Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN
SCIENZE CHIRURGICHE
Ciclo XXVIII

Settore Concorsuale di afferenza: 06/E3

Settore Scientifico disciplinare: MED/29

PhD Thesis

VALIDATION OF A WEARABLE DEVICE BASED ON AUGMENTED REALITY FOR SURGICAL MAXILLARY REPOSITIONING

VALIDAZIONE DI UN DISPOSITIVO INDOSSABILE BASATO SULLA REALTÀ AUMENTATA PER IL RIPOSIZIONAMENTO DEL MASCELLARE SUPERIORE

Presentata da: **Dr. Giovanni Badiali**

Coordinatore Dottorato — **Prof. Mauro Gargiulo**

Relatore — **Prof. Claudio Marchetti**

Esame finale Anno 2016

ABSTRACT

Aim: We present a newly designed, localiser-free, head-mounted system featuring augmented reality (AR) as an aid to maxillofacial bone surgery, and assess the potential utility of the device by conducting a feasibility study and validation. Also, we implement a novel and ergonomic strategy designed to present AR information to the operating surgeon (hPnP).

Methods: The head-mounted wearable system was developed as a stand- alone, video-based, see-through device in which the visual features were adapted to facilitate maxillofacial bone surgery. The system is designed to exhibit virtual planning overlaying the details of a real patient. We implemented a method allowing performance of waferless, AR-assisted maxillary repositioning. In vitro testing was conducted on a physical replica of a human skull. Surgical accuracy was measured. The outcomes were compared with those expected to be achievable in a three-dimensional environment. Data were derived using three levels of surgical planning, of increasing complexity, and for nine different operators with varying levels of surgical skill.

Results: The mean linear error was 1.70±0.51mm. The axial errors were 0.89±0.54mm on the sagittal axis, 0.60±0.20mm on the frontal axis, and 1.06±0.40mm on the craniocaudal axis. Mean angular errors were also computed. Pitch: 3.13°±1.89°; Roll: 1.99°±0.95°; Yaw: 3.25°±2.26°. No significant difference in terms of error was noticed among operators, despite variations in surgical experience. Feedback from surgeons was acceptable; all tests were completed within 15 min and the tool was considered to be both comfortable and usable in practice.

Conclusion: Our device appears to be accurate when used to assist in waferless maxillary repositioning. Our results suggest that the method can potentially be extended for use with many surgical procedures on the facial skeleton. Further, it would be appropriate to proceed to in vivo testing to assess surgical accuracy under real clinical conditions.

ABSTRACT (ITA)

Obiettivo: Presentare un nuovo sistema indossabile, privo di sistema di tracciamento esterno, che utilizzi la realtà aumentata come ausilio alla chirurgia ossea maxillo-facciale. Abbiamo validato il dispositivo. Inoltre, abbiamo implementato un nuovo metodo per presentare le informazioni aumentate al chirurgo (hPnP).

Metodi: Le caratteristiche di visualizzazione del sistema, basato sul paradigma *video see-through*, sono state sviluppate specificamente per la chirurgia ossea maxillo-facciale. Il dispositivo è progettato per mostrare la pianificazione virtuale della chirurgia sovrapponendola all'anatomia del paziente. Abbiamo implementato un metodo che consente una tecnica senza splint, basata sulla realtà aumentata, per il riposizionamento del mascellare superiore. Il test in vitro è stato condotto su una replica di un cranio umano. La precisione chirurgica è stata misurata confrontando i risultati reali con quelli attesi. Il test è stato condotto utilizzando tre pianificazioni chirurgiche di crescente complessità, per nove operatori con diversi livelli di abilità chirurgica.

Risultati: L'errore lineare medio è stato di 1,70±0,51mm. Gli errori assiali erano: 0,89±0,54mm sull'asse sagittale, 0,60±0,20mm sull'asse frontale, e 1,06±0,40mm sull'asse craniocaudale. Anche gli errori angolari medi sono stati calcolati. Beccheggio: 3.13°±1,89°; Rollio: 1,99°±0,95°; Imbardata: 3.25°±2,26°. Nessuna differenza significativa in termini di errore è stata rilevata tra gli operatori. Il feedback dei chirurghi è stato soddisfacente; tutti i test sono stati completati entro 15 minuti e lo strumento è stato considerato comodo e utilizzabile nella pratica.

Conclusione: Il nostro dispositivo sembra essersi dimostrato preciso se utilizzato per eseguire il riposizionamento del mascellare superiore senza splint. I nostri risultati suggeriscono che il metodo può potenzialmente essere esteso ad altre procedure chirurgiche sullo scheletro facciale. Inoltre, appare utile procedere ai test in vivo per valutare la precisione chirurgica in condizioni cliniche reali.

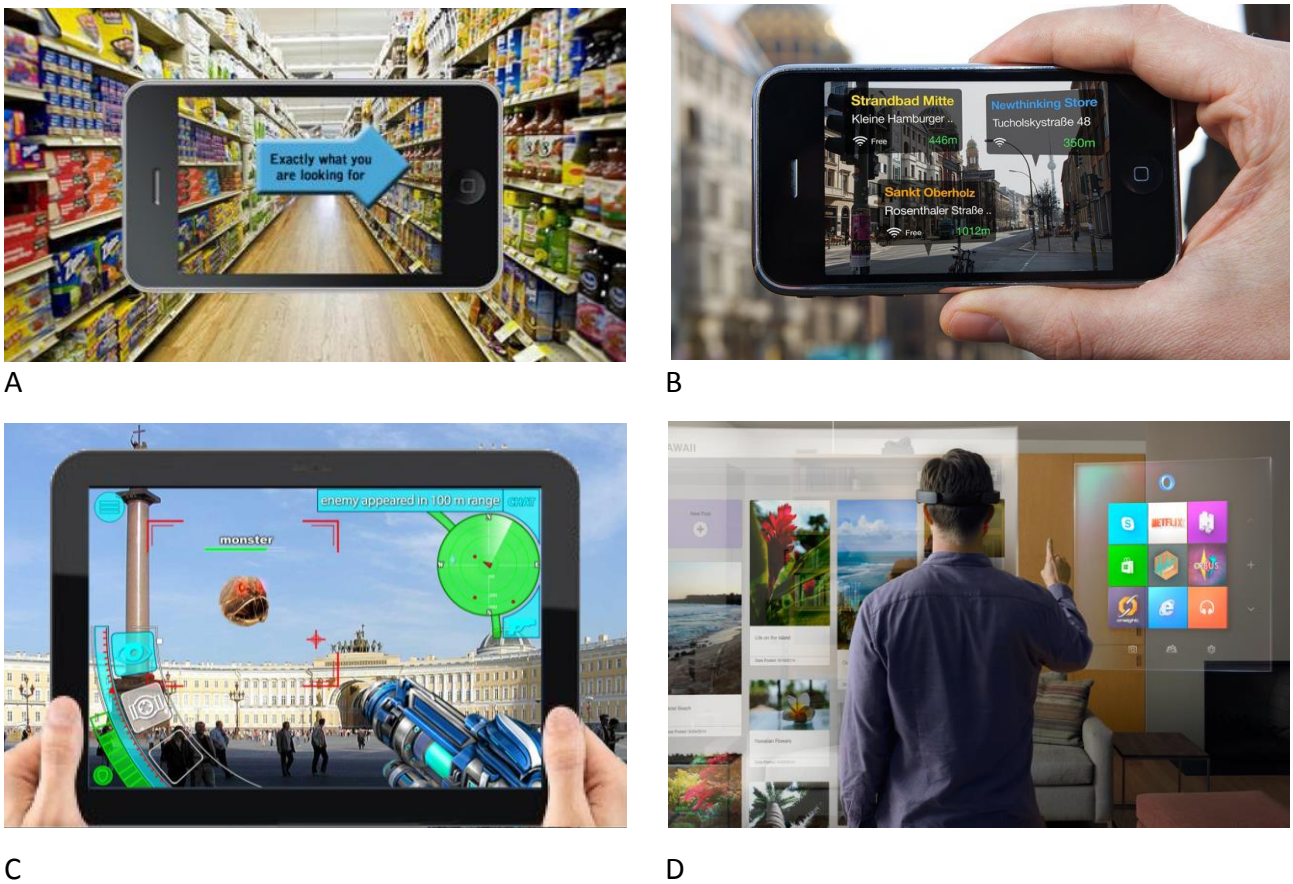
CONTENTS

1 CHAPTER 1: INTRODUCTION

1.1 WHAT IS AUGMENTED REALITY? (DEFINITION AND TAXONOMY)

We could define Augmented Reality (AR) as a real-time direct or indirect view of a physical real-world environment that has been enhanced/augmented by adding virtual computer-generated information to it [1].

Augmented Reality targets at simplifying the user's life by bringing virtual information not only to his immediate surroundings, but also to any indirect view of the real-world environment (Fig. 1.1).



