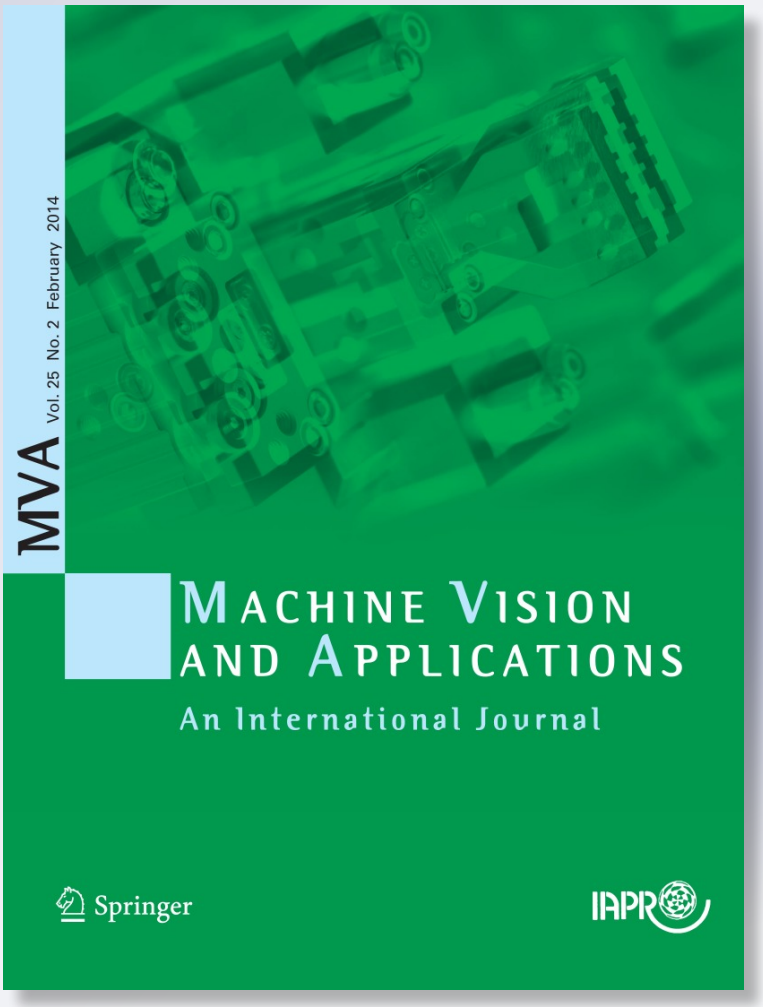
# *Structured light self-calibration with vanishing points*

**Radu Orghidan, Joaquim Salvi, Mihaela Gordan, Camelia Florea & Joan Batlle**

**Machine Vision and Applications**

ISSN 0932-8092
Volume 25
Number 2

Machine Vision and Applications (2014)
25:489-500
DOI 10.1007/s00138-013-0517-x




Springer

Machine Vision and Applications (2014) 25:489–500
DOI 10.1007/s00138-013-0517-x

ORIGINAL PAPER

# Structured light self-calibration with vanishing points

**Radu Orghidan · Joaquim Salvi · Mihaela Gordan · Camelia Florea · Joan Batlle**

Received: 4 November 2012 / Revised: 29 March 2013 / Accepted: 26 April 2013 / Published online: 15 June 2013


**Abstract** This paper introduces the vanishing points to self-calibrate a structured light system. The vanishing points permit to automatically remove the projector's keystone effect and then to self-calibrate the projector–camera system. The calibration object is a simple planar surface such as a white paper. Complex patterns and 3D calibrated objects are not required any more. The technique is compared to classic calibration and validated with experimental results.

**Keywords** Vanishing points · Self-calibration · Pattern projection · Structured light · 3D reconstruction

## 1 Introduction

Multiple view-points are needed in non-intrusive machine vision applications that deal with metric measurements of the scene such as automatic inspection, 3D reconstruction, object recognition, robot guidance for target or self-localization, reverse engineering, process control and others. The techniques used for solving the shape acquisition problem can be divided in two groups: *passive* and *active*. Passive techniques are used for obtaining depth in a scene without any physical interaction with the observed objects. The most common passive vision technique is the stereoscopy. A stereoscopic system, formed by two or more cameras, provides the necessary view-points and can be used for range estimation if two conditions are accomplished: (1) the correspondences between images are accurately determined and (2) the cameras are calibrated.

A crucial problem in stereovision is finding the corresponding features between several views of the same object. Geometric constraints, such as the ones introduced by the epipolar geometry [14], can be used for solving the correspondence problem yet at a high computational cost. Moreover, the resolution of stereovision systems is relatively low as it is dependent on the number of corresponding features which is ultimately determined by the texture of the scene objects or by the existence of singular points. On the other hand, active techniques have been introduced in order to enhance the capabilities of passive vision systems for finding correspondences. One of the most well-known techniques is the structured light(SL) [10,25] which is an active stereovision technique that alleviates the correspondence problem by illuminating the scene with a specially designed light pattern. The light reflected by the scene is measured and range estimation can be inferred. Concretely, the features projected by the pattern onto the scene objects are uniquely identified and the correspondences can be established at a high resolution. The first SL techniques were based on laser scanning [1]. In this case, the obtained images are easy to segment and these methods provide a high resolution. However, the speed and the mechanical components of such configurations introduce unacceptable restrictions for some applications such as the measuring of free moving objects. A more evolved model was introduced by the group of coded SL techniques. These techniques are projecting bi-dimensional light patterns on the scene by replacing one of the cameras by a light projector. Such techniques use temporary or spatially encoded patterns that allow a unique identification of the observed points.

The second condition that a stereoscopic system must fulfill is the calibration of its components. Many camera calibra-

R. Orghidan · M. Gordan (✉) · C. Florea · J. Batlle
Technical University of Cluj-Napoca, Cluj-Napoca, Romania
e-mail: Mihaela.Gordan@com.utcluj.ro

R. Orghidan
e-mail: radu.orghidan@com.utcluj.ro

J. Salvi
University of Girona, Girona, Spain
e-mail: joaquim.salvi@udg.edu

tion techniques [5,11,15,24,29,32] have been developed and excellent surveys [24,27] can be found in literature as it is an extensively studied field since the first machine vision applications appeared. A camera can be calibrated without the need of a calibration pattern by using the image of the points at infinity, i.e., the *vanishing points* (VPs) [4,6,12,19]. These particular points can be accurately obtained especially when dealing with scenes that meet the Manhattan world assumption [7]. The VPs extracted from mutually orthogonal directions can be used to calculate the intrinsic and the extrinsic camera parameters [14], thus, to calibrate the camera. Since the VPs can be used to derive geometric information of the camera, there have been developed many methods for their detection [2,18,23,28].

Active SL techniques use a projector which has to be calibrated, as well. The projector models can be divided into three groups [31]: the line model, the light-stripe model and the plane SL model. The last model is the most used in practice because it defines the projector as a reversed camera and, thus, camera calibration techniques can be applied. Even though the projector can be modeled using the perspective geometry, the approach to the projector calibration is different than in the case of the camera since it cannot offer an image of the scene. The lack of a scene image provided by the projector makes its calibration essentially different compared to the camera calibration. Thus, the existing camera calibration methods, relying on the image analysis and on establishing the correspondences between the image features and the scene structure, cannot be applied directly.

The calibration of a SL system is usually performed in two steps: first, the camera is calibrated; second, the calibration of the projector is performed using the calibrated camera. The calibration procedures mainly rely on the use of specially manufactured 2D or 3D calibration objects. For example, Kimura et al. [17] proposed a method for the calibration of a projector using a calibrated camera and several images of the calibration planes. Fernandez et al. [9] presented a similar method that requires, however, only a two-sided plane and is independent of the SL algorithm used for the reconstruction. In both methods the coordinates of the 3D points used for the calibration of the projector were calculated using a calibrated camera. The main drawback of this approach is that the errors introduced at the camera calibration stage are propagated to the projector's model; moreover, a previously calibrated camera is required. Martynov [20] proposed a projector calibration method that uses an uncalibrated camera as a sensor to determine the projector's image by calculating the camera–projector homography. The projector is eventually modeled as a reversed camera and is calibrated using a classical camera calibration approach. However, the method needs a series of iterations for determining the best fitted points for the projected image. Also, the method is based on the assumption that the points are accurately determined on the camera image which might be difficult if the projected pattern is distorted due to a possible misalignment between the projector and the screen.

