

THE MULTIRANSAC ALGORITHM AND ITS APPLICATION TO DETECT PLANAR HOMOGRAPHIES

M. Zuliani, C. S. Kenney, and B. S. Manjunath

Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA 93106.
{zuliani, kenney, manj}@ece.ucsb.edu

ABSTRACT

A RANSAC based procedure is described for detecting inliers corresponding to multiple models in a given set of data points. The algorithm we present in this paper (called multiRANSAC) on average performs better than traditional approaches based on the sequential application of a standard RANSAC algorithm followed by the removal of the detected set of inliers. We illustrate the effectiveness of our approach on a synthetic example and apply it to the problem of identifying multiple world planes in pairs of images containing dominant planar structures.

1 Introduction and Previous Work

Many image processing and computer vision tasks are carried out by fitting a set of data to some suitable model. As an example, starting from the detection of correspondent points in an image pair it is possible to estimate the scene epipolar geometry [1] and to register and fuse the image pair [2]. However the data available are usually noisy and rich of outliers (i.e. data that cannot be “explained” by the considered model and by the associated set of parameters). Robust statistics can be used in situations where the number of outliers does not exceed half of the data. An interesting application of this approach is described by Black et al. [3] to compute the optical flow. In situations where the number of outliers is larger than half of the data, one of the most commonly used approaches is RANSAC [4]. The basic idea proposed by Fischler et al. is to estimate the model parameters using the minimum number of data possible and then to check which of the remaining data points fit the model instantiated with the estimated parameters. This approach has been successfully applied to a wide variety of image processing and computer vision tasks, such as estimating the epipolar geometry of a scene, computing the camera matrices, registering image pairs and 3D data sets, etc. However this approach is limited by the assumption that a single model accounts for all of the data inliers. To handle multiple models (or different instances of the same model) it has been suggested to sequentially apply RANSAC and to remove the inliers from the data set as each model instance is detected

[5, 6] (sequential RANSAC). We believe that this approach is non-optimal. In fact we have observed in practice that inaccurate inlier detection for the initial (or subsequent) parameter estimation contributes heavily to the instability of the estimates of the parameters for the remaining models. Thus we feel that parallel rather than sequential multiple model detection is the safest approach; this is analogous to the related problem of data clustering in which standard iterative search methods such as Lloyd’s algorithm are based on parallel cluster detection.

The goal of this paper is two-fold. First we extend the RANSAC algorithm to deal simultaneously with multiple models in a more principled way, and we provide a theoretical analysis of the performance of the algorithm. Then we justify our approach performing line fitting on an “ambiguous” synthetic data set. Finally we validate the algorithm identifying sets of image correspondences that belong to distinct planar regions using real data sets. This is an important preliminary task in many image processing and computer vision applications, ranging from image registration to view synthesis, from augmented reality to autonomous navigation and so on so forth.

The paper is structured as follows: in Section 2 we will introduce the multiRANSAC algorithm, we will derive a suitable stopping criterion and we will analyze each building block of the algorithm. Section 3 will show how our approach performs on average better than sequential RANSAC when dealing with “ambiguous” data sets. Furthermore we will describe how multiRANSAC can be successfully used to detect multiple planar homographies starting from the point correspondences in an image pair. The conclusions and some ideas for future work can be found in Section 4.

2 The multiRANSAC Algorithm

We first introduce the notation that will be used throughout the rest of the paper. Let $D = \{\mathbf{x}_1, \dots, \mathbf{x}_N\}$ be the set of all data (inliers and outliers). The set of all the inliers D_I is generated by W different models, and has cardinality $N_I = N_{I,1} + \dots + N_{I,W}$. Let $\mathcal{M}_w(\boldsymbol{\theta}_w)$ denote the manifold of dimension k_w of all points associated to the parameter

vector $\theta_w \in \mathbb{R}^{k_w}$ for the specified model $1 \leq w \leq W$. A subset $s_w \subseteq D$ of k_w elements is called *minimal sample set* (MSS). We define the error for a point \mathbf{x} with respect to a manifold $\mathcal{M}_w(\theta_w)$ to be the distance from \mathbf{x} to $\mathcal{M}_w(\theta_w)$:

$$e(\mathbf{x}, \mathcal{M}_w(\theta_w)) \stackrel{\text{def}}{=} \min_{\mathbf{x}' \in \mathcal{M}_w(\theta_w)} d(\mathbf{x}, \mathbf{x}')$$

where $d(\cdot, \cdot)$ is an appropriate distance function. Using this error metric we define the *consensus set* (CS) $S(\theta_w)$ to be the subset of points in D such that their distance from the manifold $\mathcal{M}_w(\theta_w)$ is less or equal than a threshold δ . The notation $\{A_w\}$ indicates the collection of sets A_1, \dots, A_W .

