

## Visual Word を活用した自己位置推定の最適化手法の提案 An Efficient Self-Localization Method Using Visual Words

Hyun Juwon<sup>†</sup> 赤嶺 有平<sup>‡</sup> 根路銘もえ子<sup>§</sup>  
Juwon Hyun Yuhei Akamine Moeko Nerome

### 1. はじめに

拡張現実(AR)では、現実世界に仮想コンテンツを正確に重ね合わせることで、臨場感や没入感のある体験を提供することができる。また、ユーザーにより自然で快適な体験を提供するためには、現実世界に仮想コンテンツを正確な位置に重ね合わせて表示する必要がある。さらに、近年ではモバイル端末によるARコンテンツの需要が高まっており[1], 限られたバッテリーや処理性能下での効率的な自己位置推定を実現することが求められている。

そのため、GPS やビーコン方式などが使用されているが、屋内環境や周囲に高い障害物がある場合に電波干渉の影響や追加のハードウェアが必要であり精度や設置や維持管理に時間とコストがかかるという課題がある。この課題を解決するために広く利用されているのが画像ベースの自己位置推定手法である。

従来の画像ベースの自己位置推定手法におけるデータフィルタリング手法として、類似画像検索手法(以下画像検索手法)が用いられている[2]。画像検索手法の中でも、限られた資源を持つモバイルデバイス上でも処理可能にするため Bag of Visual Words (BoVW) の Visual Words (VWs) のヒストグラムによる画像検索手法が注目されている[3]。

しかし、このアプローチでは画像ごとのヒストグラム比較する必要があり、画像検索に起因する計算量増加が発生している。

本研究では、各 Visual Word (VW) に対してあらかじめ位置推定に必要な3次元特徴点座標や特徴量を付加し、入力画像から得られた VW を用いて直接 Perspective-n-Point (PnP) 問題を解く方式を提案する。これにより、中間的な画像検索処理を省略し、より高速かつ効率的な自己位置推定が可能になると期待される。

### 2. 提案手法

本研究では、従来の画像検索を伴う自己位置推定手法において発生していた中間的な処理を省略し、VW を利用して自己位置推定に必要な情報を直接取得する手法を提案する。

まず、画像データセットに対して OpenMVG の Structure from Motion (SfM) を用いて3次元構造を再構築し、データセットの各画像と3次元空間の対応関係を取得する。次に Scale Invariant Feature Transform (SIFT) を用いて各画像から特徴点および特徴量を抽出する。抽出された特徴量を、K-means クラスタリングを行い、各クラスタの中心点を VW として定義し、Codebook を作成する。

従来手法では、この Codebook を用いて画像ごとの VW ヒストグラムを生成し、それらのヒストグラム間の類似度を計算して類似画像を特定し、さらにその類似画像から自己位置推定に必要な3次元特徴点情報を絞り込む。しかし、本研究の提案手法では、**図1**のように各 VW に対して直接的に3次元特徴点座標および特徴量の情報を紐付けた VW データベ

