POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Development and Evaluation of a Simple Hologram-to-Patient Registration Method using a Head-Mounted Display and an External Depth Camera

Tesi di Laurea Magistrale in
Biomedical Engineering - Ingegneria Biomedica

Diego Nabor González Rubalcava, 968260

Advisor:
Prof. Emiliano Votta

Co-advisors:
Dr. Maria Chiara Palumbo

Academic year:
2022-2023

Abstract: Mixed Reality, a concept defined as the blending between virtual and physical world through the use of real and holographic elements, holds the promise of becoming a useful tool during Image Guided Surgery; latest developments in this field have achieved a number of applications which would improve the surgeons experience by providing visualization of complex and 3D anatomical structures as well as surgical instrumentation in real time through the use of Head-Mounted Displays. Nevertheless, the need for further validation to prove the accuracy of these methods is needed for its implementation inside the Operating Room. One of the fields of interest corresponds to the assessment of patient-to-hologram registration, which is the process responsible for the accurate superimposition of virtual elements over the real world anatomical structures. In this thesis, the development of a simple and marker free hologram-to-patient registration method using a Head-Mounted Display (HoloLens 2, Microsoft, Washington) in combination with an external Depth Camera (Azure Kinect, Microsoft, Washington) is presented; the selected anatomical surface for the development of the method consists of a 3D printed human head phantom with its corresponding CT virtual model. The presented work includes two methods for the acquisition of patient surfaces using a single or multiple frames in favor of an improved hologram-to-patient registration quality along with an offset correction algorithm for the adjustment of errors due to HoloLens 2 intrinsic inaccuracies. An assessment for the accuracy of both registration methods was performed using an Optical Tracking system; registered mean error distance between the head phantom and the superimposed virtual model displayed through the Head-Mounted Display was of 8.34±0.73mm before offset correction and 3.85±0.19mm after offset correction for the single frame registration method, as for the multi frame registration method the obtained results were 9.83±1.04 before offset correction and 7.15±0.87mm after offset correction. As interesting as these results seem to be, the obtained accuracy is not sufficient for surgical interventions implementation, however, the presented work sets the foundations for improvements over the developed method for future developments.

Key-words: depth camera, head-mounted display, mixed reality, point cloud, registration, singular value decomposition

1. Introduction

1.1. eXtended Reality (XR)

eXtended Reality is a concept that covers a wide spectrum of immersive technologies such virtual reality (VR), augmented reality (AR), and mixed reality (MR). Being part of a bigger concept, the subdivisions have many similarities between thus, differentiating them is not a trivial task. In [1] Milgram et al. proposed the previously mentioned classification based on the characteristics of the surrounding environment, mainly whether it is primarily real or virtual. The main feature of XR is the creation of immersive experiences that pretend to minimize the boundaries between the digital and physical worlds. In general terms, and as shown in Figure 1, we can define the three concepts as per [2]:

- Virtual Reality (VR): Technologies which aim to enhance the interaction of the user through the use of a computer-generated environment. In this case the interaction with the real world does not exist due to the fact that the devices employed for the implementation of VR applications consist of occlusive headsets, e.g., Oculus headsets (Oculus VR, California), which result in a totally immersive experience for the user. Most common applications for Virtual Reality refer to gaming, entertainment and training in different fields.
- Augmented Reality (AR): Technologies which aim to enhance the perception of the real world by superimposing virtual elements on the field of view of the real world. Recent developed applications in the Augmented Reality field are not enclosed to a specific type of device, on the contrary, different technologies have adopted its use. Ranging from cellphones to head-mounted displays, e.g., Google Glass (Google, California), and even microscopes, AR applications cover a wide range of fields.
- Mixed Reality (MR): Technologies which merge the virtual and real world to create an immersive experience. In this case the interaction is with both, virtual and real world through the use of head-mounted displays, e.g, Microsoft HoloLens (Microsoft, Washington); this type of HMDs usually include a transparent lens, like the ones used for AR, with the main difference of the inclusion of depth sensors which enable the manipulation and interaction with the holographic scene.

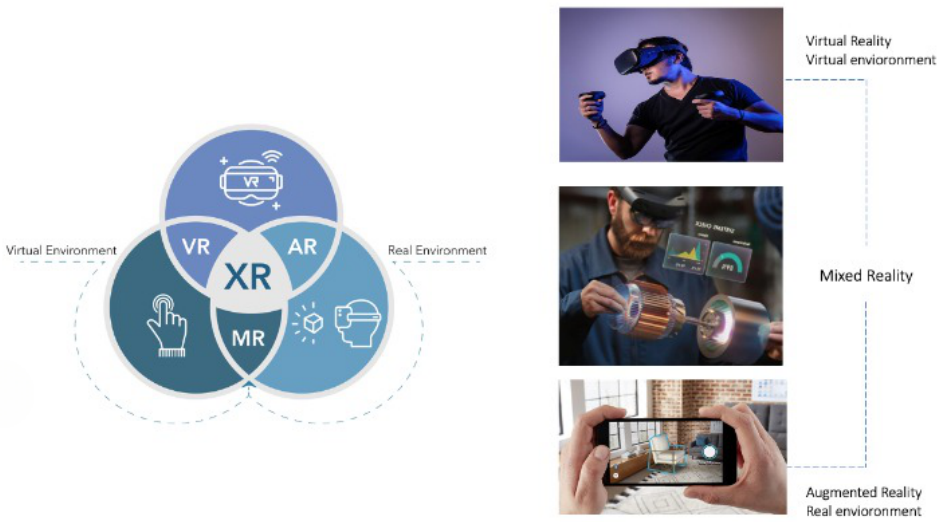


Figure 1: On the left, a general schema of the variants covered by eXtended Reality. Mixed Reality can be understood as a concept generated from the limitations that VR and AR presented. On the right, examples of applications for each XR modality; from top to bottom, an occlusive headset creating an immersive experience for the user. In the middle, a Mixed Reality application using a HMD with which the user can interact with the virtual scene. Bottom image depicts an AR in which a mobile phone is used for the superimposition of virtual elements without a direct interaction from the user.

VR was the first of the three technologies to be developed, followed by AR. While both of them remain important and relevant nowadays, the barriers that each of them presented encouraged the development of further solutions. Although VR provides a complete immersive experience, there is no relation with the real world; AR counteracts this issue by providing that contact with the physical world, however, the separation between both environments continues to be present. MR was conceived as an alternative to tackle all the previously mentioned challenges

by providing a space for interaction between real and virtual world objects which, as consequence, allows the execution of more practical scenarios.

1.2. Mixed Reality (MR)

Mixed Reality can still be considered an emerging technology; as seen before, it derived from both Virtual and Augmented Reality as an alternative to solve some of the intrinsic issues of these technologies. Some of the benefits offered by Mixed Reality are listed below.

- The immersive experience of the user is enhanced in comparison to the one provided by AR or XR thanks to the interaction with the real world in real-time.
- Mixed Reality offers a better sense of spatial awareness given the kind of devices used for its implementation. The user, in this case, can visualize both, real and virtual world at the same time.
- The increased accuracy provided by Mixed Reality when positioning virtual objects is another point to remark. Besides the improved experience for the user it provides more precise interactions.
- Even when the three modalities of XR provide a wide range of possibilities for the development of applications in different fields MR still stands as the solution with the higher flexibility due to its capacity to simulate hypothetical situations without having to leave aside the real environment in which they are developed.

It is possible to differentiate MR from AR and VR based on its four main characteristics [3]:

1. Combination of real and virtual world objects.
2. Mapping between the real and virtual objects, which allows interactions between them.
3. The previously mentioned interaction happens in real-time.
4. The use of specialized hardware (e.g., head-mounted displays) to create the virtual content.

MR goes far beyond from the concept of computer vision, it also involves signal processing, computer graphics, user interfaces, human-machine interaction, wearable computing as well as sensors and displays development; it is due to the advances in all of these fields that MR can now be applied to a different areas such as gaming, education, industry and healthcare [4].

The medical field is always in the search for implementation of new technologies which can improve the patient service and the physician experience. Mixed reality is not the exception, in recent times there have been many efforts to integrate it into diverse areas of the healthcare system [5]. Some of the most representative applications of MR in healthcare are listed below.

1. Medical Training: The limitations of medical environments, combined with the costs of specimens, has directly affected medical teaching; physicians suffer the lack of spaces and materials to practice procedures, their skill development is limited since they cannot manage to master surgical techniques in a more natural and realistic way. MR has played an important role in solving this issue; thanks to the development of virtual scenarios students and professionals can now practice procedures with the help of sophisticated and complex anatomical models, this translates in effective training time and improvement of their skill set for real world scenarios [6]
2. Remote Consultation: Healthcare has always been a complicated topic in regards to universal accessibility, nowadays there are still regions around the world in which an established healthcare system does not exist. Previous solutions to this issue included, but were not limited to, telemedicine, in which the consultations were done through pictures, videos or videocalls. MR shows signs of being a appropriate alternative by providing healthcare professionals the necessary toolset for long distance consultation and doctor-patient telecommunication. [6, 7]
3. Image Guided Surgery (IGS): MR provides the opportunity for preoperative and intraoperative visualization of patient's models superimposed with the real world. The advantages of this type of technology provide surgeons with a wider range of view allowing them to visualize farther than what the anatomical surface of patients shows. By having an extended point of view the surgery time might be reduced and the precision during procedures can be increased.[6, 8].

1.3. Image Guided Surgery

Many advantages have emerged as a consequence of the continuous development of this concept such as the drastic decrease in the use of exploratory surgery due to the increase in the use of minimal invasive surgery [8]. The history of Image Guided Surgery goes way back before the conception of MR, its roots date from the use of X-rays in the 19th century and go all the way along the development of Medical Imaging techniques such as MRI and CT. During the 19th century and the early part of the 20th century technological constrains pushed the images to be recorded and display into photographic films, it was not until the CT appeared that the idea of representing images as a set of numbers appeared. Nevertheless, IGS has been, up to these days, a topic

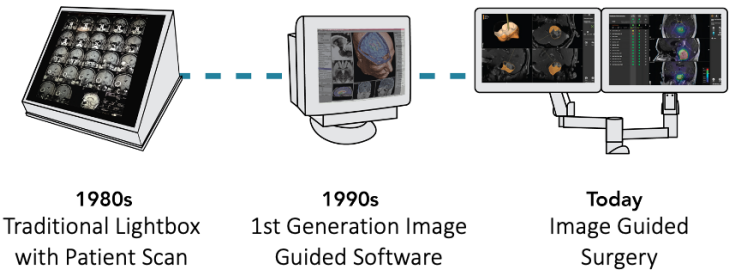


Figure 2: Image Guided Surgery through the years. From the use of Rx for surgery planning up to the latest developments of real-time tools for interventions.

that has been under constant development and evolution thanks to the constant changes in computational and display technologies. Nowadays, applications for the use of IGS have been developed for a large number of clinical procedures, between which we can find cardiovascular, intracranial, neurosurgery, orthopedics, just to name a few. Figure 2 depicts a brief graphical description of the evolution of IGS.

Even when each intervention performed with IGS is unique and different, there is a standard sequence of steps that every typical image-guided procedure follows [9].

1. Preoperative images have to be acquired. Typically, CT is the chosen imaging method selected for this purpose.
2. Surgical instruments used in the procedure are located using a trackable device. In this case there is not a clear preference, however the most common ones are optical tracking systems and electromagnetic tracking systems.
3. The patient anatomy undergoes a registration process to the preoperative image acquired in 1.
4. Surgical instrumentation position is displayed relatively to the produced anatomic image of the patient.
5. The virtually displayed instrumentation allows the surgeon to manipulate the tools with full awareness of its positions inside the patient's body.
6. Confirmation images are generated after the procedure has been completed.
   Advantages from IGS are non debatable, the benefits for both, physicians and patients are clear and the continuous improvement in this field confirms it is as one of the best techniques for minimal invasive surgery.