A B

C D

Fig. 1.1 *Various examples of Augmented Reality taken from the internet:* ***A****, AR for customer orienteering on mobile devices;* ***B****, AR for geo-localization of travel services on mobile devices;* ***C****, AR for real-environment gaming on mobile devices;* ***D****, AR for productivity on wearable glasses (Microsoft Hololens). All these examples are mock-ups, but truly representative of the eventual result.*

Milgram's Reality-Virtuality Continuum, defined by Paul Milgram and Fumio Kishino as a continuum that extents between the real environment and the virtual environment, comprises Augmented

Reality and Augmented Virtuality (AV) in between, grouping them as Mixed Reality (MR), as seen in Fig. 1.2 [2].

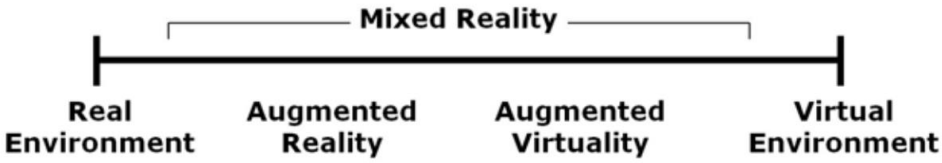
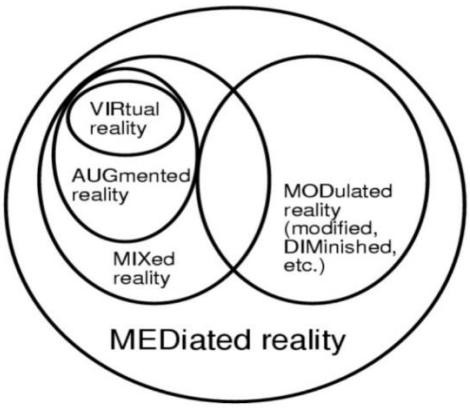


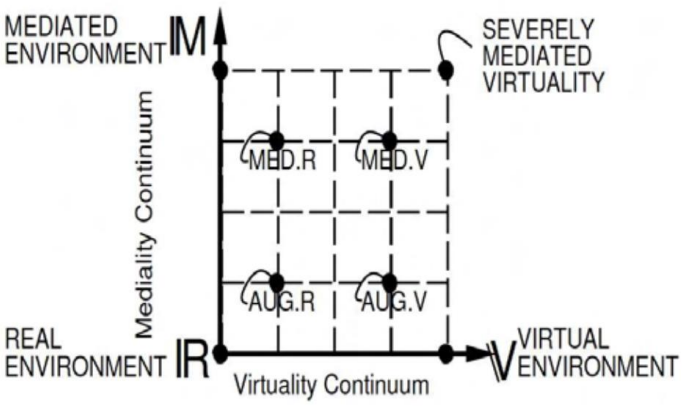
Fig. 1.2 *- Milgram's reality-virtuality continuum.*

Augmented Reality enhances the user's perception of and interaction with the real world. While Virtual Reality (VR) technology or Virtual Environment (as Milgram defines it) completely engages users in a synthetic world excluding the real world, AR technology augments the sense of reality by superimposing virtual objects and cues upon the real world in real time [1].

Milgram's Mixed Reality continuum is a one-dimensional grouping from the Real Environment to the Virtual Environment. In 1994, Mann introduced the second dimension and defined the concept of Mediated Reality as filtering or modifying the view of the real world, rather than just adding to it. Mann's concept of Mediated Reality extends the earlier definitions of AR, VR and MR as shown in Figure 1.3A [3]. Mann suggested also that a Mediality Continuum can be constructed to compliment Milgram's Mixed Reality (or Virtuality Continuum), see Figure 1.3B [4]. The vertical axis represents the amount of mediation or filtering that is being performed in the user view of the real or virtual environment.



A



B

Fig. 1.3 *- Mann's Mediated Reality (A) and Mann's Mediality/Virtuality Continuum (B).*

The Metaverse Roadmap introduces another way of classifying the AR experience [5]. According to Neal Stephenson's concept, the Metaverse is the convergence of a virtually enhanced physical reality and a physically persistent virtual space [6]. The Metaverse Roadmap is based on two key continua (i) the spectrum of technologies and applications ranging from augmentation to simulation; and (ii) the spectrum ranging from intimate (identity-focused) to external (world-focused) [7]. These are defined as follows:

• Augmentation: Technologies that add new capabilities to existing real systems;

• Simulation: Technologies that model reality;

• Intimate: Technologies focused inwardly, on the identity and actions of the individual;

• External: Technologies are focused outwardly, towards the world at large.

The technologies of Augmented Reality, Virtual Worlds, Life Logging, and MirrorWorlds can be arranged within these continua (Fig. 1.4).

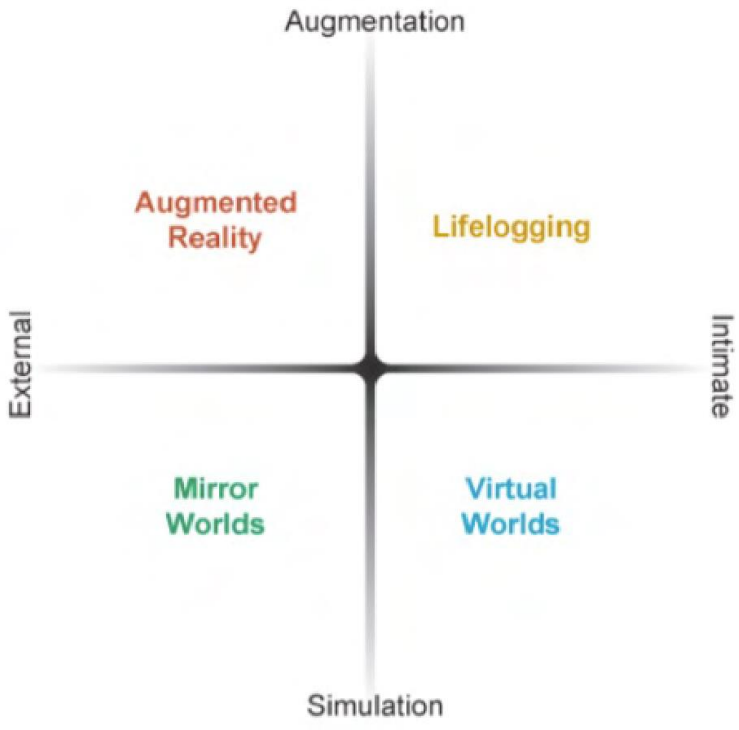


Fig. 1.4 *- The Metaverse Roadmap way of classifying the AR experience, based on two-dimensional merging of the two Augmentation-Simulation and Intimate-External Continua* [7].

1.2 AUGMENTED REALITY: A BRIEF HISTORY AND GENERAL CONCEPTS

The first advent of AR dates back to the 1950s when Morton Heilig, a cinematographer, thought of cinema as an activity that would have the ability to draw the viewer into the onscreen action by taking in all the senses in an effective manner. In 1962, Heilig built a prototype of his vision, which he described in 1955 in "The Cinema of the Future," named Sensorama, which predated digital computing [1]. Later, Ivan Sutherland invented the head mounted display in 1966. In 1968, Sutherland was the first one to create an augmented reality system using an optical see-through head-mounted display (HMD). These are the words of Ivan E. Sutherland accompanying his visionary effort to design "The Ultimate Display" Fig. 1.5 [8]:

> *"...A display connected to a digital computer gives us a chance to gain familiarity with concepts not realizable in the physical world. It is a looking glass into a mathematical wonderland...."*

