

ROBOTICS AND VIRTUAL REALITY FOR CULTURAL HERITAGE DIGITIZATION AND FRUITION

D. Calisi^{a,*}, F. Cottefoglie^a, L. D'Agostini^a, F. Giannone^a, F. Nenci^a, P. Salonia^b, M. Zaratti^a, V.A. Ziparo^a

^a Algorithmica S.r.l., Rome, Italy - *lastname@algorithmica.it*

^b ICOMOS Italia, Rome, Italy - *firstname.lastname@itabc.cnr.it*

KEY WORDS: Digitization, Robotics, Virtual Reality, Archaeological Data Management

ABSTRACT:

In this paper we present our novel approach for acquiring and managing digital models of archaeological sites, and the visualization techniques used to showcase them. In particular, we will demonstrate two technologies: our robotic system for digitization of archaeological sites (DigiRo) result of over three years of efforts by a group of cultural heritage experts, computer scientists and roboticists, and our cloud-based archaeological information system (ARIS). Finally we describe the viewers we developed to inspect and navigate the 3D models: a viewer for the web (ROVINA Web Viewer) and an immersive viewer for Virtual Reality (ROVINA VR Viewer).



Figure 1. Example of image based reconstruction, using high-definition camera pictures

1. INTRODUCTION

The conservation of cultural heritage sites is an important goal for both scientists and the general public and is a complex task that involves activities in the field of knowledge gathering, prevention and monitoring, and that requires the cooperation of several actors (i.e., surveyors and cultural heritage professionals) as well as inter- and multi-disciplinary approaches involving diagnosticians, restorers, art historians, chemists and architects.

Monitoring activities of the involved sites, in particular, play a key role as they allow to study and analyze the causes, the effects and the progression of the decay of a target heritage site. Data acquisition, processing and maintenance are crucial issues in monitoring: in a typical scenario, surveyors access heritage sites with a number of different sensors, such as laser range finders or cameras, to collect all the required data. While this is acceptable for accessible sites, although it may require a very long time and many human resources, it limits the digitization of sites that are difficult to be accessed by humans.

The data gathered in this way is then processed in several different ways with the aim of extracting consistent and meaningful information: for example, accurate models of such sites, that are often a prerequisite for prevention, maintenance, restoration, and security tasks. In this phase, a lot of manual labor is currently

needed to align scans, to create annotations, or to reconstruct textured meshed models, due to the intrinsic complexity of the algorithms involved and to the possible incompleteness of data.

The models obtained in this way are often used only for the purposes of a single monitoring task, although they have a broader exploitability, both in terms of intended uses and from a temporal point of view. This is due to the difficulty of managing and sharing these models and of taking advantage of them without specific technical tools.

In the fields of robotics, artificial intelligence and virtual reality there has recently been a significant progress in the development of robotic and automated technologies, as well as visualization techniques that can support and improve the efficiency of this task from different points of view. Leveraging on these improvements, we elaborated a paradigm which aims at improving efficiency and safety of monitoring activities and exploitability of the results through the development of novel technologies for exploration, digitization and visualization of cultural heritage sites. Our solution has been designed and developed during a FP7 European Project, called ROVINA, that started in 2013 and is now successfully concluded. This project involved four academic partners (the University of Bonn, the Rheinisch-Westfälische Technische Hochschule of Aachen, the Katholieke Universiteit of Leuven, the Sapienza University of Rome), one company (Algorithmica s.r.l.) and the non-profit International Council of Monuments and Sites (ICOMOS).

Although our system has been thought for any indoor or underground archaeological site, that is difficult to be accessed by humans, or that is complex to be digitized using the common surveying strategies, we selected catacombs as case studies. Catacombs, indeed, are peculiar sites because: i) they are rich in both geometrical and texture features, such as loculi, chapels, tunnels, frescos and epigraphs; ii) they are challenging for navigability issues; iii) they often extend for several kilometres and at multiple depth levels. For example, the Roman Catacombs of Priscilla, that has been selected as our primary test site, extend for approximately 15km over multiple floors.

