

NAIST-IS-MT9551003

修士論文

ロバスト統計を用いた背景位置合せに基づく 移動カメラ画像からの移動物体抽出

朝岡 忠

1997年2月14日

奈良先端科学技術大学院大学
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士(工学)授与の要件として提出した修士論文である。

朝岡 忠

指導教官： 横矢 直和 教授
千原 國宏 教授
竹村 治雄 助教授

ロバスト統計を用いた背景位置合せに基づく 移動カメラ画像からの移動物体抽出*

朝岡 忠

内容梗概

動画画像からの移動物体抽出は動画画像処理における重要な課題の一つである。特に移動カメラでは移動物体以外の背景にも見かけの動きが生じるために移動物体の抽出が困難となる。本論文では環境やカメラの動きが不明な状況下において移動カメラ画像から移動物体を抽出する手法を提案する。提案手法では画像間の対応点からロバスト統計を用いて背景の見かけの動きを推定し、その推定結果を基に背景位置合せ（カメラ移動による見かけの動きをキャンセル）された画像間での輝度の変化を利用して移動物体領域を抽出する。提案手法の特徴は安定な背景の動き推定を行うために M-推定と LMS 推定を組合せて用いていること、および信頼性の高い移動物体抽出を行うために背景位置合せの誤差を考慮して背景と移動物体を分離していることである。本論文では実画像を用いた移動物体の抽出実験により提案手法の有効性を示す。

キーワード

動画画像処理, 移動物体抽出, 移動カメラ, ロバスト統計, M-推定, LMS 推定, 位置合せ

*奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 修士論文, NAIST-IS-MT9551003, 1997年2月14日.

Motion Detection from Image Sequences with a Moving Camera Using Robust Statistics*

Tadashi Asaoka

Abstract

It is important to detect moving objects from image sequences in computer vision. But it is difficult to detect moving objects from image sequences obtained by a moving camera because of the apparent background motion. In this thesis, I propose a method to detect moving objects from image sequences with a moving camera. First, background motion parameters are estimated with point correspondences using robust statistics. In this estimation I use a combination of M-estimation and LMS estimation for stable estimation of apparent background motion. The backgrounds are then registered using estimated motion parameters. Finally, moving objects are extracted from registered adjacent image frames. In this motion detection I allow for errors in this registration and divide the original images into moving objects and backgrounds for reliable motion detection. Experimental results show the feasibility of the proposed method.

Keywords:

Image sequences, Motion Detection, Moving camera, Robust statistics,
M-estimation, LMS estimation, Registration

*Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT9551003, February 14, 1997.

目次

1. はじめに	1
2. ロバスト統計を用いた背景の動き推定	4
2.1 背景の動き推定に関する仮定	4
2.2 画像間の対応点探索	5
2.3 ロバスト統計の背景の動き推定への適用	6
2.4 LMS 推定による背景の動きパラメータ推定—GA による解法	7
2.5 M-推定による動きパラメータの精密化	8
3. 背景位置合せを用いた移動物体抽出	10
3.1 移動物体抽出のアルゴリズム	10
3.2 許容領域としきい値の設定	11
4. 移動物体の抽出実験	13
4.1 背景の動きが射影変換で良好に近似できる場合	14
4.1.1 実験1—理想的な状況 (ゲームのみ)	14
4.1.2 実験2—移動物体の注視 (首振りのみ)	14
4.1.3 実験3—カメラの複合移動 (ロール+首振り)	15
4.1.4 実験4—複数移動物体 (ゲーム+首振り)	15
4.2 実験5—背景の動きモデルに比較的大きな近似誤差がある場合 (従来のフレーム間差分による移動物体抽出との比較)	15
4.3 考察	16
5. むすび	23
謝辞	24
参考文献	25

目 次

2.1	特徴点抽出のためのウィンドウの設定	6
3.1	背景位置合せの誤差を考慮した非類似度	11
4.1	移動物体の抽出結果 実験1：ズームイン)	18
4.2	移動物体の抽出結果 実験2：左方への首振り)	19
4.3	移動物体の抽出結果 実験3：時計回りのロール+右方への首振り)	20
4.4	移動物体の抽出結果 実験4：ズームアウト+右方への首振り)	21
4.5	移動物体の抽出結果 実験5：カメラ視点の前方への平行移動)	22

1. はじめに

動画画像からの移動物体抽出は動画画像処理における重要な課題の1つであるが、移動カメラでは移動物体以外の背景にも見かけの動きが生じるために移動物体の抽出が困難となる。移動カメラ画像からの移動物体抽出は、例えば移動する自動車から前方をカメラで見て他の自動車や歩行者などの移動障害物を発見するような状況で必要となる。本論文では環境やカメラの動きが不明な状況下において移動カメラ画像から移動物体を抽出する手法について述べる。

これまで移動カメラ画像からの移動物体抽出に関する様々な手法が提案されてきた。代表的な手法として極座標変換を用いる方法 [1][2]、オプティカルフローを用いる方法 [3][4][5]、統計的な検定を用いる方法 [6]、背景位置合せを用いる方法 [7] などがある。

極座標変換を用いる方法では各画像に対して極座標変換（横軸に距離、縦軸に角度）を行う。カメラ運動が並進のみの場合、背景は画像上の1点（Focus of Expansion : FOE）からの放射状の動きとして観測されるので、FOEを原点とする上記の変換画像中では背景は水平成分のみの動きを持つ。したがって変換画像中で垂直成分の動きを持つ領域として移動物体が抽出できる [1][2]。しかしこの方法ではカメラと同じ方向に移動する物体は抽出困難となる。

オプティカルフローを用いる方法ではフローの違いから背景と移動物体を分離する [3][4][5]。しかしこの方法では正確なオプティカルフローが得られない場合には移動物体抽出は困難となる。オプティカルフロー推定に関してはこれまでに様々な手法が提案されている [3][4][8][9] が、正確なオプティカルフロー推定は依然として困難な問題である。

統計的な検定を用いる方法では移動物体領域に属する各点が統計的な検定により背景の動きモデルに従わない動きを持つ点として検出される [6]。しかしこの方法は複雑な数理解析を用いるために計算コストが大きいという欠点がある。

