Tampere University

Aditya Premi

MIXED REALITY TOOLSET FOR ONSITE PROCESSES IN CONSTRUCTION AND MANUFACTURING INDUSTRIES

Master of Science Thesis

Master's Thesis
Faculty of Engineering
and Natural Sciences

Examiners:
Jose L. Martinez Lastra
Luis Gonzalez Moctezuma

18 June 2022

ABSTRACT

Aditya Premi : Mixed Reality Toolset For Onsite Processes In Construction and Manufacturing Industries
Master of Science Thesis
Tampere University
Master’s Degree Programme in Automation Engineering
Major: Factory Automation and Industrial Informatics
June 2022

In traditional and labour-intensive industries, such as construction and manufacturing, every time a task must be repeated, it not only affects the schedule but also increases costs. Site visits to collect information, during the early stages of the project, are very crucial as this information is used to create plans and budgets. Furthermore, a large percentage of the existing factories and buildings are not digitally documented. Many buildings may also have outdated 2D designs. Therefore, a key problem during site planning is the collection of accurate data. Evidently, there is a lack of modern technological tools that ensure that the reality is consistent with the virtual planning data.

Augmented Reality (AR) technology can digitalise the process of site information collection and related tasks such as visualisation of 3D designs. Various other capabilities of AR, with smart glasses such as Microsoft HoloLens, include scanning, measurement, interactions of virtual objects with the real world and location aware textual notes. This thesis explores the capabilities of AR with smart glasses such as HoloLens, by first understanding the needs of the industries along with identifying ways in which AR can digitalise industry processes and developing necessary tools to enable the same.

The Unity3D engine is used to carry out the development work and the resultant tools are implemented on HoloLens. These functionalities are developed keeping in mind principles of user experience. They allow the user to capture the space around them as a 3D mesh, measure distances between objects in that space or with virtual objects that have been placed, interact with various models by moving them or even slicing them to meet the user’s requirements and creating annotations in different parts of the working space.

The developed tools were then tested in the industry case study with Fortum. Here it was verified that the newly developed tools met most of the requirements. However, each company may have their own set of requirements and may need more customised AR solutions to fix them. This observation resulted in the development of an AR toolset that provides different functionalities to the user. This enabled the leveraging of user knowledge of the situation and maximising the strength of AR technology while solving the problem.

Keywords: Augmented Reality, Smart glasses, HoloLens, Manufacturing, Construction Industry

PREFACE

This thesis was done under the guidance of Prof. Jose L. Martinez Lastra and Luis Gonzalez Moctezuma, Faculty of Engineering and Natural Sciences, Automation Technology and Mechanical Engineering, Automation Technology, Tampere University, Finland.

I would sincerely like to express my gratitude towards Prof. Lastra for guiding my thesis and providing constant support throughout my research. The support of Luis Gonzalez Moctezuma and Wael Mohammed in bringing this research to a conclusion shall also not be forgotten. I would also use this opportunity to thank Timo Pastila, Kii Oy, for providing an opportunity to perform such research work in the Finnish Industry.

Heartfelt thanks to my dear friends, Ramyakrishna Jayanthi, Abhishek Singh Gaur, Mathew Mailaparampil, Shruti Mittal, Vipul Dhingra, Jayesh Gupta and Jazib Javed, who greatly supported me in their own ways throughout the course of this thesis, and to my parents for their constant motivation. Last but not the least, I thank the reader for taking time to be an audience of this research.

Tampere, 18 June 2022

Aditya Premi

CONTENTS

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

AR	Augmented Reality
VR	Virtual Reality
MR	Mixed Reality
2D	Two Dimensional
3D	Three Dimensional
IT	Information Technology
GB	Gigabyte
RGB	Red Green Blue Colour Model
HMI	Human-Machine Interface
SDK	Software Development Kit
KPI	Key Performance Indicator
CAD	Computer-aided design
BIM	Building Information Modelling
ToF	Time of Flight
UX	User Experience
UI	User Interface
API	Application Program Interface
SDLC	Software Development Life-Cycle
IoT	Internet of Things
URL	Uniform Resource Locator

.

1. INTRODUCTION

In today's day and age there is a growing trend for an improved Human Computer Interaction and to make technology more wearable. One of the results of this trend is the advent of smart glasses [1]. Smart glasses not only revolutionise the way we interact with the environment but also enhance user experience with enriched content. Smart glasses are wearable augmented reality device that employs interactive computer systems, wherein they display virtual elements along with what lies in the user's view [2].

There are two different methods of augmenting information into head mounted displays, namely Augmented Reality (AR) and Virtual Reality (VR). AR refers to placing contextual data onto the real world. On the other hand, VR takes a user into a virtual world with no contact to the real world. Slowly as the technology is progressing, AR and VR are merging to form a new experience called Mixed Reality (MR). For industries, AR has shown a lot of promise as it can, by definition, assist the user to perform daily tasks while still being in their current work conditions.

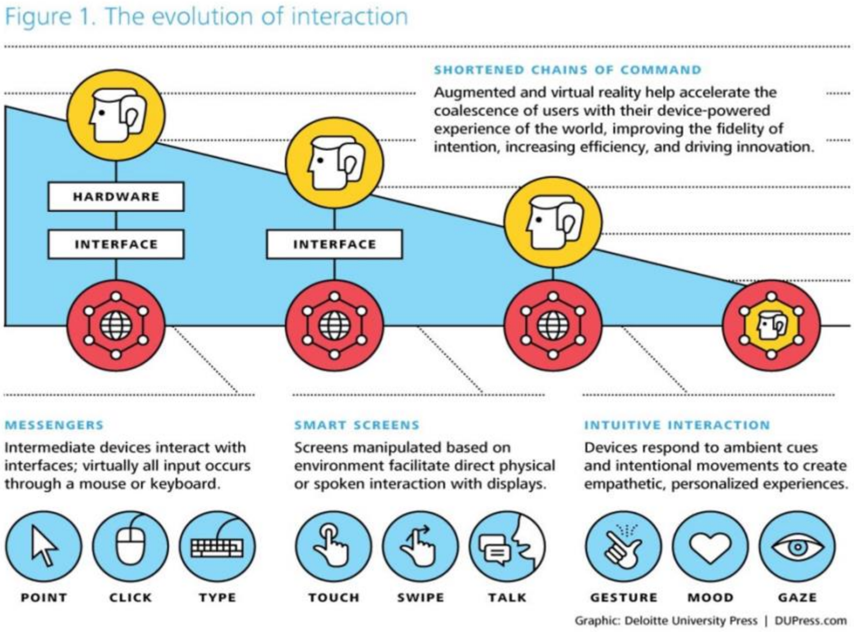
Different industries have shown different levels of acceptance to these technologies. In sales and marketing, using technologies such as VR are fast gaining popularity. However, some areas like manufacturing and construction, which rely on labour intensive workflows at site are still not leveraging the potential of such upcoming technology.

The research has been carried out at the company, Kii Oy, which is pioneering in smart glass applications. Kii Oy has been in the field of developing smart glass solutions for industries and has had a close association with various companies in the manufacturing, construction, renovation, demolition, and power industry. This association has brought to light various issues faced by companies when using traditional methods during onsite planning. Furthermore, it also has been found that there is a gap between what the smart glass industry develops and what users in various industries need. This research aims at understanding the needs of industries and creating a solution that can not only meet industries' needs but also prove to be an asset in their workflows.

1.1 Motivation

The origin of interest in Augmented Reality (AR) and Virtual Reality (VR) was during a mobile application development project, while working as a Research Assistant at Factory Automation Systems and Technologies Laboratory (FAST Lab.) in Tampere University of Technology (TUT). This led to exploration in the area of VR on mobile phones using Google Cardboard. A major outcome of further research at Kii Oy elucidated the opportunity for using this technology to introduce improvements to the existing workflows in the industry with development of industry focused applications. This thought paved the path to carry out a thesis in this domain, prompting to research different industrial processes and find a gap between AR offerings and requirements of the established industries.

It is quite intriguing that the construction and manufacturing industries have seen slower adoption of modern interaction practises for information generation, retrieval and transfer at the site as opposed to similar functions by sales. For example, in the manufacturing industry, assembly planning is a highly visual problem involving complex combinatorial operations. A major percentage of assembly-based operations are still manual except very basic pick and place operations. This highlights an important problem regarding Human-Machine Interface. This has led to a lot of prevalent issues like multiple site visits and extensive labour cost which are currently not faced by other completely digitalized industries like IT and media.



Augmented and virtual reality help accelerate the coalescence of users with their device-powered experience of the world, improving the fidelity of intention, increasing efficiency, and driving innovation.

Intermediate devices interact with interfaces; virtually all input occurs through a mouse or keyboard.

Screens manipulated based on environment facilitate direct physical or spoken interaction with displays.

Devices respond to ambient cues and intentional movements to create empathetic, personalized experiences.

Figure 1. *The evolution of interaction [3]*

Figure 1 exemplifies how the modern tools have been developing along to find new ways to solve complex problems. Digitalisation has proved to be powerful enabler to several industries, but it has always been difficult to digitalise industry processes which majorly constitute of real world interactions. In construction industry, most of the work is happening in harsh and remote locations which are normally not well suited to software and hardware developed for office environment [4]. Furthermore, construction industry is one of the least digitalised sector as per MGI's (McKinsey Global Institute) digitization index as referred in the figure below [5].

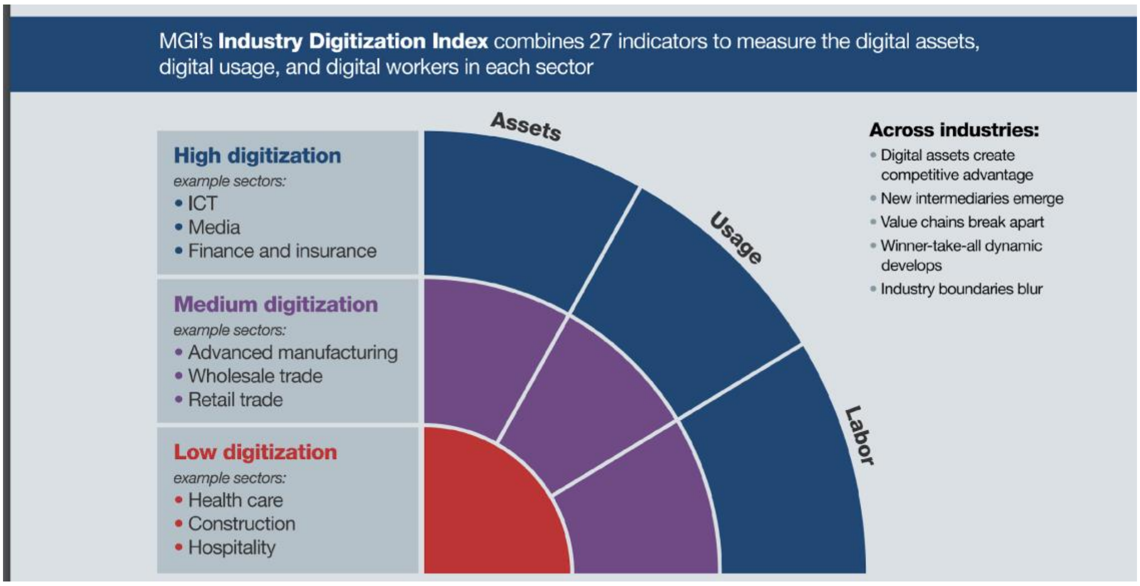


Figure 2. *MGI's Industry Digitization Index*

Hence many companies in this sector, still limit the use of IT to scheduling and designing works usually done from the comfort of an office environment. The work on the floor is performed using pen and paper, and messages between teams and contractors are often marked using painted symbols on walls.

While the benefit of the use of AR lies in its understanding of the environment and ability to interact with it, it is a promising technology which has the capability to overcome this major barrier of digitalising onsite operations [6]. Hence there is a need to realise industry problems in depth and find solutions which are practical and feasible to the manufacturing and construction industry.

1.2 Justification and Hypothesis

Harvard business review [7] has mentioned that the labour-intensive construction industry has the lowest level of digitisation. This gap at the sector level has been depicted by the Industrial Digital Index. The review also claims that for its digitalisation there is the need to provide digital tools to the workforce to enable improved efficiency and user experience.

At this point, let us consider a case where a construction engineer must undergo site visit for renovation. The ideal scenario involves the engineer to:

- Do a basic survey of the building by using tools like measurement tape.
- Mark the points of interest by using physical sprays.
- Write notes about the changes required and customer’s concerns on paper.
- Return to the office and use the collected data for designing the solution.
- Show the proposed design to the customer.

However, the actual scenario is much more complex than the above description. In most cases, the engineer has insufficient data after the first visit required for designing the solution. It requires either subsequent visits to gather the missing data or assume approximate values for the measurement. This results in redundant travelling or having an unreliable solution respectively [8]. For specific cases, Laser scanning is used which gives highly accurate representation of the environment. The process is expensive and time consuming which limits its applicability [9]. Further, the retrieval of measurement data and its storage is still mostly a manual process. For example, in the quality control process, the dimensional assessment is done through visual inspection which is a time-consuming and costly process [10]. Hence, there is high-cost impact as well as loss of time.

The second case is concerning the customised installation sector of the manufacturing industry. The initial phases of the project involve visiting the customer premises and gathering the parameters critical for installation. This includes:

- Measurement of space available for installation and taking notes of environmental conditions & installation bottlenecks.
- Make site layout using the gathered data
- Create and insert 3D design of the product in the resultant layout.
- Send the designed solution as a collection of rendered images to the customer.

As previously observed in the construction industry, many issues are experienced in this process as well. Firstly, customer is not easily satisfied by the solution because in many cases it turns out that the site requirements have not been correctly taken into consideration.

For instance, a major specialist organisation based in Finland, in the field of machines and systems for sheet metal working, experience issues while taking measurement data and planning layout for installation of their machines. During renovation, the company's customer premises consist of huge piles of debris and construction machines making it very difficult to take accurate measurements. Their customer needs them to show how their proposed layout of machines visualise in the site environment. The most critical requirement is that customer need to understand how space is being planned so that they can understand how much free space is available after the company has installed their machines. For this they use large cardboard boxes and stack them one over the other to give an idea on how much space their machines shall use. In some instances, customers even require a prototype of proposed design for proper validation. In all current solutions there is a lack of possibility for a product design to be brought to the customer environment before it is installed. Hence the customer approval phase takes long duration and multiple iterations. This greatly impacts the manufacturing industry as a lot of time and resources are spent for the bidding and planning phase of the project.

The use of Augmented Reality has shown to bring high-impact benefits such as the minimisation of operational time, avoiding user errors and costs reduction [11]. However, enabling mixed reality concepts also require compromise in usability and technology trade-offs. This brings the need of selection of the right mixed reality device / smart glass for a chosen application. This helps to optimise the parameters based on KPIs. Characteristics such as field of view, interaction, price, and display resolution play an important role in governing the choice of the smart glasses for a particular application [12]. This shows a close link with the industry acceptance of the glasses. However, the use of such glasses in industrial sites, especially smart factories must be further explored. Anna Syberfeldt et al. in their work have provided a detailed step by step method to select a glass suitable for the needed application [13]. Studies are also being carried out to understand the capabilities of various smart glasses, depending upon their technology, the availability of different applications and various SDKs to enable further development [11] [14].

The hypothesis of the thesis is that digitalisation of site work would improve workflows on site in the construction and manufacturing segments of the Finnish Industry, by improving processes associated with on-site data collection and manipulation. This digitalisation is to be achieved using a precise set of AR tools put together after understanding the standardised workflow in different companies.

1.3 Problem Statement

From both the cases of construction and manufacturing industry as previously studied, it is apparent that the traditional methods of collecting information causes a negative impact on project time, costs and quality. A lack of digitalisation comes out to be an evident reason to this problem. A further deeper investigation into this also suggests that the different needs of the industry are not in-line with the methods proposed by technology.

Research Questions:

1. What are the problems faced by the industries using current traditional methods for onsite planning?
2. Which are the most common industry methods causing the problems?
3. How can Smart glass technology provide digitalised tools to improve the existing solution?
4. What are the features of a common toolset that can serve the necessities of the construction and manufacturing industry?

This thesis tries to solve the time, cost, quality issues for the construction and manufacturing industries for onsite planning by enhancing the traditional industry methods using smart glasses as a tool.

1.4 Scope

This thesis includes research in the field of Augmented Reality and Virtual Reality and their application in the Manufacturing and Construction Industries to improve site related tasks. It details the use of Mixed Reality via smart glasses to develop solutions like 3D scanning, Slicing, Annotation and Measurement. The research also includes a case study which strengthens the findings of the thesis.

