

# Scale estimation in multiple models fitting via Consensus Clustering

Luca Magri<sup>1</sup> and Andrea Fusiello<sup>2</sup>

<sup>1</sup> Department of Mathematics - University of Milan  
Via Saldini, 50 - 20133 Milan, Italy

<sup>2</sup> DIEGM - University of Udine  
Via Delle Scienze, 208 - 33100 Udine, Italy

**Abstract.** This paper presents a new procedure for fitting multiple geometric structures without having a priori knowledge of scale. Our method leverages on Consensus Clustering, a single-term model selection strategy relying on the principle of stability, thereby avoiding the explicit trade-off between data fidelity (i.e., modeling error) and model complexity. In particular we tailored this model selection to the estimate of the inlier threshold of T-linkage, a fitting algorithm based on random sampling and preference analysis. A potential clustering is evaluated based on a consensus measure. The crucial inlier scale  $\epsilon$  is estimated using an interval search. Experiments on synthetic and real data show that this method succeeds in finding the correct scale.

**Keywords:** multi-model fitting, segmentation, scale estimation

## 1 Introduction

An ubiquitous issue in Pattern Recognition is the robust fitting of geometric structures to noisy data corrupted by outliers. This task becomes demanding when the data have arisen from multiple instances of the same structure, since it is necessary to cope with the so called *pseudo*-outlier (i.e. “outliers to the structure of interest but inliers to a different structure” [15]) straining robust estimation. If in addition the number of structures is not known in advance the problem of multiple fitting turns into a challenging model selection problem as we have to choose, among all the possible interpretation of the data, the most appropriate one. In general there is not a canonical way to judge the appropriateness of a model, which may have been proposed in order to resolve the problem of multiple fitting, most of which rely on the well known model selection principle based on the balance between model complexity and data fidelity.

In this paper we develop a novel technique for fitting multiple geometric structures avoiding the classical model selection trade-off of two terms in favor of a single term criterion. In particular we borrow from the *Consensus Clustering* technique [9] the idea that the stability of the clustering suffices in disambiguating the correct estimate of models. The rationale behind this method is that the “best” partition of the data is the one more stable with respect to input randomization. We translate this principle in the context of geometric fitting, tailoring

the Consensus Clustering strategy to T-Linkage, a clustering-based algorithm for fitting multiple instances of a model to noisy data possibly corrupted by outliers [7]. This gives rise to an automatic method for multiple fitting that will be shown to perform favorably on some simulated data and on some publicly available real dataset.

### 1.1 Related works

Two main approaches can be recognized that aim at fitting multiple models: non parametric methods and parametric ones. Among the non parametric Randomize Hough Transform [22], Multi-RANSAC [25] and FLoSS [6] rely on the inspecting of consensus sets of models. An alternative approach is represented by the so called *preference* analysis. Originally introduced by RHA [24], these preference-oriented methods reverse the role of data points and model typical of consensus analysis, examining the residuals of individual data points with respect to the models, in order to work in a conceptual space. J-Linkage [16], T-Linkage [7], in which points are represented as *preference functions*, [1] and QP-MF [23], where points are represented by the permutation that arrange the models in order of ascending residuals, belong to this category. Parametric methods commonly achieve better performances than non parametric ones and have a more general applicability. However their success depends critically onto the correct specification of the inlier threshold  $\epsilon$  (also called scale), which is usually manually tuned.

Since in many real applications selecting the correct scale is a hard problem, several solutions for automatic scale selection have been proposed. For example this problem is addressed in [12, 2] as regard the case of one model (i.e. in the case of RANSAC), whereas [8, 20, 4] treat the case of inlier noise estimation for multiple models exploiting elaborated robust statistic. Probably the approach of StaRSaC is the most closely related prior work; In [2] Choi and Medioni demonstrate that choosing the correct  $\epsilon$  enforces the stability of the parameter of the solution in the case of a single structure. We extend this result to the multiple structures scenario, reasoning on segmentation rather than on models parameters.

Multiple model fitting is usually dealt with using criteria of model selection. The classical model selection strategies consist in striking a good balance between fidelity to the data and model complexity (see e.g. [18]). Following the spirit of Occam's razor, all these methods result in minimizing an appropriate cost function composed by two terms: a modelling error and a penalty term for model complexity. This approach is taken also in [13, 3, 10, 11, 5] where sophisticated and effective minimization techniques such as SA-RCM [11], ARJMC [10] have been proposed.