*Corresponding author

2. STATE OF THE ART

Conservation of cultural heritage sites is a practice that encompasses several disciplines ranging from material theory to structural engineering. Although it is not feasible to provide a complete overview of the state of the art, in order to place our system in context, we will provide a brief introduction to the major activities involved in the conservation of cultural heritage sites: measure, documentation, classification, and diagnostics.

Measuring is a key building block of any surveying activity. There are two types of measurements: “direct” and “indirect” ones. While direct measurements are performed directly (and manually) by the surveyor on the artefact, indirect measurements, on which our system is focused, entail the construction of a digital model of the involved artefact Martinelli (2006) in order to perform measures on it. Those models are usually obtained through the use of laser scanning and/or image analysis: both methods allow for morphometric surveys Drap et al. (2003).

Laser-based systems are common for architectural conservation Barber et al. (2006) and offer a wide variety of technologies (time-of-flight, phase modulation, optical triangulation, etc.) to adapt to the specific task (e.g., surveying large environments or digitizing small artefacts) and accuracy needed. These sensors can be rather expensive but offer a high precision and direct access to 3D information Johansson (2002). Image-based systems have been shown to provide precisions similar to those of Lidar scanners, especially when a sufficient number of images close to the surfaces can be taken. They have the advantage that image/color data are available, that are perfectly aligned with the 3D data. The captured data – as for lidars – produces point clouds that need further post-processing. Image-based systems usually employ high-resolution commercial cameras and commercial photogrammetry software or, more recently, self-calibrating structure-from-motion systems Vergauwen and Van Gool (2006); Theo Moons and Vergauwen (2008) or multi-view stereo approaches. These approaches are lower cost solutions compared to lasers, but can still provide high quality reconstructions especially for what concerns appearance.

Documentation is the goal of many surveys and aims at producing digital archives of the site under observation. In the realm of cultural heritage, documentation activities are performed by public bodies that, for example, in Italy usually are superintendencies and ministries. The digital archives can host contents in many different formats, including 3D models that can be either purely geometric or can also include textures from images. When the surveys have a purely descriptive documentary purpose, i.e., they are not intended for measurement or diagnosis, 3D models can have a lower resolution but they are, in general, more visually appealing. Their main goal is to disseminate cultural heritage to broad audiences, but also support archeological interpretation Berardin et al. (2005).

Classification activities are usually tied to documentation tasks and pertain to the categorization of elements of a site into taxonomies or ontologies with different degrees of complexity. For example, in an industrial context it would be interesting to categorize rooms of a plant – and equipment therein – based on their functional properties. In a cultural heritage site, architectural components are classified on the basis of a number of different parameters such as period of construction, used materials, state of conservation: this classification is usually performed by



Figure 2. Example of Table of Deterioration.

human users, who manually tag items and portions of the environment. When data is collected on a geographical scale, the models are generally archived into Geographical Information Systems (GIS). In such a scenario, data can be queried on both geographical and qualitative levels. For example, one may look for “all the sites built before 1000 B.C. in Italy” or “all the pots made of ceramic from Germany”.

Diagnostics has the goal to collect and analyze the information about the state of conservation of the surveyed areas in order to prevent damage or perform restoration. From a practical perspective, the diagnostic activities have the purpose of generating specific deliverables. Examples of such deliverables in the context of cultural heritage are the Table of Deterioration and the Table of Materials. The Table of Deterioration is a map showing possible deteriorations such as cracks and mold (see Fig. 2 for an example). In Italy, the table follows the UNI-NorMal commission standards, that is composed by the Ente Nazionale Italiano di Unificazione (UNI) and the Commissione NORmalizzazione Materiali Lapidei (NorMaL). At a European level, the standards are dictated by CEN (i.e., WS Construction, WS Measurement, WS Material). The Table of Materials instead maps the areas of the site under survey to the different materials of which they are composed.

