Multimedia Tools and Applications (2020) 79:29265–29288
https://doi.org/10.1007/s11042-020-09097-w

Genetic algorithms and bundle adjustment for the enhancement of 3D reconstruction

B. Satouri[1] · **K. Satori**[1] · **A. El abderrahmani**[2]

Received: 18 June 2019 / Revised: 13 March 2020 / Accepted: 22 May 2020 /
Published online: 11 August 2020

Abstract

In this paper, we present a new technique of tridimensional reconstruction from a sequence of uncalibrated stereo images taken with cameras having varying parameters. At first, our system allows to recover initial coordinates of a set of 3D points. In this context, we have used our method of self-calibration based on the use of unknown 3D scene with its image projections and genetic algorithms to estimate all intrinsic parameters. After that extrinsic parameters are estimated based on classical pose estimation algorithms. Matching points and estimated value of intrinsic and extrinsic parameters are used to estimate initial 3D model that helps us in the initialization step. In order to have a reliable and relevant 3D reconstruction the proposed method is based on good and new exploitation of bundle adjustment (without camera poses initialization) technique based on Levenberg-Marquardt optimization with the aim to estimate our optimal 3D model that has special features compared to the classical case because it masks the pose parameters estimation in the optimization process. Finally, 3D mesh of the 3D scene is constructed with Delaunay algorithm and the 2D image is projected on the 3D model to generate the texture mapping. Experiments is conducted on real data to achieve demonstrate the validity and the performance of the proposed approach in terms of convergence, simplicity, stability and reconstruction quality.

Keywords 3D reconstruction · Bundle adjustment · Self-calibration · Interests points · Matching · 3D mesh

✉ B. Satouri
boutainasatori@gmail.com

K. Satori
khalidsatori@gmail.fr

A. El abderrahmani
elabderrahmani@yahoo.fr

[1] LIIAN, Department of Mathematics and Informatics, Faculty of Sciences Dhar-Mahraz, P.O.Box 1796, Atlas Fes, Morocco

[2] LIIAN, Department of Mathematics and Informatics, Larache Poly disciplinary School, Larache, Morocco

1 Introduction

3D reconstruction [1, 18–21, 28, 34, 45, 46] is one of the most popular research areas in computer vision and computer graphics, it is widely used in many fields, such as video game, animation and so on. It consists in recovering three-dimensional information from 2D images (taken from different viewpoints) or from videos. Using this technology, we can implement scene recurrence, observe the model from any viewpoints stereoscopically and perceive the world well. With the development of computer technology and the increase of digitalizing demand, we hope to break the computers information processing ability and turn it to intellectively process multi-dimensional information.

The principle of 3D reconstruction is always to extract information on the three-dimensional scene from a sequence of images taken by numerical cameras with different viewpoints with or without a priori knowledge of the scene, based on two very important methods; the first requires a known scene or scene containing geometric objects with known structure to achieve a reliable result. This category is called the camera calibration [40, 44, 58, 59] these methods use the grid 2D or 3D to calibrate the cameras. The second consists to estimate the camera parameters without any a priori knowledge about the scene using the correspondence between the primitives detected in the images. This method is called camera self-calibration who has two categories: self-calibration of cameras with constant intrinsic parameters [2, 3, 6, 7, 16, 22, 41, 51, 52, 56, 61, 68, 69] and self-calibration of cameras with varying intrinsic parameters [8, 17, 25, 27, 32, 37, 38, 43, 54, 57, 70]. In this paper, we are interested in the self-calibration methods that can calibrate the cameras with varying parameters. These methods consist in finding a relation between the intrinsic parameters and the corresponding points in the images whose goal is to obtain a nonlinear cost function. The minimization of this function using robust algorithms allows the estimation of the intrinsic parameters of the camera in each image.

Our reconstruction system (Fig. 1) provides the 3D models of any scene and with any cameras type. One of the main arguments of our approach is an improved robustness and quality of 3D reconstruction from an optimization problem based on new and robust optimization algorithm (without camera poses initialization) in comparison to more existing approaches.

Our approach is a robust and powerful method for 3D reconstruction and modeling of the unknown three-dimensional scenes based on the estimation of the varying intrinsic and extrinsic parameters. After the detection of color interests points in the images by the Harris approach for color image [26, 29, 48] and the matching of these points in each pair of images by the correlation measure *ZNCC* [11, 12], intrinsic [53] and extrinsic [14, 15] parameters are estimated to obtain initial 3d model. After that our work is based on a formalism that uses a bundle adjustment [13, 62, 64, 65] that carries some special properties which are the use of Levenberg-Marquardt minimization to optimizing intrinsic parameters of each camera but the difference is that it incorporates an update of the sensor positions using the POSIT algorithm [14, 15], at each iteration. Thereafter this approach helps us to increase the convergence rate of the Levenberg-Marquardt algorithm to 40%. The POSIT algorithm is a position estimation algorithm that does not require any initialization, it converges after four or five iterations by means and therefore it is a good choice for real-time reconstruction applications.

Our method presents a novelty: two image from the sequence of images are sufficient to estimate the cameras' intrinsic and extrinsic parameters, the use of any camera (with varying intrinsic parameters) which render the self calibration procedure most robust and the use of an unknown 3D scene, this will provide more points of learning in a viewpoint, reduces the planarity constraints and the self-calibration process is thus simplified and faster. These advantages allow us, on the one hand, to solve some problems related to the self-calibration system and, on the other hand, to work freely in the domain of self-calibration with fewer constraints.

This paper is organized as follows: In the third part, we present the camera model and matching containing three subparts: The first subpart comprises the camera model, the second is the Color Interest points 'detection. It is a preliminary step in many computer vision processes; many methods have been advanced to extract interest points. In this paper, we used Harris interest point detector for color image. The third is the *Matching*: Finding in two images of the same scene, taken at different positions, pairs of pixels which are the projections of the same point of the scene. In this phase, the detected color interest points are matched by ZNCC (Zero mean Normalized Cross Correlation) correlation measure. The most important section is related to the use of bundle adjustment of type Levenberg-Marquardt to estimate the optimal intrinsic and extrinsic parameters in section four. The five parts presents 3D Mesh of cloud points. The experiment results are discussed in the six parts, and finally, in section seven we will proceed to make a general conclusion.

2 Related works

The Reconstruction of 3D scenes is often necessary and essential in Computer Vision and especially for robotic applications. In this section we will look into the state of the art of the majority of techniques and approaches used in Tridimensional Reconstruction of the 3D scenes using the bundle adjustment for producing more realistic 3D model.

In the literature the Three-dimensional Reconstruction based on classical bundle adjustment, which is often used as a basic element in many approaches [4, 24, 30, 42, 47, 66], exists in several arrays. In the same context other practical methods are proposed: [67] a method which use epipolar and trifocal constraints to implicitly establish the relations between the cameras via the

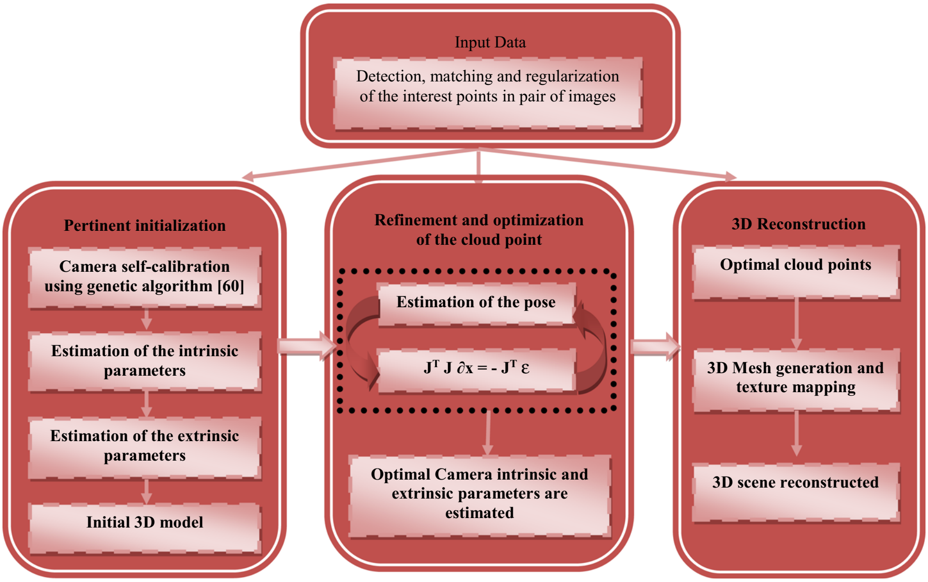

Fig. 1 The stages of our 3D reconstruction system

object structure without including the 3D points in the optimization as parameters. [55] This article based on numerical optimization including sparse Newton methods. A very effective approach was presented by [49, 63]. In [35, 50] Instead of applying the projection equations, the cost-optimized function in the suggested approach is based on the three-view geometry constraints that must be satisfied for any three-view with a common overlap area which gives a reduction significant computational complexity compared to a standard BA, as the number of unknown parameters participating in iterative optimization is much smaller.

