The Democratization of Processes and the Use of Remotely Guided Acquisition Tools for Survey 2.0

Carlo Bianchini, Luca J. Senatore, Lorenzo Catena

Abstract

The technology that ‘amplifies’ the real world as well as the technological means of data capturing, are continuing to have a development difficult to control or predict. Projects of a certain interest are born, develop and die sometimes without producing significant results nevertheless becoming phases of a fluid and dynamic process, on which new software applications and hardware systems are built. At the same time, the role of the academy seems to be often overtaken by what is developed by the citizen science. Researchers no longer seem to exclusively explore the ‘discovery’ domain, but rather work at the setup of scientifically reliable protocols for technologies often used by different communities of citizens in a way that does not conform to the specifications for which they were originally designed.

This paper focuses on these issues taking as its starting point two experiments with high innovative content that cover various stages of the process defined by the term Survey. They pertain on the one hand to a hw/sw system of data capturing, processing and communication developed for mobile devices (Tango Project) and on the other to a mixed ROV/UAV multisensor platform equipped with a LIDAR system and digital camera for data acquisition in inaccessible places (Heritagebot Project).

Keyword: Google Tango, Heritagebot, Augmented Reality, UAV, Survey.

Introduction

In the last years, the Survey domain has been experiencing a season of rapid and in some ways extraordinary change that can be assimilated to an evolutionary leap or, more precisely, to numerous evolutionary leaps all tending towards a phase of consolidation that we could name as Survey 2.0. What is the real meaning of Survey 2.0 is still hard to know, as it becomes immediately obsolete any attempt to define it. At the same time, the role of the academy seems to be often overtaken by what is developed by the so-called citizen science [1]. Researchers no longer seem to exclusively explore the ‘discovery’ domain, but rather work at the setup of scientifically reliable protocols for technologies often used by different communities of citizens in a way that does not conform to the specifications for which they were originally designed. Projects of a certain interest are born, develop and die sometimes without producing significant results nevertheless becoming phases of a fluid and dynamic process during which new software applications and hardware systems are built.

This typical phenomenon of the so-called ‘digital revolution’ influences our interaction with the real world both in terms of capturing and 2D/3D modeling. New systems and digital applications that only 20 years ago were almost unconceivable, become commonplace and develop through the simplification of processes, the miniaturization of sensors, the change of paradig-
ms. As in the case of the massive 3D capturing that, since the appearance of LIDAR (Laser Imaging Detection and Ranging) scanners and, more recently, Structure from Motion (SfM) systems that generate point clouds directly from common digital images, has substantially changed the approach to the understanding of built artefacts (architectural and archaeological).

In fact, LIDAR or SfM 3D capturing systems have quickly become a sort of standard thanks to the reliability of the instruments, their extraordinary acquisition speed, a certain user-friendliness and overall economy of use (even if not always of cost).

In this context, mobile and low-cost systems are playing an increasingly relevant role, paving the way to the democratization of what was once the realm of few specialized technicians and high-end research or professional organizations. At the same time, the use of remote-controlled robotic systems is constantly and rapidly evolving, providing interesting perspectives also for the field of architecture and archaeology. Robotics in fact allows improving the methods for data capturing and knowledge enhancing of Built Heritage (BH) especially when accessibility of sites and exploration of artifacts is dangerous or impossible to human operators. Furthermore, the miniaturization of robotic platforms components combined with their reduced weight and consumption guarantee not only higher performances but also a minimization of impacts on the objects under study.

These processes have been (and still are) so tumultuous that in some cases the technology seems to have gone even too far, proposing an offer of innovative solutions for which there is not yet a real demand. In our opinion, the Google Tango Project and the Heritagebot Project discussed in this paper are two of these examples. Assuming as pilot cases the ancient church of Santa Maria delle Vigne in Pratica di Mare near Rome and the archaeological area of Cassino, our research has therefore focused on two main tracks. On one side, within the Tango Project, we tested characteristics, performance and reliability of a mobile device for real-time 3D capturing and Augmented, Mixed and Virtual Reality (AR/MR/VR) content development. On the other, we assessed the capabilities of a remote controlled robotic prototype developed within the Heritagebot Project and particularly its data capturing sensors (LIDAR and SfM) when used in not easily accessible contexts.
A Google Tango/Google ArCore Project

Tango is an experimental project developed by Google for Android smartphones and tablets that started in 2013 and ended in 2018 being presently replaced by ARCore.

