

ロバスト統計を用いたハイパーメディア作成のための 高速移動物体追跡法

金子 敏充* 堀 修†

(株) 東芝 研究開発センター 情報・通信システム研究所†

あらまし 本稿では、変形する追跡対象を高速に追跡する手法の提案を行う。この追跡手法は、ハイパーメディア・コンテンツ作成時に移動物体にたいしてリンクを付与する際に用いることを目指したものである。提案手法はブロックマッチングをベースとした追跡であるが、予め精度の高い移動ベクトルの推定が期待できる部分領域のみを追跡するため、高速な処理が可能であるという特徴を持つ。また、追跡物体の変型モデルとしてアフィン変換に基づく変形を仮定し、部分領域のブロックマッチングの結果から変形を記述するパラメータをロバスト推定により推定する。これにより、追跡物体の変形に追従が可能になり、かつノイズにロバストな追跡が行えることを実験により確認した。

キーワード: トラッキング, ブロックマッチング, ロバスト統計, アフィン変換

A Moving Object Tracking Method for Supporting Hypermedia Contents Authoring Using Robust Statistics

Toshimitsu KANEKO and Osamu HORI

R&D Center, Toshiba Corporation

Abstract In this report, a moving object tracking method using robust statistics is proposed. It is designed for supporting hypermedia contents authoring. For fast processing, the method selects a small number of subregions for block matching. This is done by estimating the reliability of each subregion to find displacement vectors and selecting subregions with high reliability. We assume that the deformation of tracking objects can be approximated by Affine transformation. By estimating Affine parameters using robust statistics, the tracking method becomes robust against noise and occlusions.

keywords: tracking, block matching, robust statistics, Affine transformation

*E-Mail: toshimitsu.kaneko@toshiba.co.jp

†E-Mail: osamu.hori@toshiba.co.jp

†〒 210-8582 川崎市幸区小向東芝町 1

1 はじめに

ハイパーメディアは様々なメディアをハイパーリンクにより関連づけ、関連情報へのアクセスを容易にしたものである。例えば、映像中の人・車・建物などの登場物体を指定すると、それらを説明するテキストが表示されたり、音声・映像などが再生されるというものである。

このようなコンテンツの制作過程では、制作者は映像中の登場物体と関連情報とのハイパーリンクを付与する。すなわち、ある関連情報へのハイパーリンクが付与されている映像中の画素の集合をアンカー領域と呼ぶことにすると、関連情報ごとに映像中にアンカー領域を設定する。動画像の場合には、登場物体の動きを追いかけながら多くのフレームに対してアンカー領域を指定しなければならず、大変負荷の大きい作業であった。

これまでに、ハイパーメディアのアンカー領域指定支援を目的とした研究として [1] が報告されている。[1] では、テンプレートマッチングにより追跡を行っているが、移動先が一意に特定できない場合に移動先の候補を複数保持しておき、後のフレームにおいて最も信頼性の高い追跡が行えた候補を移動先として決定している。これにより単純なテンプレートマッチングの追跡性能を改善できるが、追跡物体の変形がある場合には追跡が困難である。また、処理速度も追跡領域全体についてテンプレートマッチングを行うのと同様である。[2] は監視目的の物体追跡法で、追跡領域を2つの矩形領域に分割し、それぞれブロックマッチングを行う手法である。しかし追跡不可能な変形パターンが多く、処理速度も [1] と同等である。

ハイパーメディアのアンカー領域指定支援ツールを安価に制作者に提供できることを追跡処理の条件とすると、追跡性能が高いことに加え、追跡速度が速いことがまず要求される。提案する移動物体追跡手法は、追跡領域全体についてブロックマッチングを行うよりもはるかに処理が高速で、かつ物体の変形に追従可能な手法である。提案手法はブロックマッチングをベースとするが、予め高精度なブロックマッチングが期待できる部分領域のみを追跡するため、高速な処理が可能である。また、追跡物体の変型モデルとしてアフィン変換モデルを仮定し、部分領域のブロックマッチングの結果から変形記述パラメータをロバスト推定により推定するため、追跡物体の変形に追従が可能になり、かつノイズにロバストな追跡が行える。

