Outliers in 3D Point Clouds Applied to Efficient Image-Guided Localization

Ekaterina Sirazitdinova[1], Stephan M. Jonas[1], Deyvid Kochanov[1], Jan Lensen[2], Richard Houben[2], Hans Slijp[2], Thomas M. Deserno[1]

[1]Institut für Medizinische Informatik, RWTH Aachen
[2]Applied Biomedical Systems, Maastricht, Netherlands
ekaterina.sirazitdinova@rwth-aachen.de

Abstract. In this work, the tasks of improving positioning efficiency and minimization of space requirements in image-based navigation are explored. We proved the assumption that it is possible to reduce image-matching time and to increase storage capacities by removing outliers from 3D models used for localization, by applying three outlier removal methods to our datasets and observing the localization associated with the resulting models.

1 Introduction

Nowadays, medical imaging processing is not limited to radiography or MRI but novel imaging modalities are presented including optical imaging technologies. For instance, photography is a prominent modality in wound documentation and skin lesion quantization. Furthermore, photography has been applied in other fields supporting patient. For instance, we are developing a guidance system for visually impaired and blind people that is based on optical imaging [1]. The aim is to locate a blind or visually impaired user in outdoor environments. Using structure from motion (SfM), 3D reconstructions of given tracks are created and stored in a database in the form of sparse point clouds. With a client-side App, query images are acquired and matched with the model to retrieve the precise location and orientation of the camera. High computational costs of the matching process and limited storage capacity cause the necessity of compressing 3D point clouds without loss of localization performance. It is likely that positioning accuracy can be maintained after removing outliers from 3D data.

According to the definition of Grubbs [2], an outlying observation, or outlier, is "one that appears to deviate markedly from other members of the sample in which it occurs". Outliers in a 3D point cloud may be of different nature. Firstly, they may result from errors occurring during the reconstruction process, such as inherent inaccuracies in feature detection, false matching, and errors in estimation of fundamental and projection matrices. Second, non-static environment objects (e.g., cars, chairs and tables of street cafes, advertisings, market stalls) create reconstruction noise.

In this paper, we analyse outlier removal in generated 3D point clouds for pedestrian navigation. Our hypothesis states that it is possible to maintain

H. Handels et al. (Hrsg.), *Bildverarbeitung für die Medizin 2015*, Informatik aktuell,
DOI 10.1007/978-3-662-46224-9_35,

positioning accuracy while reducing the number of outliers in a reconstructed 3D model.

2 Materials and methods

2.1 Outlier removal

We implemented outlier removal approaches of Sotoodeh [3] and Wang et al. [4].

The *density-based* approach of Sotoodeh [3] is outlier detection algorithm based on local outlier factor (LOF) [5] applied to laser point clouds. LOF is depending on the local density of the neighborhood of an object being observed. The neighborhood is defined by the distance to the k-th nearest neighbor.

The *connectivity-based* method of Wang et al. [4] is a 2-step pipeline for outlier filtering. The authors detect sparse outliers applying a scheme based on the relative density deviation of the local neighborhood and the average local neighborhood, providing a scoring strategy that includes a normalization to become independent from the specific data distribution. To remove further small dense outliers, a clustering method is used. While the density-based method runs in a linear time, the second part of the connectivity-based approach, performed by agglomerative hierarchical clustering, has the run-time complexity of $O(n^3)$.

To asses the potential of computational speedup, a *distance-based* method of outlier detection in 3D point clouds was proposed. Our approach is based on the assumption that points belonging to building wall structures are normally distributed. Thus, we apply a double-threshold scheme: firstly, we reduce the impact of infrequent points in the model, the relative distances from which to the other points in the model are comparatively big. After eliminating such points, we estimate the second filtering factor based on the global mean over mean distances of each point's neighborhood.

2.2 Dataset

Evaluation was performed on a dataset recorded at the downtown of Maastricht, the Netherlands. The dataset results from 7 walks with a recording device (iPhone 5 with acquisition application running on it) attached with a chest mount utility to the body of the person acquiring images. Within a walk an image was sequentially acquired every second. A total of 3291 images were recorded. All recordings differ in date, time and weather condition.

The route passes by several landmarks. The main characteristics of the location are a large number of pedestrians, high vehicle traffic, narrow streets and houses located close to the road.

Processing with VisualSFM [6] resulted in a dataset consisting of 17 separate models. Each model represents a reconstructed set of building walls or a single wall as a sparse 3D point cloud. The models contain from 200 to 12792 points.

