



HAL
open science

3D Structured Light Scanner on the Smartphone

Tomislav Pribanić, Tomislav Petković, Matea Donlić, Vincent Angladon,
Simone Gasparini

► **To cite this version:**

Tomislav Pribanić, Tomislav Petković, Matea Donlić, Vincent Angladon, Simone Gasparini. 3D Structured Light Scanner on the Smartphone. Proceedings of the 13th International Conference on Image Analysis and Recognition, ICIAR 2016, Jul 2016, Póvoa de Varzim, Portugal. pp.443 - 450, 10.1007/978-3-319-41501-7_50 . hal-01420732

HAL Id: hal-01420732

<https://hal.science/hal-01420732v1>

Submitted on 20 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

3D Structured Light Scanner on the Smartphone

Tomislav Pribanić¹, Tomislav Petković¹, Matea Đonlić¹, Vincent Angladon² and Simone Gasparini²

¹University of Zagreb, Faculty of Electrical Engineering and Computing, Zagreb, Croatia
{tomislav.pribanic, tomislav.petkovic.jr, matea.donlic}@fer.hr

²University of Toulouse, IRIT-INP, Toulouse, France
{vincent.angladon, simone.gasparini}@irit.fr

Abstract. In the recent years turning smartphones into 3D reconstruction devices has been greatly investigated. Different 3D reconstruction concepts have been proposed, and one of the most popular is based on IR projection of a pseudorandom dots (speckle) pattern. We demonstrate our idea how a pseudorandom dots pattern can be used and we also present an active approach applying a structured light (SL) scanning on the smartphone. SL has a number of advantages compared to other 3D reconstruction concepts and likewise our smartphone implementation inherits the same advantages compared to other smartphone based solutions. The shown qualitative and quantitative results demonstrate the comparable outcome with the standard type SL scanner.

Keywords: smartphone, 3D reconstruction, structured light, pseudorandom dots pattern

1 Introduction

Apple's iPhone started the era of modern smartphones in 2007. Other smartphone models soon followed causing the recognition of a smartphone as a visual computing powerhouse. A modern smartphone has a high-speed multi-core CPU, a 3D graphic processor, a DSP for image and video processing, a high resolution camera, a high quality color display, and quite impressive local storage capabilities. Therefore, turning a smartphone into a powerful 3D reconstruction device opens additional application and research avenues which are beyond a simple gadget, e.g. reverse engineering (digitization of complex, free-form surfaces), object recognition, 3D map building, biometrics, clothing design, and others.

3D surface reconstruction methods applicable to smartphones may be categorized into passive and active methods. Most common passive 3D reconstruction solutions require the user to go around the object carefully taking a relatively large number of images ([1], [2]) that are processed in a SLAM pipeline to extract a 3D shape. Image processing of such approaches (e.g. extracting image features and matching them across the images) is a central part and it is to a large extent typically done in the cloud requiring a network connection to upload the acquired images. On the other hand, complete on board smartphone solutions heavily rely on the additional sensors

such as accelerometer and gyroscope [3]. Alternatively, the shape from silhouettes has also been proposed, still creating relatively coarse 3D models of small-scale objects [4].

In the context of active stereo, there are solutions proposing the photometric stereo where the smartphone screen is conveniently used as a light source, however, as noted by the authors themselves, a dark environment is required ([5], [6]). Somewhat more robust solution, but at the expense of using an extra smartphone, is proposed in [7]. They used a pair of smartphones which collaborate as master and slave: the slave was illuminating the scene using the flash from appropriate viewing points while the master recorded images of the object. Considering such approaches, it become apparent that the lack of an appropriate light source on the smartphone is a substantial obstacle for implementing any form of active stereo. Project Tango by Google is perhaps one of the most well-known examples where, in order to overcome that obstacle, a custom made IR projector and IR camera are installed in a smartphone [8]. Unfortunately, added IR projector may be used only for 3D depth sensing and not much beyond. Finally, some of the more recent work proposed the use of a laser line projector attached to a smartphone [9]. Although conceptually simple, such solution, similarly to all single line laser approaches, requires many images for 3D surface reconstruction making it necessary to have an effective 3D registration tool to combine single line reconstructions. The authors in [9] impose a constraint that a marker has to be visible and tracked throughout the frames.

