TEMEL-ij ISSN 2545 - 4390
Technology | Engineering | Management | Entrepreneurship | Learning | International Journal

BASIC STEPS AND TECHNIQUES USING IN OBTAINING THE STRUCTURE FROM MOTION

Svetlana Mijakovska, Filip Popovski

Abstract: *Obtaining structure from motion (recovery three dimensional structure from video footage of scenes) has been a interesting area of research in computer vision. This problem is known as Structure from Motion (SfM). SfM is trying to recover, from segments of images the 3D structure of a scene and the position and pose (orientation) of camera at the moment of capturing each images (as a part of video sequences).*
This article presents the basic steps and techniques used in SfM (feature matching,structure and camera calculation, auto calibration and increase robustness).

Key words*: structure from motion, 3D model, feature matching, point matching, correspondences, essential matrix, fundamental matrix, tensor.*

1. INTRODUCTION

Structure from motion is used for recovered 3D models from video sequences. The purpose of structure from motion is getting cloud of 3D points in the scene, which in the processs of feature matching and meshing, give the final 3D model.
So, given the position of a feature in one image (sequence) we need to find the corresponding position of the same feature in successive image. This is correspondence problem and is based on principles of multiple view geometry.
Features can include points, lines, and higher level primitives, for example planes. Generally most common approach is based on points (Fig.1). Points features are extracted easy from images, using corner detector. There are many corner detectors,but most used is Harris corner detector [1].
When one points of interest have been identified, then is needed some similarity to compare candidate points between the images. There are more correlation techniques for similarity measure by comparing the intensity differences between pixels.
Commonly used method is Normalised Cross Correlation (NCC) [2], because this method is invariant to global changes in illumantion between the images (sequences).
Another method is KLT tracker [3], who select corner features based on the principe: a good feature can be well tracked, and the tracking should not be separated from feature extraction. This method measures the dissimilarity between feature in the first and current image and allow affine image changes.

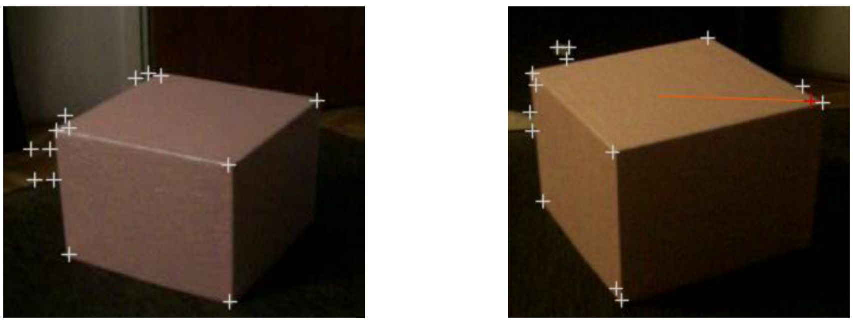

Fig. 1 *Found corner points as a feature points*

Corresponding author: Svetlana Mijakovska, Bitola, Faculty of Technical Science, e-mail: svetlana.mijakovska@uklo.edu.mk

2. STRUCTURE AND CAMERA CALIBRATION

Structure from motion is based on principes of multiple view geometry.
The projective geometry that describes the relations between two views of a scene is epipolar geometry (Fig.2), and depends only on the instrict parameters of the cameras and their pose relative to each other.

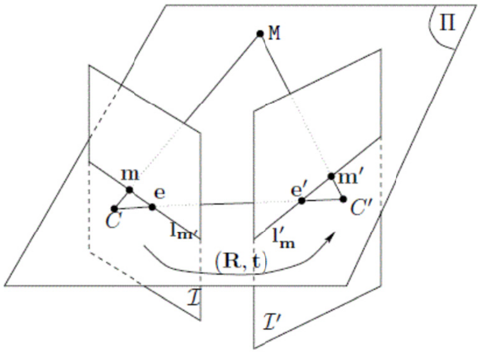

Fig. 2 *Epipolar geometry*

The matrix that encapsulated the epipolar geometry of two calibrated cameras is called *Essential matrix* [4]. Essential matrix can also be generalised to the case of uncalibrated cameras, and this matrix is *Fundamental matrix* [5]. This method is very usefull for recover projective structure of a scena and camera motion from image correspondences, without any knowledge of the camera parameters.
The three view equivalent to the fundamental matrix is the trifocal tensor, who depends only of the relative pose of the cameras and their intrinsic parameters. Three views geometry allow a overlap of two views between succesive image triplets in long sequences, and this simplifies method of stitching of a large numbers of views. (Fig.3).

