

A STRUCTURED-LIGHT APPROACH FOR THE RECONSTRUCTION OF COMPLEX OBJECTS

Ilias KALISPERAKIS, Lazaros GRAMMATIKOPOULOS, Elli PETSAS, George KARRAS*

Department of Surveying, Technological Educational Institute of Athens (TEI-A),
GR-12210 Athens, Greece

ikal@teiath.gr, lazaros@teiath.gr, petsa@teiath.gr

* Department of Surveying, National Technical University of Athens (NTUA),
GR-15780 Athens, Greece

gkarras@central.ntua.gr

Keywords: photogrammetric scanning, 3D reconstruction, triangulation, camera calibration

Abstract:

Recently, one of the central issues in the fields of Photogrammetry, Computer Vision, Computer Graphics and Image Processing is the development of tools for the automatic reconstruction of complex 3D objects. Among various approaches, one of the most promising is Structured Light 3D scanning (SL) which combines automation and high accuracy with low cost, given the steady decrease in price of cameras and projectors. SL relies on the projection of different light patterns, by means of a video projector, on 3D object surfaces, which are recorded by one or more digital cameras. Automatic pattern identification on images allows reconstructing the shape of recorded 3D objects via triangulation of the optical rays corresponding to projector and camera pixels. Models draped with realistic photo-texture may be thus also generated, reproducing both geometry and appearance of the 3D world. In this context, subject of our research is a synthesis of state-of-the-art as well as the development of novel algorithms, in order to implement a 3D scanning system consisting, at this stage, of one consumer digital camera (DSLR) and a video projector. In the following, the main principles of structured light scanning and the algorithms implemented in our system are presented, and results are given to demonstrate the potential of such a system. Since this work is part of an ongoing research project, future tasks are also discussed.

1. INTRODUCTION

Recently, we are witness to a rapidly increasing demand for 3D content in a variety of application fields and scales, which range from city modeling [1], industrial metrology, inspection and quality control [2] or robotics [3] to 3D printing and rapid prototyping, augmented reality or entertainment. Of course, among these fields cultural heritage recording and documentation occupy an outstanding position [4] (see also cited projects [5, 6]). A common representation of 3D content is through detailed 3D digital surface models (usually in the form of point clouds or 3D surface triangular meshes), rendered with photo-texture from real imagery. Ideally, the 3D models must be generated rapidly and accurately by automatic techniques. Responding to this demand are different approaches and technologies for 3D model acquisition, typically classified in the two main categories of “passive” and “active” techniques. The former methods rely on the processing of recorded ambient radiation (usually light reflectance); they include stereovision and, more generally, image based modeling [7, 8, 9, 10], but also shape from silhouettes [11] and shape from shading [12, 13]. Active methods for 3D surface reconstruction, on the other hand, employ radiation – mainly laser or light – emitted onto the object surface and triangulated with the image optical rays. Among these approaches, most common are 3D laser scanning, single-image slit-scanning and structured light scanning.

Structured light scanning (SL), which is the topic of this paper, thus rests on the principle of triangulation of optical rays. These systems consist of a video projector which projects a sequence of specific light patterns (black-and-white or grayscale stripes, coloured line patterns, specific targets etc.), while one or more digital

cameras record the deformation of the patterns projected onto the objects (which depends on the shape of the 3D surface). Suitable encoding of the projected patterns allows establishing correspondences between camera and projector pixels. Subsequently, knowledge of the interior (perspective) and the exterior geometry of projector and cameras allows 3D surface reconstruction with triangulation of the homologous optical rays.

A main drawback of existing commercial solutions (both laser and SL scanners) is high cost. Recently, however, several DIY (Do It Yourself) systems have appeared which – utilizing means like a video camera and a simple lighting system (e.g. [14]), one or two webcams and a simple linear laser [15, 16], or a webcam and a video projector (<http://mesh.brown.edu/byo3d/source.html>) – allow 3D reconstruction of small objects at very low cost. Regarding SL scanning, publications [17, 18, 19, 20], among others, have demonstrated that consumer video projectors and off-the-shelf digital cameras can also be utilized by self-developed algorithms to create SL scanning systems, which provide accurate and reliable results, directly comparable to those from high cost commercial systems.

