ISSN 0005–1144

Tomislav Pribanić, Mario Cifrek, Stanislav Peharec

Simplified Light Plane Determination during Structured Light Scanning

UDK 004.923
IFAC 5.8.7

Original scientific paper

Structured light illumination is a widely spread approach for 3D shape reconstruction. Scanning the scene of interest with hard edge stripe via video projector is a very common implementation. During the scanning light planes are being projected across the space and their positions in time are needed to triangulate 3D position of some point. Traditionally, light plane equation is readily obtained in case of calibrated video projector. This paper presents a method where light plane positions are determined without explicit projector calibration. The comparison of traditional and proposed method showed no difference in terms of reconstruction accuracy. However, there are two major advantages of proposed method: end user can faster start using the system itself and system implementation is easier primarily in software sense. Also, successful method implementation for computer graphic application is shown.

Key words: structured light, video projector calibration, 3D reconstruction

1 INTRODUCTION

3D reconstruction of surfaces is one of the most important tasks in a variety of applications and areas, particularly in the area of machine/computer vision. Some examples of a wide field of application are: reverse engineering, robot navigation, extracting 3D structure, industrial part inspection, computer graphics animation, sports/medicine etc. Very often there is a prerequisite that the employed 3D measurement technique has to be contact less [1]. Stereovision fulfils this demand. It is based on imaging a scene with camera from two or more points of view. However, an image obtained by the vision system can be simply expressed as a perspective projection of the scene in question, inherently two dimensional. The missing depth dimension can be recovered after finding pixel correspondence on different images, in order to triangulate 3D position [2]. The correspondence problem becomes feasible when surface is either textured or there are easily distinguishable points such as image corners, markers attached to certain positions of interest etc. Although applicable in a variety of applications such systems have certain inherent restrictions. Namely, we are actually reconstructing only a rather limited number of object points in space, e.g. attached markers [3], and not the entire object shape which is within the cameras field of view. Still, there is a number of applications which require full 3D shape recovery, particularly in cases of non-textured surfaces [4].

Many techniques have been proposed, both passive and active [5]. Among the active techniques a rather convenient approach is the so called structured light where an artificial source of light is usually used instead of ambient light [6, 7, 8, 9, 10]. The main advantage of using artificial source of light is that features extracted from images are better defined and hence correspondence problem is easier solved even for completely non textured surfaces. Namely, by means of artificial source of light the scene is illuminated by projecting one or more light patterns. Each image point in the projector's pattern(s) can be coded in certain way (hence, sometimes referred as coded light approach). As the result of pattern(s) projection (almost) all image points captured by the camera will also be coded in a way that correlates (i.e. corresponds) to certain image point in the projected pattern(s). Therefore, the correspondence problem can be neatly solved and triangulation is made possible.

Codification strategy can vary considerably [11]. Actually, there is no absolute strategy suitable to all kinds of applications and scene characteristics since usually some scene/measuring setup conditions are presumed [12, 13, 14]. Some strategies require colour neutral scenes (e.g. emphasized sensitivity to surface reflectance/albedo), but are usually meant for dynamic scenes and thus require projection of a single light pattern. Other strategies can be more robust, but assume strictly static objects since they project multiple patterns; furthermore, some are more sensitive to surface discontinuities etc. Generally speaking, projection strategies could be summarized in three different categories: time--multiplexing (e.g. grey code), spatial neighbourhood (e.g. De Brujin sequences) and direct coding (e.g. grey levels).

This paper focuses on the strategy which can be qualified as time multiplexing strategy. Specifically, our light source is low end video projector which projects in time vertical hard edge stripe. Each projected stripe is characterized by correspondent projectors image horizontal coordinate. Effectively, the so called planes of light, spanned by the projector's optical centre and projector's stripe edge image, are projected in space, i.e. scanned throughout. The method is already rather explored [4]. For example, apart from a light source it is similar to approach undertaken in [15]. However, it quite differs in determination of geometry between the source of light (video projector in our case) and camera itself. In other words, in order to carry out triangulation in addition to camera calibration light source (projector) parameters must also be known [16, 17]. In brief, calibrated projector allows the determination of light plane equation in space at the time instance of crossing over certain point, which can be identifiable on the captured images of the camera. We will show an alternative way for calculating light plane equations without explicitly calibrating the projector. In the following sections the entire scanning principle will be explained in more detail, along with the traditional projector calibration method and the proposed method. Also, the two methods will be compared through accuracy reconstruction evaluation and practical implementation issues.

