

アフィン動きパラメータのロバスト推定に基づく動 領域の統合

著者	今村 幸祐, 橋本 秀雄
雑誌名	映像情報メディア学会誌 : 映像情報メディア = The journal of the Institute of Image Information and Television Engineers
巻	63
号	11
ページ	1625-1629
発行年	2009-11-01
URL	http://hdl.handle.net/2297/23250

アフィン動きパラメータのロバスト 推定に基づく動領域の統合

Merging of Moving Regions Based on Robust Estimation for Affine Motion Parameters

今村 幸祐[†], 正会員 橋本 秀雄[†]

Kousuke Imamura[†] and Hideo Hashimoto[†]

Abstract Moving object extraction in video sequences has been studied from various technical viewpoints. We propose a region-merging method for moving object extraction based on robust motion estimation. An over-segmented image is obtained by means of morphological watershed algorithm, and a robust motion estimation method using feature points and an evaluation function composed of motion estimation errors is then applied. Computer simulations demonstrated the merging precision of the proposed method.

キーワード：動オブジェクト抽出, 領域統合, アフィン動きモデル, ロバスト動き推定

1. ま え が き

動オブジェクト抽出は、映像から意味のある特徴データを得る処理として、非常に重要な技術である。抽出された動オブジェクトの情報は、例えば、MPEG-4 のオブジェクト符号化や MPEG-7 における動画像検索メタデータ生成のための特徴データなど、より高効率・高機能を目指した新しいアプリケーションで活用されている。

動オブジェクト抽出は、従来より様々なアプローチから研究が行われている^{1)~8)}。しかし、良好な抽出結果を得るためには、背景が静止である等の制限条件が多く、十分な性能と機能を備えた汎用的な手法の確立には至っていない。

動オブジェクト抽出手法は、画像の空間的情報と時間的情報を用いて動オブジェクトを抽出する手法が主流であり、時空間情報に基づく分解を行う手法と、空間的情報により画像を領域分割し動き等の特徴情報によりその領域を統合する手法の二つに大別できる。前者は多数のフレームから構成される時空間情報を扱うため、多くの処理量と記憶装置が必要となるのが一般的である。また、それらの対策のために、背景画像が既知あるいはモザイク処理などにより作成可能であるという条件が付加されているものが多い²⁾⁶⁾⁷⁾。それに対して、後者の空間情報による領域分割と時間情報による統合からなる手法は、少数フレーム単位での処理が可能となるため、応用範囲はより広いと考えられる。

空間的情報による領域分割には、Watershed アルゴリズム⁹⁾が用いられる。この手法により、動オブジェクトが複数の画像領域に分割されるが、オブジェクトを剛体と仮定して、同一の動きをする領域同士を統合することで動オブジェクトの抽出を行うことができる。この際、オブジェクトの抽出結果は、領域統合の精度に大きく依存する。つまり、領域の動き検出精度と統合評価の妥当性が、最終的な動オブジェクト抽出結果に大きく影響する。また領域統合合法に限らず、正確な領域の動きの推定と動きの類似度の評価は、他の動オブジェクト抽出手法においても重要な要素である。

本論文は、空間情報により分割された領域を、動き情報により統合することで、動オブジェクト抽出を行う手法において、特に領域統合における領域の動き推定と統合評価法を提案し、精度の高い領域統合手法の確立を目的とする。提案手法では、一般的な morphological watershed 法^{10)~12)}により、オブジェクトの境界が失われない程度に過分割が抑制された領域分割画像に対して、統合を行うものとする。まず、領域の正確な動きを導出するために、信頼できる特徴点から得られたオプティカルフローに対して、ロバスト推定を適用し、領域の正確なアフィン動きパラメータを算出する。領域統合評価においては、特徴点における動き推定誤差の変化量に基づいた評価関数を用いて精度の高い統合を行うものである。