Thus, although several commercial structured light systems do exist, significant on-going research regarding open issues in the development and implementation of novel 3D scanning systems indicates that 3D shape reconstruction with SL methods is far from being considered as a fully solved or outdated problem. To name a few, innovations concern the form of projection patterns in order to either improve scanning accuracy [21, 22] or scanning in real time [23, 24, 25, 26 and 27]. In this context, in [28] a single pattern is used to record moving objects. Moreover, in [29, 30] the question of using un-calibrated camera-projector systems by solving simultaneously the problems of system calibration and registration of scans from different viewpoints is addressed. Characteristic for this active research activity is also the existence of the dedicated annual workshop PROCAMS (IEEE International Workshop on Projector-Camera Systems: www.procams.org).

Here, a structured-light approach for 3D reconstruction is proposed which is based on 3D triangulation of optical rays generated by a video projector and recorded by a high resolution digital camera. The system is calibrated in one step by projecting a coloured chess-board pattern on top of a targeted planar surface. The planar calibration board is rotated in space producing different perspectives captured by the camera. Both the targets and corners of the projected pattern are automatically identified on the images with sub-pixel accuracy, allowing precise simultaneous estimation of internal and external system parameters in a bundle adjustment. The scanning process is performed through projection of binary Gray coded pattern (horizontal and vertical stripes) onto the unknown 3D surface. In this way, matching of homologue projector and camera pixels can be performed without redundancy. Sub-pixel correspondences are also established to increase precision and smoothness of the final 3D reconstruction, which is obtained through standard photogrammetric space intersection. In section 2 the main components of our 3D scanner are described, while sections 3 and 4 give details regarding the algorithms for calibration and automatic establishment of pixel correspondences. The approach has been used for recording two small statues, and results are presented in section 5. Conclusions and future tasks are discussed in section 6.

2. SYSTEM DESCRIPTION

The hardware components of the system implemented for the tests of this contribution are:

- a Canon EOS 400D DSLR camera (resolution 3888×2592)
- a Mitsubishi XD600 DLP video projector (resolution 1024× 768)
- a calibration board (white non-reflective planar object with at least 4 black-and-white symmetrical targets printed with a laser printer).

During scanning, the camera–projector relative position has to be fixed. The system is flexible regarding its hardware components as it may incorporate any combination of consumer video projector and digital camera (provided they can be controlled by a personal computer). Additionally, it can be adapted to scan objects at different scales by changing the size of the calibration board and the distance between camera and projector (baseline) as well as by suitably adjusting the focus of both devices.

3. CALIBRATION

Essential step in 3D triangulation with SL systems is their calibration, i.e. the determination of the interior orientation (focal length, principal point position, lens distortion parameters) of video projector and digital camera as well as their (scaled) relative position in space. Typically, a camera-projector calibration is carried out in two separate steps. First, the camera interior orientation is estimated, and next projector interior and

relative orientations are found. In this context, [31] use a planar surface and a combination of printed circular control points and projected targets to perform plane-based calibration. Then, using the epipolar constraint, homologies between projector and camera pixels are established and the projector is calibrated. In [32] the camera is calibrated and then a full SL scanning of a planar object containing targets is performed to obtain correspondences between the projector and the camera pixels. This procedure is repeated with different orientations of the planar surface, and synthetic images of what the projector would capture as a virtual camera are computed and used for its calibration. Finally, [33] adopts the same technique of “virtual” projector images, while an alternative approach is also proposed, in which a calibrated stereo camera configuration is used to compute the 3D coordinates of a projected pattern on different planar object orientations; the acquired 3D coordinates are used in a subsequent step of projector calibration.

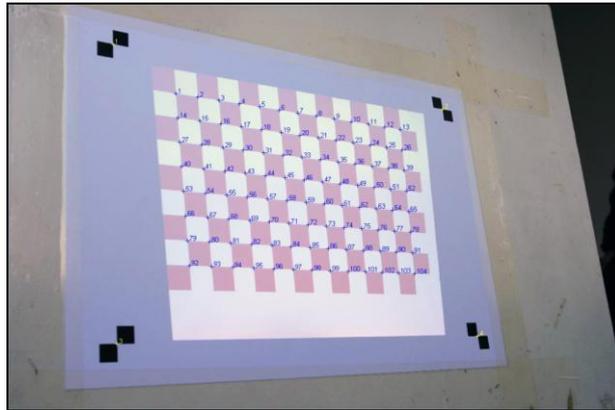


Figure 1: Typical image for camera–projector system calibration.