1.4. MR in Image Guided Surgery

Medical Imaging has been pivotal for the way in which the concept of surgery has evolved through the years, not only does it provide the surgeon with detailed and accurate 3D anatomical images but it can also be used to guide surgical instruments across the anatomical structure in real-time. The use of real-time imaging would give the surgeon the opportunity to a clearer vision of what is happening inside and outside of the anatomical structure laying in front of him [10].
The main issue with the implementation remains the long path that it must go over in order to be a common practice in the Operating Room as there is still many challenges to overcome in order to prove the feasibility of MR in surgical procedures. Even when MR specific hardware such as Head-Mounted Displays (HMDs) have been rapidly evolving and the arguments about their use in the Operating Room taking minimal invasive surgery to a whole new level, the need to prove that this benefits would result in a breakthrough in surgery remains [11]. In this context, the achievement of proper hologram-to-patient registration represents a huge step towards the implementation of MR as a common tool during surgical procedures from a technical and practical point of view; achieving proper accuracy and precision in this type of procedures remains as the first step towards real-life scenario implementations. Here, different approaches for hologram-to-patient registration are presented, the focus will be set to solutions implemented using Head-Mounted Displays (HDMs), which refers to all those methods developed for the superimposition of a hologram over the anatomical structure of the patient with the goal of visualizing the diverse components contained in the human body in a virtual way through the use a specialized piece of hardware. Until today, many different methods have been proposed with the objective of providing the most accurate registration. For the remaining part of this section, a review of those methods will be performed.
Various efforts have been made to achieve a proper hologram-to-patient registration. Being one of the first steps in the context of IGS the accuracy of this procedure results essential for a proper visualization and, therefore,

the success of the surgery might fall back on it. Different approaches have been studied to achieve the hologram-to-patient superimposition, the spectrum of variations between methods goes from the differences between the chosen algorithm to perform registration to the various experimental setups proposed.
In [12] Kuhlemann et al. propose a method for the holographic visualization of the patient's surface along the movement of a catheter inside the anatomical structure using HoloLens as their HMD and a magnetic tracking system for the selection of fiducial points for the alignment of the virtual reference frame with the real world. The results reported for this solution correspond to a root-mean-square error (point-to-point correspondence) of 4.347mm with a variance of 0.709mm. The main issue with this solution refers to the registration process due to the inaccuracies that might occur when trying to obtain the exact same points in the virtual scene and in the real world through a manual procedure. In other words, to perform the calibration the user must select some points in the real world which then will have to be selected in the virtual scene.
Another method worth mentioning is the one proposed by von Haxthausen, Chen and Ernst in [13] which is based on the use of HoloLens 2, a handheld scanner and markers for the registration process. The idea remains the acquisition of the markers in the real world for then being selected in the virtual scene in favor of performing the registration which allows the alignment between both coordinate reference frames. Although the results for this method show to be promising showing an average error of 22.3mm in the x-axis, 35.6mm for y-axis and 13.3mm for the z-axis between the real and virtual markers used for evaluation; the idea of using markers to perform the registration can be improved in order to produce a more robust algorithm which does not depend on the accuracy of the user to match the required points. Sun et al. propose a similar method in [14] with the main difference being that the calibration and registration procedure are based on the use of an optical tracking system; in this paper, the reported average mean square error of the points in the holographic scene and the phantom used for evaluation equals 1.30±0.39mm . Even when optical tracking systems can be considered as a robust solution for providing a stable reference frame to a system the issue arises due to the fact that in some cases they might become impractical in real life situations given the fact that the tracking system is based on the use of a tracking probe which has to be in the field of view of IR spectrum of the optical tracker; coupled with this, if the focus is centered around the velocity of the method then an optical tracking system might not be the best solution since the acquisition of coordinates by the user can take some time that for means of the efficiency of the system could prove to be significant.
The use of markers seems to be the most accepted method for achieving an accurate registration. Fick et al. stick to the use of markers in their proposed solution [15]; in this case the registration is performed based on the use of QR codes placed around the patient. Results for FRE (Fiducial Registration Error) which corresponds to the root mean square distance between recognized fiducial positions and their homologous virtual fiducial after registration were reported as a mean of 8.5mm. registration The purpose of the markers remains the same as other studies previously commented in this section, the acquisition in both, real and virtual scenes, allows the user to obtain an alignment of both coordinate frames. QR codes seem to be useful tools in many fields due to the amount of methods that have been developed in different platforms for their automatic detection, however, the experimental setup exposed in this study results impractical and unstable; in the method proposed in this thesis QR codes are used for the sole purpose of evaluating the performance of the algorithm while Aruco markers are used for means of a reconstruction algorithm which will be explained later in the thesis. The use of this type of tools for a crucial process such as calibration of the system could prove to be unstable.
More complex approaches have been developed in order to achieve hologram-to-patient registration. In [16] Palumbo et. al. propose a method for the automatic registration for the superimposition of a holographic model over the segmented skin model of the head of a real patient through the use of point cloud registration. The main idea behind the development of this method consists of the acquisition of the patient's point cloud using HoloLens 2 research mode for a posterior alignment with the position of the acquired surface for a final holographic superimposition. A point-to-point accuracy evaluation yielded a RMSE of 3.19±1.13mm. As interesting as the results proved to be the final conclusions made reference to the addition of a further and independent external depth camera, such as the Azure Kinect, to achieve a more robust and stable point cloud acquisition which should derive in the accuracy improvement of the complete registration process.
Many other alternatives which go out of the scope of this study since they are focused on the use of AR have been proposed such as Golab et al. in [17]; even when they cannot considered as a competing relation it is worth to mention them since they show how the use of HMDs is being widely studied and its implementation as a regular practice in the Operating Room might be getting closer.

1.5. Head-Mounted Display

After a review in the latest developments of Mixed Reality in Image Guided Surgery and having already established HMDs as a basic need for the use of MR the focus of this section will be centered around the two main devices used on the development of this thesis.
The design and development of new HMDs such as Google Glass, Oculus Quest and Epson Moveiro has been

benefited by the latest progress in the field of XR. In this context, there is also strong evidence that suggests their readiness to be used in the provision of healthcare since they hold the promise to support the delivery of different healthcare services. However, there is a specific device which stands above the rest when referring to MR application, namely HoloLens [18]. The selection of HoloLens 2 for the development of this thesis follows the line of research which suggests a superior performance compared to the other available commercial headsets. In this section, a brief description of the main features of HoloLens 2 is provided.
HoloLens 2 (Microsoft, Washington) is a mixed reality head-mounted display which, as its working modality suggests, allows the blending of digital content, e.g., holograms, with the real world. Its main advantage over other devices refers to its capacity to deliver an immersive experience to the user while allowing the interaction with the virtual environment. It functions over a holographic computer running Windows Holographic Operating System, an operating system characteristic of the HoloLens series. The main constitution of the device consists of a headset with a visor-like design equipped with see-through holographic lenses which allow the visualization of the real-world surroundings while the display of digital content is achieved through the use of an advanced optical projection system.
Other than its sophisticated system for holographic display it is important to also focus on the sensors contained within the structure of HoloLens 2. The HMD is equipped with different sensors which allow to exploit its capacities at a maximum level through the tracking of user's movements and the interaction with both, virtual and real world; Figure 3 provides a better visualization of the placement of this sensors in the device. Listed below, the key sensors with which HoloLens 2 is equipped.

- RGB cameras: Two 2MP RGB cameras are located in the front of HoloLens 2 together with an infrared camera and a depth camera. Between the many functions of this elements the most significant one refers to the tracking of movements and recognition of hand gestures.
- Depth Sensors: HoloLens 2 contains four depth sensors based on time-of-flight technology which allow the measurement of between the headset and the different objects located in the surroundings.
- IMU: Four Inertial Measurement Units track the orientation and movement of the device due to the movement of the user. This type of sensors play an essential role in the placement of the holographic elements in the virtual scene.

Other than its remarkable performance, the selection of HoloLens 2 responds to the necessity of a device with a depth camera; the presence of this specific type of device allows, through the recognition of the surroundings, the use of the information it provides to be used for the alignment of virtual elements over real objects as well as the acquisition of point clouds. This concept, known as *Registration* in the context of computer vision will result fundamental through the development of this thesis.

Figure 3: On the left, HoloLens 2 (up) and a breakdown of its main components (down). On the right, Azure Kinect (up) and a breakdown of its main components (down).

As mentioned by [16] the use of an external depth camera could prove beneficial if used in combination with HoloLens 2 in view of acquiring more precise and stable point clouds. For this purpose, Azure Kinect (Microsoft, Washington) was selected as the external depth camera for the development of the thesis. Azure Kinect is a depth-camera designed for the acquisition of 3D data; just as HoloLens 2, this device bases its working principle in the use of specialized sensors which are listed below. Figure 3 depicts a graphical description of the sensors included in Azure Kinect.

- RGB camera: A high-resolution RGB camera is included in the composition of Azure Kinect. It provides the capacity of color image capturing which can be combined with the information of the depth camera

to create more realistic images.
- Depth camera: Following the working principle implemented in HoloLens 2, Azure Kinect includes a time-of-flight based depth camera which allows the creation of 3D maps which can be used for motion tracking or, in the case of the thesis, the generation of point clouds.

Azure Kinect supposes a versatile solution to work with due to the existence of Software Development Kits (SDKs) for the creation of different applications granting access to the different components that have been already mentioned.

1.6. Motivation

In this thesis, the development of a framework for the registration of virtual and real-world objects is presented as a general method for hologram-to-patient registration which is intended to have the capacity to be applied to any kinds of surgical procedure. The presented workflow was tested with a rigid anatomy, such as the ones present during neurosurgery cases thus, it would work better for this purpose; however, the possibility to expand it towards different surgical modalities can be achieved providing the necessary adjustments. For means of this thesis the presented workflow was developed in a general manner, not directed towards as a specific neurological procedure. The main goal consists of the superimposition of a hologram corresponding to a CT derived 3D model which, in this case, corresponds to a human head phantom.
The implementation of the reported framework was achieved through the use of a Head-Mounted Display (HoloLens 2, Microsoft, Washington) and an external depth camera (Azure Kinect, Microsoft, Washington); communication between devices is performed through ROS Noetic (Stanford Artificial Intelligence Laboratory, Stanford). Further details about the technical implementation will be discussed in a separate section. The proposed solution is composed by two main sections which are briefly described below.

- Calibration phase. Given the fact that HoloLens 2 and Azure Kinect operate in different coordinate reference frames a calibration to align them in the same reference frame is needed.
- Hologram to Patient registration. Having both devices operation in the same reference frame it is then possible to perform hologram-to-patient registration in order to achieve the desired hologram superimposition over the anatomical surface. For this segment of the workframe two methods are proposed:
    1. Single Frame Registration. Based on the acquisition of one frame of the head phantom surface using Azure Kinect Depth Camera for the hologram-to-patient registration.
    2. Multi Frame Registration. Proposed method which implements the reconstruction of the head phantom surface through the use of multiple point cloud acquisiton and registration in order to improve the final hologram-to-patient superimposition.

Upon completion and implementation of the complete framework, an experimental setup was established for the evaluation of the accuracy and precision of the hologram-to-patient registration. As a consequence of the nature of the obtained results in which, virtual and real elements are combined for measurement, a custom evaluation method emerged as necessary. A complete section covering this topic is depicted later in this thesis.
Other than the use of the previously mentioned devices, the presented solution makes use of Aruco Markers and QR Codes. The first ones are used during the reconstruction process of the head phantom's surface in the Multi Frame registration method while the latter are utilized during the accuracy assessment for both registration methods.
Given the already mentioned developments in the field of Mixed Reality in Image Guided Surgery, the goal of the reported framework and the obtained results is intended to show the reliability, accuracy and simplicity of the developed solution for the superimposition of holograms over real-world objects.

2. Materials and Methods

2.1. System description and workflow

The aim of this study is to achieve a reliable and accurate method for hologram-to-patient registration. The main structure of the application consists of a combined registration procedure using HoloLens 2 and Azure Kinect; the first one working as the dedicated tool for for holographic visualization and user interaction with the virtual scene; the latter provides the surface acquisition required for the registration procedure. The system is designed to be controlled through HoloLens 2 by means of a holographic interface developed in Unity3D (Unity Technologies, version 2020.3.41f1) using the Mixed Reality toolkit library MRTK v.2.5.1, while data is processed on a workstation running Ubuntu Focal 20.04. Communication between devices is achieved using ROS Noetic, specifically the Unity Robotics Hub repository which facilitates message passing to and from Unity and the workstation where ROS is installed.

A brief description of the sequence of steps to achieve a proper hologram-to-patient registration is provided below. Figure 4 provides a schematized graphical version of the general steps of the registration process.

1. Starting from the Holographic interface, the user can choose between the two developed registration methods, Single Frame or Multi Frame, further description regarding these methods will be provided later.
2. Once the method has been selected, the calibration phase comes into place. The main purpose of this process is the calculation of a transformation matrix which would allow to align Azure Kinect and HoloLens 2 to a common coordinate reference frame so to manage to relate each surface acquired by the Azure Kinect depth camera to the holographic device. Roughly, the user starts by acquiring a point cloud of a calibration object, previously selected, using HoloLens 2 by means of the previously mentioned holographic interface. After point cloud acquisition, the system will automatically proceed with the acquisition of a second point cloud of the calibration object using Azure Kinect; upon completion, and having the 3D information of the calibration object, namely the point clouds, for both devices the main idea of the calibration process consists of the use of correspondent points in the acquired object; by being able to transform these known points in Azure Kinect and HoloLens 2 coordinate reference frames the rototranslation matrix that relates one coordinate system to the other can be obtained. Section 2.3 provides the detailed implementation regarding the Calibration procedure.
3. Following the calibration phase, the hologram-to-patient registration constitutes the next step of the process. For this purpose the user, through the use of the holographic interface, can initialize the process. As previously mentioned, two methods are presented for this segment of the process.
   - The first method consists of the registration procedure performed taking into consideration a single scan of the head phantom using Azure Kinect. The mentioned frame is registered with the CT model of the head phantom to achieve the final hologram-to-patient display. Section 2.4 provides further details regarding Single Frame Registration method.
   - The second method consists of the same principle of the previous one but performing a reconstruction of the phantom's surface through a multi frame scan using Azure Kinect and point cloud registration. The main goal of this proposed method is the improvement of the registration between the patient's surface and its CT model in order to achieve a better holographic alignment. Multi frame registration, including acquisition and reconstruction process are better explained in Section 2.5.