2. 領域分割とアフィン動きモデルを用いた領域統合

提案法を用いた動オブジェクト抽出手法は、Watershed アルゴリズムを適用して空間的な領域分割を行い、領域のアフィン動き情報を用いて領域統合することで抽出を行う。

2009 年 5 月 27 日受付, 2009 年 8 月 31 日再受付, 2009 年 9 月 14 日採録
[†] 金沢大学 理工研究域電子情報学系
(〒 920-1192 石川県金沢市角間町, 076-234-4894)

[†] School of Electrical and Computer Engineering, College of Science and Engineering, Kanazawa University
(Kakuma-machi, Kanazawa-shi, Ishikawa, 920-1192, Japan)

各処理の詳細について、以降の節で述べる。

2.1 Watershed アルゴリズムによる領域分割

Watershed アルゴリズムは、地形学をモチーフにした領域拡大法の一つであり、画像の領域分割を行う代表的な手法である^{9)13)~15)}。Watershed アルゴリズムを自然画像に適用した場合、雑音や照明条件により、非常に多くの領域に過分割されてしまうため、Morphological 処理などにより過分割の抑制を伴う手法が適用されるのが一般的である^{10)~12)}。提案手法では、一般的な morphological watershed 法により、オブジェクトの境界が失われない程度に過分割が抑制された領域分割画像に対して、統合を行うものとする。

2.2 ロバスト推定法による領域の動き検出

(1) 特徴点の選択とオプティカルフローの検出

領域のアフィン動きパラメータを求めるために、まず領域の各点における動きの平行移動成分をブロックマッチング法により求める。動きの検出は、領域すべての画素に対してではなく、正確な動き情報が検出可能な点として、その点を中心としたブロックの輝度分散が高いものから順に特徴点として選択し、動きの検出対象とする。また、特徴点が領域中で偏って配置された場合、本来の動きと異なった領域の動きが算出される可能性もあるため、特徴点間の距離が領域の面積に依じて一定以上となるように設定し、一つの領域における特徴点の最大数も制限する。本論文では、領域あたりの特徴点の最大数は 100 点、動きの検出精度は半画素精度とする。

(2) 領域のロバストな動き推定

通常の最小 2 乗法では、すべての特徴点の動き情報を均等に扱うため、雑音や孤立点のような例外値が含まれた場合、その推定精度に悪影響を与える。提案手法では、特徴点ごとに誤り度を計算し、誤り度に応じて重みを決定するロバスト最小 2 乗法¹⁶⁾を行うことで、推定精度を向上させる。

特徴点の平行移動ベクトルから、下記のロバスト最小 2 乗法を用いて、領域の動き情報を推定する。

$$F = \sum_{i \in r} w(i) [\{v_{rx}(i) - v_x(i)\}^2 + \{v_{ry}(i) - v_y(i)\}^2] \quad (1)$$

ここで、 i は領域 r 内の特徴点であり、 $(v_x(i), v_y(i))$ は特徴点 i の平行移動ベクトル、 $(v_{rx}(i), v_{ry}(i))$ は、推定されるアフィンパラメータによる特徴点 i の変移量を表す。また、 $w(i)$ は重み係数を表す。この評価関数 F を最小にするアフィンパラメータが推定値となる。

特徴点から得られた平行移動ベクトルと領域の動きから推定された特徴点の変位量との誤差から得られる誤り度 e は次式のように定義される。

$$e(i) = |v_{rx}(i) - v_x(i)| + |v_{ry}(i) - v_y(i)| \quad (2)$$

この誤り度に従って各特徴点に与える重み $w(i)$ は次式で

定義される。

$$w(i) = \begin{cases} 1 - e(i)/(\mu_e + \sigma_e) \\ 0 \end{cases} \quad (\text{if } w(i) < 0) \quad (3)$$

ここで、 μ_e は領域の誤り度の平均値、 σ_e は標準偏差を表す。

式 (1) に示すように、重み係数 $w(i)$ を各特徴点 i の誤差量に乗じた最小 2 乗法を繰り返し適用することで、信頼性の高い領域の動き推定値を得ることができる。

