CLUSAC: CLUSTERING SAMPLE CONSENSUS FOR FUNDAMENTAL MATRIX ESTIMATION

Xuanyu Xiao[1], Zongqing Lu[1], Jing-Hao Xue[2]

[1]Shenzhen International Graduate School, Tsinghua University, China
[2]Department of Statistical Science, University College London, UK

ABSTRACT

In the process of model fitting for fundamental matrix estimation, RANSAC and its variants disregard and fail to reduce the interference of outliers. These methods select correspondences and calculate the model scores from the original dataset. In this work, we propose an inlier filtering method that can filter inliers from the original dataset. Using the filtered inliers can substantially reduce the interference of outliers. Based on the filtered inliers, we propose a new algorithm called CLUSAC, which calculates model quality scores on all filtered inliers. Our approach is evaluated through estimating the fundamental matrix in the dataset kusvod2, and it shows superior performance to other compared RANSAC variants in terms of precision.

Index Terms— Inlier filter, model quality score, fundamental matrix estimation, RANSAC

1. INTRODUCTION

Fundamental matrix estimation is an essential problem in the field of computer vision. RANSAC (RANdom SAmple Consensus) [1] and its variants have been widely used in fundamental matrix estimation. Standard RANSAC randomly selects a sample subset of minimal size and fits the model, repeats these two steps, and returns the model with the highest quality. Since the publication of RANSAC, many improved variants have been proposed. There are three main ways to improve RANSAC: modify the model generation (including sampling strategy) [2, 3, 4, 5, 6], modify the model scoring strategy [7, 8, 9, 10] and modify the termination criterion [11, 12, 13].

RANSAC and its variants select the sample subsets from the original contaminated dataset to fit a model and calculate the model score on all correspondences. These approaches fail to consider and thus cannot reduce the interference of outliers in the process of model fitting. In this work, we focus on reducing the interference of outliers in the model fitting process, and propose to filter inliers from the original dataset and calculate scores of the estimated model on the filtered inliers, so as to reduce the interference of outliers.

As the *major contribution* of this paper, we propose a new method of inlier filtering based on clustering. We cluster the original dataset into two clusters based on the different characteristics of inliers and outliers. As a *second contribution*, we propose an algorithm based on the filtered inliers. The proposed algorithm CLUSAC (Clustering Sample Consensus) calculates the model quality score on all the filtered inliers. Using the filtered inliers to fit the model can reduce the interference of outliers on the model estimation.

We compare the proposed algorithm CLUSAC with other RANSAC variants in fundamental matrix estimation and find that CLUSAC can estimate the fundamental matrix better than the other compared RANSAC variants in dataset kusvod2 in terms of precision.

2. RELATED WORK

Since we proposed an inlier filtering method and calculated the model scores based on the filtered inliers, we review previous work related to model generation and model scoring.

Modify the model generation. GroupSAC [2], PROSAC [3], NAPSAC [4] and EVSAC [14] modify the sampling strategy to increase the probability of selecting an all-inlier sample earlier. GroupSAC assumes that there are groups in the dataset and uses hierarchical sampling to increase the probability of extracting all-inlier samples. PROSAC sorts the matches linearly through a similarity function and extracts the sample subset of the minimal size from the top-ranked correspondences. NAPSAC utilizes the spatial correlation between inliers and selects inliers by constructing a hypersphere. EVSAC leverages Extreme Value Theory to model the statistics of matching scores. MLESAC [5] assumes that inliers and outliers obey different distributions, estimates the model quality through maximum likelihood, and returns the model with the maximum likelihood. Based on LO-RANSAC [15], Graph-Cut RANSAC [6] exploits spatial coherence, making it a local optimization RANSAC, which alternates graph-cut and model re-fitting as the LO step. These methods try to modify the model generation under certain assumptions to select an all-inlier sample earlier, but they did not try to filter inliers from the original dataset to reduce the interference of outliers on the model estimation.

