MindSpaces

Art-driven adaptive outdoors and indoors design

H2020- 825079

D3.1

Initial data collection from sensors

Dissemination level:	PU: Public
Contractual date of delivery:	Month 17, 31 May 2020
Actual date of delivery:	Month 18, 9 June 2020
Workpackage:	WP3 Sensor data collection
Task(s):	T3.1: Emotional and cognitive sensing, T3.2: Visual sensing for behavioral analysis, T3.3: Space sensing for 3D reconstruction, T3.4: Interiors sensing for 3D reconstruction, T3.5: Social media and web data crawling
Type:	Report
Approval Status:	Approved
Version:	1.0
Number of pages:	103
Filename:	D3.1_Initial data collection from sensors_2020-06-09_v1.0_final.docx
Abstract	

D3.1 describes the basic techniques that have been taken into consideration in order to collect the appropriate data for the support of the Mindspaces use cases. More specifically, this deliverable provides the details of the methodologies that were followed in order to support: i) Emotional and cognitive sensing, ii) Visual data collection for Behavioural Analysis, iii) Space sensing for 3D reconstruction, iv) Interiors sensing for 3D reconstruction and v) Content extraction from social media and web.

Co-funded by the European Union

History

Version	Date	Reason	Revised by
V0.1	28-5-2020	Creation of ToC	Petros Alvanitopoulos, Sotiris Diplaris
V0.2	26-5-2020	Chapter 4	Panagiotis Petrantonakis, Anastasia Karatzia
V0.3	27-5-2020	Chapter 6 and 7	Ilias Kalisperakis
V0.5	06-03-2020	Complete version of the deliverable for internal review	Petros Alvanitopoulos, Sotiris Diplaris
V0.6	8-7-2020	Internal review	Tyson Hosmer
V1.0	8-7-2020	Last amendments	Petros Alvanitopoulos

Author list

Organization	Name	Contact Information
CERTH	Petros Alvanitopoulos	palvanitopoulos@iti.gr
CERTH	Panagiotis Petrantonakis	ppetrant@iti.gr
CERTH	Anastasia Karatzia	akaratzia@iti.gr
up2metric	Ilias Kalisperakis	ilias@up2metric.com
up2metric	Christos Stentoumis	christos@up2metric.com
up2metric	Andreas El Saer	andreas.saer@up2metric.com
up2metric	Stavroula Bourou	stavroula.bourou@up2metric.com
up2metric	Theodor Mandilaras	theodor.mandilaras@up2metric.com
up2metric	Panagiota Stamatopoulou	panagiota.stamatopoulou@up2metric.com
up2metric	Lazaros Grammatikopoulos	lazaros@up2metric.com

Executive Summary

The goal of this deliverable is to report the techniques, the actions that were made and the collected data from sensors towards the data collection activity included in WP3 of the MindSpaces project. The methodology that we have followed is described and the related steps regarding the data preparation, collection and business/data understanding are also connected to the PUCs and requirements.

This deliverable lists the workflow of the creation of the initial dataset by using different kind of sensors. Physiological signals, visual data for 3D reconstruction and behavioural analysis and textual data extracted from web are collected in order to support the activities for analysis of data in WP4 and WP5. The equipment, the time plan for recording and the challenges that we faced for the generation of the initial dataset are included.

The present document describes the high-level framework of CRISP-DM and how it is applied to MindSpaces project. Moreover, it emphasizes the need for multiple iterations for collection of data to ensure successful completion of the MindSpaces project. This deliverable concludes with describing the future work that will be carried out towards the delivery of the final dataset.

Abbreviations and Acronyms

API	Application programming interface
DMP	Data Management Plan
DoA	Description of Actions
EEG	Electroencephalographic
GPS	Global Positioning System
GSR	Galvanic Skin Response (GSR)
HLUR	High Level User Requirement
LSL	Lab Streaming Layer
PUC	Pilot Use Case
SDK	Software Development Kit
TCP	Transmission Control Protocol)?
VR	Virtual Reality
WP	Work package

Table of Contents

1 INTRODUCTION

This deliverable outlines the data collection techniques for the first seventeen months of the MindSpaces project. More precisely, it gives an overview of the high-level methodology used in WP3 for the initial data collection, in order to fulfil the scope of the project.

This introductory section describes the context of collaboration between the partners for the collection and creation of datasets from the start of the project until the delivery date of this deliverable.

In Section 2 the Methodology is presented. More specifically, in our approach the data collection process is an essential step of a wider procedure broken down into a set of iterative activities. For the creation of the datasets the applied agile workflow is described and the timeline is also presented.

In Section 3 a brief presentation regarding the data collection process and the relation to the user requirements is included. The High-Level User requirements (HLURs) presented in D7.1 are connected to the data collection activities that were carried out during this period.

Section 4 is devoted to the Emotional and cognitive sensing. A detailed presentation of the equipment, experiments and the methodology are included for the collection of physiological data from subjects.

Section 5 describes the collection of visual data aiming to estimate the behavioral analysis for indoor and outdoor design. The approach for each scenario of PUCs, the related equipment and the anonymization of visual data are some of the aspects described in this section.

A description of the collection of data for outdoors environments data is given in Section 6. This section includes the State-of-the-Art equipment, the mobile mapping platform Development and the data collection for 3D reconstruction in PUC1.

Section 7 includes all the details for the Interiors sensing for 3D reconstruction. A comprehensive description about the collection of data is presented. The data collection process is described, aiming to create the 3D-models of interior spaces in order to support PUC2 and PUC3.

In all Sections, 4 through 7, information about the equipment used is provided, along with the applied techniques. Also, a short description of goals and requirements is given, followed by an overview of the data collected for each Section.

In Section 8 the activities for the collection of data from Social media and Web are included. This section refers to the collection of freely available textual and visual content from open web resources and social media. The development and the different types of crawling and scraping sources are also described.

The current deliverable concludes with outlining how each data collection activity connects to the other scientific and technical objectives of the MindSpaces project and what future work needs to be followed towards deliverable D3.2, D3.3, D3.4 and challenges ahead.

2 METHODOLOGY

In this section, the high-level process for the data collection and integration to the MindSpaces platform is primarily described. The cross-industry standard process for data mining (CRISP-DM) which is the leading de facto standard for DM applications, is the methodology followed in MindSpaces. This section provides an extensive overview of the steps included in this process and how it is applied to MindSpaces project. It outlines the types of actions required and why different iterations for collection of data are important and ensure the successful completion of the MindSpaces project.

2.1 High-level methodology

The aim of WP3 is to acquire different types of sensor data that will help architects and artists to produce new designs of spaces and deploy installations. This process covers most of the project's lifetime (M3-M32). Our work is not limited to perform data collection, but the aim is to create a plan and present the most relevant data in the most usable format for the target audiences taking into consideration the different conditions and requirements per use case.

The relevance of data to the user requirements is defined in Section 3: Relation to User Requirements. In MindSpaces project the related WPs for the analysis of data are WP4 and WP5. WP4 includes the analysis of acquired data in order to create 3D models of urban and indoors spaces and the extraction of aesthetics and style information from visual content. WP5 focuses on the application of sophisticated models for emotion extraction from EEG and physiological signals and human behavior analysis from visual signals. Thus, it is important from the data collection phase to collect relevant data in an appropriate form without noise that could be used in the analysis process, in order to fulfil the objectives and meet the expectations of the project. RGB video cameras and sensors for collecting physiological signals along with crawlers and scrappers are utilized to collect data from subjects, sensing environment under-studied and related topics from web and social media. Both WP4 and WP5 leverage data provided by WP3 to train machine learning models that could interpret this information and provide valuable insights for the design of spaces. In that sense the work in WP3 constitutes the first steps which are crucial in data mining process. The CRISP-DM application warrants a framework in which the related tasks for data collection in WP3 are organized properly to ensure the smooth implementation of the project and the fulfilment of the objectives.

Before the start of WP3, an initial desk research for a proper framework for data collection process was conducted, having in mind that such a framework would be the guide to fulfil our goals. Following a comparative analysis of framework [1], [2], the CRISP-DM process model was applied as the high-level process for the definition of the workflow in WP3.

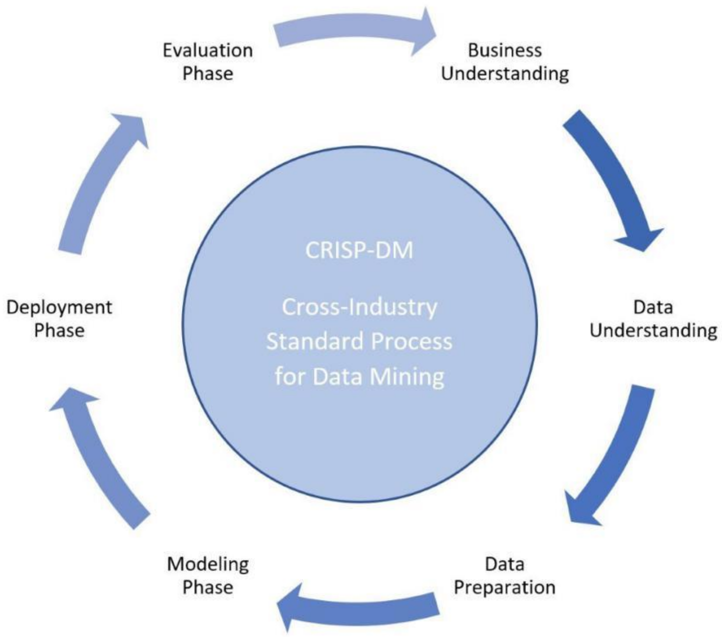

Figure 1. The CRISP-DM phases[1]

CRISP-DM provides an overview of the life cycle of a data mining project. This framework is an iterative process and includes the Business Understanding, Data Understanding, Data preparation, Modeling, Evaluation and Deployment phase. It is worth mentioning that ML algorithms depends on data. Therefore, in practice, it is essential to spend time and pay attention to the Data Understanding and Data Preparation phases for every DM project. A description of the two first steps of the CRISP-DM standard based on Wirth and Hipp [3] is given below:

"Business Understanding: This initial phase focuses on understanding the project objectives and requirements from a business perspective, and then converting this knowledge into a data mining problem definition, and a preliminary project plan designed to achieve the objectives"

"Data Understanding: The data understanding phase starts with an initial data collection and proceeds with activities in order to get familiar with the data, to identify data quality problems, to discover first insights into the data, or to detect interesting subsets to form hypotheses for hidden information."

According to this high-level overview, it is important to link all the CRISP-DM phases to the MindSpaces project. First the Business Understanding phase is taken place by taking into consideration the User Requirements received from User groups. The Data Understanding phase is followed which includes the definition of required data, the identification of challenges for data quality, the configurations for each use case and the technical requirements for supporting the data collection. Data preparation is also carried out by taking into account aspects such as the sampling of data, the format, the filtering methods

[1] Giang Nguyen et al., *Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey*, 2019

for data anonymization and noise removal and the transmission and integration of data into the MindSpaces platform. The rest of this deliverable presents a comprehensive description of the activities and workflow of the Crisp-DM process. The next section is devoted to the low-level workflow for the creation of the MindSpaces datasets.

2.2 Introduction to the iterative agile workflow process for the creation of the MindSpaces datasets

Within the MindSpaces project the development cycle is described in DoA with multiple development iterations and involves the cooperation between users and technical partners, in order to develop a platform that satisfies user expectations and project's goals. More precisely, each iteration consists of the steps below:

1. user requirements' and use cases' definition
2. design and development of the platform
3. deployment of the platform components
4. evaluation of each prototype taking into consideration use case scenarios created by artists and architects

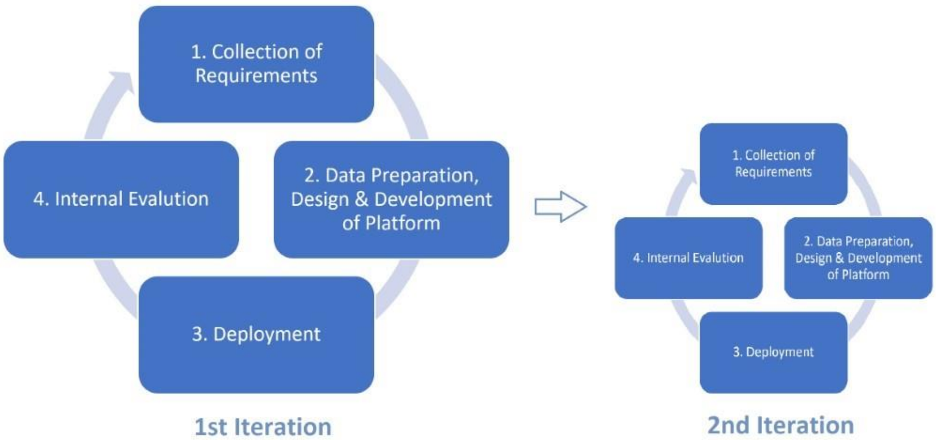

Figure 2. Description of the iterative process followed in MindSpaces

For the realisation of the data collection the two iterations are taken into consideration. The Gantt Chart below shows the high-level workflow that we followed:

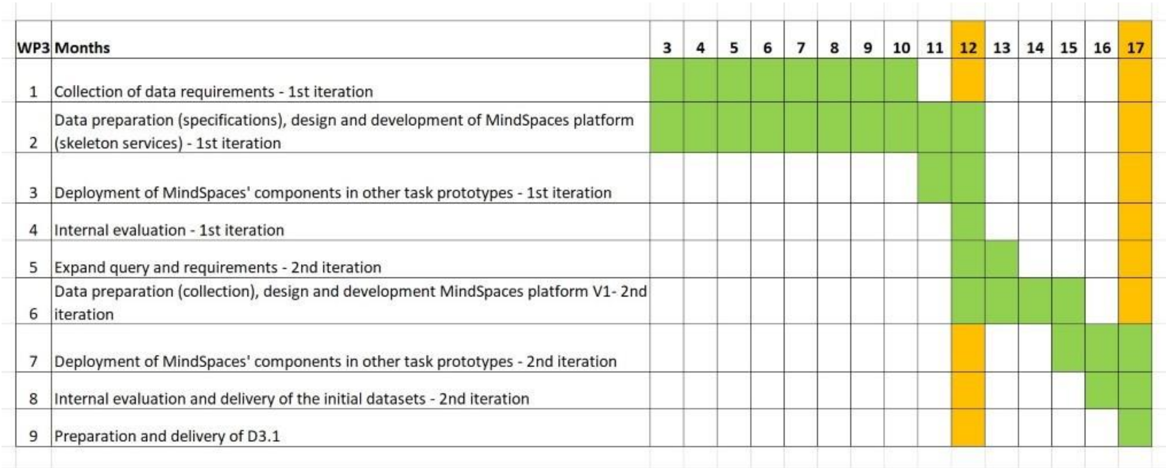

Figure 3. Gantt Chart of the actions carried out until M17. Milestone months are highlighted in orange.

MindSpaces project involves the collection of multidisciplinary data for the support of the interdisciplinary research of designing indoor and outdoor spaces. The related data collection tasks started first at M3 with the creation of crawling and scraping techniques applied to websites and social media. At M7 the remaining tasks related to visual data and physiological signals collections were started. The first phase of the data collection process ends at M17, along with this deliverable (D3.1), which includes the initial dataset that will be used for the support of PUCs and the evaluation of V1 of the MindSpaces platform. Below, we describe in more detail the iterations until M17.

The 1st iteration started at M3 and ended with the development of skeleton services in M12. From M3 to M10 we collected the data requirements for this iteration. From M3 to M12 a parallel process was running related to the specifications for data collection and the design and development of skeleton services of the MindSpaces platform. During this period several discussions were made in order to better understand and have a more concrete description of the required data which includes parameters such as 1) the areas of interest for recording, 2) the time plan for recording, 3) the format of data, 4) the management of data, 5) the licences for recording, 6) the types of subjects 7) the related scenarios per PUC, 8) the types of web sources, 9) the topics for social media 9) the selected equipment and 10) the duration of the experiments in case of physiological signals collection. By the end of the 1st iteration at M12 the initial data were collected which include textual information from websites and Twitter, visual data for behavioural analysis for PUC1 and PUC3, visual data for the 3D reconstruction of models in all PUCs and data for training models for emotional state estimation through physiological signals. The 2nd iteration started at M13. In the 2nd iteration the challenges during the development of the platform and the creation of more concrete scenarios per PUC allowed us to re-evaluate the data collection process, collect more related data and in some cases to update the existing methodology.

In general, having multiple iterations helps the consortium partners to gradually enrich the processes and this applies to the data collection actions. By starting from an initial dataset, we then continue with potential adjustments needed to the data collection methodology, before we reach to the creation of final datasets. Having just one iteration for this process would result in data collections that do not completely fulfil the project's needs, as there would be no chance for redefinition of the requirements.

Following this methodology with the multiple iterations, we can identify if the already collected data met the project's requirements and if not apply the proper adjustments to the data requirements in the next iteration.

3 RELATION TO USER REQUIREMENTS

A thorough analysis regarding the High Level User Requirements (HLURs) and their connection to the data collection activities is essential. The HLURs, in combination with Pilot Use Cases' (PUCs) expectations, play a significant role for 1) the definition of the data types needed, 2) the data quantity and quality, 3) the equipment required for the data collection, and also 4) planning the installation of sensors and the duration of recordings. In this section, we describe how MindSpaces' HLURs are related to the data collection process needed for each PUC.

The collection of the user requirements finished at M10. A short description of the user requirements list is provided in the D7.1. The following Table shows the connection of PUCs to the data collection related HLURs.

Table 1. List of HLUR related to Data Collection tasks in WP3

	HLUR Title	HLUR Description	PUC1	PUC2	PUC3
HLUR_1	"User interaction and control"	Architects can collect onsite, geolocate and aggregate biometric/behavioural data in 3D reconstructed environments.	X	X	X
HLUR_2	"Manipulation of spatial conditions"	Architects/Designers and artists can use spatial conditions and environmental attributes of spaces with the emotional state and behaviour of the users, to increase social interactions and communicate artistic concepts.	X	X	X
HLUR_3	"Data Analysis for understanding social needs and human values"	An artist/designer can use Data Analysis for understanding social needs and human values through social interaction with public/private spaces.	X	X	X
HLUR_4	"Adaptable spaces"	Citizens/office workers and seniors can have adaptable spaces indoor and/or outdoor depending on their needs.	X	X	X
HLUR_5	"Space use prediction"	An architect/designer can predict the potential uses for new spaces by analysing previous behavioural data.	X	X	X
HLUR_6	"Intelligent projects based on feedback"	An architect/designer can produce social intelligent projects based on feedback (emotional and rational).	X	X	
HLUR_11	"Novel textures based on the aesthetics of famous paintings"	Create novel and inspiring textures that are based on the aesthetics of famous paintings and other images of artwork that do not exist in current 3D modelling market.	X	X	X
HLUR_12	"Architects/designers aesthetics gallery"	Understand the aesthetics of design structures (i.e. interior objects, buildings, materials etc.) and provide it to architects and designers as a gallery.	X	X	X

Before the examination of the connection between the collection of data and the user requirements, a brief description of the related PUCs follows.

In PUC1 scenario the aim is the production of outdoors architecture and urban design proposals for an area of special cultural interest. The Tecla Sala Cultural Centre in the city of L'Hospitalet de Llobregat was selected. According to this scenario Artists and Architects will propose new installations and designs taking into consideration the citizens' feedback. The citizens' feedback involves several types of sensing data which includes 1) textual information from social media and websites, 2) Footage from the CCTVs cameras installed around the Tecla Sala area for Visual Behavior Analysis 3) Visual Data for the 3D reconstruction of space and the support of the experience of new designs in VR and 4) Physiological data for the estimation of emotional state evoked by the different versions of the re-designed space.

The goal of PUC2 is the design of better quality workplace environments. The general goals involve: 1) the understanding of how people behave and feel emotionally in relation to workplace designs 2) the examination of design parameters that could evoke a more positive emotional state 3) Designs that promote social interaction and collaboration. Different types of users are considered, such as office worker, office designer and office manager with different aspects of evaluation, which include individual experience, design evaluation and workplace performance. For the evaluation of PUC2 the McNeel's offices in Barcelona and a 3D model developed by Zaha Hadid (see Section 4.4.3) will be used for the 3d virtual workplace experience.

PUC3 intends to re-design and refurbish a senior's home in Paris, so that it is emotionally reactive, comforting, appealing and inspirational as well. Within PUC3, the senior's behaviour activity and emotional state will be analysed and connected to new designs of space and art installations. Architects' and artists' solutions will be examined in a VR environment and will be assessed by end-users. There is a relation between PUCs and HLURs, as shown in Table 2. In the following paragraphs our analysis goes one step further by connecting and describing the different data collection activities related to the HLURs and PUCs.

According to HLRU_1, for PUC1 video footage is required for visual behaviour analysis. The installed CCTV cameras at Tecla Sala are used for the collection of information related to people's behaviour around the cultural centre. For the biometric data, proper sensors will be used to get the required information from citizens while they experience the new designs of spaces in VR environment aiming to estimate the emotional state of subjects evoked by the changes of space. Concerning PUC2, visual data are also needed for capturing the employees' behaviour in the workplace under-study. For this reason, cameras were installed at McNeel offices. Moreover, while they experience the VR environment of the proposed workplace, their physiological signals will be collected. Similarly, in PUC3 indoor cameras were installed in the senior's house in order to record the types of activities from different rooms. As we can see, HLUR_1 defines the type of data that we need to collect for each PUC.

HLUR_2 involves the spatial conditions and environmental attributes of spaces. Taking this HLUR into consideration, we need to collect visual data and connect them with changes related to the environment, and examine aspects such as the social interaction and the interaction between subjects evoked by physical art installations. The video footage from multiple CCTV cameras coverage different areas of interest at Tecla Sala and the collected data will provide valuable information regarding the current behaviour of citizens and reveal the areas of maximum and lowest activity. Four indoor cameras are also available and installed at MCNEEL offices in order to receive video footage from different areas of interest. For PUC3 2 indoor cameras were installed in the living room and in the computer room. For all PUCs in order to measure the physiological signals in VR environment a 3D model was created based on the visual data collected from Up2metric.

In order to fulfil the HLUR_3, data from social media are also collected, along with visual data, for the understanding of social needs and human values through social interaction. Social media were taken into consideration and textual data from Twitter and websites were collected, along with data collection and analysis techniques.

HLUR_4 requires information about how people behave in adaptable spaces (what activities take place there, how many people are usually at a specific place/room, what social interactions exist, how their behaviour changes depending on the hour of the day etc.). This information can be extracted by analyzing the people's behaviour using visual data taken from cameras installed at the spaces of interest at different times of day.

For HLUR_5 the information from sensors about the environment under-study is also needed. The aim is the exploitation of this information by architects and designers for the prediction of the potential uses for their proposed new spaces. Related "heatmaps" can be created that will help develop simulation algorithms of human behaviour at a specific place. Architects and designers can create their proposed space taking into account the information given by the sensors and simulations about how people would use this space.

HLUR_6 indicates the use of data from social media and web. Architects and designers will use this information to produce social intelligent projects.

HLUR_11 and HLUR_12 are related to the aesthetics. They do not have a direct relation to the data collection activities of WP3. These HLURs describe the application of aesthetics in the design process (e.g. interior objects with new textures, gallery of materials, etc.). However in order to support this action, 3D scanners need to be used for the 3D-modelling of the existing interior objects, capturing the existing texture and geometry of objects.

The initial data collection for the MindSpaces project took place from M3 to M17. During these months, we had two iterations of the data collection process, according to the CRISP-DM model. For each iteration, all consortium partners were involved and contributed in order to ensure that the collected data are enough to support the objectives of each PUC. Their feedback was used to redefine the data collection requirements and properly adjust the process.

4 PHYSIOLOGICAL SIGNAL SENSING

4.1 Introduction

The Physiological Signal Sensing (PSS) module of MindSpaces is responsible for measuring and collecting users' physiological data while experiencing various MindSpaces' designs of spaces and art installations.

PSS consists of electroencephalographic (EEG) and Galvanic Skin Response (GSR) signals. Data collection will consist of two different modules as presented in Figure 4. These modules are used to collect:

1. EEG signals from the scalp,
2. Bio – measures, in particular, the GSR.

Each of these modules has its own SDK in order to acquire and process the related data. In order to acquire and synchronize the data, each module is compatible with Lab Streaming Layer (LSL) specifications, an open source library to collect and synchronize multiple digital signal sources.

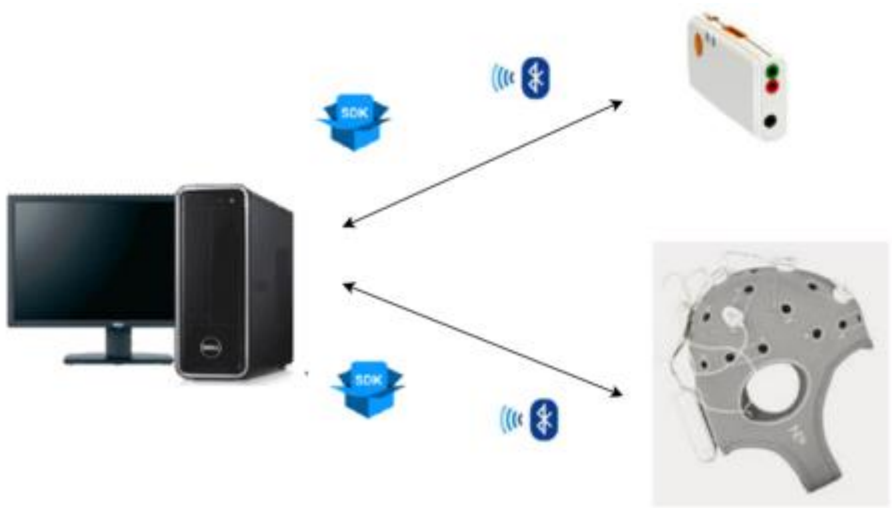

Figure 4. Physiological signal data collection modules.

All the above devices run together, controlled by a PC/laptop with the recommended requirements presented in Table 2.

Table 2. Recommended system requirements.

FEATURE	RECOMMENDED SPECIFICATIONS
OPERATING SYSTEM	Windows 7
CPU	i5-2410M 2.3 GHz
RAM	2 GB
USB PORT	3.0
GPU	Nvidia GeForce GT 520 MX

BLUETOOTH	yes

Given the aforementioned configurations, this section is focused on:

-Technical specifications for the selected EEG device,

-Technical specifications for the GSR device,

-SDKs for each module,

-Synchronization software,

-Preliminary experiments description.

