

# カメラとジャイロセンサを組み合わせた高精度な姿勢推定

小堀訓成 † , 出口大輔 † , 高橋友和 ‡ , 井手一郎 † , 村瀬洋 †

Norimasa KOBORI † Daisuke DEGUCHI † Tomokazu TAKAHASHI ‡  
Ichiro IDE † and Hiroshi MURASE ‡

†: 名古屋大学大学院大学院 情報科学研究科, {kobo, ddeguchi, ide, murase}@murase.m.is.nagoya-u.ac.jp

‡: 岐阜聖徳学園大学 経済情報学部, ttakahashi@gifu.shotoku.ac.jp

移動ロボットや自動車への実用化に向けた高精度な姿勢推定手法として、安価な単眼カメラと汎用のジャイロセンサを一体化したハイブリッドなシステムを提案する。ジャイロセンサを用いて姿勢推定を行う場合、温度などの影響によりドリフト誤差が発生するため、時間経過に対して精度が低下する。一方、カメラ画像を用いた姿勢推定手法では、ドリフト誤差が発生しない反面、オクルージョンにより継続的な姿勢推定は難しい。提案手法は、ジャイロセンサとカメラ画像をカルマンフィルタの枠組みで組み合わせることにより、上記センサの各々が持つ問題の解決を図るものである。具体的には、SIFT 特徴量を利用し、モーションステレオによって回転行列・並進ベクトルを算出し、その情報を参照値としてジャイロセンサの誤差を推定する。提案手法により姿勢の推定精度が向上することを実験により示す。

<キーワード> SIFT, モーションステレオ, エピポーラ幾何, カルマンフィルタ, ジャイロセンサ

## 1. はじめに

自動車やロボットなどの移動体において自らの姿勢を知ることは重要である。姿勢の推定にはジャイロセンサを使うのが一般的である。特に近年 MEMS (Micro Electro Mechanical Systems) タイプの振動型ジャイロセンサ[1]が出てきており、小型かつ安価になった。その一方で、課題も残存する。基板温度などの影響によってドリフト誤差は発生し、時間経過に対して誤差が蓄積してしまう。そこで本論文ではカメラ画像を利用して、ジャイロセンサの誤差を除くことを試みる。具体的には、画像から求めた回転行列・並進ベクトルを参照値として、ジャイロセンサの誤差をカルマンフィルタの枠組みで推定する。これによりジャイロセンサで求めた姿勢の蓄積誤差低減を図る。画像から回転行列・並進ベクトルを算出する際には、SIFT 特徴量[2]とモーションステレオを利用する。SIFT 特徴量は、画像の回転・拡大・縮小・照明の変化に頑健である。また微小なアフィン変換に対しては、ロバストであるため利用する。

提案するシステムでは、ジャイロセンサとカメラ画像が相互補完的な役割を担っている。その関係を図 1 に示す。ジャイロセンサは精度があまり高くないが、必ず値は出てくるためロバスト性は高く、応答性も高い。一方、画像処理は、特徴点の対応づけが正確であれば高精度であるが、オクルージョンによって姿勢が算出されない場合がある。そのためロバスト性は低い。さらに計算コストが高く、1Hz 程度の処理しかできない。よって応答性は低い。そこで、互いの欠点を双方の利点で補う手法を実現する。