2 SCANNING PRINCIPLE

A short programming script was written in order to move vertical hard edge black stripe on the white background from the left to the right side of the PC monitor (Figure 1). The same event was projected in space with commercially available low end video projector.

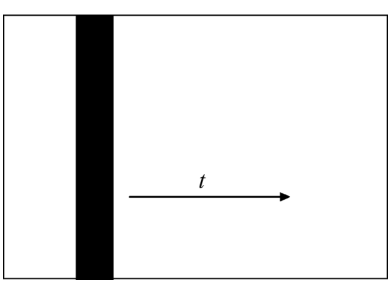

Fig. 1 Projected pattern at a time instance t

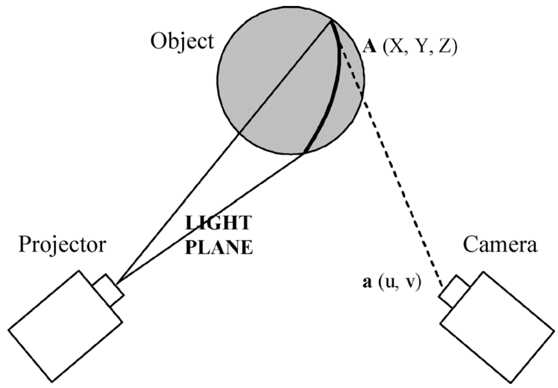

Fig. 2 Triangulation principle during the hard edge stripe scanning

At the same time, scanning was imaged by a video camera. Figure 2 describes geometrical relation at camera frame, as follows. Black hard edge stripe (Figure 1) is back projected [18] in space as a light plane (Figure 2). Light plane intersects the object of the scene and its intersection is detectable on the camera image. Similarly to the image line every detected intersection edge pixel $a(x, y)$ can be back projected in space as a line (dashed line on Figure 2). Assuming that we know the equation of the back projected line and the light plane we can rather simply find the 3D point position $A(X, Y, Z)$ for every intersection edge pixel in all frames. The above assumption is readily satisfied for the calibrated pair of a video camera and projector.

The basic principle of 3D point position determination described above may not be easily applicable for non neutral scenes where it might be demanding to determine the edge (i.e. intersection of light plane with object) on camera images due to different albedo/reflection properties [1]. Thus, the determination of light plane crossing a point in the scene can be alternatively viewed in temporal domain, rather than spatial [15]. In other words as

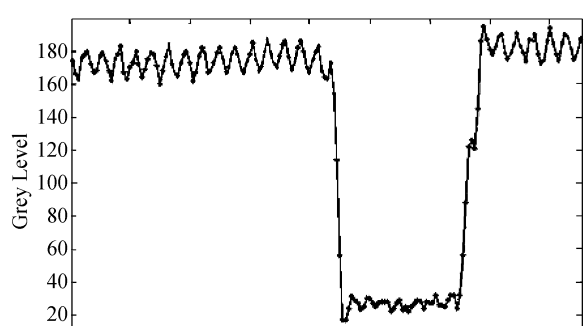

Fig. 3 Change of grey level value for pixel during light planes scanning

the scanning takes place the illumination of certain points in space changes from maximum to minimum and back to maximum. It means that the grey level value of the corresponding image pixels changes accordingly (Figure 3).

Based on Figure 3, for every pixel we can calculate its minimum, maximum and, from these two, average value. The time when light plane crosses over a certain point can be defined as the moment where the change of grey level value reaches calculated average. By means of linear interpolation we can calculate the time instance with subframe precision. As already implicated, a similar approach was implemented in [15].

