

## A MODULAR MOBILE MAPPING PLATFORM FOR COMPLEX INDOOR AND OUTDOOR ENVIRONMENTS

I. Kalisperakis<sup>1</sup>, T. Mandilaras<sup>1</sup>, A. El Saer<sup>1</sup>, P. Stamatopoulou<sup>1</sup>, C. Stentoumis<sup>1</sup>, S. Bourou<sup>1</sup>, L. Grammatikopoulos<sup>2</sup>

<sup>1</sup> up2metric P.C., Engineering - Research - Software Development, Michail Mela 21, GR-11521, Athens, Greece - (ilias, theodor.mandilaras, andreas.saer, panagiota.stamatopoulou, stavroula.bourou, christos)@up2metric.com

<sup>2</sup> Department of Surveying and Geoinformatics Engineering, University of West Attica, Ag. Spyridonos Str., Egaleo, GR-12243, Athens, Greece – lazaros@uniwa.gr

Commission I, WG I/6, WG I/7

**KEY WORDS:** Mobile, Mapping, Laser Scanning, Lidar, Data fusion, Calibration, 3D reconstruction

### ABSTRACT:

In this work we present the development of a prototype, mobile mapping platform with modular design and architecture that can be suitably modified to address effectively both outdoors and indoors environments. Our system is built on the Robotics Operation System (ROS) and utilizes multiple sensors to capture images, pointclouds and 3D motion trajectories. These include synchronized cameras with wide angle lenses, a lidar sensor, a GPS/IMU unit and a tracking optical sensor. We report on the individual components of the platform, its architecture, the integration and the calibration of its components, the fusion of all recorded data and provide initial 3D reconstruction results. The processing algorithms are based on existing implementations of SLAM (Simultaneous Localisation and Mapping) methods combined with SfM (Structure-from-Motion) for optimal estimations of orientations and 3D pointclouds. The scope of this work, which is part of an ongoing H2020 program, is to digitize the physical world, collect relevant spatial data and make digital copies available to experts and public for covering a wide range of needs; remote access and viewing, process, design, use in VR etc.

### 1. INTRODUCTION

3D mapping platforms are a core component for many and diverse workflows and are becoming increasingly useful the last years. There is a growing interest for platforms that allow accurate, but also fast and massive capturing of data for 3D mapping. Many commercial solutions are offered as products or services that target the recording of either outdoors or indoors environments. They usually rely on a combination of modern geospatial technologies such as laser scanning, GNSS navigation and photogrammetry. They come in different forms; systems mounted on cars, trolleys, backpacks, or even autonomous robots. Such systems can capture effectively large urban areas, or buildings and construction sites, but they remain very expensive due to the top-end hardware components they rely on.

Urban planning and public infrastructure management are application areas that benefit from such technologies as they require constantly updated geographic information. Mobile mapping technologies have been widely used in a variety of applications in urban areas, for mapping transportation infrastructure, utilities, buildings, vegetation and lately for autonomous vehicle driving (Shi et al., 2017). A recent survey of such applications for lidar based mobile mapping is presented in the work of Wang et al. (2019). Real Estate, and Architecture, Engineering, Construction (AEC) sectors are also adopting digitization. Building Information Modelling (BIM) and Geographic Information Systems (GIS) are becoming standard tools that handle large amounts of geospatial data.

Besides specialists, the general public is also daily consuming mobile mapping data through online tools or mobile apps like *Google Street View* (Anguelov et al., 2010). Of special interest is also the case of *Mapillary*, a collaborative alternative of Google Street View that allows users not only to access but also

to capture street level videos or image sequences with any camera and upload them on a map.

The development of new, more versatile mobile mapping systems is expected to grow due to i) an abundance of new medium/low cost sensors that are widely produced for the mobile phone and the automotive industry, and ii) constant advancement of the underlying methods and technologies from the robotics and autonomous navigation communities.

The work presented here is part of an ongoing European H2020 STARTS Research Program called “Mindspaces” that aims to utilize 3D mapping among other technologies, towards art-driven adaptive outdoors and indoors design (Alvanitopoulos et al., 2019). The scope of our research is to provide a tool that captures multiple types of relevant spatial data of the environment such as raw video footage, georeferenced imagery, pointclouds etc. These can be subsequently exploited by designers and artists to collaborate with scientists and engineers towards the creation of innovative designs and experiences.

In the following sections, after a short review of related work, we present the individual sensors and components of our platform, we describe their integration within the Robotics Operation System (ROS) and discuss processing workflows for generating 3D reconstructions from the collected data. Initial experiments are also presented.

### 2. RELATED WORK

#### 2.1 Mobile Mapping Systems

3D laser scanners, photogrammetry and surveying have been the typical means for 3D recording of physical world and manmade constructions. The scientific and technological advances during the last decade have made possible the

adaptation in everyday use of much more scalable approaches of data capturing for 3D reconstruction. In this context, during the last years, several mobile mapping platforms are available in the market (Puente et al., 2013). These fall into two main categories, i) those that are suitable for mapping of large-scale outdoor environments and ii) those suitable for indoor scenes.

### 2.1.1 Outdoor mapping

Most major players in the geospatial market offer similar systems, like UltraCam Mustang by VEXCEL (VEXCEL, 2020), RIEGL (RIEGL, 2020), LEICA Pegasus by HEXAGON (LEICA, 2020) and VIAMETRIS (VIAMETRIS, 2020). All these systems combine proprietary high-end laser scanners with high performance INS/GNSS units and optionally 360 panoramic high resolutions multi-camera rigs. The latter two offer also backpack versions of their platforms for vehicle restricted areas. Imajbox by imajing, originally designed for trains and now updated for cars is a lower cost vision-based alternative (imajing, 2020).

