STAG: Smart Tools and Apps in computer Graphics (2016)
Giovanni Pintore and Filippo Stanco (Editors)

Low Cost Handheld 3D Scanning for Architectural Elements Acquisition

D. Allegra[1], G. Gallo[1], L. Inzerillo[2], M. Lombardo[3], F.L.M. Milotta[1], C. Santagati[4] and F. Stanco[1]

[1] University of Catania, Department of Mathematics and Computer Science, Italy
{allegra, gallo, milotta, fstanco}@dmi.unict.it
[2] University of Palermo, Department of Architecture, Italy
laura.inzerillo71@gmail.com
[3] University of Catania
graphics.marc@gmail.com
[4] University of Catania, Department of Civil Engineering and Architecture, Italy
cettinasantagati2@gmail.com

Abstract

3D scanning has gone a long way since its first appearance in cultural heritage digitization and modeling. In the recent years some new low cost, fast, accurate emerging technologies are flooding the market. Envisioning the massive use of these cheap and easy to use devices in the next years, it is crucial to explore the possible fields of application and to test their effectiveness in terms of easiness of 3D data collection, processing, mesh resolution and metric accuracy against the size and features of the objects. In this study we focus the attention on one emerging technology, the Structure Sensor device, in order to verify a 3D pipeline acquisition on an architectural element and its details. The methodological approach is thought to define a pipeline of 3D acquisition exploiting low cost and open source technologies and foresees the assessment of this procedure in comparison with data obtained by a Time of Flight device.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Line and curve generation

1. Introduction

In the last years some new low cost emerging technologies have been released on the market delivering a long term dream of the practitioner of cultural heritage: fast, accurate, low cost 3D scanning with a handheld device. Envisioning the massive use of these cheap and easy to use devices in the next years, it is crucial to explore the possible fields of application thus testing their effectiveness in terms of easiness of 3D data collection, processing, mesh resolution and metric accuracy against the size and features of the objects. In this study we focus the attention on one emerging technology, the Structure Sensor device [str], in order to verify a 3D pipeline acquisition on an architectural element and its details. As case study we choose the XVIII century doorway placed in the monastery of Benedettini in Catania, in UNESCO's world heritage list. The doorway presents both planar, complex (mouldings) and sculpted surfaces and allow us to carry out several tests on different geometries. The goal is to outline a 3D pipeline following as much as possible, a low cost and open source workflow from 3D data collecting to the digital replica.

The methodological approach foresees the assessment of the 3D acquisition procedure in comparison with data obtained by a Time of Flight device in order to point out weaknesses and advantages of the hand held scanning approach in relation to other well assessed technology. Then 3D modeling issues are explored and discussed to obtain a digital replica in an open source environment suitable for architectural representation and communication purposes.

The remainder of the paper is structured as follows: Section 2 will illustrate an overview on Handheld 3D scanning technologies. Section 3 will deal with the tests on the case study with some specifications on the employed device and method description from 3D acquisition to alignment and metric accuracy verification. Then, discussion on results is reported in Section 4, while Conclusion will complete the paper.

2. Handheld 3D scanning

The 3D scanners are devices which are able to collect geometry information about a real-world object or environment. Then these information are processed in order to build a digital 3D models of the scanned elements. Nowadays, 3D scanning devices play a key

DOI: 10.2312/stag.20161372

role in many research field and applications such as industrial, prosthetics and medicine prototyping, cultural heritage preservation and documentation, etc. [ACP*10, GMS*10, STBB11, STG*12, ST13, SIP13, Rem11, BBG11, BGF10]. Since these devices works by employing many different technologies and their cost change in a wide price range, it is important to select the best solution for your own applications.

The most common technologies employed for 3D scanning are triangulation (e.g., laser triangulation or structured light) and Time of Flight (ToF).

Sensors that exploit these technologies belong to the class of the so-called active sensors. Indeed, these devices "emit" electromagnetic waves on the objects to estimates their geometrical properties. On the other hand, sensors which do not introduce waves in the environment are called passive sensors. In this latter case, the 3D acquisition could be achieved for instance by stereo vision or structure from motion [Cur].

