

(様式 5)

指導教員 承認印	
-------------	--

2019 年 3 月 25 日
Year Month Day

学位（博士）論文要旨

(Doctoral thesis abstract)

論文提出者 (Ph. D. candidate)	工学府博士後期課程 電子情報工学 専攻 (major) 平成 25 年度入学(Admission year) 学籍番号 13834302 氏名 北川 正理 印 (student ID No.) (Name) (Seal)
主指導教員氏名 (Name of supervisor)	清水 郁子
論文題目 (Title)	画像のマッチングのための計算量の小さい手法に関する研究 A Study on Methods for Fast and Compact Image Matching
論文要旨 (2000 字程度) (Abstract(400 words)) ※欧文・和文どちらでもよい。但し、和文の場合は英訳を付すこと。 (in English or in Japanese) 近年ではカメラを搭載するデバイスが多様化し、デバイスに搭載されているカメラからの画像を用いて物体の 3 次元形状を推定し、利用するアプリケーションが産業分野において開発されている。それらの中にはたとえば自動運転・運転者支援システムやスマートフォンやデジタルカメラで撮影した画像を用いた AR・VR に応用するアプリケーションが開発されており、これらで画像から物体の 3 次元形状情報を得る手法が活用されている。こうしたアプリケーションを実行するデバイスの中には計算リソースが限られ、拡張性も少ないデバイスやバックグラウンドで複数のプロセスを管理するようなデバイスが含まれており、3 次元形状を推定するプロセスを、計算量を抑えて行う必要がある場合がある。 画像から 3 次元形状を推定するためには、複数の異なる視点に設置したカメラから得た画像中の物体の同一の点を示す点を対応付けるいわゆる画像のマッチングが必要である。特に、例えば AR・VR への応用などでは、対象の密な 3 次元形状を得て 3 次元モデルを生成する必要がある。このような場合、まず特徴的な少数の点のみを画像間で対応づける。疎なマッチングでカメラのパラメータを推定し、次に画像のすべての点を画像間で対応づける密なマッチングを行い、3 次元モデルを得ることが行われる。 しかしながら先述したような計算リソースが限られた環境で使用することを想定した場合、現在提案されている手法では精度と計算量の小ささが十分に両立はされていない。例えば疎なマッチングに関して、従来手法では特徴点の輝度値の大小比較によるバイナリ記述を行う際、特徴として算出されるビット列の各ビットが独立しているために記述力が	

下がってしまうという問題や極めて局所的な範囲を記述しているために特徴点周りの輝度値のパターンの変化が乏しい領域を記述してしまうという問題があった。

一方で密なマッチングでは注目点周辺の輝度値に基づいた対象の3次元形状の変化の滑らかさを仮定してウィンドウベースのフィルタリングを行う手法が提案されているがフィルタサイズが適切でなく形状が過剰に平滑化される問題や空間計算量または時間計算量が多いフィルタリングや最適化手法を用いているという問題点がある。

したがって本研究では計算リソースの限られた環境において点の密な3次元モデルを生成するために必要である高精度で計算量の小さい疎なマッチングと密なマッチングの手法を提案する。

本論文は6章で構成される。第1章は先述した本論文の研究背景についてより詳しく述べる。

第2章では疎なマッチングと密なマッチングについて既存手法についての調査結果をまとめる。疎なマッチングについては既存手法で用いられている

輝度値の勾配を記述するアプローチと輝度値の大きさを記録するアプローチを大別し、現在知られている主な手法の特徴とそれらが持つ問題点を明らかとする。

また、同様に密なマッチングについて現在提案されているグローバルマッチング、セミグローバルマッチング、ローカルマッチングのうち計算量が小さい利点があるローカルマッチングのアプローチについて既存手法を調査する。特にローカルマッチングのアプローチのうちで高精度な Adaptive Support Weight を用いた手法の既存手法についてその特徴と問題点を明らかとする。

第3章では疎なマッチングについて本研究の提案手法について述べる。本手法は Difference of Gaussian によって算出されるスケールに対してバイアスをかけることで特徴点周りの広範囲を考慮して記述を行う。さらに特徴点周りの輝度値の大きさを関連付けて記述することで少ないビット数で特徴点周りの輝度値のパターンを頑健に記述する手法を実現する。

第4章では密なマッチングについて本研究の提案手法について述べる。本手法では ASW を用いたローカルマッチングのうちガイドドフィルタを用いた平滑化を用いるが、フィルタサイズマップと呼ばれるガイドドフィルタのウィンドウに2本以上のテクスチャ境界を含まないようなウィンドウサイズを画像の各画素に対してあらかじめ計算することで高精度なローカルマッチングを実現する。さらにそのフィルタサイズマップをメディアンフィルタと組み合わせ、視差の再割り当て時のフィルタとして用いることで従来時間計算量の大きかった重み付きメディアンフィルタと同等の結果を得ながら高速に演算を行う。さらに逐次的なコストの最適化を行うことで空間計算量を抑えたローカルマッチングを実現する。

第5章では Middlebury のデータセットを用いた実験を行い、疎なマッチングと密なマッチングについて提案手法と従来手法の違いと提案手法の特徴を考察する。疎なマッチングに関してはビット数とマッチング精度を比較することで提案手法の従来手法に対する有効性を示す。さらにビットの記述法の違いによる精度の違いを示し、提案手法の特徴を示す。密なマッチングでは他手法との精度の違いをもとに提案手法の有効性を示す。加えて計算時間および空間計算量のプロファイリングを行い、提案手法の特徴の考察を行う。

第6章ではそれまでの内容を受け、本研究をまとめるとともに今後の課題を示し、本論文の結言を述べる。

(英訳) ※和文要旨の場合(400 words)

In recent years, a lot of industrial applications using 3D object shape estimated from images captured by camera on various devices are developing. To estimate the 3D shape of the object, technique for image matching which corresponds the same object points between images is essential.

To apply matching techniques for various devices, it is important to cohabit both fast and compact calculation and high accurate because resource for computation are limited or there are many background process is running in parallel.

For 3D shape estimation from multiple images, in usual, first the method for coarse matching is applied which corresponds only feature points and estimates relative position and orientation of each camera.

Then, the dense matching is applied which corresponds all points in each image using the cameras' positions and orientations to obtain 3D model of the object. To generalize the application of 3D shape estimation using cameras on the various types of devices, the methods for both coarse and dense matching needs to be more fast and compact which has enough high accuracy.

In this paper, we propose a sparse matching method and a dense matching method which is more compact and faster than the conventional methods with the same accuracy.

In our sparse matching, binary feature are computed by using wider region around feature point than the conventional method which is invariant to rotation or scale changing. After that, binary feature vector is computed by comparing the pixel value by using all combinations of the selected points.

Our method achieves higher accuracy than the conventional methods for the same memory size because our description contains the relations between the selected points are related each other by using all combinations of the points.

In our dense matching, we introduce a filter size map of each image which has optimal window sizes of the Guided Filter based on the response of Difference of the Gaussian for each pixel for the Guided Filter based stereo matching method. By using the filter size map for the Guided Filter, the matching cost is smoothed in a region which is not including two or more texture boundaries. In addition, by optimizing matching cost sequentially, it is possible to perform calculations with reduced using memory. In the refinement step of the stereo matching, we combine a filter size map and median filter instead of weighted median filter which is used conventional method.