Modify the model scoring strategy. Standard RANSAC pre-defines an inlier-outlier threshold, and if the point-to-model residual is smaller than the threshold, the point is an inlier, otherwise, the point is an outlier. RANSAC returns the model with the most number of inliers. RANSAC variants try to modify the model scoring strategy to make the model scoring process faster and better. Preemptive RANSAC [7], $T_{d,d}$ test [8] and Bail-out test [9] are RANSAC variants that try to make RANSAC scoring faster. Preemptive RANSAC uses a preemptive scoring strategy in the scoring process, and $T_{d,d}$ test and Bail-out test terminate the scoring process early if it is sufficiently certain that the score being computed will fail to be better than the previously recorded best consensus score. Latent RANSAC [11] maps the hypothesis to a latent space and defines the hypothesis distance as the difference in the size of the correspondences residuals, and this method can evaluate the RANSAC hypothesis within a constant time. Other RANSAC variants try to modify the model quality function to select a better model. The state-of-the-art algorithm MAGSAC [10] assumes the noise σ to be a random variable, marginalizing the log-likelihood quality function to get the updated quality function without parameter σ. MAGSAC solves the problem where the inlier noise scale needs to be defined by the user. These RANSAC variants calculate the model scores on the original dataset, and thus the quality of the estimated model can be interfered by outliers in the process of estimation.

3. CLUSTERING SAMPLE CONSENSUS (CLUSAC)

The existence of outliers will affect the model fitting process. RANSAC and its variants did not try to reduce the interference of outliers in the process of model fitting. In this work, We propose an inlier filtering method and then an algorithm that calculates the model quality scores based on the filtered inliers to reduce the interference of outliers.

3.1. Inlier Filtering

Standard RANSAC randomly selects a sample subset of minimal size to fit the model and returns the model with the most number of inliers. The model obtained by RANSAC is random, hence both the estimated model θ and the residual γ of each point to the estimated model are random variables.

For each estimated model θ_i, the point-to-model residuals of inliers should be smaller than the residuals of outliers. Hence, the different characteristics of correspondences can be used to distinguish inliers and outliers and thus filter inliers.

For each estimated model θ_i, we calculate the residuals of all correspondences $p_j, j = 1, \ldots, m$. In order to filter the inliers, we estimate the model n times, and thus we can get the residual list into a $n \times m$ matrix Γ. By transposing matrix Γ, each row of the transposed matrix represents the residuals of each correspondence for different models. As the inliers

and outliers have different characteristics of residuals, we can use a clustering method to group the original dataset into two clusters, and the cluster with smaller residual can be treated as the filtered inliers.

The proposed inlier filtering algorithm is summarized in Algorithm 1. Firstly, we fit the model multiple times and calculate the residuals of each model for all correspondences. We fit RANSAC models 40 times in experiments. To obtain a better clustering result, we normalize the residuals of each model for all correspondences. Secondly, we transpose the residual list Γ, and get its principal component list ζ using principal component analysis. Using PCA to get the principal component list is because each row of the transposed matrix may be related. Thirdly, we cluster original correspondences into two groups through the list ζ. Finally, we output the cluster with the smaller residual as the filtered inliers I.

Algorithm 1 Inlier filtering.

Input:
  P - data points, σ - inlier-outlier threshold;
  n_1 - number of models, n_2 - dimension of PCA;
Output:
  I - filtered inliers
1: **for** i = 1 $\rightarrow n_1$ **do**
2:   Fit model θ_i, calculate the model's residuals for all correspondences $\gamma_{ij}, j = 1, \ldots, m$.
3:   Normalize the calculated residuals of each model θ_i.
4: **end for**
5: Transpose residual list Γ and get its n_2 principal component list ζ using PCA.
6: Cluster the original dataset into two clusters through list ζ.
7: Output the cluster with the smaller RMS residual as the filtered inliers I.

3.2. Model Quality Score

To evaluate the quality of the model, we define the quality score of the model θ as follows:

$$Q(I, \theta) = \frac{1}{L(I, \theta)}, \tag{1}$$

where

$$L(I, \theta) = \sqrt{\frac{1}{|I|} \sum_{p_i \in I} \gamma^2(p_i, \theta)} \tag{2}$$

is a loss function that calculates the model's RMS residual on all filtered inliers, in which $\gamma(p_i, \theta)$ is the residual of point p_i to the estimated model θ and I is the set of filtered inliers. A smaller loss function value means that the model fits the filtered inliers better.

CLUSAC uses Eq.(1) as the quality function to score the estimated model. As the model scores are calculated in the filtered inliers, we can reduce the interference of outliers.

