

(17) クレーン周辺の三次元地図高速作成のための特徴点の自動分類手法の開発

繁森 央一¹・須崎 純一²・小林 知生¹
米田 瑞生³・Marek Ososinski³

¹ 学生会員 京都大学大学院工学研究科 (〒615-8540 京都市西京区京都大学桂)

E-mail: {[@st.kyoto-u.ac.jp](mailto:shigemori.hisakazu.82x:kobayashi.tomoki.28s)}

² 正会員 京都大学大学院工学研究科 (〒615-8540 京都市西京区京都大学桂)

E-mail: susaki.junichi.3r@kyoto-u.ac.jp

³ 非会員 株式会社タダノ 技術研究所 (〒761-0301 香川県高松市林町 2217-13)

E-mail: {[@tadano.com](mailto:mizuki.yoneda;marek.ososinski)}

昨今では建設現場における人手不足が深刻化し、労働生産性の向上が課題となっている。その解決策の一つにクレーンの自動操縦化が挙げられており、実現のためにはクレーンブーム先端に取り付けた単眼カメラから取得した動画像を用いてクレーン周辺の三次元復元をする方法が有効であると考えられている。しかし、カメラから得られる画像には常にクレーンフックが映り込み、そのことが三次元復元の精度を著しく損なう要因となっている。そこで本研究ではクレーンフックなどの三次元復元精度を損なう要因となりうる対象物を画像上から除去することを目的に、画像上の物体を判別することを試みた。その際、三次元復元の前段階の処理であることを踏まえ、可能な限り計算コストが低くなるように画素全体ではなく特徴点ベースでの分類を試みた。

Key Words: photogrammetry, optical flow, classification, 3D mapping.

1. 背景と目的

昨今の建設現場では人手不足が問題となっており、労働生産性の向上が喫緊の課題となっている。そこでその解決策の一つとしてクレーン操縦の簡略化あるいは自動化が挙げられている。クレーンの自動操縦にはクレーンの吊荷の経路決定の際に必要なクレーン周辺の三次元地図の生成が必要とされている。現状では、クレーンブーム先端に鉛直下向きに取り付けられた単眼カメラから取得した動画像を用いてクレーン周辺の三次元地図を生成する手法が製造コストおよび実装上の観点から最も望ましいと考えられている。

三次元地図生成には静止物体から得られた特徴点を用いることで画像間のホモグラフィを計算する必要がある。

しかし、上記のように取り付けられたカメラから得られる画像には揺動するクレーンフックが常に映り続けるという特徴がある。そのためクレーンフックから取得された特徴点を用いてホモグラフィが計算されると、取得された三次元地図生成の精度が著しく損なわれてしまう。また、クレーンアームの伸び次第では画像中にアームと共に回転する操縦室が映るため、上記と同様の理由で三次元地図精度が損なわれる。そのため、静止物体とその

他の物体の特徴点を判別することが求められる。更に、仮にフックと操縦室のうち、フックの特徴点の判別ができれば、その特徴点をフックの位置の推定に用いることができる。したがって、本研究では動画像から静止物体、フック、操縦室の特徴点を判別することに注力する。

画像からある特定の物体を判別する手法には、Semantic Segmentation¹⁾のような、画素情報をもとに物体を判別する機械学習アルゴリズムが一般に用いられている。しかし、こういったアルゴリズムは計算コストが高く処理のフレームレートは 10 未満であることが多く、また学習にも数万フレームの学習データが必要とされる。本研究は三次元地図生成の前段階を対象としていること、そして時々刻々と変化する建設現場では三次元地図生成の高速化が求められていることを考慮し、限られた研究予算で目標を達成するために計算と実装コストの低いオプティカルフローを用いた物体の特徴点の分類を試みた。

オプティカルフローとは、カメラの隣接フレーム間での物体の動きを追跡する手法であり、画像から抽出した特徴点およびその周辺の画素情報をもとに特徴点を追跡する。この手法は画像の一部画素のみを扱うため、計算コストが低いという利点がある。代表的な手法には Lucas-Kanade 法²⁾があり、本研究ではこの手法を用いる。