To the best of our knowledge no smartphone-based solution considered a well-established and powerful concept for 3D shape acquisition, a structured light (SL) strategy [10]. Briefly, the SL concept involves the use of a camera-projector pair where a pattern, designed to contain a certain code, is projected and the corresponding code is then identified on the camera image(s). Through found camera-projector pixels correspondences a triangulation of 3D points is carried out. During this work we have identified a number of commercially available smartphones [11], which have an embedded pico projector. In particular we have used the Samsung Galaxy Beam smartphone [12]. Our main contribution is demonstration how such smartphones can be successfully turned into very efficient 3D scanning devices. Interestingly, Samsung (including other manufactures as well) have not considered the use of a camera-projector pair for SL scanning, since apparently smartphone's camera and pico projector typically do not share a common field of view (FOV). To redirect light rays, complex configurations of mirrors have been extensively used in all kinds of imaging systems [13]. As an additional contribution we propose the use of a simple adapter with a first surface mirror which, as will be shown, neatly resolves the FOV issue. We demonstrate the use of the proposed system using one of the most popular and robust SL scanning strategies, multiple phase shifting (MPS) [14]. We also show the implementation of a pseudorandom dots pattern for projection, basically the same type of pattern has been used in the first version of globally popular Microsoft Kinect for Xbox 360 (Kinect v1).

2 Method

We first describe the used hardware in subsection 2.1, then the imaging geometry in subsection 2.2., and finally, we describe used SL approach in subsection 2.3.

2.1 Hardware Components

Fig. 1 shows the Samsung Galaxy Beam smartphone which has a 5MP camera and a 15 lumen DLP projector. Unfortunately, camera and projector have no common FOV. To overcome this difficulty we propose using a small re-attachable adapter with a first surface mirror as shown in Fig. 2. This allows defining the geometrical and computational framework for the 3D surface scanning between a camera and a projector, as described below.



Fig. 1. Samsung Galaxy Beam: a camera and an embedded pico projector comprise an angle of 90° , having no common FOV.



Fig. 2. Smartphone on a tripod with the deflection adapter for projector attached. Note the mirror image of the projector, i.e. a virtual projector.

2.2 Camera-Projector Imaging Geometry

Fig. 3 represents a cross-section view of a smartphone positioned sideways and upgraded with a first surface planar mirror placed in the front of the projector. Smartphone's projector P_R (real) and camera C have their respective FOVs denoted with green lines which evidently do not intersect. We are interested in reconstructing a point A which is within camera's FOV, but not within the projector's FOV. The red dashed line joining the projector P_R and the point A represents a hypothetical pattern projection on A which obviously cannot happen due to insufficient FOV of the projector. However, consider a planar mirror positioned at an angle α with respect to the optical axis of the projector P_R as shown in Fig. 3. This will create a virtual projector P_V on the opposite side of a mirror, which can project a SL pattern on the point A with the corresponding projector coordinate p . In addition, the effect of mirroring will provide a mirror image A_m of the point A , which is not visible to the virtual projector P_V . The real projector P_R can be related to the point A_m through a pattern projection with the corresponding projector coordinate p_m .

Without loss of generality, let us assume that the world frame axes x_w and y_w coincide with the mirror edges, as shown in Fig. 3, and with the third axis z_w defined by the right hand rule. This allows expressing a simple spatial position relationship between \mathbf{A} and \mathbf{A}_m , i.e. they differ only in the sign of y_w coordinate. Similarly, optical centers \mathbf{P}_R of the real projector and \mathbf{P}_V of the virtual projector differ only in the sign of y_w coordinate.

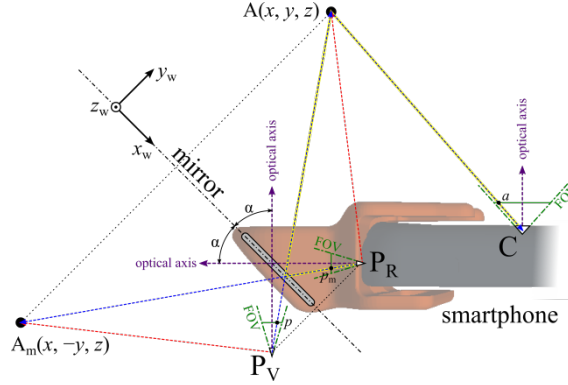


Fig. 3. Representation of 3D structured light scanner comprised of smartphone embedded camera and projector, and a planar first surface mirror (see text for more details).

