

## 科学研究費助成事業 研究成果報告書

令和 2 年 5 月 28 日現在

機関番号：32612

研究種目：基盤研究(C)（一般）

研究期間：2017～2019

課題番号：17K06439

研究課題名（和文）位相情報を用いた物体認識のための高精度位置・姿勢推定に関する研究

研究課題名（英文）High accuracy position and pose estimation for object recognition using phase information

研究代表者

池原 雅章（IKEHARA, MASAOKI）

慶應義塾大学・理工学部（矢上）・教授

研究者番号：00212796

交付決定額（研究期間全体）：（直接経費） 3,500,000円

研究成果の概要（和文）：物体検出とは、対象物体を認識し、高精度な位置・姿勢推定を行うことである。この技術は産業における自動生産ラインやロボットビジョンにおいて非常に重要である。動画像符号化におけるシフト量推定やステレオビジョン等、その応用は広範囲である。本研究課題では、位相情報を用いて、対象物体の高精度なサブピクセル精度の位置・姿勢推定を行うことを目的とした。位置的精度を追求する研究は少ないが、従来法と比較して1桁以上高い精度の推定が可能になったのみならず、一次元処理を基本とするため超高速な推定が可能になった。本研究により画像処理分野のみならず、産業応用上極めて経済効果の高い成果が期待できる。

研究成果の学術的意義や社会的意義

微細基盤形成技術が必要となるフラットパネルや太陽電池を含む多くの半導体産業において、パネル貼り合わせやマスクパターンと基盤の位置合わせのために、物体検出は極めて重要である。その市場規模は2016年には4000億円にも達すると言われ、画像処理技術の中でも最も経済効果の高い技術であり、且つ近年の重要産業を支える基盤技術である。また、近年特に注目されている拡張現実感（AR）等において、現実世界の対象物体に何らかの重畳表示を行う際に、視覚的違和感がなく安定した表示を行うためには、対象の位置・姿勢推定が重要になる。このようにサブピクセル物体検出は新たな産業への橋渡しともなる重要な技術である。

研究成果の概要（英文）：Rotation and scale estimation of images are fundamental tasks in image registration. The conventional estimation method uses log-polar transform and 1D shift estimation to estimate rotation and scale regardless of the shift of images. However, this transform requires interpolation of the frequency components, which causes estimation error. We propose a rotation and scale estimation algorithm based on the Radon transform and sub-pixel shift estimation. The Radon transform can estimate the rotation independent of the shift and can reduce the influence of interpolation error. In addition, sub-pixel shift estimation using a linear approximation of the phase component improves the precision of 1D shift estimation and achieves accurate rotation estimation. The proposed method was evaluated on test images, and the results demonstrate that the proposed method has higher accuracy compared with the log-polar transform.

研究分野：信号処理、画像処理

キーワード：物体検出 姿勢推定 サブピクセル精度

## 様式 C - 19、F - 19 - 1、Z - 19 (共通)

### 1. 研究開始当初の背景

物体検出とは、対象物体を認識し、高精度な位置・姿勢推定を行うことである。この技術は産業における自動生産ラインやロボットビジョンにおいて非常に重要であり、学術的にも画像処理における最も重要な技術の一つである。動画像符号化におけるシフト量推定やステレオビジョン等、その応用は広範囲である。本研究課題では、位相情報を用いて、対象物体の高精度なサブピクセル精度の位置・姿勢推定を行うことを目的とする。位置的精度を追求する研究は少なく、従来法と比較して1桁以上高い精度の推定を目指すことにより、画像処理分野のみならず、産業応用上極めて経済効果の高い成果が期待できる。

### 2. 研究の目的

微細基盤形成技術が必要となるフラットパネルや太陽電池を含む多くの半導体産業において、パネル貼り合わせやマスクパターンと基盤の位置合わせのために、物体検出は極めて重要である。その市場規模は2016年には4000億円にも達すると言われており、画像処理技術の中でも最も経済効果の高い技術であり、近年の重要産業を支える基盤技術である。また、近年特に注目されている拡張現実感(AR)等において、現実世界の対象物体に何らかの重畳表示を行う際に、視覚的違和感がなく安定した表示を行うためには、対象の位置・姿勢推定が重要になる。このようにサブピクセル物体検出は新たな産業への橋渡しともなる重要な技術である。大きな市場要求があるサブピクセル物体検出であるが、物体検出については多くの研究が行われてきたものの、位置・姿勢的精度を追求する技術については多くの研究がなされているとは言い難い。以下これまでに提案されてきた手法を示す。