From another perspective the 3D scanning devices can be categorized in respect to their portability. In recent years, thanks to the miniaturization and integration of the electronic and optical sensors, has been possible to produce small and compact high performance 3D scanners [KTSP14, SSHP15, PLB, VM]. Hence, we may further distinguish two kinds of devices: handheld scanners and not portable ones. Today, the emerging handheld scanners are a remarkable resource for affordable price and good performance and the convenience ensured by the portability. For the relative low-cost and usability, most of these devices became consumer electronics product, while other are still used in professional context. However, they represent a great resource in the field of Cultural heritage. Below, we report a list of the main handheld 3D scanners and their knows specs in Table 1:

Microsoft Kinect is mainly used in home videogames entertainment. However, some examples of applications of these devices to cultural heritage could be found in the works of Cappelletto [CZC16] and Remondino [Rem11]. **Scanify Fuel 3D** is a handheld device, which exploits combination of photometric and stereography techniques to acquire depth information, so it can reach a high accuracy. **Google Project Tango** is a Google device with exploits motion tracking to understand position and orientation of the device user. It is particularly suitable for augmented reality application. **Artec Eva and Artec Spider** are two semi-professional active scanners produced by Artec 3D company. The first one has a high resolution and it is suitable for small and detailed object, while Artec Eva is though for architectural elements such as doors, statue etc. An example of Cultural heritage application can be find in [AAA*15]. **Structure Sensor** is a small active scanner produced by Occipital. It exploits structured light technology to guarantee a good quality scan with a low expense. This device has been employed in the study conducted on this paper, hence more details are provided in the following Sections.

3. Case study: structure sensor scanning for architectural elements

In most cases the use of handheld scanners is limited to small objects (approximately a volume of $1m^3$), if there is the need to acquire bigger objects, then it is necessary to carry out several scans and then align them in an unique model. Thus, in architectural heritage field the use of this kind of scanner should be recommended only for architectural details (basis, capitals, pedestals). Nevertheless in our study we explore the possibility of using Structure Sensor also for bigger architectural element such us a doorway. Our goal is to provide a full low cost and open source 3D pipeline highlighting potentialities and weakness. We choose an eighteen century doorway in Benedettini monumental complex in Catania (UNESCO heritage) located in the gallery at the first floor of the monastery and it provides access to one of the cells of the friars, nowadays used as offices for the Department of Humanities of Catania University. This doorway, realized with limestone, is made by the plane surfaces of the jambs and architrave, the complex surfaces of the moldings (bed cornice, cymatium and tympanum), the sculpted decorations of the frieze and the capital. So we tested the performances of this sensor both on the details and on the overall shape of the doorway. The study is completed by a metric accuracy test that uses as ground truth a ToF scan [TCDL*14].

3.1. Employed device

In our case study we employed the Structure Sensor Scanner. Similarly to Microsoft Kinect, this device has an operative range capability from $0.4m$ to $12m$. Indeed, in the closer range from the sensor the device reaches a declared 3D point accuracy of $0.5mm$. The accuracy become smaller if the scanned object is placed over $3.5m$ or if the scanning volume is increased. Structure sensor is an infrared structured light 3D device, hence several issues are related to it: it does not work well in outdoor environment, since sunlight is a too strong source of infrared interference. However, our case study is located in an indoor environment not affected by direct sunlight interferences. Another critical issue is related to the material of the surface of the scanned objects. Infrared waves can be reflected, absorbed or distorted respectively by not opaque, black, or transparent surfaces, as glassy, plastic or polished objects. Our case study is composed in the majority by opaque materials such as the limestone in the door jamb and decorations. The handle and the label of the door are in a polish metal but they have been still acquired with just some light distortion (Figure 1). A possible third issue related to the Structure Sensor could be related to object with "poor geometry": in the acquisition phase the sensor needs a minimum amount of geometrical details of the object to be scanned. This is required for an optimal frame-by-frame mesh reconstruction. If not enough geometry is provided then the sensor will prompt an error message and the acquisition will fail. This problem occurs in case of particularly flat object. Our case study present a geometry complex enough to enable a good acquisition with the employed device.