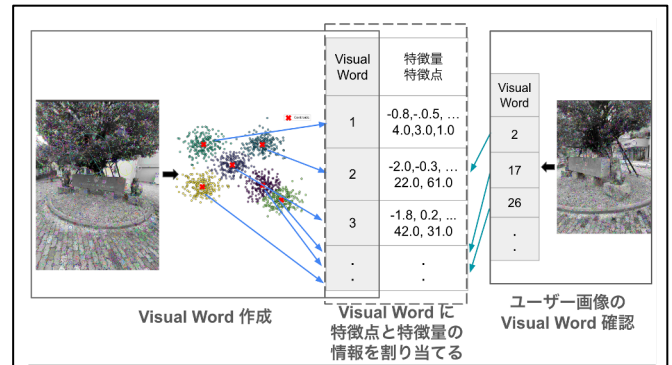


図1: Visual Word にデータ割り当て

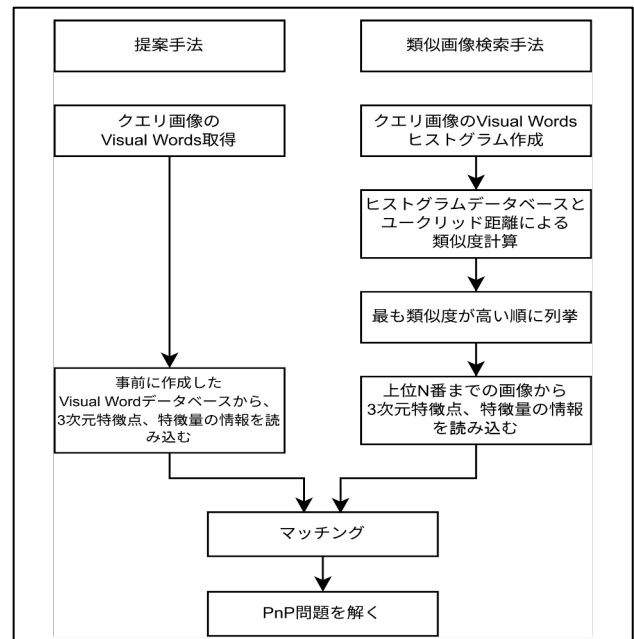


図2: 提案手法と画像検索手法の流れ

ースを構築することで、入力画像から得られた VW を用いて自己位置推定に必要な3次元情報を直接取得することが可能になる。

その後、各手法は同じように得られた特徴量を、クエリ画像の特徴量とマッチングを行い、クエリ画像の2次元特徴点に対応する3次元特徴点を絞り込む。しかし、ここには誤ったマッチング結果が含まれている可能性があるため、誤対応を除去するために Random Sample Consensus (RANSAC) を適用しながら、PnP問題を解くことで自己位置推定を行うことができる。

図2のように、提案手法は従来手法において必要であった VW ヒストグラムの生成および類似画像検索という中間プロセスを排除することで、処理の高速化および計算リソースの効率的な利用が可能になる。

<sup>†</sup> 琉球大学理工学研究科 Graduate School of Engineering and Science, Ryukyu University

<sup>‡</sup> 琉球大学工学部 Faculty of Engineering, Ryukyu University

<sup>§</sup> 沖縄国際大学経済学部 College of Economics and Environmental Policy, Okinawa International University

### 3. 実験



図 3：建物周辺の画像(上), 公園の画像(下)

本実験では, 従来の BoVW を用いた画像検索による自己位置推定手法を比較対象とし, 直接 VW にデータを割り当てることによる提案手法の処理時間および推定精度を評価した。また, 画像検索手法については検索される上位枚数を 1~10 枚に変化させた際の誤差変動および処理時間を調査し, 位置推定結果の平均値を提案手法と比較する。

モバイルデバイスにおける屋外でのリアルタイム AR コンテンツへの応用を想定し, iPhone 13 pro の広角カメラを用いて図 3 のように 2 種類のデータセット (建物周辺: 293 枚, 公園: 455 枚) を作成した。また, 実験は AMD Ryzen7 5800H, 16GB を搭載したノート PC で行った。

### 4. 実験結果

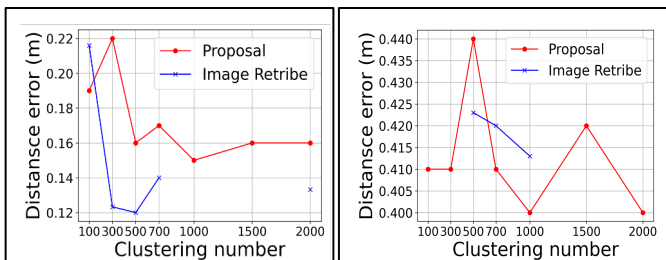


図 4：距離誤差の比較 (建物(左), 公園(右))

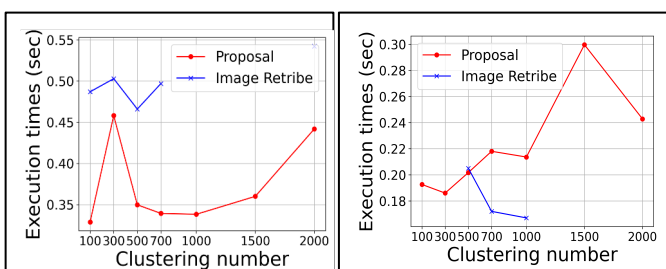


図 5：処理時間の比較 (建物(左), 公園(右))

