Comparative Study on RANSAC and Mean Shift Algorithm

Daniela Fuiorea[1], Vasile Gui[2], Dan Pescaru[3], Toma Corneliu[4]

Abstract – **This paper is a comparative study between two important robust methods: the RANSAC algorithm and the mean shift algorithm. These methods are used in an image registration technique. The purpose is to demonstrate that mean shift could replace with success the RANSAC algorithm. These techniques are analyzed and tested for performance evaluation.**

Keywords: RANSAC, mean shift, image registration, robust estimation

I. INTRODUCTION

The geometric alignment or registration of images is a fundamental task in numerous applications such as medicine, computer vision, surveillance, etc. Many computer vision algorithms include a robust estimation step where model parameters are computed from a data set containing a significant proportion of outliers. By the term *robust estimation* we mean estimation techniques which are robust with respect to the presence of gross errors in the data. In this context, gross errors are defined as observations which do not fit to the stochastic model of parameter estimation. In response to the need for robustness in statistical analysis, nonparametric methods have been widely used in computer vision problems. Nonparametric methods estimate the underlying distributions from the data without any assumptions about the structures of the distributions. Nonparametric robust methods in computer vision have a history at least two decades old using standard methods in the statistical community: for example, M-Estimators [1] and Least Median of Squares [2], [3]. At least one novel technique, RANSAC [4] was developed by vision researchers in the early days (and is still a widely used technique). Another nonparametric probability density mode location estimation method, the mean shift, is employed in different computer vision tasks (e.g., color space analysis [5], face tracking [6], and registration [7]). In [7] we proposed a mean shift based solution for robust parameter estimation in image registration.

The RANSAC algorithm introduced by Fishler and Bolles in 1981 is possibly the most widely used robust estimator in the field of computer vision. RANSAC has been applied in the context of short baseline stereo [8], [9], wide baseline stereo matching [10], [11], [12], motion segmentation [13], mosaicing [14], detection of geometric primitives [15], robust image matching [16]. RANSAC does not need to start from an initial estimate, but the solution does not take into account all the available data, thus its precision is not maximized. Like RANSAC, the mean shift estimator does not require an initial estimate. At the same time, as the (related) M-estimators, the mean shift estimator makes a better use of the available inlier samples.

The point mapping technique is a primary approach taken to register two images when the type of misalignment is unknown. The general method consists of three steps. In the first step features in the images are computed. The second step is identifying feature correspondences in pairs of images and the last step is estimating parameters of geometrical transforms optimally mapping features between pairs of images. For the last step we can use two algorithms: the RANSAC algorithm and the mean shift algorithm, as proposed in [7]. In order to assess the performances of the two last methods and to observe which one gives the best estimate, a comparative study is carried out in Chapter 4 of this paper. The registration algorithm used to test these robust techniques can be found in [17]. In Chapter 2 a short description of the RANSAC algorithm is presented, followed in Chapter 3 by a short description of the mean shift algorithm. Finally, conclusions of this work are presented in Section 5.

[1] Facultatea de Electronică şi Telecomunicaţii, Departamentul Comunicaţii Bd. V. Pârvan Nr. 2, 300223 Timişoara, e-mail daniela.fuiorea@etc.upt.ro
[2] Facultatea de Electronică şi Telecomunicaţii, Departamentul Comunicaţii Bd. V. Pârvan Nr. 2, 300223 Timişoara, e-mail vasile.gui@etc.upt.ro
[3] Facultatea de Electronică şi Telecomunicaţii, Departamentul Comunicaţii Bd. V. Pârvan Nr. 2, 300223 Timişoara, e-mail dan@cs.utt.ro
[4] Facultatea de Electronică şi Telecomunicaţii, Departamentul Comunicaţii Bd. V. Pârvan Nr. 2, 300223 Timişoara, e-mail corneliu.toma@etc.upt.ro