2.1 MultiRANSAC Search Procedure

Suppose we want to simultaneously estimate the parameters of W models, each one represented by a k_w -dimensional parameter vector θ_w . At each iteration we draw a MSS and then we calculate the corresponding CS. We then repeat the same procedure W times after removing the inliers from the data set D . The probability of picking W MSSs entirely composed by inliers given that the size of the corresponding CSs was respectively $N_{I,1}, \dots, N_{I,W}$ is given by:

$$q = \frac{\binom{N_{I,1}}{k_1} \binom{N_{I,2}}{k_2} \dots \binom{N_{I,W}}{k_W}}{\binom{N}{k_1} \binom{N-N_{I,1}}{k_2} \dots \binom{N-\sum_{w=1}^{W-1} N_{I,w}}{k_W}} \quad (1)$$

where $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ indicates the binomial coefficient. Note that this probability is minimized if the CSs are found in increasing order of cardinality, i.e. $N_{I,1} \leq \dots \leq N_{I,W}$. The probability that a MSS does not have all its members in the inlier set is $1 - q$. If we select h MSSs at random the probability that none of them consists entirely of inliers is given by $(1 - q)^h$. This value goes to 0 as h increases. We would like to pick h large enough so that the probability of failure (by which we mean that we do not find W MSSs only composed by inliers) is less than some specified value ε . This is ensured if we require h to be large enough so that $(1 - q)^h < \varepsilon$. However there is a problem: we do not know q since a priori we do not know the true number of inliers $N_{I,w}$ for each model. This problem can be solved by observing that the cardinality of the CS associated to any parameter vector θ_w is less or equal than the actual number of inliers, i.e. $|S(\theta_w)| \leq N_{I,w}$ ($|A|$ denotes the cardinality of the set A). Therefore, provided that $N_{I,1} \leq \dots \leq N_{I,W}$, we can write¹ $q(\theta) \leq q$ and consequently $1 - q \leq 1 - q(\theta)$. In other words, if h is large enough and $(1 - q(\theta))^h < \varepsilon$, then $(1 - q)^h < \varepsilon$ as well. Thus we can define an upper bound to the maximum number of iterations using the sorted cardinalities of the largest CSs found so far. In other words the iteration threshold can be fixed to:

$$\hat{T}_{iter} = \left\lceil \frac{\log \varepsilon}{\log(1 - q(\theta))} \right\rceil \quad (2)$$

¹The parameter vector θ denotes the concatenation of the parameter vectors associated to each model.

2.2 MultiRANSAC Fusion Procedure

Even though the total number of inliers at iteration h is smaller than the total number of inliers at iteration $h - 1$, it may happen that the CSs obtained from two consecutive iterations can be combined in a compatible way (i.e. without intersections) to generate an increased number of inliers. This observation can be stated more formally as follows. Let $S_w^{(h)} = S^{(h)}(\theta_w)$ be the CS associated to the w^{th} model at the h^{th} iteration. The new collection of CSs $S^{(h)}(\theta)$ is formed combining the CSs calculated at the $(h - 1)^{th}$ iteration with those obtained at the h^{th} iteration using the following greedy algorithm:

```

UPDATECS( $\{S_w^{(h-1)}\}, \{S_w^{(h)}\}$ )
1  $S_{all} \leftarrow \{S_w^{(h-1)}\} \cup \{S_w^{(h)}\}$ 
2  $S^{(h)}(\theta) \leftarrow \emptyset$ 
3 while  $w \leq W$ 
4   do  $S^* \leftarrow \text{GETMAXCARDINALITY}(S_{all})$ 
5     if  $S^*$  is disjoint from any set in  $S^{(h)}(\theta)$ 
6       then  $S^{(h)}(\theta) \leftarrow S^{(h)}(\theta) \cup S^*$ 
7          $S_{all} \leftarrow S_{all} \setminus S^*$ 
8          $w \leftarrow w + 1$ 
9 return  $\{S_w^{(h)}\}$ 

```

The algorithm picks the largest CS among those calculated at the $(h - 1)^{th}$ and h^{th} iteration provided that such set is disjoint from those already added to $S^{(h)}(\theta)$. The function GETMAXCARDINALITY returns the set with the maximum cardinality.