It is well known that, when the projector is not orthogonal with respect to the target screen, its projection is affected by a trapezoidal shape distortion known as *the keystone effect*. Moreover, the camera's image is a perspective transformation of the scene. Thus, a new distortion is added when the scene containing the pattern is captured by the camera. Even though the pattern is distorted twice, a homography between the scene and the projector can be calculated [26] using the image provided by camera. Consequently, the transformation that compensates the keystone effect can be determined and applied to the projected pattern. The resulting projection is identical to the one produced by a projector whose optical axis is normal to the screen plane. We use the vanishing points of the corrected pattern for calibrating the projector, independently of the camera parameters.

This paper proposes a calibration method for a SL configuration and brings in the following contributions:

- The calibration object is simplified becoming a white planar surface with a known size and orthogonal edges.
- The error propagation is eliminated as the camera and the projector are calibrated independently.
- The method requires very low human interaction and does not depend on the SL algorithm used.
- The corresponding VPs of the camera and of the projector are used for the first time in SL systems.

The calibrated configuration is eventually used for shape acquisition using coded SL.

The remaining of the paper is structured as follows: Sect. 2 introduces the mathematical fundamentals for the calibration of the pinhole model using VPs and presents the calibration methodology for the camera and the projector of a SL system. The calibrated model can be used for 3D reconstruction as shown in Sect. 3. The accuracy of the reconstruction is studied through experimental results presented in Sect. 4. The paper ends with the conclusions that are detailed in Sect. 5.

## 2 Calibration methodology

The camera and the projector are projective devices and can be represented using the pinhole model. A projective transformation consists of a non-singular linear transformation of homogeneous coordinates from a point $P_w = [x_{wi}\ y_{wi}\ z_{wi}\ 1]^{\mathrm{T}}$ in the world coordinate system to a point of the image $P_i = [x_i\ y_i\ 1]^{\mathrm{T}}$ up to a scale factor $\lambda_i$, as shown in Eq. (1):

$$\lambda_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \end{bmatrix} \begin{bmatrix} x_{wi} \\ y_{wi} \\ z_{wi} \\ 1 \end{bmatrix} \tag{1}$$

The projection matrix, $A_{3\times4}$, can be decomposed and written as the product of the camera matrix and the transformation matrix from the world to the camera coordinate system:

$$^{C}\mathbf{A}_W = \mathbf{K}\,[\mathbf{R} \quad \mathbf{t}] \tag{2}$$

The pinhole model considers the skew coefficient between the two image axes, denoted by $\gamma$, the aspect ratios, denoted by $k_u$ and $k_v$ and the focal distance $f$. Thus, the camera matrix $K$ has the form:

$$\mathbf{K} = \begin{bmatrix} k_u f & \gamma & u_0 \\ 0 & k_v f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

The six extrinsic parameters of $[R \quad t]$ are the three rotations and three translations corresponding to each orthogonal axis. The pinhole model is calibrated when the intrinsic and extrinsic parameters are determined.

### 2.1 Pinhole model calibration using two VPs

Two VPs, determined from a projection screen having the orthogonal edges visible in the camera image, can be used for the calibration of a pinhole device using a method inspired by Guillou et al. [12]. For simplicity, this section explains the calibration for the camera case but the method is also applied for the projector after the image rectification, as it is demonstrated latter on.

We assume, without loss of generality, that the principal point is located at the center of the image, the skewness is equal to zero ($\gamma = 0$) and the aspect ratio is equal to one, i.e., $k_u = k_v = 1$. Hence, the intrinsic and extrinsic camera parameters can be obtained by means of geometric relations using only two vanishing points.

Let us consider two coordinate systems: the world coordinate system, centered at $O_w$ and having the orthogonal axes $(x_w, y_w, z_w)$ and the camera coordinate system, centered at $O_c$ with the axes $(x_c, y_c, z_c)$. Let the camera projection center be placed at $O_c$ and the center of the image, denoted by $O_i$, be the orthographic projection of $O_c$ on the image plane. Let the two vanishing points $V_1$ and $V_2$ be the vanishing points of two axes $x_w$ and $y_w$ of the world coordinate system, as shown in Fig. 1. The coordinates of the vanishing points in the image plane are $V_1 = (v_{1i}, v_{1j})$ and $V_2 = (v_{2i}, v_{2j})$. The projection of $O_i$ on the line $(V_1V_2)$ is denoted by $V_i$.

The principal point is located at the intersection of the optical axis with the image plane, thus, its coordinates $(u_0, v_0)$ are immediately obtained. Its position is crucial [13] for further calculations implied in the calibration process.

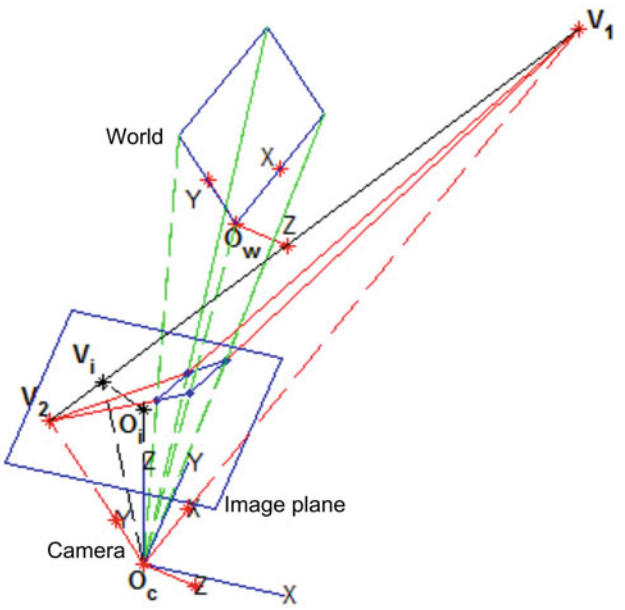



**Fig. 1** The focal distance and the orientation of the camera with respect to the world can be determined from the vanishing points

The focal distance $f$ can be calculated by considering that $O_c$ and $O_i$ are placed along the optical axis, as shown in Fig. 1, which means that:

$$f = \|O_cO_i\| = \sqrt{\|O_cV_i\|^2 - \|O_iV_i\|^2} \tag{4}$$

The distance $O_iV_i$ from the center of the image to the line at infinity, joining the two VPs, is calculated [12] as:

$$\|O_cV_i\| = \sqrt{\|V_1V_i\| \cdot \|V_iV_2\|} \tag{5}$$

The matrix $^{C}\mathbf{R}_W$ models the world-to-camera rotation. Since the two vanishing points $V_1$ and $V_2$ correspond to orthogonal axes and considering the fact that all parallel lines meet at the same VP, a new coordinate system can be determined. Therefore, the three orthogonal axes are also centered at $O_c$ and have the same orientation as the world's axes since $\mathbf{X}'_\mathbf{c} = \overrightarrow{O_cV_1}$, $\mathbf{Y}'_\mathbf{c} = \overrightarrow{O_cV_2}$ and $\mathbf{Z}'_\mathbf{c} = \mathbf{X}'_\mathbf{c} \times \mathbf{Y}'_\mathbf{c}$.

Hence, the rotation between the camera and the new coordinate systems is identical with the rotation between the camera and the world coordinate systems.

