See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/352936161

A SLAM Integrated Approach for Digital Heritage Documentation

Chapter · July 2021 DOI: 10.1007/978-3-030-77411-0_3





A SLAM Integrated Approach for Digital Heritage Documentation

Salvatore Barba^(⊠), Carla Ferreyra, Victoria Andrea Cotella, Andrea di Filippo, and Secondo Amalfitano

Dipartimento di Ingegneria Civile, Università degli Studi di Salerno, Fisciano, Italy sbarba@unisa.it

Abstract. The digital acquisition of Cultural Heritage is a complex process, highly depending on the nature of the object as well as the purpose of its detection. Even if there are different survey techniques and sensors that allow the generation of realistic 3D models, defined by a good metric quality and a detail consistent with the geometric characteristics of the object, an interesting goal could be to develop a unified treatment of the methodologies. Villa Rufolo, with its intricate articulation, becomes the benchmark to test an integrated protocol between photogrammetry, Terrestrial Laser Scanning (TLS) and a Wearable Mobile Laser System (WMLS) based on a SLAM approach. To quantify the accuracy of the latter solution, a comparison is proposed. For the case study the ZEB1, produced and marketed by GeoSLAM, is tested. Computations of cloud-to-cloud (C2C) absolute distances is performed, using stationary laser scanner (Faro Focus^{3D} X130) as a reference. Finally, the obtained results are reported, allowing us to assert that the quality of the WMLS measurements is compatible with the data provided by the manufacturer, thus making the instrumentation suitable for certain specific applications.

Keywords: Indoor mapping · Laser scanning · Photogrammetry

1 Introduction

Cultural Heritage can be defined as a living memory of our society, an irreplaceable testimony of a particular moment in human history. The guidelines for its conservation and enhancement, codified in the Athens Charter and repeatedly reiterated by subsequent documents up to the most recent UNESCO Recommendations, underline the importance of multidisciplinary and scientific approaches for the management of interventions in cultural heritage sites.

The digital survey plays a key role in their documentation; it provides an interesting and innovative scientific basis for study and research, as well as ensuring an effective dissemination approach even for a non-technical audience.

Currently, the use of new technologies for data acquisition in the architectural field has reached a wide diffusion, mainly due to the ability to digitize artifacts with great accuracy and the possibility to generate information models useful for the phases of analysis, simulation, and interpretation [1, 2]. The most widespread techniques, which have

[©] Springer Nature Switzerland AG 2021

M. Rauterberg (Ed.): HCII 2021, LNCS 12794, pp. 27–39, 2021. https://doi.org/10.1007/978-3-030-77411-0_3

now become a reference standard, are modern photogrammetry and laser scanning. Photogrammetry acquires two-dimensional images that require mathematical processing to derive 3D information. Through precise formulations based on projective or perspective geometry, it transforms the data extracted from the images into three-dimensional metric coordinates with colour information [3]. For its part, laser scanning can directly obtain the spatial position of the 3D point [4, 5] with high accuracy and without need of illumination conditions, especially useful for homogeneous surfaces where photogrammetry cannot provide reliable results. Integration is then a result that overcomes the limitations of each technique.

The main products obtained from both 3D point cloud and 2D orthoimage techniques have been used for the virtual reconstruction of cultural heritage sites, as the analysis of cave paintings [6], the creation of accurate numerical simulations, or the analysis of pathological processes [7], among others.

In addition to the wide range of advantages that these solutions can offer, the digitization of large and complex areas, especially indoor scenarios, generally involves the use of many images (in the case of photogrammetry) or scanning stations (in the case of laser scanning), resulting in time-consuming fieldwork and moreover in an important propagation of errors [8]. Hybrid solutions, such as mobile mapping systems (MMS), have emerged with great capabilities and possibilities in recent years, allowing the management of multiple sensors and the ability to operate in complex outdoor and indoor scenarios [9–13].

The main purpose of this paper is to highlight the advantages of an integrated approach to digital heritage documentation, resulting from the combination of techniques and technologies that stand out for their strengths but also for their criticalities [14]. Villa Rufolo (Fig. 1), a cultural symbol of the city of Ravello on the Amalfi coast, represents the test case of the experience conducted. It was built between the twelfth and thirteenth centuries as a family residence and a material representation of social status. It has almost unique architectural features, which blend Arab-Byzantine typologies and ornaments with elements of local culture. In the period of maximum splendour, it is said that the Villa had "more rooms than days of the year", although today it is possible to appreciate only some parts of the original construction, such as the Moorish cloister, the entrance tower, and the main tower whose 30 m in height represented the family's prestige.