### 2.1.2 Indoor mapping

For interior spaces different approaches exist. There are platforms with similar technologies like the car-mounted systems that are built on trolleys (NAVVIS, 2020), helmets (REscan, 2020) or backpacks (LEICA, 2020) (VIAMETRIS, 2020). Handheld devices like the PARACOSM PX-80 (PARACOSM, 2020) are also available but their accuracy is not directly comparable to the above systems.

Matterport (MATTERPORT, 2020) has a dedicated solution for creating digital twins for the Real Estate market. Indoor spaces are scanned via a proprietary low-cost 360 camera with depth sensors, or lately via a mobile phone and all required processes as well as hosting of data is done on a web service they provide.

For the Architecture, Engineering, Construction (AEC) market Doxel (DOXEL, 2020) provides automated solutions for quality inspection and progress tracking. They use artificial intelligence and autonomous robots that capture images and perform laser scanning surveys on a daily basis.

## 2.2 Simultaneous Localization and Mapping

Estimating the 6 Degrees of Freedom (DoF) motion trajectory of a mobile mapping platform is key to obtain georeferenced data. Direct Georeferencing from the GNSS/INS sensors is not always accurate and can fail in GPS restricted areas. Lately many systems adopt workflows from the robotics literature, like Visual Odometry or Simultaneous Localization and Mapping. A taxonomy and review of standard methods for visual odometry can be found in the well-known articles of Scaramuzza and Fraundorfer (2011a, 2011b).

Simultaneous Localization and Mapping is currently under heavy research. Current state-of-the-art SLAM algorithms exploit a broad range of data, such as images, IMUs or laser scanners and achieve remarkable results (Zhang and Singh, 2015), especially in autonomous driving scenarios, whilst maintaining near real-time performance. Visual methods of SLAM can be divided into feature-based, where features are first extracted on images (Mur-Artal et al., 2015) and direct methods that exploit all image gradients on the available images (Newcombe et al., 2011). Forster et al., (2017) proposed a semi-direct approach that combines direct methods for tracking pixels and features correspondences to refine both camera poses and

structure by bundle adjustment. In a recent publication Kuo et al. (2020) propose a generic vision based SLAM solution, which is sensor-agnostic and adapts to arbitrary multi-camera configurations. Other approaches rely on 3D point cloud to image matching using specialized descriptors (Pujol-Miro et al., 2017), as well as on constraining a SLAM algorithm given a street map background (Vysotska and Stachniss, 2017).

## 2.3 Structure from Motion

Structure from Motion (SfM), for the last two decades, is widely considered as the dominant image-based technique for automatic image alignment and 3D model generation and can be employed in workflows of processing data from mobile mapping platforms. SfM is a well-studied topic in the research community with a lot of nearly production-ready implementations (Schonberger and Frahm, 2016) and extensions, such as integration of video from aerial platforms (Leotta et al., 2016). However, it's still an open research field and new approaches have emerged, for example in robust image matching, mainly due to recent developments in deep learning (Yi et al., 2016).

## 3. SENSORS - COMPONENTS

Most mobile mapping systems share similar sensors for recording simultaneously visual information, depth, 3D point clouds, as well as the position and the orientation of the system in the world. More specifically, the proposed space sensing platform can support multiple sensors. The current implementation (Figure 1, Figure 2) consists of: i) four embedded 13MP machine vision cameras by econ-systems which can record still images or synchronized 4K video sequences, ii) a Velodyne® PUC VLP-16 LiDAR sensor which captures 3D point clouds, iii) an Xsens MTI-G-700 GPS/IMU unit that record absolute 3D positions and rotations and iv) an Intel RealSense T265 Tracking camera for relative positioning in GPS restricted areas (such as indoor scenes). Currently no lighting device is integrated in the platform.