Here a simultaneous estimation of camera and projector calibration along with their relative orientation is proposed. The implemented algorithm includes:

- The projection of a chessboard-like colour pattern (red and white tiles) onto a planar object containing at least 4 (black and white) printed targets, and the recording of these projections by the camera (Fig. 1). This is repeated for different successive orientations of the planar surface.
- The automatic detection (with sub-pixel accuracy) of targets and corners on the imaged colour pattern.
- A bundle adjustment for optimal estimation of calibration parameters (initial values are found after [34], in which our team presented a fully automatic camera calibration toolbox, now available on the Internet).

3.1 Detection of printed targets

To detect the 4, or more, printed targets (Fig. 2, above) on the calibration images, Harris corners are extracted. Subsequently, possible areas of targets are identified by applying an intensity threshold to the three RGB channels. Detected interest areas are expanded using morphological dilation, corners outside them are discarded. However, due to the symmetrical form of the chessboard-like targets, more than one interest points are assigned to each actual target. As seen in Fig. 2 (below left), the 7 peaks of the “cornerness” measure correspond to 7 Harris points extracted at the corners and the centre of the target. What differentiate the central point from the rest are the strongly symmetrical intensity values of its neighbourhood. Here, a descriptor measuring the symmetry of a window around every image pixel is computed as the inverse of the norm of the intensity differences of anti-diametric pixels with respect to the centre of the window. In order to avoid homogeneous image regions, the descriptor values are multiplied by the local standard deviation and the “cornerness” measure. Fig. 2 (below right) presents the descriptor values for a target area. In this way the central corner (which shows the highest “symmetry descriptor”) is assigned to the target. Finally, a point is detected at this position with sub-pixel accuracy. It is noted that mirrored targets are used in the left and right sides of the planar surface to allow unique identification of target ordering under different perspectives.

3.2 Detection of projected patterns

Detection of the projected colour grid is more straightforward, as there are no severe perspective distortions in different views (horizontal lines are projected nearly horizontal). The projected red-white chess-board pattern is differentiated from the printed black and white calibration targets by a threshold in the HSV colour space. Chess-board corners are first detected by normalized cross correlation template matching with a pre-

defined pattern. In this way, by connecting and labeling pixels of high correlation values, blobs are created, and then corners are estimated with sub-pixel accuracy at the centre of gravity of each blob. Finally, nodes are brought in correspondence to the respective projected pattern nodes by means of an ordering process guided by their convex hull.

