

投票法とSIFTを用いたテンプレート更新型動物体追跡

○田久保 遠鴻（立命館大学） 平井 慎一（立命館大学）

1. 緒言

対象物体をロバストに認識・追跡ができるロボットビジョンシステムが近年必要とされている。例えば、複数の動物体から対象物体のみ追跡したい場合や劣悪な環境下での非剛体物体の追跡を行いたい場合が挙げられる。

しかし、水中下においては照明変化や水自体の濁りなどの環境的な外乱、追跡したい対象物体自体が変形することによって追跡し続けることが困難な場合が生じる。特徴点追跡を用いて物体追跡を行うことを考える場合、一般的にKLT(Kanade Lucas Tomasi)法[1]が挙げられる。このKLT法が物体の移動は平行移動し、輝度変化はないという条件を仮定しているため、照明変化など環境変化や3次元的な変形に弱いという欠点が存在する。そのため、実環境で特徴点追跡をするには困難な場合がある。

そこで本研究では、スケール変化・回転変化、および照明変化に頑強で、変形に対してもロバスト性を持つSIFT(Scale-Invariant Feature Transform)[2]を用いたテンプレート更新型物体トラッキングを行う。さらに、パーティクルフィルタを用いて1枚目のテンプレートを作成する。

2. アルゴリズムの構成

本報告で提案するアルゴリズムについて述べる。まず、SIFTを用いたテンプレート更新型マッチング処理を図1、パーティクルフィルタの処理を図2に示す。

まず、カメラから取得される入力画像に対してパーティクルフィルタによる処理を行う。現在、パーティクルフィルタによる動物体の判別は色相を用いている。色相の判別により集まったパーティクルの重心点を用いて、テンプレート画像として使用する部分を決定している。このとき、切り抜く画像の大きさはROI領域の大きさであり、適当に決めることができる。

パーティクルフィルタにより作成されたテンプレート画像を用いてSIFTマッチングを行う。連続した入力画像に対してマッチングを行う。その際、2枚目以降のテンプレート画像は1枚目との対応点から重心点を導出し、次のテンプレート画像をROI領域で取得するテンプレート更新型にしている。これは対象物体の3次元的な変形に対応した追跡するためである。このとき、背景・他物体なしの状態での実画像を用いたテンプレート固定時の追跡を図3、テンプレート更新時の追跡を図4に示す。このことから、3次元的な変形に対応した動物体の追跡を行うことができる。

なお、本システムの開発環境としてVisual C++ 2008 Express Edition, OpenCV ver.2.1[3]を用い、計算処理ライブラリとしてGSL(GNU Scientific Library) ver.1.8を使用している。また、SIFTは元のソースプログラム[4]を改良し使用している。