Several alternatives have been explored for encoding model complexity. PEaRL [5] for example, optimizes a global energy function that balances geometric errors and regularity of inlier clusters, also exploiting spatial coherence. In [17], an iterative strategy for estimating the inlier-threshold, the score function, named J-Silhouette, is composed by a looseness term, dealing with fidelity, and a separation one, controlling complexity.

The idea of exploiting stability appears in the context of clustering validation. In particular in [9] the authors propose Consensus Clustering, a strategy that succeeds in estimating the number of clusters in the data with a single term model selection criterion based on stability. The next section is devoted to present the Consensus Stability approach and to summarize the T-linkage.

## 2 Background material

### 2.1 Stability and Consensus Clustering

In some cases the thorny problem of correctly tradeoff data fidelity for model complexity (a.k.a. bias-variance dilemma) can be bypassed introducing a different model selection principle based exclusively on the *stability* of models. The key idea of this method is that good models should be found among the ones that are stable with respect to small perturbations of the data. This very general principle with the necessary specifications can be applied in many contexts, and can be exploited also in the classical segmentation problem.

In [9] the authors develop this idea and present the Consensus Clustering approach to determine the correct number of clusters by maximizing the *consensus*, i.e., the agreement of clustering after perturbation of the data.

More in detail, the Consensus Clustering approach consists in assuming a clustering algorithm, for example k-means, and a resampling scheme (e.g. bootstrapping) in order to perturb the data. Then for each possible clusters number  $k = 2, 3, \dots, k_{\max}$  the data are subsampled several times and processed by the clustering algorithm. The corresponding results are described for each  $k$  by means of a *consensus matrix*  $M_k$  which is intended to capture the mutual consensus of attained clusters. The consensus matrix  $M_k$  is defined as follows: the element  $(M_k)_{ij}$  stores the number of times points  $i$  and  $j$  are assigned to the same cluster divided by the total number of times both items are selected by the resampling scheme. In other words, the consensus matrix records the proportion of clustering runs in which the two points  $i, j$  have been clustered together. For this reason  $(M_k)_{ij} \in [0, 1]$  and perfect consensus corresponds to a clean consensus matrix with all the entries equal to either 0 or 1<sup>3</sup>, whereas a deviation from this case should be explained with lack of stability of the estimated clusters. Exploiting this observation, the  $k$  that yields the cleanest consensus matrices according to an ad hoc measure, is selected as the optimal estimate of number of model.

### 2.2 T-Linkage

T-Linkage [7] is a clustering-based algorithm for fitting multiple instances of a model to noisy data (possibly) corrupted by outliers. The method is based on random sampling, and, along the same line of J-Linkage [16], follows a preference oriented approach: fixed an inlier threshold  $\epsilon$ , rather than taking models and see which points match them, T-Linkage uses the model hypotheses each point “likes better” to determine which points belong to the same cluster.

<sup>3</sup> If the data points were arranged so that points belonging to the same model are adjacent to each other, perfect consensus would translate into a block-diagonal matrix

**Choosing the scale in T-Linkage: a model selection problem.** Since T-linkage does not have any scale selection strategy, the inlier threshold  $\epsilon$  has to be manually specified by the user (as in RANSAC). If some prior knowledge of the noise in the data is available,  $\epsilon$  can be easily tuned, otherwise the scale turns out to be a free parameter onto which T-Linkage depends critically.

It is important to observe that  $\epsilon$  plays a crucial role in both the two steps of T-Linkage. At first, in conceptual representation step, the inlier threshold  $\epsilon$  explicitly defines which points belong to which model (a point belongs to a model if its distance is less than  $\epsilon$ ). If the scale is underestimated the models do not fit all their inliers; on the contrary, if the scale is overestimated, the models are affected by outliers or pseudo outliers. As regards the clustering step, points are linked together by T-Linkage until their vectorial representations are orthogonal. Here again, since  $\epsilon$  controls the orthogonality between these vectors, also the final number of models depends on this parameter.

In other words the tuning of  $\epsilon$  is a typical model selection problem: If  $\epsilon$  is too small, we are stuck in under-segmentation: multiple similar structures explain the same model in a redundant way. On the contrary, if  $\epsilon$  is too large, we run into the problem of over-segmentation obtaining fewer structures than necessary that poorly describe the data.

For these reasons by tuning the single free parameter  $\epsilon$  we are able at the same time to implicitly balance between both the complexity of the obtained models and their fidelity to the data.

In the next Sections we will show how stability can be fruitfully exploited for automatically selecting a reasonable scale.

**Refinement step and outlier rejection.** T-Linkage, as any hierarchical clustering method, is a greedy algorithm that fits models to all the data points, outliers included. In this section we propose some adjustment on T-Linkage in order to alleviate its greediness and introduce an outlier rejection criterion necessary, if outliers are present, to filter out bad models.