The major difference is in the determination of light plane equations needed for triangulation. Specifically, in [15] shadows (hence, light planes are referred as shadow planes) are cast with a waving stick and a desktop lamp and consequently the use of video projector is avoided. While this approach can be quite attractive, primarily due to the hardware simplicity of the system, it exercises certain practical/computational disadvantages. To mention some of them: firstly, in every image light plane intersection with one or two (depending whether the desktop lamp position is known or not; for more details see [15]) planes in space have to be determined in order to recover light plane equation. This may pose serious restrictions when trying to recover the shape of larger objects and/or from different points of view. Secondly, for an inexperienced user the optimal setup of scanning time, positioning of lamp etc. might take some time. Thirdly, due to the mentioned user interaction fully automatic implementation, needed for many applications, is not possible. For this reason we settled for video projector inclusion regardless of the increased complexity of the system, i.e. cost which we do not consider the major issue in case of low end equipment.

3 DETERMINATION OF LIGHT PLANE EQUATION AVOIDING EXPLICIT PROJECTOR CALIBRATION

This section deals with the core of our approach. In the above text it was clearly stressed that the calibrated video projector would assure light plane equation determination by back projection of the projector vertical image line (Figure 1, i.e. hard edge stripe). Video projector by its nature is said to be in dual relationship with camera [16, 17]. Among other things, duality is obvious through camera and projector calibration principles also. Camera calibration is done with known pairs of 3D points in space and its correspondents detected in an image. In case of projector calibration it is the other way round. While image coordinates are precisely known, its correspondents in space are yet to be determined. Usually, this is done by projecting a certain pattern on known planes in space, capturing images with previously calibrated camera(s) and finally reconstructing them. We have also performed the above mentioned (usual) video projector calibration and compared results with our approach. Several patterns of circles were projected on the white board positioned on multiple locations in space. Images were captured by two cameras (although one camera and known board position in space would be sufficient) and then their 3D projector calibration point positions were reconstructed (Figure 4).

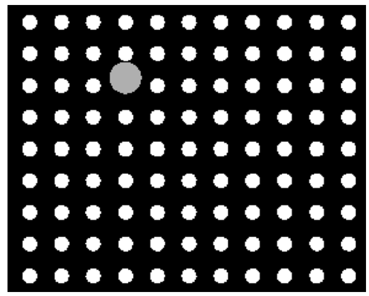

Fig. 4 One of the patterns projected during the traditional video projector calibration. The purpose of the large circle is to solve correspondence problem

We shall now consider the following entities (1). Homogeneous coordinates of a line **l** in the image plane, coordinates **x** of point on the image plane and lying on the **l**, coordinates **X** of corresponding point in space which projects to image point **x** by camera (video projector) projection matrix $\mathbf{P}_{3\times 4}$ [18] and finally coordinates of plane **p** obtained by the so called back projection of line **l**. In other words, plane **p** represents all those points **X** in space that are projected on the line **l** as image **x**.

$$\begin{aligned} \mathbf{l} &= \begin{bmatrix} l_1 & l_2 & l_3 \end{bmatrix}^T \\ \mathbf{x} &= \begin{bmatrix} w\cdot x & w\cdot y & w \end{bmatrix}^T \\ \mathbf{X} &= \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^T \\ \mathbf{p} &= \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \end{bmatrix}^T \end{aligned} \tag{1}$$

where components (x, y) and (X, Y, Z) are inhomogeneous representatives of image and space point, respectively. For given entities these relationships hold up to some scale factors (not written for clarity reasons):

$$\begin{array}{ccc} \mathbf{l}^T \cdot \mathbf{x} = 0 & \mathbf{x} = \mathbf{P}\cdot\mathbf{X} & \mathbf{p}^T \cdot \mathbf{X} = 0 \\ & \downarrow & \\ \mathbf{l}^T \cdot (\mathbf{P}\cdot\mathbf{X}) = 0 & (\mathbf{l}^T \cdot \mathbf{P})\cdot\mathbf{X} = 0 & \mathbf{p}^T = \mathbf{l}^T \cdot \mathbf{P} \\ & \downarrow & \\ & \mathbf{p} = \mathbf{P}^T \cdot \mathbf{l} & \end{array} \tag{2}$$

The above expression for a given camera (projector) projection matrix $\mathbf{P}$ and a line $\mathbf{l}$ in the camera (projector) image plane gives plane $\mathbf{p}$ in space, spanned by camera (projector) projection center and line $\mathbf{l}$ in image plane.