画像空間相関値ピーク検出

周波数空間相関値ピーク検出(POC)

輝度ベース誤差最小化

特徴点ベース誤差最小化

これらの中で、のFFTを使った位相限定相関手法が精度、処理時間とも優れており注目されている。これは移動した対象(画像)のフーリエ変換が、元の対象のフーリエ変換に位相回転項が乗算された形式になる事に基づき、2つの対象の正規化相互パワースペクトルを逆フーリエ変換することにより位相限定相関関数(POC:Phase Only Correlation)が算出され、これをsinc関数に似た理想関数でフィッティングすることにより、そのピーク位置が移動量に相当する。この手法はFFTで実現できるため非常に高速に計算でき高い精度も得られるが、雑音が重畳したり、対象物体の類似性が低い場合には十分な関数フィッティングが行えず、その精度は極端に劣化する。

本研究課題では、これらの背景の基に2つの対象の正規化相互パワースペクトルが位相回転項に相当することから、この複素数から直接位相情報を取得し、その傾きを単純な最小二乗法で求めるだけで高精度な位置情報が得られることを示す。簡単な試算によるとPOCに比べて一桁高い精度の物体認識が可能となる。さらに平行移動だけでなく、回転、スケール変換を含む姿勢情報を得ることを考える。従来はlog-polar変換を行なって回転不変のRIPOC[4]による方法が示されていたが、log-polar変換はサンプリングに粗密が生じその精度は明らかに劣化する。そこで本研究課題では、トモグラフィーで用いられているラドン変換を用いる。二つの対象にラドン変換を適用し、ラドン変換の回転軸に対して1次元の提案法を適用することにより、その位相情報から高精度に回転量が求まる。この回転量を補正して縦軸を比較することにより、スケール変換係数が求まることになる。最終的には任意の姿勢情報を得るためにアフィン変換のパラメータを求めることを目的とする。更に提案課題の有効性を示すために、物体検出の典型的な例である、ステレオ視による奥行き推定と、複数の低解像度画像から高解像度画像を得るマルチフレーム超解像に適用し、より高性能な結果が得られることを示す。

### 3. 研究の方法

以下本研究課題の目的を整理する。

正規化相互パワースペクトルから位相情報を取得し、最小自乗法により傾きを求め、物体認識のための高精度な平行移動量推定を行う。

二つの対象にラドン変換を適用し、回転軸の1次元正規化相互パワースペクトルから回転量を推定し、補正後スケール係数を求めることにより姿勢推定を行う。

より一般的なアフィン変換に対して行列分解し、各行列を個別に推定、補正を繰り返すことにより、究極の位置・姿勢推定を行う。

研究課題の高精度性を検証するために、ステレオビジョンによる対象物体の奥行き推定や複数フレーム超解像に適用し、提案手法の有効性を示す。

物体検出は工業上の大きな重要テーマであり、国内外で多くの研究が行われている。しかし位置・姿勢推定についてはその重要性にも関わらず多くの研究が行われているとは言い

難しく、国外よりも国内の研究者によって活発に研究が行われているのが現状である。これは画像という限られた解像度の中で、より高精度な位置・姿勢推定は困難であり、考えられる手法にも限界があるからかもしれない。提案法は従来法を基礎にしつつ全く新しい簡単な手法を提案するものであり、ここに大きな新規性・独創性があると考えている。この着想に至ったきっかけは、代表者の研究歴に大きく拠っている。代表者の専門は信号処理、特にフィルタ設計では大きな成果を上げており、全域通過回路の位相近似問題では、FIR フィルタの最適設計に用いられている Remez アルゴリズムを適用するという革新的な方法を提案し、世界で初めて全域通過回路の位相の最適近似に成功し、顕著な成果を上げている。これを基に、2つの全域通過回路の並列回路を構成し、位相差を考慮して設計すると通過域と阻止域を最適近似できるという画期的な提案を行っており、従来から位相に着目した様々な研究を行い大きな成果を上げている。本研究課題はこのような位相情報に着目した研究の発展であり、代表者の研究経験から大きな成果を上げられると信じている。