The vectors $\mathbf{X}'_\mathbf{c}$, $\mathbf{Y}'_\mathbf{c}$, $\mathbf{Z}'_\mathbf{c}$ are:

$$\begin{aligned} \mathbf{X}'_\mathbf{c} &= \frac{\overrightarrow{O_cV_1}}{\|\overrightarrow{O_cV_1}\|} = \left(\frac{v_{1i}}{\|\overrightarrow{O_cV_1}\|}, \frac{v_{1j}}{\|\overrightarrow{O_cV_1}\|}, \frac{f}{\|\overrightarrow{O_cV_1}\|}\right) \\ \mathbf{Y}'_\mathbf{c} &= \frac{\overrightarrow{O_cV_2}}{\|\overrightarrow{O_cV_2}\|} = \left(\frac{v_{2i}}{\|\overrightarrow{O_cV_2}\|}, \frac{v_{2j}}{\|\overrightarrow{O_cV_2}\|}, \frac{f}{\|\overrightarrow{O_cV_2}\|}\right) \\ \mathbf{Z}'_\mathbf{c} &= \mathbf{X}'_\mathbf{c} \times \mathbf{Y}'_\mathbf{c} \end{aligned} \tag{6}$$

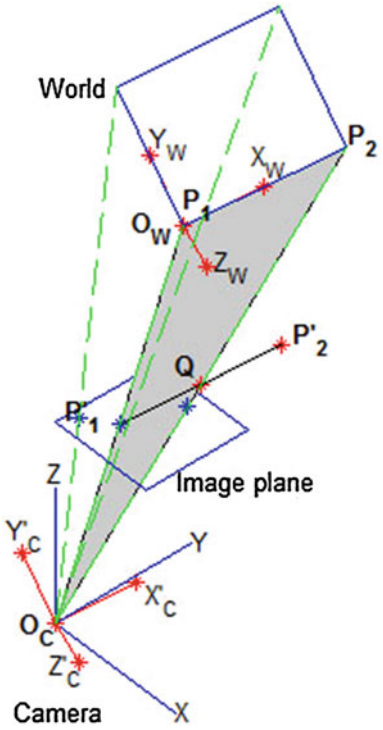



**Fig. 2** Projection of a scene segment through a pinhole camera model

Obtaining the rotation matrix ${}^{C}\mathbf{R}_{W}$ as:

$$ {}^{C}\mathbf{R}_{W} = \begin{bmatrix} \frac{v_{1i}}{\sqrt{v_{1i}^2+v_{1j}^2+f}} & \frac{v_{2i}}{\sqrt{v_{2i}^2+v_{2j}^2+f}} & z'_{cx} \\ \frac{v_{1j}}{\sqrt{v_{1i}^2+v_{1j}^2+f}} & \frac{v_{2j}}{\sqrt{v_{2i}^2+v_{2j}^2+f}} & z'_{cy} \\ \frac{f}{\sqrt{v_{1i}^2+v_{1j}^2+f}} & \frac{f}{\sqrt{v_{2i}^2+v_{2j}^2+f}} & z'_{cz} \end{bmatrix} \tag{7} $$

Finally, the translation vector $\mathbf{t}$ must be calculated.

A segment with a known size in the scene is used considering, without loss of generality, that one of its two ends is located in the origin of the world coordinate system. The segment is determined by the world points ${}^{W}\mathbf{P}_1 = [0, 0, 0]^T$ and ${}^{W}\mathbf{P}_2 = [x_{p2}, y_{p2}, z_{p2}]^T$, expressed in metric units, as shown in Fig. 2.

The segment can be aligned with its image in the system of coordinates of the camera using the rotation matrix ${}^{C}\mathbf{R}_{W}$:

$$ \left[ {}^{C}\mathbf{P}_{1m}\ {}^{C}\mathbf{P}_{2m} \right] = {}^{C}\mathbf{R}_{W} \left[ {}^{W}\mathbf{P}_1\ {}^{W}\mathbf{P}_2 \right] \tag{8} $$

The original segment is imaged by the camera through a projective transformation and two image points ${}^{I}\mathbf{P}_{1px}$ and ${}^{I}\mathbf{P}_{2px}$, represented in pixels, are obtained. In the pinhole model, the metric coordinates of any point in the image can be calculated by undoing the pixel transformation, the third coordinate being the focal distance:

$$ {}^{C}\mathbf{I}_{im} = {}^{I}\mathbf{P}_{ipx} - [u_0\ v_0]^T \tag{9} $$

We can now translate the segment on the image plane by setting its first point on its image ${}^{I}\mathbf{P}_{1m}$ and calculating the position of the second point. Thus, the translated segment is represented by the points ${}^{I}\mathbf{P}'_{1m}$ and ${}^{I}\mathbf{P}'_{2m}$:

$$ \begin{aligned} {}^{I}\mathbf{P}'_{1m} &= {}^{C}\mathbf{I}_{1m} \\ {}^{I}\mathbf{P}'_{2m} &= {}^{C}\mathbf{I}_{1m} + ({}^{C}\mathbf{P}_{2m} - {}^{C}\mathbf{P}_{1m}) \end{aligned} \tag{10} $$

The obtained segment is parallel to the original one thus forming two similar triangles $\triangle O_C P_1 P_2$ and $\triangle O_C P'_1 Q$, as shown in Fig. 2.

Taking advantage of the properties of similar triangles, we can write:

$$ \frac{\|O_c P_1\|}{\|O_c P'_1\|} = \frac{\|P_1 P_2\|}{\|P'_1 Q\|} \tag{11} $$

Therefore, the distance $D$ from the camera center to the world center can be calculated as:

$$ D = \|O_c P_1\| = \frac{\|O_c P'_1\| \cdot \|P_1 P_2\|}{\|P'_1 Q\|} \tag{12} $$

Hence, the translation vector is:

$$ \mathbf{t} = D \frac{O_c P'_1}{\|O_c P'_1\|} \tag{13} $$

Our method automatically determines the rotation about the $X$ and $Y$ axes of the calibration plane such that the VPs of the world's $XY$ axes are aligned with the camera $IJ$ axes. Then, the intrinsic and extrinsic camera parameters can be obtained by means of the geometric relations presented above. The VPs are invariant to translation, therefore, the camera translation is refined through a Levenberg–Marquardt error minimization algorithm with the initial solution given by Eq. (13).

## 2.2 Projected image rectification

The projector is calibrated by projecting the image of a checkerboard onto the same white plane used for the camera calibration. Generally, the optical axis of the projector is not perpendicular to the screen so the pattern is affected by a double distortion introduced by the keystone effect and by the camera perspective transformation. The pattern is rectified by eliminating the two distortions. Figure 3 presents the four points $P_i$ that bound the projected pattern. We consider a distortion-free projected pattern as a rectangle having the same VPs as the screen. Consequently, geometric relations can be used to calculate the positions of the points $P'_i$ that bound such a projection. Let $V_x$ and $V_y$ be the coordinates of the VPs formed in the camera image along the $X$ and $Y$ world axes, respectively. The keystone effect is removed by enforcing the co-linearity between the points $(P'_1, P'_2, V_x)$ and $(P'_4, P'_3, V_x)$ along the $X$ direction and

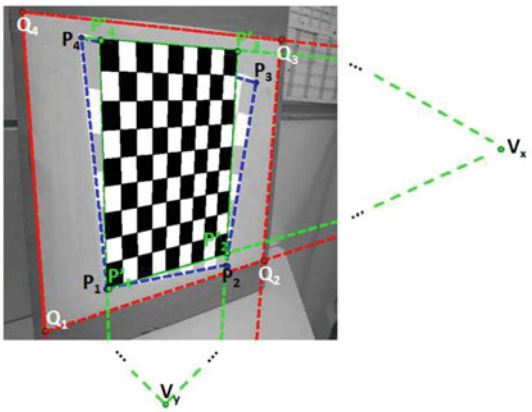



**Fig. 3** The keystone effect is removed when the VPs of the plane and of the projected image are coincident

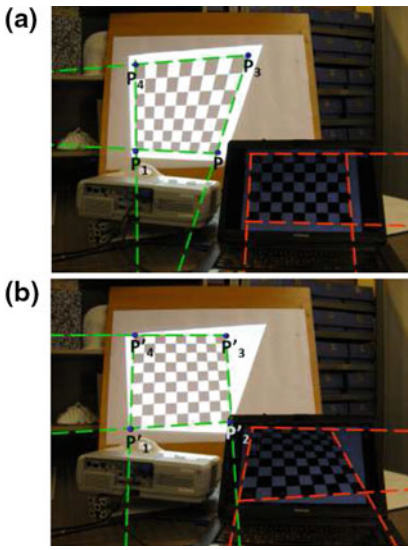



**Fig. 4** After removing the keystone effect, the VPs of the projector are identified. **a** Visible keystone effect affecting the projection. **b** The undistorted projection

between the points $(P'_1, P'_4, V_y)$ and $(P'_2, P'_3, V_y)$ along the $Y$ direction.

