Proceedings of the Federated Conference on
Computer Science and Information Systems pp. 123–130

ISBN 978-83-60810-48-4

# Camera calibration using two or three vanishing points

Radu Orghidan*, Joaquim Salvi†, Mihaela Gordan* and Bogdan Orza*
*Center for Multimedia Technologies and Distance Education,
Technical University of Cluj-Napoca, Cluj-Napoca, Romania
Email: {radu.orghidan,mihaela.gordan,bogdan.orza}@com.utcluj.ro
†Computer Vision and Robotics Group,
Polytechnic School - University of Girona, Girona, Spain
Email:joaquim.salvi@udg.edu

***Abstract*—The perspective projection models the way a 3D scene is transformed into a 2D image, usually through a camera or an eye. In a projective transformation, parallel lines intersect in a point called *vanishing point*. This paper presents in detail two calibration methods that exploit the properties of vanishing points. The aim of the paper is to offer a practical tool for the choice of the appropriate calibration method depending on the application and on the initial conditions. The methods, using two respectively three vanishing points, are presented in detail and are compared. First, the two models are analyzed using synthetic data. Finally, each method is tested in a real configuration and the results show the quality of the calibration.**

## I. Introduction

CAMERAS need to be calibrated when they are used in applications that require object reconstruction or interaction with the world. Tasks such as 3D reconstruction, object inspection, scene mapping and target or self localization require metric measurement of the scene. Just capturing images is not enough.

Explicit camera calibration means that the calibration process ends up with a set of physical parameters, obtaining a detailed model, as close as possible to a full description of the real system. A comparative review of camera calibration methods with accuracy evaluation has been published by Salvi et al. [18]. A remarkable merit of this survey is that it standardizes notation which enables the easy comparison of well-known calibration methods such as Tsai [20], Hall [8] or Faugeras [9]. The proposed models have been later improved by Zhang [21], Chen [5] or Heikkila [14].

The most common camera model, the pinhole camera, generates the image through a projective transformation from a 3D Euclidean space onto the image plane. Assuming a perfect projection, the collinearity of the points is preserved. Thus, lines in the scene are projected as lines onto the image plane. An interesting property of the projective space is that parallel lines intersect in a point on the image, unlike the familiar case of the Euclidean space where parallel lines never cross. It is convenient to say that the point of intersection of parallel lines is placed at infinity and its projection onto the image plane is called *vanishing point*(VP). In this paper we will use the short notation *VPs* for vanishing points that belong to orthogonal directions.

Many methods [17][15][2][19] have been proposed for the accurate detection of the VPs. The properties of the VPs are directly related to the focal length and the rotation of the camera with respect to the world coordinate system. Caprile [4] and Beardsley [1] were among the first to use VPs for the estimation of the internal parameters of a camera. Later on, Hartley and Zisserman [12], Cipolla et al. [6] or He [13] used VPs with the aim of extracting the camera parameters. Two similar works [10][16], presented the method for finding the intrinsic parameters using the calibration sphere obtained from several images containing two VPs. The authors explained the relation between the calibration sphere and the image of the absolute conic, used for extracting the calibration matrix [12].

The projective transformation consists of a non-singular linear transformation of homogeneous coordinates:

$$\lambda_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix} \quad (1)$$

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

$$\mathbf{P} = \mathbf{K}\,[\mathbf{R} \;\; \mathbf{t}] \quad (2)$$

The general model of the pinhole camera considers the skew coefficient between the two image axes, denoted by $\gamma$, and the aspect ratios, or the scale factors, denoted by $\alpha_u$ and $\alpha_v$. Thus, the camera matrix $K$ has the form:

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.

$$\mathbf{K} = \begin{bmatrix} \alpha_u f & \gamma & u_0 \\ 0 & \alpha_v f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

However, a simplification is often used by taking the skewness to be zero ($\gamma = 0$) and the scale factor equal to one, i.e. $\alpha_u = \alpha_v = 1$.

The six extrinsic parameters, that form the rotation and the translation matrices, are the three rotations and three translations corresponding to each orthogonal axis. The camera is calibrated when the intrinsic and extrinsic parameters are determined.

In the following, we present two camera calibration approaches that take advantage of the properties of the vanishing points. The first approach, proposed by Guillou et al. [7], uses only two vanishing points. The second one was proposed by Cipolla et al. [6] and uses three vanishing points to determine the seven parameters of the camera model. The resulting models from both methods were originally used to build architectural scene models.

The remainder of this article is structured as follows. The two methods are presented in detail in Sections 2 and 3 respectively using similar mathematical notations. Section 4 shows experimental results using synthetic and real data for each method and includes a comparison in terms of the setup complexity, of the robustness to noise and from the point of view of the possible applications. Finally, Section 5 presents the conclusions.