3年の研究期間を想定しており、前半2年間で高精度な位置・姿勢推定を行い、理論と精度、処理時間等の基礎研究を行う。3年目に得られた基礎理論を使った応用研究を行い、従来法に比べ飛躍的に性能が向上することを検証する予定である。

物体検出は産業における最も重要な要素技術であり、その精度が向上し高速にできれば、様々な工業上の波及効果が期待できる。半導体産業における位置合わせ、ロボットビジョン、衛星、航空写真における位置合わせ、医療画像応用、AR等のコンピュータビジョンにおける位置合わせ等、高精度化による経済効果は計り知れず、今後の生産技術やコンピュータビジョンを根本的に変える大きな可能性を秘めている。

#### 4. 研究成果

2つの対象画像に対してフーリエ変換を行い、正規化パワースペクトルを得る。その複素数から位相を求め、傾きを求めることにより移動量が得られる。次に姿勢推定を行うためにラドン変換を行い、回転軸の一次元正規化パワースペクトルから回転量が得られ、補正後の縦軸からスケール変換係数を求める。

具体的な方法を示す前に、本研究課題における問題設定を行う。2つの座標点  $x = [x_1, x_2]^T$ ,  $y = [y_1, y_2]^T$  に対して、次式の関係性を有しているものとする。

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = s \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = Ax + t$$

2つの対象(画像)  $f(x)$ ,  $f(y)$  が与えられた時、平行移動量  $t = [t_1, t_2]^T$ 、回転角  $\theta$ 、スケール  $s$  を求めたい。現実的な問題は画像  $f(x, y)$ ,  $g(x, y)$  が与えられた時、並進、回転、スケールを精度よく求めて位置合わせすることである。

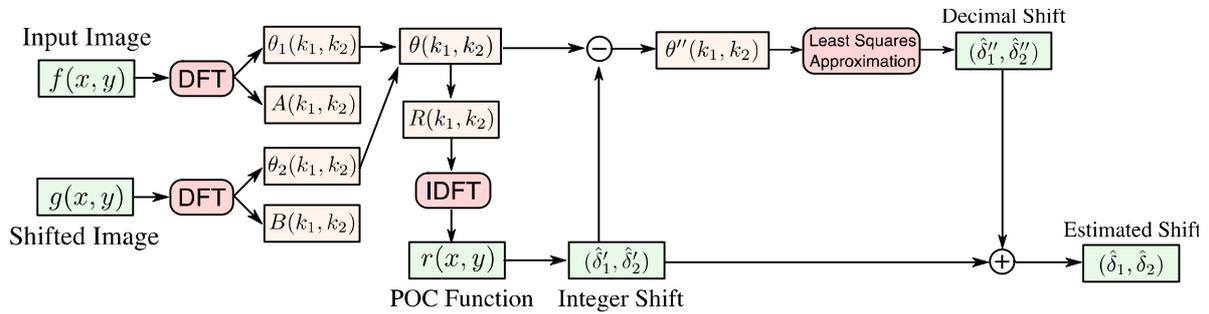
##### (1) 平行移動の推定

まず単純な問題として2つの対象が並行移動しているものとし、その移動量を求める。フーリエ変換の基礎として、 $F(\omega) = \mathfrak{F}[g(x - \tau)] = e^{-j\tau\omega} G(\omega)$  の関係が成り立つから、2つの対象の離散フーリエ変換を求め、その正規化相互パワースペクトルを

$$R(k) = \frac{F(k)\overline{G(k)}}{|F(k)\overline{G(k)}|}$$

とする。これにより位相回転項が得られるから、これを逆離散フーリエ変換して得られた  $r(n)$  を位相限定相関関数(POC)と呼び、そのピーク点が移動量に相当する。これによって整数精度の移動量が得られる。サブピクセル精度の移動量を求めるために、従来法では sinc 関数に似た理想関数を求め、 $r(n)$  に関数フィッティングして、その関数のピークを求めることにより、サブピクセル精度の移動量を推定していた。しかし対象に雑音が付加されたり、類似性が低い場合にはフィッティングそのものが困難となり、さらにフィッティングさせるために繰り返し計算が必要になり計算量は増大する。またその推定精度は一般に高くない。