The problem of multiple fitting can be regarded from two alternative points of view usually coexisting: we want to faithfully segment the data and at the same time to obtain an accurate estimate of the underlying models. Each of these two tasks can not be undertaken without the other. T-Linkage concentrates on the first task segmenting the data in the conceptual space and extracting model only at the end via least-squares fitting. Once models have been obtained, we propose to perform an additional *refinement* step: points are reassigned to their nearest model – if it has distance smaller than  $\epsilon$  – and finally models are re-estimated according to this new segmentation. In this way not only the segmentation and the model estimation step can take advantages from each other, but we also gain the benefit of mitigating the greedy behavior of T-Linkage since the final clustering depends less critically on the order in which points were merged together.

As outliers are concerned, T-Linkage is agnostic about the strategy for dichotomizing inliers and outliers and an outlier rejection criterion has to be specified at the end of the algorithm. In [7] a simple strategy has been proposed, which we refine here by proposing the following criterion. The procedure starts

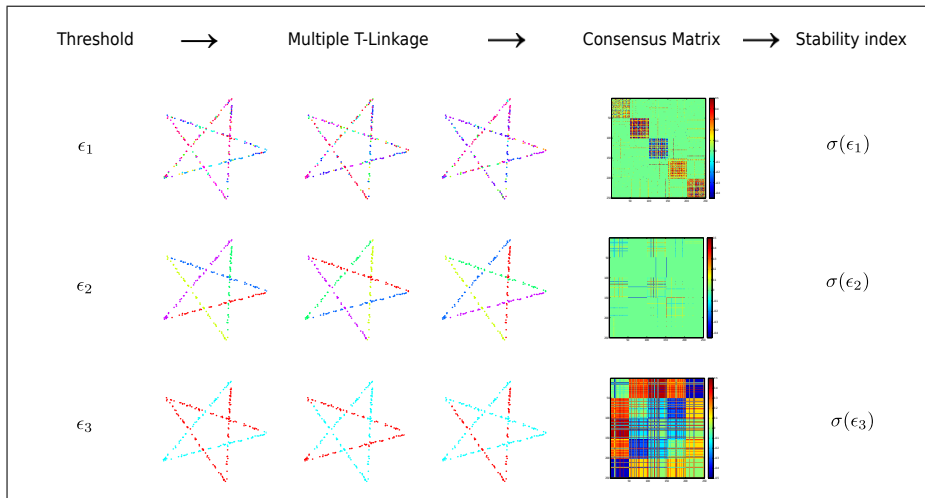


Fig. 1: The proposed method in a nutshell. Different  $\epsilon$  values are used for running multiple times T-Linkage on the perturbed data. The corresponding consensus matrices are employed for defining the stability index. The more stable clustering is selected. (Best viewed in color)

rejecting all those clusters that have less than  $|h| + 1$  elements where  $|h|$  denotes the cardinality of a minimal sample set. Then we aim at recognizing and discarding those models that “happen by chance” and do not reflect an authentic structure in the data. More precisely, under the assumption that outliers are independently distributed [14], it is possible to estimate how much it is likely that a cluster is entirely composed by outliers according to its cardinality and the model it defines. Consequently we retain only the groups with high confidence of being inliers and discard the others.

In practice at first the probability  $p$  that an outlier belongs to a specific model is estimated with Monte-Carlo simulation. Then the probability that  $k$  points belong to the same given model is computed as  $\alpha(k) = 1 - \mathcal{F}(k, n, p)$ , where  $n$  is the total number of data points, and  $\mathcal{F}$  is the binomial cumulative distribution function. For each model we compute  $k_{\min} = \alpha^{-1}(0.01)$  the minimum cardinality necessary to be not considered mere coincidence, if the considered model is supported by less than  $k_{\min}$  points is rejected as outlier.

This outlier rejection criterion differs from the one adopted in [7], since here we compute the values of  $p$  (and consequently  $k_{\min}$ ) for every specific model attained by T-Linkage at the end of the clustering, instead of estimating in advance a single probability value for a generic model. In this way our approach takes into account the fact that in general models are not all equiprobable and avoids to consider a fixed minimum cardinality and to reason about cardinality drop of clusters.

### 3 Method

In this Section we shall concentrate on a method (henceforth referred to as TLCC) for automatically fitting multiple models tailoring Consensus Clustering to T-Linkage algorithm without having a priori knowledge of the scale  $\epsilon$  conceiving a single term model selection criterion based on consensus stability.

