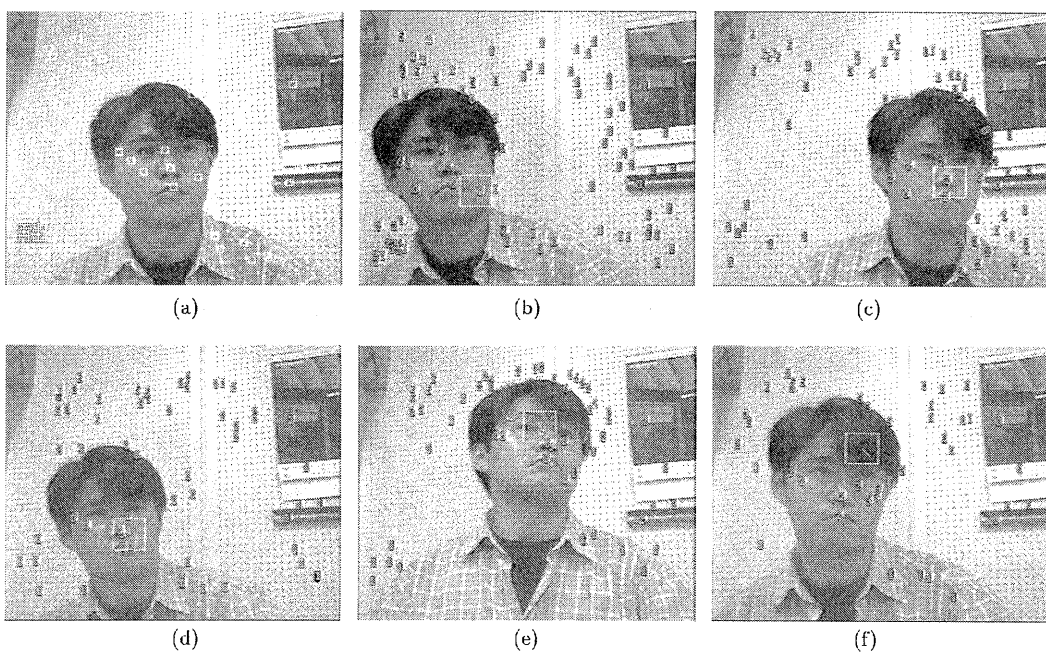


図 8: 人の動きのオンライン処理結果. (a)0 秒. (b)5 秒. (c)14 秒. (d)18 秒. (e)23 秒. (f)27 秒.

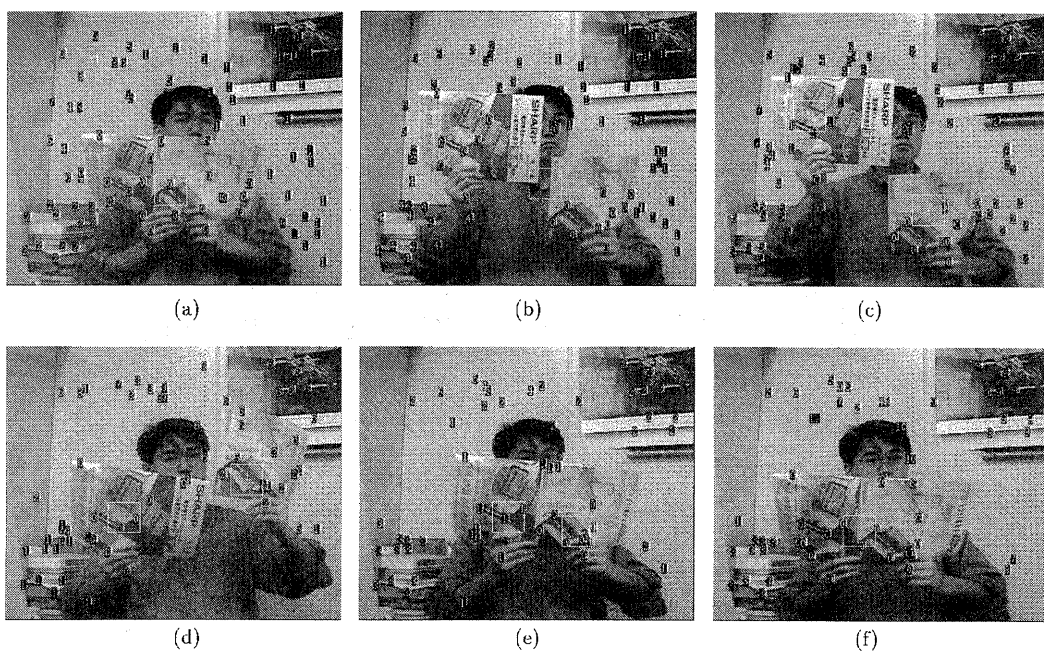


図 9: 2 冊の本の動きのオンライン処理結果. (a)9 秒. (b)27 秒. (c)28 秒. (d)41 秒. (e)75 秒. (f)77 秒.