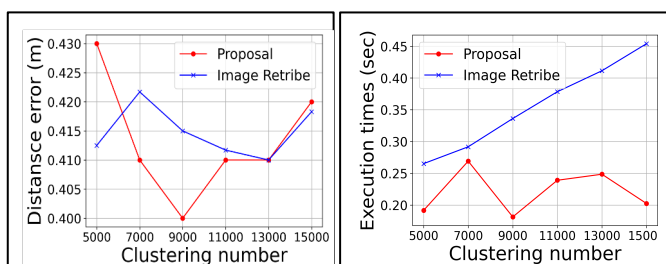


図 6：公園データセットにおける高クラスタ数設定時の距離誤差(左)および処理時間(右)の変化

実験結果は図 4 および図 5 に示す。建物周辺データセットにおいては, 提案手法の平均処理時間 0.37 秒, 平均距離誤差 0.147m であった。また, 従来手法では, 平均処理時間 0.50 秒, 平均距離誤差 0.173m であり, 平均距離誤差は 0.03m 増加したが, 平均処理時間は 26%改善された。また, 従来手法では合同一クラスタ数に設定する場合であっても, 画像検索の枚数によって自己位置推定が成功または失敗が混在していることが確認された。

公園データセットでは, クラスタ数が 2000 以下の場合, 画像検索手法は 0.18 秒を示し, 提案手法の 0.22 秒より早い, 位置推定にほとんど失敗する傾向があることがわかった。これは, クラスタ数が少ないほど処理速度は早くなるが, 検索された画像から読み込む特徴点の情報が足りないか間違っていることを示す。それに対して, 提案手法では比較的安定した処理時間および距離誤差を示すことが確認された。

多様な物体がある環境では BoVW のクラスタ数の変えることで性能向上の可能性がある [4]。そこで, クラスタ数を段階的に増加させ, 結果への影響を確認する実験を行った。その結果, 画像検索手法は, 図 6 で示したようにクラスタ数の増加につれ処理時間が増加する傾向を示している。それに対して, 提案手法ではクラスタ数が増加しても実行時間に大きな影響がない傾向を見せた。

以上より, 今回の実験での提案手法は画像検索手法に比べ安定した距離誤差を維持しながらも処理時間を改善することが確認できた。また, 多様な物体がある環境などクラスタ数を増加する必要がある場所であっても, 安定した性能が得られることがわかった。

### 5. おわりに

本研究では, モバイルデバイスを対象とした AR における自己位置推定手法として, VW に 3 次元特徴点情報を直接紐付ける手法を提案した。本実験の結果では従来の画像検索手法に比べて, 建物周辺環境での処理時間が平均で 26%改善され, 公園環境でも安定した距離誤差および処理速度を維持することが示された。また, クラスタ数が変わるとしても安定した結果が得られることがわかった。

以上により, より多様な環境での検証やクラスタ数の最適化を目的とした追加実験を行うことで, 画像検索手法の距離誤差を維持しながら, より効率的かつ安定した自己位置推定が可能になると考えられる。

#### 謝辞

本研究は JSPS 科研費 JP23K11667 の助成を受けたものです。

#### 参考文献

- [1] 米国の小売業界で普及する拡張現実 (AR) の動向。  
<https://www.jetro.go.jp/biz/areareports/2023/7d9a0c664e627a61.html>
- [2] 風間 光, 川本一彦, 岡本一志, “画像検索を用いた図書館内での自己位置推定”, 情報処理学会研究報告, Vol. 2012, CIVM-182, No. 27, 2012
- [3] 松崎 康平, 酒澤 茂之, 2014 年 情報科学技術フォーラム (FIT), 2014
- [4] 矢口 瑛貴, 久留亜沙美, 横原絵里奈, 小野 景子, “BoVW を用いた室内における空間特徴量抽出法の提案”, 第 48 回 知能システムシンポジウム, 2021