   Both methods provide the same outcome, a registration matrix which allows the alignment of the holographic model over the head phantom 3D structure.
4. For means of this thesis, an assessment section is added in order to evaluate the accuracy of the hologram-to-patient registration obtained through the developed method. An experimental setup using an Optical Tracking system was developed for such purpose. A complete walk through this method will be provided in Section 2.6.

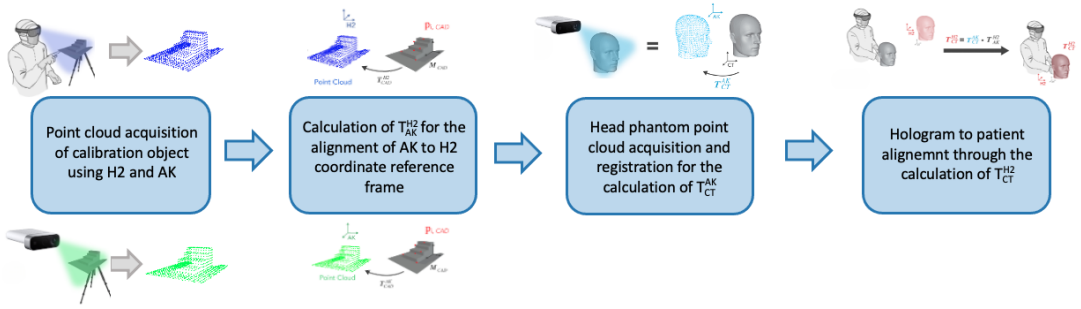


Figure 4: Workflow for the hologram-to-patient registration process. Main steps of the procedure are reported and briefly described here.

System's general workflow relies on the use of the holographic interface implemented in the HoloLens 2 for data acquisition and execution of certain commands that translate into the publication of messages to different ROS topics. As for the Ubuntu part of the application, the main purpose is to employ it as a module for data processing and controller for the functions of the Azure Kinect when necessary. Summarizing, the idea behind the solution translates to an information exchange system through the use of publishers and subscribers to trigger specific actions in both parts of the system.
Figure 5 shows the proposed experimental setup for the application. Description of each component and further explanation about the workflow will be detailed in the following sections.

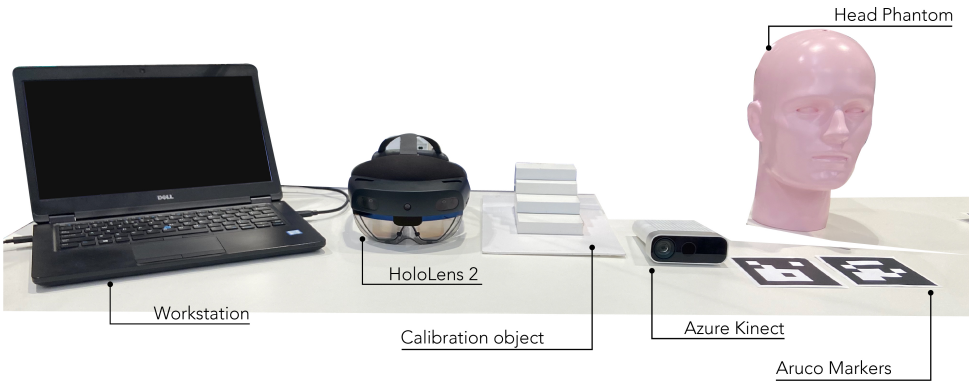


Figure 5: Hardware and physical components included in the experimental setup for the developed hologram-to-patient registration method. Workstation operates together with Azure Kinect using Ubuntu 20.04 Focal; communication with HoloLens 2 is achieved through ROS Noetic.

2.2. Implementation

HoloLens 2

As mentioned in section 2.1, Unity3D is used for the development of the application for HoloLens 2. Main features of the application are implemented using MRTK v.2.5.1, a Unity plugin that provides a number of assests and components to facilitate the development of spatial interactions. Access to the depth camera of HoloLens 2 is achieved using Research Mode, a feature that provides access to key sensors such as the previously mentioned depth sensor. Figure 6 illustrates the use of HoloLens 2 during a real case scenario.

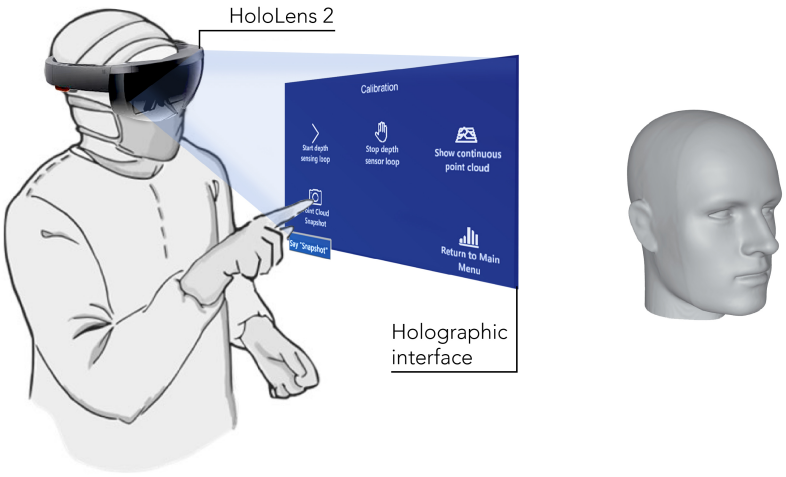


Figure 6: Example of use for HoloLens 2. The user has access to the holographic interface which includes various menus for each step of the registration process.

Azure Kinect

The main subject of this study is based on the use of point cloud registration to achieve hologram-to-patient superimposition.Azure Kinect serves the sole purpose of generating point clouds for registration and color images; this could be achieved by using the Azure Kinect Sensor SDK in the workstation, which is a set of tools that, when implemented, grant access to both, depth and color camera in compliance with the requirements of the general framework of the system.
Point cloud processing is done using Open3D, an open-source library that supports the use of 3D data. Figure 7 provides a representation of the use of Azure Kinect for the proposed framework.

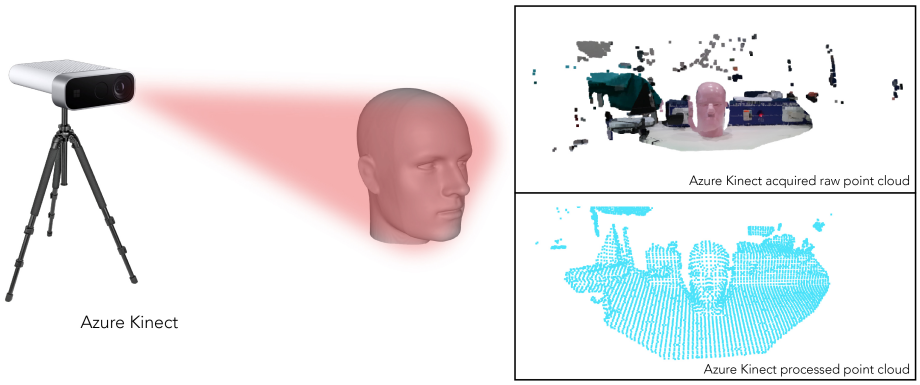


Figure 7: On the left, example of use for Azure Kinect. Here, the scanning of a patient is represented; by placing the device at the correct distance a full point cloud of a specific object might be obtained. On the right, an example of the actual acquisition of a point cloud using Azure Kinect.

ROS Noetic

Due to the constrains imposed by the design of the application the communication between HoloLens 2 and the workstation running Ubuntu is done using ROS Noetic via the Unity Robotics Hub, a repository which includes robotic tools for unity, among which we find ROS-Unity Integration. This set of tools allows the creation of a subscriber-publisher network of various topics that ended up being the main core of the system. Figure 8 shows the diagram for the communication between nodes by illustrating the publishers and subscribers for each topic employed for the development of the application. For simplicity, only standard messages are used for the development of this section.

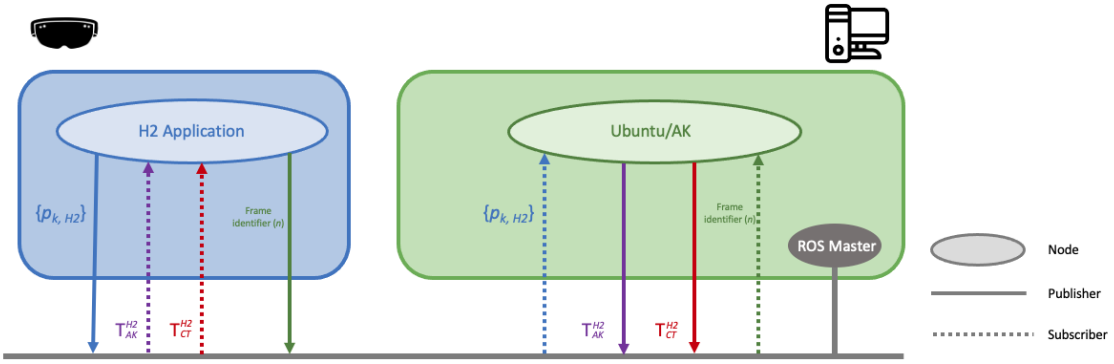


Figure 8: Schematic representation of the data communication between Workstation and HoloLens 2 through ROS Noetic. The system is based on a publisher-subscriber model.

2.3. Calibration

The calibration procedure has the purpose of relating the coordinate reference frames of HoloLens 2 and Azure Kinect. To accomplish a proper calibration the object shown in Figure 5 was selected as the calibration object. The physical object is scanned using both devices in order to generate two sets of point cloud which are then registered with the point cloud corresponding to the virtual CAD model of the head phantom, in this way, a set of predefined points on the CAD model can be referred to each of the two coordinate reference systems. Having both sets of corresponding points in both coordinate systems will allow the application of a process to find a relationship between them. It is important to point out that no matter the selection of registration modality, single or multi frame, the calibration process has to be carried out. The whole process is described below:

1. After selection of the registration modality via HoloLens 2 the user is required to capture a frame of the calibration object using the HoloLens 2 depth camera. To execute this action, the sensor loop of the depth camera has to be initialized so that the point cloud can be visible in real time through the head-mounted

display. Once the calibration object is completely covered by the continuous point cloud the user can use voice commands to capture the frame, obtaining $\{p_{k,H2}\}$ with k=1, 2, ..., N, which is a set of 3D points in HoloLens 2 coordinate reference frame representing the calibration object. $\{p_{k,H2}\}$ gets published as a Nx3 matrix for posterior processing in the workstation.

2. As a consequence of the publication of $\{p_{k,H2}\}$ the callback function for the subscriber to this topic initializes the scan of the calibration object using the depth camera of Azure Kinect; this process gives as result $\{p_{k,AK}\}$, which is a set of 3D points in the Azure Kinect coordinate reference frame.
3. Having acquired both point clouds, the process that each one undergoes is mainly the same. As the main purpose is to obtain a transformation matrix $\mathbf{T}_{\mathrm{AK}}^{\mathrm{H2}}$ that allows the whole system to operate on the Hololens 2 coordinate reference frame the idea of using the point cloud of the CAD model of the calibration object $\{p_{k,CAD}\}$ to perform a first registration emerges as a first step towards the final registration. Both sets of point clouds, $\{p_{k,AK}\}$ and $\{p_{k,H2}\}$, are registered with $\{p_{k,CAD}\}$. The outcome of these registrations allows the alignment of any point in the CAD model into each of the devices coordinate reference frames, this concept take relevance in the next steps since the obtention of $\mathbf{T}_{\mathrm{AK}}^{\mathrm{H2}}$ is based on the identification of corresponding points taken from the calibration object which require to be translated into AK and H2 coordinate reference frames. First, a Fast Global Registration Algorithm is applied in order to obtain an initial alignment between the two set of points; afterwards, a refining registration method based on Point-to-Plane ICP algorithm is performed giving as result a transformation matrix for each acquired point cloud $\mathbf{T}_{\mathrm{CAD}}^{\mathrm{AK}}$ and $\mathbf{T}_{\mathrm{CAD}}^{\mathrm{H2}}$. However, these transformation matrixes does not conclude the calculation of $\mathbf{T}_{\mathrm{AK}}^{\mathrm{H2}}$.
4. As shown in Figure 9, six points, denominated as $\{p_{i,CAD}\}$, which properly represent the main geometrical features of the calibration object are selected. $\mathbf{T}_{\mathrm{AK}}^{\mathrm{CAD}}$and $\mathbf{T}_{\mathrm{H2}}^{\mathrm{CAD}}$are used to transform $\{p_{i,CAD}\}$ into each of the devices coordinate reference frames, obtaining:

$$\{q_{i,AK}\} = T_{CAD}^{AK} \cdot \{p_{i,CAD}\} \tag{1}$$

$$\{q_{i,H2}\} = T_{CAD}^{H2} \cdot \{p_{i,CAD}\} \tag{2}$$

5. With $\{q_{i,AK}\}$ and $\{q_{i,H2}\}$ being the six surface points in Azure Kinect and HoloLens 2 coordinate reference frame respectively, singular value decomposition is applied for finding the optimal translation and rotation that yields $\mathbf{T}_{\mathrm{AK}}^{\mathrm{H2}}$.
6. Some important remarks for the calibration algorithm:
   - Singular value decomposition algorithm, being a least-square based solution, requires a minimum of three points from each point cloud as well as position correspondence between the two sets of points to provide an accurate transformation matrix.
   - $\mathbf{T}_{\mathrm{AK}}^{\mathrm{H2}}$ is only valid as long as Azure Kinect remains in the same position in which the initial registration is performed. As for HoloLens 2, the same matrix will remain valid until the application is restarted since the device's coordinate reference frame changes with every reboot.
   - HoloLens 2 works in a left-handed coordinate system while Azure Kinect and CAD models work in a right-handed coordinate system. Taking this into consideration, and since the end product has to be implemented in the head-mounted display, all the acquired point clouds were transformed to a left-handed coordinate system.