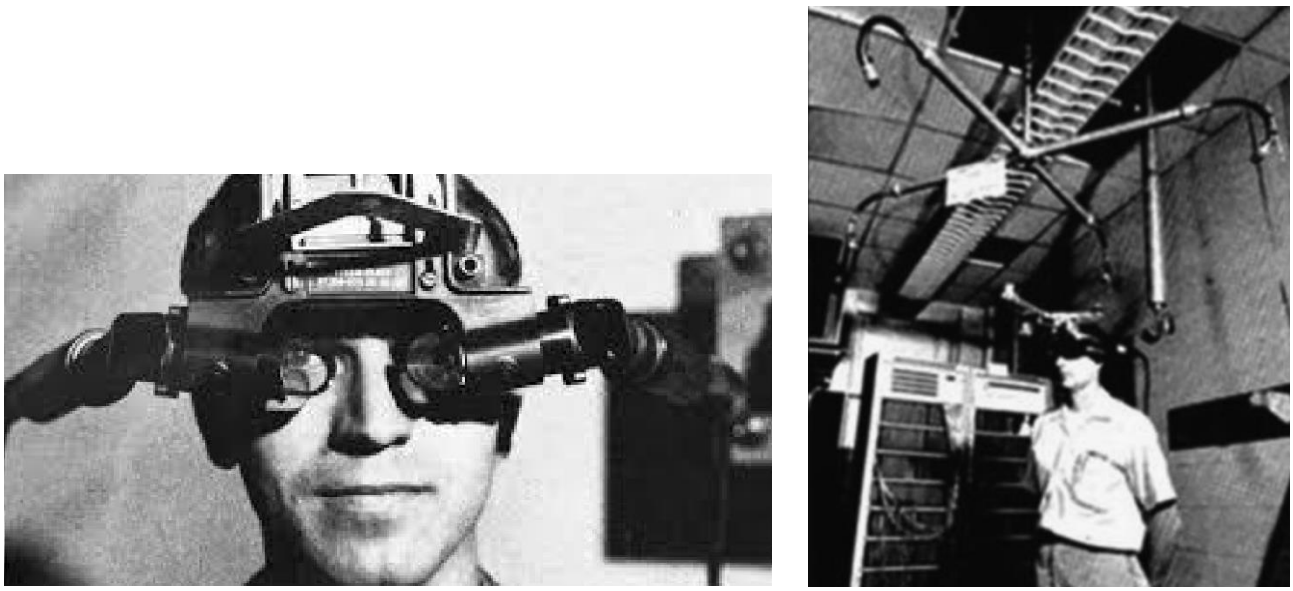


Fig. 1.5 *- Ivan Sutherland's Head Mounted Display.*

However, it was only in the 80s that this revolutionary and pioneering idea could be enforced in practice thanks to the technological advances in computer technology. The term "augmented reality" was ultimately invented in 1992 [9] to describe an experimental system for aircraft manufacturing. Later, it took few more years within the scientific community and among early adopters to establish a general background and a common language to deal with AR. Initially, most

of the research work was targeted to find the possible interactions between AR and other fields of application, to elaborate a common syntax and to determine the basic technological components of AR systems [10].

On the contrary, two studies by Azuma [11], [12] were directed at providing an insight view into augmented reality technology from an applicative standpoint. Azuma defined the three key properties of an augmented reality system:

1) It combines real and virtual content.

2) It is interactive in real time.

3) It is registered in 3D.

The two surveys by Azuma also included comprehensive overviews of all the features and limitations of most of the augmented reality systems until then. Azuma does not consider AR to be restricted to a particular type of display technologies such as HMD, nor does he consider it to be limited to the sense of sight. Azuma also included AR applications that require removing real objects from the environment, which are more commonly called *mediated or diminished reality*.

Recently in 2015, Billinghurst et al. [7] provided a detailed survey of 50 years of research and development in the field of augmented reality. Billinghurst states that the three key properties by Azuma also define the technical requirements of an AR system, namely that it has to have a display that can combine real and virtual images, a computer system that can generate interactive graphics the responds to user input in real time, and a tracking system that can find the position of the user's viewpoint and enable the virtual image to appear fixed in the real world (Fig. 1.6).

	Virtual Reality Replacing Reality	**Augmented Reality** Augmenting Reality
Scene Generation	requires realistic images	minimal rendering okay
Display Device	fully immersive, wide FOV	non-immersive, small FOV
Tracking and Sensing	low accuracy is okay	high accuracy needed

Fig. 1.6 *- Virtual Reality and Augmented Reality technology requirements as described by Billinghurst* [7].

1.3 General Design of an Augmented Reality System

According to Azuma and Billinghurst, we can broadly summarize the key components of any AR system, solely from a hardware standpoint:

• Display unit (two-dimensional or three-dimensional, wearable, hand-held or spatial)

• Computational unit (with or without dedicated graphics card)

• Tracker (embedded in the system or external to it).

Augmented Reality displays can be grouped according on where the display is located with respect to the object and the observer (Fig. 1.7) [13].

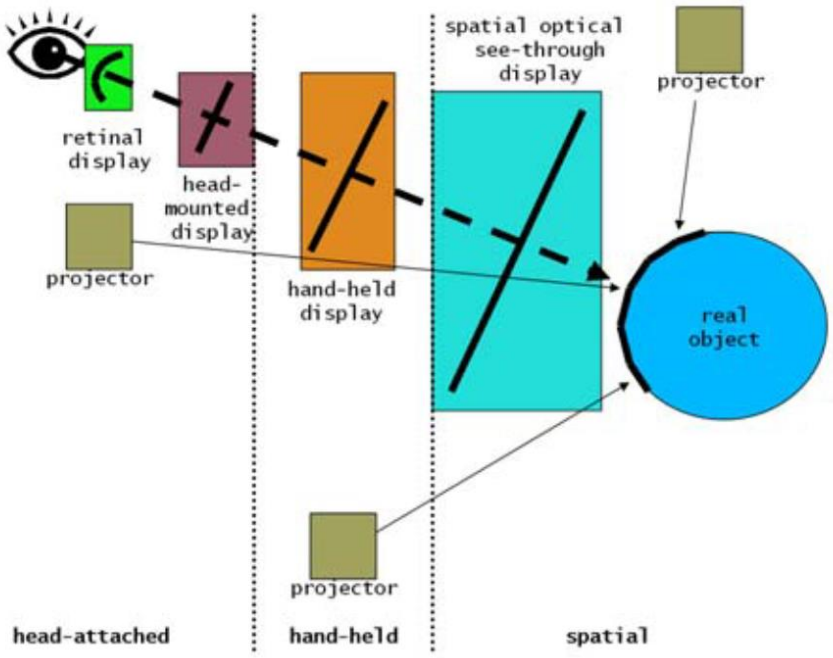


Fig. 1.7 *- Different embodiments of AR displays in relation to their distance from the real object and the observer. Courtesy from Bimber, 2006* [13].

In spatial displays, the display technology is separated from the user and it is rather cohesive into the real environment. In hand-held displays, the AR mechanism is based on the video see-through paradigm: the actual view of the world is acquired by a camera and presented on the display after

being coherently merged with the virtual content. Thus, in these systems, the augmented view is presented sharing the camera viewpoint.

Projector-based AR systems embody a valid and fascinating alternative to standard display-based devices. Projector-based tools provide a high degree of immersion and can cover large surfaces, hence providing wider field of view respect to standard display-based AR systems [14].

The ideal AR system, particularly if designed for complex applications highly demanding of dexterity, should not show any perceivable difference between the user's natural view of the reality and their augmented view. For this goal, the conditions to be satisfied are: 1) accurate registration and 2) ergonomic interaction. According to Azuma et al. [12], «The basic goal of an AR system is to enhance the user's perception of and interaction with the real world through supplementing the real world with 3D virtual objects that appear to coexist in the same space as the real world».

Thus, wearable AR solutions, head-attached according to Bimber [14], but commonly referred to as head-mounted displays (HMDs) should be considered the most ergonomic subclass of AR systems as long as they are capable to deliver a natural and self-centred viewpoint and allow the user to work hands free. Augmented Reality HMDs are now gradually entering a broad range of disciplines.

1.4 Augmented Reality as an Emerging Technology

In 1995, the Information Technology research and advisory firm Gartner® introduced the Hype Cycle curve (Fig. 1.8), giving a graphic description of the level of expectations towards an emerging technology [15]. The curve defines the early stages of an emerging technology from conception to maturity and widespread adoption. This model is essentially related to the market, but offers an immediate and concrete measure of the overall visibility of a new technology, although fairly imperfect from a methodological point of view [16]. By monitoring year by year where the AR technology is located, we can gather an interesting inside view of the degree of maturity of AR applications in the market.

As Figure 1.8 describes, the Hype Cycle is divided into stages, from *technology trigger*, crossing *peak of inflated expectations* – *trough of disillusionment* – *slope of enlightment* to the final *plateau of productivity*.

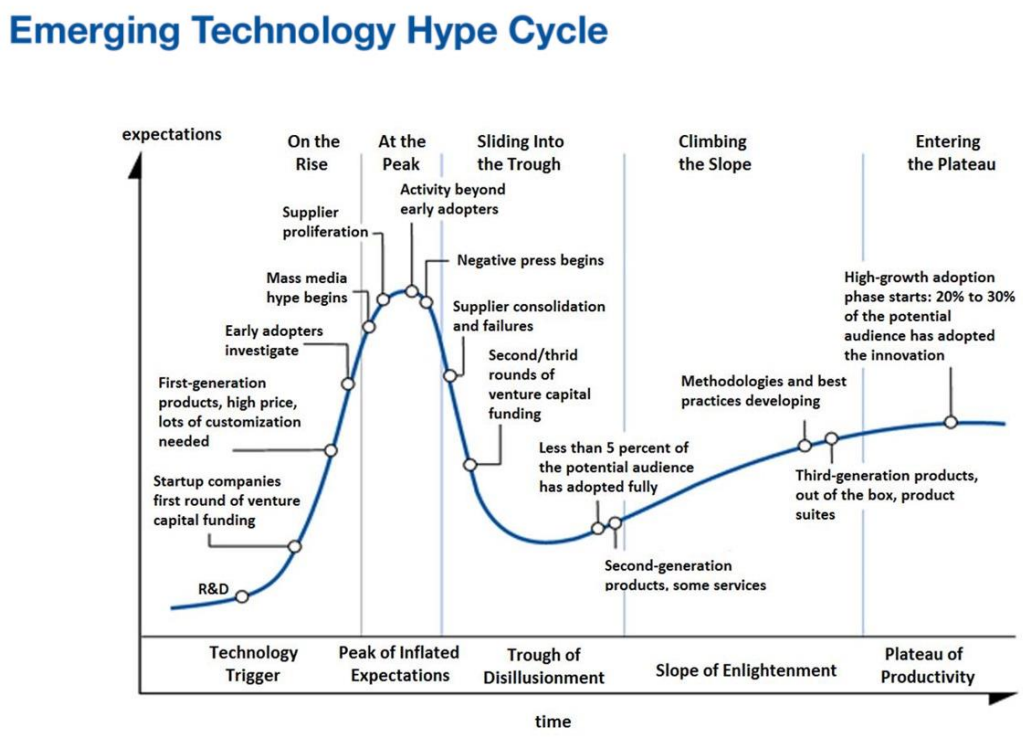


Fig. 1.8 *- Gartner® Hype Cycle for emerging technologies.*

Gartner® describes these stages as follows:

Technology Trigger: A potential technology breakthrough kicks things off. Early proof-of-concept stories and media interest trigger significant publicity. Often no usable products exist and commercial viability is unproven.