## II. Camera calibration using 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 Figure 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. Its position is crucial [11] for further calculations implied in the calibration process. Assuming that the principal point is located at the center of the image and the aspect ratio is equal to one, i.e. $\alpha_u = \alpha_v = f$, the intrinsic and extrinsic camera parameters can be obtained by means of geometric relations [7] using only two vanishing points.

### A. Intrinsic parameters calculation

The image center is considered to be coincident with the principal point. Thus, its coordinates $(u_0, v_0)$ are immediately obtained.

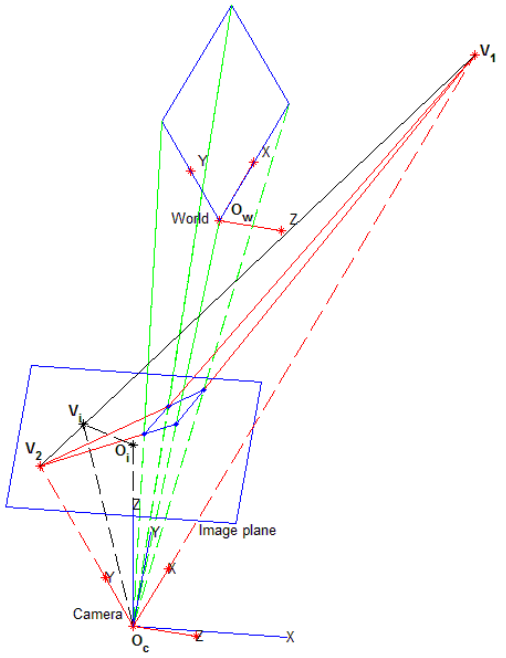


Fig. 1. The focal distance and the orientation of the camera with respect to the world can be determined from the vanishing points.

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

$$f = \|O_cO_i\| = \sqrt{\|O_cV_i\|^2 - \|O_iV_i\|^2} \quad (4)$$

Here, $O_iV_i$ is the distance from the center of the image to the horizon line determined by the two vanishing points and

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

### B. Extrinsic parameters calculation

The rotation between the world and the camera coordinate systems is expressed by the matrix $\mathbf{R}$. Taking into account that the two vanishing points $V_1$ and $V_2$ are in the direction of two orthogonal axes of the world reference system, centered at $O_w$, and that all parallel lines meet at a vanishing point, we can build a new coordinate system centered at $O_c$ and having the same orientation as the world system by considering the vectors $\mathbf{X'_c} = \overrightarrow{O_cV_1}$, $\mathbf{Y'_c} = \overrightarrow{O_cV_2}$ and $\mathbf{Z'_c} = \mathbf{X'_c} \times \mathbf{Y'_c}$.

Therefore, the rotation between the new coordinate system and the camera coordinate system is the same as the rotation between the world coordinate system and the camera coordinate system.

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

$$\begin{aligned} \mathbf{X'_c} &= \frac{\overrightarrow{O_cV_1}}{\|\overrightarrow{O_cV_1}\|} = \left(\frac{v_{1i}}{\|\overrightarrow{O_cV_1}\|}, \frac{v_{1j}}{\|\overrightarrow{O_cV_1}\|}, \frac{f}{\|\overrightarrow{O_cV_1}\|}\right) \\ \mathbf{Y'_c} &= \frac{\overrightarrow{O_cV_2}}{\|\overrightarrow{O_cV_2}\|} = \left(\frac{v_{2i}}{\|\overrightarrow{O_cV_2}\|}, \frac{v_{2j}}{\|\overrightarrow{O_cV_2}\|}, \frac{f}{\|\overrightarrow{O_cV_2}\|}\right) \\ \mathbf{Z'_c} &= \mathbf{X'_c} \times \mathbf{Y'_c} \end{aligned} \quad (5)$$

And the resulting rotation matrix $\mathbf{R}$ is:

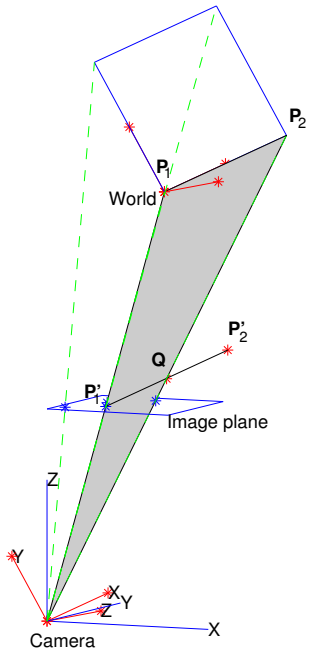


Fig. 2. Projection of a scene segment through a pinhole camera model.

$$\mathbf{R} = \begin{bmatrix} \frac{v_{1i}}{\sqrt{v_{1i}^2+v_{1j}^2+f}} & \frac{v_{2i}}{\sqrt{v_{2i}^2+v_{2j}^2+f}} & z'_{cx} \\ \frac{v_{1j}}{\sqrt{v_{1i}^2+v_{1j}^2+f}} & \frac{v_{2j}}{\sqrt{v_{2i}^2+v_{2j}^2+f}} & z'_{cy} \\ \frac{f}{\sqrt{v_{1i}^2+v_{1j}^2+f}} & \frac{f}{\sqrt{v_{2i}^2+v_{2j}^2+f}} & z'_{cz} \end{bmatrix} \tag{6}$$

The last step for the camera calibration is the calculation of the translation vector $\mathbf{t}$.

