

Spatial location in 360° of reference points over an object by using stereo vision

V. H. Flores^a, A. Martínez^a, J. A. Rayas^a, and K. Genovese^b

^a Centro de Investigaciones en Óptica, A.C.

Apartado Postal 1-948, León, Gto., 37150, México,

^bSchool of Engineering, Università degli Studi della Basilicata,
Viale dell'Ateneo Lucano, 10-85100 Potenza, Italy.

Received 11 April 2012; accepted 12 February 2013

Stereo Vision is a powerful tool used to get a 360° scan of an object in order to obtain topography details or getting the spatial position of a given set of points of interest. However, the required computational time could be slow to perform real-time measurements. In this work we present an alternative approach based on Hering's coordinate system and the use of high reflective markers, used as reference points to track an object movement. The advantage of using these markers is that their detection is faster than a full scene correlation since it is performed by matching the position of the centroids of each marker without using any pixel-by-pixel analysis.

Keywords: Hering's coordinates; stereo vision; track markers.

PACS: 42.30.Tz; 42.40.My; 42.30.Sy.

1. Introduction

The shape recovery of an object has a variety of applications. One of them is the creation of digital models of real objects in order to be recreated by using a 3D-printer [1]. Such a process consists on the tridimensional construction of pieces using powder through a layer by layer deposition to get an exact copy of the 3D CAD model. When the reproduction of a real object is entailed, a 360 degrees digital reconstruction of the object is needed to create the CAD model to give in input to the 3D-printer.

One of the methods used for shape recovery is the fringe projection technique [2,3]. In this method, a fringe pattern is projected over the object. The image of the object with the superimposed fringe pattern is captured by using a CCD camera and then processed by in order to get the fringes phase distribution. To this scope, several techniques have been proposed, such as phase shifting [4] or Fourier method [5]. By using the information on phase distribution and sensitivity vector of the optical arrangement, it is possible to recover the object topography. The main limitation of this approach is the need to have a reference plane which implies practical problems for the study of large objects [6].

In this work, we propose a system that uses stereo vision to identify the spatial location of markers collocated on a sample which will be scanned in 360°. The markers will be used to construct a coordinate system that depends only of the sample. In order to do the spatial recognition of the centroids we used stereo vision [7,8]. The algorithms used for the correlation of the markers are slow to do a real time 3D reconstruction such as the "Block Matching" method [9]. So we used a faster version of this algorithm based in the detection of the centroids of the markers so instead of making a pixel by pixel identification. This algorithm compares the coordinates of the centroids to get the information of their spatial location.

The experimental results obtained in this work will be part of the implementation of an optical system to a full-view scanning of an object based on four steps. In the first part, the localization of the markers over the object will established. The second part corresponds to the obtaining of virtual reference planes by use of at least three points. The object topography will be recovered by the technique projected fringes using the virtual reference planes. In the third part, it tracks the movement of the markers to rotate the perspective of the topography recovered while another area is profiled. Finally, the different views will be pasted to have a complete topography of an object.

The detection of the markers and the use of a sample-referenced system involve several issues like occlusions of the markers, wrong correlation between the points, deformations of the recovered shapes depending of the profiling perspective, etc.

The aim of this work is the location of the markers over the object.

2. Stereo vision

2.1. Parallel axes geometry

The stereo vision technique is very useful to get information on the depth of points of the same object. The capability of perceiving the depth is called *stereopsis* and it is obtained by capturing two images with two cameras separated by a certain distance called Stereo-system baseline. The optical system is shown in Fig. 1.

The experimental setup consists of two cameras which capture an image of the same object but from different perspectives. The cameras are separated by a distance b . Both cameras are located at a distance D from a virtual reference plane which is tangent to the highest point of the object. The

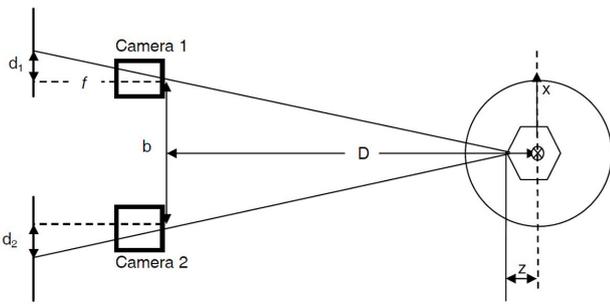


FIGURE 1. Experimental arrangement for stereo vision technique.

following equation is used to calculate the depth of any point of interest of the object with respect to the virtual reference plane [11].

$$z = D \frac{f}{d_1 - d_2} \tag{1}$$

where:

z = Depth of the point.

