International Journal of Business, Human and Social Sciences
ISSN: 2517-9411
Vol:13, No:5, 2019

The Democratization of 3D Capturing: An Application Investigating Google Tango Potentials

Carlo Bianchini, Lorenzo Catena

***Abstract*—**The appearance of 3D scanners and then, more recently, of image-based systems that generate point clouds directly from common digital images have deeply affected the survey process in terms of both capturing and 2D/3D modelling. In this context, low cost and mobile systems are increasingly playing a key role and actually paving the way to the democratization of what in the past was the realm of few specialized technicians and expensive equipment. The application of Google Tango on the ancient church of Santa Maria delle Vigne in Pratica di Mare – Rome presented in this paper is one of these examples.

***Keywords*—**Architectural survey, augmented/mixed/virtual reality, Google Tango project, image-based 3D capturing.

I. Introduction

THE wide changes triggered by the so-called "digital revolution" have deeply affected our interaction with the real world both in terms of capturing and 2D/3D modelling. While the latter has been profiting from new digital systems and applications that only 20 years ago were almost unimaginable [1], [2], the former has gone through at least two revolutions. First, the appearance of 3D scanners [6] (quickly become standard instruments for 3D capturing) and then, more recently, of image-based systems that generate point clouds directly from common digital images. [3] [4] This process is so tumultuous that in some cases technology seems to have even gone too far beyond, creating an offer of innovative solutions that still await for a use. The Google Tango Project discussed in this paper is one of these examples.

Image-based 3D capturing systems have rapidly become a real alternative to 3D scanning and by now, especially Structure from Motion (SfM) has become a reliable tool available for both researchers and professionals [5]. In this context, low cost and mobile systems are increasingly playing a key role and actually paving the way to the democratization of what in the past was the realm of few specialised technicians and expensive equipment.

Assuming the ancient church of Santa Maria delle Vigne in Pratica di Mare – Rome as a pilot case (Fig. 1), our research[1] moved along two main tracks. On one side, we have tested the features, performance and reliability of the software Google Tango, an hw/sw suite for real-time 3D capturing implemented on a mobile device (in our case the Lenovo Phab 2 Pro). On the other, we have explored the potential of this suite as a low cost capturing/modelling tool for built artefacts and for Augmented/Mixed/Virtual Reality applications [7].

Carlo Bianchini and Lorenzo Catena are with the Sapienza–University of Rome, Italy (e-mail: carlo.binchini@uniroma1.it, catena.1472501@studenti.uniroma1.it).

[1]Part of the project was developed in the framework of the degree thesis in architecture by Lorenzo Catena, under the supervision of Maria Grazia Ercolino and Carlo Bianchini.

II. The Google Tango Project

Tango is an experimental project developed by Google for Android smartphones and tablets that has started in 2013 and closed in 2018 being substituted by ARCore [7].

The main goal of the project was to test the reliability, impact and interest of users in an hw/sw package allowing the creation of AR/MR/VR content just using their smartphones or tablets. In this framework, some manufacturers (Lenovo, Asus) have released pioneer Tango devices equipped with a bundle of integrated hw components RGB (Red Green Blue) camera; time of flight camera; infrared projector; fish-eye motion camera; accelerometer and gyroscope – Fig. 2) and innovative sw features. In brief, these devices can real-time "scan" the targeted scene thanks to a mix of special inertial sensors that, without using any type of external signal, determine the position in space of the smartphone/tablet and generate point clouds and texturized meshes.

The fundamental principles of Tango platform are the following (Figs. 4-6):

- Motion Tracking - SLAM system (simultaneous localization and mapping)
- Area Learning
- Depth Perception

A. Motion Tracking - SLAM

This feature allows referencing the movement of the device to the 3D space around it. At this stage, in fact, Tango collects a large amount of data from the fish-eye camera (about 60 black and white frames per second) and from the IMU inertial system (accelerometers and gyroscope updating the position of the device up to 100 times per second). The processing of images leads to identify the "visual features" of the represented scene mostly referring to geometric discontinuities of the real environment. These would act as anchor points and allow for an accurate frame by frame "feature tracking", a crucial step of the entire process.

For each captured frame, the feature tracking and the IMU readings are continuously coupled: from the combination of these two dataset, the software identifies a 3D trajectory representing the movement of the smartphone in space.