3. THE ROVINA PARADIGM

Our project aims at improving the state of the art in measuring, documentation and classification (and thus indirectly supporting diagnosis activities) through a novel approach to surveying, data management and fruition based on three main components:

- **DigiRo**, an automated robot for collecting data with high-precision sensors, including laser scanners and cameras;
- **ARIS**, the cloud-based Archaeological Information System, to manage, share and elaborate data in the form of photo-realistic and metrically precise 3D models of the explored sites;
- **Web and VR Visualizers**, that allow to virtually navigate the 3D models through a very intuitive interface which also allows for an immersive experience.

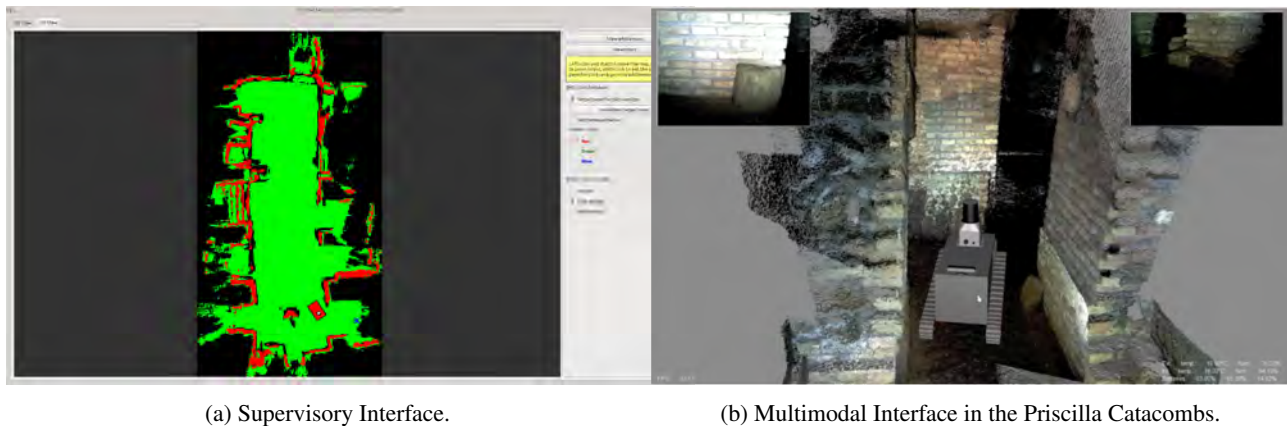


Figure 3. Mission Control interface Early Prototype.

3.1 DigiRo: the digitization robot

While there have been many attempts in the past to use robots to explore hard-to-access sites (see, e.g., Thrun et al. 2005), DigiRo is one of the first platforms specifically conceived for performing methodological surveys of unknown cultural heritage sites.

DigiRo development has been based on one hand on an iterative design process of the robotic platform and on the other hand on the integration and development of novel algorithms that extend the current state of the art in autonomous mapping and localization, 3D reconstruction and on-line analysis. During the last year of the project the robotic platform went through its third and last iteration of a continuous process that has been guided by the requirements dictated by the analyses of the environments being explored: we thus equipped a commercial tracked mobile base with a sensor suite composed by an inexpensive 3D laser range finder, three low-budget RGB-D cameras, an array of 7 high-quality RGB cameras, an inertial measurement unit, battery status monitors and thermal/humidity sensors.

In addition, the robot is also equipped with a distributed computation system that includes: two laptops, one for running the algorithms in charge of the autonomous behaviors and another for logging the data coming from the array of cameras. Three inexpensive single-board computers are also on-board, their role is to preprocess sensor data and to interface with the other hardware components of the robot.