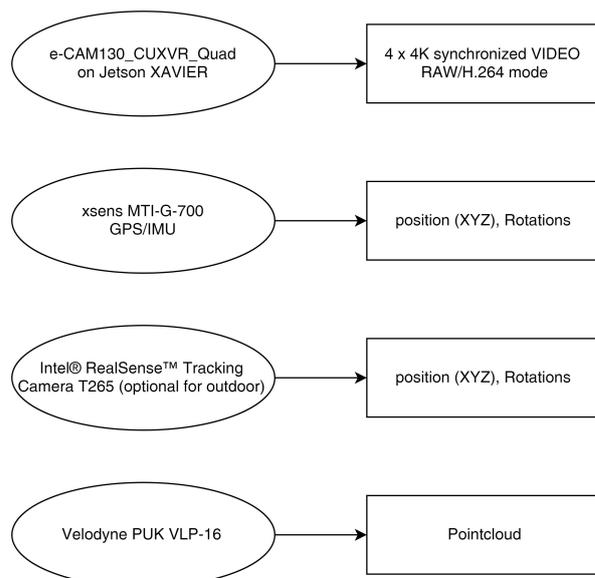


Figure 1. Sensors in the current implementation of the platform

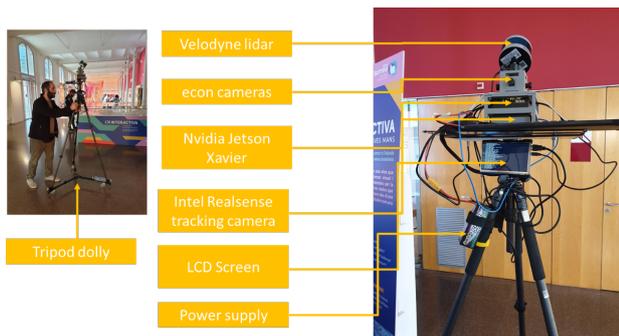


Figure 2. Prototype of the mobile mapping platform

### 3.1 Multi camera rig

For image and video capturing the platform hosts a multi-camera rig of four 13MP embedded cameras. The rig is based on a multi-camera board (e-CAM130\_CUXVR) by an OEM camera manufacturer (e-con systems), that targets the NVIDIA® Jetson AGX Xavier™ development kit. Cameras are connected on the board via high-speed 4-lane MIPI CSI-2 interface. These four camera modules can capture and stream synchronized 4K resolution (3840x2160) video sequences at 30fps in RAW or H.264 compressed format. Synchronization is achieved via an on-board Pulse Width Modulation (PWM) generator circuit that provides the necessary trigger signal for synchronous capturing. Each camera holds a wide-angle lens with 95° x 73° field-of-view (FOV). The four cameras are positioned at 90° configuration facing sideways offering a full 360° horizontal FOV. No upside camera is used in the current setup.

### 3.2 Lidar

The platform uses a Velodyne® PUC VLP-16 LiDAR sensor for pointcloud recording. The specific sensor is selected for its relatively low cost and high performance balance. It has a range of 100m, a positional accuracy of ~3cm, and 360° horizontal and 30° vertical fields of view (at 16 discrete channels). It can capture up to 600.000 points/second depending on the selected horizontal rotation velocity. It can be directly connected to a GPS/IMU device and supports data synchronization with precise GPS-supplied time via Pulse-Per-Second (PPS), in conjunction with a once-per-second NMEA GPRMC or GPWGA sentence. The lidar sensor is mounted on a ball camera tripod head on top of the platform to avoid occlusions from the other sensors and it is placed with an inclination of ~35° to capture floors and ceilings.

### 3.3 GPS/IMU

For direct Georeferencing in outdoors spaces the platform uses the Xsens® Mti-G-700 GPS/IMU Unit. It is the 4<sup>th</sup> generation motion tracker by Xsens and has built-in vibration-rejecting gyroscopes and accelerometer, a multi-GNSS receiver (GPS, GLONASS, BeiDou and Galileo) and a barometer. It measures attitude angles and accelerations and Xsens applies a Kalman Filter based sensor fusion algorithm to provide 3D position and orientation information.

### 3.4 Tracking camera

For GPS restricted areas like indoors environments the platform uses a new sensor by Intel®, the RealSense™ Tracking Camera T265. It is an embedded computer vision solution that combines two fisheye lens sensors with a combined close to hemispherical

~160° FOV, an Inertial Measurement Unit (IMU) and an Intel Movidius Myriad 2 Visual Processing Unit (VPU) that runs a proprietary Visual SLAM algorithm directly on the device. The T265 is connected and powered via USB and outputs 6DoF data at a sample rate of 200Hz.

### 3.5 Embedded PC & Laptop PC (optional)

To host the multi-camera rig, the platform utilizes an NVIDIA® Jetson AGX Xavier™ development kit that is widely used for the development of end-to-end AI robotics applications. This kit bundles a carrier board, an integrated thermal solution together with the embedded system-on-module (SoM) Jetson AGX Xavier. It combines an 8-Core ARM v8.2 64-Bit CPU, a 512-Core Volta GPU with Tensor Cores and 32 GB 256-Bit LPDDR4x Memory. It is configured to run Ubuntu 18.04.

An NVMe disk is added for storage and an LCD touch screen for control and visualization. Since this embedded system is powerful enough, our initial intention was to build the entire platform on it. This was partially achieved, except of support for the Xsens GPS/IMU unit, since no drivers were implemented for the ARM architecture. Thus, a laptop PC configured with Ubuntu 18.04 is an additional component used to include the GPS/IMU sensor (see distributed architecture implementation in Section 4.1). In an upcoming version of the platform we plan to replace the specific sensor with one compatible with the embedded PC.

### 3.6 Power supply

To power all the sensors and the embedded PC a 4 cell LiPo Battery of 5500mAh and 14.8V voltage was used. This provides enough power to run the platform for ~30min. When the platform is mounted on a car a typical 12v-220v inverter can be used instead.

### 3.7 Mounting

To combine physically all available sensors a prototype base was designed in 3D and then 3D printed (Figure 3). This design also provided a good approximation of all sensors' relative orientations (boresight alignment parameters). The base includes a typical camera mount that can be connected on a camera tripod on a dolly (Figure 2) and pushed around to perform data collections of interior environments or relatively small outdoors areas (like squares, individual buildings etc). Alternatively, it can be mounted on a car roof, via a DSLR suction cup camera mount. To further optimize the capturing process the use of a gimbal to reduce sensors shake as well as a backpack form factor version are considered for future implementations.

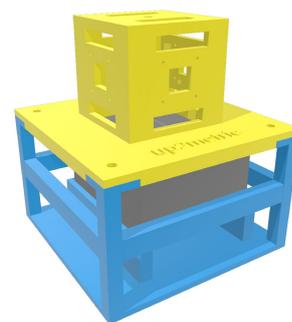


Figure 3. 3D Design of the base supporting the sensors.

