

FFTに基づくロバストな画像のレジストレーション手法の設計

学生員 李 堯希* 上級会員 藤本 博志*

Practical Algorithm Design for the FFT-based Robust Image Registration Method

Yoshi Ri*, Student Member, Hiroshi Fujimoto*, Senior Member

(2018年4月13日受付, 2018年8月29日再受付)

Image registration plays a very important role in a camera-based remote sensing system. This paper discusses a 2D fast Fourier transform (FFT)-based method, which is well known for its easy implementation, accuracy, and robustness. The problem with the FFT-based method is that its robustness and accuracy highly depend on some preprocessing such as the design of a bandpass filter and subpixel estimation. This paper provided solutions for its algorithm design and tuning issues. Finally, the evaluation based on the image dataset and the actual images demonstrated its effectiveness.

キーワード：画像レジストレーション, 位相相関, 回転不変離散フーリエ変換

Keywords: image registration, phase correlaton, RIPOC

1. 序 論

近年のイメージセンサの小型高性能化と画像処理技術の発展に伴い、イメージセンサを用いたリモートセンシング技術は産業において幅広く用いられる様になった。画像を用いたセンシングでは物体のテクスチャ情報を元に対象の回転や並進等の相対変位について詳細な情報を得ることが出来るため、対象物体の追跡やカメラを取り付けた箇所の変位の測定等に用いられている。これらの技術の根幹を支えているのは複数の画像間の幾何的な変位を計測する画像のレジストレーション技術であり、正確性やロバスト性、計算量の少なさ等が求められている。特にビジュアルサーボなど、検出結果を元に何らかの動作を行う際にはこれに加えて異常値の検出可能性も重要になる⁽¹⁾。

本稿では一般によく用いられる、カメラの観察対象の物体が剛体かつ平面に近似できるとの仮定の元での画像のレジストレーション技術について議論を行う。現在用いられる画像のレジストレーション技術はおおまかに、画素に基づく評価関数で変位を最適化する手法、局所特徴量を用いた手法、離散フーリエ変換を用いた周波数領域に着目した手法の3つに分類できる。

Lucas-Kanade 法⁽²⁾を始めとした評価関数に基づく手法では、画像間の変位のモデルを仮定した後に残差2乗和など

の評価関数を設定し、評価が最も良くなるような変位のパラメータを勾配法等の逐次的な手法で最適化して求める。適切な評価関数を設定することで照明変動等にロバストパラメータを推定できる⁽³⁾という利点がある一方で、繰り返し計算による計算コストの増大や最適化の際に局所解に陥るといった構造的な問題を抱えている。他方、局所特徴量としてはロボティクスの分野でよく用いられる Scale Invariant Feature Transform (SIFT) 特徴量⁽⁴⁾や Speed Up Robust Features (SURF) 特徴量⁽⁵⁾に基づく特徴点の検出とマッチング手法が存在する。局所的な対応関係を得られるため非平面対象などにも適用することが出来、3次元再構成等に用いられる一方で、特徴点の検出やマッチングに計算コストが多く掛かる他、誤った対応付けを元に意図しない検出結果を生むという危険性もはらんでいる⁽⁶⁾。本稿ではこれらに対して高速フーリエ変換 (FFT) に基づくレジストレーション手法について提案を行う。画像の2次元フーリエ変換を用いた手法は位相相関 (Phase Correlation)⁽⁷⁾がよく知られており、フーリエ変換の平行移動と定数倍に関する性質から画像の並進と回転、スケール変位量を検出することが出来る⁽⁷⁾⁽⁸⁾。同様のパラメータを推定する際に、位相相関を用いることのできる他の手法と比べた際の利点として次の3つが挙げられる。

計算の効率性 本手法は Lucas-Kanade 等の反復計算が必要な手法と異なり、決められたプロセス内で推定値を導き出せる。そのうち計算量の大多数を占める FFT や一部座標変換は FPGA や GPU 等のハードウェアでの並列計算が容易であり数 ms オーダーまで高速化が可能⁽⁹⁾である。

* 東京大学工学系研究科
〒277-8561 千葉県柏市柏の葉 5-1-5
The University of Tokyo
5-1-5, Kashiwanoha, Kashiwa, Chiba 277-8561, Japan

計算時間の一定性 計算に用いるプロセスが固定されているため計算時間が一定になることが期待される。局所特徴量を介した手法は画像のテクスチャや特徴点の数に依存して計算量が変動する他、外れ値除去に用いられるランダムサンプリングに基づく RANdom SAmple Consensus (RANSAC)⁽¹⁰⁾といった手法も計算時間の一定性をいっそう困難にする場合が存在する⁽¹¹⁾。

評価の容易さ 位相相関を用いた手法を用いることで得られる相関のピーク値を見ることで誤検知を検出できるため、制御系などにおける誤動作の危険性を回避できる。効率的な計算が比較的容易に実装でき、かつ最適化やマッチングの探索等のプロセスを介さないという性質は画像情報を元にしてロボットを駆動するビジュアルサーボ等のリアルタイム性が求められるアプリケーションに適している。また、変位検出時に得られる類似度を評価することにより意図しないレジストレーション結果を能動的に排除する⁽¹²⁾ことで実際のロボットの危険な動作を予防できるといった利点も持ち合わせている。

このように位相相関を元にした画像レジストレーション技術はこれらの利点を有しているものの、並進に加えて回転や拡大縮小量を含めた4自由度のパラメータ推定時には文献(7)で公開されている情報に基づく手法では十分なロバスト性を確保出来ない問題が存在する。後の青木ら⁽⁸⁾⁽¹³⁾によって適切な周波数フィルタリングや白色化等の一部指針が示されたものの、依然としてアルゴリズムの設計には幾多の手作業に依るチューニングが必要である。

