

**科学研究費助成事業 研究成果報告書**

平成 29 年 6 月 14 日現在

機関番号：14603

研究種目：基盤研究(C) (一般)

研究期間：2014～2016

課題番号：26330193

研究課題名(和文) 空撮画像を外部指標として用いた地上撮影動画像のカメラ位置・姿勢推定

研究課題名(英文) Camera pose estimation by using aerial images as an external reference

研究代表者

佐藤 智和 (Sato, Tomokazu)

奈良先端科学技術大学院大学・情報科学研究科・准教授

研究者番号：50362835

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：本研究では、地上を移動するカメラの位置・姿勢をビデオ撮影した映像から高精度に推定することを目的として、空撮画像を外部指標として利用する手法を開発した。従来の映像に基づくカメラ位置・姿勢推定手法は、映像中の特徴点の動き情報を利用することで、カメラの相対的な運動情報を推定することができるが、長距離の移動を伴うカメラ運動の推定時においては、誤差が蓄積するという問題があった。これに対して、本研究では、空撮画像中の特徴点と地上撮影画像中の特徴点を対応付けることで、この問題を解消する手法を開発した。開発した手法によって、空撮画像を活用することで、カメラ位置・姿勢の精度を大幅に向上可能であることを示した。

研究成果の概要(英文)：In this research, we propose a novel method for estimating camera poses from a video sequence using aerial images as an external reference. Conventionally, there was a problem of accumulative errors in camera pose estimation especially in the case that the target video is taken with long-distance movement. In order to solve this problem, we have used feature points on both video images taken in the ground and aerial images. With the experiments, we have confirmed that our method can successfully reduce the accumulation of errors and which can much increase the accuracy of the camera pose estimation.

研究分野：コンピュータビジョン、ジオメトリ

キーワード：動画像からの三次元復元 Structure from motion 空撮画像 Bundle adjustment コンピュータビジョン

### 1. 研究開始当初の背景

コンピュータビジョンの分野において、空撮画像を外部指標として利用することで地上撮影カメラの絶対位置・姿勢を推定しようという試みが成されている。これらの手法の多くは、対象となるシーンを平面と仮定することで、射影変換によって地上撮影画像の見えを空撮画像の見えに近づけた後に、一般的な特徴点対応付けを行い、これによりカメラ位置を推定している。しかし、実際の空撮画像上においては、同一の道路標示や繰り返しパターン、テクスチャが存在しない領域が多数存在し、地上撮影画像一枚からカメラ位置を一意に決定することは難しい。これに対して、移動撮影した数枚の地上撮影画像を射影変換し貼り合わせた上で、比較的広範囲の特徴点を用いて対応点探索を行うことでこの問題を解決しようという試みが成されている。しかし、射影変換を介した画像合成では蓄積誤差の問題が生じるため、誤差が顕在化しない程度の局所的な情報しか利用できず、曖昧性を解決するには不十分である。従って、上記の問題を解決するためには、より大局的な情報を利用することで、正しい対応点を決定する手法を開発する必要がある。

一方、画像中の特徴点の運動視差からカメラの位置・姿勢を推定する手法は、Structure from Motion(SfM)法と呼ばれ、近年はカメラ位置・姿勢推定において生じる蓄積誤差を解消するために、閉経路を用いる手法や、外部指標としてGPSやランドマークを用いる手法が提案されている。SfMから得られるカメラの運動は、上記の問題を解決する大局的な情報として利用できると考えられるが、そのままではやはり蓄積誤差の問題が生じるため、蓄積誤差の影響を回避しつつ上記の曖昧性を解消する手法の開発が求められる。

### 2. 研究の目的

本研究課題では、地表面において特異な特徴点が乏しい環境下においても正しい対応点を自動決定し、効果的に蓄積誤差を解消する新たなSfM法を開発する。

### 3. 研究の方法

本研究では、動画全体と外部指標である空撮画像中の特徴点位置の整合性を指標とするロバスト推定法を開発し、これを用いて正しい対応点を選別し、頑健なカメラ位置・姿勢推定を行う手法を開発する。これを実現するために、空撮画像と地上撮影画像間に検出される対応点の局所的整合性およびカメラ運動と対応点位置の大局的整合性の検証による誤対応排除手法の開発を行い、基本アルゴリズムを構築する。また、実環境において撮影した長距離の移動を伴う映像データセットを用いて手法の性能評価を行う。また、深層学習を用いてカメラ運動の推定の頑健性を向上させるための基礎的検証にも取り組む。

