

SIFT 特徴量に基づく Mean-Shift 探索による特徴点追跡

都 築 勇 司^{†1} 藤 吉 弘 亘^{†1} 金 出 武 雄^{†2}

KLT 法 (Kanade-Lucas-Tomasi Feature Tracker) による特徴点追跡アルゴリズムは高速であり広く利用されているが、人などの非剛体の特徴点追跡に失敗することがある。本論文では、回転・スケール変化・照明変化による画像の変化に不変な SIFT (Scale Invariant Feature Transform) 特徴量を用いた Mean-Shift 探索による特徴点追跡法を提案する。評価実験により、回転やスケール変化を含む対象物体の移動に対しても追跡可能であることを確認した。また、追跡結果の表現法として、追跡点をセグメンテーションすることなく動線を表現する方法について述べる。点の移動方向と密度により表示する色を決めることで、移動の方向と頻度を表現する。本手法により、長時間に及ぶ人の追跡結果をビジュアライゼーションし、移動の流れが表現されていることを示す。

Mean Shift-based Point Feature Tracking Using SIFT

YUJI TSUDUKI,^{†1} HIRONOBU FUJIYOSHI^{†1} and TAKEO KANADE^{†2}

The Kanade-Lucas-Tomasi feature tracker algorithm for point feature tracking is widely used because it is fast; however, it is sometimes fails to accurately track non-rigid objects such as pedestrians. We have developed a method of point feature tracking using a scale invariant feature transform (SIFT). Our approach uses mean-shift searching to track a point based on the information obtained by a SIFT. Since a SIFT feature is invariant to changes caused by rotation, scaling, and illumination, we can obtain a better tracking performance than that of a conventional approach. Using the trajectory of the points obtained by our method, it is possible to visualize traffic pedestrian traffic flow.

1. はじめに

Structure from Motion による 3 次元形状の復元¹⁾ や、移動物体のセグメンテーション²⁾ には長時間に及ぶ特徴点追跡が必要である。特徴点追跡の手法としてすでに KLT 法 (Kanade-Lucas-Tomasi Feature Tracker)^{3),4)} が提案され、多くのアプリケーションで用いられている。KLT 法は、局所領域における各点の動きは同一であると仮定し、弛緩法により目的関数を最小化する手法である。微小時間において領域は平行移動のみしかしない、照明の変化による輝度値の変化がない、という条件を仮定して移動先を求める。そのため、対象物体の運動に回転やスケール変化を含む場合や、照明の変化による輝度値の変化が激しい場合に、特徴点追跡に失敗することがある。KLT 法を改良し、回転やスケール変化などにも対応可能な手法^{5),6)} も提案されているが、目的関数を決める際に運動モデルを

仮定する必要がある、平行移動成分と回転・スケール変化成分を別々に求める必要がある。

本論文では、回転やスケール変化に対して頑健な特徴点追跡手法を提案する。提案手法は、SIFT⁷⁾ (Scale Invariant Feature Transform) により記述された特徴量を重みとして Mean-Shift 探索⁸⁾⁻¹¹⁾ することで特徴点の追跡を行う。Mean-Shift による追跡は、局所的な重み分布に対する山登り計算である。特徴量にはカラーヒストグラムが用いられる場合が多く、ヒストグラム間類似度関数に従い求められる各画素の重みの分布に対して探索を行うため、追跡対象の形状変化や部分的なオクルージョン問題に対して頑健である。文献 12) では色特徴以外に形状特徴として 3 方向のエッジを手がかりとすることで、さらに頑健な追跡を提案している。ただし、これらの特徴は領域を対象とした追跡に適したものであり、点特徴の追跡には適していない。提案手法で用いる SIFT は、回転・スケール変化・照明変化による画像の変化に不変な特徴量を画素ごとに記述することが可能であるため、Mean-Shift 探索による特徴点の追跡が可能となる。

次に、特徴点追跡結果の応用として、人の流れであ

^{†1} 中部大学大学院工学研究科情報工学専攻

Department of Computer Science, Chubu University

^{†2} カーネギーメロン大学ロボティクス工学研究所

The Robotics Institute, Carnegie Mellon University

る動線のビジュアライゼーションについて述べる．各追跡点において，SIFTにより検出されたスケールを考慮することで，移動体領域全体の動きを表現することが可能である．

2. SIFT

Lowe によって提案された SIFT は，グレースケール画像に対して，特徴の抽出に適した点（以下，キーポイントと呼ぶ）を検出し，回転・スケール変化・照明変化に不変な特徴量を記述する手法である．検出段階はスケール空間極値検出，キーポイントのローカライズからなり，記述段階はオリエンテーション割当てと，SIFT Descriptor による特徴量の抽出からなる．以下に SIFT の各処理について示す．

2.1 スケール空間極値検出

SIFT は，スケール変化による不変性を得るため，対象とする点の特徴をより表現できる近傍領域の範囲を自動的に決定する．範囲を決めるパラメータをスケールといい，キーポイント検出時に最適なスケールを決める．まず，スケール s のガウス関数 $G(x, y, s)$ と入力画像 $I(x, y)$ を畳み込んだ平滑化画像 $L(x, y, s)$ を以下の式により作成する．

$$L(x, y, s) = G(x, y, s) * I(x, y) \quad (1)$$

$$G(x, y, s) = \frac{1}{2\pi s^2} \exp\left(-\frac{x^2 + y^2}{2s^2}\right) \quad (2)$$

スケールの異なる平滑化画像間の差分処理を DOG (Difference of Gaussian) といい， $D(x, y, s)$ で表す． $D(x, y, s)$ は，以下の式より求める．

$$D_i(x, y, s) = L(x, y, s_{i+1}) - L(x, y, s_i) \quad (3)$$

次に，図 1 に示すように DOG 画像 $D(x, y, s)$ から極値の検出を行う．注目画素（図中×印）が 26 近傍領域において極値（最大値もしくは最小値）である場合，キーポイント候補点とする．また，このとき極値となる DOG 画像のスケールをキーポイントの特徴抽出に用いる．