The main objective of the project was to test the reliability, impact and interest of users in a hw/sw package that would allow for the creation of AR/MR/VR content only using their ‘smart’ devices. Both Lenovo and Asus producers joined this challenge and released innovative Tango devices with a special package of integrated hardware components (RGB camera; time of flight camera; infrared projector; fish-eye motion camera; accelerometer and gyroscope) and innovative software features. In summary, these devices can ‘scan’ the target scene in real time thanks to a mix of special inertial sensors that, without using any type of external signal, determine the position of the smartphone/tablet in space and generate point clouds and textured meshes.

The basic principles of Tango are the following:
- Motion Tracking - SLAM system (simultaneous localization and mapping);
- Learning Area;
- Perception of depth.

A. Motion Tracking - SLAM System
This function refers the movement of the device to the 3D space around it.
In this phase, in fact, Tango collects a large amount of data thanks to the fish-eye camera (about 60 black and white frames per second) and the inertial IMU (Inertial Measuring Unit) which includes accelerometers and gyroscope that update the position of the device up to 100 times per second. The processing of the images leads to the identification of the ‘salient features’ (feature tracking) of the explored scene mainly referring to the geometric discontinuities of the real environment. These in fact act as reliable elements that allow frame by frame an accurate feature tracking to which IMU readings are continuously coupled. From the combination of these two sets of data, the software identifies a 3D trajectory that represents the movement of the device in space.

B. Learning Area
This step causes the software to determine its position with respect to the surrounding environment. This pro-
cess develops in close combination with the previous one (Motion Tracking) and aims at accurately assessing the position of the device in space, a key function in both 3D capturing and AR/MR/VR applications. In brief, by visiting a previously acquired environment, the device compares new and old information by updating/refining its own calculated trajectory and automatically introducing corrections to the errors accumulated during the movement.

C. Perception of depth
The dynamic information acquired during the first two phases is finally placed into a three-dimensional metric grid generated thanks to the measurements of the infrared TOF (Time of Flight) sensor. It is in this phase, in fact, that Tango creates the 3D model of the surrounding scene, an essential step to guarantee the required metric precision of the final dataset, which from now on will become the reference framework for 3D capturing and AR/MR/VR applications.

Tango Applications

Using the combination of the integrated sensors described in the previous paragraph, Tango technology is able to correlate position information, RGB readings and data obtained from the infrared TOF sensor to generate a textured 3D mesh model of the "scanned" areas. The acquisition process is particularly simplified and fully automated, requiring only the user to choose and complete his own trajectory within the area to be acquired. Our experimentation focused on the quality and possibilities offered by this technology, testing the systems on a small pilot case (the Church Santa Maria delle Vigne in Pratica di Mare, near Rome). Our activity involved the surveying (3D capture), the metric validation of the data and the creation of 3 models/contents developed respectively in AR, MR and VR environments. The results were rather encouraging especially about the acquisition and construction of communication models built almost exclusively with the consumer Tango device we used (Lenovo Phab Pro). However, we encountered some relevant problems during the acquisition phase mainly due to the poor performance of the hardware and its storage low capacities. To overcome these bottlenecks, the capturing phase has
been divided into segments, each one carefully planned and optimized according to the capacity of the device. The acquired 3D numerical model (point clouds) was then processed to obtain the corresponding mesh surface. Once all the mesh surfaces were built, they were aligned by means of homologous points and validated with traditional mesh management software (in our case Meshlab and Rhinoceros) with respect to a reference model derived from an integrated topographic and SfM surveying campaign.

However, the Tango project was mainly conceived as a tool for ‘enriching reality’ in the form of AR/MR/VR content. In our experimentation the first two models (AR and MR) have been developed before and in some way in preparation of the VR model more oriented towards the use of the information content about the artefact. The AR and MR models provide in fact interactive information on the different historical phases of the building, merging its real appearance with three-dimensional reconstructions of missing parts or elements relevant for its stratigraphic reading. The mobile device, previously used for the development of content, has in this phase acted as a VR viewer ensuring the exploration of an immersive model with which the user can interact.

The expeditious survey conducted with the Tango system shows how the high-end technology integrated in the device allows a 3D survey of built artefacts accurate enough and at the same time easy to perform. The results are encouraging and sometimes even competitive with SfM technology. Many of the limitations found (hw performance, memory) are typical defects of an ongoing development. The hardware structure is not completely optimized for these ‘survey/modelling’ activities: first of all the limited RAM of the device forces users to perform multiple scans even for small constructions. Other problems concern a workflow not yet standardized or design choices that can be improved (for example, the integrated infrared sensor has a limited range of action and obviously cannot work properly in a too bright environment).

