

Automatic estimation of the inlier threshold in robust multiple structures fitting.

Roberto Toldo and Andrea Fusiello

Dipartimento di Informatica, Università di Verona
Strada Le Grazie 15, 37134 Verona, Italy

roberto.toldo@univr.it andrea.fusiello@univr.it

Abstract. This paper tackles the problem of estimating the inlier threshold in RANSAC-like approaches to multiple models fitting. An iterative approach finds the maximum of a score function which resembles the Silhouette index used in clustering validation. Although several methods have been proposed to solve this problem for the single model case, this is the first attempt to address multiple models. Experimental results demonstrate the performances of the algorithm.

1 Introduction

Fitting multiple models to noisy data is a widespread problem in Computer Vision. One of the most successful paradigm that sprouted after RANSAC is the one based on random sampling. Within this paradigm there are parametric methods (Randomized Hough Transform [1], Mean Shift [2]), and non parametric ones, (RANSAC [3], Residual histogram analysis [4], J-linkage [5]). In general the latter achieves better performances than the former and have a more general applicability, provided that the inlier threshold ϵ (also called *scale*), onto which they depend critically, is manually specified.

Some works [6–9] deal with the automatic computation of ϵ in the case of *one* model – i.e., in the case of RANSAC – but that are not extendible to the case of multiple models. In this paper we aim at filling this gap, namely at estimating ϵ when using a random sampling and residual analysis approach to fit *multiple instances* of a model to noisy data corrupted by outliers.

If ϵ is too small, noisy data points are explained by multiple similar well-fitted models, that is, the separation of the models is poor; if ϵ is too large, they are explained by a single poorly-fitted model, that is, the compactness of the model is poor. The "just right" ϵ should strike the balance between model separation and model compactness.

The degree of separation with respect to compactness is measured by a score very similar to the Silhouette index [10] for clustering validation. We compute the difference between the second and the first least model distance for each data point. The scale ϵ provides a normalizing denominator for the error difference. Consider some perturbation to the "just right" scale. As ϵ increases, a model recruits new data points, increasing the first error while decreasing the second error, causing the index to drop. As ϵ decreases, new models are created, decreasing the second error, causing the index to drop as well. Therefore, the "just right" scale is the one that yields the largest overall score from all the points.

We demonstrate our method in association with a baseline algorithm such as sequential RANSAC – sequentially apply RANSAC and remove the inliers from the data set as each model instance is detected – and with an advanced algorithm such as J-linkage [5], which have been recently demonstrated to have very good performances.

2 Fitting validation

The inlier threshold (or scale) ϵ is used to define which points belong to which model. A point belongs to a model if its distance from the model is less than ϵ . The points belonging to the same model form the *consensus set* of that model. We define the “just right” ϵ as the smallest value that yields the correct models.

In this Section we shall concentrate on a method for estimating such value based on *validating* the output of the robust fitting, which consist in a grouping of data points according to the model they belong to, plus some points (outliers) that have not been fitted. The criterion for discarding outliers is described in Sec. 3.

The validation of the fitting is based on the concepts of *compactness* and *separation*, drawn from the clustering validation literature. In particular, our method stems from the following observation (see Fig. 1):

- if ϵ is too small a single structure will be fitted by more than one model, resulting in a low separation between points belonging to different models;
- if ϵ is too large the points fitted by one model will have a low compactness (or, equivalently, a high looseness), which produces a sub-optimal estimate. As an extreme case, different structures might be fitted by a single model.

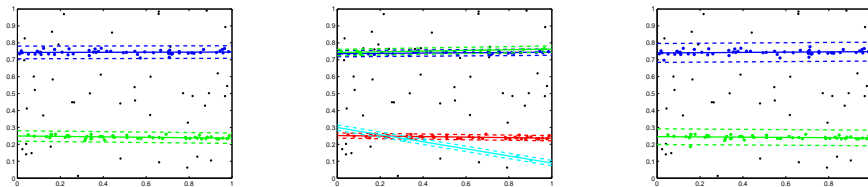


Fig. 1. From right to left: an optimal fit (according to our score), a fit with a smaller ϵ and a fit with a bigger ϵ . The dotted lines depicts the inlier band of width 2ϵ .