本稿ではFFTに基づく画像レジストレーション手法のロバスト性と精度の向上を目指し、各フィルタのハイパーパラメータのチューニングなどの設計指針について幾何的な意味付けとともに明らかにする。2章では原理を説明した後、3章で実際のアルゴリズムに用いられるいくつかの工夫とその設計手法について述べる。その後、4章でデータセットに基づく精度とロバスト性の実証を行う。

2. 位相相関を用いた変位検出の原理

位相相関による変位検出の原理はフーリエ変換の特性により導くことが出来る。Fig. 1 に一般的に知られる位相相関の手法の流れを示す。

〈2・1〉 並進量の検出 画像 f と画像 g が x, y 方向に (x_0, y_0) だけ平行移動した関係である場合、これらの2つの間には次のような関係性が成り立つ。

$$g(x, y) = f(x - x_0, y - y_0) \dots \dots \dots (1)$$

(1) 式を2次元フーリエ変換した結果、 f と g のフーリエ変換 F, G の関係は(2)式のようになる。

$$G(\chi, \psi) = e^{-j2\pi(\chi x_0 + \psi y_0)} F(\chi, \psi) \dots \dots \dots (2)$$

(2) 式より、画像 f と g のクロスパワースペクトラム R は次のように表せる。

$$R(\chi, \psi) = \frac{FG^*}{|FG^*|} = e^{j2\pi(\chi x_0 + \psi y_0)} \dots \dots \dots (3)$$

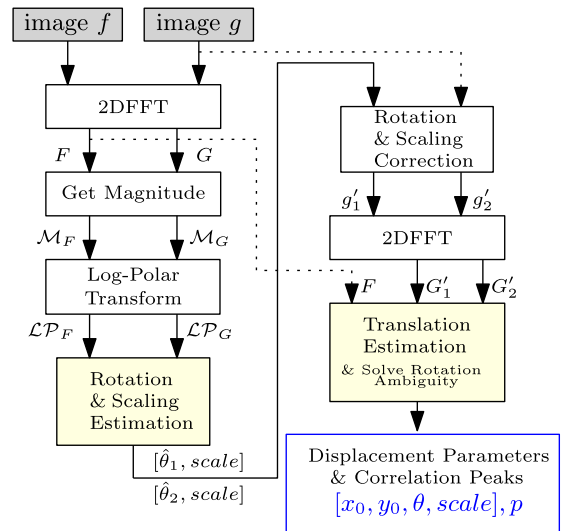


Fig. 1. The FFT based image registration algorithm flow⁽⁷⁾⁽⁸⁾.

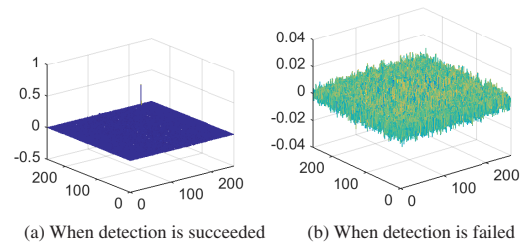


Fig. 2. Appearance of POC function r when detection is (a) succeeded or (b) failed. Existence of the peak shows the detection succeeded.

この R を2次元フーリエ逆変換することにより、座標 (x_0, y_0) において急峻なピークを持つデルタ関数 $r(x, y)$ を得る。

r は画像が繰り返しの場合はデルタ関数になるが、実際には画面端の影響があるため、急峻なピークを持つ関数である。

〈2・2〉 回転・拡大縮小量の検出 次に画像 f と画像

g が並進 (x_0, y_0) だけでなく、回転変位 θ_0 とスケール変化 κ も含んでいる場合を考える。なお今後、画像 f の座標 (x, y) における画素値を $f(x, y)$ と表す。これらの画像の関係は次のような数式で表すことが出来る。

$$g(x, y) = f(x', y') \dots \dots \dots (4)$$

$$x' = \kappa \cos \theta_0 x - \kappa \sin \theta_0 y + x_0 \dots \dots \dots (5)$$

$$y' = \kappa \sin \theta_0 x + \kappa \cos \theta_0 y + y_0 \dots \dots \dots (6)$$

(4) 式にフーリエ変換の定数倍の定理を当てはめることで回転と拡大縮小成分と並進成分を(7)式の様に分離することが出来る。

$$G(\chi, \psi) = e^{-j2\pi(\chi x_0 + \psi y_0)} F(\kappa \cos \theta_0 \chi - \kappa \sin \theta_0 \psi, \kappa \sin \theta_0 \chi + \kappa \cos \theta_0 \psi) \dots \dots \dots (7)$$

(7) 式より、画像 f, g のフーリエ変換 F と G の振幅には回転と拡大縮小の情報のみが保存されており、従って F と G

の振幅 M_F , M_G の関係は (8) 式に示す通りになる。

$$M_G(\chi, \psi) = M_F(\kappa \cos \theta_0 \chi - \kappa \sin \theta_0 \psi, \kappa \sin \theta_0 \chi + \kappa \cos \theta_0 \psi) \dots \dots \dots (8)$$

これらの振幅 M_F, M_G を対数極座標変換 (Log-Polar 変換) することで、回転とスケール変位を並進変位へと変換することが出来る。画像の原点を (cx, cy) とすると対数極座標変換後の画像 $\mathcal{LP}_F(\theta, \rho)$ と元の振幅 M_F の関係は次のように表される。

$$\mathcal{LP}_F(\theta, \rho) = M_F(cx + e^{\frac{\rho}{M}} \cos \theta, cy + e^{\frac{\rho}{M}} \sin \theta) \dots \dots \dots (9)$$