However, the advantages that the Tango suite offers to the entire field of surveying and representation of built structures are important: the low cost of the system (about 500 €), the ease of use, portability and, above all, the ability to operate in a real stand-alone condition. This last feature is particularly significant because the device does not need in every phase of the work

Fig. 4. Interface for accessing the contents of AR/MR/VR.
nothing but the power supplied by its own battery, regardless of any connection (i.e. internet, GPS).

In conclusion, we can affirm that Tango, although rudimentary, still represents an interesting innovative and ‘democratic’ tool offered to any surveyor or AR/MR/VR content developer. In particular, it allows the creation of new types of user experiences and content thanks to new functions (such as inland navigation and environmental recognition) that lead to a greater user involvement. Furthermore, the limitations that currently affect Tango in our opinion will be quickly overcome by the natural evolution of new products coming from what we might consider its heir, the ArCore suite.

With the closure of Tango, in fact, Google has started a new and more ambitious project called ArCore, initially presented as revolutionary because it would be compatible with all new Android devices and not only with some specific models as its predecessor Tango.

The idea was to extend the great potential of Tango technology (unfortunately not so great) to any smartphone with an Android version equal to or greater than 7.1 Nougat (in some cases recently corrected to 8.0 Oreo). The substantial difference between Tango and ArCore lies in the removal from the process of the depth detection (TOF sensor data) and the motion tracking (fish-eye camera data), a choice that has certainly determined a reduction of costs and an increment of compatibility with the Android devices that meet the minimum requirements of ArCore.

Unfortunately, even if the cutting of important hardware components has given impulse to the development and adoption of ArCore, still there are significant critical issues in the software, mainly because of the additional ‘calibration’ operations entrusted to the IMU system (much more than it was for Tango), which require an exchange of information between the manufacturers of smartphones and those of the various sensors (accelerometers and gyroscope) at the moment not always adequate.

Moreover, ArCore does not seem to have yet overtaken Tango also because, despite the appreciable efforts of simplification (no specific hardware is required), the platform is not sufficiently supported and powerful. For example, it is not possible to perform (at least officially) neither 3D capture nor use the Area Learning function: this means that the device is not able to recognize a previously explored environment limiting in fact the creation of AR/MR/VR experiences.

**Fig. 5. Heritagebot prototype design and technical specifications.**

**SPECIFICHE TECNICHE**

- **Dimensioni (overall size):** cm 50x50x50
- **Peso (stimato):** Kg. 5
- **Power supply:** 12 V
- **Battery capacity:** 5,2 A/h
- **Payload (max):** 2 kg
- **Walking speed:** 65 mm/sec = 234 m/h
- **Battery life (stimato):** 16/32 min in small flight 5 hours in static mode 2 hours video/sensor mode
Despite all problems that afflict the newborn ArCore, expectations are still very high, and some signs suggest the implementation soon of new features to reduce the gap with the ‘old’ Tango.

The Heritagebot Project

Robotics has so far provided relevant improvements to several fields (industrial, research, professional). More recently, though, it has increasingly influenced also recreational applications thanks to the availability of low cost equipment and the simplification of interfaces. This fact has triggered a virtuous helix of ‘democratization’ in which affordable remote-controlled platforms and lay users influence each other actually creating innovative products capable of exponentially enlarging the fields of application. This process deeply involves also the documentation and survey workflows of BH with a clear tendency to make capturing activities almost automatic, increasingly remote controlled and not affected by many of the logistic limitations we have experience so far.

Generally, robotic equipment tends to be bulky as the result of a coupled requirement: provide enough mobility and stability to the platform that hosts the sensors and ensure the energy needed to operate. For this reason, these models seem not to be adequate to perform the activities connected with the documentation and survey of BH where instead a low impact often represents a mandatory constraint in order to preserve the object and its parts.

In this framework, we took up the challenge of developing a robotic platform providing an increased manoeuvrability, a wide range of sensors and, finally, enough energy to ensure an effective operability. Together with the platform, ‘sensors’ clearly represent a crucial part of the problem for not only their mechanical or engineering implications but above all for the task they must perform in order to make the robotic station operative. Presently, we can already count on a variety of small size and low consuming devices that appear promising candidates for building up an interesting robotic platform in which different sensors could work together to perform a reliable capturing of BH. In addition, the data acquired by such a digital sensors’ system can be integrated in a single work protocol ensuring both the control of the platform itself and the capturing of the studied artefact.
Quite apart from the design issues we have been briefly describing in the previous lines, we must underline how such a robotic system intends to join that challenging ‘task force’ of equipment aiming at making the documentation and surveying of BH as much as possible complete improving the accessibility to what until now is inaccessible for various reasons.