#### 4. SENSORS INTEGRATION & DATA PROCESSING

Our platform aims to provide precise and fast 3D recording of outdoors and indoors spaces and consists of two main modules, i) the space sensing module that is responsible for data collection and ii) the 3D reconstruction module for processing all available data. The platform runs also in two separate modes, one for indoors and one for outdoors, each with some different hardware components, such as the tracking camera for indoors and the GPS/IMU sensor for outdoors.

##### 4.1 Space Sensing Module

For the integration of all sensors into a single capturing system the Robotic Operation System (ROS) (Quigley et al., 2009) was adopted. ROS is a middleware that is widely used by robotics teams both in Academia and the development of commercial products. Several ROS based opensource projects that implement sensors integration are available. Lately ROS was also proven to be a suitable platform for building mobile mapping systems to capture 3D interior (Blaser et al., 2018) and underground environments ((Blaser et al., 2019). ROS was selected since it is open source, it supports multiple programming languages (C++ and Python), it allows for low-level device control and is modular by design, making it relatively easy to add or remove devices.

ROS implements a message-passing communication architecture. A node is created for each sensor, which publishes sensor data as messages on specific topics. Topics may contain raw or processed values and each data entry inside a topic is assigned with a timestamp. Real-time processes are usually implemented as nodes that subscribe to specific topics and then publish their estimations in new topics. For offline processing, all messages are recorded on a single “bag” file. This is implemented by a “rosbag” node that subscribes to all messages from available sensors and stores them on the disk drive in a “bag” file. Then “bag” files can be reproduced for developing and testing algorithms.

ROS offers tools to monitor all recorded topics (“rqt-topic”) (Figure 4) as well as tools to visualize 2D and 3D sensor data (rviz) (Figure 5). Since all data are published as topics with timestamps, these timestamps are recorded in the “bag” file. Synchronization during offline data access or processing is usually handled by taking the data of each sensor that corresponds to the nearest timestamp or by interpolation.

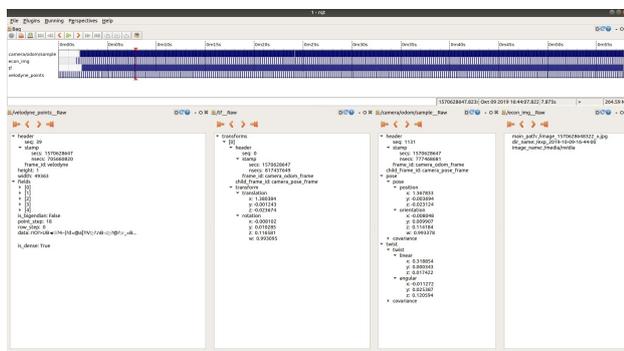


Figure 4. Data from all sensors in ROS can be monitored via “rqt-topic” tool. Data entries are usually accessed through a timeline feature.

More specifically, the proposed space sensing platform was implemented in ROS Melodic Morenia on Ubuntu 18.04. A

node was created for each sensor. Nodes communicate with the sensors and publish their data on a suitable designed topic.

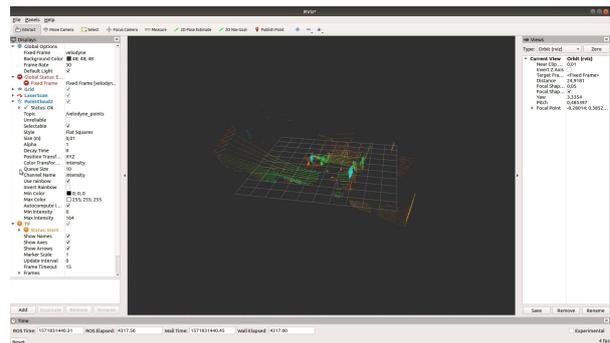


Figure 5. Visualization of the pointcloud topic from the Velodyne VLP-16 Lidar node in rviz ROS tool.

For the Velodyne® PUC VLP-16 LiDAR node, the official ROS “velodyne\_driver”<sup>1</sup> and “velodyne\_pointcloud”<sup>2</sup> packages were used. The first provides basic device handling for Velodyne lidars and publishes the raw data packets that are transmitted from the sensor through an ethernet connection. The second provides point cloud conversions. The Velodyne node publishes a “velodyne\_points (sensor\_msgs/PointCloud2)” topic which contains accumulated Velodyne points transformed in a selected frame of reference.

An official ROS package “xsens\_mti\_driver”<sup>3</sup> was also used for the Xsens® Mti-G-700 GPS/IMU Unit Node. The node publishes a “tf (geometry\_msgs/TransformStamped)” topic that contains 6 DoF orientation parameter (X, Y, Z translations and quaternion rotations) transformed in a selected frame of reference. A similar topic is published from the RealSense™ Tracking Camera T265 node that uses the ROS “realsense2\_camera”<sup>4</sup> package.

A new package was developed by our team for the multi-camera rig since no ROS compatible implementation was available. It is designed to work on the NVIDIA® Jetson AGX Xavier™, with custom made nodes and topics. The package consists of two subprograms, the “capturer” (C) and the “publisher” (C++). The first handles the cameras and captures images via v4l2 and gStreamer libraries, while the second is responsible to publish image data and metadata as a ROS topic. Initially we published image frames in a ROS topic but this approach lead to low FPS performance. In the current implementation the “capturer” app records 4 synchronised 4K videos at 30 FPS as .mkv files with H264 encoding format at a storage path defined by the “publisher” app. The latter publishes the start/end timestamps of the video sequence, as well as the timestamps and the frame\_ids of every synchronized frame that is added to the buffer of the gStreamer. Video files require ~60MB/camera/minute.

