ORIGINAL ARTICLE

Camera self-calibration with varying intrinsic parameters by an unknown three-dimensional scene

Nabil El Akkad · Mostafa Merras · Abderrahim Saaidi · Khalid Satori

Abstract This work proposes a method of camera self-calibration having varying intrinsic parameters from a sequence of images of an unknown 3D object. The projection of two points of the 3D scene in the image planes is used with fundamental matrices to determine the projection matrices. The present approach is based on the formulation of a nonlinear cost function from the determination of a relationship between two points of the scene and their projections in the image planes. The resolution of this function enables us to estimate the intrinsic parameters of different cameras. The strong point of the present approach is clearly seen in the minimization of the three constraints of a self-calibration system (a pair of images, 3D scene, any camera): The use of a single pair of images provides fewer equations, which minimizes the execution time of the program, the use of a 3D scene reduces the planarity constraints, and the use of any camera eliminates the constraints of cameras having constant parameters. The experiment results on synthetic and real data are presented to demonstrate the performance of the present approach in terms of accuracy, simplicity, stability, and convergence.

N. El Akkad · M. Merras · K. Satori
LIIAN, Department of Mathematics and Computer Science, Faculty of Science, Dhar El Mahraz, Sidi Mohamed Ben Abdellah University, B.P. 1796, Atlas, Fez, Morocco

N. El Akkad
e-mail: nabil_abdo80@yahoo.fr

A. Saaidi (✉)
LIMAO, Department of Mathematics, Physics and Computer Science, Polydisciplinary Faculty of Taza, Sidi Mohamed Ben Abdellah University, B.P. 1223, Taza, Morocco
e-mail: saaidi.abde@yahoo.fr

1 Introduction

The development of new approaches and technologies in computer vision, also called artificial vision, can improve or complement human vision in several areas. Applications of computer vision are numerous, including recognition, object localization, 3D reconstruction, mobile robotics, and medical imaging. The reasons for using computer vision are the possibilities offered by computer science: speed of processing a large volume of information, reliability, and availability. What is of interest to us in this work is the use of computer vision for three-dimensional perception of the environment, also called 3D vision. The input data to be treated are images taken by cameras from different views. A 3D vision system is designed to extract the information of a three-dimensional nature; this requires the calibration of the cameras. The parameters of the cameras can be estimated by two major methods: calibration and self-calibration. In this paper, we are interested in the self-calibration methods that can calibrate the cameras without any prior knowledge about the scene. The standard process of most of these methods is to search for equations according to intrinsic parameters and the invariants in the images, whose aim generally is to solve a nonlinear equation system. The algorithm used to solve this system requires two steps, initialization and optimization of a cost function. Self-calibration of the cameras is the main step to obtain three-dimensional coordinates of points from matches between pairs of images. Several methods of camera self-calibration with constant intrinsic parameters [1–8] and those with varying intrinsic parameters [9–20] are treated in this area.

Problems in the field of camera self-calibration are usually related to the constraints of the camera self-calibration system (the number of images used, the characteristics of the cameras, and the type of scene). These constraints limit the majority of the methods described in the literature and entail many difficulties; for example, self-calibration with a large number of images provides more equations, and its resolution becomes more complicated and requires powerful algorithms, and, finally, the execution time to optimize the solution increases. The methods developed by the use of cameras with constant intrinsic parameters are limited in cases where cameras having varying intrinsic parameters are used. Moreover, the planar scenes rarely exist in nature, which shows the importance of using 3D scenes.

A new approach will be discussed in this paper. It is an approach of camera self-calibration having the varying intrinsic parameters by the use of an unknown three-dimensional scene. After the detection of control points in the images by the Susan approach [24] and the matching of these points in each pair of images by the correlation measure *NCC* [28], the fundamental matrix can be estimated from eight matches by the Ransac algorithm [22]. This matrix is used with the projection of two points of the 3D scene in images taken by different views in order to formulate linear equations. Solving these equations allows the estimation of the projection matrices. The determination of a relationship between the two points of the 3D scene and their projections in the planes of the images i and j and the relationships between the images of the absolute conic allow the formulation of a nonlinear cost function. The minimization of this function by the Levenberg–Marquart algorithm [23] allows the estimation of the intrinsic parameters of the cameras used.

This new approach has several advantages: the use of any camera (with varying intrinsic parameters), two images only are sufficient to estimate the cameras' intrinsic parameters, and the use of an unknown 3D scene (more common in nature). These advantages allow us, on the one hand, to solve some problems mentioned in the previous paragraph and related to the self-calibration system and, on the other hand, to work freely in the domain of self-calibration with fewer constraints.

The remainder of this paper is organized as follows. The second part explains the related work. The third part presents the camera model. Camera self-calibration is explained in the fourth part. The experiment results are discussed in the fifth part, and the conclusions are presented in the sixth part.

2 Related work

This section presents a summary of some methods of cameras self-calibration. It starts with methods based on constant intrinsic parameters, and, in the third paragraph of this section, it moves to methods based on the varying intrinsic parameters. Several approaches have been discussed in the literature on constant parameters. A theoretical and practical method is presented in [1]; it is based on the projection of two circular points of an unknown planar scene in different images and the use of homographies between the images (at least five images) to estimate the unknown intrinsic parameters of cameras. The use of the Kruppa equations in [2] allows the estimation of the cameras' parameters. These equations are obtained from information contained in the images taken by a camera, without any knowledge about its movement and the structure of the 3D scene used. The minimization of the nonlinear cost function [3] allows the estimation of the camera parameters. This function is formulated from a particular movement of the camera, "a translation and a small rotation," and the estimation of the homography of the plane at infinity. The use of vanishing line is the main idea discussed in [4]. The solving of three linear equations formulated from the circles and their respective center allows determining the vanishing line. The camera's intrinsic parameters are estimated from the theory of these lines and the circular points. A simple method is discussed in [5]. It is based on the calculation of the fundamental matrix (estimated from the matches between pairs of images) and the proposal of some constraints (the pixels are square, and principal point is in the image center) to estimate the focal length of two cameras. A method in [6] uses an unknown planar scene and parallelogram for camera self-calibration. The projection matrices are determined from the projection of the parallelogram vertices on the three images used. They are used with the relationship between the images of the absolute conic and the homographies between image pairs to formulate a nonlinear cost function. The resolution of this function allows estimating the camera's intrinsic parameters. In [7], a new method is based on geometric constraints to estimate the intrinsic and extrinsic parameters of the camera. The use of geometric constraints on the first image provides the initial solution, and the minimization between the scene points and their projections in the image planes by exploiting geometric constraints on the two images allows optimizing the initial solution. The main idea presented in [8] is based on the use of an equilateral triangle. The projection of the triangle vertices in the planes of the three images allows obtaining the projections matrices. The resolution of a nonlinear cost function, which is formulated from homography matrices, projections, and the relationship between the images of the absolute conic, allows obtaining the intrinsic camera parameters.