2.3 動き推定誤差の変化量による領域統合

提案法では、隣接する 2 領域を仮に統合した場合の、統合の前後における動き推定誤差の変化量に基づいた評価関数を用いる。ある領域 s における動き推定誤差 E は次式で表される。

$$E(s) = \sum_{i \in s} [\{v_{sx}(i) - v_x(i)\}^2 + \{v_{sy}(i) - v_y(i)\}^2] \quad (4)$$

ここで、 $(v_x(i), v_y(i))$ は特徴点 i の平行移動ベクトルであり、 $(v_{sx}(i), v_{sy}(i))$ は、統合前のアフィンパラメータにより求めた特徴点 i の変位量を表す。また、領域 s, t を統合した場合の評価関数 $V(s, t)$ を、次式のように定義する。

$$\delta(s, t) = \frac{|E(s+t) - \{E(s) + E(t)\}|}{N_s + N_t} \quad (5)$$

ここで、 N_s, N_t は領域 s, t における特徴点数を表す。

領域の統合は、まず画像中のすべての隣接領域間において、評価関数の値 δ を求める。次に求めた中で最小の δ を持つ領域同士を統合する。統合後、統合による新領域とその隣接領域について評価関数を更新する。あらかじめ設定された領域数となるまで統合処理を繰り返し行い、最終的な領域統合の結果を得る。

3. シミュレーション実験

提案する領域統合法について計算機を用いたシミュレーション実験により評価を行う。

まず、動き推定の精度の評価のために、オブジェクトの形状と動きが既知の人為的に作成したテスト画像を用いて検証を行う。テスト画像は、Table Tennis のあるフレームを静止背景として、Mobile and Calendar の 1 部 (120×120 画素の矩形領域) をオブジェクトとしてはめ込み、動きとして画像の中心を原点に反時計回りに 3 度回転させ、その後左方向に 3 画素平行移動させたものである。テスト画像の領域分割結果に、各領域における検出した特徴点とその位置の動きの平行移動成分を描画したものを図 1 に示す。図 1 より、特徴点が面積に応じて均一に配置されており、各点においてほぼ正確な向きの動きが検出されている。これにより、特徴点選択の効果が視覚的に確認できる。

画像中の点 (x, y) におけるアフィンモデルによる動き (u, v) は、式 (6) によって表される。

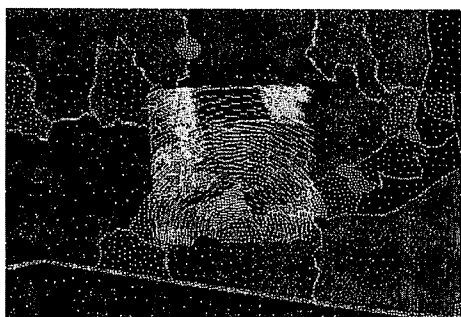


図 1 特徴点と動き情報
Feature points and motion information.

表 1 テスト画像のアフィン動きパラメータ
Affine motion parameters of test sequence.

	a	b	c	d	e	f
背景	1.0000	0.0000	0.0000	0.0000	1.0000	0.0000
オブジェクト	0.9986	-0.0523	3.0000	0.0523	0.9986	0.0000

表 2 動き推定の比較
Comparison of motion estimation.