In order to allow DigiRo to survey archaeological sites in an autonomous or semi-autonomous way, several challenges had to be addressed, and dedicated computer vision and robotics solutions have been developed. At the core of the platform intelligence there is the capability of simultaneously building a 3D map of the environment and localizing within this map (SLAM). DigiRo SLAM module extended different so-far state-of-the-art components such as g2o Kuemmerle et al. (2011), DCS Agarwal et al. (2014) and a variant of ICP Serafin and Grisetti (2015), thus obtaining a novel 3D localization and mapping approach. Compared to most existing SLAM methods, DigiRo SLAM module builds 3D maps in real-time during the survey allowing the robot to act autonomously and enabling its own decision-making based on the environment explored so far. A considerable effort has been devoted in making the approach more robust and developing an extension of DCS, originally proposed by Agarwal et al. Agarwal et al. (2014), to assessing the degree of consistency of maps Mazuran et al. (2014) and to automatically calibrating

the sensors (see also Basso et al. 2014; Tedaldi et al. 2014). In order to safely navigate in the environment, the robot uses an abstract 2D representation of the environment called traversability map Bogoslavskyi et al. (2013) that, when coupled with exploration techniques that consider the expected gain of novel information (see Stachniss and Burgard 2012 for an overview tutorial), allows for a safe and exhaustive survey of the environment.

Despite being still a research prototype, our robot has already obtained a number of achievements, showing that it has a good mobility and that it can run for a long time while processing huge amounts of data from a number of heterogeneous sensors. During the survey, while collecting data for off-line processing, the robot builds an on-the-fly mission-oriented reconstruction of the surrounding environment, that can be used either by the autonomous navigation system, but also by the surveyors to get a proper situation awareness remotely through the Mission Control Interface (MCI).

The MCI is the graphical interface used by surveyors during the missions at the cultural heritage sites¹. The MCI is composed of a multimodal interface (see Fig. 3b) and a supervisory interface (see Fig. 3a). The first one allows to visualize all the relevant data provided by the robot in an integrated way: video streams, 3D local and global reconstructions based on colored point clouds from the RGB-D sensors or from the 3D laser scanner, robot attitude, battery status, internal and external temperature and humidity, etc. The operator can change the point of view and inspect the explored environment or control the robot using a bird's-eye view. The operator can also select regions in the environment and annotate them for further analysis and classification. If the semantic segmentation module detects some interesting or known object, a marker is added to the map.

The supervisory interface, on the other hand, is used for mission control when there is little or no connectivity: a 2D representation of the environment is shown to the operator, where colors provide qualitative information about the terrain. The user can select target locations by clicking on the map, triggering the autonomous navigation behavior. When the connection with the operator is about to be lost, he can instruct the robot to explore a given region for a specified duration and report back the data collected: when the connection is available again, the robot will send back the traversability map and the markers on interesting locations, so that the operator can choose to download pictures of these locations and possibly request further analyses. DigiRo has

¹See video <http://youtu.be/16MKMsnyHJw>

already accomplished a number of successful missions and its results have been presented at the Maker Faire European Edition 2014 where it won the Maker of Merit award.

3.2 ARIS, the cloud-based Archaeological Information System

The amount of data gathered by the robot and the complexity of the reconstruction process is such that it cannot be efficiently handled by a single computer. For this reason, the data collected by the robot are uploaded to the cloud where ARIS, our information management system, processes it in order to offer a number of different services. Although there are already many interesting examples of archaeological information systems (e.g., the Arches Project² by the Getty Institute), these projects focus on descriptive artifacts where qualitative data, such as textual descriptions, is manually provided by human operators. On the contrary, our information system is focused on the management and automatic interpretation of large amounts of quantitative raw sensor data, such as laser scans, 3D images and pictures.