背景位置合せを用いる方法では、まず支配的（dominant）な動きとして背景の動きを推定する。そしてその推定結果に基づいて画像間で背景位置合せ（カメラ移動による見かけの動きをキャンセル）し、それらの画像間での変化分（たとえば輝度の変化など）を利用して移動物体を抽出する [7]。この方法では支配的な背

景の動きが2次元動きモデルなどで近似できるならば移動物体を安定して抽出することができる。多くの動画像において背景の動きを2次元動きモデルで近似することができるので、この方法を用いることによって効率的かつ安定的な実用性の高い移動物体抽出の実現が期待できる。

背景位置合せに基づく方法では画面内に背景以外の移動物体が存在する画像から背景の動きを推定することが要求される。背景と移動物体が分離されていない状況では、移動物体に属するデータが外れ値 (outlier) となるため背景の動きを正しく推定するのが困難である。寺久保ら [7] は画像間の粗い対応点を用いて背景の動きを推定している。このときデータ中に含まれる outlier をあらかじめ排除することにより背景の動きを求めている。寺久保らは outlier の判定にヒューリスティックを用いているために非常に不安定であると考えられる。

outlier を含むデータに対して outlier の影響を受けずにパラメータを推定する有効な手法にロバスト統計 [10][11] がある (ただし、outlier の割合は50%未満でなければならない)。Sawhney ら [12] は対応点を用いずに画像間の輝度値の差を直接残差としてロバスト統計を用いて背景の動きを推定する手法を提案している。Sawhney らはロバスト統計の一手法である M-推定 (M-estimation) を用いているが、M-推定は得られる解が初期値に依存するため適切な初期値が不明な場合には安定した結果が得られない。

また、寺久保ら [7] は画像間での背景位置合せを行った後、背景位置合せされた画像間での差分に基づいて移動物体の抽出を行っている。差分に基づく方法では近似誤差などに起因する背景位置合せの誤差 (ずれ) によって誤抽出を生じやすい。

そこで本論文では画像間の背景位置合せに基づいて移動物体を抽出する新たな手法を提案する。提案手法では画像間の対応点からロバスト統計を用いて背景の見かけの動きを推定し、その推定結果を基に背景位置合せされた画像間での輝度の変化を利用して移動物体領域を抽出する。提案手法の特徴は安定な背景の動き推定を行うためにロバスト統計における M-推定と最小2乗中間値 (Least Median of Squares : LMS あるいは LMedS) 推定を組合せて用いていること、および信頼性の高い移動物体抽出を行うために背景位置合せの誤差を考慮して背景と移動

物体を分離していることである。

以下、2章でロバスト統計を用いた背景の動き推定について、3章で背景位置合せを用いた移動物体抽出についてそれぞれ説明し、4章で実画像を用いた移動物体の抽出実験により提案手法の有効性を示し、最後に5章でまとめと今後の課題を述べる。

2. ロバスト統計を用いた背景の動き推定

移動カメラ画像間の背景位置合せを行うためには背景の見かけの動きを求める必要がある。本章では背景の動きを画像内の支配的な動きであると見なし、画像間の対応点からロバスト統計を用いて背景の動きを推定する方法について述べる。

2.1 背景の動き推定に関する仮定

画面内に背景以外の移動物体が存在する画像からの背景の動き推定は、移動物体などに起因する outlier を含むデータからの推定となるので困難な問題である。さらに、カメラの移動によって生じる背景の見かけの動きは複雑なのでその動きを正確に推定するのは困難な問題である。そこで本研究ではこれらの問題を解決するために一般に成り立つ以下の2つの仮定1) および2) を利用して画像間の対応点からロバスト統計により背景の動きを推定する。また、本研究の目的 (問題設定) から仮定3) が成立する。

1) データ中の outlier の割合は50%未満である

この仮定は背景の動き推定にロバスト統計を適用するための必要条件であり、この場合の outlier は移動物体上の特徴点および背景上の誤対応点である。多くの画像において移動物体の領域は背景よりも小さく、連続する画像間での対応点探索では誤対応は少ないと考えられるので、この仮定は実際的である。

2) 背景の見かけの動きは2次元動きモデル (射影変換) で近似できる

射影変換 $f: (b_x', b_y') = f(\mathbf{a}, (b_x, b_y))$ は

$$\begin{pmatrix} b_x' \\ b_y' \end{pmatrix} = \frac{1}{a_7 b_x + a_8 b_y + 1} \begin{pmatrix} a_1 b_x + a_2 b_y + a_3 \\ a_4 b_x + a_5 b_y + a_6 \end{pmatrix} \quad (1)$$

で表される。ここで、 $\mathbf{a} = \{a_1 \sim a_8\}$ は動きパラメータであり、 (b_x, b_y) 、 (b_x', b_y') はそれぞれ変換前、後の座標である。射影変換は視点を固定して視軸を変える場合の座標変換式であるが、視点移動に関しても背景が平面で近似できる場合、つまりカメラの視点移動による背景上の各点での奥行き方向の変化がカメラと背景の距離に比べて小さい場合にこの仮定は成立する。動画像では連続する画像間で

の視点移動が小さいのでこの仮定が成立しやすく、動画像処理に関する多くの研究でこの仮定が用いられている [4] [5] [6] [7] [12]。

3) 画像間で生じる背景の見かけの動きは不明である

この仮定は本研究が環境やカメラの移動が不明な状況下での移動物体抽出を目的としているからであり、これは本手法の適用範囲を拡張するものである。

2.2 画像間の対応点探索

本研究では画像間の対応点を用いて背景の動きを推定する。ここでは、連続する2つの画像（基準画像と参照画像）間での対応点を残差逐次検定法（Sequential Similarity Detection Algorithm : SSDA） [14] を用いて求める。

対応点を信頼性良く求めるためにはコーナー点やスポット状の点などの特徴点を用いて対応づけを行う必要がある。また、背景の動きを信頼性良く推定するためには対応点が画像全体に一様に分布しているのが望ましい。そこで基準画像上に $M \times N$ 個のウィンドウを等間隔に配置（図 2.1）し、各ウィンドウ内からインタレストオペレータ [13] の出力値が最大である点 $\mathbf{p}_{i,j}$ ($i = 1 \sim M, j = 1 \sim N$) をそれぞれ1点ずつ特徴点として抽出する。インタレストオペレータでは画像上の小領域における方向性の分散が大きな点（コーナー点やスポット状の点など）が特徴点として抽出される。この特徴点を SSDA を用いて参照画像上の点 $\mathbf{q}_{i,j}$ と対応づける。