D = Distance between the two camera planes and the virtual reference plane.

f = Focal distance of the camera.

d_1 and d_2 = Distance from the optical axis of the camera and the image of the point of interest on the sensor plane.

Each image point in the image of the two camera sensors has a given position (namely d_1 and d_2) with respect to the camera optical axis. The two distances d_1 and d_2 directly depend on the spatial location of the points. The difference between these distances is called *disparity*.

The depth of the object points of interest is obtained from the disparity of the two image points in the two images captured by the cameras. By using this approach it is possible to obtain a point cloud which defines the shape of the object. To find the disparity between the two images of each point of interest, two homologous images need to be individuated by performing the so-called “matching operation”. In this work we used a simplified version of the Block Matching algorithm according to the flux diagram shown in Fig. 2.

The algorithm of block matching consists on capturing a reference image and comparing with the other image in order to get the more similar areas between the images and calculate the disparities of corresponding points. The result is a set of blocks with disparity information for each pixel of the scene.

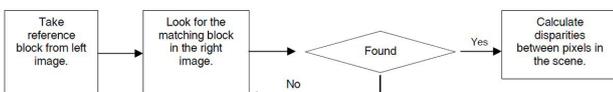


FIGURE 2. Flux diagram of the Block Matching Algorithm.

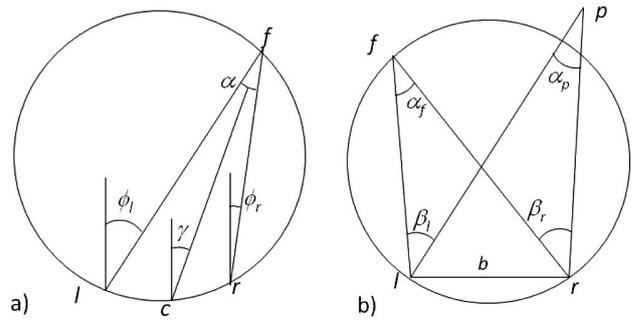


FIGURE 3. Convergent axes arrangement. f : fixation point; l and r : view points; c : cyclopean point; p : point of interest; α : vergence; γ : version, b : baseline distance. α_f and α_p are the vergence of the fixation point and the point of interest. β_l and β_r are the angles between the fixation point and the point of interest.

In this work, we use an algorithm that detects the centroid of each marker for each pair of images. These centroids are used as points of interest and their disparities allow obtaining the spatial location of each marker. Conversely to the standard approach, we only make the comparison of two vectors with the information of the centroids of each image and then locate the marker. This process is faster than a pixel-by-pixel-based process.

2.2. Hering Coordinates

In Fig. 3 the variables describing a binocular head vision system are shown. The presented stereo vision corresponds to a convergent axes arrangement. In this case we assume that the cameras are epipolar to the (x, z) plane. The viewing direction of the cameras, φ_l and φ_r , are defined as the angle between the optical axis and the axis defined by the camera position and fixation point respectively. Positive angles are measured clockwise.

Instead of using these viewing directions (φ_l and φ_r) we will use the quantities such as [8]:

$$\alpha = \varphi_l - \varphi_r \tag{2}$$

$$\gamma = \frac{1}{2}(\varphi_l + \varphi_r) \tag{3}$$

These quantities are known as the Hering-vergence (α) and Hering-version (γ).

Vergence value is 0 when cameras axes are parallel. For each vergence, $\alpha < 0$, a circle can be drawn through the nodal points and the intersection point of the two camera axes. This circle is called the Vieth-Müller circle (VMC) of vergence α , this circle is shown in Fig. 3a, and has the radius of:

$$R = \frac{b}{2 \sin \alpha} \tag{4}$$

Version is defined as the average (or cyclopean) viewing direction of the two cameras, in a more formal way, when both cameras have a fixation point f with some vergence α , the half-way spot of the baseline between the two nodal

points (cameras) is called cyclopean point and the visual direction from this point to the fixation point is the Hering-version.