To this end, ARIS is capable of automatically generating accurate 3D models and to automatically classify data into semantic classes through the use of beyond-state-of-the-art Artificial Intelligence technologies. Indeed, 3D scans and images can be interpreted more effectively within aggregate 3D models than on their own. In ARIS we compute 3D reconstructions of two different types: 3D point clouds and textured 3D meshes. For example, Figure 1 shows a small portion of textured 3D meshed reconstruction of a catacomb that ARIS has computed from high-definition photos by using state-of-the-art photogrammetry approaches Vergauwen and Van Gool (2006); Theo Moons and Vergauwen (2008).

ARIS has been designed as a collaborative platform aiming at promoting cooperation among users with different types of expertise, during their conservation or analysis activities. To this end, it provides some useful facilities, such as dashboard, chat sessions and messaging systems in order to promote this social behavior.

The data archive in ARIS is organized in sites, each of which contains a number of datasets, that can be uploaded by users and pre-processed in the cloud, in order to provide search and visualization capabilities. Search can be performed in two ways: the first is provided through a classic GIS-like interface that allows to search for datasets using location-specific information (e.g., latitude and longitude) and showing them in a layer over a geographic map.

The second way to search through the data is using semantic queries: users can specify a query as a conjunction of criteria across a number of user-defined semantic layers (e.g., “all the epigraphs which are in the region of Lazio, which are made of marble and which are from the II Century”). These semantic queries are possible thanks to user-defined taxonomies and to an automated classification mechanism that are offered by ARIS. After defining the taxonomies, users can annotate a small number of images that are given as examples to ARIS machine learning technologies that will learn from those examples and then automatically annotate the rest of the datasets. Our approach to automatic classification combines Random-Forest-based classification with Conditional Random Fields Hermans et al. (2014) and has won the IEEE ICRA14 Best Vision Paper Award.

²<http://archesproject.org>



Figure 6. ROVINA VR Viewer prototype tested by visitors of the Maker Faire Rome 2015

3.3 3D Visualizers

The 3D model generated by ARIS is suitable for a multitude of applications. In addition to being functional to support documentation and monitoring activities, it has the great advantage of being visualizable on a plethora of devices (PC, smartphone, tablet, VR).

We developed a WebGL widget, the ROVINA Web Viewer. This is a tool to display and navigate the model in a web page. It can be used within ARIS or embedded in any other website. The user is then able to access the model via any web-capable device.

As the model is both accurate and captivating, we identified two main end-user branches.

The first branch of public is that of the technical personnel. Viewing and handling the metrically reliable model can support the planning and verification of surveys and interventions. A typical example is that of producing a graphically realistic digital version of a statue before and after a restoration work. The digital version of the statue not only serves as a comparison, but it also allows for infinitely detailed (yet easily consultable) annotations, and serves as a history diary of the statue itself.

The second branch of usage is that of the virtual touristic purposes. We leveraged on the recent spread of low-cost VR devices for smartphones (i.e. Google Cardboard, Samsung Gear VR and similar) by producing a mobile app that makes use of this technology, proposing an affordable, interactive and immersive experience.

We then presented the developed applications at major public events (i.e. Digital Heritage EXPO 2015, Maker Faire Rome 2015). The apps met a notable interest from the community.

In particular, getting in touch with people helped imagining a possible future usage. As high-end devices as “Oculus Rift”, “Playstation VR” and “HTC Vive” have recently made it to the market, bringing to the large public the opportunity of extremely detailed and smooth visualization experience of very large environments, it is envisionable the development of powerful VR applications that put the power of these tools to the service of technical personnel, who need high model detail, accuracy and longer time of use.

