Comparison of algorithms for elimination the outliers in 3D modelling from video

Svetlana Mijakovska[1], Filip Popovski[2], Roberto Pasic[3] and Ivo Kuzmanov[4]

Abstract – **In this paper, we make a comparison of different algorithms for the elimination of outliers in the process of 3D modeling from video. The process of 3D modeling from video uses specific algorithms for detecting and matching key points. The first step in this process is obtaining a cloud of points, then matching the key points, obtaining the network, and converting the network into a polygonal model. There are two types of points: inliers (points that describe the model) and outliers (unnecessary points). From the practical examples, we got results and conclusions for comparison of Random sample consensus (RANSAC), MEstimator Sample Consensus (MSAC), and Maximum Likelihood Sample Consensus (MLESAC) algorithms. Using different algorithms for eliminating the outliers in practical experiments we get different quality of the final 3D model.**

Keywords – **Inliers, outliers, key points, 3D model, Maximum Likelihood Sample Consensus (MLESAC), MEstimator Sample Consensus (MSAC), Random sample consensus (RANSAC).**

I. Introduction

The data information in 3D modelling from video is get from video. After that, we got a lot of points because data is processed into a point cloud data. That points must be connected with some method [1]. There are two types of points: inliers (points that describe the model) and outliers (unnecessary points).

The 3D reconstruction is composed of four main tasks:

First step is feature detection and matching. This step is from different images, finding out the same features and match them.

Second step is structure and motion recovery. In this step the scene structure and motion are recovered. In this step we get these features: orientation, position, and parameters of the camera at capturing positions and 3D coordinates of detected features.

Third step is stereo mapping. In this step, dense matching map is created.

Forth step is modelling. This step is making realistic model of scene. This includes building mesh models and mapping textures.[2]

Recovering motion information of the camera and the structure of the scene is goal of second task. That is structure and motion recovery. Because there are only a few features, reconstruction is possible with using projective reconstruction. Getting projective reconstruction can be from basic matrix. Or can be get from focal tensor. Focal tensor and fundamental matrices can be detected by the projective reconstruction. Data can be match using one of listed algorithms:

- Random sample consensus (RANSAC),
- MEstimator Sample Consensus (MSAC) and
- Maximum Likelihood Sample Consensus (MLESAC) [3].

The motion information of the camera and the structure of the scene are recovered in structure and motion recovery step. Information about motion is: intrinsic parameters of the camera at the captured views, orientation and position. From the 3D coordinates of features, structure information is captured. Because, video sequence is composed of many images, we must use multiple view geometry i.e. 3D reconstruction from multiple views. [4].

If there is knowledge of feature correspondence, reconstruction is possible with projective reconstruction. Projection matrices can be obtained in many ways, from a fundamental matrix or a focal tensor.

In Hartley and Zisserman [5] are discussed evaluations, methods and implementation. If there are outliers in the input, i.e. feature correspondences, than must be used robust methods like Random sample consensus (RANSAC), The Least Mean Square (LMS). There are two tasks in stereo mapping:

1. Rectification – preparing the data and align the pairs that are corresponded along the same scan line of images. That means all corresponding points will have same y-coordinate in two images.
2. Dense stereo mapping – faster searching and matching over the whole image [2].

The Least Mean Square (LMS) is an adaptive algorithm, which uses a gradient-based method of steepest decent. LMS algorithm uses the estimates of the gradient vector from the available data. LMS incorporates an iterative procedure that makes successive corrections to the weight vector in the direction of the negative of the gradient vector which eventually leads to the minimum mean square error.

Mapping texture on the model is the final step. Triangulation is used in this step. Depth maps are generated from points of each stereo map.[5] Those maps are important because they are used to construct the mesh of the scene. Using the texture from images, can be built finally textured model.

II. Review of the algorithms for elimination the outliers

There are different algorithms for elimination the unnecessary points – outliers in 3D modelling from video. The 3D model is creating from inliers – the points that describe the 3D model.