1.5 Challenges and Limitations

Every new technology has its own learning curve. As augmented reality is still not widely available, not many individuals have had a first-hand experience of the same. Using smart glasses is not normalised. The device feels ergonomically uncomfortable to some users. Also, working with the glasses on may feel intrusive to some, and as it collects data online, some have expressed that using such a device may be perceived as an intrusion of privacy. All these factors result in possible resistance by site workers and engineers to use this technology.

The applicability of the proposed solution is limited by the initial costs associated with implementing such solutions. This is because the hardware is currently viewed to be expensive, and this can become a barrier for doing the first pilot projects even though the solution shall be cost efficient in the longer term [15].

In both industries, there have been challenges with respect to the 3D models. One such challenge faced was that of data security. In the manufacturing industry, the designs are the most priced data of the manufacturer, which makes the companies hesitant to share this data. Also, they object to uploading this data on public cloud platforms, thereby increasing the cost of the solution, or limiting the number of models that can be accessed by an application. On the other hand, the construction industry has models with file sizes that are a few GB. When data of this size is imported into Unity3D, the model is reconfigured into smaller parts by the platform itself to be processed. Also, smart glasses have small processors, which cannot work with such large 3D models. This rises a need to optimise the 3D content to suit the platform and smart glass environment, a task which is beyond the scope of this thesis.

Another technological challenge is in the integration of data. Different software are used to create and manipulate 3D models like AutoCAD, SiemensNX etc. The preference of software differs from company to company, and the output of 3D formats vary. Examples include like .obj, .ifc, .fbx and so on. This requires the solution to be able to integrate data in these different formats and the absence of the same affects the adoptability of the solution drastically. Finally, while performing this thesis, the augmented reality devices provided by Kii Oy limited the AR devices on which the solution could be tested.

1.6 Document description

The first chapter of the thesis briefly introduces the reader to the needs of the industry and technology used to meet them. It informs the reader, the motivation behind this research, the problem statement which the thesis aims to solve, the scope of the research and thereby its limitations. The second chapter familiarises the reader further in the area of research by bringing together the state of art being carried out in different fields that aid in formulation of the methodology to find the solution to the problem statement. The third chapter focuses on the methodology employed to solve the problem. It informs the reader of the approach taken to find the solution and details the technology used in doing so. The fourth chapter provides a description of the AR based solution, the algorithms used and the possibilities of the developed solution. Chapter five outlines a case study at Fortum Oy, where the developed solution was employed, and the results were recorded to be analysed. The sixth chapter shows the results of the research, including some data from the case study and some independent experimentation to validate the developed solution. Chapter 7 discusses the results and what it implies, especially in context to the problem statement identified at the start of the research. Chapter 8 and 9 provide a conclusion to the research and suggest possibilities of future work in this research to enhance the solutions further and formulate new methods to solve similar problems.

2. STATE OF ART

2.1 Mixed Reality

Mixed reality is an environment where real objects exist and interact with virtual objects. This interaction is made possible by software tools that implement features that are often difficult or impossible to replicate in the real world. There are several applications of this technology, and it also saves time, effort, and cost.

2.2 Applications in Industries

Mixed reality finds many applications in many industries. For instance, the digitised models of the site environment are not only useful in architecture but also for use cases such as facility management in manufacturing and energy industry and also in robotics for rescue and inspection [16]. Also, in gaming environments rather than creating models from scratch, it is believed that scanned models of real environments are believed to improve the visual impact of the game and it helps the gamer feel more involved in the gameplay.

The focus of the thesis is on the manufacturing and construction industry. The industry in the current times face challenges on the fronts of being highly labour intensive and face high losses per error made. The thesis tends to illustrate some applications of mixed reality that can prove vital on these fronts and help in changing the industrial dynamics in the future.

2.2.1 Manufacturing

The manufacturing industry is one of the industries known for being highly labour intensive. The industry planning includes site planning and design. For site planning, one of the indispensable tasks for the onsite workforce is measurement. Measurements are needed to check if different items can be fit into the given area and what is the best possible fit for them. Also, the distances between objects which are not yet present need to be calculated often to effectively space out and maximise the use of the space. One method of doing this is to mark out the area each item occupies on the floor of the site or make measurements of the site area and then use modelling software to create 3D layouts. This lays the foundation to plan factory floors. Using these 3D layouts, indoor factory setup can be achieved by deploying desktop software solutions for importing and placing 3D models of the machines to check and measure if the space can accommodate the machines. Further it helps to validate whether the setup is sensible and if there can be a more efficient utilisation of the given space. Finally, it also helps to check if all the safety needs are being compiled correctly. This limits the work to be performed off site, i.e. in the office where the software is available.

Retrofitting in the context of the manufacturing industry means to upgrade a manufacturing site by replacing machines, modifying assemblies, or adding parts to the current site. Hence the site design needs to be upgraded and its planning involves accurate understanding of existing space [8]. In many cases, laser measurement tools are used for precise measurement. This is because in retrofits, the site visit happens when the place to be renovated is still being occupied, and the renovators have to make measurements really fast. Also, once the measurements have been made, it is very difficult to go back and make them again. Hence quick measuring tools are preferred. Sometimes the measurements include diagonals, areas, and curved edges. Also, in some cases there is a need to make measurements between things that exist and things that shall be placed, like measuring the distance between a door frame and a new balcony that shall be constructed. These measurements are generally not made but are calculated later depending on the plans and available measures.

In the past, the companies in the manufacturing sector have faced challenges to make accurate estimation of factory spaces which do not have updated designs. Shop floor design means placing heavy machinery and tools along with workbenches at different places within the site. Thus, the whole process can become labour intensive and the need for precise planning arises. When a factory layout has to be planned, rather than just referring to plans, it is always preferred to have a 3D model of the available space. This is to make it possible to visualise all the machines in the space and plan the area to meet the safety and ergonomic requirement while in planning to reduce rework later. Making layouts by taking discrete measurements needs a lot of manual work and multiple iterations. In this context, where company's factories are not having 3D designs, indoor scanning turns out to be an effective tool for making as-built designs.

Likewise, while installing new equipment into an existing place, a scanned model is very useful as it will include other information about changes in the space which may not have been updated in the plan documentation. When logistics are planned it is extremely important to know if the equipment being moved can fit through a particular space. This is so that the tools needed to move the equipment can be planned and decisions regarding location of containers in a warehouse etc. can be properly made. These become necessary to maintain logistics estimation based on area and unit sizes. Scanning the location becomes very useful in such an environment as the scanned model can be used to design the logistic movement to prevent any sudden problems at site. This is not only used by factory warehouses but also for automobile assembly lines to plan size of work cells, corridors, space for robots to ensure easy and quick movement of equipment.

The mixed reality can help in creating virtual environments which become a walkthrough. An important industrial application of this is to create training environments that closely resemble the real environment so that new users can learn and prepare themselves for the real environment before entering it. This is to reduce any hazard and to replicate the real environment remotely, thus reducing the cost of training.

Scanning of objects and environments for not only measurement but for the model itself is also used in the industries. Another example is seen in traditional industries. As more and more digitalisation is taking place, these industries are facing a growing need to scan objects which do not have 3D models so as to create their soft twins in their IT systems. Though there seems to be a lot of options in finding the right hardware and technology to meet most applications needing a 3D scan, it does not go without its share of drawbacks.

Using annotations in the automotive manufacturing industry is an area that is gaining popularity. Use of annotations to notify the driver about the space available behind the vehicle while reversing the vehicle is a well-known example. Smartphone based applications like the Hyundai Virtual Guide can be used to explain the meaning of various symbols on the dashboard, status of the vehicle can be seen as well as manuals and how to use videos can be displayed [17].

All these applications show how important digital annotations can be, but in practice to achieve them there are some hindrances. When there are no models available, like in the case of 360 videos, annotating in 3D space becomes difficult and requires experienced or trained personnel to do it.

Most of the devices that make it possible to gather 3D data at site like laser scanners do not provide a facility to add any annotations online. This makes the process cumbersome and introduces the possibility of some information going missing between when the data was captured and when the annotation was performed. When planning logistics, to move large machines or items, measurements need to be made to access the right strategy. Most of these measurements are still made by measuring tape, which may sometimes be tedious.

The above challenges find solutions in mixed reality. Creating virtual environments along with onsite measurements, corrections and adding annotations are some of the many resolutions that mixed reality brings forth. Thus, it becomes very useful in the manufacturing sector.

2.2.2 Construction

Construction companies taking up renovation projects especially are requesting for such tools to know the true picture of the existing space as compared to blueprints to capture any changes that may have occurred since when the building was first constructed. Also, such scans make it possible for them to make extra measurements after the site has been visited once, without needing to go back. During the execution, scanning the space helps keep a tab on the exact progress of construction work making planning and expediting more accurate.

Companies which demolish large buildings also find a need for such tools. It helps them understand the scope of work, enabling accurate bids and measuring cost. Further, the volume of debris expected can be calculated, which helps them in project management extensively. Similarly, even real estate planning is greatly benefited by indoor scanned models. It is used by large convention centres, universities, and offices to better utilise the space they have and create comfortable environments for the people using them.The applications for tools that can accurately measure length are endless. Here we shall explore a few where in the use of smart glasses could be a potential advantage.

The construction industry requires number of measurements to be done regularly. From measuring lengths of wall to heights of pillars to measuring platform heights, check floor and ceiling flatness, to compare walls for parallelism and perpendicularity, to sometimes even measuring curved surfaces. The measurements of the site area enable creation of as-built 2D layouts with the help of CAD software like Autodesk AutoCad. This can be used to plan large conference spaces to maximise participant interaction or a school to promote a healthy atmosphere for children to learn, alike.

These measurements require high precision, but also should not take too much time. Most of the measurements are done by measuring tapes or laser detectors, which either require the person taking the measurements to move large distances or cause errors because the device is not held parallel.

When construction is in progress, site evaluations and inspections to track construction progress are common. These inspections involve a supervisor making notes and taking measurements to check the quality of work and see if any possible issues have come up. These measurements are also mostly made by measuring tapes, taking up a lot of time from the supervisor’s schedules, effectively meaning they are costly.

In companies that carry out demolition, one very critical measurement is to measure the volume of debris. Volume is critical not only for measuring work required for handling such volume but also for calculating volume of valuable materials like metals. Hence it helps to give a fair estimation on what should be the cost of demolition. This is done by measuring the volume of different walls, columns, floors etc. This needs not only area measurement but a device that can multiply the area with the thickness of these structural elements. Most of the measurement tools are incapable of doing this, which is why such companies rely on extra software to make such calculation since site measurements have been made. This needs many measurements to be taken at site, and if this is not possible, it usually results in them making bids which may not be correctly estimated.

Another area where in measurement plays a very important role is in installation of pipes and cables. These materials are ordered by length, and it is very critical to ensure the correct length of material is used at the right place to avoid wastage. Hence, before every installation measurement are made extensively to check if there is enough material and there are no errors. Sometimes the places where such installations have to be made are not easy to reach, which make the entire process very difficult and introduces the possibility of errors.

In video-based e-learning, it is beneficial to have tools that annotate videos along with addition of questions and comments [18]. Annotated data is used in almost every industry as it is very effective in adding context to visual information. In the field of architecture, such data is used to denote lengths, colours, texture/finishes and material being used on different surfaces like floor, ceiling, walls, roof etc. It also is used to denote lengths, provide information like floor, area, and other such details. Special doors and passageways like those used for safety are marked out clearly.

In renovation projects, annotations are very important to make sure information is passed on from one department to another without any confusion. Also, as the tasks that have to be carried out differ from place to place, details of such information are usually printed and pasted on a door at the site. In cases where complete walls have to erased or the flooring has to be replaced, the physical walls are marked with paint to act as an indicator for the workers to carry out the required work. Our visit to Fira Oy’s renovation site in Helsinki provided us with this knowledge and the unexpectedly heavy dependence of annotation by the renovation industry.

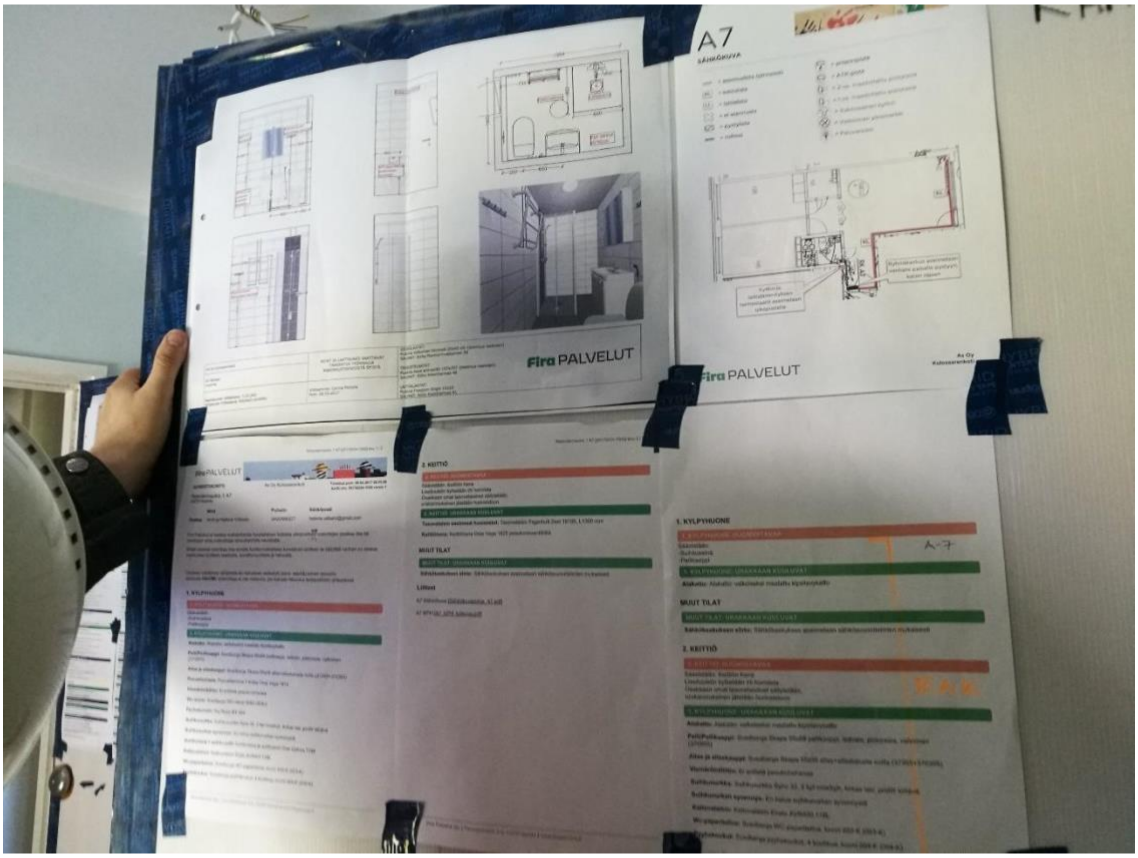


Figure 3. *Role of annotation in renovation*

In layout or site planning applications, annotations are used not only to add textual information in the form of labels but also several symbols are used to mark out different areas or features. This may include the used of different images, signs, characters, colours or types of lines. Similar methodologies for annotations are also observed in engineering and design drawing applications, though the meaning of the symbols may be completely different within the two applications. Visual annotations like warning signs are used on different HMI screens across industries to notify the user or operator or bring their attention to a particular event or situation.

Annotations are used to even make particular notes that need to be used later, when one accesses information and may need to visit it again later, or for someone else to refer. In the field of education and training this becomes very important. Not only can the teacher provide guides and instructions, but the students can use this tool to add notes or comments about what they find necessary. This makes the learning process more intuitive and easier. With the advent of using multimedia to make the process of earning easier, photos, videos and voice overs are used to ensure the necessary information is relayed.

Often annotations turn out to be asynchronous, like an additional desktop work which can be performed only on certain software like videoANT or Videonotes for videos or those that can read the 3D models or captured data and allow annotations for model-based information. This needs the software to attach the annotation to a particular position with respect to the visual data and save it in a format such that it can be reproduced without error. This calls for additional work hours and investment.

2.3 Virtual-Real World Interaction

Reality is, in essence, a representation of the world which can be felt through sensory organs. While the dimension of reality has been extended with technology, there is an apparent gap of testing all senses while interacting with reality. This limited sensory reality or virtual reality finds a major overlap with the real world and thus, becomes essential to solving multiple problems in multiple industries. This interaction between the virtual and real world is possible by the features which improve the overlap between virtual world and reality. These features are needed in different scenarios for different industries. Some of these actions have led to development of dedicated hardware or software. These include 3D scanning, slicing, annotation and measurement. Let us understand each of these features in detail.

