HOMOGRAPHY ESTIMATION BASED ON ERROR ELLIPTICAL DISTRIBUTION

Lulu Mao[1], Haijiang Zhu[1], Fuqing Duan[2]

[1] *Beijing University of Chemical Technology* [2]*Beijing Normal University*

ABSTRACT

How to estimate accurately the homography is always a challenging problem in computer vision. In the reported literature, the measurement error of the image points is usually assumed to obey isotropic Gaussian distribution. However, real data very seldom follows this assumption. This paper proposes an estimation of homography under the assumption of image point errors following elliptical distribution, which is more coincident with real data. In the proposed method, the adaptive-scale elliptical residual kernel consensus (ASERKC) robust estimator is used to filter out inliers which are utilized to compute homography. Then, the elliptical weighted L-M (EW L-M) algorithm is optimized the homography. The experimental results show that the proposed method may present a more accurate homography. Especially when we applied it to incremental structure-from-motion (SFM), we find that the exact homography matrix is useful to select a better initial image pairs which can help obtain a more complete 3D points cloud.

***Index Terms*—** Homography, Elliptical distribution, Adaptive-scale elliptical residual kernel consensus, Elliptical weighted L-M

1. INTRODUCTION

In the field of computer vision, the homographic matrix represents the corresponding relationship in any two images of the same planar surface in space [1]. Homography estimation is required in many computer vision tasks like image stitching [2], 3D reconstruction [3], camera calibration [4], scene understanding [5] etc. So estimating an accurate homography is a fundamental and crucial issue [6], which attracts a large number of researchers.

A number of algorithms [7][8][9] for estimating the homography were proposed in previous studies. Feature-based homography estimation is a hot research topic now [20]. The Harris corner [10], SIFT [11], SURF [13] is a commonly feature used in homography estimation. In order to make the process of homography estimation more robust, LMeds [14], RANSAC [15], MLESAC [16], ASRC [17], ASKC [18] are utilized to deal with the data containing outliers. In addition, many deep neural networks methods, like [30] [31] [32], have emerged.

Most of algorithms assume that the measurement error obeys isotropic Gaussian distribution. But real data very seldom satisfy this assumption [19]. The covariance weighted MLESAC (CW MLESAC) [19] was proposed to solve the problems in which the measurement error obeys anisotropic Gaussian distribution. However, the non-Gaussian distribution problem still remains to be solved.

In probability and statistics, elliptical distribution generalizes the multivariate normal distribution. And it includes a lot of symmetric distribution like the multivariate t-distribution, multivariate stable distribution. And it is commonly used in statistics and mathematical economics [21][22][23][24]. However, elliptical distribution is rarely applied to estimate homography up to now.

In this paper, we assume that the measurement error of image feature points follows the elliptic distribution. This assumption is very reasonable because the elliptical distribution contains multivariate normal distribution, which makes it not conflict with traditional assumptions. At the same time, the elliptical distribution also contains many other distributions which make it closer to the actual distribution of error.

Based on the assumption above, we present a novel adaptive-scale elliptical residual kernel consensus (ASERKC) robust estimator, which integrates CW MLESAC and ASKC algorithm. In addition, we also propose an elliptical weight L-M (EW L-M) method to optimize results of ASERKC. The experiment shows our method can obtain more accurate homography matrix than state-of-art methods such as LMeds + L-M, RANSAC+L-M, MLESAC+L-M, ASKC+L-M, CWMLESAC+CW L-M and Unsupervised Deep Homography [32].

The paper is structured as follows: Section 2 describes ASERKC robust estimator and EW L-M algorithm in detail. And the experiment results on both synthetic and real data sets are presented in Section 3. Finally, a conclusion is provided in Section 4.

2. PROPOSED METHOD

In this section, a brief introduction of elliptical distribution is given in the first part. Then the proposed method is introduced in details.

2.1. Elliptical Distribution

The definition of elliptical distribution is as follows:
Definition: A random vector X has an elliptical distribution if its characteristic function ϕ satisfies the following functional equation (for every column-vector t):

$$\phi_{X-\mu}(t) = \phi\left(t^T \Sigma t\right) \tag{1}$$

Where μ is the location parameter and Σ is nonnegative-definite matrix which is proportional to the covariance matrix if the latter exists [25][26]. The notation $X \sim E_p(\mu, \Sigma, \phi)$ is commonly used to indicate that X obeys elliptical distribution.