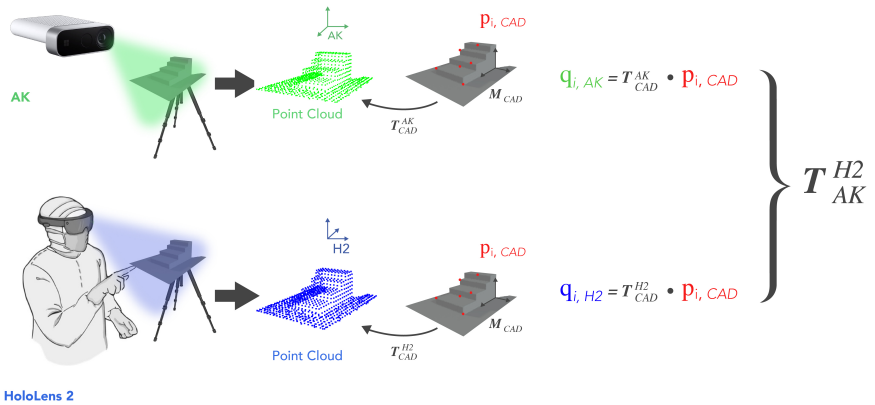


Figure 9: Calibration phase. This step of the process produces $\mathbf{T}_{\mathrm{AK}}^{\mathrm{H2}}$; the correct localization of the six representative points in both, AK and H2 coordinate reference frames, is crucial for the calculation of the desired matrix through the use of singular value decomposition.

2.4. Single frame registration

Upon completion of the calibration process, the next step consists of the registration of the patient's surface with its CT model in order to obtain a transformation matrix $\mathbf{T}_{\mathrm{CT}}^{\mathrm{H2}}$ which yields a proper hologram-to-patient alignment. Figure 10 represents this part of the process step by step. The sequence of steps for achieving this are listed below.

1. Having obtained $\mathbf{T}_{\mathrm{AK}}^{\mathrm{H2}}$ it is possible now to transform any set of points in Azure Kinect coordinate system into HoloLens 2 coordinate reference frame. Having already established that Azure Kinect remains in a fixed position it is possible to infer that any scanned point with it and transformed to HoloLens 2 reference system will remain fixed when visualized through the head-mounted display. Following this hypothesis and by using the patient's CT model for the registration the next equation yields the final transformation matrix for hologram-to-patient alignment:

$$T_{CT}^{H2} = T_{CT}^{AK} \cdot T_{AK}^{H2} \tag{3}$$

2. Through HoloLens 2 holographic interface the user can proceed with the patient's surface point cloud generation using Azure Kinect and subsequent registration. Once the proper command is selected the message for point cloud acquisition initialization with Azure Kinect is published and its callback function is executed; this process is technically the same one executed during the calibration procedure, main change is the scanned object. Once the acquisition is completed the output can be denominated as $\{p_{i,PatAK}\}$.
3. In order to obtain $\mathbf{T}_{\mathrm{CT}}^{\mathrm{AK}}$ a point cloud of the patient's CT model is generated assuming its origin to be located in $\{0,0,0\}$, this point cloud will be denominated as $\{p_{i,CT}\}$. Registration between $\{p_{i,PatAK}\}$ and $\{p_{i,CT}\}$ is performed using the same procedure employed during calibration, a Fast Global Registration algorithm provides an initial alignment of the points and is refined using a Point-to-Plane ICP algorithm which produces the transformation matrix $\mathbf{T}_{\mathrm{CT}}^{\mathrm{AK}}$.
4. With $\mathbf{T}_{\mathrm{CT}}^{\mathrm{AK}}$ and $\mathbf{T}_{\mathrm{AK}}^{\mathrm{H2}}$ a simple matrix multiplication is enough to obtain $\mathbf{T}_{\mathrm{CT}}^{\mathrm{H2}}$, as explained in the first point of this section.
5. Once the calculation of $\mathbf{T}_{\mathrm{CT}}^{\mathrm{H2}}$ is completed, this transformation matrix is published as a 4x3 matrix. HoloLens 2, having the subscriber for that topic, receive that information and, once the user selects the option to display the hologram $\mathbf{T}_{\mathrm{CT}}^{\mathrm{H2}}$ replaces the rotation and translation information of the CT model, previously loaded to the head-mounted display, placing the CT hologram over the patient's surface achieving in this was the originally expected hologram-to-patient registration.

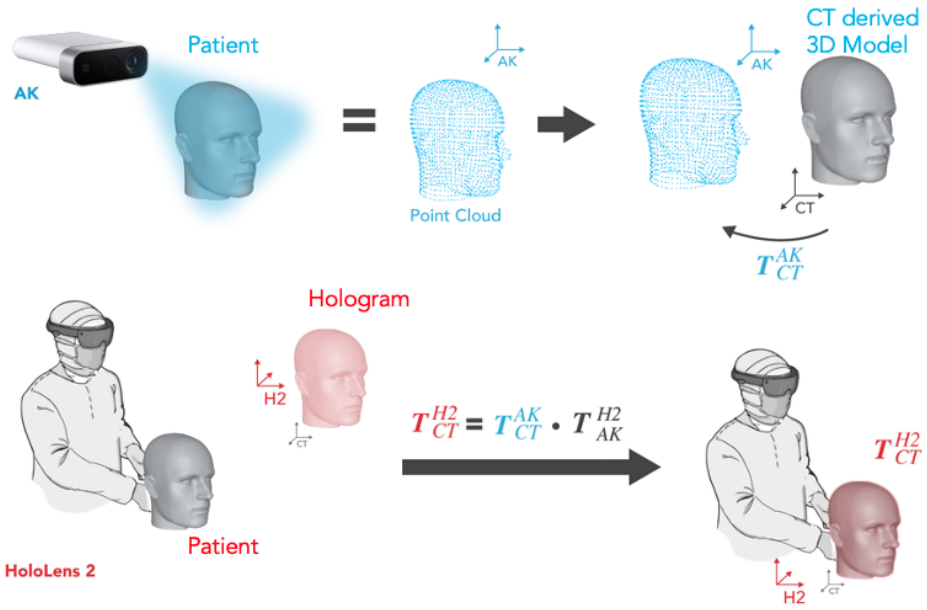


Figure 10: Single Frame Registration. Process conceived to obtain $\mathbf{T}_{\mathrm{CT}}^{\mathrm{H2}}$. In this case the acquisition of the point cloud through Azure Kinect is pivotal for the final hologram-to-patient superimposition. It is important to recall the change of coordinate systems that had to be performed due to AK and CT operating in right-handed while H2 does it in a left-handed manner.

2.5. Multi frame registration

Alternatively to the single frame registration method, a second modality for the registration between $\{p_{i,ObjAK}\}$ and $\{p_{i,CT}\}$ was developed. The purpose behind the development of this solution responds to the idea of improving the registration process for the patient's surface; in this context, the collection of different views from the same object would suppose the acquisition of a wider area which, theoretically, would result in an improved registration quality. One of the key concepts when evaluating the registration of two point clouds, and that will be further discussed in the following sections, is the number of correspondent points; for two sets of point clouds $\{p_{i,Source}\}$ and $\{p_{i,Target}\}$ which underwent registration, this parameter indicates the amount of points from the source point cloud that found a correspondent point in the target point cloud within a given threshold. To some extent, the correspondence set shows to be a good indicator of the quality of the registration procedure; the higher the number of points the more accurate the registration will be. The aim of this second proposed procedure was to develop a method that yields a more accurate registration process.
The main idea for this proposed method consists of the reconstruction of the patient's surface using multiple point clouds stitched together using Aruco markers identified in each acquisition. The reconstruction is performed by acquiring multiple frames, using Azure Kinect, of the body to be registered and reconstructing it based on the registration of the position of the Aruco Markers located around the subject. Figure 11 provides a clearer representation of the employed setup and the movements that have to be performed with Azure Kinect.

1. The working principle of this method is based on the workflow implemented for the single frame registration; starting from the assumption that $\mathbf{T}_{\mathrm{AK}}^{\mathrm{H2}}$ has been calculated the premise that any set of points can be transformed into HoloLens 2 coordinate reference frame remains valid. In this case, however, Azure Kinect does not remain in a fixed position since it has to be moved for the acquisition of different faces of the object to be registered. An important remark that has to be noted is that one of the main conditions for a successful registration is that the position of the camera of the last acquired frame has to match the position of the camera during the calibration procedure. Since the goal remains the hologram-to-patient registration, the same equation applies for this scenario:

$$T_{CT}^{H2} = T_{CT}^{AK} \cdot T_{AK}^{H2} \tag{4}$$

   For means of this section, the patient's surface will be referred to as the reconstructed object.
2. Using HoloLens 2 holographic interface and after Multi Frame Registration modality has been selected the user can proceed with the acquisition of the frames using Azure Kinect. The holographic menu gives the option for frame acquisition by using the proper button or by using voice commands; once the button is triggered HoloLens 2 publishes a message with the frame number that is being acquired; at the same time, the subscriber to the frame number topic initializes the callback function in charge of the frame acquisition.
3. As mentioned before, the idea of the whole surface reconstruction is based on the overlapping of the generated frames by applying transformation matrixes obtained through the recognition and posterior registration between the point clouds corresponding to the position in space of a N number of Aruco markers placed around the object in the scene in subsequent frames. The main challenge for this process emerged due to the fact that the detection of Aruco markers can only be performed in 2D images, thus, a translation to 3D space is required:
   - First, the 2D image of the scene is generated using Azure Kinect. This image undergoes processing to detect the position of the Aruco markers present in the frame by providing the coordinates for the center of the marker; the number of markers that will be used through the entire process depends on the amount of markers detected during the acquisition of the first frame. All the acquired frames must contain the same amount of Aruco markers as the first acquired frame, otherwise, the capture must be repeated.
   - Taking into consideration that RGB and Depth Camera of Azure Kinect do not operate within the same coordinate reference frame nor in the same coordinate system it is necessary to perform a transformation and place $\{p_{i_n,ObjAK}\}$, generated using the RGB camera, into the Depth camera coordinate reference frame. The previous transformation has the purpose of finding an equivalence between the pixels from the 2D image and the point cloud in order to find the position of the Aruco markers in the 3D space. In some cases, the equivalence between the 2D and 3D space might not be achieved due to the nonexistence of a point within $\{p_{i_n,ObjAK}\}$ that matches the position of the Aruco marker in the color image, in those cases where not equivalence could be determined the resulting coordinate corresponds to a point in $\{0, 0, 0\}$ and the frame acquisition has to be repeated. Figure 11 shows the main idea behind the transformation process from RGB to Depth Camera.
   - After a successful remapping process $\{p_{i_n,Aruco}\}$ is produced containing all the Aruco marker positions in the 3D space for the acquired frame. A frame acquisition can be considered as successful upon arrival to this point and the user might proceed with the acquisition of the next frame.

The minimum number of frames required for the reconstruction process is set to three, however, the maximum number of frames has no restrictions. The design of the algorithm requires the acquisition of frames to be performed in a sequential clockwise or counterclockwise manner. Figure 12 shows the suggested movement pattern for a successful reconstruction.