### 4. 研究成果

本研究では、(1)局所的な整合性の検証による対応点の誤対応の排除および、(2)大局的な整合性の検証による類似環境に対する誤対応の排除、を実現する手法を開発し、(3)これらの効果を実シーンの映像を用いて評価した。また、(4)深層学習によって特徴の乏しいシーンにおいてもカメラ運動を推定することが可能であるか検証した。

(1) 図1の上段・中段に示すように、繰り返しパターンが多く存在する環境において、一般的な対応点探索を実行すると、多数の誤対応が生じるため、そこから正しい対応点のみを抽出することは困難である。これを解決するために、我々は、局所的な整合性の検証と大局的な整合性の検証を行うことで、効果的に誤対応を排除する手法を開発した。局所的な整合性の検証においては、特徴点周辺のテクスチャから主回転方向およびテクスチャスケールの情報を抽出し、これらが2枚の画像間で整合していることをRandom Samplingベースの手法によって検証する。図1下段は、まず局所的な整合性を検証することで、画像中の誤対応を大幅に抑制した結果を示している。

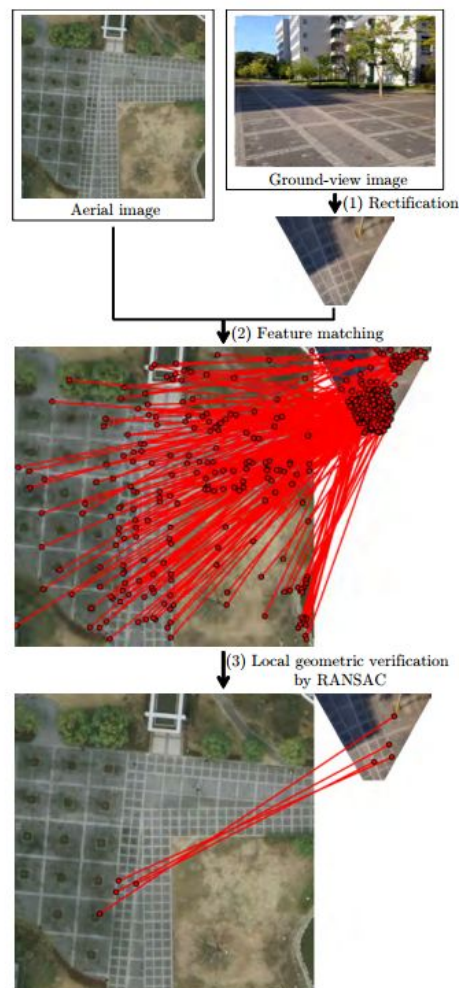


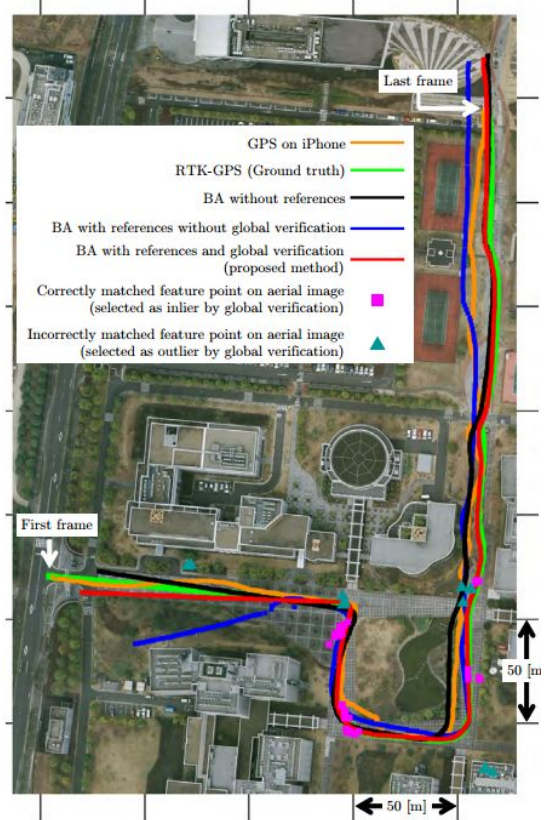
図1 局所的な整合性の検証による誤対応の排除

(2) 図 2 に示すような繰り返しパターンが多く存在する環境においては、局所的な整合性のみによる検証では、異なる地点の空撮画像と地上撮影画像が対応付けられる場合がある。