2.3 Preparation of test models

To evaluate our initial hypothesis, we selected as a reference a model from our dataset allowing for the best automatic alignment to the real world coordinates. We aligned all models to the OpenStreetMap [7] inspired by the approaches of Strecha et al. [8] and Untzelmann et al. [9]. The selected model contains 11650 points and 374 cameras.

This model was then reconstructed again by 10-fold cross-validation: all images used in the reference model were randomly partitioned into 10 subsamples of equal size. For each new reconstruction, a newly selected single subsample containing 10% of original images was used as test data, the remaining 90% of images were used to reconstruct a model.

2.4 Testing process

To test the hypothesis, the following sequence of steps was applied to 8 test reconstructions:

1. Align each model to the map to estimate their scaling factors relatively to the real world coordinate system.
2. Align the test reconstruction to the reference reconstruction. For that, we apply the estimated scaling parameters to the test and the reference models. Roughly estimate translation between the models by calculating the difference between the models' centroids.
3. Refine translation and rotation by applying the Iterative Closest Point (ICP) algorithm [10].
4. Estimate a position of each test image not used for the reconstruction like it was described in [1].
5. Use the corresponding positions of the reconstructed images from the reference model to estimate the localization error of each image. The error is calculated as the distance between the estimated position and the reference position in 2D (as we localize the user in 2D, the z-component is omitted).
6. Apply the three outlier removal methods to the aligned test reconstruction. Repeat the two previous steps with the resulting models.

2.5 Performance measures

We evaluate the performance of localization distinguishing between efficiency and quality indicators.

Efficiency indicators refer to performance in terms of time and space and estimate matching time T_{m} (in seconds) and model size S_{m} (in KB) accordingly. In order to show the changes in performance caused by the application of a certain outlier removal method, we introduce the parameters for changes in matching time ΔT_{m0j} and space requirements ΔS_{m0j}, defined as

$$\Delta T_{\mathrm{m0j}} = \frac{T_{\mathrm{m0}} - T_{\mathrm{mj}}}{T_{\mathrm{m0}}} \times 100\% \tag{1}$$

$$\Delta S_{\mathrm{m0j}} = \frac{S_{\mathrm{m0}} - S_{\mathrm{mj}}}{S_{\mathrm{m0}}} \times 100\% \tag{2}$$

where $j = 1, ..., 4$ corresponds to a model in a test case. A *test case* contains four models: one model before outlier removal and three after different outlier removal methods applied.

Quality indicators, i.e. matching rate R and matching error E, describe localisation performance associated with a certain model.

Let n be a total number of test images associated with a certain tested model. Given a test image contained in the reference model, an image is considered as *matched* if it is possible to reconstruct its position p in the tested model. Accordingly, n_{m} is the total number of matched images in the model. A match is considered as a *correct* match if the positioning error, estimated as a distance between a reconstructed position p and its corresponding position p_0 in the reference model, is less than a threshold τ

$$\|p_0 - p\| < \tau \tag{3}$$

where we set $\tau = 1.6$ m (2-3 human steps).

The number of correct matches n_{c} is estimated as

$$n_c = \sum_{\mathrm{i=1}}^{n_{\mathrm{m}}} [\|p_{\mathrm{0i}} - p_{\mathrm{i}}\| < \tau] \tag{4}$$

The matching rate R is then calculated as the ratio of the number of correct matches n_{c} and the total number of images n

$$R = \frac{n_{\mathrm{c}}}{n} \times 100\% \tag{5}$$

The matching error E is the average value of all positioning errors of the correct matches

$$E = \frac{\sum_{\mathrm{i=1}}^{n_{\mathrm{m}}} \|p_{\mathrm{0i}} - p_{\mathrm{i}}\| \left(\|p_{\mathrm{0i}} - p_{\mathrm{i}}\| < \tau \right)}{n_{\mathrm{c}}} \tag{6}$$

Finally, we estimate the weighted error E_w as

$$E_{\mathrm{w}} = wE \tag{7}$$

where w is the corresponding weighting coefficient of a certain model. For each j-th model in a test case, where $j = 1, ..., 4$, the coefficient w_{j} is calculated as follows

$$w_{\mathrm{j}} = 1 - \frac{R_{\mathrm{j}} - \min\{R_1, ..., R_4\}}{100\%} \tag{8}$$

The ICP alignment of a test model to the reference model might contain an error up to 1 m. Thus, the absolute values of localization measurements might

Table 1. Results.