ROS natively supports a distributed architecture where sensors can run across multiple machines, which communicate through a local Network via a talker/listener logic. All nodes are configured to use a single ROS Master app (“roscore”), the address of which is defined by an environmental variable (“ROS\_MASTER\_URI”). Although the initial plan was to build the space sensing module on a single machine (NVIDIA® Jetson AGX Xavier™) where all sensors would be connected,

<sup>1</sup> [http://wiki.ros.org/velodyne\\_driver](http://wiki.ros.org/velodyne_driver)  
<sup>2</sup> [http://wiki.ros.org/velodyne\\_pointcloud](http://wiki.ros.org/velodyne_pointcloud)  
<sup>3</sup> [http://wiki.ros.org/xsens\\_mti\\_driver](http://wiki.ros.org/xsens_mti_driver)  
<sup>4</sup> [http://wiki.ros.org/realsense2\\_camera](http://wiki.ros.org/realsense2_camera)

due to incompatibility of the GPS/IMU sensor with ARM processors, the system was actually built following two alternative architectures (Figure 6). A single machine mode, when GPS/IMU is not used (for example in indoors environments) and a distributed one that supports the GPS/IMU device. In the latter configuration all sensors are connected on a laptop pc, except for the camera-rig, which by design requires to run on the Nvidia Jetson Xavier embedded computer.

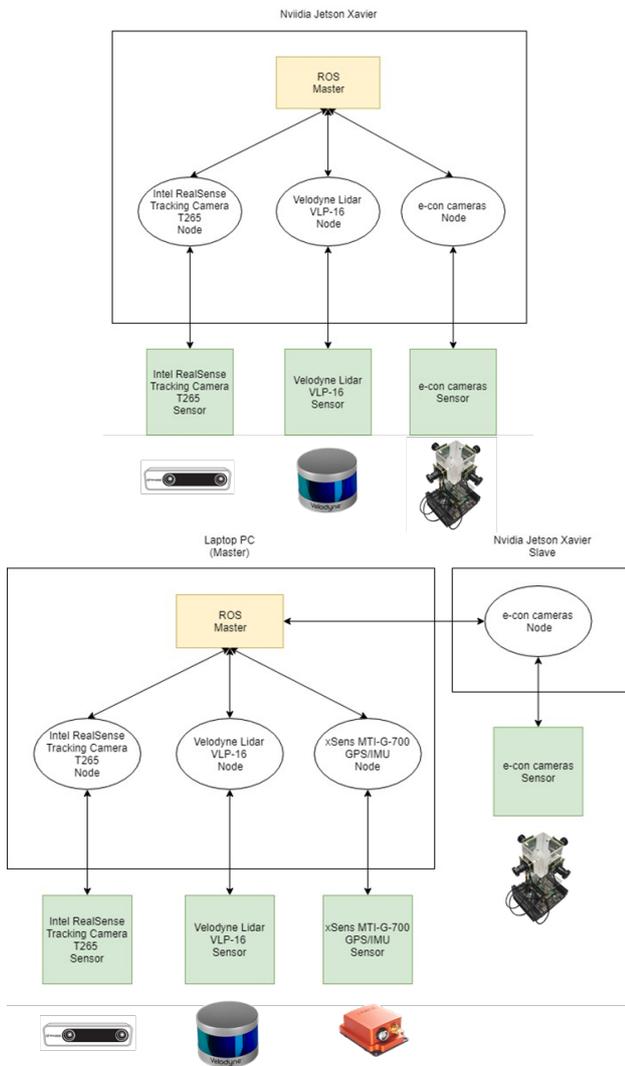


Figure 6. Robotic Operation System (ROS) integration of mobile mapping platform. Single Machine architecture (up) Distributed architecture (down).

The space sensing module is executed by a script that launches all processes (Figure 7). A basic GUI for touch screens (Figure 8) was also developed. It has tools to set the capturing parameters and start/stop the capturing session. Tools to inspect sensors connectivity and to assist the refocusing of each camera are also included.

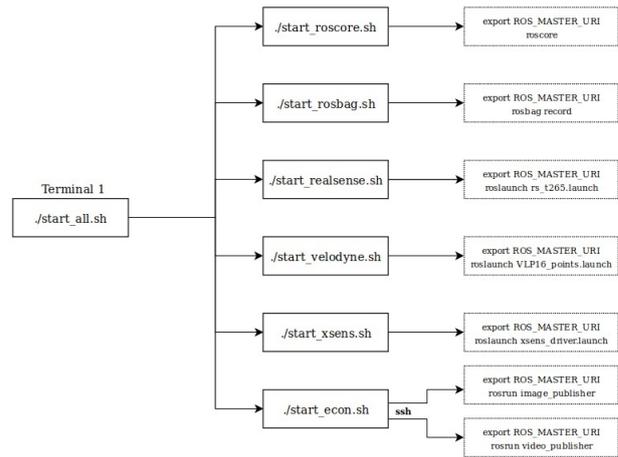


Figure 7. ROS subprograms of the space sensing module

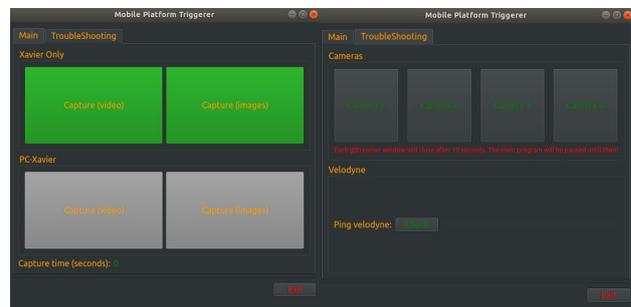


Figure 8. Basic user interface to set capturing parameters, start/end data capture and test/troubleshoot sensors connectivity

