Abstract—In the past, 3D shape reconstruction process was based on passive stereo which do not require direct control of any illumination source, instead relying entirely on light. Nowadays, 3D shape reconstruction is based on active stereo which replace one camera with a projector. The projector plays an important part in solving the correspondence problem. It projects coded patterns on the scanned object. By capturing the deformed pattern using cameras, the correspondences between image pixels and projector (columns-rows) can be found easily. To do that, the projector must be calibrated. In this work, the problem of projector calibration is solved by passive stereo and triangulation. Our system consists of two cameras, projector, and planer board. A checkerboard pattern is projected on the board and then captured by the two cameras. Using triangulation, the corresponding 3D points of the projected pattern is computed. In this way, having the 2D projected points in the projector frame and its 3D correspondences (calculated using triangulation) the system can be calibrated using a standard camera calibration method. A data projector has been calibrated by this method and accurate results have been achieved.

Index Terms—Correspondences, projector calibration, 3D reconstruction, triangulation.

I. INTRODUCTION

This paper addresses the problem of projector calibration which is critical step in any active vision systems, and particularly in active optical scanners. Passive optical scanners do not require direct control of any illumination source. One of the most widely used passive 3D imaging systems is stereo vision system. Like the human visual system, stereoscopy use triangulation to estimate the positions of 3D scene points. First, the 2D projection of a given point is identified in each camera. Using the calibration parameters of each camera, a single 3D line is drawn from each camera’s center of projection through the 3D point. The depth of the point is recovered by the intersection of these two lines. The main problem in these systems is correspondences matching among the various viewpoints. Matching algorithms development remains an open and challenging problem in the field [1]. By using controlled illumination, active optical scanners could overcome the correspondence problem. Solving the correspondence problem in many active systems is done by replacing one of the cameras, in a passive stereoscopic system, with a controllable illumination source. A projector can be used as a source of illumination, but it must be calibrated. A data projector can be seen as a dual of a camera.

In practice, projector calibration is more complicated than camera because projectors cannot capture the surface that they illuminate so a camera must be used to make the correspondence between the 2D projected points and the 3D illuminated points and also it’s difficult to obtain the co-ordinates of the 3D points because the calibrating pattern is projected and not attached to the world coordinate frame.

In this work, we present a method of projector calibration which based on passive stereo and triangulation. It aims to be both accurate and easy-to-use. This method has been implemented as an extension to Bouguet Camera Calibration Toolbox [2].

The remainder of this article is organized as follows. Section II discusses the related work. Sections III through VI describe our system. Our results are described in Section VII. Finally, Section VIII gives the conclusion.

II. RELATED WORK

Projector calibration has received increasing attention, because of the emergence of lower-cost digital projectors. A projector is simply the "inverse" of a camera, where 2D points on an image plane are mapped to outgoing light rays passing through the center of projection. Camera and projector calibrations are the necessary steps in any active computer vision systems, and therefore, various approaches and methods have been proposed to calibrate projectors.

One class of these approaches projects a calibration pattern onto a plane, “the wall”, captures it by a camera, and then goes through the standard calibration work flow. It makes use of the idea which is based in considering the projector as an inverse camera which maps 2D image intensities into 3D rays. This idea is not new, and it has been explained by several authors [3]–[4]. Sergio Fernandez et al. proposed a plane-based calibration method of a projector-camera system. A checkerboard pattern is projected on a plane which contains another printed checkerboard. They recover 3D position for each projected corner using ray-plane intersection [5]. Ivan Martynov et al. also proposed a projector calibration method by inverting the standard camera calibration workflow. The calibration procedure requires a single camera, which does not need to be calibrated. The camera works as the sensor whether projected dots and calibration pattern landmarks, such as the checkerboard corners, coincide. The 3D position for the projected dots is recovered by adjusting the projected dots to coincide with the landmarks and the final coordinates are used as inputs to a camera calibration method [6].