Each method presented in the previous paragraph is based on a new idea, which is different from the ideas developed by other methods, but each method has specific advantages and disadvantages. For example, one of the disadvantages of the method presented in [1, 6], and [8] is the use

of a planar scene. Moreover, the method [1] requires at least five images to calibrate the camera used. The weak point of the method [3] is the use of a simple rotation of the camera instead of using arbitrary rotations. Among the drawbacks of the method [5], there is the proposal of some constraints on the self-calibration system (the pixels are square, and the principal points are at the center of the images). In addition to that, all the methods presented in the previous paragraph are based on the self-calibration of camera with constant parameters. These methods are limited in cases where cameras having varying parameters are used.

This section presents some methods of camera self-calibration with varying intrinsic parameters. A theoretical and practical method is presented in [9]; it is based on the recovery of metric reconstruction from a set of images taken by cameras, and the absence of the skew factor is sufficient to estimate the intrinsic parameters of different cameras. The camera's intrinsic parameters are estimated in [10] by simplifying the Kruppa equations by the singular value decomposition of the fundamental matrices and uncertainty related to these matrices. A method of camera self-calibration with varying and unknown focal length (other parameters are considered known) is treated in [11]. It is based on the complete derivation of critical motion sequences to estimate the focal lengths of different cameras. A new method of camera self-calibration with varying focal length is discussed in [12]. The intrinsic camera parameters and those related to the observed Euclidean structure are estimated from an image sequence of an object whose Euclidean structure is unknown, by solving a nonlinear cost function, but the initialization step of the focal length leads to some problems. To solve these problems, a new formulation of the cost function independently from the focal length is proposed. The use of a constant movement between images taken from different views of an object rotating around a single axis is the main idea discussed in [13]. A nonlinear cost function is defined from the relationship between the projection matrices and fundamental matrices (determined from the matched points), and its minimization provides the camera parameters. The idea discussed in [14] is based on the transformation of the image of the absolute dual quadric to get the camera parameters. The solution obtained from the estimation of elements of the dual absolute quadric is very sensitive to noise because these elements have a large difference of magnitude. To solve this problem, the authors propose a transformation to obtain the same magnitude for all these elements, and they show that the solution becomes more stable by this transformation. A new self-calibration method based on the relative distance is presented in [15]. A nonlinear cost function is defined from invariant relative distance and homography matrix (which transforms the projective reconstruction to the metric reconstruction), whose elements depend on the intrinsic camera parameters. Its minimization provides the

camera parameters and the 3D structure of the scene. The idea discussed in [16] is based on the quasi-affine reconstruction; the camera intrinsic parameters are estimated after this reconstruction, which is performed from the homography matrix of the plane at infinity, and constraints on the image of the absolute conic. In [17], a method with a positive tri-prism is based on circular points. The coordinates of these points are determined from the properties of the positive tri-prism, and after determining the coordinates of the circular points in the images and the vanishing points of the tri-prism corners, the intrinsic camera parameters can be estimated linearly. A new multistage approach of low complexity is presented in [18]. The intrinsic camera parameters are obtained by deriving an optimization function, which is expressed according to these parameters, and the refinement of the estimated parameters is realized by a multistage algorithm. A method based on a circle is presented in [19]. The camera parameters are estimated from the use of at least four images. The projection of two points of the planar scene in different images allows estimating the four projection matrices. To obtain a nonlinear cost function, these matrices are exploited with the images of the absolute conic and homographies. The minimization of this function allows the estimation of the intrinsic parameters of the cameras used.

The advantage of the methods presented in the previous paragraph is the use of cameras characterized by varying parameters (the use of any cameras). However, each method has specific disadvantages. For example, in [9] the authors propose the absence of skew factor ($= 0$). The authors of [11] assume that the principal point, the scale factor, and the skew factor are known. The method of [13] is based on a constant movement (and not a random one) of cameras when taking images. Therefore, it is limited in cases of any movement. The authors of the article [16] use some constraints on the image of the absolute conic. The method [18] requires at least four images to estimate the camera parameters.

A detailed study up to 2003 on methods of camera self-calibration with constant and varying intrinsic parameters is presented in [20].

3 Camera model

Figure 1 shows the pinhole model used in this work. This model projects the scene in the image planes; a 3×4 matrix characterizes this model; for the camera i, it is defined by $M_i(R_i \ t_i)$ with a 3×4 matrix $(R_i \ t_i)$ containing extrinsic parameters R_i, the rotation matrix, and t_i, the translation vector of camera in space; M_i is a 3×3 matrix containing the intrinsic parameters and is expressed as follows:

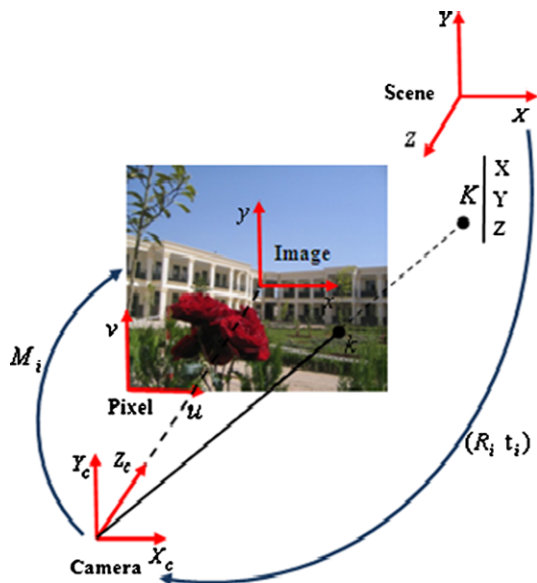$$M_i = \begin{pmatrix} f_i & s_i & x_{0i} \\ 0 & \varepsilon_i f_i & y_{0i} \\ 0 & 0 & 1 \end{pmatrix}, \tag{1}$$

Fig. 1 Pinhole camera model

where f_i is the focal length, ε_i is the scale factor, s_i is the skew factor, and (x_{0i}, y_{0i}) represent the coordinates of the principal point in the images.

The passage of the scene reference to the camera reference (Fig. 1) is made by the 3×4 matrix $(R_i\ t_i)$, and the passage of the camera reference to the image reference (Fig. 1) is performed by the 3×3 matrix M_i. A point k of the image is the projection of a point K of the 3D scene. This projection is conducted by the following formula: $k \sim M_i(R_i\ t_i)K$, where $\sim$ indicates equality up to multiplication by a nonzero scale factor. $(x\ y\ 1)^T$ and $(X\ Y\ Z\ 1)^T$ are the homogeneous coordinates of the points k and K, respectively.