4.2 Hardware equipment

Equipment selection is the first and most important step in the setup of PSS data collection. It is necessary to choose EEG and GSR recording devices that will serve the purpose of the project. In MindSpaces project, apart from indoor recordings in a controlled environment, i.e., a lab, that are necessary for the acquisition of PSS data collection, out-of-the-lab recordings are also necessary to take place for the evaluation of the MindSpaces Pilot Use Cases (PUCs). In addition, the various MindSpaces designs and installations will take place in a virtual environment. Hence, it is necessary to choose lightweight recording devices that the user will be able to wear in combination with the virtual reality headset.

4.2.1 EEG devices

There are many EEG recording devices released by different companies. Every one of them offers different capabilities in the recording of EEG signals. Below there is a brief description of all the EEG recording devices that were taken into consideration in order to select the most suitable one for MindSpaces PSS data collection.

Emotiv Epoc+ [4] (Figure 5) is a 14 channel mobile EEG device, designed for scalable and contextual human brain research and advanced brain-computer interface applications. It connects to the computer wirelessly through Bluetooth, or through USB. Its battery life can hold up to 12 hours if using USB receiver and up to six hours if using Bluetooth. Its dynamic resolution is 14 bits and the available bandwidth is 0.16 to 43 Hz. It consists of 14 EEG sensors representing channels AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8 and AF4 and two reference channels. Electrodes consist of Saline soaked felt pads. It is supported by multiple different platforms (Windows, Mac, iOS, Android) and provides its own SDK, which allows connecting the device to the computer, displaying and managing the EEG data streams and recordings. Finally, Emotiv EPOC+ offers some facial expression and performance metrics detection and can integrate data from external sources and share data with external systems, like LSL library.

Figure 5. Emotiv EPOC device and its components

Emotiv Insight [5] (Figure 6) is a five channel mobile EEG recording device designed for self-quantification, brain-computer interfaces, and research. Emotiv Insight boasts advanced electronics that are fully optimized to produce clean, robust signals. It connects to the computer wirelessly using Bluetooth or through USB. It consists of a rechargeable battery with a life of up to eight hours using the USB receiver and up to four hours if using Bluetooth. Emotiv Insight has a dynamic resolution of 14 bits and the available bandwidth is 0.5 to 45 Hz. Emotiv Insight consists of five EEG channels: AF3, AF4, T7, T8, Pz and two reference channels. Sensors are made of Hydrophilic semi-dry polymer. Like Emotiv Epoc+, Emotiv Insight is supported by the same platforms, has its own SDK and offers performance metrics detection. Furthermore, it can integrate data from external sources, such as LSL library, which allows real-time data stream.

Figure 6. Emotiv Insight device

Muse 2 [6] (Figure 7) is a multi-sensor meditation device that provides real-time feedback on the brain activity, heart rate, breathing, and body movements. Muse does not transmit data while plugged into USB, but wirelessly through Bluetooth. Its battery life can hold up to five hours of continuous use. Muse consists of seven calibrated EEG brain sensors, two on the forehead, two behind the ears plus three reference sensors that detect and measure the activity of the brain. In addition, it has added pulse oximetry breath and heart sensors that are located on the front, right-hand side of the forehead. Gyroscope and accelerometer body sensors are found behind the ears. Muse provides its own SDK (Muse Monitor) which allows manipulation of data. Furthermore, it can connect to external sources and share data with external systems.

Figure 7. Muse 2 device

Enobio [7] (Figure 8) is a mobile, comfortable and wireless device that can capture EEG signals with high precision. Enobio is one of the most precise systems in its class, because of its dynamic resolution (24 bits, 0.05 uV) and sampling rate. Device's available bandwidth is 0 to 125 Hz. It establishes a wireless connection to the computer via Bluetooth and allows the setup of eight, 20 or 32 electrodes in 39 different head positions. Enobio allows the set-up of the EEG channel recordings with dry electrodes (Figure 9), solid gel or gel-based electrodes (Figure 10). It also provides its own SDK (NIC2) in order to be able to connect to the Enobio device and manage it through the computer. It consists of a rechargeable battery with a life of 6.5 hours with wireless data transmission (range of 10 meters from the computer running the software NIC2). It also offers basic and advanced modes in order to design and monitor experiments. In addition, it can integrate data from external sources and share data with external systems, like LSL library.

Figure 8. Enobio device

Figure 9. Enobio's dry electrodes

Figure 10. Enobio's solid gel (left) and gel-based (right) electrodes

B-Alert X series [8] (Figure 11), is a wireless EEG system for mobile data acquisition and analysis. B-Alert X series system is equipped with 10 or 24 channel input electrodes. EEG signals are delivered wirelessly to the computer via a Bluetooth connection with a range of up to 10 meters. The system provides more than eight hours of operation time. The headset is flexible and can be resized in order to fit adolescent (age 6) through adult heads. B-Alert X series also provides multiple options for data synchronization.

Figure 11. B- alert X series device

4.2.2 **GSR devices**

Galvanic Skin Response is a measure of skin conductance. Skin conductance levels depend on the amount a person sweats; the more a person sweats, the higher the conductance levels will be. GSR signals consist of two components, a phasic and a tonic one. The tonic signal is a slowly varying signal, while the phasic component is the fastest changing part. Phasic component is related to external stimuli or non-specific activation. More particularly, if a person carries out a task that requires concentration, or thinks about something exciting or stressful, it will cause a noticeable response on the GSR signal. Other events, like a sudden noise (phone ringing) in a quiet environment can also cause GSR responses. As with EEG recording devices, there are various GSR recording devices and every one of them offers different capabilities in the recording of GSR signals. Below the two lightweight devices investigated for acquiring GSR data in MindSpaces PSS data collection are presented.

NeuLog 217 [9] (Figure 12) is a wireless device that includes two GSR probes attached by means of durable rubber-coated wires and two white Velcro finger connectors. Connection to the computer can be done via either Bluetooth or Wi-Fi with an available SDK. NeuLog 217 comprises of several smaller connected pieces (module NUL-217 for GSR, NUL-208 for HR, Wifi-200) thus becoming an inconvenient and uncomfortable device.

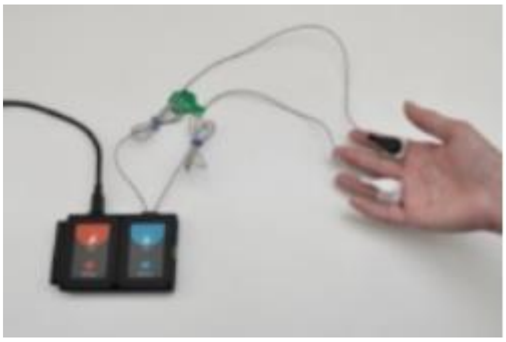

Figure 12. Neulog 217 device

Shimmer3 GSR + UNIT [10] (Figure 13) is a portable, wireless device that can measure the Galvanic Skin Response of the subject’s skin. In response to internal and external stimuli, sweat glands become more active, increasing moisture content on the skin, allowing electrical current to flow by changing the balance of positive and negative ions in the secreted fluid. Shimmer3 GSR + UNIT is able to measure this balance. Shimmer3 has a frequency range of 15.9 Hz and connects to the computer wirelessly via Bluetooth. It has a rechargeable battery of 450 mAh and an integrated 2 GB micro SD card slot for temporary storage. It is a portable device that does not cause any discomfort to the users since it utilizes only two sensors to acquire GSR data. Moreover, it is compatible with external applications and supporting hardware accessories.

Figure 13. Shimmer3 GSR + UNIT device

4.2.3 Hardware considerations related to MindSpaces

A suitable recording device should fulfill certain conditions according to the aims and needs of MindSpaces project about wearability, ease of use and efficient recording. The development of an effective and reliable data collection experience requires the consideration of several factors, including:

- the ability to obtain high quality signals from a wide range of people with different head or hand shapes and sizes, different types and lengths of hair and different properties of the scalp;
- an acceptable number of electrodes;
- sensor attachment;
- user comfort issues and coexistence with a virtual environment headset;
- other practical issues such as simplicity of application, portability and cost.

Enobio headset fulfills the above requirements and can be a good choice, in terms of quality/price ratio. The flexible head cap design of Enobio device allows adapting the headset electrodes to the specific needs of the project, by using disposable electrodes for the lab experiments and dry electrodes for the out-of-the-lab recordings. This flexibility regarding the number and type of electrodes makes it easy to use and tolerable for the subjects to wear. In addition, Enobio provides flexibility between the number of channels and head position of the electrodes. Furthermore, it is wireless, mobile, precise and compatible with other external systems and sources. For these reasons, Enobio is chosen for EEG data collection for MindSpaces project at this stage.

Shimmer3 GSR + UNIT device seems the most suitable device for the needs of MindSpaces project for GSR data collection, thus we select this device to acquire GSR signals from subjects. It is a wireless, portable, easy to use device that will not cause any discomfort to the user in combination with the Enobio device and the virtual reality headset, since it has small dimensions (65x32x12 mm) and light weight (30g).

4.2.4 SDK

As already mentioned, each one of the selected devices for the MindSpaces project provides an SDK that allows the connection and exchange of data between the recording devices, in this case Enobio, Shimmer3 GSR + UNIT and the PC/laptop.

Enobio provides Neuroelectrics Instrument Controller (NIC2) software. NIC2 is an integrated environment for management of Enobio devices from the computer. It offers basic and advanced modes to design and monitor any experiment involving EEG. It is compatible with Windows Vista/7/8/10 and Mac OS X operating systems and it requires a processor of 1.6 GHz and 2 GB of RAM memory. It can operate in two modes, an online which allows the performance and recording of interventions and experiments, and an offline. The offline mode permits the review of previously recorded files.

NIC2 communicates with the PC/laptop via Bluetooth and data are sent through the application displayed in Figure 14. Once the device is paired to the computer, it will appear in “My devices” section of the NIC2 environment. If nothing appears in this section, clicking the “SCAN FOR DEVICES” button will refresh it and finally the Enobio device will appear. On the “Settings” section, user can lock/unlock TCP connections to and from NIC2, activate double blind mode, enable synchronizer, activate line noise filter to remove main line artifacts from EEG data etc. The connection and transmission of data will start after selecting the “NE-ENOBIO8” on “My devices” section and by clicking the “USE THIS DEVICE” button.

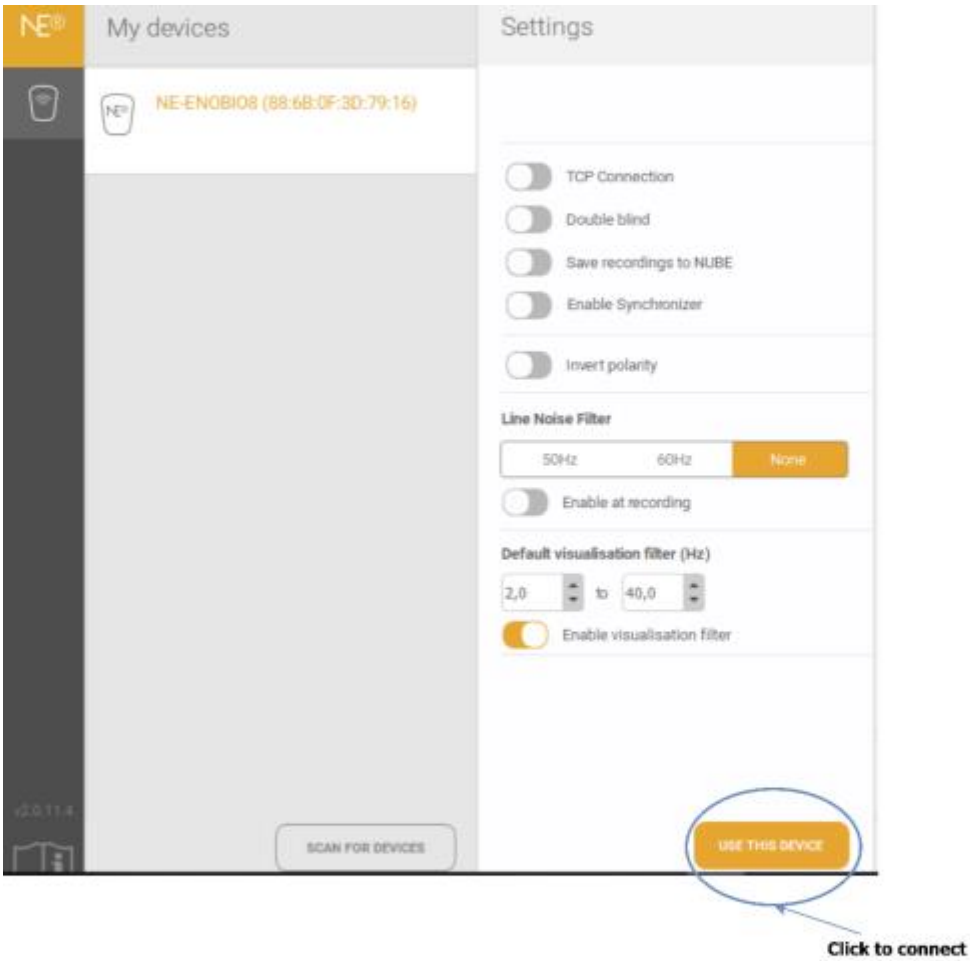

Figure 14. NIC2 interface to connect to Enobio device

Shimmer3 GSR + UNIT provides an API library that allows the GSR data to be streamed in real time from the device to the computer via Bluetooth. Communication with Shimmer is established using Bluetooth Serial Port Profile. This API provides the necessary building blocks in order to interact with a Shimmer device. It works with both Windows and Linux. An interface was available upon the Shimmer API that is compatible with LSL library. The necessary components are:

- ShimmerCapture.dll
- SDKShimmer.dll
- ShimmerDevice.exe

By running the "ShimmerDevice.exe" application, the Shimmer device communicates with the computer via Bluetooth and data are sent through the application displayed in Figure 15. Through the "Com Port" bar the port that sets up the link with the device is chosen and by clicking the "Link" button, connection with the Shimmer3 GSR + UNIT device is established.

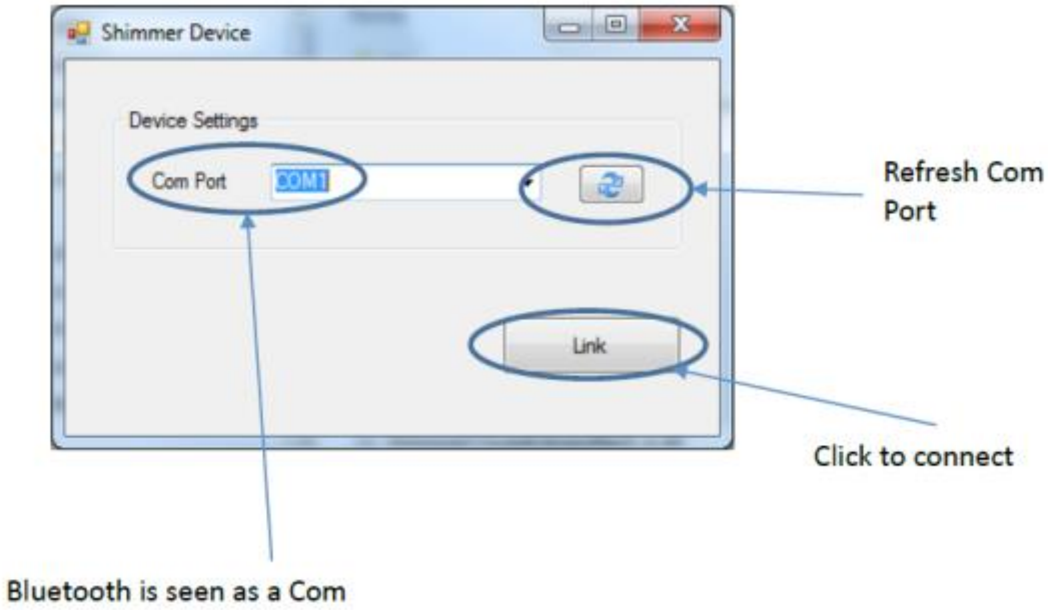

Figure 15. Shimmer Device interface to connect the Shimmer3 GSR+ unit to computer

4.3 Synchronization

In order to synchronize the different types of physiological signals for MindSpaces PSS data collection, an external tool is used. This tool is the LabRecorder software [11]. The LabRecorder is the default recording program that comes with LSL [12]. It allows recording all streams on the lab network (or a subset) into a single file, with time synchronization between streams. The file format used by the LabRecorder is XDF [13]. This new open general-purpose format was designed concurrently with LSL and supports all features of LSL streams. There are importers for MATLAB (included with the distribution), EEGLAB, BCILAB, Python, and MoBILAB. The source code has been compiled with Microsoft Visual Studio

2017, producing the executable file "LabRecorder.exe". Bellow there is a brief overview of its usage.

LabRecorder displays a list of currently present device streams under "Record from Streams" (Figure 16). If a device has been turned on while the LabRecorder is already running, by clicking the "Refresh" button the list can be updated (this takes ca. 2 seconds).

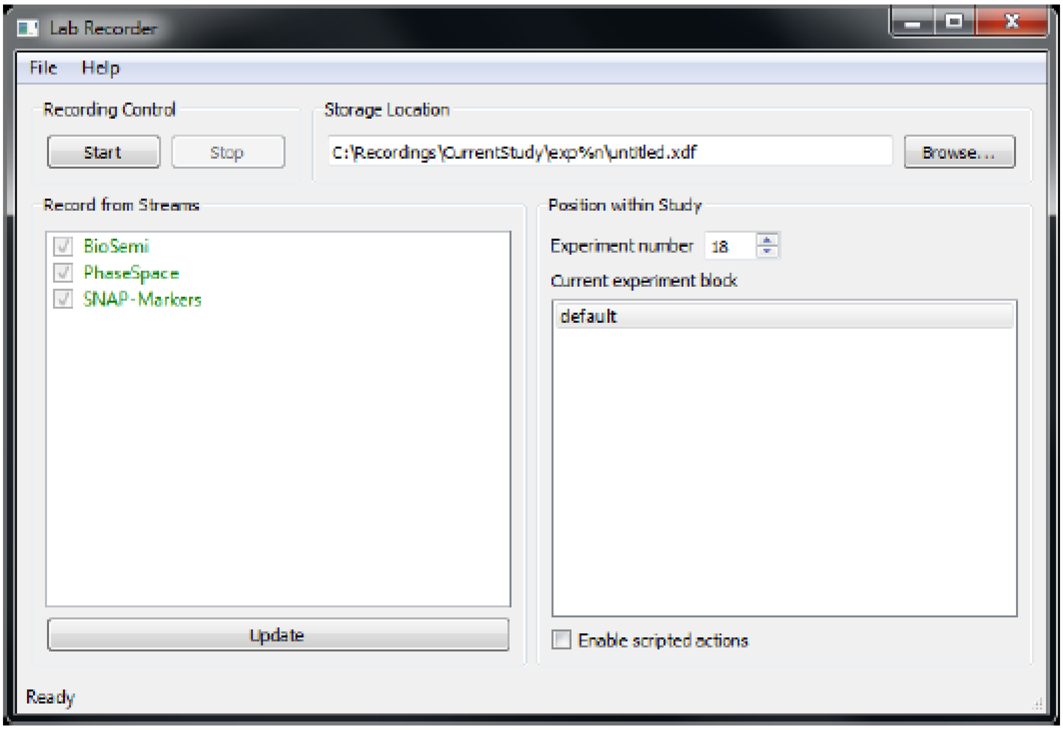

Figure 16. LabRecorder Interface

The LabRecorder displays the online streams with green color. The desired stream for recording can be selected by checking the box next to it. By default, the online streams will be initially checked. The file name (or file name template) where the recording will be stored, is defined in the "Storage Location" tab. If the file already exists, it will be renamed (the string _oldX will be appended where X is the lowest number that is not yet occupied by another existing file). This way, it is impossible to accidentally overwrite data. The "Start" button initiates a recording and the "Stop" button terminates it. Once the user clicks the "Start" button, the status bar will start displaying the time since the beginning of the recording and the current file size (e.g. Recording (00:00:07; 16756kb)…). The file size will grow slowly.

4.4 Preliminary experiments

Preliminary experiments took place in the context of MindSpaces project, to obtain an initial physiological signal dataset related to Pilot Use Case 2 (PUC2) and test the equipment and the recording techniques. Purpose of PUC2 is to redesign workplaces according to users' emotional, cognitive and environmental responses to provoke worker engagement, inspiration, interaction and productivity, while improving functionality. Thus, the aim of the experiments is to elicit emotional states via different working environment configurations using screenshots from a 3D application developed and provided for Mindspaces by Zaha Hadid for the immersive first person controlled experience of different design configurations of 3d workplace models. Using ZH's application, configurations of design elements can be changed and adjusted dynamically while a person is experiencing them in real time. These configurations were based on design parameters defined for PUC2. A more detailed description regarding design parameters for the preliminary experiments can be found in section 4.4.4. A full description regarding the design parameters for every PUC can be found in D5.3. In these preliminary experiments no VR equipment was used.

4.4.1 Emotional models

The term *emotional state* refers to the psychological and physical state in which emotions interact with behaviour and are evaluated in a common context. Emotional states can be described with a variety of models. Among all these models, two are the most prevalent ones to model an emotional state and therefore to recognize it: the discrete model and the two dimensional valence - arousal space model [14].

In the discrete model, according to Ekman, emotions can be clustered in six basic categories (happiness, sadness, surprise, anger, disgust and fear). Other emotion categories, such as jealousy, boredom, terror, excitement, satisfaction etc. were added later by Ekman, extending the list of the basic emotions [15].

The two dimensional valence – arousal model was initially suggested by Russel [16]. In the two dimensional valence - arousal space model, emotions are not presented as discrete states but as continuous ones and are represented in the two – dimensional space, characterized by two metrics, valence and arousal.

Valence reflects the level of pleasure. It ranges from negative to positive values and these values represent the "bad" feelings (e.g., angry, sad) and the "good" feelings (e.g., happy, relaxed), respectively. Arousal is a measure of excitement. It reflects a person's physiological reaction to a stimulus and it ranges from low to high values [15].

Most studies adopt the two dimensional valence – arousal model, since it is more suitable for a variety of feelings and the basic six emotions, defined in the discrete model, can also be represented through the two dimensional valence – arousal model as displayed in Figure 17. The two dimensional valence – arousal model is used to describe the emotional states in the MindSpaces' physiological signal sensing.

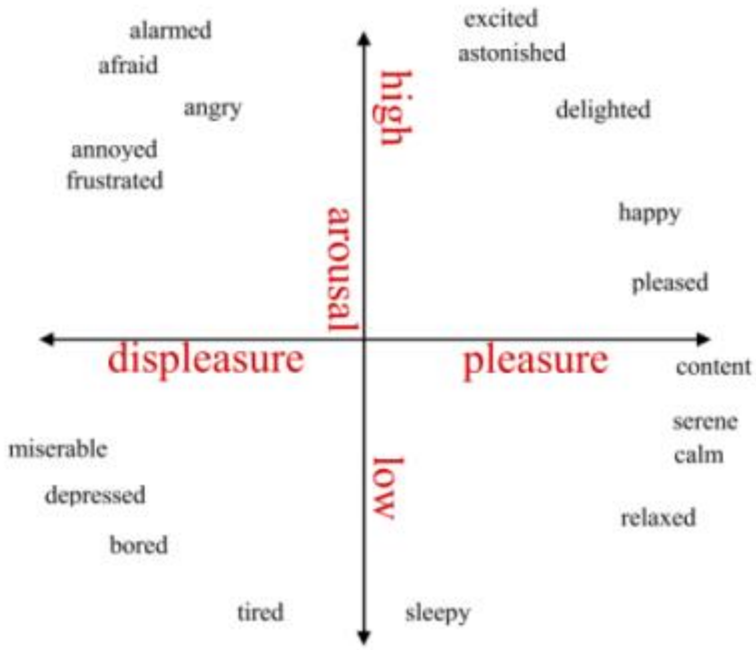

Figure 17. Mapping discrete emotions to the two dimensional valence – arousal space

Valence and arousal levels during an experiment can be evaluated through Self-Assessment Manikin (SAM) questionnaires. SAM questionnaire is an emotion assessment tool that uses graphic scales, depicting cartoon characters expressing two emotion elements: valence and arousal. SAM questionnaire is usually filled after the user is familiar with the stimuli [17].

SAM is displayed in Figure 18. First row is used to evaluate valence, cartoon character on the far left represents negative valence, i.e. negative emotions (Low Valence) while the cartoon character on the far right represents positive valence or positive emotions (High Valence). The second row is used to evaluate arousal, cartoon character on the far left represents low arousal (e.g. calm), while cartoon character on the right represents high arousal (e.g. excited).

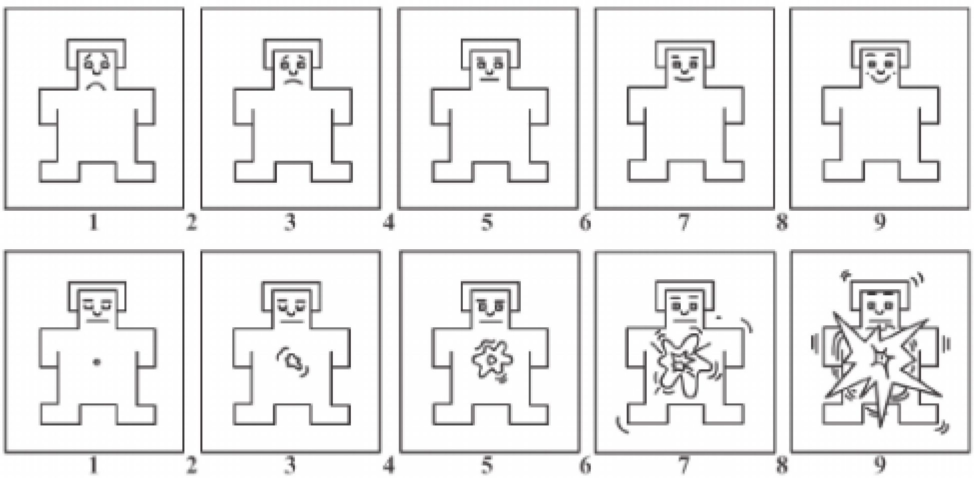

Figure 18. SAM questionnaire

4.4.2 **Emotion elicitation**

The first step in the field of emotion recognition from EEG and other physiological signals (e.g. GSR) is the evocation of emotions. Emotions can be subject-elicited (related to specific memories of the subject) or event-elicited (the stimuli is related to specific events).

In this section, we will describe the initial step, the evocation of event-elicited emotions and the experimental procedures that usually take place for signal acquisition. The emotional stimuli is very important in the field of emotion recognition in order to effectively evoke emotions to the subjects. This emotional stimuli can be visual, acoustic or audio-visual.

Images can be considered as the most typical example regarding visual emotional stimuli. International Affective Pictures System (IAPS) is the most popular such database [18]. However, its extensive use has reduced the impact that images have had on users as their knowledge on them has increased [16]. In addition, the limited number of IAPS images and their specific themes have challenged its effectiveness for studies that wanted to focus on a particular emotion or for experiments that required many trials of the same type [16]. Due to these reasons, other databases are also available, such as the Geneva Affective Picture Database (GAPED) [19] and the Nencki Affective Picture System (NAPS) [20].