Let us consider a segment of known length in the scene, having the first of its two end points placed at the origin of the world. Without loss of generality, the center of the world can be chosen at any point in the scene. The segment is determined by the world points $\mathbf{P}_1 = [0,0,0]^T$ and $\mathbf{P}_2 = [x_{p2}, y_{p2}, z_{p2}]^T$, represented in metric units, as shown in Figure 2.

Since the rotation matrix $\mathbf{R}$ is known, we can align the segment with its image in the camera coordinate system:

$$\begin{bmatrix} \mathbf{P}_{1m}^C \\ \mathbf{P}_{2m}^C \end{bmatrix} = \mathbf{R} \begin{bmatrix} \mathbf{P_1} \\ \mathbf{P_2} \end{bmatrix} \tag{7}$$

The original segment is imaged by the camera through a projective transformation resulting in two image points $\mathbf{P}_{1px}^I$ and $\mathbf{P}_{2px}^I$, represented in pixels. 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:

$$\mathbf{I}_{im}^C = \mathbf{P}_{ipx}^I - [u_0 \ v_0]^T \tag{8}$$

We can now translate the segment on the image plane by setting its first point on its image $\mathbf{P}_{1m}^I$ and calculating the position of the second point. Thus, the translated segment is represented by the points $\mathbf{P}'_1$ and $\mathbf{P}'_2$:

$$\begin{aligned} \mathbf{P}'_1 &= \mathbf{I}_{1m}^C \\ \mathbf{P}'_2 &= \mathbf{I}_{1m}^C + (\mathbf{P}_{2m}^C - \mathbf{P}_{1m}^C) \end{aligned} \tag{9}$$

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 Figure 2.

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

$$\frac{\|O_c P_1\|}{\|O_c P'_1\|} = \frac{\|P_1 P_2\|}{\|P'_1 Q\|} \tag{10}$$

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

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

Hence, the translation vector is:

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

## III. Camera Calibration Using Three Vanishing Points

This approach uses three VPs determined from orthogonal directions in the scene. We assume that three vanishing points can be determined in the image from known patterns, such as two orthogonal checkered patterns. VPs detection methods in unstructured scenes are not discussed here as this topic falls beyond the subject of this paper.

### A. Intrinsic parameters calculation

In the current approach, we consider the principal point to be located at the center of the image, the skewness is zero ($\gamma = 0$) and the scale factor is equal to one, i.e. $\alpha_u = \alpha_v = f$. Thus, the camera matrix presented in equation (3) has a simplified form:

$$\mathbf{K} = \begin{bmatrix} f & 0 & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{12}$$

Determining the location of the principal point is straight forward as the image size is known. The only intrinsic parameter to be calculated is the focal distance.

Let three vanishing points corresponding to mutually orthogonal directions be projected onto the image plane, through the following homography:

$$\begin{bmatrix} \lambda_1 u_1 & \lambda_2 u_2 & \lambda_3 u_3 \\ \lambda_1 v_1 & \lambda_2 v_2 & \lambda_3 v_3 \\ \lambda_1 & \lambda_2 & \lambda_3 \end{bmatrix} = P_{3\times 4} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \tag{13}$$

The three vanishing points can be expressed, up to scale, as:

$$\begin{array}{l} V_1 = \lambda_1 \left[ \begin{array}{ccc} u_1 & v_1 & 1 \end{array} \right]^T \\ V_2 = \lambda_2 \left[ \begin{array}{ccc} u_2 & v_2 & 1 \end{array} \right]^T \\ V_3 = \lambda_3 \left[ \begin{array}{ccc} u_3 & v_3 & 1 \end{array} \right]^T \end{array} \quad (14)$$

Considering the decomposition of the projection matrix, shown in equation (2), we can write:

$$\left[ \begin{array}{ccc} V_1 & V_2 & V_3 \end{array} \right] = \mathbf{K} \left[ \mathbf{R} | \mathbf{t} \right] \left[ \begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{array} \right]$$

Taking into account the effect of the multiplication of the homogeneous points at infinity with the translation vector, we obtain:

$$\left[ \begin{array}{ccc} V_1 & V_2 & V_3 \end{array} \right] = \mathbf{KR} \quad (15)$$