	Outliers removed	Benefit in computational time	Benefit in storage requirments	Loss in the accuracy of localization
	P_r	ΔT_{m0j}	ΔS_{m0j}	ΔE_{w0j} (cm)
Density-based	33.3%	31.1%	28.97%	8
Connectivity-based	20.4%	17.6%	19.24%	4
Distance-based	10.2%	8.8%	10.1%	*1*

not be precise. However, as we always use the same alignment within a test case, estimation of relative errors is possible. Thus, our final quality indicator is

$$\Delta E_{\mathrm{w0j}} = E_{\mathrm{w0}} - E_{\mathrm{wj}} \tag{9}$$

where E_{w0} is the weighed localization error associated with the reference model, and E_{wj} $(j = 1, ..., 3)$ are the corresponding weighed errors in localization using the models after the outlier removal methods applied.

We apply Student's t-test to the entire sample of positioning errors to see whether the changes in positioning performance are significant or not.

3 Results

We achieved the following performance: on average, the density-based method classified the biggest number of points (33.3% of the initial number) as outliers, while the smallest result was obtained by the distance-based method (10.2%) (Tab. 1).

In all cases the reduction of outliers leads to significant improvement in matching time T_{m} and has a positive impact on model's size S_{m}, comparing to the performance associated with a model before outlier removal. The benefits in matching time ΔT_{m0j} and storage requirements ΔS_{m0j} are proportional to the number of points P_{r} removed from the model (Tab. 1).

In the worst case, the probability to locate an image with a precision up to 1.6 m was 70%. The absolute error values were below 0.56 m for all of the cases. The average localization error resulted as the lowest (0.51 m) for our outlier removal method. At the same time, the relative weighted localization error tended to increase for the methods classifying a greater number of points as outliers. The Student's t-test resulted in the probabilities of 0.28, 0.2, 0.39 for the distance-based, the connectivity-based, and for the density-based approaches, respectively.

4 Discussion

The results have shown that image-based localization achieves a significantly higher positioning precision than the one reached by modern consumer-level

GPS sensors (34 m [11]). Average error of localization is 0.56 m including the biggest detected loss in quality of 8 cm after outlier removal. Furthermore, this value additionally accumulates an error gained in the process of alignment to the reference model, which we are unable to extract from the final result. Comparing our results to the average GPS error of 34 m, we consider the loss in quality of 8 cm as reliable and acceptable. The Student's t-test confirms our conjecture classifying those losses as insignificant. Together with the fact that the conducted experiment has shown obvious benefits of outlier removal in terms of matching time and space requirements, it makes us believe that our initial hypothesis holds.

References

1. Kochanov D, Jonas S, Hamadeh N, et al. Urban positioning using smartphone-based imaging. Proc BVM. 2014; p. 186–191.
2. Grubbs FE. Procedures for detecting outlying observations in samples. Technometrics. 1969;11(1):1–21.
3. Sotoodeh S. Outlier detection in laser scanner point clouds. In: International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences XXXVI-5; 2006. p. 297–302.
4. Wang J, Xu K, Liu L, et al. Consolidation of low-quality point clouds from outdoor scenes. Comput Graph Forum. 2013;32(5):207–10.
5. Breunig MM, Kriegel HP, Ng RT, et al. LOF: identifying density-based local outliers. Proc ACM SIGMOD. 2000; p. 93–104.
6. Wu C. Towards linear-time incremental structure from motion. Proc 3DV. 2013; p. 127–134.
7. OpenStreetMap. OpenStreetMap contributors; 2014. [Online; accessed 2014 Jul 21]. Available from: https://www.openstreetmap.org/.
8. Strecha C, Pylvänäinen T, Fua P. Dynamic and scalable large scale image reconstruction. Proc IEEE CVPR. 2010; p. 406–413.
9. Untzelmann O, Sattler T, Middelberg S, et al. A scalable collaborative online system for city reconstruction. Proc IEEE ICCVW. 2013; p. 644–651.
10. Zhang Z. Iterative point matching for registration of free-form curves and surfaces. Int J Comput Vision. 1994 Oct;13(2):119–152.
11. Modsching M, Kramer R, ten Hagen K. Field trial on GPS accuracy in a medium size city: the influence of built-up. Proc 3rd Workshop on Positioning, Navigation and Communication. 2006; p. 209–218.