During the scanning procedure video projector projects hard edge stripe which is in our case black (white) vertical line on white (black) background (Figure 1). It is reasonable to assume that the projector driven by PC produces linear scanning in time and therefore, knowing the total projection time T and the projected image resolution W it is possible to predict equation of line $\mathbf{l}$ in projector image coordinates at a time instance t.

$$\mathbf{l}(t) = \begin{bmatrix} 1 & 0 & -\frac{W}{T}\cdot t \end{bmatrix}^T. \qquad (3)$$

Combining (2) and (3) we would be also able to find plane equation $\mathbf{p}$ at a time instance t. However we would first need to calibrate video projector in order to find out matrix $\mathbf{P}$ (2). The question is whether we can take advantage of the linear time scanning and develop similar equation as (3) for plane $\mathbf{p}$ and therefore successfully avoid explicit video projector calibration. Let us assume that we know plane equations $\mathbf{p}_1$ and $\mathbf{p}_2$ at times t_1 and t_2, respectively and we want to find plane equation $\mathbf{p}(t)$ at an arbitrary moment t. From (2) and (3) we can write:

$$\mathbf{p}_1 = \mathbf{P}^T \cdot \begin{bmatrix} 1 \\ 0 \\ -\frac{W}{T}\cdot t_1 \end{bmatrix}$$

$$\mathbf{p}_2 = \mathbf{P}^T \cdot \begin{bmatrix} 1 \\ 0 \\ -\frac{W}{T}\cdot t_2 \end{bmatrix}$$

$$\mathbf{p}_2 - \mathbf{p}_1 = \mathbf{P}^T \cdot \begin{bmatrix} 0 \\ 0 \\ \frac{W}{T} \end{bmatrix} \cdot (t_1 - t_2) \qquad (4)$$

$$\mathbf{p}(t) - \mathbf{p}_1 = \mathbf{P}^T \cdot \begin{bmatrix} 0 \\ 0 \\ \frac{W}{T} \end{bmatrix} \cdot (t_1 - t).$$

It can be seen that the equation for $\mathbf{p}(t)$ can be rearranged as:

$$\mathbf{p}(t) = \mathbf{p}_1 + \frac{\mathbf{p}_2 - \mathbf{p}_1}{(t_2 - t_1)} \cdot (t - t_1). \qquad (5)$$

Evidently, (5) shows that the initial goal for calculating the plane equation at a time instance t, without knowing projection matrix $\mathbf{P}$, is accomplished. The only thing to be explained is how to conveniently find $\mathbf{p}_1$ and $\mathbf{p}_2$ at times t_1 and t_2, respectively.

We will describe how to find plane equation $\mathbf{p}_1$ at an instance t_1 since plane equation at other time instances can be found in analogue manner. For a given projector horizontal resolution W we can pick any hard edge stripe position w_1 ($w_1 < W$), which will be reached at some t_1 during the scanning period of T, according to (3). For the chosen w_1 we can project hard edge stripe on the white board (otherwise used for traditional projector calibration during this work), obtain its image and reconstruct its spatial position. Its reconstructed spatial position L_1 is the line which is intersection of plane $\mathbf{p}_1$ (at instance t_1) and the white board placed in space on location one (Figure 5). From basic geometry it is known that plane is defined by minimum of three non collinear points or a line and a point not lying on the line. We know so far the line L_1 lying on the plane $\mathbf{p}_1$ and therefore we need at least one more point or another line L_2 lying on the plane $\mathbf{p}_1$ as well. By placing the white board on Location 2 in space and projecting the same hard edge stripe from w_1 we can obtain another line L_2 which also belongs to plane $\mathbf{p}_1$ (Figure 5). Knowing lines L_1 and L_2 we can easily find components of plane $\mathbf{p}_1$.