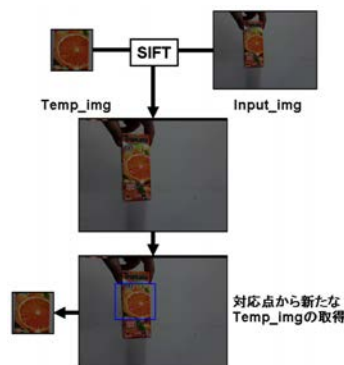


図1 SIFTを用いたテンプレート更新型マッチング処理

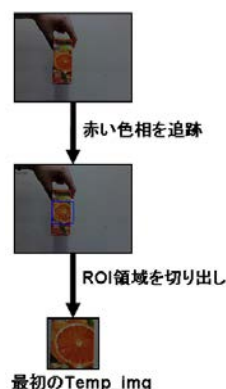


図2 パーティクルフィルタの処理

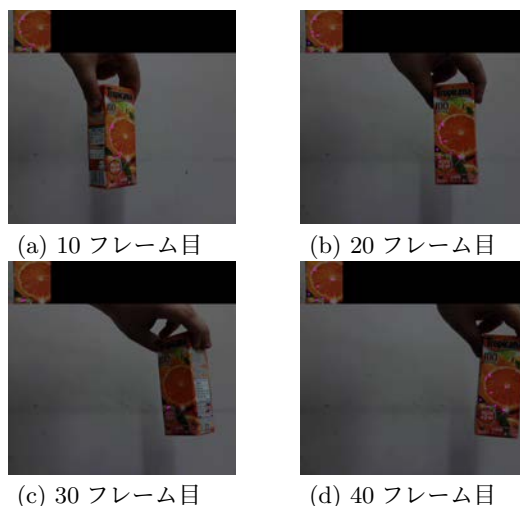
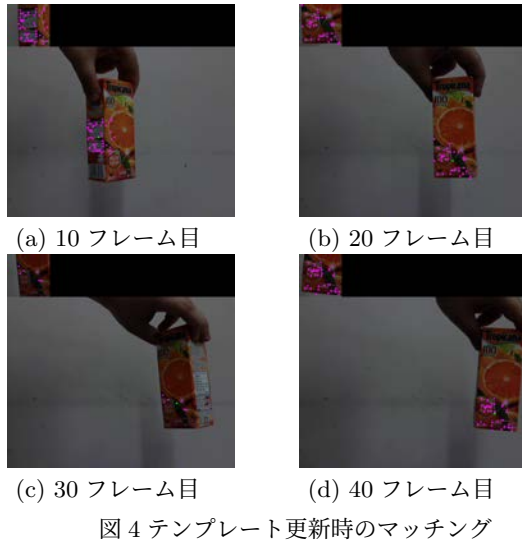


図3 テンプレート固定時のマッチング



3. 投票法を用いた重心点の導出

重心点をテンプレート切り出しの基準とするアルゴリズムでは 2 フレーム目以降に使用するテンプレートの切り出しを行う場合、対象物体以外の外れ点が考慮されていない。ROI 領域にテンプレート画像を切り抜くことは背景とのマッチングを減らすことができる。しかし、それは ROI 領域内に対象物体を正確に捉える必要がある。

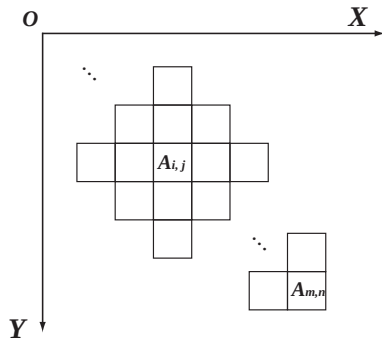
そこで本報告で投票法を提案する。投票法のモデルを図 5 に示す。まず、入力画像 A を $m \times n$ 個のブロックに分割する。このとき、 A は

$$A = \begin{bmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{bmatrix} \quad (1)$$

で示される。ブロック $A_{i,j}$ に属する対応点の個数を $P_{i,j}$ で表す。このとき、

$$p_{i^*,j^*} = \max[p_{i,j}] \quad (2)$$

となる A_{i^*,j^*} を抽出する。このブロックに属する対応点の座標の平均値を (x^*, y^*) とし、投票法で導出した対象物体の座標として扱う。



4. 実画像による外乱影響の比較

本実験では重心点と投票法について複雑背景下と複数物体下の追跡への影響を比較した。本実験で撮影に使用している Web カメラは共通して Logicoool HD Webcam C270 を使用する。また、対象物体はオレンジのジュースパックを用いる。テンプレート画像はジュースパックのみを切り出した画像を共通で使用する。また、入力画像の大きさは $640 \times 480[\text{pixel}]$ 、テンプレート画像切り出しの ROI 領域の大きさを $80 \times 80[\text{pixel}]$ とする。

4.1 複数物体下の影響

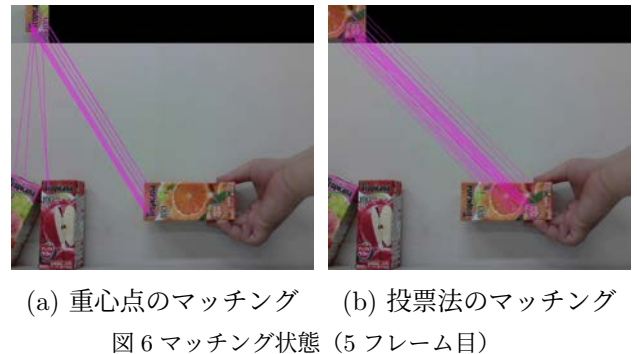
本実験では対象物体と同様のラベル部分を持つ物体を複数個置き、複数物体下での影響を確認した。重心点を用いたマッチングが失敗したときについて双方のマッチング状態を図 6 に示す。また、重心点を用いて切り出した結果を図 7 に、投票法を用いて切り出した結果を図 8 に示す。

重心点を用いた切り出しは 5 フレーム目までしかマッチングが行えず、テンプレート画像の切り出しに失敗している。これは、設置した対象物体と同様のラベルを持つ複数物体の部分にマッチングしたために外れ点に引きずられた結果である。単純にすべての対応点の平均だけでは対象物体を追跡することが困難である。一方、投票法を用いて切り出した結果は他の対象物体との対応点に引きずられることなく、対象物体をテンプレート画像内に捉えることができる。それは双方のマッチング状態を見ることでもいえる。

4.2 複雑背景下の影響

本実験では背景に新聞紙を用いて、複雑な背景下での影響を確認した。重心点を用いて切り出した結果を図 9 に、投票法を用いて切り出した結果を図 10 に示す。

双方とも対象物体であるジュースパックの切り出しに成功している。重心点を用いたの結果は対応点の平均を基準に切り出しを行っているが、今回は背景に外れ点に乗ることが少なかった。投票法を用いた結果のテンプレート画像はフレームごとに切り出しの基準となる位置がふらついているように見える。これは、切り抜かれたテンプレート画像との対応点のピークが立つ領域の中が逐次変化するためである。目的である対象物体の追従ができていることが確認できる。



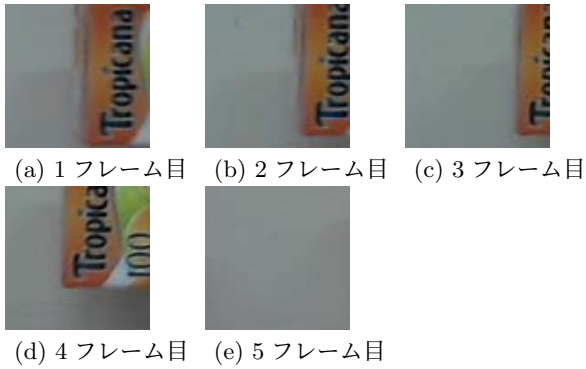


図7 重心点を用いて切り出した結果

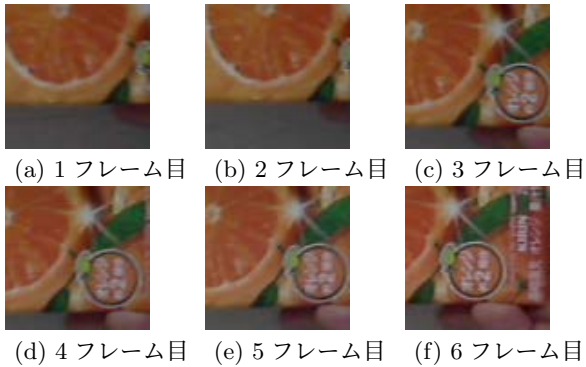


図8 投票法を用いて切り出した結果



図9 重心点を用いて切り出した結果

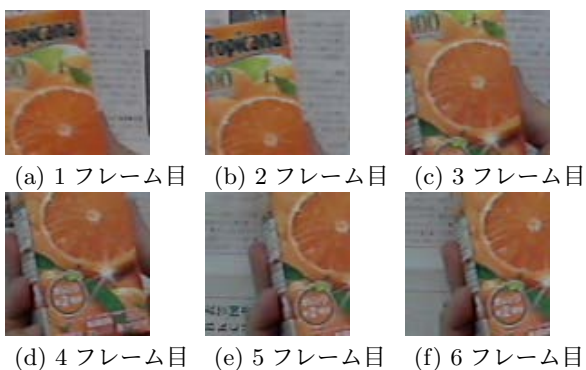


図10 投票法を用いて切り出した結果

5. 結言

今回の実験では投票法を用いた手法を提案し、実画像において重心点との比較を行った。比較の結果、投票法を用いた手法の方が、ROI 領域内に対象物体を捉えることができる。これは外れ点に引きずられることなく対象物体をテンプレート画像として切り出せているといえる。よって、投票法を用いた手法にすることでよりロバストな ROI 領域を定めることができ、対象物体を追跡できているといえる。これからは、パーティクルフィルタで移動量による対象物体の判別部分の改良や最終的な目的である水中下での動物体追跡の実験を行い、その影響を確認することで水中下での魚追跡の実現を目指す。

謝辞

本研究を進めるあたり、大変多くのご助力、ご教授を賜りました大日本スクリーン製造株式会社の皆様に厚くお礼申し上げます。

参考文献

- [1] Carlo Tomasi and Takeo Kanade, "Detection and tracking of point features", Technical report, CMU-CS-92-132, 1991
- [2] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", Journal of Computer Vision, 2004
- [3] "OpenCV2.1 リファレンスマニュアル", <http://opencv.jp/opencv-2.1/c/>
- [4] Rob Hess: "A C implementation of a SIFT image feature detector", <http://web.engr.oregonstate.edu/~hess/index.html>,