[36] This approach is called beam adjustment without structure. The number of unknowns in the resulting system is greatly reduced. The author uses the accumulation of stress residues as an approach to reduce the number of lines in the Jacobin matrix.

In [10] this article presents a discussion on rotation averaging and distance measures in SO (3) based on three main problems: single rotation averaging, multiple-rotation averaging and conjugate rotation averaging, which relates a pair of coordinate frames. In [31] the authors presented an algorithm that retrieves the 3D model of human faces from two calibrated images. After the estimation of scattered 3D point clouds an adjustment of the morbid model is performed to have a dense 3D reconstruction of the face.

[39] A new method based on the global nonlinear optimization of continuous scene depth rather than discrete pixel disparities.

In [5] the author describes a new method for Euclidean reconstruction based on the estimation of the projective depths from the resolution of a nonlinear equation system for the initialization of the optimization system as well as the use of the Kruppa equations for estimating the camera parameters.

3 Pinhole camera model

In this work, the pinhole camera model is used to project a 3D scene Point $A = (X, Y, Z, 1)^T$ in the image plane $a = (x, y, 1)^T$. This projection is represented by the following formula: $\lambda a = PA$.

With: λ is a nonzero scale factor and P is the projection matrix with $P = K\,[R \mid t]$.

For the camera g, the projection matrix is defined by K_g (R_g t_g) with a matrix (R_g t_g) containing extrinsic parameters. R_g the rotation matrix, and t_g the translation vector of camera in space, Kg is a matrix containing the intrinsic parameters and is expressed as follows:

$$K_g = \begin{pmatrix} f_g & \tau & u_{0g} \\ 0 & \varepsilon_g f_g & v_{0g} \\ 0 & 0 & 1 \end{pmatrix} \tag{1}$$

f_g	is the focal length
ε_g	is the scale factor
τ	is the skew factor
$(u_{0g}\ v_{0g})$	represent the coordinates of the principal point in the images.
R_g	is the rotation matrix defined by:

$$R_g = R_X\left(\varphi_g\right) R_Y(\phi_g) R_Z(\psi_g) \tag{2}$$

The three components of rotation matrix, according to the three axes, are defined as follows:

$$R_X\left(\varphi_g\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & cos\varphi_g & -sin\varphi_g \\ 0 & sin\varphi_g & cos\varphi_g \end{pmatrix} \tag{3}$$

$$R_Y\left(\phi_g\right) = \begin{pmatrix} cos\phi_g & 0 & sin\phi_g \\ 0 & 1 & 0 \\ -sin\phi_g & 0 & cos\phi_g \end{pmatrix} \tag{4}$$

$$R_Z\left(\psi_g\right) = \begin{pmatrix} cos\psi_g & -sin\psi_g & 0 \\ sin\psi_g & cos\psi_g & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{5}$$

With: φ_g, ϕ_g, ψ_g are the three Euler angles.

Therefore, from the previous formulas we obtain:

t_g is the translation vector:

$$t_g = \left(t_x, t_y, t_z\right)^T$$

4 Proposed method

We present a 3D reconstruction system (algorithm 1) based on the good and new exploitation of bundle adjustment. It takes as input two uncalibrated stereo images captured by a camera with varying parameters. As output, it determines the optimal camera parameters (intrinsic and extrinsic) and the three-dimensional structure.

Our work is based on the use of a minimization process which is Levenberg–Marquardt algorithm. In the fact, we use it as a minimization step after obtaining pertinent result in the initial estimation step.

The global algorithm of our 3D reconstruction system is presented as follows:

Algorithm 1: Camera parameters and 3D model Recovery

Input: two uncalibrated images
Output : optimal camera parameters(intrinsic and extrinsic) and the 3D model
Begin

1. Interest point detection and matching
2. Initialization step
   2.1. Estimation of the intrinsic parameters
   2.2. Estimation of the extrinsic parameters
   2.3. Recovery of initial 3D point coordinates from the intrinsic and extrinsic parameters estimated and some matches result between these two images
3. bundle adjustment based on Levenberg-Marquardt algorithm including extrinsic parameters estimation for refinement and optimization of initial 3D cloud point

End

4.1 Interests points

Interest points are characteristic points of the image that are particularly holders of information. The detection of these points in stereoscopic images is an essential step in the field of computer vision and especially in the Three-dimensional reconstruction. It is to match the projections of the same entity in the scene. We begin by extracting the corners points with the color Harris detector [26, 29, 48] that exists in the literature:

4.1.1 Harris corner detection for color images

A color image provides more information than a grey-scale image, and this additional information should help in reducing the ambiguity in corner detection.

The Harris corner detector introduced in [26, 29, 48] provides a corners measure for image data. It is calculated based on a second moment matrix N describing the gradient distribution in the local neighbourhood of a point as:

$$R = \det(N_{color}) - \alpha\ trace^2(N_{color}) \tag{6}$$

where α=0.04 (Harris parameter response) and N_{color} is the 2 × 2 matrix given by:

$$N_{color} = \begin{pmatrix} R_x^2 + G_x^2 + B_x^2 & R_xR_y + G_xG_y + B_xB_y \\ R_xR_y + G_xG_y + B_xB_y & R_y^2 + G_y^2 + B_y^2 \end{pmatrix} \tag{7}$$

With R, G and B are three components of color such as Red, Green and blue. According to the comparisons made by Gouet and Boujemaa [26], the above detector appears to be the most stable among the popular color interest point's detectors with regard to illumination changes, noise, rotation and viewpoint changes.

4.1.2 Correlation measure

After the detection of color interest points by Harris algorithm [26, 29, 48], appeal is made to one of matching methods which are looking for the pixels that are similar to the ones using correlation measure. The search of what is locally corresponding is carried out in an area of research. Very many measures have been proposed to take into account the various difficulties encountered during this step. In this paper, we chose the extent of correlation ZNCC (*Zero mean Normalized Cross Correlation*) [11, 12]. The robustness of these functions is classified according to their capabilities to match in the case of large displacement and change of brightness.

To eliminate the false matches detected in a pair of images, we will regularize all interest points by estimating the fundamental matrix F_{gd} between the *left* and right image using the RANSAC algorithm [23].

4.2 Initialization of 3D reconstruction

Initialization step is crucial for bundle adjustment algorithms. Indeed, when the initial error is small, the method converges. The purpose of initialization step is to constitute the vector of estimated parameters and to start 3D reconstruction. These components are the intrinsic parameters of the different cameras and the parameters associated with the 3D reconstructed model. It consists in:

1 Camera self-calibration: estimation of intrinsic parameters from two images using genetic algorithm.
2 estimation of extrinsic parameters based on Posit algorithm.
3 Retrieving a set of initial 3D points from matched interest points.

4.2.1 Camera self-calibration

In this step, we use our method of camera self-calibration having varying intrinsic parameters from a sequence of two images of an unknown 3D object [53]. Our method is based on the formulation of a linear cost function from the determination of a relationship between two points of the scene with their opposite relative to the axis of abscise and their projections in the