最小 2 乗法		ロバスト最小 2 乗法	
すべての点	特徴点	すべての点	特徴点
1.3928	0.6114	0.7912	0.2984

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} ax + by + c \\ dx + ey + f \end{pmatrix} \quad (6)$$

ここで、 a, b, d, e のパラメータは拡大・縮小および回転を表すパラメータ、 c, f は平行移動を表すパラメータである。テスト画像における背景とオブジェクトの正確なアフィン動きパラメータを表 1 に示す。これを用いて、領域内のすべての点または特徴点のみから得られたオブティカルフローから、最小 2 乗法とロバスト最小 2 乗法により求めた領域の動きの推定精度を比較する。表 1 に示す正確なアフィンパラメータと得られた領域の動き推定値の誤差（マハラノビス距離）の平均値を表 2 に示す。表 2 より、特徴点のみを用いることで推定誤差を削減することができている。またロバスト処理を行うことで、さらに推定精度の向上が実現している。これより、提案法を用いることで、より正確な領域の動き推定が可能となったことが確認できる。

また、ロバスト最小 2 乗法における動き推定誤差の収束性を示す。図 2 は、テスト画像におけるロバスト推定時の繰り返し処理回数に対する評価関数値 F の特徴点あたりの平均値を示したものである。図 2 より、1 回目の繰り返し処理で初期値から大きく推定値の補正が行われ、2 回目以降の変動値は減少傾向にあることを示している。したがって、数回の繰り返し処理でほぼ収束しているといえる。ここで、 F の値が単調に減少していないのは、繰り返し処理における動きパラメータの推定値の変動により、推定値からの誤差の平均値と標準偏差に基づいて決定される各特徴点の重み係数 $w(i)$ が繰り返し処理中は一定とならないためである。

次に式 (5) に示す動き推定誤差の変位量に基づく評価関

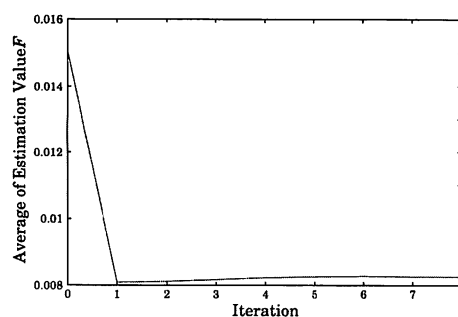


図 2 ロバスト動き推定誤差の収束性
Convergence of robust motion estimation error.

数について検証する。ここでは比較対象として、多くの画像分割に用いられている文献 15) の評価関数と輝度と動き情報の両方を考慮した文献 8) の評価関数を、前章のロバスト動き推定により求めた動きを用いた関数に修正したものを用いる。それぞれの評価関数を式 (7) と式 (8) に示す。

$$\begin{aligned} \delta_H(s, t) &= \frac{S_s \cdot S_t}{S_s + S_t} (C_s - C_t)^T M (C_s - C_t) \\ &= \frac{S_s \cdot S_t}{S_s + S_t} \cdot D^2(C_s, C_t) \end{aligned} \quad (7)$$

ここで S は領域の画素数を表し、 C は領域の代表アフィンパラメータを表す。また M はパラメータセットの共分散行列を表す。つまり、 D^2 はパラメータ間のマハラノビス距離を表す。

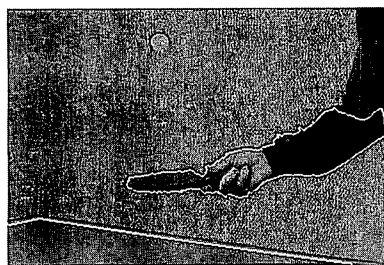
$$\begin{aligned} \delta_I(s, t) &= W \cdot \left\{ \frac{\sum_{i \in s} \{\mu(l_{s+t}) - l_s(i)\}^2}{S_s} \right. \\ &\quad \left. + \frac{\sum_{i \in t} \{\mu(l_{s+t}) - l_t(i)\}^2}{S_t} \right\} / \sigma_l^2 \\ &\quad + (1 - W) \cdot \left\{ \frac{\sum_{i \in s} D^2(V_{s+t}(i), V_s(i))}{N_s} \right. \\ &\quad \left. + \frac{\sum_{i \in t} D^2(V_{s+t}(i), V_t(i))}{N_t} \right\} \end{aligned} \quad (8)$$