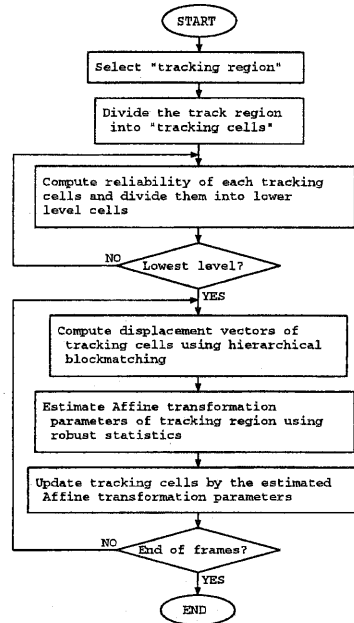


図 1: 物体追跡処理のフロー

2 準備と処理の概要

提案手法では、階層化された画像を用いて追跡を行う。階層レベルを $l = 0, 1, \dots, L$ とし、階層レベル数を $L + 1$ とする。階層レベル l の画像は縦横 $1/2^{L-l}$ にサブサンプルされた画像である。追跡は追跡物体中の部分領域であるセルを単位として行う。セルは各階層レベルに対して用意されており、セルのサイズは $W_w \times W_w$ であるが、これは原画像の $2^{L-l}W_w \times 2^{L-l}W_w$ の領域に対応している。また、レベル l のセルは下位のレベル $l+1$ の4つのセルに分割することができる。

図 1 は追跡処理全体の流れを示したものである。処理は初期処理と追跡処理とからなる。まず初期処理について説明する。

初期処理では、まずユーザーの手入力により追跡領域 (tracking region) を指定する。追跡領域はどのような形でも良いが、追跡物体 (tracking object) 内部の領域として指定する。ここでは簡単に指定できるように、矩形で追跡領域を指定するものとする。

次に、指定された追跡領域をレベル 0 のセルに分割する (図 2(a))。このときの初期セルの領域は原画像の $2^L W_w \times 2^L W_w$ の矩形領域とする。指定

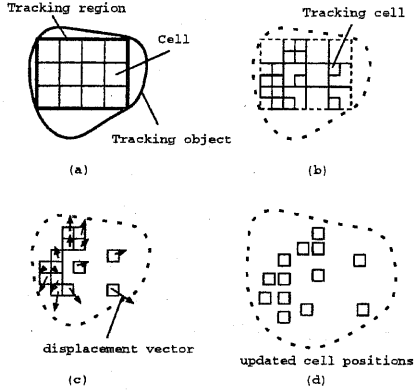


図 2: 物体追跡処理の例

した追跡領域のサイズが $2^L W_w$ の倍数であるとは限らないが、ここでは簡単のため $2^L W_w$ の倍数であるものとする。分割された各セルについては、下位レベルのセルへの分割と、実際に追跡を行う追跡セル (tracking cell) の選択を行う。これは 3 で説明するように、ブロックマッチングの信頼性が高いという基準で行う。以上の処理により、最終的に N 個のレベル L の追跡セルと、これらを含む上位レベルの追跡セルを選択する (図 2(b))。

次に追跡処理について説明する。まず、追跡セルについてのみ、次フレームとの間で 4.1 で説明する階層的ブロックマッチングを実行する (図 2(c))。相関の基準は SSD (Sum of Squared Distance) を用いる。このブロックマッチングにより各追跡セルの移動ベクトルおよびその信頼性を得る。移動物体の変型・移動がアフィン変換に従うとの仮定のもとで、移動ベクトル及び信頼性を用いてアフィン変換のパラメータを推定する。この推定には 4.3 で説明するロバスト統計の M 推定を用いる。

最後に、アフィンパラメータ推定の信頼度がしきい値よりも高ければ、パラメータに基づいて追跡セルの位置を更新する (図 2(d))。これは次フレームでの追跡セルの初期位置となる。また、追跡物体の領域は全追跡セルを囲む領域となる。

以上の追跡処理を繰り返し、追跡セルの位置を更新していくことにより物体の追跡を行う。

3 初期処理

初期処理では追跡処理の高速化を図るため、セルの中から実際に追跡を行う追跡セルを選択する。

このとき、信頼性の高い追跡が行えるようにするため、ブロックマッチングにおける移動ベクトル検出精度の高いセルを選択する。

3.1 SSD による移動ベクトルの最尤推定

時刻 t のフレームの座標 (x, y) における画素値を $f_t(x, y)$ とし、時刻 $t+1$ における (x, y) の対応点 (x', y') の画素値との差分が平均 0、分散 σ_p^2 の正規分布に従っていると仮定する。すなわち、 $\Delta = f_{t+1}(x', y') - f_t(x, y)$ の生起確率 $p(\Delta)$ が

$$p(\Delta) = \frac{1}{\sqrt{2\pi}\sigma_p} \exp\left\{-\frac{\Delta^2}{2\sigma_p^2}\right\}$$