The International Affective Digital Sounds (IADS) database utilizes acoustic emotional stimuli. IADS is developed by the same organization that developed IAPS and it is equivalent to IAPS, with the only difference that uses sound as emotional stimuli instead of images [21]. Apart from IADS, Oxford Vocal Sounds Database (oxVol) [22] is also popular on evoking emotions with acoustic stimuli.

Finally, many studies use video, which is an audio-visual stimuli to evoke emotions. The most popular databases regarding video stimuli are FilmStim Database, LIRIS-ACCEDE database, Mahnobhci Database, Emotional Movie Database (EMD), MIT Dataset [23] and DEAP database [24].

In MindSpaces project, on the preliminary experiments conducted so far, we are using visual emotional stimuli. Due to the nature of the project, no database that utilizes visual emotional stimuli of indoor spaces content is available; hence, we extracted images from the Unity built 3D application / 3D model developed by Zaha Hadid to use as emotional trigger (see 4.4.3).

4.4.3 **Zaha Hadid Interactive 3D Experience Model - The stimulus**

As already mentioned, a 3D application was developed by Zaha Hadid for Mindspaces enabling users and subjects of experiments to have an immersive and interactive experience of 3d workplace environments while configurations of the parameters in the environment are changed in real-time by a researcher on a separate screen through its interface. The virtual model simulates a 3D working environment and contains 13 different design choices of the type and layout of the desks in space along with a range of options for artificial and natural light, color, contrast, and materiality used as emotional stimuli.

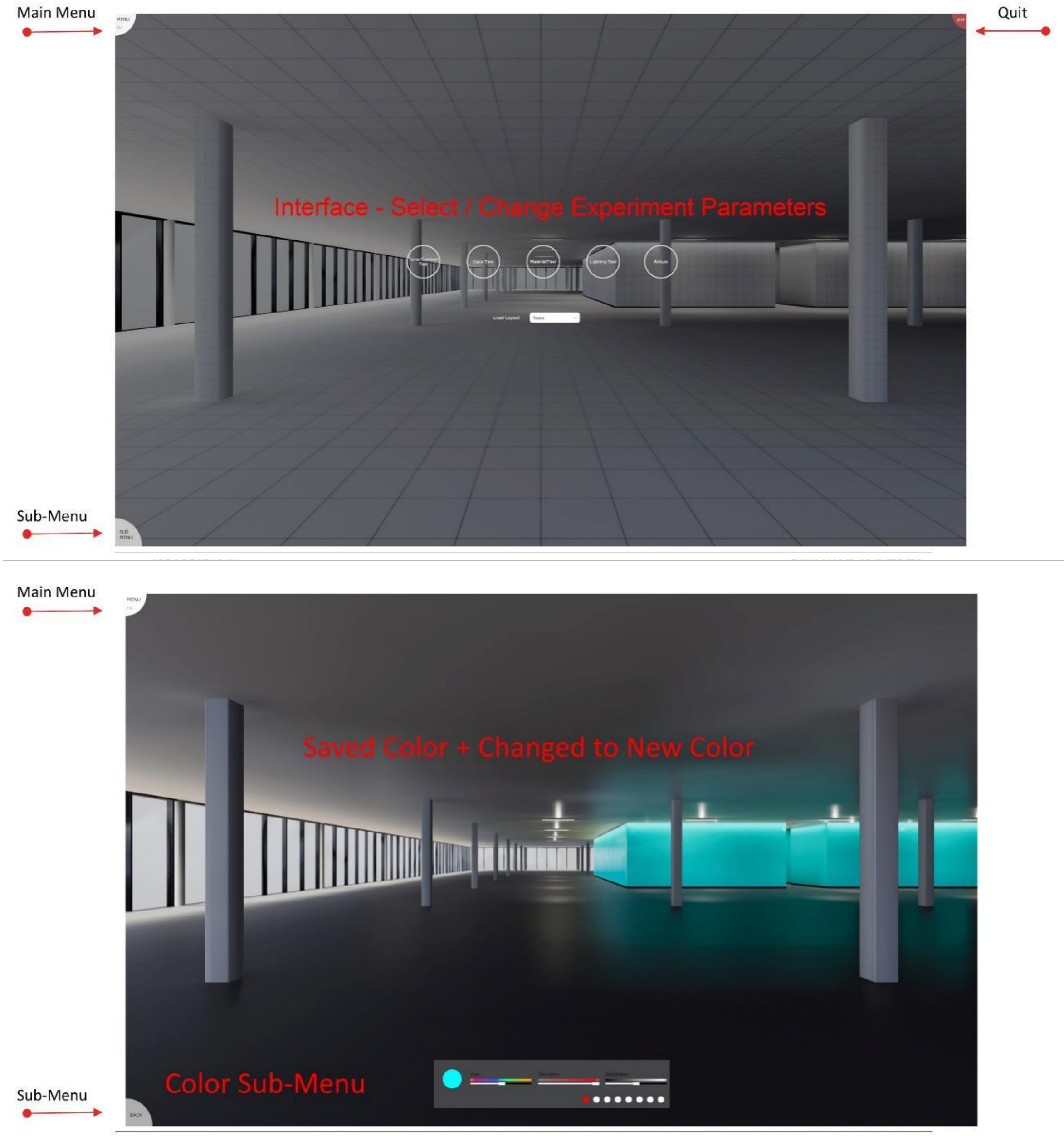

Figure 19. Interface menus allowing selection and changing of workplace environment parameters

The model currently contains 13 different layouts of workplaces. Each layout defines a different working space and thus, can change the experience a user has in a space. In every layout, desks have different shapes; round, square etc. Furthermore, they are placed in different configurations in space, i.e. placed in groups of four, facing the wall or windows etc. In addition, the desk number differs in every layout. Some of them are more crowded, while others have less desks and more open space.

Among the 13 different choices of the layouts, some have distinctive differences between each other, while others have minor ones. An example of two different layouts appears in Figure 20. This figure indicates that the available layouts of the model create a different working environment. It presents two layouts with distinct differences; desks have different shapes and different order in space.

Figure 20. Different layouts of the ZH 3D model with distinct differences

Figure 21 presents two similar layouts. Only difference between them is the surrounding walls.

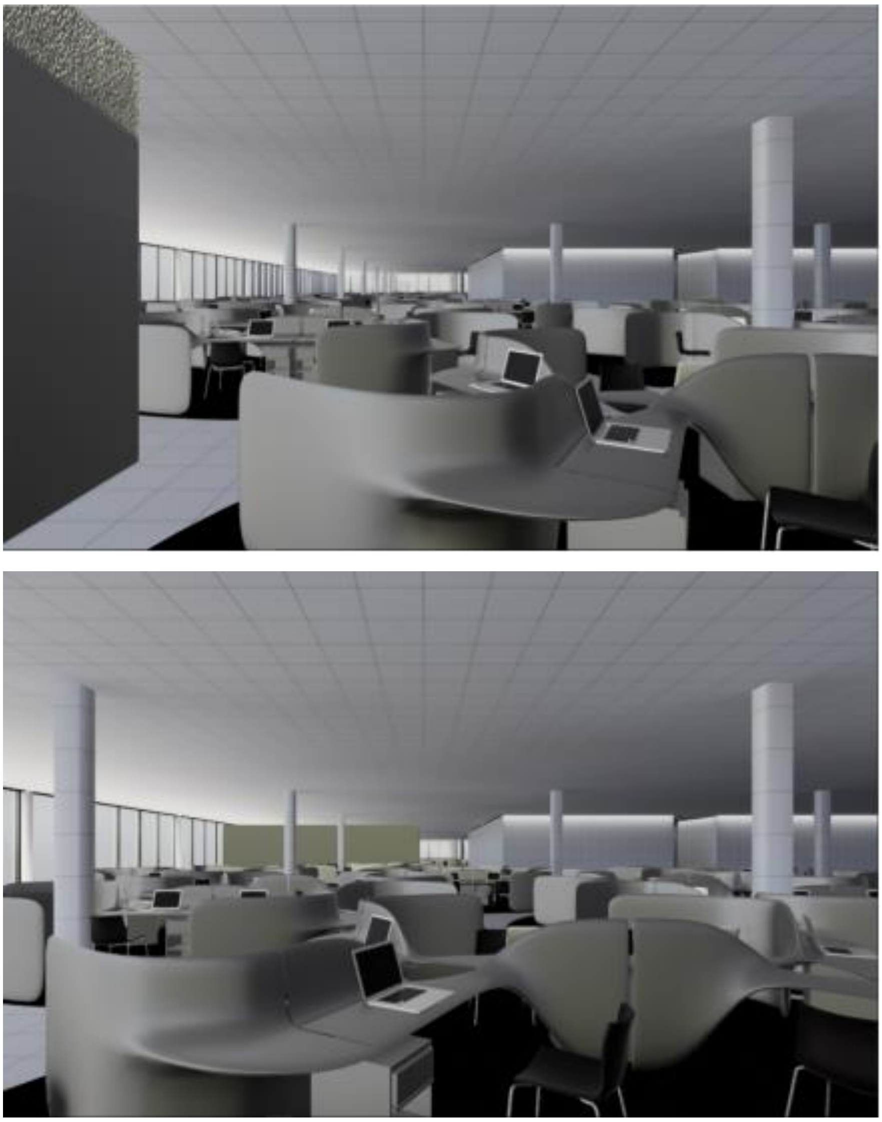

Figure 21. Different layouts of the ZH model with minor differences

In order to avoid long-lasting experiments, we decided not to use all the available layouts, so we selected a set to use during recordings. Thus, we used layouts that have distinctive differences between them and can represent different office spaces.

According to [25], valence levels are affected by presence or absence of windows and the size of windows (if present), symmetry of objects, existence of spatial alignment and design of spaces. Arousal levels are related to density and openness of spaces. We selected six layouts that seemed to have more differences between them, in terms of number of desks, shape of desks and layout of desks in space and discarded the ones with minor differences between them. The different layouts will form the Office Layout (OL) instances. Thus, in total, six OLs will be displayed to the subjects.

Apart from the different office spaces (defined by layouts), the ZH application has several other parameters that can be changed in every layout through an interface, contributing to the experience a user can have in each OL. These parameters control lighting, materials, color, and contrast on walls, the ceiling and the floor. Regarding lighting parameter, the user can experience direct or indirect, artificial (light bulbs) or natural (windows) lighting. Different textures on the floor, ceiling and walls (wood, marble etc.) can be used on the materials parameters. Finally, color parameter refers to the different colors (yellow, purple, red etc.) on the walls of each OL.

4.4.4 Experimental procedure

We decided to test the selected layouts by changing one parameter every time and estimate the emotional responses of the subjects. More specifically, based on [26] we will examine the lighting parameter since direct light increases arousal and indirect light decreases it. Another parameter we change is color. It has been shown [27], [28] that red color provokes anger, yellow provokes sadness while blue and orange describe calm and pleasure feelings, respectively. Finally, regarding the materials and textures, authors in [29] claim that wooden environments tend to produce less tension than the non-wooden environments.

Thus, three parameters were tested, lighting, materials and color. In each one of these parameters, different values will be used regarding different parts of the environment, e.g., walls, ceiling etc. Lighting is divided in two concepts, natural and artificial light and each concept can take two values. Second parameter is materials, which is divided in three concepts, floor, ceiling and walls, and each concept can take two values, and finally color, which has only the wall concept and can take four values. Thus, every OL will have 14 different architectural scenarios (ASC) in total. The different architectural scenarios are presented in Table 3:

Table 3. Parameters and their values. We examine three parameters, lighting, materials and colors.

Parameters	Concept	Value
Lighting	Natural	Side
		Middle top
	Artificial	Top
		Side
Materials	Floor	Wood flooring 044
		Black_03
	Wall	Wood fine 011
		Marble_calacatta_002
	Ceiling	Ceiling 02
		Ceiling 04
Color	Wall	Red: RGB = (255, 0, 0)
		Blue: RGB = (0,176, 255)
		Orange: RGB = (255, 94, 0)

		Yellow: RGB = (255, 211, 0)

The instances in Figure 22 represent two different architectural scenarios of the same OL, screenshotted from the 3D model and used in the experiment, with different values on the materials parameter. On the top figure, the floor has black tiles while on the bottom figure the floor is wooden.

Figure 22. Example of two instances of the 3D model used as emotional stimuli with different materials on the floor

In order to evoke emotions to the subjects, using images as the emotional stimuli, there is a certain protocol followed in the literature, regarding the experimental procedure [30], [31], [32], [33]. The subject wears the recording device (or devices if more than one physiological signals are measured) and sits on a comfortable seat in front of a screen. Usually, images are sequentially previewed, separated by a black screen, a fixation cross screen or counting down frames. The black screen phase offers time for the subjects to relax. The countdown phase, or the fixation cross screen phase is used to alert the subject to concentrate on what will follow (e.g., the emotional trigger). After the projection of each image, the subject can proceed to the self-assessment phase (SAM questionnaire), where it can evaluate each picture in the two dimensional valence-arousal space according to her/his feelings. The experimental procedure for the preliminary experiments conducted for MindSpaces project follows this experimental protocol.

There are 84 architectural scenarios in total, as there are six layouts and 14 different architectural scenarios on each layout. Each one of the 84 architectural scenarios corresponds to an experimental trial. In every trial, the display of Black Screen (BS), fixation cross and architectural scenario (image from a screenshot of the 3D model) lasts for 17 seconds. Additional time is provided in order for the subject to fill in the SAM questionnaire. SAM assessments of the subjects will be considered as the ground truth in future analysis of the physiological signals. Each trial lasts for 17 seconds plus the necessary time to complete SAM. The block diagram of the experiment protocol is displayed in Figure 23.

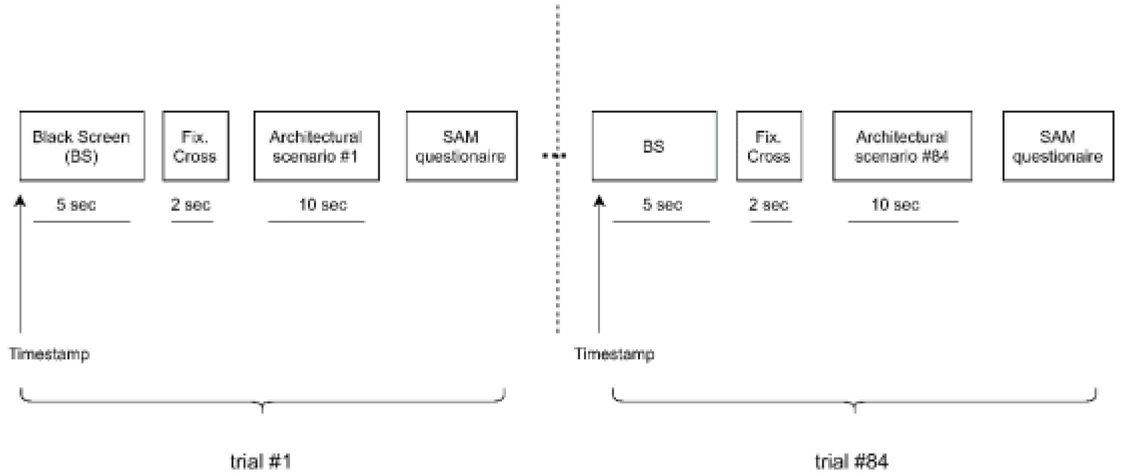

Figure 23. Block diagram explaining the experimental protocol between the trials

In order for the experimental procedure to be more pleasant and relaxing for the user, we decided to break the experiment into three sessions. Based on this separation, each session consists of 28 trials. Between the sessions, subjects are able to take a break and relax for three minutes. One minute before the break ends, a message is displayed on the screen to notify the subject in order to get ready for the next session. Since the experiment consists of three sessions, two breaks of three minutes each are included. The block diagram in Figure 24 describes the procedure between the different sessions and the breaks.

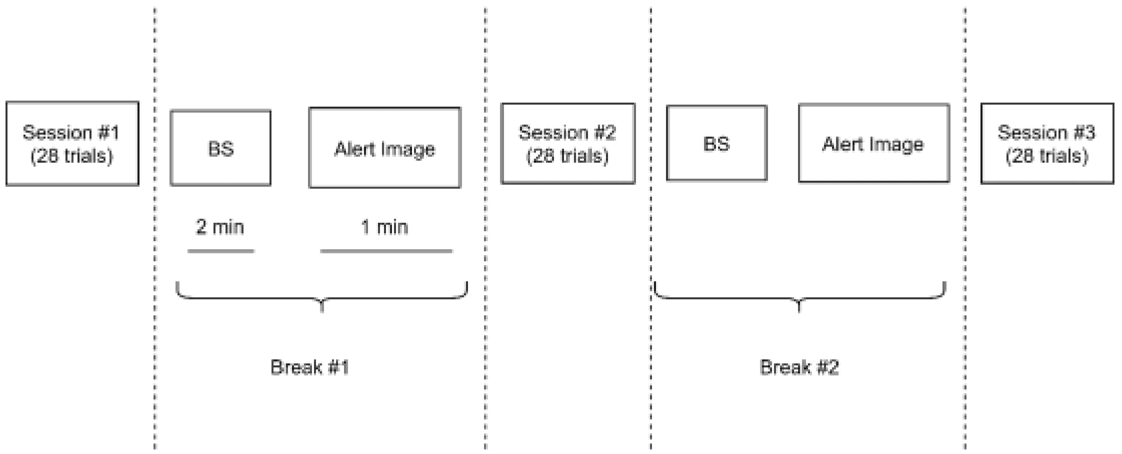

Figure 24. Block diagram explaining the overall experiment protocol between sessions

4.4.5 EEG device configuration

Enobio headset offers three different headcap options. Each one of them uses a different number of electrodes to capture EEG signals. The first one uses eight channels, the second one is using 20 and the third one is using 32 EEG channels. In all three of these headcaps, electrode positions are flexible.

For the MindSpaces experiments, we use the Enobio 8 device that uses eight electrodes in order to capture EEG signals. This number of channels can efficiently capture EEG signals related to emotions, without making the subject feel uncomfortable wearing it in combination with the virtual reality headset (in future experiments) and the GSR device. For MindSpaces project, based on [30], [34], [35] and [36], and the fact that Enobio 8 will be used in combination with a virtual reality headset, the selected EEG channels are: FC1, FC2, AF3, AF4, F3, F4, F7, F8 (see Figure 25).

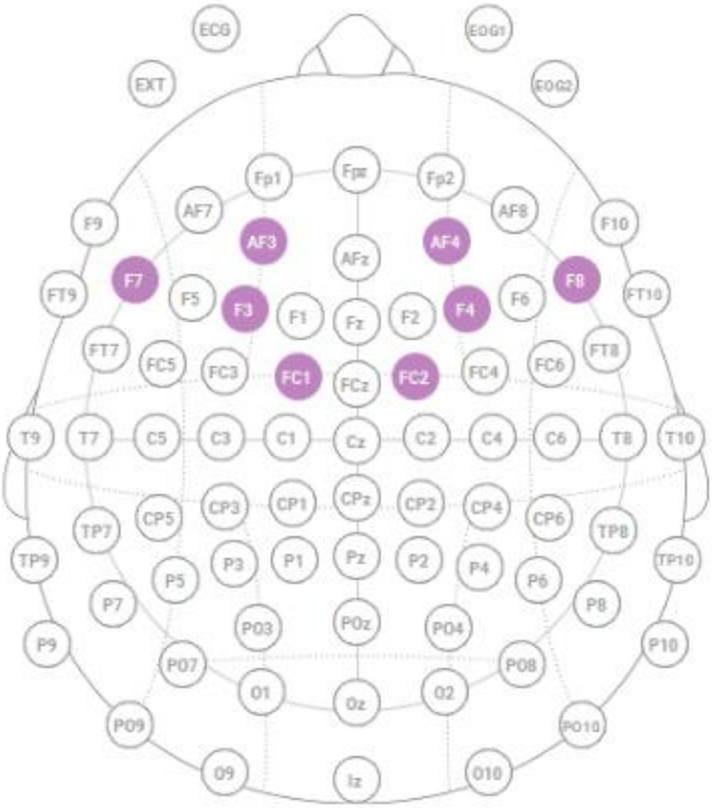

Figure 25. Electrode positions for Enobio headset

The correct headset setup is necessary in order to efficiently capture EEG signals. A short description of the Enobio 8 headset setup is described below:

- If using dry electrodes (see Figure 9), place them in the selected electrode position. If using solid gel or gel-based electrodes (see Figure 10), place the screwable spare part of the EEG monitoring electrodes in the selected electrode positions.
- Place the headcap on the subject's head. Using both hands, slide the headset down from the top of the head so that Fp1, Fpz, Fp2 electrodes abstain two fingers distance

from the subject's eyebrows and the Fpz electrode is in the imaginary line above the nose.

- If using solid gel or gel-based electrodes, place solid gel or gel into the screwable spare part respectively and then screw the electrode.
- Place the electrode clips of the electrode cable in every electrode.
- Open Enobio 8 device and pair it to the computer.
- Open NIC2 software, connect to the device and check electrode impedance. If all electrodes are marked with color green in the interface, then connection is established and EEG collection can start.

4.4.6 GSR device configuration

Shimmer ADC can convert a voltage, provided by the hardware, to a 12-bit number. This number represents the external skin conductance value. Skin resistance can be derived through skin conductance value. Shimmer3 GSR + UNIT design can resolve skin resistance levels from 10kΩ to 4.7MΩ (100uS to 0.2uS).

It is recommended for Shimmer3 GSR + UNIT to use snap connector Ag/AgCl electrodes and the surface area should not exceed 1 cm^2. Apart from Ag/AgCl electrodes, re-usable Velcro strap electrodes could be another option. In addition, if positioned correctly, sticker electrodes would also work.

One option on electrode positioning is the following: One electrode should be placed on the palmar surface of the medial phalange and the other on the palmar surface of the distal phalange.

An alternate option is placing one electrode on the index finger and the second electrode on the middle finger as displayed on Figure 26. This electrode positioning option is the one adopted for the MindSpaces experiments.

Figure 26. Electrode positioning for GSR

As a safety measure, GSR electrodes were not applied while the Shimmer3 unit is in a USB dock or multi-charger.

4.4.7 Calibration

For the EEG configuration, calibration consists simply by checking the impedance of each electrode. Generally, when the impedances are below 5 kΩ, it is considered to be a successful EEG signal acquisition procedure. Through NIC2 software interface, it is easy to check the impedance on the electrodes. If the impedance was above 5 kΩ, we corrected the contact between the electrodes and the scalp by adding more gel (in the case of gel-based electrodes) or by pressing/moving the electrodes for better contact to the skin (mainly in the case of dry electrodes).

For Shimmer3 GSR + UNIT device, no calibration is required.

4.5 Dataset and pre-processing

Any device that uses electrodes as sensors is sensitive to noise and artifacts. There will be an artifact present in the signals, every time there is movement at the site of the electrodes. Artifacts can be regarded as due to physiological sources (body movements) or to non-physiological sources (external environment).

In EEG signals, someone can encounter main noise (at 50 or 60 Hz) and physiological artifacts (i.e., eye blinks). Unfortunately there is not any way to prevent the creation of such kind of artifacts, so signal processing techniques are applied to EEG signals in order to remove them.

After acquiring EEG signals, digital signal processing techniques can be applied in order to simplify subsequent processing operations without the loss of relevant and important information. The most common EEG processing techniques can be distinguished in frequency and spatial filters. In order to enhance the quality of the signals, various filters can be employed. These filters can belong in both categories and are able to remove artifacts and noise [37], [38], [39].

For the pre-processing of the EEG signals captured during the preliminary experiments, we used frequency filters. A tenth-order bandpass Butterworth filter was applied to EEG signals between 8 – 30 Hz in order to maintain alpha and beta frequency bands to remove noise and superimposed artifacts from several sources [40]. Alpha and beta frequency bands have been chosen because artifacts related to ocular activity (blinking, eye movements etc.) perspiration (sweat glands) and respiration (inhale, exhale), effect in lower frequencies and can be confused with delta and theta frequency bands [41], while muscle activity artifacts (swallowing, chewing etc.), effect in higher frequencies such as gamma frequency band [42]. After visual inspection, no additional artifact removal considered necessary in order to take place at this stage.

In GSR signals, artifacts can take the form of a high frequency noise-like component. This may happen if the electrodes are not tightly attached to the skin and lose contact. In order to avoid their generation, electrodes are attached as securely as possible to the skin. In addition, if electrodes are tightly attached to the skin, apart from artifact creation, they will measure a higher conductance level. A low pass filter should be applied to the data in order to remove the high frequency noise that can be attributed to moving objects and other noise components. A cut off frequency as low as 1-10 Hz can be used without affecting the data of interest due to the slowly changing nature of the GSR responses [43], [44].

4.6 Participants

Before physiological signal data acquisition, subjects provided the details of their personal information by completing a questionnaire. This questionnaire consists of information regarding their age, gender and nationality, if there are right or left handed and finally, their level of education and whether they had any education specific in art field since MindSpaces is an art-driven project. A Personal Identification Code (PIC), a unique identification number generated for each participant, characterized every questionnaire (see D9.1). The form of the questionnaire can be found in Appendix I.

Each subject also read an information sheet regarding the project and signed a consent form before the beginning of the experiment. Consent form can be found in D9.1. Regarding preliminary experiments, 13 subjects (nine males, three females) participated in the experiments. Subjects were all Greek adult employees working in an office (age range: 26 – 42, mean = 32, SD = 4.5 years). All of the subjects were highly educated but did not have any specific education in art. Nine of them were right – handed while only four of them were – left handed. None of the participants suffered from any neurological illness or psychiatric disorder. Participants were instructed to rate all of the trials according to their real feelings while watching the figures instead of meeting the expected standards of their daily mood.

4.7 Data collection issues

Enobio8 and Shimer3 GSR and UNIT are wireless devices that connect to the PC/laptop through Bluetooth. Hence, some connectivity issues might appear during experiments. As already mentioned, 13 subjects participated in the preliminary experiments on MindSpaces project. Due to connectivity issues, one subject had to be discarded entirely (both EEG and GSR signals), resulting in 12 subjects while for one more subject only GSR signals had to be discarded. In sum, apart from mild connectivity issues for two subjects the proposed PSS data collection framework proved to be reliable for MindSpaces endeavors.

4.8 Future considerations

This section described the equipment selected for MindSpaces project, the synchronization between different types of signals and some preliminary experiments conducted so far for the project. The preliminary experiments used visual emotional stimuli (images instead of VR setting) and 13 subjects participated providing an initial dataset of physiological signals.