#### 4.2 3D Reconstruction Module

Every capturing mission with the space sensing module collects multiple types of data, from the connected sensors, which are stored into a “bag” file. All data are organised based on their timestamps. For any given time point or period it is possible to retrieve the corresponding data (ie image frames, pointclouds and 6DoF motion trajectories) and apply 3D reconstruction workflows. Since the work presented here is part of an ongoing research, several alternative approaches are into consideration before concluding to an optimal workflow.

More specifically, for the 3D reconstruction module of our platform we relied on existing software libraries, such as *ORB-SLAM2* (Mur-Artal et al., 2015) and *Google Cartographer* (Hess et al., 2016) for vision-based and lidar-based SLAM respectively, *AliceVision* & *Meshroom* (Jancosek and Pajdla, 2011 and Moulon et al., 2012) for Structure-from-Motion and *Open3D* (Zhou et al., 2018) for pointcloud processing.

Direct georeferencing from the specific GPS/IMU sensor or the tracking camera is not preferable due to their limited accuracy. However initial tests have shown that the provided trajectories from these sensors can be used to assist vision-based or lidar based SLAM algorithms. The latter provide more accurate estimations of the platform’s motion trajectory and initialization of orientation parameters for the individual pointclouds and the image frames. Global maps in the form of registered pointclouds are also provided but are most of the times sparse, incomplete, and noisy. In most cases though, the 3D models can be further improved by means of Structure-from-Motion Solutions.

Since the four cameras of the platform capture images at high FPS rates, each data collection mission consists of several millions of image frames. Using the SLAM provided image orientations, key poses of the multi-camera rig are selected so that they capture the area of interest with sufficient overlap and leave no gaps. Only these key frames are used in a Structure-from-Motion workflow through *Meshroom* open-source software framework. *Meshroom* implements a self-calibration bundle adjustment solution that supports Camera Rig Calibration. This allows for optimal estimation of the four cameras interior orientation parameters along with their relative orientation. This also leads to more accurate and consistent 3D reconstruction results. Finally, dense 3D point clouds are generated via Multi View Stereo 3D reconstruction algorithms.

In GPS deprived areas, where absolute orientation is a requisite, relative path and reconstruction estimations can be updated by means of Ground Control Points (GCPs) measured through standard Surveying techniques.

It must be mentioned that since all sensors were placed on a custom designed 3D printed case with known dimensions, good approximations of all sensors relative orientations were a-priori available. The effect of small misalignments was handled by the SLAM and bundle-adjustment solutions. An approach that we plan to further investigate is to update those relative orientation parameters by matching in 3D space the individual motion trajectories provided from the different sensors.

## 5. EXPERIMENTS

To demonstrate the effectiveness of a first prototype of the platform two data collection experiments are presented.

### 5.1 Indoors Environment

During the development of both the space sensing and the 3D reconstruction modules of the platform several experiments were conducted inside our office space. The presented survey corresponds to a single loop of the platform around an open desk area of  $\sim 90\text{m}^2$ . Figure 9, shows four image frames from the multi-camera rig.

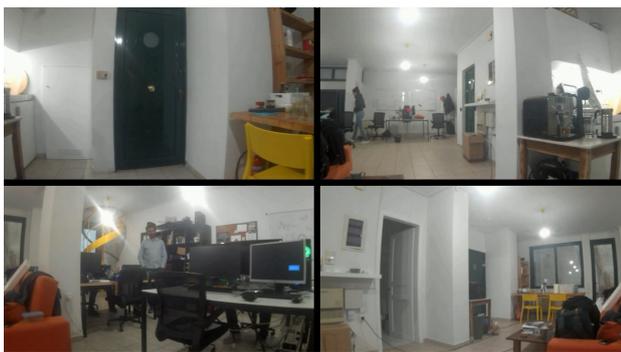


Figure 9. Example frames from an indoor environment mapping session of an office space, in Athens, Greece

In this experiment the 6 DoF motion trajectory from the RealSense™ T265 Tracking Camera was used to initialize a Structure-from-Motion Solution, with Camera Rig Calibration. A subset of the multi-camera rig image frames was used (Figure 10) (images at a given distance interval). Figure 11 shows an accumulated pointcloud from individual scans of the Velodyne® PUC VLP-16 LiDAR. The 3D translations and rotations of each individual scan were interpolated from the synchronized trajectory of the tracking camera.

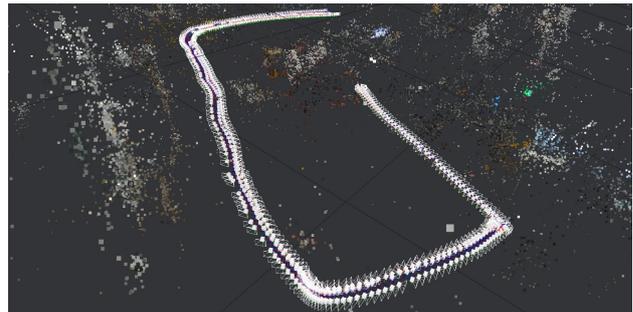


Figure 10. Estimation of mobile mapping platform trajectory by means of SfM. The solution was initialized by the RealSense™ T265 Tracking Camera.

