PROJECTOR-CAMERA CALIBRATION FOR 3D RECONSTRUCTION USING VANISHING POINTS

A. Radu Orghidan, C. Mihaela Gordan, D. Aurel Vlaicu *

Technical University of Cluj-Napoca, Romania

B. Joaquim Salvi

University of Girona, Spain

ABSTRACT

This paper presents a novel method for the calibration of a structured light configuration, which is a particular case of a stereoscopic system in which one of the cameras is replaced by a projector. The method uses *vanishing points* to remove the keystone effect and to extract the parameters of the projector and of the camera. The calibration is performed using a simple flat surface with orthogonal edges and no texture. The result is a simplified calibration technique which does not need complex calibration objects and that can be used in 3D reconstruction applications. The model obtained is validated through depth estimation measurements in a real scene.

***Index Terms*—** Vanishing points, Self-calibration, Pattern Projection, Structured Light, 3D Reconstruction.

1. INTRODUCTION

Visual 3D models are traditionally being used in object inspection and reconstruction, scene mapping and target or self localization. Nowadays, there is an increasing demand for 3D content in the video games industry, virtual reality and communications. Therefore, methods to perform metric measurements in a non-intrusive manner have been investigated for many years. The structured light (SL) techniques [10] are well-known solutions for 3D data acquisition. The typical SL configuration is formed by a stereoscopic system in which one of the cameras is replaced by a projector. The role of the projector is to actively introduce landmarks in the scene in order to solve the correspondence problem. Therefore, SL are also known as active acquisition techniques.

The calibration of the projector, as part of a SL configuration, is a critical issue since it is needed for performing the triangulation. The projector can be seen as a reverse pinhole camera [13, 5]. However, camera calibration methods cannot be directly applied for the projector calibration since they have to cope with the lack of an image of the scene. Besides, if the projector is not orthogonal to the projection screen, a specific distortion, called *the keystone effect*, appears. The traditional calibration of a SL system involves a previously calibrated camera and the use of manufactured objects with known geometry and location. The camera is used for calculating the 3D position of the projected features. This is also the main drawback of this methodology since the approximations of the camera model propagate to the projector's model. A better solution would be to calibrate independently the camera and the projector. Recently, Martynov [7] described a projector calibration method using an uncalibrated camera as a sensor to determine the projector 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 image taken with the camera, briefly called in the remaining of the paper camera image, which might be difficult if the projected pattern is distorted due to a possible misalignment between the projector and the screen.

The camera used to provide the input for the calibration of the projector introduces a perspective transformation of the scene. Thus, a new distortion is added to the camera image of the projected pattern. Despite this double distortion of the pattern, a screen-projector homography can be inferred [11] from the camera image and the projector image can be obtained. The projector image is the image projected by the projector and can be considered as the image taken by a camera with the same intrinsic parameters as the projector and placed at exactly the same location as the projector. A perspective transformation can be determined and applied to the projected pattern in order to compensate the keystone effect. Thus, a pinhole model can be obtained for the projector.

A convenient way to calibrate a pinhole model without using a predefined calibration object is to use the image of the points at infinity, known as *vanishing points* (VPs), which can be accurately extracted from the scene structures [12] when working in man-made environments. The properties of the

*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.

VPs obtained from orthogonal directions [4, 2] are directly related to the focal length and the rotation of the camera with respect to the world coordinate system. Therefore, the VPs encapsulate important information about the pinhole model and there are many available methods [9, 6, 1, 12] for the accurate detection of the VPs.

In this paper, we present a novel calibration method of a projector, as essential component of a SL system. First, the projector projects a rectangular image on a screen with orthogonal edges. The camera takes the image of the screen containing also the projected pattern affected by the keystone effect. Then, a new projector image is obtained by transforming the original image in order to remove the keystone effect. The resulting projector image presents VPs along the two orthogonal edges of the screen. The VPs are eventually used to calibrate the projector.

The calibration algorithm has the following steps:

1. Remove the keystone effect,
2. Determine the VPs of the projector image,
3. Calibrate the projector using its VPs.

In the following, the mathematical background for the calculation of the VPs of the projector and the calibration methodology are presented in section 2. The calibrated model is validated by using it for 3D reconstruction as explained in section 3. The paper ends with the conclusions that are detailed in section 4.