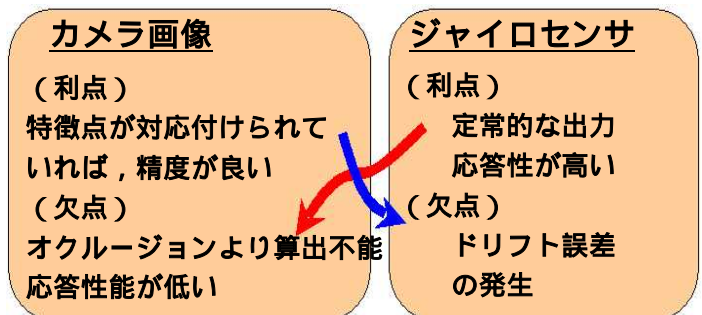


図 1 カメラとジャイロセンサの融合による相互補完

## 2. 関連研究

ジャイロセンサとカメラ画像をカルマンフィルタで組み合わせた先行研究はいくつか存在する[3][5][6][7]。

Min-geun Song らは、時刻の異なる画像間でエッジ点の対応を行い、角速度（観測量）を求め、ジャイロセンサから求まる角速度（状態量）とカルマンフィルタによって融合させた[3]。ジャイロセンサとカメラはサンプリング周期が異なり、さらに画像処理には時間がかかるため、時間的遡及を考慮したカルマンフィルタの構築を主に議論している。しかし、リアルタイム処理はシステム設計の問題であり、負荷の高い画像処理は専用のハードウェアで行うなど方法は多々ある。むしろ、画像間の対応が取れない場合の処理や画像間の対応からどの程度まで姿勢変化の推定精度（解像度）が取れるかを検討する方が実用面では重要である。しかし、文献[3][5][6][7]をはじめ、先行研究にそれらの議論を見ない。その点、本提案手法は、軍事/航空機の高精度ポジショニングで利用されているGPS/INS[4]の手法を参考に、“ジャイロセンサの誤差”を状態量としたカルマンフィルタを組むことで、断片的に画像間の対応が取れない場合にも対応する。また画像から求める姿勢変化は、ジャイロセンサを利用して基礎行列を導出するため、1制御周期あたり0.1度程度までの解像度で検出できる。文献[3]では1制御周期あたり5度程度の変化でしか実験が行われていない。

またSIFT特徴量を利用した点も関連研究と比較する。S.You らはAR（拡張現実）システムにおいて、画像で検出容易なパターンをあらかじめ用意しておき、パターンの見え方によって姿勢を算出、カルマンフィルタの枠組みでジャイロセンサの誤差補正をおこなった[6]。また、H.Rehbinder らはラインをランドマークとして、同様の実験を実施した[7]。事前のランドマーク登録やライン検出は、使用エリアを限定する。その点、SIFT特徴量はエリアを限定しない。さらに、微小なアフィン変換に対して対応ができる。短時間あたりの姿勢変化を推定する上で都合が良く、精度（解像度）も期待できる。ジャイロセンサとカルマンフィルタの枠組みで、SIFT特徴量を利用した点も先行研究では類がない。

## 3. 提案手法

### 3.1 概要

カメラ画像とジャイロセンサの相乗効果を具現化

するための手法を図2に示す。特徴としては、(A)画像処理からジャイロセンサへの作用(図2点線)、(B)ジャイロセンサから画像処理への作用(図2破線)、が存在し、相互作用のループを構成する。

(A)は、SIFT特徴量を利用して時刻の異なる画像間の対応付けを行い、モーションステレオの原理より回転行列 $R_{img}$ と並進ベクトル $h_{img}$ を算出する。 $R_{img}$ を観測量とし、状態量をジャイロセンサの誤差 $\delta\omega, b$ としてカルマンフィルタによって誤差を推定する。推定した誤差をセンサ値から取り除くことにより姿勢の推定精度を向上させる。さらに画像から動作の停止判定を行い、停止の場合、ジャイロセンサの出力をゼロとすることでドリフト誤差を低減させる。一方(B)は、ジャイロセンサの出力を使い、画像処理で求める基礎行列の精度を高める。特徴点の対応付けで誤対応を除去、さらに初期値として基礎行列の逐次計算に利用した。(A)と(B)ループによってジャイロセンサは誤差を除去し、画像処理も姿勢変化(基礎行列)の精度を高めることができる。提案手法は、図2に示したように、画像処理による観測量の導出部、カルマンフィルタ(KF)による誤差推定部、姿勢算出部、の3つのブロックから構成される。以下、説明する。

### 3.2 画像処理による観測量の導出

具体的な手順は、SIFT特徴量の検出と、以下4つの処理、(1)特徴点のマッチング(2)ジャイロセンサを利用した誤対応の除去(3)停止判定処理(4)基礎行列の算出、から構成される。