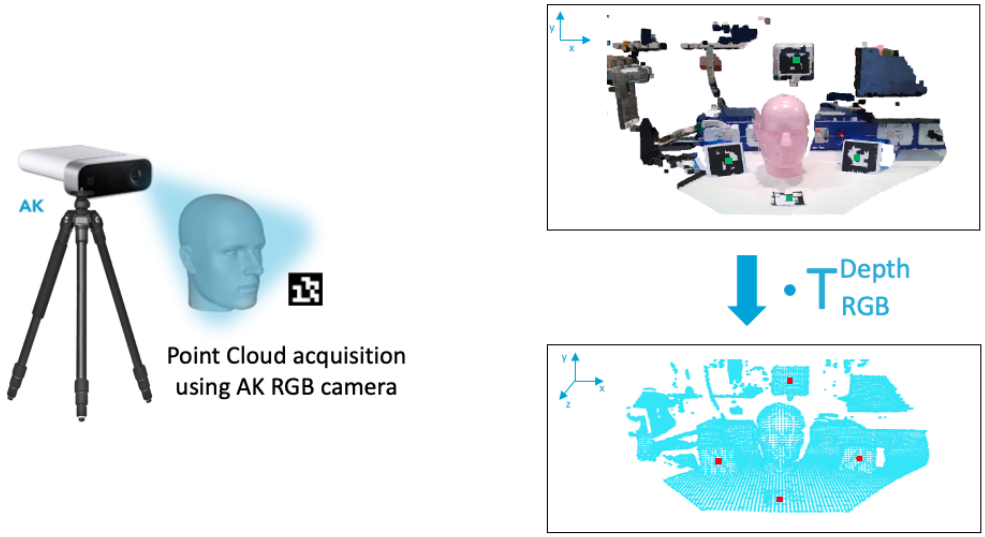


Figure 11: On the left, a practical example of point cloud acquisition using the head phantom and an Aruco marker. On the upper right, a 2D point cloud acquired using the RGB camera of Azure Kinect; green points indicate the detected position of the Aruco markers in 2D. On the lower right, the acquired 2D point cloud is transformed into depth camera's coordinate reference frame by adding the z coordinate, with corresponds to the depth information; Aruco markers in this point cloud are indicated in red.

4. After the frame acquisition is concluded and using the holographic interface the user can proceed with the reconstruction process. Up to this point a N amount of $\{p_{i_n,ObjAK}\}$ and $\{p_{i_n,Aruco}\}$, the first one being the point cloud of the scene of the N acquired frame while the latter corresponds to the point cloud containing the position of the Aruco markers in the 3D space, have been generated and stored. The reconstruction process makes use of both sets of point clouds but the algorithm revolves around $\{p_{i_n,Aruco}\}$.
    - Reconstruction process is also based on the concept of registration. Having a N number of frames the goal of this process is to find $N-1$ transformation matrixes that will place $\{p_{i_n,ObjAK}\}$ in the coordinate reference frame of the last acquired frame to perform an overlapping of all the acquired frames so that all the faces of the object which are not visible during the multiple acquisitions can be included in the final reconstructed object.
    - The transformation matrixes are obtained through the standard process that has been employed so far, having $\{p_{i_n,Aruco}\}$ and $\{p_{i_N,Aruco}\}$ a Fast Global Registration algorithm gives a first alignment of the points. A posterior refinement is achieved using a Point-to-Plane ICP algorithm which produces the transformation matrix $\mathbf{T}_{\mathrm{n}}^{\mathrm{N}}$, where N refers to the final position of the Azure Kinect while n is any position previous to the arrival to the final one.
    - Since $\{p_{i_n,Aruco}\}$ is composed only by points related to the position of Aruco markers in the 3D space the reconstruction is performed using the second generated point cloud $\{p_{i_n,ObjAK}\}$ which corresponds to the complete frame of the scene. Taking into consideration that $N-1$ transformation matrixes have to be produced then the complete set of point clouds $\{p_{i_n,ObjAK}\}$ have to undergo a total of $N-1$ transformations before having all of positioned in the main coordinate reference frame of the system, which we have already defined as the position of the last acquired frame which also matches the position of the Azure Kinect during the calibration procedure. To better explain this concept, a registration with three frames can be exemplified:

$$\{p_{i0}^{2},_{ObjAK}\} = T_{F_0}^{F_2} \cdot \{p_{i_0,ObjAK}\} \tag{5}$$

$$\{p_{i1}^{2},_{ObjAK}\} = T_{F_1}^{F_2} \cdot \{p_{i_1},_{ObjAK}\} \tag{6}$$

For this example each one of the acquired frames besides the last one underwent a transformation so that the point cloud can be translated into $\{p_{i_N},_{ObjAK}\}$, which refers to the last position of the acquisition and has the exact same coordinate reference frame as $\{p_{i_N},_{Aruco}$Being three the number of frames and, considering the previously made statement, the total number of transformation matrixes is two.

- Once every $\{p_{i_n},_{ObjAK}\}$ has been transformed to the coordinate reference frame of $\{p_{i_N},_{ObjAK}$ the following step consists of performing a sum of all the point clouds to obtain a single one containing the totality of the frames acquired. The resulting point cloud can be defined as:

$$\{p_i,_{Rec}\} = \{p_{i0}^{N},_{ObjAK}\} + \{p_{i1}^{N},_{ObjAK}\} + ... + \{p_{iN},_{ObjAK}\} \tag{7}$$

- Final step regarding the reconstruction process corresponds to the outliers removal from $\{p_i,_{Rec}$ using a DBSCAN density-based clustering algorithm [19], which takes into consideration $\varepsilon$and $minPts$ for the removal. $minPts$ refers to the amount of points necessary to consider a region as dense and $\varepsilon$is a calculated parameter that defines if the region contains the enough amount of points to be considered as a cluster or as noise.

5. Upon completion of the reconstruction process, the workflow continues to be the exact same as the one used for the single frame registration. From here on the reconstructed object can be denominated again as patient's surface. The point to reach is still $\mathbf{T}_{\mathrm{AK}}^{\mathrm{CT}}$ using a point cloud of the patient's CT model, $\{p_{i,CT}\}$, which is generated assuming its origin to be located in $\{0, 0, 0\}$. Registration between $\{p_{i,Rec}\}$ and $\{p_{i,CT}\}$ is performed using the same procedure employed during calibration, a Fast Global Registration algorithm gave a first alignment of the points and it was refined using a Point-to-Plane ICP algorithm which produces the transformation matrix $\mathbf{T}_{\mathrm{CT}}^{\mathrm{AK}}$.
6. With $\mathbf{T}_{\mathrm{CT}}^{\mathrm{AK}}$ and $\mathbf{T}_{\mathrm{AK}}^{\mathrm{H2}}$ a simple matrix multiplication was enough to obtain $\mathbf{T}_{\mathrm{CT}}^{\mathrm{H2}}$, as explained in the first point of this section.
7. Once the calculation of $\mathbf{T}_{\mathrm{CT}}^{\mathrm{H2}}$ is completed, this transformation matrix is published as a 4x3 matrix. HoloLens 2, having the subscriber for that topic, receive that information and, once the user selects the option to display the hologram $\mathbf{T}_{\mathrm{CT}}^{\mathrm{H2}}$ replaces the rotation and translation information of the CT model, previously loaded to the head-mounted display, placing the CT hologram over the patient's surface achieving in this was the originally expected hologram-to-patient registration.

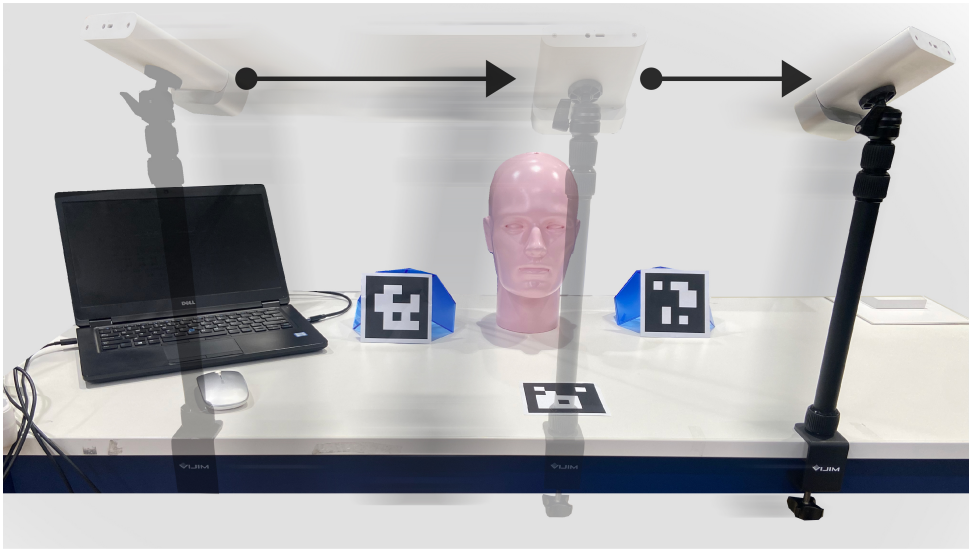


Figure 12: Multi Frame Registration setup. A N number of Aruco Markers can be placed around to object to be acquired, in this case the head phantom. The number of frames that can be acquired depends on the user and, potentially, the size of the object. The main constrain for the current setup is the need to match the last acquisition position with the position used for the calibration object point cloud generation.

The initial consideration was to use an offset of 0.5mm to analyze the behavior of the correction through each iteration. The stopping condition of the offset correction is given by the increment of the error in the six fiducial

points with respect to the previous iteration. In this section, results for the offset correction for both developed modalities are presented.

2.6. Registration Accuracy Assessment

Due to the nature of the expected results, which are a combination of coordinates, an experimental setup, illustrated in Figure 13, was developed to assess the accuracy of the hologram-to-patient registration and superimposition in a quantitative way. An optical tracker system (NDI Polaris Vicra) with a trackable probe was selected as the designated hardware to track the position of the phantom in the 3D space. The assessment of the registration accuracy can be subdivided into two different procedures.

Figure 13: Setup for registration accuracy assessment. An Optical Tracking System (NDI Polaris Vicra) with a trackable probe was used together with a QR code. As long as the probe is in range with the IR field of the system's receptor its location in the Optical Tracking system reference frame can be retrieved.

The idea behind the accuracy assessment consisted of using representative points in the surface of the human head model in both, the real and virtual worlds, to be compared through the acquisition of their coordinates after the hologram-to-patient registration has been obtained in order to have a quantitative measurement of the difference in distance between the holographic and the physical positions.
The complete framework of the registration accuracy assessment follows the below explained sequence.

1. Taking into consideration that the optical tracking system and HoloLens 2 work with different coordinate reference frames the necessity to develop a method to place both of them in the same reference frame emerged; to achieve this the QR shown in Figure 14 was used. HoloLens 2 has the capacity to detect that specific type of patterns and provide the 3D coordinates of the upper right corner of the printed code in its own coordinate reference frame; on the other hand, thanks to the versatility of the optical tracking system, the coordinates of any point in space with respect to its respective reference frame can be acquired by placing the tip of the trackable probe over the point of interest. Knowing this information and by acquiring the same set of points in both systems it is possible to apply the Singular Value Decomposition algorithm to both sets of points to produce a transformation matrix $\mathbf{T}_{\mathrm{OT}}^{\mathrm{H2}}$ that transforms any given point contained in $M_{OT,QR}$ from the optical tracking system to HoloLens 2 coordinate reference frame or vice versa, as expressed by:

$$\{q_{i,H2}\} = T_{OT}^{H2} \cdot \{p_{i,OT}\} \tag{8}$$

(a) QR Code position acquisition using Optical Tracking System. (b) QR Code position acquisition using HoloLens 2.

Figure 14: Experimental setup for the first part of the registration accuracy assessment. By acquiring the position of the upper left corner of the QR code in different positions using both, the Optical Tracking System and HoloLens 2, singular value decomposition algorithm can be applied to retrieve $\mathbf{T}_{\mathrm{OT}}^{\mathrm{H2}}$.

2. Regarding the assessment of the actual superimposition of the hologram over the 3D printed model and, as mentioned before, six points were selected on the surface of the 3D model to be used as target points for the acquisition of the coordinates in both, virtual and real world; these points represent characteristic positions in the head that are beneficial for the placement of the tip of the trackable probe since they are very representative of the structure, thus, misplacement of the probe is less likely to happen. For the acquisition of the points in HoloLens 2 coordinate reference frame a set of six spheres was added to the holographic model in the positions of interest so that, whenever the hologram-to-patient registration is performed, the position of the six spheres can be retrieved for comparison with the positions acquired. After obtaining of the two sets of point clouds $M_{H2,Head}$ and $M_{OT,Head}$ the latter is aligned to the HoloLens 2 coordinate reference frame for analysis purposes by applying:

$$\{q_{i,H2,Head}\} = T_{OT}^{H2} \cdot \{p_{i,OT,Head}\} \tag{9}$$

Once $M_{H2,Head}$ and $M_{OT,Head}$ have been aligned to the same coordinate reference frame it is possible to perform a comparison between them. Figure 15 depicts the experimental setup employed for the second part of the registration accuracy assessment.

(a) Fiducial points position acquisition using Optical Tracking System. (b) Fiducial points position acquisition using HoloLens 2.

Figure 15: Experimental setup proposed for the second part of the registration accuracy assessment. Six fiducial points were selected for acquisition using both, Optical Tracking System and HoloLens 2. By aligning both sets of points to a common coordinate reference frame it is possible to perform a quantitative comparison.