2.3 The MultiRANSAC Algorithm

The multiRANSAC procedure can be summarized as follows:

```

MULTIRANSAC( $D, W, \varepsilon$ )
1  $h \leftarrow 0, \hat{T}_{iter} \leftarrow \infty$ 
2 while  $h \leq \hat{T}_{iter}$ 
3   do  $h \leftarrow h + 1$ 
4     for  $w \leftarrow 1$  to  $W$ 
5       do  $(s_w, \theta) \leftarrow \text{GETMSS}(D, W)$ 
6          $(S_w, \theta) \leftarrow \text{GETCS}(D, s_w)$ 
7          $D \leftarrow D \setminus S_w$ 
8        $(\{S_w^{(h)}\}, \theta) \leftarrow \text{UPDATECS}(\{S_w\}, \{S_w^{(h-1)}\})$ 
9        $\hat{T}_{iter} \leftarrow \frac{\log \varepsilon}{\log(1 - q(\theta))}$ 
10 return  $(\{S_w^{(h)}\}, \theta)$ 

```

The function GETMSS chooses a MSS from the data set D . The function GETCS estimates the parameter vector θ_w and calculates the correspondent CS. Finally the maximum number of iterations \hat{T}_{iter} is updated according to the criterion outlined in Section 2.1.

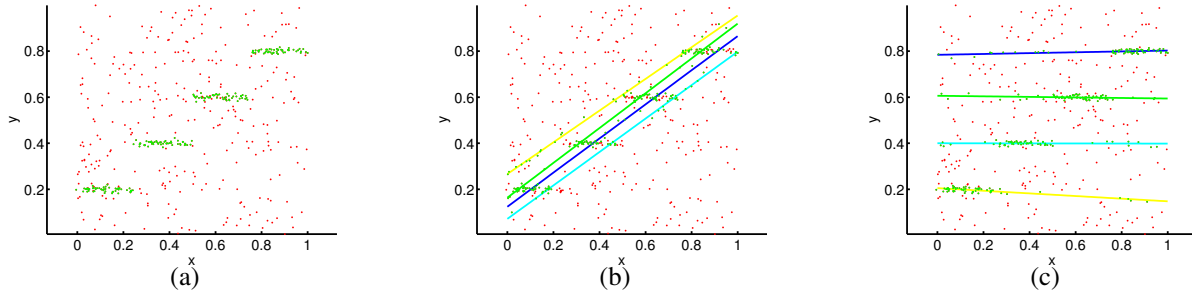


Fig. 1. This figure is best viewed in color. (a) The *stair* data set is composed by four horizontal segments with 50 points each and 300 true outliers. (b) An example of bad fit and (c) an example of good fit.

3 Experiments

3.1 Detecting Lines

The purpose of this subsection is to support experimentally our claim that inaccurate inlier detection for the initial (or subsequent) parameter estimation contributes heavily to the instability of the estimates of the parameters for the remaining models. Consider the *stair* data set in Figure 3.1(a). The inliers are composed by 4 groups of 50 points (forming 4 horizontal segments) corrupted by Gaussian noise. The final data set is obtained adding the “true” outliers: 300 points uniformly distributed in a square domain. This way each model “sees” 150 pseudo-outliers and 300 true outliers. To compare the performance of the algorithms we define the set of *correctly detected inliers* for the model w as $C_w \stackrel{\text{def}}{=} \hat{D}_{I,w} \cap D_{I,w}$ where $\hat{D}_{I,w}$ is the detected set of inliers, and the *percentage of correctly detected inliers* as:

$$c_w \stackrel{\text{def}}{=} \frac{|C_w|}{|D_{I,w}|} 100$$

Table 1 shows the results of the multiRANSAC approach versus the sequential RANSAC approach averaged over 50 instances of the stair data set for varying values of the noise standard deviation. The average value of c is given by:

$$c_{av} \stackrel{\text{def}}{=} \sum_{h=1}^{50} \sum_{w=1}^W c_w^{(h)}$$

In general multiRANSAC performs better than sequential RANSAC, especially when the standard deviation of the noise that corrupts the inliers becomes larger. This happens at the price of an increased computational complexity: expression (2) produces larger values than in the sequential RANSAC case. To cope with this problem in our experiments we stopped the iterations after the best CSs were left unchanged for 5000 iterations.

3.2 Using multiRANSAC to Detect Homographies

In this section we will describe how the multiRANSAC algorithm can be used to identify homographies between an

Table 1. multiRANSAC vs. Sequential RANSAC

σ	multiRANSAC		sequential RANSAC	
	c_{av}	N_{iter}	c_{av}	N_{iter}
$5.5 \cdot 10^{-3}$	95.96%	7499	93.80%	2016
$6.0 \cdot 10^{-3}$	95.22%	7626	86.00%	2049
$6.5 \cdot 10^{-3}$	90.08%	8194	67.14%	2080
$7.0 \cdot 10^{-3}$	90.13%	8699	47.92%	2103
$7.5 \cdot 10^{-3}$	86.01%	8110	37.29%	2103

image pair. Related work can be found in [5, 6], where the authors apply RANSAC sequentially to detect the homographies relating points in a pair of images. Consider two images and a set of point correspondences $\mathbf{x}_i \leftrightarrow \mathbf{x}'_i$ for $1 \leq i \leq N$. If two points belong to a planar region then they are related by a linear projective transformation called homography: $\mathbf{x}'_i = H\mathbf{x}_i$, where $H \in \mathbb{R}^{3 \times 3}$ has 8 degrees of freedom (scale does not matter). Therefore any homography can be estimated from 4 point correspondences (provided that the configuration of the points is not singular), and therefore $k_w = 4$. The parameter vector $\boldsymbol{\theta}_w$ is composed by the first 8 components of $\text{col}(H)$. Following [1] we define the manifold error to be the symmetric transfer error $e(\mathbf{z}, \mathcal{M}_w(\boldsymbol{\theta}_w)) = d(H_w\mathbf{x}, \mathbf{x}')^2 + d(\mathbf{x}, H_w^{-1}\mathbf{x}')^2$ where $\mathbf{z} \stackrel{\text{def}}{=} [\mathbf{x}^T \quad \mathbf{x}'^T]^T$. As suggested in [6], the MSSs are formed in such a way that points that are neighbors are selected with higher probability. More precisely, if a point \mathbf{x}_i has been selected, then a point \mathbf{x}_j will be selected with the following probability:

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

where Z is a normalization constant. The value of σ is chosen so that the probability of selecting a point \mathbf{x}_j that has distance \bar{d} from \mathbf{x}_i is equal to \bar{P} . Note that the denominator of (1) is constructed assuming that we sample uniformly

²For any matrix A the notation $\text{col}(A)$ denotes the vector obtained from the column scan of the matrix A .