4 Camera self-calibration

4.1 Detection and matching of control points

Control points in the images can be detected by several methods [24–27]. This article uses the Susan algorithm [24].

The matching of control points is an important step to calibrate the cameras used. Several studies [28–30] are developed in this area. This approach uses the mapping by Normalized Cross Correlation NCC [28].

4.2 Estimation of the projection matrices

Let K_1 and K_2 denote the two points on the 3D scene. Considering a plane π that contains these two points, let us denote the Euclidean reference by $\mathcal{R}(O\ X\ Y\ Z)$ such that the center O of the reference coincides with the midpoint

of segment $[K_1 K_2]$ and $Z \perp \pi$. The homogeneous coordinates of the two points K_1 and K_2 (Fig. 2) in the reference $\mathcal{R}(O\ X\ Y)$ are given as follows:

$$K_1 = (r \cos \emptyset, r \sin \emptyset, 1)^T \quad \text{and}$$

$$K_2 = (-r \cos \emptyset, -r \sin \emptyset, 1)^T$$

where $r = K_1 K_2/2$, and $\emptyset$ is the angle between the line $(K_1 K_2)$ and the x-axis (X).

Considering two homographies H_i and H_j that can project the plane π in images i and j, therefore, the projection of the two points K_1 and K_2 can be given by the following expressions:

$$k_{im} \sim H_i K_m, \tag{2}$$

$$k_{jm} \sim H_j K_m, \tag{3}$$

where $m = 1, 2$ and k_{im}, k_{jm}, respectively, represent the points in the images i and j that are the projections of the two vertices K_1 and K_2 of the 3D scene, and H_n $(n = i, j)$ represent the homography matrices that are expressed as follows:

$$H_n \sim M_n R_n \begin{pmatrix} 1 & 0 & \\ 0 & 1 & R_n^T t_n \\ 0 & 0 & \end{pmatrix}, \quad n = i, j. \tag{4}$$

Expressions (2) and (3) can be written as follows:

$$k_{im} \sim H_i E K_m', \tag{5}$$

$$k_{jm} \sim H_j E K_m', \tag{6}$$

where $E = \begin{pmatrix} r \cos \emptyset & 0 & 0 \\ 0 & r \sin \emptyset & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and

$$K_m' = \begin{pmatrix} q \\ q \\ 1 \end{pmatrix} \quad \left\{ \begin{matrix} m = 1 & \Leftrightarrow & q = 1 \\ m = 2 & \Leftrightarrow & q = -1 \end{matrix} \right\}.$$

Let us represent the projection matrices by

$$P_n \sim H_n E, \quad n = i, j, \tag{7}$$

where P_i and P_j represent the projection matrices of the two points K_1' and K_2' in images i and j (Fig. 2).

Formula (7) gives:

$$P_j \sim H_{ij} P_i, \tag{8}$$

where

$$H_{ij} \sim H_j H_i^{-1}; \tag{9}$$

H_{ij} is the homography between the images i and j.

Expressions (5), (6), and (7) give:

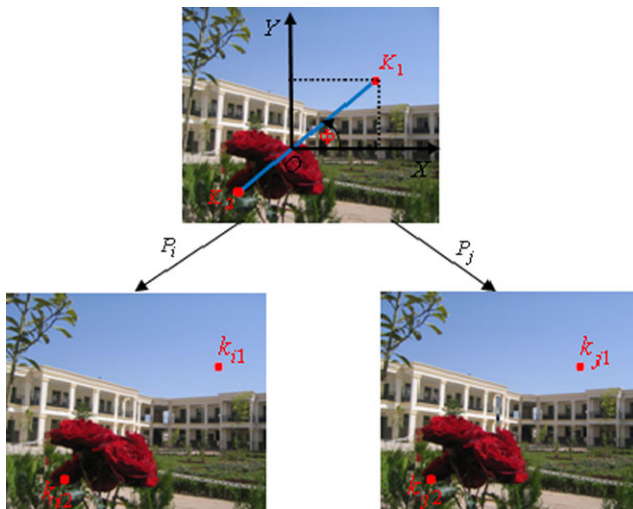$$k_{im} \sim P_i K_m', \tag{10}$$

$$k_{jm} \sim P_j K_m'. \tag{11}$$

Fig. 2 The projection of the two points K_1 and K_2 in the planes of images i and j

In addition, expressions (8) and (11) yield the following formula:

$$k_{jm} \sim H_{ij} P_i K'_m, \tag{12}$$

and expression (12) leads to

$$e'_j k_{jm} \sim e'_j H_{ij} P_i K'_m. \tag{13}$$

Knowing that $F_{ij} \sim e'_j H_{ij}$, where F_{ij} is the fundamental matrix between the images i and j, it is estimated from eight matches between these images by the Ransac algorithm [22],

$$e'_j = \begin{pmatrix} 0 & -e_{j3} & e_{j2} \\ e_{j3} & 0 & -e_{j1} \\ -e_{j2} & e_{j1} & 0 \end{pmatrix} \quad \text{and}$$

$e_j = (e_{j1} \; e_{j2} \; e_{j3})^T$ is the epipole of the image j, which is estimated from the fundamental matrix by the formula $F_{ij}^T e_j = 0$.

Therefore,

$$e'_j k_{jm} \sim F_{ij} P_i K'_m. \tag{14}$$

Expression (10) gives

$$k_{i1} \sim P_i K'_1 \tag{15}$$

and

$$k_{i2} \sim P_i K'_2. \tag{16}$$

Therefore, the latter two relations give four equations with eight unknowns, which are the P_i elements.

Expression (14) gives:

$$e'_j k_{j1} \sim F_{ij} P_i K'_1 \tag{17}$$

and

$$e'_j k_{j2} \sim F_{ij} P_i K'_2. \tag{18}$$

Therefore, each of expressions (17) and (18) gives two equations with eight unknowns that are the P_i elements.

Expressions (15), (16), (17), and (18) give eight equations with eight unknowns that are the P_i elements (each of these expressions gives two equations with eight unknowns that are the P_i elements).

So, the P_i parameters can be estimated from these eight equations with eight unknowns.

Expression (8) gives

$$e'_j P_j \sim e'_j H_{ij} P_i. \tag{19}$$

Therefore,

$$e'_j P_j \sim F_{ij} P_i. \tag{20}$$

The above expression gives eight equations with eight unknowns which are the P_j elements. Therefore, the P_j parameters can be estimated from these eight equations with eight unknowns.