2.3.1 3D Scanning

3D scanning as a process by which a three dimensional representation of an object or space can be created. The output of such a scan is often a point cloud. The point cloud data consists of a collection of Cartesian representations of the surfaces that have been scanned, sometimes along with their RGB colour data. Though there is no standard format available for creating point clouds, it follows the convention of having the first 3 columns representing the X, Y and Z coordinates, the next 3 columns representing the corresponding RGB data values. Additionally, some point clouds have intensity data and surface face data.

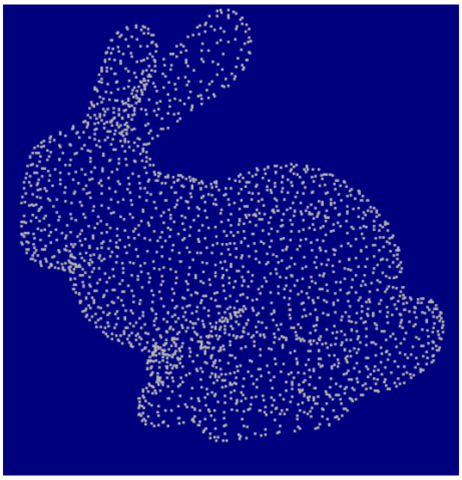


Figure 4. *A 2D representation of a simple point cloud*

Once this data is available, it can be used to create 3D models. This can be done by using any of the different methods of surface reconstruction [19]. Although the process may differ from method to method, the general idea is to first convert the data into polygon or triangular meshes or CAD models, which can then be used in a variety of software, such as AutoCad, Unity3D, blender, OBJ Viewer, MeshLab, GLC Player, JavaView amongst others.

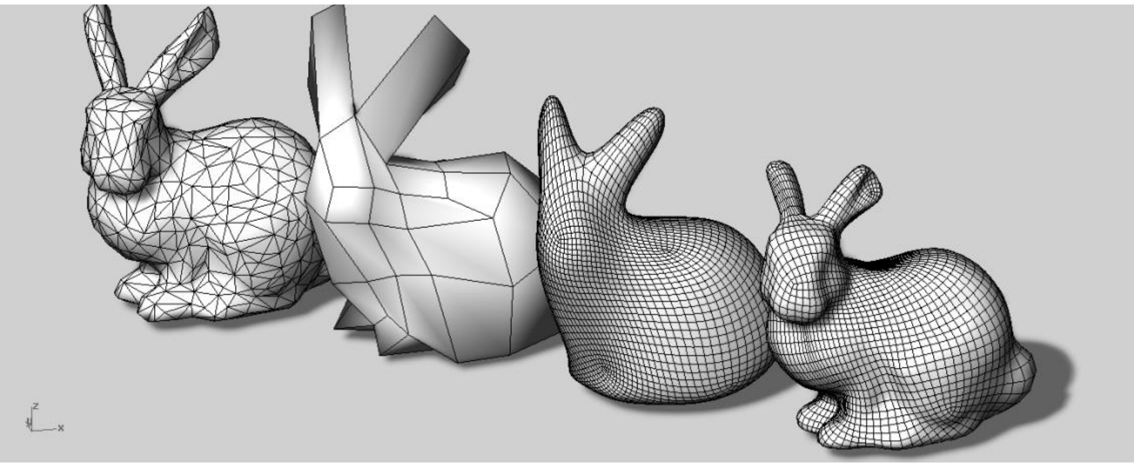


Figure 5. *Examples of triangular and polygonal mesh*

Different technologies are used today to create 3D scans. Some of them are mentioned below:

a. Laser Scanners: They employ the method of time of flight. The time between when a laser is released to when it is detected after having been reflected off a target helps determine the distance of the target from the scanner. Lasers have high directionality, which makes it easier to detect with precision the target of the reflected light. To scan an object, the scanners tend go around it, and to scan indoor locations the scanner moves inside the space while changing its angle of incident light to scan the entire area [20][21].

It is a task that is cumbersome and requires a process to be followed, the process of laser scanning is mundane, slow and repetitive [8]. There is no online computation unit on with the scanners. Hence, the technician working on the scan cannot comprehend or manipulate the scans. Also, online visualisation is not available, making it very difficult to know if some areas or spots have not been scanned completely. Secondly, sometimes format conversion is also a difficult process specially when the data gathered is so dense. As the technology depends on the reflection of light, glass and reflective surface may give rise to problems, requiring tuning methods to be applied to process the scan correctly.

b. 360 Video: In this fast pacing world, new forms of capturing the real environment has come to surface. As a special case of photogrammetry, the process starts with making a 360 video of the site using a suitable camera. There is existing software in market which can use this 360 data and generate a 3D model of the captured data. Kii has observed that manipulation of data extracted from 360 cameras is cumbersome and can take days and even up to a week to a single 3D scan of medium size room.

c. Photogrammetry: it is a process in which visual information in the form of photographs are used along with image processing tools to create a map or determine distance between or size of objects.

This process needs the technician to capture photos from multiple places along a path, and the path to be recorded as well. A combination of the path and the photos are used to create the final model, which may give rise to large errors if the angle of capturing images changes even by 5 degrees. As depth is calculated only by imaging technique and there are no depth sensors coupled, large errors may occur if there is any small inaccuracy in the initial scale setup. For example, if the initial scales are not exactly perpendicular, the resulting calculations shall compound the error. With a large dependence on consecutive frames, it requires the frames to move smoothly and clearly, thereby making it difficult to reduce the time taken for the process to happen. Texture/colour is often required to make it easy for understanding the created models. Photogrammetry provides this, while other methods of scanning have to create tools for selective colouring. Unexplored spaces create holes in the models which reduces the quality of experience in all the methods. High quality algorithms are needed to fill theses gaps in post processing work. Also, the models created are mostly only surface models, providing absolutely no information about the material being scanned or the material thickness.

Various companies use different hardware which may have different methods of working or varied algorithms, and finally result in different quality output as per below table.

Table 1: List of major companies that provide hardware for photogrammetry

Company	Hardware	Others
Mosini Caviezel SA	GarminVIBR™Elite	Around 100 images per project
Zeiss	Comet Photogrammetry	Recommended measurement volume between 100 x 100 x 100 mm³ up to 10 x 10 x 10 m³
Zebicon	Atos Plus	12 and 29-megapixel cameras for accurate measurement
Creaform	MAXSHOT NEXT™|ELITE	Accuracy is of 0.015 mm/m

d. Indoor Scanners: There are broadly two types of indoor scanners, Terrestrial and Mobile. The terrestrial models have the scanner at a fixed location, whereas the mobile scanners capture data from multiple locations and angles [22]. Some of the most popular commercial indoor mapping solutions include Matterport, NavVis, Zebedee, Stencil and Leica Pegasus: Backpack. Aalto VILMA, FGI Slammer and the Würzburg backpack are under research [23].

e. Collaborative scanning: This was an interesting method developed by a student team from the University of Washington. It breaks down a large area into smaller areas, each of which is scanned by different people present in that area. Multiple scans thus obtained are then added into one, making it possible to scan very large indoor areas in short times. However, two major problems faced are finding the correct points of overlap to combine the scans and maintaining the same quality of scan throughout. As each person is using a different piece of hardware, there arises the possibility of different resolutions being captured. Also, some areas may be over scanned while some may be under scanned.

f. Lidar: LiDaR stands for Light imaging, Detection and Ranging. It works on the same principle as radar, using laser in place of radio waves. The important components of a lidar system are the laser which are usually between 600 to 1000 nm, photodetectors which are usually PMTs or photodiodes, scanner optics which scan the resultant image from the detector and the positioning system to calculate the position and orientation when mobile lidar systems are used.

g. IR sensors: IR sensors are also used similarly like the lidar system to detect depth and use this data to create a depth map or 3D scan of an object or space. Google Tango is a leading hardware using this technology to create detailed 3D indoor maps. The infrared camera on Project Tango is a part of the OV4682 3D depth sensor, built by Omnivision. It provides the depth information, which when combined with the visual information gives a realistic 3D picture of any closed space. However, the range is limited to 4 metres, and as IR is absorbed by most biological surfaces like plants and animals, it causes many issues with accuracy in outdoors. Owing to this, Google has discontinued its support for tango in support for ARCore.

Research has been carried out to combine some of the techniques mentioned above to scan or model indoor environments like RGB-D mapping by Peter Henry et al [24] where the point cloud information and the visual RGB data are used together to create detailed 3D models which also contain their visual, colour information by frame by frame adding the scanned data and aligning it with the existing scan until a loop closure is detected.

Companies use indoor scanning for several applications. For example, while installing new equipment into an existing place, a scanned model is very useful as it will include other information about changes in the space which may not have been updated in the plan documentation. This is not only used by industries to move big equipment but also by hospitals to plan the size of corridors, doors, and rooms to ensure easy and quick movement of machines and beds. In addition to applications such as measurement, scanning is also used to create a digital model which can serve additional use cases. For instance, there has been a recent demand in the gaming industry to use scanned models of real environments rather than creating models from scratch as it is believed to improve the visual impact of the game and helps the gamer feel more involved in the gameplay.

2.3.2 Annotation

Annotation is the process of adding additional information, mostly in the form of textual content coupled with pointers/markers to existing visual data. Most people are aware of this from instruction manuals found in user guides of various items bought, textbooks explaining new concepts with diagrams and such. As many individuals are habituated to seeing visual data, the demand for annotations to accompany any 3D data is a natural progression. This allows the viewer to reorganise available information clearly at a later instance as shown below in Figure 6.

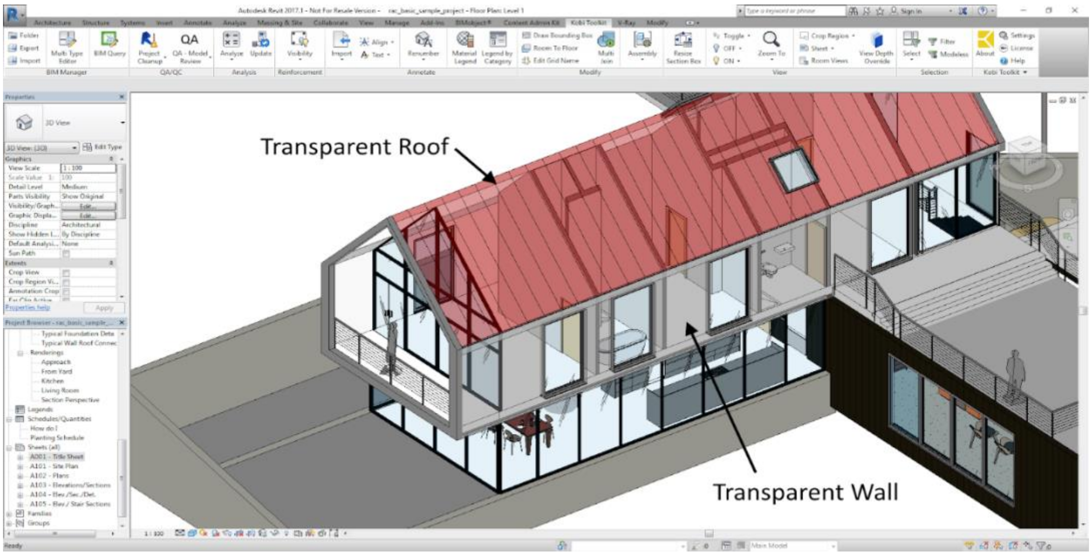


Figure 6. *Annotated 3D Model*

Also, while trying to decentralise work in industries such as construction, renovation, demolition, or planning, site visits are often accompanied by long reports and photos. Replacing these with videos that permit annotations makes the process faster and easier to understand.

360 videos are used to capture an area and then annotations are added to it. This process is relatively easy as the technology to do the same is easily available. This is primarily because while developing the 360-viewing technology, one major bottleneck faced was the difficulty in navigation, which was solved by the jump motion, where upon selecting a particular point in the view, the user could jump to that point and thereby navigate around as shown below in Figure 7.

Figure 7. *An example of a 360 video annotation*

This enabled different points to be marked or be separated from the rest of the data, and the same technology was easily extended to make it possible to annotate the captured 360 image. Alternatively, few companies have adopted another approach of using time frames of a 360 video as reference to navigate. For example, Finwe Ltd. has LiveSync Platform where a tablet app can be used to add annotations to a 360 video captured at site and later visualise on multiple devices. Kii has been working on a project wherein annotated content in 360 videos can be magically seen onsite using smart glass. Figure 8 shows the captured path of 360 video and also its view at the instance of time where the mouse cursor is pointing on the path. This work has been technically made possible by calculating roll, pitch and yaw of the annotated content and then using this data, throwing virtual raycasts from smart glass in the real world to find the target for showing annotation data. Also, now more options are being explored to add annotations more easily, like the use of a masking video and addition of annotations in subtitle like formats [25].

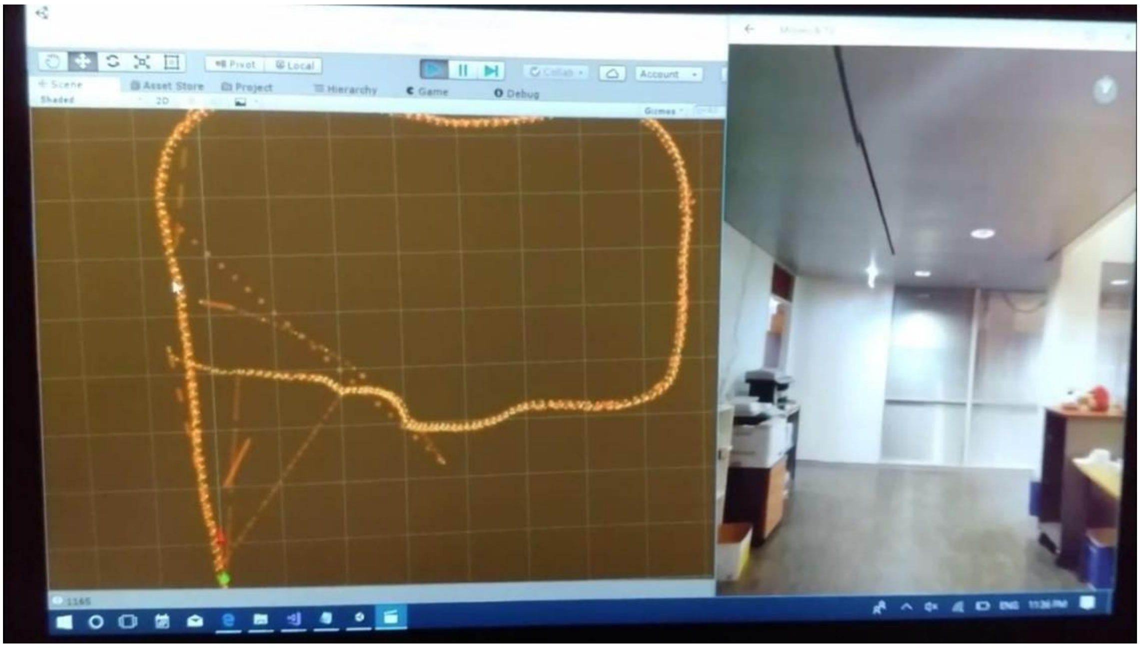


Figure 8. *Camera path of captured 360 video and the 360 view w.r.t the mouse pointer*

BIM models are building information models. These models are not only a 3-dimensional architectural representation of a building, but they also contain additional information like the structural design, the electric lines, the ventilation design, plumbing design etc. All this data is available in layers on top of each other, making it possible to get a holistic view of the entire building plan. It integrates different drawings into one. The diagram also has a lot of data that is edited at different stages of the building's lifecycle, by the various groups using the same source of data. This makes coordination and communication easy. As one can guess BIMS also come with their share of textual data to add more context to the different models.

