Reconstruction of three-dimensional scenes from a sequence of images which only one is calibrate

K. KARIM; M. Es-sabry; N. EL AKKAD; M. Merras; A. Saaidi and K. SATORI

LIIAN, Department of Computer Science, Faculty of Science, Dhar El Mahraz, Sidi Mohamed Ben Abdellah, University, B.P 1796, Atlas, Fez, Morocco.

Abstract. In this work, we propose a method for reconstructing three-dimensional scenes from a sequence of images which only one image is calibrate. The system is initialized by a calibrated image. The projection matrices of the other images in the sequence are estimated from the first projection matrix (obtained by the calibration of the camera), and the matching between the images is determined by solving a system of linear equations. The intrinsic and extrinsic parameters of the camera used are estimated from the Cholesky decomposition of the projection matrix. Finally, obtaining the 3D scene is based on the Euclidean reconstruction of the detected and matched control points in the pairs of images.

Keywords: 3D reconstruction; matching; Projection matrix; Calibration and Autocalibration.

1 Introduction

3D reconstruction is a technique for obtaining a three-dimensional representation of an object from a sequence of images taken by different views ([27], [28], [29], [30], [31], [32]). Most 3D reconstruction techniques use calibration or autocalibration methods. Calibration methods use a known object (a calibration pattern 2D or 3D...) to find the cameras parameters, the projection of some points (in the image planes) which characterize the objects permit to solve a system of linear equations. Among the calibration methods, we have those presented in ([1], [2], [3]). The methods of autocalibration of the cameras are based on the estimation of the camera's parameters but without any knowledge a priori on the stage. Among these methods, we have those treated in ([4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14], [20], [21], [22], [23], [24], [25], [26]). The cost functions of these methods are generally non-linear, and they are formulate according to the invariants in the images and camera parameters.

In this work, we treat a three-dimensional scene reconstruction method from a sequence of images with a single image calibrated. Our system is initialized from a calibrated image.

The rest of this work is organize as follows: the second part presents some notations used in this work. The third part presents the pinhole model of the camera. Detection and matching of interest points will presented in the fourth part. The fifth part deals with the determination of the projection matrices and the autocalibration of the cameras. Experiments on real data will presented in Part 6 and the conclusion to this work is presented in Part 7.

2 Notations

M_1 is a point in the 3D scene whose homogeneous coordinates are $\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$

P_{i1} et P_{j1} are the projection matrices of point M_1 on the image planes respectively i et j

m_{i1} et m_{j1} are the projections of the point M_1 on the image plane respectively i et j

3 Camera model

We use in this work the pinhole model of camera **Figure1.** which represents geometrically the perspective projection; this model is used to project the scene in the image plane. It is characterized by a 3×4 matrix
R_i : The matrix of rotation, t_i the translation vector of the camera and K_i represents the matrix of intrinsic parameters such as:

$$\begin{pmatrix} f_i & s_i = 0 & u_{0i} \\ 0 & \varepsilon_i f_i & v_{0i} \\ 0 & 0 & 1 \end{pmatrix}$$

f_i : is the focal length, ε_i is the scale factor (we consider that the pixels are square, so $\varepsilon_i = 1$) , s_i : l is the image skew, and $(u_{0i} \quad v_{0i})$: are the image coordinates of the principal point.
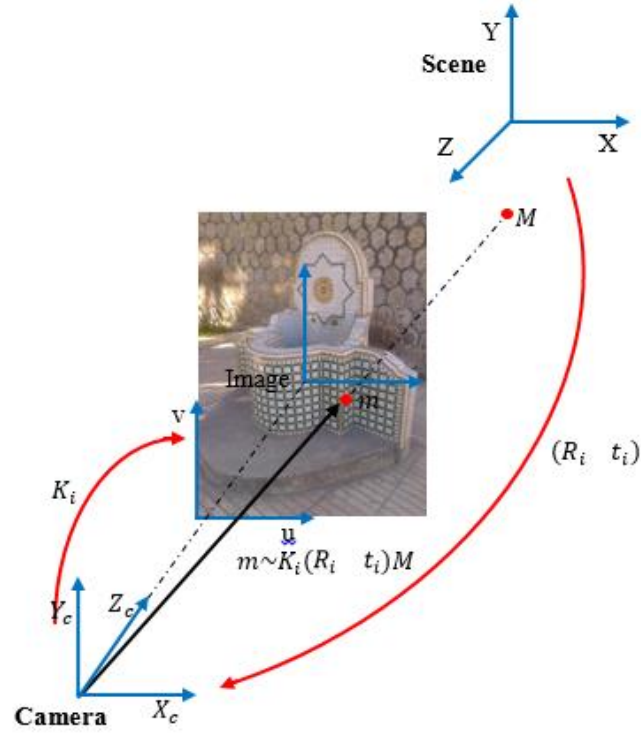
Fig. 1. Pinhole model.