Using the camera matrix $\mathbf{K}$, as shown in expression (12), the rotation matrix $\mathbf{R}$ can be written as,

$$\mathbf{R} = \left[ \begin{array}{ccc} \frac{\lambda_1(u_1-u_0)}{f} & \frac{\lambda_2(u_2-u_0)}{f} & \frac{\lambda_3(u_3-u_0)}{f} \\ \frac{\lambda_1(v_1-v_0)}{f} & \frac{\lambda_2(v_2-v_0)}{f} & \frac{\lambda_3(v_3-v_0)}{f} \\ \lambda_1 & \lambda_2 & \lambda_3 \end{array} \right] \quad (16)$$

Using the orthogonality property of the rotation matrix and applying it to its first two columns, we obtain

$$\lambda_1 \lambda_2 (\frac{(V_1 - O_i)(V_2 - O_i)}{f^2} + 1) = 0 \quad (17)$$

Since $\lambda_i \neq 0$, the focal distance can be calculated as:

$$f = \sqrt{|(O_i - V_1)(V_2 - O_i)|} \quad (18)$$

### *B. Extrinsic parameters calculation*

The extrinsic parameters are part of the rotation matrix $\mathbf{R}$ and the translation vector $\mathbf{t}$. The rotation matrix, presented in equation (16), can be calculated if the scaling factors $\lambda_i$ are determined. In order to calculate them, equation (15) can be rearranged by separating the scaling factors $\lambda_i$ and using the multiplication of the homogeneous points at infinity with the translation vector:

$$\left[ \begin{array}{ccc} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ 1 & 1 & 1 \end{array} \right] \left[ \begin{array}{ccc} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{array} \right] = \mathbf{KR} \quad (19)$$

Multiplying equation (19) on both sides by $(\mathbf{KR})^T$, and considering the orthogonality constraint of the rotation matrix, we obtain,

$$\left[ \begin{array}{ccc} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ 1 & 1 & 1 \end{array} \right] \left[ \begin{array}{ccc} \lambda_1^2 & 0 & 0 \\ 0 & \lambda_2^2 & 0 \\ 0 & 0 & \lambda_3^2 \end{array} \right] \left[ \begin{array}{ccc} u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \\ u_3 & v_3 & 1 \end{array} \right] = \mathbf{K}\mathbf{K}^T \quad (20)$$

A matrix $\mathbf{Q}$ can be defined as,

$$\mathbf{Q} = \mathbf{K}\mathbf{K}^T = \left[ \begin{array}{ccc} f^2 + u_0^2 & u_0 v_0 & u_0 \\ u_0 v_0 & f^2 + v_0^2 & v_0 \\ u_0 & v_0 & 1 \end{array} \right] \quad (21)$$

The vector containing the scaling factors $\lambda_i$ can be separated by rearranging equations (20) and (21):

$$\left[ \begin{array}{ccc} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ u_1^2 & u_2^2 & u_3^2 \\ v_1^2 & v_2^2 & v_3^2 \\ u_1 v_1 & u_2 v_2 & u_3 v_3 \end{array} \right] \left[ \begin{array}{c} \lambda_1^2 \\ \lambda_2^2 \\ \lambda_3^2 \end{array} \right] = \left[ \begin{array}{c} u_0 \\ v_0 \\ f^2 + u_0^2 \\ f^2 + v_0^2 \\ u_0 v_0 \end{array} \right] \quad (22)$$

The scale factors can be calculated by singular value decomposition on the system of equations (22) and the rotation matrix can be determined.

Note that, if the scaling factors are known, the left hand side of equation (20) is determined and the intrinsic parameters can be directly calculated by identifying their values in equation (21).

When three vanishing points, obtained from mutual orthogonal directions in the scene, are available, an alternative method for calculating the coordinates of the principal point $(u_0, v_0)$ can be used by finding the orthocenter of the triangle formed by the vanishing points [12][6].

The translation vector $\mathbf{t}$ is the vector pointing from the camera origin to the world origin and is given by the last column of the projection matrix. The projection of the world coordinate system is obtained from equation (1), setting the values $x_i = 0, y_i = 0, z_i = 0$ for an randomly chosen origin point. The translation obtained from a single view without additional information of the scene will be up to scale, with $\lambda_i$ having an arbitrary value.

If additional information is available, such as the length of a segment or the coordinates of points in the scene, the translation vector can be accurately extracted.

Let $\mathbf{r}_i$ be the $i^{th}$ row of the rotation matrix and $\mathbf{P}_{wi} = (x_i, y_i, z_i, 1)^T$ be a point of the scene, projected onto the image plane.

$$\left[ \begin{array}{c} \lambda_i u_i \\ \lambda_i v_i \\ \lambda_i \end{array} \right] = \mathbf{K} \left[ \mathbf{R} | \mathbf{t} \right] \left[ \begin{array}{c} x_i \\ y_i \\ z_i \\ 1 \end{array} \right] \quad (23)$$