If the point $P_1 = P'_1$, then the positions of the other three points are calculated from the intersections of the following lines:

$$\begin{aligned} P'_2 &= (P'_1, V_x) \bigcap (P_2, V_y) \\ P'_4 &= (P_4, V_x) \bigcap (P_1, V_y) \\ P'_3 &= (P'_4, V_x) \bigcap (P'_2, V_y). \end{aligned} \tag{14}$$

The effect of the image rectification is illustrated in Fig. 4. The keystone effect, clearly visible in Fig. 4a, appears when projecting a checkerboard with orthogonal edges on a flat surface. After removing the distortion, the corrected image is projected and the distortion-free pattern appears as shown in Fig. 4b. Note that the VPs of the projector appear in the projector image.

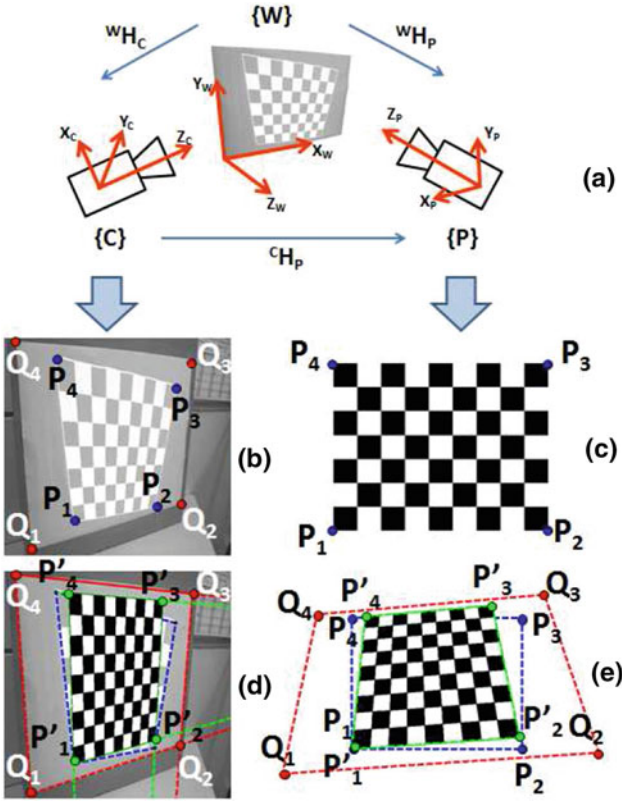



**Fig. 5** Flowchart of the calibration of a SL system

Let us consider that the camera and the projector reference systems, see Fig. 5a, are placed at $\{C\}$ and $\{P\}$, respectively. Both devices are randomly oriented and point towards a planar surface aligned with the $XY$ plane of the world coordinate system $\{W\}$. The homography ${}^W H_C$ between the world screen plane and the camera image plane, containing the points $S_i = [x_{si}\ y_{si}\ 1]^T$ and $Q_i = [u_i\ v_i\ 1]^T$, respectively, can be calculated from the system of Eq. (15) using at least four points.

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} h_{11}\ h_{12} \\ h_{21}\ h_{22} \end{bmatrix} \begin{bmatrix} x_{si} \\ y_{si} \end{bmatrix} \tag{15}$$

Similarly, the homography ${}^C H_P$ between the projector and the camera image is determined. Thus, estimating the homography between the screen and the projector becomes straight forward, as shown in relation (16):

$${}^W H_P = {}^W H_C \cdot {}^C H_P \tag{16}$$

Figure 5b shows the camera image containing the camera calibration plane, defined by the points $Q_i$ and the projected pattern, bounded by the points $P_i$. The projector's keystone effect is removed, independently of the camera's parameters, by applying a transformation to the original pattern, shown in Fig. 5c. After the transformation, the resulting projection has the same VPs as the target screen. Considering the constraints

imposed by the screen's VPs, the points $P_i'$ are calculated on the camera image, see Fig. 5d. Using the homography ${}^{P}H_C = inv({}^{C}H_P)$, the new points $P_i'$ on the projector are determined. Figure 5e shows the projector image containing also the estimated position of the points $Q_i$ and $P_i$. Thus, the pattern can be rectified by a homography derived for a given projector–screen configuration.

However, applying the homography directly will produce oversampled or subsampled images. New patterns can be designed when structured patterns, as the checkerboard, are projected. Using the homography ${}^{W}H_P$ to relate the edges of a rectified projection on the screen with the projector image, a new pattern can be generated as shown in the Fig. 5d, e. Besides, in the case of a non-structured pattern, high-quality images can be obtained by adding an interpolation method to the transformation.

The distortion of the projector's view, shown in Fig. 5e, strongly binds with the projector–camera–screen configuration. Thus, the VPs of the projector's distorted image can be used for calibrating the pinhole model of the projector.

Briefly, the calibration algorithm has the following steps:

1. Determine the VPs of the camera using the target screen,
2. Generate the rectified pattern,
3. Determine the projector's VPs,
4. Calibrate the camera and the projector using their respective VPs by using the method explained in Sect. 2.1.

## 3 Structured light for 3D reconstruction

SL techniques are used for overcoming the limitations of passive stereovision due to the lack of correspondences in the case of scenes with non-textured objects. In this work, we project a coded pattern built using a spatial neighborhood coding technique based on a De Bruijn sequence. We used the optimized De Bruijn pattern designed by Pages et al. [22] using 4 hue levels and 64 colored vertical slits separated by black bands. This technique permits the identification of the correspondences in one-shot and is suitable for the reconstruction of both static and moving objects. A De Bruijn sequence of order $m$ in conjunction with $n$ different symbols forms a single dimensional structure of length $n \cdot m$ containing unique instances of substrings of length $m$. The background is removed by using a threshold for the low values of the luminance.

In the existing SL configuration the decoding process uses horizontal scan lines. The intensity peaks resulting from the center of each stripe are located by applying a second derivative filter to the signal obtained from the scan line. The intensity differences are enhanced by the filter and the detection is obtained with sub-pixel accuracy. The segmentation is performed using a binary rule: the regions where the second derivative is <1 are set to 0 and the others are set to 1. The result of the segmentation is illustrated in Fig. 6. Finally, the observed stripes are matched with the projected pattern, operation known as *pattern decoding*. The color of a stripe can be precisely classified among the 4 levels of Hue composing the projected pattern. The key to the decoding strategy lies in identifying the hue value of each stripe in a given scan-line. The matching between the projected and the perceived stripes is solved using the colors of two neighbor stripes, i.e., with a window of size equal to 3 (sequence of order 3). A large number of correspondences can be detected by applying the process iteratively for all the image rows that contain the object. The ideal case occurs when all the stripes in the scan-line are correctly identified and the sequence is decoded. However, in reality, not all the stripes are visible or some of them are incorrectly labeled. Such classification errors produce outliers that eventually decrease the global accuracy of the reconstruction. The correspondence between the projected stripes and the detected ones is obtained through a dynamic programming algorithm [30], using the RGB components of the stripes.