4.3 Self-calibration equations

A nonlinear cost function will be defined in the main part of this work from the determination of the relationships between the images of the absolute conic (ω_i and ω_j) and from the relationships between two points (K_1, K_2) of the 3D scene and their projections (k_{i1}, k_{i2}) and (k_{j1}, k_{j2}) in the planes of the images i and j, respectively. The different relationships are based on some techniques of projective geometry. The defined cost function will be minimized by the Levenberg-Marquart algorithm [23] to estimate the ω_i and ω_j elements and, finally, by the intrinsic parameters of the cameras used.

Expression (10) gives

$$\lambda_{im} k_{im} = P_i K'_m, \tag{21}$$

where

$$P_i = \begin{pmatrix} P_{i11} & P_{i12} & P_{i13} \\ P_{i21} & P_{i22} & P_{i23} \\ P_{i31} & P_{i32} & P_{i33} \end{pmatrix}, \qquad k_{im} = \begin{pmatrix} x_{im} \\ y_{im} \\ 1 \end{pmatrix},$$

and

$$\lambda_{im} = q P_{i31} + q P_{i32} + P_{i33}.$$

λ_{im} is a nonzero scale factor that is used to switch from expression (10) to expression (21) (the transition between equality with a scale factor $\sim$ to precise equality $=$). The value of λ_{im} is determined from expression (21).

Therefore, formula (21) leads to

$$\lambda_{im} k'_{im} = P_i K''_m, \tag{22}$$

where

$$k'_{im} = \begin{pmatrix} x_{im} & \frac{P_{i12}}{\lambda_{im}} & \frac{P_{i13}}{\lambda_{im}} \\ y_{im} & \frac{P_{i22}}{\lambda_{im}} & \frac{P_{i23}}{\lambda_{im}} \\ 1 & \frac{P_{i32}}{\lambda_{im}} & \frac{P_{i33}}{\lambda_{im}} \end{pmatrix}, \qquad K''_m = \begin{pmatrix} q & 0 & 0 \\ q & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Expression (22) gives

$$P_i \sim k'_{im} K_m^{''-1} \tag{23}$$

and the same for P_j:

$$P_j \sim k'_{jm} K_m^{''-1}. \tag{24}$$

Expressions (4) and (7) give

$$P_n \sim M_n R_n \begin{pmatrix} 1 & 0 \\ 0 & 1 & R_n^T t_n \\ 0 & 0 \end{pmatrix} E, \quad n = i, j. \tag{25}$$

The previous formula leads to

$$M_n^{-1} P_n \sim R_n \begin{pmatrix} 1 & 0 \\ 0 & 1 & R_n^T t_n \\ 0 & 0 \end{pmatrix} E, \quad n = i, j. \tag{26}$$

Formula (26) gives

$$P_i^T \omega_i P_i \sim \begin{pmatrix} E'^T E' & E'^T R_i^T t_i \\ t_i^T R_i E' & t_i^T t_i \end{pmatrix}, \tag{27}$$

where $\omega_i = (M_i M_i^T)^{-1}$ is the image of the absolute conic, and

$$E' = \begin{pmatrix} r\cos\emptyset & 0 \\ 0 & r\sin\emptyset \\ 0 & 0 \end{pmatrix}, \tag{28}$$

and the same for P_j:

$$P_j^T \omega_j P_j \sim \begin{pmatrix} E'^T E' & E'^T R_j^T t_j \\ t_j^T R_j E' & t_j^T t_j \end{pmatrix}. \tag{29}$$

Expressions (23) and (27) give

$$\begin{aligned} \left(k'_{im} K_m^{''-1}\right)^T &\omega_i \left(k'_{im} K_m^{''-1}\right) \\ &\sim \begin{pmatrix} E'^T E' & E'^T R_i^T t_i \\ t_i^T R_i E' & t_i^T t_i \end{pmatrix}. \end{aligned} \tag{30}$$

The previous expression gives

$$\begin{aligned} \left(k'_{i1} K_1^{''-1}\right)^T &\omega_i \left(k'_{i1} K_1^{''-1}\right) \\ &\sim \begin{pmatrix} E'^T E' & E'^T R_i^T t_i \\ t_i^T R_i E' & t_i^T t_i \end{pmatrix}, \end{aligned} \tag{31}$$

And:

$$\begin{aligned} \left(k'_{i2} K_2^{''-1}\right)^T &\omega_i \left(k'_{i2} K_2^{''-1}\right) \\ &\sim \begin{pmatrix} E'^T E' & E'^T R_i^T t_i \\ t_i^T R_i E' & t_i^T t_i \end{pmatrix} \end{aligned} \tag{32}$$

The expressions (31) and (32) give:

$$\begin{aligned} \left(k'_{i1} K_1^{''-1}\right)^T &\omega_i \left(k'_{i1} K_1^{''-1}\right) \\ &\sim \left(k'_{i2} K_2^{''-1}\right)^T \omega_i \left(k'_{i2} K_2^{''-1}\right) \end{aligned} \tag{33}$$

Let

$$C = \begin{pmatrix} c_{11i} & c_{12i} & c_{13i} \\ c_{12i} & c_{22i} & c_{23i} \\ c_{13i} & c_{23i} & c_{33i} \end{pmatrix}$$

denotes the matrix corresponding to $(k'_{i1} K_1^{''-1})^T \omega_i \times (k'_{i1} K_1^{''-1})$ and

$$D = \begin{pmatrix} d_{11i} & d_{12i} & d_{13i} \\ d_{12i} & d_{22i} & d_{23i} \\ d_{13i} & d_{23i} & d_{33i} \end{pmatrix}$$

denotes the matrix corresponding to $(k'_{i2} K_2^{''-1})^T \omega_i \times (k'_{i2} K_2^{''-1})$.