ここで、 M は対数極座標変換における情報の圧縮具合を決定する定数である。この対数極座標変換された画像において、元の画像の回転と拡大縮小はそれぞれ変換後の直交軸 θ, ρ 方向の並進変位として表されるため、ここから前節の位相相関を適用することによって回転と拡大縮小量を抽出することが出来る。

その後求めた回転量と拡大縮小量を用いて、画像 g の回転と拡大縮小を補正した画像 g_{rev} を生成することで、 g_{rev} と f の間には並進変位のみが存在することから並進検出手法を用いて並進を含めた最終的な変位を取得できる。

3. 提案するアルゴリズム

FFT に基づいた画像変位の検出原理を2章で示したが、実際のシステムに適用するにあたって必要なロバスト性等を確保するためにはいくつかの重要な事項が存在する⁽¹³⁾。本稿で新たに提案する〈3・4〉節のフィルタリングと〈3・6〉節のサブピクセル推定処理を含めた全体のアルゴリズムの流れは Fig. 3 の様になる。これら全ての処理の意義を明らかにすることで使用者が状況に応じて柔軟にチューニング可能なアルゴリズムを提案する。

〈3・1〉 窓関数 2次元FFTを行う際に窓関数を用いて画面端の影響を低減することで検出のロバスト性を格段に向上させることが出来る。窓関数の選択によってピーク値の形状が変化することが知られているため窓関数の選択もチューニング対象になりうるが、有意な違いを検証できなかったため本論文では他の多くの文献で用いられる下記の二次元ハニング窓 $w(n)$ を採用する。

$$w(n) = \frac{1}{4} \left(1 + \cos \left(\pi \frac{y}{H} \right) \right) \left(1 + \cos \left(\pi \frac{x}{W} \right) \right) \dots \dots \dots (10)$$

ここで、 H と W はそれぞれ画像の縦と横のサイズを表す。

〈3・2〉 白色化 (8) 式において振幅 M_F を計算する際、低周波の振幅は高周波のそれに比べて指数的に大きいので白色化を行うことで高周波の情報を活かす事ができる。青木⁽¹³⁾らの文献に倣い、以下のような白色化処理を施す。

$$M_F(\chi, \psi) = \log(|F(\chi, \psi)| + 1) \dots \dots \dots (11)$$

〈3・3〉 対数極座標変換におけるスケール定数 (9) 式

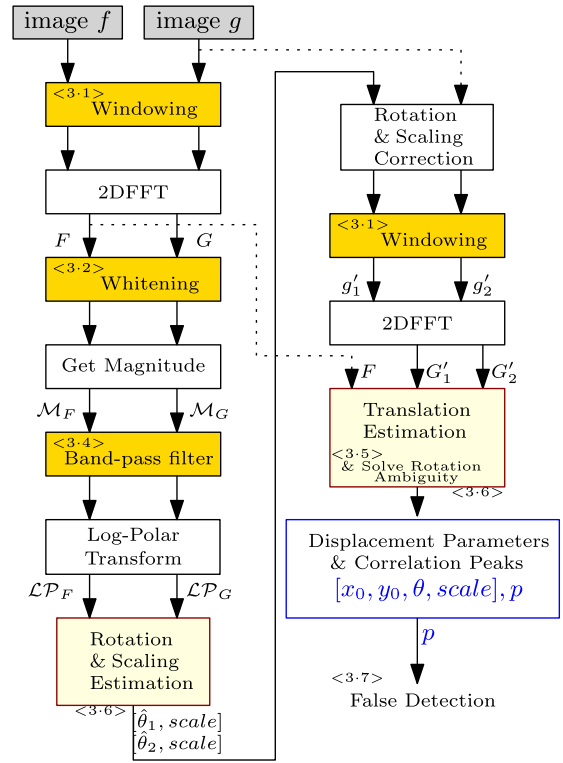


Fig. 3. The actual process flow of the proposed registration method. The yellow boxes contains the tricks described in this paper.

の対数極座標変換において、定数 M の決定は精度やロバスト性に直結する非常に重要な要素である。定数 M が必要な理由は、対数極座標変換では半径 r の円周上のピクセルが変換後のリニアスケールで $\log r$ という小さな値に圧縮されて変換後の画像が極端に小さくなってしまふ為である。

M をある程度大きくとることで座標変換後の情報量を保つ事ができるが、ある程度以上は線形補間による情報欠落もあるため有意な精度改善をもたらさない。元画像 M_F の中心から $\frac{N}{\sqrt{2}}$ 離れた4隅のピクセルは変換後の ρ 軸に対して $M \log \frac{N}{\sqrt{2}}$ に投影されるため、設計時における条件は変換前の端点に変換後に画面内に収まるように $M \log \frac{N}{\sqrt{2}} < N$ を満たすように決定する。本稿では試行錯誤の結果、文献(8)にて用いられた値 $M = \frac{N}{\log N}$ を採用した。

その結果得られる結果は $(\Delta\theta, \Delta\rho) = (\theta_0 \frac{N}{2\pi}, M \log \kappa) = (\theta_0 \frac{N}{2\pi}, N \log_N \kappa)$ と書くことが出来る。

〈3・4〉 振幅に対するバンドパスフィルタ 〈3・3〉節と同様に回転と拡大縮小の推定における重要な作業に周波数領域の情報のフィルタリングが挙げられるが、このフィルタのカットオフの値も幾何的な考察から有意に設計することが出来ることを提案する。文献(8)や文献(14)において、対数極座標変換を用いることで振幅 M_F, M_G にハイパスフィルタを乗じることで精度を上げる手法が用いられている。

この幾何的な意味として Fig. 5 に示すように対数極座標変換をした際、半径 r の円周上に存在する概算 $2\pi r$ 個のピ

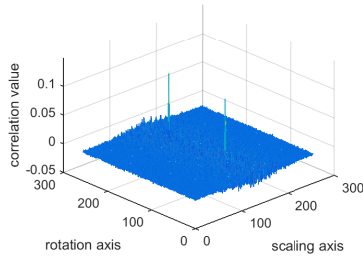


Fig. 4. An example of the correlation peaks for the rotation and scaling estimation between $F_{LP}(l_1, l_2)$ and $G_{LP}(l_1, l_2)$. Two peaks shows there are an ambiguity in rotation detection.