4. CONCLUSIONS

We discussed the main problems regarding digital reconstruction nowadays, which are mainly summarizable in:

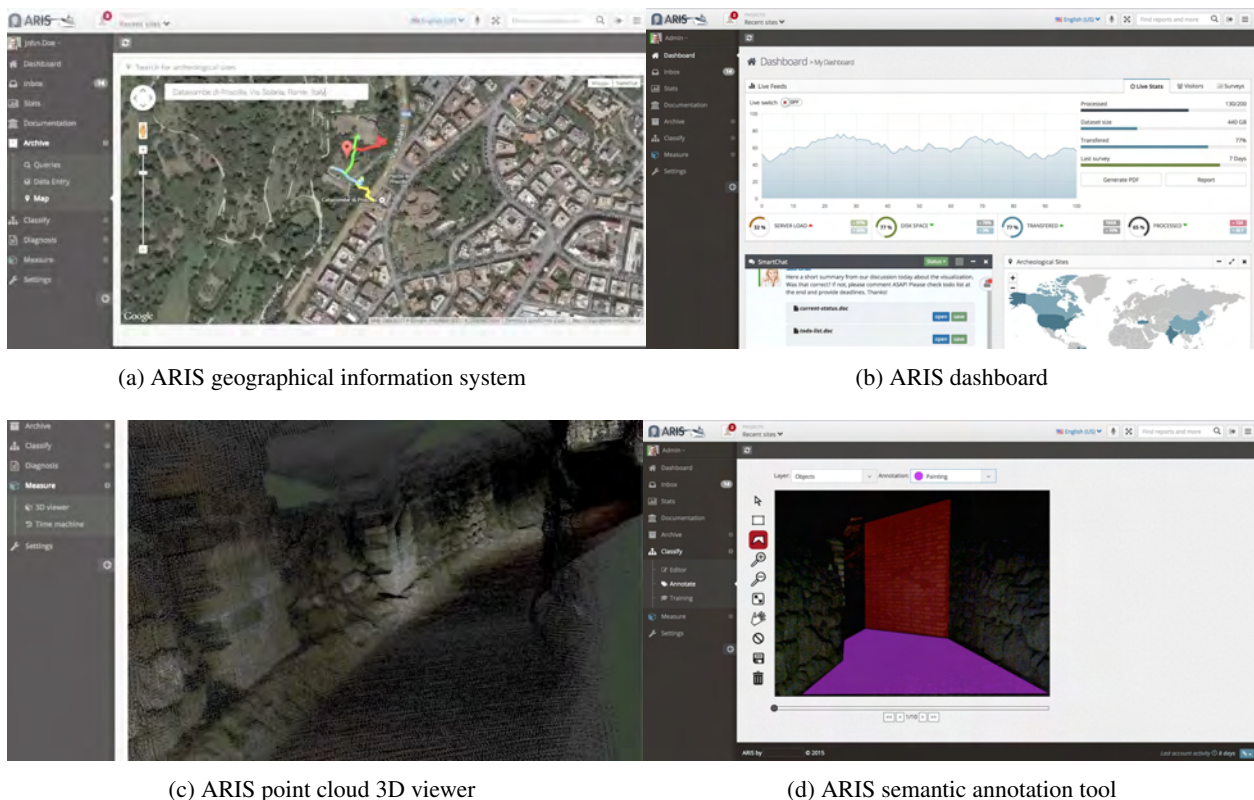


Figure 4. ARIS web interface

- high cost, in terms of time and resources, of the various phases involved in the process
- accessibility of the sites of interest
- reusability and ease of distribution of the products of the digitization

We then presented the solution proposed, designed and developed during the FP7 European Project "ROVINA", whose life time span has ranged from 2013 to 2016. The project introduced a strong automation component at several layers in the process, from the involvement of a robotic platform for acquiring the data to the automation of the processing and distribution of these data and the products of their exploitation.

We highlighted and described the technological components included and eventually made an overview of the public reaction to the project, which revealed an opportunity of putting the technology once more at the service of culture and restoration.

ACKNOWLEDGEMENTS

This work has partly been supported by the European Commission under contract number FP7-600890-ROVINA.