Then, the following system of equations is obtained:

$$\lambda_i u_i = \alpha_u r_1 P_{wi} + u_0 r_3 P_{wi} + \alpha_u t_1 + u_0 t_3 \quad (24a)$$
$$\lambda_i v_i = \alpha_v r_2 P_{wi} + v_0 r_3 P_{wi} + \alpha_v t_2 + v_0 t_3 \quad (24b)$$
$$\lambda_i = r_3 P_{wi} + t_3 \quad (24c)$$

which leads to,

$$r_3 P_w u_i + t_3 u_i = f r_1 P_w + u_0 r_3 P_w + f t_1 + u_0 t_3 \quad (25a)$$
$$r_3 P_w v_i + t_3 v_i = f r_2 P_w + v_0 r_3 P_w + f t_2 + v_0 t_3 \quad (25b)$$

That can be expressed in a compact form as,

$$\begin{bmatrix} f & 0 & u_0 - u_i \\ 0 & f & v_0 - v_i \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix} = \begin{bmatrix} r_3 P_w (u_i - u_0) - f(r_1 P_w) \\ r_3 P_w (v_i - v_0) - f(r_2 P_w) \end{bmatrix} \tag{26}$$

The components of the translation vector can be calculated by stacking equations (26) for several pairs of image and scene points and solving the resulting system using singular value decomposition. Finding three vanishing points requires at least six points, placed on three mutually orthogonal axes in the scene, which can also be used for the calculation of the translation vector.

## IV. Experimental results

A set of experiments have been conducted in order to explore the robustness to noise of the implemented methods. The advantage of working with a synthetic environment is that absolute ground truth values can be obtained. In real scenes the noise is usually present at the image level, therefore, Gaussian noise has been gradually added on the image and the camera was calibrated using the affected images.

Knowing the position of the VPs, the camera model can be estimated using the calibration methods presented previously. This step was done iteratively for increasing gaussian noise levels. For each level, 50 iterations have been worked out in order to obtain a result as close as possible to the typical behavior of each method.

The effects of the noise on the calibration model were measured by calculating the error on three outputs: the image, the intrinsic parameters and the extrinsic parameters.

The image error was calculated as the distance between the reference points and the re-projected ones using the resulting calibrated model. The intrinsic parameters, namely $\alpha_u$ and $\alpha_v$, and the extrinsic parameters, i.e. the rotation and the translation between the camera and the world reference systems, were compared with the reference ones.

### A. Synthetic camera calibration using two VPs

The synthetic setup for the calibration using two VPs is presented in Figure 3. The coplanar points, placed in the world reference system, were projected using the ideal camera model. From the resulting image, the two VPs are extracted as shown in Figure 4. When the noise level increases, the VPs start moving from the original position. Fig. 5 shows that $V_1$ is more affected by the noise than $V_2$ because it is located at a larger distance from the image center.

### B. Synthetic camera calibration using three VPs

Camera calibration with three VPs has been analyzed using a setup formed by a virtual camera pointing to a cloud of 3D points, as shown in Figure 6. The 3D points belong to two orthogonal planes in the world reference system and were imaged by the modeled camera. The three VPs are extracted from the image as shown in Figure 7. The position of the image center is also illustrated as the orthocenter of the triangle

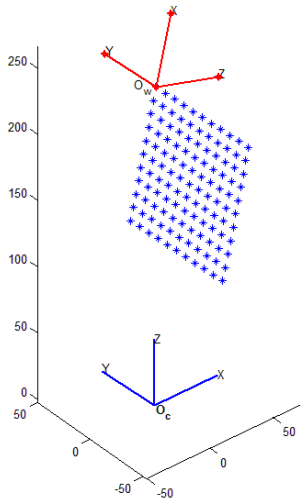


Fig. 3. Experimental setup for camera calibration using two VPs

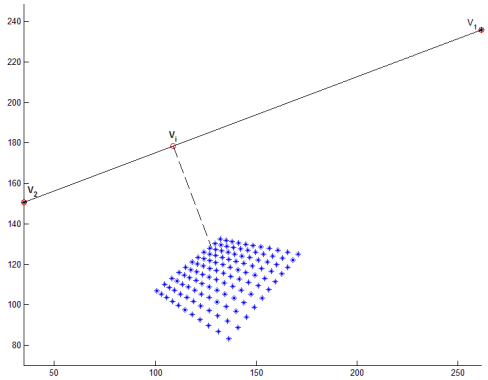


Fig. 4. VPs from two orthogonal directions.

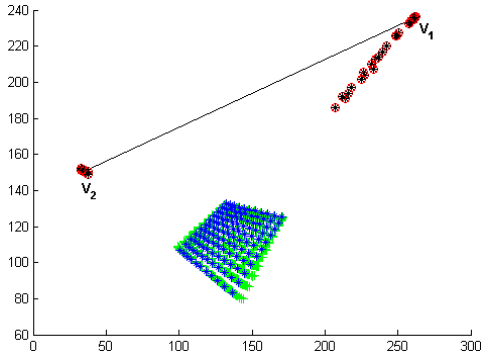


Fig. 5. VPs deviation due to the noise.