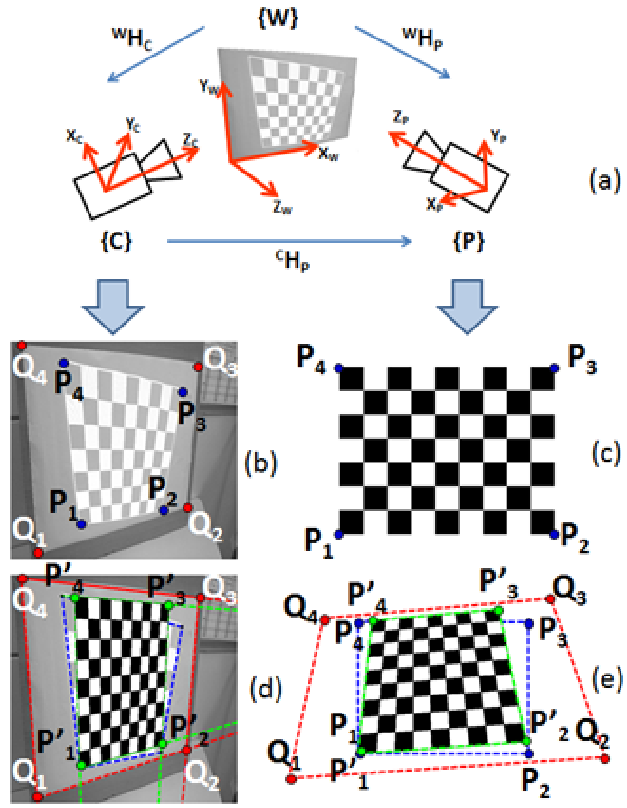


Fig. 1. Flowchart of the calibration of the SL system.

2. CALIBRATION METHODOLOGY

The steps of the calibration process are illustrated in Fig. 1. The SL configuration is outlined in Fig. 1(a). The first step is to remove the distortion of the projector, distortion visible in Fig. 1(b), introduced by the nonorthogonality of the projector to the screen. This goal is reached by transforming the original projector image, shown in Fig. 1(c), such that VPs of the observed pattern are aligned with the VPs of the screen, see Fig. 1(d). Then, the the two VPs of the resulting projector image, represented in Fig. 1(e), are used for the calibration of the projector, modeled as a reverse pinhole camera.

2.1. Estimation of the projector's VPs

A rectangular image is projected onto a screen located in front of the projector. The optical axis of the projector is not orthogonal to the screen. The projected pattern is observed by the camera and is affected by a double distortion resulting from the combination of the keystone effect and the camera perspective transformation. Figure 2 illustrates the four projected points P_i that bound the distorted pattern. A projector with the optical axis perpendicular to the screen would produce a rectangular image having the same VPs as the screen. Therefore, we can determine the positions of the points P_i' of such a projection. Let us denote by V_x and V_y the VPs of the camera image along the X and Y world axes, respectively. The condition for the keystone effect removal is to enforce the co-linearity between the points (P_1', P_2', V_x) and (P_4', P_3', V_x) along the X direction and between the points (P_1', P_4', V_y) and (P_2', P_3', V_y) along the Y direction. Let the point $P_1 = P_1'$. Then, the remaining three points can be obtained from the intersections of the following lines:

$$\begin{aligned} P_2' &= (P_1', V_x) \bigcap (P_2, V_y) \\ P_4' &= (P_4, V_x) \bigcap (P_1, V_y) \\ P_3' &= (P_4', V_x) \bigcap (P_2', V_y). \end{aligned} \tag{1}$$

The transformation that can compensate for the two distortions can now be calculated as a homography relating the two sets of points, $P_i = [x_i, y_i]^T$ and $P_i' = [x_i', y_i']^T$:

$$\begin{bmatrix} x_i' \\ y_i' \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_i \\ y_i \end{bmatrix}. \tag{2}$$

At this point, the correction of the projector image can be obtained by applying the transformation calculated from the system of equations (2). However, since the projector doesn't provide an image of the scene, we can't identify the edges of the screen directly onto the projector image, thus, the vanishing points V_x and V_y are not available, yet. Therefore, another

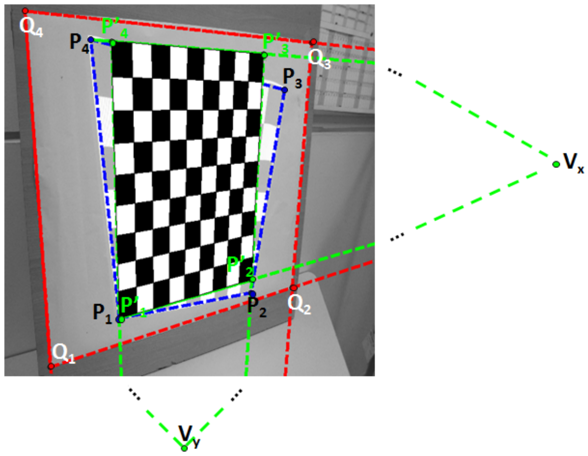


Fig. 2. Removal of the keystone effect from the projected pattern.

step is required: finding the correspondence between the camera image and the projector image.