The two transformation matrices, ${}^{C}\mathbf{A}_W = [a_{ij}]$ and ${}^{C}\mathbf{B}_W = [b_{ij}]$, obtained from the calibration of the camera and of the projector, respectively, together with the decoded points $\mathbf{Q}_{ij} = [q_i \; q_j]^T$ of the camera image and their corresponding stripes from the projected pattern $\mathbf{P}_k$ are used to calculate the location of the 3D points in the scene, ${}^{W}\mathbf{P}_i = [x_i \; y_i \; z_i]^T$, as expressed by the system of Eq. (17)

$$\begin{cases} {}^{I}\mathbf{Q}_{ipx} = {}^{C}\mathbf{A}_W \; {}^{W}\mathbf{P}_i \\ {}^{I}\mathbf{P}_{ipx} = {}^{C}\mathbf{B}_W \; {}^{W}\mathbf{P}_i \end{cases} \tag{17}$$

which can be rearranged, see (18), and solved for ${}^{W}\mathbf{P}_i$ using SVD.

$$\begin{bmatrix} a_{14} - a_{34} \cdot q_i \\ a_{24} - a_{34} \cdot q_j \\ b_{14} - b_{34} \cdot p_k \end{bmatrix} \begin{bmatrix} x_i \\ y_j \\ z_i \end{bmatrix} = \mathbf{M} \tag{18}$$

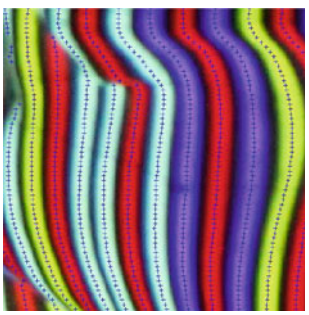


**Fig. 6** Detail of stripe segmentation using intensity peaks

where,

$$\mathbf{M} = \begin{bmatrix} a_{31} \cdot q_i - a_{11} & a_{32} \cdot q_i - a_{12} & a_{33} \cdot q_i - a_{13} \\ a_{31} \cdot q_j - a_{21} & a_{32} \cdot q_j - a_{22} & a_{33} \cdot q_j - a_{23} \\ a_{31} \cdot p_k - a_{11} & a_{32} \cdot p_k - a_{12} & a_{33} \cdot p_k - a_{13} \end{bmatrix} \quad (19)$$

## 4 Experimental results

The robustness to noise of the proposed calibration method was explored through a set of experiments using synthetic data. Working in a synthetic environment has the advantage that ground truth values are available and the deviation of the calculated parameters can be accurately determined. The synthetic setup used for our experiments contains a camera, a projector and a planar surface used as projection screen, as shown in Fig. 7. The world points were projected using a linear pinhole model. The camera image and the projector's rectified pattern are pictured in Fig. 8 that also contains the screen edges as a reference.

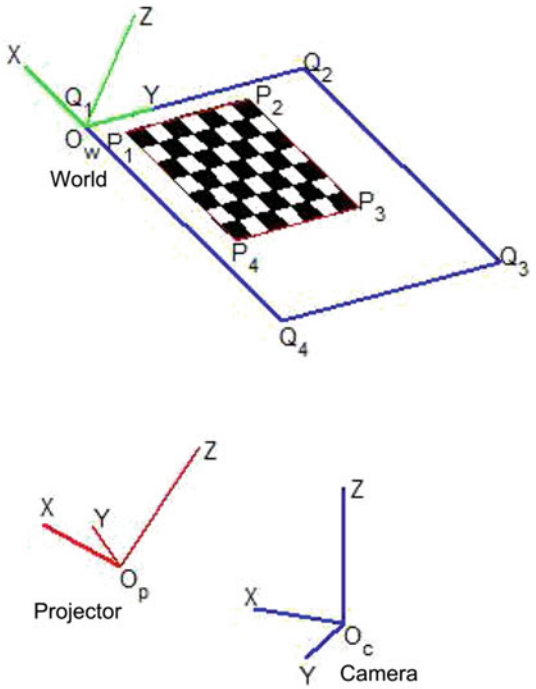



**Fig. 7** Experimental synthetic setup for calibration

In real situations, the noise is present mostly at the image level, hence, we added different levels of Gaussian noise on the 2D pixel coordinates of the synthetic images. The VPs corresponding to two orthogonal directions were extracted for each noise level, as shown in Fig. 9. Based on the position of the VPs the camera parameters were estimated. Fifty iterations have been applied at each noise level and the average values were considered for each parameter. The reprojection errors and the corresponding standard deviation for the calibrated camera and projector were calculated for different noise levels and their values, expressed in pixels, are presented in the leftmost image of Fig. 10. The righthand side of the same figure contains the 3D reconstruction error, expressed as a percentage of the size of a reference cube. The camera focal distance and its translation with respect to the world reference system were compared with the reference ones, as shown in Fig. 11. Since both the camera and the projector are calibrated independently using the same method, the relation noise-error is similar in both calibrations.

Further tests of the proposed calibration have been conducted by comparing its performance with an existing calibration toolbox (Figs.12, 13, 14), namely, the ProCam [8] implemented by Hurtos et al. [16] and later on improved by Fernandez [9]. The toolbox uses the well-known Bouguet's camera calibration toolbox [3] which implements Zhang's calibration [32].

Bouguets calibration method uses a planar checkerboard grid successively placed in front of the camera at several locations. Once the camera is calibrated using the corners of the printed checkerboard pattern, the position of the plane with respect to the camera can be recovered and the 3D locations