2. 使用データと得られるオプティカルフローの特徴

(1) 使用データ

動画は、クレーン実機を用いて撮影したものをを用いる。撮影高度は約 10 m で、クレーンがアームを約 100° 旋回運動させる、解像度 1280×720, 30 fps, 約 34 秒の 1028 フレームの画像で構成される動画となっている。

(2) オプティカルフローの特徴

使用データから抽出される特徴点の対象物には、フック、静止物体、操縦室の三種類が存在する。また、各対象物のオプティカルフローにはそれぞれ異なった次の特徴がある。まずフックのオプティカルフローはフックが揺動するため、左右方向に細かく動く軌跡を描く。次に静止物体のオプティカルフローはクレーンの旋回に合わせて、画像下側を回転中心とした円弧の軌跡を描く。最後に操縦室のオプティカルフローは、操縦室がクレーン旋回と連動して動くため画像上では固定点として軌跡が描かれる。

上記の特徴に加え、オプティカルフロー自体の特徴として追跡対象物に他の物体が重なった場合、誤追跡が生じるなどの特徴もある。そのため今回の動画では揺動するフックが静止物体と重なることによる誤追跡や操縦室が静止物体と重なることによる誤追跡が観測されている。今回は前者をフック、後者を操縦室の分類に含めることとする。

以上より、本研究ではこれらの 3 種類の対象物から得られたオプティカルフローを以下の図-1 のように分類する手法を提案する。

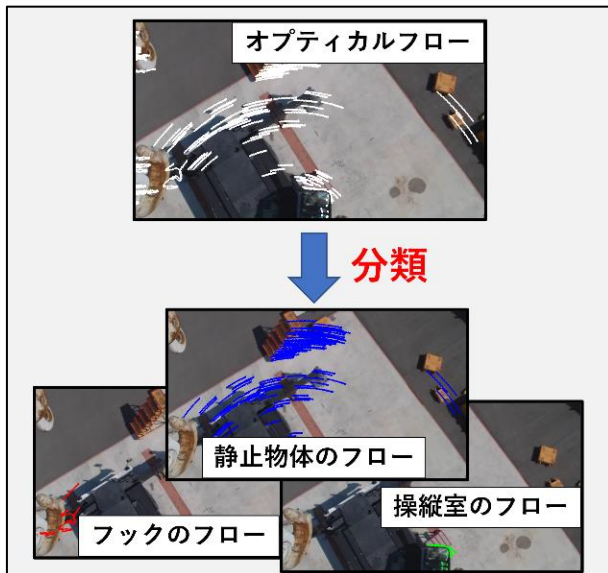


図-1 分類のイメージ

3. 提案手法

(1) 概要

本研究ではオプティカルフローを用いたフック、静止物体、操縦室の判別手法として、図-2 の分類フローを提案する

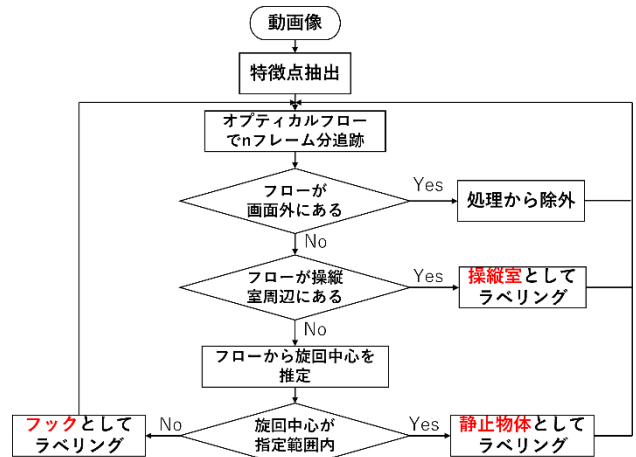


図-2 提案手法のフローチャート

以下ではそれぞれの過程について説明する。

(2) 特徴点抽出

画像中からオプティカルフローの初期点である特徴点を抽出する。

(3) オプティカルフロー n フレームの軌跡を取得

オプティカルフローの軌跡を利用して特徴点の分類を行うために、ある程度長い軌跡が情報として必要と考えた。そのため動画を n フレームごとに分けて分類を実行することで n フレーム分のオプティカルフローを分類に利用する。