このような問題を解消するために、大局的な整合性情報として、Structure from motion 法によって推定されるカメラパスと、(1)で得られた特徴点から推定されるカメラ位置が整合することを確認し、これによって誤った位置に対応付けられたカメラ位置・姿勢の推定結果を排除する(図 2)。また、これを実現するために、空撮画像中の特徴点を用いた拡張 Bundle adjustment 法を開発し、これを Random Sampling ベースの手法と組み合わせることで、誤推定結果の検出を可能とした。

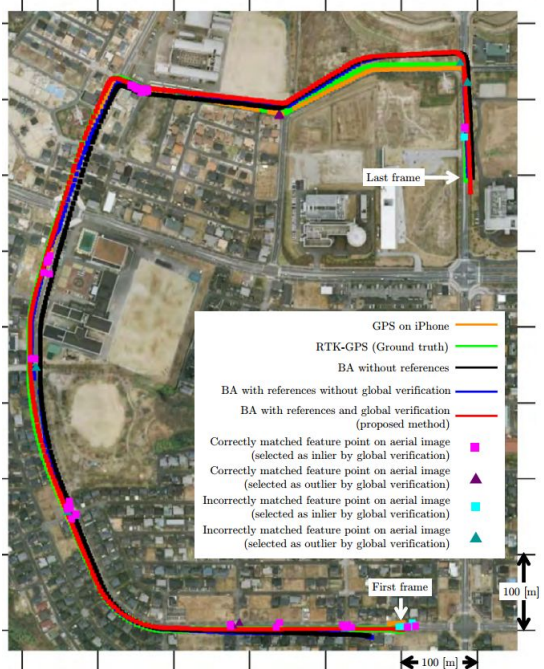


図 2 大局的な整合性の検証による誤推定結果の排除

(3) 図 3(a), (b) に示す 2 種類の屋外環境下において、開発した手法の有効性を確認した。(a) は、地表面にテクスチャの繰り返しパターンが多数存在する環境であり、市販のスマートフォンで撮影した映像を対象にカメラ位置・姿勢の推定を行った。(b) は、地表面のテクスチャが乏しい道路環境であり、車載したカメラより取得した映像を対象に推定精度の評価を行った。両環境ともに、高精度な位置計測が可能な RTK-GPS を用いてカメラ位置を計測し、これを真値として映像からのカメラ位置推定の精度を検証した。本実験では、提案手法と、空撮画像を用いない手法、提案手法に置いて大局的整合性を検証しない手法、を比較し、カメラ位置の推定精度を定量評価した(図 4)。これによって、(a), (b)、いずれの環境においても開発した手法を用いることで、空撮画像中の特徴点と地上撮影画像中の特徴点を正しく対応づけることが可能であり、対応点を得られた地点周辺においてカメラ位置の推定精度を大幅に向上させることができることを確認した。