4 Interest points and matching

4.1 Interest points

There are several methods to detect the interest points, among these methods we have those presented in ([15], [16]). We based in this work on the Harris detector [16]. The basic idea of Harris detector is to use the self-correlation function to determine the places where the signal changes in the two directions simultaneously.

$$G = \sum_{u,v} W(u,v) \otimes \begin{pmatrix} I_u^2 & I_u I_v \\ I_u I_v & I_v^2 \end{pmatrix} \tag{1}$$

With:

$$I_u = \frac{\partial I(u,v)}{\partial u}; \; I_v = \frac{\partial I(u,v)}{\partial v} \tag{2}$$

$$W(u,v) = \frac{1}{2\pi\sigma^2} Exp\left(-\frac{u^2+v^2}{\sigma^2}\right) \tag{3}$$

Harris proposes to calculate a measure based on the determinant and the trace of G by the following formula:

$$H = \det(G) - \lambda\, trace(G) \quad avec\ \lambda \in [0.04 \quad 0.06] \tag{4}$$

If $H > 0$: Area containing an interest point.

4.2 Matching of interest points

The matching of interest point is an important step to calibrate the cameras's used; several works ([17], [18], [19]) are developed on this field. The correlation measure used in this work is the Zero mean Normalized Cross Correlation (ZNCC). [19]

$$ZNCC(s_1, s_2) = \frac{\sum_{i=-N}^{N}\sum_{j=-S}^{S}(M-E)(H-F)}{\sqrt{\sum_{i=-N}^{N}\sum_{j=-S}^{S}(M-E)^2 \sum_{i=-N}^{N}\sum_{j=-S}^{S}(H-F)^2}} \tag{5}$$

With:

$$M = I_1(u_1+i, v_1+j) \tag{6}$$
$$H = I_2(u_2+i, v_2+j) \tag{7}$$

$$E = \frac{1}{(2N+1)(2S+1)} \sum_{i=-N}^{N}\sum_{j=-S}^{S} I_1(u_1+i, v_1+j) \tag{8}$$

$$F = \frac{1}{(2N+1)(2S+1)} \sum_{i=-N}^{N}\sum_{j=-S}^{S} I_2(u_2+i, v_2+j) \tag{9}$$

And $(s_1, s_2) = ((u_1, v_1), (u_2, v_2))$ are the coordinates of the pixels considered.

5 Estimation of the projection matrices and autocalibration of camera

5.1 Estimation of the projection matrices

We know that the projection of the point M_1 of the 3D scene on the image plane i is written in the following form:

$$m_{i1} \sim P_{i1} M_1 \tag{10}$$

Knowing that P_{i1} is obtained from calibration of the camera in the first view (image).

The previous equation contains two equations with 3 unknowns, which are the coordinates of M_1.

The projection of the point M_1 of the 3D scene on the image plane j is written in the following form:

$$m_{j1} \sim P_{j1} M_1 \tag{11}$$

The previous equation contains two equations with 14 unknowns. 11 unknowns for P_{j1} and the 3 unknowns of M_1.

To determine the coefficients of P_{j1}, we need 12 matches to generate a system of 48 linear equations with 47 unknowns.