The nonnegative-definite matrix can be replaced by the covariance matrix because the ratio between them is a constant [21]. There are a lot of methods [19][27] to estimate the covariance matrix for different features.

2.2. Adaptive-scale Elliptical Residual Kernel Consensus

The elliptical distribution is a broad family which contains many other distributions. The nonparametric method which was used in ASKC method can greatly reduce the impact of different distributions. But it did not concern the distribute information. The proposed method integrates CW MLESAC and ASKC algorithms to estimate an accurate result.

The ASERKC estimator can be written as:

$$\hat{H} = \arg\max_{\hat{H}} \frac{1}{N} \sum_{i=1}^{N} \frac{1}{h_{\hat{H}}} K\left(\frac{er_{i,\hat{H}}}{h_{\hat{H}}}\right) \tag{2}$$

Where $\hat{H}$ is the estimated homography. N is the number of data points. $h_{\hat{H}}$ is the bandwidth that varies with $\hat{H}$. And $er_{i,\hat{H}}$ is the elliptical residuals, which can be written as:

$$er_{i,\hat{H}}^2 = \varepsilon_i^T U \begin{bmatrix} \frac{\sqrt{\lambda_1}}{\lambda_1\left(\sqrt{\lambda_1} +_1 \sqrt{\lambda_2}\right)} & 0 \\ 0 & \frac{\sqrt{\lambda_2}}{\lambda_2\left(\sqrt{\lambda_1} +_1 \sqrt{\lambda_2}\right)} \end{bmatrix} U^T \varepsilon_i \tag{3}$$

Here, $\varepsilon_i = X_i^{'} - \hat{X}_i^{'}$ is the traditional residual. λ_1, λ_2 and U are the eigenvalues and eigenvector matrix of the covariance matrix respectively. The elliptical residual formula represents an ellipse, while the eigenvalues of covariance is the square of the long and short axes.

As shown in Fig.1, the elliptic residual formula can be considered as a combination of the following two steps: Firstly, the traditional residuals are rotated by the eigenvector matrix U. Then the weights, which are determined by the eigenvalues, are added to the result of first step. The purpose of the first step is to eliminate the correlation between the variables in the traditional residuals. And the second step, on the one hand, can eliminate the scale in different axes. On the other hand, with a small weight can increase the tolerance of the residuals in the long axis direction, which avoid the inliers with larger errors being wrongly classified as outliers.

Then Just like the ASKC, a TSSE procedure was used to find the inliers and their scale. In order to estimate the scale, TSSE method need to find the peak and valley in residual density. As shown in Fig.2, the traditional residual formula does not consider the distribution of the measurement error. So the boundary between inliers and outliers is always vague. But the elliptical residuals can make the valley become more clear than traditional residuals. So the proposed method can always select good inliers.

In addition, it is worth to mention that our method does not need a specific error tolerance which makes it more convenient to apply.

2.3. Elliptical Weight Levenberg—Marquardt

The point transposed by ASERKC's result is closer to the real position but may be farther from the measuring point. L-M and CWL-M are more concerned with the distance between the estimated point and the measurement point. So they can't exploit the potential of ASERKC.

To solve this, we modified the objective function as:

$$E = \arg\min_{H} \sum_{i=1}^{N} \left[\frac{\sqrt{\lambda_2}\left(u_1^T \cdot \left(X_i^{'} - HX_i\right)\right)^2}{\lambda_1\left(\sqrt{\lambda_1} + \sqrt{\lambda_2}\right)} + \frac{\sqrt{\lambda_1}\left(u_2^T \cdot \left(X_i^{'} - HX_i\right)\right)^2}{\lambda_2\left(\sqrt{\lambda_1} + \sqrt{\lambda_2}\right)} \right] \tag{4}$$

We named the new method as EW L-M because it adds the elliptic weight to traditional objective function. Experiments show that the combination of ASERKC and EW L-M can estimate a more accurate result.