2.2 キーポイントのローカライズ

検出されたキーポイント候補点の中にはエッジ上の点が含まれており，これらの点は開口問題に影響を受けやすいという問題がある．そこで，キーポイント候補点の中から，主曲率によりエッジ上の候補点を削除し，安定したキーポイントに絞り込む．キーポイント候補点における 2 次元ヘッセ行列 \mathbf{H} を次式により計算し，主曲率を求める．

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (4)$$

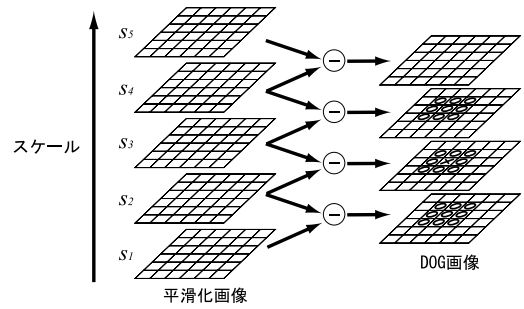


図 1 極値検出の流れ

Fig. 1 Local extrema detection.

この \mathbf{H} の第 1 固有値を α ，第 2 固有値を β とし，その比率を γ とすると以下の式が成り立つ．

$$\alpha = \gamma\beta \quad (5)$$

よって，

$$\frac{\text{Tr}(\mathbf{H})^2}{\text{Det}(\mathbf{H})} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(\gamma + 1)^2}{\gamma} \quad (6)$$

となり，固有値を求めずに主曲率を得ることができる．この値をしきい値処理することにより，不要なキーポイントの除去を行う．また，低コントラストの候補点はノイズの影響を受けやすいため，こちらも削除する．ある点 $\mathbf{x} = (x, y, s)^T$ での，DOG を $D(\mathbf{x})$ としこれをテイラー展開すると以下の式ようになる．

$$D(\mathbf{x}) = D + \frac{\partial D}{\partial \mathbf{x}} \mathbf{x} + \frac{1}{2} \mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2} \mathbf{x} \quad (7)$$

極値を $\hat{\mathbf{x}}$ とし， \mathbf{x} に関する導関数をとる．

$$\hat{\mathbf{x}} = -\frac{\partial^2 D^{-1}}{\partial \mathbf{x}^2} \frac{\partial D}{\partial \mathbf{x}} \quad (8)$$

この式を 0 とし，式 (7) に代入すると，以下の式が得られる．

$$D(\hat{\mathbf{x}}) = D + \frac{1}{2} \frac{\partial D}{\partial \mathbf{x}} \hat{\mathbf{x}} \quad (9)$$

これは，極値 $\hat{\mathbf{x}}$ のコントラストを表しており，この値の絶対値が低い場合， $\hat{\mathbf{x}}$ は候補点から除去する．

2.3 オリエンテーション割当て

検出された各キーポイントに対して，オリエンテーションを割り当てる．キーポイントが検出された平滑化画像の各画素の勾配 $m(x, y)$ とその勾配方向 $\theta(x, y)$ を，以下の式より求める．

$$m(x, y) = \sqrt{f_x(x, y)^2 + f_y(x, y)^2} \quad (10)$$

$$\theta(x, y) = \tan^{-1} \frac{f_y(x, y)}{f_x(x, y)} \quad (11)$$

$$\begin{cases} f_x(x, y) = L(x+1, y) - L(x-1, y) \\ f_y(x, y) = L(x, y+1) - L(x, y-1) \end{cases} \quad (12)$$

求めた勾配の大きさ m と勾配方向 θ (図 2 左側)

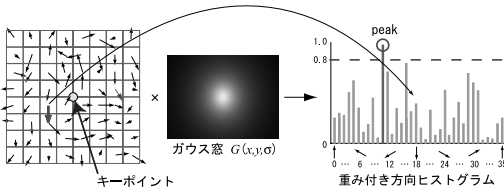


図 2 重み付き方向ヒストグラム
Fig. 2 Weighted orientation histogram.

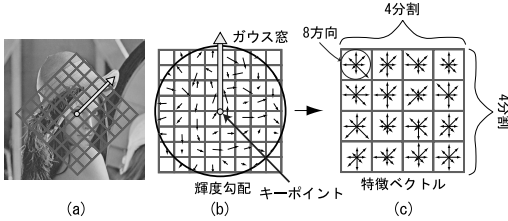


図 3 Keypoint descriptor.
Fig. 3 Keypoint descriptor.

と、キーポイントのスケールに基づくガウス分布(図 2 中央)を掛け合わせた値より、36 方向に分割した重み付き方向ヒストグラムを作成する(図 2 右側)。次に、作成した 36 方向のヒストグラムから、最大値の 80%以上になる方向をキーポイントのオリエンテーションとして割り当てる。図 2 では、キーポイントに割り当てられるオリエンテーションは 1 方向のみであるが、80%以上のピークが複数ある場合、それらすべてがキーポイントのオリエンテーションとなる。たとえば、キーポイントがコーナー上の点である場合、1 つのキーポイントに対して 2 方向のオリエンテーションが割り当てられることになる。

2.4 SIFT Descriptor による特徴量抽出

キーポイントの周辺領域を、図 3 (a) に示すように割り当てられたオリエンテーション方向に回転する。このように、キーポイントの向きにあわせて特徴量を記述するため、回転に対する不変性が得られる。特徴量の記述の際には、図 3 (b) に示すように、ガウス窓を用いて中心付近が強くなるように重み付けをする。このガウス窓の大きさは、キーポイントが検出された DOG 画像の平滑化スケールによって決定される。画像サイズが 2 倍であればスケールも 2 倍となり、構造的に同一の範囲における特徴量の記述を行うことになる。これにより、スケール変化に対する不変性が得られる。次に、分割した領域ごとに方向ヒストグラムを作成する。図 3 (c) のように、領域を 4 × 4 の 16 に分割し、それぞれに 8 方向のヒストグラムを作成するため、128 次元のベクトルの特徴量を持つことになる。

