

平成 30 年 6 月 26 日現在

機関番号：14401

研究種目：若手研究(B)

研究期間：2016～2017

課題番号：16K16100

研究課題名(和文) 動的大規模環境における高速なカメラ位置推定手法

研究課題名(英文) A fast camera localization in large dynamic environment

研究代表者

間下 以大(Mashita, Tomohiro)

大阪大学・サイバーメディアセンター・准教授

研究者番号：00467606

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：拡張現実等やロボティクスで用いられるカメラの自己位置姿勢推定は時間等の変化によって照明条件が変化すると、精度が低下する場合がある。本研究ではコンピュータグラフィックスを用いて照明条件のシミュレーションを用い、あらかじめ様々な照明条件の画像作成し、実画像と合わせて機械学習を行うことによって照明条件の変化に頑健な手法を開発した。

研究成果の概要(英文)：Camera localization methods, which is used in augmented reality and robotics, decreases its accuracy due to variations of lighting environment. To obtain robustness for lighting variations, we utilized a lighting environment simulation based on computer graphics and applied the data generated by the simulation to machine learning. By this approach, we developed a camera localization method which is robust for variations of lighting environment.

研究分野：コンピュータビジョン

キーワード：カメラ位置姿勢推定

1. 研究開始当初の背景

カメラの画像や RGBD 画像の特徴点対応付けに基づく自己位置推定問題は拡張現実感 (Augmented Reality, AR) で重要な課題とされ、盛んに研究されている。これは自己位置情報が位置に関連した情報の提示や動作を行う基本的な情報となるためである。自己位置が必要となる環境は、屋内外を問わず様々であり、GPS を用いない手法は GPS が受信できない環境でも利用できるため、汎用性が高いことから重要な課題と言える。この問題で、屋外や人が実際に生活している空間を対象とした場合、太陽の位置や天候その他の光源環境が変化する、建物が変わったり人や車など多数の移動物体が存在したりするなどの理由で正しい対応点が得られず、その結果正しい推定が得られない、あるいは推定精度が低下する。また、対象となる環境が広大になると大規模化によって探索対象となる特徴点の数が大きくなり、処理時間の増加と推定精度の低下という問題が発生する。

このような問題に対して、申請者は「広域実環境を対象とした自己位置回復手法」や「光源環境が変化した場合の画像特徴量の評価」、「自己位置推定のための画像特徴データベースの構築」などの研究を行ってきた。これらの研究成果から、図 1 に示す様に、光源環境の変化がカメラの位置姿勢推定の性能を低下させること、その問題に対して様々な光源環境の条件を含んだ DB が有効であることがわかった。しかし、光源環境の変化に対応するためには巨大な DB が必要であり、DB サイズに比例して対応点探索の時間が増大するため、実用的でないことがわかった。そして、自己位置推定の研究では、統計的パラメトリックモデルが DB と同程度の性能を維持しつつ、計算時間の増加を抑えることがわかった。この手法は様々な条件を含む特徴点群を特徴量空間中の共分散と平均で表し、マハラノビス距離で評価することで多様な条件に対応しつつ、DB の巨大化による計算時間の増大を抑

える手法である。これによって、単純な対応点探索では入力に対して DB 内の全ての点との距離を計算する必要があるのに対して、代表点との距離計算のみで済むため、条件の変化にともなう計算量の爆発が抑えられる。しかし、計算時間については、環境の規模、すなわち特徴点の位置の数に比例して大きくなるため、更に効率的な対応点探索手法が必要である。拡張現実感における自己位置推定はトラッキングがロストしたときの初期フレームにのみ必要で、それ以降は連続的にトラッキングを継続する手法がすでに確立されているため、数秒程度で推定できれば実用的と言える。

上記研究の成果を整理すると、大規模かつ動的な環境に対応するには相応の DB が必要であり、実用的な時間で自己位置の推定を行うには、DB の巨大化によって引き起こされる計算時間と推定精度のトレードオフを解決する新たな対応点探索手法が必要と言える。大規模かつ動的な環境に対応した特徴点 DB を実環境から直接得るには、実際に様々な変化が発生している状況を撮影する必要がある。すなわち、大規模かつ長期間の計測が必要になるため現実的でない。この問題に対して、申請者の研究室ではカリフォルニア州立サンタバーバラ校の Tobias Höllerer 教授と協力し、CG モデルを利用した DB の構築に関する実験を行ってきた。この成果は研究業績 4 に含まれている。このような CG モデルを用いた物体追跡は小規模な物体で有効であること [2] が示されているが、光源環境の変化や物体の移動に対するロバスト性を向上させるために CG モデルを利用した自己位置推定手法は申請者の知る限り提案されていない。