2.7. Hologram Offset Correction

Even when considering the intrinsic error corresponding to the registration between HoloLens 2 and the Optical Tracking System it is important to mention that a shifting factor is still present in the final display of the

hologram with respect to the head phantom, as already studied by Gu et. al. in [20]. Expanding on this concept, due to the intrinsic properties of the depth sensor of HoloLens 2 the reflectivity of some materials has been proved to affect the performance of the device. In this case, it is important to recall that the depth camera of HoloLens 2 is used during the calibration phase.
Considering that this situation will be present in both developed modalities of hologram-to-patient registration the proposed solution to counteract the distance difference between the displayed hologram and the head phantom in the physical world consists of the implementation of an offset correction procedure applied directly to the positioning produced by HoloLens 2. The assumption is that by applying an offset correction using the direction and orientation of the HoloLens 2 depth camera, which can be acquired as an internal parameter of the device, in combination with a correction factor the error in the final superimposition can be minimized.
The concept of offset correction consists of the acquisition of the orientation of the depth camera, this provides a x, y and z values which are directly related to the positioning of the depth camera in relation to the real world; this parameters will be denominated as $P_{depth,H2}$ A correction algorithm was developed in which through a n number of iterations the error between the position of the hologram and the head phantom was gradually reduced. The basic structure of the algorithm is described below.

1. The offset correction was performed using the two sets of fiducial points described in Section 2.6, one corresponding to the holographic model and one corresponding to the points in the physical world; the assumption made was that any correction performed over the distance between the two sets of fiducial points would translate in a general improvement in the positioning of the hologram with respect to the head phantom.
2. The Euclidean distance between each pair of corresponding fiducial points was calculated. The six calculated Euclidean distances were then added to obtain the *totalerror*.
3. Considering a correction factor of 0.05mm, denominated as $C_{correction}$, the offset correction was performed. The correction itself was performed over each coordinate of each one of the fiducial points acquired from the head phantom hologram by subtracting the corresponding coordinate in $P_{depth,H2}$ multiplied by the correction factor as expressed in

$$\{p_{i,corrected}\} = \{p_{i,H2}\} - (P_{depth,H2} \times C_{correction}) \tag{10}$$

   therefore, three corrections were performed over each point, one for x, one for y and one for z.
4. Once the correction factor had been applied to each of the six fiducial points, the Euclidean distance between each pair corresponding points was calculated once again; as before, the results of the six points were added to obtain what will be denominated as corrected error.
5. The process ends with a comparison between the total error and the corrected error. If the corrected error is smaller than the total error then a new iteration of the algorithm is executed using the corrected positions of the fiducial markers as the new reference for the error calculation. On the other hand, if the corrected error is larger than the total error then the algorithm stops since that situation is an indicator that no improvement is being achieved in the minimization of the error between the two sets of points.

3. Results and Discussion

3.1. Registration between Optical Tracker and HoloLens 2

In this section, results for the assessment of registration between Optical Tracker system and HoloLens 2 are presented for evaluation purposes. As mentioned in the previous section, the idea behind this quantitative evaluation consists of acquiring the position of the upper left corner of a QR code in at least three different positions using the Optical Tracking system and HoloLens 2 in order to perform singular value decomposition algorithm so to obtain $\mathbf{T}_{\mathrm{OT}}^{\mathrm{H2}}$.For this assessment, four positions of the QR code were acquired using both devices; this process was replicated 10 times.
Once the points were acquired, $\{p_{i,OT}\}$ was aligned to HoloLens 2 coordinate reference frame through $\mathbf{T}_{\mathrm{OT}}^{\mathrm{H2}}$. Once both sets of point clouds were aligned to the same coordinate reference frame a quantitative evaluation of the registration accuracy was performed by obtaining the error between each pair of points using the Euclidean distance from the OT point to the HL pair, as expressed by:

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \tag{11}$$

The mean error between the four sets of points was calculated for each test performed. Results are shown in Figure 16.

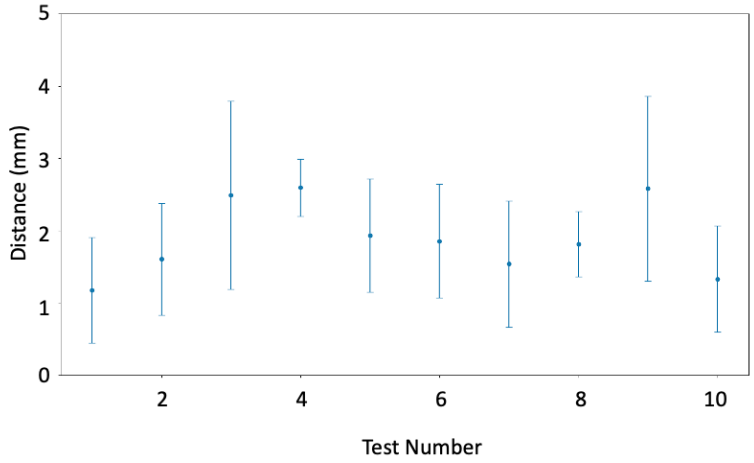


Figure 16: Mean with standard distribution for each trial of the registration between Optical Tracking system and HoloLens 2 using QR code.

As observed in the graph, the registration error oscillates between 1mm and 2mm, calculation of the mean yields an error of $1.728 \pm 0.407mm$. Although the calculated error might seem negligible it is important to take it into consideration for the following sections since the whole assessment of the registration will be based on the use of $\mathbf{T}_{\mathrm{OT}}^{\mathrm{H2}}$.

3.2. Single frame registration

Following the procedure explained in Sections 2.3 and 2.4 the single frame registration procedure was performed. For the evaluation of this process 11 tests were conducted through the execution of the complete workflow. For the evaluation of the hologram-to-patient registration the experimental setup proposed in Section 2.6 was employed; after the registration process was completed the positions of the six fiducial points in both, hologram and head phantom, were acquired. For purposes of the alignment of both sets points in HoloLens 2 coordinate reference frame $\mathbf{T}_{\mathrm{OT}}^{\mathrm{H2}}$ was obtained through the already described method. Here, results for the performance of the algorithm and each participating device are presented.

Calibration and registration parameters	**Azure Kinect**	**HoloLens 2**	**Hologram-to-Patient**
Time for Point Cloud generation [s]	4.85±0.99	——	4.26±0.61
Point Cloud registration time [s]	0.17±0.03	0.14±0.02	0.38±0.16
Iterations before convergence	1.36±0.34	4.27±1.76	3.19±0.27
Registration fitness [%]	0.97±0.02	0.97±0.01	0.57±0.01
Registration RMSE [mm]	0.004±0.00004	0.003±0.00004	0.007±0.00005
Point set correspondence size [points]	925±10.15	901.36±8.50	1130±18
Time to complete process [s]	5.43±1.19	0.27±0.07	4.94±0.89

Table 1: Registration and calibration parameters measured for each subprocess of Single Frame Registration method. For each parameter mean and standard deviation are reported.

Regarding the performance of Azure Kinect during calibration process results are reported in Table 1 on its second column. Some aspects can be commented in this section; as reported, the difference between the time for the point cloud generation and the time to complete the registration seems minimal, however, this is directly related to the number of iterations to reach convergence since, the more iteration required by the algorithm the longer the completion time; the reported results confirms this, the time to complete the process, 5.43±1.19s, is in fact very close to the time that Azure Kinect took to generate the point cloud due to the need of only around one iteration to reach convergence. For this section of the process, and due to the use of the calibration object for the registration, the mean fitness reported is predictably high reaching almost full fitness, this fact can be attributed to the quality of the point cloud acquisition since, having the capacity to cover most of the shape

of the calibration object, the high correspondence between the generated point clouds from Azure Kinect and the CAD model is very much achievable. As for the RMSE of the registration the ideal situation would yield a value close to zero which, in this case, confirms the quality of the registration process.
Table 1 on its third column displays the results for the performance of HoloLens 2 during its calibration process. Registration time is noticeably low, which shows consistency and provides support to deem the algorithm as fast given the results obtained for the same parameter during Azure Kinect calibration. Here, the fitness percentage is also quite high along with the correspondence set. It is complicated to speak about the performance of HoloLens 2 in this context since, as already known, all the processing is performed by the workstation, therefore this analysis does not necessarily correspond to the performance of HoloLens 2. Moreover, and as mentioned before, the obtained data proves to be useful when determining whether the algorithm can be reliable in terms of speed.
Results for the performance of Azure Kinect in the hologram-to-patient registration are reported in Table 1 on the fourth column. In this section, the generation of the point cloud time stays in line with the previously obtained results during the calibration process, a slight difference can be denoted by is deemed as negligible. When focusing on the registration time an increment is observed together with a decrease in the registration fitness; both of this changes follow a very intuitive explanation. Given the fact that Azure Kinect cannot scan all the faces of the head phantom in one single acquisition, due to its geometrical shape and size, then only a smaller portion of the phantom can be acquired. When registering with the CAD model point cloud of the phantom, which corresponds to a complete geometrical version of the phantom, the correspondence area will be smaller, giving as result a lower fitness. The increase in number of iterations can be explained taking from starting point the chosen registration algorithm; as mentioned in previous sections, an initial alignment is provided followed by a refinement algorithm based on ICP registration. The whole process relies on the initial guess for the first alignment and, given the fact that in this case the correspondent area is much smaller than the calibration object, the registration might not be successful due to the refinement not finding a proper point cloud alignment, thus, more registration iterations are required in order to achieve convergence.
One of the key points for the assessment of the registration accuracy was the use of fiducial points. As explained in Section 3.1, the use of these points for the registration accuracy allowed for a quantitative evaluation since of the distance between the acquired points using the Optical Tracking System and the ones acquired through HoloLens 2. Results for this evaluation are shown in Figure 17, for this section we will refer specifically to the blue boxes; the mean distance of each marker across the 11 performed tests is depicted in the graph. According to the illustrated data the hologram superimposition presents a shifting of 8.34±0.73mm for each marker, which can be translated to a general shifting of the whole structure. As mentioned in previous sections, the reason behind the distance difference between the two point clouds can be attributed to an intrinsic offset of the depth camera of HoloLens 2, for this reason, offset correction results are presented in comparison to the results before the correction. Green boxes in Figure 17 correspond to the distance difference after offset correction between the fiducial markers acquired with the Optical Tracking system and the ones acquired using HoloLens 2; it is noticeable that the distance between both sets of points was reduced, the difference in this case equals 3.85±0.19mm. Provided data shows a clear improvement in terms of error reduction after offset correction was performed.

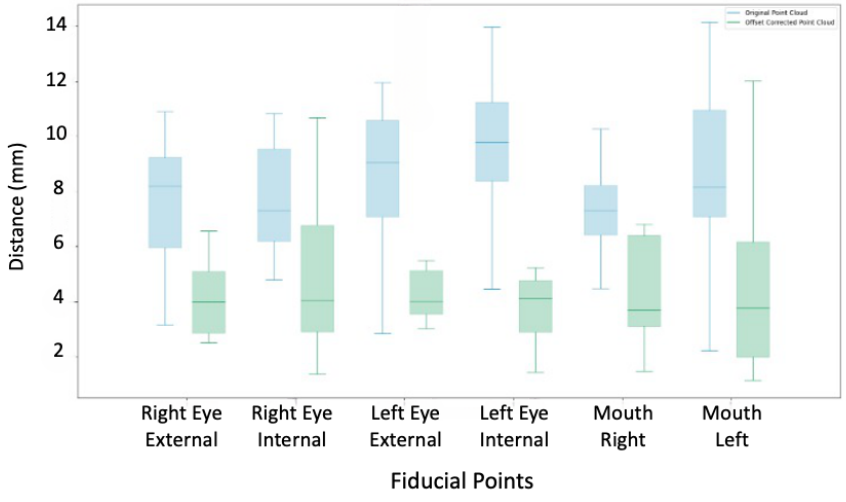


Figure 17: Single frame registration fiducial points difference between real world and virtual environment positions.

Due to the variance in the errors between tests the correction algorithm did not perform the same number of iterations every single time, a probable reason can refer to the different angles with which the H2 point clouds were acquired during the calibration phase. In fact, and as visible in Figure 18, the mean number of iterations underwent by the 11 tests performed was 15, which, translated to offset corrected yields a value of 0.75mm. The depicted curve refers to the total distance error of the fiducial points with their holographic homologous.