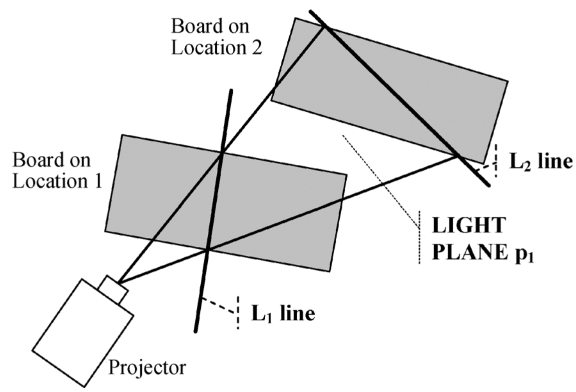

Fig. 5 Calculation of plane $\mathbf{p}_1$ at instance t_1 through reconstruction of lines L_1 and L_2, which are intersections of plane $\mathbf{p}_1$ with board placed on location one and two, respectively. Camera(s) are left out for clarity

4 METHOD EVALUATION

In order to test performance of the previously explained proposed method for calculating plane position at different time instances, we did the following. The white board used for traditional projector calibration was placed on four locations in space inside the cameras and projector calibration volume. On each location the surface of the board was scanned with hard edge black stripe. Then, the changes of grey level value vs. time (Figure 3) were detected for rather large number of cameras' image points, lying in space on the board. For these image points their spatial positions were reconstructed in two ways (Figure 6). First, by taking advantage of traditionally calibrated projector and second, by utilizing proposed method. Once the points were reconstructed, for each location we randomly chose approximately a 100 of equally spaced points in order to calculate plane equation of the board placed at a certain location in space. Ideally, all

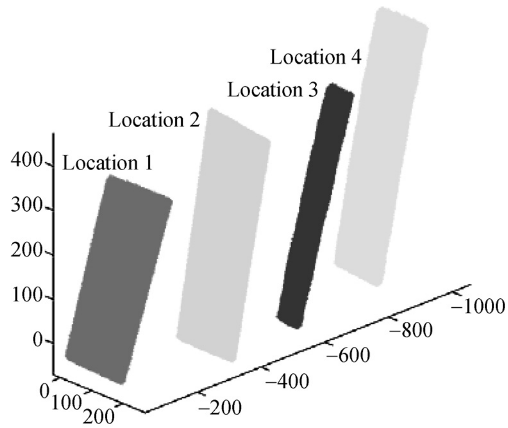

Fig. 6 Clouds of points obtained from scanning a board on four different locations in space and reconstructing them with proposed method

Table 1 **With calibrated projector: Mean distances and root mean square values of distances between reconstructed points and plane in space on which points ideally lay on**

Location	Number of points	Left camera		Right camera	
		mean	rms	mean	rms
		mm	mm	mm	mm
1st	25115	1.22	1.47	0.70	0.88
2nd	25151	1.99	2.48	0.92	1.13
3rd	35571	2.03	2.39	1.21	1.46
4th	35320	2.60	3.12	1.68	2.05

Table 2 **With proposed method: Mean distances and root mean square values of distances between reconstructed points and plane in space on which points ideally lay on**

Location	Number of points	Left camera		Right camera	
		mean	rms	mean	rms
		mm	mm	mm	mm
1st	25115	1.24	1.49	0.71	0.90
2nd	25151	2.06	2.53	0.93	1.13
3rd	35571	2.06	2.43	1.21	1.46
4th	35320	2.60	3.14	1.69	2.05

the reconstructed points should lie on the board, hence reconstructed plane. Thus, for the rest of the points mean distances and root mean square values (RMS) of distances between points and plane were calculated. Table 1 and Table 2 summarize the results obtained for two cameras (denoted left and right due to their spatial orientation with respect to the projector) and for traditional and proposed method, respectively. As it can be seen, there is practically no difference between the two approaches in terms of acquired mean and root mean square values.

In order to further demonstrate the applicability of proposed method it was used for 3D facial reconstruction. More specifically, different facial ex-