[1]Svetlana Mijakovska, [2]Filip Popovski, [3]Roberto Pasic and [4]Ivo Kuzmanov are with the Technical Faculty of Bitola, address: ul. Ivo Ribar Lola bb, 7000 Bitola, North Macedonia, E-mail: svetlana.mijakovska@uklo.edu.mk; filip.popovski@tfb.uklo.edu.mk; roberto.pasic@uklo.edu.mk; ivo.kuzmanov@tfb.uklo.edu.mk

978-1-6654-8500-5/22/$31.00 ©2022 IEEE

Connecting inliers into 3D model and eliminating the outliers can be made by using of these algorithms: RANSAC, MLESAC and MSAC.

Random sample consensus (RANSAC) is algorithm which continuously generate hypothetical solutions estimated from randomly selected, minimal data sets and testing each solution. That means the solution can be computed from the smallest sample and the likelihood of a sample containing distance is minimized. With testing many samples, measured the support for each sample and from corresponding points, the final model is estimated. So, this estimator is robust and capable method of dealing with datasets contaminated by large numbers of outliers [1].

Because RANSAC attempts to minimise the number of outliers and maximise the number of inliers, in effect the penalty for outliers is constant and the penalty for inliers is zero and (Table I).

Table I. Difference between RANSAC and MSAC

RANSAC	MSAC
$\rho(e^2) = \begin{cases} 0, & \text{for } e^2 < T^2; \\ \text{constant} & \text{for } e^2 \geq T^2. \end{cases}$	$\rho(e^2) = \begin{cases} e^2, & \text{for } {}^2 < T^2; \\ T^2 & \text{for } {}^2 \geq T^2. \end{cases}$

In Table I. $\rho(e^2)$ is robust error and T is threshold. The estimation is highly dependent on the threshold T and this not optimal. If the threshold was high enough, then all points would be inliers. And all solutions would have the support that is same.

A M–estimator known as MSAC (M–estimator SAmple Consensus) solves this problem with giving outliers a fixed penalty related to the threshold and scoring inliers according to their error. This is modification of RANSAC. It provides a number of hypothesised models that all contain the same number of inliers. In this case, the model with the best fit will score lowest [10].

Torr and Zisserman [6] proposed the Maximum Likelihood Sample Consensus (MLESAC) algorithm. It is a version of RANSAC that is probabilistic. Using mixture model, it maximises a likelihood. For inliers, the distance of the data from the proposed model is assumed to be Gaussian and for outliers, uniformly distributed with the mixing parameter being determined by expectation-maximisation [10].

After applying Maximum Likelihood Sample Consensus (MLESAC), the nonlinear minimization is conducted using the method that is modification of the GaussNewton [7]. In this method, in the minimization all the points are included and the effect of outliers are removed. This method allows outliers to be reclassed as inliers during the minimization itself without incurring additional computational complexity. In that way this method reduces the number of false classifications [10].

III. PRACTICAL EXAMPLES OF ALGORITHM COMPARASION

Using the program Voodoo Camera Tracker [7] in practical examples we compared previously described algorithms: RANSAC, MSAC and MLESAC for increase the robustness and their speed and RMSE.

The standard deviation of the prediction errors is Root Mean Square Error (RMSE). Prediction errors or residuals are a measure of how far data points are from the regression line, respectively it tells how the data is concentrated around the line of best fit. This deviation, Root mean square error, is used in many different fields (climatology, forecasting, and regression analysis) [12].

From Voodoo Camera Tracker (where we selected corner detector and algorithm for matching), we obtained information that are used in Video Trace [9] (program for 3D modelling from video). We use computer with Intel(R) Core (TM) i7-9750HF CPU and RAM 8.00 GB.

Practically, we obtained the following 3D models. (Fig.1 – Fig.8). The comparison of the algorithms for each model is presented in the Tables (Table II – Table V).