#### (1) 特徴点のマッチング

まず時刻の異なる画像間で特徴点のマッチングを行う。マッチングは、SIFT特徴量の最近傍探索処

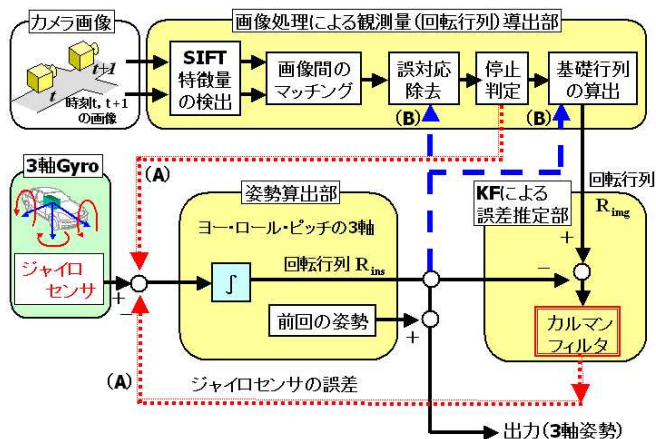


図2 提案手法のシステム構成

理によって行う。特徴間の類似度を表す尺度として、特徴空間中での距離を用い、距離が最小になるものからマッチングの候補とする。高次元ベクトルでの特徴量の探索は複雑であるが、k-d 木探索[8]を基にした Best-Bin-First (BBF)アルゴリズム[9]を用いることにより、効率的に精度よく最近傍の対応点を見つけることができる。

また SIFT 特徴量の対応付けがどの程度までのアフィン変換に対応できるのかも明確にしておく。これは、画像処理から得られる姿勢変化のダイナミックレンジになるため事前検討が必要である。回転テーブルにカメラを、1m 前方に対象物を置き、ヨー方向のみ 5 度刻みで回転させた。その際の対応例を図 3 に示す。5 度変化で対応数 87 個、30 度変化で対応数 46 個であった。対応数は落ちるもの、十分である。制御周期を 0.1s とすれば、角速度 300 度/s となり、広い範囲での動作に対応できると分かる。

#### (2) 慣性センサを利用した誤対応の除去

SIFT 特徴量の対応付けと誤対応の一例を図 4 に示す。マッチングは特徴量空間中での距離を見ているため、特徴点によっては実際の 3 次元空間中では大きく異なった対応が発生してしまう。そこで、このような誤対応を除去するため、ジャイロセンサを利用する。具体的にはまず、ジャイロセンサの測定値を積分することで回転行列  $R_{ins}$  を求める。並進ベクトル  $h_{ins}$  は (1,0,0) とした。これは、座標系を図 5 のようにとり、光軸方向が進行の中心であるため、近似になる。時刻  $t+1$  の特徴量を  $R_{ins}$  と  $h_{ins}$  より時刻  $t$  の画像に射影する。射影された点の  $N$  pixel 四方以内に時刻  $t$  の画像の特徴点があれば対応点として利用する。それ以外は、誤対応として対応点群から削除する。この方法は、処理負荷が軽く、実装も容易である。ここで  $N$  はカメラの幾何モデルを元に、前回の誤差推定量を反映した姿勢変化量を用い、以下の式より決定する。

$$N = \frac{(\text{変化量} + 0.025) \times \text{焦点距離}}{\text{ピクセルサイズ}} \quad (1)$$

#### (3) 停止判定処理

基礎行列算出の前処理として停止判定を行い、停止と判定された場合は基礎行列を求めないようにする。停止判定では、時刻  $t, t+1$  の対応付けされた画像上での特徴点位置を比較し、そのノルムを計算する。そして、全特徴点のノルムの平均値が、1.0 pixel