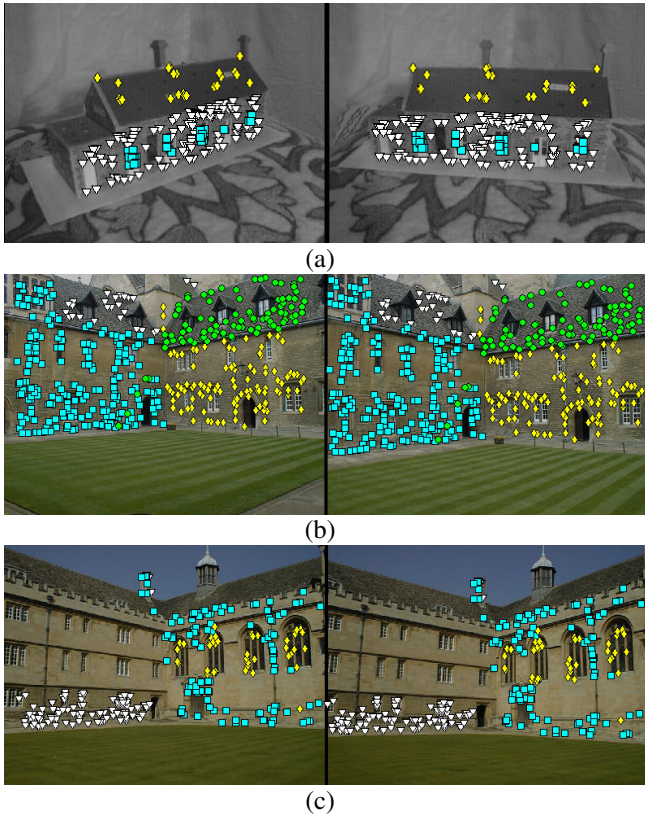


Fig. 2. This figure is best viewed in color. Markers with the same color and shape indicate features that belong to the same planar region. The image pairs are: (a) House, (b) Merton and (c) Wadham.

over the entire data set. This is not true if we adopt the sampling procedure described above. Therefore in (1) we replace the value N with:

$$\hat{N} \stackrel{\text{def}}{=} 1 + \frac{1}{N} \sum_{\mathbf{x}_i \in D} \left| \left\{ \mathbf{x}_j \in D : \sum_j P(\mathbf{x}_j | \mathbf{x}_i) \geq T_P \right\} \right|$$

where the points \mathbf{x}_j in the summation are sorted according to their distance from \mathbf{x}_i (closer points have smaller indices). Therefore \hat{N} is the average number of neighbors of each point. In our experiments we choose $T_P = 0.95$.

The results obtained using multiRANSAC to estimate planar homographies are summarized in Table 2 and displayed in Figure 3.2. To estimate the homographies we used

Table 2. multiRANSAC vs. Sequential RANSAC

Image pair	W	multiRANSAC		seq. RANSAC	
		N_I	N_{iter}	N_I	N_{iter}
House	3	150	14361	149	1802
Merton	4	437	11604	448	2004
Wadham	3	280	10024	287	1503

the normalized DLT algorithm [1]; the noise standard deviation was always set to one pixel. Qualitatively the grouping of the point features is compliant with the planar structures present in the image pairs. Note that sequential RANSAC in two cases found a larger number of inliers: this however does not necessarily mean that the quality of the homography estimates is better.

4 Conclusions and Future Work

The classic RANSAC algorithm has been extended to the problem of simultaneous parameter estimation of multiple models in data sets with a high percentage of outliers. Experimental results on synthetic data seem to confirm that a parallel approach produces more stable estimates than a sequential approach. The algorithm has also been validated estimating multiple planar homographies on real image pairs (taken with non calibrated cameras).

Future research efforts will be focused on automatically determining the optimal number of models W for a given data set, on developing better strategies for the fusion of the CSs and on the exploitation of parallel architectures (the estimation of the parameters of different models can be threaded on different processors). We are also interested in proposing stopping criteria that are tighter than the iteration bound (2): for this purpose we are currently trying to stochastically model the number of iterations between two updates of the CSs.

Acknowledgements: We would like to thank Dr. Bober and Dmitry Fedorov for helpful discussions and the Visual Geometry Group at Oxford University for the real sequences data. This work was supported in part by an ONR grant #N00014-04-1-0121.

5 References

- [1] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, second edition, 2003.
- [2] C. Kenney, B. Manjunath, M. Zuliani, G. Hoyer, and A. Van Nevel, "A condition number for point matching with application to registration and post-registration error estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 11, pp. 1437–1454, November 2003.
- [3] M. J. Black and P. Anandan, "A framework for the robust estimation of optical flow," in *Fourth International Conf. on Computer Vision*, Berlin, Germany, May 1993, pp. 231–236.
- [4] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Comm. of the ACM*, vol. 24, pp. 381–395, 1981.
- [5] E. Vincent and R. Laganière, "Detecting planar homographies in an image pair," in *2nd International Symposium on Image and Signal Processing and Analysis*, Pula, Croatia, June 2001, pp. 182–187.
- [6] Y. Kanazawa and H. Kawakami, "Detection of planar regions with uncalibrated stereo using distribution of feature points," in *British Machine Vision Conference*, Kingston upon Thames, London, September 2004, vol. 1, pp. 247–256.