In detail, this paper evaluates the suitability of a wearable mobile laser system (WMLS) for the digitalization of a complex indoor environment belonging to a cultural heritage building, as well as for the generation of cartographic products required for its conservation and restoration. This wearable system combines laser scanning technology and an inertial measurement unit (IMU) in portable equipment that can be handled by an operator while walking through the cultural heritage site. This sensor acquires point clouds on the move, thanks to the Simultaneous Localization and Mapping algorithms (SLAM) [10, 14], without needing the support of a global navigation satellite system (GNSS).

During this evaluation, possible integration strategies are examined and tested to guarantee a result that is compatible with the levels of accuracy and definition required by the survey, trying to optimise the procedure and make it applicable in other contexts



Fig. 1. Villa Rufolo, general view of the Palace with southern gardens.

as well. The most innovative aspect of the proposed application is the integration of data from diachronic surveys, differentiated not only by the technique used but also by a different time of detection. The joining element that makes their fusion possible is, however, the presence of a GNSS support network that also makes it possible to control the quality of the output model.

2 Materials and Methods

2.1 Equipment

GeoSLAM ZEB1. The ZEB1 consists of a 2D time-of-flight laser range scanner rigidly coupled to an inertial measurement unit (IMU) mounted on a spring. The motion of the scanning head on the spring provides the third dimension required to generate 3D information. A simultaneous localization and mapping (SLAM) algorithm combines the 2D laser scan data with the IMU data to generate accurate 3D point clouds, employing a full SLAM approach [15, 16]. Regarding accuracy, the manufacturer declares a value of 0.1% in relative accuracy for a 10-min scan, with a single loop closing. Special thanks are due to engineer Mariella Danzi, who kindly provided us with the instrumentation. It is worth mentioning that the acquisitions were performed in May 2019, when more advanced versions of the instrument, able to offer greater control over error propagation, were not yet available.

DJI Phantom 4. The UAV has an approximate weight of 1.4 kg. Its camera is equipped with a 12 MP Sony Exmor sensor (size 6.3×4.7 mm, pixel size 1.56μ m), and a wide-angle lens with a 4 mm focal length and FOV (Field of View) of 94°. The camera is integrated in a gimbal to maximize the stability of the images during the movements.

Faro Focus^{3D}X130. The X130 is a stationary laser scanner of the Continuous Wave -Frequency Modulation (CW-FM) type. The system can measure with great precision the direction of pointing, in addition to a distance meter that emits continuous light radiation. This one, thanks to a coding of the frequency modulated light signal that allows the identification of a phase shift between the emitted wave and the recorded one, guarantees the indirect calculation of the time of flight and therefore of the distance. In terms of error, this solution guarantees a systematic component of ± 2 mm and an accidental component of ± 0.5 mm.

2.2 Survey Design and Detection

WMLS. Systems that use SLAM algorithms to generate final models require special attention when planning acquisitions. More than any other solution, in fact, the quality of the produced data is highly dependent on how the acquisition campaign is conducted. It is important to inspect the site of interest to identify critical areas not detected during planning and remove any obstacles along the way. In addition to the focus on poorly referenced environments, transition areas and forward speeds, it should be remembered that full SLAM systems, such as the one employed, require the self-intersection of paths to ensure an appropriate redistribution of the accumulated errors. At least one loop must be closed, although it is advisable to plan routes with several self-intersections. Geometric features extracted by the algorithm are significant if the ratio of their size and their range is approximately 1:10 (e.g., a feature must be textgreater 0.5 m in size for a distance of 5 m). In addition, if there are not sufficient features along the direction of travel, the SLAM algorithm cannot correctly determine forward motion. In general, it is better to do circular loops rather than "there and back" loops where the path simply doubles back on itself. It is important to scan the closed loop regions carefully to ensure that the key features are scanned from a similar perspective. It may be necessary to turn around to return to a region from another direction. This is a crucial feature in poor environments. With these concepts in mind, the acquisition campaign is organised in this way (Fig. 2):

- *First path*, which includes the management offices on level 0 and the auditorium on level 1.
- *Second path*, covering the entrance at level 0, the west garden, the central courtyard and the east garden.
- *Third path*, which covers storage areas and rooms closed to the public located on level 1, as well as part of the east garden.
- *Fourth path*, which takes place on the underground level, going through the lower part of the Moorish cloister up to the exhibition rooms in the current "theatre".
- *Fifth path*, that covers all the "museum" rooms and the auditorium on level 1.