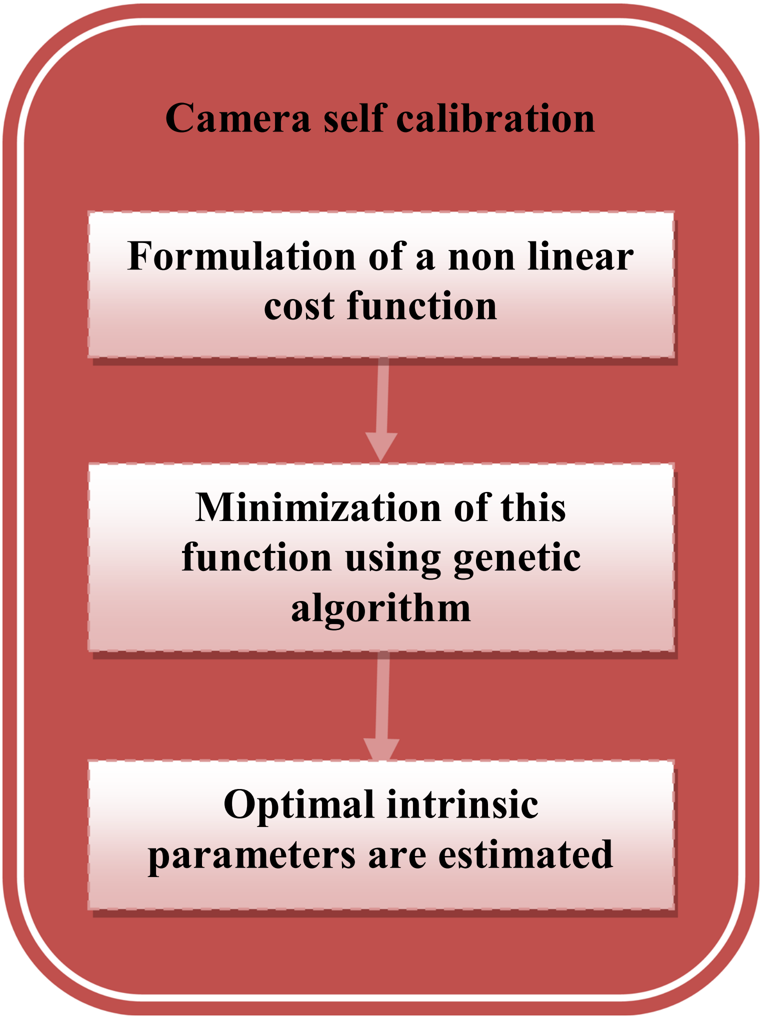

Fig. 2 The stages of our approach for Camera self-calibration [53]

image planes. The resolution of this function by the genetic algorithms [33] allows the estimation of the intrinsic parameters of the cameras used (Fig. 2).

As a result, we can estimate intrinsic parameters for the cameras that will be used in the initial reconstruction of the 3D scene.

4.2.2 Estimation of the extrinsic parameters

The classical pose estimation algorithm (POSIT) determines the pose of an object relative to the camera from a set of 2D image points of minimum 4 non-coplanar feature points and our 3D points coordinates (Fig. 3).

4.2.3 Recovering initial 3D point coordinates

The coordinates of 3D points are obtained from the matching result between the two images and the intrinsic [53] and extrinsic parameters (Posit algorithm) of the cameras estimated.

$$a_{gk} = P_g A_k = K_g \left[R_g | t_g\right] A_k \tag{8}$$

$$a_{dk} = P_d A_k = K_d [R_d | t_d] A_k \tag{9}$$

With P_g and P_d are respectively, the projections matrices of the camera. We pose:

$$P_p = \begin{pmatrix} P_{11p} & P_{12p} & P_{13p} & P_{14p} \\ P_{21p} & P_{22p} & P_{23p} & P_{24p} \\ P_{31p} & P_{32p} & P_{33p} & P_{34p} \end{pmatrix} p = g, d \tag{10}$$

From relations (8) and (9), we deduce the following relationship:

$$\begin{cases} x_{dvk} = \dfrac{P_{11g}X_k + P_{12g}Y_k + P_{13g}Z_k + P_{14g}}{P_{31g}X_k + P_{32g}Y_k + P_{33g}Z_k + P_{34g}} \\ y_{gk} = \dfrac{P_{21g}X_k + P_{22g}Y_k + P_{23g}Z_k + P_{24g}}{P_{31g}X_k + P_{32g}Y_k + P_{33g}Z_k + P_{34g}} \\ x_{dk} = \dfrac{P_{11d}X_k + P_{12d}Y_k + P_{13d}Z_k + P_{14d}}{P_{31d}X_k + P_{32d}Y_k + P_{33g}Z_k + P_{34d}} \\ y_{dk} = \dfrac{P_{21d}X_k + P_{22d}Y_k + P_{23d}Z_k + P_{24d}}{P_{31d}X_k + P_{32d}Y_k + P_{33g}Z_k + P_{34d}} \end{cases} \tag{11}$$

These equations allow obtaining the following system of linear equations:

$$\mathrm{N_{gdk}} \left(\mathrm{X_k Y_k Z_k}\right)^{\mathrm{T}} = \mathrm{G_{gdk}} \tag{12}$$

With:

$$N_{gdk} = \begin{pmatrix} P_{11g} - x_{gk}P_{31g} & P_{12g} - x_{gk}P_{32g} & P_{13g} - x_{gk}P_{33g} \\ P_{21g} - y_{gk}P_{31g} & P_{22g} - y_{gk}P_{32g} & P_{23g} - y_{gk}P_{33g} \\ P_{11d} - x_{dk}P_{31d} & P_{12d} - x_{dk}P_{32d} & P_{13d} - x_{dk}P_{33d} \\ P_{21d} - y_{gk}P_{31d} & P_{22d} - y_{dk}P_{32d} & P_{23d} - y_{gk}P_{33d} \end{pmatrix} \tag{13}$$

$$G_{gdk} = \begin{pmatrix} x_{gk}P_{34g} - P_{14g} \\ y_{gk}P_{34g} - P_{24g} \\ x_{dk}P_{34d} - P_{14d} \\ y_{gk}P_{34d} - P_{24d} \end{pmatrix} \tag{14}$$

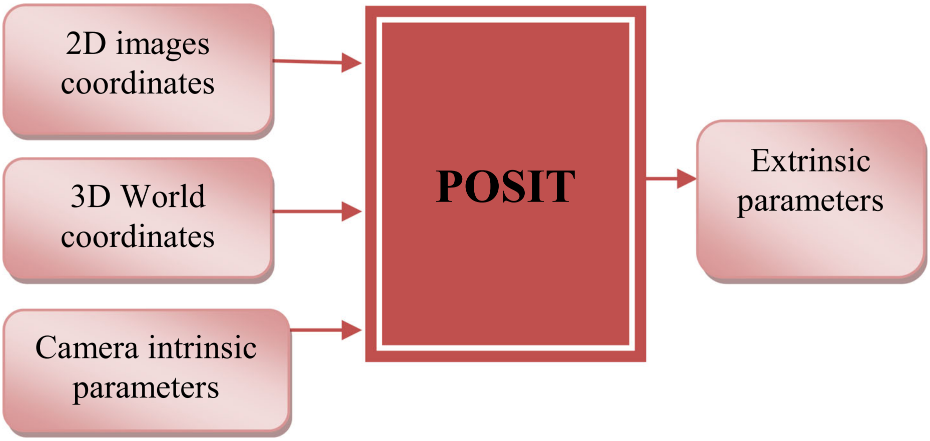

Fig. 3 Posit algorithm

The solution of (12) is obtained by:

$$(X_k Y_k Z_k)^T = \left(N_{gdk}^T N_{gdk}\right)^{-1} N_{gdk}^T G_{gdk} \tag{15}$$

After the resolution of Eq. (15) we obtain the initial 3D cloud point of the observed scene.

The different entities previously estimated (intrinsic, extrinsic parameters and the 3D point coordinates) are refined by minimizing the reprojection error of the cloud points using a new bundle adjustment based on Levenberg-Marquardt algorithm in order to obtain a best cloud point and to correct errors due to the linear solution.

4.3 Bundle adjustment

Our optimization is a new bundle adjustment [13, 60, 62, 64, 65] for refining the initial 3D model based on the minimization of a nonlinear criterion (Algorithm 2). The strong point of our approach is that the intrinsic parameters and the extrinsic parameters are not treated separately but in the same optimization loop as well as the use of the derivatives of the minimized function in order to define the evolution of the estimated parameter set. This strategy is the core of behavioral differences between methods based on the criterion's unique value or its derivatives with respect to the parameters.

As we have already said, we use Levenberg-Marquardt minimization and which differs to classical step in the calculation of the vector (ε), calculating the Jacobin matrix (J) and then at initialization step. Practically we want to minimize a cost function () for the set of points and images:

$$F = \sum_{i,k} \left\| a_k^i - P\left(K_i, POSIT\left(k_i, A_{3D,K}\right)_i\right) A_{3D,K} \right\|^2 \tag{16}$$

With:

$a_k^i = \left(x_k^i, y_k^i, 1\right)^T$Position of the K-th 2D point in image i.

PProjection matrix.

POSITRefers to calculating function of camera position (Relative to the intrinsic parameters and known 3D points).

$K_i$intrinsic camera parameters.

$A_{3D,k}$position of the K-th 3D point in scene.