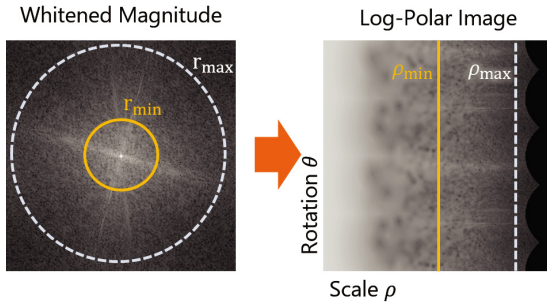


Fig. 5. The band-pass filter appearance in the whitened magnitude M_F (left) and Log-Polarized Image \mathcal{LP}_F (right).

クセル情報が変換後の1辺 N 個のピクセル情報と対応するため、変換前に中心付近にある低域の情報は大きく拡大され情報量の小さい箇所になるためである。一方で、変換前の画面端にあたる高周波域はS/N比が低いことが知られているためノイズに対する頑健性を得るためにある程度のローパスをかける必要がある。従って、これら低域と高域の情報を適切にカットするバンドパスフィルタの設計が必要である。このフィルタのカットオフ値は画像のサイズによって異なるため、本稿では幾何的に意味をもった具体的なチューニングの設計指針を示す。

本稿のハイパスフィルタの設計では、変換前の円周上のピクセル数 $2\pi r$ と変換後のピクセル数 N の比 α をチューニングパラメータとし、カットオフの半径 r_{min} を

$$2\pi r_{min} = \alpha N \dots \dots \dots (12)$$

から計算する。 α が1の時、変換前のピクセル比と変換後のピクセル比が同等になり、 α を小さく設計すればするほどより低域の情報を取り入れることになる。

一方でFig. 5における変換後の右端の部分を取り取るローパスフィルタの設計は次のようにチューニングパラメータ β を用いて設計される。

$$r_{max} = \beta \frac{N}{2} \dots \dots \dots (13)$$

ここで β が1の時、このローパスフィルタは変換前の画像に内接する円より外側にある領域をカットし、 β を小さくすれば小さくするほどローパスの帯域を下げることになる。

また、 $\rho = M \log r$ という関係式から、変換後の座標

系におけるフィルタのカットオフは $\rho_{min} = M \log \frac{\alpha N}{2\pi}$ と $\rho_{max} = M \log \beta \frac{N}{2}$ と表せる。試行錯誤の結果 $N = 256$ の場合には $\alpha = 0.5$, $\beta = 0.8$ 程度が適当であり、他の解像度においても実用に耐えうる結果を得た。シミュレーションでの結果、 α と β を小さくすればより低域の情報を用いるのでノイズへの頑健性等を得る反面、多くの情報量を持つ高域を切り捨てるので精度が低下するというトレードオフが確認できた。

〈3・5〉 回転量と拡大縮小の補正 回転量と拡大縮小量に関する変位の情報は画像の振幅に含まれるが、振幅の対象性により実際に回転量のピークを検出するとFig. 4の様に本来の角度と180度分ずれた角度の2つのピークが確認される。従ってFig. 3の右側に示す流れに従って、求めた2通りの回転量と拡大縮小量を用いて画像 g の回転と拡大縮小を補正し、補正した g'_1 と g'_2 においてそれぞれ参照画像 f との並進量推定を行う。その後Fig. 2に示す様に検出される2通りの位相相関のピーク値の大小により正しい回転角度を判別する。

〈3・6〉 サブピクセルマッチング 非理想環境における位相相関の相関値のピークはsinc関数に似た形状になりピクセル単位で離散化されているが、ピクセル間の情報を補間あるいは評価関数に対する最適化手法等を用いてピーク位置を推定することによってピクセル分解能を超えたマッチングが可能である⁽¹⁵⁾。文献(8)や文献(15)ではそのピークの形状をモデル化して観測情報を用いてフィッティングを行っているが、この非線形関数へのフィッティングは繰り返し計算を必要とする他にノイズの影響を受けやすい⁽¹⁶⁾といった問題がある。したがって本稿では幾多あるサブピクセルマッチングのうち最も簡潔な実装ができ、性能も良いとされる⁽¹⁶⁾ピーク値の周囲5×5ピクセルにおける重心の値をサブピクセル推定値として採用する。本稿における新たな提案として、重心計算をする際にピーク値のある比率 c より相関値の小さいピクセルを0とみなしてノイズの影響を低減し、さらなる精度の向上を得ることが4章のデータセットを用いた検証で示されている。本論文では試行錯誤的に $c = 0.1$ と設定した。

〈3・7〉 検出の可否の判定 最終的に得られる値は、変位 $[x_0, y_0, \theta, \kappa]$ とFig. 2に示されるpoc関数のピーク値 p である。このピーク値は検出成功時は0.1–1.0前後の値になり、失敗時は0.01前後かそれ以下の値になるため、適切なしきい値 p_{th} を設定することで成否を容易に分離できる。一般にテクスチャの強い明暗のはっきりした画像ほどピーク値が高くなる傾向にあるため、この値は画像ごとによって設定されるべきだが本検証では一律に $p_{th} = 0.06$ とした。

4. データセットによる精度・ロバスト性検証

アルゴリズムの実装をMATLABとPythonを用いて行い、変位した画像と対応する変位の値を記録したデータセットを作成した上で検出性能を検証した。性能の評価項目として、検出が成功した画像群の並進・回転・スケール量の

