ステレオと単眼を用いた屋外動的低特徴環境における 高精度ビジュアルオドメトリシステム*

Visual Odometry System Using Stereo and Monocular Camera in Outdoor Dynamic Untextured Environment

○ 石井朗 (明治大)*¹ 学 坂井敦 (明治大)*²

学 三橋雅仁 (明治大)*³ 正 黒田洋司 (明治大)*⁴

Akira ISHII, Atsushi SAKAI, Masahito MITSUHASHI and Yoji KURODA

Meiji University, Department of Mechanical Engineering

In this research, we propose a state-of-the-art 6-Degrees-Of-Freedom(6DOF) Visual Odometry(VO) system which fused stereo VO with effective feature sampling, VO using normal vector information of a ground plane and monocular VO. Stereo VO is used to do 6DOF motion estimation in untextured dynamic environments. Next, VO using ground plane information is used to improve the accuracy of roll and pitch angle estimation. Furthermore, monocular VO is introduced to solve the problems in stereo processing. Finally, Unscented Kalman Filter is adopted to fuse these VO to accomplish robust and accurate localization. We present and evaluate experimental results for our system over 730 meters run in a challenging outdoor environment and compare it with ground truth. Proposed VO system can localize the robot's position within 3% error in untextured outdoor environment.

Key Words : Visual Odometry, Robot Vision, Localization

1. 緒 言

自律移動ロボットにとって、正確な位置・姿勢を推定する技術は必要不可欠である。本研究では、その得られる情報の多種、多様性から、ロボットの可用性を著しく向上できると期待されるビジョンセンサを使用する。我々の最終的な目標は、屋外のあらゆる環境に対応した 6DOF の VO システムを実現することである。そのための様々な問題のうち、屋外動的低特徴環境における問題について考える。低特徴環境とは、濃度変化が少ない舗装された道路などを含む環境の事を指す。このような環境は VO にとってチャレンジングな問題であり、もしこのような環境で VO が実現できれば、より多くの屋外環境に対応できると我々は信じている。

屋外動的低特徴環境における問題は、VO 自体の問題と低特徴環境における問題とに分ける事ができる。VO 自体の問題として、特徴点検出と追跡の安定性、蓄積誤差がある。それぞれの問題を解決するために様々な手法が提案されているが、屋外動的低特徴環境における長距離自己位置推定に対する有効性はあまり証明

Fig. 1 The problems in untextured environments

されていない。

次に低特徴環境における問題について考える。このような低特徴環境で VO を実現するには前稿⁽¹⁾で述べたように多くの問題がある。図 1 に示すように、遠方や移動物体などの特徴的物体に対して特徴点が集中してしまうことである。

このように、屋外動的低特徴環境では様々な問題が発生する。そこで、前稿⁽¹⁾では、効率的な特徴点サンプリング手法(図 2)を用いた VO システム (StereoVO) を提案した。この手法によって、動的低特徴環境においても、高精度な自己位置推定を実現した。その後ロールやピッチ角の誤差を補正するために、道路平面の法線ベクトル情報を用いた VO(PlaneVO) と StereoVO を統合したシステムを提案した⁽²⁾。この手法によって、

*1*2*3*4 明治大学理工学部機械工学科
Email: akiraishii11@gmail.com*¹

Fig. 2 Effective feature sampling

Fig. 3 Over exposure

ロール・ピッチ角の推定精度を向上する事ができ、結果ロボストで精度の高い自己位置推定を実現した。しかし、上記二つの手法を用いてもなお、問題が残る。それはステレオカメラ固有の問題であり、例えばオクルージョン、白飛びなど光の影響(図3)、両画像の明るさの違いなど外乱によるロボスト性が低いことである。そこで、本論文では、今までの手法に加えて単眼 VO (MonoVO) を導入する。すなわち、

1. StereoVO: 低特徴環境で不定性のない6自由度運動推定を行う。
2. PlaneVO: 点情報のみでなくロボットから見える平面の法線ベクトルの変化からロールとピッチ角の推定精度の改善。
3. MonoVO: 外乱に対するロボスト性の向上。
4. Sensor Fusion using Unscented Kalman Filter: 上記三つの手法から得られる情報を統合する。の4つを組み合わせた VO システムを提案する。

本論文では、このシステムの有効性を示すために、屋外動的な低特徴環境でロボットを数百メートル走行させ、その軌跡と Ground Truth と比較し精度を検証した。

2. ステレオビジュアルオドメトリ

StereoVO は、ステレオ画像列を解析する事でカメラの(ロボット自体の)移動量と姿勢を推定するプロセスである。しかし、屋外環境で StereoVO を行うにはいくつかの問題がある。低特徴環境で特徴点を抽出すると、限られた特徴的物体に集中することや、画像内の手前などに偏ることがよくあり、これは推定結果に致命的なエラーをもたらす。これに対して、画像を領域分割し、各領域からの特徴点を選ぶことで低特徴だけでなく動的な環境においても StereoVO にとって効果的な特徴点サンプリングを可能とした。他にも StereoVO 自身の問題として特徴点のミスマッチングやミストラッキングがある。これらのアウトライアの影響を減らし、ロボストな運動推定を可能とするために Three-point algorithm と RANSAC の枠組みを用いる。最後に、逐次的に移動量を計算する StereoVO の蓄積誤差を軽減するためにキーフレーム調整を導入する。