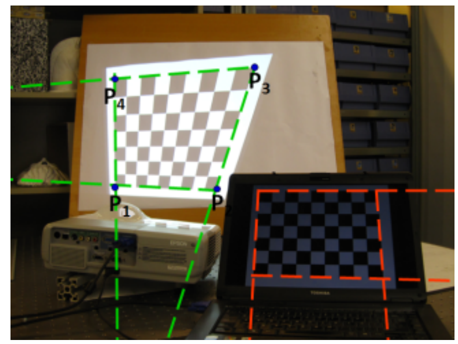
Let us consider that the camera and the projector reference systems, shown in Fig. 1(a), are placed at $\{C\}$ and $\{P\}$, respectively. The two devices point towards a planar surface aligned with the XY plane of the world coordinate system $\{W\}$.

The homography ${}^{W}H_C$ between the screen plane and the camera image can be calculated, if at least four points are available. Similarly, the homography ${}^{C}H_P$ between the images of the projector and the camera is determined. Thus, the homography between the screen and the projector can be calculated:

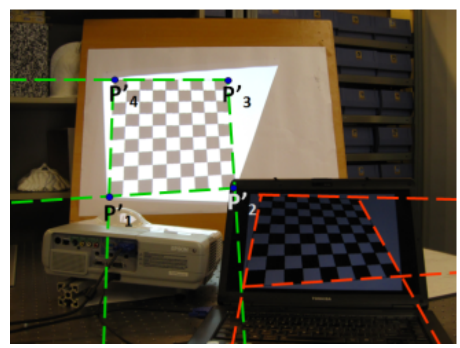
$$ {}^{W}H_P = {}^{W}H_C \cdot {}^{C}H_P. \tag{3} $$

Considering the VPs of the screen, the points P'_i are calculated on the camera image, see Fig. 1(d). Using the homography ${}^{P}H_C = inv({}^{C}H_P)$, the points P'_i can be also located on the projector image. Figure 1(e) shows the image of the projector containing the estimated position of the points Q_i and P_i. Thus, the transformation of the pattern can be modeled as the homography (2) that will rectify any image projected in the particular projector-screen configuration. Such a transformation is useful for the projection rectification but has the drawback that the original image is either oversampled or subsampled resulting in a loss of quality. When dealing with structured patterns, such as the checkerboard, this drawback can be overcome by generating a new pattern instead of applying a deformation on the original pattern. The homography ${}^{W}H_P$ can be used to define the edges of an image on the screen and then relate them with the projected image. Figs. 1(d) and (e), illustrate a checkerboard pattern that was generated instead of deforming the original one. The transformed image of the projector, shown in figure 1(e), contains two VPs that are directly related with the relative position of the projector with respect to the camera and the screen. Therefore, the VPs of the projector can be used for the calibration of the projector's pinhole model.

Figure 3 shows the effect of the image transformation using the method described previously. A checkerboard with orthogonal edges is projected and the resulting pattern is visibly affected by the keystone effect, as shown in Fig. 3 (a). The new image, obtained after the transformation, is projected and the distortion free pattern is visible on the target screen shown in Fig. 3 (b).



(a) Uncorrected image



(b) Corrected image

Fig. 3. The VPs of the projector are revealed after removing the keystone effect.

2.2. Pinhole model calibration using two orthogonal VPs

Two VPs from orthogonal directions, determined using the orthogonal edges of the screen, can be used for the calibration of a pinhole device using a method similar to the one described by Guillou et al. [2].

The world coordinate system is set to be centered at O_w and has the orthogonal axes (x_w, y_w, z_w). The camera coordinate system is located 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. 4. The coordinates

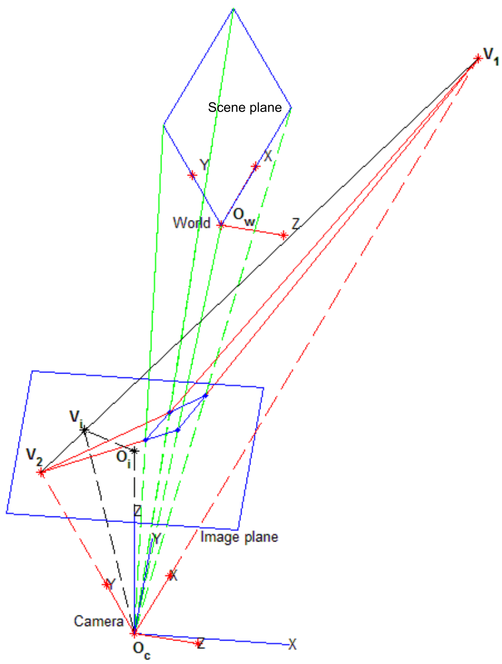


Fig. 4. The focal distance and the orientation of the camera with respect to a plane in the scene can be determined from the vanishing points.