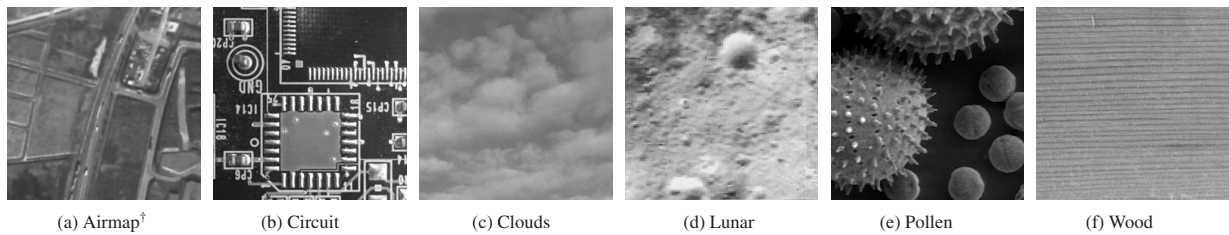


Fig. 6. Reference images used for the evaluation.

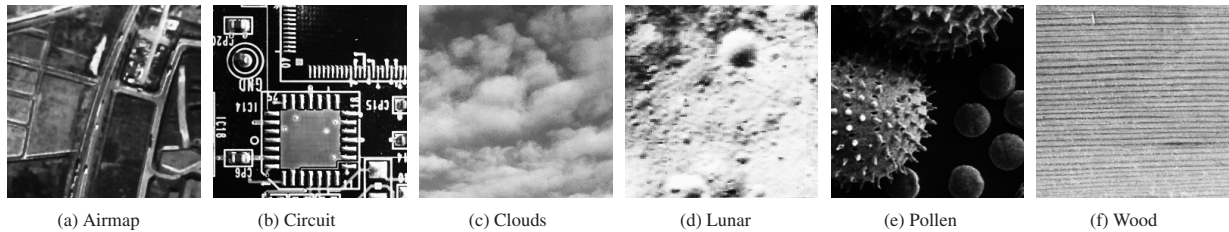


Fig. 7. Contrast changed images by Eq. (14) used for the evaluation.

平均推定誤差（以降精度と呼ぶ）と検出が成功したと判定された割合（以後検出率と呼ぶ）を用いる。

本手法にて用いられた Python と OpenCV を用いたソースコード[†]とデータセット^{††}は公開されており、本検証は Windows の Python3.5 環境で動作を確認している。

〈4・1〉 データセットの作成 本手法の検証のために画像の変形量を推定するためのデータセットを作成した。元となる高解像度の画像を加工した後、その中の中心 256×256 の領域内にある画像を抜き出して参照画像とし、Fig. 6 の画像群に示す。その後ランダム生成した変位 $\xi_{cmp_i} = [dx_i, dy_i, \theta_i, \kappa_i]$ ($i = 1, 2, 3 \dots N$) をもった比較画像を Affine 変換を用いて元の画像から適宜切り取り生成した。

Fig. 6 に示す通り、データセットは UAV や衛星の自己位置同定、基板加工器や顕微鏡画像などのアプリケーションを考慮して航空写真^{†††}、回路基板、雲、月表面^{††††}、花粉^{†‡}、木目の 6 つの画像を含めて作成されている^{†‡‡}。本論文に載せている画像例はテクスチャの強弱や細かさ、コントラストの違いなどを考慮して選定されている。

ロバスト性の検証においては、Fig. 7 に示す明度の変化を模擬してコントラストを調整した画像群を用いてそれぞれ同様の比較と検討を行った。その際、明るい箇所がより明るく、暗い箇所がより暗くなるように以下の関数を用いて変換前の輝度 i を変換後の輝度 i_{new} へと変換した。

$$i_{new} = \frac{255}{1 + \exp\left(-\gamma \frac{i - 128}{255}\right)} \dots \dots \dots (14)$$

なお、Fig. 7 の計算の際には変数を $\gamma = 15$ とした。

計算時間はプログラム実行時の計算負荷に依存したが、FFT を用いた手法と SIFT を用いた手法とも 1 つの画像あたり 50 – 100 ms 前後、ORB を用いた手法は 10 – 30 ms 前後で動作した。

〈4・2〉 既存の手法群との性能比較 前章で示した各工夫の有無に依る影響を検証する。比較として、〈3・2〉節で白色化を用いなかった場合、〈3・4〉節で高周波の領域を除去しなかった場合、〈3・6〉節でサブピクセル推定の重心計算にてカットオフの値を設定しなかった場合の 3 つを比較対象に精度や検出率の比較を行った。〈3・1〉節にて窓関数を適用しなかった場合ほとんどの画像に対してマッチングが成立しなかったため、この項目の比較は省略している。評価指標として提案したしきい値に基づいて評価された検出の成否の割合と各変位の平均誤差の 2 つを用いた。

Table 1 にデータセットのそれぞれの画像に対する各手法の検出率とその平均誤差を示す。白色化を施さない場合、提案手法と比較して精度と検出率の双方が下がってしまうことがわかる。一方で高周波数の領域の情報を除去しない場合は精度に若干の向上が見られるものの、検出率が大幅に下がってしまっている。さらにサブピクセル推定時に一定値以下の値を無視する場合としない場合にて精度に 1.1 – 2 倍前後の開きが生じることがデータセットによる検証で示された。

次に、Table 2 にも明度の変化を含めた際の検出率の変化を示す。一様に検出率は下がっているが、結果として検出時の平均誤差はほとんど変化していないことが示されており、誤り検出が正しく行われていることを示している。明度に対するロバスト性に関して、Fig. 7(c) などのテクスチャを保存している画像ではあまり影響を受けないが、Fig. 7(d)

[†] 作成した変位推定手法のライブラリ <https://github.com/YoshiRi/ImRegPOC>

^{††} 作成したデータセット <https://drive.google.com/drive/folders/1bs0N55Xe4KzFZBimSqyq8bSr1KwqwcW>

^{†††} Google Map より引用

^{††††} NASA のサイト <http://wms.lroc.asu.edu/lroc> より取得

^{†‡} フリー画像 Pixabay <https://pixabay.com/ja/photos/pollen/> より引用

^{†‡‡} 脚注のない画像は著者が iPhoneSE のカメラで撮影した

Table 1. Comparison with proposed method and other method without particular processing in Fig. 3.