The described imaging geometry can be calibrated using any standard calibration technique for 3D scanner (camera/projector) calibration. In particular, we have adopted an approach from [15].

2.3 3D Reconstruction using Structured Light

The basic principle of SL approach can be summarized as follows: a projector projects a certain number of images on the object of interest. The projected images have a particular structure, a code, which can be decoded in the acquired camera images. 3D position can be triangulated from the decoded SL code.

Among more than a dozen different SL patterns we have chosen one of the time multiplexing strategies, a well-known phase shifting (PS) method [10]. PS consists of projecting a number ($N \geq 3$) of periodic sine patterns, shifted by some period amount. The patterns are sequentially projected with a projector on the object of interest, are recorded by the camera, and are then processed in order to compute a wrapped phase map. Due to the periodic nature of sine patterns, the wrapped phase map does not provide a unique code, rather it is said that the code is wrapped within $[-\pi, +\pi]$ interval. One way to unwrap the wrapped phase map and recover the SL code is to project additional PS patterns having a different number of periods compared to the first set. Such multiple phase shifting (MPS) procedure provides two wrapped values φ_{w1} and φ_{w2} . Computing the unwrapped phase Φ_{UW} from the φ_{w1} and φ_{w2} and extracting the SL code can be done in number of different ways; we have followed the algorithm described in [14]. In brief, the algorithm in [14] emphasizes the fact that the unwrapped phase can be computed using either of φ_{R1} and φ_{R2} as:

$$\Phi_{UW} = k_1 \cdot \lambda_1 + \varphi_{w1} = k_2 \cdot \lambda_2 + \varphi_{w2} \quad (1)$$

where λ_1 and λ_2 are wave lengths corresponding to the number of periods of the first and second sine pattern, respectively, and where k_1 and k_2 are integers of full sine periods needed to reach the same unwrapped value using the wrapped values φ_{w1} and φ_{w2} , respectively. In a nutshell, the algorithm of [14] computes first all feasible pairs of (k_1, k_2) given some chosen values for λ_1 and λ_2 . Next, it chooses a specific pair (k_1, k_2) that yields the smallest discrepancy when computing the unwrapped phase Φ_{UW} using Eq. (1).

Multiple-shot methods (like PS method) produce superior reconstruction than single-shot methods in terms of resolution and accuracy but have a substantially longer acquisition times and are mostly unsuitable for dynamic scenes without using an expensive high-speed hardware. To tackle this issue, we also describe our proposal, based on the projection of a Kinect-like pseudorandom dots pattern, for extracting a depth map from a single shot. During the scanning procedure, a single image of the pseudorandom dots pattern is continuously projected using smartphone's projector and a deflection adapter, and a video sequence of the illuminated moving object is acquired using smartphone's camera. Then, each frame of the captured video sequence is matched with the reference image of the pseudorandom dots pattern. Due to limited space we omit here details how to extract SL code from a random dots pattern, instead we refer the interested reader to one alternative such as [16].

3 Results

3.1 Application Example: On the construction of a shoe insole

To demonstrate the applicability of the proposed smartphone scanner we have examined one practical application. During a shoe insole production it is customary to ask an individual to step in a special type of foam, leaving his/her imprint of the sole (foot, Fig. 4 a)). Next, the imprint is 3D scanned and, if needed, modified by the physician before it is carved out by a CAM milling machine. We note that a general purpose SL scanner or even a particularly designed foot scanner is normally used to

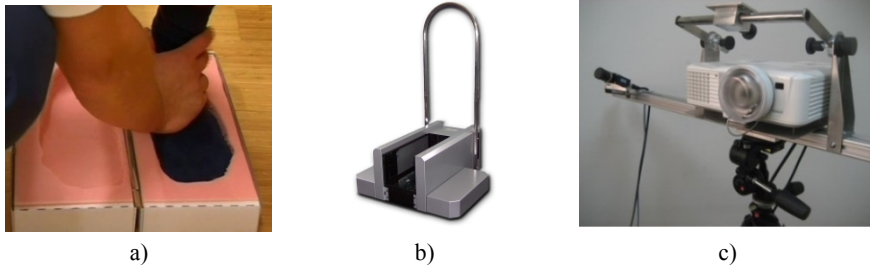


Fig. 4. a) Taking patient's sole imprint. b) Example of specially designed foot scanner; may be used only for foot scanning. c) Standard scanner consisting of industrial camera and projector fastened on a tripod; may be used as a general purpose scanning device.