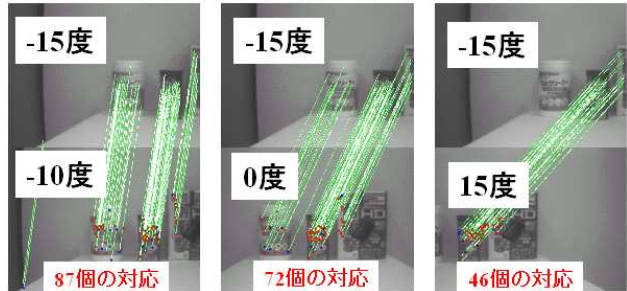


図 3 角度変化と SIFT の対応付け

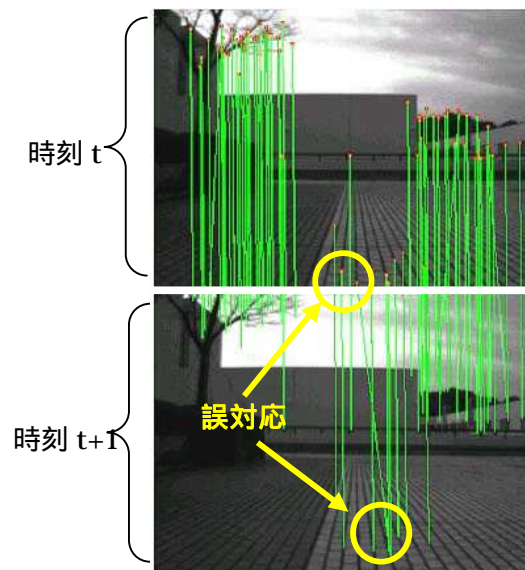


図 4 SIFT の対応付け 誤対応の例

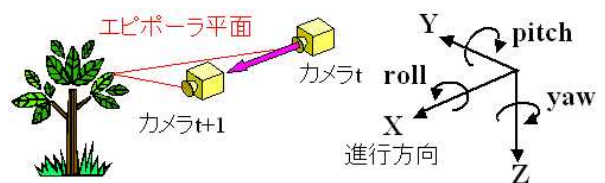


図 5 実験に使用した座標系

以下であれば停止と判定する。さらに停止と判定した場合、ジャイロセンサ値とその推定誤差にゼロを代入する。

#### (4) 基礎行列の算出

基礎行列の精度が、画像処理から得られる姿勢変化の精度（解像度）になる。微小な動作をどこまで表現できるのかを確認したい。今回、図 5 に示したようにカメラの光軸が進行方向になっているため、エピポーラ平面での三角形が直線になりやすく、精度が出にくい。進行方向に対して光軸を傾けた設置であれば改善されるが、進行方向と合わせた方がカメラを他アプリケーションに利用でき都合が良い。

そこで基礎行列の導出を工夫する。導出方法には、最小自乗法の枠組みによる解析的算出法と最急勾配法などによる逐次計算がある[10][11]。解析的算出の場合、固有値計算の計算誤差が無視できない。そこで逐次計算法である Levenberg-Marquardt 法[10]を用いた。これは最急勾配法とニュートン法を融合した方法で、特徴点の場所のピクセル誤差に対して一番安定した手法であったため利用した。評価関数  $C$  は特徴点と対応するエピポーラ線までのユークリッド距離の自乗和として計算する(文献[10]参照)。基礎行列の初期値は、ジャイロセンサから求めた  $R_{ins}$ ,  $h_{ins}$  とする。こうすることでより安定した精度の保持と計算時間の短縮ができる。また、より精度を高めるため、RANSAC[12]を利用し、最後に信頼性評価を行いカルマンフィルタの観測量に用いて良いかを決める。図6に以上のフローを記載した。

事前検討としてシミュレーションを実施した。特徴点の場所に、最大  $\pm 1.0$  ピクセルの誤差を、さらに初期値として利用するジャイロセンサ値にも最大  $\pm 0.004$  rad/sec の誤差を一様乱数で加え試験した。結果、画像処理から得られる精度誤差は  $0.0000357$  rad/sec で、積算3度の誤差に25分まで耐久できる仕様である。

### 3.3 カルマンフィルタによる誤差の推定

状態量の設計と観測方程式の立式を説明する。更新則や詳細な計算方法については、文献[13]を参考にされたい。またカルマンフィルタについては文献[14][15]を参考にした。

#### (1) 状態量の設計

ジャイロセンサの値を「真値+誤差」の形で分け、誤差量  $b$  を状態量  $x$  (6次元)として推定する。式(2)における  $\hat{\omega}$  は真値、 $\omega$  は測定値を表す。

$$\begin{bmatrix} \omega_{roll} \\ \omega_{pitch} \\ \omega_{yaw} \end{bmatrix} = \begin{bmatrix} 1+b_{roll} \\ 1+b_{pitch} \\ 1+b_{yaw} \end{bmatrix} \begin{bmatrix} \hat{\omega}_{roll} \\ \hat{\omega}_{pitch} \\ \hat{\omega}_{yaw} \end{bmatrix} + \begin{bmatrix} \delta\omega_{roll} \\ \delta\omega_{pitch} \\ \delta\omega_{yaw} \end{bmatrix} \quad (2)$$

$b$  はオフセットを意味し、 $b$  はスケールファクタの誤差を意味する(図7参照)。オフセットは電源の揺らぎなどで生じ、スケールファクタは検出した電荷を増幅する場合などで発生しやすい[16][17]。半導体は温度依存性があるため、常時推定できる方法が望ましい。図8に温度変化を与えた際のジャイロセンサの出力結果(静止時)を示す。静止時のた