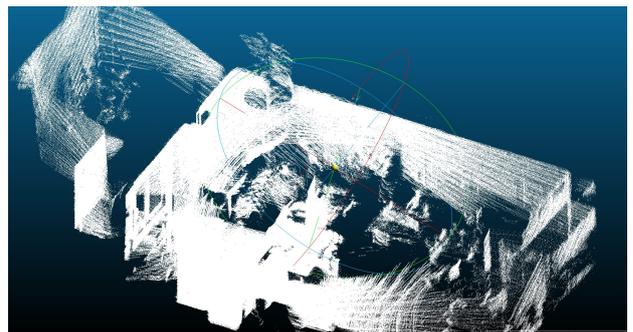


Figure 11. Registered pointcloud from individual Velodyne scans

### 5.2 Outdoors Environment

A second, more complete data capture mission was carried out in the area around the cultural centre of Tecla Sala which is situated in the City of L' Hospitalet, in Barcelona, Spain. This is the first Pilot Use Case of the "Mindspaces" H2020 Research Program. The tripod dolly was moved slowly around the cultural centre building to ensure that there is sufficient overlap between scanlines and image frames. The whole survey with the mobile mapping platform lasted  $\sim 15$  minutes.



Figure 12. Tecla Sala cultural centre area overview (image from Google Earth)

In this experiment a vision-based SLAM solution was used to estimate the motion and rotation trajectory of the platform and then an automatically selected subset of the collected image dataset was fed to the Structure-from-Motion workflow (Figure 14). A dense point cloud was also computed via Multi View Stereo Dense Reconstruction (Figure 15).



Figure 13. Four synchronized frames from an outdoor environment mapping in Tecla Sala in City of L' Hospitalet



Figure 14. Estimation of mobile mapping platform trajectory by means of SLAM and SfM



Figure 15. Final dense pointcloud from the Tecla Sala 3D survey (SfM).

## 6. CONCLUDING REMARKS

In this contribution we presented a first implementation of a modular mobile mapping platform that is based on commercial hardware components and open source software libraries. The integration of all sensors was carried out with the Robotics Operation System which allows for easy additions, changes and updates of the platform's components.

Several improvements are under consideration. A first one is the replacement of the GPS/IMU sensor with one compatible with ARM CPUs. This will allow the platform to run exclusively on the NVIDIA® Jetson AGX Xavier™ development kit and thus minimize its dependency on hardware, its size and its overall portability. The mounting of the platform on a camera gimbal is also considered to facilitate data collection sessions and obtain more stabilized data. The 3D reconstruction module requires further development and all workflows need to be thoroughly tested and evaluated with respect to their effectiveness, accuracy, and performance on well organised experiments.

A final more general remark has to do with a well-known restriction of mobile mapping systems, which is the inability to capture spatial information that is not directly visible from street-level (i.e. building roofs, backyards etc). Occlusions due

to obstacles such as buildings, parked cars or trees lead also to unavoidable gaps. To get complete digital copies of complex spaces mobile mapping missions need to be combined with either existing geospatial data from open databases, either with aerial missions from drones. This was initially planned for the Tecla Sala Pilot Use Case but was not realised because of a general ban of drone flight missions in the specific area. However, a second Pilot Use Case is currently under preparation in a more suitable area where licence to perform both mobile and aerial mapping missions is granted to the consortium. This will allow us to present soon the potential of combining mobile mapping with UAV photogrammetry.

## ACKNOWLEDGEMENTS

The presented work is partially funded by the European Commission as part of the H2020 Programme, under the contract number 825079-RIA.

## REFERENCES

Alvanitopoulos, P., Diplaris, S., de Gelder, B., Shvets, A., Benayoun, M., Koulali, P., Moghnieh, A., Shekhawat, Y., Stentoumis, C., Hosmer, T., Anadol, R., Borreguero, M., Martin, A., Sciama, P., Avgerinakis, K., Petrantonakis, P., Briassouli, A., Mille, S., Tellois, A., Fraguada, L., Sprengel, H., Kalisperakis, I., Cabanas, N., Nikolopoulos, S., Skouras, S., Vogler, V., Zavraka, D., Piesk, J., Grammatikopoulos, L., Wanner, L., Klein, T., Vrochidis, S., Kompatsiaris, I., 2019. MindSpaces: Art-driven Adaptive Outdoors and Indoors Design, in: Digital Presentation and Preservation of Cultural and Scientific Heritage International Conference, DiPP 2019.

Anguelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., Ogale, A., Vincent, L., Weaver, J., 2010. Google street view: Capturing the world at street level. *Computer (Long Beach, Calif.)* 43, 32–38.

Blaser, S., Cavegn, S., Nebiker, S., 2018. Development of a portable high performance mobile mapping system using the robot operating system. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* 4, 13–20.

Blaser, S., Nebiker, S., Wisler, D., 2019. Portable image-based high performance mobile mapping system in underground environments - system configuration and performance evaluation. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* 4, 255–262.

DOXEL, 2020. DOXEL Website <https://www.doxel.ai/>. (accessed 01/05/2020).