Digitizing the entire scene has taken 2 h, about 1/2 of the time needed to complete the TLS campaign, but covering a larger area.

Aerial Photogrammetry. To control the metric error, 14 GCP are detected on the arena floor by a Geomax Zenith 25 used in nRTK mode. The accuracy of planimetry is below

1 cm and 2.5 cm for altimetry. For the acquisition of the frames, two flight are prepared, both automatic and with double grid: a first one for the acquisition of nadir photogrammetric images and a second one, with the optical axis tilted about 45°, to survey the vertical walls and any shadow cones.

The flight lines are designed using the DJI Ground- Station software package. For all the grids, the UAV is set to a target altitude of 16 m above take off point - Torre Maggiore - (46 m from Duomo Square) and horizontal ground speed of 4.0 m \cdot s⁻¹. The height is calculated in the DJI Ground- Station software using elevation data derived from Google Earth. Parallel flights lines are programmed to have an image overlap of 60% as well as sidelap of 60%, setting the proper camera parameters (dimensions of the sensor, focal length, and flight height). In the nadir flights, 93 and 94 images are acquired for the first (from North to South) and second (from West to East) grid, respectively. Two other flights, with the camera tilted at 45° on the horizontal plan, are carried out acquiring, respectively, other 92 and 163 photos. The image acquisition is planned bearing in mind the project requirements - a Ground Sampling Distance (GSD) of about 1 cm - and, at the same time, with the aim of guaranteeing a high level of automation in the following phases.



Fig. 2. Registered SLAM paths, highlighted with different colors: the first in purple, the second in pink, the third in red, the fourth in grey and the fifth in green. (Color figure online).

TLS. A survey design should be first defining the positions of TLS stations, so that the whole object coverage at requested spatial resolution could be guaranteed. The instrument has been set to have a resolution of 6 mm–10 m. As far as the quality of the acquisition is concerned, for each point of the cloud three measurements are made to define the distance from the station as the average of the above measurements. Great

attention has also paid to the problem of environmental occlusions, due to the presence of vegetation. In the first instance, 5 scans are acquired in the west garden, 4 at the level of the central courtyard, and 5 in the east garden. The need to connect these three parts has required additional stations to ensure the completeness of the model: 3 of them are placed between the west garden and the courtyard and 3 between the latter and the east garden. With the features listed above, the TLS acquisition campaign has taken almost 5 h, with a single-scan time of approximately 12 min.

2.3 Processing: The Generation of the 3D Point Cloud

High-level, point-based approach to data fusion has been opted for, where all raw data streams are kept separate and processed independently. Only at the end the resulting point clouds are merged to obtain a complete 3D model. The processing of the raw data, coming from the acquisition phase, starts from the registration of the point clouds [17, 18] (Fig. 3).

WMLS Data Registration. Due to the low overlap between the five acquisition paths, a progressive registration approach based on an ICP pairwise algorithm is employed, each time choosing the reference scan. All preceded by a manual raw alignment. The maximum value of the RMSE on all the registration pairs, defined above, is about 1.74 cm. The idea of defining a pipeline for the registration of WMLS data that does not resort to TLS scans arises from the observation that it is not always possible to have homologous models deriving from different systems to perform accuracy checks. For this reason, the workflows for the two systems are kept separate until comparison. The individual paths are maintained disjoined (but rigidly bound to each other) for a comparison with the final TLS model, after mutual registration with ICP algorithm.