There are 11 unknown for P_{j1} matrix and 36 unknowns for the 3D points (because we have 12 matches, therefore we have 3D points and each 3D point has 3 coordinates «unknowns »).

By solving the system, we will obtain the projection matrix P_{j1} and so on until the end of the image sequence.

5.2 Autocalibration of camera

5.2.1 Extraction of intrinsic parameters :

The projection matrix of the points of the 3D scene on the image plane is defined as follows:

$$P_j \sim (K_j R_j \quad - K_j R_j t_j) \tag{12}$$

We put $\overline{P}_j$ the 3×3 sub-matrix composed of the first three columns of P_j. We have:

$$\overline{P}_j \sim K_j R_j \tag{13}$$

If we multiply each side of the equation with its transpose:

$$\overline{P}_j \overline{P}^T_j \sim K_j R_j R^T_j K^T_j \tag{14}$$

Since R_j is orthogonal: $R_j R^T_j = I$. Then:

$$\overline{P}_j \overline{P}^T_j \sim K_j K^T_j \tag{15}$$

We see on the left a symmetric and positive definite matrix.

Its Cholesky decomposition gives an upper triangular matrix B_j with:

$$\overline{P}_j \overline{P}^T_j \sim B_j B^T_j \tag{16}$$

If we compare this equation with the preceding equation, while reminding us that K_j is upper triangular, we then find that B_j is indeed the calibration matrix K_j sought or almost: the element $(3, 3)$ of K_j must be equal to 1. We then multiply B_j with the appropriate scalar in order to satisfy this constraint. Then we can extract the intrinsic parameters of B_j (or K_j).

5.2.2 Extraction of extrinsic parameters

Once the intrinsic parameters have been extracted, the calculation of the extrinsic parameters is no longer very difficult. We have :

$$P_j \sim (K_j R_j \quad - K_j R_j t_j) \tag{17}$$

Then :

$$K^{-1}_j P_j \sim (R_j \quad - R_j t_j) \tag{18}$$

We put A_j the 3×3 sub-matrix consisting of the first three columns of $K^{-1}{}_j P_j$. So we have $A_j \sim R_j$ We can make this equation exact (make disappear the symbol $\sim$) by multiplying A_j with an appropriate scalar λ:

$$\lambda \, A_j = R_j \tag{19}$$

How to choose λ ?

The equality of the matrices implies the equality of their determinants, from which we obtain:

$$\lambda^3 \det(A_j) = \det(R_j) = 1 \tag{20}$$

So, we choose $\lambda = \sqrt[3]{1/\det(A_j)}$

Now :

$$\lambda \, K^{-1}{}_j P_j = (R_j \quad -R_j t_j) \tag{21}$$

(Note the exact equality: =). The first three columns of the left matrix then give us directly the rotation matrix R_j. Once R_j determined, the calculation of the vector t_j is trivial.

$$\lambda R^{-1}{}_j \, K^{-1}{}_j P_j = (I_{3*3} \quad -t_j) \tag{22}$$

6. Experimentations

In this phase, we took a differents views of a sequence of 19 480×640 images of a three-dimensional scene unknown by a CCD camera knowing that the first image is calibrate.

Figure 2.shows two of the 19 images of the used sequence.The interest points in these two images are detected by the Harris algorithm and represented in **Figure 3.** by the green color. The matching between this pair image are determined by using the ZNCC correlation measure and are shown in **Figure 4.**