scan such imprint (Fig. 4 b) and c)). Therefore we have scanned the same imprint using a type of SL scanner as shown in Fig. 4 c) and using our 3D smartphone scanner. Due to a typically higher demand for precision and accuracy in such applications, in this particular experiment we have used the MPS method [14].

We have compared the 3D point clouds of the proposed method and a standard scanner. For a qualitative comparison, Fig. 5 d) strongly suggests that two clouds of points overlap nicely, after being registered to the same coordinate system. In addition, to provide a quantitative measure of agreement between registered point clouds, we provide the absolute mean distance between corresponding points from the various registered views at the final stage of the registration process. We have considered two cases: first when the smartphone scanner point cloud was registered to the standard scanner and second when the standard scanner point cloud was registered to the smartphone scanner (Table 1). Figures in the table strongly indicate that the 3D output of two types of scanners is basically the same.

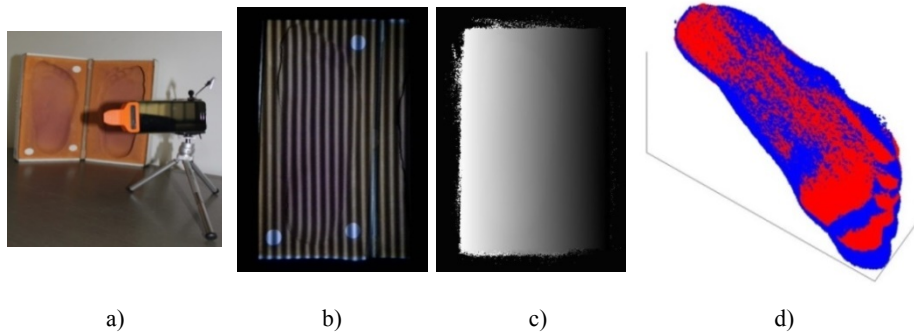


Fig. 5. a) The proposed 3D system during 3D scanning. b) Projected sine pattern during MPS. c) Recovered unwrapped phase map. d) Shape of the foot sole. Two clouds of points, red and blue, initially reconstructed from two different types of scanners and afterwards registered in the common coordinate system. Both clouds successfully overlap.

Table 1. Absolute mean distances (expressed in millimeters) between two corresponding points (nearest neighbor) from the various registered views.

	View 1	View 2	View 3	View 4	View 5
smartphone to standard scanner (mm)	0.57	0.54	0.50	0.67	0.57
standard scanner to smartphone (mm)	0.53	0.51	0.52	0.62	0.61

3.2 Moving Female Face

Projecting a more than one pattern, as explained in the previous section, typically requires an object to be still. To show feasibility of our solution for moving objects too, we demonstrate solution of projecting a single image pattern. For this experiment we have recorded a female moving her head under the proposed pseudorandom dots pattern. Fig. 6 shows four selected frames from this recording. The input sequence

only contains images with the pseudorandom dots pattern making it impossible to retrieve a speckle-free texture.



Fig. 6. Four frames from a movie of a moving female face. 3D reconstruction was performed using the pseudorandom dots pattern. Rows contain: a) input frames; b) textured 3D surface where input image is used as texture; c) 3D surface textured as depth map.

4 Discussion and Conclusion

The shown results clearly demonstrate that a smartphone with an embedded pico projector can be turned into a very powerful 3D SL scanner. We have successfully implemented one of MPS variants. We note that PS is generally regarded as state of the art strategy when it comes to the highest demands in 3D SL scanning of static objects. In addition, acknowledging in recent years an increasing number of widely affordable 3D devices based on projecting laser speckle pattern, such as Kinect v1, we have also presented an approach using a pseudorandom dots pattern on a smartphone. Since it projects a single pattern, such algorithm is capable of scanning moving object too, as shown on Fig. 6, in the case of moving female model.