3. SIFT 特徴を用いた Mean-Shift 追跡

SIFT は、回転・スケール変化・照明変化による画像の変化に不変な特徴量を画素ごとに記述することが可能である。また、この特徴量は近傍ピクセルでは類似度が高くなるという傾向がある。そこで、本論文では SIFT 特徴量から得られる重み分布の重心へと移動する Mean-Shift による特徴点追跡法を提案する。

3.1 SIFT 特徴量を用いた Mean-Shift アルゴリズム

提案手法は、Mean-Shift アルゴリズムにより、キーポイントの移動先を決定する。提案手法の追跡処理の流れを図 4 に示す。

3.1.1 初期追跡点検出

初期フレームの画像より初期追跡点を検出する。初期追跡点には、SIFT で検出されるキーポイントを用いる。各点には、SIFT で記述された 128 次元のベクトル $\mathbf{v} = (v_0, \dots, v_{127})$ とスケール s を参照用データとして保持する。

3.1.2 Mean-Shift 探索

SIFT は、スケールによって特徴を抽出する範囲を変えるため、追跡対象の画像上での大きさが変化する場合は、適応してスケール値を更新する必要がある。そこで、提案手法の Mean-Shift 探索では、画像空間の移動量 $\Delta \mathbf{x} = (\Delta x, \Delta y)$ とスケール空間の移動量 Δs を求める。SIFT 特徴量から求めた重み分布を用いて、画像空間とスケール空間の移動量を交互に求め、現フレームでの最適な位置とスケール値に移動するよう計算を繰り返す。

step1 画像空間の Mean-Shift

追跡中心 \mathbf{x} の周辺画素 $\mathbf{x}_i (i = 0, \dots, N)$ および参照用スケール s の SIFT 特徴ベクトルと \mathbf{x} における参照用ベクトル \mathbf{v} との距離より重み $\omega(\mathbf{x}_i, s)$ を求める。

$$\omega(\mathbf{x}_i, s) = \exp\left(\frac{-d(\mathbf{x}_i, s)^2}{2\sigma_d^2}\right) \tag{13}$$

$$d(\mathbf{x}_i, s) = \|\text{SIFT}(\mathbf{x}_i, s) - \mathbf{v}\| = \sqrt{\sum_{k=0}^{127} (v_{i,k} - v_k)^2} \tag{14}$$

$$\text{SIFT}(\mathbf{x}_i, s) = \mathbf{v}_i = (v_{i,0}, \dots, v_{i,127}) \tag{15}$$

σ_d は、SIFT の類似度を重みに変換するカーネル関数のパラメータである。SIFT(\mathbf{x}_i, s) は、座標 \mathbf{x}_i においてスケール s で SIFT 特徴ベクトルを出力する関数である。求めた重み $\omega(\mathbf{x}_i, s)$ を用いて画像空間上の移動量 $\Delta \mathbf{x}$ を求める。前フレームの追跡領域周辺に注目した探索を行うため、式 (17) に示す画像空間の

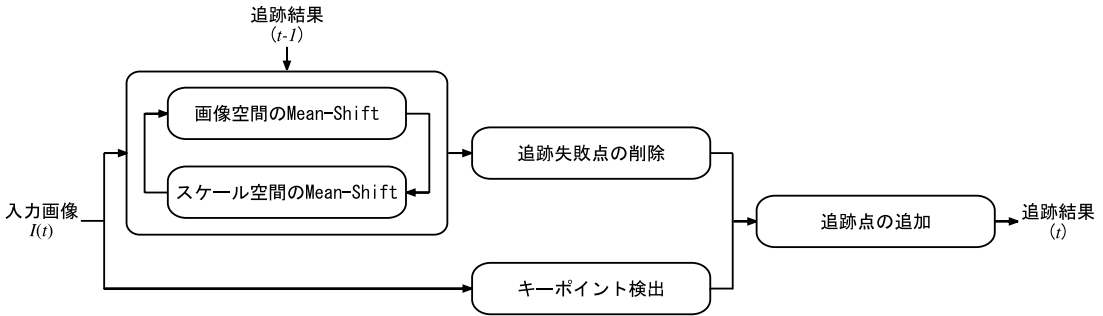


図4 提案手法の流れ

Fig. 4 Process of keypoint tracking using an image sequence.

カーネル関数 $K_{loc}(\mathbf{x}, \sigma_{xy})$ を用い、移動量 $\Delta\mathbf{x}$ を次式より求める。

$$\Delta\mathbf{x} = \frac{\sum_{i=0}^N K_{loc}(\mathbf{x}_i - \mathbf{x}, \sigma_{xy}) \omega(\mathbf{x}_i, s) (\mathbf{x}_i - \mathbf{x}_0)}{\sum_{i=0}^N K_{loc}(\mathbf{x}_i - \mathbf{x}, \sigma_{xy}) \omega(\mathbf{x}_i, s)} \quad (16)$$

$$K_{loc}(\mathbf{x}, \sigma_{xy}) = \exp\left(\frac{-(x^2 + y^2)}{2\sigma_{xy}^2}\right) \quad (17)$$

移動量 $\Delta\mathbf{x}$ から、 $\mathbf{x}' = \mathbf{x} + \Delta\mathbf{x}$ により、追跡点座標をシフトする。

step2 スケール空間の Mean-Shift

次に、Step1 より求めた移動後の座標 \mathbf{x}' でのスケール空間における移動量 Δs を求める。 Δs は追跡対象の大きさの変化に合わせる値であり、参照用スケール s と追跡対象の大きさの変化率を掛けた値が、推定すべき値となる。参照用スケール s が大きければ Δs の値も大きくなるため、スケール空間での探索をする場合、参照用スケール s に合わせて探索範囲を変えなくてはならない。