Estimation Results	Comparison of Estimation Error in Same Contrast Situation					
	Matching Rate[%], Average Error $\Delta[x, y, \theta, \kappa]$ (pix, pix, deg, scale)					
Templates	airmap	circuit	clouds	lunar	pollen	wood
Proposed	0.17	0.092	0.52	0.14	0.17	0.66
	96%	98%	96%	91%	96%	97%
	0.14	0.15	0.39	0.16	0.17	0.67
	0.11	0.11	0.11	0.12	0.075	0.097
No whitening	0.0067	0.0069	0.0077	0.0067	0.0062	0.0074
	0.27	0.12	0.53	0.20	0.29	0.68
	80%	93%	84%	90%	92%	95%
	0.20	0.16	0.38	0.22	0.27	0.65
No low-pass	0.16	0.14	0.12	0.14	0.11	0.19
	0.010	0.0073	0.010	0.010	0.0053	0.0077
	0.13	0.033	0.51	0.074	0.091	0.58
	65%	35%	71%	36%	59%	44%
No Filtered Center of Gravity	0.086	0.063	0.35	0.091	0.096	0.64
	0.080	0.077	0.078	0.071	0.092	0.068
	0.0054	0.0048	0.0055	0.0055	0.0047	0.0058
	0.27	0.18	0.58	0.17	0.26	0.66
No Filtered Center of Gravity	96%	98%	88%	89%	98%	95%
	0.26	0.21	0.34	0.25	0.31	0.67
	0.20	0.25	0.31	0.25	0.29	0.19
	0.0099	0.0088	0.012	0.010	0.010	0.0096

Table 2. Comparison on processed images with proposed method and other method without particular processing in Fig. 3. Success rate for the each estimation method and its average error within each parameter.

Estimation Results	Comparison of Estimation Error in the Different Contrast Situation					
	Matching Rate[%], Average Error $\Delta[x, y, \theta, \kappa]$ (pix, pix, deg, scale)					
Templates	airmap	circuit	clouds	lunar	pollen	wood
Proposed	0.18	0.095	0.51	0.15	0.24	0.67
	86%	89%	94%	79%	68%	94%
	0.16	0.15	0.39	0.17	0.19	0.67
	0.13	0.13	0.097	0.18	0.16	0.12
No whitening	0.0074	0.0074	0.0078	0.0078	0.0059	0.0077
	0.26	0.11	0.55	0.18	0.38	0.68
	74%	93%	85%	79%	69%	95%
	0.20	0.17	0.37	0.20	0.29	0.66
No low-pass	0.14	0.14	0.10	0.18	0.16	0.20
	0.010	0.0077	0.010	0.0094	0.010	0.0076
	0.10	0.035	0.51	0.085	0.11	0.50
	33%	12%	52%	5%	17%	28%
No Filtered Center of Gravity	0.094	0.063	0.30	0.10	0.10	0.63
	0.080	0.077	0.097	0.16	0.081	0.074
	0.0059	0.0048	0.0058	0.0090	0.0046	0.0061
	0.26	0.19	0.57	0.22	0.36	0.67
No Filtered Center of Gravity	83%	87%	83%	66%	68%	93%
	0.25	0.20	0.33	0.18	0.31	0.70
	0.19	0.24	0.33	0.44	0.39	0.24
	0.0098	0.0090	0.011	0.010	0.0087	0.010

の月の表面画像のようなテクスチャを一部消失した画像に対しては大きく検出率が下がっている。この事実や実際に検出成功した画像と失敗した画像を比較することで、FFTに基づくレジストレーション技術は比較する画像のテクスチャの豊富さに左右されることが実験的に確認された。

これらの比較からも分かる通り、FFTを用いた画像のレジストレーションにおいて提案した各工夫がロバスト性と精度に影響を与えていることが明らかであり、特に〈3・4〉節のバンドパスフィルタの設計の影響が大きいことがTable 1の1行目と3行目の検出率を比較することでわかる。〈3・4〉節の周波数のフィルタの適切なカットオフ周波数は画像のサイズや画像の種類に左右されるが、〈3・4〉節の α や β の幾何的意味付けによってより直感的なフィルタ設計が可能になる。また、今回新たに提案されたサブピクセル推定時に

一定値以下の値を切り捨てる手法は実装が容易である一方で、精度に明らかな向上が見られるため重心を用いたサブピクセル推定において実用的な選択肢の一つを示したと言える。

花粉の画像や回路基板などのテクスチャの強い画像において、提案法の検出率が元の手法より劣っている箇所が存在するがこれは精度のやや悪い検出を切り捨てるようなしきい値 p_{th} を用いているためであり、検出率が高くなっている手法は一方で精度が下がっていることから画像により p_{th} が適切にチューニングされる必要があることを示している。

〈4・3〉特徴点ベースの手法との精度比較 次に Table 3と Table 4にて特徴点を用いたレジストレーション手法との比較を示す。特徴点ベースの手法における代表的な

Table 3. Comparison between FFT based method and feature point based method.