References

- Agarwal, P., Grisetti, G., Tipaldi, G., Spinello, L., Burgard, W. and Stachniss, C., 2014. Experimental analysis of dynamic covariance scaling for robust map optimization under bad initial estimates. In: *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, Hong Kong, China.
- Barber, D. M., Dallas, R. W. and Mills, J. P., 2006. Laser scanning for architectural conservation. *Journal of Architectural Conservation* 12(1), pp. 35–52.
- Basso, F., Pretto, A. and Menegatti, E., 2014. Unsupervised intrinsic and extrinsic calibration of a camera-depth sensor couple. In: *Proc. of: IEEE International Conference on Robotics and Automation (ICRA)*, pp. 6244–6249.
- Beraldin, J.-A., Picard, M., El-Hakim, S. F., Godin, G., Valzano, V., Bandiera, A. et al., 2005. Combining 3 d technologies for cultural heritage interpretation and entertainment. In: *Proc. SPIE*, Vol. 5665, pp. 108–118.
- Bogoslavskyi, I., Vysotska, O., Serafin, J., Grisetti, G. and Stachniss, C., 2013. Efficient traversability analysis for mobile robots using the kinect sensor. In: *Proceedings of the European Conference on Mobile Robots (ECMR)*.
- Drap, P., Sgrenzaroli, M., Canciani, M., Cannata, G. and Seinturier, J., 2003. Laser scanning and close range photogrammetry: Towards a single measuring tool dedicated to architecture and archaeology. In: *CIPA XIXth International Symposium*, pp. 1–6.
- Hermans, A., Floros, G. and Leibe, B., 2014. Dense 3d semantic mapping of indoor scenes from rgb-d images. In: *Robotics and Automation (ICRA), 2014 IEEE International Conference on*, pp. 2631–2638.
- Johansson, M., 2002. Explorations into the behaviour of three different high-resolution ground-based laser scanners in the built environment. In: *Proceedings of the CIPA WG 6 International Workshop on Scanning for Cultural Heritage Recording*.
- Kuemmerle, R., Grisetti, G., Strasdat, H., Konolige, K. and Burgard, W., 2011. g2o: A general framework for graph optimization. In: *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China, pp. 3607–3613.



Figure 5. (Left picture) navigation through the Catacombs of Priscilla 3D virtual reconstruction made during ROVINA; (right picture) photos of the same area inside the real environment.

- Martinelli, M., 2006. Passato e futuro del 3d archeologico, dalle foto stereoscopiche al computer per l'architettura antica. *Arkos* (16), pp. 18–23.
- Mazuran, M., Tipaldi, G., Spinello, L., Burgard, W. and Stachniss, C., 2014. A statistical measure for map consistency in slam. In: *Proc. of the IEEE Int. Conf. on Robotics & Automation (ICRA)*, Hong Kong, China.
- Serafin, J. and Grisetti, G., 2015. NICE: Dense normal based point cloud registration. In: *To appear in Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, Hamburg, Germany.
- Stachniss, C. and Burgard, W., 2012. Particle filters for robot navigation. *Foundations and Trends in Robotics* 3(4), pp. 211–282. Published 2014.
- Tedaldi, A., Pretto, A. and Menegatti, E., 2014. A robust and easy to implement method for imu calibration without external equipments. In: *Proc. of: IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3042–3049.
- Theo Moons, L. V. G. and Vergauwen, M., 2008. 3d reconstruction from multiple images part 1: Principles. *Foundations and Trends in Computer Graphics and Vision* 4(4), pp. 287–404.
- Thrun, S., Thayer, S., Whittaker, W., Baker, C., Burgard, W., D. Ferguson, D. H., Montemerlo, M., Morris, A., Omohundro, Z., Reverte, C. and Whittaker, W., 2005. Autonomous exploration and mapping of abandoned mines. *IEEE Robotics and Automation*.
- Vergauwen, M. and Van Gool, L., 2006. Web-based 3d reconstruction service. *Mach. Vision Appl.* 17(6), pp. 411–426.