



# **Performance of RANSAC Techniques under Classical and Robust Methods**

R. Muthukrishnan<sup>1</sup>, E.D. Boobalan<sup>2</sup>, R.Reka<sup>3</sup>

Assistant Professor, Department of Statistics, Bharathiar University, Coimbatore, India<sup>1</sup>

Research Scholar, Department of Statistics, Bharathiar University, Coimbatore, India<sup>2&3</sup>

**ABSTRACT:** Robust statistics deals with deviations from the assumptions on the model and is concerned with construction of statistical procedure which still reliable and reasonably efficient in a neighbourhood of the model. In computer vision, a robust estimator should be able to correctly find the fit when outliers/noise occupies a high percentage of the data. RANSAC is most commonly used robust estimator in computer vision task. It is a non-deterministic algorithm for estimation of a mathematical model from observed data which contains outliers. This paper presents the performance of RANSAC algorithmic techniques under the various methods LS, LMS, LTS and M estimators along with the results of the simulation study which is carried out in MATLAB.

**KEYWORDS:** Robust Statistics – RANSAC - MATLAB

## **I. INTRODUCTION**

The origin of robust statistics was step in to the eighteenth century, when the first rules for the rejection of outliers were developed. Following that in the next century, these rules were formalized and techniques for robustly estimating 'means' were used. Later on, estimators that down weight outliers were developed, during the first half of the 20th century, robustness of statistical testing was considered. [1] and [21] demonstrated the importance for robust methods and their work was considered as a breakthrough in robust. After few years later, [7] and [6] put forward the foundation of comprehensive theory of robust statistics.

The goal of robust statistics is to reduce the impact outliers and their methods try to fit the data as a bulk one, which assumes that the good observation outnumbers the outliers. By using the residuals, outliers can be identified that are large in the robust analysis. An important task afterwards is to ask what has caused these outliers and they should not be ignored, but have to be analysed and interpreted.

Robust statistics are used in computer vision for the past three decades and the most popular today are linked to old methods that are proposed to solve specific image understanding or pattern recognition problems. In that, some of them were rediscovered in past few years. Computer vision techniques have numerous practical applications, in which some of them being character recognition, industrial inspection, medical image analysis, remote sensing navigation, robot navigation, scene modelling, surveillance human identification activity analysis etc.

The main goal of computer vision is to mimic human visual perception. In the broader sense an equivalent task had been performed when the robustness of a computer vision algorithm is judged against the human observer performance. Even though the usual information is carried only by a small subset of data and/ or is significantly different from an already stored representation, robustness has the ability to extract the visual information of relevance for a specific task.

In this paper an attempt has been made to study the performance of RANSAC techniques through the simulated data with error tolerance under the various classical and robust methods such as LS, LMS, LTS and M-estimators. The various most widely used RANSAC techniques are briefly discussed in the section 2 and Section 3 deals with the classical and robust methods which are used to study the performance of RANSAC techniques. The result obtained through the simulation study is summarized in the section 4 and the discussion is in the last section.



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

## II. RELATED WORK

The RANSAC algorithm is one of the most popular robust estimators in computer vision and their structure was proposed by [5]. It is simple and most effective techniques that are commonly applied to the task of estimating the parameters of model that may be contaminated by outliers. RANSAC is also a non-deterministic algorithm for the estimation of a mathematical model from observed data that contains outliers. The speed of RANSAC depends on two factors. The first one is the level of contamination determines the number of random samples that have to be taken to guarantee a certain confidence in the optimality of the solution and the next one is the time utilized to evaluate the quality of each of the hypothesized model parameters that are proportional to the size of the data set. Typically, numerous erroneous model parameters obtained from contaminated samples are evaluated and such kinds are consistent with only a small fraction of the data. The original RANSAC algorithm works in a hypothesis testing frame work, it randomly samples a minimum set of data point from which model parameters are computed and verified against all the data points to determine a fitting score and this process is repeated until a termination criterion is observed. The RANSAC finds with a certain probability, all inliers and the correspondence model by repeated drawing random samples. RANSAC, a robust estimator which gives a correct estimation of a scale of inliers that can be robust to more than 50% outliers and their performance is largely depends on the given scale or error tolerance.