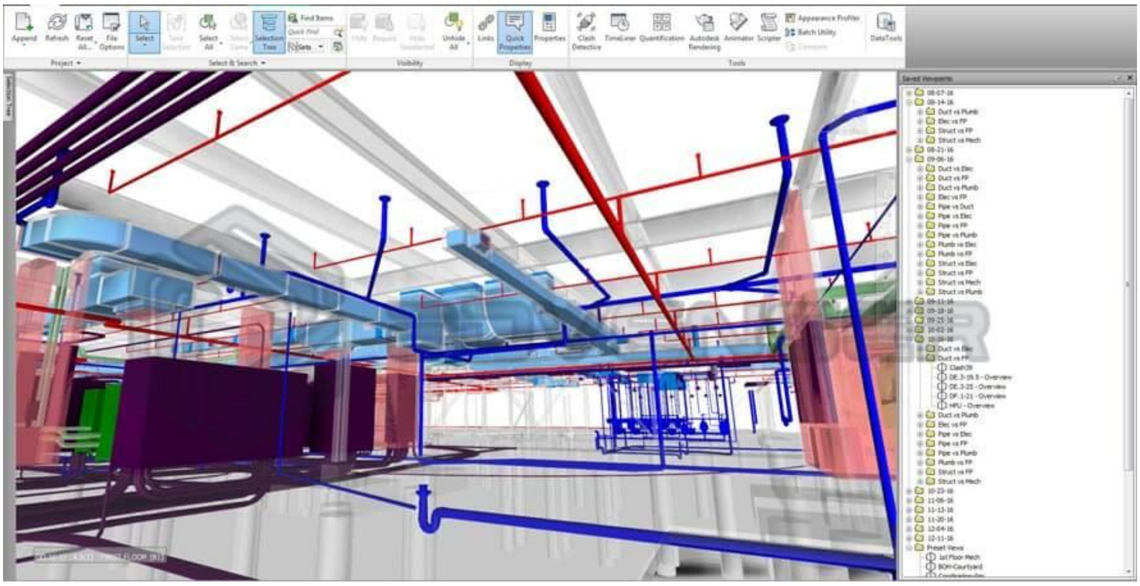


Figure 9. *Building Information Modelling (BIM)*

These are the most popular methods employed, depending on the application. Using Augmented Reality tools via smartphones to understand the meaning behind symbols or text is gaining popularity. This option is extensively popular as the cameras in smartphones are very good and make it easy to run image processing algorithms on cloud, just displaying the results on the device. Also on smartphones, using different languages is not a challenge and usability is easy for most users as they have had some experience with smartphones.

The greatest problem in adding annotations to the real world comes in locating the exact point in space where the information needs to be added. Most of the research or methods used to realise this require multiple planar surfaces to be identified to understand the real world.

2.3.3 Measurement

It is needed to measure the distance between different points in space. The scale of measurement varies from application to application in different industries. In this thesis the measurements are considered between the range of 20 cm to 100 m, typically used in real life macro measurements.

The most primary methods of measuring length were to compare the unknown length against standardised reference lengths. This method gave rise to various simple physical objects used as the reference length such as long bars or scales, measuring tapes, measuring chains to name a few. Some of these tools, like the scale and tape are still extensively used in the industry. As the need to measure longer distances arose, carrying bulky chains became difficult. Soon the measuring wheel gained popularity for its simple design and added mobility. Other methods like stadimeter which use optics to measure height became popular for their small size. These methods depend on analog computing and physical comparison. However, in today's day and age there is a variety of digital methods that have been developed to measure distance.

The virtual scale is a very common example of digitising an age-old method of length measurement by comparison to a reference length. Here a device like the mobile phone or laptop screen show a scale, with markings like a regular ruler, which are created depending on the resolution and size of the display. Such rulers are available on a number of websites but are not practical for industry use. The industry most commonly relies on Time of Flight (ToF) methods like laser measurement or IR measurement [26].

Laser measurement involves use of a laser beam and detector. The laser beam is released in pulses towards the target which gets reflects it. This reflected beam is detected, and the time taken for the beam to return is measured. Using the principal of time of flight, the distance of the target can be measured. The accuracy of this measurement depends on the type of laser being used, the sensitivity of the detector and the clocking system. These detectors are usually small in size and have a digital readout to provide the measurement in different units of length [27].

Table 2: List of major companies that provide laser measurement hardware

Product	Range	Accuracy	Remarks
Leica DISTO D2	300 ft	1/16 inch	Automatic end pipe for hard to reach corners
DEWALT DW03050	165 ft	1/16 inch	Indirect measurement using Pythagoras
Bosch GLM 35	120 ft	1/16 inch	Continuous measurement,
Tuirel T100	100 mtr	2 mm	Length width height and clearance measurement
Suaoki S9	198 ft	1/16 inch	Change the reference points for difficult spaces
Tacklife Advanced	131 ft	1/16 inch	Measures area, distance, length and volume
Eastvalley Laser Measuring Tool	328 ft	1/13 inch	Has autocorrect and error report technology

IR based measurement tools are also available. Their principle of operation is the same but in place of a beam of laser, IR radiation is used. IR and UV based measurement are also used in closed spaces like tanks to measure volume and change in volume.

Using smartphones to measure distances more accurately has been made possible by Apple Inc. and Google LLC. They have brought forward ARKit and ARCore, platforms that are used to build augmented reality applications on smartphones. These platforms use photogrammetry along with the motion and orientation sensors of a smartphone to understand the surroundings, recognise surfaces, identify unique points and provide this data to different apps. Using this information, apps like MeasureIT, AR MeasureKit, Air Measure, ARCore Measure can measure the distance between two points with a smartphone.

Model based measurement is another method of measuring sizes and distances completely digitally. In this method, a model of the actual space is generated by scanning the space and then the required measurements are made from the model rendered from the scans. This also makes it possible to integrate different models into one another, like a model of a table inside a room. This makes it possible to make remote measurements between items without having to physically place them. software like SketchUp, Rhinoceros3D, Autodesks 3ds Max, Vertex BD are used to perform such tasks [28][29].

Measurement tools also come in handy when planning, for example, large conference spaces to maximise participant interaction or a school to promote a healthy atmosphere for children to learn, alike. As mentioned earlier, the most popular method of measuring distances or lengths is the tape. However, it is a traditional method which creates no digital records. This creates the need for the information to be fed into a system for later extraction. Also, the measurement is limited by the size of the scale, and larger lengths require heavy and bulky scales.

The measurement wheel, be it traditional or digital, requires the user to physically run the ruler wheel along the path that needs to be measured. This is an advantage while measuring the lengths of curves or non-uniform surfaces, but it also increases the chances of inaccuracies while measuring straight lines as the line followed may not be always orthogonal.

Laser measurement is one of the most popular methods of digital measurements. It however has some limitations. It needs the start and end points to be in line of site. Another problem faced arises because of manual error, when the user does not hold the device parallel to the axis that needs measurement. The laser measurement devices have proved to be most popular for the floor to ceiling measurements and to check perpendicularity of walls. They are particularly useless in the case of measuring curved surfaces or arches.

Figure 10. *Laser measurement*

The model-based method of measurement is dependent on the accuracy of the scanned model and application used to render it. Also, the processing capacity of hardware used in rendering the model may affect the accuracy of measurement. This method provides easy measurement once the site visit is completed as it usually needs heavy computing, making direct measurements on site difficult and sometimes impossible. As measurement is not online, it may make related work like estimating volume of debris in construction industry or measuring possible cable length in cable laying operations almost impossible.

Another problem often faced by the industry is to make measurements in areas that are difficult to reach because of their shape or because they are in a hazardous environment. In hazardous environments, to make measurements with hardware that requires installation once the area has been sealed off is an expensive procedure as it needs the area to be made inert before any installation procedure can take [25]. The solution used in such cases are Intrinsically Safe (IS) mobile or handheld devices based on non-contact methods like model-based measurement or ToF.

2.3.4 3D Model Interaction

Scale, Rotate and Translate are the three primary functionalities which are used to create an interactive experience when any 3D models are viewed. Depending upon the device used to render and view the model, the interaction may change, however these primary functions are always observed.

Scale

Scale is a property used when visualising images or models. It changes the scale of the item is being represented with respect to its actual size. Most applications that have a visual output on a digital screen allow some form of scaling to happen. On computer screens this is done using the zoom function, which is a representation one ‘moving closer or away’ from the screen. Similar action can be done on a phone screen, generally using the pinch to zoom function or double tap function.

Other applications like image editors and viewers zoom into or out of a picture by changing the pixel density. This is a method of digital zoom and in practice does not scale the image. This is because the true image is not changed, only the way it is viewed is changed.

When a 3D model is scaled, its true size is changed with respect to the environment. It effectively means the models now occupies different area or volume in the same scene and is not being looked from closer or further away. It can be thought of as the model having grown or shrunk. This can be done in a number of software like unity, unreal, AutoCAD.

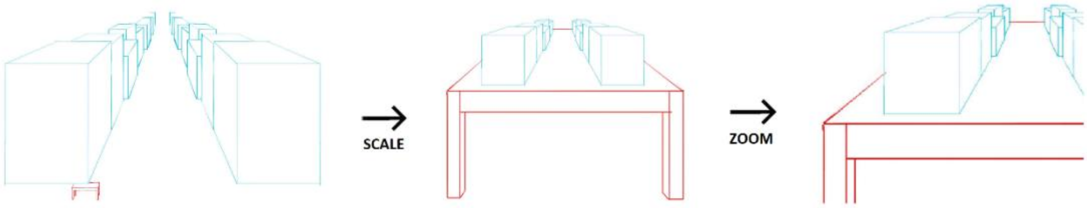


Figure 11. *Zoom versus Scale*

This property is extremely useful when the visualisation area available is small, but the actual model is very large. A classic example of it is making a map, which is a scaled representation of a much larger area. Scaling also helps to change the perception, when an object is seen in a scale much larger than its true scale, a lot more detail about the same can be understood. This is done very often in trainings, to explain more complex and intricate workings. This also helps to keep focus and remove unwanted information. Viewing models in their true scale helps understand how they would be in the real environment without carrying the physical product with them. This proves very helpful in marketing and design applications, to see and show the customers what the new installation would look like.

One major problem in changing the scale is that it is heavy on computation. It requires the object to be segregated from the environment and then be changed. This is not something a lot of visual rendering software can do. The UX for scale on a smartphone or similar device is not very easy to establish as a lot of users can confuse it with a more commonly known function of zoom. If the visualising area is small, it becomes even more difficult to see the changes caused by such a function.

Rotation

It is used to change the orientation of the object being rotated with respect to the viewer. In most cases, only the object i.e. the 3D model is rotated. However, in certain cases the environment can also be rotated along with the object, effectively providing the user a different point of view of the same scene. This section discusses cases where just 3D models are rotated and not the entire concept broadly.

Various software on desktops, laptops, phones that can render the models and scenes that are to be viewed. Different UX are designed depending on the device being used. On a computer it may be a click and drag by the mouse once the rotate mode is selected or by clicking a button to rotate about different axis. On a phone the same may be achieved by tilting the phone or by tapping on the object and dragging the finger.

In mixed reality, rotation helps to view a model from more points of view without the person having to move himself. This enables better, more detailed understanding of the model and the environment. Education, design, planning, marketing need this tool more often than thought of.

The major issues faced are in the interaction which is usually multistep and cumbersome. Also, to rotate a model about multiple axes requires some computational power and is not supported by all applications.

Translation

This feature lets the user move the model in the space available in the scene. It can help simulate movement, place the model in different locations and help create a realistic representation of the space. It also makes it possible to view the model clearly, by moving it around, like one would inspect an item in the real world.

Various software on desktops, laptops, smartphones can render the models and scenes. Different UX are designed depending on the device being used. On a computer it may be a click and drag by the mouse once the translate mode is selected or by clicking a button to move along different axes. On a phone the same may be achieved by tapping on the object and dragging the finger.

Almost all fields that require models to be visualised would need this functionality, as it allows for better visualisation. It is useful in modelling, visualisation, placement, layout designing, marketing, etc. In training applications, different parts of a machine can be removed or assembled using this feature, where the trainee can move around the parts to understand the machine and learn how it works.

Similar to rotation, the interaction may not be smooth especially on small screens. In software that are not very powerful in computing, such basic features give rise to errors where in one object passes through or hides inside another.

Depth perception also becomes a problem as one may find it difficult to estimate what distance the object is from the rendering camera and thereby create a confusion between a moving away and moving up transition. Also, some applications are not capable of providing movement along all axes. Though this may be useful in some cases (e.g. no upward movement to simulate gravity) it is usually causing a hindrance.

2.3.5 Mesh Manipulation

This feature allows an existing model to be sliced into pieces. This feature though is not very obvious in its use was requested by many customers. Slicing a 3D model can be done on specialised software like AutoCad or Unity, which have a strong engine that can calculate the result of such an action and a computer that is capable of supporting the same. It has to calculate the result of a plane intersecting through the model, not only its surface but also any internal details that may be present and create an accurate representation of the same.

This is useful in creating cleaner versions of scanned models. When a model is scanned from the environment, more often than not the scanner picks up parts of the surrounding environment into the 3D model. These parts have to be snipped off and hence such a tool is useful. Also this helps to understand what is inside an object without actually destroying it. However, owing to its need for high computational power, only desktop solutions are available for the same. This makes it not possible to clean, redundant data captured on site. It leads to longer time being spent on cleaning up the model than the time needed to capture it.

2.3.6 Mixed Reality Capture

This is a method of capturing augmented data along with its actual real environment. These features are those which help make the experience more universal. They play an important role in increasing the usability and adoption of new methods of working. This is because such features help better integration with existing frameworks and make the user feel more comfortable by familiarity.

Most solutions that can perform this use some form of 3D model overlaying techniques on a background of the real world and then make a 2D mapping of the same. In a pure augmented reality application like in a smartphone or smart glass, this is done using the camera and the screen. The camera helps understand the real world, the device helps compute the output of the augmented data, and the screen displays it to the user.

In other methods, image processing is done on a capture of the real world. Here the information to be shown over the real world is added later and the resultant is then provided. This helps store and share results that have been created by the user. It can serve as a record of what was envisioned during planning, explain to others what must be done while using a mixed reality application, and reproduce the visual information as seen through the smart glasses to other users.

On phones, any screen recording application when run while using the AR application can effectively produce a mixed reality capture. Most smart glasses have limited capabilities or no capabilities to capture such data. Some of the devices which do have shown stability issues. Further, smartphones tend to run slowly, or the screen painting starts to jump when screen recording apps are used. (Away from marker will not work)

2.3.7 Cloud Technologies

Upload

Uploading data to a server is an extremely important functionality in today's day and age. This is because once a computer can connect itself to a server, it can have access to a large amount of data that cannot be stored in the device's internal memory. Uploading data frees up internal memory, while making the data available for future use and to other devices that have authorised access to the server. Connecting a device to the internet would make the hardware perform new tasks by providing it with computational power of a cloud support system [30].

Some of the most popular forms of data that are uploaded in the industries that are under consideration are images and videos. These may include engineering and design drawings, photographs, layout designs, and how to use videos. Another major source of data is textual information in the form of a PDF file. These include a variety of information from contracts to manuals and a variety of documentation.

Apart from these very common forms of data, others include 3D models, location coordinates, status reports and special messages like alarms or alerts. Depending on the server being used, there are several APIs and support applications that can be used to ensure files are uploaded securely and safely. A variety of devices can upload their data to a server. One method used to achieve this is using RESTful services. Another method is using Publish - Subscribe methods with the help of a broker.

Some devices do not directly connect themselves into the network. Instead they use another device capable of such tasks, sending their data to the intermediate device, which in turn sends the data forward. A number of industries need data from different departments to be uploaded onto the same cloud platform. This enables cloud applications to integrate the complete data to create a realistic representation of the project or tasks being carried out.

The construction industry requires material status, personnel information, work status, schedules, documentation, plans, drawings, 3D models, BIM models to be uploaded and updated from time to time, sometimes on a daily basis. Most of this work is done on a desktop or laptop computer, usually from the comfort of an office. The process involves information being transferred from the site to the office, and from the office to the data storage platform.

Logistics deal with slightly different data. Their concerns include shipment details, location of materials, status of carriers, schedules of shipments and possible routes. The more often this data is made available to the routing applications, more efficient can the routing be made. In recent years the importance of tracking the fleet to improve efficiency has been tremendous.

Interior designers and floor planners use software to plan the area and create models of what the final look would be. To share this information with clients and other third parties who do not have access to the software, they use walkthrough videos and representational photographs of the model, which are then sent to others for approval. Also, they have data about the site, like measurements, existing plans, customer needs, photographs etc. collected over multiple site visits which are manually uploaded and compiled, so as to be used over the period of the project. Layout planning and site planning also follows a similar execution strategy, but as the teams involved are usually larger, exchange of data becomes more important.

Remote collaborations are possible only when the data from one side is uploaded for the other to see. Mostly the data sent here is audio or visual information, with some text data being sent sometimes. Trainings and education rely highly on information exchange over different mediums and platforms.

With the fast advancements in networking, it is not very difficult to upload data in areas where a fast and reliable connection to the internet is established. Security, which once used to be a problem, is now not much of a concern thanks to the numerous advancements in cyber security. However, one problem that still remains in most industrial sites is the availability of hardware that is small enough to be comfortably carried to site, powerful enough to carry out the required tasks and then upload the data then and there, thus not only saving time, energy, improving efficiency, reducing chances of data leaks but also making practises like remote collaboration fast and meaningful.