Estimation Results	Comparison of Estimation Error in Same Contrast Situation					
	Matching Rate[%], Average Error $\Delta[x, y, \theta, \kappa]$ (pix, pix, deg, scale)					
Templates	airmap	circuit	clouds	lunar	pollen	wood
FFT based method (Proposed)	0.17	0.092	0.52	0.14	0.17	0.66
	96%	98%	96%	91%	96%	97%
	0.14	0.15	0.39	0.16	0.17	0.67
	0.11	0.11	0.11	0.12	0.075	0.097
SIFT based method	0.0067	0.0069	0.0077	0.0067	0.0062	0.0074
	0.35	0.30	0.69	0.32	0.31	0.95
	100%	100%	74%	100%	100%	93%
	0.33	0.31	0.69	0.32	0.34	0.95
ORB based method	0.013	0.0084	0.16	0.0096	0.0080	0.038
	0.00026	0.00012	0.0029	0.00015	0.00012	0.00057
	0.97	0.49	NaN	0.66	1.5	0.60
	100%	99%	0%	99%	96%	24%
ORB based method	0.73	0.407	NaN	0.54	1.0	0.48
	0.61	3.9	NaN	4.1	1.6	0.31
	0.017	0.0057	NaN	0.0077	0.019	0.0052

Table 4. Comparison between FFT based method and feature point based method. Success rate for the each estimation method and its average error within each parameter.

Estimation Results	Comparison of Estimation Error in the Different Contrast Situation					
	Matching Rate[%], Average Error $\Delta[x, y, \theta, \kappa]$ (pix, pix, deg, scale)					
Templates	airmap	circuit	clouds	lunar	pollen	wood
FFT based method (Proposed)	0.18	0.095	0.51	0.15	0.24	0.67
	86%	89%	94%	79%	68%	94%
	0.16	0.15	0.39	0.17	0.19	0.67
	0.13	0.13	0.097	0.18	0.16	0.12
SIFT based method	0.0074	0.0074	0.0078	0.0078	0.0059	0.0077
	0.57	0.30	1.3	0.47	0.33	0.95
	86%	100%	69%	97%	100%	93%
	0.78	0.31	0.98	0.38	0.47	0.95
ORB based method	0.63	0.025	0.84	0.13	0.14	0.038
	0.019	0.00043	0.017	0.003	0.0030	0.00057
	0.13	0.033	NaN	2.7	3.7	0.78
	65%	35%	0%	71%	79%	19%
ORB based method	0.086	0.063	NaN	1.4	2.2	0.61
	0.080	0.077	NaN	1.6	10.9	0.28
	0.0054	0.0048	NaN	0.020	0.056	0.0040

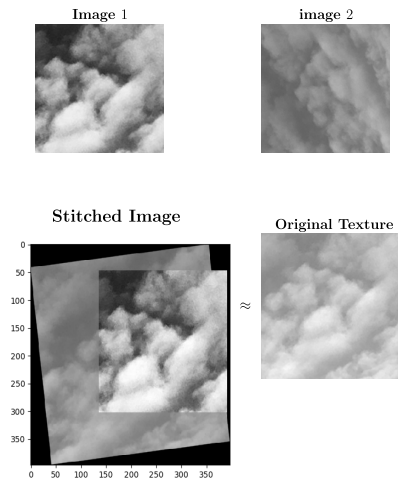


Fig. 8. One of the image matching example with the proposed method, while other methods can not achieve registration.

特徴量としてロバストな手法としてロボティクスで多く用いられる SIFT⁽⁴⁾と高速性から実時間アプリケーションで多用される ORB⁽¹⁷⁾を選択した。これらの特徴点を用いた変位推定にあたって、外れ値の影響を低減するためランダムサンプリングに基づく外れ値除去手法の RANSAC⁽¹⁰⁾を用

い、残った対応点の数が4つ以上の場合に検出成功とした。

ORBを用いた手法は Fig. 6(a)や Fig. 6(b)に示すテクスチャの強い画像に対しては高速かつ正確に変位を計測可能であるが、一方で明度の変化に弱く、Fig. 6(c)の Fig. 6(f)などのテクスチャのぼやけた画像を得意としないことがわかる。SIFTは全体的にロバスト性と精度を兼ね備えているが、テクスチャの曖昧な画像に対しては十分な特徴点が取得できず、正しい検出結果を得られない場合があった。データセットのうち提案手法でのみ変位を取得できた画像の一例を Fig. 8に示す。精度の比較では提案手法の誤差は特徴ベースの手法と異なり、並進1ピクセル未満、回転1deg未満の高精度を保っており、検出失敗時のデータをうまく判別できていることがわかる。

位相相関は原理上比較する2つの画像のオーバーラップ領域がある程度以上大きくなる必要があるが、SIFTやORBなどの特徴点ベースの手法の場合は特徴点が豊富すなわちテクスチャが豊富であればより狭い領域同士でもマッチングできる可能性があり、Fig. 6(e)の画像における検出率の差異につながっていると思われる。しかし、冒頭に示したとおり、高速化やハードウェア実装上での利点や誤検出の検知などの面では位相相関ベースの提案手法の方が有利となる場合もあるため使用シーンに応じてこれらの手法を柔

軟に切り替える必要がある。

5. ま と め

本稿では画像のFFTに周波数領域の情報を用いた変位検出手法について各処理の物理的・幾何的意味付けを明らかにした上で、精度・ロバスト性を向上させるフィルタ設計指針とサブピクセル推定法を含む一連のアルゴリズムの設計指針を検討した。実際の画像に対しては対象のもつテクスチャや照明条件、画像サイズに応じてバンドパスフィルタなどのいくつかのパラメータを変更させる必要があるが、提案したアルゴリズムにおいてはそれらを直感的に解釈することが可能である。