Algorithm 2 : Levenberg-Marquardt

Input : P_0 : initial parameters estimated (intrinsic parameters, initial 3D points)
Output : P_S: solution of minimization (optimal camera parameters , optimal 3D points)
Begin
$i \leftarrow 0$, $\lambda \leftarrow 10^{-3}$, $P_i \leftarrow P_0$
While stopping condition is not reached **do**
1. $P_i \leftarrow P_i + \delta P_i$;
2. Calculate Jacobin matrix $J(P_i)$ using numerical derivatives;
3. Calculate Hessien matrix $H(P_i)$;
$H(P_i) \leftarrow J^T(P_i)\, J(P_i)$;
4. Calculate vector of errors $\varepsilon(P_i)$;
5. Calculate vector of gradient $g(P_i)$;
$g(P_i) \leftarrow J^T(P_i)\, \varepsilon(P_i)$;
6. Solve follow system $[\, H(P_i) + \lambda\, \mathrm{diag}(H(P_i))\,]\, \delta P_i = -\, g(P_i)$;
7. **if** $\| \varepsilon(P_i + \delta P_i) \| < \| \varepsilon(P_i) \|$ **then**
$\lambda \leftarrow \lambda / 10$;
go to (1);
else $\lambda \leftarrow \lambda * 10$;
go to (6);
end
end
End

After the reliable initialization of the intrinsic parameters and 3D model using initial 3d cloud point and intrinsic parameters.

estimated the Optimization step is performed in two essential steps outlined below:

Algorithm 3: Calculation of error vector

Input: P_i : (intrinsic parameters, 3D points)
Output : $\varepsilon(P_i)$
Begin
1. Calculate the pose of each sensor at each iteration:
$$Rt_i = POSIT\,(P_i)$$
2. Calculate the 2d projection of each primitive in the images:
$$\widetilde{a}_{k,i} = \Pr ojection(P_i, Rt_i)$$
1. Storing the measurement in error vector:
$$\varepsilon(P_i) = \left\| a_{k,i} - \widetilde{a}_{k,i} \right\|^2$$
End

1 Calculation of the Jacobin matrix using numerical derivatives of measurement with respect to the parameters to be estimated. The particularity from the classic case is the use of an algorithm of position at each iteration. Camera position and 3D model estimation are done simultaneously.

2 Calculation of the error vector that presents the difference between detected primitives from images and estimated ones. To calculate it, we based on Posit [14, 15] algorithm

because it does not require initialization and thus makes it possible to carry out the bundle adjustment without knowing the attitude of the cameras at the beginning. (Algorithm 3)

Algorithm 4 : Delaunay triangulation using circles

```
Input : LCP: List of cloud point
Output : LIP : list of ideal points (the points in order along the curve)
Combinations( LCP , 3) : LT List of triangles;
LIP : empty
Begin
    For Triangle in LT do
        Flag ← true;
        Calculate the coefficients of the two mediators
        Calculate Circle Center;
        Calculate Radius;
      For P in LCP do
          if P is not in triangle then
             Calculate distance;
             if Distance < Radius then
                Flag ← false;
                Break;
             end
          end
       end
       if Flag is true then
          Add the triangle in LIP ;
      end
    end
    Return LIP
End
```

After the optimization and the refinement of cloud point the next step is to produce a model from this set of cloud point in space, that is to say, obtain a description of the scene as more realistic surfaces. This description allows particularly making the mesh of the cloud point and the texture mapping to obtain a realistic 3D model.

4.4 Triangulation of the 3D cloud point and texture mapping

The triangular surface meshes are the most used for the representation of objects in three-dimensional space. They are fast becoming the standard representation for modeling geometric objects thanks to their simplicity and efficiency.

In this paper we use Delaunay triangulation [9] algorithm to generate the mesh of the point cloud (Algorithm 4) it is performed to process the cloud point obtained from stereo matching and then a 2D image is projected orthogonally on the 3D model to generate the texture mapping. All images taken by the camera are the source texture to texture mapping.

To every triangle of surface model, each vertex's texture coordinates can be obtained from the image coordinates of its matched point in the texture image, and the texture coordinates of internal points can be calculated by linear interpolation of the vertexes' texture coordinates. The texture image is mapped automatically to a model in this way. Finally, the 3d model of the scene is generated.

Fig. 4 Three sequence of two real images

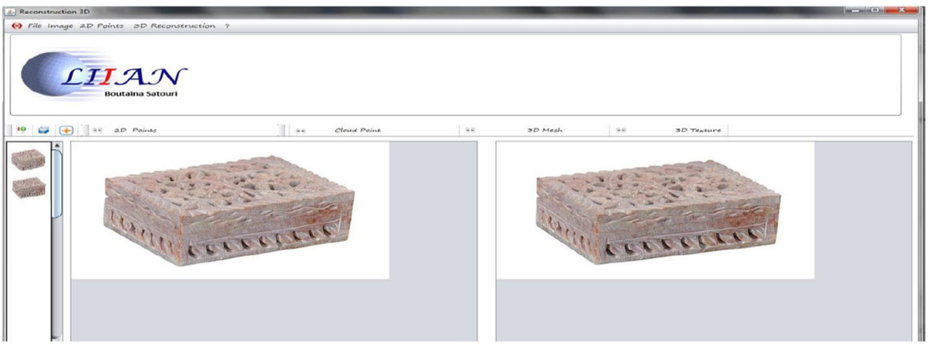

Fig. 5 Loading 2D images

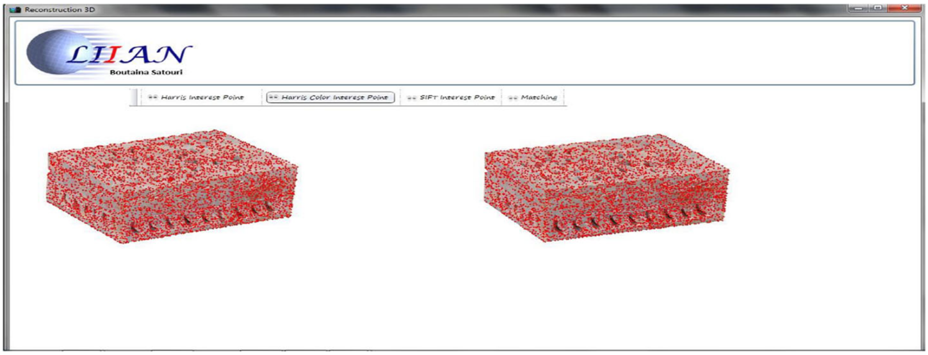

Fig. 6 Result of interest point

4.5 Experimental results

In this section, we tested on synthetic and real data the efficiency and performance of the work realized in this paper based on a sequence of stereoscopic images taken by a camera which turns around the object or scene.

The interest points are detected in the images by the Harris detector for color image and matched by the correlation measure ZNCC. These matches are regularized by the RANSAC algorithm. The determination of the relationship between the points matching and the cameras parameters allows formulating a system of nonlinear equations. This system is converted to a nonlinear cost function which is minimized by genetic algorithm in order to finding an initial solution of the intrinsic parameters. These parameters are used with extrinsic parameters estimated by Posit algorithm and matching points to obtain an initial 3D cloud points. The new and good exploitation of bundle adjustment technique based on Levenberg-Marquardt optimization are used with the aim to estimate our optimal 3D scene that has special features compared to the classical case because it masks the pose parameters in the optimization process using classical pose estimation algorithms.

The different algorithms (Harris, ZNCC, RANSAC, Posit, Levenberg-Marquardt algorithm, 3d reconstruction, Delaunay, …) used in this work are implemented by the Java Programming language and the OpenCV library and executed on a core i3 computer with 2 GHz of the processor and 4 GB of the RAM.

Fig. 7 Result of matching between two images after removing false matches

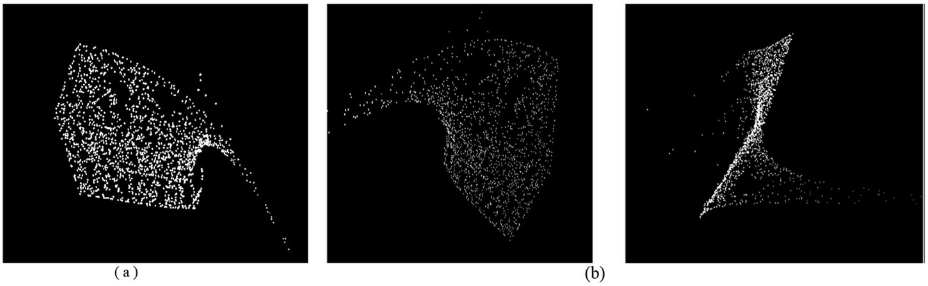

Fig. 8 **a** initial 3D Model (**b**) Two views of initial 3D Model after some movement

4.5.1 Results of real data

To show the quality, efficiency and the robustness of the proposed method, several images sequences are tested. We present the results of three image sequences of objects (Fig. 4) (each image sequence contains two real world image of object).