(4) 画面外のフローを除外

特徴点の追跡時に、追跡対象物が画面外へ出るとフローも画面外へと出てしまう。この場合、ホモグラフィの計算にこの特徴点を使うことができないため、あらかじめ分類処理から除外する。

(5) 操縦室周辺のフローを分類

今回使用した動画には、画像の下端中央部にアームの旋回と連動して動く操縦室が常に映り込んでいる。しかし操縦室は画像内の特定の領域にしか存在しないため、Region of Interest (ROI)を画像下部に設定し、ROI において背景差分法を適用することで操縦室領域を抽出できると考える。ただし、今回は簡単のため操縦室領域を手動で定める。操縦室領域を手動で指定したのち、領域とフローが被っているものに関してこれを操縦室の特徴点としてラベリングする。

(6) フローから回転中心を推定

ここまでで未分類のフローにはフック、静止物体

それぞれから得られたものが含まれる。そこでこれらを分類するために各フローによって巡回中心を推定する。ここで巡回中心とは、クレーンの巡回運動の中心を指す。巡回中心は以下の手順で求める。

1. n フレーム分のフローを、 $\frac{n}{2}$ フレーム目を境に二等分する。
2. 分割したフローそれぞれの端点同士で線分を作り、その線分に対する垂直二等分線を引く。
3. 最後に、引いた二本の垂直二等分線の交点を求め、その点を巡回中心の推定値とする。

以上の手順を全てのフローに適用し、それぞれの巡回中心の推定位置を求める。

(7) 巡回中心の位置に基づく分類

(6)により推定した巡回中心は、静止物体のフローによって推定したものならば、画像の操縦室よりも下方の領域に位置し、フックによって推定したものならばその他の領域に位置するはずである。

その仮定に基づき、今回は手動で操縦室よりも下部に存在する巡回中心の領域を指定し、その領域内に中心が存在するものを静止物体、その他をフックとしてラベリングする。分類を終えると再び(3)に戻り、動画が終わるまで以上の過程を繰り返す。

4. 結果

(1) 結果

3章で提案した手法を CPU: Intel(R) Core(TM) i7-9700 3.0 GHz, OS: Windows10 Pro, メモリ: 32GB の環境のもと python3.9.5 と opencv4.5.3 で実装した。

そして実際に使用データに適用したところ、処理

速度は 204 fps となり、リアルタイムでの処理を実現した。その分類結果を表-1 に示す。更に、結果の例として 2 セット目の時の巡回中心位置を示した図-3 と 4 セット目の時の巡回中心位置を示した図-4 を示す。なお、今回は 3-(2)の特徴点抽出には Shi-To mashi 法³⁾が実装された *goodFeaturesToTrack* 関数を用い、そのパラメータを *maxCorners: 255, qualityLevel: 0.1, minDistance: 7, blockSize: 7* と設定した。また、3-(3)における n は 90 に設定した。そのため 1028 フレームの使用動画に対して、90 フレーム毎のオプティカルフローは合計で 11 セット取得された。そして同じく 3-(3)における特徴点の追跡手法には Lucas-Kanade 法²⁾が実装された *calcOpticalFlowPyrLK* 関数を用い、パラメータには *winSize: (15, 15), maxLevel: 4, criteria* の終了条件として回数を 10、精度を 0.03 と設定した。さらに、3-(5)における操縦室の領域を左上頂点座標:(500, 590)、右下頂点座標:(780, 720)とする矩形、3-(7)における巡回中心の領域を左上頂点座標:(400, 720)、右下頂点座標:(700, 920)とする矩形に設定した。

(2) 精度評価

精度評価には各クラスの予測結果の正解との適合率を示す *Precision*、各クラスの予測結果がどれくらい正解データを網羅できているかを示す *Recall*、そしてそれらの調和平均を示す *F1 Score* を用いた。

また、各クラスで得られるデータ数に差があることを考慮し、精度評価には多クラス分類における精度評価指標として *Accuracy* を用いた。また正解データには使用データから取得したオプティカルフローを手動で分類したものをを用いた。