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.

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 using only two vanishing points. 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 [3] for the calculations implied in the calibration process.

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. 4, which means that:

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

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

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

The rotation between the world and the camera coordinate systems is expressed by the matrix ${}^C\mathbf{R}_W$. Taking into account that (a) 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 (b) all parallel lines meet at a vanishing point, it follows that we can build a new coordinate system centered at O_c that has 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 $\mathbf{X_c'}$, $\mathbf{Y_c'}$, and $\mathbf{Z_c'}$ are:

$$\begin{aligned} \mathbf{X_c'} &= \frac{\overrightarrow{O_cV_1}}{\|\overrightarrow{O_cV_1}\|} = \left[\frac{v_{1i}}{\|\overrightarrow{O_cV_1}\|} \quad \frac{v_{1j}}{\|\overrightarrow{O_cV_1}\|} \quad \frac{f}{\|\overrightarrow{O_cV_1}\|} \right]^T \\ \mathbf{Y_c'} &= \frac{\overrightarrow{O_cV_2}}{\|\overrightarrow{O_cV_2}\|} = \left[\frac{v_{2i}}{\|\overrightarrow{O_cV_2}\|} \quad \frac{v_{2j}}{\|\overrightarrow{O_cV_2}\|} \quad \frac{f}{\|\overrightarrow{O_cV_2}\|} \right]^T \\ \mathbf{Z_c'} &= \mathbf{X_c'} \times \mathbf{Y_c'} \end{aligned} . \tag{6}$$

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

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

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. The segment is determined by the world points ${}^W\mathbf{P}_1 = [0, 0, 0]^T$ and ${}^W\mathbf{P}_2 = [x_{p2}, y_{p2}, z_{p2}]^T$.

The segment can be aligned with its perspective projection in the camera coordinate system using the rotation matrix ${}^C\mathbf{R}_W$:

$$\begin{bmatrix} {}^C\mathbf{P}_{1m} \\ {}^C\mathbf{P}_{2m} \end{bmatrix} = {}^C\mathbf{R}_W \begin{bmatrix} {}^W\mathbf{P}_1 \\ {}^W\mathbf{P}_2 \end{bmatrix}. \tag{8}$$

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

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

The segment can be translated 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 new 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 segment and two similar triangles are formed: $\triangle O_C P_1 P_2$ and $\triangle O_C P'_1 Q$. It follows that:

$$\frac{\|O_cP_1\|}{\|O_cP'_1\|} = \frac{\|P_1P_2\|}{\|P'_1Q\|}. \tag{11}$$

Therefore, the distance D from the camera to the world is:

$$D = \|O_c P_1\| = \frac{\|O_c P_1'\| \cdot \|P_1 P_2\|}{\|P_1' Q\|}. \quad (12)$$

Hence, the translation vector is:

$$\mathbf{t} = D \frac{O_c P_1'}{\|O_c P_1'\|}. \quad (13)$$

As a preliminary step to the calibration using the VPs, the rotation about the X and Y axes of the calibration plane must be determined such that the VPs of the world's XY axes are aligned with the camera axes. Then, the intrinsic and extrinsic camera parameters can be obtained by means of the geometric relations presented above. Since the VPs are invariant, the camera translation is, in our implementation, refined through a Levenberg-Marquardt error minimization algorithm with the initial solution given by Eq. (13).

3. EXPERIMENTAL RESULTS

The calibration is evaluated by performing one-shot reconstructions of objects placed in front of the SL system formed by a camera and the projector calibrated using the proposed method. The setup is composed of a Canon EOS 30D camera, a Epson EMP-400W projector, and a laptop with an Intel Pentium Dual CPU at 2.16GHz with 2GB RAM. The SL configuration is placed in front of a planar screen, with an arbitrary orientation with respect to the camera and the projector.

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 shown in Figs. 5(a) and (b).