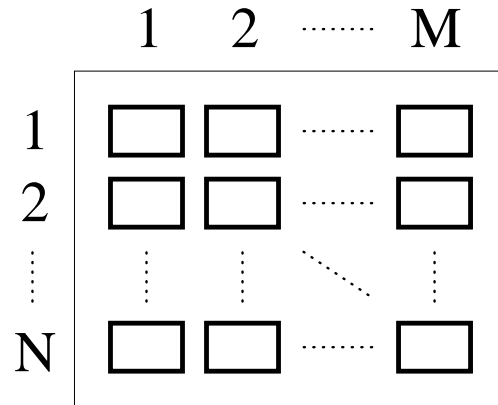


図 2.1 特徴点抽出のためのウィンドウの設定

2.3 ロバスト統計の背景の動き推定への適用

本研究では画像間の背景位置合せを行うために前節 2.2 で得られた対応点を用いて背景の動きパラメータ \mathbf{a} を推定する。このとき上述の対応点には outlier (おもに移動物体上の特徴点) が不可避免的に含まれてしまうため、最も一般的な最小 2 乗法では outlier の影響を受けて正しいパラメータが得られないという問題がある。本研究ではこの問題を解決するためにロバスト統計を用いる。

ロバスト統計には連続的な手法である M-推定と離散的な手法である LMS 推定の 2 つの代表的な手法がある。M-推定は高精度な推定が可能であるが、繰返し法を用いて解を求めるために本研究のように適切な初期値が不明な場合には適用することができない。一方の LMS 推定は乱数発生的な解法によって推定できるので初期値が不要である。しかし LMS 推定は非線形の最適化問題となるため最適解を見つけるのが難しく、一般的には準最適解しか求めることができない。そこで本研究では前述の粗い対応点を用いて、まず LMS 推定により背景の動きパラメータを求め、得られた準最適解を初期値として M-推定を行うことによりパラメータを精密化する。以下では、対応点 $\mathbf{p}_{i,j}$ と $\mathbf{q}_{i,j}$ の距離を残差 $r_{i,j}$

$$r_{i,j}(\mathbf{a}) = |\mathbf{p}_{i,j} - f(\mathbf{a}, \mathbf{q}_{i,j})|$$

$$= \sqrt{(p_x - q_x')^2 + (p_y - q_y')^2} \quad (2)$$

として推定を行う。ここで、 (p_x, p_y) 、 (q_x', q_y') はそれぞれ $\mathbf{p}_{i,j}$ 、 $f(\mathbf{a}, \mathbf{q}_{i,j})$ の座標である。

2.4 LMS 推定による背景の動きパラメータ推定—GA による解法

本研究では画像間の背景位置合せに用いる背景の動きパラメータを求めるために、まず初期値の設定が不要な LMS 推定を用いて画像間の対応点から背景の動きパラメータを推定する。この LMS 推定は次式で定義される。

$$L(\mathbf{a}) = \text{median} \{r_{i,j}^2(\mathbf{a})\} \rightarrow \min \quad (3)$$

LMS 推定ではこのように残差の 2 乗の中間値を最小とすることにより outlier の影響を受けずに動きパラメータ \mathbf{a} を推定することができる。一般的にはランダムサンプリングを用いた推定が行われる [5]。この推定は確率的な操作であり、正しい推定が行われる確率がデータ中の outlier の割合、標本の要素の数および標本のサンプリング回数に依存してしまう。そのため、処理の安全性を確保するためには相当大きなサンプリング回数が必要となり、効率が悪化するという問題がある [5]。

本研究では遺伝的アルゴリズム (Genetic Algorithm : GA) [15] を用いて LMS 推定を行う。GA は従来手法では解決が困難であった様々な最適化問題に対して準最適解を効率的に求めることができる有効な手法であり、また本研究では連続する画像間での動きパラメータ推定を対象としているために解の探索範囲をある程度制限することができるので GA により安定的かつ効率的な推定が期待できる。さらに、GA の特徴として並列処理に適していることがある。現行システムでは並列処理を用いていないが、将来的には並列処理によりさらなる高速化が期待できる [15]。以下では、GA を用いた LMS 推定による背景の動きパラメータの推定方法について説明する。

(1) 染色体の定義

動きパラメータ $\mathbf{a} = \{a_1 \sim a_8\}$ を個体 k の染色体 C_k とする。

$$C_k = (a_1, a_2, a_3, a_4, a_5, a_6, a_7, a_8) \quad (4)$$

(2) 適応度の定義

各個体の適応度 F_k を

$$F_k = \frac{1}{\text{median} \{r_{i,j}^2(\mathbf{a}_k)\}} \quad (5)$$

によって計算する。ここで \mathbf{a}_k は染色体 C_k の表す動きパラメータである。

(3) 遺伝規則の設定

つぎのようにして次世代の個体群を生成する。まず現世代の個体群を適応度の大きい順に並べる。その後、下位の一定の割合の個体を淘汰して消滅させる。そして淘汰されなかった上位の個体の中からランダムに選んだペアを交差させて（ここでは2点交差を使用）新しい個体を生成することにより個体の総数を一定に保つ。最後に各個体に対して突然変異を実行する。

(4) 推定処理の実行

本研究ではランダムに生成させた初期個体群を一定の回数の世代交代でもって進化させ、進化終了時において最大適応度をもつ個体の染色体である動きパラメータ \mathbf{a}' を解とする。

2.5 M-推定による動きパラメータの精密化

本研究では信頼性の高い背景位置合せを行うために、前節 2.4 の LMS 推定によって得られた動きパラメータ \mathbf{a}' を初期値として M-推定を行うことにより動きパラメータを精密化する。M-推定では以下のような繰返し重みつけ最小 2 乗 (Iteratively Reweighted Least Squares : IRLS) 法 [10] によってパラメータ推定を行う。