に従うと仮定する。また、セル内の追跡物体の変形は無視できるものとする。

時刻 t のフレームにおける番号 i のセルが時刻 $t+1$ において $\mathbf{r} = (r_x, r_y)^T$ だけ移動したとするときの尤度 $L_i(\mathbf{r})$ の対数は

$$\begin{aligned} -\log L_i(\mathbf{r}) &= \sum_j -\log p(\Delta_{ij}(\mathbf{r})) \\ &= \sum_j \left\{ \frac{\Delta_{ij}^2(\mathbf{r})}{2\sigma_p^2} + \log \sqrt{2\pi}\sigma_p \right\} 1 \end{aligned}$$

となる。ここで添字 j はセル内の画素番号であり、セル i の画素 j の座標を (x_{ij}, y_{ij}) として $\Delta_{ij}(\mathbf{r}) = f_{t+1}(x_{ij} + r_x, y_{ij} + r_y) - f_t(x_{ij}, y_{ij})$ である。ブロックマッチングにおける画像間の距離尺度として 2 乗距離の総和 SSD を用いれば、セル i の移動ベクトルが \mathbf{r} のときの SSD は $SSD_i(\mathbf{r}) = \sum_j \Delta_{ij}(\mathbf{r})^2$ となり、これは $-\log L(\mathbf{r})$ と定数項およびスケール $1/2\sigma_p^2$ を除いて一致する。従って、良く知られているように SSD 最小化と尤度最大化は一致する。

セル番号 i の SSD を最小化する移動ベクトル \mathbf{r} を $\mathbf{r}_{\min i}$ と表記する。そして SSD 分布を $\mathbf{r}_{\min i}$ だけ平行移動した座標形での分布を $d_i(\mathbf{e}_i) = SSD_i(\mathbf{r}_{\min i} + \mathbf{e}_i)$ により表す。このとき、追跡セル全体を用いた移動ベクトルの集合 \mathbf{e}_i の最尤推定は

$$\sum_i d_i(\mathbf{e}_i) = \min \quad (2)$$

となる \mathbf{e}_i である。

3.2 追跡セルの選択

精度の高い移動ベクトルの検出が可能なセルを選択することを考える。移動ベクトルの検出精

度が高いということは、得られる推定値が真値に近い、すなわち、推定値の分散が小さいということである。

推定値のばらつきの評価尺度としては、真値における対数尤度関数の曲率 $-\nabla^2 \log L(\mathbf{e})$ がある。直感的には、この曲率は真値からのずれにどれだけ敏感であるかを示している。この曲率のノイズ分布に対する平均値 $I(\mathbf{e}) = -E[\nabla^2 \log L(\mathbf{e})]$ は、平均的な推定精度を示しており、フィッシャー情報行列として知られている量である。最尤推定を行う場合には、推定量が達成可能な最小の分散 $I(\mathbf{e})^{-1}$ が漸近的に達成可能されるので、 $I(\mathbf{e})$ を推定精度の目安とするのが自然である。

セル i の画素 j の座標を $\mathbf{x}_{ij} = (x_{ij}, y_{ij})$ とする。時刻 t から時刻 $t+1$ の間に \mathbf{x}_{ij} の画素に生じたノイズを $\epsilon(\mathbf{x}_{ij})$ とし、真の移動ベクトル $\bar{\mathbf{r}}_i$ から $\bar{\mathbf{e}}$ だけずれた画素について

$$\begin{aligned} \Delta_{ij}(\bar{\mathbf{r}}_i + \bar{\mathbf{e}}) &= f_{t+1}(\mathbf{x}_{ij} + \bar{\mathbf{r}}_i + \bar{\mathbf{e}}) - f_t(\mathbf{x}_{ij}) \\ &= f_t(\mathbf{x}_{ij} + \bar{\mathbf{e}}) - f_t(\mathbf{x}_{ij}) + \epsilon(\mathbf{x}_{ij}) \end{aligned}$$

となる。 $\epsilon(\mathbf{x}_{ij})$ は $\bar{\mathbf{e}}$ に独立であることから、(1)および(3)より

$$\begin{aligned} I(\bar{\mathbf{e}}) &= -E[\nabla^2 \log L(\bar{\mathbf{e}})] \\ &= \frac{1}{2\sigma_p^2} \nabla^2 \left\{ \sum_j (f_t(\mathbf{x}_{ij} + \bar{\mathbf{e}}) - f_t(\mathbf{x}_{ij}))^2 \right\} \end{aligned}$$