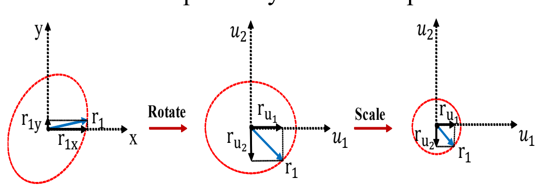
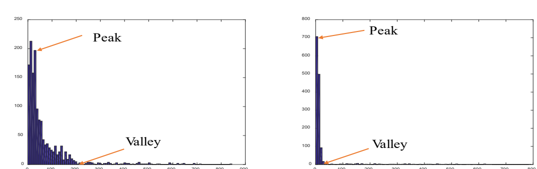


Fig.1. Workflow of the elliptical residuals



(a) (b)

Fig.2. Peaks and Valleys of different residuals: (a) the traditional residual (b) the elliptical residual

3. EXPERIMENT AND RESULTS

In order to analyze the accuracy and robustness of the proposed method, both synthetic data and real image were used in the experiment. The performance of the proposed method is compared with state-of-art methods like LMeds + L-M, RANSAC+L-M, MLESAC+L-M, ASKC+L-M and CWMLESAC+CWL-M. Then we applied it to the process of initial image pairs selection in the incremental structure-from-motion (SFM) system.

3.1. Evaluation Criteria

In this paper, two criteria were used to evaluate the accuracy of the estimated homography. The first is the root mean squares error (RMSE) of the re-projection error which is defined as follows:

$$RMSE = \sqrt{\sum_{i=1}^{N} \left\| \overline{X_i^{'}} - H_{est}\overline{X_i} \right\|_2^2 \Big/ N} \tag{5}$$

Where $\left(\overline{X_i}, \overline{X_i^{'}}\right)$ are the true corresponding points. And N is the number of correspondence. H_{est} is the estimated homography.

The second is homographic 2-norm [28] which is defined as:

$$L_2 = \left(\sum_{r \in S} \left\| r^{H_{est}} - r^{H_{true}} \right\|_2^2 \right)^{1/2} \tag{6}$$

Here r represents the point in the first image. S is the projection region of a plane in the first image. H_{est} is the estimated homography. H_{true} represents the ground truth.

Both two criteria focus on the difference between the estimated homography and the ground truth. The first is used to evaluate the results of the synthetic dataset and the second is for the real image experiment results. Because the synthetic dataset can provide the ground truth of correspondence but real image can't.

3.2. Synthetic Dataset Experiment

The purpose of using synthetic dataset was to analyze the accuracy of the proposed algorithm under different inlier ratios, different noise levels and different error distribution.

In this paper, the synthetic dataset is generated in a similar way to [19]. But the bivariate Cauchy noise were also added to the image points which is different with [19]. Some results of Gaussian noise and Cauchy noise are given in Fig.3and Fig.4, respectively.

As shown in Fig.3 and Fig.4, the error of LMeds is extremely large when the inlier ratios are less than 50%. This is because LMeds is unable to handle the data that have more than 50% outliers. As can be seen from the Fig.3, errors of CW MLESAC + CW L-M and ASERKC + EW L-M always have a lower magnitude than others. This happens because both two methods considered the distribution information of measurement error. But under the Cauchy distribution (as shown in Fig.4), ASKC+L-M's result is better than CWMLESAC + CWL-M when the inlier ratio is more than 0.7. This demonstrates that the assumption of CWMLESAC + CWL-M can't properly handle the case that the measurement error does not follow Gauss distribution. Although our hypothesis is broader than the Gaussian distribution, the proposed method can still obtain a better result when the measurement error obeys the Gaussian distribution as shown in Fig.4. This is because elliptical distribution is any member of a broad family of probability distributions. And the proposed method can automatically fit any ellipse distribution. Besides, it can be seen that the accuracy of ASERKC + EWL-M is always higher than CWMLESAC + CWL-M and ASKC+L-M at Cauchy distribution.