On the other hand, by trigonometry the Hering α, γ coordinates can be turn into Cartesian x, y, z coordinates by using this transformation:

$$\begin{aligned}
 H(\alpha, \gamma) &= \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{b}{2 \sin \alpha} \begin{bmatrix} \sin 2\gamma \\ 0 \\ \cos \alpha + \cos 2\gamma \end{bmatrix} \\
 &= R \begin{bmatrix} \cos \varphi_r \sin \varphi_l - \cos \varphi_l \sin \varphi_r \\ 0 \\ 2 \cos \varphi_r \cos \varphi_l \end{bmatrix} \quad (5)
 \end{aligned}$$

The assumption of $y = 0$ implies that the cameras are located at the same height between them, so it can be said that there is no vertical disparity between the two perspectives captured by the cameras. In the practice the problem is to locate the vergence and version of each point of interest with respect to the fixation point. The viewing angles β_l and β_r are used to calculate the disparity and the eccentricity of each point.

$$\delta = \beta_l - \beta_r \quad (6)$$

$$\eta = \frac{1}{2}(\beta_l + \beta_r) \quad (7)$$

A positive disparity d denotes that the interest point is located front of the fixation point, and a negative one means that is located back of the fixation point. On the other hand

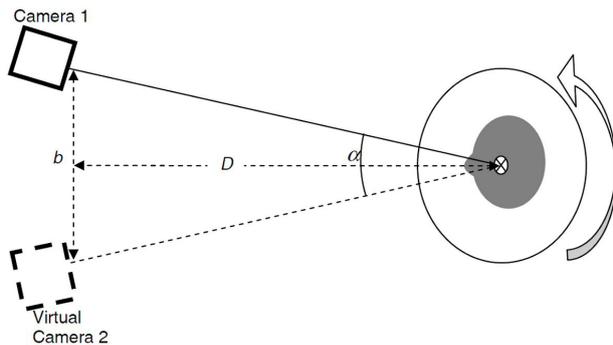


FIGURE 4. Optical arrangement with a rotating mount used to emulate stereo vision.

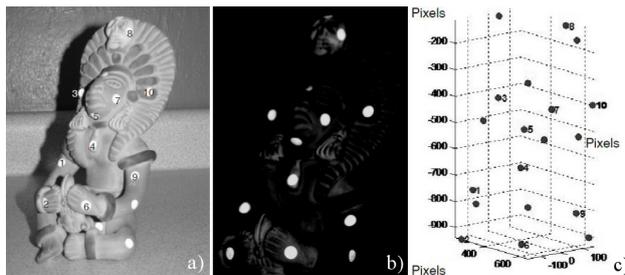


FIGURE 5. a) Sample with markers, b) Image captured by the camera c) Detected points around the sample shown in a 3D map.

the eccentricity determinates if the interest point is located right or left of the reference point.

3. Experimental

3.1. Optical system with one camera

In this experiment the stereo system was emulated as shown in Fig. 4 by using only one monochromatic FireWire CCD camera of 8 bits that has a resolution of 659×494 pixels with a lens with 8 mm of focal length and rotating plate with an accuracy of 1' per step. The sample was rotated in order to have two views of the object to locate the markers.

Images of the sample were captured sequentially by rotating the mount at intervals of 8 degrees. Once the different views of the test object were taken, we consider pairs of images to have the two necessary perspectives to perform stereo correlation. The angles of vergence and version were measured from the angle of the mount rotation used in the optical arrangement showed in Fig. 4. The sample was scanned at 360° . The accuracy of the method depends of the angular displacement between the views of the object.

To avoid the uncertainty of the measurement due to the depth of focus, the diaphragm of the camera was closed and we used a high frequency lamp to illuminate the sample and to improve the brightness of the high reflective markers. Figure 5 shows: a) the scanned object with the markers, b) the image captured by the camera with the bright markers and c) the graphic of some detected points. In Fig. 5a), the spots with no labels belong to points of hidden perspectives.

This methodology is very practical for a scan at 360° and it is easy to make the mapping of the centroids because we know the angle of rotation of the object. However this method cannot be implemented in real time due to the need of rotating the sample to simulate stereo vision.

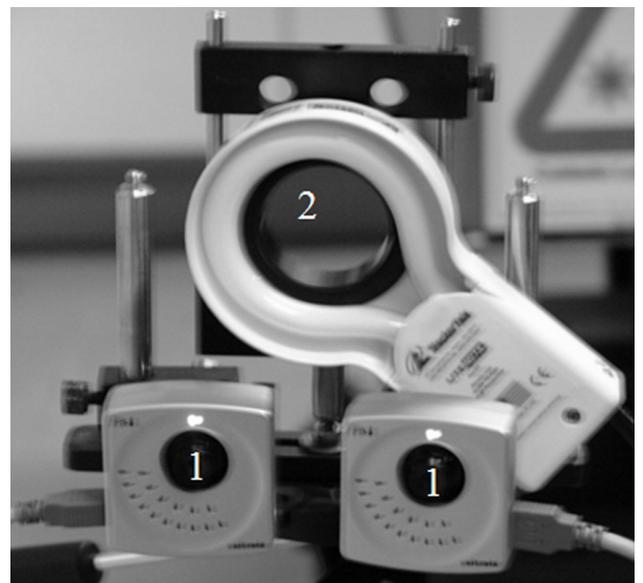


FIGURE 6. Optical system for stereo vision: 1) CCD cameras, 2) High frequency lamp.