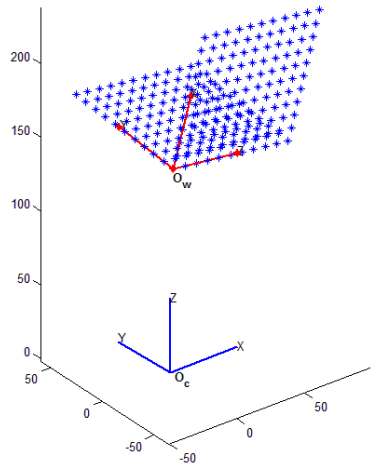


Fig. 6. Experimental setup for synthetic camera calibration using three VPs

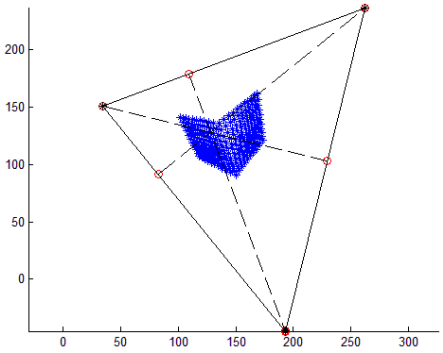


Fig. 7. VPs from three orthogonal directions and the image center.

formed by the three VPs. The noise level affects the VPs because the equations of the parallel lines change and so does their crossing point. When dealing with three VPs, their positions change proportionally with the distance from the image center, as shown in Figure 8.

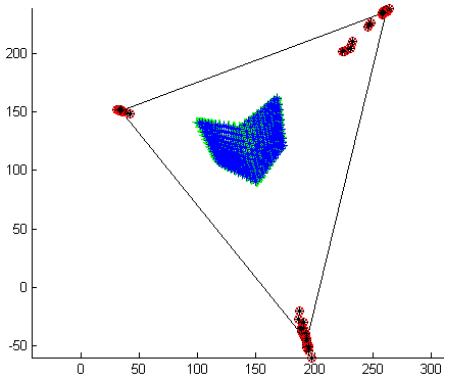


Fig. 8. VPs deviation due to the noise.

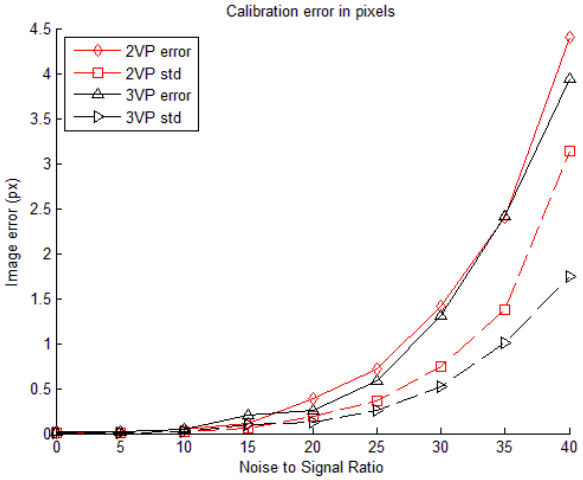


Fig. 9. Projection error using the camera model calibrated from the noisy images.

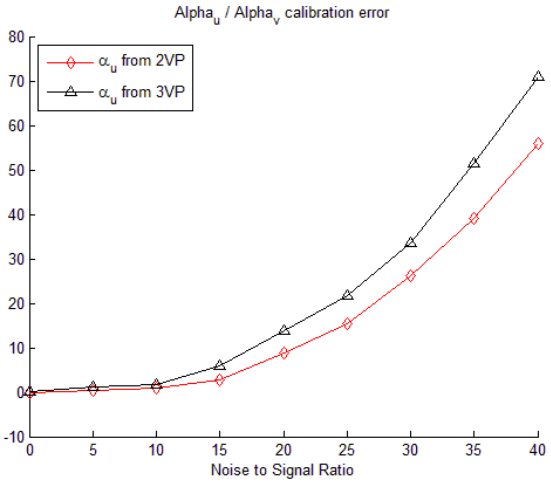


Fig. 10. Intrinsic parameters estimation error using the camera model calibrated from the noisy images.

### C. Noise robustness analysis

The outputs of the calibration process have been evaluated for both calibration methods and the evolution of the errors, as a function of the noise, was compared. Fig. 9 shows the projection error obtained by using the two camera models calibrated from the noisy images. Fig. 10 illustrates the evolution of the intrinsic parameters. The extrinsic parameters are also increasingly deviating from the ideal values, as shown in Figure 11. As expected, in all cases the error increases as the noise level grows. Both calibration methods have a similar behavior up to a certain level of noise. However, the method using two VPs shows better performance for a high amount of noise. This is due to the fact that the calibration plane has been previously chosen as the best initial solution among the 12 possible configurations.