そこで、探索を行う空間の軸をスケール値ではなく、スケールの変化率 S とし、 ΔS を探索する。画像空間の場合と同様に、式 (13) を重みとして以下の式により変化率を求める。

$$\Delta S = \frac{\sum_{j=0}^M K_{scale}(S_j - 1, \sigma_s) \omega(\mathbf{x}', s S_j) (S_j)}{\sum_{j=0}^M K_{scale}(S_j - 1, \sigma_s) \omega(\mathbf{x}', s S_j)} \quad (18)$$

$$K_{scale}(S, \sigma_s) = \exp\left(\frac{-S^2}{2\sigma_s^2}\right) \quad (19)$$

ここで S_j ($j = 0, \dots, M$) は、たとえば $S_j = (\dots, 0.9, 1.0, 1.1, \dots)$ のような、1.0 を中央として等間隔で増加する数列である。これは参照用スケール s に対する倍率であり、1.0 は参照用スケール s をそのまま用いる場合である。ただし $S_0 > 0$ である。カー

ネル関数への入力に $S - 1$ を用いるのは、現在の値である $S = 1$ においてカーネル関数の出力が最大となるようにするためである。求めた変化率 ΔS から、 $s' = s \Delta S$ により、スケールを更新する。

step3 繰返し処理

Step1, Step2 の処理を $|\Delta\mathbf{x}| < \epsilon_{xy}$, $|\Delta S - 1.0| < \epsilon_s$ を満たすまで各空間の Mean-Shift 探索を交互に繰り返す。最適な追跡点座標とスケールを求める。 ϵ_{xy} と ϵ_s はしきい値である。

3.1.3 追跡失敗点の削除

移動後の座標 \mathbf{x} での SIFT の特徴ベクトルを求め、式 (14) により求めた参照用ベクトルとの距離 $d(\mathbf{x}, s)$ がしきい値以上の場合、追跡に失敗したと判断し削除する。

3.1.4 追跡点の追加

新たに画像内に入力した移動体の追跡を行うために、特徴点の追加を行う。初期追跡点と同様に、キーポイントとして検出された点を新たな追跡点として登録する。ただし、追跡中の点と同じ位置が検出される場合もあるため、各追跡点から一定以上の距離があるキーポイントのみを追加する。

3.2 スケール空間の探索例

スケールは特徴量の記述に用いるピクセルの範囲を示す値である。図5は拡大していく画像に対して追跡を行った際の特徴点とスケールの追跡結果である。図5(a)における画像中の白い円が特徴点の座標を、青い円がスケールの大きさを表している。画像の大きさが変化しても、同一の範囲を選択している。また、図5(b)のグラフは画像の拡大率とスケールの変化率を表している。画像の拡大率の傾きが1.0であるのに対し、推定値を最小二乗法でフィッティングした際の傾きは0.91であるため、画像の拡大率とほぼ同じ割合で提案手法により推定したスケールも変化していることが分かる。提案手法は、画像空間とスケール空間

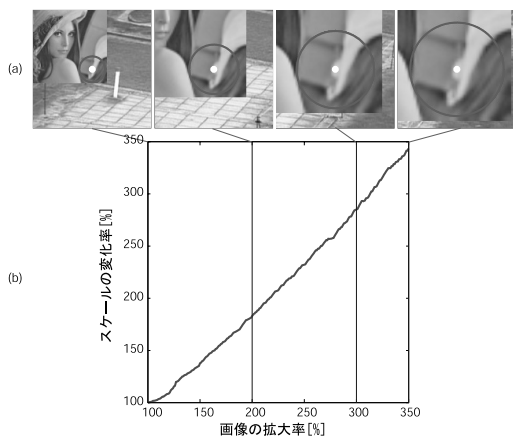


図 5 特徴点とスケールの追跡結果

Fig. 5 Tracking example of location and scale.

の Mean-Shift 探索を繰り返し行うため、特徴点のスケールと移動量を同時に求めることができる。

3.3 提案手法の位置づけ

SIFT は対応点探索に非常に有効な手法のため、時系列画像に適用することで特徴点の追跡が可能となると考えられる。また、文献 14) のように、特徴点どうしの対応を用いて追跡を行う手法も存在する。しかし、これらは特徴点の検出時に対応すべき点が検出されない場合、追跡可能な時間は短いものとなるため、長時間の追跡が可能となる点はわずかとなる。このような手法は少数であっても高精度な追跡点を求める場合に有効であるが、多数の追跡点を求めるのは不向きである。提案手法は、すべての点の移動先を求めることが可能なため、多数の点を長時間追跡できる。また、文献 13) ではエピソード拘束を用いて誤追跡を防止している。提案手法は、これらの拘束を用いる前段階の処理であると位置づけており、状況に適した制限を加えることで、より高精度な追跡が可能になると考えられる。また、KLT 法^{3),4)} や、それに基づく追跡手法^{5),6)} は追跡点としてコーナー点が選ばれやすい。これに対し、提案手法は SIFT により検出された特徴点を追跡する。SIFT による特徴点検出では、局所的にはテクスチャがない領域内の点を選択される場合もある。このような点はスケール値が大きく、広い領域を用いて特徴量を記述する。局所領域においてはテクスチャがなくても、エッジなどを含む広い領域を用いることで、有効な特徴量を記述できる。そのため、提案手法は従来手法では追跡できない点を追跡することが可能である。

4. 評価実験

合成画像を用いた実験により提案手法の有効性を

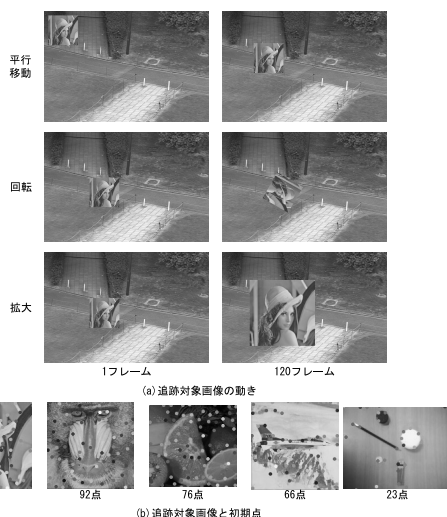


図 6 実験用画像の例

Fig. 6 Example of experimental image.