3.3. Real Image Experiment

In order to test the effect of our method in the real images, several pictures which come from the Oxford VGG Affine Covariant Regions dataset [28] were used. The Oxford VGG

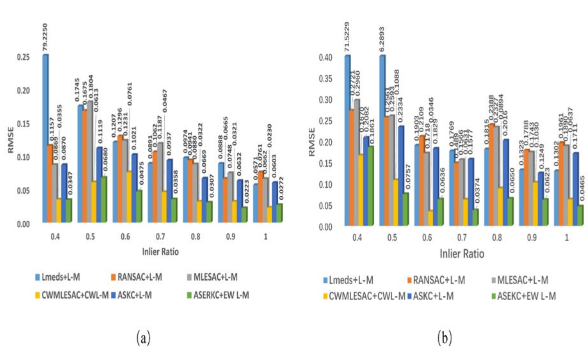


Fig.3. The RMSE of LMeds + L-M, RANSAC+L-M, MLESAC+L-M, ASKC+L-M, CWMLESAC+CWL-M and ASERKC+EW L-M under Gaussian noise in different inlier ratios and noise level is (a) 0.1 (b) 0.5

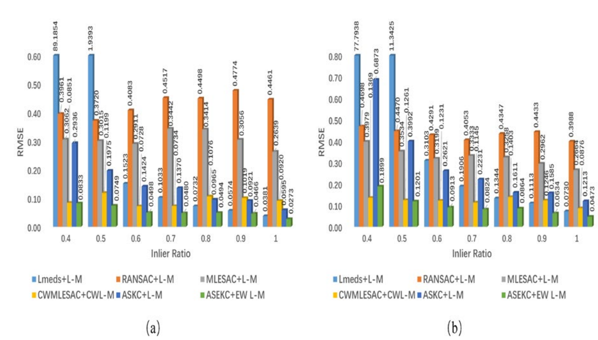


Fig.4. The RMSE of LMeds + L-M, RANSAC+L-M, MLESAC+L-M, ASKC+L-M, CWMLESAC+CWL-M and ASERKC+EW L-M under Cauchy noise in different inlier ratios and noise level is (a) 0.1 (b) 0.2

Affine Covariant Regions dataset contains viewpoint and scale changes. So it is widely used to evaluate the performance of homography estimation algorithms.

The real image experimental process is as follows:

First, the SIFT algorithm was adopted to extract the feature points. And the covariance matrix of each feature point was estimated by the method of [18]. Then, the KNN algorithm was used to get the initial correspondence. After that, the LMeds + L-M, RANSAC+L-M, MLESAC+L-M, CW MLESAC+CW L-M, ASKC+L-M and ASERKC+EW L-M were used to estimate the homography respectively. In order to compare the results estimated by Unsupervised deep homography, we add the pictures used in real image experiments to the test-set. Then compare the results with proposed methods. Results are shown in Fig.5.

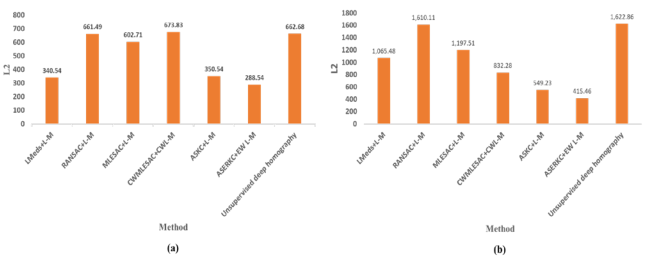


Fig.5. The homographic 2-norm L2 of LMeds + L-M, RANSAC+L-M, MLESAC+L-M, ASKC+L-M, CWMLESAC+CWL-M and ASERKC+EW L-M in real image pairs (a) graffiti 1to2 (b) graffiti 1to3

As can be seen from Fig.5, the CWMLESAC + CWL-M can't continue to give a precise result compared to other methods in the real image. But ASKC+L-M can do well. This is because ASKC focuses on the residual information of the image and has no limitation on the distribution of the measurement error. Therefore, it can obtain a better result in the real image experiment. However, it is clear that the proposed method ASERKC+EW L-M is superior to ASKC. Because we estimate it based on a more reasonable error distribution. In addition, it can be seen from the Fig.5 that the results of Unsupervised deep homography are not as good as those traditional methods. This may be due to the fact that the traditional method can still do well based on the high-precision SIFT feature points.