Fig. 2. Two images of the three-dimensional scene

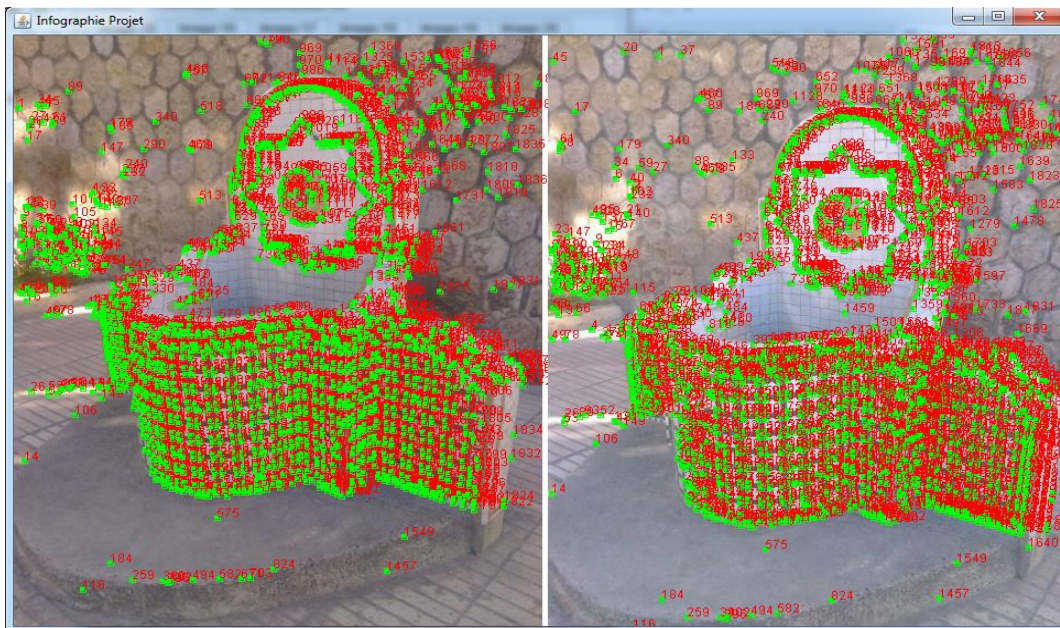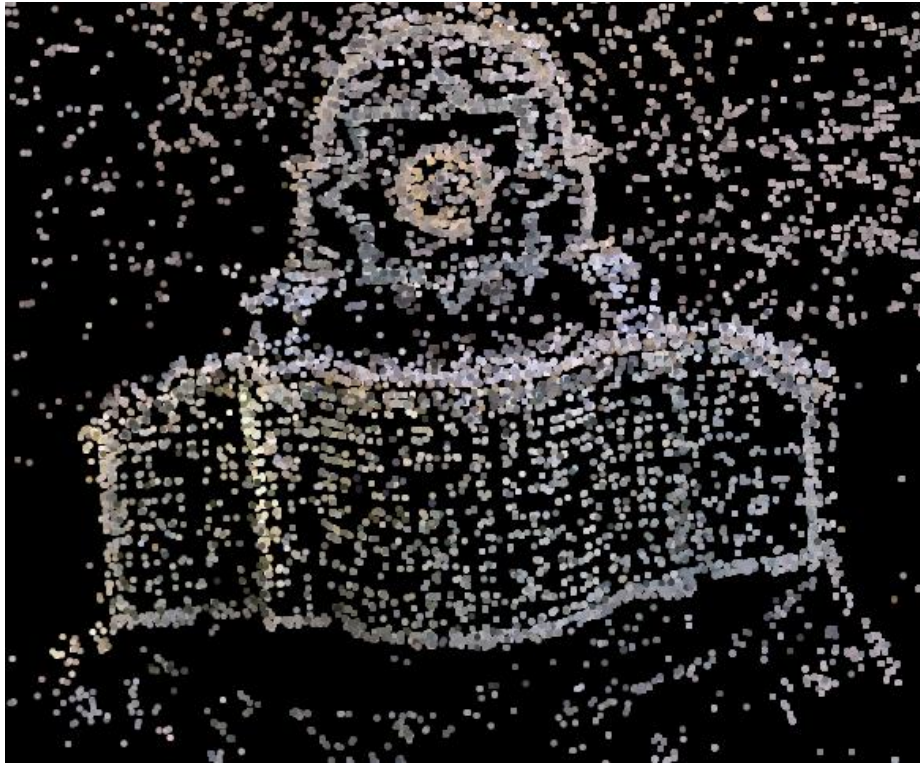**Fig. 3.** The interest points detected in two images.

Fig. 4. The matchingsbetween the couple of two images.