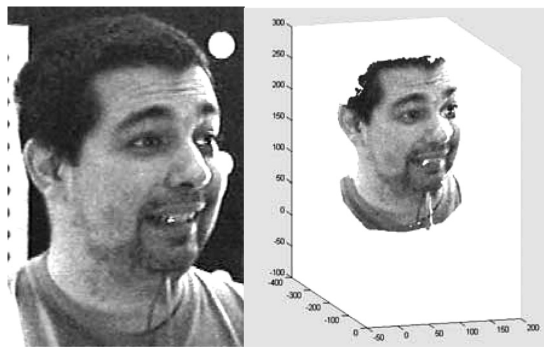

Fig. 7 Camera image and its 3D reconstruction of ***happy*** *mood*

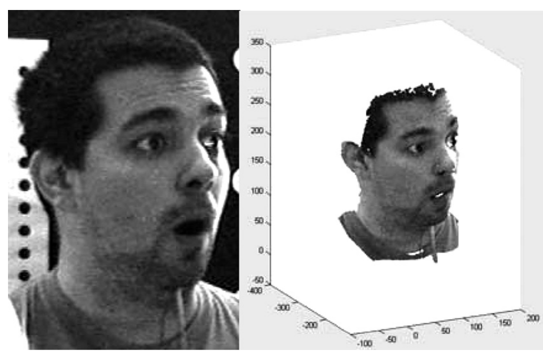

Fig. 8 Camera image and its 3D reconstruction of ***surprised*** *mood*

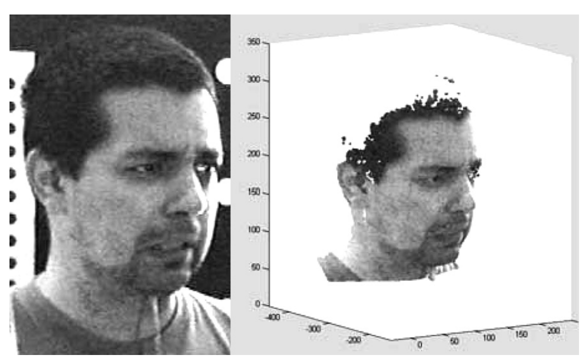

Fig. 9 Camera image and its 3D reconstruction of ***angry*** *mood*

pressions were captured and reconstructed in 3D. The acquired results were used as an intermediate step for another application not discussed in this paper. However, in brief, its goal was to classify a person's mood based on captured facial images, i.e. their 3D reconstruction. Three different types of a person's mood imaged by the camera and then 3D reconstructed are shown below as representative results. Figure 7 shows camera captured image and its 3D reconstructed counterpart respectively, in case of a happy, lively mood. Similarly, Figure 8 demonstrates a surprised, astonished facial expression. Finally, angry mood is shown in Figure 9.

5 CONCLUSION

Among all possible structured light implementations, i.e. pattern projections, this paper has considered probably one of the simplest-scanning with hard edge stripe. Since the scanning was done via video projector there arises a natural need for video projector calibration. Similarly to camera calibration, it is subject to many errors and above all it is a step that end user would like to simplify as much as possible. This paper has demonstrated that it is possible to find plane equations during the scanning even without knowing projector parameters. It has been shown that in terms of undertaken reconstruction accuracy, proposed method is neither superior nor inferior to traditional approach. However, instead of performing traditional projector calibration, all that needs to be done is place a flat surface in space (white board), project the same hard edge stripe on several locations and reconstruct their spatial positions. Obviously, this is far more convenient from both, computational (implementation) and procedural point of view. The latter is frequently one of the major concerns to the end user, since many calibration methods (algorithms) are sensitive to placing and manipulating with calibration devices. Finally, several different types of facial expressions were reconstructed in 3D space. The utilization of obtained results for further processing in computer graphics is additional guarantee of proposed method applicability.

REFERENCES

[1] R. Klette, K. Schluns, A. Koschan, **Computer Vision: Three-Dimensional Data from Images.** Springer Verlag, 1998.

[2] O. Faugeras, **Three-Dimensional Computer Vision.** MIT Press, 1993.

[3] P. Allard, I. Stokes, Jean-Pierre, **Three Dimensional Analysis of Human Movement.** Human Kinetics, Champaign, 1995.

[4] ..., **Simple 3D**; http://www.simple3d.com/

[5] R. A. Jarvis, **A Perspective on Range-finding Techniques for Computer Vision.** IEEE Trans. Pattern Analysis Mach. Intell., 5:122{139, March 1983.

[6] F. W. DePiero, M. M. Trivedi, **3-D Computer Vision Using Structured Light: Design, Calibration, and Implementation Issues.** Advances in Computers 43: 243–278.

[7] D. Scharstein, R. Szeliski, H**igh-Accuracy Stereo Depth Maps Using Structured Light.** IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2003), volume 1, pages 195–202, Madison, WI, June 2003.