Both the simulation experiment and the actual image experiment show that it is more reasonable to assume that the measurement error obeys the elliptic distribution, and the new method based on this assumption is more effective.

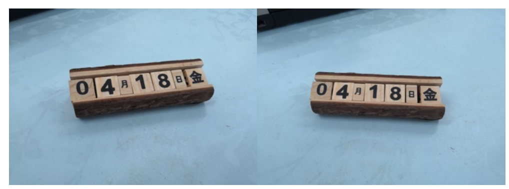
3.4. Reconstruction Experiment

In the incremental SFM, the selection of initial image pairs is extremely important. Most of SFM techniques need to compute the homography matrix to find a wide baseline image pair. Therefore, accurate homography matrix can really help to obtain a more suitable initial image pair and improve the accuracy of the 3D structures.

The OpenMVG C++ library [29] provides a vast collection of structure-from-motion techniques which make it easier to use. In this section, the proposed method is used to replace the part of the homography matrix estimation in the OpenMVG library. Then compared it with the method used in the library. 26 photographs of wooden calendars were used in this experiment. The results are shown in Fig.6 and Fig.7.

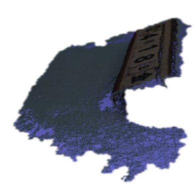
As can be seen in Fig.6, the baseline of the initial image pair selected by our method is significantly larger than the image pair selected by the traditional method. Besides, as red part shown in Fig.7, the 3D points cloud calculated by our method is more complete.

(a)

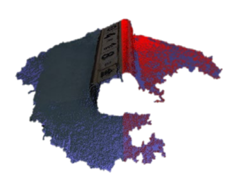


(b)

Fig.6. Initial image pair in structure-from-motion (a) choosed by OpenMVG library method (pairs 1 and 2), (b) choosed by ASERKC+EW L-M method (pairs 1 and 3).



(a)



(b)

Fig.7. The results of structure-from-motion using different homography estimation method (a) The OpenMVG's method (b) The ASERKC + EW L-M method

4. CONCLUSION

In this paper, we proposed the adaptive-scale elliptical residual kernel consensus(ASERKC) robust estimator and elliptical weight L-M (EW L-M) method to estimate the homography for the problem that the measurement error follows elliptical distribution. Compared with the state-of-the-art, the proposed method can estimate a more accurate model and can deal with different kind of distributions. We also demonstrate the effectiveness of the proposed method for incremental SFM techniques.

5. REFERENCES

[1] Chaofu Wu. Mathematical methods in computer vision. science press, 2008.

[2] Bonny, Moushumi Zaman, and M. S. Uddin. "Feature-based image stitching algorithms." International Workshop on Computational Intelligence IEEE, 2017:198-203.

[3] XU Lian-xia. "3D Reconstruction of Skin Surface from Image Sequence." Journal of Langfang Teachers University 83.7(2017):1415-1421.

[4] Jacobsen, K, and M. Gerke. "Sub-Camera Calibration of a Penta-Camera." EuroCOW 2016, the European Calibration and Orientation Workshop 2016.

[5] Aliakbarpour, Hadi, et al. "Heterogeneous Multi-view Information Fusion: Review of 3-D Reconstruction Methods and a New Registration with Uncertainty Modeling." IEEE Access PP.99(2016):1-1.

[6] Kukelova, Zuzana, et al. "Radial distortion homography." Computer Vision and Pattern Recognition IEEE, 2015:639-647.

[7] Chum, O, and J. Matas. "Homography estimation from correspondences of local elliptical features." International Conference on Pattern Recognition IEEE, 2012:3236-3239.

[8] Huang, Haifei, H. Zhang, and Y. M. Cheung. "Homography Estimation from the Common Self-Polar Triangle of Separate Ellipses." Computer Vision and Pattern Recognition IEEE, 2016:1737-1744.

[9] Liu, S, et al. "BB-Homography: Joint Binary Features and Bipartite Graph Matching for Homography Estimation." Circuits & Systems for Video Technology IEEE Transactions on 25.2(2014):239-250.