The result of the reconstruction of the scene after the implementation of the different algorithms (Harris, ZNCC, Ransac ...) by the Java programming language is represented in **Figure 5.** Below.

Fig. 5. Cloud points of the reconstructed 3D scene.

7. Conclusion

In this work, we have treated an approach to reconstruct three-dimensional scenes from a sequence of images which only one image is calibrate. The projection matrices of the others are obtained by solving a system of linear equations containing 16 equations at 14 unknowns. The interest points are detected by the Harris algorithm and matched by the ZNCC correlation method, and they are used with the projection matrices of the scene in the image planes to determine the point cloud coordinates , and finally to reconstruct the scene.

References

[1]   Zhang, Z.: A Flexible New Technique for camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22, No.11, pp. 1330-1334, 2000.
[2]   Gurdjos, P., Crouzil, A.,Payrissat, R.: Another Way of Looking at Plane-Based Calibration: The Centre Circle Constraint. ECCV, 2002.
[3]   Boufama, B., Habert, A.: Registration and tracking in the context of augmented reality. ICGST-GVIP Journal, vol. 5 Issue 3, March 2005.

[4]   Boudine, B., Kramm, S., El Akkad, N., Bensrhair, A., Saaidi, A., Satori, K.: A Flexible Technique based on Fundamental matrix for Camera Self-Calibration with Variable Intrinsic Parameters from two views. J. Vis. Commun. Image R. 39, 40–50. (2016)

[5]   Manolis, I., Lourakis, A., Deriche, R.: Camera self-calibration using the kruppa equations and the SVD of the fundamental matrix: the case of varying intrinsic parameters. Technical Report 3911, INRIA, 2000.

[6]   Sturm, P.: Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. Image and Vision Computing, 20(5-6): 415-426, 2002.

[7]   Malis, E., Capolla, R.: Camera self-calibration from unknown planar structures enforcing the multi-view constraints between collineations. IEEE Transactions on Pattern Analysis and Machine Intelligence, 4(9), September 2002.

[8]   Gurdjos, P., Sturm, P.: Methods and Geometry for Plane-Based Self-Calibration. CVPR, pp. 491-496, 2003.

[9]   Peijun, L., Jiaoying, S., Zhou, J., Longchao, J.: Camera Self-Calibration Using the Geometric Structure in Real Scenes. In Proceedings of the Computer Graphics International, 2003.

[10]   Elsayed, E.: A Survey of Camera Self-Calibration. In Proceedings of the IEEE Conference on AVSS, 2003.

[11]   Merras, M., El akkad, N., Saaidi, A., Nazih, G., Satori, K.: A New Method of Camera Self-Calibration with Varying Intrinsic Parameters using an Improved Genetic Algorithm. 8th International Conference on Intelligent Systems: SITA'13 (IEEE). , Rabat, Morocco, 08 and 09 May 2013.

[12]   El akkad, N., Merras, M., Saaidi, A., Satori, K.: Camera Self-Calibration with Varying Intrinsic Parameters by an Unknown Three-Dimensional Scene. The Visual Computer (Springer).30(5), 519-530 (2014).

[13]   Zetao, J.: Self-calibration of Varying Internal Camera Parameters Algorithm Based on Quasi-affine Reconstruction. Journal of Computers, Vol. 7, No. 3, March 2012.

[14]   Zhao, Y., Hu, X., Lv, X., Wang, H.: Solving the Camera Intrinsic Parameters with the Positive Tri-prism Based on the Circular Points. Information Technology Journal 11(7):926-930, 2012.

[15]   Trajkovic, M., Hedley, M.: Fast corner detection. Image and Vision Computing, 16:75–87, 1998.

[16]   Harris, C., Stephens, M.: A combined Corner et Edge Detector. 4th Alvey vision Conference. pp. 147-151, 1988.