となり、真値におけるフィッシャー情報行列は

$$I(\bar{\mathbf{0}}) = \frac{1}{2\sigma_p^2} \nabla^2 \left\{ \sum_j (f_t(\mathbf{x}_{ij} + \bar{\mathbf{e}}) - f_t(\mathbf{x}_{ij}))^2 \right\} \Big|_{\bar{\mathbf{e}}=\mathbf{0}}$$

となる。

$\nabla^2 \{ \sum_j (f_t(\mathbf{x}_{ij} + \bar{\mathbf{e}}) - f_t(\mathbf{x}_{ij}))^2 \}$ は時刻 t のフレーム内におけるセル i のSSD曲面の曲率を示している。SSD分布は離散値として観測されるので、この観測値から最小二乗法による曲率の当てはめを行う。これには[3]により最適なフィルタが与えられているので、これを用いる。

追跡セル選択の基準としては、真の移動ベクトルにおける曲率が最小となる方向の曲率の逆数を用いる。すなわち、フィッシャー情報行列 $I(\bar{\mathbf{0}})$ の2つの固有値を λ_0, λ_1 としたとき、 $1/\max(\lambda_0, \lambda_1)$ をセルの信頼度と定義する。これは、推定値の分散の最悪値が小さいほど信頼性が高くなるということになる。このような信頼度の定義の基で、追跡セルの選択を以下のように行う。

1. $l = 0$ とする。また、レベル0のセル全てを追跡セルの候補とする。
2. レベル l のセルのうち、追跡セルの候補となっているセルのSSD曲面をレベル l のサブサンプリングされた画像を用いて求める。SSD曲面よりセルの信頼度を求め、信頼度の高い上位 N 個のセルを追跡セルと決定する。
3. $l = L$ なら4へ。そうでなければ、レベル l の追跡セルを分割してレベル $l+1$ のセルを生成し、これらを追跡セルの候補とし、 $l = l+1$ として2を行う。
4. $0 \leq l \leq L-1$ となるレベル l の追跡セルについて、レベル L の追跡セルを下位レベルに含まない追跡セルを追跡セルから取り除く。

4 追跡処理

ブロックマッチングは処理が高速であり、実装が容易である反面、単純な平行移動をモデルとしているために変形により誤差を生じやすいことや、勾配法で用いられているような隣接した点の移動ベクトルの類似性に関する拘束条件がないので、不連続な移動ベクトルが得られたり外れ値が生じるという欠点がある。追跡処理では、高速性を重視してブロックマッチングをベースとした処理を行うが、この際、移動ベクトルのばらつきを抑えることができる階層的ブロックマッチングを行う。

また、アフィン変換パラメータはブロックマッチングの結果得られた移動ベクトルを用いて推定する。このとき、外れ値やオクルージョンの影響を極力なくするため、ロバスト統計[4][5]を用いた推定法を採用する。

4.1 階層的ブロックマッチング

階層的ブロックマッチングでは、上位レベルの画像により大きな移動量を求め、下位レベルの画像で上位レベルのブロックマッチングの結果を修正して細かな移動ベクトルを決定する。このようにすることで、外れ値が少なくなめらかに変化する移動ベクトルが得られる。各レベルでのブロックマッチングに用いるブロックサイズを $W_w \times W_w$ 、探索範囲を $W_s^{(l)} \times W_s^{(l)}$ とする。ブロックサイズが同じでも、上位のレベルでは下位レベルよりも大きい領域に対応する。以下に、階層的ブロックマッチングの処理手順を示す。

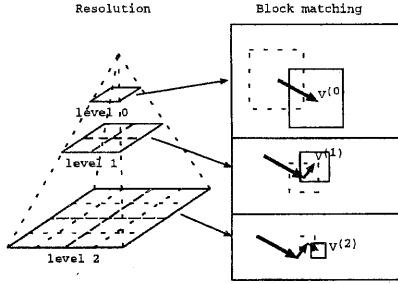


図 3: 階層的ブロックマッチング ($L=2$ のとき)