Until the end of the MindSpaces project, more experiments are scheduled to be conducted, using virtual environments as emotional stimuli instead of images. Virtual reality models will be developed from MindSpaces partners for each Pilot Use Case (PUC1, PUC2, and PUC3) and specific experiments will be conducted for each PUC with multiple subjects participating in each experiment. On these experiments, subjects will be able to navigate through the environment while different parameters (lighting, materials etc.) will change resulting in different experiences regarding the virtual space. During these experiments, the goal is for more subjects to participate, providing a dataset that will allow the development of highly efficient emotional state recognition algorithms.

Apart from additional experiments, new equipment will be investigated that can combine a virtual reality headset along with EEG signal acquisition, e.g., Looxidvr [45].

5 VISUAL SENSING FOR BEHAVIORAL ANALYSIS

5.1 Face Blurring Application to Collected Visual Data

Visual data collected for the MindSpaces project in many cases contain personal data, such as people's faces. More specifically, for PUC1 the CCTV cameras installed at the region of Tecla Sala may include personal data from citizens. In PUC3 the senior appears in the video footage, so it is important to anonymize data taking into consideration no to affect the performance of the visual behavioural analysis. In PUC2, we have also the employees of McNeel offices in Barcelona and the face blurring technique is considered in order to anonymize data and protect people. Maintaining the anonymity of the people that appear in the videos was a prerequisite for their participation of subjects in MindSpaces experiments. The people that participated at the experiments in PUC2 and PUC3 have provided their consent for their participation. In PUC2, an NDA was signed between CERTH, McNeel and MU describing the data collection and management of visual data. In PUC3, an information sheet was provided to the senior describing the research objectives and the entire process of visual data collection. In PUC1, which is a public space case, CERTH collaborated with L'Hospitalet partner in order to provide all the documents needed to receive the approval and get the footage from the installed CCTV cameras.

In general, the collection of visual data is not a simple process, since it involves sensitive data, thus it was important to respect people's privacy and follow all the proper actions. In D9.2 there is a detailed description regarding the documents and the actions that we had to take.

MindSpaces, in order to protect people's anonymity, uses a face blurring algorithm to hide their faces in the collected videos. Face blurring process consists of two phases. For the first phase, we use a deep learning network for the detection of faces in each frame of the video. Then, we blur the regions of the frame, where faces are detected.

For the detection of faces we used a Faster Region-based Convolutional Neural Network (Faster R-CNN)[2]. Faster R-CNN is a famous deep learning network used for object detection. It is the evolution of R-CNN[3] and Fast R-CNN [46] and, as its name reveals, it is much faster than its ancestors. Figure 27 shows the architecture of Faster R-CNN.

[2] http://papers.nips.cc/paper/5638-faster-r-cnn-towards-real-time-object-detection-with-region-proposal-networks.pdf

[3] https://arxiv.org/pdf/1311.2524.pdf

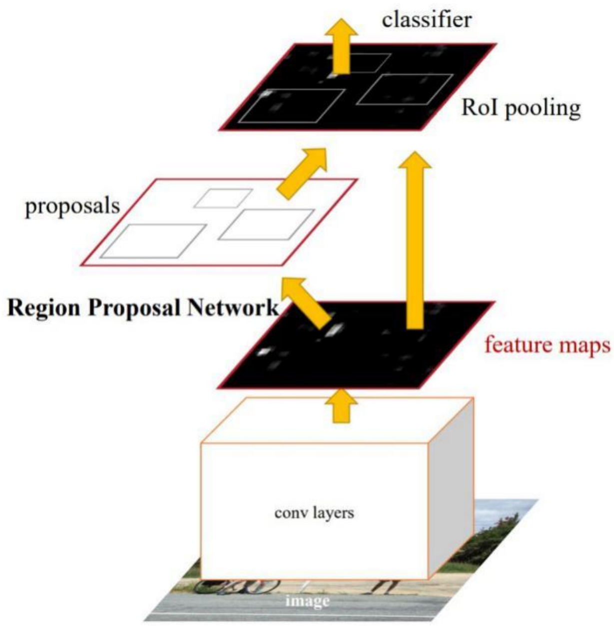

Figure 27. The Faster R-CNN architecture

Faster R-CNN comprises two modules. The first is a deep fully convolutional network (VGG16, ResNet101, etc.), named Region Proposal Network (RPN). The RPN takes an image as input and outputs a set of rectangular proposed regions that probably contain objects of interest. The second module is a Fast R-CNN detector [46] that gets these proposed regions as input and makes the detection. The entire system is a single, unified network for object detection and can be trained end-to-end using backpropagation and Stochastic Gradient Descent (SGD)[47].

For the Faster R-CNN that we used for the face detection process, we have chosen a VGG16 architecture for the RPN, pre-trained on the ILSVRC-2012-CLS image classification dataset[4]. The above dataset contains 150000 images, collected form Flickr and other search engines, labelled with 1000 object categories. We trained the entire Faster R-CNN on the WIDER FACE dataset[5] and then fine-tuned it using Large-scale CelebFaces Attributes (CelebA) dataset[6]. The WIDER FACE dataset is a benchmark dataset for face detection. It contains 32203 images with 393703 faces labelled in them. There is a variety in the scale, pose and occlusion percentage of the labelled faces. The CelebA dataset consists of 202599 celebrities' face images. We used also this dataset so that we expose out face detection network to a larger number of face images. Both datasets used are free to use. We used left-right flipping as a data augmentation technique.

4 http://www.image-net.org/challenges/LSVRC/2012/

5 http://shuoyang1213.me/WIDERFACE/

6 http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html

The face detection network creates a bounding box around each detected face in an image/video frame. The next step is the blurring process. The regions within the bounding boxes are blurred using OpenCV and Gaussian blurring. Figures 28 and 29 show results of our face blurring algorithm for a video captured outside Tecla Sala and in McNeel offices in Barcelona respectively.

Figure 28. Face blurring example in video captured outside Tecla Sala (PUC1)

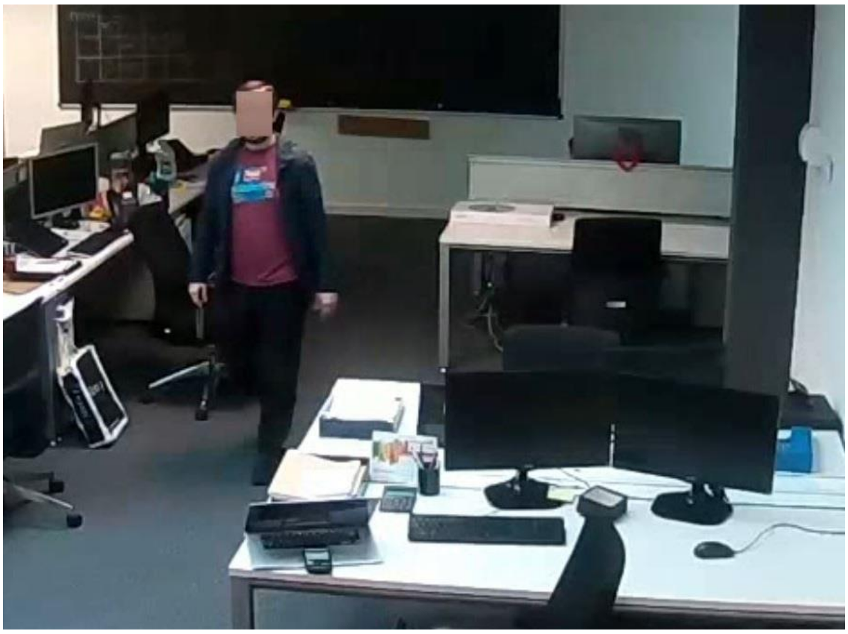

Figure 29. Face blurring example in video captured in McNeel offices in Barcelona (PUC2)

Due to the large volume of data, especially for PUC1, the face blurring process is time-consuming. In order to reduce the time needed, we initially check if there is any face detected in a video by analysing one frame per 30 seconds. Once face is detected, we proceed to the face blurring process, analysing every frame of the video. The face blurring algorithm runs on GeForce RTX 2080 Ti Nvidia GPU. Face detection and blurring takes about 0.2 seconds per frame.

Until M17 of MindSpaces project, we have applied our face blurring algorithm for all the videos captured from the senior’s home in Paris (PUC3) and from a subset of the videos collected from the Tecla Sala region (PUC1) that will be used in the evaluation of the 1st prototype.

5.2 Equipment for Visual Data Collection

Proper equipment is very important for the collection of data, as it can severely affect the data quality. Each PUC has its own equipment needs. For example, in some cases we need to capture outdoor footage (PUC1), while in others indoor footage (PUC2, PUC3). Moreover, we need to be able to record at day or night time (with good lighting conditions or with no lighting). The estimation of the 3D coordinates of space from visual data is also an important parameter that has to be considered for the proper installation of cameras.

For PUC1, until M17, we have used the CCTV cameras that are already installed at Tecla Sala. We obtained visual data from eight CCTV cameras that are installed at different points of the Tecla Sala area.

Two outdoor D-Link DCS-4703E cameras are also available, for better coverage of the region of Tecla Sala and they will be used in the future to enrich data. D-Link DCS-4703E is a high-definition professional surveillance and security outdoor video camera. It uses a high-sensitivity 3-megapixel progressive scan CMOS sensor to deliver truly superior quality video under a wide range of lighting conditions. The all-in-one rugged design makes this camera ideal for reliable deployment in surveillance.

Figure 30. The D-Link DCS-4703E camera for PUC1

For PUC2, the visual equipment contains 4 D-link DCS‐8525LH indoor cameras, two provided by CERTH and two by ESeniors to capture the video footage needed from McNeel offices in Barcelona.

For PUC3, CERTH provided two D-Link DCS‐8525LH indoor cameras that were installed at senior's home in Paris. D-Link DCS‐8525LH is a high-definition Pan & Tilt Wi-Fi indoor video camera. It provides also night vision functionality and motion detection.

Figure 31. The D-Link DCS-4703E indoor camera for PUC2 and PUC3

Table 4 below presents the equipment and the related specifications.

Table 4. Visual Equipment for Behavioural Analysis

Equipment	Specifications	Usage
D-Link DCS‐8525LH	• Indoor • Connectivity:Wired, Wireless • Resolution: Full HD 2 megapixel • Image sensor: 1/2.7" 2-megapixel progressive scan CMOS • Maximum Video Resolution: 1920 x 1080 (16:9) • Lens: 2.39 mm, F2.0, Fixed • Infrared-Cut Removable Filter • Night Vision • Night vision range: 5 m • Motorised Pan/Tilt • 2-way Audio • Built-in Microphone • Built-in Speaker • External card slot: MicroSD SDHC (max 32 GB) • Features:Amazon Alexa & Google Assistant, Cloud Recording, MicroSD Card slot • Video Format:H.264 • Mobile Stream • Digital Zoom: 4x • Motion Detection • Sound Detection	• 4 cameras of this type to be used for PUC2, 2 of them provided by ESeniors and 2 provided by CERTH • 2 cameras of this type used for PUC3 provided by CERTH
D-Link DCS-4703E	• Outdoor • Bullet • Wired PoE	• 2 cameras to be used for PUC1 provided by CERTH

	• Resolution: 3 megapixel • Image sensor: 1/3" 3-megapixel progressive scan CMOS • Maximum Video Resolution: 1920 x 1080 (16:9), 2048 x 1536 (other) • Angle of view (H/V/D): 70o/40o/80o • Lens: 3.6 mm, F1.8, Fixed • Infrared-Cut Removable Filter • Night Vision • Night vision range: 20 m • Wide Dynamic Range • LowLight • Privacy Masks: 3 zones • Features:ONVIF • Video Format:H.264, MJPEG • Multi Stream • Mobile Stream • Digital Zoom: 10x • Event Recording • Email notification • Recording to NAS	
Tecla Sala CCTV footage coverage of the area	• Resolution: 352 x 288 • Frame rate: 25 frames/sec	• Used 8 channels that capture video footage from different points of Tecla Sala

5.3 Visual Data Collection for PUC1

For capturing of video footage from the Tecla Sala region, we initially used the already installed CCTV cameras. There are 8 channels available. We initially received video footage from channels 1, 5, 6, 7 and 8. Figure 32 shows the exact points, where the CCTV cameras that correspond to these specific channels are installed. In blue colour we can see the area covered by channels 1, 5, 6, 7 and 8.

Figure 32. The points where CCTV cameras of the channels of our interest are installed in Tecla Sala area

Below, screenshots from CCTV cameras are presented.

Figure 33. Screenshots from CCTV cameras of Tecla Sala (channels 1,5, 6, 7 and 8)

The visual data consists of around 1350 video files and the total size of these files is around 330 GB. The duration of recordings contains visual information from 30/11/2019 to 18/12/2019. The mean duration of a video file is about 1 hour. The resolution of each video is 352 x 288 and the frame rate is 25 frames per second. The video format is mp4. In order to create a basic dataset to work on, we started the analysis of footage from channels 5, 6 and 8 focusing on the timeslot from 8:00 to 23:00.

In our future plan, two additional outdoor cameras will be used to capture citizens' activities and have a better coverage of the Tecla Sala area when physical art installations will be placed.

The provided data is stored in CERTH's FTPS server of MindSpaces project. The access to this server is permitted only for authorized users. More specifically, only those who work on the visual analysis part of the project could use this data. In addition, the physical access to the data in the server room is controlled.

5.4 Visual Data Collection for PUC2

The equipment installation at the McNeel offices in Barcelona is completed. As mentioned in section 5.2 of this deliverable, 4 cameras are available for PUC2, 2 provided by CERTH and 2 from ESeniors. Figure 33 presents a 3D model of the McNeel offices and the red circles correspond to the points where the 4 cameras is decided to be installed for our first data collection activity.

Tests have also been conducted to ensure that the equipment installed works properly and have the best coverage of area. However, the plans for the actual visual footage capturing for PUC2 changed due to the COVID-19 crisis. Once the McNeel employees return to their normal working status, the recording will start, so that we will be able to collect the data needed for this PUC.

For additional coverage of the areas, 2 more cameras may be installed, as shown by the yellow circles in Figure 33, after receiving and evaluating the first data.

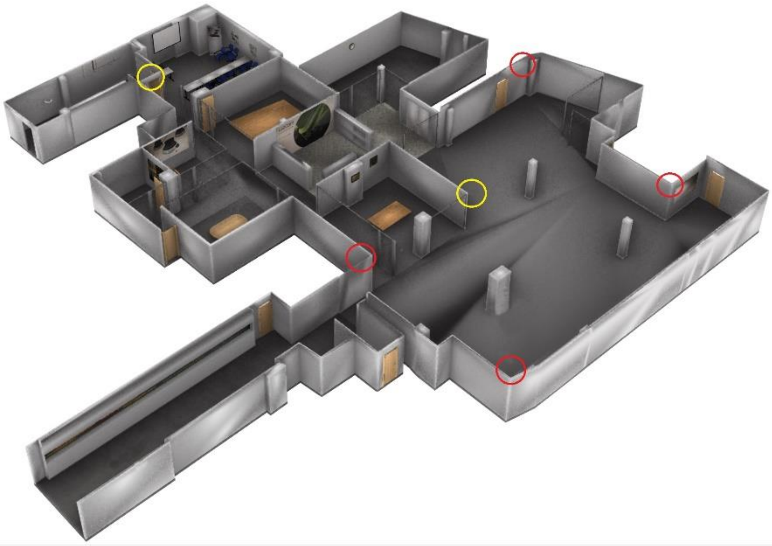

Figure 34. Red Circles: the points where the cameras are install in the McNeel offices in Barcelona
Yellow Circles: the points where additional cameras are considered to be installed

5.5 Visual Data Collection for PUC3

For the collection of visual data for PUC3, two cameras were installed at the senior's home in Paris. The installation involved two phases. At the 1st phase, the cameras had been installed to the red points shown in Figure 34. However, the red points in Living Room 1 and 2 could not fully cover all the activities in the desired areas. So, at the 2nd phase of the equipment installation, the camera of Living Room 1 was moved to a higher position (green point) and the camera in Living Room 2 was finally installed in another corner of the room (green point). Moreover, we have increased by one hour the daily timeslot of recording.

The collection of data for the PUC3 is completed. We have collected 550 videos of around 8 GB. The video resolution is 1920 x 1080 and the frame rate is 25 frames per second. The video format is mp4. The duration of recording of visual data in the senior's home lasted four weeks, starting from 22/11/2019 and ended at 18/12/2019.

The same procedure for the management of visual data is followed. The face blurring technique was applied to the collected videos. The initial data was deleted and only the anonymised data is kept to CERTH's FTPS server of MindSpaces project.

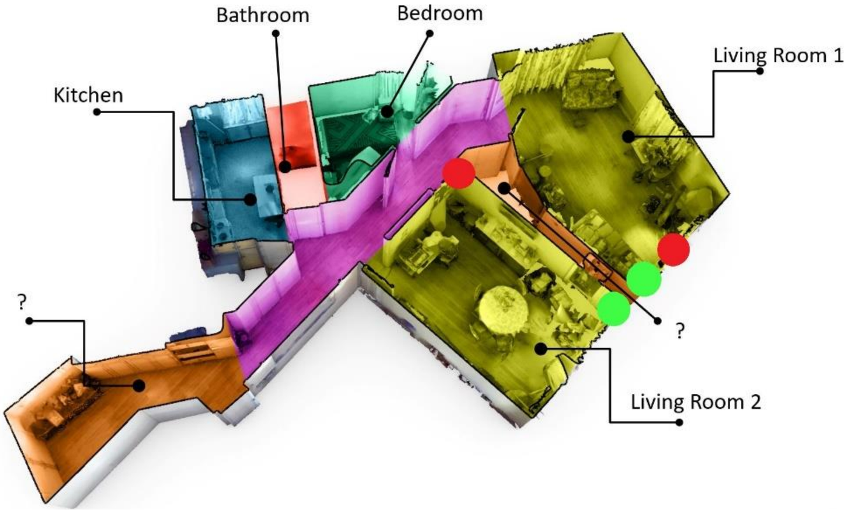

Figure 35. Red Circles: the points where the cameras were initially installed
Green Circles: the final positions of the cameras

Table 5 summaries the information of the collected data for PUC1 and PUC3.

Table 5. Summary of the collected data in PUC2 and PUC3

	Quantity of video files	Size	Video Format	Video Resolution	Recordings' Duration
PUC1 Data	1350	~ 300 GB	mp4	352 x 288	3 weeks
PUC3 Data	550	~ 8 GB	mp4	1920 x 1080	4 weeks

6 SPACE SENSING FOR 3D RECONSTRUCTION

This section contains the information about the space sensing tools and equipment used and developed in Mindspaces, along with the applied techniques for the collection of data for 3D reconstruction of outdoor spaces. A relevant state-of-the-art and commercial systems review is also included together with a description of the data collection mission in Tecla Sala Cultural Centre for the needs of PUC1.

The scope of our research is to provide tools and techniques that capture 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.

According to DoA, the space sensing for 3D reconstruction of Outdoor environment is involved in TO1 (Data Collection) and TA1.3. (Visual data collection) technological objectives and activities of the MindSpaces project. It is expected to contribute to the following Key Results:

KR13. Visual data, point clouds from laser scanner and depth images [TRL 7 to 8]:

MindSpaces will use *industrial tools*[7][8][9] to acquire visual data, point clouds and depth images from interior and exterior environments. MindSpaces foresees to bring an improvement on the use of these technologies and set them up to be deployed in real case scenarios.

KR14. Agile system architecture [TRL 3 to 7]:

Currently there is no product combining all the services foreseen by MindSpaces. The integration of these services in a common platform and testing them in an operational environment in 3 realistic use cases is expected to bring the *technology readiness of the platform close to 6 or 7.*

6.1 State-of-the-art and commercial systems review (SoA-R)

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 outdoors environments. They usually rely on a combination of modern geospatial technologies such as laser scanning, GNSS navigation and photogrammetry. They are usually big structures, mounted on top of cars. Such systems can capture effectively large urban areas, road networks, public infrastructures but they remain expensive due to the top-end hardware components they rely on.

[7] https://www.faro.com/products/construction-bim-cim/faro-focus/

[8] https://structure.io/

9 https://store.stereolabs.com/products/zed/

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 [48]. A recent survey of such applications for lidar based mobile mapping is presented in the work of Wang et al. [49]. 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 a daily consumer of mobile mapping data through online tools or mobile apps like *Google Street View* [50]. 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.

6.1.1 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 [51]. Here we focus on those that are suitable for mapping of large-scale outdoor environments.

Most major players in the geospatial market offer similar systems, like UltraCam Mustang by VEXCEL[10], RIEGL[11], LEICA Pegasus[12] by HEXAGON and VIAMETRIS[13]. 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[14] by imajing, originally designed for trains and now updated for cars is a lower cost vision-based alternative.

[10] https://www.vexcel-imaging.com/ultracam-mustang/

[11] http://www.riegl.com/products/mobile-scanning/

[12] https://leica-geosystems.com/products/mobile-sensor-platforms/capture-platforms

[13] https://www.viametris.com/vms3d

[14] https://imajing.eu/mobile-mapping-technologies/sensors/

6.1.2 Sensors

Most mobile mapping systems combine similar sensors for simultaneous data collection (images and videos, depth, 3D point clouds and 3D motion trajectories). Typically, those systems consist of multiple cameras with lenses, depth sensors, lidar sensors, a GPS/IMU unit for outdoors systems and a tracking sensor for indoor systems. For the development of the custom mobile mapping platform several sensors were considered, based on their specs and price.

Cameras

• e-CAM130_CUXVR - Multiple Camera Board for Jetson AGX Xavier (**https://www.e-consystems.com/nvidia-cameras/jetson-agx-xavier-cameras/four-synchronized-4k-cameras.asp**)

• Leopard Imaging multi-camera systems (**https://leopardimaging.com/product-category/nvidia-jetson-cameras/nvidia-agx-xavier-camera-kits/**)

• Leopard Imaging for Jetson AGX Xavier (**https://www.leopardimaging.com/uploads/LI-XAVIER-KIT-IMX477M12-X_datasheet.pdf**)

• Multi-Camera FMC Module (**http://www.zedboard.org/product/multi-camera-fmc-module**)

• Arduino Panorama Photography with ArduCAM (**http://www.arducam.com/arduino-panorama-photography-arducam/**)

• IVPort V2 Raspberry Pi Camera Module V2 Multiplexer

(**https://ivmech.com/magaza/en/development-modules-c-4/ivport-v2-raspberry-pi-camera**module-v2-multiplexer-p-107)

• Action camera rig by Mapillary (**https://help.mapillary.com/hc/en-us/articles/115001471709-**

Multiple-camera-setups#%E2%80%9Csony-rig%E2%80%9D)

Visual Tracking

• Intel® RealSense™ Tracking Camera T265 (**https://www.intelrealsense.com/tracking-camera**t265/)

GPS/IMU

• XSENSE (MTI-G-700-3A5G4) GPS/IMU (**https://www.xsens.com/products/mti-g-710/**)

• u-blox F9 high precision GNSS module (**https://www.u-blox.com/en/product/zed-f9p**module)

• simpleRTK2B (**https://www.ardusimple.com/simplertk2b/**)

Depth Sensors

• ZED Camera (**https://www.stereolabs.com/**)

• Intel Realsense depth cameras (https://software.intel.com/realsense)

• Azure Kinect DK (https://www.microsoft.com/en-us/p/azure-kinect-dk/8pp5vxmd9nhq)

• Structure Sensor (https://structure.io/)

• Terabee 3Dcam (https://www.terabee.com/shop/3d-tof-cameras/terabee-3dcam/)

• TeraRanger One (https://www.terabee.com/shop/lidar-tof-range-finders/teraranger-one/)

Lidar

• Velodyne Puck (https://www.velodynelidar.com/vlp-16.html)

• Oyster OS-1 Lidar sensor (https://www.ouster.io/product-os1)

• RPLIDAR A3M1 360° Laser Range Scanner (https://www.slamtec.com/en/Lidar/A3)

• LeddarTech LIDAR Modules (https://leddartech.com/lidar/lidar-modules/)

• R-Fans 360 Lidar (http://www.isurestar.com/en/rfan-51.html)

• Laser Bear Honeycomb™ by Waymo (https://waymo.com/lidar/)

• Livox Mid-40/Mid-100 LIDAR sensor(https://www.livoxtech.com/mid-40-and-mid-100)

In section 6.1.2 the sensors that were selected for the implementation of the first version of the mobile mapping platform are described in more detail.

6.1.3 **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 [52] and extensions, such as integration of video from aerial platforms [53].

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 [54].

6.1.4 **Open source tools**

During the last years, several open source projects are active and provide state-of-the-art tools for various tasks such as SLAM, SFM and 3D processing. Below we present a list of the most commonly used tools in the Computer Vision and Robotics communities.

SLAM

• UcoSLAM (http://www.uco.es/investiga/grupos/ava/node/62)

• OrbSLAM (https://github.com/raulmur/ORB_SLAM2)

• Cartographer (https://github.com/cartographer-project/cartographer)

• SVO 2.0: Fast Semi-Direct Visual Odometry for Monocular, Wide Angle, and Multicamera Systems (http://rpg.ifi.uzh.ch/svo2.html)

• LDSO: Direct Sparse Odometry with Loop Closure

(https://vision.in.tum.de/research/vslam/ldso)

• OpenVSLAM: a Versatile Visual SLAM Framework

(https://github.com/xdspacelab/openvslam)

SFM

• AliceVision Photogrammetric Computer Vision Framework (https://alicevision.github.io/)

• Meshroom (https://github.com/alicevision/meshroom)

• COLMAP (https://colmap.github.io/)

• MicMac (https://micmac.ensg.eu/index.php/Accueil)

• OpenSfM (https://github.com/mapillary/OpenSfM)

• Regard3D opensource SfM program (http://www.regard3d.org/)

3D Processing

• The Computational Geometry Algorithms Library CGAL (https://www.cgal.org/)

• OpenMesh A generic and efficient polygon mesh data structure (https://www.openmesh.org/)

• OpenFlipper (https://www.openflipper.org/)

• Open3D A Modern Library for 3D Data Processing (http://www.open3d.org/)

• Geogram A programming library of geometric algorithms (http://alice.loria.fr/software/geogram/doc/html/index.html)

• GEOS - Geometry Engine, Open Source (http://trac.osgeo.org/geos)

• geometry3Sharp (https://github.com/gradientspace/geometry3Sharp)

GPS

• RTKLIB: An Open Source Program Package for GNSS Positioning (http://www.rtklib.com/)

Integration

• rospy (http://wiki.ros.org/rospy)

• IMU_Cam_LiDAR_driver (https://github.com/yangyulin/IMU_Cam_LiDAR_driver)

6.2 Mobile mapping platform development (MMPD)

In this section we describe the development of the mobile mapping platform till M17. Although typically the related task – T3.3 was starting, according to the overall schedule, at M7, the integration and the compilation of the system has started since the beginning of the project by selecting the appropriate equipment and hardware based on the preliminary user and technical requirements, setting and integrating the hardware and testing open source and proprietary in house s/w.