In the case of T-Linkage we do not have to select the number of clusters (that is automatically determined by T-Linkage clustering) but we shall concentrate on the scale  $\epsilon$  which, as explained in Section 2.2, is a sensitive input parameter that implicitly tunes the balance between the complexity of the obtained clusters and their fidelity to the data.

The estimation of  $\epsilon$  is iteratively laid out as follows. At first the interval search  $[\epsilon_L, \epsilon_R]$  has to be defined, ensuring that the correct  $\epsilon$  belongs to the interval. For this reason a sound choice of  $\epsilon_L$  is a small scale value that surely over-segments the data, whereas  $\epsilon_R$  has to give rise to under-segmentation (for example it can be estimated fitting a single model to all the data point and taking the maximum of their residuals). For each  $\epsilon$  value belonging to the interval search, T-Linkage is run  $t$  times  $t = 1, \dots, t_{\max}$  on the data properly perturbed.

Rather than bootstrapping in advance the raw data as in [9], we perturb their representation in the conceptual space inside T-Linkage by bootstrapping the generated hypothesis. After the data have been processed we obtain  $t_{\max}$  clustering outputs for each  $\epsilon$  value. The intuition is that, at the correct scale, there will be consistency between the partitions produced by T-linkage. For each scale the consistency of the partitions is hence tabulated via the consensus clustering matrix  $M_\epsilon$  introduced in Section 2.1.

Now we measure the consensus stability of each matrix boiling down each  $M_\epsilon$  to a single consensus stability value  $\sigma$  per scale. If we were to plot a histogram of the entries of  $(M_\epsilon)_{ij}$ , perfect consensus would translate into two bins centered at 0 and 1 and, in general, a histogram skewed toward 0 and 1 indicates good clustering. With this idea in mind, consider the following change of variable:

$$F(x) = \begin{cases} x & \text{if } x < 0.5 \\ x - 1 & \text{if } x \geq 0.5. \end{cases} \quad (1)$$

$F$  redistributes the entries of  $M_\epsilon$  from the  $[0, 1]$  range to the interval  $[-0.5, 0.5]$ . The effect is to rearrange the histogram symmetrically around the origin. In this way stable entries are concentrated around 0 whereas unstable ones are accumulated at the tails of the histogram. For this reason measuring how far the entries of  $F(M_\epsilon)$  are spread out accounts for the consensus stability of a given scale  $\epsilon$ . For this purpose we propose to employ the variance<sup>4</sup> of the vectorized upper triangular part of  $F(M_\epsilon)$  and define a *consensus stability index* as

$$\sigma(\epsilon) = \text{Var}(\text{vech}(F(M_\epsilon))), \quad (2)$$

where  $\text{vech}$  returns the vectorization of the upper triangular matrix it receives in input. Then, assuming to deal with authentic multiple structures, the scale

<sup>4</sup> We also tested the entropy and other dispersion indices with comparable results.

is selected between the tested  $\epsilon$  that segment the data in almost two clusters. Within these  $\epsilon$  we retain as correct the smallest one obtaining the lower score of  $\sigma$ :

$$\epsilon^* = \min \left( \underset{\epsilon: \# \text{ cluster} > 1}{\text{argmin}} \sigma(\epsilon) \right). \quad (3)$$

The most stable solution (the one obtained with  $\epsilon^*$ ) is then returned.

The procedure can be summarized in Algorithm 1.

---

**Algorithm 1** TLCC
 

---

**Input:** the set of data points  $X$ ;  
 an interval search  $[\epsilon_L, \epsilon_H]$ ;  
**Output:** scale  $\epsilon^*$ ;  
 clusters of point belonging to the same model.

Generate hypotheses  $H$ ;  
**for**  $\epsilon \in [\epsilon_L, \epsilon_H]$  **do**  
  **for**  $t = 1, \dots, t_{\max}$  **do**  
     $\tilde{H} = \text{Bootstrapping}(H)$ ;  
  **end for**  
   $C_t = \text{T-Linkage}(X, \epsilon, \tilde{H})$ ;  
   $M_\epsilon = \text{Consensus Matrix}(C_1, \dots, C_{t_{\max}})$ ;  
  Compute  $\sigma(\epsilon)$ ;  
**end for**  
 $\epsilon^* = \min(\text{argmin}_{\epsilon: \# \text{ cluster} > 1} \sigma)$ ;  
 $C^* = \text{T-Linkage}(X, \epsilon^*, H)$ ;

---

As regards the computational complexity of this method, if  $c$  is the execution time of T-linkage,  $k_1$  the threshold values tested and  $k_2$  the number of bootstrapping trials, the total execution time of TLCC is  $k_1 k_2 c$  to which the time needed for computing the consensus matrices has to be added. Even if the number of bootstrap iterations is small ( $k_2 = 4$  in our experiments suffices in providing good results), there is space for improvement by replacing exhaustive search on the interval  $[\epsilon_L, \epsilon_R]$  with a suitable (direct) minimization strategy.