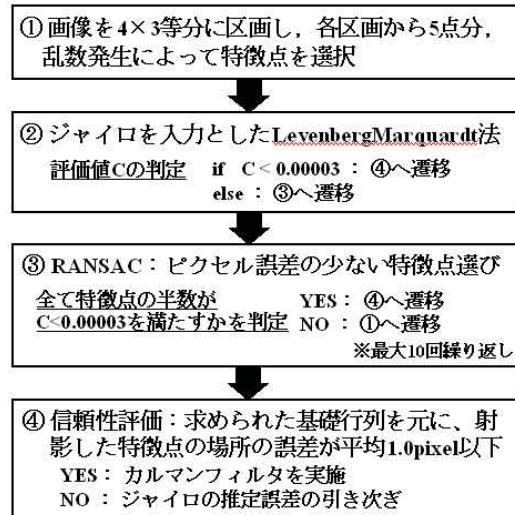


図6 高精度な基礎行列の導出フロー

めゼロが真値である。  $\pm 0.004$  rad/s の白色ノイズがあり(高周波成分)、温度の上昇と共に緩やかに誤差が増える(低周波成分)。白色ノイズは時間平均をとれば0であるため、ローパスフィルタで除くことができる。カットオフ周波数が57Hzであるが、取り除けない部分もソフト処理で除ける。一方、低周波成分の誤差は、完全にDC成分でない限り、除くのが難しい。温度などの外的要因で緩やかな変化があるので、これをカルマンフィルタで除く。

#### (2) 観測方程式の立て方

観測方程式は、画像から求めた  $R_{img}$  と  $h_{img}$  の各要素を観測量  $s$  (12次元)とし、 $s$  がジャイロセンサで求めた  $R_{ins}$  と  $h_{ins}$  の各要素と等しいという拘束条件で設計する。拘束式は、 $s$  の非線形関数にもなるため、 $s$  周辺でテーラー展開して線形化する。拡張カルマンフィルタの枠組みで設計する[13]。

### 3.4 誤差を考慮した姿勢推定

システムとしての出力は姿勢である。センサの測定値から式(2)で設計した誤差を引き、ジャイロセンサ値を積算したものが姿勢である(図2参照)。

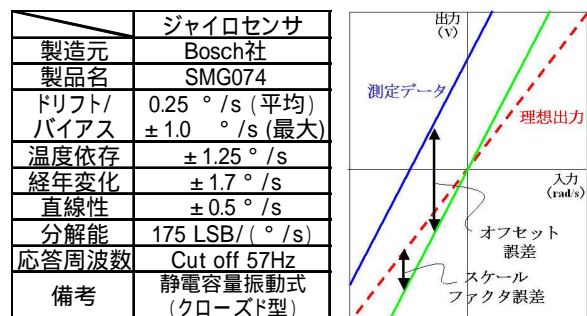


図7 使用したジャイロセンサと状態量

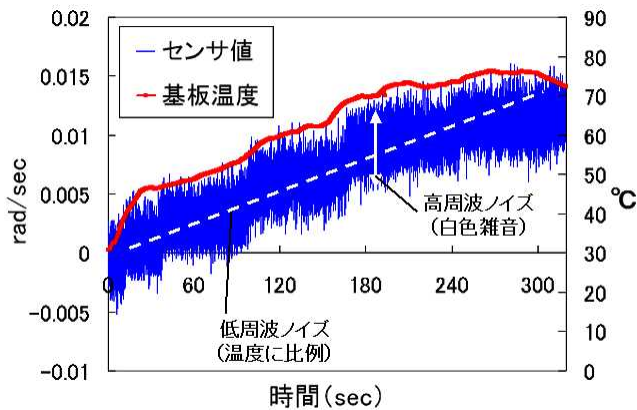


図8 温度とジャイロセンサの誤差

## 4. 実験

### 4.1 実験条件・手順

車両の上にジャイロセンサ ( Bosch 製: SMG074 ), 評価用に光ファイバジャイロ ( FOG ) ( 日本航空電子工業製: JM3403 ) を搭載し, カメラを組み付け, 1500mm/s 以下の移動速度で走行, 10Hz の周期で計測した. 走行路面は石畳である. FOG の値を真値 ( 精度: 1H で 3° の誤差 ) として, 提案手法によってジャイロセンサ ( SMG074 ) の誤差の補正を試みた. また今回はデータを計測後, オフラインにての検証結果になる.

