Image-Based Localization Using Context

Charbel Azzi   University of Waterloo, ON, Canada
John Zelek   University of Waterloo, ON, Canada
Daniel Asmar   American University of Beirut, Beirut, Lebanon
Adel Fakih   University of Waterloo, ON, Canada

Abstract

Image-based localization problem consists of estimating the 6 DoF camera pose by matching the image to a 3D point cloud (or equivalent) representing a 3D environment. The robustness and accuracy of current solutions is not objective and quantifiable. We have completed a comparative analysis of the main state of the art approaches, namely Brute Force Matching, Approximate Nearest Neighbour Matching, Embedded Ferns Classification, ACG Localizer(Using Visual Vocabulary) and Keyframe Matching Approach. The results of the study revealed major deficiencies in each approach mainly in search space reduction, clustering, feature matching and sensitivity to where the query image was taken. Then, we choose to focus on one common major problem that is reducing the search space. We propose to create a new image-based localization approach based on reducing the search space by using global descriptors to find candidate keyframes in the database then search against the 3D points that are only seen from these candidates using local descriptors stored in a 3D cloud map.

1 Introduction

Image Based Localization (IBL) addresses the problem of estimating the 6 DoF camera pose in an unknown environment given a query image and a representation of the scene. In the famous SLAM systems the camera pose is estimated relative to an online built 3D map where the approximate location is roughly known and it is corrected based on the last measurement by tracking which makes it prone to drift errors. In IBL there is no information about the initial location which makes IBL mainly a localization system where the camera is pose is estimated with respect to a 3D offline map where the scale is known. Thus no tracking is needed which makes IBL more resistant to drift errors when large-scale scenes are considered.

We have performed a comparative study of the main state of the art approaches, namely Brute Force Matching, Approximate Nearest Neighbour Matching [4], Embedded Ferns Classification [1], ACG Localizer(Using Visual Vocabulary) [5] and Keyframe Matching Approach [2]. The objective was to first uncover the specifics of each of these techniques and thereby understand the advantages and disadvantages of each of them. These approaches have many shortcomings in terms of accuracy and computational performance mainly in search space reduction, clustering, feature matching and sensitivity to where the query image was taken.

We focus on reducing the search space problem as mean to solve the IBL problem. Most of the work that focus on reducing the search space introduces minor contributions by trying to improve the best systems such as [5], whereas [3] tries to tackle the problem by creating a new search space system yielding a new localization system which uses MPEG descriptor to generate artificial images to cover the space.

Sattler et al. [5] is the best state of the art approach. It aimed to accelerate the Keypoint Matching step through reducing the search space by clustering features into visual words. However our comparison methodology proved that this approach has looses information due to quantization effect. We propose a new IBL system focused on solving the search space problem in two stages: (1) We start by performing a new keyframes matching approach using global image descriptors to find a constellation of keyframes in the database. (2) Then we perform a 2D-3D matching against the map's 3D points that are only seen from candidates keyframes returned from the keyframe approach.

2 Equations

Algorithm 1 presents a description of our proposed algorithm.

Algorithm 1: GIST IBL Algorithm

1. Get the GIST for each KF(Keyframe) + the 3D pts and all the Kf's each pt is visible in from VSFM map + the camera transformation estimates from VSFM
2. Take a query image Q and extract its GIST
3. **for** *all database KFs* **do**
4.   Compute the cost $C(Q, KF_i) =$ GIST distance between Q and KF_i
5.   **if** $C(Q, KF_i) < N_{(min)threshold}$ **then**
6.   Qualify KF_i for localization
7.   **else if** $C(Q, KF_i) > N_{(max)threshold}$ **then**
8.   Discard KF_i
9. ▷ **Match the query to the 3D pts coming from the qualified KFs:**
10. Take the 3D pts viewed only in the qualified KFs
11. Perform a 2D-3D match between the query and those 3D pts
12. Image Registration: Reject outliers via RANSAC and ratio test. If enough Inliers are found then Image qualifies to the Pose Estimation otherwise discard the image
13. Pose Estimation

---

3 Figures

Table 1 shows the preliminary results of our system. The testing was done on two standard datasets provided by Microsoft. Each set is composed of 4000 keyframes and 1000 query images. The results were compared to FLANN(fast approximate nearest neighbor) which is considered as ground truth to test against in IBL.

Dataset	% Error R Mean/SD		R Error (Deg) Mean/SD		T Error (Deg) Mean/SD		Time(s)	
	GIST Approach	Flann	GIST Approach	Flann	GIST Approach	Flann	GIST Approach	Flann
Chess	0.093/0.117	0.098/0.124	0.32/0.501	0.3304/0.583	0.2635/0.280	0.2769/0.355	0.059	0.14
Heads	0.156/0.157	0.163/0.194	0.229/0.237	0.233/0.241	0.729/0.764	0.746/0.792	0.036	0.116

Fig. 1: Results on Chess and Heads Datasets from Microsoft.

4 Conclusion

In this work, we completed a comparative study which revealed that IBL is still an unsolved problem. Thus we presented a new IBL system using context. The results shows that our system outperformed the standard FLANN in terms of both accuracy and computational time. We are currently working on improving our system and making it more robust by applying a new pose graph approach to it.

References

[1] Michael Donoser and Dieter Schmalstieg. Discriminative feature-to-point matching in image-based localization.

[2] Ben Glocker, Jamie Shotton, Antonio Criminisi, and Shahram Izadi. Real-time rgb-d camera relocalization via randomized ferns for keyframe encoding.

[3] Iris Heisterklaus, Ningqing Qian, and Artur Miller. Image-based pose estimation using a compact 3d model. In *Consumer Electronics??? Berlin (ICCE-Berlin), 2014 IEEE Fourth International Conference on*, pages 327–330. IEEE, 2014.

[4] Marius Muja and David G Lowe. Scalable nearest neighbor algorithms for high dimensional data. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 36(11):2227–2240, 2014.

[5] Torsten Sattler, Bastian Leibe, and Leif Kobbelt. Improving image-based localization by active correspondence search. In *Computer Vision–ECCV 2012*, pages 752–765. Springer, 2012.