Peak of Inflated Expectations: Early publicity produces a number of success stories — often accompanied by scores of failures. Some companies take action; many do not.

Trough of Disillusionment: Interest wanes as experiments and implementations fail to deliver. Producers of the technology shake out or fail. Investments continue only if the surviving providers improve their products to the satisfaction of early adopters.

Slope of Enlightenment: More instances of how the technology can benefit the enterprise start to crystallize and become more widely understood. Second- and third-generation products appear from technology providers. More enterprises fund pilots; conservative companies remain cautious.

Plateau of Productivity: Mainstream adoption starts to take off. Criteria for assessing provider viability are more clearly defined. The technology's broad market applicability and relevance are clearly paying off.

In 2010, only six years back from now, AR was climbing almost at the top of the bell-shaped part of the curve, the *peak of inflated expectations* (Fig. 1.9 left), being one of the new-big-next-things about forthcoming technologies. According to the most recent updates, published by Gartner® in 2015, AR has fell into the *trough of disillusionment* (Fig. 1.9 right), meaning that the excessively enthusiastic stance of the initial stages, which has boosted the early investments of the industrial pioneers and innovators, has resulted in commercial products failing to meet the goals.

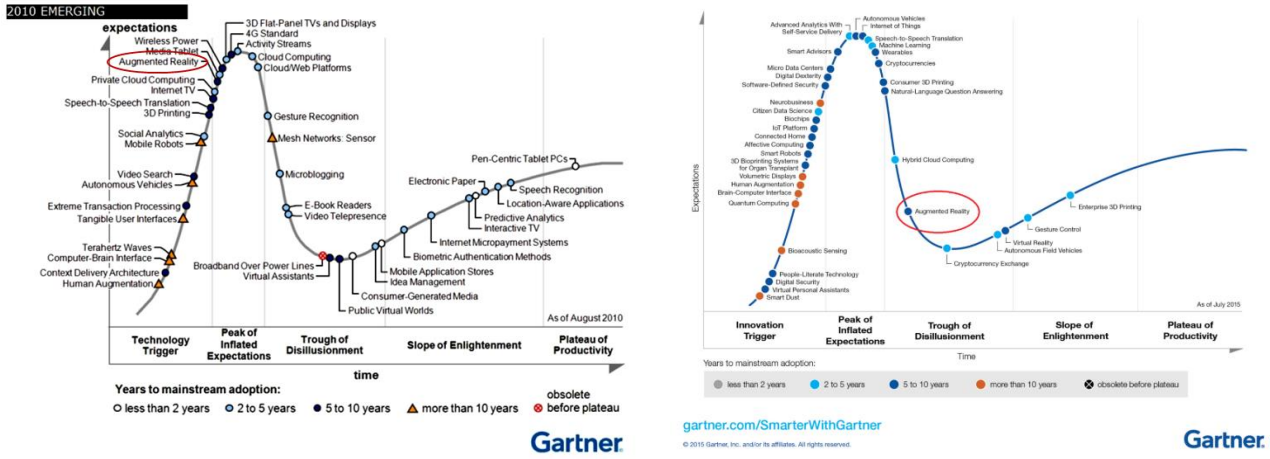


***Fig. 1.9** – Comparison between 2010 and 2015 ranking on the Gartner® Hype Cycle*

This raises a question: is AR a bubble destined to burst? An answer is reported in the PhD thesis written by Eng. Fabrizio Cutolo (EndoCAS Lab, University of Pisa) [17], which is a member of the engineering team involved in this study and co-author of all the work presented in these pages. Indeed, an answer could be suggested by Moore's adaptation [18] of Rogers' bell curve [19], a model about the lifecycle and market penetration of a new technology. Moore's curve considers the presence of a gap (chasm) between early adopters and early majority of customers (Figg. 1.10 - 1.11).