Structure sensor can only acquire depth information by itself. In order to add some texture information an external RGB camera is needed. Structure sensor also needs an external computation unity to process acquired data. Usually Structure is attached or connected to an Apple iPad exploiting a wired connection and in this way textures can be acquired exploiting the standard RGB camera of the tablet. Although this is the most common way to use the sensor due to its practical aspects, we do not choose this acquisition method, since the final model is decimated before the exportation from the

Table 1: *Specs of the described handheld scanners*

Sensor	Accuracy	Resolution	Acquisition Speed	Texture
Kinect V1	n.a.	n.a.	30 fps	Yes
Kinect V2	n.a.	n.a.	30 fps	Yes
Asus Xtion PRO Live	n.a.	n.a.	n.a.	Yes
Scanify Fuel 3D	0.35 mm	n.a.	10 fps	Yes
Google Project Tango	n.a.	n.a.	n.a.	Yes
Artec Eva	0.1 mm	0.1 mm	2,000,000 per second	Yes (standard ver.)
Artec Spider	0.05 mm	0.1 mm	1,000,000 per second	Yes
Structure Sensor	0.5 mm	1.0 mm	30/60 fps	Yes (with iPad)

Figure 1: *View of the Structure Sensor behavior on three different materials (from left to right): wood, metal, limestone.*

tablet resulting in a too low quality mesh. Exploiting proper software like Skanect [ska], it is also possible to connect the Structure and the tablet to a computer through a wireless network, or just the Structure with a wired one. In the latter case we do not acquire any color information reaching a real-time acquisition of the case study. Note that texture is not really needed to estimate the mesh (e.g. the geometry) of the scanned object. Latency during acquisition is an issue that must be taken into account: sensor could lose or estimate a wrong alignment through consecutive acquisition instants, introducing noise or, in the worst case, requiring to restart the whole acquisition.

3.2. Method description

As said, the resolution of the final mesh is strictly related to the scanning volume. The case study is a door with an height of almost $3m$ and a width of almost $2.5m$ resulting in a scanning volume too large in order to obtain a quality sufficiently good. For this reason we decide to set a scanning volume of $1m^3$ and to subdivide the acquisition of the door into several single acquisitions. We acquired a total of 23 parts, starting from the bottom left position until the top right. Note that the acquisition range depends also on the sensor, so the scanning volume could be set larger than the reported $1m^3$, but with limited precision. The 23 parts have been carefully acquired with at least the 30% of overlapping between each other. This redundant information is required to correctly perform the alignment process of the subparts into the full model. We process and align the meshes exploiting the software Meshlab [CCC*08]. We acquired highly detailed meshes, with an average number of $600K$ vertices and $1M$ faces. We perform a preprocess phase to reduce the noise, as some isolated face or vertex, using the Quadric Edge Collapse Decimation of Meshlab. We discard the 80% of the points in each mesh without any visual-perceptible loss of details. Then, using the Point Glue tool of Meshlab we perform all the required alignment and saved the final model of the case study in the common OBJ format.

3.3. Comparison with Time of Flight 3D scanning

In this subsection we deal with the experimental results we obtained during the visual and metric accuracy tests. As ground truth we use a ToF mesh model. The pipeline followed is by the time used in literature [Rem11, SIP13, IS13, GRAB16, RM15] and foresees the alignment of the different models in the same reference system and the calculation of the distance between the meshes by means of Hausdorff distance algorithm application [CRS98]. Considering the performances of the handheld scanner and the purposes of this paper we consider both two details (a capital and frames and mouldings of the jams and entablature) and the overall doorway.

During ToF laser scanner acquisition (using a HDS 3000 by Leica Geosystem) we decide to carry out three scans: one frontal and two lateral and we choose a scan step very dense (about $2mm$) to have a very detailed point cloud. In these cases, as reported in previous literature works [CCDS09], the size of the noise exceeds the sampling rate so that it hides most of the details: in the following meshing phase it is mandatory to apply a specific combination of surface reconstruction and smoothing algorithms in order to avoid spikes meshes. In Meshlab we carry out the merging of the scans into a unique model, then we apply the pipeline employed in Ref. [CCDS09] by testing and choosing the parameters that better smoothed the surfaces without losing details.