3.3. Termination Criterion

We calculate the model score of each model on the filtered inliers. This score represents the fitness between the estimated model and the filtered inliers. By predefining a model score threshold, we terminate the iteration process if the model score of the estimated model is larger than the predefined score threshold. To avoid the issue due to an overly-small predefined model score threshold, we set the maximum number of iterations. When the number of iterations is larger than the maximum number of iterations, terminate the iteration process, and output the model with the highest quality score.

The standard termination of RANSAC is as follows [10]:

$$k(\theta, \sigma, P) = \frac{\ln(1 - \eta)}{\ln(1 - (\frac{|I(\theta, \sigma, P)|}{|I|})^m)}, \quad (3)$$

where k is the iteration number, η is the manually set confidence, $|I(\theta, \sigma, P)|$ is the inlier number of so-far-the-best model, and m is the size of the minimal sample.

We use the maximum number of iterations of the standard RANSAC as a reference and set a maximum iteration threshold greater than k.

3.4. CLUSAC Algorithm

The CLUSAC algorithm is summarized in Algorithm 2. Firstly, we filter inliers from the original correspondences using Algorithm 1. Then, we randomly select a sample subset of the minimal size on the filtered inliers I and fit the model. For each estimated model, we calculate its RMS residual on all filtered inliers and the model's quality. If the score of the estimated model is larger than the predefined score threshold or the iteration number reaches the maximum number of iteration, we output the model with the highest quality.

Algorithm 2 CLUSAC

Input:
    P - data points, k - maximum iteration number;
    σ - inlier-outlier threshold, τ - model quality threshold;
 n_1 - number of models, n_2 - dimension of PCA;

Output:
    θ - model parameters
1: Filter inliers $I(P, \sigma, n_1, n_2)$ (Algorithm 1)
2: **for** i = 1 → k **do**
3:     Randomly select sample subset of the minimal size on filtered inliers I, and fit the model.
4:     Calculate the estimated model's RMS residual on all filtered inliers I, get the model's quality Q_i.
5:     **if** $Q_i > \tau$ **then**
6:         break.
7:     **end if**
8: **end for**
9: Output the estimated model θ with the highest quality.

---

4. EXPERIMENTAL RESULTS

Dataset. The performances of CLUSAC and RANSAC variants are compared in dataset kusvod2[1]. These image pairs were previously used in RANSAC variants [3] [6] [16] [17]; it contains 16 image pairs for fundamental matrix estimation. For a fair comparison, we used the point pairs obtained by matching the SIFT descriptors of MSER's in LO$^+$-RANSAC [16]. This dataset has two parts of point pairs: the first part is used to fit the model, the second part is used to calculate the precision of the estimated model, which contains about ten pairs of manually annotated ground-truth point. We checked all these ground-truth point pairs in kusvod2 to ensure that these ground-truth points are reliable. Fig.1 shows the scene Shout with eleven manually annotated ground-truth point pairs.
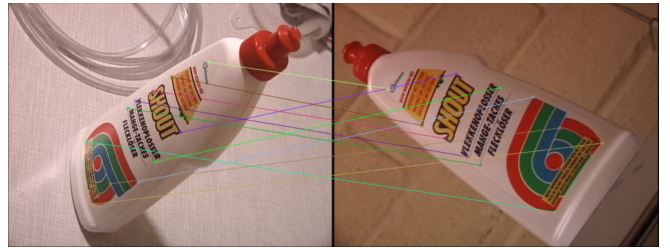
Fig. 1. The scene Shout with ground-truth point pairs. We checked all these point pairs in kusvod2 to ensure that they are reliable.

Compared RANSAC variants. The compared RANSAC variants are RANSAC [1], LO-RANSAC [15], MSAC [18], LO-MSAC [16], MSAC+Lsq [16], LO'-RANSAC and LO$^+$-RANSAC [16]. The inlier-outlier threshold related parameter σ of all compared RANSAC variants was set to 0.3. The state-of-the-art algorithm MAGSAC [10] also used the dataset kusvod2 to estimate the fundamental matrix. The average RMS Sampson distance of 16 image pairs of kusvod2 in MAGSAC is 0.38, which is much smaller than the experimental result in LO$^+$-RANSAC [16]. This is because MAGSAC merged the two parts of point pairs in the dataset kusvod2 and augmented the manually annotated ground truth correspondences as follows[2]: First, the implied model is estimated from the manually selected inliers. Second, the inliers of the ground-truth model are selected. In contrast, LO$^+$-RANSAC did not include those manually annotated ground truth point pairs in the estimation process, nor does our experiments. Hence, we do not include MAGSAC for comparison.