Another important class of the methods, including those referred to as Auto-calibration methods. These methods do not need a physical calibration target. Most auto-calibration
methods can only estimate the extrinsic parameters [7] or require a calibrated camera [3], but recently many automatic methods have been proposed. For example, the method by Draneni et al. [8] assumes a plane projection geometry, “the wall”, and that one of the projector poses is “roughly frontal”. These methods are attractive choices because of their automatic processing, but there is always a need for highly accurate calibration in the structured light and active vision systems. The auto-calibration methods can solve the extrinsic parameters, but the intrinsic parameters should be solved by the inverted camera approach which uses a physical calibration target, since this is accurate and should be done just once. Furukawa and Kawasaki proposed a technique which uses structured light projection to calibrate the projector [9]. The correspondences are obtained using Gray code patterns, and the projector’s intrinsic and extrinsic parameters are estimated using the epipolar constraints. The calibration depends on the non-linear optimization of an objective function, which needs good initial values of both intrinsic and extrinsic parameters.

Our system uses two cameras in the calibration stage instead of one camera to increase the accuracy of our system. Adding the second camera will not increase the cost of the system. Because there are many systems such as [10] and [11] which use two cameras in the 3D reconstruction stage, but they do not use them in the calibration stage. In the 3D reconstruction stage, every camera with the projector will reconstruct parts of the scanned object which not seen from the other camera. Merging these parts together will reconstruct the whole object in only few scans.

III. SYSTEM OVERVIEW

Our method consists of three major steps (See Fig. 1). The first step is pattern displaying and capturing. In this step, the projector displays a checkerboard pattern on the white board in full screen mode. The two cameras capture this pattern and store the images. The white board is placed in different positions and the above step is repeated. Section IV.A describes the capturing step in detail.

The second step is corners extraction and correspondences matching. In this step, the 2D corners in every pair of images are extracted so, we have many sets of correspondences between the two cameras. Section IV.B describes the corners extraction step in detail. Finally, the 3D points of the right/left correspondences are reconstructed using triangulation and the calibration step is done. Section V describes the reconstruction step in detail, while section VI describes the calibration step.

IV. CORRESPONDENCES MATCHING

This section describes how to get the correspondences points and it consists of two steps.

A. Pattern Displaying and Capturing

Our system consists of two calibrated cameras (left-right), white board, and projector (See Fig. 2). The two cameras are calibrated using Zhang’s method [12], a flexible new technique to easily calibrate a camera. It only requires the camera to observe a planar pattern shown at a few (at least two) different orientations. This procedure consists of a closed-form solution, followed by a nonlinear refinement based on the maximum likelihood criterion. Zhang’s technique has been tested using both computer simulation and real data, and accurate results have been obtained. For more details about Zhang’s algorithm, see [12]. This algorithm was implemented in Matlab Camera Calibration Toolbox [2] by Jean-Yves Bouguet and C++ in Intel OpenCV library [13].

![Pattern displaying and capturing](image1)

![Camera extraction and correspondences matching](image2)

![Compute the 3D points of the projected pattern using triangulation](image3)

![Projector calibration](image4)

Fig. 1. Overview of our pipeline.

![Pattern displaying and capturing](image5)

![Camera extraction and correspondences matching](image6)

![Compute the 3D points of the projected pattern using triangulation](image7)

![Projector calibration](image8)

Fig. 2. (a) The white board and the world coordinate system. (b) The two cameras and the projector.

![Extracted corners](image9)

![Extracted corners](image10)

![Extracted corners](image11)

Fig. 3. (a) Checkerboard pattern in projector frame. (b) The left view. (c) The right view. (d) The left extracted corners. (e) The right extracted corners.
These two libraries are probably the most widely used tools for camera calibration nowadays. The projector displays a checkerboard pattern (see Fig. 3a). Fig. (3b, 3c) show this pattern falling on the white board and seen from the two cameras. Move this board in many positions and capture the pattern with the two cameras. Now, we have many pairs of images (left-right) and we are ready for the next step.

B. Corners Extraction

After capturing the patterns, the 2D corners in all (left-right) image pairs are extracted. First, the right image is displayed and the four extreme corners on the projected checkerboard pattern are clicked clockwise or counter clockwise starting with any corner. When the left image is displayed, the same clicking mechanism must be used. The corners of all pairs are extracted by the corner extraction engine used in Bouguet Calibration Toolbox. Fig. (3d, 3e) show the extracted corners. (i.e., the correspondences).

V. RECONSTRUCTION BY TRIANGULATION

This is the most important step in our system. The 3D coordinates values of every left-right images corners extracted in the previous step can be constructed using triangulation. In this section we explain models describing the image formation process, leading to the development of reconstruction equations allowing the recovery of 3D points by geometric triangulation.