表-1 提案手法によるオプティカルフローの 90 フレーム毎の分類結果

		予測 (1セット)			予測 (2セット)			予測 (3セット)			予測 (4セット)			予測 (5セット)			予測 (6セット)		
		静止物	フック	操縦室	静止物	フック	操縦室	静止物	フック	操縦室	静止物	フック	操縦室	静止物	フック	操縦室	静止物	フック	操縦室
正解	静止物	0	140	5	10	125	3	45	82	4	91	22	3	102	3	1	44	1	5
	フック	0	5	0	0	12	0	1	18	0	2	28	0	1	36	0	0	49	0
	操縦室	0	0	15	0	0	15	0	0	14	0	0	14	0	1	13	0	1	10
Precision		-	0.03	0.75	1.00	0.09	0.83	0.98	0.18	0.78	0.98	0.56	0.82	0.99	0.90	0.93	1.00	0.96	0.67
Recall		0.00	1.00	1.00	0.07	1.00	1.00	0.34	0.95	1.00	0.78	0.93	1.00	0.96	0.97	0.93	0.88	1.00	0.91
F1 Score		-	0.07	0.86	0.14	0.16	0.91	0.51	0.30	0.88	0.87	0.70	0.90	0.98	0.94	0.93	0.94	0.98	0.77
Accuracy		0.12			0.22			0.47			0.83			0.96			0.94		

		予測 (7セット)			予測 (8セット)			予測 (9セット)			予測 (10セット)			予測 (11セット)		
		静止物	フック	操縦室	静止物	フック	操縦室	静止物	フック	操縦室	静止物	フック	操縦室	静止物	フック	操縦室
正解	静止物	20	5	3	10	5	3	5	5	2	0	10	0	0	13	0
	フック	1	48	0	0	42	0	0	29	0	0	29	0	0	30	0
	操縦室	0	0	7	0	2	6	0	0	6	0	2	6	0	3	6
Precision		0.95	0.91	0.70	1.00	0.86	0.67	1.00	0.85	0.75	-	0.71	1.00	-	0.65	1.00
Recall		0.71	0.98	1.00	0.56	1.00	0.75	0.42	1.00	1.00	-	1.00	0.75	-	1.00	0.67
F1 Score		0.82	0.94	0.82	0.71	0.92	0.71	0.59	0.92	0.86	-	0.83	0.86	-	0.79	0.80
Accuracy		0.89			0.85			0.85			0.74			0.69		

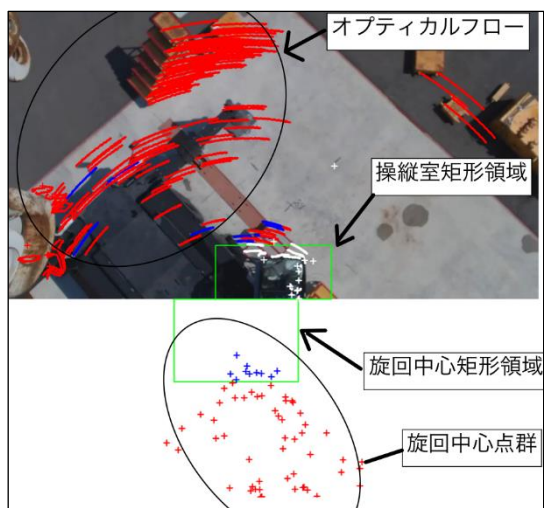


図-3 2セット目での旋回中心点群とフローの分類結果。
十字点が旋回中心，曲線がオプティカルフローを示す。それぞれ青色は静止物体，赤色はフック，白色は操縦室に分類されたことを意味する。