The design and the architecture of the mobile mapping platform was selected to be modular so that it 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.

The mobile mapping platform presented here is not a fixed system, but instead it evolves during the lifespan of the project. A 2nd acquisition platform development phase will take place from M27 to M34 to redevelop, correct and refine the platform according to the feedback from the MindSpaces platform.

In the following sections, details are given on the selected sensors, their integration, the data they collect and the processing of the acquired data.

6.2.1 **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 36, Figure 37) 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.

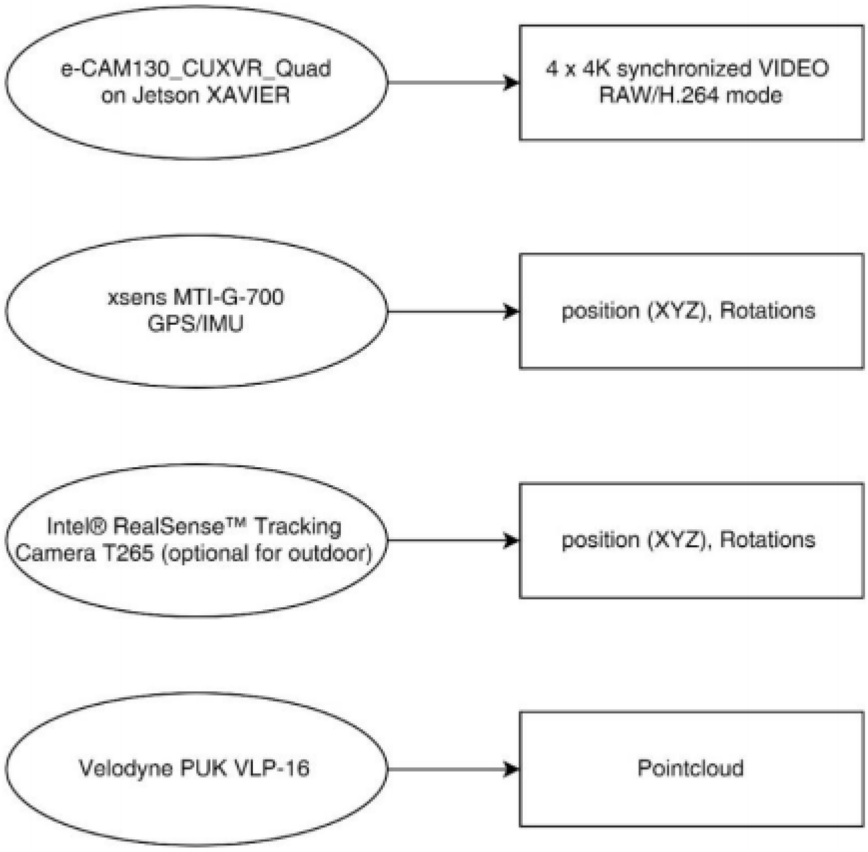

Figure 36. Sensors in the current implementation of the platform for outdoors environments

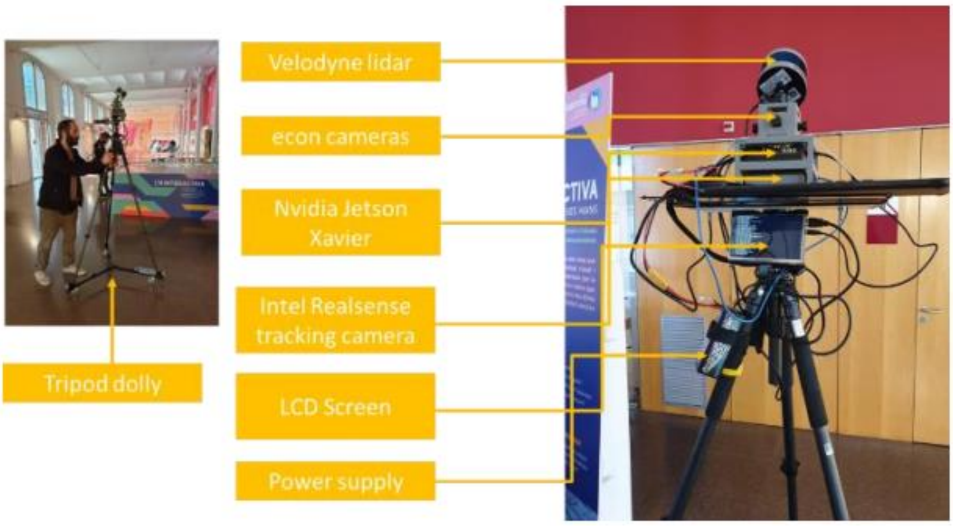

Figure 37. Prototype of the mobile mapping platform

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^{o} x 73^{o} field-of-view (FOV). The four cameras are positioned at 90^{o} configuration facing sideways offering a full 360o horizontal FOV. No upside camera is used in the current setup.

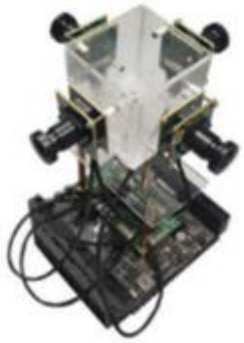

Figure 38. The multi-camera rig

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^{o} horizontal and 30o 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 GPGGA 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^{o} to capture floors and ceilings.

Figure 39. Velodyne® PUC VLP-16 LiDAR sensor

GPS/IMU

For direct Georeferencing in outdoors spaces the platform uses the Xsens® Mti-G-700 GPS/IMU Unit. It is the 4th 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.

Figure 40. Xsens® Mti-G-700 GPS/IMU Unit

Tracking camera

For GPS restricted areas like narrow roads, areas under trees or 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 ~160o 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.

Figure 41. by Intel®, the RealSense™ Tracking Camera T265

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. In an upcoming version of the platform we plan to replace the specific sensor with one compatible with the embedded PC.

Figure 42. by Intel®, the RealSense™ Tracking Camera T265

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.

Mounting

To combine physically all available sensors a prototype base was designed in 3D and then 3D printed (Figure 40). 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 37) and pushed around to perform data collections of relatively small outdoors areas (like squares, individual buildings etc) or interior environments. 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 43. 3D Design of the base supporting the sensors

6.2.2 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.

Space Sensing Module

For the integration of all sensors into a single capturing system the Robotic Operation System (ROS) [55] 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 [56] and underground environments [57]. 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 44) as well as tools to visualize 2D and 3D sensor data ("*rviz*") (Figure 45). 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.

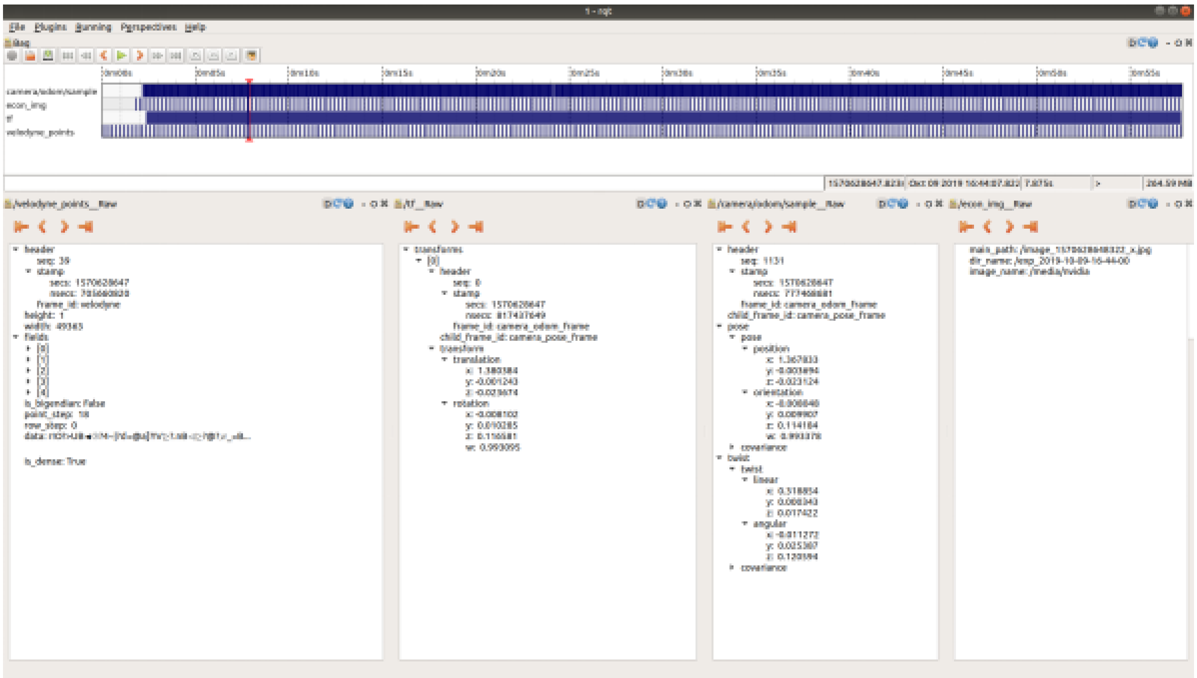

Figure 44. Data from all sensors in ROS can be monitored via “rqt-topic” tool. Data entries are usually accessed through a timeline feature

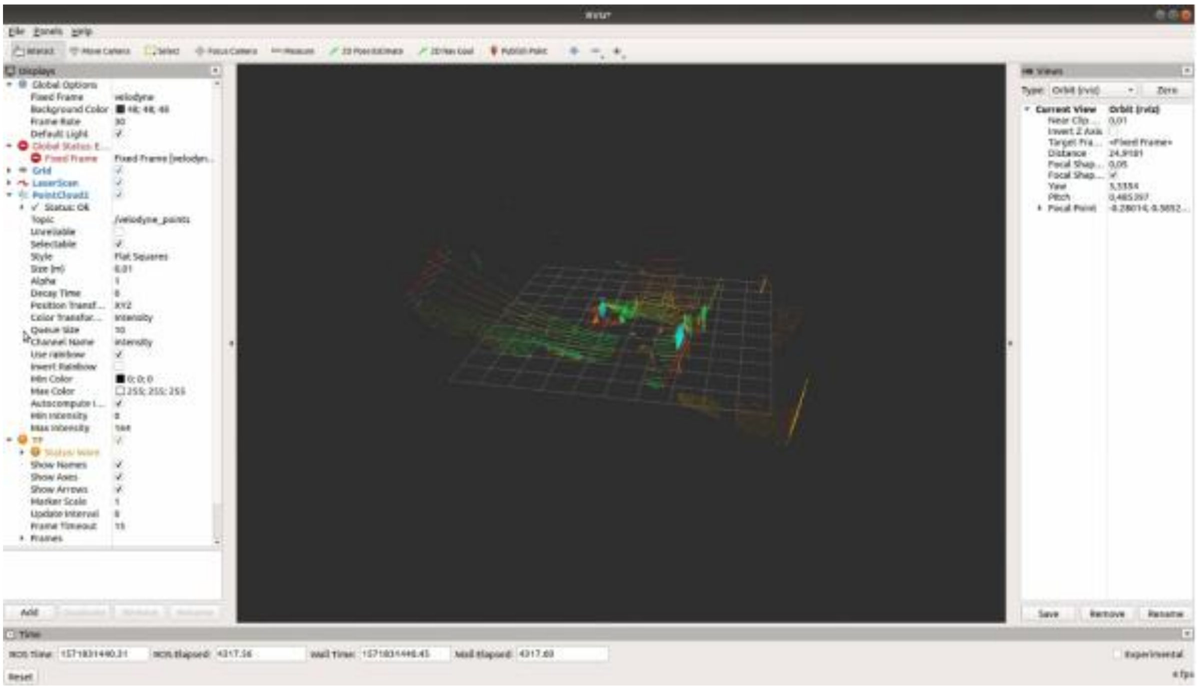

Figure 45. Visualization of the pointcloud topic from the Velodyne VLP-16 Lidar node in rviz ROS tool

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.

For the Velodyne® PUC VLP-16 LiDAR node, the official ROS "*velodyne_driver*"[15] and "*velodyne_pointcloud*"[16] 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*"[17] 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*"[18] 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, due to incompatibility of the GPS/IMU sensor with ARM processors, the system was actually built following a distributed architecture (Figure 46). 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.

15 http://wiki.ros.org/velodyne_driver

16 http://wiki.ros.org/velodyne_pointcloud

17 http://wiki.ros.org/xsens_mti_driver

18 http://wiki.ros.org/realsense2_camera

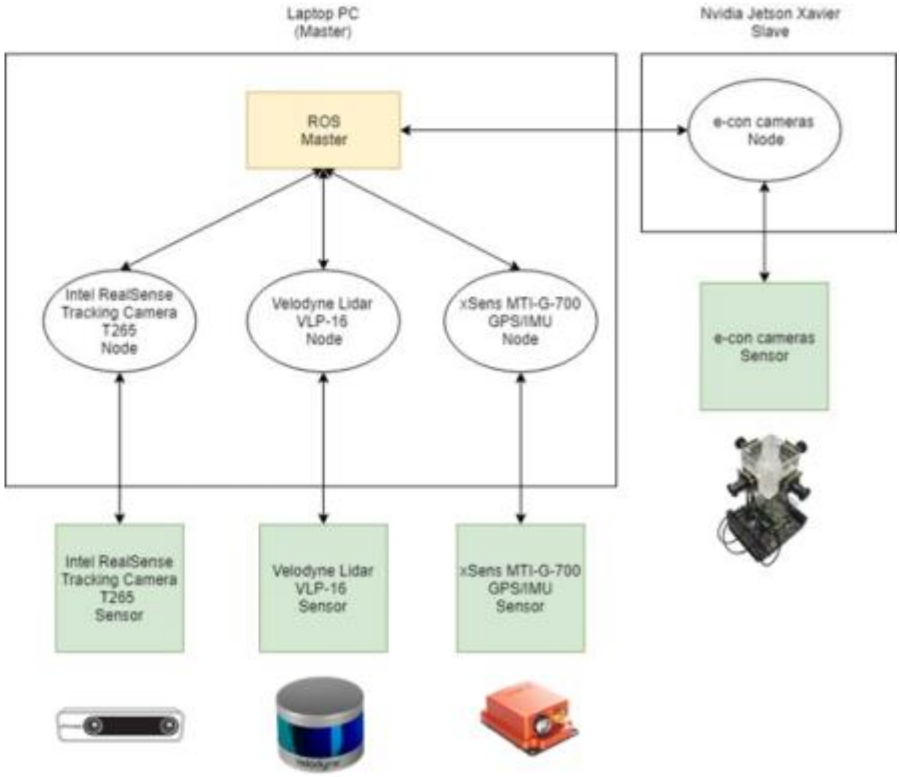

Figure 46. Robotic Operation System (ROS) integration of mobile mapping platform. Distributed architecture

The space sensing module is executed by a script that launches all processes (Figure 47). A basic GUI for touch screens (Figure 48) 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.

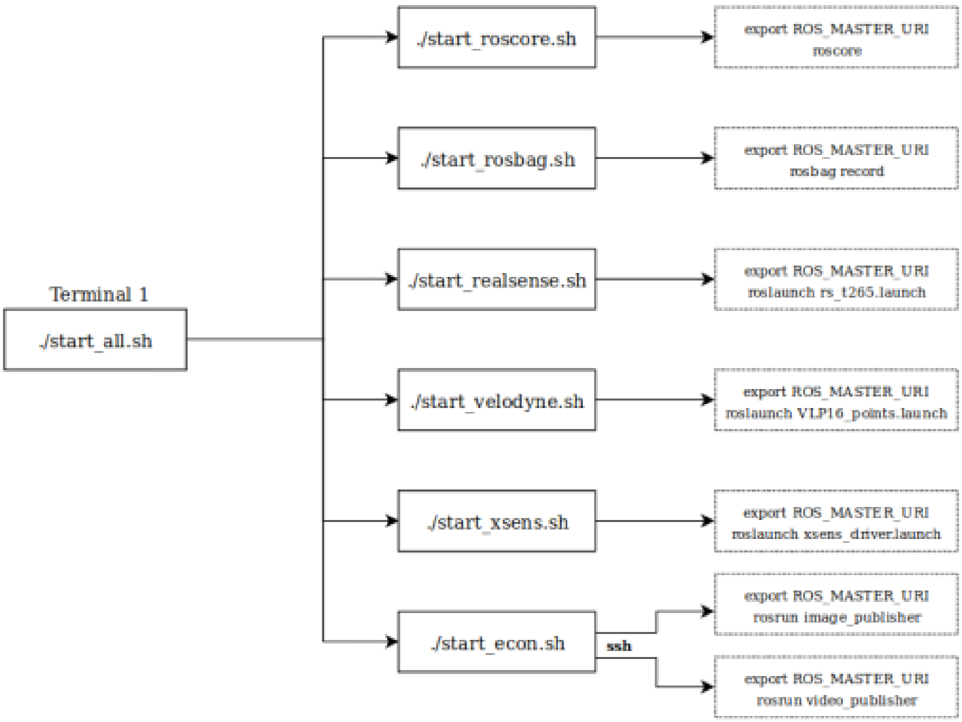

Figure 47. ROS subprograms of the space sensing module

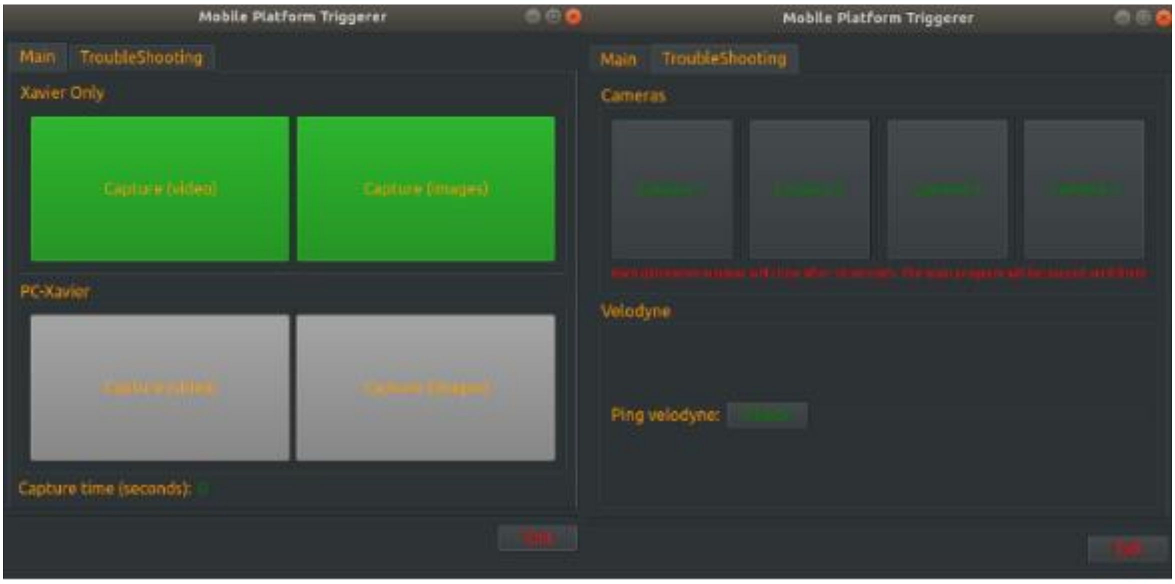

Figure 48. Basic user interface to set capturing parameters, start/end data capture and test/troubleshoot sensors connectivity

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 organized based on their timestamps. For any given time point or period it is possible to retrieve the corresponding data (ie image frames, pointclouds and 6 DoF 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* [58] and *Google Cartographer* [59] for vision-based and lidar-based SLAM respectively, *AliceVision & Meshroom* [60] and [61] for Structure-from-Motion and *Open3D* [62] 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 are placed on a custom designed 3D printed case with known dimensions, good approximations of all sensors relative orientations are a-priori available. The effect of small misalignments is 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.

6.2.3 **SWOT Analysis**

In D8.3 deliverable about Market Analysis and Industrial Requirements a SWOT analysis was carried out for the mobile mapping platform. For completeness, this analysis is also presented here.

Table 6. SWOT analysis

Strengths	Weaknesses
• Easier way to develop 3D environment than traditional tools • Image and point cloud-based • Portability	• Not connected to major CAD softwares, or BIM authoring tools like Revit or AutoCAD • Not online • No automatic plan extraction
Opportunity	**Threats**
• Value for money since based on low-cost equipment • The environment is rapidly growing with new and emerging applications • Rise of Digitization in industry	• Similar platforms developed the last decade and already in market

6.3 **Data collection (DC)**

Under the task T3.3, up2metric is responsible to set up data collection missions for space sensing for 3D reconstruction. This includes the scanning of the outdoors environment with the use of drones, cameras, 3D scanners and a custom-built mobile mapping platform, in order to produce reliable 3D-models. The objective of this service is to record images, video sequences, pointclouds and 3D motion trajectories. The methodology (Figure 49) to achieve this contains:

- Optimal mission planning to ensure the necessary spatial, temporal and radiometric resolution of initial data.
- Horizontal, oblique and vertical imagery
- Employment of a custom mobile mapping platform for the acquisition of images, lidar data and GPS / IMU data

The expected results of this service are synchronized spatio-temporal:

- Data of aerial and ground images, geolocations, point clouds, inertial info
- Data from terrestrial laser scanner with basic photo-texture

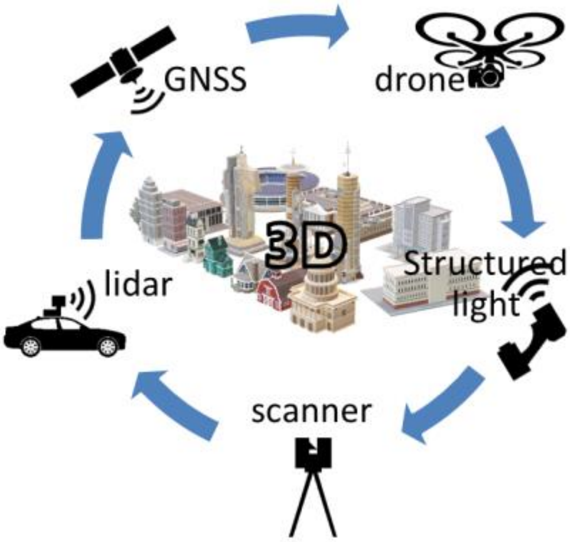

Figure 49. Methodology for Space Sensing for 3D reconstruction

6.3.1 PUC 1 - Outdoors urban environments

The first data collection mission took place at 4-6 November 2019, in the area around the cultural centre of Tecla Sala (Figure 50) which is situated in the City of L' Hospitalet, in Barcelona, Spain. This is the first Pilot Use Case of the Mindspaces Project. This PUC is selected as a case of an urban area of special cultural interest (i.e. city square, old market, riverside, etc.), as well as for the proximity and knowledge of the selected place to project partner(s), Hospitalet de Llobregat City Council and the Espronseda Center for Art & Culture. It is a central area in L'Hospitalet, which holds several projects on contemporary visual arts regarding training, creation, production and exhibition. This area will be the spatial base of work that will host the artists installations and will also serve as a final exhibition space of the result of the work led by the artists.

The task of MindSpaces partner up2metric was to collect all the required data to produce 2D and 3D documentation of the urban environment that architects (ZH, AUTH) and artists (ESP, MoBen, AN) will work on. The initial plan was to use a commercial drone (UAV) to acquire vertical and oblique images of the area of interest. Aerial imagery would be combined with street view images acquired by the developed mobile mapping platform, yielding optimal image orientations as well as a more elaborate modelling of the complicated building and surrounding area (3D mesh of high geometric accuracy and visual quality). Optimal flight missions were designed, and the relevant flight licenses were requested from the responsible authorities. Unfortunately, this plan was not realised because of a general ban of drone flight missions in the specific area.

Figure 50. Tecla Sala cultural centre area overview (image from Google Earth)

Instead, a photocapture mission was planned with a DSLR camera with a fixed interior orientation (Figure 51, Table 7), mounted on a specifically designed telescopic carbon fibre pole, which allowed to take pictures from different heights, from street level up to 5m height, at 1m intervals. Images were taken with high horizontal and vertical overlap (more than 80%). At each camera position 3 images were taken with different orientations, one parallel to the wall of the building and two at ~45 degrees, facing left and right. This step is essential to strengthen the image block geometry but also decrease the risk of occlusions created by the object's complexity (Figure 53). Approximately 6000 images were taken (~40 GB).

It must be noted, that this alternative terrestrial, instead of aerial photocapture, resulted in incomplete data (since it was not possible to capture the roof of the building) and significantly more labour work, since Structure from Motion algorithms typically underperform or even fail in such conditions. This said, the post-process has showed, till now, that for specific paths in VR (hence not for flying above the roof) the results will be more than satisfactory.

Complementary to the photogrammetric mission a 3D survey with a terrestrial 3D laser scanner (Figure 52, Table 7) was also planned and carried out. In order to achieve minimum occlusions and sufficient overlapping for the individual scans, a careful positioning of the laser scanner was applied (Figure 54). This resulted in ~100 individual scans (20GB) of ~3mm resolution (Figure 55, Figure 56).

Figure 51. The Nikon 550 D DSLR camera with a 17mm wide-angle lens was used for the collection of visual data for 3D reconstruction

Figure 52. Faro focus 3D laser scanner used for the 3D survey

Table 7. Laser Scanner and DSLR Camera Specifications

	Laser Scanner Specifications
ERROR	2mm at 50m
NOISE	0.3 mm at 10m distance, regarding objects of 90% reflectance (noise compression mode)
BEAM DIAMETER	3mm at exit
SCAN DENSITY	2mm
	DSLR RGB Camera Specifications
RESOLUTION	18MP (5184 x 3456 pixel)
PIXEL SIZE	4.3µm
FOCAL LENGTH	17 mm

Figure 53. Sample images from the image capture survey. Images were captured at multiple heights and different camera orientations.

Figure 54. Planning of scan stations for the 3D scanning survey by means of a terrestrial laser scanner

Figure 55. 3D scanning in Tecla Sala

Figure 56. Sample pointclouds from two different scan stations

Finally, a survey with the developed mobile mapping platform was also carried out. The platform was mounted on a tripod dolly, which 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 ~15 minutes.

Figure 57. Four synchronized frames from the multi rig camera of the mobile mapping platform

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 57). A dense point cloud was also computed via Multi View Stereo Dense Reconstruction (Figure 58).