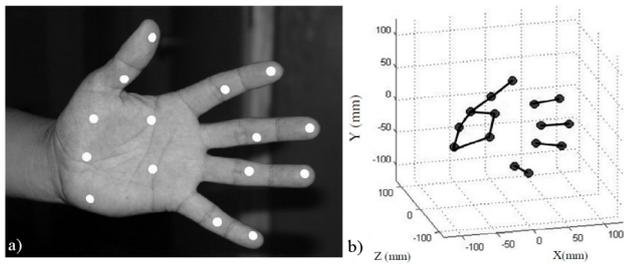


FIGURE 7. a) Hand used as test object, b) 3D plot of the spatial position of each of the markers. The coordinate (0, 0, 0) of the system corresponds to the fixation point of the system.

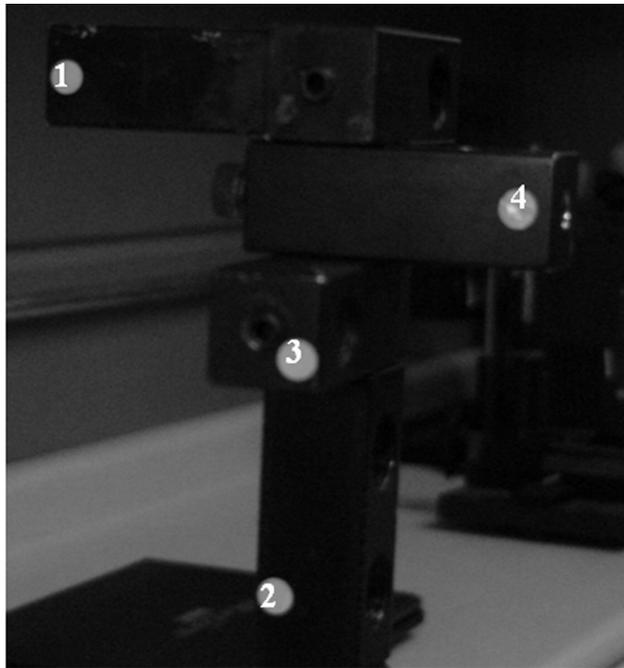


FIGURE 8. Stacked blocks which have been marked and numerated for measurement of depth between them.

3.2. Optical system with two cameras

Figure 6 shows the implemented system. In this case, two cameras (with the same specifications mentioned in Sec. 3.1) simultaneously captured two different perspectives of the object.

For the follow-up testing and identification of markers, we used the hand of a volunteer which was placed in front of the optical system for locating the markers. The test object is shown in Fig. 7a. The parameters of the geometry of the arrangement such as view angles and the distance from the cameras plane to the fixation point were considered in the program to monitoring of the hand. One given position of the hand was captured by the CCD camera. The digitization of the markers placed onto the hand is shown in Fig. 7b. It can be observed the spatial distribution of the markers. This is a special case in which the points of the test object are near to the plane containing the fixation point.

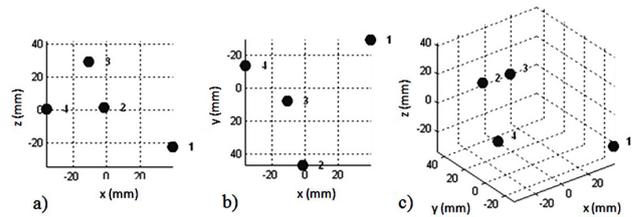


FIGURE 9. Markers location plots: a) xz-plane, b) xy-plane, c) 3D plot.

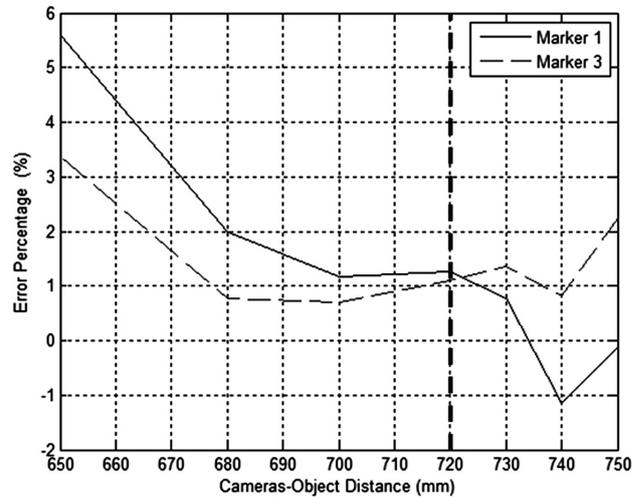


FIGURE 10. Graphic of error in the spatial localization of markers 1 and 3.