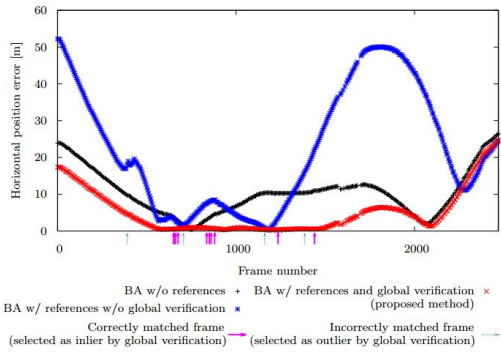


(a) 実験環境 1

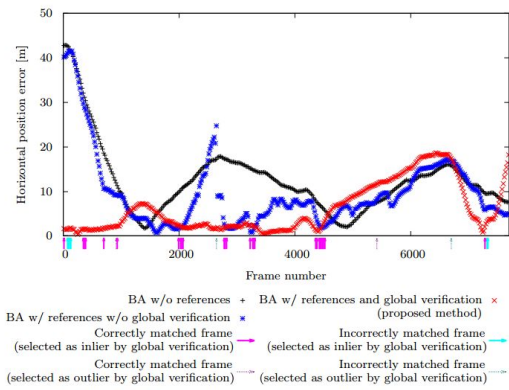


(b) 実験環境 2

図 3 実験環境および各手法によって推定されたカメラパスの比較



(a) 実験環境 1 の結果

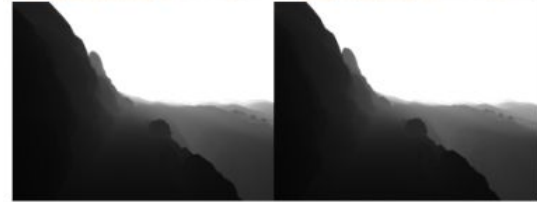


(b) 実験環境 2 の結果

**図 4 各手法に対するカメラ位置の推定誤差の比較 (赤が提案手法)**

(4) 上記の手法においても、テクスチャの乏しいシーンにおいては、地上撮影画像と空撮画像の対応づけが困難であり、対応が得られない箇所については推定精度を向上させることが難しい。これを打開するために、深層学習を用いたカメラ運動の推定によって、特徴点を直接用いることなく、頑健にカメラ運動を推定する手法を開発するための基礎的検証を行った。ここで、深層学習においては、いかに大量の学習データを収集するかが問題となるが、本研究では VR 空間内で仮想カメラを移動させ、CG により図 5 に示すような映像をレンダリングすることによって、自動で大量の学習データを生成する手法を開発した。またここでは奥行き画像も同時に生成することで、カメラ運動推定精度の向上を試みた。

開発した深層学習のフレームワークを用いることで、図 6 に示すように 2 枚の連続画像から奥行き画像を推定できることを確認した。しかし、現在までのところ単純な CNN 型ネットワークによるカメラの運動推定では、良い推定結果が得られておらず、今後初期値の工夫等によって学習を促進させ、良い結果を得る手法を開発することが必要である。



**図 5 深層学習に用いた CG レンダリング画像と対応する奥行き画像**



(a) 推定結果 (b) 真値

**図 6 深層学習によって推定された奥行き画像と真値の比較**

## 5. 主な発表論文等

### 〔雑誌論文〕(計 1 件)

H. Kume, T. Sato, and N. Yokoya: "Bundle adjustment using aerial images with two-stage geometric verification", Computer Vision and Image Understanding, Vol. 138, pp. 74-84, Sep. 2015.  
DOI: 10.1016/j.cviu.2015.05.003

### 〔学会発表〕(計 4 件)

橋岡 佳輝, 大谷 まゆ, 中島 悠太, 佐藤 智和, 横矢 直和: "DNN を用いたカメラの 6 自由度相対運動推定", 情報処理学会 研究報告, CVIM-206-13, 2017 年 3 月 9 日~3 月 10 日, 国立情報学研究所(NII), 東京都千代田区.

宮本 拓弥, 武原 光, 佐藤 智和, 河合 紀彦, 横矢 直和: "航空写真を用いた Visual SLAM の蓄積誤差軽減手法", 2016 年電子情報通信学会総合大会講演論文集, Vol. D-12-61, 2016 年 3 月 15 日, 九州大学 伊都キャンパス, 福岡県福岡市.

H. Kume, T. Sato, and N. Yokoya: "Sampling-based bundle adjustment using aerial images as external references", 画像の認識・理解シンポジウム(MIRU), SS1-29, 2014 年 7 月 29 日~7 月 31 日, 岡山コンベンションセンター, 岡山県岡山市.

糸 秀行, 佐藤 智和, 穴井 哲治, 武富 貴史, 高地 伸夫, 横矢 直和: "GPS・航空写真の併用による動画像からのカメラ位置・姿勢推定", 平成 26 年度第 6 回動体計測研究会, 2014 年 12 月 5 日, 東京大学, 東京都文京区.

### 〔図書〕(計 1 件)

糸 秀行, 佐藤 智和, 武富 貴史, 横矢 直和, 穴井 哲治, 高地 伸夫: "GPS 測位情報の併用による動画像からのカメラ位置・姿勢推定の高精度化", 画像ラボ, Vol. 25, No. 12, pp. 61-69, Dec. 2014.

〔産業財産権〕  
該当なし

〔その他〕  
ホームページ等  
<http://yokoya.naist.jp/>

## 6. 研究組織

### (1) 研究代表者

佐藤 智和 ( SATO, Tomokazu )  
奈良先端科学技術大学院大学・情報科学研究科・准教授  
研究者番号: 50362835