## 4 Experimental results

This section is devoted to evaluating the proposed method on both simulated and real data, proving that consensus stability  $\sigma$  can be exploited as a single term model selection criterion for automatically fit multiple structures. The *misclassification error* (ME) is employed for assessing clustering results, defined as the percentage of misclassified points, where a point is misclassified when it is assigned to the wrong model, according to the ground-truth.

First we compare TLCC with T-linkage + “oracle”, where the “oracle” guesses always the optimal scale according to the ME, in the interval search:

$$\epsilon_{\text{opt}} = \underset{\epsilon \in [\epsilon_L, \epsilon_R]}{\text{argmin}} \text{ME}(\epsilon), \quad (4)$$

in other words  $\epsilon_{\text{opt}}$  is the global minimum of ME. For each experiments we compare the  $\text{ME}(\epsilon^*)$  achieved by the scale  $\epsilon^*$  estimated by TLCC with the  $\text{ME}(\epsilon_{\text{opt}})$  of the *optimal* scale.

Finally we also compare indirectly TLCC with several other methods on some real datasets that have been used in the literature [21, 19].

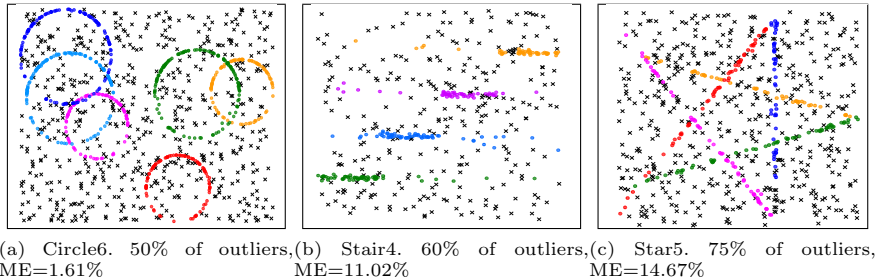


Fig. 2: *Synthetic examples*: models returned by TLCC are color coded, black crosses are outliers.

#### 4.1 Experiments with simulated data

Some synthetic experiments are carried on in order to asses the proposed approach. In particular, as shown in Fig. 2, we address the problem of fitting circles (Fig. 2a) and lines (Fig. 2b, 2c) to noisy data contaminated by gross outlier. The scale selected by TLCC always corresponds to the optimal one; the outlier rejection criterion works properly filtering out bad models with different percentage of outliers.

#### 4.2 Experiments with real data

In this section we deal with three applications of geometric multi model fitting on real data: video motion segmentation, two views motion segmentation and two views plane segmentation.

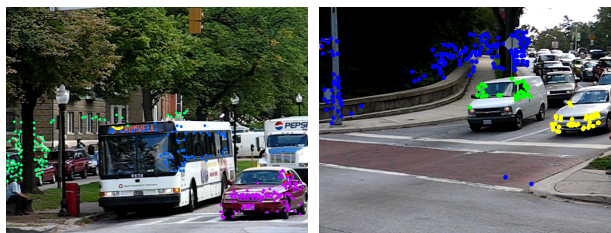
**Video motion segmentation.** In video motion segmentation the input data consist in a set of features trajectories across a video taken by a moving camera, the goal consists in recovering the multiple rigid-body motions contained in the dynamic scene.

Segmentation of motion in a video can be seen as a subspace segmentation problem under the modeling assumption of affine cameras as explained in [19]. We evaluate TLCC on the seven Traffic sequence (Fig. 2) with three motion of the Hopkins 155 motion dataset [19]. All the trajectories are inherently corrupted by noise, but no outliers are present. Our algorithm succeeds for all the sequences in estimating optimal segmentations as shown in Tab. 1, for the estimated scale always achieves the same ME as the optimal one.



Table 1: ME (%) for video **motion segmentation**.