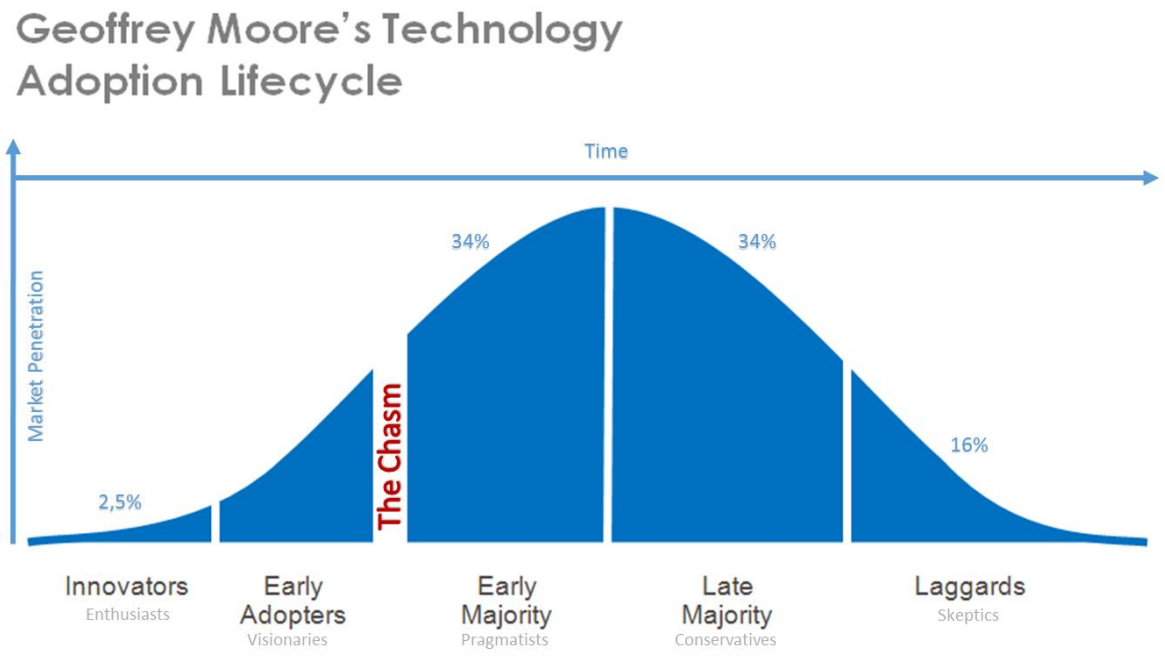


Fig. 1.10 *– Rogers' technology adoption life cycle modified by Moore.*

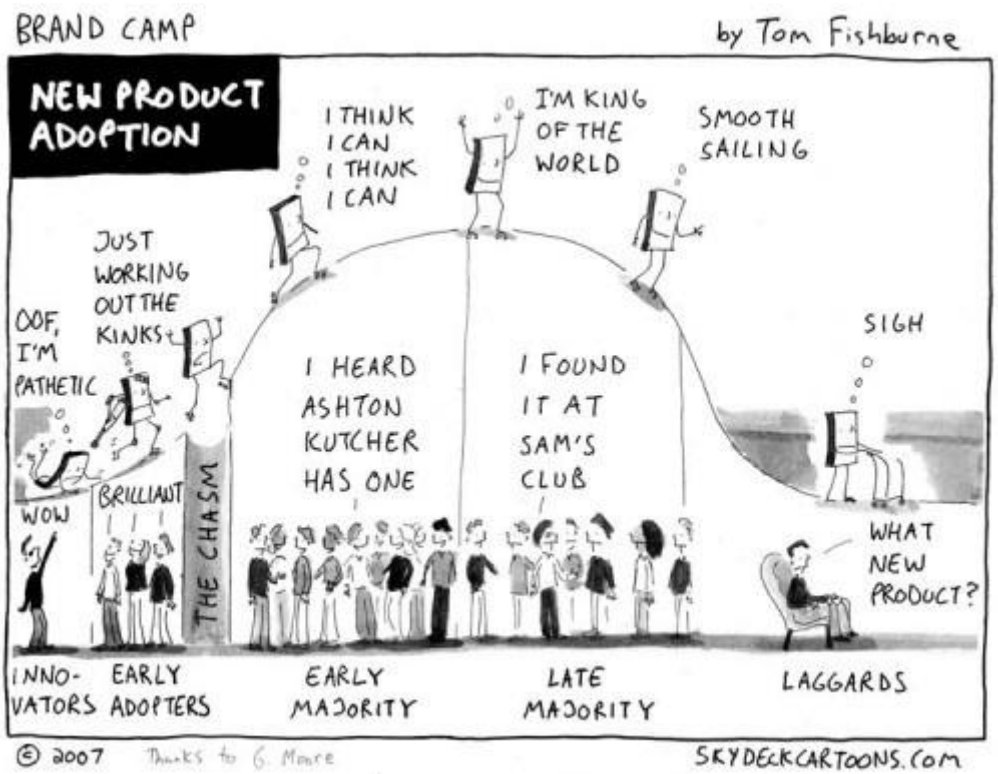


Fig. 1.11 *– Moore's curve depicted as a funny cartoon.*

It should be stated that an emerging technology may also never cross the chasm and reach widespread diffusion. Therefore, the Hype cycle plots the *expectations* towards an emerging technology, whereas the Moore's curve plots the *adoption rate* among consumers. Accordingly, we can hypothesize that most of the Hype cycle is concluded before the Moore's *chasm*. The possibility that a novel technology manages to survive the *trough of disillusionment* and enters the *plateu of productivity* matches with its widespread diffusion, namely with its capacity of crossing the *chasm* [17]. Merging the two graphical models (Hype Cycle and Technology Adoption Life Cycle) we can obtain a broad idea on the issues that AR, as emerging technology, has encountered while entering the market (Fig. 1.12).

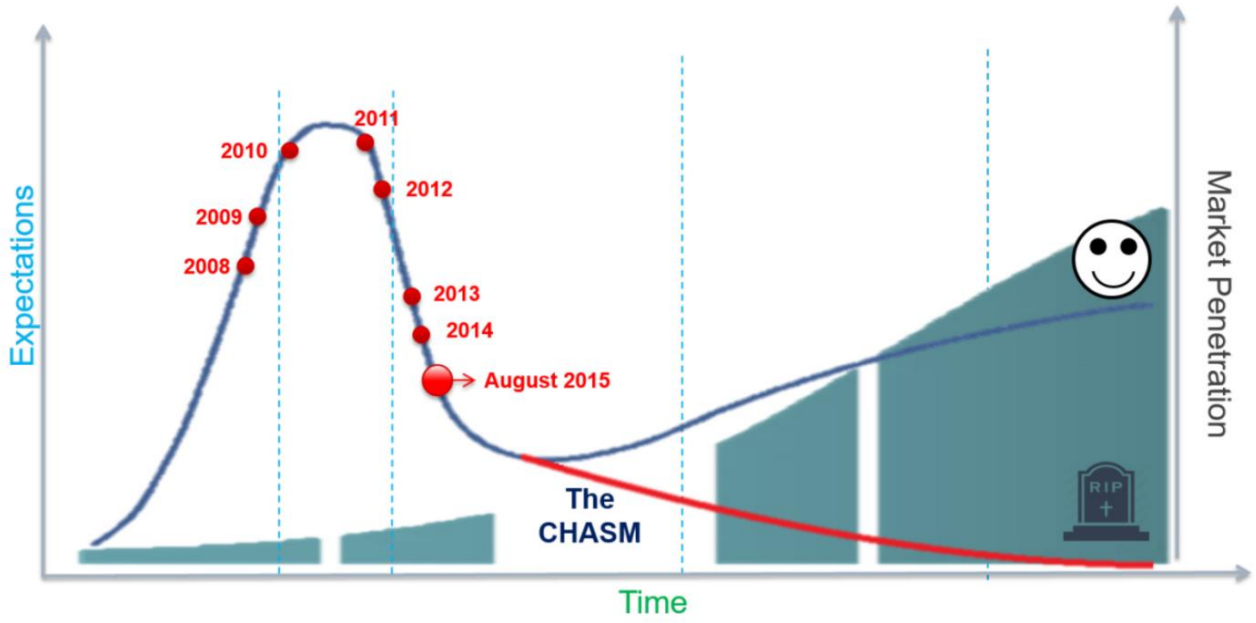


Fig. 1.12 *– The combination of the two models: Hype Cycle and Technology Adoption Life Cycle, as presented in Cutolo, 2015* [17]*, showing the two different paths an emerging technology could take crossing the chasm.*

Thus, if we go back to the question (is AR a bubble destined to burst?) the answer is: we still do not know, but it largely depends on the field of application. Indeed, AR social networking applications as well as location-based apps or games, especially for smartphones or tablets (hand-held displays, according to Bimber), are already used by a wide mass of early adopters and not anymore by a small amount of technology visionaries. Likewise, AR applications in multimedia and entertainment have

found a consistent application, for example in sports broadcasting (spatial displays, according to Bimber).

Yet, the first applications of AR were in the medical and in the military field [20]. Nevertheless, most of the applications were not properly established, and early adopters and researchers in these fields are still struggling to prove their efficacy and reliability (i.e. we are still far from crossing the *chasm*) [17].

2 CHAPTER 2: AUGMENTED REALITY IN MEDICINE AND SURGERY

In 2006, Charles B. Wilson (Department of Neurological Surgery, University of California, San Francisco, CA) published on the British Medical Journal a fine analysis about «Adoption of new surgical technology» [21]. The insightful subheading sounded like a warning: «Surgeons and patients seeking improved treatment often forget that a new technique is not necessarily a better one». Wilson claims that a new surgical technology offering the promise of improved patient care is certainly attractive. «Intrigued, and with an intuitive certainty, surgeons — cheered on by their patients — may adopt new technologies, despite little evidence of either their efficacy or their superiority over existing procedures». Wilson describes that factors responsible of the adoption and diffusion of a new technology fall into two categories: characteristics of the technology itself and contextual factors that promote it (Fig. 2.1).

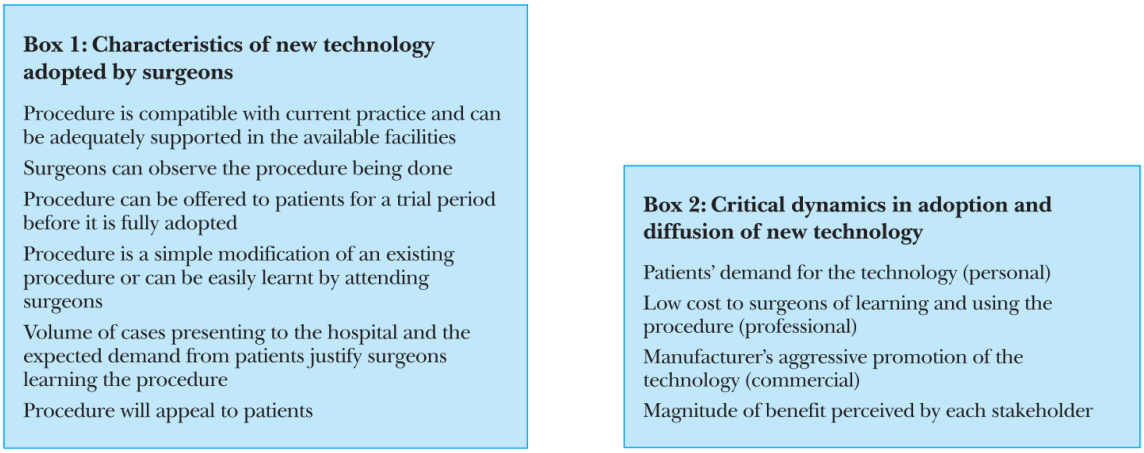


Fig. 2.1 *– Factors affecting adoption and diffusion of a new technology in surgery according to Wilson* [21].

Wilson also cites Roger's curve (in a slightly different version) [19], suggesting that in the medical/surgical field the introduction of a new technology encounters the same stages of any other field of application. This concept could be intuitive, but it is actually not foregone, because in the medical/surgical field the final user and the final beneficiary of a new technology are not the same person: the first is the surgeon, the second is the patient. Moreover, the medical/surgical field is usually supposed to deal with a high level of evidence before adopting new procedures and the decision about the supremacy of a new technology should not be entrusted to the market.

Augmented Reality does not elude these models, but does not come as a completely new thing. Indeed, AR can be considered an improvement in the field of Image-Guided Surgery, which has

received a broad amount of attention and evoked a significant mass of research in the last twenty years.

2.1 Augmented Reality as a Tool for Image-Guided Surgery

Surgeons and physicians have always simulated real anatomy on patients' surface as an aid to their diagnostic, therapeutic or surgical acts (Fig. 2.2).