(a) Feature tracking in MonoVO. (b) An example of the comparison result of yaw estimation error

Fig. 4 Results of feature tracking and VO in MonoVO

3. 路面情報を用いたロール・ピッチ角の推定

前節で述べた StereoVO では、回転と並進を含む運動をした場合、移動量を算出する際に、回転と並進を高精度に分解する事が難しく、局所解に陥りやすい。特に、市街地や遊歩道などでロボットを動かす事を考えた場合、その走行可能領域はおおよそ平面であり、高さ情報が少ないため、高さ方向に誤差が生じやすい。結果として、ロールやピッチ角に誤差が生じる。そこで本研究では、路面領域の法線ベクトルの変化からロール・ピッチ角を推定する事でこの問題を解消する。特徴点の三次元点情報だけでなく面情報を用いることで点情報に含まれる誤差が軽減できるというメリットがある。さらに、法線ベクトルの相対的な変化からロール・ピッチ角を推定しているため、カメラの設置誤差の影響も軽減できる。

本研究では、道路平面の法線ベクトルを抽出するために、ステレオ画像を用いて、ステレオ間のホモグラフィ行列を算出することで推定する。このホモグラフィ行列は平面領域から最低4組の対応点が得られれば計算する事が出来る。すなわち道路平面領域から特徴点が抽出されなければならない。そこで本研究では、アウトライアの影響をなくし高精度にホモグラフィ行列を推定するために、StereoVO で述べた特徴点サンプリング方法と投影誤差を最小化する評価関数を用いた Four-point algorithm と RANSAC を用いる。時系列のステレオ画像から推定したホモグラフィ行列と二つのカメラの内部パラメータ行列を用いて、毎フレーム特異値分解によって法線ベクトルを算出し、差分を取ることでロールとピッチ角の変化量を得る。

4. 単眼ビジュアルオドメトリ

これまで述べた StereoVO と PlaneVO は両者ともステレオ画像を用いる必要があるため、ステレオカメラ固有の問題である、オクルージョンや左右の画像での明るさの違いなどの問題が生じる。このような現象が多く続くと VO による計算は困難になる。特にロボットが回転運動をしている時にこのような現象が生じる

と、致命的な誤差が発生する。そこで、この問題を解消するために本研究では、単眼カメラによる角度推定を導入する。

本研究では、単眼カメラから角度推定を行うために、エピポーラ幾何によって理論を導くことができる Fundamental 行列 (厳密には Essential 行列) を用いる。F 行列は、二視点から得られた二枚の画像間において対応点が最低 8 点得ることができれば、線形的に計算することが可能である。この F 行列は、カメラの内部パラメータとカメラ間の運動のみを含む行列であるという性質を持つ。したがって、これを分解することで、二視点間の運動を抽出する。

しかし、F 行列の推定は一般的には安定しないと報告されている。そこで本研究では、次の三つの手法を取り入れて F 行列の推定を高精度に行う。

- 分解能 0.1 ピクセルでの特徴点の検出と追跡。
- 先に述べたホモグラフィ行列との区別をつけ、かつ角度推定の誤差を少なくするために、カメラから遠い領域からの特徴点抽出。
- 物理的に意味のある量の最適化による Eight-point algorithm + RANSAC を用いた外れ値の除去。

推定した F 行列とカメラ内部行列から E 行列を算出する。この E 行列は Rank2 の行列でカメラ間の運動パラメータのみの情報を含む行列である。E 行列を特異値分解して回転と並進に分解し運動パラメータを得る。図 4 に MonoVO における特徴点追跡の様子と StereoVO と MonoVO によるヨー角の推定誤差の比較結果の一例を示す。図 4(b) より、StereoVO より MonoVO の方が高精度な推定が行えており MonoVO の有効性がわかる。

5. UKF によるデータ統合

本章では、StereoVO から得られる状態量と、PlaneVO から得られるロール、ピッチ角の変化情報、さらには MonoVO から得られるヨー角の変化情報を統合する手法について述べる。本研究では、入力値 (StereoVO) と観測値 (PlaneVO, MonoVO) を Unscented Kalman Filter (UKF) を利用して確率的に推定することによって、高精度で安定した自己位置推定を実現する⁽³⁾。このようなデータ統合の手法としては Extended Kalman Filter (EKF) と Particle Filter (PF) が一般的な手法であるが、UKF はこれらの手法の欠点を解決する利点を有している。EKF とは異なり、非線形なモデルを使用した場合でも、シグマ点という確率分布の代表点を利用して確率計算を行うため、高精度な推定が可能である。本論文のように 6DOF の自己