### 4.2 結果・考察

表 1 に提案手法によって測定された姿勢精度とジャイロによって測定された姿勢精度をヨー, ロール, ピッチの各軸に対する平均誤差として示す. 提案手法では, いずれの軸も 900 フレーム間 ( 90sec ) での時間平均にして 3 度程度の改善を果たした. さらにヨー方向のグラフを図 9 に示す. ジャイロセンサ ( SMG074 ) は徐々にずれていき, 最終的には 900 フレーム ( 90sec ) 後, 0.1rad 程度 ( 約 6 度 ) の誤差があるのに対し, 提案手法は常時ほぼ FOG ( 真値 ) に近い値を示した. これより提案手法による精度改善が確認できる.

また図 10 に画像のみで姿勢算出を実施した場合の結果を示す. 途中大きなずれが生じている. これは基礎行列を十分な精度で求めることができなかったためである. しかし上記のような場合は, 画像からの回転行列を参照せず, カルマンフィルタを適用しない. 推定した誤差とジャイロセンサ値から姿勢を導出する. またオクルージョンなどの影響で, 特徴点の対応ができない場合なども同様である. カルマンフィルタを適用するのは, 図 6 に示した基礎行

表 1 姿勢精度 ( 平均誤差 ) の比較

900フレーム間の平均誤差(度)	yaw	roll	pitch
従来手法(ジャイロのみ)	3.9382	5.3427	9.8859
提案手法	0.6791	1.9909	6.7544

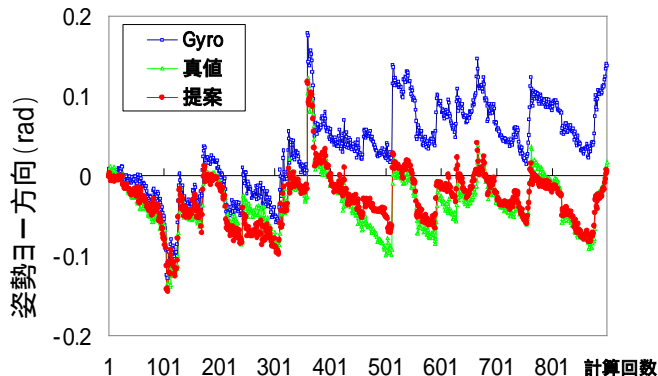


図9 提案手法の効果

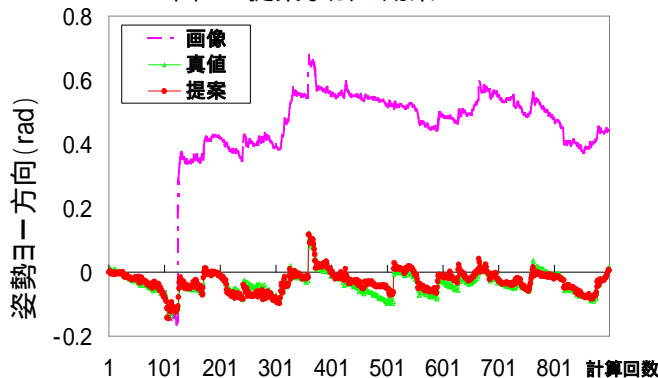


図10 カメラ画像による姿勢

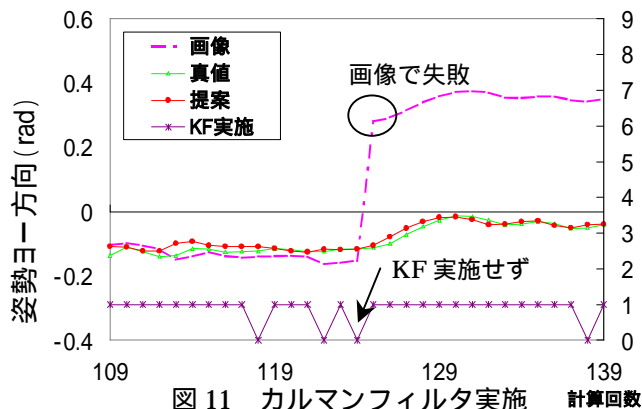


図11 カルマンフィルタ実施

列導出の際に信頼性評価をクリアした場合のみである. 実際, 900 フレーム間でカルマンフィルタ ( KF ) を適用したのは 685 フレームである. 図 11 は KF の実施を示す. 『 KF 実施 』 は, 1 がカルマンフィルタの実施を, 0 が実施しなかった場合を意味する. 画像で失敗した際, カルマンフィルタは実施されていない. 断片的に画から回転行列が求めれば良いことが分かる. 画像からの対応が不十分な場合でもロバストに働くことが分かる.

