The Sensitivity of RANSAC Iterations to Measurement Error and Facet Elongation

Your name 1 (student number 1) Your name 2 (student number 2)
Your name 3 (student number 3) Your name 4 (student number 4)

September 13, 2019

Abstract

The theoretical estimate on the number of iterations that RANSAC needs to find a plane with a certain probability depends on the number of inliers, the number of outliers, and this probability. It assumes that the inliers lie exactly on the plane, and hence were measured with no measurement error. In this short paper we will examine by experiments on synthetic data how the number of iterations is affected when there is measurement error. In urban reconstruction the model to be found is not a full plane but some facet on a plane, often a rectangle. With measurement error, the elongation of this facet may also affect the number of iterations, which we also examine experimentally.

1 Introduction

To determine parameters of an unknown model, a technique called RANdom SAmple Consensus (RANSAC) can sometimes be used. It takes observed data and assumes that the model parameters may already be determined from just a few observations. RANSAC is an iterative procedure that tries to find the best-fitting model by trying many, and determining the support for each model from all observations. The model (parameters) with most support is returned. RANSAC was introduced by Fischler and Bolles [1]. The particular use of RANSAC for building reconstruction was studied by Schnabel et al. [2], among others.

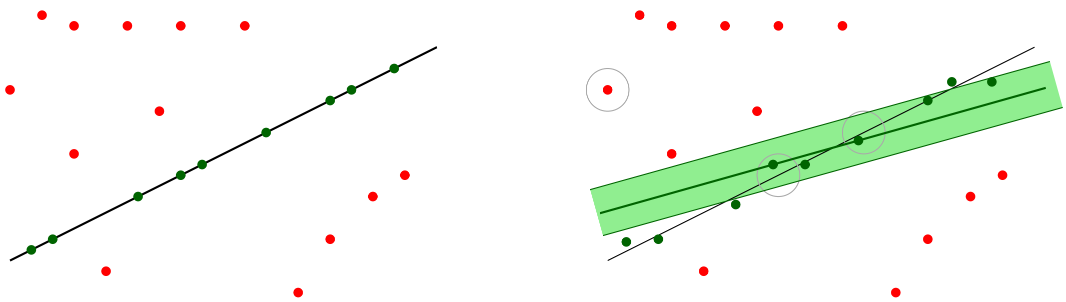

Figure 1: Left, a situation where all inliers lie on the model and any two inliers make a model that will get support of all points on the model. Right, the same situation but with 5 cm error where two inliers will not get support from several other inliers, even with a distance threshold of 5 cm.

When RANSAC is used to determine the plane in a 3D point set P that contains the maximum number of points of P, we have to set a number of iterations that makes us sufficiently certain that the best plane we have found so far is in fact the plane with the most points. If P has n points, and the number of points inside the plane with the most points is k, we call k/n the inlier ratio. To have a probability p that we found the plane with the most points after r iterations, we have the expression $p = 1 - (1 - (k/n)^3)^r$. This expression can be rewritten to determine the value of r needed in the algorithm. It is valid if there is no measurement error, but data acquired with LiDAR will typically have up to 5 cm of measurement error. Hence the expression is incorrect in practice, see Figure 1 for an illustration. We will examine how r depends on the measurement error and on the facet elongation. To this end, we will run experiments on synthetic data.

We phrase the research questions precisely:

1. Does the number of iterations required to find the best plane increase when measurement error is present, and how does it increase in the amount of error?

2. When measurement error is present, does the facet elongation influence the number of iterations that is needed, and in what way?

For both questions we assume that the probability with which we want to find the best plane is fixed to a constant. We aim to know r for $p = 0.95$.

2 Experiments: set-up and results

To make our synthetic data correspond to a possible scenario in the real World, we choose our point set P to lie in a box of $20 \times 20 \times 20$ meters. We will place $n - k$ points uniformly at random in the box, and another k points uniformly at random on the model. Then we perturb every one of the k points on the model by taking a ball with radius b around it, and randomly replacing the point inside its ball, again with uniform distribution. This models measurement error.

The set-up has four aspects that can be varied to generate the test data: the values k and n, the shape of the model, and the radius b of the ball.

2.1 Experiment 1: dependency on imprecision

Our first experiment ignores the shape aspect of the model. The model will always be a rectangle of 6 meters long and 4 meters high placed vertically in the middle of the box. We will use the following parameters:

- $n = 8000$;
- $k = 500$ or $k = 1000$;
- The radius b of the ball is 0–5 cm with a step size of 1 cm.

This leads to 12 different parameter combinations—called settings from now on—that all need to be repeated sufficiently often to get an estimate on the number of iterations r. To run RANSAC, we need one more parameter, namely the distance to the plane (model) that is allowed for a point to be counted as support for that plane. We will choose this distance to be 4 cm in all 12 settings.

Notice that we may find a plane with support higher than k. The $8000 - k$ randomly placed points could lie such that several of them support the plane as well, on top of the k points. We will say that we found the plane when we get a support of at least 90% of k, so either 450 or 900, depending on the choice of k. We can then stop this test and record the number of iterations needed.

For every setting we do the following. We run RANSAC and count how many iterations are needed until we find a plane with the desired support. We do this 200 times and then choose, as the estimate of r, the smallest number of iterations so that 190 of the runs found a plane with sufficient support within r iterations (and 10 of the runs needed more than r iterations).

Results *Perform the experiments and write this part. Choose a suitable way to present the outcome of the experiments. Make sure to summarize; a table containing* $12 \times 200 = 2400$ *values is not useful. Choose a visualization that helps to answer the research questions and is informative. The total length may be at most half a page of text to explain tables/figures/graphs/plots, and at most one page for these tables/figures/graphs/plots (and do not make them unreadably small). Some further notes: (1) Since we are not measuring time, the choice of programming language and computer is irrelevant. (2) The choice of* 10 *out of* 200 *runs taking more time comes from* $10/200 = 0.05$, *the maximum allowed probability that we do not find the plane in* r *iterations. (3) Do you understand all ways in which the* $n - k$ *outliers influence the result? (4) The choice "at least 90% of* k*" is not motivated. Probably 80% or 95% will work just as well, but we are not sure.*

2.2 Experiment 2: dependency on imprecision and model elongation

Our second experiment takes the elongation of the model into account as well. We will use the same parameters and ranges, except that we replace the fixed rectangle of 6×4 meters by a range of different models. These are:

- *Choose a range of shapes that allows you to analyze the dependency of the number of iterations needed on model elongation.*

Write the motivation for your choice of models in one or two paragraphs.

Results *Perform the experiments and write this part. Choose a suitable way to present the outcome of the experiments. Make sure to summarize. Choose a visualization that helps to answer the research questions and is informative. Total length: at most half a page of text to explain tables/figures/graphs/plots, and at most one page for these tables/figures/graphs/plots (and do not make them unreadably small).*

3 Evaluation and discussion

Write this section. Address the research questions and answer them as far as the experimental results justify this. Observe trends relating the number of iterations to the ball radius b, *and the number of iterations to the facet elongation and* b *together. Use at most half a page of text. Be succinct.*

4 Conclusions

Summarize the paper in one short paragraph, stating what was investigated and what was learned. Then write another paragraph describing what next could be done to answer the research questions better, and if applicable, give new interesting research questions that showed up because of this research.

Note: Don't change the font size or margins. If your paper is more than seven pages, you did not follow the specifications.

References

[1] Martin A. Fischler and Robert C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[2] Ruwen Schnabel, Roland Wahl, and Reinhard Klein. Efficient RANSAC for Point-Cloud Shape Detection. *Computer Graphics Forum*, 26(2):214–226, 2007.