Forster, C., Zhang, Z., Gassner, M., Werlberger, M., Scaramuzza, D., 2017. SVO: Semidirect Visual Odometry for Monocular and Multicamera Systems. *IEEE Trans. Robot.* 33, 249–265.

Hess, W., Kohler, D., Rapp, H., Andor, D., 2016. Real-time loop closure in 2D LIDAR SLAM. *Proc. - IEEE Int. Conf. Robot. Autom.* 2016-June, 1271–1278.

imajing, 2020. imajing imajbox Website <https://imajing.eu/mobile-mapping-technologies/sensors/>. (accessed 01/05/2020).

- Jancosek, M., Pajdla, T., 2011. Multi-view reconstruction preserving weakly-supported surfaces. Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. 3121–3128.
- Kuo, J., Muglikar, M., Zhang, Z., Scaramuzza, D., 2020. Redesigning SLAM for Arbitrary Multi-Camera Systems.
- LEICA, 2020. Leica Pegasus Website <https://leica-geosystems.com/products/mobile-sensor-platforms/capture-platforms>. (accessed 01/05/2020).
- Leotta, M.J., Smith, E., Dawkins, M., Tunison, P., 2016. Open source structure-from-motion for aerial video. 2016 IEEE Winter Conf. Appl. Comput. Vision, WACV 2016.
- MATTERPORT, 2020. MATTERPORT Website <https://matterport.com/>. (accessed 01/05/2020).
- Moulon, P., Monasse, P., Marlet, R., 2012. Adaptive Structure from Motion with a contrario model estimation 1 Introduction 2 Structure from Motion the classical pipeline. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 7727 LNCS, 1–14.
- Mur-Artal, R., Montiel, J.M.M., Tardos, J.D., 2015. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. IEEE Trans. Robot. 31, 1147–1163.
- NAVVIS, 2020. NAVVIS Website <https://www.navvis.com/>. (accessed 01/05/2020).
- Newcombe, R.A., Lovegrove, S.J., Davison, A.J., 2011. DTAM: Dense tracking and mapping in real-time, in: 2011 International Conference on Computer Vision. pp. 2320–2327.
- Nüchter, A., Bleier, M., Schauer, J., Janotta, P., 2017. Improving google's cartographer 3D mapping by continuous-time slam. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 42, 543–549.
- PARACOSM, 2020. PARACOSM Website <https://paracosm.io/px-80-overview>. (accessed 01/05/2020).
- Puente, I., González-Jorge, H., Martínez-Sánchez, J., Arias, P., 2013. Review of mobile mapping and surveying technologies. Meas. J. Int. Meas. Confed. 46, 2127–2145.
- Pujol-Miro, A., Ruiz-Hidalgo, J., Casas, J.R., 2017. Registration of images to unorganized 3D point clouds using contour cues. 25th Eur. Signal Process. Conf. EUSIPCO 2017 2017-Janua, 81–85.
- Quigley, M., Conley, K., Gerkey, B.P., Faust, J., Foote, T., Leibs, J., Wheeler, R., Ng, A.Y., 2009. ROS: an open-source Robot Operating System, in: ICRA Workshop on Open Source Software.
- REScan, 2020. RESCAN 360 Website <https://rescan360.com/>. (accessed 01/05/2020).
- RIEGL, 2020. RIEGL Laser Measurement Systems Website <http://www.riegl.com/hc/products/mobile-scanning/produktdetail/product/scanner/56/>. (accessed 01/05/2020).
- Scaramuzza, D., Fraundorfer, F., 2011a. Visual odometry Part I: The First 30 Years and Fundamentals. IEEE Robot. Autom. Mag. 18, 80–92.
- Scaramuzza, D., Fraundorfer, F., 2011b. Visual Odometry [Tutorial]. IEEE Robot. Autom. Mag. 18, 80–92. <https://doi.org/10.1109/mra.2011.943233>
- Schonberger, J.L., Frahm, J.-M., 2016. Structure-from-Motion Revisited, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 4104–4113.
- Shi, W., Alawieh, M.B., Li, X., Yu, H., 2017. Algorithm and hardware implementation for visual perception system in autonomous vehicle: A survey. Integr. VLSI J. 59, 148–156.
- VEXCEL, 2020. UltraCam Mustang Website <https://www.vexcel-imaging.com/ultracam-mustang/>. (accessed 01/05/2020).
- VIAMETRIS, 2020. VIAMETRIS Website <https://www.viametris.com/>. (accessed 01/05/2020).
- Vysotska, O., Stachniss, C., 2017. Improving SLAM by Exploiting Building Information from Publicly Available Maps and Localization Priors. Photogramm. Fernerkundung, Geoinf. 85, 53–65.
- Wang, Y., Chen, Q., Zhu, Q., Liu, L., Li, C., Zheng, D., 2019. A survey of mobile laser scanning applications and key techniques over urban areas. Remote Sens. 11, 1–20.
- Yi, K.M., Trulls, E., Lepetit, V., Fua, P., 2016. LIFT: Learned Invariant Feature Transform Motivation & Contributions. Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics).
- Zhang, J., Singh, S., 2015. Visual-lidar odometry and mapping: Low-drift, robust, and fast. Proc. - IEEE Int. Conf. Robot. Autom. 2015-June, 2174–2181.
- Zhou, Q.-Y., Park, J., Koltun, V., 2018. Open3D: A Modern Library for 3D Data Processing. arXiv Prepr. arXiv1801.09847.