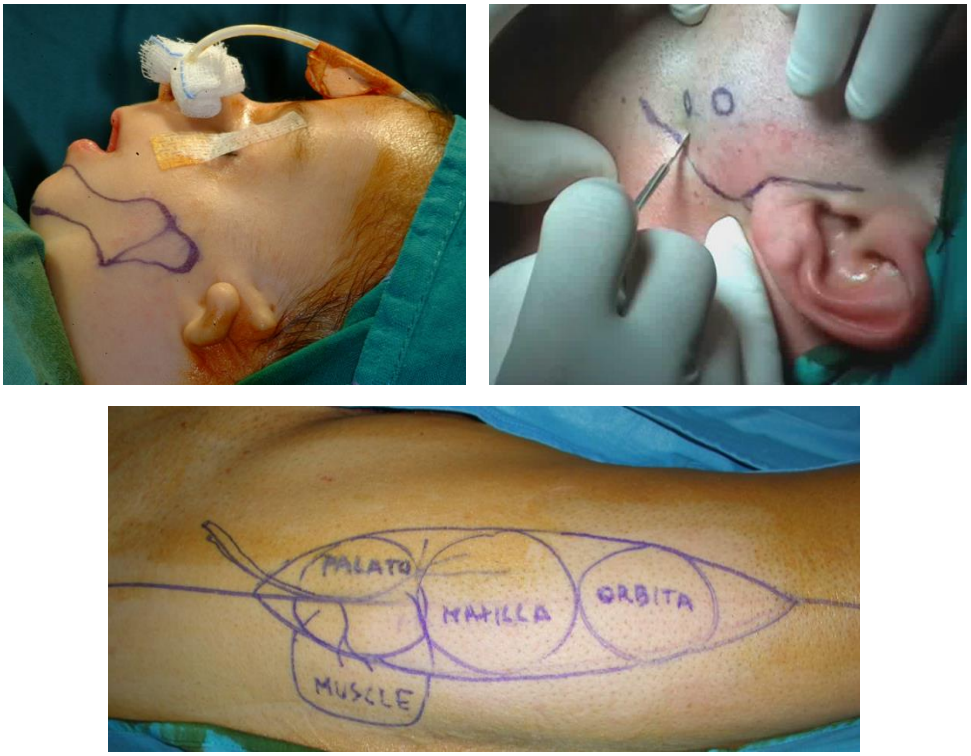


***Fig. 2.2** – Surgeons' historical way to depict internal anatomy on patients' surface as an aid to surgery, representing real (dismorphic) anatomy (upper-left), tumor location (upper-right) or surgical planning (bottom).*

Physicians have always imagined the ability to see into a living human system and to transfer the three-dimensional complexity of the human body into a comprehensive visual representation (Fig. 2.3).

Fig. 2.3 *– How surgeons imagine to see anatomy through superficial tissues of the patient.*

Nowadays, many new medical imaging modalities are available. This, together with the need to reduce the invasiveness of the surgical procedures, have encouraged the research for new 3D visualization modalities of patient-specific virtual reconstructions of the anatomy, both acting as surgical guidance or as tools for surgical planning or alternatively for diagnosis [22], [23].

The terms image-guided surgery (IGS) and computer-aided surgery (CAS) refers to this well-known concept that includes several technologies, each with a peculiar and significant use in surgery. Three-dimensional virtual surgery and navigation are the two main areas of interest in IGS. These technologies has undoubtedly increased accuracy of treatment planning and performing [24]–[31] and has undergone a remarkable diffusion, thanks to the production of many experimental and commercial software platforms and navigation systems [25], [32]–[35].

In this context, the idea of integrating in situ the surgeon's perceptive efficiency with the aid of new AR visualization modalities has become a dominant topic of academic and industrial research in the medical domain since the 90's. AR visualization is indeed considered capable to provide the surgeon with the ability to access the radiological images and surgical planning contextually to the real patient anatomy.

For surgeons, AR could represent a way to get back to the right perspective, where the surgeon can merge all the digital information into the surgical field, instead of looking at a monitor placed far from the action (Fig. 2.4).

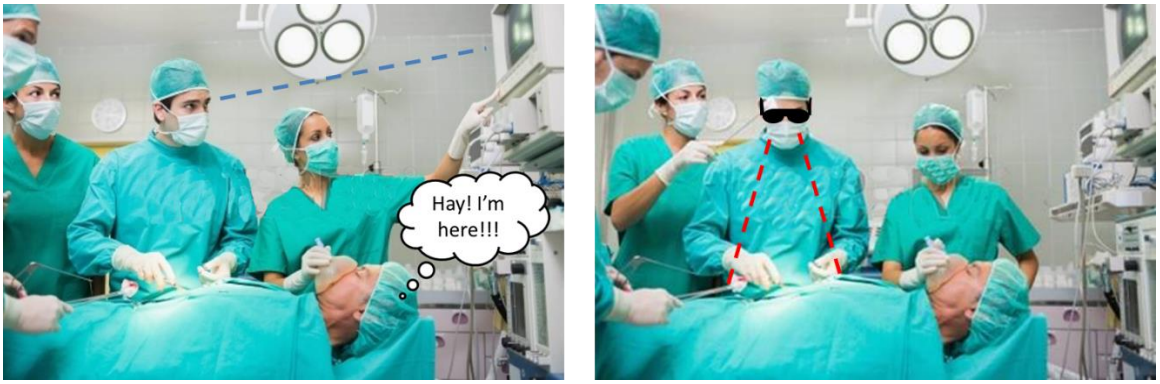


Fig. 2.4 *– The return to the right surgical perspective (pictures modified from the internet).*

However, there are still some reasons why such systems are not yet routinely used in the medical workflow, despite those ambitious targets [36]. Most of the systems were (and still are) developed as proof-of-concept prototypes, mostly conceived for research purposes more than for their immediate translation into real and reliable clinical conditions. Most experimental systems did not consider the operational restrictions imposed by the surgical context and subsequently did not satisfy the surgeons' practical needs. This is because most of these systems have lacked of a systematic evaluation within a clinical context. Moreover, the added benefit is for less experienced surgeons or surgical residents. For these reasons, it is still rare to see the integration of mixed reality systems and technologies integrated into real clinical environments and workflows [37]. In fact, the basic condition for a good reception of a new technology in the operating room is related to its capacity of being smoothly integrated into the workflow of the intervention, without affecting and disturbing the user during the rest of the procedure [38].

In the last two decades, the attempts to introduce AR in the OR, even if with severe experimental restrictions, have been numerous in several different surgical specialties. A review of the literature, according to specialty, is following. The presented literature is roughly limited to the last five years, in order to include only the most recent advances in the field.

However, AR display modalities that have been tested or applied in surgery do not differ from Bimber's paradigm: spatial display (external monitor), hand-held (tablets or projectors), retinal displays (head mounted displays), as described by Fig. 2.5 [39].

Fig. 2.5 *– Different display modalities for AR described by literature according to Nicolau* [39].

2.2 AUGMENTED REALITY IN GENERAL SURGERY

In the field of general surgery, Augmented Reality has been tested mostly in the context of laparoscopic surgery. The developing of accurate preoperative planning platforms for soft tissue organs has played an central role in this technical progression (Fig. 2.6) [40]. In this background, the hepatic-duodenal-pancreatic area is preeminent.

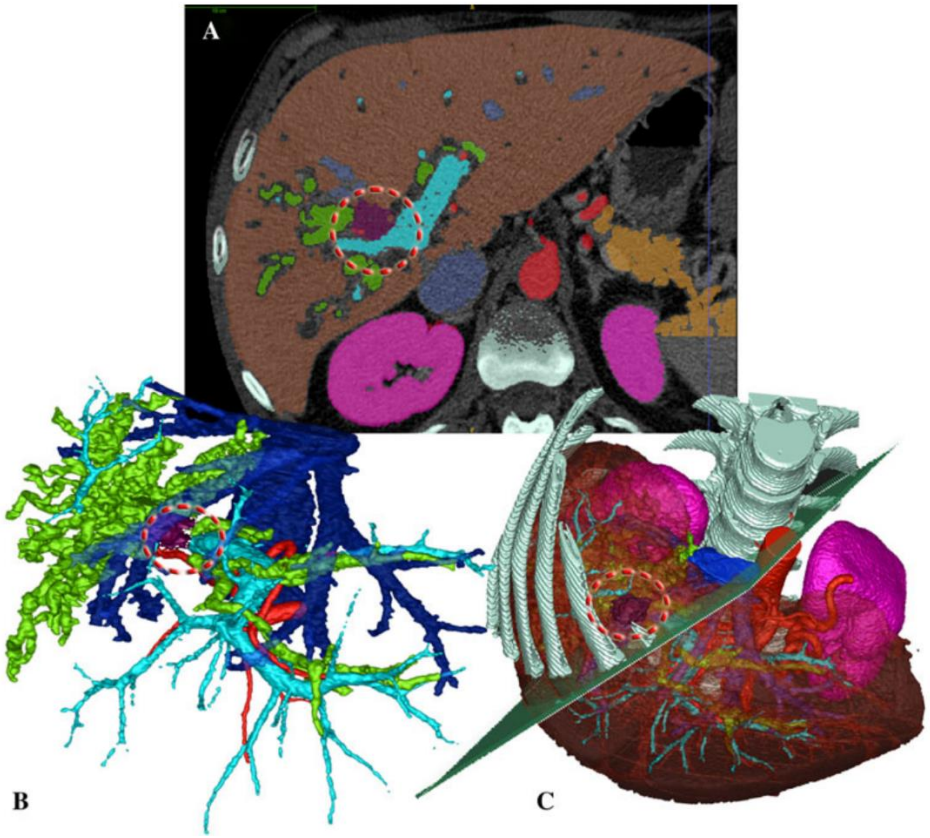


Fig. 2.6 *– Computed tomography (CT) slice with superimposed segmentation. From Ferrari et al.* [40].