II. BRIEF REVIEW OF THE RANSAC

RANSAC is an abbreviation for "RANdom SAmple Consensus". It is an iterative method to estimate parameters of a mathematical model from a set of observed data which contains outliers. The structure of the RANSAC algorithm is simple but powerful. Repeatedly, subsets are randomly selected from the input data and model parameters fitting the sample are computed. The size of the random samples is the smallest sufficient for determining model parameters. Therefore, the chance of contamination with outlier data is minimized. In a second step, the quality of the model parameters is evaluated on the full data set. Different cost functions may be used for the evaluation, the standard being the number of inliers, i.e. the number of data points consistent with the model. The process is terminated when the likelihood of finding a better model becomes low. To find the model, a threshold t is used, which is the number of compatible points used to imply that the correct model has been found.

The strength of the method stems from the fact that, to find a good solution, it is sufficient to select a single random sample not contaminated by outliers. Depending on the complexity of the model (the size of random samples) RANSAC can handle contamination levels well above 50%, which is commonly assumed to be a practical limit in robust statistics [18].

The speed of RANSAC depends on two factors. Firstly, 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. Secondly, the time spent evaluating the quality of each of the hypothesized model parameters is proportional to the size N of the data set. Typically, a very large number of erroneous model parameters obtained from contaminated samples are evaluated. Such models are consistent with only a small fraction of the data.

III. BRIEF REVIEW OF THE MEAN SHIFT

Given a sample of N d-dimensional data points, $\mathbf{x}_i$, drawn from a distribution with multivariate probability density function $p(\mathbf{x})$, an estimate of this density at $\mathbf{x}$ can be written as [4]:

$$\hat{p}_{\mathbf{H}}(\mathbf{x}) = \frac{1}{N} \sum_{i=1}^{N} K_{\mathbf{H}}(\mathbf{x} - \mathbf{x}_i) \qquad (1)$$

where

$$K_{\mathbf{H}}(\mathbf{x}) = |\mathbf{H}|^{-1/2} \, K_{\mathbf{H}}\left(\mathbf{H}^{-1/2}\mathbf{x}\right) \qquad (2)$$

is the kernel function depending on the symmetric positive definite $d \times d$ matrix $\mathbf{H}$, called bandwidth matrix. Frequently $\mathbf{H}$ has a diagonal form or even the form $\mathbf{H} = h^2 \mathbf{I}$, assuming the same scale h for all dimensions, i.e. a single scale parameter and an isotropic estimator, K_h. A radially symmetric estimator can be generated starting from a 1D kernel function K_1 as:

$$K^R(\mathbf{x}) = \alpha K_1(\|\mathbf{x}\|), \qquad (3)$$

where α is a strictly pozitive constant chosen such that the kernel function integrates strictly to 1. The profile of the radially symmetric kernel is defined as:

$$K^R(\mathbf{x}) = c_{k,d} k(\|\mathbf{x}\|^2), \qquad (4)$$

with $c_{k,d}$ a normalization constant.

Starting from any location $\mathbf{y}$, a gradient ascent *mean shift* algorithm can be used to find the location of the maxima of the estimated PDF closest to the starting location. This can be simply done by iterating the equation

$$\mathbf{y}_{j+1} = \frac{\sum_{i=1}^{n} \mathbf{x}_i g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)}{\sum_{i=1}^{n} g\left(\left\|\frac{\mathbf{y}_j - \mathbf{x}_i}{h}\right\|^2\right)}, \quad j = 1,2,... \qquad (5)$$

where

$$g(x) = -k'(x) \qquad (6)$$

until convergence. The proof of the convergence can be found in [19]. More, in practice the convergence is very fast, typically only two or three iterations being needed.

IV. COMPARATIVE STUDY

To test the performances of the two presented robust methods and to observe which method is providing a good estimate for the parameters of the similarity transformation, a comparative study was made on image pairs containing a common field of view, obtained for different camera positions and orientations. An example is given in Fig. 1.