2. 研究の目的

本研究計画では、大規模かつ光源環境や物体の位置が動的に変化する実環境での自己位置推定を実現するため、大規模動的環境に対応

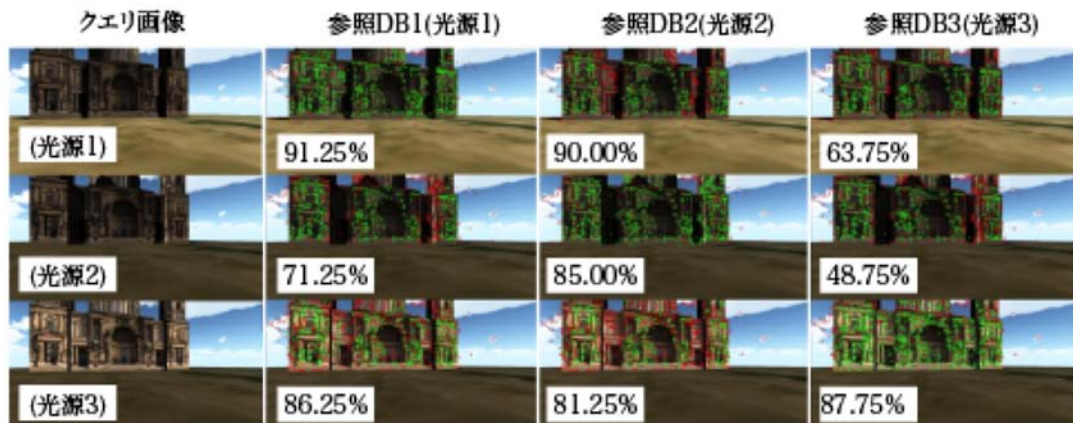


図 1 : 3 種類の光源環境によるそれぞれの DB (DB1, DB2, DB3) を参照した際の、各環境で撮影した画像について特徴点の対応付けの正誤を調べた例。緑色の点は正しく対応付けられた点、赤色の点は誤対応。

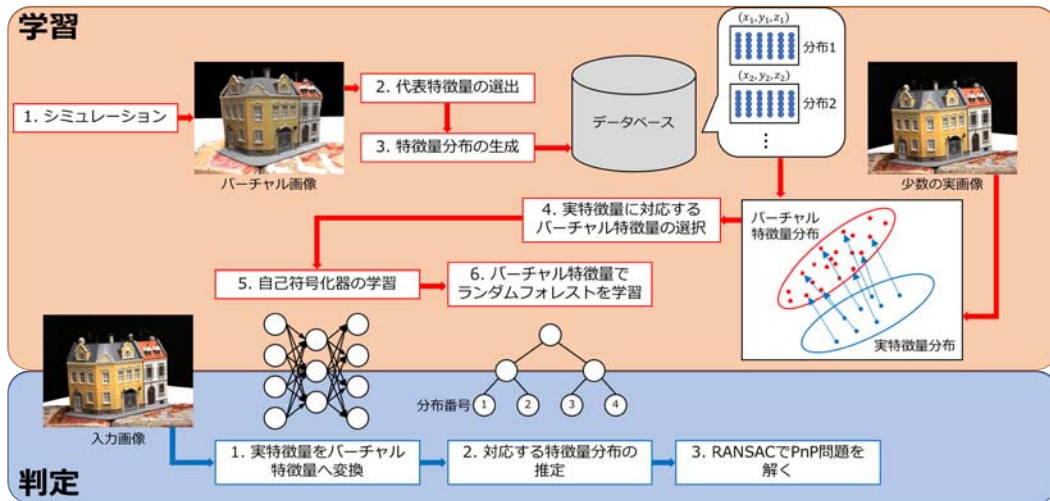


図 2：提案手法の学習時、判定時の流れ

可能な特徴点 DB の構築手法と DB の巨大化によって引き起こされる計算時間と推定精度のトレードオフを解決する新たな対応点探索手法を開発する。

3. 研究の方法

DB 構築では、最低限の実環境の計測と CG モデルを用いた環境変化のシミュレーションによって DB を構築する手法を開発する。ここでは、シミュレーションと実環境の乖離の解決が課題になると予想される。本研究ではこの乖離を実環境に対するシミュレーションの統計的バイアスとみなし、少数の実環境の計測を利用して補正する手法の開発を行う。対応点探索手法では、DB 内の特徴点の分布に対して機械学習を行うことで各特徴点の分布モデルを作成し、各特徴点の尤度を求めることで対応点を求める。