B. Area Learning

This phase leads the Tango device to determine its own position with respect to the surrounding environment. This process develops in tight combination with the previous one

(Motion Tracking) and aims at accurately determining the 3D position of the device, a fundamental function both in 3D capturing and the application of augmented reality. In brief, when visiting a previously acquired environment, the device compares the new information with the old one and updates and refines its calculated trajectory, automatically correcting those errors accumulated during the movements.

Fig. 1 The church of Santa Maria delle Vigne in Pratica di Mare – Rome

Fig. 2 The Lenovo Phab Pro 2 Tango device

Fig. 3 The Google Tango device used as AR/MR/VR viewer

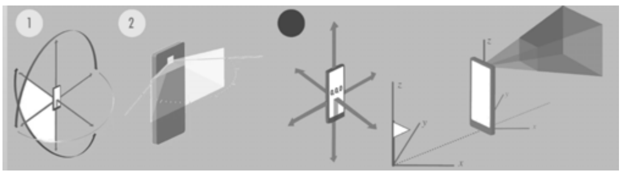

Fig. 4 The first principle: Motion Tracking - SLAM system (simultaneous localization and mapping)

C. Depth Perception

The dynamic 3D information acquired during the first two steps is finally framed into a metric scaffolding generated by the measurements of the infrared Time of Flight (ToF) Sensor. In this phase, in fact, Tango creates a 3D model of the surrounding scene, an essential step to ensure the due metric accuracy to the final dataset while using the device for 3D acquisition or augmented reality applications.

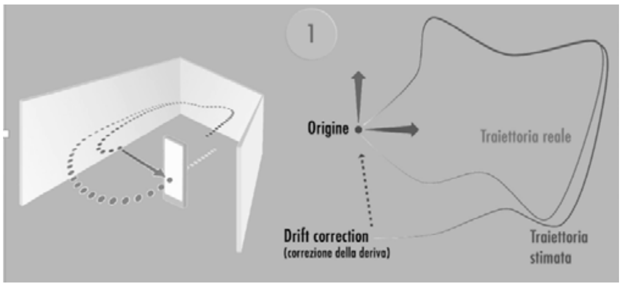

Fig. 5 The second principle: Area Learning

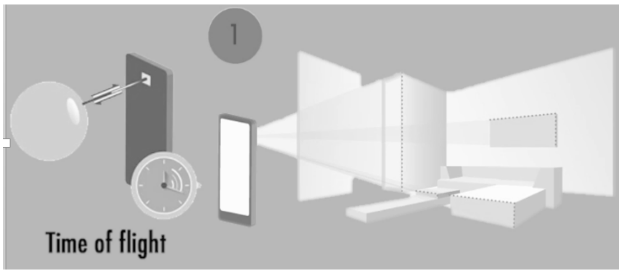

Fig. 6 The third principle: Depth Perception

III. Testing Google Tango

Exploiting the combination of the built-in sensors described in the previous paragraph, Tango technology can correlate positional information, RGB readings and data obtained from the infrared TOF sensor so to generate a textured 3D mesh model of the "scanned" areas. The acquisition process is particularly simplified and completely automated, actually requiring the user only to choose and complete his own track within the area to be captured.

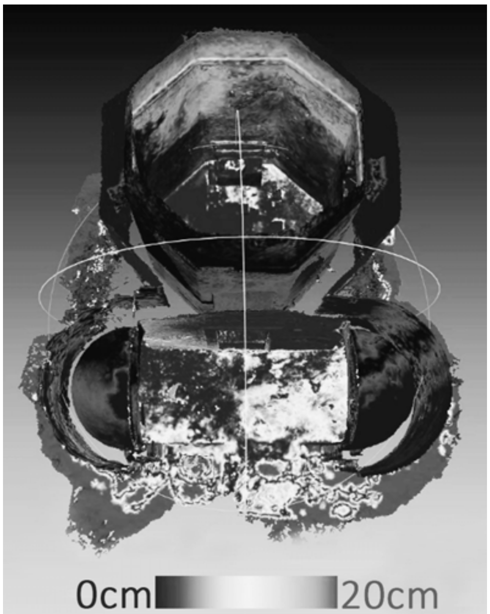

Fig. 7 The model generated from the Tango device acquisition

Fig. 8 The AR/MR/VR applications

Our experimentation concentrated on the quality and the possibilities offered by this technology testing the systems on a small pilot case (the Church "Santa Maria delle Vigne" in Pratica di Mare – Rome). Our activity has entailed the surveying (3D capturing), the metric validation of outputs (Fig. 7) and the creation of 3 models/contents respectively developed in AR/MR/VR environments (Fig. 8).[2] The results were quite encouraging especially vis-à-vis the capturing and construction of models for communication that were completed almost using only the first Tango consumer device. Nevertheless, we have reported some major bottlenecks actually affecting the capturing phase due essentially to hardware constraints and in particular the memory limitations of the device. To overcome these problems the acquisition phase was divided into segments each one accurately planned and optimized on the capabilities of the smartphone used.