A. Perspective Projection and Pinhole Model

The pinhole model is simple and popular geometric model for cameras or projectors. It composed of a plane and a point external to that plane. The plane is called the image plane and the point is called the center of projection (see Fig. 4a). In a camera, every 3D point (except the center of projection) determines a unique line passing through the center of projection. If this line is not parallel to the image plane, then it must intersect the image plane in a single image point. In mathematics, this mapping from 3D points to 2D image points is called a perspective projection. The geometry of a projector can be described with the same model because of the fact that light traverses this line in the opposite direction. That is, given a 2D image point in the projectors image plane, there must exist a unique line containing this point and the center of projection (since the center of projection cannot belong to the image plane). In summary, we can say that the projector is a camera inverse which means that light travels away from a projector along the line connecting the 3D scene point with its 2D perspective projection onto the image plane [10].

1) The ideal pinhole camera

In the ideal pinhole camera shown in Fig. 4b, the center of projection is at the origin of the world coordinate system, with coordinates (0, 0, 0)\(^T\), and the point \(q\) and the vectors \(v_1\) and \(v_2\) are defined as

\[
\begin{bmatrix}
v_1 \\
v_2 \\
q
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}.
\]

Note that not every 3D point has a projection on the image plane. An arbitrary 3D point \(p\), with coordinates \((q^1, p^1, p^2)\) belongs to this plane if \(p^3 = 0\), otherwise it projects onto an image point with the following coordinates.

\[
\begin{align*}
&u^1 = p^1 / p^3 \\
&u^2 = p^2 / p^3
\end{align*}
\]

The relation between the coordinates of a point and the image coordinates of its projection can be described in many ways; for example, the projection of a 3D point \(p\) with coordinates \((p^1, p^2, p^3)\), has image coordinates \(\mu = (u^1, u^2, 1)\) if, for some scalar \(p^3 \neq 0\), we can write

\[
\lambda \begin{bmatrix}
u^1 \\
u^2 \\
1
\end{bmatrix} = \begin{bmatrix}
p^1 \\
p^2 \\
p^3
\end{bmatrix}.
\]  

(1)

2) The general pinhole camera

It is not necessarily that the center of a general pinhole camera is placed at the origin of the world coordinate system and it may be oriented. However, it does have a camera coordinate system attached to the camera, in addition to the world coordinate system (see Fig. 4c). A 3D point \(P\) has world coordinates described by the vector \(p_w = (p^1, p^2, p^3)\) and camera coordinates described by the vector \(p_c = (p^1, p^2, p^3)\). These two vectors are related by a rigid body transformation specified by a translation vector \(T \in \mathbb{R}^3\) and a rotation matrix \(R \in \mathbb{R}^{3 \times 3}\), such that

\[
p_c = Rp_w + T.
\]

In camera coordinates, the relation between the 3D point coordinates and the 2D image coordinates of the projection is described by the ideal pinhole camera projection (i.e., (1)), with \(\lambda \mu = p_c\). In world coordinates this relation becomes

\[
\lambda \mu = Rp_w + T. 
\]  

(2)
The parameters \( R, T \) are the extrinsic parameters of the camera; describe the location and orientation of the camera with respect to the world coordinate system. These parameters translate the coordinate of a point from the world coordinate to the camera coordinate. Equation (2) assumes that the unit of measurement of lengths on the image plane is the same as for world coordinates, that the distance from the center of projection to the image plane is equal to one unit of length, and that the origin of the image coordinate system has image coordinates \( U_i = 0 \) and \( V_i = 0 \). In practice, none of these assumptions hold. For example, lengths on the image plane are measured in pixel units, and in meters or inches for world coordinates, the distance from the center of projection to the image plane can be arbitrary, and the origin of the image coordinates is usually on the upper left corner of the image. In addition, the image plane may be tilted with respect to the ideal image plane. To overcome these limitations of the current model, a matrix \( K \in \mathbb{R}^{3 \times 3} \) is introduced in the projection equations to describe intrinsic parameters as follows.

\[
\lambda u = K (R p_v + T).
\]  