Sequence	TLCC	Optimal
<b>cars10</b>	3.07	3.07
<b>cars2B</b>	0	0
<b>cars2_06</b>	0	0
<b>cars2_07</b>	0	0
<b>cars3</b>	0	0
<b>cars5</b>	0	0
<b>cars9</b>	0	0

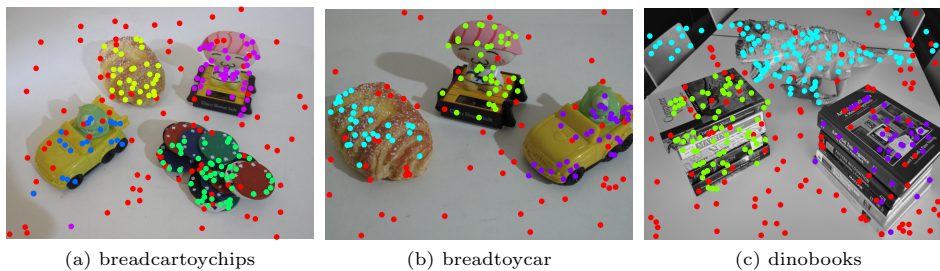


(a) Cars10

(b) Cars9

Table 2: Sample results of TLCC in video **motion segmentation** (point membership is color coded).

**Two-views segmentation.** In two views segmentation experiments, given two images of the same scene composed by several objects moving independently, the aim is to recognize and segment the motions by fitting fundamental matrices. The datasets used in these experiments consist of matching points in two uncalibrated images corrupted by gross outliers.



(a) breadcartoychips

(b) breadtoycar

(c) dinobooks

Fig. 3: Sample results of TLCC in two-view **motion segmentation** (point membership is color coded, red dots are points rejected as outliers)

We tested our method on image pairs correspondences taken from the AdelaideRMF dataset[21].

According to Tab. 3 TLCC succeeds in estimating the optimal  $\epsilon$  in six cases (marked in bold) and misses the global optimum in two cases, for which we plot the ME and the stability index in Figures 5a and 5b. It can be appreciated that the profile of the ME is fairly flat near the optimum, and that the minimum of the stability index is fairly close to the optimum of ME anyway.

Our conjecture for such a behavior is that the models have mutual intersections (or close to), and the ME does not measure properly the quality of a clustering.

Using the data reported in [11] we are able to compare indirectly TLCC with other state of the art algorithms inspired to the classical model selection approach, results are presented in Tab. 4.

In all but two cases TLCC achieves the best result, and it is the best algorithm if the mean ME is considered. These two cases are reported in Fig. 3 where it can be appreciated that the resulting segmentation is reasonable anyway.

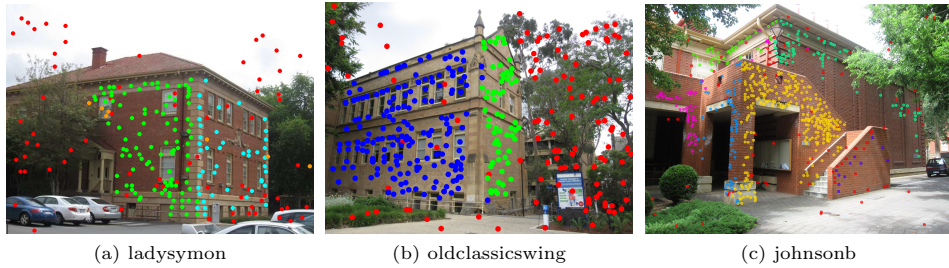


Fig. 4: Sample results of TLCC in two-view **plane segmentation** (point membership is color coded, outliers are colored in red)

Table 3: ME (%) for two-view **motion segmentation**.

Sequence	TLCC	Optimal
biscuitbookbox	2.71	0.39
<b>breadcartoychips</b>	5.19	5.19
<b>breadcubechips</b>	2.17	2.17
<b>breadtoycar</b>	4.27	4.27
<b>carchipscube</b>	1.22	1.22
cubebreadtoychips	4.46	3.50
<b>dinobooks</b>	13.86	13.86
<b>toycubecar</b>	3.03	3.03

Table 4: ME (%) comparison for two views **motion segmentation**.

	PEARL	QP-MF	FLOSS	ARJMC	SA-RCM	TLCC
biscuitbookbox	4.25	9.27	8.88	8.49	7.04	<b>2.71</b>
breadcartoychips	5.91	10.55	11.81	10.97	<b>4.81</b>	5.19
breadcubechips	4.78	9.13	10.00	7.83	7.85	<b>2.17</b>
breadtoycar	6.63	11.45	10.84	9.64	<b>3.82</b>	4.27
carchipscube	11.82	7.58	11.52	11.82	11.75	<b>1.22</b>
cubebreadtoychips	4.89	9.79	11.47	6.42	5.93	<b>4.46</b>
dinobooks	14.72	19.44	17.64	18.61	<b>8.03</b>	13.86
toycubecar	9.5	12.5	11.25	15.5	7.32	<b>3.03</b>
Mean	7.81	11.21	11.68	11.16	7.07	<b>4.62</b>

**Plane segmentation.** In the third case (*plane segmentation*) the setup is similar to the previous one: given two uncalibrated views of a scene, the aim is to recover robustly the multiplanar structures fitting homography to points correspondences. Results on the dataset taken from the AdelaideRMF dataset [21], are collected in Tab. 5.