Photogrammetric Process. Data treatment is performed by Agisoft Metashape, 1.5.3 build 8469 version. Its workflow is based on four steps: Align Photos, Build Dense Cloud, Build Mesh and Build Texture. At the first step an algorithm evaluates the camera internal parameters (Focal Length, position of the principal point, radial and tangential distortions), the camera positions for each photo and the Sparse Cloud. In the next phase, a greater pixel number is re-projected for each aligned camera, creating the Dense Cloud. In the Build Mesh step, it is possible to generate a polygonal mesh model based on the dense cloud data. Finally, the polygon model is textured in the Build Texture step. The outputs of the photogrammetric model, necessary for further documentation studies and data integration with active sensors, are a nadir orthophoto of the entire villa and the dense point cloud. The extracted point cloud has more than 48 million points, with average errors of about 2.8 cm (Fig. 4).

TLS Scan Registration. These data are used to create a reference network (our ground truth), indispensable for performing a quality assessment on the WMLS. The clouds are characterized by a high degree of overlap and for this reason are registered employing a global bundle adjustment procedure, accomplished after a top view-based preregistration. Given the set of scans, the algorithm searches for all the possible connections between the pairs of point clouds with overlap. For each connection, a pairwise ICP is performed and the best matching point pairs between the two scans are saved.



Fig. 3. High-level integrated point cloud from different sources: SLAM in grey, photogrammetry in pink and TLS in yellow. (Color figure online).



Fig. 4. Orthophoto showing the homogeneous distribution of the 14 GCPs, with an elevation difference between P4 and P9 of about 19 meters (by Marco Limongiello).

A final non-linear minimization is run only among these matching point pairs of all the connections. The global registration error of these point pairs is minimized, having as unknown variables the scan poses [19]. To define a common reference system for all the items produced, the coordinates of the 14 GCPs used for photogrammetric orientation are imported as external references into the TLS station registration process. The RMSE on the registration is about 3.4 mm and the maximum value is 32.4 mm.

2.4 Accuracy Assessment

The comparison between homologous models, produced with the TLS technique and the SLAM approach, is performed according to different modalities. Before proceeding, it may be useful to identify the error components when comparing point clouds:

- The component of position of the cloud, depending on the technology used in the acquisition phase.
- The component of registration among the point clouds, depending on the technique used to define a common reference system.
- The errors depending on occlusions and on the process of discretization of the survey.

Once the error components have been defined, some algorithms are selected to perform the comparison. The main ones are presented below.

Direct C2C Comparison with Closest Point Technique. This method is the simplest and fastest direct 3D comparison method of point clouds, as it does not require gridding or meshing of the data, nor calculation of surface normal vectors. For each point of the analysed cloud, a closest point can be defined in the reference. In its simplest version, the surface change is estimated as the distance between the two points [20]. An improvement may consist of local modelling of the reference mesh. This technique is also used in cloud matching techniques such as the ICP. The difference lies in the fact that the computation of the closest point for the ICP is performed only on sample used to construct the matching pairs. In the case of a comparison, instead, all the points of the analysed cloud are considered. This type of distance is sensitive to the cloud roughness, outliers and point spacing. For this reason, the technique is developed for rapid change detection on very dense point clouds (like out TLS model) rather than accurate distance measurement.

3 Results

3.1 C2C Absolute Distance Computation

This analysis involves the selection of the SLAM path that shows the greatest overlap with the fused TLS cloud, used as a reference for the subsequent C2C comparison. For the case study, no local modelling is used due to the high surface density of the TLS model, but simply the absolute C2C distance is calculated. The designated route is number two, which runs from the west garden to the east garden via the central courtyard.

The purpose of this operation is to verify whether, along the acquisition path, drifts from the reference cloud had occurred, typical of solutions based on SLAM approach. To this end, three portions are extracted from route two (Fig. 5): A - east garden, starting and ending point of the path; B - central courtyard; C - east garden. A registration of the whole route on the TLS model, with ICP, would produce a partial compensation of possible drifts. It was therefore decided to limit the registration to portion A, which is rich in geometric elements easily recognisable by the SLAM algorithm and therefore less subject to deformations. The path selection itself is not random. The designated path is in fact characterised by a very limited number of loops, with a prevalence of "round trips" on the same path, a possible cause of drift (Fig. 6).



Fig. 5. Identification of portions extracted from the second path to perform the comparison. To avoid the partial compensation of path deformations, produced by a global registration on the TLS model, it was preferred to perform the alignment using only the portion A.