Fig.1 Example of a vase model

Table II. Comparison of the speed and RMSE parameter of the algorithms for elimination of unnecessary points (outliers)

Algorithm	Speed	RMSE
MSAC	35s	0,723
RANSAC	35s	0,989
MLESAC	30s	0,699

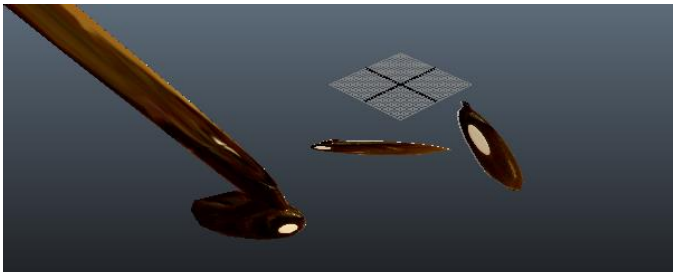

Fig.2 Comparison of different algorithms (MSAC, RANSAC and MLESAC) for a practical example of a vase

Fig.3 Example of a star model

Table III. Comparison of the speed and RMSE parameter of the algorithms for elimination of unnecessary points (outliers)

Algorithm	Speed	RMSE
MSAC	48s	0,352
RANSAC	50s	0,347
MLESAC	45s	0,346

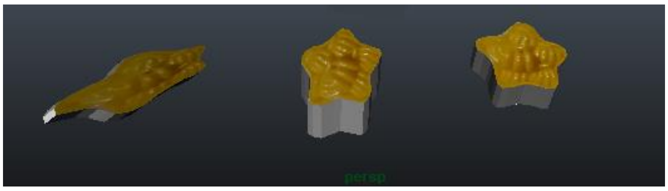

Fig.4 Comparison of different algorithms (MSAC, RANSAC and MLESAC) for a practical example of a star

Fig.5 Example of a cylinder model

Table IV. Comparison of the speed and RMSE parameter of the algorithms for elimination of unnecessary points (outliers)

Algorithm	Speed	RMSE
MSAC	100s	1,515
RANSAC	120s	2,600
MLESAC	20s	0,767

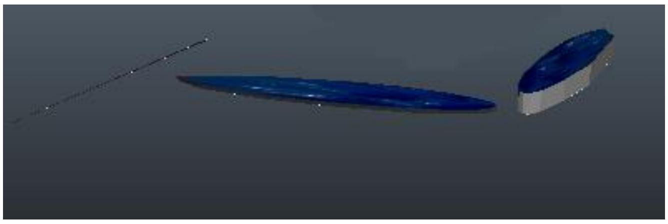

Fig.6 Comparison of different algorithms (MSAC, RANSAC and MLESAC) for a practical example of a cylinder

Fig.7 Example of a cube model

Table V. Comparison of the speed and RMSE parameter of the algorithms for elimination of unnecessary points (outliers)

Algorithm	Speed	RMSE
MSAC	6s	0,341
RANSAC	5s	0,387
MLESAC	5s	0,334

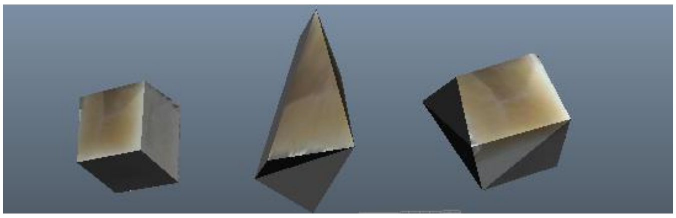

Fig.8 Comparison of different algorithms (MSAC, RANSAC and MLESAC) for a practical example of a cube

Practical examples have shown that the MLESAC algorithm improves the accuracy of existing methods, i.e the RANSAC algorithm, in terms of dealing with unnecessary points (outliers). Although the RANSAC algorithm is the original algorithm for eliminating unnecessary points (outliers), the main difference between RANSAC and MLESAC is that they have a different curve of the loss function near error zero. The RANSAC algorithm has a zero loss in the specified range, but the MLESAC algorithm has a loss increase (Fig. 9). In other words, the RANSAC algorithm does not take into account the quality of the required points (inliers), while the MLESAC algorithm takes into account the quality of the required points (inliers). For this reason, better results give the MLESAC algorithm (even in the case of higher noise) than the RANSAC algorithm. The difference of the loss functions is presented in the following graph Fig.9.