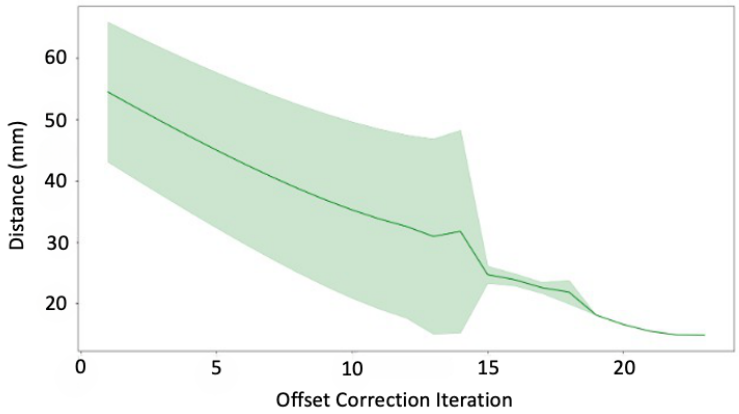
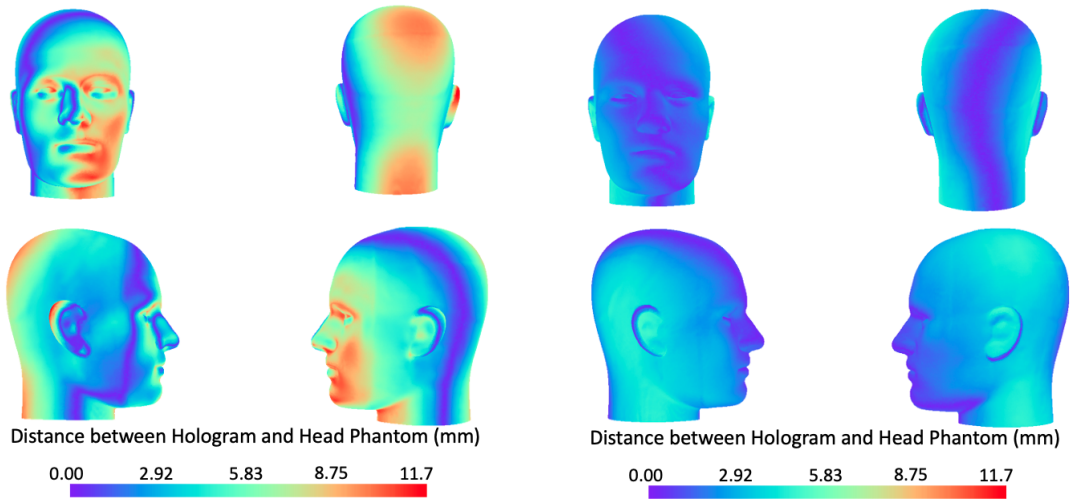


Figure 18: Single Frame registration offset correction. Through each iteration the cumulative distance between fiducial markers and their virtual homogeneous decreases.

In order to provide a clear idea of the actual result of the whole registration process along with the correction offset procedure, Figure 19 provides a colormap representation of the final distance between the real world position of the head phantom, obtained through the use of the Optical Tracking System, and the hologram in the virtual scene displayed by the HoloLens 2. The mean distance between the 3D printed head phantom and its holographic representation seen through HoloLens 2 across the 11 trials was calculated. Point-to-point accuracy evaluation yielded 8.27±1.21mm as the mean difference between the head phantom and the visualized hologram. It is possible to see that the largest differences in distance between both point clouds are located in the back of the head, being the upper part and the frontal plane the regions where the distance minimizes.



(a) Distance before offset correction.

(b) Distance after offset correction.

Figure 19: Colormap of the mean distance between the positioning of the hologram with respect to the real world position of the head phantom across the 11 performed tests. On the left, results before offset correction are shown; on the right, results after offset correction are displayed.

Also in Figure 19 distance between hologram and head phantom after offset correction are presented. Many facts can be commented about this part of the results; firstly, the homogeneity in the frontal part of the head phantom representation shows the improvement of the registration in this area of the phantom, in other words, the offset correction algorithm appears to present the best results in this zone of the model. The main issue

continues to be the posterior part of the phantom since the distance between points reaches its maximum there. Altogether, offset correction for the Single Frame registration method can be deemed as successful; recalling that there is another intrinsic error corresponding to the registration between HoloLens 2 and the Optical Tracking System it can be said that a way to achieve an almost fully accurate alignment can be achieved.

3.3. Multi frame registration

Testing of the multi frame registration method followed the same workframe proposed for the single frame method evaluation. 11 tests were conducted and, as previously explained, the design of the algorithm supposes the acquisition of three different frames for the point cloud generation of the head phantom. In this case, the main focus of the reported results is centered around the performance of the algorithm, no results for the calibration part nor the performance of HoloLens 2 are reported.
For this section, in order to avoid redundancy in the results, parameters for the calibration process of HoloLens 2 and Azure Kinect are not reported. Table 2 depicts the performance of Azure Kinect during hologram-to-patient section. Taking into consideration that these results refer to the use of the reconstructed surface the first value that pops up in comparison to the single frame method is the fitness of the registration process; one of the main goals of this proposed method was to increase the size of the correspondence set in order to improve the registration outcome. Having a more complete point cloud covering a wider area of the object to be registered is the reason behind this improvement.
It is also possible to notice that the number of registration iterations remains stable. Due to the higher number of points used for the registration in comparison to the previous method the first alignment provided by the fast global registration algorithm produces a better outcome, thus, the refinement achieves a better registration after just one iteration. Point cloud registration time stays in line with the previously reported results, this provides further confirmation in regards to the reliability of the registration algorithm.

Registration parameters	**Hologram-to-Patient**
Point Cloud registration time [s]	0.16±0.06
Registration iterations before convergence	1±0
Registration fitness [%]	0.76±0.01
Registration RMSE [mm]	0.006±0.00001
Point set correspondence size [points]	1498±23
Time to complete process [s]	0.70±0.26

Table 2: Azure Kinect Hologram to Patient registration results.

Table 5 provides the results related to the acquisition and reconstruction of the head phantom surface. Considering that each test consisted of the acquisition of three different frames the mean time for this process shows a very foreseeable result which complies with the previously obtained results. As for the reconstruction time, its elevated value can be attributed to the inclusion of the filtering process for the removal of outliers in the final point cloud. Even when this specific procedure can be consuming in terms of computational time it remains an important part for the registration process.

Acquisition and reconstruction parameters	**Azure Kinect**
Frame acquisition and point cloud generation time [s]	16.58±0.69
Reconstruction time [s]	43.60±8.79

Table 3: Azure Kinect frame acquisition and reconstruction results.

Same evaluation for the fiducial points was performed, the mean distance of each marker between the hologram and the acquired positions using the Optical Tracking system across the 11 performed tests is presented in Figure 20. First, referral to the blue boxes corresponding to the results prior to applying offset correction is made. Just like the previous evaluation using the single frame method, a general tendency of displacement can

be observed for the six markers which translates, once again, in the shifted placement of the hologram over the head phantom. Mean distance across the six markers is equal to 9.83±1.04mm. For this method, it can be observed that the displacement distance increased with respect to the previous section; even when a good reconstruction process has been performed the constrains of the manual movement of the Azure Kinect for the frame acquisition emerge as the first explanation for this situation, the difficulty to place the camera in the exact same place as its position during the calibration procedure has a huge impact in the final registration due to the modification of the reference frames. However, and as mentioned before, up to a certain extent this misplacement can be also attributed to HoloLens 2 intrinsic inaccuracies. Regarding the green boxes of Figure 20, they make reference to the obtained results for each fiducial marker after offset correction was performed over the point cloud. Even when a slight improvement can be observed it cannot be deemed as enough to consider HoloLens 2 depth camera offset as the responsible for the relatively large differences between holographic and physical fiducial points. Mean distance between hologram and head phantom fiducial points equals 7.15±0.87mm in this case. In fact, at this point it is clear that the reconstruction process might have to biggest impact for the obtained results.

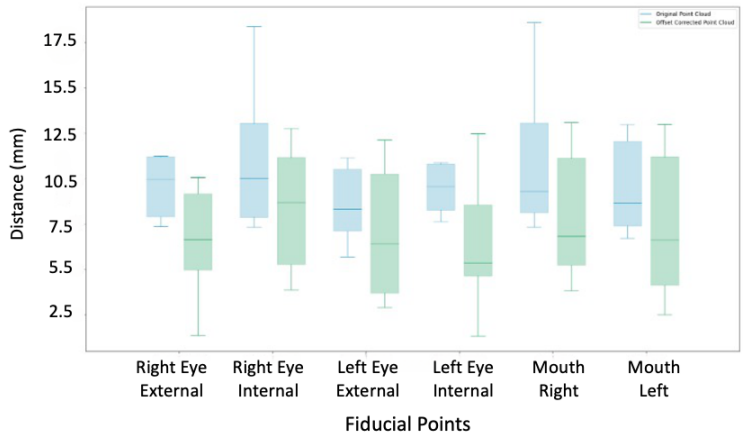


Figure 20: Multi frame registration fiducial points difference between real world and virtual environment positions.

Results corresponding to the performance of the offset correction algorithm are presented in Figure 21. It can be noticed that in this case the mean number of iterations corresponds to 13, which, translated to millimeters yields an offset correction of 6.5mm; just as in the previous method, the number of iterations does not remains constant through all of the tests. As for the curve behavior, the characteristic improvement presented also for the Single Frame registration method remains present. Even when both methods appear to benefit from the offset correction process it has to be mentioned, once again, that the performance of the Multi Frame registration method is far worse than the Single Frame method.

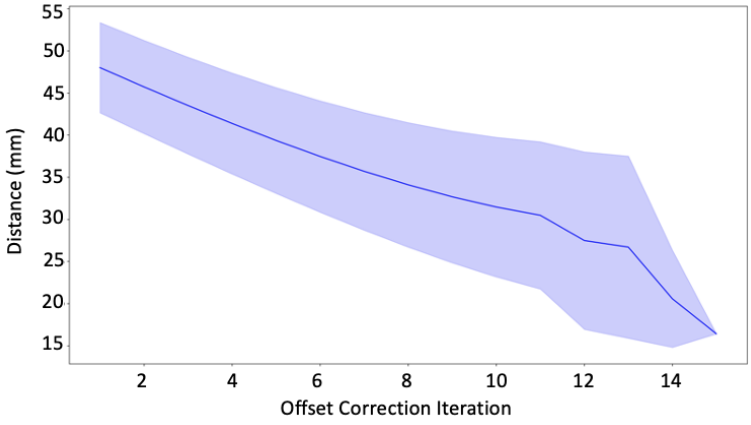
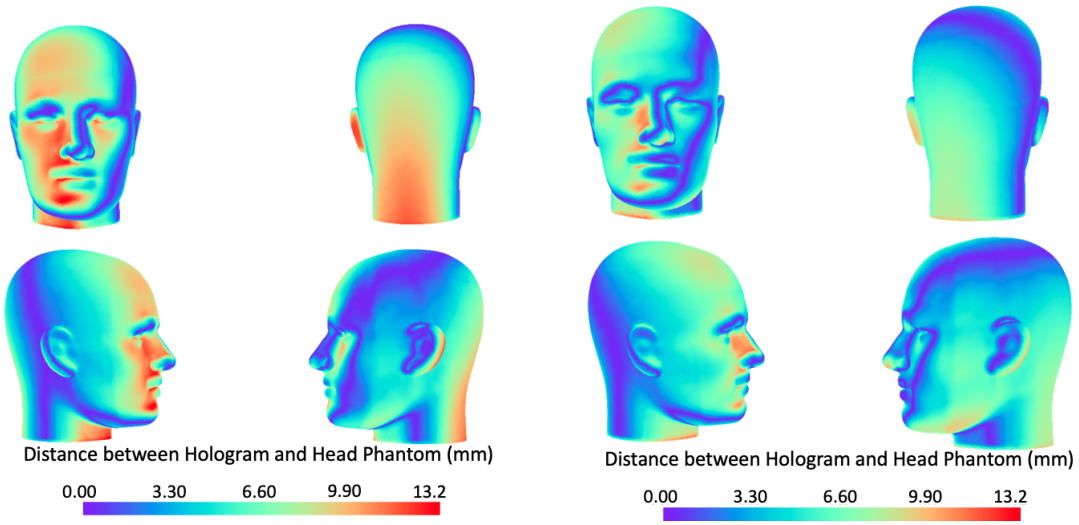


Figure 21: Multi Frame registration offset correction. Through each iteration the cumulative distance between fiducial markers and their virtual homogeneous decreases.

Final analysis for this section corresponds to the colormap of the distance between the positioning of the real and virtual world is provided in Figure 20. Here the shifting of the produced hologram with respect to the head phantom position in the real world becomes more evident; the largest difference can be appreciated in the back and front part of the model while on the sides the point cloud seems to have a better alignment. For the Multi Frame registration method an average distance of 10.17±1.81mm for point-to-point evaluation accuracy. Once again, the dominant factor for the obtained results continues to be the reconstruction process which, even when increasing the fitness of the registration and the number of correspondent point could have deformed the shape of the head phantom in some areas resulting in the evident distance differences. Regarding the results obtained after offset correction, even when a slight improvement in the distance between hologram and head phantom can be observed the difference continues to be too significant to deem the offset correction as the solution. An important remark in this case, however, is the fact that the mean point-to-point accuracy assessment indicates yields an error of 8.81±1.47mm; this value indicates that the general correction for the complete representation of the head phantom was not as significant as the one achieves for the Single Frame method. The assumption here remains that the quality of the reconstruction has a direct impact over the registration quality and, therefore, the distance between corresponding points.



(a) Distance before offset correction.

(b) Distance after offset correction.

Figure 22: Colormap of the mean distance between the positioning of the hologram with respect to the real world position of the head phantom across the 11 performed tests for the Multi Frame registration method. On the left, results before offset correction are shown; on the right, results after offset correction are displayed.