を示す。

4.1 実験方法

定量的に各手法の追跡精度を比較するため、合成画像による実験を行う。背景画像に、追跡対象画像を貼り付け、平行移動、回転、拡大の動きを持つ合成画像を生成する。ここで、追跡対象画像は 5 種類のものを用いる。各動きに対し、180 フレーム \times 5 種類で 900 フレームの画像を使用する。図 6 (a) に合成した画像の例を示す。提案手法と KLT 法により追跡を行い、追跡結果を比較する。本実験では追跡性能のみを評価するため、提案手法、KLT 法それぞれの手法における追跡の初期点を検出し、両手法とも失敗点の削除は行わない。また、KLT 法の各種パラメータは <http://www.ces.clemson.edu/stb/klt/> で公開されているソースコードのものを使用し、SIFT の各種パラメータは元論文⁷⁾ で示されている値を使用した。図 6 (b) に提案手法での各追跡対象画像における初期点とその数を示す。

4.2 実験結果

初期点と、実験用画像の合成に用いた変換行列より真値を求め、全フレームの各追跡点との追跡実験結果との誤差（ユークリッド距離）を計算する。この誤差があるしきい値以下の誤差で追跡した点の割合を追跡成功率とし、特徴点追跡性能の評価に用いる。図 7 に追跡対象 5 種類のシーケンスでの全フレーム結果から求めた追跡成功率を示す。横軸がしきい値、縦軸が追跡成功率である。また、しきい値を 5 ピクセルとした際の追跡成功率を表 1 に示す。提案手法は、平行移動、回転において KLT 法よりも高い追跡成功率を得

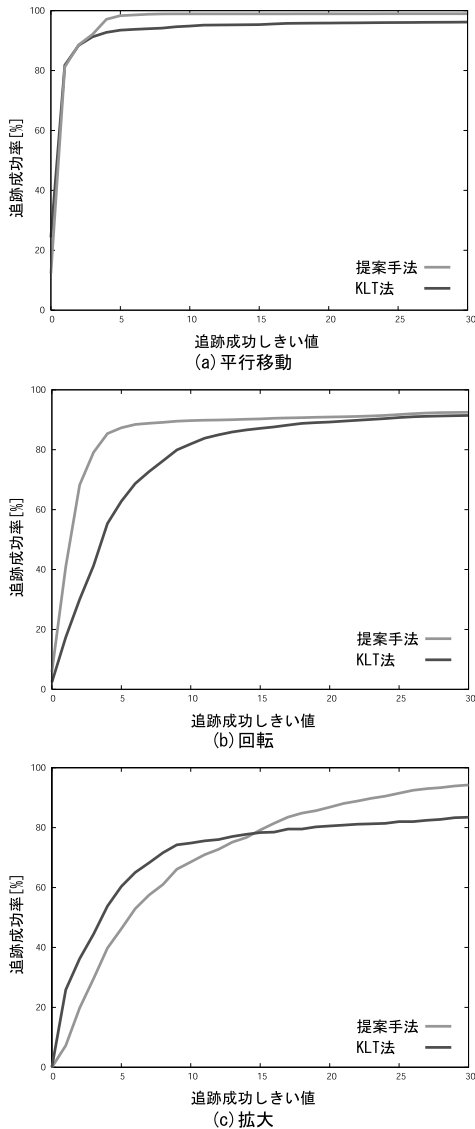


図 7 実験結果

Fig. 7 Experiment result.

表 1 しきい値 5 ピクセルにおける追跡成功率 [%]
 Table 1 Tracking success rate [%] in threshold 5.

	平行移動	回転	拡大	平均
提案手法	98.3	87.3	46.7	77.3
KLT 法	93.4	62.7	61.5	72.5

た．平行移動や回転に対して SIFT 特徴は不変であるため，従来法である KLT 法より精度の良い追跡が可能であることが分かる．

一方，拡大は，しきい値を 5 ピクセルとすると提案手法よりも KLT 法の追跡精度が良い．しかし，しきい値を高くすると（15 ピクセル以上）提案手法の方が追跡成功率が高くなるという傾向が見られた．本実験

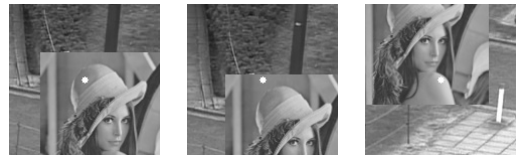


図 8 テクスチャのない領域内の特徴点例

Fig. 8 Examples of feature point which has less texture.

で用いた KLT 法で選択される追跡対象点はコーナー点が多く存在し，拡大による変化の影響を受けにくい．ため，KLT 法でも追跡が可能である．ただし，追跡に失敗する点が多数存在するため，しきい値を高くしても追跡成功率はほとんど上昇しない．一方，提案手法は，逐次，適応的にスケールを探索するため，拡大後も構造的に同じ範囲で特徴を抽出することができる．しかし，SIFT で選ばれる特徴点は物体のエッジなどの点のみではなく，図 8 に示すようなテクスチャのない領域中の点も選択されることもある．このような特徴点はスケールの値が大きく，かなり広い領域から特徴量を抽出するため，特徴点周辺の SIFT 特徴量が非常に近い値をとる．そのため，まったく同じ点を追跡することが困難である場合がある．しかし，特徴点周辺に高い重み分布が発生するため，真値に近い位置を特徴点として追跡することが可能である．ただし，スケールの値によっては追跡対象の領域のみではなく，背景領域も特徴抽出に含まれることがある．この場合は，追跡すべき点で特徴記述を行っても，背景とのずれにより大きく異なる特徴量となり，追跡に失敗する．