Devices like Google Tango have shown some promise in scanning and creating 3D models that can be uploaded to the cloud. Most smartphones make it possible to upload visual, audio and textual data. However, to upload more specific file formats, smartphones may not always be the correct choice of device as they may corrupt the file which cannot be read.

Download

Similar to uploading data, it is a necessity to download data to access it in today's world. One cannot be contained just to the data available on one’s device, most of the information needed to carry out daily tasks in the industry today need a connection to the world wide web at least. Even to perform a task as trivial as checking emails requires data to be downloaded. Downloading data is a flip to uploading data and works mostly on the same technology. Different applications propose different methods to restrict data downloading as a method of security, to ensure no information reaches unauthorised personnel.

Downloading 3D models plays a very important role in designing and layout applications. It can help visualise how an item or machine may fit into the space. Logistics use models of the shipment to check for collision during loading and unloading activities or movement within closed spaces. Similar models are used in the training applications to explain to the trainees how the machines look and how to operate on them. Collaboration work relies highly on them, where the expert can show on a model how a task has to be carried out for reference. Also, downloading 3D models can help to compare what the planned output is, enabling workers and supervisors on site in construction and renovation projects to get a detailed idea as to what needs to be done. Product fits at site can be verified and confirmed using models of the product before placing an order for the same. Marketing and sales teams also find such data very useful.

Downloading data from sensors helps to monitor the environment in factories. This data when downloaded by supervisors and maintenance in-charge even remotely proves to be beneficial as they can monitor what is happening and make informed decisions when needed.

The roadblocks faced by devices to download data is that the hardware must be equipped with the correct applications and capabilities in terms of processing capabilities to read the downloaded data and present it in a usable form to the user. File formats should be compatible, and devices internal memory should be large enough to support files to be downloaded. The hardware should have the capability to connect to the internet, or to another device which can, from where the required data can be accessed.

3. METHODOLOGY

3.1 The research approach

The constructive research approach intends to solve real world problems by producing innovative construction thereby contributing to the theory of the discipline where it is applied. [31]. The main objective of this research approach is to identify and solve real functional issues.

Once a relevant problem is identified, a deep understanding of the topic area is obtained, both theoretically and practically. An innovative solution is then designed, implemented, and tested. The scope of the solution is studied further, along with its theoretical contribution. The constructed solution is operationalised to evaluate its workability and appropriateness [32].

Therefore, the constructive research approach is deemed suitable not only for developing solutions to the problems identified in the earlier chapters of this thesis, but also applicable for feasibility analysis and suggesting improvements.

3.2 Software Development Life-Cycle (SDLC)

There are several SDLC models such as the traditional Waterfall approach, the Spiral methodology, DevOps, V-Model, Prototyping and Agile. Due to the operational nature of this thesis, the Agile methodology is the most suitable approach in developing a software solution that involves a constructive research approach as it enables continuous improvement and constant collaboration at every stage along with ensuring the quality of the developed solution.

The Cambridge dictionary defines a prototype as an example of something, such as a product from which all later forms are developed. The Prototyping methodology enables the early construction of prototypes without full functionality. This enables gathering feedback through continuous testing. In this thesis, a methodology has been created by using Agile principles and Prototyping research method.

A prototype of the software application is created using the Agile methodology and verified through piloting in the industry. A wireframe could also be created simulating the User Interface (UI). However, at this state, verification of the features is deemed more important. Hence, improvements to the UI are outside the scope of this thesis and could be considered as a topic of further research. The desired output of this thesis is the development of digital tools which aid in interaction with the physical world for solving real world problems and to validate it through piloting and case study.

3.3 Research steps

3.3.1 State of art review

Chapter 1 introduces the processes and technologies used in the construction and manufacturing industries, for measuring, marking and for interacting with the site environment. It also introduces the challenges with understanding the physical environment due to the limited perspective of 2D diagrams.

The state of art, presented in Chapter 2, explains the terminologies, such as annotation and measurement, within the context of the applied environment. It also goes deeper into the limitations arising from the usage of the currently employed methods. Further, it also explains the relevant technologies currently in use like 3D scanning, laser measurement and its associated uses and limitations [49]. Finally, it delves into the existing capabilities of smart glasses and how it can benefit the industry.

Onsite training is proven to be more effective than classroom teaching. This is due to the fact that humans have a tendency to learn faster in a 3D environment that is close to reality. A detailed study of the 3D model interaction, in the state of art, helped to realise the advantages of learning by using 3D models and various other use cases.

In the literature review, an initial study has been carried out to gain better understanding about the different methods of working in various industries like construction, demolition, renovation, automobile, logistic planning, education and training, interior design, factory layout planning, real estate planning and ship building amongst others.

3.3.2 Problem investigation and user requirement analysis

This chapter provides a description of work processes, the methods employed to carry them out, areas of their application and the usual problems faced while carrying them out. Below table shows basic description of the existing process and features. The proposed solution along with the verification process is also listed below. These are compiled based on the information provided in the previous chapters.

Table 3: Problem exploration and Requirement analysis

	Task	Purpose	Current solution	Problem	Proposed solution	Verification
Measurement	Measurement of volume, area and distances between objects	To check if objects can fit in the location measure debris	Manual or laser measurement, gathering information from 2D plans	Outdated or incomplete information in the 2D plans, Additional travel.	Generate a 3D model of the object and place it virtually in the desired space after indoor scanning	Piloting on site
Annotation	Annotating, in the form of textual information and labels	Denote physical attributes, share information between personnel/departments	Printed and pasted on information boards, doors etc Physical marking software tools for placing annotations	Confusion in the meanings of symbols. Paper and data may be lost Software tools, used for annotation, are cumbersome and not user-friendly	A tool that supports 3D scanning, slicing, annotation and measurement in mixed reality	Piloting on site
Measurement in curved surfaces	Measurement in curved or difficult to reach surfaces	Collect data and save them	Laser measurement	Inaccurate data High probability of manual error	Model-based measurement	Piloting on site
Scaling	Scale an image	Visualise an image in a larger size	Use zoom function	The actual image is not modified but the change is in only the way it is viewed	Develop and use scale function	Manual software testing
Rotation	Rotate a model	To view the object from a different point of view within the same scene	Various software on desktops and mobile devices provide this feature	A multistep and cumbersome interaction is needed. Also, many applications do not support it.	Develop a powerful tool to rotate a 3D model on multiple axes	Manual software testing
Translation	Move the model	To inspect the item as in the real world	Various software on desktops and mobile devices provide this feature	The interaction is cumbersome and there may be problems with depth perception	An user-friendly tool is developed that enables smooth translation of models within the space	Manual software testing
Mesh manipulation	Slice an existing model	Models scanned from the environment may have parts of the surrounding environment attached	Slicing a 3D model is possible with specialised software	Only desktops can run specialised software with high computation power	A tool with capability to slice models is developed	Manual software testing
Mixed reality capture	Capture augmented data within real environment	Test the usability of a model within an environment	Smartphones with screen recording and AR applications can produce mixed reality capture.	Most smartphones run slow and some have stability issues. Most smart glasses have limited or no capabilities to capture mixed reality data.	A stable tool that can be used with smart glasses is developed	Manual software testing

The above table laid out the necessary features of the software solution and the requirements of the hardware. The work processes which have the potential to be carried out by using smart glasses were identified. It was observed that the use cases required to execute them were quite similar. Additionally, the smart glass capabilities which are required for enabling these processes were found to have striking similarities.

It is concluded that the features should be developed in an app-based framework in a hands free wearable device. Using smart glass serves as a user-friendly platform through which 3D information and tools can be accessed in a mixed reality environment to interact with a 3D world and also be connected to the Internet of Things (IoT) environment [33]. Hence, the smart glass is chosen as the most suitable mobile and wearable device onto which the features can be implemented.

3.3.3 Proposed methodology

The following steps are proposed based on the research methodology:

1. Evaluation and selection of suitable smart glass
2. Setting the development environment
3. Implementation of features
4. Verification through piloting
5. Suggestions for improvement

The next chapter details the implementation of the proposed research methodology.

4. EXECUTION

This thesis studies the problems faced by industries regarding site measurements and processes, and based on the state of art and evaluation of various smart glasses, a solution consisting of digital tools, is implemented in the form of Artificial Reality toolkit for Microsoft HoloLens that was developed using the Unity3D game engine which was tested and piloted at Fortum.

4.1 Smart glasses evaluation

Smart glasses can help overcome most of the problems described in chapter 2 by creating powerful software applications. HoloLens are a popular smart glass developed by Microsoft. The research carried out by the Industrial Design Dept of the University of Lapland [34] showed that factory workers have a very positive response towards the use of smart glasses on the factory floor. This is because the required information is available at the location, and it keeps the factory workers hands free to perform other tasks.

A study into the possible hardware that could be used was carried out to understand which device would prove to be the best fit to solve most industry problems. The analysis was done based on ergonomics, capabilities, software support and price.

A concise version of the same analysis is presented below:

Table 4: List of high-end AR smart glasses

Name	FOV (degrees)	Price	Sensors	Cameras	Connectivity	Battery	Tethering
Daqri	44° Diagonal FOV	4995$	Depth sensor camera	RGB 1080p HD Camera, 166° Diagonal Wide-Angle Fisheye Lens	Wifi, BLE	Built in lithium ion battery 5800 mAh	Not Required
HoloLens	About 30°	3000-5000 €	IMU, ambient light	Depth camera	Wifi, BLE, Micro-USB Speakers	2.5-5h	Not Required
Vuzix M300	20°	1,650 €	3-degree of freedom head tracking, 3 axis gyro, 3 axis accelerometer, 3 axis mag/integrated compass, GPS	13 megapixel stills 1080p video Image stabilized auto-focus Flash/scene illumination	microUSB 2.0 HS Wi-Fi b/g/n/ac – Dual-B 2.4/5 Ghz MIMO 2x2 BT 4.1/2.1+EDR	100 mha, options for up to 5000mha rechargeable battery packs , 2 – 12 hours	No
Athers	50°, Resolution 1800 x 720	4,000$	9 axis IMU: accelerometer, gyroscope, magnetometer	3D depth and dual 4 MP RGB cameras	WiFi, BT4.1, GPS, optional 4G LTE	3100 mAh, 8 hours	No. However a pocket PC is required
Meta2	90°, Resolution 2560*1440	950$	Sensor array for hand interactions and positional tracking.	720p front-facing camera	9 foot cable for video, data, and power (HDMI Version 1.4b)	NA	Yes. Requires a modern computer with Windows 8 or 10

HoloLens has two major advantages. Having been developed by a giant like Microsoft, the investment being pumped into this product is notably large, and further development is certainly on the cards. The second advantage is the ease of development. It uses tools like Visual Studio and Unity Game Engine, where the number of developers already comfortable on these platforms are huge. Further, Microsoft has been heavily working to provide a powerful development kit, so the app development is easier. Also, it becomes relatively easy to leverage existing projects that have been developed on Unity, even if they were targeted for different devices. This is the reason HoloLens was selected to be the device upon which the development was to be carried out.

4.2 HoloLens in detail

Being one of the most versatile and easy to use of the available smart glasses, and with no loss of speed shown towards its future development, HoloLens was picked as a preferred choice. The hardware, weighing a little over half a kilogram, comes packed with a variety of sensors and computational units. They include the following [35]:

Sensors

- 1 IMU
- 4 environment understanding cameras
- 1 depth camera
- 1 2MP photo / HD video camera
- Mixed reality capture
- 4 microphones
- 1 ambient light sensor

Human Understanding

- Spatial sound
- Gaze tracking
- Gesture input
- Voice support

Processors

- Intel 32 bit architecture with TPM 2.0 support
- Custom-built Microsoft Holographic Processing Unit (HPU 1.0)

Memory

- 64GB Flash
- 2GB RAM

HoloLens provides a very good experience to the viewer where he can see 3D objects added to the environment, he is present in, as opposed to him being transformed into a different environment like in any Virtual Reality Headset like the HTC Vive or the OculusRift. This is made possible by using its complex lens system for which Microsoft has also applied to a few patents. The essential way by which a HoloLens can control the amount of light entering it is by using a dimming module. It enables various levels of dimming, which is used during initial contrast setup, and helps in adjusting the brightness with respect to ambient lighting. This works by sandwiching a monochromic electrochromic cell compound and an insulator between the lens system and connected to a conductor. To control the amount of light being passed through, the amount of current being passed through the conductor needs to vary [36].

The lens system in the HoloLens plays a very important role in creating 3D holograms. They are made by 3 layers of glass, one each of red, blue, and green, for each eye. A small projection unit on top of these glass layers projects light onto them. This light then gets scattered by the tiny grooves on these layers, making the shapes visible to our eye. A combination of these images causes the brain to perceive 3D images or holograms at different virtual distances [37].

To ensure that the image quality is not reduced, and the layers get the correct image data, HoloLens uses a special mechanism that is able to have the image output of the exit pupil (a beam of light having least cross-section while containing complete image data) outside the projection system. This image is then coupled to the glasses, thereby providing better image quality as against a more common system where the image is decoupled using extra optics after the exit pupil, which results in reduced image light strength by the time it would reach the glasses [38].

Once the hologram is visible to the user, he or she can place it at a certain virtual distance in the scene. The HoloLens sensors measure the user's position relative to the room and calculate the position and the angle of the hologram from him/her and project a suitable image, enabling the user to walk around, under or even through a hologram to better experience it [37].

The sensor that primarily makes this possible is the depth sensor. The depth sensor used here is a Time of Flight based IR sensor, much like the Kinect V2 sensor. However, it is important to note that the size and power requirement have been greatly reduced [39][40]. So, the HoloLens depth sensor shoots very small light bursts into the scene at intervals of time to measure the depth. However, it is believed that the power saving as compared to the Kinect V2 is achieved by selecting a different frequency than Kinect.

The exact frequency has not been disclosed by Microsoft yet. To add to this, HoloLens also relies on its environment sensing cameras, which capture images of the surroundings, and these help the HoloLens to understand its position and location better. Hence Microsoft warns users beforehand that the experience may not be very good in a dark room because theses environment cameras are unable to pick up sufficient data though the IR camera would face no problems [39][41][42].

The data thus collected is mapped into a grid of x and y pixels, each having a value of corresponding to depth d from the HoloLens. Each frame a new depth map is computed and is compared to the previous depth maps. Thus, it becomes possible to find a set of pixels that are not static and have a very low depth (usually less than one meter). These pixels are clubbed together to be considered as the cursor position of a gesture. The HoloLens has gestured input method known as airtap which is similar in functionality to a mouse click on the mouse or trackpad or a touch input on a touch screen. It works by first keeping the index finger and thumb finger apart with the remaining fingers closed, followed by pinching the index and thumb fingers together to register the tap. Therefore the gestures in HoloLens have a start and final position. Similarly, it has a menu gesture that includes opening your fingers like a flower, equivalent to pressing the windows button on a pc, and a pinch and drag option that is a continuous gesture like a click and drag or swipe feature we are aware of. It also accepts inputs from a clicker, which some users may find easy [43][44].

4.3 Development environment

Different software tools can be used to develop augmented reality solutions. The two major game engines supporting AR development for cross-platform development are Unity3d and Unreal Engine. Alternatively, for mobile devices, Google and Apple has also provided native support for AR development on Android and iOS respectively. There are not yet built any standards for Augmented Reality and that also gives a possibility of new frameworks/ engine being developed currently. Some of them include AR.js, WebARonARCore, WebARonARKit. WebXR Viewer by Mozilla.

Unity Engine and Unreal Engine are clearly ahead among other platforms, reason being that others are either on an experimental level or are not cross-platform. The two are closely competing technologies having their own pros and cons as shown in Table 5. Kii had chosen Unity Engine for development of the industrial applications. One of the major reasons is that Unity supports a huge range of devices and all the major AR platforms have been having Unity support. Although unreal engine is known for its high-end graphics, Unity Engine has an edge over Unreal engine in Augmented Reality. The pros of Unreal Engine are subdued by the fact that most AR devices including MS HoloLens, Daqri, Vuzix etc do not have a powerful processor and in cases like HoloLens, Unreal engine does not support its development yet. However, in Virtual Reality, Unreal Engine may be a preferred engine in the use cases where crystal clear graphics are required.

Table 5: Unity Engine vs Unreal Engine for Industries

Unity Engine	**Unreal engine 4**
Cross-platform engine supporting 21+ platforms	Supports mainly PC and console games
C#, UnityScript, Boo	C++
Community and official Support	Tutorials mostly for designers
Better option for beginners	Only for Pro
Designer and developer friendly	Designer friendly only
Detailed documentation	Basic documentation
Good Performance for 2D and 3D apps	Better powered for 3D games

4.4 Features implementation

Once the hardware was decided, the development environment was fixed, the problems were present, the solution had to be developed. Each problem was tackled differently, using the capabilities of the device, while keeping in mind its limitations, the solutions for each of the features started taking shape. This section describes the solutions, the main constituents of it using some code snippets and images.