さらに実際のデータセットに基づく検証から、各工夫が精度やロバスト性にどの程度寄与するか、また他の分野でよく用いられるSIFT特徴点などの手法との比較と使い分けについて明らかにした。Fig. 6(c)のような曖昧なテクスチャを持った画像には依然として提案した手法の優位性が存在し、ピーク値による誤り判別はビジュアルサーボなどの動作生成時の切り替えやパノラマ画像作成⁽¹⁸⁾時の最適化の重み導出などの際に有用である。

文 献

- (1) Y. Ri and H. Fujimoto: "Proposal of visual servoing using phase-only-correlation (POC)", IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society, pp.5068–5073 (2015)
- (2) S. Baker and I. Matthews: "Lucas-Kanade 20 years on: A unifying framework", International Journal of Computer Vision, Vol.56, No.3, pp.221–255 (2004)
- (3) A. Dame and E. Marchand: "Second-order optimization of mutual information for real-time image registration", IEEE Transactions on Image Processing, Vol.21, No.9, pp.4190–4203 (2012)
- (4) D.G. Lowe: "Object recognition from local scale-invariant features", Proceedings of the Seventh IEEE International Conference on Computer Vision, Vol.2, No.8, pp.1150–1157 (1999)
- (5) H. Bay, T. Tuytelaars, and L. Van Gool: "SURF: Speeded up robust features", Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol.3951, LNCS, pp.404–417 (2006)
- (6) C.D. Kuglin and D.C. Hines: "The phase correlation image alignment method", Proceeding of the IEEE International Conference on Cybernetics and Society, pp.163–165 (1975)
- (7) B.S. Reddy and B.N. Chatterji: "An FFT-based technique for translation, rotation, and scale-invariant image registration", IEEE Transactions on Image Processing, Vol.5, No.8, pp.1266–1271 (1996)
- (8) K. Takita, T. Aoki, Y. Sasaki, T. Higuchi, and K. Kobayashi: "High-Accuracy Subpixel Image Registration Based on Phase-Only Correlation", IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, Vol.E86-A, pp.1925–1934 (2003)
- (9) K. Matsuo, T. Hamada, M. Miyoshi, Y. Shibata, and K. Oguri: "Accelerating Phase Correlation Functions Using GPU and FPGA", 2009 NASA/ESA Conference on Adaptive Hardware and Systems, pp.433–438 (2009)
- (10) M.A. Fischler and R.C. Bolles: "Random Sample Consensus: A Paradigm for Model Fitting with", Communications of the ACM, Vol.24, pp.381–395 (1981)
- (11) K. Tanaka and E. Kondo: "Incremental RANSAC for online relocation in large dynamic environments", Proceedings - IEEE International Conference on Robotics and Automation, Vol.2006, pp.68–75 (2015)
- (12) T. Hatakeyama, T. Numata, and Y. Iwatani: "Indoor hovering flight of a nano-scale helicopter by markerless visual servo control", 2014 IEEE International Conference on Robotics and Biomimetics, IEEE ROBIO 2014, pp.1880–1885 (2014)
- (13) 青木孝文・伊藤康一・柴原琢磨・長嶋 聖:「位相限定相関法に基づく高精度マシンビジョン—ピクセル分解能の壁を越える画像センシング技術を目指して」, IEICE Fundamentals Review, Vol.1, pp.30–40 (2007)
- (14) G. Tzimiropoulos, V. Argyriou, S. Zafeiriou, and T. Stathaki: "Robust FFT-based scale-invariant image registration with image gradients", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.32, No.10, pp.1899–1906 (2010)
- (15) H. Foroosh, J.B. Zerubia, and M. Berthod: "Extension of phase correlation to subpixel registration", IEEE Transactions on Image Processing, Vol.11, No.3, pp.188–199 (2002)
- (16) A. Alba, J.F. Viguera-Gomez, E.R. Arce-Santana, and R.M. Aguilar-Ponce: "Phase correlation with sub-pixel accuracy: A comparative study in 1D and 2D", Computer Vision and Image Understanding, Vol.137, pp.76–87 (2015)
- (17) E. Rublee, V. Rabaud, K. Konolige, and G. Bradski: "ORB: An efficient alternative to SIFT or SURF", Proceedings of the IEEE International Conference on Computer Vision, pp.2564–2571 (2011)
- (18) Y. Ri and H. Fujimoto: "Drift-free Motion Estimation from Video Images using Phase Correlation and Linear Optimization", The 15th International Workshop on Advanced Motion Control, p.N/A (2018)

李 堯 希 (学生員) 2017年3月東京大学大学院新領域創成科学研究科先端エネルギー工学専攻博士前期課程(修士課程)卒業。同年4月より同大学大学院工学系研究科電気系工学専攻博士後期課程に在学。ビジュアルサーボ、モーションコントロール、センサフュージョンに関する研究に従事。



藤 本 博 志 (上級会員) 2001年3月東京大学大学院工学系研究科電気工学専攻博士課程修了。博士(工学)。同年長岡科学技術大学工学部電気系助手。2002年2003年、米国Purdue大学工学部機械工学科客員研究員。2004年横浜国立大学大学院工学研究院講師。2005年同助教授、2007年同准教授。2010年東京大学大学院准教授。制御工学、モーションコントロール、マルチレート制御、ナノスケールサーボ、電気自動車の運動制御、モータとインバータの高性能制御、ビジュアルサーボに関する研究に従事。2001年および2013年にIEEE Trans. IE 最優秀論文賞、2010年Isao Takahashi Power Electronics Award、2010年計測自動制御学会著述賞、2016年永守賞大賞などを受賞。IEEE 上級会員、計測自動制御学会、日本ロボット学会、自動車技術会各会員。