1) 推定値 \mathbf{a}^m (m は繰返し回数を表す) の初期値として \mathbf{a}' を与える。

$$\mathbf{a}^0 \leftarrow \mathbf{a}' \quad (6)$$

2) 各対応点ごとに残差 $r_{i,j}(\mathbf{a}^m)$ に基づいて重み $w_{i,j}(\mathbf{a}^m)$ を計算する。ここでは重み関数として収束性の良い Tukey の Biweight 関数 [11] を用いている。

$$w_{i,j} = \begin{cases} (1 - (r_{i,j}/\sigma)^2)^2 & \text{if } r_{i,j} \leq \sigma \\ 0 & \text{if } r_{i,j} > \sigma \end{cases} \quad (7)$$

ここで、

$$\sigma = 1.4826 \times \text{median} \{ r_{i,j} \} \quad (8)$$

である。各データに重み $w_{i,j}$ を与えることにより真値から外れたデータ (outlier) の影響を排除することができる。

3) 次の最小2乗法を実行し、推定値を更新する。

$$L(\mathbf{a}^{m+1}) = \sum_{i=1}^M \sum_{j=1}^N w_{i,j}(\mathbf{a}^m) r_{i,j}^2(\mathbf{a}^{m+1}) \rightarrow \min \quad (9)$$

4) 重み $w_{i,j}$ が変化しなくなるまで2) および3) の処理を繰り返し、収束値 $\hat{\mathbf{a}}$ を得る。

本研究では移動物体を抽出するために、以上のようにして得られた背景の動きパラメータ $\hat{\mathbf{a}}$ を用いて画像間の背景位置合せを行う。

3. 背景位置合せを用いた移動物体抽出

本研究では前章で得られた背景の動きパラメータを用いて画像間の背景位置合せを行うことにより背景の見かけの動きをキャンセルし、背景位置合せされた画像を用いて移動物体を抽出する。本章では背景位置合せされた画像から移動物体を抽出する方法について述べる。

3.1 移動物体抽出のアルゴリズム

信頼性の高い移動物体抽出を行うためには背景の動きモデルの近似誤差などに起因する背景位置合せの誤差（ずれ）を吸収しなければならない。そこで本研究では画像を $S_x \times S_y$ 画素の小ブロックに分割し、以下のようなブロック単位での解析により移動物体領域を抽出する。

まず2章で得られた背景の動きパラメータ $\hat{\mathbf{a}}$ を用いて基準画像 I_α と参照画像 I_β との間で背景位置合せを行い、その後これらの画像間でブロックを単位として非類似度を計算する。もし背景位置合せに誤差がないならば、画像 I_α 中のブロック $B(m, n)$ が背景領域の場合、 $B(m, n)$ と対応する位置の画像 I_β 中の領域 $R(m, n)$ との非類似度 D （高速化のため次式で評価）

$$D = \sum_{m=1}^{S_x} \sum_{n=1}^{S_y} |B(m, n) - R(m, n)| \quad (10)$$

は小さく、逆に移動物体領域の場合大きくなるのでしきい値処理により移動物体領域を検出できる。しかし本研究の場合、背景位置合せに近似による誤差が本質的に存在する。そこで領域 $R(m, n)$ の周辺で非類似度を求めることを考え、周辺領域（許容領域） A における非類似度 $D(u, v)$

$$D(u, v) = \sum_{m=1}^{S_x} \sum_{n=1}^{S_y} |B(m, n) - R(m + u, n + v)| \quad (11)$$

のうち1つでもしきい値 T 以下のものがあれば $B(m, n)$ は背景領域であるとし、そうでなければ移動物体領域とする。ただし、 $(u, v) \in A$ である（図 3.1）。

$$B(m, n) : \begin{cases} \text{背景領域} & \text{if } \exists u \exists v D(u, v) \leq T \\ \text{移動物体領域} & \text{otherwise} \end{cases}$$

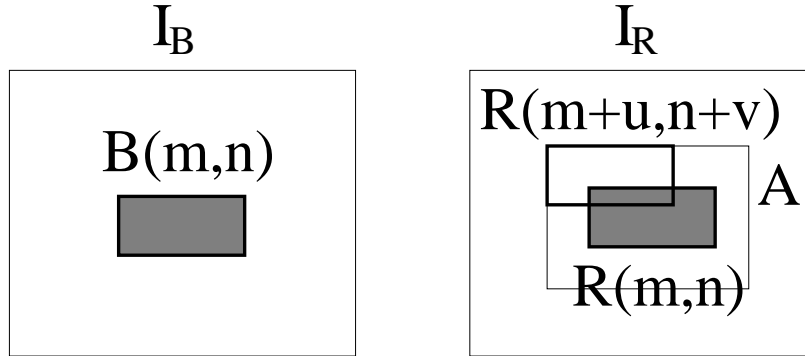


図 3.1 背景位置合せの誤差を考慮した非類似度

3.2 許容領域としきい値の設定

前節のアルゴリズムにおいて問題となるのが許容領域 A としきい値 T の大きさであり、移動物体領域を正しく抽出するためにはこれらを適切に設定する必要がある。本研究ではロバスト統計を用いて許容領域を設定し、窪田らの誤差モデル [4] を用いてしきい値を設定する。

許容領域が小さいと背景を移動物体として誤検出する恐れがある。逆に大き過ぎると移動物体を背景として検出洩れする割合が大きくなる。そこで位置合せの誤差分布がガウス性であると仮定してロバスト統計から推定標準偏差 $\hat{\sigma}$ を

$$\hat{\sigma} = 1.4826 \times \text{median} \{ r_{i,j}(\hat{\mathbf{a}}) \} \quad (12)$$

とし [11]、位置合せ誤差の最大値 (許容値) r_{max} を推定標準偏差 $\hat{\sigma}$ の 2.5 倍

$$\begin{aligned} r_{max} &= 2.5 \times \hat{\sigma} \\ &= 2.5 \times 1.4826 \times \text{median} \{ r_{i,j}(\hat{\mathbf{a}}) \} \end{aligned} \quad (13)$$

に設定し [11]、半径が r_{max} の円内を許容領域 A とする。

また、しきい値として固定値を用いた場合、適切なしきい値の設定が困難なために汎用性が損なわれる。そこでしきい値 T を画像のテクスチャに基づいてブロックごとに設定する。