### A. *M-estimator Sample Consensus (MSAC) and Locally optimized MSAC (LOMSAC)*

The M-estimator sample consensus (MSAC), a generalization of the RANSAC estimator that is introduced by [20], mainly used to robustly estimate multiple new relations from point correspondences. MSAC adopts the same sampling strategy as RANSAC to generate reputed solutions but chooses the solution that maximizes the likelihood rather than just the number of inliers. One of the problems in RANSAC is that if the threshold  $T$  for considering inlier is set too high then the robust estimator can be very poor. Maximum likelihood estimation sample consensus (MLE-SAC) is a case of MSAC that yields a modest to large benefit to all robust estimations with absolutely no additional computational burden. The maximum likelihood error allows suggesting an improvement over MSAC. By locally optimization this MSAC changed into LOMSAC.

### B. *R-RANSAC (RANDOMIZED RANSAC)*

A new randomized version of R-RANSAC proposed by [2] that increases the speed of model parameter estimation under a broad range of conditions, due to the randomization of the hypotheses evaluation step. The sample contaminated by outliers was shown that it was sufficient to test only a small number of data points to conclude with high confidence. The emphasized ideas of the R-RANSAC are that the most model parameter hypothesis evaluated is influenced by outliers. To avoid such erroneous models, a statistical test have to be perform on only a small number of data points. The hypotheses generation step proposes a model that the data points are either good or bad, the property of good is a hidden state that is not directly observable events, but it is statistically linked to it. The observable events are "data a point are or is not consistent with the model".

The speed up of the R-RANSAC depends mainly on the probabilities of the two types of errors that committed in the pre-test, the rejection of an uncontaminated and the acceptance of contaminated model.

R-RANSAC used random sampling more thoroughly in contrast to them. It uses small amount of data, when R-RANSAC counts inlier candidates, most RANSAC based approaches are need to tune variables such as the number of iteration. The variable should be adjusted again when the given data are modified.

### C. *N Adjacent Points Sample Consensus (NAPSAC)*

Author [11] introduced the N Adjacent points sample consensus estimator. The algorithm of N Adjacent Points Sample Consensus (NAPSAC) focused on the efficiency of the sampling strategy. The idea is that inliers tend to be closer to one another than outliers, and the sampling strategy can be modified to exploit the proximity of inliers. This might bring efficient robust recovery of high dimensional models, where the probability of drawing an uncontaminated sample becomes very low even for data sets with relatively low contamination of outliers.



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

## D. Locally Optimized RANSAC (Lo-RANSAC)

Lo-RANSAC proposed by author [3]. It has been observed that to find an optimal solution, the number of samples drawn in RANSAC is significantly higher predicted from the mathematical model. This is due to the incorrect assumption, that a model with parameters computed from an outlier-free sample is consistent with all inliers. The Locally optimized RANSAC based on the observation that virtually all models estimated from an uncontaminated minimal sample contain large fraction of inliers within their support. An optimization process starting from the so-far-the-best hypothesized model is therefore inserted into RANSAC. The algorithm was extensively tested on the problem of estimation of the two view relations from image point correspondences. Lo-RANSAC improves hypothesis quality by performing a local optimization step on the current inlier set every time a new best hypothesis is found. Other approaches aim to reject bad hypotheses early on.

## III. CLASSICAL AND ROBUST METHODS

### A. Least Square Estimator

Linear regression analysis is an important tool in most applied science including computer vision. The least squares (LS) method is one of the most widely used to estimates the parameter in linear regression methods and it has been used in many scientific fields for a long time.

The classical linear model can be described in the following form

$$y_i = x_{ij}\theta_1 + \dots + x_{ip}\theta_p + e_i \quad (i = 1, \dots, n)$$

The ordinary least squares regression estimator can be written as

$$\hat{\theta} = \arg \min_{\hat{\theta}} \sum_{i=1}^n r_i^2$$

The LS estimator is highly efficient and achieves optimum results under Gaussian distributed noise. Although the LS method has the advantages of low computational cost and high efficiency, it is extremely sensitive to outliers.

### B. LMS and LTS Estimator

In the context of regression analysis many methods have been proposed, the two methods Least Median Square (LMS) and Least Trimmed Square (LTS) are still alive and widely used. [15] author Proposed the LMS and it is based on the idea of replacing the sum in least sum of squares formulation by a median. It finds the parameters to be estimated by minimizing the median of squared residuals corresponding to the data points.