Rotation 0°   Rotation 60°, Scale 70%

Rotation 0°	Rotation 45°, Scale 80%
Rotation 0°	Rotation 45°, Scale 0%
Rotation 0°	Rotation 60°, Scale 120%
Rotation 0°	Rotation 90°, Scale 80%
Rotation 0°	Rotation 45°, Scale 80%

Fig. 1. Images used in the registration algorithm

The threshold t used in the RANSAC algorithm is an unspecified parameter. To test the performance of the algorithm, an optimal threshold has to be found. The level of this threshold highly influences the elimination of points. If the threshold is too high, many outliers could be chosen and considered being inliers, allowing many false matches. This results in increased processing time and increased percentage of outliers, which in turn affects the registration accuracy. Conversely, a low threshold results in missing valid correspondences. Therefore, choosing an optimal level of the threshold has to be addressed.

To find the optimal threshold, we carried out eight experiments. Each experiment was realized for six images (figure 1) for different constant values – 1, 2 and 3 – of the threshold T. A different number of combinations between corresponding and mismatched points were used. The total number of points in the experiments was set to: 6, 7, 10, 15, 21, 25 and 35. The mean shift algorithm is tested for the same

number of points as the RANSAC, in the second series of experiments. Both methods are working with unknown information about the number of corresponding or mismatched points.

The experiments are organized in two groups. The small sets of 6, 7 and 10 points were used in the first stage of the experiments. The mean square error of the parameter vector components for these experiments is presented in Table I, II, III and IV. The point sets between 15 and 35 were used in the second stage of the experiments. Given the higher number of points, the results of these tests for all parameter solution vector components are given in graphical form in Fig. 2.

TABLE I
MEAN SQUARE ERROR FOR RANSAC – THRESHOLD 1

Number of points	RANSAC			
	Tx	Ty	Scale	Angle
3 correct – 3 false	5.1239	41.1440	0.1302	0.0041
3 correct – 4 false	24.3461	132.908	0.3041	1.4116
4 correct – 6 false	157.041	78.6692	0.2384	1.4861
6 correct – 4 false	41.1761	162.312	0.1831	1.4781

TABLE II
MEAN SQUARE ERROR FOR RANSAC – THRESHOLD 2

Number of points	RANSAC			
	Tx	Ty	Scale	Angle
3 correct – 3 false	36.9889	133.087	0.4155	1.3192
3 correct – 4 false	94.3188	202.099	0.3319	2.0442
4 correct – 6 false	94.0327	80.8905	0.1846	0.9050
6 correct – 4 false	1.1350	1.5727	0.0071	0.0089

TABLE III
MEAN SQUARE ERROR FOR RANSAC – THRESHOLD 3

Number of points	RANSAC			
	Tx	Ty	Scale	Angle
3 correct – 3 false	49.7444	173.672	0.4912	1.5493
3 correct – 4 false	98.5106	202.133	0.4232	2.0506
4 correct – 6 false	32.5009	74.142	0.0580	0.6647
6 correct – 4 false	1.3544	1.7269	0.0070	0.0072

TABLE IV
MEAN SQUARE ERROR FOR MEANSHIFT

Number of points	MEANSHIFT			
	Tx	Ty	Scale	Angle
3 correct – 3 false	0.9358	1.2581	0.0310	0.0197
3 correct – 4 false	1.1969	1.3238	0.0273	0.0109
4 correct – 6 false	8.1557	2.0958	0.0506	0.0411
6 correct – 4 false	1.4601	1.4834	0.0158	0.0108

Regarding the RANSAC algorithm, the results from Tables I to III, are presenting the mean square error for parameters of the similarity transform. The error for horizontal and vertical translation is different from a threshold to another. The values are very high for every choice of the threshold (1, 2 and 3), starting from 1.13 to 202. The smallest values of the mean square error for horizontal and vertical translation, are present in the case were the threshold is 2 or 3 and the number of points is 10 (the number of corresponding points is bigger than the number of mismatched points). When T=1, a substantially higher error is obtained. The value of the mean square error for 10 points is 41.1761.