The matrix \( K \) has the following form

\[
\begin{pmatrix}
fs_1 & fs_\theta & 0^1 \\
0 & fs_2 & 0^2 \\
0 & 0 & 1
\end{pmatrix}
\]

where \( f \) is the focal length (i.e., the distance between the center of projection and the image plane). The parameters \( s^1 \) and \( s^2 \) are the first and second coordinate scale parameters, respectively. Note that such scale parameters are required since some cameras have non-square pixels. The parameter \( S_\theta \) is used to compensate for a tilted image plane. Finally, \((s^1, s^2)^T\) are the image coordinates of the intersection of the vertical line in camera coordinates with the image plane. This point is called the image center or principal point. All intrinsic parameters (i.e., the matrix \( K \)) are independent of the camera pose. The matrix \( K \) can be estimated once through a calibration procedure because it describes physical properties related to the mechanical and optical design of the camera. We can normalize image plane measurements in pixel units by multiplying the measured image coordinate vector by \( k^{-1} \), so that the relation between a 3D point in world coordinate and its 2D image coordinate is described by (2) \[10\].

**B. The Mathematics of Triangulation**

Under the pinhole camera model, each corner in the left image creates a ray (the unique line containing this image point and the center of projection), and also the corresponding corner in the right image. The intersection of these two rays is the 3D value related to these corners so; we calculate the 3D values of all the corners of the projected 2D checkerboard pattern. Given a 2D point correspondence \( x_i \) (a corner in the left image); \( x_j \) (the corresponding corner in the right image) in homogeneous coordinates, \( P_i \) and \( P_j \) are the two projection matrices for the left and right cameras respectively, the 3D point location \( X \) is given as follows

\[
\lambda_i x_i = P_i X \\
\lambda_j x_j = P_j X
\]

We can now build the cross-product of each point with both sides of the equation and obtain

\[
x_i \times P_i X = [x_i \times] P_i X = 0 \\
x_j \times P_j X = [x_j \times] P_j X = 0,
\]

where we used the skew-symmetric matrices \([x_i \times]\) to replace the cross product

\[
a \times b = [a \times] b = \begin{pmatrix} 0 & -a_z & a_y \\ a_z & 0 & -a_x \\ -a_y & a_x & 0 \end{pmatrix} b
\]

where \( a = (a_x, a_y, a_z)^T \) and \( b = (b_x, b_y, b_z)^T \).

Each 2D point provides two independent equations for a total of three unknowns. We can therefore solve the over constrained system by stacking the first two equations for each point in a matrix \( A \) and computing the least-squares solution for \( AX = 0 \) which can be easily solved by Singular Value Decomposition (SVD) \[14\]. Applying SVD to yields the decomposition \( A = U D A \). The homogeneous least-squares solution corresponds to the least singular vector, which is given by the last column of \( V \). Now, we have the 3D coordinates of the projected checkerboard pattern (see Fig. 5) and the 2D corners of the projected pattern in the projector frame can be easily obtained by extracting the corners of the image projected by the projector. Finally, the system is ready for calibration.
VI. PROJECTOR CALIBRATION

Modeling the projector as camera inverse lets us use any camera calibration methods to calibrate our projector like Zhang's method. Camera (and also projector) calibration requires estimating the parameters of the general pin-hole model presented in section V.A.1 This includes the intrinsic parameters, being focal length, principal point, and the scale factors, as well as the extrinsic parameters, defined by the rotation matrix and translation vector mapping between the world and camera coordinate systems. In total, 11 parameters (5 intrinsic and 6 extrinsic) must be estimated from a calibration sequence. In practice, a lens distortion model must be estimated as well. We use Bouguet Camera Calibration Toolbox routines to calibrate our projector and accurate results are obtained. Section VII shows these results in details.

VII. RESULTS

This section shows the results we obtained when testing the system. The setup used for the tests was the same one used for the calibration step as in Fig. 2b. It consisted of a data projector (Optoma EH020) with a resolution of 1024×768 pixels, two cameras (Sony nxcam) and a frame grabber (DeckLink Studio) digitizing images at 1920×1080 pixels with 24 bits per pixel (RGB). The method runs on an Intel Core2 Duo CPU at 3.00GHz. We used Bouguet’s calibration toolbox to calibrate our cameras. In order to see the performance of our method, we use the reprojection error function available in the Bouguet’s calibration toolbox. As it can be seen from Fig. (7a, 7b), our method is more accurate than Gabriel Falcao et al. [15] method. Our method’s average standard deviation of the error is [0.19877 0.33484] pixels while [0.85371 0.78253] pixels for Gabriel Falcao et al. method.