In fact, the layered nature of buildings that have grown over time, their physical response to years, centuries and millennia, or even the transformations produced by catastrophic events, are all pieces of information crucial to design a respectful and effective intervention on BH. The possibility of consciously making critical choices about the ‘value’ of an architectural or archaeological artefact often comes up against the objective limits imposed by the difficulty of exploring particular ‘hostile’ environments. What is accessible and visible is in fact already analysable in high detail, but often the objects of study have parts precluded to the conventional procedures of acquisition both for reasons of safety of the operator (collapse, toxicity etc.) and because they are too small for the instruments to physically access them.

Just to give an example, let us consider an artefact seriously at risk for collapse. Its documentation and survey would thus be precluded to human operators unless a complex system of consolidation structures would be setup. Nevertheless, many narrow spaces (such as subfloors or tunnels) would be excluded anyway due to practicability reasons. A robotic platform equipped with the ‘right’ sensors and a ‘several degrees of freedom’ movement capacity would be in this case the solution to overcome the mentioned above information gap.

The Heritagebot Project [2] was launched to meet these needs. It has aimed at constructing a remotely controlled mobile platform, highly innovative for its ability to acquire digital data autonomously and in locations otherwise inaccessible.

Applications using the Heritagebot remote control prototype

A major part of the Heritagebot Project has been devoted to the analysis of the state of the art of acquisition technologies and to what extent these technologies would support the creation of models for communication. The first phase of the activities concerned the assessment of which sensors would fit at best the requirements of the prototype to be developed. The market proposes different solutions that can be classified ac-

Fig. 7. FSM procedure in inaccessible location: frame of the processing phase and point cloud.
cording to their acquisition technology. In general, sen-
sors can be divided essentially into two classes: LIDAR
(2D or 3D) based on TOF laser technology and sensors
capable of reacting at different electromagnetic wave-
lengths both in the visible section of the spectrum (the
common photographic sensors) and in the non-visible
ultraviolet and infrared section (the latter used by the
so-called depth cameras including the one mounted on
the Tango device mentioned in the first part). The two
technologies acquire partial information from objects in
the form of digital data that after processing can provide
metric and/or perceptive information.
In order to evaluate the quality of the data acquired
through the individual sensors, we used a comparative
method using on the one side laboratory acquisitions of
reference objects and on the other readings recorded in
real contexts. Besides, the larger and bulkier technologies
(3D laser scanning and topography) not easy to use in
inaccessible environments but with certified metric qua-

ty, have represented the benchmark for the validation
of the results acquired during the test of the sensors.
Furthermore, working according to specific sections or
plans, we assessed not only the metric quality of data
but also the performance of the sensors in terms of
operational distance, the average deviation per unit of
measurement considered, the ability to acquire RGB or
infrared data with respect to the brightness or sensitivity
of the detector (a key factor for modeling using SfM
software).
The final comparison showed a comforting picture, even
if not definitive. Although each sensor considered indi-
vidually shows evident limits, however this problem was
easily overcome by integrating different sensors. In con-
clusion, the most promising solution to obtain digital data
useful for critical or descriptive processing is probably a
single device equipped with several sensors all interlaced
(e.g. referred to a common geometric matrix) and thus
able to provide consistent digital data of the analysed
object.

Procedure based on Structure from Motion

The capturing procedure implied to usage of SfM. We
carried out two different tests in the Cassino archae-
ological area (the first outdoor and the second under-
ground) using an action camera mounted on a Parrot
**Mini Drone** remotely controlled.

The first test intended to capture a part of the cavea of the ancient theatre of Cassino. The flight was quite successful because the photos taken from the drone allowed the software to recognize a great number of points and eventually the creation of a dense point cloud. After this first test, we performed a new micro capturing campaign with the same method (mini drone, action camera and SfM) but this time focussing on a structure positioned under the theatre scene where are located some archaeological finds. In this case, though, the photogrammetric software has recognized a much smaller number of points with direct negative influence on the quality of the final point cloud. These tests gave us the opportunity to identify some bottlenecks of the process, e.g. the level of illumination that underground was beyond the operational limits of the sensor used. However, the test lead us to the easy (in this case) solution of this problem by integrating in the system a macro flash for the camera.