What is measured. To compare the precision of the estimated fundamental matrix, we calculate the error of the esti-

---

Table 1. CLUSAC's RMS Sampson distance compared with RANSAC variants. The best is in bold.

| scenes	RSC	LO-RSC	MSC	LO-MSC	MSC+LSq	LO'-RSC	LO$^+$-RSC	CLUSAC
Booksh	3.22	1.82	3.05	1.77	2.81	2.83	1.77	**1.29**
Box	49.46	62.28	50.16	62.58	54.71	56.78	58.52	**7.63**
Castle	4.58	0.95	4.29	0.94	3.54	2.35	**0.81**	3.76
Corr	0.51	0.21	0.48	0.18	0.37	0.31	**0.18**	0.47
Graff	2.74	3.03	2.69	3.09	2.63	2.15	3.08	**0.91**
Head	0.78	0.32	0.78	0.31	0.40	0.30	0.31	**0.21**
Kampa	16.54	13.89	14.27	12.46	14.10	9.12	12.52	**1.75**
Kyoto	2.27	0.87	2.25	0.81	1.64	1.07	0.78	**0.71**
Leafs	8.19	3.87	7.94	3.88	5.86	3.18	3.85	**1.39**
Plant	23.84	23.44	21.24	20.93	21.31	16.63	20.93	**7.85**
Rotunda	1.31	0.47	1.30	0.52	0.62	0.44	**0.43**	2.27
Shout	1.77	0.86	1.72	**0.82**	1.62	1.34	0.82	1.46
Valbonne	30.21	29.29	29.46	28.56	29.67	28.03	28.54	**25.82**
Wall	2.26	0.45	2.23	0.42	1.06	0.53	0.49	**0.21**
Wash	1.05	0.30	1.04	**0.27**	0.39	0.28	**0.27**	6.74
Zoom	2.49	0.89	2.39	0.86	1.92	1.52	0.86	**0.64**

mated model on manually annotated ground-truth correspondences. The error is the RMS Sampson distance in pixels; the Sampson distance [19] is defined as follows:

$$\frac{\left(x_i'^T F x_i\right)^2}{(F x_i)_1^2 + (F x_i)_2^2 + (F^T x_i')_1^2 + (F^T x_i')_2^2}, \quad (4)$$

where x_i and x_i' are point pairs and F is the estimated fundamental matrix. A smaller error indicates a better precision; however, small differences in small values are insignificant [16].

Results and analysis. CLUSAC uses the normalized eight-point method [20] to calculate the fundamental matrix, thus we draw a minimal sample subset of size eight in each iteration. The parameter σ is set to 0.3 as other compared RANSAC variants, the model score threshold τ is set to 1.5, and the maximum iteration number is set to 1000. The error of CLUSAC and compared RANSAC variants are shown in Table 1, with the best in bold.

It can be seen from Table 1 that CLUSAC is superior to other compared RANSAC variants in terms of precision. The RMS Sampson distance of the fundamental matrix estimated by CLUSAC is smaller than other RANSAC variants in most of the scenes in kusvod2. Especially in the scenes Box, Kampa, and Plant, the RMS Sampson distances are much smaller than other compared RANSAC variants, which means that CLUSAC can estimate the fundamental matrix much better in those scenes. For example, in the scene Box, the RMS Sampson distance of CLUSAC is 7.63, while most of the other compared RANSAC variants are more than 50. CLUSAC is the worst on scenes Rotunda and Wash; however, both the compared RANSAC variants and CLUSAC have small errors on these scenes, and small differences between small errors are not significant.
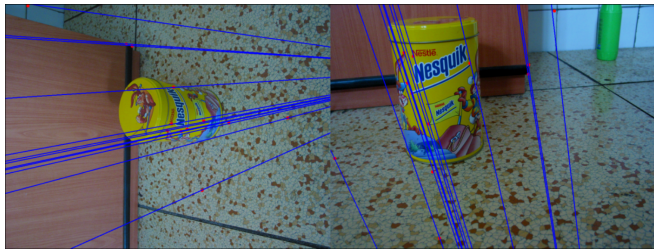
Fig. 2. The epipolar lines and the ground-truth points in the scene Box. Errors: RMS Sampson distance = 7.63 pixels.