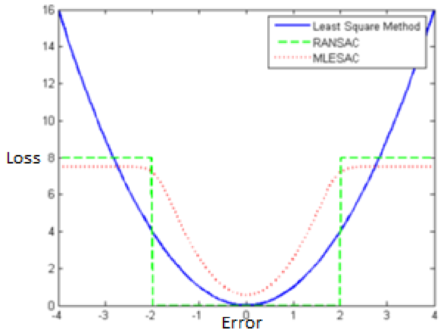

Fig.9 Graph of loss functions in Least Square, RANSAC and MLESAC

IV. Conclusion

From the previous tables for the speed characteristics and the Root Mean Square Error (RMSE) parameter, as well as from the generated 3D models of the different models (examples), it can be seen that the best results are given by the Maximum Likelihood Sample Consensus (MLESAC) algorithm for elimination of unnecessary points (outliers).

Because the value of the Root Mean Square Error (RMSE) parameter when using the MEstimator Sample Consensus (MSAC), Random sample consensus (RANSAC) algorithms is high, the shape of the resulting 3D model is not the same as the shape of the input data model.

REFERENCES

[1] Allyn Joy Calcaben, Jemwel Autor, & Jefferson Butch OberoComputer, 2018, Vision: Feature matching with RANdom Sample Consensus, University of the Philippines Visayas, April 2018.

[2] S.Mijakovska, I.Nedelkovski, F.Popovski, 2014, Generating 3D model from Video, Advanced Computing: An International Journal (ACIJ), Vol.5, No.5/6, November 2014.

[3] S.Mijakovska, I.Nedelkovski, F.Popovski, 2014, Overview of the process of 3D modelling from video, International Journal of Engineering Sciences & Emerging Technologies, Dec. 2014, ISSN: 2231 – 6604 Volume 7, Issue 3, pp: 680-686 ©IJESET.

[4] S. Mijakovska, R. Pasic, I. Kuzmanov, Influence of Corner Detectors in the Process of 3D Modeling from Video; International Journal of Scientific and Engineering Research, Volume 7, Issue 9, Sep 2016, pages 228-232, ISSN 2229-5518.

[5] Svetlana Mijakovska, Filip Popovski, Roberto Pasic and Ivo Kuzmanov "Triangulation Method in Process of 3D Modelling from Video"; Asian Journal of Social Science and Management Technology, 3 (6). pp. 16-21. ISSN 2313-7410.

[6] P. Torr and A. Zisserman. MLESAC: A New Robust Estimator with Application to Estimating Image Geometry. Computer Vision and Image Under- standing, 78(1):138–156, 2000.

[7] P. E. Gill and W. Murray. Algorithms for the solution of the nonlinear leastsquares problem. SIAM J Num Anal, 15(5):977–992, 1978.

[8] http://www.softsea.com/review/Voodoo-Camera-Tracker.html#:~:text=Voodoo%20Camera%20Tracker%20is%20a,camera%20parameters%20for%20video%20sequences.

[9] Van den Hengel, A., Dick, A. Thormahlen, T. Ward, B. and Torr, P. 2007. VideoTrace: Rapid Interactive Scene Modeling from Video.

[10] Oliver Daniel Cooper, Robust Generation of 3D Models from Video Footage of Urban Scenes, University of Bristol, March 2005.

[11] https://www.analyticsvidhya.com/blog/2021/02/new-approach-for-regression-analysis-ransac-and-mlesac/

[12] https://www.statisticshowto.com/probability-and-statistics/regression-analysis/rmse-root-mean-square-error/