Nicolau et al. (2011) presented a fine review about AR in laparoscopic surgery [39]. Nicolau's hypothesis to support AR introduction is founded on the limitations given by the laparoscopic minimally invasive technique. The first main difficulty is the loss of the direct sense of touch, replaced by force feedback through a minimally invasive instrument: surgeons can hardly distinguish hard tissues from soft ones and they cannot feel the blood pulse. The second main difficulty is the loss of depth perception because most endoscopic cameras are monocular. The third main difficulty is related to the limited field of view of the endoscopic camera (usually 70% vs. 160% for the human eye) which does not consent controlling all organs and instrument movements at the same time. Furthermore, the paper highlights that AR can also be used to show to novice surgeons the movement that should be done, by superimposing virtual instruments sketching a specific task on the endoscopic view.

After that work, many surveys demonstrated the efforts in developing more precise methods based on AR. It should be remarked that for abdominal surgery, the main limitation of AR is given by the variability and deformability of soft tissue anatomy. Literature on this topic is mainly addressed to overcome this issue and to demonstrate improvements in accuracy.

Nguyen et al. (2013) described a marker-less camera tracking method for registering 3D model on top of endoscopic image in Endoscopic Retrograde Cholangiopancreatography (ERCP) [41]. They also introduced an adaptive pose estimation which handles the lack of textures of endoscopic images of duodenum. The preliminary result showed a promising real-time tracking performance with fusion image. Even when there is some occlusion as some water was sprayed into duodenum or the catheter is inserted, the visual tracking method still showed a robust and accurate result (Fig. 2.7).

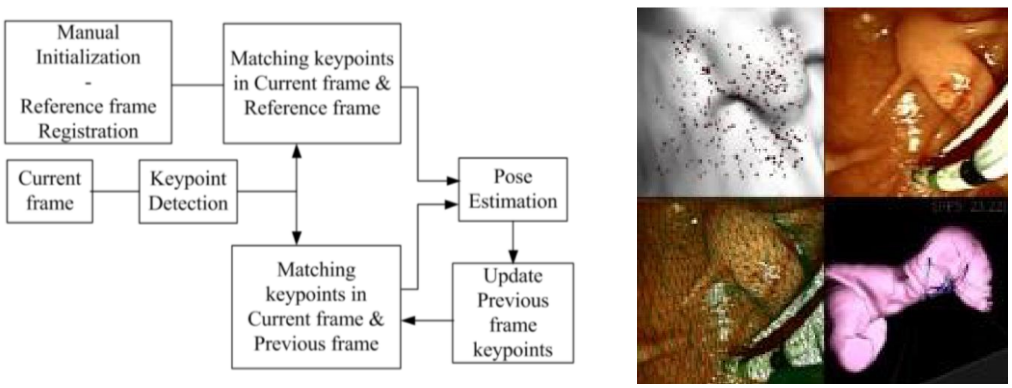


Fig. 2.7 *- Overview of camera tracking procedure (left) and an example of pose extimation (right)* [41].

Marzano et al. (2013) presented AR navigation as a tool to improve safety of the surgical dissection in a case of artery-first pancreatico-duodenectomy (PD) (Fig. 2.8) [42].

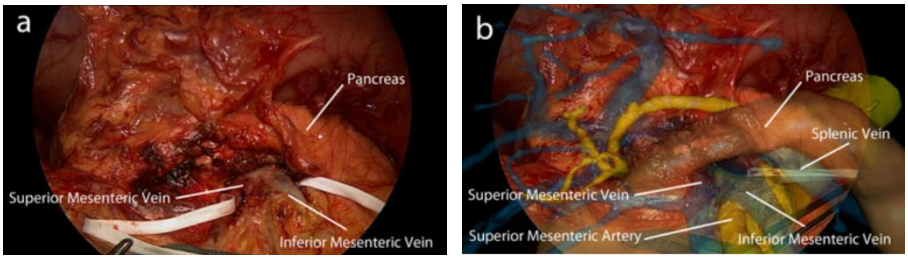


Fig. 2.8 *- Overview of camera tracking procedure (left) and an example of pose extimation (right)* [42].

Katić et al. (2013) presented a new system providing context-aware AR for laparoscopic surgery [43]. Context-aware systems provide visualizations of preeminent data just in time without handicapping the surgeon with a too much crowded images (Fig. 2.9). Indeed, abundant visualizations pose a risk as long as they distract surgeons from the surgical field, especially since visualizations are commonly overlaid on top of areas of interest. Context-aware systems have a certain understanding of the situation, enabling them to automatically derive the current information needs of the surgeon. The basic idea is to monitor the progress of the surgery with sensors, e.g. intraoperative data from medical devices in the operating room (OR), positional information from tracking systems or analysed endoscope images.

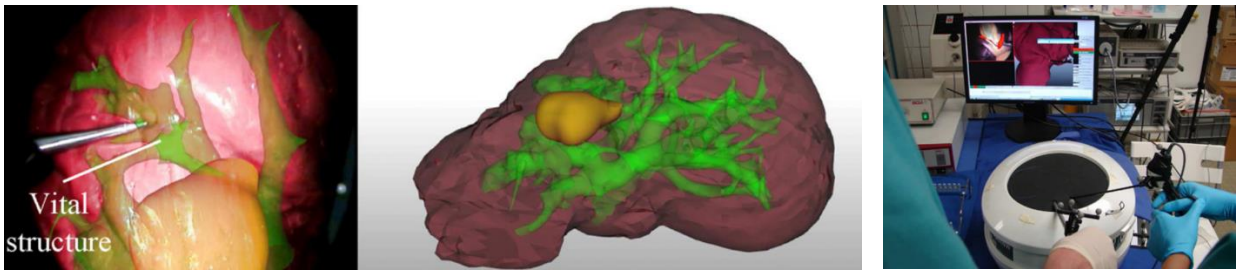


Fig. 2.9 *– Context-aware Augmented Reality according to Katić* [43].

Another interesting system proposed in this field is the one by Lopez-Mir et al. (2013) [44], who also depicted a concrete point of view about advantages and disadvantages of different visualization modalities (Figg. 2.10 - 2.11).

Device	Advantages	Disadvantages
Projector	Direct vision of 3D model onto the patient, free movement of the surgeon in the scene, and easy hand-eye training.	Low resolution, ambient light in the operating theater that reduces projection sharpness, heat and air ventilation that are dangerous in sterile environments, shadow effect, and setup just above the scene that entails an associated risk.
AR glasses	High grade of immersion and good mobility of the surgeon in the scene but more reduced in comparison to the projector.	Possibility of dizziness, decrease of reality perception and loss of depth, additional material in the surgeon's field of view, and feeling of stress until surgeon gets used to it.
Monitor	Harmless for patient and surgeon, free movement in the scene (the screen can be placed farther from the doctor), easy to sterilize, and high resolution.	Hand-eye training since the surgeon must look up to see the monitor.

Fig. 2.10 *– Advantages and disadvantages of different visualization devices* [44].

Fig. 2.11 *– Display visualisation according to Lopez-Mir* [44].

Onda et al. (2014) brought an AR in the OR, presenting a clinical survey (a series of 7 patients) where AR was used as an aid for identification of inferior pancreaticoduodenal artery during pancreatico-duodenectomy [45]. Results were promising (Fig. 2.12).