In Fig.2 we plot the epipolar lines of the estimated fundamental matrix and the manually annotated ground-truth point pairs in the scene Box. We can find that most of the epipolar lines are near the ground-truth point pairs, which clearly indicates that the fundamental matrix estimated by CLUSAC is excellent in this scene.

5. CONCLUSIONS

In this paper, we propose a method of inlier filtering, and based on the filtered inliers, we propose the algorithm CLUSAC, which calculates the model quality scores on all filtered inliers. By using the filtered inliers to fit the model and calculate the model scores, we can reduce the interference of the outliers on the process of model fitting. The proposed algorithm is evaluated through estimating the fundamental matrix in the real-world dataset kusvod2, and the result demonstrates that the proposed algorithm is superior to other compared RANSAC variants in terms of precision.

6. 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*, vol. 24, no. 6, pp. 381–395, 1981. 1, 3

[2] Kai Ni, Hailin Jin, and Frank Dellaert, "GroupSAC: Efficient consensus in the presence of groupings," in *ICCV*, 2009, pp. 2193–2200. 1

[3] Ondrej Chum and Jiri Matas, "Matching with PROSAC - progressive sample consensus," in *CVPR*, 2005, vol. 1, pp. 220–226. 1, 3

[4] D. R. Myatt, P.H.S. Torr, Slawomir J. Nasuto, J. Mark Bishop, and R. Craddock, "NAPSAC: High noise, high dimensional model parameterisation - it's in the bag," in *BMVC*, 2002, pp. 458–467. 1

[5] P.H.S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Computer Vision and Image Understanding*, vol. 78, no. 1, pp. 138–156, 2000. 1

[6] Daniel Barath and Jiri Matas, "Graph-Cut RANSAC," in *CVPR*, 2018, pp. 6733–6741. 1, 3

[7] David Nister, "Preemptive RANSAC for live structure and motion estimation," *Machine Vision and Applications*, vol. 16, no. 5, pp. 321–329, 2005. 1, 2

[8] Jiri Matas and Ondrej Chum, "Randomized RANSAC with td, d test," *Image and Vision Computing*, vol. 22, no. 10, pp. 837–842, 2004. 1, 2

[9] D. P. Capel, "An effective bail-out test for RANSAC consensus scoring," in *BMVC*, 2005. 1, 2

[10] Daniel Barath, Jiri Matas, and Jana Noskova, "MAGSAC: Marginalizing sample consensus," in *CVPR*, 2019, pp. 10197–10205. 1, 2, 3

[11] Simon Korman and Roee Litman, "Latent RANSAC," in *CVPR*, 2018, pp. 6693–6702. 1, 2

[12] Ben J Tordoff and David William Murray, "Guided-MLESAC: Faster image transform estimation by using matching priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1523–1535, 2005. 1

[13] Rahul Raguram, Ondrej Chum, Marc Pollefeys, Jiri Matas, and Jan-Michael Frahm, "USAC: A universal framework for random sample consensus," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 2022–2038, 2013. 1

[14] Victor Fragoso, Pradeep Sen, Sergio Rodriguez, and Matthew Turk, "EVSAC: Accelerating hypotheses generation by modeling matching scores with extreme value theory," in *ICCV*, 2013, pp. 2472–2479. 1

[15] Ondrej Chum, Jiri Matas, and Josef Kittler, "Locally optimized RANSAC," in *Joint Pattern Recognition Symposium*, 2003, pp. 236–243. 1, 3

[16] Karel Lebeda, Jiri Matas, and Ondrej Chum, "Fixing the locally optimized RANSAC," in *BMVC*, 2012. 3, 4

[17] Daniel Barath, "Five-point fundamental matrix estimation for uncalibrated cameras," in *CVPR*, 2018, pp. 235–243. 3

[18] P.H.S. Torr and A. Zisserman, "Robust computation and parametrization of multiple view relations," in *ICCV*, 1998, pp. 727–732. 3

[19] Richard I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge university press, 2003. 4

[20] Richard I. Hartley, "In defense of the eight-point algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 6, pp. 580–593, 1997. 4