A first consideration that can be done, in terms of visual accuracy of the 3D reconstructions, is that the Structure Sensor single scan models are more detailed and less noisy with respect to ToF reconstructions. This is in line with the kind of used sensor. The comparison between the three models (two details and the overall doorway) and their corresponding ToF scans was carried out in Meshlab. As for the two details, the alignment between Sensor

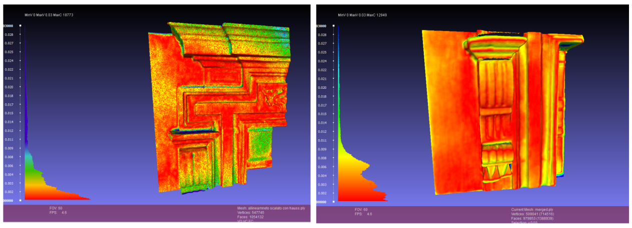

Figure 2: *Hausdorff distance and subsequent quality histogram between TOF model and Structure Sensor model of two chosen details.*

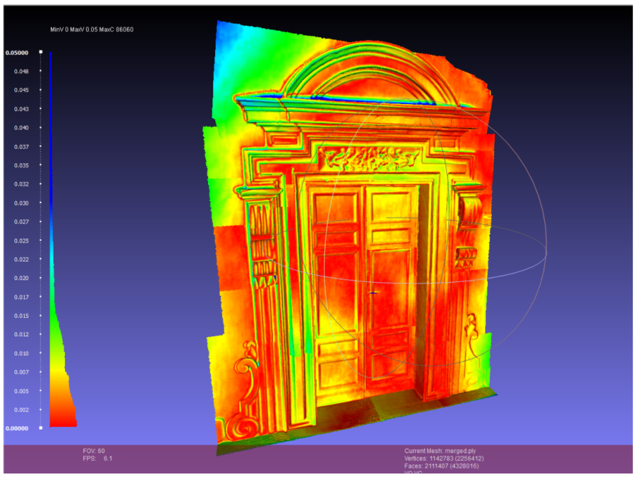

Figure 3: *Hausdorff distance and subsequent quality histogram between TOF and Structure Sensor models.*

Structure model/ToF model involved an alignment error of 3*mm*. The range calculation interval for Hausdorff distance is $0-30mm$. The second test involves the overall model of the doorway. A detailed visual analysis of the Structure Sensor model reveals some mismatches in the overlapping areas. These alignment errors could be interpreted as fallacies of the alignment step probably due to boundary geometric inconsistencies of the single scans. In order to take into account these mismatches, we calculate Hausdorff distance with the following range values $0-50mm$.

The experimental results are shown in Table 2. Furthermore, it is very interesting to read the trend of the histogram and observe the distribution of the distances between the two meshes directly on the 3D model (Figures 2 - 3), where the red color means the minimum distance between the two meshes and the blue means the maximum one.

Table 2: *Experimental results. All values are expressed in mm.*

Model	Hausdorff Range	Mean	RMS
Capital	$0-30$	4.344	6.879
Entablature	$0-30$	4.775	6.797
Overall Model	$0-50$	9.619	14.104

4. Conclusions

In this paper we defined a low cost indoor procedure facing the criticalities of the handheld 3D scanner Structure sensor for architectural elements acquisition. Furthermore, the metric accuracy test highlighted the reliability of this sensor for the details acquisition. Indeed, as shown in Table 2, the Mean distance computed on details is lower than 5*mm*, comparable with the usual ToF accuracy. On the other hand, the Mean distance computed on the whole model is 9.6*mm*, due to the severe amount of noise introduced by alignment process. These results demonstrate that this sensor can obtain high quality 3D models of architectural details, useful to integrate ToF scannings and make the digitalization of the cultural heritage easier and faster with affordable economical efforts.

References

[AAA*15] Alberghina M. F., Alberghina F., Allegra D., Paola F. D., Maniscalco L., Milotta F. L. M., Schiavone S., Stanco F.: Archaeometric characterization and 3d survey: new perspectives for monitoring and valorisation of morgantina silver treasure (sicily). In *In Proceedings of 1st International Conference on Metrology for Archaeology* (2015). 2

[ACP*10] Arcifa L., Calì D., Patanè A., Stanco F., Tanasi D., Truppia L.: Laser scanning and 3d modelling techniques in urban archaeology: the excavation of "st. agata al carcere" church in catania. *Virtual Archaeology Review 1*, 2 (2010). 2

[BBG11] Bandiera A., Beraldin J., Gaiani M.: Nascita ed utilizzo delle tecniche digitali di 3d imaging, modellazione e visualizzazione per l'architettura e i beni culturali. 81–134. 2