ここで W は重み係数であり、 μ は領域内画素値 $l(i)$ の平均値、 σ^2 はその分散を表す。

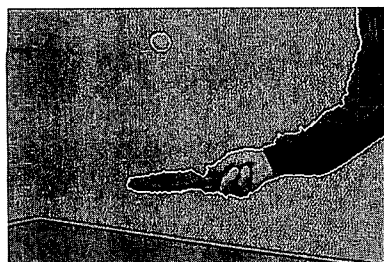
これらの評価関数における動きの評価は、評価関数 (7) が代表ベクトル同士の誤差、評価関数 (8) が統合による推定誤差の量に基づくものであるといえる。それに対し、提案法の評価関数は、統合前後の推定誤差の変位量に基づくものである。

自然画像 Table Tennis と Flower における各評価関数による動オブジェクト抽出の結果を図 3 と図 4 に示す。いずれも最終的な領域数が 3 となるまで処理を行い、領域分割結果の境界線を原画像上に描画した。初期領域数は、Table Tennis が 13、Flower が 121 領域である。

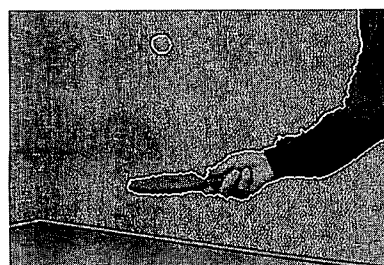
図 3 より、評価関数 式 (7) を用いた場合は、球や上腕の部分が背景と統合され、卓球台が一つの領域となった。式 (7) における統合領域の面積積の影響で小さい球の領域は早い段階で背景領域と統合されてしまった。それに対し、評価関数 式 (8) と提案法の場合には、球と腕の部分を正確に



(a) 評価関数 式 (7) (代表ベクトルの誤差)



(b) 評価関数 式 (8) (統合による推定誤差)



(c) 提案法

図 3 統合結果の比較 (Table Tennis)
Comparison of merging result (Table Tennis).



(a) 評価関数 式 (7) (代表ベクトルの誤差)



(b) 評価関数 式 (8) (統合による推定誤差)



(c) 提案法

図 4 統合結果の比較 (Flower)
Comparison of merging result (Flower).

統合することができた。このとき、評価関数 式 (8) の重み係数 W は 0.3 にしたが、0.6 などの高い数値に設定した場合、手首と腕の輝度差の影響で手首が背景に統合されるという影響が見られた。

また、図 4 の結果より、評価関数 式 (7) と式 (8) による統合では、前景、後景、樹木の分割を行うことができなかった。しかし、提案法ではそれらを別領域として抽出することができ、ロバスト性が高いことが確認できた。一部の空の領域はほとんど輝度値の変化がなく、ロバスト推定を用いても正確な動きが得られないため、提案法においても誤った統合がみられた。

これらの実験結果により、統合前後の推定誤差の変位量に基づく提案手法は、動きによる統合を他の評価関数と比較して正確に行うことができることを確認できた。

4. む す び

本論文では、動オブジェクト抽出を目的として、領域の動きのロバストな推定法と動きによる領域統合のための評価関数を提案した。シミュレーション実験により、提案手法を用いることで領域は正確に統合され、その結果、良好な動オブジェクト抽出結果が得られることが確認できた。今後は、他手法との組合せ効果の検証や統合を自動的に適切な領域数で終了する手法の検討を行う。