After loading each couple of images (Fig. 5), the detection of interest points (Fig. 6) and the matching points (Fig. 7) are performed, respectively, by the Harris detector for color image and the correlation measure ZNCC. The determination of the relationship between two points of three-dimensional scene and their projections in the planes of different images allows finding intrinsic and extrinsic parameters. These parameters are used with matching points to obtain a set of initial 3D cloud points (Fig. 8). The new exploitation of bundle adjustment technique based on Levenberg-Marquardt optimization are used with the aim to estimate our optimal 3D scene (Fig. 9) that has special features compared to the classical case because it masks the pose parameters in the optimization process using classical pose estimation algorithms. A Delaunay algorithm is applied to generate the mesh (Fig. 10), and the texture mapping (Fig. 11) is made from two dependent viewpoints.

a Results of the first sequence
b Results of the second sequence
c Results of the third sequence

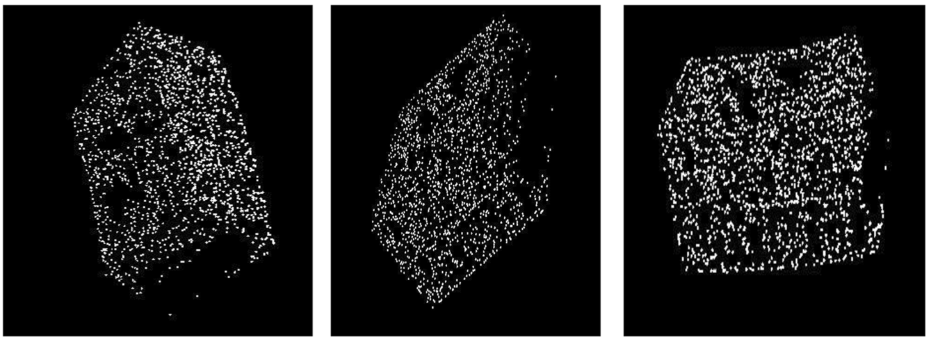

Fig. 9 Three view of 3D cloud points reconstruction from two view (1594 Cloud points)

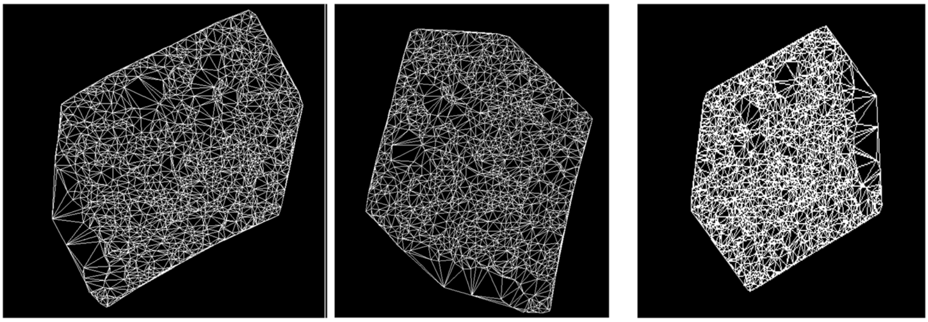

Fig. 10 3D Mesh generated by Delaunay algorithm for triangulation

According to the results of the three sequences (reconstructed) shown in the Figs. 9, 10, 11, 12 and 13 we can conclude that our approach gives good results of 3D reconstruction of observed scenes despite the choice of scenes with specific features (which are difficult to be all taken into consideration). This reconstruction quality is due, on the one hand to the use of a robust method of camera self-calibration (the estimation of all intrinsic camera parameters using genetic algorithm), and on the other hand, to the optimization of coordinates of 3D cloud points by a simple and reliable approach.

4.5.2 Comparison and interpretation

In this part, we test the accuracy and performance of the proposed approach with the simulations of a sequence containing ten images of an unknown three-dimensional scene that is taken by a CCD camera from different views. After we estimated the cameras parameters using the proposed algorithm and we calculated the 3D points cloud. As a comparison, we implemented five state-of-the-art methods: our method, the method of Jiang [38], the method of Elakkad [18], the method of Merras [45] and the method of Hu [34]. The quality of three-dimensional reconstruction of scene obtained by the five methods is determined from the reprojection error. Jiang approach [38] is based on quasi-afine reconstruction for linear self-calibration. For it, the authors assume that the skew factor is zero in the first view, which allows them to estimate the other parameters (four in the first view and five of each other views). Elakkad approach [18] and Merras approach [45] are two methods based on the formulation of a nonlinear cost function by determining the relationship between 2D points of

Fig. 11 3D model of the object after texture mapping from two views

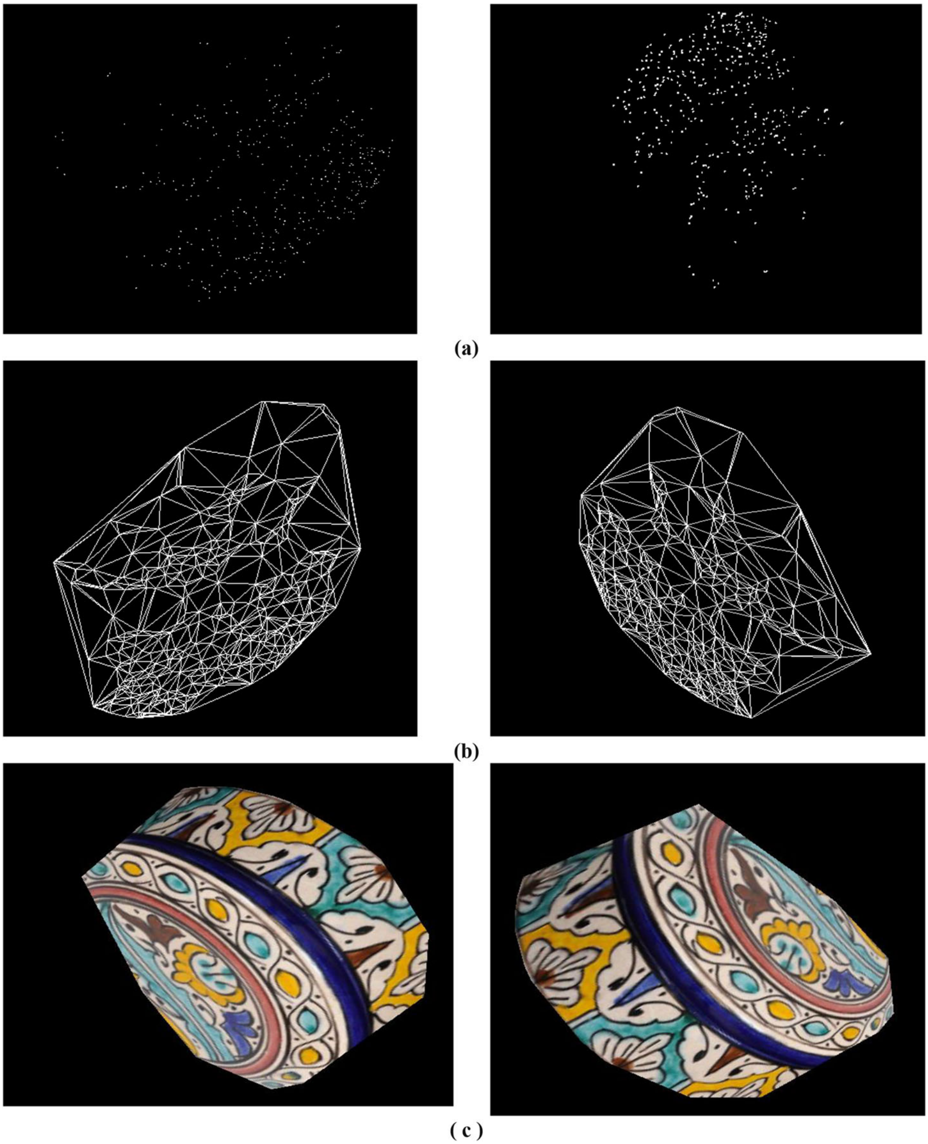

Fig. 12 Modeling and 3D reconstruction results of the second sequence. **(a)** Two views of 3D reconstruction (414 Cloud points), **(b)** Mesh generated by Delaunay algorithm, **(c)** two views of the object after texture mapping

the images and the cameras parameters; the optimization of this function by a genetic algorithm makes finding the optimal cameras parameters. Hu approach [34] uses three images for linear self-calibration to estimate the constant cameras parameters which will be, thereafter, refined by minimizing the overall back-projection errors across the three images within a tiny-scale bundle adjustment framework.

Table 1 in the same way the figures (Figs. 14, 15) show that our method gives satisfactory results compared to the other methods [18, 34, 38, 45]. This is due to the good initialization based on the use of our approach for self-calibration from two uncalibrated images, using

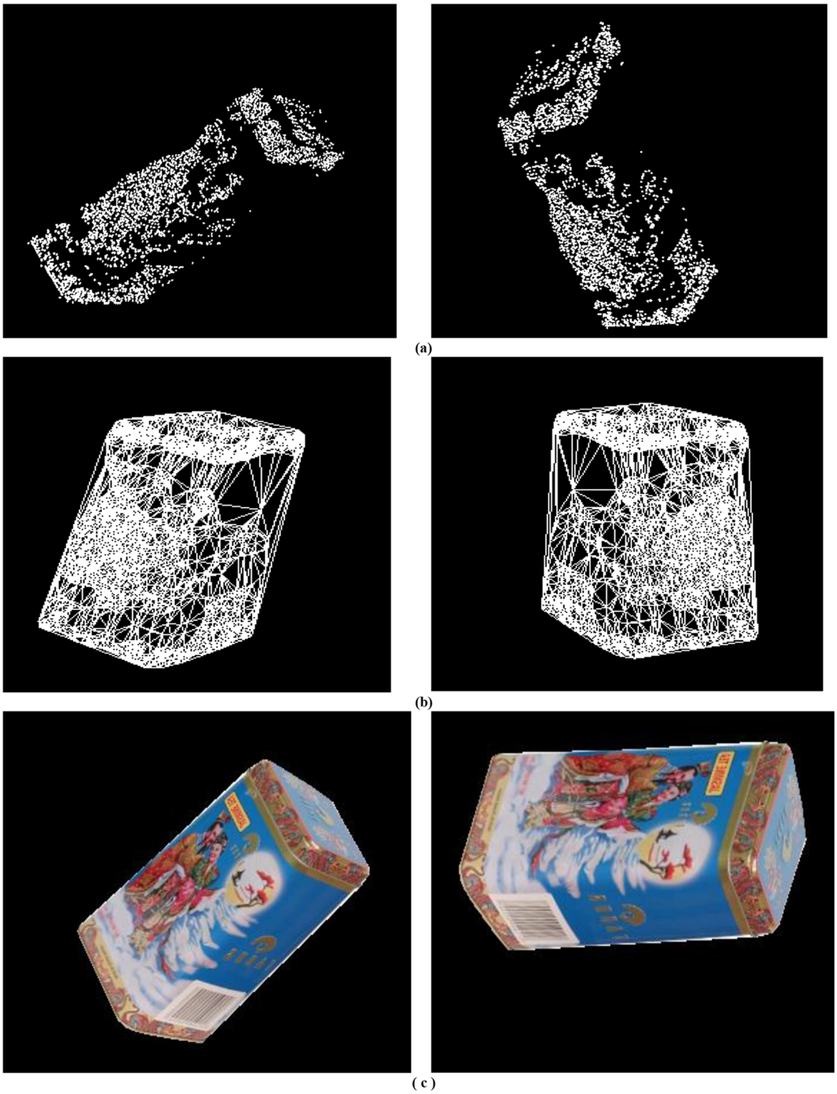

Fig. 13 Modeling and 3D reconstruction results of the third sequence. **(a)** Two views of 3D reconstruction (2548 Cloud points), **(b)** Mesh generated by Delaunay algorithm, **(c)** two views of the object after texture mapping

genetic algorithm, which allows us estimating a set of initial 3D points of the scene, and also to the recovery of optimal 3D points based on the new and good exploitation of bundle adjustment. In regard to the computation time according to the image number, our approach is faster than Merras [45] approach. The proposed method allows having quality results, in a time close to Merras [45] approach. This latter applies the genetic algorithm to find the optimal 3D point cloud, on the other hand we use a bundle adjustment technique based on Levenberg-Marquardt optimization which is much rapid related to the reliable and better initialization. Like the Hu [34] approach, the optimal solution depends on the initialization step, if initialization is close from the optimum, the algorithm converges quickly to a global solution, and thus the convergence time decreases. Our approach is less fast than Elakkad approach [18] but

Table 1 Comparison results

Approaches	Sequences	Reprojection error (%)	Execution time (s)
Our approach	First sequence	0,14	19,3
	Second sequence	0,10	14
	Third sequence	0,11	16
Elakkad approach [18]	First sequence	0,23	18
	Second sequence	0,17	13
	Third sequence	0,21	14
Merras approach [45]	First sequence	0,15	25
	Second sequence	0,12	23
Hu approach [34]	Third sequence	0,10	26
	First sequence	0,13	17
	Second sequence	0,9	15
	Third sequence	0,10	19
Jiang approach [38]	First sequence	0,35	53
	Second sequence	0,29	41
	Third sequence	0,31	45

the results indicate that our reconstruction system is more accurate than the other two systems Elakkad [18] and Jiang [38] thanks to the good and new exploitation of bundle adjustment.

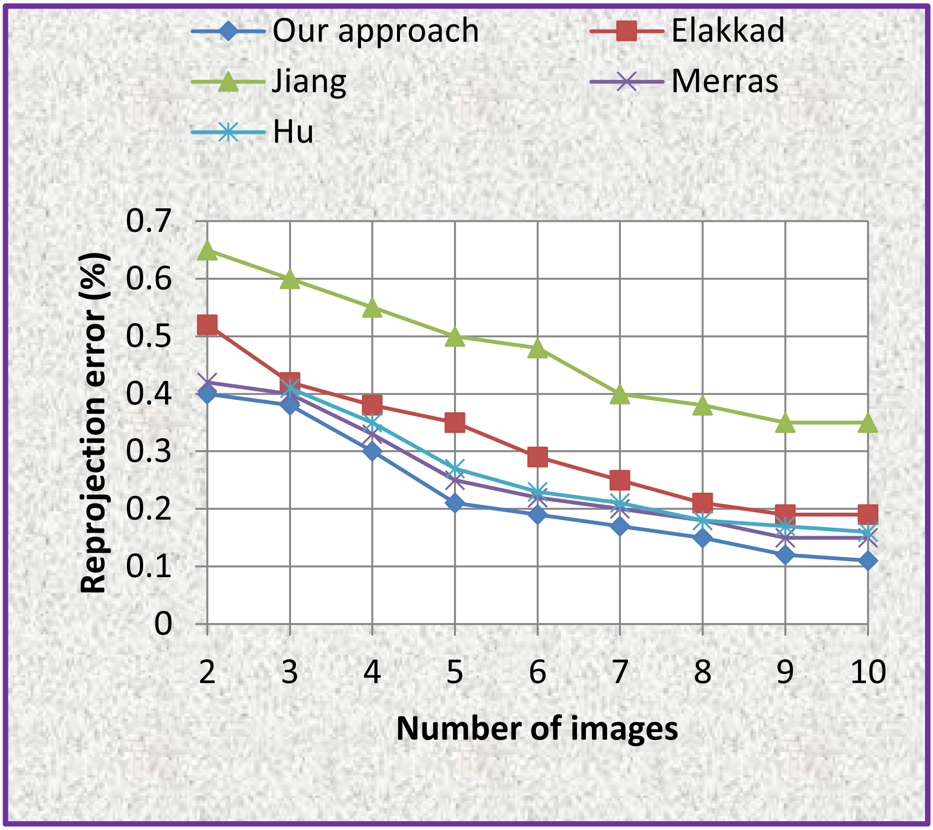

Fig. 14 Reprojection error of the recovered 3D points according to the

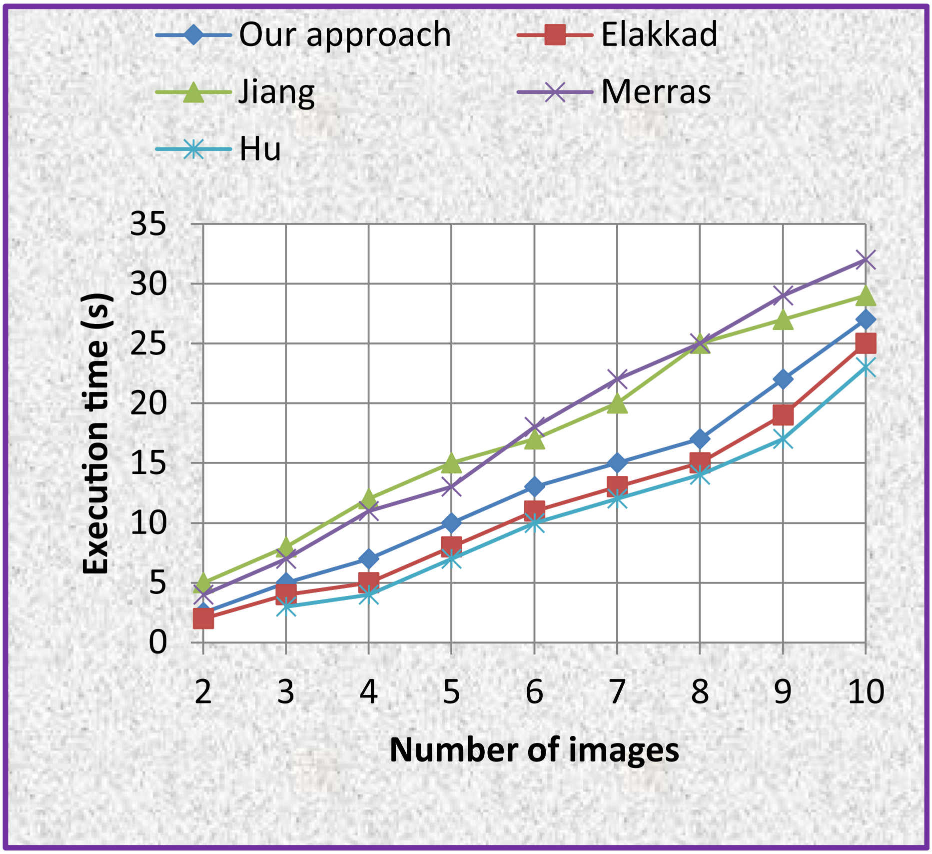

Fig. 15 Computation time according to number of images for different number of images.approaches

The comparison between our method and a textured reconstruction method based on camera self calibration and Genetics algorithms [18] present a big difference from our results (Fig. 16 (b)) and result obtained by Elakkad approach (Fig. 16 (b)). Visually, our results are reliable and more real thanks to the right choice of triangulation method and texture mapping. Our system is robust and can achieve a very satisfactory reconstruction quality.

(a) (b)

Fig. 16 Comparison of textured 3D reconstruction between our method (**b**) and an example of used images in reconstruction (**a**) for method presented in [18]

5 Conclusion

All in all we have described a new approach for building 3D models for any complex scenes from two stereoscopic images taken from different views by uncalibrated cameras has been presented here. The proposed technique consists at modeling 3D scenes without any prior knowledge on the scenes based on the combination between genetic algorithm and bundle adjustment. We suggest a method based on genetic algorithms and using the relationships between two points of the 3D scene and their projections in the image planes for reliable initialization of the intrinsic parameters. Based on these last results the extrinsic parameters are estimated using Posit algorithm in order to obtain the initial 3d cloud points. The minimization of the reprojection errors using a new bundle adjustment (without camera poses initialization) based on Levenberg-Marquardt algorithm provides the optimal intrinsic and extrinsic camera parameters. At the end, the resolution of a linear system formulated from the camera parameters and the matching between a pair of images produced a set of optimal 3d cloud point. The realistic 3D model of the scene is obtained after the mesh generation and the texture mapping.

Robustness of our methodology apparent in the use of complex 3D scenes without any prior knowledge and finding the 3D coordinates of the cloud point in the scene while assessing their accuracy through the optimization of the reprojection error. The different three-dimensional reconstruction examples presented here, as well as the comparison study between the proposed approach and state-of-the-art modeling methods, illustrates the robustness, quality, and reliability of our 3D reconstruction system.

References

1. M.A. Ameller, A. Bartoliet, and L. Quan, “Minimal metric structure and motion from three affine images,” In Proc of the Fifth Asian Conference on Computer Vision, Melbourne, Australia, Volume 1, p:356–361, Jaunuary 2002
2. Armstrong M, Zisserman A, Hartley R (1996) Self-calibration from image triplets. In: Buxton B, Cipolla R (eds) Proceedings of the 4th European conference on computer vision, Cambridge, England, volume 1064 of lecture notes in computer science, pages 3.16. Springer- Verlag, April
3. Baataoui, A., El batteoui, I., Saaidi, A., Satori, K. (2012) Camera self calibration by an equilateral triangle. Int J Comput Appl., 29–34
4. P. Beardsley, P. Torr, and A. Zisserman. 3D Model Acquisition From Extended Image Sequences. In *Proc. of ECCV’96*, pages 683–695, 1996
5. Blumenthal-Barby DC, Eisert P (2014) High resolution depth for binocular image-based modeling. Comput Graph 39:89–100
6. Brooks MJ, de Agapito L, Huynh DQ, Baumela L (1996) Direct methods for self-calibration of a moving stereo head. In: Buxton B, Cipolla R (eds) Proceedings of the 4th European conference on computer vision, Cambridge, England, volume 1065 of lecture notes in computer science, pages 415.426. Springer-Verlag, April
7. M.J. Brooks, W. Chojnacki, and L. Baumela. Determining the ego-motion of an uncalibrated camera from instantaneous optical flow. Journal of the Optical Society of America, 14(10), October 1997
8. Cao X, Xiao J, Foroosh H, Shah M (2006) Self-calibration from turn-table sequences in presence of zoom and focus. Comput Vis Image Underst 103(2):227–237
9. Cazals F, Giesen J (2004) Delaunay triangulation based surface reconstruction: ideas and algorithms. Technical report, RR-5393, INRIA
10. Cefalu A, Haala N, Fritsch D (2016) Structureless bundle adjustment with self-calibration using accumulated constraints. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, pp 3–9

11. Chambon, S., Crouzil, A. (2004) “Correlation measures for color images ”. In traitement du signal 2004_volume 21_numéro spécial L'image numérique couleur
12. Chambon S, Crouzil A (2011) Similarity measures for image matching despite occlusions in stereo vision. Pattern Recogn 44(9):2063–2075
13. Cornou, S., Dhome, M., & Sayd, P. (2002). Bundle adjustment: a fast method with weak initialisation. In Proceedings of the British machine vision conference 2002
14. David P, Dementhon D, Duraiswami R, Samet H (2004) SoftPOSIT: simultaneous pose and correspondence determination. Int J Comput Vis 59(3):259–284
15. L. S. Davis, D.F. Dementhon. – Model-Based Object in 25 Lines of Code,” International Journal of Computer Vision, pp. 15(2) :123–141, 1995
16. A. El abderrahmani, A. Saaidi and K. Satori (December 2010) Robust technique for self-calibration of cameras based on a circle. ICGST-GVIP, Volume 10, Issue 5
17. El akkad, N., Saaidi, A., Satori, K. (2012) Self-calibration based on a circle of the cameras having the varying intrinsic parameters. In: Proceedings of IEEE International Conference on Multimedia Computing and Systems, pp. 161–166
18. El Akkad, N., et al. (2016). Reconstruction of 3D scenes by camera self-calibration and using genetic algorithms. 3D research
19. El Hazzat S, Saaidi A, Satori K (2014) Euclidean 3D reconstruction of unknown objects from multiple images. J Emerg Technol Web Intell 6(1):59–63
20. El hazzat S., Akkad, N., Saaidi, A., Satori, K..: 3D reconstruction system based on incremental structure from motion using a camera having the varying intrinsic parameters. Visual Computer (2018)
21. El hazzat, S., Akkad, N., Saaidi, A., Satori, K (2018) Enhancement of sparse 3D reconstruction using a modified match propagation based on particle swarm optimization. Multimedia Tools and Applications
22. A.Elabderrahmani, A.Saaidi and K.Satori. Planar self-calibration with less constraint. IJCST Vol. 2, Issue 2, June 2011
23. Fischler MA, Bolles RC (1981) Random sample consensus: aparadigm for model fittingwith applications to image analysis and automated cartography. Commun ACM 24(6):381–395
24. A. Fitzgibbon and A. Zisserman. Automatic Camera Recovery for Closed or Open Image Sequences. In *Proceedings of ECCV'98*, pages 311–326, 1998
25. Gao, Y., Radha, H. (2004) A multistage camera self-calibration algorithm. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 537–540
26. Gouet V, Boujemaa N (2002) About optimal use of color points of interest for content-based image retrieval. Internal Report, INRIA Rocquencourt
27. P.Gurdjos and P.Sturm (2003) Methods and geometry for plane-based self-calibration. CVPR, pp. 491–496
28. M. Hanet and T. Kanade, “Multiple motion scene reconstruction from uncalibrated views,” In Proc of the 8th International Conference on computer Vision, Vancouver, Canada, July 2001
29. Harris C, Stephens M (1988) A combined corner and edge detector. In Alvey Vision Conference
30. R. Hartley. Euclidean Reconstruction from Uncalibrated Views. In J. Mundy and A. Zisserman, editors, *Applications of Invariance in Computer Vision*, volume 825 of Lecture Notes in Computer Science, pages 237–256. Springer-Verlag, 1993
31. Hartley R, Trumpf J, Dai Y, Li H (2013) Rotation averaging. Int J Comput Vis 103(3):p267–p305
32. Hemayed, E.E. (2003) A survey of camera self-calibration. In: Proceedings of the IEEE Conference on Advanced Video and Signal BasedSurveillance, pp. 351–357
33. J. H. Holland (1975) Adaptation in natural and artificial systems. University Bibliographieof Michigan Press: Ann Arbor
34. Hu Z, Zhao B (2019) Minimal vision system for linear camera self-calibration and 3D reconstruction using translation platform. Multimed Tools Appl 78:25731–25751. https://doi.org/10.1007/s11042-019-7666-5
35. Indelman, V., 2012a. Bundle adjustment without iterative structure estimation and its application to navigation. Position location and navigation symposium (PLANS), 2012 IEEE/ION
36. Indelman, V., Roberts, R., Beall, C., & Dellaert, F., 2012b. Incremental light bundle adjustment. Proceedings of the British machine vision conference (BMVC 2012), pp. 3-7
37. Jiang, Z., Liu, S.: (2011) The self-calibration of varying internal camera parameters based on image of dual absolute quadric transformation. In: Information and Automation, Communications in Computer and Information Science, vol. 86, pp. 452–461. Springer, Berlin
38. Jiang Z, Liu S (2012) Self-calibration of varying internal camera parameters algorithm based on quasi-affine reconstruction. J Comput 7(3):774–778
39. Le V, Tang H, Cao L, Huang T (2010) Accurate and efficient reconstruction of 3D faces from stereo images. In: 17th IEEE international conference on image processing (ICIP), pp 4265-4268, 2010
40. Liangfu Li, Zuren Feng, Yuanjing Feng. Accurate calibration of stereo cameras for machine vision. JCS&T Vol. 4 No. 3, October 2004