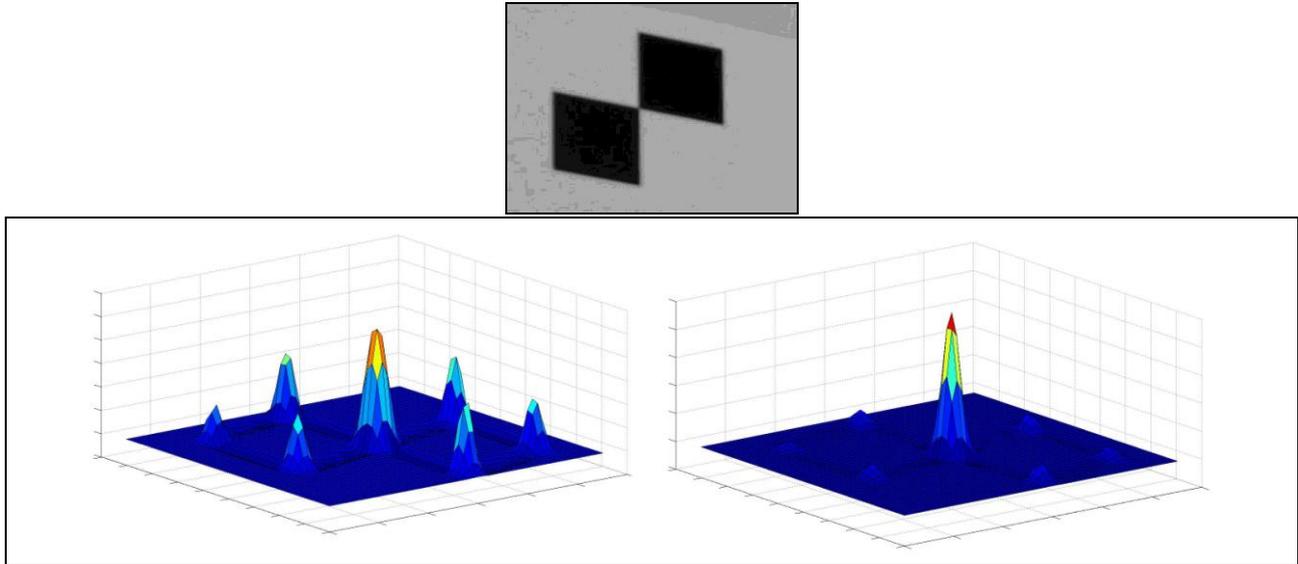


Figure 2: Calibration target (above), Harris ‘cornerness’ measure (below left), ‘symmetry descriptor’ (right).

3.3 Bundle adjustment

Through the detection procedure, point matches are established between the camera and projector frames. During the calibration process the planar object changes orientation and position in 3D space against a fixed camera-projector system. Of course, this is equivalent to a rigid body movement of the structured light system against a plane fixed in space. In this sense, the 4 targets correspond to 4 fixed points with known coordinates on a plane. The corners of the projected pattern in every successive frame correspond to different object points, all lying on the plane defined by the 4 targets. Thus, a bundle adjustment solution is feasible. Targets serve as full control points and grid corners as tie points with two unknown plane coordinates. Unknown are also the 6 camera-to-projector orientation parameters and the 10 (2×5) interior orientation parameters of the camera and the projector. In all calibration tests the standard error of the adjustment was below 0.2 pixels. Table 1 shows the calibration data for the experimental application described in Section 5.

Table 1: Calibration results (from 6 images)

$\sigma_o = \pm 0.08$ pixel		
	<i>camera</i>	<i>projector</i>
c_x (pix)	4620.85 ± 1.87	2177.01 ± 1.16
x_o (pix)	-43.90 ± 1.02	16.02 ± 0.51
y_o (pix)	3.25 ± 0.60	-462.96 ± 0.62
$k_1 (\times 10^{-09})$	-2.84 ± 0.07	-7.32 ± 0.07
$k_2 (\times 10^{-16})$	1.29 ± 0.18	-48.71 ± 2.13
<i>scaled relative orientation</i>		
B_x (cm)	30.27 ± 0.01	
B_y (cm)	-8.78 ± 0.00	
B_z (cm)	-11.40 ± 0.02	
ω ($^\circ$)	-5.01 ± 0.02	
φ ($^\circ$)	30.40 ± 0.02	
κ ($^\circ$)	-2.04 ± 0.01	

4. MATCHING CAMERA AND PROJECTOR PIXELS

Crucial step in SL systems is, of course, the establishment of correct correspondences between projector and camera pixels, since obviously the accuracy of this matching procedure affects directly the accuracy of 3D

reconstruction. Our approach implemented so far is based on [20] and uses successive projections of binary Gray-code patterns (Fig. 3), i.e. black-and-white vertical and horizontal stripes of variable width (see Fig. 4). Each projection is recorded by the camera, and dark and light areas are identified on the image. Since each projector pixel is characterized by a unique sequence of black and white values, identification of the sequence of dark and light values for each camera pixel directly allows establishing camera–projector pixel homologies. In particular, $\log_2(n)$ patterns are required to uniquely model n different labels. Thus, for a 1024×768 projector 10 patterns are needed for each direction. In order to determine whether an image pixel corresponds to a dark or a light projected area in a more robust way, the inverse of each Gray-code pattern is also projected. Consequently, a total of 40 different patterns are used. A pixel is characterized as illuminated with white colour from a specific pattern if the difference of its intensity values corresponding to successive normal and negative patterns is positive. The rest of the pixels are assigned to dark values (Fig. 5). Finally, pixels with absolute differences less than a threshold (e.g. 4 gray values) can be rejected as outliers.

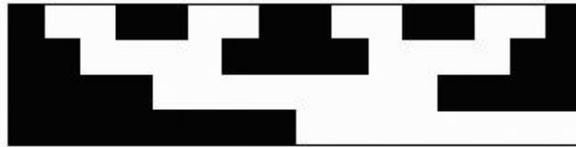


Figure 3: Binary Gray code.