4. Conclusions

In this thesis, a method for hologram-to-patient registration through the combined use of Azure Kinect and HoloLens 2 was presented. Upon completion of the development of this workframe and having performed the pertinent analysis a comparison with the initial aim of the development can be performed.
The main objective of the development of this project was to achieve a simple, fast, marker free and accurate system which allowed to perform a proper hologram-to-patient registration. Besides the mentioned aim the focus was to implement a computational efficient structure following the state-of-the-art advances in the field of Mixed Reality in Image Guided Surgery.
Starting from the development and testing of the single frame registration method, there are many aspects than can be pointed out. First of all, the required time to complete a registration process stands as one of the striking aspects, if the correct conditions are met the hologram-to-patient registration can be achieved in a relatively short time. Simplicity of the implementation represents another interesting aspect to highlight, contrary to what other solutions propose, the workflow is designed to facilitate the use of the interface to the user as much as possible; from a very intuitive calibration process to a patient registration with very little intervention from the user, the complete process can be deemed as user-friendly. Regarding the accuracy of the system and taking into consideration the explained intrinsic factor that might interfere it can be established that, even when improvable, the proposed solution can be seen as a reliable tool for the hologram superimposition. All

this affirmations, however, are subject to the testing of the solution by external users along with further testing with different anatomical structures which can provide heterogeneity to the obtained results and which can test the performance and applicability of the developed work to different surgical fields.
Along with the positive aspects mentioned above there are some considerations that have to be taken into account. Firstly, and due to the phase of development in which the application is, the placement of the camera will still play a crucial role in the outcome of the algorithm; an extremely precise alignment is not required, however, a positioning with a wide field of view might prove beneficial for the final registration. Another point that needs to be mentioned is the fact that, after scanning, the patient cannot be moved since the registration is performed using a single frame which produces a static hologram. All the previously mentioned points result in constrains when thinking about applying the developed solution in real-life scenarios. Finally, and as denoted in previous sections, when dealing with the point cloud acquisition for the patient's surface it is important to consider that the amount of covered surface will have a direct impact in the quality of the registration, position of the camera, once again, results as an important fact to consider.
Focusing now on the multi frame registration method there are some noticeable aspects to remark. First of all, the increase of the correspondence set along with the fitness when performing the hologram-to-patient registration supports the idea of an improvement in the registration procedure in comparison with the single frame registration algorithm; in this specific area, the developed solution seems to fulfill its purpose by improving the fitness of the registration process. However, there are many aspects which have plenty of room for improvement. Firstly, for this study and for the testing of the solution the Azure Kinect displacements where performed in a manual way; being a constrain related to the phase of development of the solution, it should be remarked that these type of movements result in another source of uncertainty for the system. These movements, in combination with the almost impossible task of placing the device in the exact same position as the one used during the calibration procedure will derive in the inconsistencies obtained. The use of Aruco Markers appears to be a good solution for the reconstruction of the patient's surface reconstruction due to its easiness of use and the wide range of pre built functions that have been developed for working with them; in this specific case and, due to the environmental limitations, only four of them were used. Hypothetically, the use of more Aruco Markers in the scene would not only benefit the reconstruction process but also the registration procedure.
Regarding the offset correction performed during the study, further analysis and testing needs to be performed in order to have a better idea of how to implement a real-time solution for this issue. Initial idea for the suppression of this factor could be the elimination inside the developed application for HoloLens 2, however, and as already mentioned, further testing should be performed to find out the exact impact of this inaccuracy.

4.1. Limitations and future work

Having providing the previous comments in regards to this solution it can be established that until further testing, this application holds the promise of becoming a useful tool for the hologram-to-patient registration. The main limitations of the system refer to a lack of validation using other anatomical structures which would prove its efficiency when exposed to different scenarios.
As mentioned before, the static nature of the registration might limit the performance of the solution in real life scenarios due to the many variables that are present in an Operating Room which make it difficult to guarantee that the patient will remain in a single position during the whole surgical procedure.
Regarding the multi frame registration method, it has been mentioned that the inaccuracy caused due to the movement of the camera in a non precise manner has a huge impact in the quality of the registration. The clear solution for this issue would be the further development of a device which carries the camera around the patient, with fixed acquisition positions in order to have a more stable acquisition procedure and, therefore, a better reconstruction which would massively improve the hologram superimposition. Upon refinement of this procedure the possible outcomes could suggest an exponential improvement in the hologram-to-patient registration.
With all the mentioned points, the positive outcomes of this study suggest that this solution holds the promise of setting a strong foundation for the development of commercial solutions for real-life scenarios. Room for improvement is clear and evident, but so is the theoretical scope of this workflow.

Abstract in lingua italiana

Realtà mista, un concetto definito come la fusione tra il mondo virtuale e quello fisico attraverso l'uso di elementi reali e olografici, promette di diventare uno strumento utile durante la chirurgia guidata dalle immagini; gli ultimi sviluppi in questo campo hanno raggiunto una serie di applicazioni che migliorerebbero l'esperienza dei chirurghi fornendo la visualizzazione di strutture anatomiche complesse in 3D e di strumentazione chirurgica in tempo reale attraverso l'uso di display montati sulla testa. Tuttavia, la necessità di un'ulteriore convalida

per dimostrare l'accuratezza di questi metodi è necessaria per la sua implementazione all'interno della sala operatoria. Uno dei campi di interesse corrisponde alla valutazione della registrazione paziente-ologramma, che è il processo responsabile dell'accurata sovrapposizione di elementi virtuali sulle strutture anatomiche del mondo reale. In questa tesi viene presentato lo sviluppo di un metodo di registrazione ologramma-paziente semplice e privo di marcatori utilizzando un display montato sulla testa (HoloLens 2, Microsoft, Washington) in combinazione con una videocamera di profondità esterna (Azure Kinect, Microsoft, Washington); la superficie anatomica selezionata per lo sviluppo del metodo è costituita da un fantoccio di testa umana stampato in 3D con il corrispondente modello virtuale CT. Il lavoro presentato include due metodi per l'acquisizione delle superfici del paziente utilizzando uno o più fotogrammi a favore di una migliore qualità di registrazione ologramma-paziente insieme a un algoritmo di correzione dell'offset per la regolazione degli errori dovuti alle imprecisioni intrinseche di HoloLens 2. È stata eseguita una valutazione dell'accuratezza di entrambi i metodi di registrazione utilizzando un sistema di tracciamento ottico; la distanza di errore media registrata tra il fantoccio della testa e il modello virtuale sovrapposto visualizzato tramite il display montato sulla testa era di 8,34±0,73 mm prima della correzione dell'offset e di 3,85±0,19 mm dopo la correzione dell'offset per il metodo di registrazione a fotogramma singolo, per il metodo a più fotogrammi i risultati ottenuti sono stati 9,83±1,04 prima della correzione dell'offset e 7,15±0,87 mm dopo la correzione dell'offset. Per quanto interessanti possano sembrare questi risultati, l'accuratezza ottenuta non è sufficiente per l'attuazione degli interventi chirurgici, tuttavia, il lavoro presentato pone le basi per miglioramenti rispetto al metodo sviluppato per sviluppi futuri.

Parole chiave: display montato sulla testa, nuvola di punti, realtà mista, registrazione, scomposizione di un valore singolo, telecamera di profondità

Acknowledgements

It has been almost three years since I began this chapter of my life, through ups and downs, success and frustration I have, somehow, managed to finally get to this point. Upon completion of this thesis and my studies there is nothing else I would like more than to acknowledge all those who have been part of this amazing journey. Firstly, I would like to thank Prof. Emiliano Votta for giving me the chance to be a part of this research group and allowing me to work in such an interesting topic. I would also like to thank my supervisor, Maria Chiara Palumbo, for the support, patience and kindness through the development of this project; thanks for guiding me through a topic of which I had very little to no idea, it is not every day that you get the chance to work with such a brilliant person. Thank you Mom, Dad, Paulina and Mariana for encouraging me, for standing by my side and for keeping my heart beating for 29 years. Finally, infinite thanks to my one and only, Alexa, thanks for choosing this path together which, I am sure, we will remember for the rest of our lives; thanks for the love, thanks for the patience, thanks for cheering me up and rooting for me even when I felt like I could not go anymore, not even a lifetime would be enough to give you everything back.

References

[1] Paul Milgram, Haruo Takemura, Akira Utsumi, and Fumio Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Telemanipulator and telepresence technologies*, volume 2351, pages 282–292. Spie, 1995.

[2] Mafkereseb Kassahun Bekele, Roberto Pierdicca, Emanuele Frontoni, Eva Savina Malinverni, and James Gain. A survey of augmented, virtual, and mixed reality for cultural heritage. *J. Comput. Cult. Herit.*, 11(2), mar 2018.

[3] Somaiieh Rokhsaritalemi, Abolghasem Sadeghi-Niaraki, and Soo-Mi Choi. A review on mixed reality: Current trends, challenges and prospects. *Applied Sciences*, 10(2):636, 2020.

[4] Enrico Costanza, Andreas Kunz, and Morten Fjeld. *Mixed reality: A survey*. Springer, 2009.

[5] Dhaval Sahija. Critical review of mixed reality integration with medical devices for patientcare. *2022*, 10, 2022.

[6] Hong-zhi Hu, Xiao-bo Feng, Zeng-wu Shao, Mao Xie, Song Xu, Xing-huo Wu, and Zhe-wei Ye. Application and prospect of mixed reality technology in medical field. *Current medical science*, 39:1–6, 2019.

[7] Fazliaty Edora Fadzli, Ajune Wanis Ismail, Mohamad Yahya Fekri Aladin, and Nur Zuraifah Syazrah Othman. A review of mixed reality telepresence. In *IOP Conference Series: Materials Science and Engineering*, volume 864, page 012081. IOP Publishing, 2020.

[8] Marta Kersten-Oertel, Pierre Jannin, and D Louis Collins. The state of the art of visualization in mixed reality image guided surgery. *Computerized Medical Imaging and Graphics*, 37(2):98–112, 2013.

[9] Kevin Cleary and Terry M Peters. Image-guided interventions: technology review and clinical applications. *Annual review of biomedical engineering*, 12:119–142, 2010.

[10] Stijn Keereweer, Jeroen DF Kerrebijn, Pieter BAA Van Driel, Bangwen Xie, Eric L Kaijzel, Thomas JA Snoeks, Ivo Que, Merlijn Hutteman, Joost R Van Der Vorst, J Sven D Mieog, et al. Optical image-guided surgery—where do we stand? *Molecular Imaging and Biology*, 13:199–207, 2011.

[11] Maria Chiara Palumbo. *Unveiling the Invisible: mixed reality based approaches for surgical empowerment*. PhD thesis, Politecnico di Milano, 2023.

[12] Ivo Kuhlemann, Markus Kleemann, Philipp Jauer, Achim Schweikard, and Floris Ernst. Towards x-ray free endovascular interventions–using hololens for on-line holographic visualisation. *Healthcare technology letters*, 4(5):184–187, 2017.

[13] Felix von Haxthausen, Yenjung Chen, and Floris Ernst. Superimposing holograms on real world objects using hololens 2 and its depth camera. In *Current Directions in Biomedical Engineering*, volume 7, pages 111–115. De Gruyter, 2021.

[14] Qichang Sun, Yongfeng Mai, Rong Yang, Tong Ji, Xiaoyi Jiang, and Xiaojun Chen. Fast and accurate online calibration of optical see-through head-mounted display for ar-based surgical navigation using microsoft hololens. *International journal of computer assisted radiology and surgery*, 15:1907–1919, 2020.

[15] T Fick, JAM van Doormaal, EW Hoving, L Regli, and TPC van Doormaal. Holographic patient tracking after bed movement for augmented reality neuronavigation using a head-mounted display. *Acta Neurochirurgica*, 163:879–884, 2021.

[16] Maria Chiara Palumbo, Simone Saitta, Marco Schiariti, Maria Chiara Sbarra, Eleonora Turconi, Gabriella Raccuia, Junling Fu, Villiam Dallolio, Paolo Ferroli, Emiliano Votta, et al. Mixed reality and deep learning for external ventricular drainage placement: A fast and automatic workflow for emergency treatments. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part VII*, pages 147–156. Springer, 2022.

[17] MR Golab, PJ Breedon, and M Vloeberghs. A wearable headset for monitoring electromyography responses within spinal surgery. *European Spine Journal*, 25:3214–3219, 2016.

[18] Arrigo Palumbo. Microsoft hololens 2 in medical and healthcare context: state of the art and future prospects. *Sensors*, 22(20):7709, 2022.

[19] Anant Ram, Sunita Jalal, Anand S Jalal, and Manoj Kumar. A density based algorithm for discovering density varied clusters in large spatial databases. *International Journal of Computer Applications*, 3(6):1–4, 2010.

[20] Wenhao Gu, Kinjal Shah, Jonathan Knopf, Nassir Navab, and Mathias Unberath. Feasibility of image-based augmented reality guidance of total shoulder arthroplasty using microsoft hololens 1. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 9(3):261–270, 2021.