## 5. まとめ

画像から回転行列・並進ベクトルを算出し、カルマンフィルタによってジャイロセンサの誤差補正を行う手法を提案した。提案手法は、(A)画像処理からジャイロセンサへの作用、(B)ジャイロセンサから画像処理への作用、の2つの相互作用のループによってロバスト性、精度向上の相乗効果を実現した。画像間の対応が取れない場合や画像間の対応からどの程度の姿勢変化の推定精度が得られるかの検討を行った。さらに、システム全体の性能評価を行った。

今後は、ジャイロセンサの誤差とカルマンフィルタでの推定量を周波数解析する。本手法がどの周波数領域で最も効果が出るのかを確認していきたい。

## 参考文献

- [1] Reinhard Neul, Udo-Martin Gómez, Kersten Kehr, Wolfram Bauer, Johannes Classen, Christian Döring, Ermin Esch, Siegbert Götz, Jörg Hauer, Burkhard Kuhlmann, Christoph Lang, Michael Veith, and Rainer Willig, "Micromachined Angular Rate Sensors for Automotive Applications," IEEE Sensors Journal, Vol. 7, No. 2, pp.302-309, Feb 2007
- [2] D. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," Proc. of International Journal of Computer Vision (IJCV), 60(2), pp.91-110, 2004.
- [3] Min-geun Song, Jinseong Park, Youngjin Park and Youn-sik Park, "Fusion Filter for Orientation Estimation of Biped Robot," Motion and Vibration Control, pp.285-294, Springer Netherlands, 2009.
- [4] Eun-Hwan Shin, Naser El-Sheimy, "Accuracy Improvement of Low Cost INS/GPS for Land Applications," Proceedings of the 2002 National Technical Meeting of the Institute of Navigation, pp.146-157, Jan. 2002.
- [5] Major M. Veth, J. Raquet, "Fusion of Low-Cost Imaging and Inertial Sensors for Navigation," Proceedings of the ION GNSS, 2007
- [6] S. You and U. Neumann, "Fusion of Vision and Gyro Tracking for Robust Augmented Reality Registration," Proceedings of the Virtual Reality 2001 Conference, pp. 71-78, 2001
- [7] Henrik Rehbinder, and Bijoy K. Ghosh, "Pose Estimation Using Line-Based Dynamic Vision and Inertial Sensors," IEEE Transactions on Automatic Control, Vol. 48, No. 2, pp.186-199, 2003
- [8] R. セジウィック, "領域探索," アルゴリズム C 探索・文字列・計算幾何, pp.205-220, (株)近代科学社, 1996.
- [9] J.S. Beis and D.G. Lowe., "Shape indexing using approximate nearest-neighbour search in high-dimensional spaces," Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, pp.1000-1006, 1997.
- [10] 徐剛, 辻三郎, "エピポーラ幾何," 3次元ビジョン, pp.61-77, 共立出版(株), 2005.
- [11] 金谷健一, "3次元剛体運動の画像解析," 画像理解 3次元認識の数理, pp.79-114, 森北出版(株), 2005.
- [12] Fischler M.A and Bolles R.C, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," Communications of the ACM, Vol.24, No.6, pp.381-395, June 1981.
- [13] 小堀訓成, 出口大輔, 高橋友和, 井手一郎, 村瀬洋, "カメラとジャイロセンサを組み合わせた移動ロボットの高精度な姿勢推定," 画像の認識・理解シンポジウム (MIRU), pp.1757-1763, 2009.
- [14] 加藤寛一郎, 最適制御入門, レギュレータとカルマン・フィルタ, 東京大学出版会, 1987.
- [15] 金井喜美雄, ビークル, 坪内孝司(編), pp.142-173, (社)計測自動制御学会, 2003.
- [16] 前田龍太郎(他), MEMSのはなし, 日刊工業新聞社, 2008.
- [17] 坂本修(他), ジャイロ活用技術入門, 多摩川精機(編), (株)工業調査会, 2002.