4. 研究成果

(1) パイプライン

提案手法のパイプラインを図 2 に示す。まず既存手法[3]と同様にシミュレーションを行うことで多様な光源環境下のバーチャル画像を取得する。そのバーチャル画像の特徴量の中から Kurz らの手法[4]でカメラ視点変動に頑健な代表特徴量を選出し、3次元座標が近い特徴量をグループ化して特徴量分布を作成して DB を構築する。DB の特徴量と少数の実画像から得られる特徴量を用いて自己符号化器の学習データを生成し、学習を行う。ランダムフォレストは DB の特徴量を学習データにして学習を行う。判定時は入力画像の特徴量を自己符号化器によってバーチャル画像の特徴量に変換し、ランダムフォレストで特徴量マッチングを行う。そのマッチングの中から RANSAC[5]を用いて誤対応を排除し、PnP 問題を解くことによって自己位置推定を行う。

(2) 自己符号化器の応用による乖離問題の解決

自己符号化器とは入力 x と出力 y が同じに

なるように学習された順伝搬型ニューラルネットワークである。自己符号化器は次元圧縮やディープニューラルネットワークの事前学習にも用いられる。自己符号化器の応用に入力画像に人工的にノイズを付与したり変形したりするなどの前処理を加えて、前処理をかける前の画像が出力画像に得られるように学習するデノイジング (Denoising) がある。本研究では自己符号化器のデノイジングでの使い方を参考にして、実画像から得られた特徴量をバーチャル画像から得られた特徴量に変換するように学習させる。以降、実画像の特徴量を「実特徴量」と呼び、バーチャル画像の特徴量を「バーチャル特徴量」と呼ぶ。シミュレーションで得られた特徴量分布はシミュレーション空間における 3次元座標 p_k とバーチャル特徴量列で構成される。自己符号化器はまずバーチャル特徴量同士を学習データとして事前学習させ、その後、実特徴量とそれに対応するバーチャル特徴量を学習データとして学習させる。事前学習を行うのは、バーチャル特徴量を自身へ変換させるときの誤差関数を最小にする重みと、実特徴量をバーチャル特徴量に変換させるときの誤差関数を最小にする重みは近い位置にあると予想されるからである。このように学習させると自己符号化器の出力はバーチャル特徴量と同じように扱うことができる。

(3) 実特徴量に対応するバーチャル特徴量の選択

自己符号化器の学習には実特徴量に対応するバーチャル特徴量が必要である。ここで実画像と、各実画像に対応する現実空間における 3次元座標を画像座標へ投影する投影行列が既知と仮定する。これらから実特徴量を同一 3次元点に対応するもの同士でグループ化し、バーチャル特徴量分布に対応づけられる。しかし実特徴量がどのバーチャル特徴量に真に対応するかは不明であるため、特徴量の対応付けが問題になる。単純に最近傍のバーチャル特徴量を選択すると図 3 左のように分布

の位置がずれていた場合、境界付近へ写像されるため、後段の特徴量マッチングが困難になる。そこで本手法では実特徴量分布とバーチャル特徴量に白色化を行って特徴量の対応付けを行う。白色化によって平均と分散が正規化された空間において最近傍の特徴量を求めれば、図 3 右のように各実特徴量はバーチャル特徴量分布に満遍なく写像される。さらに K 最近傍法を用いて 1 つの実特徴量につき K 個のバーチャル特徴量を対応させることで、自己符号化器の学習データをより多く確保できる。また K の値を次式のようにバーチャル特徴量分布 D_k が持つバーチャル特徴量数 n_k に応じて調整する。

$$K = \lceil r\sqrt{|D_k|} \rceil$$

ここで r は比率を表すパラメータである。前段の代表特徴量の選出により、 $|D_k|$ が大きいほどカメラ視点変動に頑健な特徴量であると言えるため、マッチングしやすい特徴量を重点的に学習することができる。

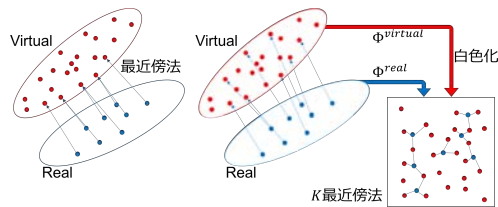


図 3：白色化による対応付け

(4) ランダムフォレストによる識別性能の向上

本手法ではランダムフォレストを用いて特徴量マッチングを行う。ユークリッド距離やマハラノビス距離を用いた特徴量マッチングと比較して、ランダムフォレストによるマッチングでは特徴量分布の形状が複雑な場合でも正しい決定境界を学習してマッチングすることが期待できる。光源変動によって特徴量分布の形状は複雑になると想定されるので、既存手法 [1] と比較して光源変動に対してより頑健になると期待できる。しかし実画像の低コントラストな領域から特徴点を検出する場合、特徴量記述子の閾値を弱く設定する必要がある。そのようにした場合 DB に存在しない特徴点も検出されるが、ランダムフォレストを用いるといずれかの特徴量分布にマッチングする。これらのマッチング結果を RANSAC へ入力しないようにするため、ランダムフォレストで得られたマッチングのうち M 個を選出する。ランダムフォレストは入力の特徴量分布に対応する確率を求めることができる。その最大尤度で全てのマッチングを降順に並び替え、上位 M 個のマッチングを採用する。