窪田らの誤差モデルを用いると、ブロックを $(\Delta x, \Delta y)$ ずらしたときの元のブロックとの非類似度はテーラー展開し、高次の微小項を無視することで

$$\begin{aligned} & \sum_{x=1}^{S_x} \sum_{y=1}^{S_y} |E(x, y) - E(x + \Delta x, y + \Delta y)| \\ &= \sum_{x=1}^{S_x} \sum_{y=1}^{S_y} |E_x(x, y)\Delta x + E_y(x, y)\Delta y| \end{aligned} \quad (14)$$

と近似できる。ここで、 $E(x, y)$ 、 $E(x + \Delta x, y + \Delta y)$ はそれぞれブロック内の点 (x, y) 、 $(x + \Delta x, y + \Delta y)$ における輝度値であり、 E_x 、 E_y はそれぞれ x 、 y 方向の微分値である。この式から背景領域のブロック $B(m, n)$ に対して画像の量子化によって生じ得る非類似度 $D(u, v)$ の最大値（許容値）が求められるので、これをしきい値 T とする。

$$T = \max_{|\Delta x|=0.5, |\Delta y|=0.5} \left\{ \sum_{x=1}^{S_x} \sum_{y=1}^{S_y} |E_x(x, y)\Delta x + E_y(x, y)\Delta y| \right\} \quad (15)$$

なお、このしきい値 T は画像のノイズを全く考慮していないので、特に画像間での輝度の変化に対して弱い。本研究では非類似度 $D(u, v)$ を求めるときブロックごとに画像を正規化してから求めるようにしている。画像の正規化は各点の輝度値からブロック内の平均輝度値を引くことにより行う。

以上のように許容領域 A 内で得られる非類似度 $D(u, v)$ としきい値 T との大小比較を行い、背景位置合せされた画像から移動物体を抽出する。

4. 移動物体の抽出実験

提案手法の有効性を示すためにカメラを移動させながら撮影した実画像を用いて移動物体の抽出実験を行った。計算機はワークステーション (DEC alpha) を使用し、カメラは市販のビデオカメラ (SONY, Handycam PRO) を使用した。実験では 6×6 個の特徴点 ($M = N = 6$) を用いて背景の動き推定を行った。SSDA による対応点探索には 15×15 画素のテンプレートを使用し、対応洩れが起こらないように探索範囲は大きめにとった (20×20 画素)。移動物体の抽出処理は画像を 9×6 画素のブロック ($S_x = 9, S_y = 6$) に分割し、連続する 2 フレーム間で行った。なお、画像の周辺部の一定領域 (端から 20 画素程度) はあらかじめ背景領域として処理した。

また、背景の動き推定における GA のパラメータはそれぞれ個体数 30、世代交代数 20、淘汰率 40%、突然変異率 5%、染色体長 18 ビットである。染色体長の内訳はパラメータ $a_1 \sim a_6$ それぞれ 3 ビットであり、各パラメータの範囲はそれぞれ

$$\begin{cases} a_1, a_5 = 1 + 0.1 \times d_1 (d_1 \in [-4, 3], \text{整数}) \\ a_2, a_4 = 0.1 \times d_2 (d_2 \in [-4, 3], \text{整数}) \\ a_3, a_6 = 2 \times d_3 (d_3 \in [-4, 3], \text{整数}) \end{cases} \quad (16)$$

である。なお、GA による LMS 推定では探索範囲を減らすために a_7, a_8 を 0 に固定 (背景の動きをアフィン変換で近似) し、あらかじめ染色体から取り除いて推定を行った。ただし、後続の M-推定では全てのパラメータ $a_1 \sim a_8$ について推定を行った。これらの GA パラメータは全て経験的に定めた。

以下では、4.1 で背景の動きが射影変換で良好に近似できる場合の本手法の有効性を確認するためにカメラの視点移動が比較的小さな様々な移動カメラ画像に対する移動物体の抽出実験、4.2 で背景の動きモデルに比較的大きな近似誤差がある場合でも本手法によってある程度の誤差を吸収して移動物体を抽出できることを示すためにカメラの視点移動が比較的大きな移動カメラ画像に対する移動物体の抽出実験をそれぞれ行った。

4.1 背景の動きが射影変換で良好に近似できる場合

背景の動きが射影変換で良好に近似できる場合、本手法によって移動物体を良好に抽出できることを確認するために様々な環境やカメラの移動に対して実験を行った。なお、以下の実験結果では1列目、2列目にそれぞれ原画像（ 360×243 画素、白黒256階調）および本手法によって抽出された移動物体領域を示す。

4.1.1 実験1—理想的な状況（ズームのみ）

視点移動を伴わないカメラ操作（ズーム、パン、チルト、ロール）の場合、背景の動きは射影変換で誤差無しに表現できる。この実験ではそのような理想的な例としてカメラをズームインさせながら撮影した画像列（15フレーム/秒）に対して行った実験結果の一部を図4.1に示す。ただし、カメラを手を持って撮影したため厳密には手ぶれによる微小な視点移動が存在する。画像間での輝度の変化が小さいために起こったと思われる抽出洩れおよび画像のノイズが原因と思われる誤抽出が一部で発生しているが、本手法により移動物体（歩行者およびリヤカー）がほぼ正しく抽出されている。

4.1.2 実験2—移動物体の注視（首振りのみ）

カメラに左方への首振りを与えて移動物体を注視しながら撮影した画像列（10フレーム/秒）に対して行った実験結果の一部を図4.2に示す。この実験では移動物体を注視しながら撮影したため原画像中での移動物体の見かけの動きが背景の動きよりも小さい。このような場合には背景の動きの初期値として適切な値を与えるのが非常に困難なため、M-推定のみでは正確な背景位置合せが難しいと思われる。本手法では背景位置合せが正しく行われて移動物体（歩行者）がほぼ正しく抽出されている。ただし、画像間での輝度の変化が小さいために起こったと思われる抽出洩れおよび画像のノイズが原因と思われる誤抽出が一部で発生している。