Distance measurement is carried out to evaluate the performance of our calibration method. An 8×8 checkerboard pattern (49 corners) is projected on the whiteboard. The 3D coordinates corresponding to these corners calculated using both camera-camera triangulation and camera-projector triangulation are obtained. The whiteboard is moved in five different positions and in every position the 3D coordinates of the pattern corners are calculated. Table I shows our results of this experiment compared with the results obtained by calibrating the projector using Gabriel Falcao et al. method.

Finally, in order to demonstrate the applicability and efficiency of the proposed technique a 3D reconstruction of an LCD screen using structured light has been performed. The algorithm proposed by Posdamer and Altschuler [16] was implemented. The reconstructed LCD screen can be observed in Fig. 6.

VIII. CONCLUSION

In this paper, we have described an easy and accurate projector calibration method which based on passive stereo and triangulation. The simplicity of the method comes from considering the projector as an inverse camera and thus making the calibration of a projector the same as that of a camera for which there already exists well and accurate established methods. This method has been implemented as an extension to Bouguet Camera Calibration Toolbox and makes extensive use of its functions.
Projector calibration is more complicated than camera calibration because projectors cannot capture the surface that they illuminate, so we must use a camera to make the correspondence between the 2D projected points and the 3D illuminated points. Since the calibrating pattern is projected and not attached to the world coordinate system, it is difficult to retrieve the co-ordinates of the 3D points. Projecting a checkerboard pattern on a white board and capturing it from two points of view helps us to solve the problem and compute the 3D co-ordinates of the projected pattern corners by using Triangulation. In order to verify the correctness and the accuracy of our calibration method, a simple reconstruction of an LCD screen has been performed. Our method is more accurate than Gabriel Falcao et al. method. The use of two cameras increases the accuracy of our method. The two cameras will help us in the 3D shape reconstruction stage. Every camera with the projector will reconstruct parts of the scanned object which not seen from the other camera. The whole object will be constructed in only few scans by merging these parts together. So, adding the second camera will not increase the cost of our system.

REFERENCES


Yousef B. Mahdy was born at Luxor, Egypt in 1956. He received the B.Sc. and M.S. degrees in electrical engineering, both from Assiut University, Assiut, Egypt, in 1979 and 1983, respectively, and the Ph.D. degree in Computer Engineering, Assiut University, Egypt under scientific channel system with Michigan Tech. University and University of Missouri, Rolla, USA, 1991. From 1988 to 1990, he worked as Visiting Scholar at the EE Dept. of MTU, Houghton, USA.

From 1990 to 1991, he worked as Visiting Scholar at the EE Dept. of University of Missouri, Rolla, USA. From 1991 to 1997, he worked as Assistant professor of Computer Engineering, Assiut University, Egypt. From 1997 to now, he is an Associate professor of Computer Engineering, Assiut University, Egypt. He is now Dean, Faculty of Computers and Information, Assiut University. His major research interests include computer architecture, fault tolerance, vision systems, parallel computing.

Khaled F. Hussain was born at Assiut, Egypt in 1972. He received the B.Sc. (with distinction and honors) and M.S. degrees in electrical engineering, both from Assiut University, Assiut, Egypt, in 1994 and 1996, respectively, and the Ph.D. degree in computer science from the University of Central Florida, Orlando, in 2001. From 2002 to 2006, he worked as a Visiting Assistant Professor with the University of Central Florida. From 2006 to 2007, he worked as an Assistant Professor with the Department of Electrical Engineering, Faculty of Engineering, Assiut University. Since 2007, he has been with the Department of Computer Science, Faculty of Computers and Information, Assiut University, and is currently an Executive Director with the Multimedia Laboratory, Faculty of Computers and Information, Assiut University. His major research interests include computer vision, computer graphic, augmented reality, and computer animation.

Mostafa A. Abdel-Majid was born in Minia, Egypt 1985. He received the B.Sc. (with distinction and honors) in Computer Science, faculty of Computers and Information, Assiut University, 2006. He is now a post graduate student at faculty of Computers and Information, Assiut University and Demonstrator, Department of Computer Science, Faculty of Computers and Information, Assiut University. Scientific His major research interests include interests: 3D scanners, gpu computing computer vision, and computer graphic.