The “just right” ϵ has to strike the balance between these two effects. This idea is captured by the following index – which resembles the Silhouette index [10] – for a given point i :

$$r_i = \frac{b_i - a_i}{\epsilon} \quad (1)$$

where a_i is the distance of i to the model it belongs to (looseness) and b_i is the distance of i to the second closest model (separation). The global score, function of ϵ , is the average of r_i over all the points. We claim that the “just right” ϵ is the one that yields the

maximum global score. Indeed, imagine to start with the “just right” ϵ . If we decrease it, a new model is created which causes the average separation b_i to drop. If ϵ is increased such that at least one new point is added to the model, this point will increase the average looseness a_i and decrease the average b_i , resulting in a decrease of the score (see Fig. 2). A score with a similar behavior (but slightly worse performances with very small ϵ) is the ratio b_i/a_i which resembles the ratio matching used in [11].

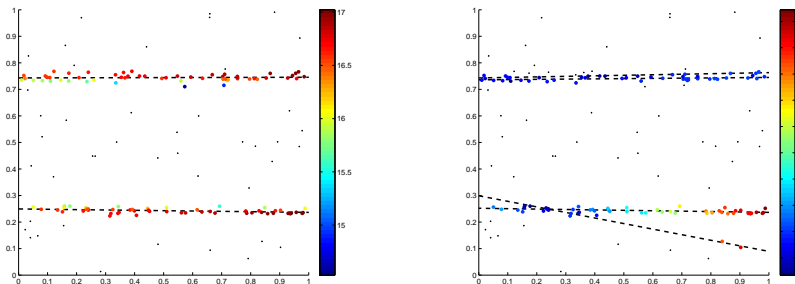


Fig. 2. This figure is best viewed in color. The left plot depicts a “good fit”, whereas the right plot depicts a “bad fit” (the upper structure is fitted by two lines very close to each other). The color of the points encodes the score values, which are consistently lower in the “bad fit”.

The algorithm for finding the optimal ϵ is iterative: the fitting algorithm is run several times with different ϵ and the value that obtained the higher score is retained.

3 Outliers rejection

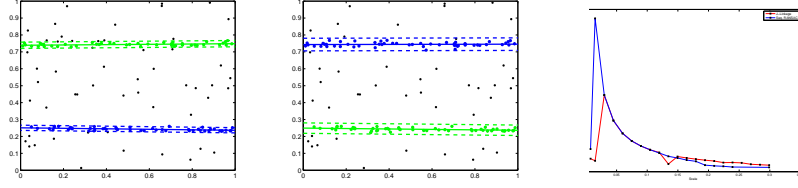
Sequential RANSAC and J-linkage – like clustering algorithms – in principle fit all the data. Bad models must be filtered out a-posteriori. Usually it is assumed that the number of models is known beforehand, but this assumption is too strong and we prefer not to rely on it. Thresholding on the cardinality of the consensus set may be an option, but this does not work if the number of data points supporting the different model instances are very different. A better approach would exploit the observation that outliers possess a diffused distribution whereas structures are concentrated. In particular, we used here a robust statistics based on points density.

Let e_i be distance of point i to its closest neighbor belonging to the same model. According to the X84 rejection rule [12], the inliers are those points such that

$$e_i < 5.2 \text{ med}_i |e_i - \text{med}_j e_j|. \quad (2)$$

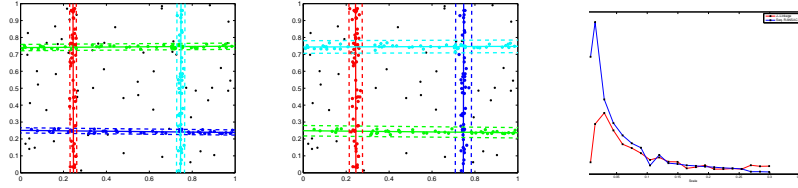
where med is the median. The models that are supported by the majority of outliers are discarded. Points that are classified as outliers but belongs to a “good” model are retained.

As the robust fit algorithm is executed repeatedly in order to estimate the best ϵ , the models that are fitted change, and hence changes the inlier/outlier classification.



(a) Seq. RANSAC optimal configuration. (b) J-Linkage optimal configuration. (c) Score for different ϵ values.