[8] G. Sansoni, S. Corini, S. Lazzari, R. Rodella, F. Docchio, **Three Dimensional Imaging Based on Gray-code Projection: Characterization of the Measuring Algorithm and Development of a Measuring System for Industrial Applications.** Applied Optics, 36(19):4463{4472, 1997.

[9] M. Trobina, **Error Model of a Coded-Light Range Sensor.** Technical Report BIWI-TR-164, Communication Technology Laboratory, Image Science Group, ETH Zurich, 1995.

[10] N. L. Chang, L. Nelson, **Efficient Dense Correspondences Using Temporally Encoded Light Patterns.** IEEE International Workshop on Projector-Camera Systems in conjunction with ICCV 2003, 12 October Nice, France 2003.

[11] J. Salvi, J. Pagés, J. Batlle, **Pattern »Codification Strategies in Structured Light Systems«.** Pattern Recognition 37(4), pp. 827–849, April 2004.

[12] O. Hall-Holt, S. Rusinkiewicz, **Stripe Boundary Codes for Real-time Structured-light Range Scanning of Moving Objects.** In The 8^{th} IEEE International Conference on Computer Vision, pages II: 359–366, 2001.

[13] E. Horn, N. Kiryati, **Toward Optimal Structured Light Patterns.** Image and Vision Computing, 17(2):87–97, February 1999.

[14] D. C. Douglas. Hung, **3D Scene Modelling by Sinusoid Encoded Illumination.** Image and Vision Computing, 11: 251–256, 1993.

[15] J-Y. Bouguet, P. Perona, **3D Photography on Your Desk.** IEEE Proceedings of International Conference of Computer Vision, Bombay India January 1998, 43–50.

[16] T. Koninckx, L. Van Gool, **High-speed Active 3D Acquisition Based on a Pattern-specific Mesh.** SPIE's 15^{th} annual symposium on electronic imaging – videometrics VII, vol. 5013, pp. 26–37, January 21–22, 2003, Santa Clara, California, USA.

[17] R. Raskar, P. Beardsley, **A Self Correcting Projector.** Mitsubishi Electric Research Laboratories, TR-2000-46, January 2002.

[18] R. Hartley, A. Zisserman, **Multiple View Geometry in Computer Vision.** Cambridge University Press, September 2000.

Pojednostavljenje određivanja položaja ravnine svjetla za vrijeme skeniranja strukturiranim svjetlom. Primjena tzv. strukturiranog svjetla (SL) je vrlo popularna metoda kod 3D rekonstrukcije. Jedna od jednostavnijih implementacija SL-a obuhvaća projiciranje uskog tamnog (svijetlog) vertikalnog »prozora« na svijetloj (tamnoj) pozadini pomoću videoprojektora. Zbog projiciranja tzv. ravnine svjetla udaraju u pojedine točke u prostoru i položaj tih ravnina je preduvjet za 3D triangulaciju dotične točke. Tradicionalni pristup triangulaciji zahtijeva i kalibraciju 3D videoprojektora. Ovaj rad predlaže metodu gdje se do položaja ravnina dolazi bez eksplicitne kalibracije videoprojektora. Usporedba tradicionalnog pristupa i predložene metode pokazala je da nema razlike po pitanju točnosti 3D rekonstrukcije. Međutim, dvije prednosti predložene metode su, kao prvo, da korisnik može jednostavnije i brže prijeći na uporabu sustava za neposrednu 3D rekonstrukciju. I kao drugo, sama implementacija sustava je jednostavnija, posebice u softverskom smislu. Konačno, demonstrirana je uspješna uporaba predloženog sustava (metode) za jednu aplikaciju u računalnoj grafici.

Ključne riječi: strukturirano svjetlo, kalibracija videoprojektora, 3D rekonstrukcija

AUTHORS' ADDRESSES

Tomislav Pribanić
Mario Cifrek
Faculty of Electrical Engineering and Computing/Department of Electronic Systems and Information Processing, Zagreb, Croatia
tomislav.pribanic@fer.hr, mario.cifrek@fer.hr,

Stanislav Peharec
Biomechanics laboratory/PEHAREC Polyclinic, Pula, Croatia
peharec@peharec.com

Received: 2006-11-08