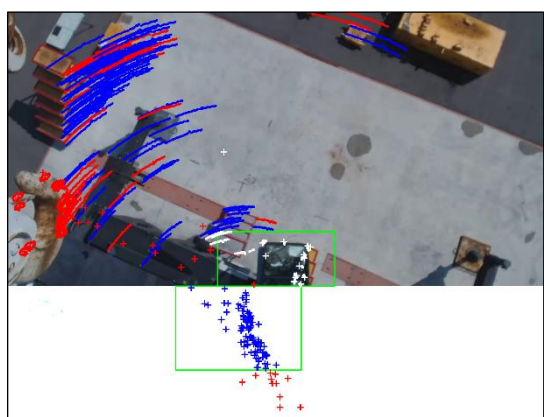


図-4 4セット目での旋回中心点群とフローの分類結果

5. 考察

4章の結果では，1, 2, 3, 10, 11セット目における *Accuracy* の値が 0.75 以下であったのに対し，4~9セット目における *Accuracy* の値は 0.8 以上となり，動画中盤の方が動画序盤と終盤よりも精度良く分類されていた。また，静止物体の *Precision* はどのセットにおいても 0.95 以上と高く，ホモグラフィ計算に適用が十分可能な精度が得られたと考える。

以下に 1, 2, 3 セットでの精度低下要因と 10, 11 セットでの精度低下要因について分けて考察する。

まず 1, 2, 3 セットについて考察する。図-3 と図-4 のように 1, 2, 3 セット目の旋回中心は 4 セット目よりも画像の下方に位置していた。そのため本来は静止物体に分類される点がフックに分類された。この結果は，動画序盤ではクレーンは加速運動をしているのに対し，動画中盤ではクレーンがほとんど等速で旋回しているというクレーンの運動状態の違いに要因があると考えられる。動画はクレーンの振動によって上下に細かく振動しており，得られるフ

ローも振動していた。つまり今回の場合，動画序盤の方が動画中盤よりも旋回が低速になるためオプティカルフローの軌跡に対する振動の影響がより顕著になったと考えられる。その結果今回使用した，3-(6)における旋回中心の推定手法の精度に影響が生じたことが原因と考えられる。しかしながら，クレーンの使用を考えると加減速は必ず考慮しなければならない運動である。そのため今回のような動画序盤におけるフローの分類精度を高める必要がある。

次に，10, 11 セットについても静止物体に分類されるべき特徴点がフックへ分類された。この理由として 10, 11 セットではクレーンが停止状態へと移行しており，カメラがほとんど静止していたため，推定された旋回中心が各特徴点の座標とほぼ一致し，指定領域内に含まれなかったからだと考えられる。

6. 結論

本研究では，建設現場における労働生産性の向上を目的としたクレーンの自動操縦化のためのクレーン周辺の三次元地図の生成の前段階の準備として，動画像中に映るフック，静止物体，操縦室のオプティカルフローでの分類手法について提案した。

結果として提案手法では 1) 静止物体以外の物体の特徴点が静止物体の特徴点へと誤分類されることがほとんどなかったこと，2) フックの分類に関してはクレーンの旋回速度が低い時に旋回中心の推定精度が低くなり静止物体がフックへと分類されてしまうこと，3) 停止時にはフローが固定点になるため円弧を利用した旋回中心の推定ができず，静止物体がフックに分類されてしまうこと，が明らかとなった。

今後の課題として，クレーンの動き出しや停止時のフローの分類方法を検討することが挙げられる。また，今回は手動で設定した操縦室や旋回中心の矩形領域に関しても，クレーンブームの伸長によって領域の設定を変える必要がある。そのため操縦室領域の抽出に対しては背景差分法を用い，旋回中心の抽出に対しては旋回中心のクラスタリングを行い，点数が最も多いクラスタをもとに旋回中心領域を決定するといった，動的に旋回中心領域を変更する手法を導入することも検討する必要がある。加えて，今回は画面外のフローを除外したまま処理を継続したが，画面外のフロー割合が大きくなった場合に特徴点の再抽出を行う条件を加えることで静止物体の特徴点数を増加させることも検討したい。

参考文献

- 1) Jonathan Long, Evan Shelhamer, Trevor Darrel, UC Berkeley: Fully Convolutional Networks for Semantic Segmentation, 2014
- 2) Bruce D. Lucas, Takeo Kanade: An Iterative Image Registration Technique with an Application to Stereo Vision. Proc. DARPA Image Understanding Workshop, pp.121-130, 1981
- 3) Jianbo Shi, Garlo Tomasi: Good Features to Track, CVPR94, 1994