**Second Test LIDAR**

A further test on the potential of the prototype has been carried out with the use of a LIDAR sensor and combining the acquired data with those obtained with the SfM technique. This activity was carried out again in an inaccessible area in the Roman archaeological site of Cassino: the tunnel that passes from the centre of the amphitheatre under the outer wall to the southeast of the building, probably originally devoted to the drainage of water. The choice of this site came from the need to test the instruments in a context showing many of the challenges that characterize an area with low accessibility: the spontaneous vegetation, the narrow space available, the unevenness of the ground and, more generally, of the entire surrounding surfaces. The testing environment was accurately mapped with 184 photos, taken not only inside the tunnel but also outside to allow the combination with other relevant data of the building. In the following phase, 183 photos were correctly aligned, obtaining a point cloud composed of more than 27 million points. To simulate the actual operation of our prototype (able to move along the x,y,z axis and to make acquisitions both in flying and ground mode) the tests were carried out by making several scans on parallel planes to map the entire tunnel using the LIDAR sensor. The processing of the data necessary to convert the polar coordinates (the way in which the instrument acquires the data) in cartesian xyz (crucial to read the file as a point cloud) was used by an accelerometer installed on the acquisition platform in order to provide the relative movements. Considering that our LIDAR sensor was recording the coordinates only on a plane, the z coordinate has been set to zero for all scans and then the vertical sections have been rotated and aligned according to the zenith while the acquired plane has been left horizontal at different heights. The alignment processes were performed in a CAD environment thanks to the compatibility between Recap 360 (for the management of point clouds) and AutoCad.

The alignment process produced a satisfactory result, both for the density of the points acquired and for the metric uncertainty of the measurement. Moreover, unlike the photogrammetric process, the LIDAR does not suffer from the lack of light inside the tunnel and this allows an easier detection of those areas that, without additional light, would not be detectable with traditional cameras. From this data, numerous characteristic sections have been created: a longitudinal section that underlines the altimetric variations and different transversal sections in the key points where it is possible to notice the different solution used for the ceiling of the gallery.

**Conclusions**

The important activity aimed at creating models of manufactured products is certainly influenced directly by the systems used for the capturing with a clear preference for those easy to transport, highly automated, multisensor, as miniaturized as possible but still able to provide scientifically reliable data and results. At the same time, the diffusion of intelligent devices allows us to offer users new ways of enriching knowledge by creating models able to outline a more involving and efficient user/reality interaction and, eventually, to reveal new forms of communication and diffusion of culture. Augmented, Mixed and Virtual Reality applications have in fact massively rekindled the interest of users (and the experiential leverage has certainly been one of the key elements for their success) only after these technologies
have been incorporated into low-cost mobile devices. In fact, even if for years there have been systems able to simulate immersive or interactive effects, these were quite expensive and required dedicated installations. The same phenomenon has concerned robotics until a number of experiences in the field of 3D printing and ‘play’ robotics have released, especially thanks to the internet, numerous examples to the wide public. With the miniaturization of acquisition and visualization technologies, this scenario has radically changed, opening the way to the involvement of a wider audience, given the easy access to ready-to-use tools, simply by downloading a specific application and exploiting the potential of a smartphone. Both Tango’s and Heritagebot’s experience show that the technology is now mature to change the global approach to data acquisition and AR/MR/VR content production. It is now a concrete and widespread option for acquiring geometric data, interacting with environments, even inaccessible ones, and designing a new way of perceiving reality with the help of intelligent devices. A new leap, therefore, towards the democratization of health or natural resource management by fostering collaborations between community citizens and scientific institutions <https://it.wikipedia.org/wiki/Citizen_science> (accessed 2019, May 26).

Notes

[1] Citizen Science is defined in different ways because it has multiple origins and derives from different concepts. The Oxford English Dictionary in 2014 defined it as “scientific activity conducted by members of the general public in collaboration with scientists or under the direction of professional scientists and scientific institutions”. However, Citizen Science, as Bonney, Cooper and Ballard (Bonney et al. 2016, p. 1) point out, also refers to something else when it allows ordinary citizens to ask questions and provide answers on important scientific issues or to direct public attention to environmental issues, public health or natural resource management by fostering collaborations between community citizens and scientific institutions <https://it.wikipedia.org/wiki/Citizen_science> (accessed 2019, May 26).

[2] The contribution is part of the products of a FILAS Lazio Region research carried out by the Department of Economics of the University of Cassino and southern Lazio and involving researchers from the laboratories DART, LARM, IMPRENDILAB and FINLAB.

Authors

Carlo Bianchini, Dipartimento Storia, Disegno e Restauro dell’Architettura, Sapienza Università di Roma, carlo.bianchini@uniroma1.it
Luca J. Senatore, Dipartimento Storia, Disegno e Restauro dell’Architettura, Sapienza Università di Roma, luca.senatore@uniroma1.it
Lorenzo Catena, Dipartimento Storia, Disegno e Restauro dell’Architettura, Sapienza Università di Roma, catena.1472501@studenti.uniroma1.it
Reference list