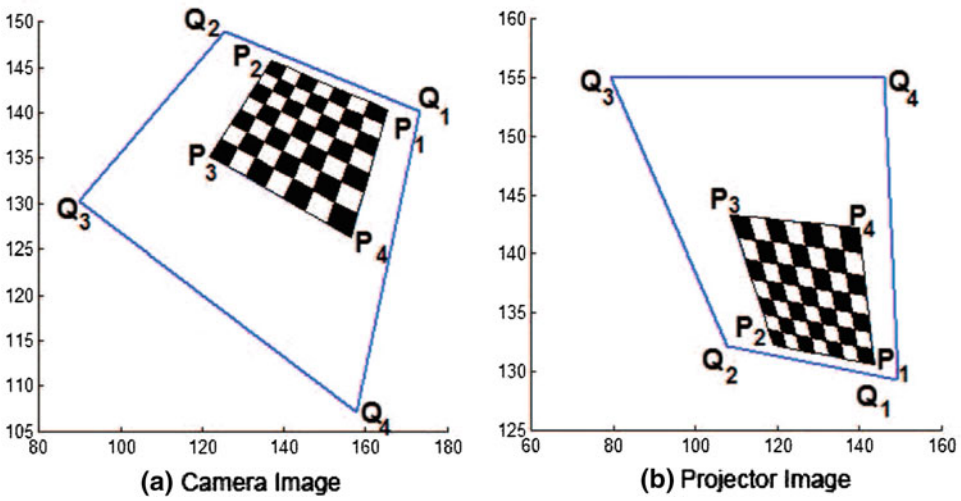



**Fig. 8** Synthetic camera and projector images

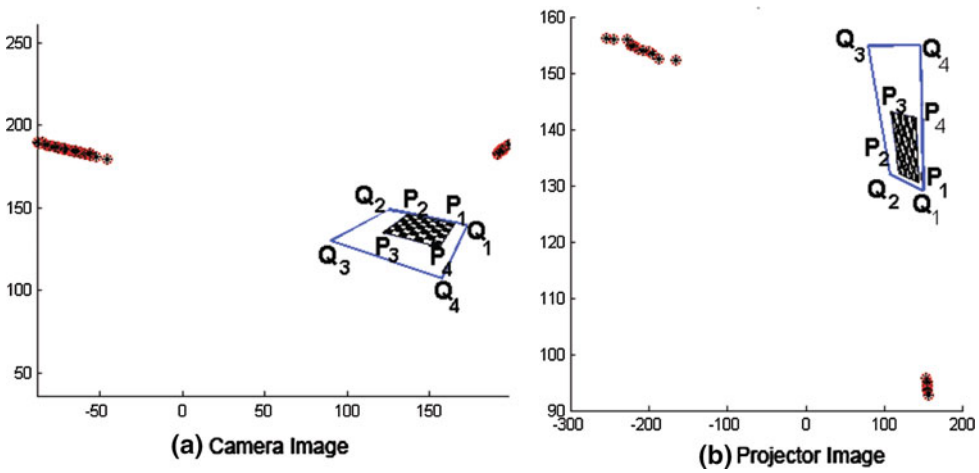



**Fig. 9** The VPs are deviated by the noise that affects the image points

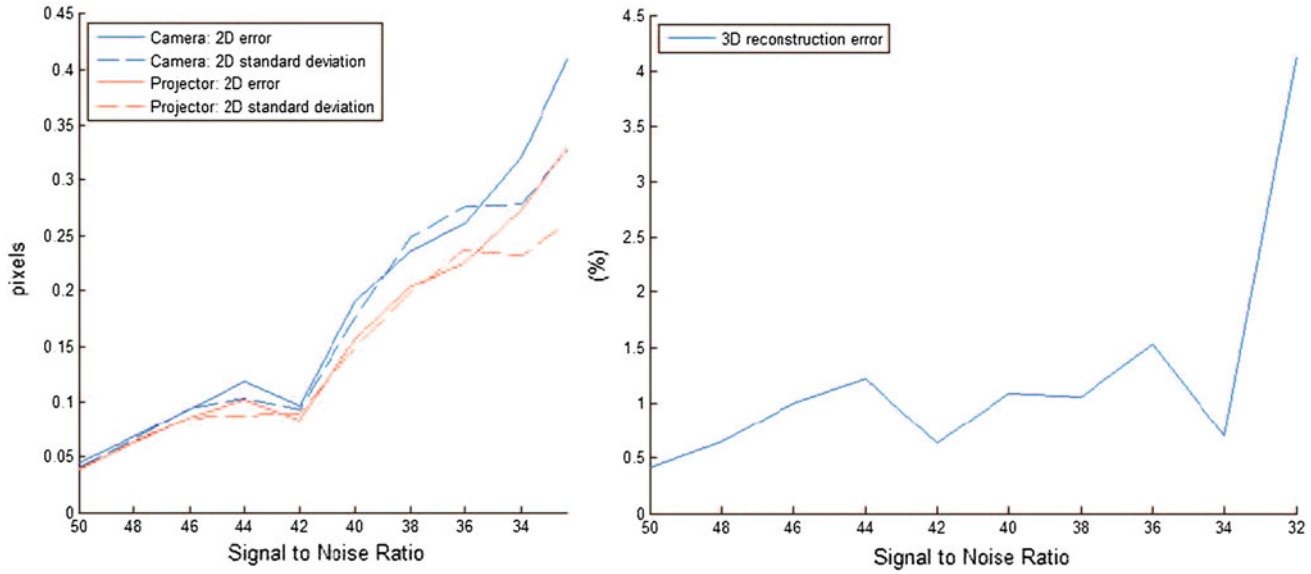



**Fig. 10** 2D reprojection error and depth estimation error using camera and projector calibration affected by different noise levels

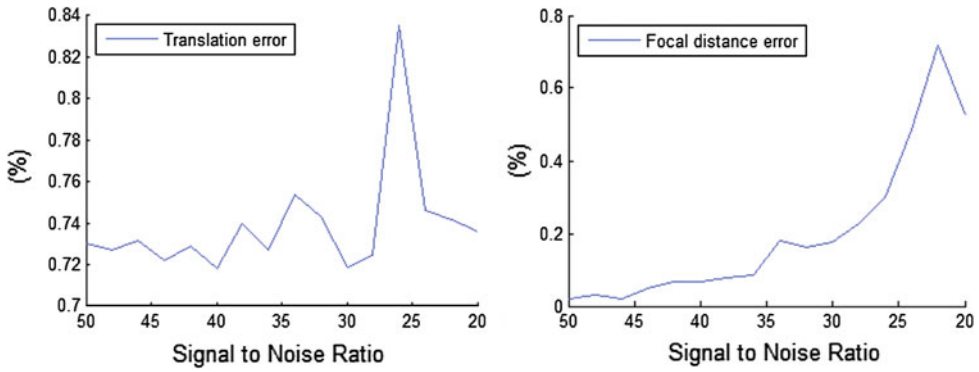



**Fig. 11** Translation and focal distance variation for the camera calibration affected by different noise levels

of the projected pattern's features can be calculated. Thus, the projector can be calibrated. The comparison tests were performed using a linear pinhole model, i.e., by setting the distortion parameters to zero in the Bouguet's method. The pixel reprojection error obtained using the ProCam toolbox were in the range of $\mu_{\text{cam}} \in [1.3697, 2.7041]$ pixels for the camera and $\mu_{\text{proj}} \in [3.2556, 3.7661]$ pixels for the projector. The same calibration planes were used and the reprojection error was calculated using our method obtaining an error in the range $\mu_{\text{cam}} \in [1.2085, 2.5498]$ pixels for the cam-

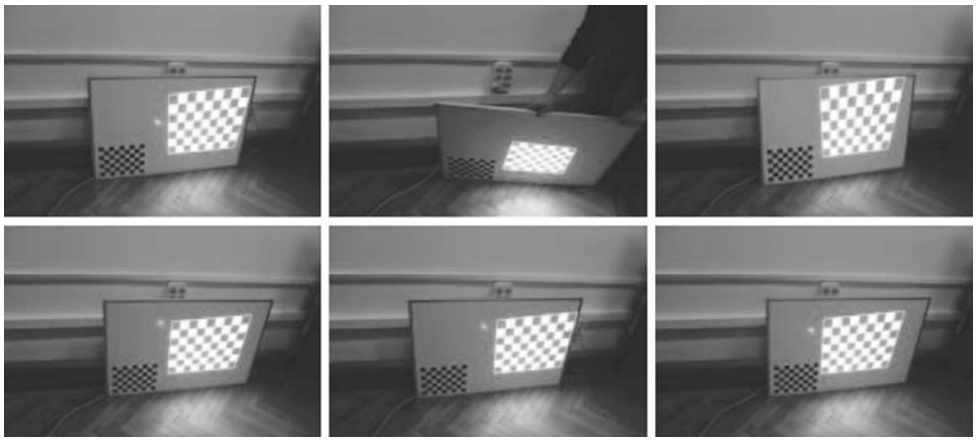


**Fig. 12** Several positions of the calibration plane with the projected pattern

**Table 1** 2D reprojection error in an experiment with real data for 10 calibration planes using the proposed method

| Camera | | | | Projector | | | |
|---|---|---|---|---|---|---|---|
| Pos. | Err. (px.) | Pos. | Err. (px.) | Pos. | Err. (px.) | Pos. | Err. (px.) |
| 1 | 2.0259 | 2 | 1.4648 | 1 | 1.1640 | 2 | 1.0825 |
| 3 | 2.5498 | 4 | 2.0387 | 3 | 0.4111 | 4 | 3.3566 |
| 5 | 1.4428 | 6 | 1.9825 | 5 | 0.1553 | 6 | 2.7940 |
| 7 | 1.2085 | 8 | 1.9976 | 7 | 1.4465 | 8 | 0.1090 |
| 9 | 2.5127 | 10 | 1.2706 | 9 | 0.2120 | 10 | 1.3327 |