Figure 4: Examples of black and white vertical and horizontal projection patterns.

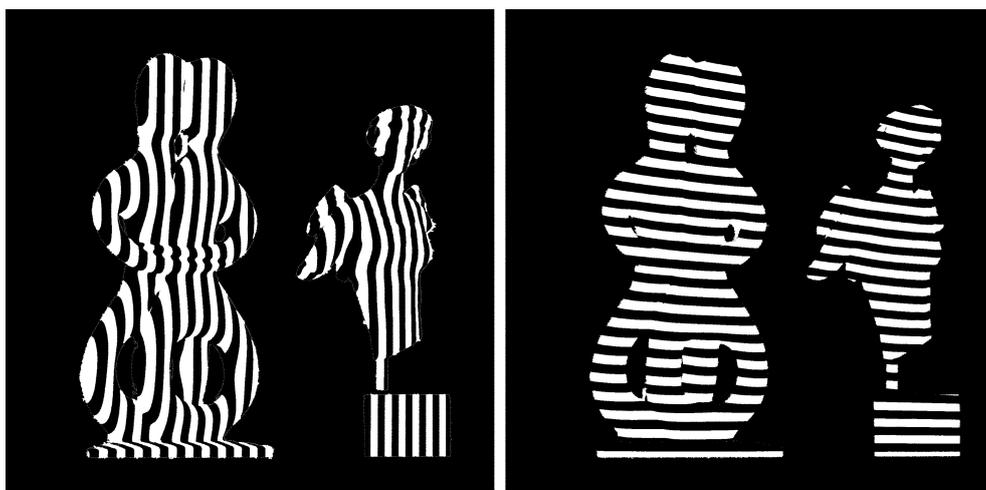


Figure 5: Detection of illuminated and dark areas.

Due to differences in the camera and projector resolutions (cameras have usually higher resolution) several camera pixels may be assigned to the same projector pixel. This results in 3D reconstructions with discrete steps and strong moiré-like artifacts (Fig. 6, left). Thus, to obtain more accurate and smooth 3D reconstruc-

tions each camera pixel must be associated with a unique sub-pixel point on the projector frame. Different approaches to obtain such sub-pixel correspondences exist in literature (for a taxonomy of state-of-the-art methods see [35]). Here, we have adopted the approach of [20] who, after establishing correspondences at pixel level, interpolate the integer projector coordinate values by means of a 1D averaging filter (7×1 pixels) in the prominent pattern direction. In our implementation 2D orthogonal convolution windows (11×7 , 15×7) are used for averaging in order to obtain smoother results and consistency among different scan-lines (Fig. 6, centre and right). Once pixel (or sub-pixel) matches are established, the 3D position of depicted object points is computed through simple triangulation of the corresponding optical rays. Colour values of these points are also directly available from the camera images; thus a 3D coloured point cloud can be reconstructed (Fig. 7).



Figure 6: Reconstruction improvement by sub-pixel correspondences.



Figure 7: Examples of coloured 3D point clouds.

5. EXPERIMENTAL RESULTS

To demonstrate the effectiveness of the SL system implemented to this stage, a scanning of two small statues (~ 20 cm and ~ 12 cm in height) was performed. In each case 11 separate scans were carried out, and the reconstructed point clouds were registered with respective RMS deviations of $60 \mu\text{m}$ and $80 \mu\text{m}$. Finally, 3D mesh models were created for each object, seen in Figs. 8 and 9.

6. FUTURE TASKS

In this contribution a realization of a structured light system with consumer hardware components (a video projector and a digital camera) was described. Algorithmic details were discussed regarding system calibration and matching of pixels among projector and camera frames. Experimental results were also presented showing the potential of such low-cost systems. Innovations of the proposed approach can be found in the calibration process which is performed in one step for camera and projector, considerably simplifying the use

of such a system. Future tasks of our ongoing research include introduction of a second camera; investigation of alternative methods for obtaining sub-pixel accuracy; automatic detection of occlusions, hole-filling and combination of structured light techniques with dense-stereo matching algorithms in search for higher accuracy; automatic registration of scans acquired from different viewpoints for obtaining full 3D representation of scanned objects; and, finally, investigation of the potential of un-calibrated camera-projector systems.