[10] Harris C. A combined corner and edge detector. Proc Alvey Vision Conf, 1988, 1988(3):147-151.

[11] Lowe, David G. "Distinctive Image Features from Scale-Invariant Keypoints." International Journal of Computer Vision 60.2(2004):91-110.

[12] Bay, Herbert, T. Tuytelaars, and L. V. Gool. "SURF: Speeded Up Robust Features." European Conference on Computer Vision Springer-Verlag, 2006:404-417.

[13] Peter J. Rousseeuw. "Least Median of Squares Regression." Publications of the American Statistical Association 79.388(1984):871-880.

[14] Fischler, Martin A., and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. ACM, 1981.

[15] Torr, P. H. S, and A. Zisserman. "MLESAC: A New Robust Estimator with Application to Estimating Image Geometry." Computer Vision & Image Understanding 78.1(2000):138-156.

[16] Wang, Hanzi, and D. Suter. "Robust Fitting by Adaptive-Scale Residual Consensus." 3023(2004):107-118.

[17] Wang, Hanzi, D. Mirota, and G. D. Hager. "A Generalized Kernel Consensus-Based Robust Estimator." IEEE Transactions on Pattern Analysis & Machine Intelligence 32.1(2009):178-184.

[18] Zeisl B, Georgel P F, Schweiger F, et al. Estimation of Location Uncertainty for Scale Invariant Feature Points. British Machine Vision Conference, BMVC 2009, London, UK, September 7-10, 2009. Proceedings. 2009.

[19] Zhao, Chunyang, and H. Zhao. Accurate and robust feature-based homography estimation using HALF-SIFT and feature localization error weighting. Academic Press, Inc. 2016.

[20] Kollo, and D. Von Rosen. Advanced Multivariate Statistics with Matrices. Springer Netherlands, 2005.

[21] Gupta, Arjun K.; Varga, Tamas; Bodnar, Taras (2013). Elliptically contoured models in statistics and portfolio theory (2nd ed.). New York: Springer-Verlag. doi:10.1007/978-1-4614-8154-6. ISBN

[22] Killiches, Matthias. Elliptical Distributions. AV Akademikerverlag, 2013.

[23] Sun, Ying, P. Babu, and D. P. Palomar. "Robust estimation of structured covariance matrix for heavy-tailed distributions." IEEE International Conference on Acoustics, Speech and Signal Processing IEEE, 2015:5693-5697.

[24] Stamatis Cambanis, Steel Huang, and Gordon Simons. "On the theory of elliptically contoured distributions." Journal of Multivariate Analysis 11.3(1981):368-385.

[25] Frahm, Gabriel, M. Junker, and A. Szimayer. "Elliptical copulas: applicability and limitations." Statistics & Probability Letters 63.3(2003):275-286.

[26] Orguner, Umut, and F. Gustafsson. "Statistical Characteristics of Harris Corner Detector." Statistical Signal Processing, 2007. Ssp '07. Ieee/sp, Workshop on IEEE, 2007:571-575.

[27] Mikolajczyk, K, and C. Schmid. "A performance evaluation of local descriptors." IEEE Trans. on 2013:257.

[28] Je, Changsoo , and H. M. Park . Homographic p-norms. Elsevier Science Inc. 2015.

[29] Pierre Moulon, et al. "OpenMVG: Open Multiple View Geometry." (2016):60-74.

[30] F. E. Nowruzi, R. Laganiere and N. Japkowicz, "Homography Estimation from Image Pairs with Hierarchical Convolutional Networks," 2017 IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, 2017, pp. 904-911.

[31] Daniel DeTone, Tomasz Malisiewicz, Andrew Rabinovich, "Deep Image Homography Estimation", RSS Workshop on Limits and Potentials of Deep Learning in Robotics, arXiv:1606.03798 [cs.CV]

[32] T. Nguyen, S. W. Chen, S. S. Shivakumar, C. J. Taylor and V. Kumar, "Unsupervised Deep Homography: A Fast and Robust Homography Estimation Model," in IEEE Robotics and Automation Letters, vol. 3, no. 3, pp. 2346-2353, July 2018.