4.2.1 スケール変化に対する評価

提案手法は画像空間における座標の探索だけでなく，スケールの探索を行うため，スケール変化に対する追跡の評価を行う．1 フレーム間における追跡対象画像の拡大率を 10%，20%，30%としたシーケンスに対する追跡の，全点のスケール変化率の平均を図 9 に示す．

追跡対象の拡大率が増加するほど，追跡点のスケールの値が真値よりも小さくなる傾向にある．これは，追跡対象の拡大に追従しきれないためだと思われる．

4.3 非剛体の追跡

Mean-Shift では，アフィン変形などのモデルを仮定していないため，非剛体の動きに対する追跡が期待できる．図 10 に，屋外で撮影した人の動きに対する提案手法と KLT 法による追跡結果を示す．各点は 50 フレームの軌跡を表している．目視により追跡成功と判定した各特徴点の追跡フレーム数を求め，評価を行う．このとき，静止物体上の追跡点も，正しい位置に静止していれば成功として判定する．図 11 に，追跡フレーム数の分布を示す．提案手法は KLT 法よりも

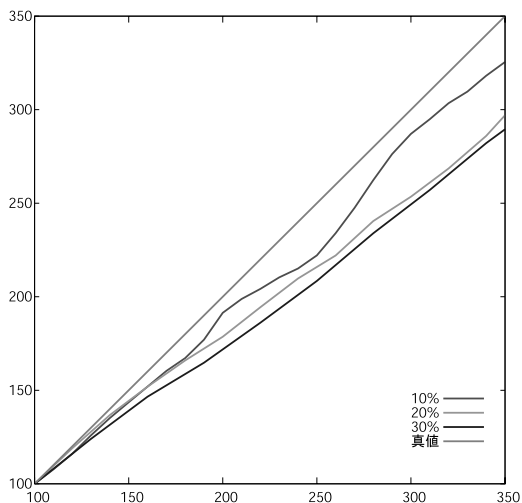


図 9 スケール変化の様子
Fig. 9 Scale change.

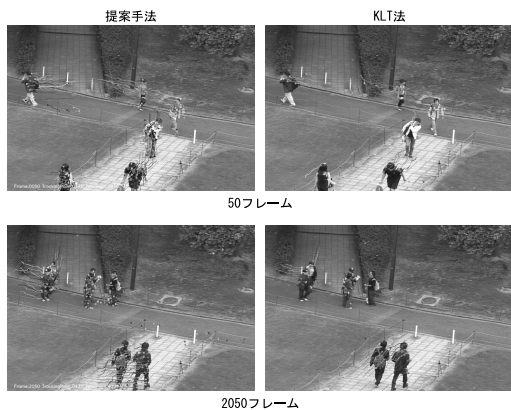


図 10 人の追跡結果

Fig. 10 Examples of feature point tracking for images of pedestrians.

多くの点を長い時間にわたり追跡可能であることが確認できる。

提案手法は、Mean-Shift と SIFT を組み合わせた手法であるため、それぞれ単独での追跡性能と比較し、手法を組み合わせることでの効果を確認する。図 12 はカラーヒストグラムを用いた Mean-Shift による 50 フレーム間の追跡結果である。このとき追跡点は SIFT により得られる点とし、カラーヒストグラム作成のウィンドウサイズは 10×10 の固定とした。図 12 の結果は、図 10 の提案手法や KLT 法と比較し、追跡がほとんどできていないため、カラーヒストグラムを用いた Mean-Shift は特徴点の追跡に適していないといえる。

また、図 13 は SIFT の対応点を用いた 2 フレーム

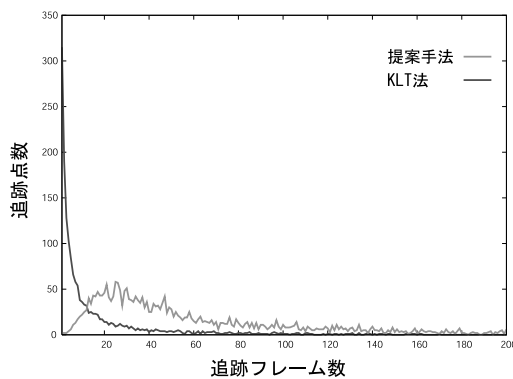


図 11 追跡フレーム数の分布

Fig. 11 Distribution of total number of tracking frames.



図 12 Mean-Shift による追跡結果

Fig. 12 Tracking example by mean-shift.



図 13 SIFT による追跡結果

Fig. 13 Tracking example by SIFT matching.

の結果である。各フレームで検出された特徴点どうしで、最も特徴量のユークリッド距離が近いものを対応点とした。時系列データを用いる場合は、画像空間での探索範囲を限定することで信頼性のある追跡点を選ぶことができるが、SIFT の性能のみを比較するため、画像空間での制限は加えていない。多くの点で誤対応が発生しており、長時間における追跡は不可能である。

カラーヒストグラムを用いた Mean-Shift による追

跡が、特徴点の追跡に適していないのはウィンドウサイズが固定であることや、色のわずかな違いに敏感であるためと思われる。これに対し、提案手法では SIFT を用いることで、各点における最適な周辺の領域を用いて特徴を記述することができる。また、SIFT は照明変化にも頑健であるため、色のわずかな違いの影響を受けにくい。SIFT の対応点では、特徴点の検出の時点で対応する点が見つからない場合や、同一点であっても、わずかな形状の変化により前フレームとは大きく異なるスケールで検出される場合などが誤対応の原因として考えられる。また、特徴量間のユークリッド距離のみを用いるため、画像空間での距離が大きく異なる場合でも対応点としてしまう。提案手法は画像空間での距離を反映させることが可能な Mean-Shift を用い、SIFT で検出されていないピクセルに対しても走査を行うため、非剛体においても長時間かつ多数の点の追跡が可能となる。また、前フレームでのスケールをもとに画像空間での移動量と、スケールの変化量を求めるため、大きく異なるスケールによる特徴量の変化を抑制できる。