In five cases (marked in bold) the proposed method estimate an optimal scale according to ME.

For the *johnsonb* image pairs the attained segmentation by TLCC is slightly less accurate than the optimal one, however from Fig. 5c, where the ME and the stability index are shown, it can be appreciated that the value achieved by TLCC correspond to a plateau of ME. The segmentation produced by TLCC is presented in Fig. 4. Notice that the actual global optimum of ME can be conditioned by arbitrary tie-breaking of disputed points between models.

Tab. 6 compares TLCC with state of the art methods (results for all the methods but TLCC are taken from [11]). Our method achieves in all cases, but one, the best ME and a reasonable segmentation and it scores first on the average.

In summary, results show that TLCC places in the same range as the state of the art competing algorithm adopting a classical two-term model selection strategy, with a free balancing parameter. Experiments show that this method succeeds in estimating the scale parameter of T-linkage and provide evidence

Table 5: ME (%) for **plane segmentation**.

Sequence	TLCC	Optimal
<b>johnsona</b>	3.12	3.12
johnsonb	8.33	8.81
<b>ladysymon</b>	6.17	6.17
<b>neem</b>	4.78	4.78
<b>oldclassicswing</b>	1.65	1.65
<b>sene</b>	0.42	0.42

Table 6: ME (%) comparison for **plane segmentation**.

	PEARL	QP-MF	FLOSS	ARJMC	SA-RCM	TLCC
johnsona	4.02	18.5	4.16	6.88	5.9	<b>3.12</b>
johnsonb	18.18	24.65	18.18	21.49	17.95	<b>8.33</b>
ladysymon	<b>5.49</b>	18.14	5.91	5.91	7.17	6.17
neem	5.39	31.95	5.39	8.81	5.81	<b>4.78</b>
oldclassicswing	<b>1.58</b>	13.72	1.85	1.85	2.11	<b>1.65</b>
sene	0.80	14	<b>0.80</b>	<b>0.80</b>	<b>0.80</b>	<b>0.42</b>
Mean	5.91	20.16	6.05	7.62	6.62	<b>4.08</b>

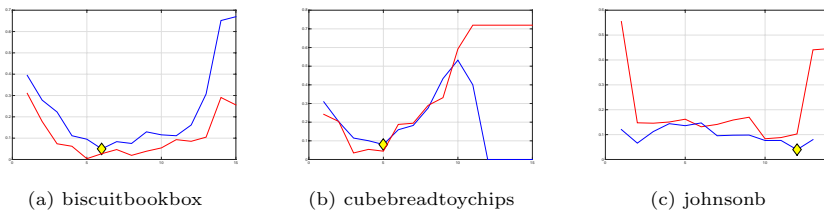


Fig. 5: Stability index  $\sigma$  (blue) and ME (red) as a function of the scale  $\epsilon$  parameter for some image pairs of the **motion segmentation** (5a, 5b) and **plane segmentation** experiments (5c). The estimated scale is marked with a diamond on the  $\sigma$  curve.

that stability has a minimum in the “right” spot, ideally the same spot where the misclassification error (ME) achieves its minimum.

## 5 Conclusion

In this paper we presented an automatic approach aimed at estimating the scale in the context of multiple structure geometric fitting. Our method exploits a single term model selection strategy relying on the principle of stability, thereby avoiding the tricky trade-off between data fidelity and model complexity. Experimental evaluation on both simulated and real data provides evidence that consensus stability succeeds in producing accurate and reliable multiple models.

*Acknowledgments* The use of the AdelaideRMF dataset<sup>5</sup> and the Hopkins 155 motion dataset<sup>6</sup> is here acknowledged.