The least median square estimate is as follows:

$$\hat{\theta} = \arg \min_{\hat{\theta}} \text{med}_i r_i^2$$

The LMS method has excellent global robustness and high breakdown point. However, the relative efficiency of the LMS method is poor when Gaussian noise is present in the data. Rousseeuw noted that the LMS has a very low convergence rate [16]. To compensate the deficiency Rousseeuw improved the LMS method by carrying out a weighted least square procedure after the initial LMS fit. The LTS method uses h data point (out of n) to estimate the parameters. The aim of LTS estimator is to find the h subset with smallest least squares residuals and use the h subset to estimate parameters of models. The LTS have less sensitive to local effects than LMS and also better statistical efficiency than LMS. The implement of the LTS method also uses random sampling because the number of all possible h subsets grows fast with n. Like LMS, the efficiency of LTS can also be improved by adopting a weighted least squares refinement as the last stage.

$$\hat{\theta} = \arg \min_{\hat{\theta}} \sum_{i=1}^h (r^2)_{i:n}$$

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

### C. M-Estimator

The maximum-likelihood type estimators (M-estimators) are well known among the robust estimators. The theory of M-estimator was firstly developed by Huber. They minimize the sum of a symmetric, positive-definite function of the residuals with a unique minimum at zero.

The M-estimator of the parameter are obtained by converting the minimization

$$\min \sum_i \rho(r_i)$$

into a weighted least squares problem which is handled by available sub routines. The weights depend on the assumed  $\rho$  function and the data.

M-estimators are generalizations of MLEs and least squares. In particular, the M-estimate of a is

$$\hat{a} = \arg \min_a \sum_{X_i \in \mathcal{X}} \rho(r_{i,a} / \sigma_i)$$

Where  $\rho(u)$  is a robust loss function,  $\sigma_i^2$  is the variance.

## IV. SIMULATION STUDY

The simulation study has been performed to study the behaviour of various RANSAC methods under classical and robust methods such as LS, LMS, LTS and M-estimators. The data were generated with the model  $ax + by + c = 0$  where  $a^2 + b^2 = 1$  using MATLAB software with fixed threshold  $t = 0.2$  and various sample sizes  $[n=100, 500, 1000]$  with error also consider here. The study has been performed in the context of number of inliers estimated by using the various RANSAC techniques and it is summarized in the TABLE I.

TABLE I  
ESTIMATED NUMBER OF INLIERS

Methods	Least Square				Least Median Square				Least Trimmed Square				M – estimator				
	$\epsilon$	0.0	0.1	0.15	0.2	0.0	0.1	0.15	0.2	0.0	0.1	0.15	0.2	0.0	0.1	0.15	0.2
N = 100																	
RANSAC	93	85	81	75	100	91	88	85	100	96	89	84	100	93	88	85	85
MSAC	91	83	79	74	100	89	87	84	100	95	87	82	100	92	87	84	84
RRANSAC	95	82	82	75	100	91	87	85	100	96	88	83	100	93	85	85	85
NAPSAC	89	82	79	71	100	89	86	83	100	95	88	82	100	92	83	82	82
LOMSAC	93	85	81	74	100	91	87	84	100	95	89	83	100	92	87	84	84
LORANSAC	92	83	80	73	100	91	87	84	100	95	89	83	100	93	87	84	84
N = 500																	
Real model	500	450	425	400	500	450	425	400	500	450	425	400	500	450	425	400	400
RANSAC	476	436	414	391	500	459	442	419	500	457	442	416	500	456	439	413	413
MSAC	472	432	411	387	500	459	441	418	500	455	439	414	500	456	438	412	412
RRANSAC	476	433	413	388	500	460	441	419	500	457	442	416	500	456	439	414	414
NAPSAC	460	430	410	383	500	458	438	406	500	451	440	414	500	446	436	371	371
LOMSAC	470	436	414	389	500	459	439	417	500	455	440	414	500	455	438	413	413
LORANSAC	473	435	413	390	500	460	439	418	500	456	440	414	500	455	439	412	412
N = 1000																	
Real model	1000	900	850	800	1000	900	850	800	1000	900	850	800	1000	900	850	800	800
RANSAC	947	864	835	778	1000	917	880	832	1000	913	877	833	1000	918	882	824	824
MSAC	944	850	833	771	1000	914	878	829	1000	912	875	829	1000	917	881	821	821
RRANSAC	948	860	835	772	1000	917	879	830	1000	914	877	830	1000	917	883	824	824
NAPSAC	940	856	829	764	1000	916	869	827	1000	910	875	807	1000	915	881	822	822
LOMSAC	944	857	836	779	1000	919	878	829	1000	913	876	831	1000	917	881	821	821
LORANSAC	941	857	834	745	1000	919	878	829	1000	913	877	834	1000	917	881	821	821

It is observed from the table except least square almost all other robust method estimated the number of inliers exactly when  $\epsilon = 0.0$ . Number of inliers estimated by LS methods is around 90-95%. When  $\epsilon$  increase LS method get



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 2, Issue 7, July 2014

mostly affected. While considering various RANSAC techniques, performance of NAPSAC is very poor to estimate the number of inliers under various models. No of inliers estimated by RANSAC and R-RANSAC methods provides the better result compare with other methods.