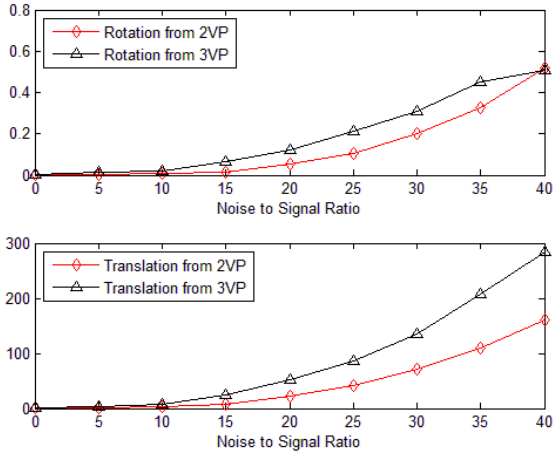


Fig. 11. Extrinsic parameters estimation error using the camera model calibrated from the noisy images.

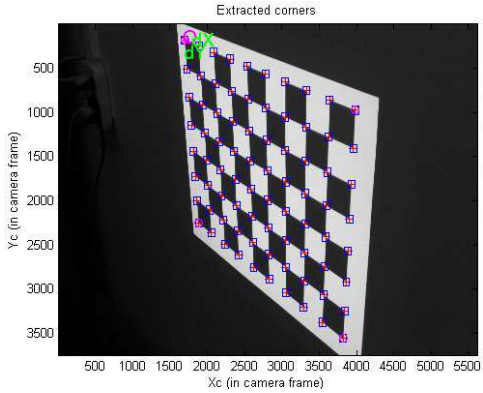


Fig. 12. Grid points extraction from the calibration pattern.

### D. Real camera calibration using two VPs

The 2 VPs calibration has been tested using a real camera. The position of the VPs was calculated from the images in three steps. First, the points of the pattern have been detected using a function of the camera calibration toolbox of Bouguet [3], see Figure 12. Then, parallel lines were fitted to the points on the direction of the VP. At last, the intersection of all the lines was calculated by solving the overdetermined system of line equations using singular value decomposition and obtaining their optimal crossing point as shown in Figure 13.

The camera was calibrated using a pattern from a projector as part of a structured light reconstruction experiment. The obtained reconstruction of a hand is shown in Figure 14 and proves that the calibration was correct.

### E. Real camera calibration using three VPs

The camera calibration using three VPs was applied for the reconstruction of a 3D cube build using Google SketchUp. Two images of the cube were taken, see Figure 15, and the camera's parameters and position with respect to the world was calculated using the three vanishing points resulting from

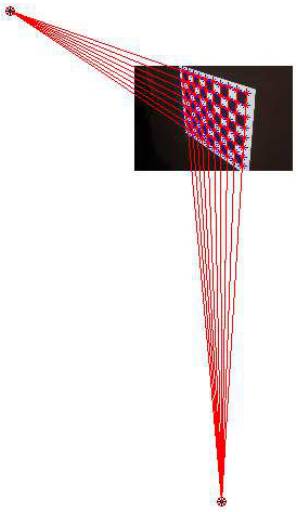

Fig. 13. VPs detection from a planar pattern.

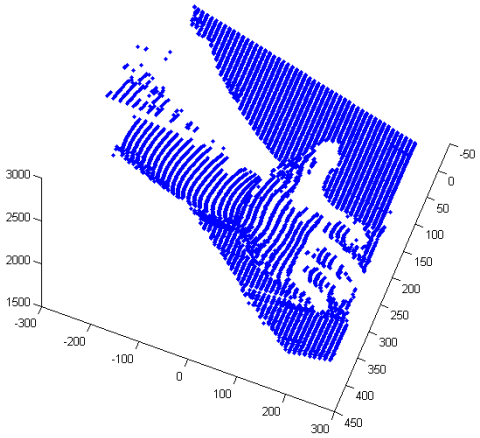

Fig. 14. Hand reconstruction using a structured light system.

the cube. Using the correspondences and the camera model, the object reconstruction was obtained from the two images and the texture was mapped correspondingly, as shown in Figure 16.

## V. Conclusion

Two camera calibration methods, based on vanishing points, have been presented in detail in this paper, using a standardized

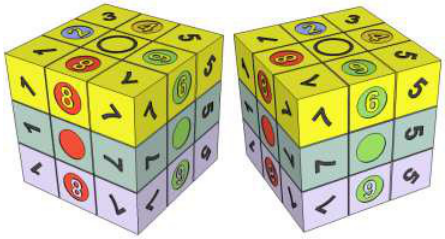

Fig. 15. Two images of a cube presenting three vanishing points.