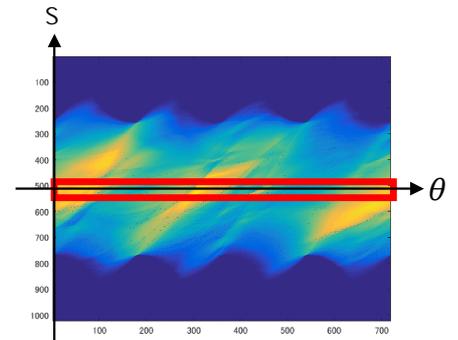
本研究課題では、正規化相互パワースペクトル  $R(k)$  が位相回転項に相当する複素数であるから、ここで直接位相を算出する。ただし位相の値域は  $-\pi \leq \theta(\omega) < \pi$  であるから、単純に計算すると不連続性が生じ、その傾きを求めることができない。これが2次元になるとさらに複雑になり、連続にするためにアンラップという操作が必要になるが、これだけで研究テーマになる程困難な問題である。これを解決するために、POC から整数精度 ( $\ell$ ) の移動量は得られるので、これに相当する位相回転項あらかじめ  $R(k)$  から除算し、少数部分の位相回転項だけから位相を算出する。周波数は  $-\pi \leq \omega < \pi$  としても、少数部分 ( $p$ ) の位相は  $-\pi \leq p\omega < \pi$  となり、煩雑なアンラップも不要になる。ここで  $p$  は1以下の少数である。結局得られた位相は理想的には周波数に対して線形となるので、単純な最小自乗法を使って  $p$  が求められ、結局  $\ell + p$  が移動量となる。この際雑音や非類似性によって線型性が崩れる可能性があるが、重要な周波数領域に重みをかけたり、外れ値を除くことによってロバストで高精度な移動量推定が可能になる。下に提案法のアルゴリズムのブロック線図を示す。簡単な試算によると POC に比べ、一桁以上高い精度の推定が可能になる。



## (2) 姿勢推定 (回転・スケール変換からアフィン変換)

次に回転・スケール変換を含む姿勢推定を行う。従来 POC の改良版として Log-Polar 変換を用いて、回転及びスケールの推定を可能にした回転不変位相限定相関(RIPOC)が提案されている。しかし Log-Polar 変換はサンプリングに粗密が生じ、この時点で誤差が生じる。従って回転・スケール量にも誤差が生じ、高精度な姿勢推定は不可能である。そこで均一なサンプリングを行うことができ、トムグラフィィーで用いられているラドン変換を適用する。ラドン変換は対象画像  $f(x, y)$  に対して

$$\mu(s, \theta) = - \int_{-} f(s \cos\theta - t \sin\theta, s \sin\theta + t \cos\theta) dt$$



で定義され、右図の結果が得られる。ラドン変換は回転角  $\theta$  に

対して周期的であり、同一の対象が回転すると、そのラドン変換は、回転量分だけ水平にシフトする。従って原理的には  $s = 0$  の横軸に対して、1次元の提案法を適用すれば、高精度な回転量推定が可能になると考えられる。次にこの回転量を補正すると、縦軸に関する相似画像が得られるから、その比がスケールに相当する。このようにラドン変換を用いることにより、よりロバストで高精度な姿勢推定が可能となる。

## (3) 全体構造と比較

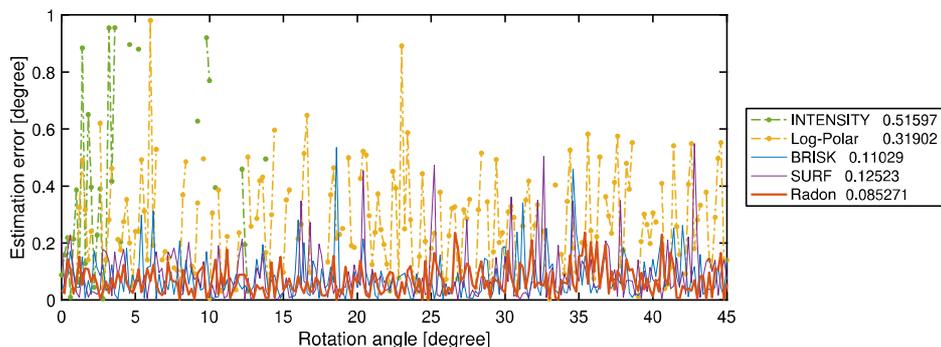
これらの成果をもとに end-to-end の処理を行うために調整を行い、最終的に以下の姿勢推定フローを構築した。