## V. CONCLUSION

In this paper the behaviour of various RANSAC techniques such as RANSAC, R-RANSAC, MSAC, LO-MSAC, NAPSAC, LO-RANSAC discusses through a simulation study under various models LS, LMS, LTS and M-estimator. The study concluded that the robust models LMS, LTS and M-estimator provides better results when compare to least square with or without error. Among various RANSAC techniques RANSAC and R-RANSAC provides better solutions. One can apply this various RANSAC techniques and select proper robust models such as LMS, LTS, M-estimator and find better solution in the context of studying image processing application

## REFERENCES

1. Box, G.E.P., "Non- normality and tests of variance", *Biometrika*, pp.40, 318–335, 1953.
2. Chum O., and Matas J., "Randomized RANSAC with Td,d test", In *Proceedings of the 13th British Machine Vision Conference (BMVC)*, pp.448–457, 2002.
3. Chum O., Matas J., and Kittler J., "Locally optimized RANSAC", *Proceedings of DAGM-Symposium*, pp.236–243, 2003.
4. Chum O., and Jiri Matas, "Optimal randomized RANSAC", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(8), pp.1472–1482, 2008.
5. Fischler M.A., and Bolles R.C., "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", *Communication ACM*, 24(6), pp.381-395, 1981.
6. Hampel, F. R., "The influence curve and its role in robust estimation", *J.Am. Stat. ASOC.*, 69, 383-393, 1974.
7. Huber, P.J., "Robust estimation of a location parameter", *Ann.Math.Stat*, 35, 73-101, 1964.
8. Huber P.J., "Robust Statistics", Wiley New York, 1981.
9. Jureckova J. and Jan picek, "Robust statistical methods with R", Chapman & Hall/CRC, 2006.
10. Matas J., and Chum O., "Randomized RANSAC with sequential probability ratio test", *Proceedings of the 10th IEEE International Conference on Computer Vision (ICCV)*, 2005.
11. Myatt D.R., Torr P.H.S., Nasuto S.J., Bishop J.M., and Craddock R., "NAPSAC: High noise, high dimensional robust estimation", *Proceedings of the 13th British Machine Vision Conference (BMVC)*, pp.458-467, 2002.
12. Ondrej Chum, Jiri Matas, and Stepan Obdrzalek, "Enhancing RANSAC by generalized model optimization", *Proceedings of the Asian Conference on Computer Vision (ACCV)*, 2004.
13. Raghav Subbarao and Peter Meer, "Beyond RANSAC: User independent robust regression", *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop (CVPRW)*, 2006.
14. Rahul Raguram, Jan-Michael Frahm, and Marc Pollefeys, "A comparative analysis of RANSAC techniques leading to adaptive real-time random sample consensus", *Proceedings of the European Conference on Computer Vision (ECCV)*, 2008.
15. Rousseeuw P.J., "Least Median of Squares Regression", *Journal of the American Statistical Association* 79: pp.871-880, 1984.
16. Rousseeuw P.J., and Leroy A., "Robust Regression and outlier detection", John Wiley & Sons, New York, 1987.
17. Rousseeuw P. J., and Croux C., "Alternatives to the Median Absolute Deviation", *Journal of the American Statistical Association* 88(424): pp.1273-1283, 1993.
18. Stewart C.V., 'Robust Parameter Estimation in Computer Vision', *SIAM Review*, 41(3):pp.513-537, 1999.
19. Sunglok Choi, Taemin Kim and Wonpil Yu, "Performance Evaluation of RANSAC Family", *Proceedings of the British Machine Vision Conference (BMVC)*, 2009.
20. Torr, P. and Zisserman, A., "MLESAC: A New Robust Estimator with Application to Estimating Image Geometry", *Computer Vision and Image Understanding*, 78(1), 138-156, 2000.
21. Tukey, J.W., "A survey of sampling from contaminated distributions", *Contributions to Probability and Statistics*. Stanford University Press, Stanford, California, 1960.
22. Tarsha-Kurdi F., Landes T., and Grussenmeyer P., "Extended RANSAC algorithm for Automatic detection of building Roof planes from LIDAR Data", *The Photogrammetric Journal of Finland*. Vol. 21, pp.97-109, 2008.