Fig. 3. Two lines example.



(a) Seq. RANSAC optimal configuration. (b) J-Linkage optimal configuration. (c) Score for different ϵ values.

Fig. 4. Four lines example.

4 Localized sampling

The assumption that outliers possess a diffused distribution implies that the a-priori probability that two points belong to the same structure is higher the smaller the distance between the points [13]. Hence minimal sample sets are constructed in a way that neighboring points are selected with higher probability. This increases the probability of selecting a set of inliers [13, 5]. Namely, if a point \mathbf{x}_i has already been selected, then \mathbf{x}_j has the following probability of being drawn:

$$P(\mathbf{x}_j|\mathbf{x}_i) = \begin{cases} \frac{1}{Z} \exp\left(-\frac{\|\mathbf{x}_j - \mathbf{x}_i\|^2}{\sigma^2}\right) & \text{if } \mathbf{x}_j \neq \mathbf{x}_i \\ 0 & \text{if } \mathbf{x}_j = \mathbf{x}_i \end{cases} \quad (3)$$

where Z is a normalization constant and σ is chosen heuristically. Given an estimate α of the average inlier-inlier distance, the value of σ is selected such that a point at a distance α has a given probability P to be selected (we set $P = 0.5$).

$$\sigma = \frac{\alpha}{\sqrt{-\log(P) - \log(Z)}}. \quad (4)$$

The value of α is iteratively estimated as the outlier/inlier classification changes.

The normalization constant can be approximated as:

$$Z \simeq (1-\delta)e^{-\frac{\omega^2}{\sigma^2}} + \delta e^{-\frac{\alpha^2}{\sigma^2}} \quad (5)$$

where ω is the average inlier-outlier distance and δ is the lowest inlier fraction among the models.

The required number of samples that give a desired confidence of success is derived in [5].

5 Summary of the method

The method that we propose can be seen as a meta-heuristic that is able to set automatically all the thresholds, perform outlier rejection and validate the final fitting. Any algorithm based on random sampling, from RANSAC to J-linkage, could fit.

The interval search for ϵ , $[\epsilon_L, \epsilon_H]$ must be specified.

Algorithm 1

1. Set $\alpha = \epsilon_H$, $\omega = 2\alpha$.
 2. for $\epsilon = \epsilon_H$ to ϵ_L
 - (a) Compute σ , using α and ω (Eq. 4);
 - (b) Run multiple-models fitting (e.g. sequential RANSAC, J-linkage) using σ for sampling and ϵ for consensus;
 - (c) Identify outliers (Sec. 3);
 - (d) Compute the average score (Eq. 1);
 - (e) Compute α, ω ;
 3. end
 4. Retain the result with the highest score.
-

6 Experiments

We tested our scale estimation method with two different multiple models fitting algorithm based on random sampling: Sequential RANSAC and J-Linkage. The scale is initialized to a huge value such that in the first step only one model arises. Subsequently the scale is decreased by a constant value. We used the same number of samples generated (10000) for all the experiments.

The examples reported consist of both synthetic and real data. In the synthetic ones (Fig. 3 and Fig. 4) we used two and four lines in the plane, respectively. The inlier points of the lines are surrounded by 50 pure outliers. Gaussian noise with standard deviation of 0.1 is added to coordinate of each inlier point.

In these experiments our meta-heuristic always proved able to converge toward a correct solution, provided that the underlying fitting algorithm (namely RANSAC or J-linkage) produced at least one correct fit.

In the real 3D data example (Fig. 5 and Fig. 6) planes are fitted to a cloud of 3D points produced by a Structure and Motion pipeline [14] fed with images of a castle and a square, respectively.

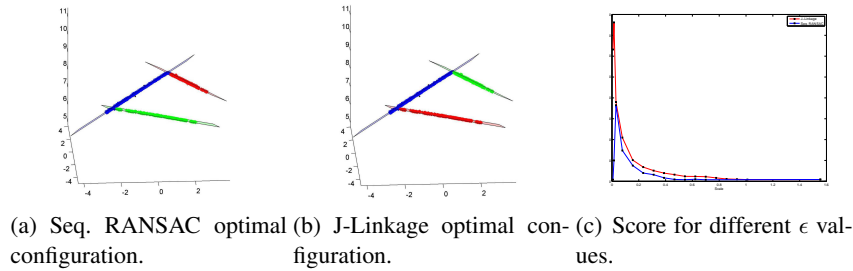


Fig. 5. Example of planes fitting in 3D from real data (castle).