Therefore, the formula (33) gives:

$$\begin{cases} c_{12i} = 0, \quad d_{12i} = 0, \quad \dfrac{d_{11i}}{d_{13i}} = \dfrac{c_{11i}}{c_{13i}} \\ \dfrac{d_{13i}}{d_{22i}} = \dfrac{c_{13i}}{c_{22i}}, \quad \dfrac{d_{22i}}{d_{23i}} = \dfrac{c_{22i}}{c_{23i}}, \quad \dfrac{d_{23i}}{d_{33i}} = \dfrac{c_{23i}}{c_{33i}} \end{cases} \tag{34}$$

The previous expression gives:

$$\begin{cases} c_{12i} = 0 \\ d_{12i} = 0 \\ d_{11i} c_{13i} - c_{11i} d_{13i} = 0 \\ d_{13i} c_{22i} - c_{13i} d_{22i} = 0 \\ d_{22i} c_{23i} - c_{22i} d_{23i} = 0 \\ d_{23i} c_{33i} - c_{23i} d_{33i} = 0 \end{cases} \tag{35}$$

The expressions (24) and (29) give:

$$\begin{aligned} \left(k'_{jm} K_m^{''-1}\right)^T &\omega_j \left(k'_{jm} K_m^{''-1}\right) \\ &\sim \begin{pmatrix} E'^T E' & E'^T R_j^T t_j \\ t_j^T R_j E' & t_j^T t_j \end{pmatrix} \end{aligned} \tag{36}$$

The previous expression gives:

$$\begin{aligned} \left(k'_{j1} K_1^{''-1}\right)^T &\omega_j \left(k'_{j1} K_1^{''-1}\right) \\ &\sim \begin{pmatrix} E'^T E' & E'^T R_j^T t_j \\ t_j^T R_j E' & t_j^T t_j \end{pmatrix}, \end{aligned} \tag{37}$$

and

$$\begin{aligned} \left(k'_{j2} K_2^{''-1}\right)^T &\omega_j \left(k'_{j2} K_2^{''-1}\right) \\ &\sim \begin{pmatrix} E'^T E' & E'^T R_j^T t_j \\ t_j^T R_j E' & t_j^T t_j \end{pmatrix}. \end{aligned} \tag{38}$$

Expression (28) gives

$$E'^T E' = \begin{pmatrix} r^2 \cos^2\emptyset & 0 \\ 0 & r^2 \sin^2\emptyset \end{pmatrix}. \tag{39}$$

Let

$$A = \begin{pmatrix} a_{11j} & a_{12j} & a_{13j} \\ a_{12j} & a_{22j} & a_{23j} \\ a_{13j} & a_{23j} & a_{33j} \end{pmatrix}$$

denote the matrix corresponding to $(k'_{j1} K_1^{''-1})^T \omega_j \times (k'_{j1} K_1^{''-1})$.

Therefore, from formulas (37) and (39) we have

$$a_{12j} = 0. \tag{40}$$

Let

$$B = \begin{pmatrix} b_{11j} & b_{12j} & b_{13j} \\ b_{12j} & b_{22j} & b_{23j} \\ b_{13j} & b_{23j} & b_{33j} \end{pmatrix}$$

denote the matrix corresponding to $(k'_{j2} K_2''^{-1})^T \omega_j (k'_{j2} K_2''^{-1})$.

Then, from the expressions (38) and (39) we get

$$b_{12j} = 0. \tag{41}$$

Equations (37) and (38) give

$$(k'_{j1} K_1''^{-1})^T \omega_j (k'_{j1} K_1''^{-1}) \\ \sim (k'_{j2} K_2''^{-1})^T \omega_j (k'_{j2} K_2''^{-1}). \tag{42}$$

The previous expression gives

$$\begin{cases} \dfrac{b_{11j}}{b_{13j}} = \dfrac{a_{11j}}{a_{13j}}, & \dfrac{b_{13j}}{b_{22j}} = \dfrac{a_{13j}}{a_{22j}}, \\[2mm] \dfrac{b_{22j}}{b_{23j}} = \dfrac{a_{22j}}{a_{23j}}, & \dfrac{b_{23j}}{b_{33j}} = \dfrac{a_{23j}}{a_{33j}}. \end{cases} \tag{43}$$

Therefore, formulas (40), (41), and (43) give

$$\begin{cases} a_{12j} = 0, \\ b_{12j} = 0, \\ b_{11j} a_{13j} - a_{11j} b_{13j} = 0, \\ b_{13j} a_{22j} - a_{13j} b_{22j} = 0, \\ b_{22j} a_{23j} - a_{22j} b_{23j} = 0, \\ b_{23j} a_{33j} - a_{23j} b_{33j} = 0. \end{cases} \tag{44}$$

Expressions (31) and (37) show that the first line and columns of the matrices $(k'_{i1} K_1''^{-1})^T \omega_i (k'_{i1} K_1''^{-1})$ and $(k'_{j1} K_1''^{-1})^T \omega_j (k'_{j1} K_1''^{-1})$ are identical, which gives

$$\frac{c_{11i}}{c_{22i}} = \frac{a_{11j}}{a_{22j}}. \tag{45}$$

Therefore,

$$c_{11i} a_{22j} - a_{11j} c_{22i} = 0. \tag{46}$$

Expressions (32) and (38) show that the first line and columns of the matrices $(k'_{i2} K_2''^{-1})^T \omega_i (k'_{i2} K_2''^{-1})$ and $(k'_{j2} K_2''^{-1})^T \omega_j (k'_{j2} K_2''^{-1})$ are identical, which gives

$$\frac{d_{11i}}{d_{22i}} = \frac{b_{11j}}{b_{22j}}. \tag{47}$$

Therefore,

$$d_{11i} b_{22j} - b_{11j} d_{22i} = 0. \tag{48}$$

Expressions (35), (44), (46), and (48) give a system of fourteen equations with ten unknowns that are expressed as follows:

$$\begin{cases} c_{12i} = 0, \\ d_{12i} = 0, \\ d_{11i} c_{13i} - c_{11i} d_{13i} = 0, \\ d_{13i} c_{22i} - c_{13i} d_{22i} = 0, \\ d_{22i} c_{23i} - c_{22i} d_{23i} = 0, \\ d_{23i} c_{33i} - c_{23i} d_{33i} = 0, \\ a_{12j} = 0, \\ b_{12j} = 0, \\ b_{11j} a_{13j} - a_{11j} b_{13j} = 0, \\ b_{13j} a_{22j} - a_{13j} b_{22j} = 0, \\ b_{22j} a_{23j} - a_{22j} b_{23j} = 0, \\ b_{23j} a_{33j} - a_{23j} b_{33j} = 0, \\ c_{11i} a_{22j} - a_{11j} c_{22i} = 0, \\ d_{11i} b_{22j} - b_{11j} d_{22i} = 0. \end{cases} \tag{49}$$

The following nonlinear cost function will be minimized to solve the previous nonlinear equations:

$$\min_{\omega_{i,j}} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \left(\Delta_i^2 + \Gamma_i^2 + \alpha_i^2 + \beta_i^2 + \gamma_i^2 + \theta_i^2 + \vartheta_j^2 \right. \\ \left. + \zeta_j^2 + \eta_j^2 + \varphi_j^2 + \rho_j^2 + \psi_j^2 + \chi_{ij}^2 + \delta_{ij}^2 \right), \tag{50}$$

where $\Delta_i = c_{12i}$, $\Gamma_i = d_{12i}$, $\alpha_i = d_{11i} c_{13i} - c_{11i} d_{13i}$,

$$\beta_i = d_{13i} c_{22i} - c_{13i} d_{22i},$$

$$\gamma_i = d_{22i} c_{23i} - c_{22i} d_{23i},$$

$$\theta_i = d_{23i} c_{33i} - c_{23i} d_{33i}, \qquad \vartheta_j = a_{12j},$$

$$\zeta_j = b_{12j}, \qquad \eta_j = b_{11j} a_{13j} - a_{11j} b_{13j},$$

$$\varphi_j = b_{13j} a_{22j} - a_{13j} b_{22j},$$

$$\rho_j = b_{22j} a_{23j} - a_{22j} b_{23j},$$

$$\psi_j = b_{23j} a_{33j} - a_{23j} b_{33j},$$

$$\chi_{ij} = c_{11i} a_{22j} - a_{11j} c_{22i},$$

$$\delta_{ij} = d_{11i} b_{22j} - b_{11j} d_{22i},$$

and n is the number of images.