Once all the mesh surfaces covering the entire object had been built, they were aligned by means of homologous points

[2] Lenovo Phab2 Pro. Integrated sensor system: RGB Camera; Time of Flight Camera (TOF); IR Projector; Fish-eye Motion Camera; Accelerometer and Gyroscope.
Suite software: Constructor Developer Tool" v.1.4.2. Presently we can use also other software as "RTAB-Map" (Mathieu Labbé) that constructs point clouds, or previous "Open Constructor" now "3D Scanner for Arcore" that produces a texturized mesh model similarly to the app used in this project.

and validated using traditional mesh management software (such as Meshlab and Rhinoceros) by comparison with a reference model acquired with an integrated topographic/SfM approach.[3] As we said, though, the Tango project has been designed mainly as a tool for the "enrichment of reality" in the form of Augmented, Mixed and Virtual Reality content. In our experimentation, the first two models (AR and MR) have been developed before and somehow in preparation of the VR one, more oriented this last towards the fruition of artefact's informative content. AR and MR models actually provide interactive information on the different historical phases of the building merging its real appearance with three-dimensional reconstructions of missing portions or elements relevant for its stratigraphic reading. The mobile device, previously used for content development, here acted as a VR viewer ensuring the exploration of an immersive model with which the user was able to interact.

The on-the-fly survey conducted with the Tango system demonstrates how the high-end technology integrated in the device allows for a user-friendly but accurate 3D surveying of built artefacts. The results are encouraging and at times competitive with SfM technology. Many of the limitations encountered are typical defects related with an in progress development. The hardware framework is not completely optimized for such “survey/modelling” activities: first of all the on board RAM obliges to perform multiple scans even for small constructions. Other issues deal with a not yet standardized workflow or to improvable design choices (i.e. the integrated infrared sensor has a limited range of action and obviously cannot work efficiently enough in a too lighted environment). Nevertheless, the advantages that Tango suite offers to the entire sector of survey and representation of built artefacts are important: the low cost of the system (about 500€), its user-friendliness, its portability and, above all, its stand-alone capacity. This last feature is crucial because the device needs in any working phase nothing else than the power provided by its own battery regardless any internet or GPS connection.

In conclusion, we can affirm that Tango, though quite rudimental, still represents an additional innovative tool to improve the possibilities offered to developers (and even to generic users). In particular, it allows for the creation of new types of user experiences and AR/MR/VR content thanks to new functions (such as inland navigation and environmental recognition) leading to an enhanced level of user's involvement. On the other hand, the limitations presently affecting Tango in our opinion will be quickly overcome by the natural evolution of new products starting from what we could consider its heir, the ArCore suite.

IV. Beyond Tango: The ArCore Suite

ArCore was initially presented as revolutionary because it would be compatible with all new Android devices and not only with some specific models as Tango. The idea was to extend the great potential of Tango technology (later revealed not so great unfortunately) to any smartphone with an Android version equal to or greater than 7.1 Nougat (recently corrected to 8.0 Oreo in some cases).

The substantial difference between Tango and ArCore lies in the removal of the depth sensing (TOF sensor data) and motion tracking (fish-eye camera data) from the process. This has allowed a reduction of costs and a constantly increasing compatibility with all devices that, updated to the new version of Android, respect the minimum ArCore requirements.[4]

Unfortunately, even if the removal of important hardware components is improving the development and adoption of ArCore, significant bottlenecks persist in the software especially due to a "calibration" operation charging the IMU system much more than it was for Tango. This process in fact requires an exchange of information between the manufacturer of the smartphone and the companies that produce the various sensors (accelerometers and gyroscope). Furthermore, so far ArCore is not comparable to Tango because, although simplification efforts are appreciable (no specific hardware required), the platform is not enough supported and comprehensively less powerful. It is not possible to perform (at least officially) neither 3D capturing nor to use the Area Learning function. This means that the device is not able to recognize a previously scanned environment and therefore it is impossible to create the experiences of augmented/mixed reality. The reciprocal accurate positioning of reality-based and virtual 3D models is in fact fundamental for the creation of AR/MR content. Despite all these problems that affect the new-born ArCore, expectations are still very high and some signals suggest the implementation in the near future of new functions aiming at reducing the gap with the “old” Tango.