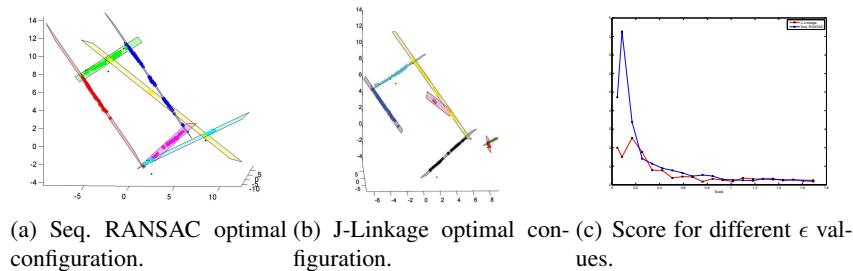


Fig. 6. Example of planes fitting in 3D from real data (square).

Finally, in the last two datasets (Fig. 7 and Fig. 8), SIFT features [11] are detected and matched in real images and homographies are fitted to the set of matches.

In the last three cases our scale estimation produced qualitatively different results when applied to sequential RANSAC or J-linkage, as not only the value of the optimal ϵ are different, but also estimated models are different. The J-linkage result is more accurate, but the output of sequential RANSAC is not completely wrong, suggesting that our meta-heuristic is able to produce a sensible result even when applied to a weak algorithm like sequential RANSAC.

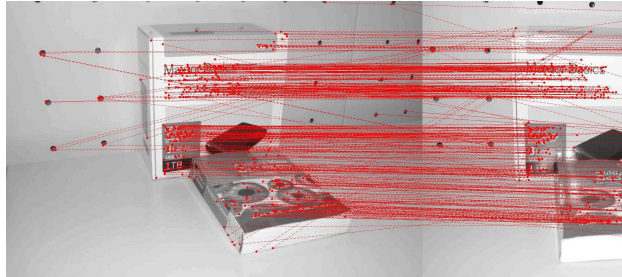
7 Conclusions

In this paper we demonstrated a meta-heuristic for scale estimation in RANSAC-like approaches to multiple models fitting. The technique is inspired by clustering validation, and in particular it is based on a score function that resembles the Silhouette index. The fitting results produced by sequential RANSAC or J-linkage with different values of scale are evaluated and the best result – according to the score – is retained. Experimental results showed the effectiveness of the approach.

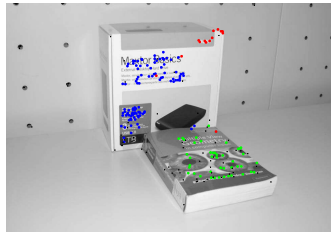
Future work will aim at improving the strategy for finding the optimal scale and smoothly cope with the special case of a single model instance.