4.4.1 Scanning

The scanning process by HoloLens is performed by the depth sensors and the colour cameras. The depth cameras provide the distance information, and the colour camera adds precision by edge detection. When a HoloLens device is switched on in a new environment, the HoloLens scans the environment and saves it in the device. Microsoft refers to this scanned mesh as spaces. In case, HoloLens has been connected to a WIFI, it maps the saved scan to the wi-fi existing at that place. This saved mesh is used as a reference so that the device can realise its location in a particular environment by comparing what it is currently seeing with the saved mesh. Hence HoloLens naturally operates with the help of scanning. However, it does not let the user access this mesh. Though, there is a manual method to extract a scanned mesh using web-based Windows developer portal, but it does not allow any control on how to scan the mesh and do any manipulation.

This led Kii to explore the ways how the device can be used to scan and convert it into a 3D format recognizable by design softwares. Microsoft provides two types of API's for understanding spaces: LowLevel API (LLAPI) and High Level API (HLAPI). HLAPI is used to extract planes and other shapes from the environment. For example, it can help to recognise the floor, wall and even the shape of a sofa. LLAPI on the other hand provides a granular control over the environment under consideration.

Algorithm

The algorithm consists of the following steps:

1. Using LLAPI to access the scanned surface segments
2. Extracting the vertices, normals and faces when the application is running
3. Converting it to a JSON format with a defined format
4. Using REST API to POST the json to Kii's cloud
5. A web server capable of converting json files stored at Kii's server into OBJ files.
6. HTML provides a button to enable downloading the mesh in OBJ format.

Roadblocks

One of the major problems faced during this overall process has been to extract the 3D data of the mesh being scanned. This is because LLAPI does not give direct access to the complete mesh. The reason is that Unity cannot process 3D models with more than 65000 vertices. Secondly, HoloLens converts the whole environment into segments. A small to medium sized room, when scanned can have approximately 30-60 segments of mesh. The segmentation done helps to optimise the operations as at any instance of time, the whole mesh is not always needed. Studying the different classes exposed by LLAPI helped to understand the issue better and debugging was done through several iterations by visualising the solution in HoloLens.

Another major issue was to convert the JSON file containing just the 3D coordinate data into a standard 3D format. For this, a thorough study was done to understand 3D formats and libraries which perform format conversions. There was no standard solution available. While studying the open source javascript library Three.js, it helped a deeper understanding on how formats are converted, and this facilitated the generation of algorithm for conversion of JSON into OBJ format.

Solution

The work resulted in a solution where one can scan its surroundings in multiple resolutions as low, medium and high. The data was transferred in a secure JSON format to the cloud. This also provides the possibility in future to integrate the scanned mesh with other 3D and 2D data like existing 3D models placed in the environment, annotations etc. There are several advantages of generating a scanned 3D model. During any site visit, there is a need to capture the details of the site, so that it can be later used to do planning and design based on the site. For example, one needs to do offline measurement and currently there is no efficient way to do that. Scanning makes it possible as all the dimensions of the site are well preserved. This helps to make better customised installation, planning, design and reduces chance of error.

4.4.2 Measurement

Fortum has been testing measurement tools developed in 2 power plants and done benchmarking with laser measurement tools of range 40 metres. It has been observed that accuracy of HoloLens measurement application is approximately the same as that of laser measurement tools. Further, Fortum has been able to measure up to 70 m by using the developed application in multiple sites.

Various new benefits were realised while using HoloLens for site measurement:

1. It is possible to measure between real and virtual objects that are placed in the scene.
2. It is possible to measure between curved surfaces.
3. Multi point measurement is also possible with minimum user interaction.
4. Area measurement can also be achieved.
5. Offline measurement is possible.
6. No line of sight is required. Hence it is useful for sites with debris.

The table below lists the measurement readings using HoloLens application, under varying lighting conditions and for different distances.

Table 6: Measurement table with readings and error

1.5 metre

Sr No	Lighting	Surface	Reading	Error (mm)
1	Dark	Floor	1.489386	10.6
2	Dark	Floor	1.512694	12.7
3	Dark	Floor	1.495897	4.1
4	Dark	Floor	1.503399	3.4
5	Dark	Wall	1.540832	40.8

6	Dark	Wall	1.449931	50.1
7	Dark	Wall	1.485218	14.8
8	Dark	Wall	1.48335	16.7
9	Dark	Other	1.51539	15.4
10	Dark	Other	1.485901	14.1
11	Dark	Other	1.514734	14.7
12	Dark	Other	1.48578	14.2
13	Well lit	Floor	1.503242	3.2
14	Well lit	Floor	1.504747	4.7
15	Well lit	Floor	1.495644	4.4
16	Well lit	Floor	1.507931	7.9
17	Well lit	Wall	1.493343	6.7
18	Well lit	Wall	1.490672	9.3
19	Well lit	Wall	1.482227	17.8
20	Well lit	Wall	1.498051	1.9
21	Well lit	Other	1.49616	3.8
22	Well lit	Other	1.50587	5.9
23	Well lit	Other	1.49399	6.0
24	Well lit	Other	1.49805	1.9

3 metre

Sr No	Lighting	Surface	Reading	Error (mm)
1	Dark	Floor	2.995438	4.6
2	Dark	Floor	2.998473	1.5
3	Dark	Floor	3.00036	0.4
4	Dark	Floor	3.00059	0.6
5	Dark	Wall	2.984072	15.9
6	Dark	Wall	2.982108	17.9

7	Dark	Wall	2.970455	29.5
8	Dark	Wall	2.983206	16.8
9	Dark	Other	3.02668	26.7
10	Dark	Other	3.015813	15.8
11	Dark	Other	2.990445	9.6
12	Dark	Other	3.018194	18.2
13	Well lit	Floor	3.002176	2.2
14	Well lit	Floor	3.004222	4.2
15	Well lit	Floor	3.000997	1.0
16	Well lit	Floor	3.004135	4.1
17	Well lit	Wall	2.986997	13.0
18	Well lit	Wall	2.982259	17.7
19	Well lit	Wall	2.992783	7.2
20	Well lit	Wall	2.991853	8.1
21	Well lit	Other	1.496157	3.843
22	Well lit	Other	1.505873	5.873
23	Well lit	Other	1.493991	-6.009
24	Well lit	Other	1.498054	-1.946

4.4.3 Annotations

Annotations in the context of mixed reality, refers to an event where media information eg. text, picture or video is added to the real world when perceived from a human eye. Text is chosen as the type of media and this feature has been implemented by using Hololens development kit for various functions including the capability to use voice commands. An additional Unity plugin was also used for enabling the use of a keyboard in mixed reality. The feature to be able to add text annotations in the real world works based on the following workflow:

1. Write a text annotation using a keyboard in the mixed reality environment using gaze tracking to aim at the right character and click using air tap or button click by Holoclicker.
2. Give a voice command so that text gets copied and stays in the real world.
3. Walking to different points of interests and repeating the same process above.

Figure 12. *Mixed reality annotation tool*

This annotated environment can be later saved in a mixed reality capture video for future use. However, once the experience is created with multiple annotations in the real world, the annotation can only be visualised in the same session. In the current scope, once the application is closed, the textual information cannot be retrieved again.

In the future, this feature could be combined with the uploading feature so that the information can be stored. Further, for saving the experience in the mixed reality environment, there needs to be a marker for eg. a QR code placed in the physical environment and scanned, in order to save the data w.r.t. a fixed reference. This also enables downloading the annotations and bringing them back to the real world by loading the data using the same marker. Other than physical markers, virtual markers can also be used. With this, the feature can be further developed using spatial anchors in Holotoolkit which can allow the same annotations to be visible from different devices.

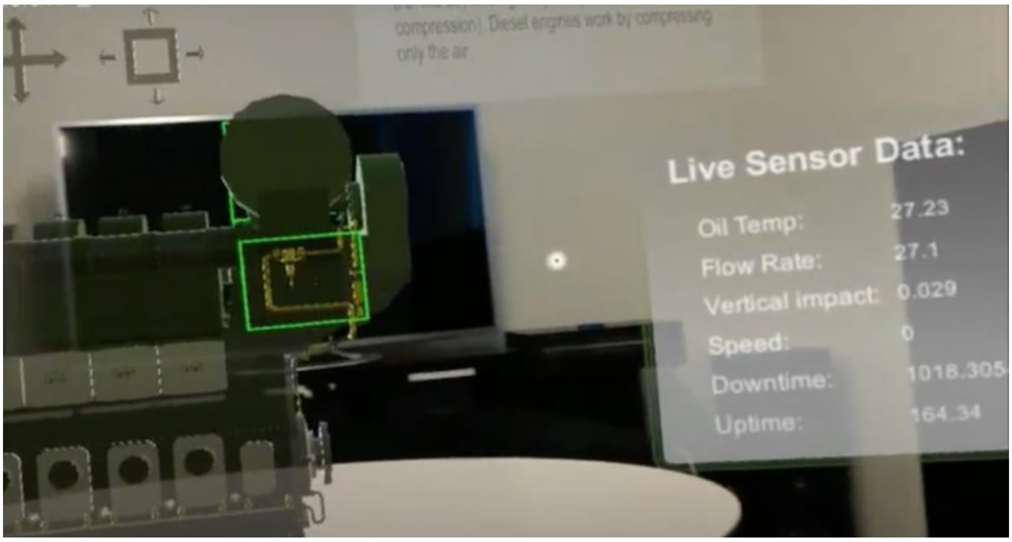
4.4.4 Upload

The feature was developed at Kii for the case study for uploading a scanned model to Kii cloud service. The scanned mesh was saved as vertices, faces and normals in JSON format and uploaded to Kii Cloud by using POST method (REST API). The POST method was exposed using Kii infrastructure. Later a simple HTML file was created to download the model in obj format. The recommended size of scanned mesh is 60,000 vertices based on recommendation by Unity Engine developer guide, even though the uploading feature itself has no defined limit. This limit allows the mesh to be downloaded again later and load properly in Unity based smart glass app without significant drops in frame rate of the app.

4.4.5 Download

Download feature allows retrieval of the information from the internet to the smart glass device for further use. This feature was developed for two scenarios:

- Download Scanned mesh - This feature allows users to download scanned mesh and visualise it in front of the user in a mixed reality environment.
- Download Sensor data - The use case for which this demo feature was created was to demonstrate real-time access of sensor information of machine part in a smart glass for a particular maintenance case. Main task was to access Live sensor data from ThingSee sensor accessible real-time from Kii cloud infrastructure. These were the values that could be measured real-time:
    a. Temperature
    b. Humidity
    c. Vertical impact
    d. Speed
    e. Latitude
    f. Longitude



Live Sensor Data:
Oil Temp: 27.23
Flow Rate: 27.1
Vertical impact: 0.029
Speed: 0
Downtime: 1018.305
Uptime: 164.34

Figure 13. *Live sensor data*

A text box encases the sensor data and shall appear by air tapping the green box as a trigger as indicated in above figure. For further development, the feature could be also extended to download any 3D design from a cloud library to visualise any 3D model.

4.4.6 Scale

The scaling feature is intended to change the size of the 3D model being used. This feature in any basic 3D modelling tool and is useful for end user in mixed reality experience to immerse oneself better by changing the scale. In the use case a diesel engine is being studied, there are two ways to scale:

- Voice actions - Use ‘Scale Up’ voice command for increasing the size by 100% and ‘Scale Down’ for vice versa. ‘Scale Reset’ voice command resets the default scale selected in the application.
- Scale button - Air tapping the Scale button in the menu triggers the diesel engine to scale in real life size.

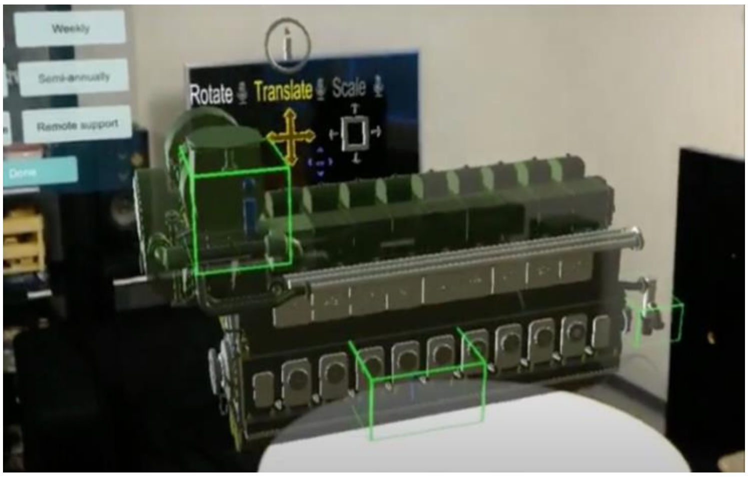


Figure 14. *Default scale*

Figure 15. *Real life-size scale*

4.4.7 Rotation and translation

The movement functionalities are among the primary functions needed in most solutions. It involves rotating or translating objects or models in the view. To achieve this functionality different approaches have been formulated, depending upon how the 3D model is placed. The first method developed is for when the 3D model is suspended in air.

Solution

Load application → Rectangular fitbox appears in the centre of the field of view. It is fixed relative to the camera of the unity engine. This mean s while wearing the smart glass if we are moving the fitbox shall also move with us. The 3D model under consideration appears in front of the user by doing an air tap after moving to the intended place. This is how a user can place a 3d model for the first time and later he can adjust its transform by using rotate and translate.

For Rotation of a 3D mode, the user needs to first select the rotate mode by air tapping the rotate button which is present in the menu above the initial location of the 3D model. On selection of rotate mode user hears a sound along with change in colour of the selected buttons as feedback.

The user uses gaze to interact with the hologram and the real world. Gaze refers to the centre of the field of view of user based on the orientation of the user's head. After this the user gazes at the 3D model and this triggers the onGaze function. This trigger serves as a feedback for the rotation function, and which provides access to the rotate function the 3D model/ game object. Whenever we use the tap and drag gesture the 3D model starts rotating. The drag has to be done towards the left or right side for enabling the 3D model to be rotated clockwise or counter clockwise about vertical y axis. The speed of rotation is proportional to the amplitude of the drag. The proportionality factor can be adjusted to change the sensitivity of rotation. The final angle of rotation is the same as the angle at which the drag motion was completed i.e., the pinch gesture was broken. The model can be rotated multiple times, in any direction as long as the rotate mode is selected, and the model is being gazed at.

On selection of translate mode user hears a sound along with change in colour of the selected buttons as feedback. The interactions in this mode are also similar to the rotate mode. Once the user gazes at the 3D model and performs the pinch and drag gesture, the model translates. The displacement is proportional to the drag vector calculated in the pinch and drag gesture action. The proportionality factor can be adjusted to adjust the sensitivity of the gesture. This makes it possible to translate the 3d model anywhere in the visible space. Also once the pinch gesture has been performed, the user can move without dragging. This causes the translate function to calculate the displacement vector like in a drag gesture and translate the 3D model correspondingly.

The second method is used when the model has to be placed on horizontal surfaces and can be extended to walls or ceiling. This method needs far less user interaction in terms of hand gestures and relies more heavily on the users gaze, making it more intuitive and easy to use.

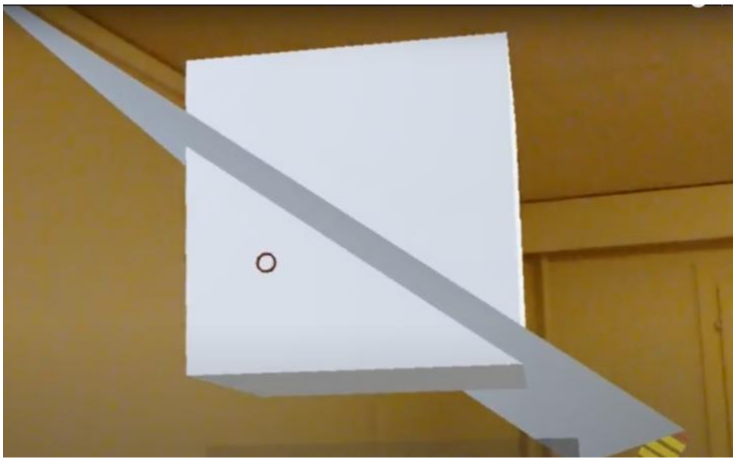
When we switch on the HoloLens and connect the device with WIFI, HoloLens starts 3D mapping the environment around us. The mesh details improve as we scan for longer lengths of time and as more area is discovered, the size of the mesh increases. The device links the resulting 3D mesh with the wi-fi and saves it as space. Microsoft Holotoolkit provides a high-level API that is capable of recognising various surfaces in the environment and categorising them as floor/ceiling/walls or horizontal surfaces. The application uses this high-level API to recognise floor and other horizontal surfaces.