First, the camera image is segmented and the stripes from a set of rows are detected. The stripe segmentation is performed using the derivative of the luminance profile for each row. The center of each stripe is detected with sub-pixel accuracy using a peak detector. The color of a stripe can be precisely classified among the four levels of hue composing the projected pattern. Then, 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. A large number of correspondences can be detected by applying the process iteratively for all the image rows that contain the object.

Figure 5 shows two examples of reconstructions performed with the calibrated SL configuration. Figures 5(a) and (c) show the image of a hand with the projected pattern and its reconstruction, respectively. The scene to be reconstructed includes the background plane, the forearm, and the palm with the fingers. The moon statue, presented in Fig. 5(b), was reconstructed and the background plane was filtered out. The forehead and the chin can be seen clearly in Fig. 5(d).

The accuracy of the reconstruction is estimated using a planar surface, such as the background reconstructed in Fig. 5(c), placed at approximatively 1.5m from the camera. The 3D points are automatically filtered in order to eliminate the outliers. A number of 5509 points remained after removing the Delaunay triangles with an area larger than a given threshold. A plane, fitted to the cloud of points, is considered as the reference plane and the distance from each point to the plane is calculated. The average distance from the points to the planar surface is around 3% of the distance from the projector to the planar surface and the standard deviation is 4.8%.

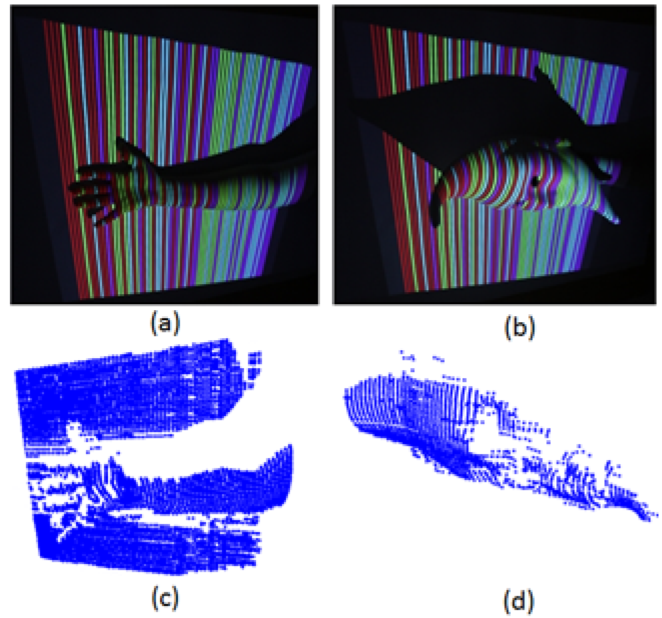


Fig. 5. Forearm and moon statue reconstructed using the calibrated SL system.

4. CONCLUSIONS

This paper presents a novel method for the calibration of a projector using vanishing points. The calibration setup is very simple since there is no need for specially tailored calibration objects. Despite the simplicity of the calibration method, it provides a resolution of about 3% of the measured range even for single shot reconstructions. The accuracy proves that the method is suitable for 3D computer vision applications that need good depth estimation and a fast self-calibration.

A toolbox containing the proposed calibration method is available for public use [8].

5. REFERENCES

[1] 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.

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

[3] R. I. Hartley and R. Kaucic. Sensitivity of calibration to principal point position. In *Proc. Seventh European Conference on Computer Vision*, pages 433–446. Springer, 2000.

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

[5] M. Kimura, M. Mochimary, and T. Kanade. Projector calibration using arbitrary planes and calibrated camera. *CVPR*, 2007.

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

[7] Ivan Martynov, Joni-Kristian Kämäräinen, and Lasse Lensu. Projector calibration by "inverse camera calibration". In *Scandinavian Conf. on Image Analysis (SCIA)*, 2011.

[8] Radu Orghidan and Joaquim Salvi. Toolbox for the calibration of a structured light system. `users.utcluj.ro/~orghidan`, 2012.

[9] 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. 9-10:647–656, January 2002.

[10] J. Salvi, S. Fernandez, T. Pribanic, and X. Llado. A state of the art in structured light patterns for surface profilometry. *Pattern Recognition*, 43(8):2666–2680, 2010.

[11] Rahul Sukthankar, Robert G. Stockton, and Matthew D. Mullin. Smarter presentations: Exploiting homography in camera-projector systems. *Proceedings of International Conference on Computer Vision*, 1:247–253, 2001.

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

[13] X. Zhang and L. Zhu. Projector calibration from the camera image point of view. *Optical Engineering*, 48:117–208, 2009.