4.1.3 実験3—カメラの複合移動（ロール+首振り）

カメラに時計回りのロールと右方への首振りを同時に与えながら撮影した画像列（30フレーム/秒）に対して行った実験結果の一部を図4.3に示す。この実験ではカメラに複雑な動きを与えたが、本手法により移動物体（自動車）がほぼ正しく抽出されている。ただし、画像間での輝度の変化が小さいために起こったと思われる抽出洩れおよび画像のノイズが原因と思われる誤抽出が発生している。特に自動車の中央部分において大きな抽出洩れが生じているが、これは自動車がほぼ水平に移動したために中央部分における輝度の変化が小さかったためと思われる。

4.1.4 実験4—複数移動物体（ズーム+首振り）

カメラにズームアウトと右方への首振りを同時に与えながら複数の移動物体を撮影した画像列（30フレーム/秒）に対して行った実験結果の一部を図4.4に示す。この実験では画面内に互いに交差する複数の移動物体が存在するが、本手法によりそれぞれの移動物体（2台の自動車）がほぼ正しく抽出されている。ただし、画像間での輝度の変化が小さいために起こったと思われる抽出洩れおよび画像のノイズが原因と思われる誤抽出が一部で発生している。

なお、これらの実験において1回の移動物体抽出に要した計算時間は対応点探索（画像の取込みも含む）に約25秒、背景の動き推定に約0.4秒、移動物体抽出に約3秒、計約28秒である。

4.2 実験5—背景の動きモデルに比較的大きな近似誤差がある場合 従来のフレーム間差分による移動物体抽出との比較）

カメラが視点を比較的大きく平行移動する場合、背景の実際の動きと仮定の動き（射影変換）との間に比較的大きな誤差が発生する。この実験ではそのような例として、カメラを移動ステージの上に乗せて約10cmずつ前方へ平行移動させながら撮影した画像列に対して行った。なお、初期位置から背景（壁）までの距離は約5mであり、背景の奥行きは約1.5mである。

実験結果の一部を図 4.5 に示す。左上が原画像 (640 × 486 画素、白黒 256 階調) である。背景に見かけの動きが生じている様子を右上に示す。これは次フレームとの原画像同士の差分画像である。次フレームとの間で背景位置合せが行われた様子を左下に示す。これは背景位置合せを行った後の差分画像を 2 値化したものである。白い部分が 2 値化によって抽出された部分であり、2 値化のしきい値は移動物体の領域が良好に抽出されるように実験的に 30 と定めた。この実験では背景位置合せに比較的大きな近似誤差が存在する状況で行ったため、背景上にも大きな抽出領域が生じている。本手法によって抽出された移動物体領域を右下に示す。本手法により背景位置合せの誤差が吸収されて移動物体 (人物) のみがほぼ正しく抽出されている。ただし、画像間での輝度の変化が小さいために起こったと思われる抽出洩れおよび画像のノイズが原因と思われる誤抽出が一部で発生している。

なお、この実験において 1 回の移動物体抽出に要した計算時間は対応点探索 (画像の取込みも含む) に約 37 秒、背景の動き推定に約 0.4 秒、移動物体抽出に約 12 秒、計約 49 秒である。

4.3 考察

2.1 節の仮定 1) および 2) が成立する場合、4.1 節の実験結果から本手法によって移動カメラ画像から移動物体をほぼ良好に抽出できることが確認できた。特に 4.1.2 節の実験 2 では初期値の設定が難しい状況で行ったため従来の M-推定に基づく背景の動き推定手法では安定な推定が困難であると考えられるが、本手法では安定な推定が可能であった。さらに背景の動きモデルに比較的大きな誤差がある場合、従来のフレーム間差分に基づく移動物体の抽出手法では大きな誤抽出が発生すると考えられるが、4.2 節の実験結果から本手法によってある程度の誤差なら吸収して移動物体をほぼ良好に抽出できることが確認できた。ただし、4.1, 4.2 の一部において抽出洩れ (実験 4 における自動車の中央部分などに) や誤抽出 (実験 3 のフレーム 9 や実験 5 のフレーム 1 などに) が生じた。

環境やカメラ移動の変化にもかかわらずほぼ良好な抽出結果が得られた理由として

- 1) M-推定と LMS 推定を組合せて背景の動きを推定したこと
- 2) 背景位置合せの誤差を考慮して背景と移動物体を分離したこと
- 3) 画像のテクスチャに基づいて移動物体抽出時のしきい値を設定したこと

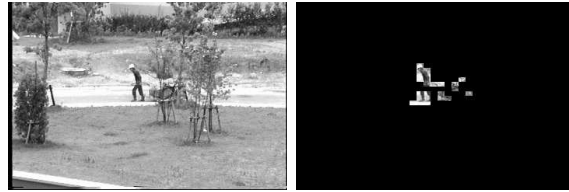
などが考えられる。なお、2.1節の仮定 1),2) が成立しない場合、つまり本手法が原理的に旨くいかない場合として、1) 画面内に多数の移動物体が存在して背景上の特徴点よりも移動物体上の特徴点の方が数多く抽出される場合や 2) 屋内環境のような背景の凹凸が激しい環境内においてカメラが視点を大きく変更する場合などが考えられる。

抽出洩れは輝度値の差に基づく非類似度によって移動物体の抽出を行っているために起こる輝度値の変化が小さい部分での抽出洩れによるところが大きい。対策としてはエッジなどの他の画像特徴から得られた結果との統合が考えられる。

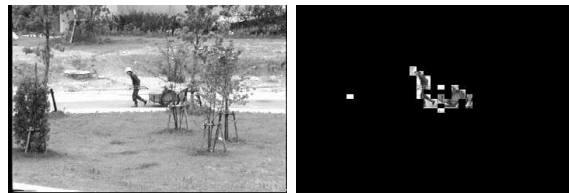
誤抽出のおもな原因としては画像のノイズが考えられる。本手法では画像のノイズを全く考慮せずに移動物体抽出時のしきい値を定めているからである。しかし誤って抽出された領域のほとんどが孤立した微小領域なのでノイズ成分として除去可能である。

また、本研究では背景の動き推定に用いる対応点を SSDA によって探索しているため各対応点には画像の量子化により最大 0.5 (画素) の誤差があり、それが移動物体抽出の信頼性 (背景の動きおよび位置合せ誤差の推定精度) に大きく影響していると考えられる。そこで、移動物体抽出の信頼性を上げるためにエッジの交点として特徴点の位置を定める方法によりサブピクセル単位の精度で対応点を求めることが考えられる。

さらに、実験に使用した GA のパラメータおよびフレーム間隔 (タイムスケール) は移動物体領域が良好に抽出されるように経験的に定めたものであり、適切な GA パラメータおよびタイムスケールの設定手法の確立については今後検討する必要がある。



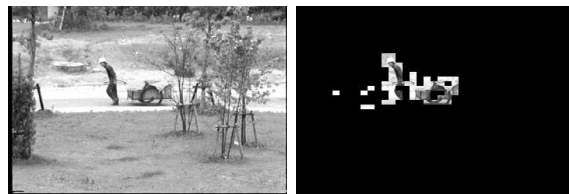
フレーム 1



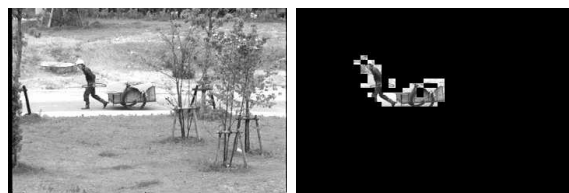
フレーム 6



フレーム 11



フレーム 16

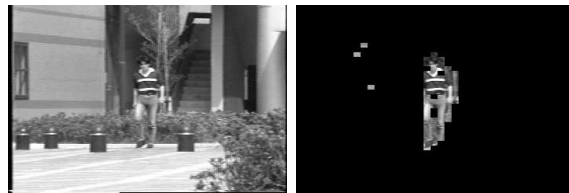


フレーム 21

図 4.1 移動物体の抽出結果 (実験 1: ズームイン)



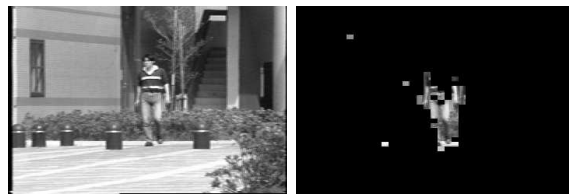
フレーム 1



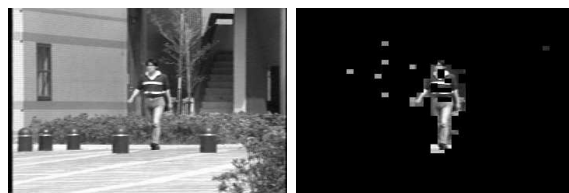
フレーム 3



フレーム 5

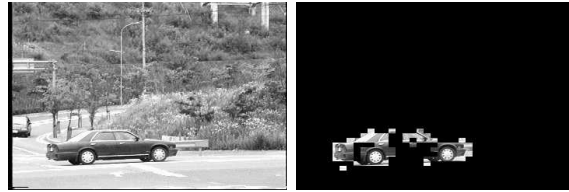


フレーム 7

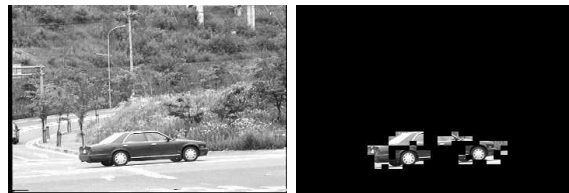


フレーム 9

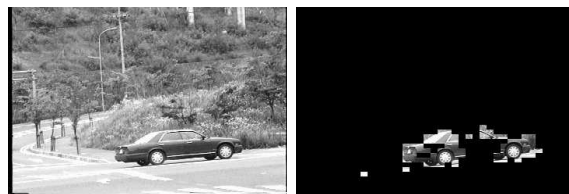
図 4.2 移動物体の抽出結果 (実験 2: 左方への首振り)



フレーム 1



フレーム 6



フレーム 11

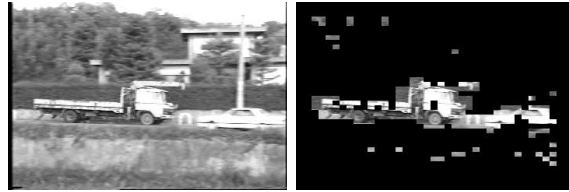


フレーム 16

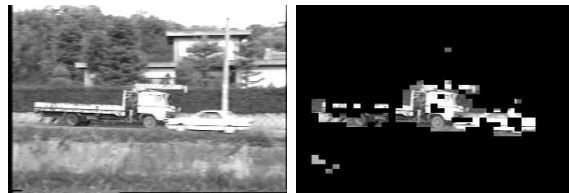


フレーム 21

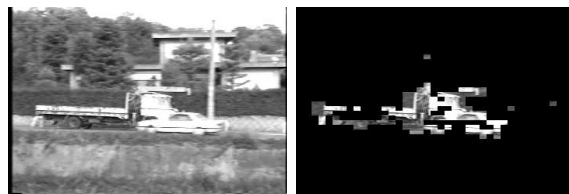
図 4.3 移動物体の抽出結果 (実験 3: 時計回りのロール+右方への首振り)



フレーム 1



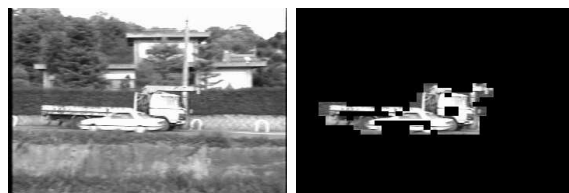
フレーム 3



フレーム 5



フレーム 7



フレーム 9

図 4.4 移動物体の抽出結果 (実験 4: ズームアウト+右方への首振り)

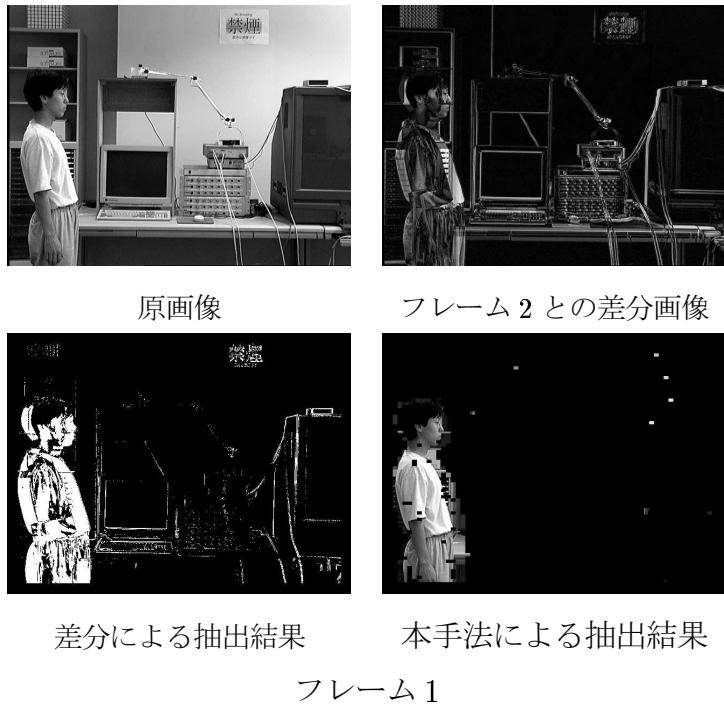


図 4.5 移動物体の抽出結果 (実験 5: カメラ視点の前方への平行移動)

5. むすび

本論文ではロバスト統計を用いた背景位置合せに基づいて移動カメラ画像から移動物体を抽出する手法を提案し、多種類の実画像に対して移動物体の抽出実験を行った。提案手法では M-推定と GA による LMS 推定を組合せて連続フレーム間の対応点から背景の動きを推定することにより背景位置合せを行い、対応点の距離（残差）からロバスト統計により評価された背景位置合せの誤差を考慮して背景位置合せされた画像上で背景と移動物体の分離を行った。実験により様々な環境やカメラ移動に対して安定に画像間の背景位置合せができることを確認した。さらに、本手法では背景位置合せの誤差をある程度吸収して安定に移動物体が抽出できることを実験により確認した。

本手法の問題点としては画像間での輝度の変化が小さな部分において移動物体の抽出洩れが生じることである。今後は様々な状況下での適用を可能とするための背景の動きモデルの複雑化、実時間での動作を可能とするための各種処理（対応点探索、背景の動き推定、移動物体抽出）の高速化などが課題である。

謝辞

日頃から有益な御指導・御助言を頂いたソフトウェア基礎講座の横矢直和教授、竹村治雄助教授、山澤一誠助手、岩佐英彦助手に感謝します。

また、本論文をまとめるにあたり有益な御助言を頂いた像情報処理学講座・千原國宏教授に感謝します。

最後に、実験用の画像データを提供して頂いた住友電気工業株式会社の寺久保敏氏並びに本研究を行うにあたり御協力を頂いたソフトウェア基礎講座諸氏に感謝します。

参考文献

- [1] J.Frazier and R.Nevatia, “Detecting Moving Objects from a Moving Platform”, *Proc. IEEE Robotics and Automation*, pp.1627-1633 (1992).
- [2] D.Nair and J.K.Aggarwal, “Detecting Unexpected Moving Obstacles that Appear in the Path of a Navigating Robot”, *Proc. ICIP*, Vol.II of III, pp.311-315 (1994).
- [3] 太田直哉, “信頼性情報をもったオプティカルフローからの形状復元とその移動物体検出への応用”, *信学論*, Vol.J76-D-II, No.8, pp.1562-1571 (1993).
- [4] 窪田進, 下辻成佳, “ハフ変換を用いたカメラパラメータの推定及び動画像からの移動物体の分離”, *画像の認識・理解シンポジウム (MIRU'96) 講演論文集*, pp.II-121-126 (1996).
- [5] 梅山伸二, “ロバスト回帰に基づく多重オプティカルフローの推定について”, *信学技報*, PRMU96-108 (1996).
- [6] P.H.S.Torr and D.W.Murray, “Statistical Detection of Independent Movement from a Moving Camera”, *Image and Vision Computing*, Vol.11, No.4, pp.180-187 (1993).
- [7] 寺久保敏, 横矢直和, 岩佐英彦, 竹村治雄, “2段階アフィン変換を利用した移動カメラ画像からの動物体の抽出”, *信学技報*, PRU95-181 (1995).
- [8] S.Hsu, P.Anandan and S.Peleg, “Accurate Computation of Optical Flow by Using Layered Motion Representations”, *Proc. ICPR*, Vol.A, pp.743-746 (1994).
- [9] M.Bober and J.Kittler, “Robust Motion Analysis”, *Proc. CVPR*, pp.947-952 (1994).

- [10] 佐藤嘉伸, “画像処理における統計モデル利用に関する最近の動向: ロバスト推定法と MDL 基準”, *Medical Imaging Technology*, Vol.12, No.1, pp.48-58 (1994).
- [11] P.J.Rousseeuw and A.M.Leroy, *Robust Regression & Outlier Detection*, John Wiley & Sons, New York (1987).
- [12] H.S.Sawhney and S.Ayer, “Compact Representations of Videos Through Dominant and Multiple Motion Estimation”, *IEEE Trans. PAMI*, Vol.18, No.8, pp.814-830 (1996).
- [13] H.P.Moravec, “Visual Mapping by a Robot Rover”, *Proc. 6th IJCAI*, pp.598-600 (1979).
- [14] D.J.Barnea and H.F.Silverman, “A Class of Algorithms for Fast Digital Image Registration”, *IEEE Trans. Computers*, Vol.C-21, pp.179-186 (1972).
- [15] 安居院猛, 長尾智晴, “ジェネティックアルゴリズム”, 昭晃堂 (1993).
- [16] 朝岡忠, 横矢直和, 竹村治雄, 山澤一誠, “ロバスト統計を用いた背景位置合せに基づく移動カメラ画像からの移動物体抽出”, 信学技報, PRU96-145 (1997).
- [17] 朝岡忠, 横矢直和, 竹村治雄, 山澤一誠, “ロバスト統計を用いた背景位置合せに基づく移動カメラ画像からの移動物体の抽出”, 情処学第 53 回全大講演論文集, 第 2 分冊, 5N-5 (1996).