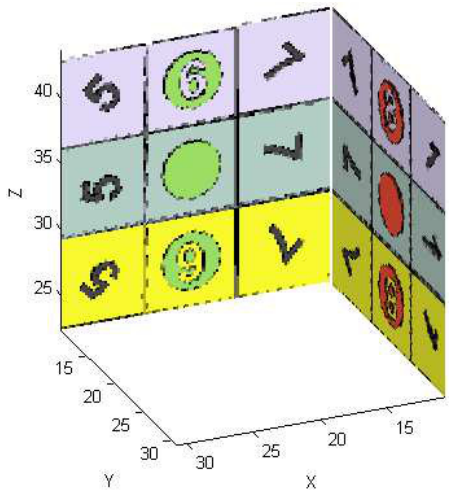


Fig. 16. Reconstruction of a cube from two images with the camera calibrated using three VPs.

mathematical formalization for both of them. Our goal was to analyze the performance of these methods and to highlight their advantages and disadvantages.

In terms of the calibration setup complexity, the method using only two VPs is clearly easier to use since we only need a simple planar pattern capable of producing two VPs in orthogonal directions. However, the camera must be carefully oriented with respect to the calibration pattern in order to avoid obtaining VPs at infinity. Moreover, special care must be taken when defining the world reference system as the two VPs can be placed along any two axes. Therefore, this calibration method is more suitable for controlled configurations in which it is possible to have a good initial estimation of the orientation of the camera with respect to the world.

Regarding robustness to noise, which has been checked in perfectly controlled simulation conditions, the two methods have similar performances up to a reasonable level of noise with the two VPs method performing a little better thanks to more rigorous initial constraints.

One reason to prefer a method against another is from the point of view of the possible applications. Sometimes it is impossible to have a perfectly determined 3D structure in the scene and the 2D plane can be found easier. On the other hand, if it is possible to determine a 3D structure we can benefit from a more robust calibration and, thus, the resulting model is more accurate.

## REFERENCES

[1] Paul Beardsley and David Murray. Camera calibration using vanishing points. *The British Machine Vision Conference (BMVC)*, pages 416–425, 1992.

[2] Michael Bosse, Richard Rikoski, John Leonard, and Seth Teller. Vanishing points and three-dimensional lines from omni-directional video. *The Visual Computer*, 19:417–430, 2003.

[3] J. Bouguet. Camera calibration toolbox for matlab, 2004.

[4] B. Caprile and V. Torre. Using vanishing points for camera calibration. *Int. J. Computer Vision*, 4(2):127–140, 1990.

[5] X. Chen, J. Davis, and P. Slusallek. Wide area camera calibration using virtual calibration object. *Proc. CVPR 00*, pages 2520–2527, 2000.

[6] R. Cipolla, T. Drummond, and D. Robertson. Camera calibration from vanishing points in images of architectural scenes. *BMVC99*, pages 382–391, 1999.

[7] Guillou E., Meneveaux D., Maisel E., and Bouatouch K. Using vanishing points for camera calibration and coarse 3d reconstruction from a single image. *The Visual Computer*, 16:396–410, 2000.

[8] Hall E.L., Tio J.B.K., McPherson C.A., and F.A. Sadjadi. Measuring curved surfaces for robot vision,. *Computer*, 15:42, 1982.

[9] O. D. Faugeras and G. Toscani. The calibration problem for stereo. *Proceedings of IEEE Computer Vision and Pattern Recognition*, pages 15–20, 1986.

[10] L. Grammatikopoulos, G. Karras, and E. Petsa. Camera calibration combining images with two vanishing points. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences.*, 35(5):99–104, 2004.

[11] R. Hartley and R. Kaucic. Sensitivity of calibration to principal point position. pages 433–446, 2002.

[12] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision – 2nd Edition*. Cambridge University Press, 2004.

[13] B.W. He and Y.F. Li. A novel method for camera calibration using vanishing points. *14th International Conference on Mechatronics and Machine Vision in Practice*, pages 44 – 47, 2007.

[14] J. Heikkila and O. Silven. A four-step camera calibration procedure with implicit image correction. *CVPR*, page 1106ij1112, 1997.

[15] J. Kogecka and W. Zhang. Effcient computation of vanishing points. *IEEE International Conference on Robotics and Automation*, 1:223–228, 2002.

[16] Dong Hoon Lee, Kyung Ho Jang, and Soon Ki Jung. Intrinsic camera calibration based on radical center estimation. *The 2004 International Conference on Imaging Science, Systems, andTechnology,*, pages 7–13, 2004.

[17] Carsten Rother. A new approach for vanishing point detection in architectural environments. *Journal Image and Vision Computing (IVC; Special Issue on BMVC 2000)*, vol. 20, no. 9-10:pp. 647–656, January 2002.

[18] J. Salvi, X. Armangué, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 7:35, 2002.

[19] J.P. Tardif. Non-iterative approach for fast and accurate vanishing point detection. *IEEE 12th International Conference on Computer Vision*, pages 1250–1257, 2009.

[20] R. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, 3:323, 1987.

[21] Z. Zhang. A flexible new technique for camera calibration. msr-tr-98-71. Technical report, Microsoft Research, Redmond, USA, 1998.