On opening the app, the 3D model is placed on any horizontal surface present in the linked space that has an area large enough to accommodate the bottom surface of the collider of the 3D model. Once the users gaze intercepts the model, the onGaze function is triggered much like the previous case. Once triggered it waits for an air tap from the user. This gesture acts as the final trigger and causes the 3D model to ‘unlatch’ from the horizontal surface, float a few centimetres above it and follow the user’s gaze along the highest horizontal surface present below the user’s gaze. If the surface is large enough to accommodate the 3D model, the area where the model would be placed is highlighted in green to notify the user, and if not then it would be highlighted in red. The user can use this information to decide where the model should be placed. On finding a suitable place, the user can perform the air tap gesture again. This ‘latches’ or places the model on the highlighted surface.

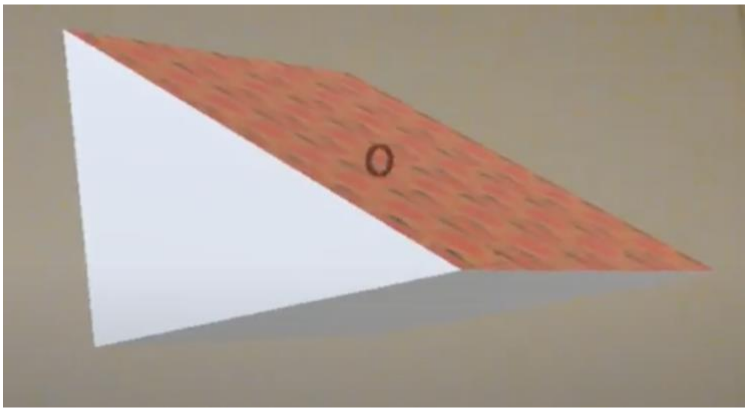
While the 3D model is following the user’s gaze, the user always sees the same face of the model as seen when the air tap to unlatch the model was performed. To rotate the model on the same spot with respect to its environment, the user has to simply walk around the model while looking at it. To purely translate it, the user can move sideways or walk up and down the area while holding his or her head still. Also moving his or head up and down causes an effect similar to walking up or down with the head being held still.

4.4.8 Slicing

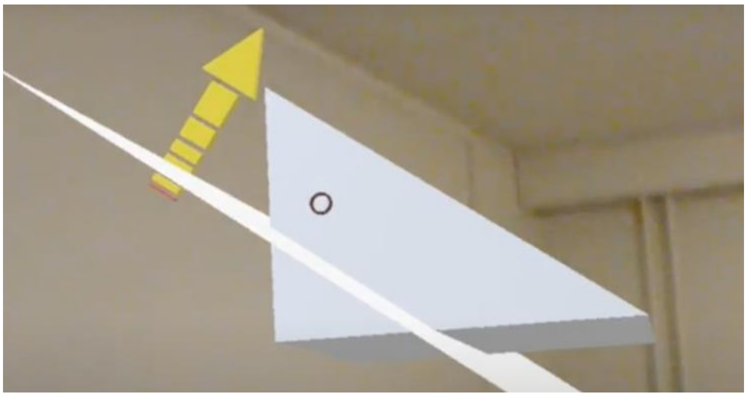
For development of slicing feature, a unity plugin was used based on which the 3D models could be trimmed using a slicing plane. The new solution calculates the intersection of the polygons and calculates the final mesh based on the slicing plane. The slicing tool thus developed could delete the meshes used for trimming scanned mesh and quickly extract a 3D model of a particular machine out of a scanned mesh. Below figures provide a demonstration how the tool thus developed could trim 3D models. For the intended use case, the slicing plane shall always appear in front of the camera and shall appear stationary with respect to the head, so for using this slicer, one has to move and aim where to place the intersecting plane inside the 3D model by placing the head in the appropriate orientation and then use a voice command saying “Slice”. This command shall trigger the intended function and using the same steps a proper model can be extracted from any scanned scene.



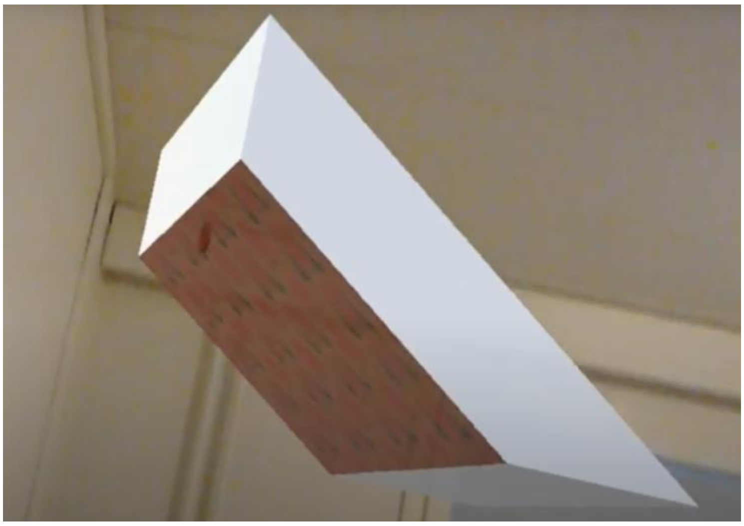
Step 1



Step 2



Step 3



Step 4

Figure 16. *Steps demonstrating AR slicing tool*

4.4.9 Mixed reality capture

This developed feature is greatly dependent on the device capability. Hololens provide a possibility to capture mixed reality capture at any point of time while you are using the Hololens. However, to create in-app experience, Camera Capture UI API was used to trigger beginning of the video capture combining both real environment as well as the 3D models spatially located in front of the camera. This feature can be hence used with other tools to save those experiences in the form of a video.

5. CASE STUDY – FORTUM

The idea behind this pilot project is to be able to scan, modify and upload the environment around us. It also gives the possibility to extract 3D models of machines / structures from the scanned environment. The application is made for Microsoft HoloLens. The experience is voice assisted as well as via Air-tap. One of the key features is that with just 4 taps, one can scan and upload the whole mesh which was just created. Further, HoloClicker could be used alternatively.

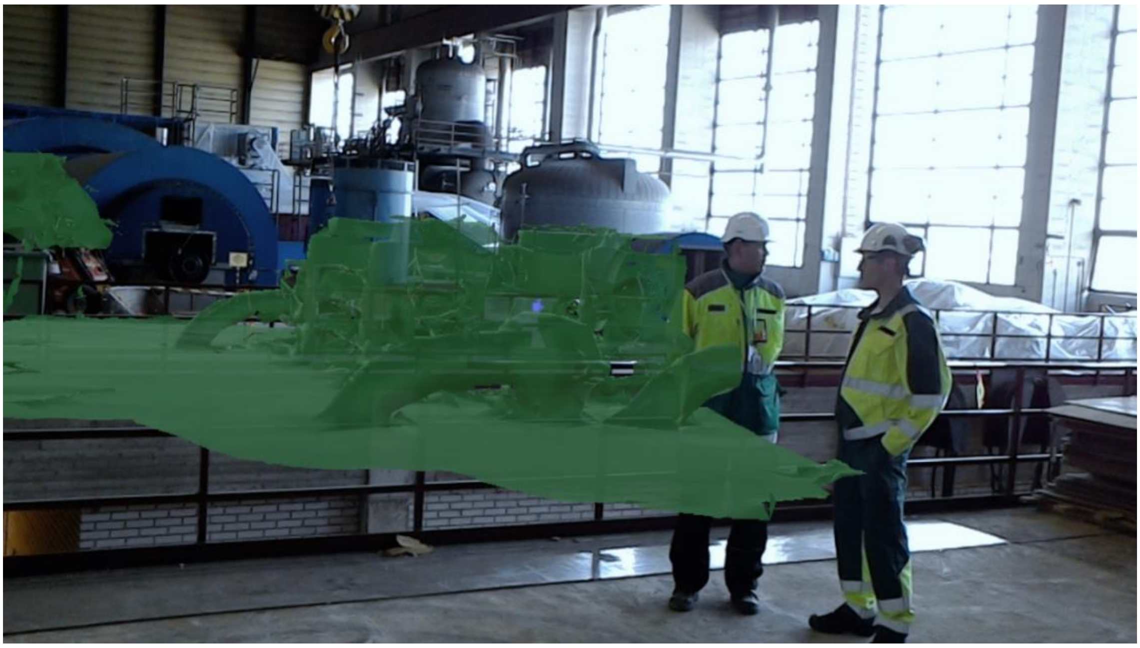
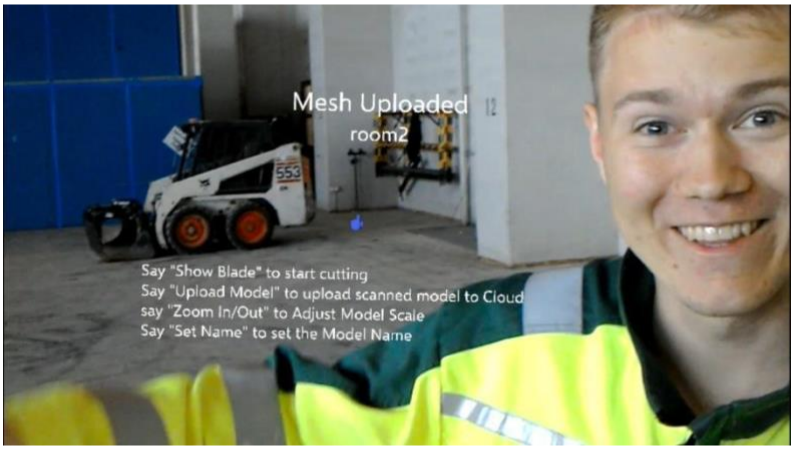


Figure 17. *Scanned mesh of power plant equipment at the demolition site*

Figure 18 shows a snapshot of the voice-based UI. The voice assisted experience is much more customisable and at any stage user is provided with instructions on what commands are applicable at that point of time.



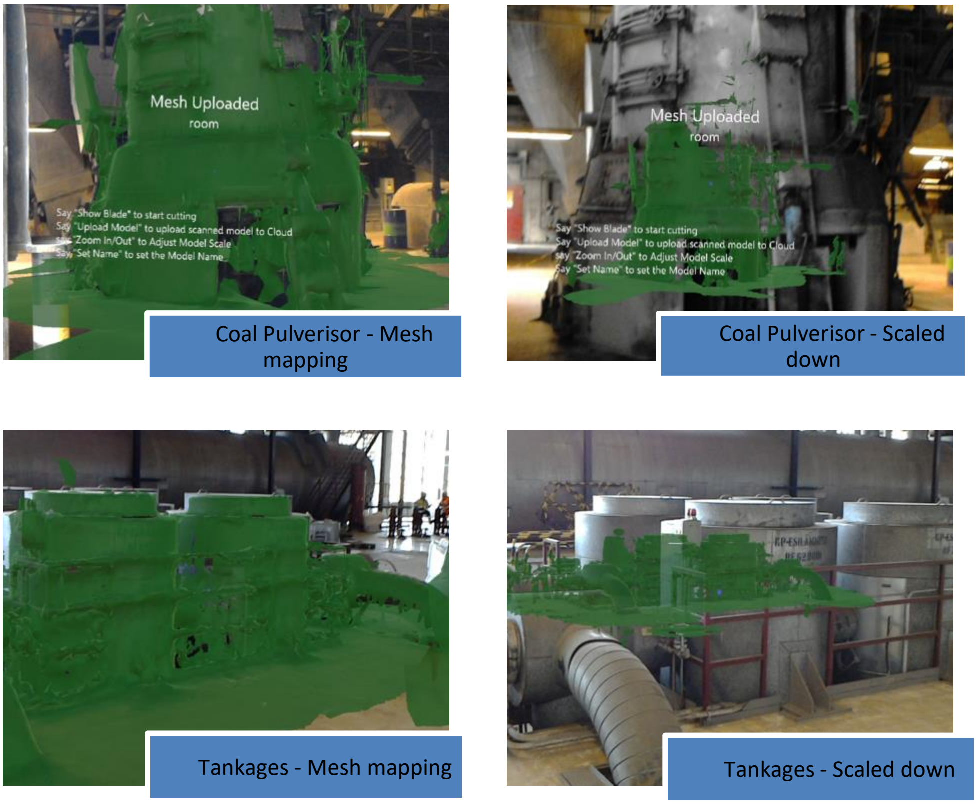
Mesh Uploaded
room2

Say "Show Blade" to start cutting
Say "Upload Model" to upload scanned model to Cloud
say "Zoom In/Out" to Adjust Model Scale
Say "Set Name" to set the Model Name

Figure 18. *Voice-based UI*

Environment can be scanned in multiple resolutions. The scanning (Exhaust in this case) has also worked good in outside environments. However, based on observation, it is still recommended to use indoor environments as the HoloLens display of holograms is not bright enough for usage in direct sunlight.

Just after the scanning, the mesh shows in the green colour and can be scaled up or down either for viewing or for cutting. Here the scan has been made in low level. In Figure 19, the left images indicate how the mesh maps to the real world. After the scan is finished, by using HoloLens relevant 3D model can be either extracted through cutting or the raw scan can directly be uploaded to the cloud server.



Coal Pulverisor - Mesh mapping

Coal Pulverisor - Scaled down

Tankages - Mesh mapping

Tankages - Scaled down

Figure 19. *Mapping scanned meshes to the real world*

Scanning the Coal Pulveriser was also a challenging task. There was very less light and only 10 minutes was available for scanning this machine. This also shows HoloLens performance in darker regions as well. Secondly, yet again, it was proved that powerful experiences can be created in a little time. The below figure also shows how intricate shapes have also been captured by HoloLens. Further, we were able to climb using the ladder for scanning coal pulveriser. This shows that the device can scan vertically more than its depth sensing range (i.e. 5 m) as well if the user can traverse vertically.

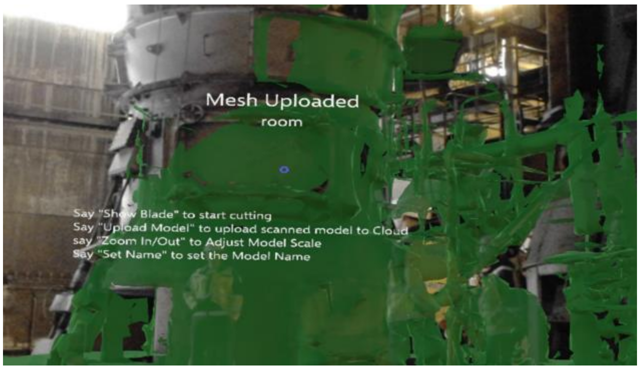


Figure 20. *HoloLens performance in dimly lit areas*

Here another key feature of the app was used where one can use a virtual keyboard to not only assign ID to the 3D model before uploading but also use it for annotation. Tagging info in the real world is also an interesting use case for Fortum and Kii is continuing collaboration with them. Downloaded 3D models after pilot project are retrievable both in the HoloLens as well as exported as OBJ file. Below images shows the exhaust, which is scanned and later visualised in a new location.