A mean distance value of 2.67 cm is obtained for portion A (corresponding to the mean absolute error - MAE - value). The value of the MAE, which defines the accuracy in terms of magnitude, it progressively increases to 14.81 cm for portion B and finally 32.73 cm for portion C. This shows a progressive accumulation of drift, which can be attributed to various causes. Some of the scanned environments are repetitive and characterised by architectural elements with very similar dimensions (such as the gallery that connects the west garden to the central courtyard) and this factor can greatly condition the performance of the SLAM algorithm.



Fig. 6. C2C distance for the analysed portions, with accumulation of error along the path.

4 Conclusions

In this paper, the quality of point clouds acquired by a wearable mobile mapping system, the handheld GeoSLAM ZEB1, is tested in a controlled environment. Quantitative and qualitative analysis of the point clouds are performed using the point cloud of a Faro Focus^{3D} X130 as reference. The main purpose of the study is not, in fact, to define a general integration methodology, but to understand how SLAM fits into a process that includes established techniques such as photogrammetry and TLS. To do so, it is necessary to go through the evaluation of errors, in this case the drift that the mobile system may manifest along the acquisition path.

To further corroborate the proposed analysis, for this case study, data from a first version of the instrument were used, which can only be processed with a raw algorithm that cannot make use of external control points to check the results (Fig. 7).



Fig. 7. Cloud section integrated with optimized SLAM, TLS, and photogrammetric data.

The computed MAE, which defines the accuracy in terms of magnitude, is compatible with the limits of accuracy declared by the manufacturer for a 10-min path with a single loop closure, equal to 30 cm. Numerical calculations are always accompanied by a visual inspection to prove the results.

The main novelty of the proposed experiment is found precisely in the methodology of checking and quantifying, through the comparison of point clouds, the limitations that afflict the SLAM algorithms if they are not controlled by external elements, which may be coordinates of points detected by other techniques. For this aspect in particular, the paper is addressed to technicians working in the field of surveying and the AEC (Architectural, Engineering and Construction) industry. The integration of techniques and technologies characterised by different accuracies is in fact a very topical issue. It is worth remembering that rigorous data acquisition and processing produce a strong impact also on the activity of the other stakeholders involved in heritage documentation, constituting the basis for future decision making in the process of managing the building's life cycle.

Currently, digital innovations, and in particular SLAM, provide greater agility in the process and make it possible to generate 3D restitutions that are more accurate, faster, interoperable, and combinable between different methods. With this application it was possible to systematise the results produced through the application of three different digital survey techniques (WMLS, photogrammetry, TLS), applied at different times and by different operators.

The applications of SLAM are not only limited to heritage documentation in combination with other techniques but, considering the distinctive features of the methodology, can range from expeditious mapping in emergency conditions to the production of models for the dissemination of content to a non-technical audience, for example the generation of a digital environment through the application of VR/ and A/R where the user is fully immersed.

Therefore, future work will focus on improving the statistical analysis of errors and on the creation of a structured database from the fusion of different sensor data by performing an analytical review of the latest advances and potentialities in this field.

References

- Garcia-Gago, J., Gomez-Lahoz, J., Rodríguez-Méndez, J., González-Aguilera, D.: Historical single image-based modeling: the case of gobierna tower, Zamora (Spain). Remote Sens. 6(2), 1085–1101 (2014)
- Allegra, V., Di Paola, F., Lo, M., Vinci, C.: Scan-to-BIM for the management of heritage buildings: the case study of the castle of maredolce (Palermo, Italy). ISPRS - Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci. XLIII-, 1355–1362 (2020). https://doi.org/10. 5194/isprs-archives-XLIII-B2-2020-1355-2020
- Remondino, F., El-Hakim, S.: Image-based 3D modelling: a review: image-based 3D modelling: a review. Photogram. Rec. 21(115), 269–291 (2006). https://doi.org/10.1111/j.1477-9730.2006.00383.x
- Francois, B.: Review of 20 years of range sensor development. J. Electron. Imaging 13(1), 231–243 (2004). https://doi.org/10.1117/1.1631921
- Guidi, G., Beraldin, J.A.: Acquisizione 3D e modellazione poligonale: dall'oggetto fisico al suo calco digitale. Poli. Design 2004, 189 (2004)
- Torres-Martínez, J.A., Sánchez-Aparicio, L.J., Hernández-López, D., González-Aguilera, D.: Combining geometrical and radiometrical features in the evaluation of rock art paintings. Digital Appl. Archaeol. Cul. Heritage 5, 10–20 (2017). https://doi.org/10.1016/j.daach.2017. 04.001
- Sánchez-Aparicio, L.J., Del Pozo, S., Ramos, L.F., Arce, A., Fernandes, F.M.: Heritage site preservation with combined radiometric and geometric analysis of TLS data. Autom. Constr. 85, 24–39 (2018). https://doi.org/10.1016/j.autcon.2017.09.023