(5) 評価実験

提案手法の実特徴量をバーチャル特徴量へ変換する自己符号化器を用いた場合と、学習データに実特徴量を用いずにバーチャル特徴量を自身に変換するように学習した自己符号化

器を用いた場合で比較する。図 4 は同一画像に対する実特徴量ありの自己符号化器を用いた場合と実特徴量なしの自己符号化器を用いた場合のマッチング結果である。各図の左側が入力画像であり、右側が DB 内の特徴量分布にマッチングした位置を表している。実特徴量なしのマッチング結果を見るとほとんどのマッチングが失敗している。このことからバーチャル特徴量で学習されたランダムフォレストの決定境界で、実特徴量を正しく分割することができないことが分かる。従って、実特徴量の分布とバーチャル特徴量の分布との間にずれが存在すると予想され、そのずれを白色化と実特徴量を用いた自己符号化器によって解決することができたと考えられる。

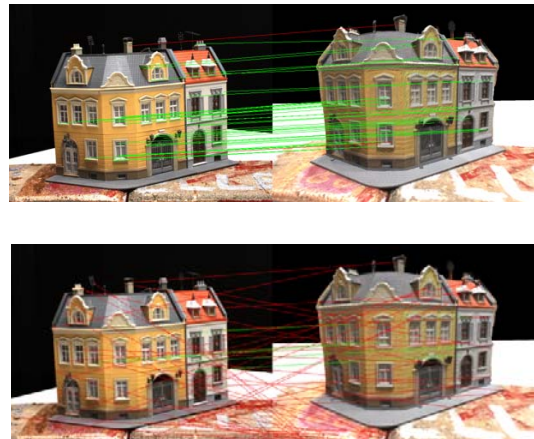


図 4：実特徴量の有無による差。(上) 実特徴量有り (下) 実特徴量無し

<引用文献>

1. C. Arth et al. "Instant Outdoor Localization and SLAM Initialization from 2.5D Maps" IEEE Trans. on Visualization and Computer Graphics, Vol. 21, No. 11, pp. 1309{1318, 2015.
2. S. Hinterstoisser et al. "Model based training, detection and pose estimation of texture-less 3D objects in heavily cluttered scenes," Asian Conf. on Computer Vision, 2012.
3. 工藤 彰, アレクサンダー プロプスキ, トビアス ヘレガー, 間下 以大, 清川 清, 竹村 治雄, "光源環境の変化にロバストな自己位置推定のための画像特徴データベースの構築", 情報処理学会研究報告, Vol. 2015-CVIM-199, No. 16, 神戸, Nov 2015.
4. D. Kurz et al., "Representative Feature Descriptor Sets for Robust Handheld Camera Localization," ISMAR, pp. 65{70, 2012.
5. M. Fischler et al., "Random Sample Consensus: A Paradigm for Model

Fitting with Applications to Image Analysis and Automated Cartography,” Communications, Vol. 24, No. 6, pp. 381-395, 1981.

5. 主な発表論文等
(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 1 件)

1. Tomohiro Mashita, Alexander Plopski, Akira Kudo, Tobias Höllerer, Kiyoshi Kiyokawa, and Haruo Takemura, “Camera Localization under a Variable Lighting Environment using Parametric Feature Database based on Lighting Simulation,” Transactions of the Virtual Reality Society of Japan, Vol. 22, No. 2, pp. 177-187, June 30, 2017.

[学会発表] (計 3 件)

1. Alexander Plopski, Tomohiro Mashita, Akira Kudo, Tobias Höllerer, Kiyoshi Kiyokawa, Haruo Takemura, “Improving Localization under Varying Illumination,” IEEE International Symposium on Mixed and Augmented Reality (ISMAR2016), Sep 2016.
2. 正満 創太, 間下 以大, Photchara Ratsamee, 浦西 友樹, 清川 清, 竹村 治雄, “シミュレーションと自己符号化器を用いた光源変化に頑健なカメラ位置姿勢推定”, 画像の認識理解シンポジウム(MIRU 2017), 広島, Aug 9, 2017
3. 間下 以大, Plopski Alexander, 工藤 彰, Höllerer Tobias, 清川 清, 竹村 治雄, “屋外拡張現実における自己位置推定のためのマハラノビス距離を用いた特徴点マッチング手法”, 第 19 回画像の認識・理解シンポジウム (MIRU2016), 浜松, Aug 2016.

6. 研究組織

(1) 研究代表者

間下 以大 (MASHITA, Tomohiro)

大阪大学・サイバーメディアセンター・准教授

研究者番号 : 00467606

(4) 研究協力者

Alexander Plopski

Tobias Höllerer