1. レベル 0 の追跡セルに移動ベクトルの初期値 $\mathbf{v}^{(-1)} = \mathbf{0}$ を付与する。 $l=0$ とする。
2. $\mathbf{v}^{(l-1)}$ が付与されたレベル l の追跡セルに対し、 $\mathbf{v}^{(l-1)}$ を中心とする $W_s^{(l)} \times W_s^{(l)}$ を探索範囲とし、レベル l の画像において SSD によるブロックマッチングを行う。その結果得られた移動ベクトル $\mathbf{v}^{(l)}$ をレベル $l+1$ の画像に移動ベクトルの初期値として付与する。
3. $l=L$ なら終了。そうでなければ $l=l+1$ として 2 へ。

4.2 物体変型のアフィン変換近似モデル

大きさの変化や追跡物体の回転等、変型にロバストな追跡を行うため、物体の変型をアフィン変換で近似したモデルを用いる。変型のアフィン変換パラメータ \mathbf{a} を $\mathbf{a} = (a_0, a_1, a_2, a_3, a_4, a_5)^T$ とする。このとき、移動前の追跡物体上の点の位置を $\mathbf{x}_i = (x_i, y_i)$ 、ブロックマッチングの結果得られた移動ベクトルが $\mathbf{v}_i^{(L)} = (u_i, v_i)$ であったとするとアフィン変換変型モデルによる移動先の点の位置 \mathbf{r}_i およびブロックマッチングにより得られた移動ベクトルだけ移動した点の位置 $\mathbf{r}_{\min i}$ は

$$R_i = \begin{pmatrix} x_i & y_i & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_i & y_i & 1 \end{pmatrix}$$

として $\mathbf{r}_{\min i} = \mathbf{x}_i + \mathbf{v}_i^{(L)}$ 、 $\mathbf{r}_i = R_i \mathbf{a}$ となる。また、パラメータを \mathbf{a} と見積もったことによる誤差 \mathbf{e} は $\mathbf{e} = \mathbf{r}_{\min} - \mathbf{r}$ となり、SSD での推定残差は $d(\mathbf{e})$ である。パラメータ \mathbf{a} の最尤推定では (2) を最小とする \mathbf{a} を推定する。

ここで、各追跡セルの SSD の分布 $d_i(\mathbf{e}_i)$ をどのように与えるかが問題となる。実際に算出され

た SSD を用いて $d_i(\mathbf{e}_i) = SSD_i(\mathbf{r}_{\min i} + \mathbf{e}_i)$ とした場合、 $\sum_i d_i(\mathbf{e}_i)$ を最小化するためには、全ての \mathbf{a} の値を試さなければならない。これを避けるために、次節で $d_i(\mathbf{e}_i)$ を \mathbf{e}_i の適当な関数に近似する。

4.3 ロバスト推定によるアフィンパラメータ推定

アフィン変換のパラメータ \mathbf{a} の尤度関数を $L(\mathbf{a}) = \prod_i p(\mathbf{e}_i/\sigma_i)$ とおけば、最尤推定法は

$$-\log L(\mathbf{a}) = \sum_i -\log p(\mathbf{e}_i/\sigma_i) = \min$$

となる。 σ_i は各セルごとの \mathbf{e}_i のばらつきのスケールである。これを $\psi(\mathbf{e}_i/\sigma_i) = -\log p(\mathbf{e}_i/\sigma_i)$ を用いて

$$\sum_i \psi(\mathbf{e}_i/\sigma_i) = \min \quad (4)$$

と表記する。この解は、 a_j で偏微分して 0 とおくことで得られる。すなわち、 ρ を ψ の導関数として、 $\mathbf{e}_i = \mathbf{r}_{\min i} - R_i \mathbf{a}$ を用いて

$$\sum_i \rho_j(\mathbf{e}_i/\sigma_i) R_{ij} = 0, \quad j = 0, 1, \dots, 5 \quad (5)$$

$$\rho_j(\mathbf{e}_i/\sigma_i) = \frac{\partial}{\partial e_{ij}} \psi(\mathbf{e}_i/\sigma_i)$$

を連立させて解くことができる。ただし、

$$(R_{i0}, R_{i1}, R_{i2}, R_{i3}, R_{i4}, R_{i5}) = (x_i, y_i, 1, x_i, y_i, 1),$$

$$e_{ij} = \begin{cases} e_{ix}, & j = 0, 1, 2 \\ e_{iy}, & j = 3, 4, 5 \end{cases}$$

最小二乗法は、

$$\psi(\mathbf{e}_i/\sigma_i) = \|\mathbf{e}_i/\sigma_i\|^2, \quad \rho_j(\mathbf{e}_i/\sigma_i) = e_{ij}/\sigma_i$$