One of the everlasting challenges in SL is constructing a so called hybrid pattern [17]. The hybrid pattern is expected to be suitable for reconstruction of both static and moving objects within the same captured frame where a different (de)coding is applied depending on object movement thus allowing higher quality 3D data for static objects. Inspired by some of the latest research in that respect ([18], [19]), our future work will be directed towards the design and implementation of hybrid patterns on the proposed smartphone system.

Acknowledgment

This work has been supported in parts by the Croatian Science Foundation's funding of the project IP-11-2013-3717. We also acknowledge the support of Croatian-French Program "Cogito", Hubert Curien partnership, funding the project "Three-dimensional reconstruction using smartphone".

References

1. 123D Catch. <http://www.123dapp.com/catch>. Access date: November 2015
2. Trnio. <http://www.trnio.com>. Access date: November 2015
3. Tanskanen, P., Kolev, K., Meier, L., Camposeco, F., Saurer, O., Pollefeys, M.: Live Metric 3D Reconstruction on Mobile Phones. In IEEE ICCV 2013, pp. 65–72 (2013)
4. Hartl, A., Gruber, L., Arth, C., Hauswiesner, S., Schmalstieg, D.: Rapid Reconstruction of Small Objects on Mobile Phones. In IEEE Conf. CVPR Workshops 2011, pp.20–27 (2011)
5. Wang, C., Bao, M., Shen, T.: 3D model reconstruction algorithm and implementation based on the mobile device. *J. Theor. Appl. Inf. Technol.*, vol. 46(1), pp. 255–262, (2012)
6. Trimensional. <http://www.trimensional.com>. Access date: November 2015
7. Won, J. H., Lee, M. H., Park, I. K.: Active 3D shape acquisition using smartphones. In IEEE Conf. CVPR Workshops 2012, pp. 29–34 (2012)
8. Project Tango. <https://www.ifixit.com/Teardown/Project+Tango+Teardown/23835>. Access date: November 2015
9. Slossberg, R., Wetzler, A., Kimmel, R.: Freehand Laser Scanning Using Mobile Phone. *Proc. of the British Machine Vision Conference*, pp. 88.1–88.10, (2015)
10. Salvi, J., Fernandez, S., Pribanić, T., LLado, X.: A state of the art in structured light patterns for surface profilometry. *Pattern Recognition*, vol. 43, pp. 2666–2680 (2010)
11. List of projector phones. https://en.wikipedia.org/wiki/Projector_phone. Access date: November 2015
12. Samsung Galaxy Beam. <http://www.samsung.com/global/microsite/galaxybeam/feature.html>. Access date: November 2015
13. Reshetouski, I., Ihrke, I.: Mirrors in Computer Graphics, Computer Vision and Time-of-Flight Imaging. In *Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications. Lecture Notes in Computer Science*, vol. 8200, pp. 77–104 (2013)
14. Pribanić, T., Mrvoš, S., Salvi, J.: Efficient multiple phase shift patterns for dense 3D acquisition in structured light scanning. *Image and Vision Computing*, vol. 28, pp. 1255–1266 (2010)
15. Zhang, Z.: A Flexible New Technique for Camera Calibration. *IEEE Transactions PAMI*, vol. 22(11), pp. 1330–1334 (2000)
16. Mellroy, P., Izadi, S., Fitzgibbon, A.: Kinectrack: Agile 6-DoF tracking using a projected dot pattern. *International Symposium on Mixed and Augmented Reality*, pp. 23–29 (2012)
17. Ishii, I., Yamamoto, K., Doi, K., Tsuji, T.: High-speed 3D image acquisition using coded structured light projection. In *Intelligent Robots and Systems, IEEE/RSJ Int. Conf. on*, pp. 925–930 (2007)
18. Zhang, Y., Xiong, Z., Yang, Z., Wu, F.: Real-Time Scalable Depth Sensing with Hybrid Structured Light Illumination. *IEEE Trans. on Image Processing*, vol.23,pp.97–109, (2104)
19. Petković, T., Pribanić, T., Đonlić, M.: The Self-Equalizing De Bruijn Sequence for 3D Profilometry. *Proc. of the British Machine Vision Conference*, pp. 155.1–155.11, (2015)