- Guidi, G., Beraldin, J.-A., Ciofi, S., Atzeni, C.: Fusion of range camera and photogrammetry: a systematic procedure for improving 3-D models metric accuracy. IEEE Trans. Syst. Man Cybern. Part B (Cybernetics) 33(4), 667–676 (2003). https://doi.org/10.1109/TSMCB.2003. 814282
- Al-Hamad, A., El-Sheimy, N.: Smartphones based mobile mapping systems. Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci. XL-5, 29–34 (2014). https://doi.org/10.5194/isp rsarchives-XL-5-29-2014
- Limongiello, M., Ronchi, D., Albano, V.: BLK2GO for DTM generation in highly vegetated area for detecting and documenting archaeological earthwork anomalies. In: Proceeding of Metroarchaeo - Metrology for Archaeology and Cultural Heritag (2020)
- di Filippo, A., Sánchez-Aparicio, L.J., Barba, S., Martín-Jiménez, J.A., Mora, R., González-Aguilera, D.: Use of a wearable mobile laser system in seamless indoor 3D mapping of a complex historical site. Remote Sens. 10(12), 1897 (2018). https://doi.org/10.3390/rs1012 1897
- Ortiz-Coder, P., Sánchez-Ríos, A.: An integrated solution for 3D heritage modeling based on videogrammetry and V-SLAM technology. Remote Sens. 12(9), 1529 (2020). https://doi.org/ 10.3390/rs12091529
- Di Stefano, F., Gorreja, A., Malinverni, E.S., Mariotti, C.: Knowledge modeling for heritage conservation process: from survey to HBIM implementation. Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci. XLIV-4/W1-2020, 19–26 (2020). https://doi.org/10.5194/isprs-arc hives-XLIV-4-W1-2020-19-2020
- Mandelli, A., Achille, C., Tommasi, C., Fassi, F.: Integration of 3D models and diagnostic analyses through a conservation-oriented information system. Int. Arch. Photogram. Remote Sens. Spatial Inform. Sci. XLII-2/W1-2017, 497–504 (2017). https://doi.org/10.5194/isprsarchives-XLII-2-W5-497-2017
- Bosse, M., Zlot, R., Flick, P.: Zebedee: design of a spring-mounted 3-D range sensor with application to mobile mapping. IEEE Trans. Rob. 28(5), 1104–1119 (2012). https://doi.org/ 10.1109/TRO.2012.2200990
- Grisetti, G., Kümmerle, R., Stachniss, C., Burgard, W.: A tutorial on graph-based SLAM. IEEE Intell. Transp. Syst. Mag. 2(4), 31–43 (2010). https://doi.org/10.1109/MITS.2010.939925
- 17. Besl, P., McKay, N.: Method for registration of 3-D shapes, vol. 1611. Robotics '91. SPIE, (1992)
- Chen, Y., Medioni, G.: Object modelling by registration of multiple range images. Image Vis. Comput. 10(3), 145–155 (1992). https://doi.org/10.1016/0262-8856(92)90066-C
- Santamaría, J., Cordón, O., Damas, S.: A comparative study of state-of-the-art evolutionary image registration methods for 3D modeling. Comput. Vis. Image Underst. 115(9), 1340–1354 (2011). https://doi.org/10.1016/j.cviu.2011.05.006
- Lague, D., Brodu, N., Leroux, J.: Accurate 3D comparison of complex topography with terrestrial laser scanner: application to the Rangitikei canyon (N-Z). ISPRS J. Photogramm. Remote. Sens. 82, 10–26 (2013). https://doi.org/10.1016/j.isprsjprs.2013.04.009