5. 特徴点追跡の応用

本章では、提案手法による特徴点追跡の結果を用いた、人の流れである動線のビジュアライゼーション手法について述べる。提案手法は、多数の点を長時間にわたり追跡可能であり、各点がスケールの値を持つ。これらの特性を利用した応用として、人の動線のビジュアライゼーションを行う。オプティカルフローのビジュアライゼーションは¹⁴⁾で行われているが、この方法は提案手法とは異なり、フローの長さや位置を示すものである。また、背景差分により簡単に移動体領域を検出することが可能であるが、背景差分のみからは移動方向を得ることができないため、追跡処理が必要となる。この場合、追跡失敗の影響が領域全体に及ぶため、ビジュアライゼーションに大きな誤りが発生しやすい。本手法は、特徴点の追跡結果に基づきビジュアライゼーションを行うため、数点の追跡失敗が発生した場合でも、他の点の追跡結果により誤りの表示を軽減できる。

5.1 ビジュアライゼーション手法

動線のビジュアライゼーションには、移動方向と頻度を読み取ることが重要である。しかし、人の特徴点追跡では、手や足の動きなどの全体の動きとは異なる細かな動きが発生しやすい。このような点をビジュアライゼーションに用いると、冗長な情報が多くなり見にくい表示となる。そこで、フレーム間で移動方向

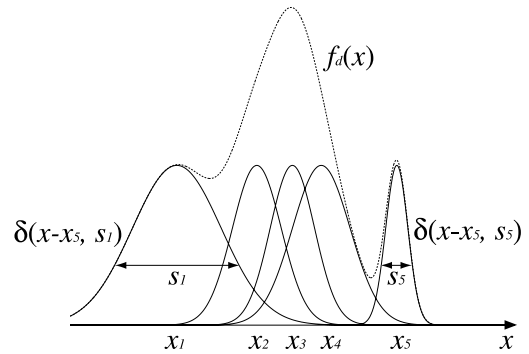


図 14 点の密度による色の強さ

Fig. 14 Color strength by density of points.

に急激な変化がない点を用いるために、以下の基準により追跡点を選択する。

$$\cos \theta = \frac{\mathbf{v}_t \cdot \mathbf{v}_{t-1}}{|\mathbf{v}_t| |\mathbf{v}_{t-1}|} > th \quad (20)$$

$$\mathbf{v}_t = (\mathbf{x}_t, \mathbf{x}_{t-1}), \mathbf{v}_{t-1} = (\mathbf{x}_{t-1}, \mathbf{x}_{t-2}) \quad (21)$$

移動方向に大きな変動がなければ $\cos \theta$ は 1 に近い値となる。本手法では、 $th = 0.9$ としたときの追跡点を用いる。選択された追跡点の動きを表現するために、移動方向に対応した色相から色を決定する。移動方向 d は 8 方向とし、座標 \mathbf{x} における方向 d の明度 $f_d(\mathbf{x})$ を以下の式より求める。

$$f_d(\mathbf{x}) = \sum_{t=1}^T \sum_{i=1}^N \delta(\mathbf{x} - \mathbf{x}_i^t, s_i) \quad (22)$$

T は総フレーム数、 N は追跡点数である。 \mathbf{x}_i^t はフレーム t における i 番目の追跡点の座標であり、Parzen Window Function である δ はガウス分布である。このときガウス分布の標準偏差として追跡点のスケール s_i を用いる。スケールは、テクスチャの少ない領域で大きくなる傾向がある。これは、テクスチャの少ない領域では、狭い範囲の画素から十分な特徴を得られないためである。テクスチャが少ない領域では、追跡点の密度が低くなるが、定数倍したスケールを用いてガウス分布を決定することで、密度の低さを補う。図 14 に示すように、追跡点の密度が高い領域は移動方向に対応した色が強く表現されることになる。また、様々な移動方向が同じ程度の頻度で発生する場所では、色が混ざり合い白く表現される。

図 15 に、あるフレームにおける追跡点を用いた、人のビジュアライゼーションの例を示す。図 15 (b) における追跡点の大きさはスケールを表している。

このように、移動体領域をセグメンテーションを必要とせず、特徴点とそのスケールを用いることで対象

物体の大まかな領域を表現することができる．

5.2 ビジュアライゼーション結果

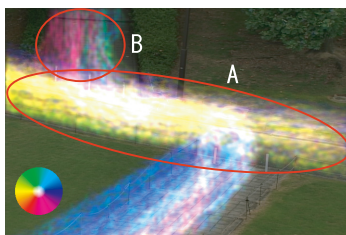
図 16 (a) に、約 1 時間 (100,000 フレーム) の追跡結果から動線をビジュアライゼーションした結果を示す．



図 15 スケールを考慮した表現

Fig. 15 Visualization by using scale information.

す．図 16 (a) おける左下の円は移動方向の色を示している．この例では、図中 A の領域では左に横切る人が多く、B の領域では移動方向が左右で反対となっていることが分かる．このように、長時間にわたる特徴点追跡の結果をビジュアライゼーションすることで、定常的な流れを確認することが可能となる．また、動線表示の対象時間を、約 2 秒間隔 (60 フレーム) とし、提案手法による追跡結果と、KLT 法による追跡結果を用いたビジュアライゼーション結果を図 16 (b), (c), (d) に示す．KLT 法における表示のためのスケール値は一定とした．提案手法では多数の点を長時間追跡可



(a) 約1時間の追跡結果



図 16 ビジュアライゼーション結果

Fig. 16 Visualization result.