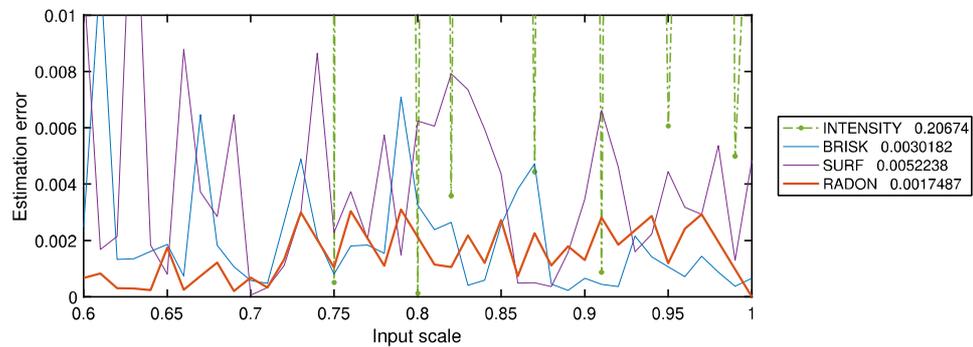
対象画像をフーリエ変換し、ラドン変換を適用することで回転量推定を行う。

対象画像そのものにラドン変換を適用し、推定回転量を考慮してスケール推定を行う。

これらのパラメータを用いて対象画像を正規化し、移動量推定を行う。

提案課題と、従来法 SURF や BRISK などの特徴ベースの方法や、極座標変換に基づく方法と性能比較を行った結果を下図に示す。下図上は回転角の推定誤差、下図下はスケールの推定誤差を示しており、推定精度、評価誤差ともに従来法を大幅に上回ったのみならず、処理速度についても 1次元処理が基本となるため、高速な処理が可能となった。





本研究成果により、サブピクセル精度の姿勢空いての必要な応用分野に大きな貢献ができるものと思われる。このように当初の目標は十分にクリアできたものと思われる。時間的な制約により、アフィン変換による任意姿勢の推定はまだ不十分であるが、本研究を利用して拡張可能であり、応用研究の開拓についても今後の課題となるが、引き続き研究を継続していきたい。

5. 主な発表論文等

〔雑誌論文〕 計2件（うち査読付論文 2件/うち国際共著 1件/うちオープンアクセス 1件）

1. 著者名 Takanori Fujisawa and Masaaki Ikehara	4. 巻 Vol.7
2. 論文標題 High-Accuracy Image Rotation and Scale Estimation Using Radon Transform and Sub-Pixel Shift Estimation	5. 発行年 2019年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 22719-22728
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

1. 著者名 Ryo Fujimoto, Takanori Fujisawa and Masaaki Ikehara	4. 巻 Vol. E101-A
2. 論文標題 Sub-Pixel Shift Estimation of Image Based on the Least Squares Approximation in Phase Region	5. 発行年 2018年
3. 雑誌名 IEICE Transactions on Fundamentals of Electronics Communications and Computer Sciences	6. 最初と最後の頁 267-272
掲載論文のDOI（デジタルオブジェクト識別子） なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

〔学会発表〕 計4件（うち招待講演 0件/うち国際学会 2件）

1. 発表者名 Yosuke Ueki, Takuro Yamaguchi and Masaaki Ikehara
2. 発表標題 Adaptive Guided Upsampling for Color Image Demosaicking
3. 学会等名 European Signal Processing Conference (EUSIPCO 2018) (国際学会)
4. 発表年 2018年

1. 発表者名 Yusuke Nakahara, Takuro Yamaguchi and Masaaki Ikehara
2. 発表標題 single image super-resolution with limited number of filters
3. 学会等名 IEEE Global Conference on Signal and Information Processing (GlobalSIP 2018) (国際学会)
4. 発表年 2018年

1. 発表者名 藤沢貴典, 藤本 遼, 池原雅章
2. 発表標題 ラドン変換と非整数シフト推定を用いた画像の回転量推定の高精度化
3. 学会等名 映像メディア処理シンポジウム
4. 発表年 2017年

1. 発表者名 Takanori Fujisawa and Masaaki Ikehara
2. 発表標題 High-Accuracy Image Rotation and Scale Estimation Using Radon Transform and Sub-Pixel Shift Estimation
3. 学会等名 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考