Figure 8: Different views of the first rendered 3D model (obtained from 11 scans).

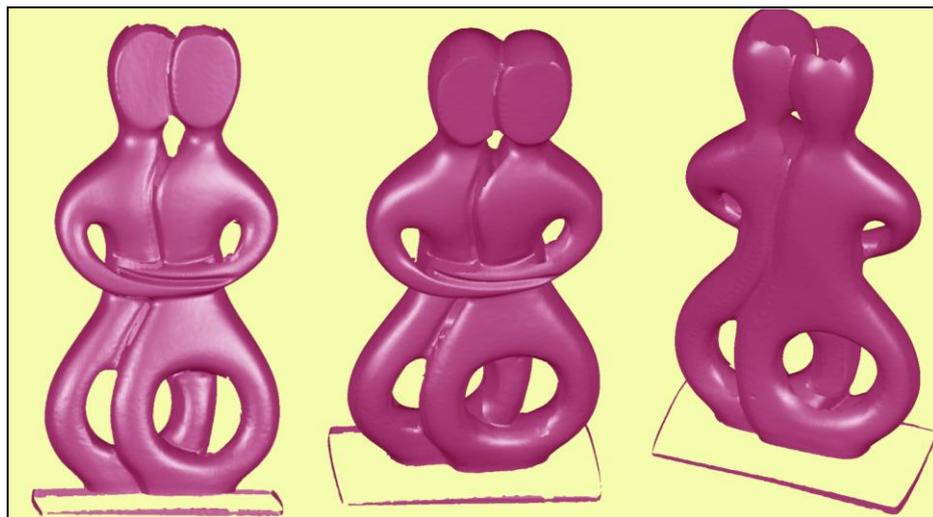


Figure 9: Different views of the second rendered 3D model (obtained from 11 scans).

7. REFERENCES

- [1] Xiao, J., Fang, T., Zhao, P., Lhuillier, M., Quan, L.: *Image-based street-side city modeling*, ACM Transaction on Graphics (TOG), 28(2009) 5.
- [2] Teutsch, C.: *Model-based Analysis and Evaluation of Point Sets from Optical 3D Laser Scanners*, PhD Thesis, Magdeburger Schriften zur Visualisierung, Shaker Verlag, 2007.
- [3] Claes, K., Bruyninckx, H.: *Robot positioning using structured light patterns suitable for self calibration and 3D tracking*, Proceedings of the 2007 International Conference on Advanced Robotics, Jeju, Korea.
- [4] Cignoni, P., Scopigno, R.: *Sampled 3D models for CH applications: A viable and enabling new medium or just a technological exercise?* ACM Journal on Computing and Cultural Heritage, 1(2008)1.
- [5] CyArk Project. <http://archive.cyark.org/about>
- [6] 3D-COFORM. <http://www.3dcoform.eu/>
- [7] Seitz, S.M., Curless, B., Diebel, J., Scharstein, D., Szeliski, R.: *A comparison and evaluation of multi-view stereo reconstruction algorithms*, Proceedings CVPR, 1(2006), 519-528.
- [8] Strecha, C., Fransens, R., van Gool, L.: *Combined depth and outlier estimation in multi-view stereo*,

Proceedings CVPR, 2(2006), 2394- 2401.