能であり各点がスケールの値を持つため、KLT法に比べ、より対象物の形状を表現できていることが分かる。短時間における表示では、図16(b),(d)のAの領域とは異なる右方向への移動が確認できる。提案手法は、多数の特徴点の追跡結果を用いて可視化を行うため、いくつかの点が追跡に失敗し間違った色となる場合でも、全体としては正しい色付けを行うことが可能である。

5.3 ビジュアライゼーションの応用

提案手法は、ビジュアライゼーションに使用するフレーム数を変えることで異なる目的に使用することが可能である。長時間を使用した場合の応用として、マーケティングへの利用が考えられる。混雑した店舗内を対象として、自動的にマーケティングデータを取得するシステムを考える場合、人同士の密集や、柵により人の全身が写らないことなどから、人領域をセグメンテーションすることは困難である。提案手法は、特徴点追跡の結果のみから大まかな動きの流れを表現可能であるため、このような状況に適している。また、マーケティングは個人の細かな動きよりも、多くの人の動きの流れの方が重要であるため、長時間の結果を1枚の画像に表現することができる提案手法は有効である。また、短時間の結果を用いた場合の画像を連続で表示することで、プライバシーを保護しながら移動の様子を表示することが可能となる。

6. おわりに

SIFT特徴を用いたMean-Shiftによる特徴点追跡手法について提案した。提案手法はSIFT特徴の頑健性により、平行移動、回転に対して精度良く追跡できることを確認した。追跡対象が拡大する場合においても、SIFTのスケールを適応的に更新することで、追跡が可能である。また、人の追跡を例に、特徴点追跡結果のビジュアライゼーション法を示した。点の密度とスケールにより色の強度を変えることで、セグメンテーションを必要とせず、人の流れを表現可能であることを確認した。今後は、定常的な流れとは異なる動きの自動検出に取り組む予定である。

参考文献

- 1) Tomasi, C. and Kanade, T.: Shape and Motion from Image Streams under Orthography: A Factorization Method, *Int. Journal of Computer Vision*, Vol.9, No.2, pp.137-154 (1992).
- 2) Costeira, J.P. and Kanade, T.: A multibody factorization method for independently moving objects, *Int. Journal of Computer Vision*,

Vol.29, No.3, pp.159-179 (1998).

- 3) Tomasi, C. and Kanade, T.: Detection and tracking of point features, Technical report, CMU-CS-91-132 (1991).
- 4) Shi, J. and Tomasi, C.: Good Features to Track, *1994 IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, pp.593-600 (1994).
- 5) 深尾隆則, 金出武雄: 2段階特徴点追従アルゴリズム, 情報処理学会研究報告, 2003-CVIM-141, pp.103-110 (2003).
- 6) 深尾隆則, 金出武雄: オクルージョンに対してロバストな2段階特徴点追従アルゴリズム, 情報処理学会研究報告, 2003-CVIM-141, pp.111-118 (2003).
- 7) Lowe, D.G.: Distinctive image features from scale-invariant keypoints, *Int. Journal of Computer Vision*, Vol.60, No.2, pp.91-110 (2004).
- 8) Comaniciu, D. and Meer, P.: Mean shift analysis and applications, *IEEE Int'l Conf. Comp. Vis.*, pp.1197-1203 (1999).
- 9) Comaniciu, D., Ramesh, V. and Meer, P.: Real-time tracking of non-rigid objects using mean shift, *IEEE Computer Vision and Pattern Recognition*, pp.II:142-149 (2000).
- 10) 川中大樹, 中島 平, 岡谷貴之, 出口光一郎: 平均値シフト法を用いた複数物体追跡の研究, IPSJ SIG Technical Report CVIM, No.138(2) (2003).
- 11) 末岡千枝, 市川清人, 三田雄志, 堀 修: Mean Shift を用いたリアルタイム顔画像追跡手法に適した特徴量の提案, 画像の認識・理解シンポジウム (MIRU 2006) 論文集, pp.376-381 (2006).
- 12) She, K., Bebis, G., Gu, H. and Miller, R.: Vehicle Tracking Using On-Line Fusion of Color and Shape Features, *IEEE Conference on Intelligent Transportation Systems* (2004).
- 13) Pollefeys, M., Gool, L.Van, Vergauwen, M., Verbiest, F., Cornelis, K., Tops, J. and Koch, R.: Visual modeling with a hand-held camera, *International Journal of Computer Vision*, Vol.59, No.3, pp.207-232 (2004).
- 14) Brox, T., Bruhn, A., Papenberger, N. and Weickert, J.: High Accuracy Optical Flow Estimation Based on a Theory for Warping, *ECCV2004*.

(平成19年5月11日受付)

(平成19年11月21日採録)

(担当編集委員 椋木 雅之)



都築 勇司

2006 年中部大学工学部情報工学科卒業。同年より同修士課程に在籍。コンピュータビジョンに関する研究に従事。



藤吉 弘亘 (正会員)

1997 年中部大学大学院博士後期課程修了。1997~2000 年米カーネギーメロン大学ロボット工学研究所 Post-doctoral Fellow。2000 年中部大学講師を経て 2004 年より同大学准教授。2006 年米カーネギーメロン大学ロボット工学研究所客員研究員。工学博士。計算機視覚、動画像処理、パターン認識・理解の研究に従事。2005 年度ロボカップ研究賞。電気学会、IEEE 各会員。



金出 武雄 (正会員)

1974 年京都大学電子工学科博士課程修了(工博)。1976 年同助教授。1980 年カーネギーメロン大学計算機科学科・ロボット研究所高等研究員。1985 年同教授。1992~2001 年カーネギーメロン大学ロボティクス研究所所長。1993 年カーネギーメロン大学 U.A. and Helen Whitaker 記念教授。1998 年カーネギーメロン大学 U.A. and Helen Whitaker 記念全学教授。2001 年産業技術総合研究所デジタルヒューマン研究センターセンター長(非常勤)。計算機視覚、マルチメディア、センサー、医用ロボット、自律ヘリコプター等の研究に従事。National Academy of Engineering のメンバー、IEEE フェロー、ACM フェロー、American Association of Artificial Intelligence フェロー、C&C 賞、Joseph Engelberger 賞、JARA 賞等受賞。