The cost function (50) is minimized by using the algorithm in [23]. This algorithm requires an initialization step; therefore, the camera parameters are initialized as follows. Pixels are square; thus, $\varepsilon_i = \varepsilon_j = 1$, $s_i = s_j = 0$, the principal point is in the image center; therefore, $x_{0i} = y_{0i} = x_{0j} = y_{0j} = 256$ (because the images used are of size 512×512), and focal lengths f_i and f_j are obtained by solving the system of equations (49) after the replacement of the parameters $x_{0i}, y_{0i}, \varepsilon_i, s_i, x_{0j}, y_{0j}, \varepsilon_j, s_j$ in this system.

5 Experiment results

5.1 Simulations

In this section, a sequence of twelve 512×512 images of a checkerboard pattern is simulated to test the performance and robustness of the present approach. After the detection of control points by the Susan algorithm [24], the matches between each pair of images are determined by the correlation function NCC [28]. The pattern is projected in images taken from different views with Gaussian noise of standard deviation σ, which is added to all image pixels. The projection of the pattern points in the image planes allows formulating the linear equations, and the solution of these equations gives the projection matrices. The determination of a relationship between the two points of the pattern and their projections in the images i and j and the relationships between images of the absolute conic can define a nonlinear cost function. The minimization of this function by the Levenberg–Marquart algorithm [23] allows estimating the intrinsic parameters of the cameras used.

Figure 3, given below, shows the relative errors on x_0 according to the number of images by the present method, the methods of Zhang [21], Wang [31], Jiang [16], and Triggs [1].

Figure 4, presented below, shows the relative errors on y_0 according to the number of images by the five methods.

Figure 5, presented below, shows the relative errors on f according to the number of images by the five methods.

Figure 6, presented below, shows the relative errors of ε according to the number of images by the five methods.

Figure 7, presented below, shows, in the present method, the relative errors on the coordinates of the principal point, focal length, scale factor, and skew factor, according to the Gaussian noises\sigma (standard deviation).
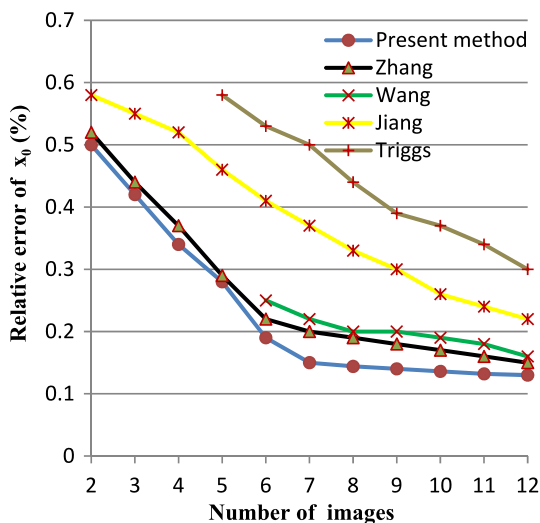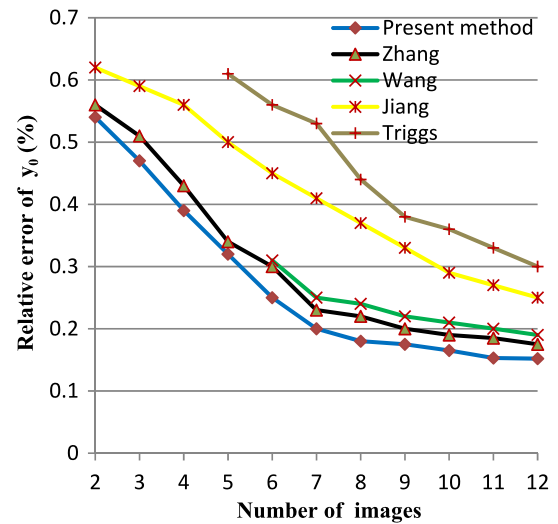

Fig. 4 The relative errors of y_0 according to the number of images

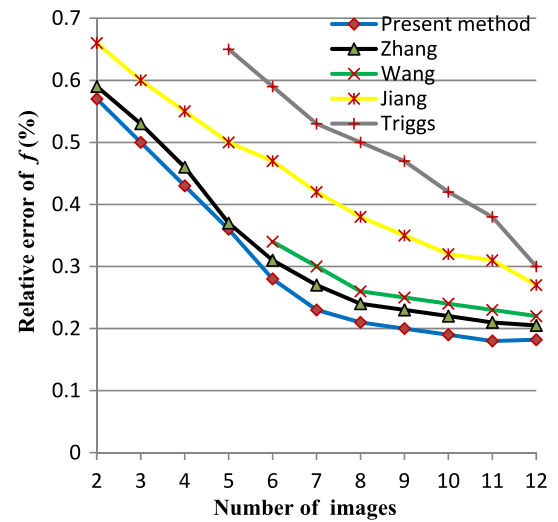**Fig. 5** The relative errors of f according to the number of images

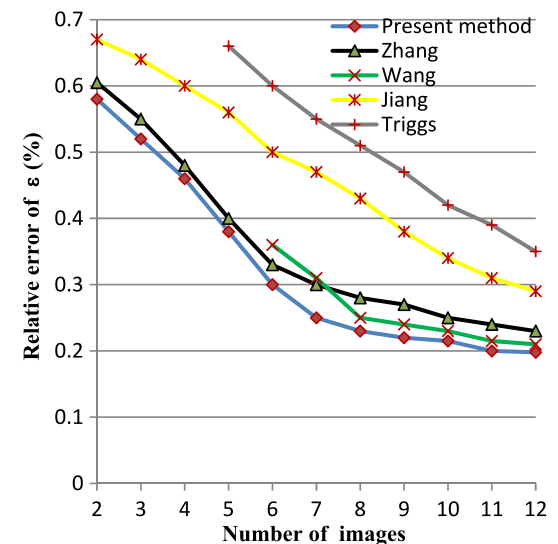**Fig. 3** The relative errors of x_0 according to the number of images

Fig. 6 The relative errors of ε according to the number of images

Figure 8, presented below, shows the execution time according to the number of images by the five methods.