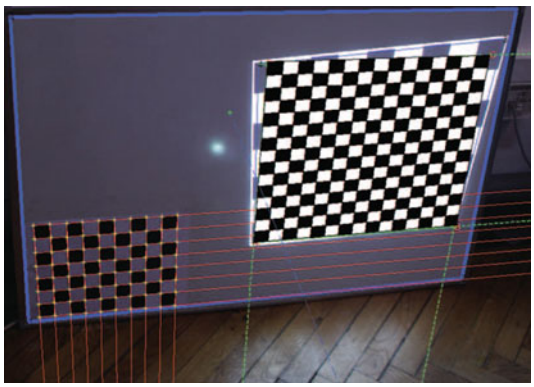


**Fig. 13** VPs detection using a printed pattern placed on the calibration screen

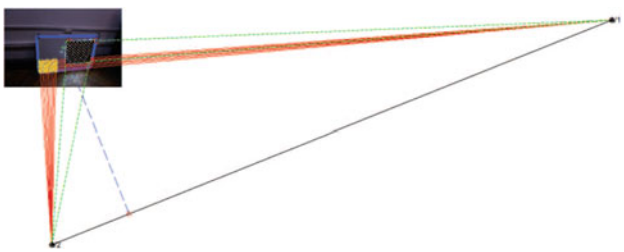


**Fig. 14** The VPs of the camera and of the rectified projector image are identical

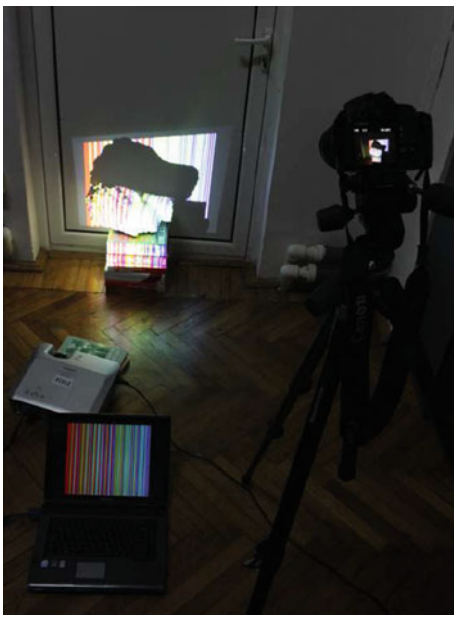


**Fig. 15** SL setup placed in an ad hoc configuration in front of an office door

era and $\mu_{\text{proj}} \in [0.1090, 3.3566]$ pixels for the projector, as presented in Table 1 for each plane.

The presented calibration was evaluated by performing one-shot reconstructions of objects placed in front of the SL system. The correspondence problem is solved using a color-encoded light pattern. The multi-slit pattern is a De Bruijn sequence built using 4 colors and 64 stripes, as detailed in Sect. 3.

The experimental setup was composed by a Canon EOS 500D camera, a Hitachi CP-X260 LCD projector and a laptop with an Intel Pentium Dual CPU at 2.16 GHz with 2 GB RAM. The components were placed in an ad hoc configuration, as shown in Fig. 15.

The accuracy of the reconstruction was estimated using a planar surface. The distance from the 3D points and an ideal plane, fitted to the cloud of points, was calculated and was

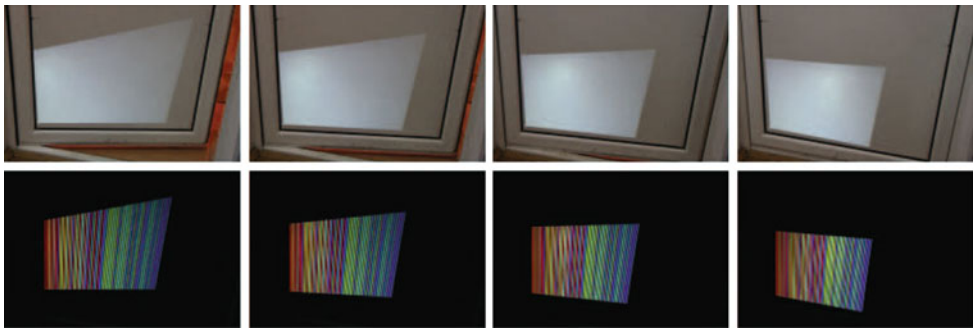


**Fig. 16** 3D reconstruction of a plane with different orientations using the calibrated SL system

**Table 2** 3D reconstruction of the planar surface of a door

| Pos. | Err. % | Std. % | VP cam.(u/v) 1.0*e*+005(px.) | VP proj.(u/v) 1.0*e*+005(px.) |
|---|---|---|---|---|
| 1 | 3.55 | 1.37 | [0.768, 0.031] | [0.032, 0.007] |
| | | | [0.013, 0.165] | [−0.025, −0.864] |
| 2 | 3.25 | 1.18 | [2.272, 0.144] | [0.040, 0.006] |
| | | | [0.013, 0.162] | [−0.091, −2.777] |
| 3 | 3.25 | 1.20 | [−1.896, −0.173] | [0.053, 0.006] |
| | | | [0.012, 0.161] | [−0.036, −1.371] |
| 4 | 3.20 | 1.21 | [−0.531, −0.069] | [0.082, 0.005] |
| | | | [0.013, 0.164] | [−0.039, −1.297] |

The columns contain the current position, the mean error and the standard deviation represented with respect to the distance from the SL system to the target plane and the values of the VPs for the camera and the projector, calculated for each position

considered to be the reconstruction error. For our experimental purposes, the SL system was placed at approximatively 1.5 m from an office door that was opened at different orientation angles, as shown in Fig. 16. Naturally, the projector's pattern deformation is proportional with the inclination of the ad hoc screen. The camera and projector VPs were calculated for each position and the system calibration was performed accordingly. The 3D plane reconstruction of the surface was obtained and the outliers were automatically eliminated by removing the Delaunay triangles with the area higher than a given threshold. A plane was fitted to the cloud of points and the distance from each point to the plane was calculated. The reconstruction error, the standard deviation and the values of the VPs for the camera and the projector for each position of the door are presented in Table 2. The average distance from the points to the planar surface was estimated as a percentage of the distance from the SL system to the plane of the door.

Finally, the qualitative accuracy of our calibrating method is depicted in Fig. 17 in which the head of a ceramic object is reconstructed. The shape of the face, the nose, the chin and even small details such as the eye ball can be noticed clearly.

## 5 Conclusion

This paper describes a new projector–camera self-calibration method introducing the vanishing points to remove the error propagation present in previous calibration methods and to simplify the calibrating pattern to an ubiquitous white plane.

The calibration setup is very simple since there is no need of specially tailored calibration objects. Thus, the process is easy to handle even by inexpert users. For example, as shown in this work, a normal office door or a white board on a wall can be used for calibrating the SL system.

Despite the calibration simplicity, the method provides accurate results. The tests made in the presence of the noise proves that the method performs well for acceptable noise levels and that a good 3D reconstruction can be obtained, see Fig. 11. The proposed method produced good results when compared with other available calibration methods. Our technique has been compared to ProCam toolbox. Procam needs several images to perform the calibration. Our method needs only one image for the calibration and pro-

**Fig. 17** Head reconstruction using the calibrated SL system: the camera image, the 3D cloud of points and the surface representation

vides good accuracy for each position of the calibration plane. Another advantage of the proposed method is that it uses the screen itself as a calibration pattern compared with other methods that require two patterns to be shown simultaneously on the same plane which introduces constraints on the 3D movement and lead to a non-uniform error distribution as the printed calibration pattern does not cover the whole surface.

The mean error obtained experimentally was around 3 % of the distance from the SL to the target screen with a low standard deviation, similar to other more complex linear calibrating techniques, despite the simplicity of the proposed technique. The measured accuracy proves that the method is suitable for computer vision applications that need good depth estimation and a fast self-calibration. The method works for single shot reconstructions which means that it can be used for dynamic scenes. Moreover, the calibrated model can be applied in a flexible manner for a variety of applications as it is not tied to a specific SL algorithm. A toolbox containing the proposed calibration method is available for public use [21].

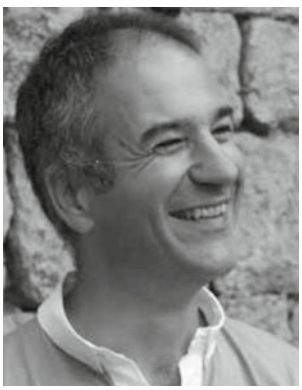
**Acknowledgments** This paper was supported by the project "Development and support of multidisciplinary postdoctoral programmes in major technical areas of national strategy of Research, Development, Innovation" 4D-POSTDOC, contract no. POSDRU/89/1.5/S/52603, project co-funded by the European Social Fund through Sectoral Operational Programme Human Resources Development 2007–2013.