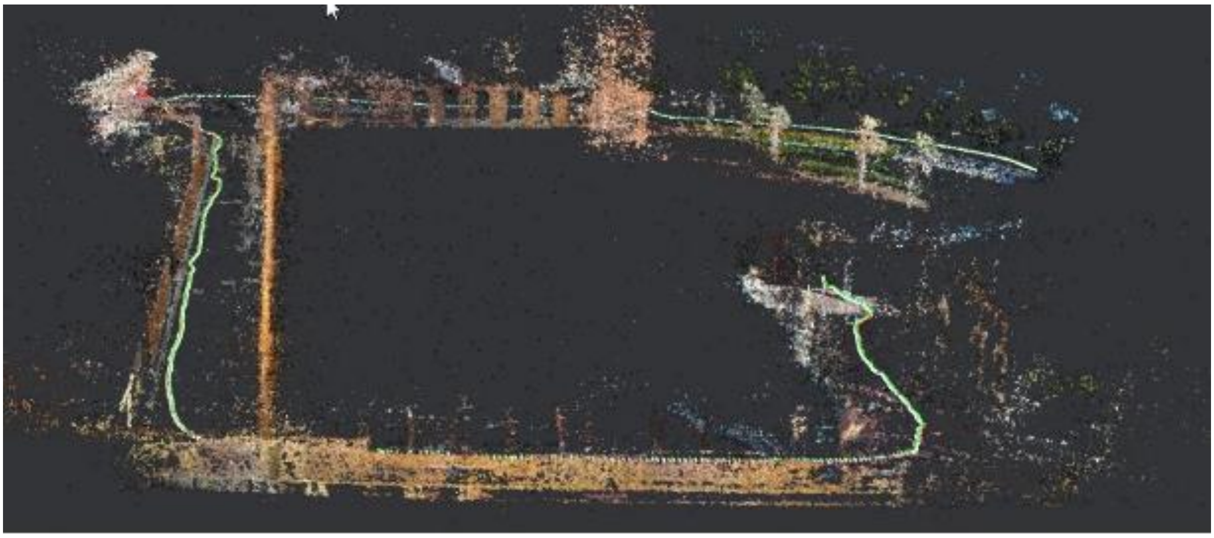

Figure 58. Estimation of mobile mapping platform trajectory by means of SLAM and SfM

Figure 59. Final dense pointcloud from the Tecla Sala 3D survey (SfM)

6.4 Quality Assessment of Data Collection (QADC)

A quality assessment of all acquired data was performed on-field, after each photo-capture, 3D scanning and mobile mapping mission. All image and video data were validated visually to ensure that they are clear, in-focus, sharp and that there is enough overlap between them. A draft 3d model was also created on very low resolution to ensure that data are enough for 3D model completeness. The 3D pointclouds from the 3D scanner were also validated visually and an initial registration was estimated to verify that there are no unmodelled areas. The “bag” files from the mobile mapping platform were also inspected visually, using “rqt-topic” and rviz ROS dedicated tools to ensure that data recorded from all sensors were stored successfully with valid timestamps.

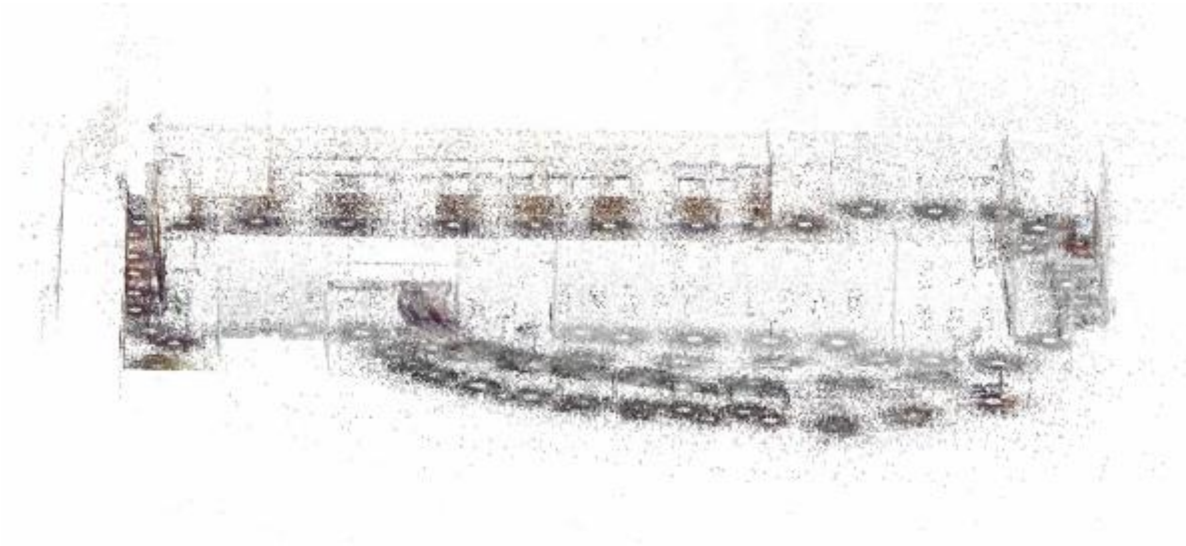

Figure 60. Initial registration of individual pointclouds for quality assessment

The data collection was considered to reach the initial expectations for this stage of the project. Although it was not feasible to collect drone footage for the selected PUC1, a large number of images and 3D pointclouds together with mobile mapping data (lidar and ego-motion camera trajectories) was collected from Tecla Sala cultural centre area. The data is of good quality to allow dense 3D mapping, texturing and object segmentation. It can also be available as raw data to the artists. It must be noted, that during data collection special effort was given to minimize capturing images or pointclouds that included faces or people. All data will be filtered and in case faces or people are included in the collected data, they will automatically be excluded from further analyses.

A second assessment will be based on the final 3D model which will be recontructed via SfM and SLAM alogrithms.

7 INTERIORS SENSING FOR 3D RECONSTRUCTION

This section contains the details about the Interiors sensing for 3D reconstruction. A relevant state-of-the-art and commercial systems review is also included together with a description of the data collection mission in the office space of McNeel and in a senior's apartment in Paris for the needs of PUC2 and PUC3 respectively. Information about the development of the indoor scanning system and a custom built structured light 3D scanner are also given.

According to DoA, the interiors sensing for 3D reconstruction is involved, similar to Space sensing for 3D reconstruction of outdoor spaces, in TO1 (Data Collection) and TA1.3. (Visual data collection) technological objectives and activities of the MindSpaces project. The differences lie at the data capturing platform and at the processing algorithms. It is also expected to contribute similarly to the following Key Results:

KR13. Visual data, point clouds from laser scanner and depth images [TRL 7 to 8]: MindSpaces will use *industrial tools* to acquire visual data, point clouds and depth images from interior and exterior environments. MindSpaces foresees to bring an *improvement* on the use of these technologies and set them up to be *deployed in real case scenarios*.

KR14. Agile system architecture [TRL 3 to 7]: Currently there is no product combining all the services foreseen by MindSpaces. The integration of these services in a common platform and testing them in an operational environment in 3 realistic use cases is expected to bring the *technology readiness of the platform close to 6 or 7*.

7.1 State-of-the-art and commercial systems review (SoA-R)

7.1.1 Mobile Mapping Systems

Lately there is a constantly increasing need for fast, accurate and reliable 3D recording of interior spaces and many companies offer dedicated mobile mapping systems and services to address them. Such systems can capture effectively large buildings and construction sites or small detailed objects. However, most of them are either focused on specific application scenarios, or the most accurate and versatile systems although reliable, they remain expensive due to the high-end proprietary components they rely on. There are platforms with similar technologies like the car-mounted systems that are built on trolleys (NAVVIS[19], Applanix[20]), helmets (Rescan[21]) or backpacks (Leica[22], Viametris[23]). Handheld devices like the

[19] https://www.navvis.com/

20 https://www.applanix.com/products/timms-indoor-mapping.htm

21 https://rescan360.com/

22 https://leica-geosystems.com/products/mobile-sensor-platforms/capture-platforms

23 https://www.viametris.com/

PARACOSM PX-80[24] are also available but their accuracy is not directly comparable to the above systems.

Matterport[25] 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[26] 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.

7.1.2 **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, like indoors environments. 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 [63], [64].

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 [65], 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 [58] and direct methods that exploit all image gradients on the available images [66]. Forster et al. [67] 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. [68] 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 [69], as well as on constraining a SLAM algorithm given a street map background [70].

7.1.3 **Targeted Sectors**

The main industry sectors that can benefit from a mobile mapping platform for indoor environments, thus are potential users and clients for these modules are:

- Real Estate
- Digital Factories
- Surveying and Mapping
- City modelling

24 https://paracosm.io/px-80-overview

25 https://matterport.com/

26 https://www.doxel.ai/

- Geospatial
- Construction (architect firms, construction firms etc.)
- Autonomous vehicles

Some of the needs that emerge in the above sectors and require the 3D modelling of the actual space are:

Building Information Modelling (BIM) to document the physical world, but also the reality that has changed since space construction, so nothing is as planned; the target is to enhance the full life cycle management of an infrastructure.

Virtual Reality (VR) for reality-based training of employees in difficult and dangerous environments, such as industrial spaces, or road infrastructure. This can also assist people for training them for fast response.

Augmented Reality (AR) for assisting the workers in a factory to their everyday tasks, or the visitors in a shopping centre. AR has also started to gain momentum for streaming enhanced information on top of the real city world to citizens and visitors of a city.

Navigation in indoor and outdoor environment.

Fast and informed response under emergency situation requires the accurate knowledge of a space, so that the first responders and the remote handles can move around the space. This is important for designing and implementing successful evacuation plans. This is of increasing importance in places of high people density, such as stadiums and shopping centres.

7.2 Indoor scanning system development (ISS-V1)

In this section we describe the development of the indoor scanning system till M17. This system is based on the mobile mapping platform for outdoors environments described in Section 6.2. By design the platform is modular so that it can be modified to address effectively both outdoors and indoors environments. It is built on the Robotics Operation System (ROS) and is capable to use multiple sensors to capture images, pointclouds and 3D motion trajectories. The version of the platform that targets indoor spaces utilizes synchronized cameras with wide angle lenses, a lidar sensor, a tracking optical sensor and a depth camera, which was not employed in the outdoors version. The GPS/IMU unit is not used since no GPS signal is available in indoor environments. The architecture is also different since all sensors are compatible with the embedded pc and thus no laptop is required. A 2nd acquisition platform development phase will take place from M27 to M34 to redevelop, correct and refine the platform according to the feedback from the MindSpaces platform.

Since both platforms share most of the sensors and the integration as well as the sensing and data processing modules are similar in the following sections, details are given only on the additional sensors and the changes in the architecture of the integration. A detailed description of the overall development of the platform is given in Section 6.2.

7.2.1 Sensors - Components

The current implementation (Figure 61) 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 Intel RealSense T265 Tracking camera for relative positioning in GPS restricted areas (such as indoor scenes) and iv) a Microsoft® Kinect v.2 Depth Camera which captures depth maps and 3D pointclouds.

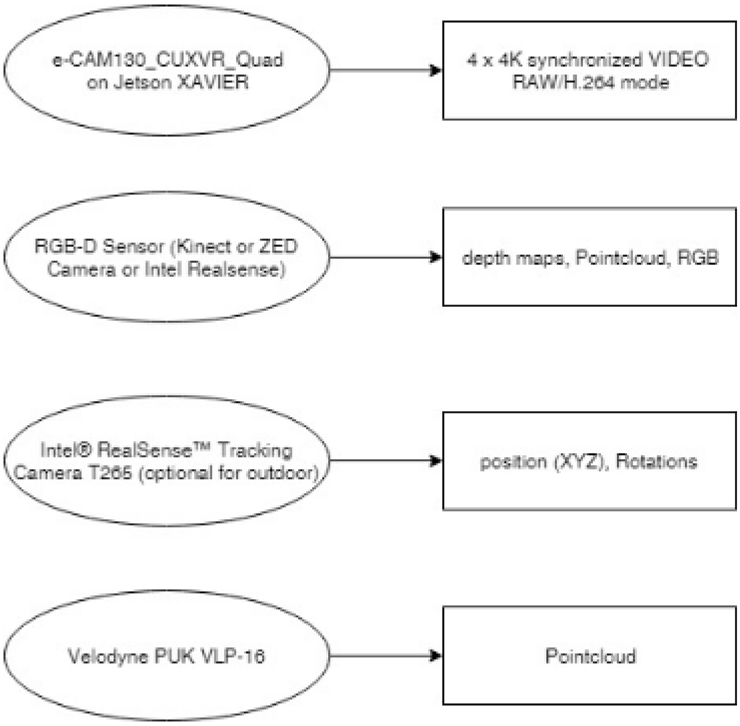

Figure 61. Sensors in the current implementation of the platform for indoors environments

Depth camera

For recording depth information in indoor environments, the platform uses the Microsoft® Kinect v.2 Depth Camera. It is a is a Time-of-Flight (TOF) camera device with a 512 × 424 CMOS IR TOF sensor. It also combines an RGB camera of 1920 x 1080 pixels resolution. The operative measuring range varies from 0.5 to 4.5m. It has a horizontal field of view of 70^o and vertical 60^o. The Microsoft® Kinect v.2 is connected and powered via USB and outputs depth and visual data at 30 fps.

Figure 62. Microsoft® Kinect v.2 Depth Camera

The integration of the Kinect v.2 Depth Camera into the platform is still under development due to instabilities of the dedicated ROS driver and since the specific camera is currently discontinued and not properly supported by Microsoft we consider to replace it, with either the newer version (Kinect Azure DK[27]) or the Intel® RealSense™ Depth Camera D435i[28]. A single machine architecture is adopted for the mobile mapping platform, when GPS/IMU is not used, which is the case in indoors environments (Figure 63).

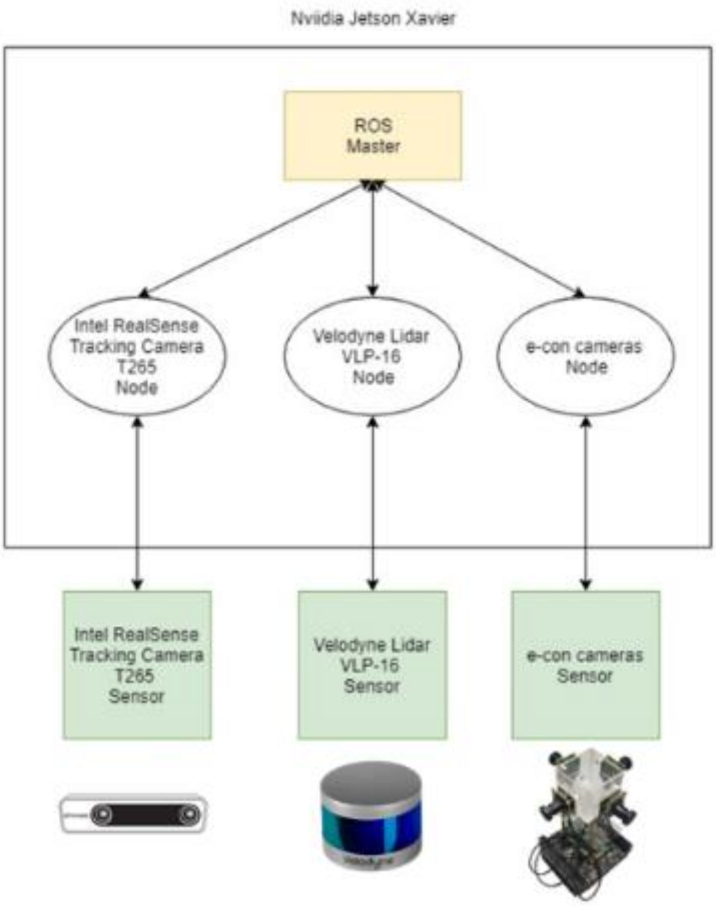

Figure 63. Robotic Operation System (ROS) integration of mobile mapping platform. Single Machine architecture

7.3 Structured light system development (SLS-V1)

To capture in 3D small decorative objects and artefacts, a structured light system was developed by up2metric based on a prototype system design that uses off-the-shelf components [71]. Structured light scanning relies on the projection of different light patterns, by means of a video projector, on 3D object surfaces, which are recorded by one or more digital cameras. Automatic pattern identification on images allows reconstructing the shape of recorded 3D objects via triangulation of the optical rays corresponding to projector and camera pixels.

[27] https://azure.microsoft.com/en-us/services/kinect-dk/

[28] https://www.intelrealsense.com/depth-camera-d435i/

7.3.1 System description

More specifically the system consists of the following hardware components:

• a Canon EOS 400D DSLR camera (resolution 3888×2592) (Figure 64)

• a Mitsubishi XD600 DLP video projector (resolution 1024× 768) (Figure 64)

• a calibration board (white non-reflective planar object with at least 4 black-and-white symmetrical targets printed with a laser printer) (Figure 65).

During scanning, the camera–projector relative position is fixed. The system is flexible regarding its hardware components as it may incorporate any combination of consumer video projector and digital camera. Additionally, it can be adapted to scan objects at different scales by changing the size of the calibration board and the distance between camera and projector (baseline) as well as by suitably adjusting the focus of both devices.

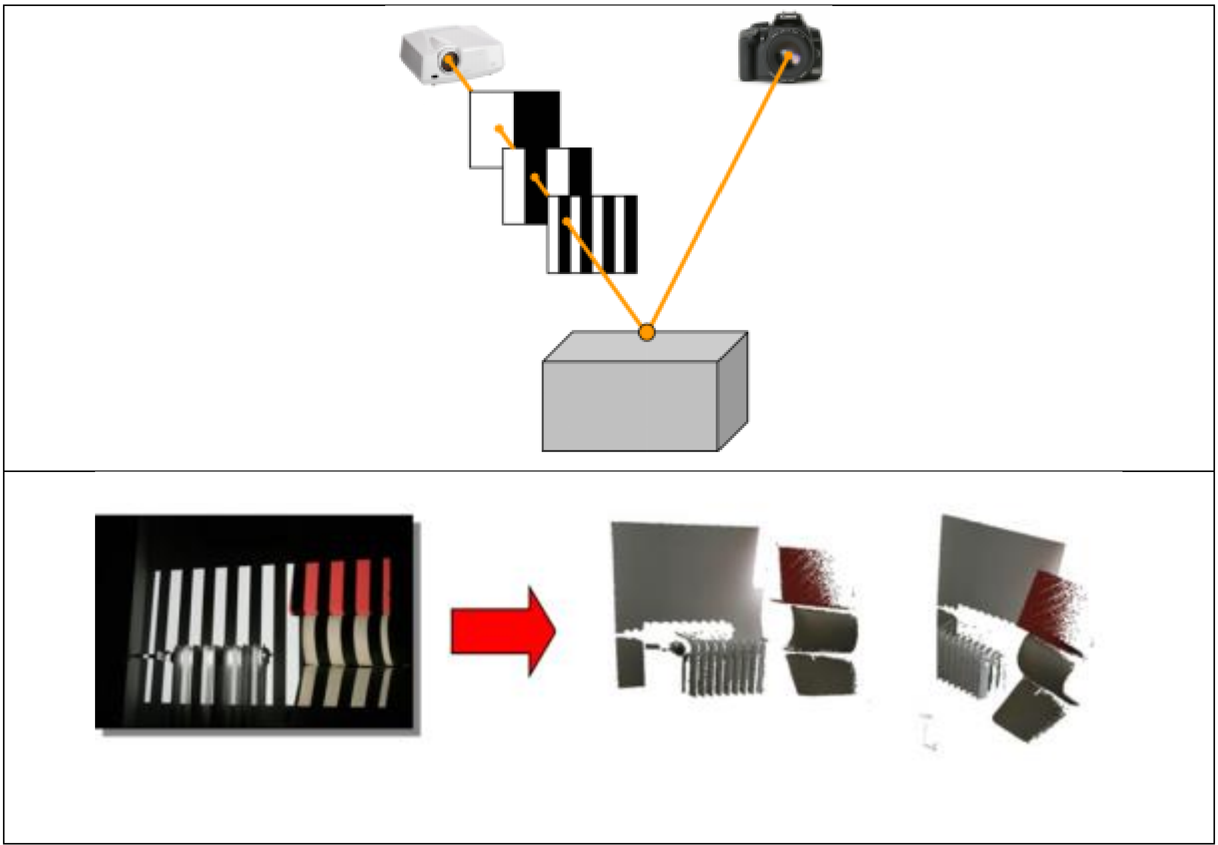

Figure 64. Overview of Structured Light system for Detailed Artifacts

7.3.2 System calibration

Essential step in 3D triangulation with SL systems is their calibration, i.e. the determination of the interior orientation (focal length, principal point position, lens distortion parameters) of video projector and digital camera as well as their (scaled) relative position in space. Typically, a camera-projector calibration is carried out in two separate steps. First, the camera interior orientation is estimated, and next projector interior and relative orientations are found.

Here a simultaneous estimation of camera and projector calibration along with their relative orientation is performed. The implemented algorithm includes:

- The projection of a chessboard-like color pattern (red and white tiles) onto a planar object containing at least 4 (black and white) printed targets, and the recording of these projections by the camera (Figure 65). This is repeated for different successive orientations of the planar surface.
- The automatic detection (with sub-pixel accuracy) of targets and corners on the imaged color pattern.
- A bundle adjustment for optimal estimation of calibration parameters.

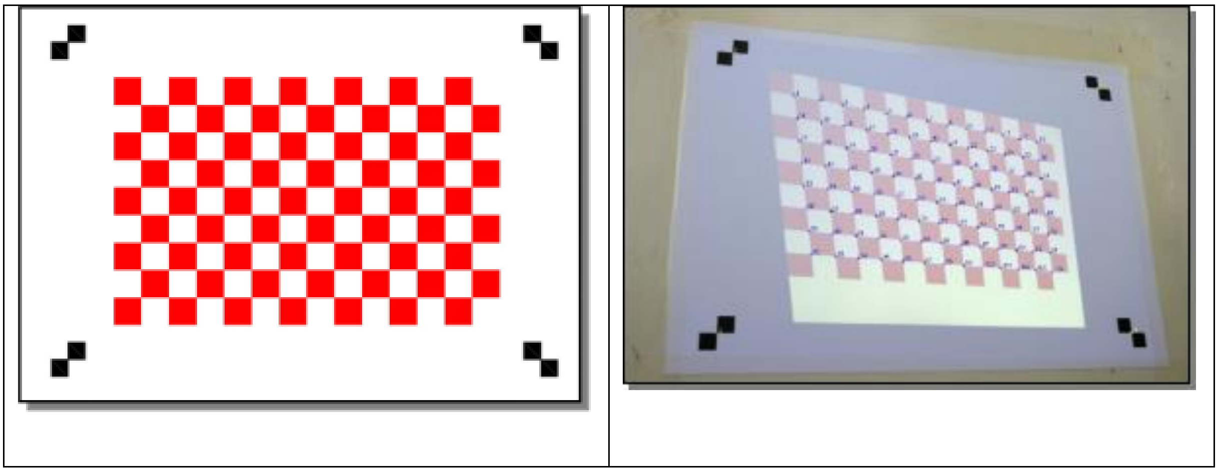

Figure 65. Calibration of camera – video projector system

7.3.3 **Matching camera and projector pixels – triangulation**

Crucial step in SL systems is the establishment of correct correspondences between projector and camera pixels, since the accuracy of this matching procedure affects directly the accuracy of 3D reconstruction. Our approach implemented so far is based on [72] and uses successive projections of binary Gray-code patterns (Figure 66), i.e. black-and-white vertical and horizontal stripes of variable width. Each projection is recorded by the camera, and dark and light areas are identified on the image. Since each projector pixel is characterized by a unique sequence of black and white values, identification of the sequence of dark and light values for each camera pixel directly allows establishing camera–projector pixel homologies. In order to determine whether an image pixel corresponds to a dark or a light projected area in a more robust way, the inverse of each Gray-code pattern is also projected. A pixel is characterized as illuminated with white colour from a specific pattern if the difference of its intensity values corresponding to successive normal and negative patterns is positive. The rest of the pixels are assigned to dark values. Finally, pixels with absolute differences less than a threshold (e.g. 4 gray values) can be rejected as outliers.

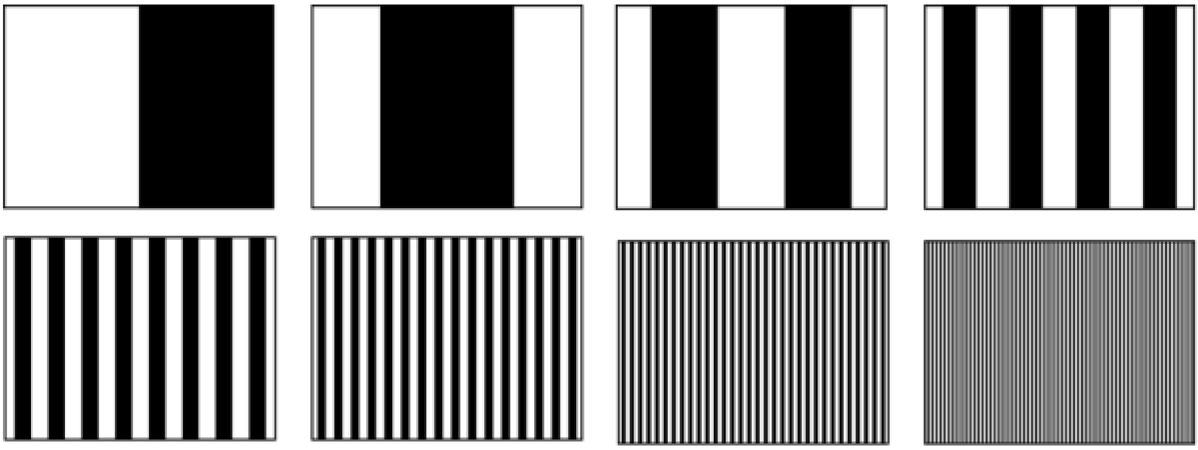

Figure 66. Binary Gray Code

Due to differences in the camera and projector resolutions several camera pixels may be assigned to the same projector pixel. This results in 3D reconstructions with discrete steps and strong artifacts. Thus, to obtain more accurate and smooth 3D reconstructions each camera pixel must be associated with a unique sub-pixel point on the projector frame. After establishing correspondences at pixel level, the integer projector coordinate values are interpolated by means of a 1D averaging filter (for example 7×1 pixels) in the prominent pattern direction. In our implementation 2D orthogonal convolution windows (11×7, 15×7) are used for averaging, to obtain smoother results and consistency among different scan-lines. Once pixel (or sub-pixel) matches are established, the 3D position of depicted object points is computed through simple triangulation of the corresponding optical rays. Multiple individual scans can be registered and a complete 3D mesh model can be obtained from the unified pointcloud (Figure 67).