A simple reading of Figs. 3, 4, 5, and 6 shows that the relative errors of the coordinates of the principal point, focal length, and scale factor calculated by the present method decrease almost linearly if the number of images is between two and six; they decrease slowly if the number of images is between six and eight; and they become almost stable if the number of images exceeds eight. Therefore, if the number of images increases, the relative errors decrease, which means that this method gives higher precision. But the mathematical calculations become somewhat complex, and, in practice, the execution time of the program becomes high because the increase in image number implies the increase of parameter number to be estimated, which implies the increase of the equation number, and which finally shows the increase of execution time. It is clearly seen in Fig. 8 because this figure shows that when the number of images increases,
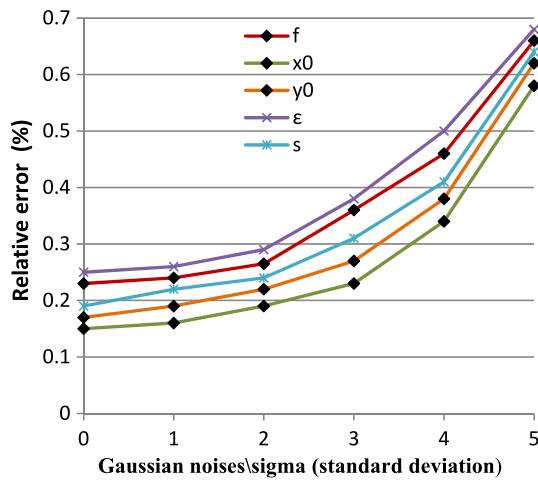

Fig. 7 The relative errors of x_0, y_0, f, ε, and s according to Gaussian noises\sigma (standard deviation)
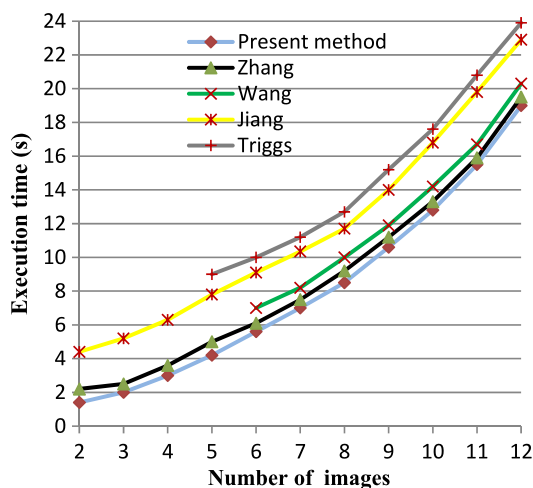

Fig. 8 The execution time according to the number of images

Fig. 9 Two images of an unknown 3D scene

the execution time of the different methods increases, and finally it shows the effect of the use of a large image number.

A Gaussian noise with a standard deviation of 1 pixel (so that $\sigma \leq 5$ pixels) is added to all image pixels to show the performance and robustness of the present method. After calculating errors of the coordinates of the principal point, focal length, scale factor, skew factor, and their presentations (Fig. 7) according to noise, it may be concluded that the errors increase slowly if $0 < \sigma \leq 3$ pixels, but these errors increase quickly if σ exceeds three pixels. This shows that noise has little effect on this method if $\sigma \leq 3$ pixels; on the other hand, it has more influence if $\sigma > 3$ pixels.

In order to show the performance and robustness of the algorithms presented in this paper, the simulation results are compared to those obtained by several efficient methods of Zhang [21], Wang [31], Jiang [16], and Triggs [1]. Af-

Fig. 10 The control points are shown in red in the two images

Fig. 11 The matches are shown and numbered in red in the two images

ter reading the results obtained by these different methods (Figs. 3, 4, 5, and 6), it appears that the relative errors corresponding to the intrinsic parameters estimated by the present method are lower than those obtained by Triggs and Jiang, and they are closer to those calculated by Zhang and Wang. This shows that the present approach gives satisfactory results compared to the method presented by Triggs and Jiang, and it gives similar results to those obtained by the methods of Zhang and Wang. Furthermore, the method of Wang requires at least six images, and the method of Triggs requires at least five images to calibrate the camera. However, two

images are sufficient to calibrate the cameras used by the present method.

5.2 Real data

This section deals with the experiment results of the different algorithms (Susan, *NCC*, Ransac, Levenberg–Marquardt, etc.) implemented by the Java object-oriented programming language. These experiments were conducted on eight 512×512 images of an unknown three-dimensional scene, taken by a CCD camera having varying intrinsic parameters. Two images (among eight) are shown in Fig. 9. The control points and the matches (obtained by the Ransac algorithm)

Table 1 The results of the intrinsic camera parameters estimated by the three methods

		f	ε	s	x_0	y_0
The present method	Image 1	1170	0.93	0.03	257	261
	Image 2	1167	0.95	0.02	262	258
	Image 3	1176	0.94	0,01	260	263
	Image 4	1174	0.91	0.02	259	260
Wang	Image 1	1172	0.91	0.04	259	262
	Image 2	1165	0.94	0.03	263	260
	Image 3	1174	0.95	0.02	262	264
	Image 4	1175	0.92	0.03	260	261
Jiang	Image 1	1152	1	0	250	259
	Image 2	1144	0.78	0.16	245	247
	Image 3	1145	0.75	0.19	240	247
	Image 4	1151	0.72	0.18	248	249

between these two images are shown respectively in Figs. 10 and 11, and the intrinsic parameters estimated by three methods (the present method, Wang's [31], and Jiang's [16]) are shown in Table 1 below.

Figure 10 bellow shows the control points obtained by the Susan algorithm in the two images. 691 control points were detected in the first image, and 694 control points were detected in the second image.

The control points detected by de Susan algorithm are matched between two images by the correlation function *NCC*. The result obtained contains false matches. To eliminate them, the authors of this paper regularized all the matches by the Ransac algorithm. Figure 11 bellow shows the 132 matches obtained by this algorithm.