41. Liu, P., Shi, J., Zhou, J., Jiang, L (2003) Camera self-calibration using the geometric structure in real scenes. In: Proceedings of the Computer Graphics International, pp. 262–265
42. Lourakis M, Argyros A (Aug. 2005) Efficient. Causal Camera Tracking in Unprepared Environments Computer Vision and Image Understanding Journal 99(2):259–290
43. Manolis I.A. Lourakis and R. Deriche. (2000) Camera self-calibration using the kruppa equations and the SVD of the fundamental matrix: the case of varying intrinsic parameters. Technical Report 3911, INRIA
44. Xiaoqiao Meng, Hua Li and Zhanyi Hu. A new easy camera calibration technique based on circular points, bmvc2000
45. Merras M, Saaidi A, El Akkad N, Satori K (2018) Multi-view 3D reconstruction and modeling of the unknown 3D scenes using genetic algorithms. Soft Comput 22:6271–6289. https://doi.org/10.1007/s00500-017-2966-z
46. Pollefeys M, Koch R, Gool LV (1999) Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. Int J Comput Vis 32(1):7–25
47. M. Pollefeys, L. V. Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch. Visual Modeling With a Hand-Held Camera. *IJCV*, 59(3):207–232, Sep./Oct. 2004
48. P. Ram and S. Padmavathi, “ Analysis of Harris Corner Detection For Color Images,” In International conference on Signal Processing, Communication, Power and Embedded System (SCOPES)-2016
49. Rodriguez, A. L., López-de-Teruel, P. E., & Ruiz, A., 2011a. Reduced epipolar cost for accelerated incremental SfM. Computer vision and pattern recognition (CVPR), 2011 IEEE conference on
50. Rodriguez, A. L., López-de-Teruel, P. E., & Ruiz, A., 2011b. GEA optimization for live structureless motion estimation. Computer vision and pattern recognition (CVPR), 2011 IEEE conference on
51. Saaidi A, Halli A, Tairi H, Satori K (2008) Self-calibration using a particular motion of camera. WSEAS Tran Comput Res 3(5):295–299
52. Saaidi, A., Halli, A., Tairi, H., Satori, K. (2009) Self-calibration using a planar scene and parallelogram. In: ICGST-GVIP, pp. 41–47
53. B.Satouri, A. Elabderrahmani, H. Tairi and K.satori. Camera Self-Calibration with Varying Intrinsic Parameters by an Unknown Three-Dimensional Scene. International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 7, No. 9, 2016
54. Shang Y, Yue Z, Chen M, Song Q (2012) A new method of camera self-calibration based on relative lengths. Inf Technol J 11(3):376–379
55. Steffen, R., Frahm J.-M., & Förstner W., 2010. Relative bundle adjustment based on trifocal constraints. Trends and topics in computer vision. Springer Berlin Heidelberg, 2012. pp. 282–295
56. Sturm P (2000) A case against Kruppa’s equations for camera selfcalibration. IEEE Trans Pattern Anal Mach Intell 22:1199–1204
57. Sturm P (2002) Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. Image VisComput 20:415–426
58. Sun J, Hongbin G (2009) Research of linear camera calibration based on planar pattern. World Acad Sci Eng Technol 60
59. Tarek A. Al-Saeed1, Nahed H. Solouma2, and Said M. El-Sherbiny3. Retinal Motion Detection and 3D Structure Recovery From Two Perspective Views. GVIP Special Issue on Medical Image Processing, March, 2006
60. Torr PHS, Murray DW (1997) The development and comparison of robust methods for estimating the fundamental matrix. Int JComput Vis 24:271–300
61. Triggs B (1998) Autocalibration from planar scenes. In: Proceedings of the 5th European Conference on Computer Vision, pp. 89–105
62. Triggs B, McLauchlan P, Hartley R, Fitzgibbon A (1999) Bundle adjustment – a modern synthesis. In Proceedings of the International Workshop on Vision Algorithms: Theory and Practice, pages 298–372
63. Triggs B, McLauchlan PF, Hartley RI, Fitzgibbon AW (2000) Bundle adjustment—a modern synthesis. In vision algorithms: theory and practice. Springer, Berlin Heidelberg, pp 298–372
64. B. Triggs, P. McLauchlan, R. Hartley, A. Fitzgibbon. (2000) Bundle adjustment - a modern synthesis. Vision Algorithms: Theory et Practice,” International Workshop on Vision Algorithms
65. Triggs, P. McLauchlan, R. Hartley, A. Fitzgibbon (2000) Bundle adjustment - a modern synthesis. Vision Algorithms: Theory et Practice. International Workshop on Vision Algorithms
66. G. Wang, Q. M. J. Wu. Perspective 3D Euclidean reconstruction with varying camera parameters. IEEE Transactions on Circuits and Systems for Video Technology Vol. 19, N°. 12, pp. 1793–1803. 2009
67. Wiles C, Brady M (1996) Ground plane motion camera models. In: Buxton B, Cipolla R (eds) Proceedings of the 4th European conference on computer vision, Cambridge, England, volume 1065 of lecture notes in computer science, pages 238.247. Springer-Verlag, April
68. Zhang W (2005) A simple method for 3D reconstruction from two views. In: GVIP 05 Conference

69. Zhao Y, Lv XD (2012) An approach for camera self-calibration using vanishing-line. Inf Technol J 112: 276–282
70. Zhao Y, Hu X, Lv X, Wang H (2012) Solving the camera intrinsic parameters with the positive tri-prism based on the circular points. Inf Technol J 11(7):926–930

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Boutaina Satouri received the bachelor's and master's degrees from SMBA Fez University in 2003 and 2009 respectively. He is currently working toward the PhD degree in the LIIAN Laboratory at SMBA Fez University. His current research interests include camera self-calibration and 3D reconstruction.

Khalid Satori received the PhD degree from the National Institute for the Applied Sciences INSA at Lyon in 1993. He is currently a Professor of Computer Science at SMBA-Fez University. He is the Director of the LIIAN Laboratory. His research interests include real-time rendering, image based rendering, virtual reality, biomedical signal, camera self-calibration, genetic algorithms and 3D reconstruction.

Abdellatif El Abderrahmani received the bachelor's and master's degrees from SMBA-Fez University in 2002 and 2007 respectively. He is currently a Professor of Computer Science at Larache Poly Disciplinary School University. His current research interests include camera self-calibration, 3D reconstruction and real-time rendering.