## References

1. Agin, G.J., Binford, T.O.: Computer description of curved objects. In: International Joint Conferences on, Artificial Intelligence. Stanford, pp. 629–640 (1973)
2. Bosse, M., Rikoski, R., Leonard, J., Teller, S.: Vanishing points and three-dimensional lines from omni-directional video. Vis. Comput. **19**, 417–430 (2003)
3. Bouguet, J.: Camera calibration toolbox for matlab. Cambridge University Press, Cambridge (2004)
4. Caprile, B., Torre, V.: Using vanishing points for camera calibration. Int. J. Comput. Vis. **4**, 127–140 (1990)
5. Chen, X., Davis, J., Slusallek, P.: Wide area camera calibration using virtual calibration object. Proc. CVPR, pp. 2520–2527 (2000)
6. Cipolla, R., Drummond, T., Robertson, D.: Camera calibration from vanishing points in images of architectural scenes. In: BMVC99, pp. 382–391, University of Cambridge, Cambridge (1999)
7. Coughlan, J.M., Yuille, A.L.: Manhattan world: compass direction from a single image by bayesian inference. In: International Conference on Computer Vision (ICCV), pp. 941–947. MIT Press, Cambridge (1999)
8. Falcao, G., Hurtos, N., Massich, J., Fofi, D.: Projector-camera calibration toolbox, Tech. rep., University of Burgundy, Le2i UMR CNRS 6306-IUT Le Creusot http://code.google.com/p/procamcalib (2009). Accessed 12 Mar 2013
9. Fernandez, S., Salvi, J.: Planar-based camera-projector calibration. In: IEEE 7th International Symposium on Image and Signal Processing and Analysis, pp. 633–638. Dubrovnik (2011)
10. Forster, F.: A high-resolution and high accuracy real-time 3d sensor based on structured light. In: Proceeding of 3th International Symposium on 3D Data Processing, Visualization and Transmission, pp. 208–215. Padova, Italy (2006)
11. Furukawa, Y., Ponce, J.: Accurate camera calibration from multi-view stereo and bundle adjustment. Int. J. Comput. Vis. **84**, 257–268 (2009)
12. Guillou, E., Meneveaux, D., Maisel, E., Bouatouch, K.: Using vanishing points for camera calibration and coarse 3D reconstruction from a single image. Vis. Comput. **16**, 396–410 (2000)
13. Hartley, R.I., Kaucic, R.: Sensitivity of calibration to principal point position. In: Proceeding of Seventh European Conference on Computer Vision, pp. 433–446. Springer (2000)
14. Hartley, R.I., Zisserman, A.: Multiple view geometry in computer vision, 2nd edn. Cambridge University Press, Cambridge (2004)
15. Heikkila, J., Silven, O.: A four-step camera calibration procedure with implicit image correction, pp. 1106–1112, CVPR, USA (1997)
16. Hurtos, T., Falcao, G., Massich, J.: Plane-based calibration of a projector camera system. Master's thesis, VIBOT master, University of Girona, Girona (2008)
17. Kimura, M., Mochimary, M., Kanade, T.: Projector calibration using arbitrary planes and calibrated camera, pp. 1–2. CVPR, USA (2007)
18. Kogecka, J., Zhang, W.: Efficient computation of vanishing points. IEEE Intern. Conf. Robot. Autom. **1**, 223–228 (2002)
19. Li, B., Peng, K., Ying, X., Zha, H.: Simultaneous vanishing point detection and camera calibration from single images, pp. 151–160, ISVC' 10, USA (2010)
20. Martynov, I., Kämäräinen, J.K., Lensu, L.: Projector calibration by "inverse camera calibration". In: Scandinavian Conference on Image Analysis (SCIA) (2011)
21. Orghidan, R., Salvi, J.: Toolbox for the calibration of a structured light system. users.utcluj.ro/∼orghidan (2012)
22. Pages, J., Salvi, J.: A new optimised de bruijn coding strategy for structured light patterns. 17th International Conference on Pattern Recognition, vol. 4, pp. 284–287, Spain (2004)
23. Rother, C.: A new approach for vanishing point detection in architectural environments. J. Imag Vis. Comput. **20**(9–10), 647–656 (2002)
24. Salvi, J., Armangué, X., Batlle, J.: A comparative review of camera calibrating methods with accuracy evaluation. Pattern Recognit. **35**(7), 1617–1635 (2002)
25. Salvi, J., Fernandez, S., Pribanic, T., Llado, X.: A state of the art in structured light patterns for surface profilometry. Pattern Recognit. **43**(8), 2666–2680 (2010)
26. Sukthankar, R., Stockton, R.G., Mullin, M.D.: Smarter presentations: Exploiting homography in camera-projector systems. Proc. Intern. Conf. Comput. Vis. **1**, 247–253 (2001)
27. Sun, W., Cooperstock, J.: An empirical evaluation of factors influencing camera calibration accuracy using three publicly available techniques. Mach. Vis. Appl. **17**(1), 51–67 (2006)
28. Tardif, J.: Non-iterative approach for fast and accurate vanishing point detection. In: IEEE 12th International Conference on Computer Vision, pp. 1250–1257, Kyoto (2009)
29. Tsai, R.: A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. IEEE J. Robot Autom. **3**(4), 323–344 (1987)
30. Zhang, L., Curless, B., Seitz, S.M.: Rapid shape acquisition using color structured light and multi-pass dynamic programming. In: Proceedings of the 1st International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT), pp. 24–36. Padova (2001)
31. Zhang, X., Zhu, L.: Projector calibration from the camera image point of view. Opt. Eng. **48**, 117–208 (2009)
32. Zhang, Z.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. **22**(11), 1330–1334 (2000)

## Author Biographies

**Radu Orghidan** received his Ph.D. degree in Information Technologies in 2006 from the University of Girona, Spain, and the B.Sc. degree in Automation and Informatics Engineering in 2001 from the Technical University of Cluj-Napoca (TUCN), Romania. Currently, he is a postdoctoral researcher in the computer vision field at the TUCN. His research is focused on 3D perception using stereoscopic techniques, 3D interactive tools, image processing, multimedia technologies and eLearning.



**Joaquim Salvi** graduated in Computer Science at the Technical University of Catalonia in 1993, received the DEA (M.Sc.) in Computer Science in July 1996 and the Ph.D. in Industrial Engineering in 1998 both at the University of Girona, Spain. He is Professor of Computer Vision at the Computer Architecture and Technology Department and at the Computer Vision and Robotics Group, University of Girona; he was a visiting professor at the Ocean Systems Lab, Heriot-Watt University (UK). His current interests are in the field of computer vision and mobile robotics, focused on visual SLAM, structured light, stereovision, and camera calibration. Currently, he is the director of the Polytechnic School of the University of Girona.

**Mihaela Gordan** received the Ph.D. degree in Electronics and Telecommunications Engineering and the M.Sc. degree in Applied Electronics Engineering in 2004 and 1996, respectively, from the Technical University of Cluj-Napoca, Romania. Currently, she is an Associate Professor at the same university in the Department of Communication, Faculty of Electronics, Telecommunications and Information Technology. She is an expert in digital image processing and analysis, multimodal image analysis, fuzzy logic, fuzzy image processing and machine learning for image analysis and understanding.

**Camelia Florea** received the B.Sc. and Ph.D. degrees in Electronics and Telecommunications Engineering in 2005 and 2009, respectively, from the Technical University of Cluj-Napoca, Romania. Currently, she is an Assistant Professor at the same university in the Department of Communication, Faculty of Electronics, Telecommunications and Information Technology. Her research interests are: signal processing and analysis, images and video sequences compression, compressed domain data processing, multimedia technologies, artificial intelligence and fuzzy logic.

**Joan Batlle** Ms.C. in Physics by the Universitat Autonoma de Barcelona, Ph.D. in Computer Engineering by the Universitat Politècnica de Catalunya-Barcelona. Professor of Computer Science at the University of Girona, teaching matters concerning Computer Vision Systems, and advanced Technologies. He was Rector of the University of Girona from 2002 to 2005.