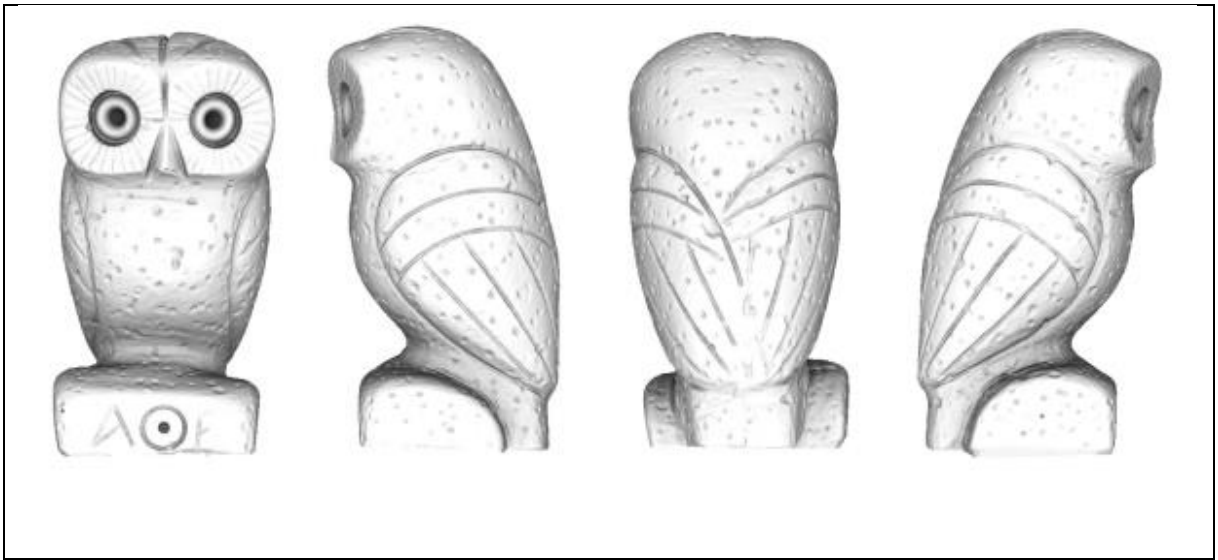

Figure 67. Example of individual pointclouds from different scan positions (up). Final 3D mesh model reconstructed by means of structured light scanning (down)

7.4 Data collection (DC)

Under the task T3.4, up2metric is responsible to set up data collection missions for interiors sensing for 3D reconstruction. This includes the scanning of the indoor environments with the use of terrestrial laser scanners, such as the FARO Focus 3D phase-shift laser scanner, capturing dense overlapping image sets with DSLR cameras, and data collection with the custom-built mobile mapping platform, in order to produce reliable 3D-models. For smaller interior design objects like decorations, small furniture, artwork etc. a custom built structured light 3D scanner is employed. The objective of this service is to record, visual, range, location, and inertial data for reconstructing photorealistic 3D models of indoor scenes.

The expected results of this service are synchronized spatio-temporal:

- Data of images, geolocations, point clouds, inertial info
- Data from terrestrial laser scanner with basic photo-texture
- Structured light data of small artifacts if needed from the artists and architects

7.4.1 PUC 2 - Inspiring workplaces

For PUC2, a data collection mission took place at 4-6 November 2019, in the office facilities of MindSpaces partner McNeel. This workspace will be used as a testing environment for designing friendly, emotionally sensitive, and functional interior workspaces and interior objects. The task of MindSpaces partner up2metric was to collect all the required data to produce a complete and geometrically accurate 3D documentation of the original workspaces, which will be re-designed by architects (ZH, AUTH) and artists (ESP, MoBen, AN). ZH's Agent Based Parametric Semiology life process modeling will be used to simulate and test social behavior within proposed workplace designs.

The data collection included several overlapping laser scans to capture 3D pointclouds of all office spaces. The Faro Focus 3D terrestrial 3D laser scanner was used (Figure 52, Table 7) and a total number of 50 individual 3D scans (Figure 34) were recorded (5GB). A parallel photocapture mission was planned and took place with a Canon 550D DSLR camera (Figure 51, Table 7), in order to provide sufficient visual information for photorealistic texture mapping of the reconstructed 3D models. The camera was mounted on a tripod dolly and was moved around the office space to ensure that all surfaces were photographed. Around 1500 overlapping images (Figure 69) (8GB) were collected.

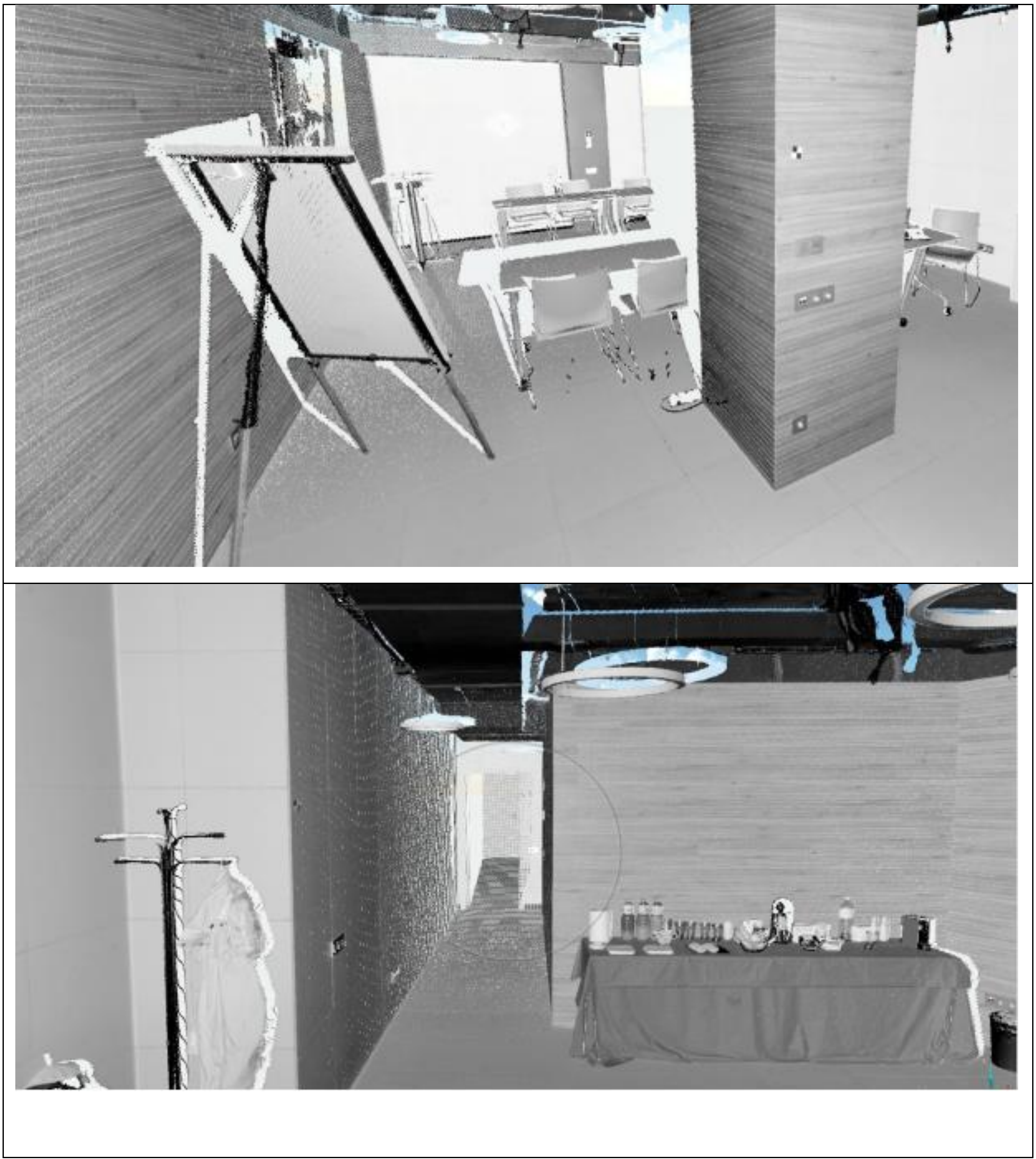

Figure 68. Sample pointclouds from the 3D scanning survey

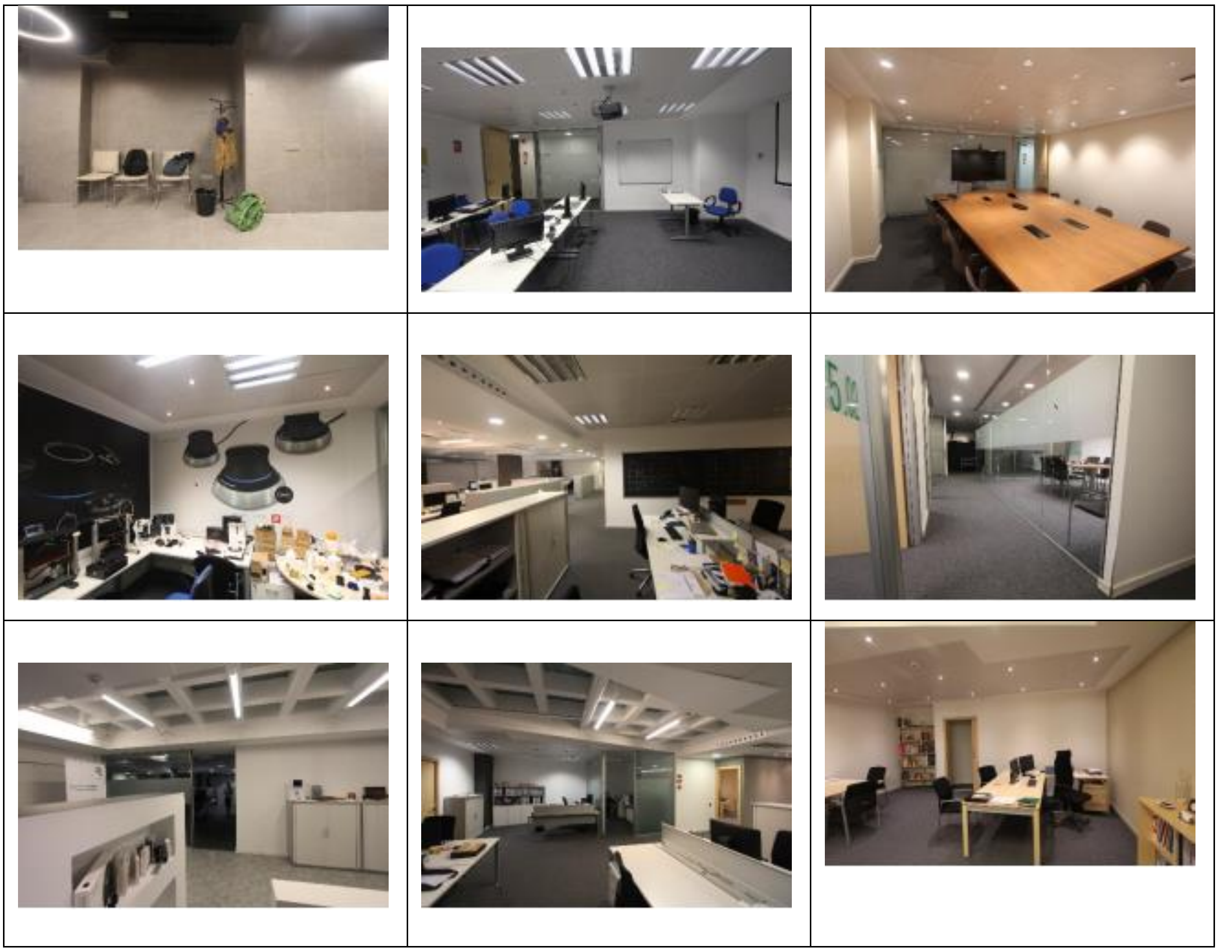

Figure 69. Sample images from the image capture survey

7.4.2 **PUC 3 - Emotionally-sensitive functional interior design**

For PUC3, a data collection mission was carried out on 3 July 2019, in the apartment of a senior person in Paris. The specific residence was selected by MindSpaces partner E-Seniors. The goal was to create a 3D model of the apartment and to capture in 3D, characteristic small objects and furniture of special importance for the senior resident. Then artists, designers and architects will propose redesigns and refurbishments to make domestic environments emotionally and functionally senior-friendly, as well as designing objects and spaces that evoke positive cognitive and emotional experiences and memories, and following design trends and aesthetic values likely to be appreciated by the elderly living there.

The data collection included 20 overlapping laser scans (Figure 70) with the Faro Focus 3D terrestrial 3D laser scanner (Figure 52, Table 7). A parallel survey was carried out in the living room with a first beta version of the mobile mapping platform using only two sensors:

a. the Intel RealSense T265 Tracking camera for relative positioning and
b. the Microsoft® Kinect v.2 Depth Camera to acquire depth images and pointclouds (Figure 72).

A photographic survey of small objects and selected furniture was also performed to provide visual data for 3D reconstruction by means of multi-view and single-image approaches.

Figure 70. 3D scanning of the PUC3 residence

Figure 71. Capturing of small objects

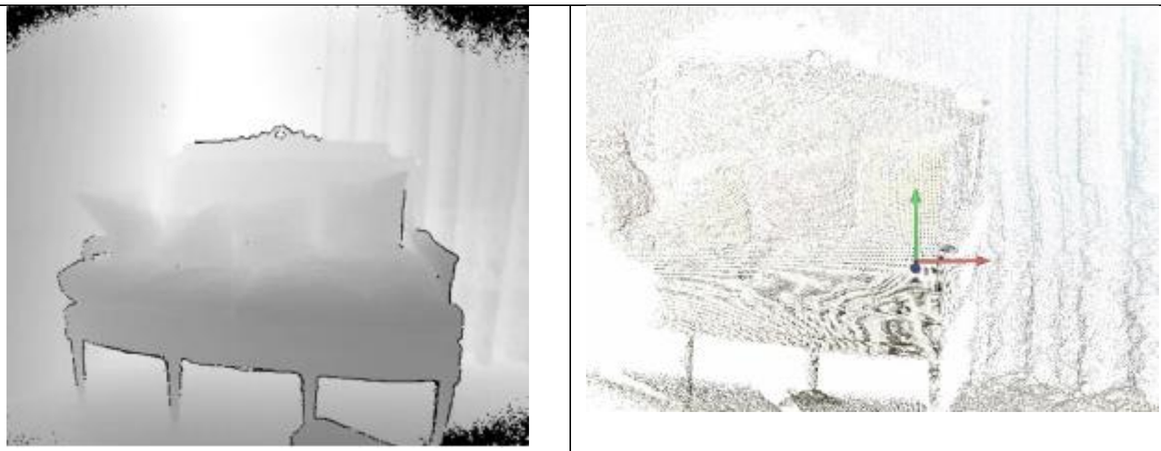

Figure 72. Depth image and respective 3D point cloud from Microsoft Kinect v2 depth sensor

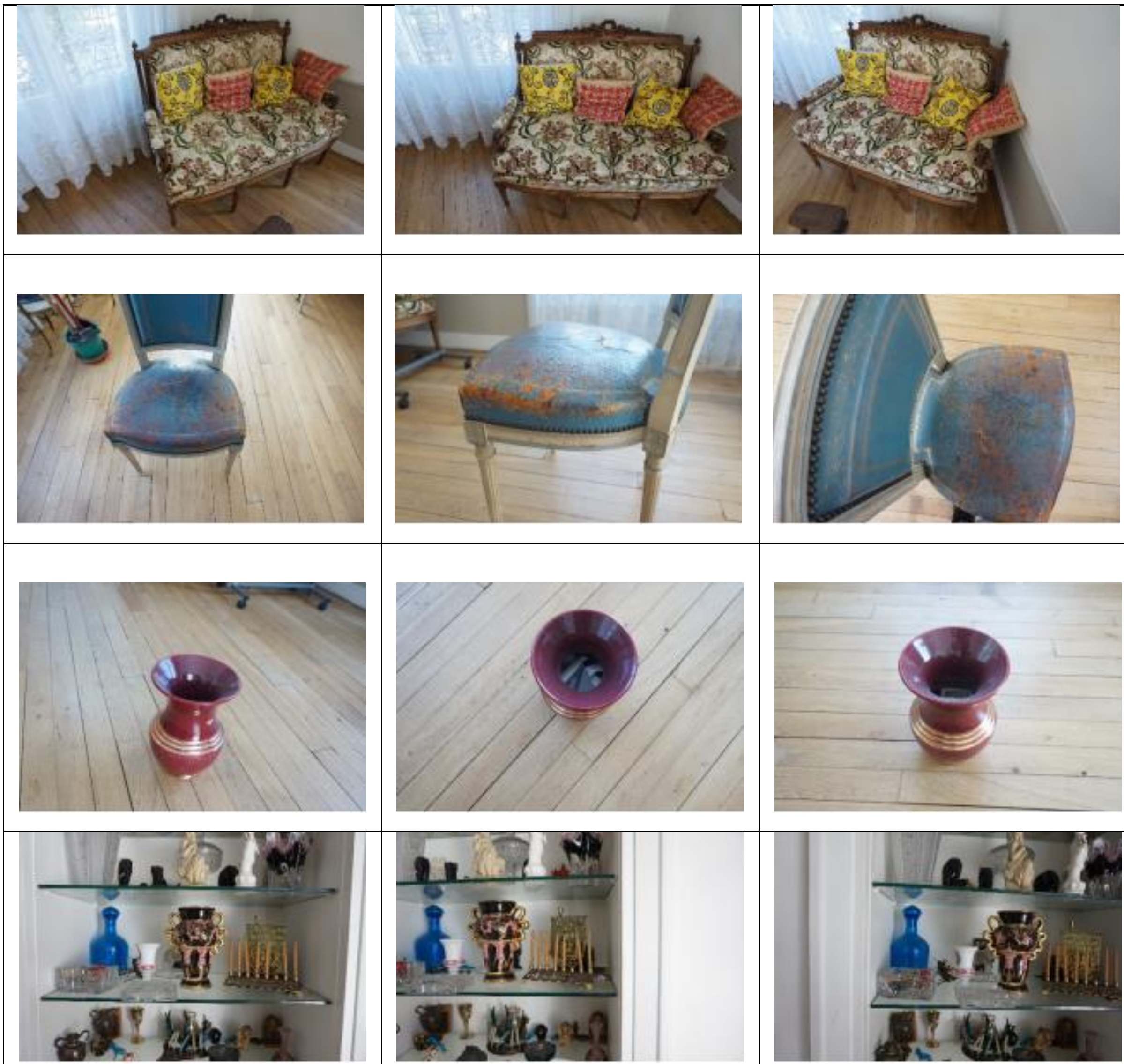

Figure 73. Image samples of furniture and small decorative objects that were captured in PUC3

7.5 Quality Assessment of Data Collection (QADC)

A quality assessment of all acquired data was performed on-field, after each photo-capture, 3D scanning and mobile mapping mission. All image and video data were validated visually to ensure that they are clear, in-focus, sharp and that there is enough overlap between them. A draft 3d model was also created on very low resolution to ensure that data are enough for 3D model completeness. The 3D pointclouds from the 3D scanner were also validated visually and an initial registration was estimated to verify that there are no unmodelled areas. The "bag" files from the mobile mapping platform were also inspected visually, using "rqt-topic" and rviz ROS dedicated tools to ensure that data recorded from all sensors were stored successfully with valid timestamps.

The data collection was considered to reach the initial expectations for this stage of the project. A large number of images and 3D pointclouds was collected from McNeel workplace and from an elder's residency. The data is of good quality to allow dense 3D mapping, texturing and object segmentation. It can also be available as raw data to the artists. It must be noted, that during data collection special effort was given to minimize capturing images or pointclouds that included faces or people. All data will be filtered and in case faces or people are included in the collected data, they will automatically be excluded from further analyses.

A second assessment will be based on the final 3D model which will be recontructed via SfM and SLAM alogrithms.

8 SOCIAL MEDIA AND WEB DATA CRAWLING

Vast amounts of structured and unstructured data are being generated at rapid rates by numerous online entities. Consequently, usability issues arise regarding the storing procedures and management of such data, rendering analysis and indexing of data deriving from social media platforms and websites a truly challenging task. Abundance of information, noise injections, different formats, multiple languages and further issues dictate for an intelligent crawler, specifically developed to address the project and user requirements.

At the moment there are two kinds of sources of data to be crawled for MindSpaces demands: Status updates from a social media platform and web pages. The MindSpaces crawlers extract freely available textual content from open web resources. The tools are based on existing open source crawlers and scrappers to collect content.

For targeted websites, scraping is performed to get only meaningful textual content regarding original posts and user comments where applicable. For social media, the initial platform selected was Twitter in order to extract public tweet posts. Finally, the collected textual content is utilized for text analysis tasks.

8.1 Framework

The scope of the tools developed for the project is to collect useful content which can be then processed by WP4 modules. It is the first component of a pre-processing pipeline that collects content to be utilized by the design tool (WP6) to demonstrate information in a visualized manner. The exact modules that are incorporated in this framework are depicted in Figure 74.

Initially, we define the web entry points and search queries relied on pre-existing online resources. The web entry points are URL addresses of online web domains (7 out of 7 websites whose original posts were extracted whereas 6 out of 7 occasions user comments wele additionally scrapped) whereas the search queries are formed by textual keywords which are passed on to the twitter API from the configuration file created. We provide multiple resource filtering techniques based on: language selection (english, spanish, catalan), duplicate removals, language matching among quotes, retweets and original posts, and merging of text to be analyzed in case of intricate quoted posts and retweets formations. Currently, the extracted raw content is then saved to a mongoDB instance inside the data storage. The MindSpaces crawler development involves the task T3.5.

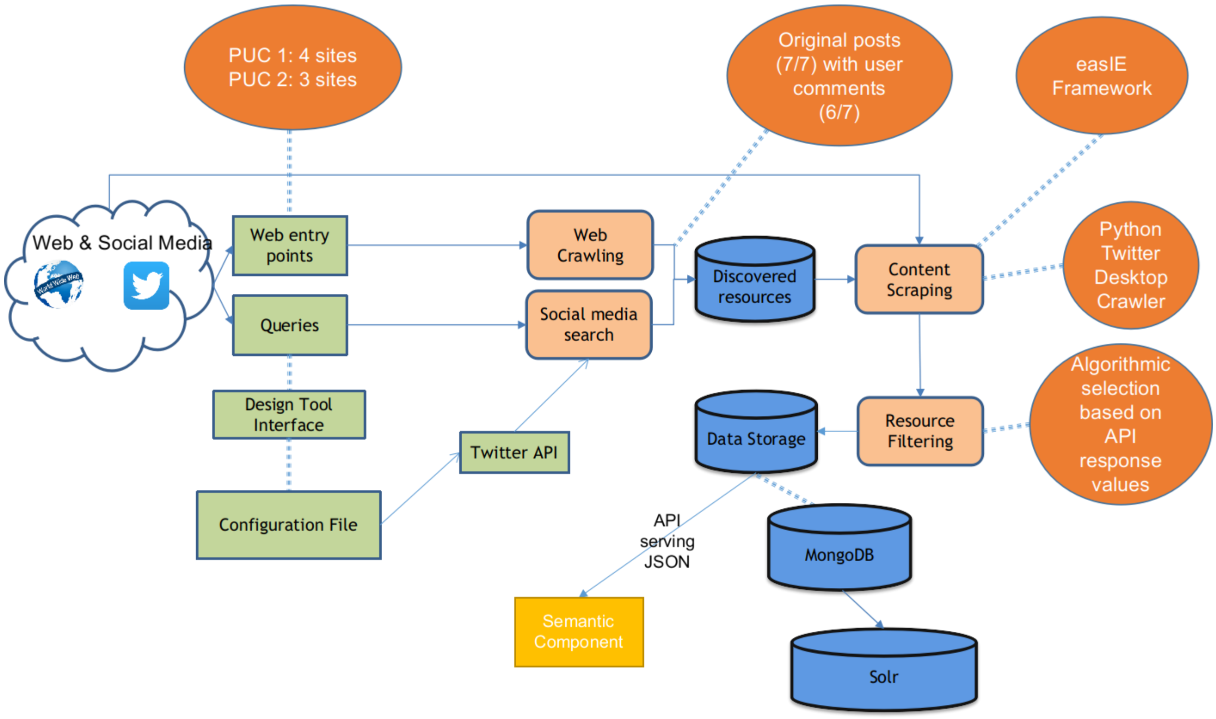

Figure 74. The MindSpaces Crawler

8.2 Information Retrieval

The entire web contains a plethora of available public data regarding the textual nature that can be used to extract sentiments based on aspects of a certain topic. This kind of data can be found in numerous types of web resources such as standard thematic websites and social networking platforms. It facilitates the artist to comprehend general public opinion about certain topics when constructing installations of art.

Towards automating the process of information retrieval, in subsequent subsections the MindSpaces crawling and scrapping components are further elaborated.

8.2.1 Information Retrieval from Social Media

The initial approach is to utilize the twitter crawler for pilot use case 1 only, but keeping in mind that if circumstances arise in case of future needs, it is able to support other use cases and scenarios also if dictated by the artists. The framework was developed to encompass additional functionalities apart from mere keyword searching which to this moment remain unexploited, but are offered in case of a specialized scenario deriving from the artists in future.

The crawler functions in a dynamic manner, being an always online service. The scrapping is performed in a round robin fashion where if one collection has been completed for the time being, it is paused and the crawler continues scrapping the following collection, and generally performs such recursive cycles among collections in an iterative never-ending approach unless interrupted by the user. It is also worth mentioning that as one API twitter key is provided at the moment, there is a possibility to reach maximum content bandwidth where again the crawler automatically pauses for as long as the Twitter API mandates.

PUC 1

For pilot use case 1, CDH after an extensive research on artistic topics about culture and art in urban environments assembled a list of collection of keywords in both Spanish and Catalan. They were fed into the configuration file of the crawler, maintaining the grouped structure when fetching raw data via the twitter API and storing it inside a NoSQL mongoDB instance in the data storage, and are shown in Tables 8 and 9:

Table 8. Collection of keywords in Catalan

Main topics	Keywords
Ordenació de la ciutat	Accessibilitat, arquitectura, ciutat, entorn urbà, espai urbà, estètica, industrial, medi ambient, mobilitat, sostenibilitat, Tecla Sala, urbanisme, modernitat, territori
Art	Artista, arts visuals, creació artística, obra d'art, patrimoni, emergent, art urbà
Impacte social	Benefici social, comportament social, desenvolupament cultural, innovació social, interacció social, participació,
Sensacions	Benestar, inspiració, sentiments, acústic, experimentació,
Llenguatges	Ciència, coneixement, creativitat, cultura, LHCultura, digital, difusió, disseny, exposició, indústria creativa, innovació, instal·lació artística, investigació, neurociència, realitat virtual, recerca, tecnologia
Persones	Ciutadà, ciutadania, col·lectiu vulnerable, comunitat, dona, gènere, gent gran, joves, públic
Convivència	Cohesió social, identitat, integració, multicultural, economia, pertinença, socialització

Table 9. Collection of keywords in Spanish

Main topics	Keywords
Ordenación de la ciudad	Accesibilidad, arquitectura, ciudad, entorno urbano, espacio urbano, estética, industrial, medio ambiente, movilidad, sostenibilidad, Tecla Sala, urbanismo, modernidad, territorio

Arte	Artista, artes visuales, creación artística, obra de arte, patrimonio, emergente, arte urbano
Impacto social	Beneficio social, comportamiento social, desarrollo cultural, innovación social, interacción social, participación
Sensaciones	Bienestar, inspiración, sentimientos, acústico, experimentación
Lenguajes	Ciencia, conocimiento, creatividad, cultura, LHCultura, digital, difusión, diseño, exposición, industria creativa, innovación, instalación artística, investigación, neurociencia, realidad virtual, tecnología
Personas	Ciudadano, ciudadanía, colectivo vulnerable, comunidad, mujer, género, gente mayor, jóvenes, público
Convivencia	Cohesión social, identidad, integración, multicultural, economía, pertenencia, socialización

8.2.2 Information Retrieval from Websites

There were plenty of options regarding the appropriate selection of Websites. In order to be able to extract information from them though, we targeted at choosing well-structured constructions and that is because the semi-automated process of scrapping demanded such to function properly and aggregate data, thus avoided freely or anarchistly manufactured websites. We further developed the eaSIE framework to address targeted needs, an open-source tool; creation of CERTH. We mainly focused on scrapping original posts from websites along with user comments in a dynamic process referred as deep crawling, resulting in receiving every public available content through iterative visits in each web page of a website. Deep crawling is achieved by inserting the suitable CSS selectors, about the central page table, the next item and the next page among others, inside the configuration file of the framework.