とおく \mathbf{a} の推定方法である。これはブロックマッチングの結果が真の移動ベクトルから \mathbf{e} だけずれる確率が等方性の正規分布

$$p(\mathbf{e}_i/\sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left\{-\frac{\|\mathbf{e}_i\|^2}{2\sigma_i^2}\right\}$$

に従っているときの最尤推定法 (4), (5) 式

$$\psi(\mathbf{e}_i/\sigma_i) = \frac{\|\mathbf{e}_i\|^2}{2\sigma_i^2} + \frac{1}{2} \log 2\pi + \log \sigma_i,$$

$$\rho_j(\mathbf{e}_i/\sigma_i) = e_{ij}/\sigma_i$$

と一致する。従って、観測結果 $d_i(\mathbf{e}_i)$ を

$$d_i(\mathbf{e}_i) \approx \|\mathbf{e}_i/\hat{\sigma}_i\|^2 + \hat{\beta}_i$$

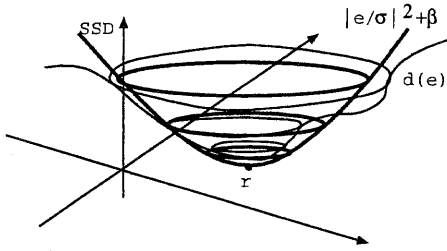


図 4: SSD 曲面の 2 次関数への近似

と近似すれば、近似的に 3.1 のノイズモデルにおける最尤推定が行える。よって、 $\hat{\sigma}_i, \hat{\rho}_i$ は $r_{\min i}$ の周辺の SSD 値を用いた最小二乗法により推定する。

ロバスト統計の M 推定法は最小二乗法を一般化し、 ψ または ρ を任意の関数に拡張したものである。Tukey は M 推定法として Biweight 推定法と呼ばれる方法を提案している [5]。これは

$$\rho_j(z) = \begin{cases} z(1 - (z_j/c)^2)^2 & |z_j| < c \\ 0 & \text{otherwise} \end{cases}$$

を用いる方法である。定数 c は 5 ~ 9 に選ぶのが良いとされている。以下に Biweight 法によるアフィン変換パラメータ \mathbf{a} の解法を示す。

1. 初期解を重み付き最小二乗法で求める。これを $\hat{\mathbf{a}}^{(0)}$ とする。
2. $i-1$ 回目の推定パラメータ $\hat{\mathbf{a}}^{(i-1)}$ を用いた推定値との残差

$$|\nu_i| = \left| \left(r_{\min i} - R_i \hat{\mathbf{a}}^{(i-1)} \right) / \hat{\sigma}_i \right|$$

を計算し、残差の推定量を $|\nu_i|$ の中央値 $s = \text{median}(|\nu_i|)$ として求める。

3. 調節重み関数のスケール s により規格化された重み w_i を (6) 式から求める。

$$w_i = \begin{cases} \left(1 - \left(\frac{|\nu_i|}{cs} \right)^2 \right)^2 & |\nu_i| < cs \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

4. 重み w_i を用いた重み付き最小二乗法

$$\sum_i \left| \left(r_{\min i} - R_i \mathbf{a}^{(i)} \right) / \hat{\sigma}_i \right|^2 w_i = \min \quad (7)$$

で i 回目の推定パラメータ $\hat{\mathbf{a}}^{(i)}$ を求める。

5. $i < M$ ならば 2 に戻る。 $i = M$ なら終了。

繰り返し回数 M は 6 程度でよいとされている。

また、(7) 式の重み付き最小二乗法は

$$\begin{aligned} \mathbf{Q} &= (\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_5)^T, \\ \mathbf{q}_j &= \left\{ \sum_i \left(\frac{R_i^T R_i}{\hat{\sigma}_i^2} \frac{\partial \mathbf{a}}{\partial a_j} w_i \right) \right\}^T, \\ \mathbf{t} &= (t_0, t_1, \dots, t_5)^T, \\ t_j &= \sum_i \left(\frac{r_{\min i}^T R_i}{\hat{\sigma}_i^2} \frac{\partial \mathbf{a}}{\partial a_j} w_i \right) \end{aligned}$$

として、

$$\hat{\mathbf{a}} = \mathbf{Q}^{-1} \mathbf{t}$$

により解くことができる。

M 推定の結果から、推定したパラメータ $\hat{\mathbf{a}}$ のばらつきを推定することができる。サンプル数 N 、パラメータ数 6 として

$$\text{Cov}(\hat{\mathbf{a}}) = K^2 \frac{1/(N-6) \sum_{ij} \rho_j^2(\hat{e}_i/\hat{\sigma}_i)}{(1/N \sum_{ij} \rho_j'(\hat{e}_i/\hat{\sigma}_i))^2} \left(\sum_i \frac{R_i^T R_i}{\sigma_i} \right)^{-1},$$

$$K = 1 + \frac{6 V[\rho']}{N E[\rho']^2}$$

ただし、 $E[\cdot]$ および $V[\cdot]$ はそれぞれ平均、分散を表しており、

$$E[\rho'] \approx \frac{1}{n} \sum_{ij} \rho_j'(\hat{e}_i/\hat{\sigma}_i),$$

$$V[\rho'] \approx \frac{1}{n} \sum_{ij} (\rho_j'(\hat{e}_i/\hat{\sigma}_i) - E[\rho'])^2$$

により推定する。この共分散が大きいときには、アフィン変換モデルが当てはまらないことになるため、テンプレート更新、追跡の中断、追跡領域の再指定等の判定に利用することができる。ここでは $\text{Cov}(\hat{\mathbf{a}})$ の固有値 $\lambda_0, \lambda_1, \dots, \lambda_5$ の逆数の最小値を $\hat{\mathbf{a}}$ の信頼性と定義し、これがしきい値よりも大きいときにのみテンプレートの更新を行うこととする。テンプレートの更新はセルの中心座標をアフィン変換 $\hat{\mathbf{a}}$ により変換した先を新たなセルの中心とすることにより行う。

5 実験

提案手法の追跡性能を確かめるため、テスト映像を用いた実験の結果を示す。ここで用いたのはサイズが $352 \times 240, 30$ フレーム / 秒の映像で

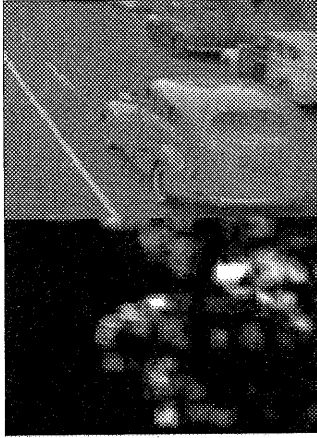


図 5: テスト映像1(フレーム0)の一部(上)と、各画素を中心とするサイズ 8×8 のセルのブロックマッチング信頼性

ある。また、実験に用いたのは Intel 社製 PentiumII266MHz を CPU とする計算機である。テスト映像は 30 フレーム / 秒であるが、これを 1/3 に間引いて処理を行ったため、実質的には 10 フレーム / 秒のフレームレートである。このフレーム間引きの条件の下で、ソフトウェアによるリアルタイム処理が可能であった。なお、用いたパラメータは $N = 50$, $W_w = 4$, $W_s^{(0)} = 31$, $W_s^{(1)} = 3$, $W_s^{(2)} = 3$, $M = 6$ である。

図 5 上はテスト映像 1 の初期フレームの一部である。提案追跡手法では図 5 のフレームを用いて追跡セルを決定する。図 5 下は、各画素を中心とするサイズ 8×8 のセルを構成した場合に、これらセルを用いてブロックマッチングを行ったときの移動ベクトル推定精度を示したものである。輝度の高い画素ほど信頼性が高いことを示している。点や角を含む領域は高い信頼性を示している一方、左下部分などの直線においては、直線の垂直方向の推定精度は高いものの平行方向の推定精度が低い。そのため、全体として信頼性が低くなっていることがわかる。

図 6 は映像中の車両を追跡領域として指定したときに 8×8 のセル 25 個を選択した結果を示している。外側の大きな矩形が指定した追跡領域を示しており、小さな矩形が選択された追跡セルである。

図 7 はテスト映像 2 における 2 フレームを示し

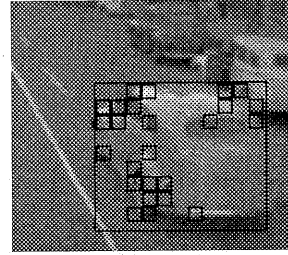


図 6: ブロックマッチングの信頼性の高い上位 25 セルを選択した結果. セルサイズは 8×8 , 大きな矩形は指定された追跡領域を示す。

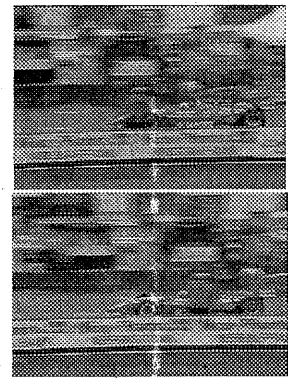


図 7: テスト映像 2(上: フレーム 6, 下: フレーム 9)