The scale estimate presents a mean square error close to 0 for *T*=2 and *T*=3, when the number of points is 10 (the number of corresponding points is bigger than the number of mismatched points). The values obtained are 0.0071 for *T*=2 and 0.0072 for *T*=3. When *T*=1, the value of the mean square error for 10 points is 0.1831.

The angle estimate presents a mean square error close to 0 for *T*=2 and *T*=3, when the number of points is 10 (4 corresponding points and 6 mismatched points, 6 corresponding points and 4 mismatched points). For *T*=1 these values are bigger.

From these tables we can observe that the errors for *T*=2 are the best ones, having the smallest values.

In table IV, the mean square error for all the parameters of the similarity transform for the mean shift algorithm is presented. These errors are smaller than the errors provided by the RANSAC algorithm in all cases of the threshold.

In figure 2 the errors for all parameter solution vector components are presented. The number of points used in the experiments is between 15 and 35.
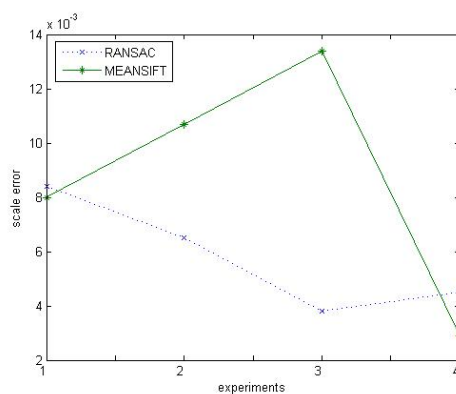
The mean shift algorithm performs better for the vertical translation estimation error (a). The present errors are smaller than the ones given by the RANSAC. For the scale, angle and horizontal translation estimation error (b, c, d), both methods present similar errors.
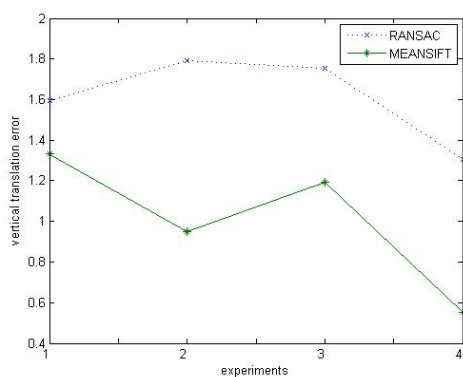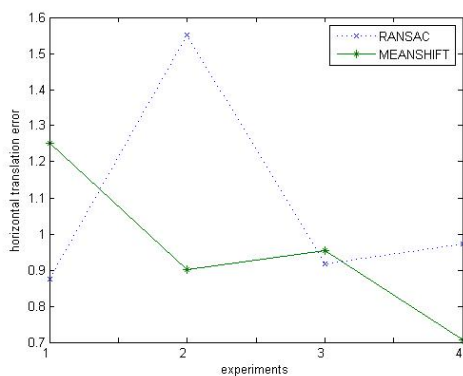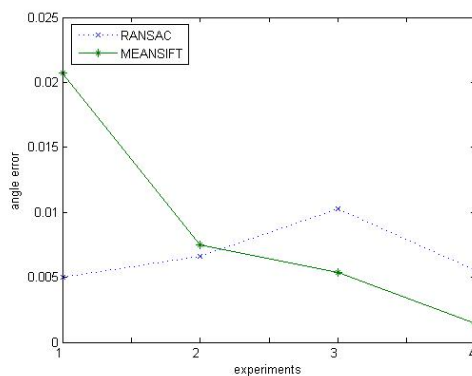
(c)

(d)

Fig. 2. (a) Vertical translation error, (b) Horizontal translation error, (c) Scale estimation error, (d) Angle estimation error

(a)

(b)

V. CONCLUSIONS

This paper is a comparative study between two important robust methods: the RANSAC algorithm and the mean shift algorithm. In computer vision, the most widely robust estimator used in the image registration is RANSAC. The mean shift has proven to be successful in image segmentation. Therefore, demonstrating that mean shift could replace with success RANSAC algorithm is an important task.