[17]   Chambon, S., Crouzil, A.: Similarity measures for image matching despite occlusions in stereo vision. Pattern Recognit. 44(9), 2063–2075 (2011)

[18]   Mori, M., Kashino, K.: Fast template matching based on normalized cross correlation using adaptive block partitioning and initial threshold estimation. In: Proceedings of IEEE International Symposium on Multimedia, pp. 196–203 (2010)

[19]   Mattoccia, S., Tombari, F., Di Stefano, L.: Fast full-search equivalent template matching by enhanced bounded correlation. IEEE Trans. Image Process. 17(4), 528–538 (2008)

[20]   El akkad, N., Merras, M., Baataoui, A., Saaidi, A., Satori, K.: Camera self-calibration having the varying parameters and based on homography of the plane at infinity. Multimedia Tools and Applications (Springer). DOI: 10.1007/s11042-017-5012-3, 2017.

[21]   Mostafa, M., Elakkad, N., Saaidi, A., Gadhi Nazih, A., Satori, K. Robust method for camera calibration with varying parameters using hybrid modified genetic simplex algorithm. Journal of Theoretical and Applied Information Technology. Vol. 51 No.,pp. 363-373, 2013.

[22]   Baataoui, A., El akkad, N., Saaidi, A., Satori, K., Masrar, M.: Spherical Self-Calibration of Cameras Characterized by Varying Intrinsic Parameters. International Journal of Software Engineering and Its Applications. Vol. 8, No. 7, pp. 23-40, 2014.

[23]   Baataoui, A., El akkad, N., Saaidi, A., Satori, K., Masrar, M.: Robust Method for Camera Self-Calibration by an Unkown Planar Scene. Machine Graphics & Vision. Vol. 23, No. 1, pp. 37-57, 2014.

[24]   El akkad, N., Saaidi, A., Satori, K.: Self-calibration based on a circle of the cameras having the varying intrinsic parameters. The 3rd International Conference on Multimedia Computing and Systems: ICMCS'12 (IEEE). Tangier, Morocco, May 10-12, 2012.

[25] El akkad, N., Merras, M., Saaidi, A., Satori, K.: Camera self-Calibration with Varying Parameters from Two views. Wseas Transactions on Information Science and Application, Vol. 10, No. 11, pp. 356-367, 2013.

[26] El akkad, N., Merras, M., Saaidi, A., Satori, K.: Robust Method for Self-Calibration of Cameras Having the Varying Intrinsic Parameters. Journal of Theoretical and Applied Information Technology. Vol. 50, No. 1, pp. 57-67, 2013.

[27] Merras, M., Saaidi, A., El akkad, N., Satori, K.: Multi-view 3D reconstruction and modeling of the unknown 3D scenes using genetic algorithms. Soft computing (Springer). DOI : 10.1007/s00500-017-2966-z, 2017.

[28] El hazzat, S., Merras, M., El akkad, N., Saaidi, A., Satori, K.: 3D reconstruction system based on incremental structure from motion using a camera with varying parameters. The Visual Computer (Springer). DOI : 10.1007/s00371-017-1451-0, 2017.

[29] El akkad, N., El hazzat, S., Saaidi, A., Satori, K.: Reconstruction of 3D scenes by camera self-calibration and using genetic algorithms. 3D Research (Springer). Vol. 7, No. ¬6, pp. 1-17, 2016.

[30] KARIM, K., El akkad, N., Satori, K.: Reconstruction of the 3D scenes from the matching between image pair taken by an uncalibrated camera. The 3rd International Conference on Big Data, Cloud and Applications: BDCA'18. Kenitra, Morocco, (Springer) doi.org/10.1007/978-3-319-96292-4_35, April 04-05, 2018.

[31] Merras, M., El akkad, N., El hazzat, S., Saaidi, A., Nazih, G., Satori, K.: Method of 3D Mesh Reconstruction from Sequences of Calibrated Images. The 4th International Conference on Multimedia Computing and Systems. ICMCS'14 (IEEE). Marrakesh, Morocco, April 14-16 2014.

[32] El hazzat, S., Merras, M., El akkad, N., Saaidi, A., Satori, K.: Enhancement of sparse 3D reconstruction using a modified match propagation based on particle swarm optimization. Multimedia Tools and Applications (Springer). DOI: https://doi.org/10.1007/s11042-018-6828-1, 2018.