- [9] Furukawa, Y., Ponce, J.: *Accurate, dense, and robust multi-view stereopsis*, Proc. CVPR, 2007, 1-8.
- [10] Vu, H.H., Keriven, R., Labatut, P., Pons, J.-P.: *Towards high-resolution large-scale multi-view stereo*, Proc. CVPR, 2009.
- [11] Mercier, B., Meneveaux, D.: *Shape from silhouette: Image pixels for marching cubes*, Journal of WSCG, 13(2005), 112-118.
- [12] Prados, E., Faugeras, O.: *Perspective shape from shading and viscosity solutions*, Proc. 9th IEEE ICCV, vol. II, Nice, France, October 2003, 826-831.
- [13] Tankus, A., Sochen, N., Yeshurun, Y.: *A new perspective on shape from-shading*, Proc. 9th IEEE ICCV, vol. II, Nice, France, October 2003, 862-869.
- [14] Bouguet, J.-Y., Perona, P.: *3D photography on your desk*, Proc. ICCV 1998, 43-50.
- [15] Winkelbach, S., Molkenstruck, S., Wahl, F.M.: *Low-cost laser range scanner and fast surface registration approach*, Proc. DAGM '06, Lecture Notes in Computer Science, vol. 4174, Springer, 2006, 718-728.
- [16] Prokos, A., Karras, G., Petsa, E.: *Automatic 3D surface reconstruction by combining stereovision with the slit-scanner approach*, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII, Part 5, 2010, 505-509.
- [17] Gühring, J.: *Dense 3-D surface acquisition by structured light using off-the-shelf components*, Videometrics and Optical Methods for 3D Shape Measurement, 2001, 220-231.
- [18] Rocchini, C., Cignoni, P., Montani, C., Pingi, P., Scopigno, R.: *A low cost 3D scanner based on structured light*, EUROGRAPHICS 2001, 20 (2001)3.
- [19] Tchou, C.: *Image-Based Models: Geometry and Reflectance Acquisition Systems*, UC Berkeley, M.Sc. Thesis, 2002.
- [20] Scharstein, D., Szeliski, R.: *High-accuracy stereo depth maps using structured light*, Proc. IEEE CVPR, 1(2003), 195-202.
- [21] Fechteler, P., Eisert, P., Rurainsky, J.: *Fast and high resolution 3D face Scanning*, Proc. ICIP, 2007.
- [22] Peng, T., Gupta, S.K.: *Model and algorithms for point cloud construction using digital projection patterns*, Journal of Computing and Information Science in Engineering, 7(2007)4, 372-381.
- [23] Koninckx, T., Griesser, A., van Gool, L.: *Real-time range scanning of deformable surfaces by adaptively coded structured light*, In: 3-D Digital Imaging & Modelling, Banff Canada, 2003, 293-300.
- [24] Peisen, S., Zhang, H.S.: *Fast three-step phase-shifting algorithm*, Applied Optics, 45(2006)21, 5086-5091.
- [25] Chen, S.Y., Li, Y.F., Zhang, J.: *Vision processing for realtime 3-D data acquisition based on coded structured light*, IEEE Transactions on Image Processing, 17(2008)2, 167-176.
- [26] Liu, K., Wang, Y., Lau, D.L., Hao, Q., Hassebrook, L.G.: *Dual-frequency pattern scheme for high-speed 3-D shape measurement*, Optical Express, 18(2010), 5229-5244.
- [27] Zhang, S.: *Recent progresses on real-time 3-d shape measurement using digital fringe projection techniques*, Optics and Lasers in Engineering, 40(2010), 149-158.
- [28] Schmalz, C., Angelopoulou, E.: *A graph-based approach for robust single-shot structured light*, IEEE International Workshop on Projector-Camera Systems, 2010.
- [29] Furukawa, R., Kawasaki, H.: *Uncalibrated multiple image stereo system with arbitrarily movable camera and projector for wide range scanning*, Proc. 5th 3DIM '05, 2005, 302-309.
- [30] Aliaga, D., Xu, Y.: *Photogeometric structured light: a self-calibrating and multi-viewpoint framework for accurate 3D modeling*, Proc. IEEE CVPR, 2008.
- [31] Gao, W., Wang, L., Hu, Z.: *Flexible calibration of a portable structured light system through surface plane*, Acta Automatica Sinica, 34(2008)11, 1358-1362.
- [32] Zhang, S., Huang, P.S.: *Novel method for structured light system calibration*, Optical Engineering 45(2006)8.
- [33] Knyaz, K.L.: *Multi-media projector – single camera photogrammetric system for fast 3D reconstruction*, International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVIII, Part 5(2010), 343-348.
- [34] Douskos, V., Grammatikopoulos, L., Kalisperakis, I., Karras, G., Petsa, E.: *FAUCCAL: An open source toolbox for fully automatic camera calibration*. XXII CIPA Symposium on Digital Documentation, Interpretation & Presentation of Cultural Heritage, Kyoto, Japan, October 2009.
- [35] Salvi, J., Fernandez, S., Pribanic, T., Llado, X.: *A state of the art in structured light patterns for surface profilometry*, Pattern Recognition, 43(2010)8, 2666-2680.