たものである。図 7 上ではレーシングカーの前輪前付近に遮蔽された部分がある。この遮蔽物は図 7 下ではレーシングカーの中央部分に移動している。図 8 上は上記 2 つの画像間で階層的ブロックマッチングを行った結果得られた移動ベクトルを図示したものである。遮蔽物の影響により、右半分の移動ベクトルが大きな外れ値となっている。図 8 下は図 8 上の移動ベクトルからロバスト推定によりアフィン変形パラメータを推定し、その結果得られた移動ベクトルを示したものである。遮蔽物により生じた外れ値に影響されることなく、レーシングカーの変形・移動が推定されていることがわかる。

図 9 はテスト映像 3 の追跡結果を示したものである。追跡結果は追跡領域を囲む楕円で表示している。テスト映像 3 では追跡物体の変形が大き

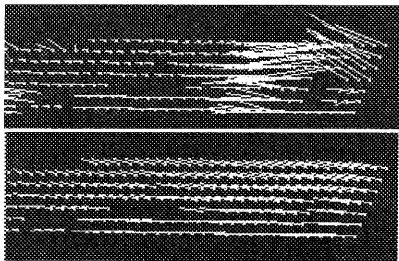


図 8: テスト映像 2 のフレーム 3 からフレーム 6 への移動ベクトルを階層的ブロックマッチングを用いて推定した結果 (上) と, ロバスト推定により推定されたアフィン変形パラメータにより移動ベクトルを求めた結果 (下)

く, 図 9 中においては追跡領域の左側が追跡物体の傾きにより隠れてしまうなど, ブロックマッチングをベースとした追跡がしにくい映像である。しかし, 図 9 下において追跡領域が多少小さくなってはいるものの, 本手法において追跡が可能であることがわかる。

6 まとめと課題

本報告では, 映像中の登場物体のアンカー領域指定を支援するための物体追跡手法の提案を行った。この手法の特徴は, 信頼性の高い追跡セルを選択することにより高速に処理が可能であること, 物体の変形をアフィン変換に近似し, パラメータ推定にロバスト統計を用いることで安定した物体の変形への追従が可能なことである。

今後の課題としては, セルのブロックマッチングにおける信頼度を変形の度合いを考慮するように拡張することが考えられる。これには, 変形の度合いを予め決定しておき, 信頼度を見積もるという方法と, セル内の誤差分布から変形の度合いを推定する方法がある。しかし, 変形の度合いを予め正しく決定しておくことは難しく, また推定値を反映させるには繰り返し処理が必要になるという問題がある。

また, セルの信頼度及び SSD 曲面の方向による偏りを持たせたまま利用することにより, 追跡性能の改善が期待できる。しかし, この場合算出される式が複雑になるため, 適当な近似を用いないと処理時間が大きくなるという問題がある。

さらに, テンプレートを順次更新していくため

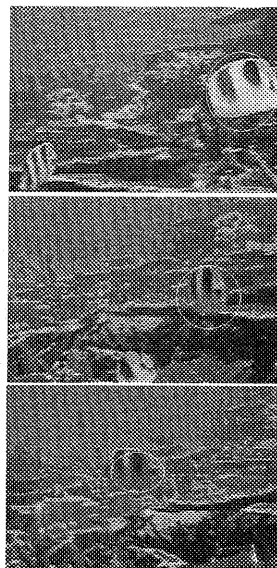


図 9: テスト映像 3 の追跡結果 (上: フレーム 36, 中: フレーム 78, 下: フレーム 108)

に生じる推定誤差の蓄積をどう抑えるかという課題がある。

参考文献

- [1] 神田準三郎, 脇本浩司, 阿部博信, 田中聡. “動画像ハイパーメディアシステムにおけるアンカー自動設定方法とその評価”. 信学技報, PRMU96-146, 1997.
- [2] 伊藤渡, 上田博唯. “物体の見せかけの大きさ変動に追従できるテンプレートマッチング法”. 信学技報, PRMU96-43, 1996.
- [3] P.R.Beaudet. “rotationally invariant operators”. In *Proceedings of the 4th ICPR*, pages 579-583, 1978.
- [4] P.J.Huber. “*Robust Statistics*”. Wiley, New York, 1981.
- [5] 中川徹, 小柳義夫. “最小二乗法による実験データ解析”. 東京大学出版会, 1982.