V. Conclusions

The future is now on the doorstep and our challenging task is to let it in so to enhance our vision about the world around us.

The survey activities finalized at the creation of models of built artefacts certainly are directly influenced by the systems used for capturing with a clear preference for those easy to transport, highly automated, multi-sensor, as much as possible miniaturized but nonetheless capable of providing scientifically reliable data and outcomes. At the same time, the widespread diffusion of technologically advanced devices allows us to offer to users new methods for knowledge enriching through the creation of models capable of outlining a more engaging and efficient user/reality interaction and, eventually, disclosing new forms of communication and dissemination of culture.

Augmented/Mixed and Virtual Reality applications have in fact massively rekindled the interest of users (and the experiential lever has certainly represented one of the key elements for their success) only once these technologies have

[3] For this phase was used the Agisoft Photoscan software and a Leica TCR 4500 Total station.

[4] The officially supported list of devices that meets Google standards is available at the link: https://developers.google.com/ar/discover/supported-devices and is continuously updated.

been incorporated into low-cost mobile devices. In fact, even though systems capable of simulating immersive or interactive effects have existed for years, these approaches were quite expensive and required dedicated installations. These facts have proven both relevant obstacles for their wide diffusion. With the miniaturization of acquisition and visualisation technologies, this scenario has radically changed paving the way to the involvement of a wider audience given the easy access to tools ready for use simply downloading a specific application and exploiting the potential of a smartphone.

Both Tango closed experience and its successor ArCore show that technology is now mature enough to change the overall approach to data acquisition and AR/MR/VR content production. It is by now a commonplace option to acquire geometric data, interact with environments and design a new way of perceiving reality all with the help of smart devices, all essential steps towards the democratization of knowledge.

Acknowledgments

The AR content has been developed together with the start-up *Vertical*.

References

[1] Bianchini, C. Survey 2.0: new technologies, new equipment, new surveyors. In Giandebiaggi P., Vernizzi C. (Eds.), *Italian Survey & International Experi*ences, 2014, 763-768..

[2] Bartolomei, C., Bianchini, C., & Ippolito, A. Low cost documentation for the knowledge of the city. In *Proceedings of the 4th Annual International Conference on Architecture and Civil Engineering* (ACE 2016). 10.5176/2301-394X_ACE16.5.H.

[3] Bianchini C., Ippolito A., Senatore, L.. J., Borgogni, F., Capiato, E., Capocefalo, C., Cosentino, F. Laser scanner surveys vs low-cost Surveys. A methodological comparison. In Di Giamberardino P., Iacoviello D., Natal J. R., Tavares J. M. (eds). *CompIMAGE 2012 Computational Modeling of Objects Presented in Images: Fundamentals, Methods and Applications* 3rd edition. London: Taylor & Francis Group, 2012, pp. 453-457. ISBN: 978-0-415-62134-2 (Hbk), 978-0-203-07537-1 (eBook).

[4] Guidi G., Remondino F., Russo M., Menna F., Rizzi A. 3d Modeling Of Large And Complex Site Using Multi-Sensor Integration and Multi-Resolution Data. In *VAST 2008: 9th International Symposium on Virtual Reality, Archaeology and Cultural Heritage*. Aire-la-Ville, Switzerland: Eurographics Association, 2008, pp. 85-92. ISBN: 978-3-905674-14-9.

[5] De Luca Livio. La Fotomodellazione architettonica. Palermo: Dario Flaccovio Editore, 2011. 264 p. ISBN: 978-88-579-0070-4.

[6] Bianchini C.. Laser Scanning X. In Chiavoni, E., Paolini, P. (). Metodi e tecniche integrate di rilevamento per la costruzione e fruizione di modelli virtuali 3D dell'architettura e della città. Ricerca Cofin 2004, coordinatore nazionale Mario Docci. Roma: Gangemi Editore, 2007, pp. 24-31. ISBN: 88-492-1413-4