PUC 1

For pilot use case 1, the selection of relevant websites was conducted by CDH where successfully original posts were crawled along with user comments in every occasion and are shown below:

1) https://elfar.cat/

It is a digital archive of a monthly newspaper circulated in the broader area of L'Hospitalet whose topics include a wide variety of interests for the citizens of that area.

2) https://www.llobregatdigital.cat/

It is a local portal from L'Hospitalet with news, interviews and opinions about culture, economy, society, health and politics.

3) **http://www.estrellalh.com/**

Estrella de L'Hospitalet is a proximity means of communication centered in the city of L'Hospitalet de Llobregat (Barcelona). It is an initiative of the association Foment de la Informació Crítica (FIC-LH), with the aim of promoting knowledge and exchange of information among the citizens of L'Hospitalet.The staff of the newsroom consists of journalists and non-journalists who for now, they work on the project in their spare time

4) **http://localmundial.blogspot.com/**

It is a personal blogspot from Manuel Domínguez that presents topics about history, society, politics and economy for the region of L'Hospitalet.

PUC 2

For pilot use case 2, the selection of relevant websites was conducted by ZHA where successfully original posts were crawled along with user comments in every occasion but one (**https://www.archdaily.com/**) and are shown below:

1) **https://www.dezeen.com/**

Dezeen is the world's most popular and influential architecture and design magazine, and the winner of numerous awards for journalism and publishing. Their mission is to bring to the audience a carefully edited selection of the best architecture, design and interiors projects and news from around the world.

2) **https://www.danieldavis.com/**

This website is a researcher's personal point of views with interests in two main themes: how technology influences architecture, and how architecture influences people.

3) **https://www.archdaily.com/**

Arcdaily.com began as a platform to collect and spread the most important information for architects seeking to build a better world. Today, it is an ever-evolving tool for anybody who has a passion and determination to shape the world around them.

8.3 Datasets Created

Inside this subsection, the datasets which were created by the utilization of the MindSpaces Crawling frameworks are presented. Every collected resource serves the data collection needs as were defined from the beginning of the project to this reporting period. All aggregated datasets reside inside files, for the moment, inside an ftp server at the disposal of CERTH and addressed the immediate need for resources for the textual analysis algorithms development.

8.3.1 Twitter Dataset

The Twitter dataset creation for pilot use case 1 was based on the collection of keywords described thoroughly in subsection 8.2.1. Apart from sole collections of keywords, we also tested and applied several geolocation restrictions (20 Km bounding box, 4Km radius) around the broader area of tecla sala de l' Hospitalet resulting in additional datasets. Unfortunately, we did not investigate further the geolocating potentialities due to the geolocation API announcement about deprecation and withdrawal from the general Twitter API. The fundamental twitter crawling services were active for an estimated duration of more than a week and resulting entities were stored following the same JSON format as was fed by the Twitter API. An exemplary response of the twitter API without expanded indentations is shown below:


```
{  "created_at":"Wed Jul 03 13:13:19 +0000 2019",
  "id":1146406567203528704,
  "id_str":"1146406567203528704",
  "full_text":"Starwood sella la compra de cinco edificios de oficinas, en Madrid y Barcelona,\u00a0a Oaktree por 153 millones https://t.co/KlUOB1kg2J @Ejeprime #oficinad",
  "truncated":false,
  "display_text_range":[],
  "entities":{},
  "metadata":{},
  "source":"<a href=\"http://twitter.com/download/android\" rel=\"nofollow\">Twitter for Android</a>",
  "in_reply_to_status_id":null,
  "in_reply_to_status_id_str":null,
  "in_reply_to_user_id":null,
  "in_reply_to_user_id_str":null,
  "in_reply_to_screen_name":null,
  "user":{},
  "geo":null,
  "coordinates":null,
  "place":{
},
  "contributors":null,
  "is_quote_status":false,
  "retweet_count":0,
  "favorite_count":0,
  "favorited":false,
  "retweeted":false,
```

"possibly_sensitive":false,
"lang":"es"}

8.3.2 Website Dataset

The scrapped website datasets creation for pilot use cases 1 and 2 was based on scrapping textual content from pre-defined provided websites by the user group. Intentions were to farm general opinions about relevant topics in original posts along with user feedback as provided in textual user comments, where applicable, abiding also by the anonymization principles and procedures. The dataset format followed that of a JSON object, same for both pilot uses cases and below an exemplary structure is presented:

{
"website":[
{
"article_title":"El municipio estrena el mural contra el estigma en salud mental"
}
],
"metrics":[
[
{
"name":"text_article",
"citeyear":2019,
"source":"https://elfar.cat/art/28405/el-municipio-estrena-el-mural-contra-el-estigma-en-salud-mental",
"type":"textual",
"value":"Un total de 32 artistas de toda Europa han elaborado, de manera conjunta con los usuarios de Benito Menni CASM, un mural contra el estigma en salud mental, que decorará la pared exterior del centro de Hermanas Hospitalarias de Sant Boi. Se trata de la segunda edición del proyecto ARTabbo, que tiene el objetivo de luchar contra los estereotipos que todavía sufren las personas con enfermedades mentales, a través de la población joven, que es aquella que puede contribuir a modificar estos prejuicios. Así, chicos y chicas de Rumanía, Holanda, Suecia y Chipre han convivido, del 1 al 9 de julio, con usuarios del Hospital Benito Menni con una enfermedad mental o discapacidad intelectual, a fin de confeccionar la creación artística. Esta iniciativa es una acción de educación no-formal, organizada por el Departamento de Juventud y Cultura del Ayuntamiento de Sant Boi de Llobregat y por Benito Menni CASM, y financiada por Erasmus+ y Acción Clave 1- Intercambios juveniles. La obra, que se llevó a cabo el pasado fin de semana, se hizo pública ayer frente al Hospital Benito Menni, en un acto que contó con la presencia de la alcaldesa de Sant Boi, Lluïsa Moret, entre otras autoridades municipales, y con la superiora y el director gerente de Benito Menni CASM, Sor Teresa Íñiguez y el doctor Joan Orrit, respectivamente. La pintura imita el estilo y los motivos que aparecen en las baldosas de las casas de los diferentes países participantes en ARTaboo, con la intención de mostrar el recinto psiquiátrico como un hogar y no como un lugar estigmatizante. Grupos de personas residentes han participado en la iniciativa conviviendo con los grupos de artistas y aportando ideas al proyecto artístico. La actuación se ha hecho en el muro de la calle Benito Menni, la calle peatonal que separa las instalaciones de Benito Menni, Complejo Asistencial en Salud Mental y las del Parc Sanitari Sant

Joan de Déu. El año pasado ya se pintó justo enfrente, durante la primera edición del proyecto, el muro de este otro centro hospitalario. Además de la elaboración del mural, durante su estancia en Sant Boi, los artistas han podido disfrutar de un programa de actividades, la mayoría al aire libre, acompañados de los usuarios y colaboradores de Benito Menni CASM, así como del resto del vecindario. "
        },
        {
          "name":"comment",
          "citeyear":2019,
          "source":"https://elfar.cat/art/28405/el-municipio-estrena-el-mural-contra-el-estigma-en-salud-mental",
          "type":"other",
          "value":[
            "Estic d'acord amb el predecessor",
            "Si me parece ,genial la idea del mural las personas con enfermedad mental es incomprendida .. Un saludo.."
          ]
        }
      ]
   ]
}

8.3.3 Summary

Apart from the qualitative description of the datasets formed during the crawling procedure, in the following table quantitative metrics are presented, such as volume of compressed data in *.zip files and exact dates of crawling and/or scrapping.

	PUC 1 Websites	PUC 2 Websites	PUC 1 Twitter
Total Volume in MBs	26,22 MBs	146,8 MBs	19.400 MBs
Dates	9/9/2019	24/2/2020	25/6/2019 - 3/7/2019

8.4 Conclusion

In this chapter, we have presented, in detail, the initial methods with which the crawling services have collected, accumulated and created the datasets, vital for the development of the textual analysis algorithms. As the first pre-processing component in a complex pipeline it deemed important to have had the initial datasets ready as soon as possible for the entire work flow to keep evolving rapidly.

The Twitter crawling tool incorporates additional unused functionalities, depending in various Twitter APIs except for the "search" API, that may prove to be useful in next iterations. The modifications of the tool occurred in pair with UPF to better satisfy the needs of the entire pipeline. The full live integrated pipeline necessitated for the crawler to be an always online service, wrapped with gRPC (google remote procedure call) protocol in order to dispatch data in a robust synchronous manner to the text analysis service. Further datasets will be created for the pilots and stored in a mongoDB instance, indexed by a SoLR instance, all located inside the data storage and finally visualized by the design tool on demand by the user.

Regarding future steps to expand and enrich our modules we intend to investigate on how to crawl efficiently and automatically further resources, both social media platforms and websites provided by each user per each PUC, in a meaningful way towards the creation of targeted datasets to fulfil the needs of the project properly.

9 CONCLUSION

This deliverable provides an extensive insight into the past 17 months of work conducted by the MindSpaces partners in WP3 of the project. It incorporates the feedback and the outcome from discussions between the involved partners. The description of the high-level methodology and its application to the MindSpaces project is presented. The processing techniques, the sensing equipment, the different types of data and the context of data collection activities per PUC are also included.

Over the next months, the data collection process will continue having as a goal to enrich the collection of data for each use case. The new data will help us to train and improve the performance of the MindSpaces models and also gain new insights from subjects regarding their emotional and behavioural state while they experience different designs of spaces and art installations.

References

[1] Azevedo, Ana Isabel Rojão Lourenço, and Manuel Filipe Santos. 2008. “KDD, SEMMA and CRISP-DM: A Parallel Overview.” IADS-DM

[2] Kurgan, Lukasz A., and Petr Musilek. 2006. “A Survey of Knowledge Discovery and Data Mining Process Models.” The Knowledge Engineering Review 21 (1): 1–24

[3] Wirth, R., & Hipp, J. (2000, April). CRISP-DM: Towards a standard process model for data mining. In Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining (pp. 29-39). London, UK: Springer-Verlag.

[4] “**https://www.emotiv.com/epoc/**.” .

[5] “**https://www.emotiv.com/insight/**.” .

[6] “**https://choosemuse.com/muse-2/**.” .

[7] “**https://www.neuroelectrics.com/solutions/enobio**.” .

[8] “**https://www.advancedbrainmonitoring.com/products/b-alert-x-series**.”

[9] “**https://neulog.com/gsr/**.” .

[10] “**https://www.shimmersensing.com/products/shimmer3-wireless-gsr-sensor**.” .

[11] “**https://github.com/labstreaminglayer/App-LabRecorder**.” .

[12] “**https://github.com/sccn/labstreaminglayer**.” .

[13] “**https://github.com/sccn/xdf**.” .

[14] M.-K. Kim, M. Kim, E. Oh, and S.-P. Kim, “A review on the computational methods for emotional state estimation from the human EEG.,” Comput. Math. Methods Med., vol. 2013, p. 573734, 2013.

[15] S. Wang and Q. Ji, “Video affective content analysis: a survey of state-of-the-art methods,” IEEE Trans. Affect. Comput., vol. 6, no. 4, pp. 1–1, 2015.

[16] J. A. Russell, “A circumplex model of affect,” J. Pers. Soc. Psychol., 1980.

[17] “**https://www.allaboutux.org/self-assessment-scale-sam**.” .

[18] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, “International Affective Picture System (IAPS): Instruction Manual and Affective Ratings,” Tech. Rep. A6, 2005.

[19] E. S. Dan-Glauser and K. R. Scherer, “The Geneva affective picture database (GAPED): A new 730-picture database focusing on valence and normative significance,” Behav. Res. Methods, 2011.

[20] A. Marchewka, Ł. Żurawski, K. Jednoróg, and A. Grabowska, “The Nencki Affective Picture System (NAPS): Introduction to a novel, standardized, wide-range, high-quality, realistic picture database,” Behav. Res. Methods, 2014.

[21] R. A. Stevenson and T. W. James, “Affective auditory stimuli: Characterization of the International Affective Digitized Sounds (IADS) by discrete emotional categories,” Behav. Res. Methods, 2008.

[22] C. E. Parsons, K. S. Young, M. G. Craske, A. L. Stein, and M. L. Kringelbach, “Introducing the oxford vocal (OxVoc) sounds database: A validated set of non-acted affective sounds from human infants, adults, and domestic animals,” Front. Psychol., 2014.

[23] Y. Baveye, E. Dellandréa, C. Chamaret, and L. Chen, “LIRIS-ACCEDE: A video database for affective content analysis,” IEEE Trans. Affect. Comput., 2015.

[24] S. Koelstra et al., “DEAP: A Database for Emotion Analysis ;Using Physiological Signals,” IEEE Trans. Affect. Comput., vol. 3, no. 1, pp. 18–31, 2012.

[25] S. Ergan, Z. Shi, and X. Yu, “Towards quantifying human experience in the built environment: A crowdsourcing based experiment to identify influential architectural design features,” J. Build. Eng., 2018.

[26] J. Marín-Morales et al., “Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors,” Sci. Rep., vol. 8, no. 1, pp. 1–15, 2018.

[27] W. Tantanatewin and V. Inkarojrit, “The influence of emotional response to interior color on restaurant entry decision,” Int. J. Hosp. Manag., vol. 69, no. November 2017, pp. 124–131, 2018.

[28] K. Yildirim, M. L. Hidayetoglu, and A. Capanoglu, “Effects of Interior Colors on Mood and Preference: Comparisons of Two Living Rooms,” Percept. Mot. Skills, vol. 112, no. 2, pp. 509–524, 2011.

[29] X. Zhang, Z. Lian, and Y. Wu, “Human physiological responses to wooden indoor environment,” Physiol. Behav., vol. 174, pp. 27–34, 2017.

[30] P. C. Petrantonakis and L. J. Hadjileontiadis, “Emotion recognition from EEG using higher order crossings,” IEEE Trans. Inf. Technol. Biomed., 2010.

[31] R. M. Mehmood and H. J. Lee, “A novel feature extraction method based on late positive potential for emotion recognition in human brain signal patterns,” Comput. Electr. Eng., 2016.

[32] A. Bhardwaj, A. Gupta, P. Jain, A. Rani, and J. Yadav, “Classification of human emotions from EEG signals using SVM and LDA Classifiers,” in 2nd International Conference on Signal Processing and Integrated Networks, SPIN 2015, 2015.

[33] M. Mohammadpour, S. M. R. Hashemi, and N. Houshmand, “Classification of EEG-based emotion for BCI applications,” in 7th Conference on Artificial Intelligence and Robotics, IRANOPEN 2017, 2017.

[34] P. Ackermann, C. Kohlschein, J. Á. Bitsch, K. Wehrle, and S. Jeschke, “EEG-based automatic emotion recognition: Feature extraction, selection and classification methods,” in 2016 IEEE 18th International Conference on e-Health Networking, Applications and Services, Healthcom 2016, 2016.

[35] O. Bazgir, Z. Mohammadi, and S. A. H. Habibi, “Emotion Recognition with Machine Learning Using EEG Signals,” 2019.

[36] N. Jatupaiboon, S. Pan-ngum, P. Israsena, and et al., “Real-Time EEG-Based Happiness Detection System,” Sci. World J., vol. 2013, no. 3, pp. 1–12, 2013.

[37] Y. Li, K. K. Ang, and C. Guan, “Digital Signal Processing and Machine Learning,” 2009.

[38] A. V Oppenheim and R. W. Schafer, “Discrete Time Signal Processing 2nd Edition,” Book. 1998.

[39] J. Proakis, Digital signal processing(4th Edition). 1975.

[40] M. Fatourechi, A. Bashashati, R. K. Ward, and G. E. Birch, “EMG and EOG artifacts in brain computer interface systems: A survey,” Clinical Neurophysiology. 2007.

[41] J. A. Urigüen and B. Garcia-Zapirain, “EEG artifact removal - State-of-the-art and guidelines,” J. Neural Eng., 2015.

[42] S. D. Muthukumaraswamy, "High-frequency brain activity and muscle artifacts in MEG/EEG: A review and recommendations," Frontiers in Human Neuroscience. 2013.

[43] W. Wen, G. Liu, N. Cheng, J. Wei, P. Shangguan, and W. Huang, "Emotion recognition based on multi-variant correlation of physiological signals," IEEE Trans. Affect. Comput., vol. 5, no. 2, pp. 126–140, 2014.

[44] J. Bakker, M. Pechenizkiy, and N. Sidorova, "What's your current stress level? Detection of stress patterns from GSR sensor data," in Proceedings - IEEE International Conference on Data Mining, ICDM, 2011.

[45] "https://looxidlabs.com/looxidvr/." .

[46] R. Girshick, "Fast R-CNN," in IEEE International Conference on Computer Vision (ICCV), 2015.

[47] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. Backpropagation applied to handwritten zip code recognition. Neural computation, 1989.

[48] W. SHI, M. B. ALAWIEH, X. LI AND H. YU, "ALGORITHM AND HARDWARE IMPLEMENTATION FOR VISUAL PERCEPTION SYSTEM IN AUTONOMOUS VEHICLE: A SURVEY," INTEGRATION, THE VLSI JOURNAL, VOL. 59, NO. MAY, PP. 148-156, 2017.

[49] Y. WANG, Q. CHEN, Q. ZHU, L. LIU, C. LI AND D. ZHENG, "A SURVEY OF MOBILE LASER SCANNING APPLICATIONS AND KEY TECHNIQUES OVER URBAN AREAS," REMOTE SENSING, VOL. 11, NO. 13, PP. 1-20, 2019.

[50] D. ANGUELOV, C. DULONG, D. FILIP, C. FRUEH, S. LAFON, R. LYON, A. OGALE, L. VINCENT AND J. WEAVER, "GOOGLE STREET VIEW: CAPTURING THE WORLD AT STREET LEVEL," COMPUTER, VOL. 43, NO. 6, PP. 32-38, 2010.

[51] I. PUENTE, H. GONZÁLEZ-JORGE, J. MARTÍNEZ-SÁNCHEZ AND P. ARIAS, "REVIEW OF MOBILE MAPPING AND SURVEYING TECHNOLOGIES," MEASUREMENT: JOURNAL OF THE INTERNATIONAL MEASUREMENT CONFEDERATION, VOL. 46, NO. 7, PP. 2127-2145, 2013.

[52] J. L. SCHÖNBERGER AND J. FRAHM, "STRUCTURE-FROM-MOTION REVISITED," IN 2016 IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), 2016.

[53] M. J. LEOTTA, E. SMITH, M. DAWKINS AND P. TUNISON, "OPEN SOURCE STRUCTURE-FROM-MOTION FOR AERIAL VIDEO," 2016 IEEE WINTER CONFERENCE ON APPLICATIONS OF COMPUTER VISION, WACV 2016, 2016.

[54] K. M. YI, E. TRULLS, V. LEPETIT AND P. FUA, "LIFT: LEARNED INVARIANT FEATURE TRANSFORM MOTIVATION & CONTRIBUTIONS," LECTURE NOTES IN COMPUTER SCIENCE (INCLUDING SUBSERIES LECTURE NOTES IN ARTIFICIAL INTELLIGENCE AND LECTURE NOTES IN BIOINFORMATICS), 2016.

[55] M. QUIGLEY, K. CONLEY, B. P. GERKEY, J. FAUST, T. FOOTE, J. LEIBS, R. WHEELER AND A. Y. NG, "ROS: AN OPEN-SOURCE ROBOT OPERATING SYSTEM," IN ICRA WORKSHOP ON OPEN SOURCE SOFTWARE, 2009.

[56] S. BLASER, S. CAVEGN AND S. NEBIKER, "DEVELOPMENT OF A PORTABLE HIGH PERFORMANCE MOBILE MAPPING SYSTEM USING THE ROBOT OPERATING SYSTEM," ISPRS ANNALS OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES, VOL. 4, NO. 1, PP. 13-20, 2018.

[57] S. BLASER, S. NEBIKER AND D. WISLER, "PORTABLE IMAGE-BASED HIGH PERFORMANCE MOBILE MAPPING SYSTEM IN UNDERGROUND ENVIRONMENTS - SYSTEM CONFIGURATION AND PERFORMANCE EVALUATION," ISPRS ANNALS OF THE PHOTOGRAMMETRY, REMOTE SENSING AND SPATIAL INFORMATION SCIENCES, VOL. 4, NO. 2/W5, PP. 255-262, 2019.

[58] R. MUR-ARTAL, J. M. MONTIEL AND J. D. TARDOS, "ORB-SLAM: A VERSATILE AND ACCURATE MONOCULAR SLAM SYSTEM," IEEE TRANSACTIONS ON ROBOTICS, VOL. 31, NO. 5, PP. 1147-1163, 2015.

[59] W. HESS, D. KOHLER, H. RAPP AND D. ANDOR, "REAL-TIME LOOP CLOSURE IN 2D LIDAR SLAM," PROCEEDINGS - IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, VOLS. 2016-JUNE, PP. 1271-1278, 2016.

[60] M. JANCOSEK AND T. PAJDLA, "MULTI-VIEW RECONSTRUCTION PRESERVING WEAKLY-SUPPORTED SURFACES," PROCEEDINGS OF THE IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, PP. 3121-3128, 2011.

[61] P. MOULON, P. MONASSE AND R. MARLET, "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), VOL. 7727 LNCS, NO. PART 4, PP. 1-14, 2012.

[62] Q.-Y. ZHOU, J. PARK AND V. KOLTUN, "OPEN3D: A MODERN LIBRARY FOR 3D DATA PROCESSING," ARXIV PREPRINT ARXIV:1801.09847., 2018.

[63] D. SCARAMUZZA AND F. FRAUNDORFER, "VISUAL ODOMETRY [TUTORIAL]," IEEE ROBOTICS & AUTOMATION MAGAZINE, VOL. 18, NO. 4, PP. 80-92, 2011.

[64] D. SCARAMUZZA AND F. FRAUNDORFER, "VISUAL ODOMETRY PART I: THE FIRST 30 YEARS AND FUNDAMENTALS," IEEE ROBOTICS AND AUTOMATION MAGAZINE, VOL. 18, NO. 4, PP. 80-92, 2011.

[65] J. ZHANG AND S. SINGH, "VISUAL-LIDAR ODOMETRY AND MAPPING: LOW-DRIFT, ROBUST, AND FAST," PROCEEDINGS - IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, VOLS. 2015-JUNE, NO. JUNE, PP. 2174-2181, 2015.

[66] R. A. NEWCOMBE, S. J. LOVEGROVE AND A. J. DAVISON, "DTAM: DENSE TRACKING AND MAPPING IN REAL-TIME," IN 2011 INTERNATIONAL CONFERENCE ON COMPUTER VISION, 2011.

[67] C. FORSTER, Z. ZHANG, M. GASSNER, M. WERLBERGER AND D. SCARAMUZZA, "SVO: SEMIDIRECT VISUAL ODOMETRY FOR MONOCULAR AND MULTICAMERA SYSTEMS," IEEE TRANSACTIONS ON ROBOTICS, VOL. 33, NO. 2, PP. 249-265, 2017.

[68] J. KUO, M. MUGLIKAR, Z. ZHANG AND D. SCARAMUZZA, "REDESIGNING SLAM FOR ARBITRARY MULTI-CAMERA SYSTEMS," 2020.

[69] A. PUJOL-MIRO, J. RUIZ-HIDALGO AND J. R. CASAS, "REGISTRATION OF IMAGES TO UNORGANIZED 3D POINT CLOUDS USING CONTOUR CUES," 25TH EUROPEAN SIGNAL PROCESSING CONFERENCE, EUSIPCO 2017, VOLS. 2017-JANUA, PP. 81-85, 2017.

[70] O. VYSOTSKA AND C. STACHNISS, "IMPROVING SLAM BY EXPLOITING BUILDING INFORMATION FROM PUBLICLY AVAILABLE MAPS AND LOCALIZATION PRIORS," PHOTOGRAMMETRIE, FERNERKUNDUNG, GEOINFORMATION, VOL. 85, NO. 1, PP. 53-65, 2017.

[71] I. KALISPERAKIS, L. GRAMMATIKOPOULOS, E. PETSA AND G. KARRAS, "A STRUCTURED-LIGHT APPROACH FOR THE RECONSTRUCTION OF COMPLEX OBJECTS," IN GEOINFORMATICS FCE CTU, 6, 2011.

[72] D. SCHARSTEIN AND R. SZELISKI, "HIGH-ACCURACY STEREO DEPTH MAPS USING STRUCTURED LIGHT," IN IEEE CVPR, 2003.

ANNEX I

This section presents the questionnaire that subjects had to fill before conducting the experiments.

subject N.

1. Surname:
2. Age:
3. Gender:
4. Nationality:
5. Left handed / Right handed:
6. Education:
7. Occupation:
8. Education in art/fine art:

<u>Voluntary participation</u>

Participation is voluntary and there will not be any monetary remuneration. You are able to withdraw your agreement to participate in this research project at any time, without having to give any reasons and without incurring disadvantages as a result.

<u>Confidentiality and anonymity</u>

Confidentiality is ensured with all data raised within the context of this research project. No personal information will be disclosed to individuals who are not members of the responsible research team. The data collected can only be published in anonymized form that will render impossible the identification of your person, your family, your employer and your place of residence. The data collected will be published only in anonymized form, i.e. without your name or address.

<u>Participant</u>

Place, date, signature:

<u>Project manager/contact person</u>

Name, first name:

Place, date, signature: