

(23) 単眼カメラからのクレーン周辺の 三次元地図作成手法の開発

小林 知生¹・須崎 純一²・繁森 央一¹
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: {kobayashi.tomoki.28s; shigemori.hisakazu.82x}@st.kyoto-u.ac.jp

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: {mizuki.yoneda; marek.ososinski}@tadano.com

建設現場での人手不足が深刻化する昨今において、クレーン自動操縦などの技術が期待されているが、その実現のためにはクレーン周辺の三次元地図を精確かつ高速に作成することが必須とも言える。本研究では、ステレオマッチングや視差画像生成技術を利用し、クレーンの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基に三次元地図を生成することを試みた。その際、視差画像に対する中心投影から正射投影への変換や視差の閾値の自動設定、複数枚の視差画像の重ね合わせなどの処理を行うことで精確性向上を試みた。

Key Words: photogrammetry, stereo matching, disparity map, semi-global matching, 3D mapping.

1. 背景と目的

昨今の建設現場においては慢性的な人手不足が問題となっており、解決策としてクレーン操縦の補助あるいは完全自動操縦が注目されている。それらの技術を用いて吊り荷を運搬することを考えるとクレーン周辺の三次元地図を作成することが必須ともいえる。またその際、建設現場では状況が刻一刻と変化するため、高速性も重要となる。

高速性に長けた三次元復元手法には、Simultaneous Localization And Mapping (SLAM)¹⁾が知られる。代表的なSLAMアルゴリズムの一つであるORB-SLAM2²⁾やORB-SLAM3³⁾を用いて、車載カメラで撮影された動画像から高速で地図を作成する事例⁴⁾など幅広く報告されている。繁森ら⁵⁾は、クレーンのブーム先端に真下に向けて取り付けられた単眼カメラで撮影された動画像にORB-SLAM2を適用し、三次元地図の作成を試みた。既存研究でのカメラの位置は自由で不規則であり、また航空写真測量やドローン測量では画像が重なるように直線を往復して撮影が行われる。しかし、クレーン先端に取り付けられたカメラは同じ半径の円軌道上を動く。この独特な軌跡のためにORB-SLAM2での三次元地図作成には得

られる点数や安定性に限界があることが判明している。

そこで本研究では近いフレーム間で視差画像を作成するSemi-Global Matching (SGM)⁶⁾の手法に注目し、作成された視差画像同士を適切に連結することで三次元復元を行った。ここで、視差画像とは、ステレオ平行化された2枚の左右画像において、左画像上の各画素に対応する右画像上の画素を検出し(ステレオマッチング)、それらに対応する画素の左右方向の座標のずれ(視差)を基にして、物体とカメラとの距離を色の濃淡などで表した画像である。用いる画像の全ての画素に対して視差値が計算されるためSLAMなどを用いた手法に比べて密な点群が得られるというメリットがある。さらに、同じ高度から撮影された画像を基に生成された視差画像はお互いを連結しやすいため、今回の(地面にほぼ水平な)円軌道からの動画像に適していると考えた。なお株式会社タダノによると、クレーンの自動操縦に必要とされる三次元地図の精度は5~10cm、要求される処理速度は10分以内であるので、本研究ではこれを目標とした。

2. 使用データ

今回用いる動画は、シミュレータを用いて撮影した。地物の大きさやテクスチャ、配置などを任意に調節できるため、クレーン実機を用いて撮影された動画に比べて各種精度評価がしやすいと考えたためである。今回はシミュレータソフトの内、Gazebo⁷を用いた。なお今回は提案手法の有効性を確認するためにシミュレータを利用したが、今後は実画像を用いて更に実用性を試していく予定である。

3. 本手法の概要

(1) 概要

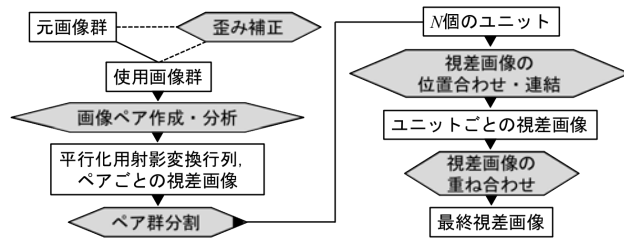


図-1 提案手法のフローチャート

本研究の提案手法では図-1の流れで処理を行った。

(2) 使用画像およびユニットの生成

まず、元動画画像から1秒に n フレームの割合でフレームを抽出し元画像群とする。今回、レンズに歪みは与え

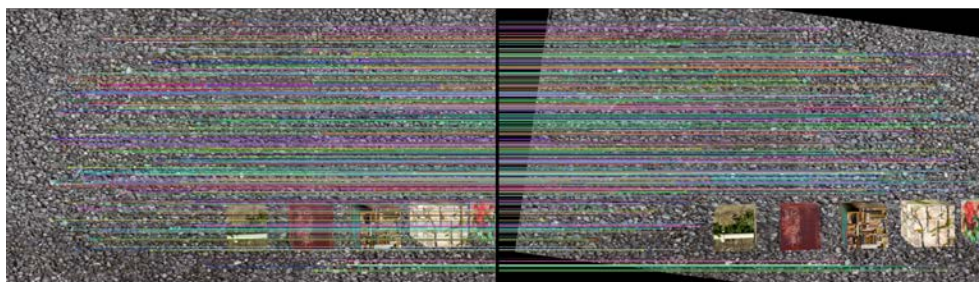


図-2 ステレオ平行化後のペア画像対応点描画の例

次に左右画像より視差画像を算出する。今回、視差画像を算出する手法としては、前述のSGMの上位互換であるSemi-Global Block Matching (SGBM)を用いた。SGBMは、隣り合う画素におけるノイズによる視差の不連続を軽減することができる手法である。

ここで、使用画像はある一点から撮影した中心投影の画像であり、それに伴って視差画像も中心投影の画像となっている。しかし、理想の最終視差画像は全ての地点を真上から見たような画像、すなわち

なかったため、元画像をそのまま使用画像群とした。

今回は、全画像から直接1枚の視差画像を生成するのではなく、まず全画像から複数枚の全体視差画像を生成し、それらを各地点の視差値について平均値を取りながら重ね合わせるという手法を用いた。複数枚の視差画像を重ね合わせる段階でノイズを除去できると考えたためである。本手法では、全画像を N 個の小さな画像群(ユニットと定義)に分割して考えた。具体的には、全使用画像の通し番号を n 、ユニットの通し番号を m としたとき、 $n \equiv m \pmod{N}$ となるように各ユニットに使用画像群を分配した。

(3) 画像ペアの作成・分析

視差画像を算出する際の入力画像である2枚の画像(時系列的に先の画像を左画像、後の画像を右画像とする)はお互いにステレオ平行化されている必要がある。ここで、ステレオ平行化とは、左画像上の任意の特徴点の y 座標が、それに対応する右画像上の特徴点の y 座標と等しくなるよう右画像を変換する処理のことである。この処理にはホモグラフィ行列による射影変換を利用した。今回扱う画像では、画像中での地面部分の占有面積が大きい。そのため、地面部分の対応点の座標を左右画像で一致させ、高さのある地物上の対応点には画像水平方向(x 軸方向)に視差を生じさせるようなホモグラフィ行列が計算される。元の左画像と、算出したホモグラフィ行列を用いて右画像を射影変換した画像について、改めて対応点を検出および描画した結果の例が図-2である。一連の処理によってペア画像がステレオ平行化されていることが分かる。

正射投影の画像である。よって、視差画像同士の連結前に、各視差画像を中心投影から正射投影に変換する必要がある。再投影時に視差値を画像中心に向けてずらす再投影幅 x は以下の式で求まる。

$$x = \frac{dl}{b' + d} \quad (1)$$

ただし x :再投影幅[px], l :画像主点からの距離[px], d :視差[px], b' :主点間距離[px]である。

以上の視差画像算出処理において、ペアごとに作成した視差画像の例、および各視差値を再投影した正射投影の視差画像が図-3、図-4である。

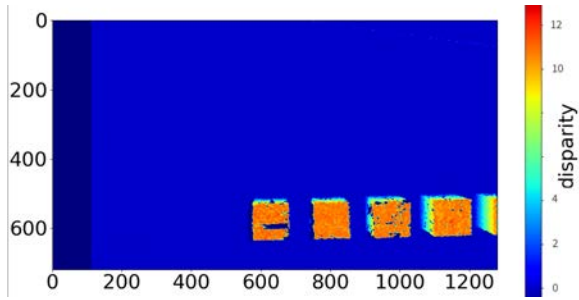


図-3 SGBMによる視差画像の例

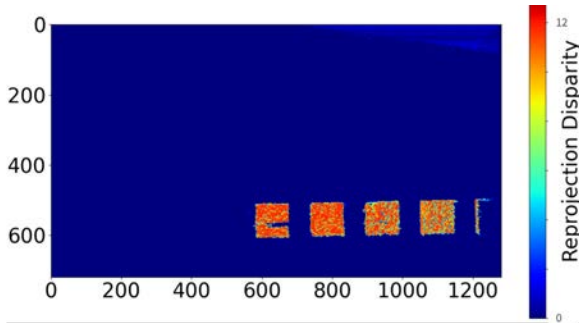


図-4 再投影した視差画像の例

(4) 視差画像の連結

各視差画像を連結するためには、それぞれの視差画像の位置合わせを行う必要がある。これにはペア画像の平行化時に算出したペア画像間のホモグラフィ行列を用いた。この変換行列によって高さが0の地点に関しては左右画像におけるy座標が同じで、かつ視差が0となるように変換されていることを考えれば、このホモグラフィ行列は視差画像同士の位置合わせに利用できる。これを利用してユニット内の全ての視差画像をそのユニットの先頭の視差画像に位置合わせした。

次に、位置合わせした視差画像を重ね合わせてユニットごとに1枚の視差画像を算出する。このとき視差画像同士が複数枚重なる地点の画素については複数の視差の候補値が存在することになるが、閾値を設けながら平均値を計算し、当該画素の視差値とした。

最後に、ユニットごとの視差画像をN枚重ね合わせることで最終視差画像を算出した。この処理により、ユニットごとの視差画像を算出した際に閾値と平均値によって取り除けなかったノイズを取り除く、あるいは軽減することができると考えた。

4. 結果

(1) 精度評価方法

本研究では作成された視差画像の精度を、図-5に示す箇所注目しながら、水平・鉛直方向に分けて定量的に評価した。なお、画像画素と実世界長の

縮尺は未知である。よって視差画像上での立方体地物の水平長さおよび視差と、それらに対応する真値(3 m)から縮尺 s_H および s_V を設定し、それを他の二つの地物に適用して計算した推定値と真値の誤差Eを算出した。

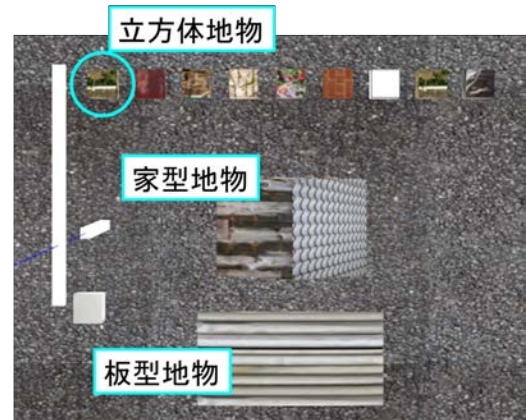


図-5 精度評価の注目箇所

(2) 結果

ここでは元動画画像1秒あたり5枚のフレームを使用し、それらを5つのユニットに分割した際の結果を示す。なお処理環境はCore i7, メモリ 32GB, Ubuntu 18.04.6, python 3.8.8, OpenCV 4.5.3である。提案手法にて算出した最終視差画像を図-6に示す。この視差画像の作成に要した時間は約379秒であった。また、最終視差画像の精度評価結果を表-1に示す。

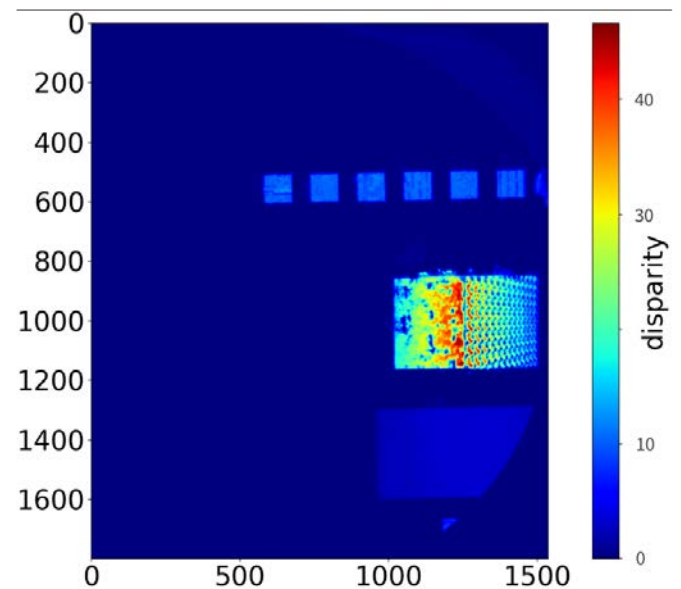


図-6 最終視差画像

表-1 各精度評価指標

水平方向	s_H (mm/px)	31.25	鉛直方向	s_V (mm/px)	284.2
	E_{2H} (mm)	31.3		E_{2V} (mm)	1485.0
	E_{3H} (mm)	592.7		E_{3V} (mm)	81.2
	E_{4H} (mm)	664.6			