To obtain the measurement error, we measured an object whose dimensions are known. The object was built by stacking some blocks, Fig. 8. The distances between the blocks were measured with a caliper. The measurements obtained with optical system shown in Fig. 6, are compared with the target positions measured with the caliper. The difference between the two values is considered the error in the measurement.

Figure 9 shows the markers positions detected by the stereo system. The whole object was collocated in different positions with respect to the fixation plane. Figure 10 shows the measurement error for markers 1 and 3 which are the farthest to the fixation point. It is observed that the error in the spatial localization of the markers increases as much as the markers depart from fixation plane.

As shown in Fig. 10, the error is larger for the markers 1 and 3 due to the fact that they are localized farther from the fixation point plane. This test was done by moving the sample in different positions, from 650 mm to 750 mm far from the cameras. The line in 720 mm indicates the distance where the fixation plane is located. It can be seen in the graphic that the error is reduced when the markers are located near the fixation point. The localization of the markers 1 and 3 with respect to fixation point plane is of 750 mm and 700 mm respectively. The maximum error was approximately 5.6% which is associated for marker 1.

Since we used a pin-hole camera model, the error expressed in Fig. 10 is associated to the error introduced by

distortion of the lenses and to the distance of the points with respect of the fixation point. The alternatives to reduce the error are to correct the image distortion caused by the lens and to use an iterative algorithm⁶ to correct the position of the markers with the respect of their distance of the fixation point.

4. Conclusions

We have implemented the stereo vision technique based on the Hering coordinate system for point detection and depth measurement in a scene.

A scanning was performed at 360° of markers placed onto an archaeological piece by using a camera and rotating the object. The stereo system that uses a single camera has the disadvantage that cannot be used for real-time measurements.

Qualitative results were obtained in the experiment of tracking the movement of a hand. In this case the optical system comprises two cameras. High reflectance markers were

placed on the test object and their centroids were spatially detected by using the stereo system. The technique detects the trajectory of each point in real time.

To evaluate the error of the stereo vision technique that uses two cameras, the results were compared with reference values obtained with a caliper. A maximum error was obtained for the points localized beyond the reference plane that contains the fixation point. This system has a volume of work of approximately 100 mm around the fixation point with a maximum error of 5.6%.

The detected markers will be used to create virtual reference planes in profilometry techniques where is no possible to have a physical reference plane.

Acknowledgements

This research is part of a bilateral project sponsored by CONACYT-MAE between Mexico and Italy, reference number 146523. VHFMM wants to thank to CONACYT for his scholarship.

-
1. J. A. Alonso, *Sistemas de Prototipado Rápido* (Electronic Version, 5-8 2001). <http://webs.uvigo.es/disenoindustrial/docs/protorapid.pdf>
 2. K. J. Gasvik, *Optical Metrology* (Wiley, Great Britain, 2003). p. 180-188.
 3. J. A. Rayas, R. Rodríguez-Vera, and A. Martínez, "Three-dimensional micro-topography by Talbot-projected fringes", *Proc. of SPIE* 6046(1), 60461 y-1 - 60461 y-6 (2006).
 4. D. Malacara, M. Servín, and Z. Malacara, *Interferogram Analysis For Optical Testing* (Marcel Decker, USA, 1998). p. 247-278.
 5. M. Takeda, H. Ina, and S. Kobayashi, *J. Opt. Soc. Am.* **72** (1982) 156-160.
 6. A. Martínez, J. A. Rayas, H. J. Puga, and K. Genovese, *Opt. and Lasers Eng.* **48** (2010) 877-881.
 7. K. Reinhard, K. Schlüns, and A. Koschan, *Computer Vision* (Springer-Verlag, Singapore, 1998). p. 129-176.
 8. P. Geibler, T. Dierig, and H. A. Mallot, *Computer Vision and Applications* (Academic Press, USA, 2000). p. 398-413.
 9. J. Stam, "Stereo Imaging with CUDA", NVIDIA, (1998).
 10. L. Gordon, "3D Scanning to Go", Z Corporation.
 11. J.F. Cárdenas-García, H. and Yao, and S. Zheng, *Opt. and Lasers Eng.* **22** (1995) 192-213.