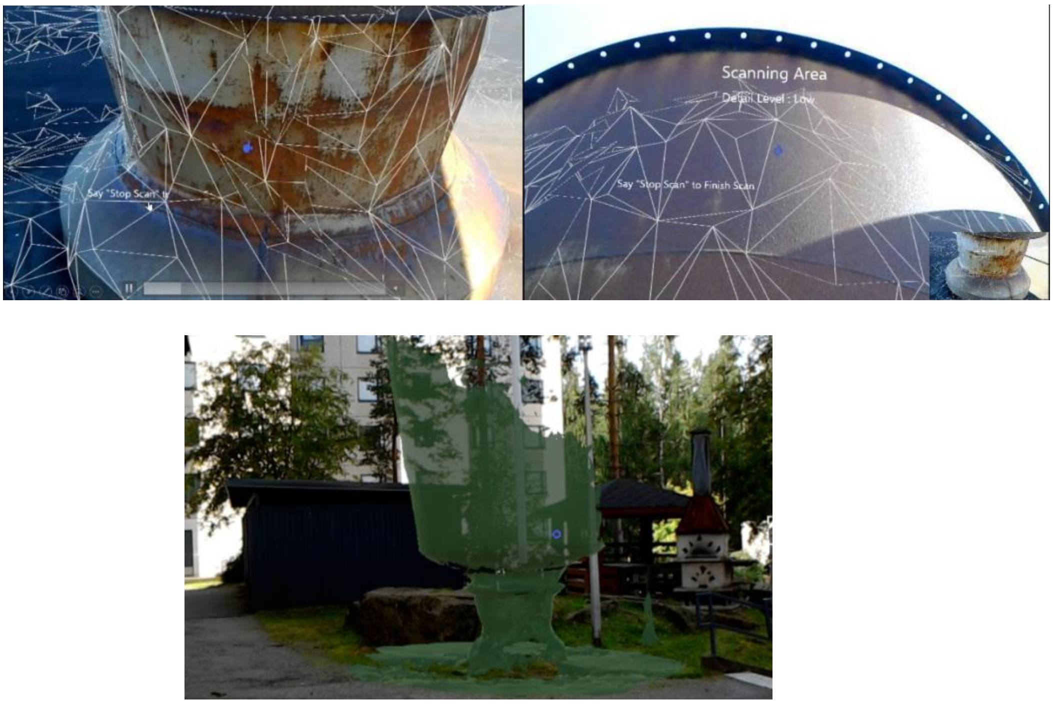


Figure 21. *Uploading a scanned mesh and downloading it in AR*

6. RESULTS

These solutions once developed were tested through piloting. Also, the industry was exposed to these tools and feedback was collected from experts to understand if the solutions met expectations of the industry. During the period multiple applications were developed. Most important ones are mentioned below:

Training / Learning in 3D

Figure 22. *Diesel engine model for training*

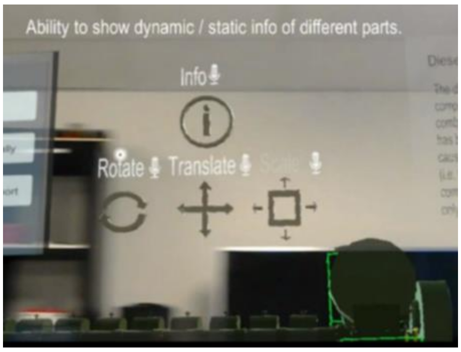


Figure 23. *UI of the training demo*

This application is a powerful demonstration of how we can learn in 3D. It allows users to manipulate a 3D model by positioning it in any place in any orientation. It also allows the user to scale up or down in discrete steps or visualise in real-life size. The green box visible in the figure is an interactive point to read information about different parts of the diesel engine. Further, it could also show dynamic information such as live sensor data.

Mixed Reality Toolbox

An alpha version of the toolbox has been made supporting scanning 3D model, slicing, scaling up and down, adding annotation, mixed reality capture and uploading / downloading the scanned 3D model. This work is state of art and till now, no one has yet attempted to combine these many numbers of tools inside one UI.

Figure 24. *Alpha version of the mixed reality toolset*

7. DISCUSSION

Though the results were promising, and we did feel that the acceptance for the same could be high, we realised that there was always a small gap between what we developed and what the user at the site needed. Each solution was a little different from the other, which meant that we had to customise and build different solutions for different applications. It was then decided that it would be wiser to develop a single toolkit-like solution which gave the user the power to pick tools that were needed to efficiently complete the task at hand.

This led to conceptualising a plug and play Mixed Reality Toolbox for industries by which one can interact with the surroundings in a natural way and gain superpowers of the mixed reality. The following schematic shows a 2D representation of the concept.

Figure 25. *Concept – Plug and Play Mixed Reality Toolbox*

The solution is being able to integrate multiple tools like rotate, translate, scale, scan and measure to be able to solve industry problems. Our study shows that the same tools, if integrated properly, can be used to solve different use cases of multiple industries like construction, marine, automobile and manufacturing.

8. CONCLUSIONS AND FUTURE WORK

Smart glasses have proven to possess beneficial capabilities which can be leveraged to minimise the gaps in the construction and manufacturing industry. The limited research carried out to see user acceptance in an industrial environment has come up with positive results [48]. This gives hope that building solutions built on smart glass hardware have long term benefits as a result of digitalisation. Many solutions have been developed, to suit different applications, which use smart glasses and other supporting technologies to provide a wholesome experience.

This field is new and ever growing and it is the right time for industries to adopt such solutions. There is a lot of space for improvement on developments in the solution created by Kii Oy as well. For example, currently scanning does not support textures. This means that it’s a single-colour scan. Having coloured 3D models shall help to recognise the objects in the scanned mesh in an easier manner. In another case, while doing measurement, it shall be a nice feature to be able to easily select the corners of the rooms. This feature can be developed using artificial intelligence using the gaze cursor position, position of walls, floors and ceilings as the input data. Creating simplistic 3D models of the site is another interesting direction wherein the model shall just consist of walls, ceilings, floors and few critical components like columns etc. This shall take site planning to the next level. For those interested, there are also other projects which are trying to use the abilities of smart glasses to make new and improved solutions. One major initiative includes working on a plug and play solution for placing 3D models in real environment for HoloLens and android.

The Society of Devices toolkit uses HoloLens along with Microsoft Kinect to create a collaborative environment, where multiple HoloLens can view the same holograms and the Kinect sensor helps to create more gestures than those available for HoloLens. This toolkit makes it possible for multiple devices to be connected to the same message server, making it possible for them to communicate and interact together [45].

Simon Mayer and Gabor Soros developed a system using smart glasses and a smartwatch to provide for better interaction possibilities. The smart glasses provide a method to select an object or model using image and surface recognition, whereas the smart watch provides different user interactions like rotate, translate, grab for the selected object. The state of the object is then updated depending on the interaction.

This method utilises different wearable computers and puts together the unique abilities of each to provide maximum functionality [46].

Using Kinect and GearVR to explore different methods to interact with virtual objects [47]. There is increased interest in indoor mapping and it is driving a demand for increased efficiency in the era of Building Information Model [50]. It is safe to say that smart glasses are here to stay. They make computing easy and the results more intuitive to use. If companies step in now to collaborate with smart glass developers, they can be the flag bearers to how this technology will develop and revolutionise the industrial processes.

REFERENCES

[1] Steve Mann, My “Augmediated” Life, IEEE Spectrum, IEEE Spectrum, "Vision 2.0" IEEE Spectrum, Volume 50, Issue 3, pp. 42-47

[2] Rauschnabel, Philipp A., Brem, Alexander, Ro, Young K., Augmented Reality Smart Glasses: Definition, Conceptual Insights, and Managerial Importance, The University of Michigan-Dearborn, College of Business, 2015. Available: https://www.researchgate.net/publication/279942768_Augmented_Reality_Smart_Glasses_Definition_Conceptual_Insights_and_Managerial_Importance

[3] Augmented and virtual reality go to work, Deloitte, 2016, Available: https://www2.deloitte.com/us/en/insights/focus/tech-trends/2016/augmented-and-virtual-reality.html

[4] Jan Koeleman, Maria João Ribeirinho, David Rockhill, Erik Sjödin, and Gernot Strube, Decoding digital transformation in construction, McKinsey and Company, 2019, Available: https://www.mckinsey.com/business-functions/operations/our-insights/decoding-digital-transformation-in-construction

[5] James Manyika, Sree Ramaswamy, Somesh Khanna, Hugo Sarrazin, Gary Pinkus, Guru Sethupathy, Andrew Yaffe, Digital America: A Tale Of The Haves And Have-Mores, 2015, Available: https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/digital-america-a-tale-of-the-haves-and-have-mores

[6] R. Yang, The Study and Improvement of Augmented Reality Based on Feature Matching, 2011 IEEE 2nd International Conference on Software Engineering and Service Science (ICSESS), Beijing, 15-17 July 2011, pp. 586- 589

[7] P. Gandhi, S. Khanna, S. Ramaswamy, Which Industries Are the Most Digital (and Why)?, Harvard Business Review, 2016, Available: https://hbr.org/2016/04/a-chart-that-shows-which-industries-are-the-most-digital-and-why

[8] Katharina Pentenrieder, Christian Badet, Augmented Reality-based factory planning - an application tailored to industrial needs, Dissertation, TU Munich 2008, Available: https://mediatum.ub.tum.de/doc/652443/652443.pdf

[9] Pica, Donato & Abanda, Fonbeyin, Emerging BIM-3D-Laser Scanning Integration in Construction Practice, 2019, p. 499

[10] Min-Koo Kim, Jack C.P. Cheng, Hoon Sohn, Chih-Chen Chang, A framework for dimensional and surface quality assessment of precast concrete elements using BIM and 3D laser scanning, Automation in Construction, Volume 49, Part B, 2015, Pp. 225-238, ISSN 0926-5805, Available: https://doi.org/10.1016/j.autcon.2014.07.010

[11] Martínez Landa H., Problems Analysis and Solutions for the Establishment of Augmented Reality Technology in Maintenance and Education, Tampere University of Technology. Publication; Vol. 1275, 2015

[12] Buti Al Delail, Chan Yeob Yeun, Recent advances of smart glass application security and privacy, Internet Technology and Secured Transactions, (ICITST) 2015 10th International Conference, pp. 65-69, 2015

[13] Anna Syberfeldt, Oscar Danielsson, Patrik Gustavsson, Augmented Reality Smart Glasses in the Smart Factory: Product Evaluation Guidelines and Review of Available Products, IEEE Access (Volume: 5), pp 9118 - 9130, 2017

[14] Stuart Elder, Alex Vakaloudis, A technical evaluation of devices for smart glasses applications, Internet Technologies and Applications (ITA), IEEE Explore 05, 2015

[15] Martin Kurze, Axel Roselius, Smart Glasses: An open environment for AR apps, IEEE Xplore: 22 November 2010

[16] Hartmut Surmann, Andreas Nüchter, Joachim Hertzberg, An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitalization of indoor environments, Robotics and Autonomous Systems 45, Pp. 181–198, 2003

[17] Miles Johnson, Hyundai virtual guide introduces augmented reality to the owner's manual, 2015, available: http://www.hyundainews.com/en-us/releases/2080

[18] Dominique De Guchtenaere, The use of video in online learning: asynchronous annotation tools, ICT for learning language 6th edition

[19] Berger, Matthew; Tagliasacchi, Andrea; Seversky, Lee M.; Alliez, Pierre; Guennebaud, Gaël; Levine, Joshua A.; Sharf, Andrei; Silva, Claudio T; A Survey of Surface Reconstruction from Point Clouds, Computer Graphics Forum, 2016

[20] Indoor-Outdoor Seamless Modelling, Mapping and Navigation Remote Sensing and Spatial Information Sciences, Volume XL-4/W5, 2015

[21] Peter Biber, Sven Fleck, Florian Busch, Michael Wand, Tom Duckett, Wolfgang Strasser, 3D modeling of indoor environments by a mobile platform with a laser scanner and panoramic camera, IEEE Xplore: 06 April 2015

[22] R. Kaijaluoto, A. Kukko, J. Hyyppa, Precise indoor localisation for mobile laser scanner, The International Archives of the Photogrammetry

[23] Ville V. Lehtola, Harri Kaartinen, Andreas Nüchter, Risto Kaijaluoto, Antero Kukko, Paula Litkey, Eija Honkavaara, Tomi Rosnell, Matti T. Vaaja, Juho-Pekka Virtanen, Matti Kurkela, Aimad El Issaoui, Lingli Zhu, Anttoni Jaakkola and Juha Hyyppä, Comparison of the Selected State-Of-The-Art 3D Indoor Scanning and Point Cloud Generation Methods , Remote sensing , MDPI, 2017

[24] Peter Henry, Michael Krainin1, Evan Herbst1, Xiaofeng Ren and Dieter Fox, RGB-D mapping: Using Kinect-style depth cameras for dense 3D modeling of indoor environments, The International Journal of Robotics Research, Pp. 647–663

[25] João Meira, João Marques, João Jacob, Rui Nóbrega, Rui Rodrigues, António Coelho, A. Augusto de Sousa, Video annotation for immersive journalism using masking techniques, 2017

[26] Laserscanning Europe GmbH, Laser scanning is the modern way of 3D surveying,Available: http://www.laserscanning-europe.com/en/laser-scanning

[27] Ian Wright, Quality Basics: How Does 3D Laser Scanning Work?, Available: http://www.engineering.com/AdvancedManufacturing/ArticleID/12390/Quality-Basics-How-Does-3D-Laser-Scanning-Work.aspx, 2016

[28] Andreas Nuchter, Hartmut Surmann, Kai Lingemann, and Joachim Hertzberg, Semantic Scene Analysis of Scanned 3D Indoor Environments, Proceedings of the Eighth International Fall Workshop on Vision, Modeling, and Visualization (VMV'03)

[29] Oliver Wulf, Bernardo Wagner, Fast 3D scanning methods for laser measurement systems, Institute for Systems Engineering, University of Hannover, Germany, 2003

[30] Thad Stanner, The challenges of wearable computing: part2, IEEE MICRO, 2001 pp 54-67

[31] Kari Lukka, The Constructive Research Approach, Publications of the Turku School of Economics and Business Administration, pp. 83, 2003

[32] Adekunle, O. , The constructive research approach in project management research, International Journal of Managing Projects in Business, Issue 4, 2011

[33] Jacek Rumiński, Adam Bujnowski, Tomasz Kocejko, Jerzy Wtorek, Alexey Andrushevich, Martin Biallas, Rolf Kistler, Performance Analysis of Interaction between Smart Glasses and Smart Objects Using Image-Based Object Identification, International Journal of Distributed Sensor Networks, vol. 2016, pp. 1, ISSN 1550-1329

[34] Jani Väyrynen, Mari Suoheimo, Ashley Colley, Jonna Häkkilä, Exploring Head Mounted Display Based Augmented Reality for Factory Workers, Proceedings of the17th International Conference on Mobile and Ubiquitous Multimedia, 2018

[35] HoloLens hardware details, Microsoft, 2021, Available: https://developer.microsoft.com/en-us/windows/mixed-reality/hololens_hardware_details

[36] Head-mounted display with electrochromic dimming module for augmented and virtual reality perception, Microsoft technology licensing, LLC., Patent Publication Number WO2016028828, Patent International Application Number PCT/US2015/045779

[37] Sophie Charara, Microsoft HoloLens: Everything you need to know about the $3,000 AR headset, Available: https://www.wareable.com/microsoft/microsoft-hololens-everything-you-need-to-know-about-the-futuristic-ar-headset-735 , May 2017

[38] Projection optical system for coupling image light to a near-eye display, Microsoft Corp., Patent Publication Number 2948813, Patent Application Number 14704438

[39] Microsoft Kinect hardware, May 2017, Available: https://developer.microsoft.com/en-us/windows/kinect/hardware

[40] Nick Statt, Microsoft's HoloLens explained: How it works and why it's different, January 2015, Available: https://www.cnet.com/news/microsoft-hololens-explained-how-it-works-and-why-its-different/

[41] Discussion on Hololens Forum. Available: https://forums.hololens.com/discussion/413/hololens-depth-sensor-and-technology

[42] Multiplayer gaming with head-mounted display, Microsoft Corporation, Patent Application Number 13/361,798 , Patent Publication Number US20130196757 (Line 0028 of description)

[43] Touchless input, Microsoft Corporation, Patent Application Number US 13/651,187, Patent Publication Number US20140104168

[44] Gestures, Microsoft, Available: https://developer.microsoft.com/en-us/windows/mixed-reality/gestures

[45] Cooper Davies, Omar Addam, Jade White, Fatemeh Hendijani Fard, Alec McAllister, Frank Maurer, Adam Saroka, A Toolkit for Building Collaborative Immersive Multi-Surface Application by published in ISS '16 Proceedings of the 2016, ACM International Conference on Interactive Surfaces and Spaces, Pp 485-488

[46] Simon Mayer, Gabor Soros, User Interface Beaming -- Seamless Interaction with Smart Things Using Personal Wearable Computers, Published in the 11th International Conference on Wearable and Implantable Body Sensor Networks Workshops on 16-19th June 2014

[47] Perlin, K, Future Reality: How emerging technologies will change language itself, IEEE Computer Graphics and Applications, 36.3, 2016, Pp. 84-89

[48] Jani Väyrynen, Mari Suoheimo, Ashley Colley, Jonna Häkkilä, Exploring Head Mounted Display Based Augmented Reality for Factory Workers, Proceedings of the 17th International Conference on Mobile and Ubiquitous Multimedia, 2018

[49] Elli Angelopoulou and John R. Wright, Jr., Laser Scanner Technology, Technical Reports (CIS) University of Pennsylvania, June 1999

[50] Jaehoon Jung , Sanghyun Yoon, Sungha Ju and Joon Heo, Development of Kinematic 3D Laser Scanning System for Indoor Mapping and As-Built BIM Using Constrained SLAM, Sensors 2015