After comparing the results on the synthetic data, the results of the present approach on real data (the two images shown in Fig. 8 and the six other images) are compared to those obtained by Wang [31] and Jiang [16] on the same data. The reading and the analysis of the intrinsic camera parameters presented in Table 1 and those estimated in the four other images show that the results of the present approach are similar to those obtained by Wang, and they are a little different from those obtained by Jiang. Furthermore, the present approach is also compared with two other methods of Zhang [21] and Triggs [1]. The results of the present approach are almost identical to those obtained by Zhang, and they are a little different from those obtained by Triggs. Therefore, this approach provides a robust performance, and it is very close to the other well-established methods. In addition, this method has several advantages: the use of any camera, two images are sufficient to estimate the intrinsic camera parameters, and the use of an unknown 3D scene.

6 Conclusions

In this paper, a robust method of camera self-calibration by an unknown three-dimensional scene is presented. This new method is based on the determination of a relationship between two points in the 3D scene and their projections in the planes of the images i and j and between the relationships between the images of the absolute conic. These relationships give a nonlinear cost function, and the minimization of this function provides the intrinsic parameters of the cameras used. The positive points that characterize this approach are diverse. Among these, there are the use of any camera (with varying intrinsic parameters), a 3D scene, and sufficiency of two points of the scene and two images to achieve the camera's self-calibration procedure. The robustness of this method in terms of simplicity, accuracy, stability, and convergence is shown by the results of the experiments and the simulations conducted.

References

1. Triggs, B.: Autocalibration from planar scenes. In: Proceedings of the 5th European Conference on Computer Vision, pp. 89–105 (1998)
2. Sturm, P.: A case against Kruppa's equations for camera self-calibration. IEEE Trans. Pattern Anal. Mach. Intell. **22**, 1199–1204 (2000)
3. Saaidi, A., Halli, A., Tairi, H., Satori, K.: Self-calibration using a particular motion of camera. WSEAS Tran. Comput. Res. **3**(5), 295–299 (2008)
4. Zhao, Y., Lv, X.D.: An approach for camera self-calibration using vanishing-line. Inf. Technol. J. **112**, 276–282 (2012)
5. Zhang, W.: A simple method for 3D reconstruction from two views. In: GVIP 05 Conference (2005)
6. Saaidi, A., Halli, A., Tairi, H., Satori, K.: Self-calibration using a planar scene and parallelogram. In: ICGST-GVIP, pp. 41–47 (2009)
7. Liu, P., Shi, J., Zhou, J., Jiang, L.: Camera self-calibration using the geometric structure in real scenes. In: Proceedings of the Computer Graphics International, pp. 262–265 (2003)
8. Baataoui, A., El batteoui, I., Saaidi, A., Satori, K.: Camera self-calibration by an equilateral triangle. Int. J. Comput. Appl., 29–34 (2012)
9. Pollefeys, M., Koch, R., Gool, L.V.: Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters. In: Sixth International Conference on Computer Vision, pp. 90–95 (1998)
10. Manolis, I.A.L., 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)
11. Sturm, P.: Critical motion sequences for the self-calibration of cameras and stereo systems with variable focal length. Image Vis. Comput. **20**, 415–426 (2002)
12. Gurdjos, P., Sturm, P.: Methods and geometry for plane-based self-calibration. In: Proceedings CVPR, pp. 491–496 (2003)

13. Cao, X., Xiao, J., Foroosh, H., Shah, M.: Self-calibration from turn-table sequences in presence of zoom and focus. Comput. Vis. Image Underst. **103**(2), 227–237 (2006)
14. Jiang, Z., Liu, S.: 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 (2011)
15. Shang, Y., Yue, Z., Chen, M., Song, Q.: A new method of camera self-calibration based on relative lengths. Inf. Technol. J. **11**(3), 376–379 (2012)
16. Jiang, Z., Liu, S.: Self-calibration of varying internal camera parameters algorithm based on quasi-affine reconstruction. J. Comput. **7**(3), 774–778 (2012)
17. Zhao, Y., Hu, X., Lv, X., Wang, H.: Solving the camera intrinsic parameters with the positive tri-prism based on the circular points. Inf. Technol. J. **11**(7), 926–930 (2012)
18. Gao, Y., Radha, H.: A multistage camera self-calibration algorithm. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 537–540 (2004)
19. El akkad, N., Saaidi, A., Satori, K.: 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 (2012)
20. Hemayed, E.E.: A survey of camera self-calibration. In: Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, pp. 351–357 (2003)
21. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. **22**(11), 1330–1334 (2000)
22. Torr, P.H.S., Murray, D.W.: The development and comparison of robust methods for estimating the fundamental matrix. Int. J. Comput. Vis. **24**, 271–300 (1997)
23. Moré, J.J.: The Levenberg–Marquardt algorithm: implementation and theory. In: Watson, G.A. (ed.) Numerical Analysis. Lecture Notes in Mathematics, vol. 630, pp. 105–116. Springer, Berlin (1977)
24. Smith, S.M., Brady, J.M.: SUSAN—a new approach to low level image processing. Int. J. Comput. Vis. **22**(1), 45–78 (1997)
25. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**, 91–110 (2004)
26. Bouda, B., Masmoudi, Lh., Aboutajdine, D.: A new grey level corner detection based on electrostatic model. In: ICGST-GVIP, pp. 21–26 (2006)
27. Harris, C., Stephens, M.: A combined corner and edge detector. In: Proceedings of the Fourth Alvey Vision Conference, pp. 147–151 (1988)
28. Chambon, S., Crouzil, A.: Similarity measures for image matching despite occlusions in stereo vision. Pattern Recognit. **44**(9), 2063–2075 (2011)
29. 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)
30. 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)
31. Wang, G., Wu, Q.M.J.: Perspective 3-d Euclidean reconstruction with varying camera parameters. IEEE Trans. Circuits Syst. Video Technol. **19**(12), 1793–1803 (2009)

Nabil El Akkad received the bachelor's and master's degrees from SMBA-Fez University in 2006 and 2010 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, 3D reconstruction, genetic algorithms and real-time rendering.

Mostafa Merras received the bachelor's and master's degrees from SMBA-Fez University in 2006 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 calibration and self calibration by the genetic algorithms.
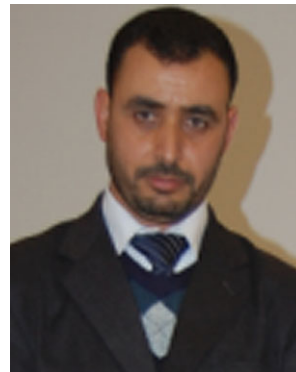
Abderrahim Saaidi received the PhD degree from SMBA-Fez University in 2010. He is currently a professor of computer science at SMBA-Taza University. He is a member of the LIIAN and LIMAO Laboratories. His research interests include camera self calibration, 3D reconstruction, genetic algorithms and real-time rendering.

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.