## References

1. Chin, T., Wang, H., Suter, D.: Robust fitting of multiple structures: The statistical learning approach. In: Proc. Int. Conf. Comp. Vis. pp. 413–420 (2009)

<sup>5</sup> available online at <http://cs.adelaide.edu.au/~hwong/doku.php?id=data>

<sup>6</sup> available online at <http://www.vision.jhu.edu/data/hopkins155>

2. Choi, J., Medioni, G.G.: Starsac: Stable random sample consensus for parameter estimation. In: Proc. Conf. on Comp. Vis. and Patt. Rec. IEEE (2009)
3. DeLong, A., Veksler, O., Boykov, Y.: Fast fusion moves for multi-model estimation. In: Proc. Europ. Conf. Comp. Vis. pp. 370–384 (2012)
4. Fan, L., Pylvänäinen, T.: Robust scale estimation from ensemble inlier sets for random sample consensus methods. In: Proc. Europ. Conf. Comp. Vis. pp. 182–195 (2008)
5. Isack, H., Boykov, Y.: Enegy-based geometric multi-model fitting. *Int. J. Comp. Vis.* 97(2), 123–147 (2012)
6. Lazic, N., Givoni, I.E., Frey, B.J., Aarabi, P.: FLoSS: Facility location for subspace segmentation. In: Proc. Int. Conf. Comp. Vis. pp. 825–832 (2009)
7. Magri, L., Fusiello, A.: T-linkage: A continuous relaxation of J-linkage for multi-model fitting. In: Proc. Conf. on Comp. Vis. and Patt. Rec. (June 2014)
8. Mittal, S., Anand, S., Meer, P.: Generalized projection based m-estimator: Theory and applications. In: Proc. Conf. on Comp. Vis. and Patt. Rec. (2011)
9. Monti, S., Tamayo, P., Mesirov, J., Golub, T.: Consensus clustering: A resampling-based method for class discovery and visualization of gene expression microarray data. *Machine Learning* 52(1-2), 91–118 (2003)
10. Pham, T.T., Chin, T.J., Yu, J., Suter, D.: Simultaneous sampling and multi-structure fitting with adaptive reversible jump mcmc. In: Neural Information Processing Systems. pp. 540–548 (2011)
11. Pham, T.T., Chin, T.J., Yu, J., Suter, D.: The random cluster model for robust geometric fitting. In: Proc. Conf. on Comp. Vis. and Patt. Rec. (2012)
12. Raguram, R., Frahm, J.M.: Recon: Scale-adaptive robust estimation via residual consensus. In: International Conference on Computer Vision. pp. 1299–1306 (2011)
13. Schindler, K., Suter, D., Wang, H.: A model-selection framework for multibody structure-and-motion of image sequences. *Int. J. Comp. Vis.* 79(2), 159–177 (2008)
14. Stewart, C.V.: MINPRAN: A new robust estimator for computer vision. *IEEE Trans. on Patt. Anal. and Mach. Intell.* 17(10), 925–938 (1995)
15. Stewart, C.V.: Bias in robust estimation caused by discontinuities and multiple structures. *IEEE Trans. on Patt. Anal. and Mach. Intell.* 19(8), 818–833 (1997)
16. Toldo, R., Fusiello, A.: Robust multiple structures estimation with J-linkage. In: Proc. Europ. Conf. Comp. Vis. pp. 537–547 (2008)
17. Toldo, R., Fusiello, A.: Automatic estimation of the inlier threshold in robust multiple structures fitting. In: Proc. Int. Conf. Image An. Proc.. pp. 123–131 (2009)
18. Torr, P.H.S.: An assessment of information criteria for motion model selection. *Proc. Conf. on Comp. Vis. and Patt. Rec.* pp. 47–53 (1997)
19. Tron, R., Vidal, R.: A benchmark for the comparison of 3d motion segmentation algorithms. In: Proc. Conf. on Comp. Vis. and Patt. Rec. (2007)
20. Wang, H., Suter, D.: Robust adaptive-scale parametric model estimation for computer vision. *IEEE Trans. on Patt. Anal. and Mach. Intell.* pp. 1459–1474 (2004)
21. Wong, H.S., Chin, T.J., Yu, J., Suter, D.: Dynamic and hierarchical multi-structure geometric model fitting. In: Proc. Int. Conf. Comp. Vis. (2011)
22. Xu, L., Oja, E., Kultanen, P.: A new curve detection method: randomized Hough transform (RHT). *Patt. Rec. Lett.* 11(5), 331–338 (1990)
23. Yu, J., Chin, T., Suter, D.: A global optimization approach to robust multi-model fitting. In: Proc. Conf. on Comp. Vis. and Patt. Rec. (2011)
24. Zhang, W., Kosecká, J.: Nonparametric estimation of multiple structures with outliers. In: Proc. Europ. Conf. Comp. Vis. vol. 4358, pp. 60–74 (2006)
25. Zuliani, M., Kenney, C.S., Manjunath, B.S.: The multiRANSAC algorithm and its application to detect planar homographies. In: Proc. Int. Conf. Image Proc. (2005)