Practical tests demonstrate that for a small number of points, the mean shift algorithm presents much better results than RANSAC. When the number of points is bigger both methods have similar results. A conclusion is that mean shift could replace RANSAC with success.

As a future work we plan to extend this research and compare the mean shift algoritm with the latest versions of RANSAC.

REFERENCES

[1] M. Black and A. Rangarajan, "On the unification of line processes, outlier rejection, and robust statistics with applications in early vision". *International Journal of Computer Vision*, 19(1), 1996, pp. 57-92.

[2] P. Meer, D. Mintz, D. Kim, and A. Rosenfeld, "Robust regression methods in computer vision: A review", *International Journal of Computer Vision*, 6, 1991, pp. 59-70.

[3] A. Bab-Hadaishar and D. Suter, "Robust optic flow computation", *International Journal of Computer Vision*, 29(1), 1998, pp. 59-77.

[4] M.A. Fischler and R.C. Bolles, "Random consensus: A paradigm for model fitting with applications to image analysis and automated cartography", *Communications of the ACM*, 24(6), 1981, pp. 381-395.

[5] D. Comaniciu, P. Meer, "Robust Analisys of Future Spaces: Color Image Segmentation", Proc. *IEEE Conf on Comp. Vis. And Pattern Recognition*, Puerto-Rico, pp. 750-755, 1997.

[6] G.R. Bradski, "Real Time Face and Object Traking as a Component of a perceptual User Interface", Proc *IEEE Workshop on Applications of Comp. Vis.*, Princeton, pp. 214-219, 1998.

[7] D. Fuiorea, D. Pescaru, V. Gui, T. Corneliu, "Feature based 2D image registration using mean shift parameter estimation", Buletinul Ştiinţific al Universităţii "Politehnica" din Timişoara, Tom 51(65), 2006.

[8] P. H. S. Torr, "Outlier Detection and Motion Segmentation", PhD thesis, Dept. of Engineering Science, University of Oxford, 1995.

[9] P.H.S. Torr, A. Zisserman, and S.J. Maybank, "Robust detection of degenerate configurations while estimating the fundamental matrix", *CVIU*, 71(3):312–333, September 1998.

[10] P. Pritchett and A. Zisserman, "Wide baseline stereo matching", *Proc. International Conference on Computer Vision*, pp. 754–760, 1998.

[11] Tinne Tuytelaars and Luc Van Gool, "Wide baseline stereo matching based on local, affinely invariant regions", *Proc. 11th British Machine Vision Conference*, 2000.

[12] F. Schaffalitzky and A. Zisserman, "Viewpoint invariant texture matching and wide baseline stereo", *Proc. 8th International Conference on Computer Vision, Vancouver, Canada*, July 2001.

[13] P. H. S. Torr, "Outlier Detection and Motion Segmentation". PhD thesis, Dept. of Engineering Science, University of Oxford, 1995.

[14] Tinne Tuytelaars and Luc Van Gool, "Wide baseline stereo matching based on local, affinely invariant regions", *Proc. 11th British Machine Vision Conference*, 2000.

[15] J. Clarke, S. Carlsson, and A. Zisserman, "Detecting and tracking linear features Efficiently", Proc. 7th BMVC, pp. 415–424, 1996.

[16] A. Leonardis, H. Bischof, and R. Ebensberger, "Robust recognition using eigenimages", Technical Report TR-47, PRIP, TU Wien, 1997.

[17] D. Fuiorea, D. Pescaru, V. Gui, T. Corneliu, "Using Registration Algorithms for Wireless Sensor Network Node Localization", Proceedings of the 4th International Symposium on Applied Computational Intelligence and Informatics, May 17-18, 2007, pp. 209-214, ISBN 1-4244-1234-X.

[18] P. Rousseeuw and A. Leroy, "Robust Regression and Outlier Detection. Wiley", 1987.

[19] V. Lepetit and P. Fua, "Monocular Model-Based 3D Tracking of Rigid Objects: A Survey". In *Foundations and Trends in Computer Graphics and Vision* Vol. 1, No 1 (2005) 1–89.