最後に、最終視差画像を基にして三次元復元を行った三次元点群を図-7に示す。ただし地面からの距離が大きい点が明るい色になるように色を調節した。

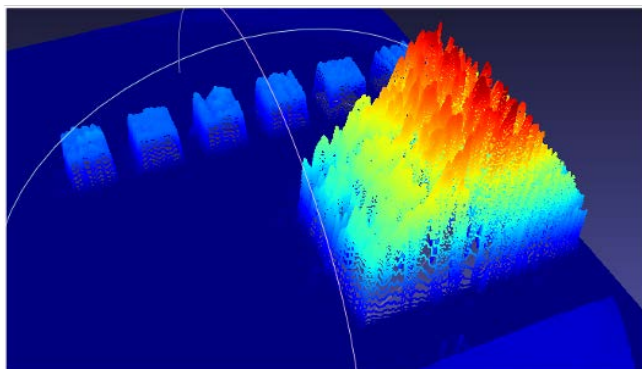


図-7 三次元復元の結果

5. 考察

まず処理時間については約 379 秒であった。画像ペアの解析に多くの時間を要しているため、使用する画像総数と全体の処理時間に大きな相関があると考えられる。この処理時間は許容範囲内ではあるものの、今後さらなる短縮を目指して検討を重ねる必要がある。

次に水平方向の精度については、家型地物と板型地物に関して目標の 5~10 cm を達成することはできなかった。三つの精度指標の数値から、最終視差画像のうち、ユニット後半部分（視差画像右側中央から右下部分にかけて）に近づくにつれてスケールが小さくなっていると考えられる。この原因としては、射影変換の繰り返し作用による誤差の蓄積が挙げられる。今回各視差画像を各ユニットの先頭の視差画像に位置合わせしたが、その際ユニット後半の視差画像は、前半部分の視差画像に比べると、射影変換の回数が多い。射影変換とは、そもそも長方形を台形に変形させるような変換であり、特徴点からのホモグラフィ行列の算出の際にわずかなスケールの変化が発生していると考えられ、そのスケール変化が蓄積し、結果としてユニット後半のスケールのずれに繋がったと考えられる。

最後に鉛直方向の精度については、板型地物では目標の 5~10 cm を達成したものの、家型地物では達成できなかった。家型地物での誤差の原因としては、地物のテクスチャ特性による同一平面上での視差値のばらつきが挙げられる。今回、家型地物の屋根部分のテクスチャについては、クレーンから遠い部分（視差画像右側）において瓦の画像を用いた規則的な模様を設定した。すると規則的な模様部分では、特徴量が大きく画像ペア間でマッチングしやすい部分とそうでない部分が存在し、結果として同一平面上であるにも関わらず、視差値についてもパターン模様状になってしまった。そのため、視差値が実際の地物の高さに対してばらついてしまったと考えら

れる。なお、今回は家型地物の屋根部分の最高地点の視差値を手動で計測しようとしたため、周辺のバラついた視差値のうち視差値の大きい部分を選択したために、精度指標も大きく計算されてしまったと考えられる。

6. 結論

本研究では、建設現場における人手不足や高齢化を背景とし、SGM を用いた視差画像作成および連結という手法でクレーン周辺環境の三次元地図の作成に取り組んだ。その際、シミュレータから得られた動画像を用いて手法の有効性を評価し、準リアルタイム（10 秒~10 分）での作成、および 5~10 cm の精度を目標とした。

本研究では、1)SGM の手法を用いることでクレーン周辺の視差画像の準リアルタイム作成が可能であること、2)地物のテクスチャや視差画像の位置合わせ精度が、作成される視差画像の水平および鉛直方向の精度に大きく影響を与えること、が分かった。

今後の課題として、今回は地物のテクスチャが原因で同一平面上であっても波打った視差値が計算される結果となったことを踏まえ、テクスチャに頑強なマッチング手法を探さる必要がある。また、今回ペア画像のステレオ平行化には射影変換を用いたが、ユニット後半の画像に対して繰り返し射影変換を行うことによってスケールの誤差が蓄積することを考慮すると、射影変換に換わる平行化の方法を模索する必要がある。今回はシミュレータ動画像を用いて提案手法の有効性を評価したが、今後は実画像に対しての適用も進めていく予定である。

参考文献

- 1) MathWorks: SLAM, <<https://jp.mathworks.com/discovery/slam.html>>, (入手 2022.6.6) .
- 2) RaulMur-Artal, Juan D.Tardos: ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras, IEEE Transactions on Robotics, 2017.
- 3) C.Campos, R.Elvia et al: ORB-SLAM3:An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM, IEEE Transactions and Robotics, 2021.
- 4) German Ros: Visual SLAM for Driverless Cars: A Brief Survey, IEEE Workshop on Navigation, Perception, Accurate Positioning and Mapping for Intelligent Vehicles, 2012.
- 5) 繁森央一ほか：ORB-SLAM2 を用いたクレーン周辺の三次元地図のリアルタイム作成，土木計画学研究会委員会，第 63 回土木計画学研究会発表会・講演集，2021.
- 6) Hirschmuller, H. : Stereo processing by semi-global matching and mutual information, IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2), pp.328-342, 2008.
- 7) Gazebo: 公式 HP, <<https://gazebosim.org/home>>, (入手 2022.6.6).