【文 献】

- 1) T. Brox, A. Bruhn, and J. Weickert: "Variational Motion Segmentation with Level Sets", European Conference on Computer Vision (ECCV), Springer, LNCS, **3951**, pp.471-483 (2006)
- 2) L.-H. Chen, Y.-C. Lai, C.-W. Su, and H.-Y.M. Liao: "Extraction of Video Object with Complex Motion", Pattern Recognition Letters, **25**, 11, pp.1285-1291 (2004)
- 3) W. Yang, K.-F. Loe, T. Tan, and W. Jian-Kang: "Spatiotemporal Video Segmentation based on Graphical Models", IEEE Trans. Image Process., **14**, 7, pp.937-947 (2005)
- 4) S.-Y. Chien, Y.-W. Huang, and L.-G. Chen: "Predictive Watershed: A Fast Watershed Algorithm for Video Segmentation", IEEE Trans. Circuits & Syst. Video Technol., **13**, 5, pp.453-461 (2003)
- 5) T. Schoenemann, and D. Cremers: "Near Real-time Motion Segmentation Using Graph Cuts", Pattern REcognition (Proc. DAGM), Springer, LNCS, **4174**, pp.455-464 (2006)
- 6) 堺田 慎一, 苗村 昌秀, 金次 保明: "背景差分法と時空間 watershed による領域成長法を併用した動画像オブジェクト抽出", 信学論, **J84-D2**, 12, pp.2541-2555 (2001)
- 7) 浜崎 唯史, 吉田 俊之: "3 次元 Multivalued Watershed 法を用いた動画像物体トラッキング", 映像学誌, **58**, 5, pp.695-704 (2004)
- 8) 池田 研二, 今村 幸祐, 橋本 秀雄: "画像の時間・空間情報を利用した動領域セグメンテーションとトラッキング", 映像学誌, **56**, 8, pp.1349-1353, (2002)
- 9) L. Vincent and P. Soille: "Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulations", IEEE Trans. Pattern Anal. Machine Intell., **13**, 6, pp.583-598, (1991)
- 10) D. Cortez et al.: "Image Segmentation Towards New Image Representation Methods", Signal Processing: Image Communication, **6**, pp.485-498 (1995)
- 11) D. Wang: "A Multiscale Gradient Algorithm for Image Segmen-

- tation Using Watersheds", Pattern Recognition, 30, 12, pp.2043-2052 (1997)
- 12) D. Vincent: "Morphological Grayscale Reconstruction in Image Analysis: Applications and Efficient Algorithm", IEEE Trans. Image Process., 2, 2, pp.177-201 (1993)
 - 13) 趙延軍, 長谷山美紀, 北島秀夫: "色エッジおよび並列実行の領域統合を用いた Watershed 画像領域分割法の改善", 信学論, J89-D, 4, pp.836-849 (2006)
 - 14) J. Cousty, G. Bertrand, M. Couprie, and L. Najman: "Fusion Graph: Merging Properties and Watersheds", Journal of Mathematical Imaging and Vision, 30, 1, pp.87-104 (2008)
 - 15) K. Haris, S.N. Efstratiadis, N. Maglaveras, and A.K. Katsaggelos: "Hybrid Image Segmentation Using Watersheds and Fast Region Merging", IEEE Trans. Image Process., 7, 12, pp.1684-1699 (1998)
 - 16) T. Papadimitriou, K. Diamantaras, M.G. Strintzis and M. Roumeliotis: "Video Scene Segmentation Using Spatial Contours and 3-D Robust Motion Estimation", IEEE Trans. Circuits & Syst. Video Technol., 14, 4, pp.485-497, (2004)



いまむら こうすけ
今村 幸祐 1995 年, 長崎大学工学部電気情報工学科卒業。1997 年, 同大大学院工学研究科修了。2000 年, 同大学大学院海洋生産科学研究科修了。同年, 金沢大学工学部情報システム工学科助手。2005 年より, 同大大学院自然科学研究科講師。博士(工学)。主として画像符号化に関する研究に従事。



はしもと ひでお
橋本 秀雄 1968 年, 大阪大学工学部卒業。1970 年, 同大大学院修士課程修了。1975 年, 同大大学院博士課程修了。同年, 電電公社(現 NTT)電気通信研究所入所。以来, 画像符号化方式, 画像通信方式の研究開発に従事。現在, 金沢大学大学院自然科学研究科教授。工学博士。正会員。