[BGF10] Benedetti B., Gaiani M., F.Remondino: *Modelli digitali 3D in archeologia: il caso di Pompei*, edizioni della normale, pisa ed. 2010. 2

[CCC*08] Cignoni P., Callieri M., Corsini M., Dellepiane M., Ganovelli F., Ranzuglia G.: Meshlab: an open-source mesh processing tool. *Eurographics Italian Chapter Conference 2008* (2008), 129–136. 3

[CCDS09] Callieri M., Cignoni P., Dellepiane M., Scopigno R.: Pushing time-of-flight scanners to the limit. In *VAST* (2009), pp. 85–92. 3

[CRS98] Cignoni P., Rocchini C., Scopigno R.: Metro: measuring error on simplified surfaces. In *Computer Graphics Forum* (1998), vol. 17, Wiley Online Library, pp. 167–174. 3

[Cur] Curless B.: From range scans to 3d models. *Computer Graphics 33*, 4. 2

[CZC16] Cappelletto E., Zattunigh P., Cortelazzo G. M.: 3d scanning of cultural heritage with consumer depth cameras. *Multimedia Tools and Applications 75*, 7 (2016), 3631–3654. `doi:10.1007/s11042-014-2065-4`. 2

[GMS*10] Gallo G., Milanese F., Sangregori E., Stanco F., Tanasi D., Truppia L.: Coming back home. the virtual model of the asclepius roman statue from the museum of syracuse (italy). *Virtual Archaeology Review 1*, 2 (2010), 93–97. 2

[GRAB16] Gaiani M., Remondino F., Apollonio F. I., Ballabeni A.: An advanced pre-processing pipeline to improve automated photogrammetric reconstructions of architectural scenes. *Remote Sensing 8*, 3 (2016), 178. 3

[IS13] Inzerillo L., Santagati C.: Il progetto del rilievo nell'utilizzo di tecniche di modellazione dense stereo matching. 3

[KTSP14] Kolev K., Tanskanen P., Speciale P., Pollefeys M.: Turning mobile phones in 3d scanner. *Computer Vision and Pattern Recognition* (2014). 2

[PLB] Pears N., Liu Y., Bunting P. (Eds.): *3D Imaging, Analysis and Applications*. 2

[Rem11] Remondino F.: Heritage recording and 3d modeling with photogrammetry and 3d scanning. *Remote Sensing 3*, 6 (2011), 1104–1138. `doi:10.3390/rs3061104`. 2, 3

[RM15] Russo M., Manferdini A. M.: Integrated multi-scalar approach for 3d cultural heritage acquisitions. *Handbook of Research on Emerging Digital Tools for Architectural Surveying, Modeling, and Representation* (2015), 337. 3

[SIP13] Santagati C., Inzerillo L., Paola F. D.: Image-based modeling techniques for architectural heritage 3d digitalization: limits and potentialities. *International Archives of Photogrammetry, Remote Sensing and Spatial and Information Sciences XL-5*, W2 (2013). 2, 3

[ska] Skanect Website http://skanect.occipital.com/. 3

[SSHP15] Schöps T., Sattler T., Häne C., Pollefeys M.: 3d modeling on the go: Interactive 3d reconstruction of large-scale scenes on mobile devices. *3D Vision 2015* (2015). 2

[ST13] Stanco F., Tanasi D.: Beyond virtual replicas: 3d modeling and maltese prehistoric architecture. *Journal of Electrical and Computer Engineering* (2013). `doi:10.1155/2013/430905`. 2

[STBB11] Stanco F., Tanasi D., Buffa M., Basile B.: Augmented perception of the past: The case of the telamon from the greek theater of syracuse. In *In Proceedings of Multimedia for Cultural Heritage (MM4CH)* (2011). 2

[STG*12] Stanco F., Tanasi D., Gallo G., Buffa M., Basile B.: Augmented perception of the past. the case of hellenistic syracuse. *Journal of Multimedia 7*, 2 (2012). 2

[str] Structure Sensor Website http://structure.io/. 1

[TCDL*14] Toschi I., Capra A., De Luca L., Beraldin J., Cournoyer L.: On the evaluation of photogrammetric methods for dense 3d surface reconstruction in a metrological context. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2* (2014), 371–378. 2

[VM] Vosselman G., Maas H.-G. (Eds.): *Airborne and Terrestrial Laser Scanning*. 2