References

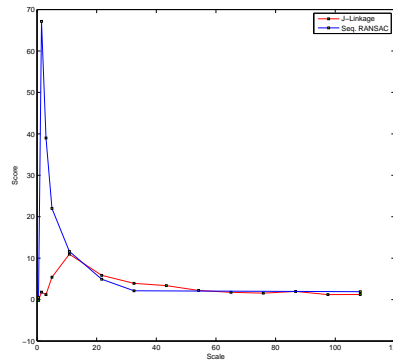
1. Xu, L., Oja, E., Kultanen, P.: A new curve detection method: randomized Hough transform (RHT). *Pattern Recognition Letters* **11**(5) (1990) 331–338
2. Subbarao, R., Meer, P.: Nonlinear mean shift for clustering over analytic manifolds. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, York, USA (2006)* 1168–1175
3. Fischler, M.A., Bolles, R.C.: Random Sample Consensus: a paradigm model fitting with applications to image analysis and automated cartography. *Communications of the ACM* **24**(6) (June 1981) 381–395
4. Zhang, W., Kosecká, J.: Nonparametric estimation of multiple structures with outliers. In: *Workshop on Dynamic Vision, European Conference on Computer Vision 2006*. Volume 4358 of *Lecture Notes in Computer Science.*, Springer (2006) 60–74
5. Toldo, R., Fusiello, A.: Robust multiple structures estimation with j-linkage. In: *ECCV (1)*. Volume 5302 of *Lecture Notes in Computer Science.*, Springer (2008) 537–547
6. Fan, L., Pylyänäinen, T.: Robust scale estimation from ensemble inlier sets for random sample consensus methods. In: *ECCV (3)*. Volume 5304 of *Lecture Notes in Computer Science.*, Springer (2008) 182–195
7. Wang, H., Suter, D.: Robust adaptive-scale parametric model estimation for computer vision. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(11) (2004) 1459–1474
8. Chen, H., Meer, P.: Robust regression with projection based m-estimators. In: *9th International Conference on Computer Vision*. (2003) 878–885
9. Torr, P.H.S., Murray, D.W.: The development and comparison of robust methods for estimating the fundamental matrix. *International Journal of Computer Vision* **24**(3) (1997) 271–300
10. Rousseeuw, P.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**(1) (1987) 53–65
11. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* **60**(2) (2004) 91–110
12. Hampel, F., Rousseeuw, P., Ronchetti, E., Stahel, W.: *Robust Statistics: the Approach Based on Influence Functions*. Wiley Series in probability and mathematical statistics. John Wiley & Sons (1986)
13. Myatt, D.R., Torr, P.H.S., Nasuto, S.J., Bishop, J.M., Craddock, R.: Napsac: High noise, high dimensional robust estimation - it's in the bag. In: *British Machine Vision Conference*. (2002)
14. Farenzena, M., Fusiello, A., Gherardi, R., Toldo, R.: Towards unsupervised reconstruction of architectural models. In Deussen, O., Saupe, D., Keim, D., eds.: *Proceedings of the Vision, Modeling, and Visualization Workshop (VMV 2008)*, Konstanz, DE, IOS Press (October 8-10 2008) 41–50



(a) SIFT Matches



(b) Seq. RANSAC optimal configuration (points membership is color coded). (c) J-Linkage optimal configuration (points membership is color coded).

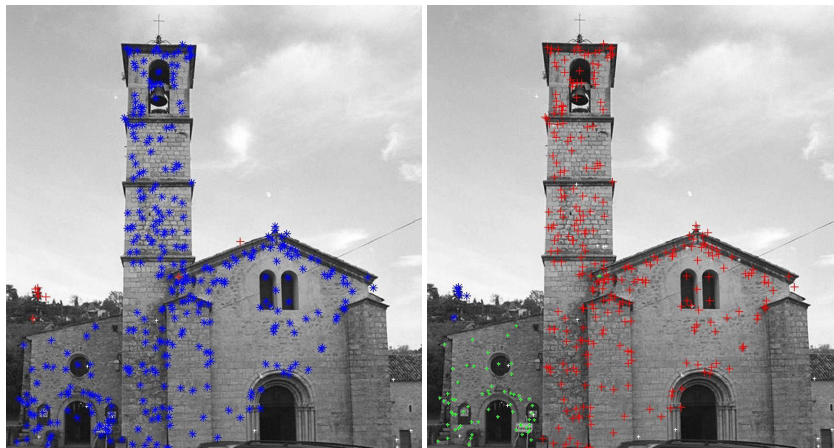


(d) Score for different ϵ values.

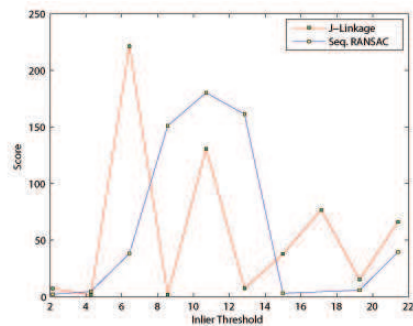
Fig. 7. Example of homography fitting (books).



(a) SIFT Matches



(b) Seq. RANSAC optimal configuration (points membership is color coded). (c) J-Linkage optimal configuration (points membership is color coded).



(d) Score for different ϵ values.

Fig. 8. Example of homography fitting (Valbonne).