Fig. 5 Experimental condition

位置推定をする場合、モデルの非線形性が高くなるため、これは大きな利点であるまた PF とは異なり、確率計算を直接行うため、計算コストが低いという利点もある。

しかし、UKF を利用してデータの統合を行うだけでは高精度な自己推定を行うことは難しい。例えば、歩行者などの動的物体や複雑な形状の物体が観測範囲内に存在している時、推定される路面情報には局所的な大きなノイズが生じてしまう。この情報をそのまま利用した場合、自己位置推定の結果に大きな誤差が生じてしまう。

したがって、本研究では推定結果の局所的なノイズの影響を減少させるために Normalized Innovation Squared (NIS) test の手法を取り入れる⁽⁴⁾。観測値の差のマハラノビス距離を計算し、その結果が閾値以下の場合はその観測値を採用し、閾値以上の時は前の観測値を使用する。この手法を用いることにより、動的・低特徴屋外環境においても、高精度な自己位置推定を実現する。

6. 実 験

提案する VO システムの有効性を検証するために、本研究の所有する屋外自律移動ロボット (図 5(a)) から取得した画像データセットをもとに実験を行った。カメラは Poinr Grey Research 社製の Flea2、レンズは興和社製の広角低ひずみレンズを使用し、ベースラインは約 33cm である。ステレオカメラの距離精度は約 5m 離れた場所を観測したとき、 $\pm 10\text{cm}$ (誤差率約 2%) である。カメラと PC 間のインターフェースは IEEE1394 で、画像は VGA で取得した。PC は Intel Core2Duo 2.00GHz、画像処理ライブラリとして OpenCV を用いた。ロボットは約 0.3m/s で走行し、我々のシステムの実行周期はおよそ 3Hz である。また、Ground Truth として、DGPS(A-100: Hemisphere 社製) と IMU (NAV420CA-100: Crossbow 社製) を用いた。DGPS の観測値の精度は約 50cm (95%)、IMU の推定精度は 3.0deg 以内で

Fig. 6 Estimated 6DOF motion of experiment

Fig. 7 Comparing estimation error of roll, pitch and yaw angle

ある。実験は明治大学生田校舎において、図 5(b) に示すようにロボットを約 730m 走行させて行った。図に示されるような環境は、道路が一様な模様である低特徴、かつ人や車などが多く存在する動的環境である。

図 6 に 6 自由度の自己位置推定結果とそれを X-Y 平面で見た結果を示す。青い線が DGPS、緑色の線が StereoVO のみ、赤い線が本研究で提案する VO システムの軌跡をそれぞれ示している。図を見てわかるように、StereoVO による推定精度は著しく悪い。反対に我々の提案する VO システムの場合は最終的な位置誤差が多少あるものの、自己位置推定の結果が著しく改善されていることがわかる。特に、図 6(b) を見てわかるように Z 軸方向の誤差が提案する VO システムの場合極端に減少している。また、図 6(b) においても StereoVO の軌跡に比べて提案する VO システムの軌跡は GPS の軌跡に近くなっている事がわかる。総移動量に対する最終的な位置誤差率は、ground truth がわかる X-Y 平面で考えた場合、StereoVO が約 10% であるのに対して、提案する手法は約 2.6% と改善された。

次に、図 7 に各時刻における推定された角度の誤差を示す。ここでも、StereoVO のみ（緑）と提案する VO システム（赤）とで比較を行った。ロールとピッチ角においては、すべての時刻において誤差が StereoVO より下回っている。

以上の実験結果から、我々の提案する VO システムの有効性が示された。

7. 結言

本研究では、カメラを用いた屋外動的な低特徴環境における高精度な長距離自己位置推定システムを提案した。StereoVO とその欠点を補う PlaneVO, MonoVO により推定された 6 自由度の移動量を UKF と NISTest を導入し統合することで、高精度な 6 自由自己位置推定を達成した。その有効性は実験により確認された。

文 献

- (1) Y. Tamura, M. Suzuki, A. Ishii, Y. Kuroda, "Visual Odometry with Effective Feature Sampling for Untextured Outdoor Environment", *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct, 10-15, 2009.
- (2) Y. Tamura, A. Sakai and Y. Kuroda, "Visual Odometry using Ground Plane in Dynamic Untextured Environment", *Frontiers of Computer Vision, Proc. of FCV2010*, pp.159-164, 2010.2.
- (3) R. van der Merwe, "Sigma-Point Kalman Filters for Probabilistic Inference in Dynamic State-Space Models", *PhD thesis, OGI School of Science & Engineering at Oregon Health & Science University*, April 2004.
- (4) S. Sukkarieh, E. Nebot, and H. Durrant-Whyte, "A High Integrity IMU/GPS Navigation Loop for Autonomous Land Vehicle applications", *IEEE Trans. Robot. Automat.*, vol. 15, pp. 572-578, June, 1999.