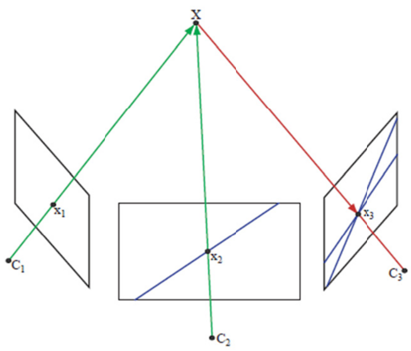

Fig. 3 *Trifocal tensor*

There is also quadrifocal tensor – geometry of four viewa. But, most cases of reconstruction from contain many more than four views and required techniquies is N-view geometry [6]. Fundamental matrix are calculated separately for overllapping pairs (triplets) of all images in the sequences. Onother, approach is to register all calculated tensors into the same coordinate frame.

3. AUTO CALIBRATION

For the projective reconstruction of a scene features the intrinsic parameters of the cameras must be recovered. By placing two constraints on some of intrinsic parameters, method of auto-calibration can recover the intrinsic parameters and upgrade projective reconstruction to metric reconstruction. Most methods of auto-calibration are based on the conceps of absolute conic, which means that its relative position to moving camera is constant and images in any view depends only on the intrinsic parameters of the camera.

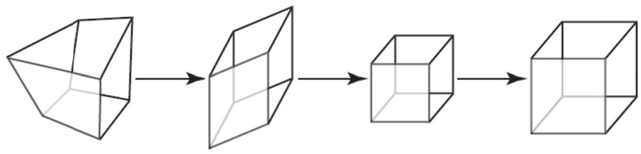

Fig. 4 Various geometric strata – projective, affine, metric, Euclidean

Fig.4 presents the relationshiop between the various geometric strata with their degrees of freedom:

- Projective – 15 degrees of freedom (dof),
- Affine – 12 degrees of freedom (dof),
- Metric 7 degrees of freedom (dof),
- Euclidean – 6 degrees of freedom (dof).

4. INCREASE ROBUSTNESS

Process of obtaining structure from motion are extremly sensitive to the presence of outlies and noise. Because of that, there are more techniques that are used for increase robustness, such a LmedS [7], RANSAC [8] and guided matching, where estimation og epipolar geometry is used to find more correspondences. These tehniques are eliminate outlies (points that do not belong to the object) and there is possibility for adding interest points, which increases the quality of an object of the scene. Usage of these techniques are present in [9], [10], [11].

5. CONCLUSION

Purpose of this paper is to give an overview of the basic steps and techniques in obtaining the structure from motion. The automatic recovery of the three dimensional structure from video of a scene are useful for creating 3D models, which can be use in many areas, as animation, computer games, virtual reality. Process of obtaining structure from motion is composed of several steps: feature matching,structure and camera calculation, auto calibration and increase robustness. Each of these steps is a complex process, starting from finding points of interest, eliminating outlies points, matching points of interest and getting the structure of the motion. For increasing the quality of the structure, techiques of auto-calibration and increase robustness are used.

REFERENCES

Journal Papers:

[1] Harris, C., Stephens, M., (1988). A combined corner and edge detector. *In Proceedings, 4th Alvey Vision Conference*, p. 147-151.
[2] Duda, R., Hart, P., (1973). Pattern Classification and Scene Anslysis. *John Wiley & Sons.*
[3] Shi, J., Tomasi, C., (1994). Good features to track. *IEEE Conference on Computer Vision and Pattern Recognition,* p.920-925.
[4] Longuet-Higgins, H.C., (1981). A computer algorithm for reconstruction a scene from two projections. *IEEE Transactions on Robotics and Automation,* p.133-135.
[5] Faugeras, O., Luong, Q., Maybank, S., (1992). Camera self-calibration: Theory and experiments. *European Conference on Computer Vision,* p.321-334.

[6] Beardsley, P., Torr, P., Zisserman, A., (1996). 3D Model Acquisition from Extended Image Sequences. *ECCV,* p.683-695.
[7] Rousseeuw, P.J., Leroy, A. M., (1987). Robust regression and outlier detection. *John Wiley & Sons.*
[8] Fischler, M., Bolles, R., (1981) Random sample consensus: a paradigm form model fitting with applications to image analysis and automated cartography. *Comunications of the ACM,* p.381-395.
[9] Mijakovska,S., Nedelkovski,I., Popovski, F., (2014). Generating 3D model from Video. *Advanced Computing: An International Journal (ACIJ), Vol.5, No.5/6*.
[10] Mijakovska,S., Nedelkovski,I., Popovski, F., (2014). Overview of the process of 3D modelling from video. *International Journal of Engineering Sciences & Emerging Technologies, ISSN: 2231 – 6604 Volume 7, Issue 3*, p. 680-686.
[11] Mijakovska,S., Pasic,R., Kuzmanov, I.,(2016). Influence of Corner Detectors in the Process of 3D Modeling from Video. *International Journal of Scientific and Engineering Research, ISSN 2229-5518 Volume 7, Issue 9,* p. 228-232.