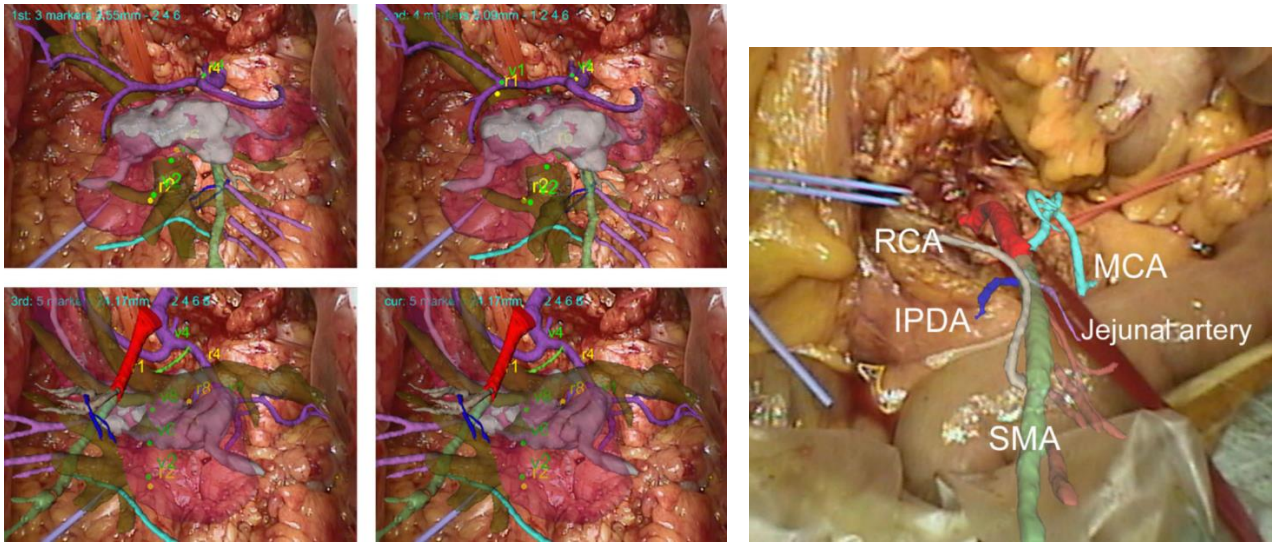


Fig. 2.12 *– AR for identification of inferior pancreaticoduodenal artery during pancreatico-duodenectomy according to Onda et al.: (left) overlaid images of the pancreas, arteries, veins, and tumor; (right) (AR)-based navigation image during pancreaticoduodenectomy* [45].

Pessaux et al. (2014) tested an AR-based system in conjunction with a robotic platform (DaVinci™, Intuitive Surgical, Inc., Sunnyvale, CA) in a small series of patients (Fig. 2.13) [46], [47].

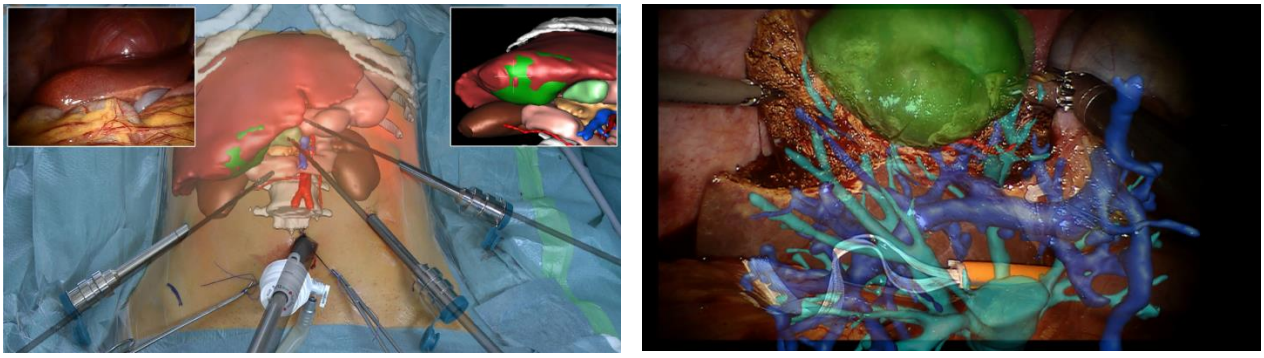


Fig. 2.13 *– AR in conjuction with a robotic tool, according to Pessaux* [47].

In 2015 and 2016, other interesting publications in this field have been produced by Okamoto et al. [48], [49], Edgcumbe et al. [50], Hallet et al. [51], Haouchine et al. [52], Koreeda et al. [53], Wild et al. [54]. Among these, it is interesting to mention the work by Haouchine et al. [52], which is addressed to presents a method for real-time augmented reality of internal liver structures during minimally invasive hepatic surgery, based on non-rigid registration of hepatic soft-tissues (Fig. 2.14). This is possible by capturing the 3D tissue deformation of an organ from intraabdominal images and updating the registration during surgical acts thanks to a specific biomechanical model and a derived computed algorithm (Fig. 2.15). The aim is to overcome the deformability of soft tissue and the subsequent difficulty to accurately register their position in real-time AR.

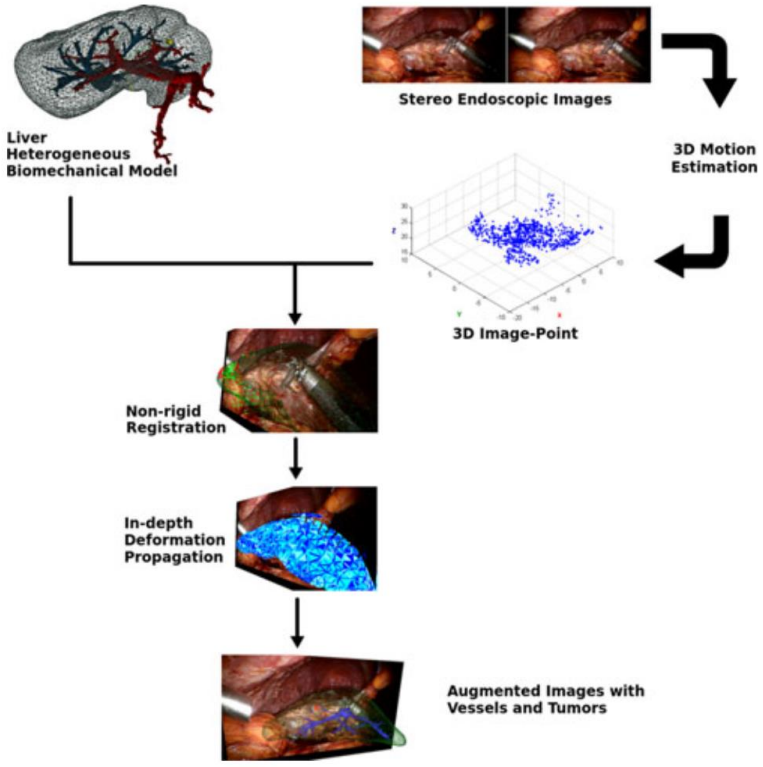


Fig. 2.14 *– The computational flow of the method: The biomechanical model guided by the 3D image-points recovered from the intra-abdominal image pair permits to propagate partial tissue deformations to vessels and tumors* [52].

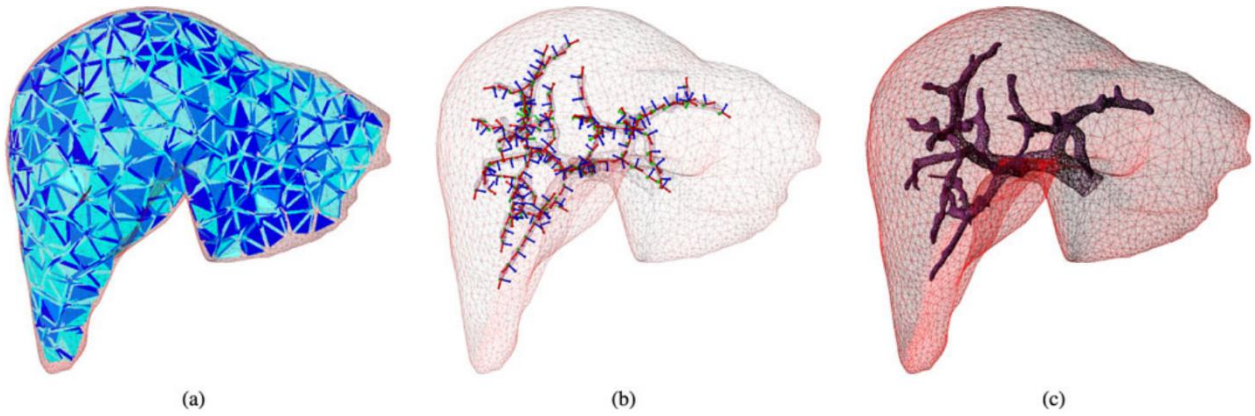


Fig. 2.15 *– 3D heterogeneous biomechanical model of the liver: (a) The volumetric mesh composed of tetrahedra. (b) The beams generated along the vessels described in Section 4. (c) The heterogeneous liver model including the vascular network shown in wire-frame* [52].

2.3 AUGMENTED REALITY IN ORTHOPAEDIC SURGERY

In Orthopaedic Surgery navigation technology is broadly used and studied. This is because orthopaedic surgery usually deals with rigid structures like bone, which is easily trackable. It is simple to understand that AR has effortlessly won the interest of the researchers in this field. Some recent examples are following.

In 2013, Abe et al. developed a novel AR guidance technique to visualize the needle insertion point on the patient's skin and the 3D trajectory path for percutaneous vertebroplasty, using a head mounted display (Fig. 2.16) [55]. Also Wu et al. (2014) worked on AR for spinal surgery [56].

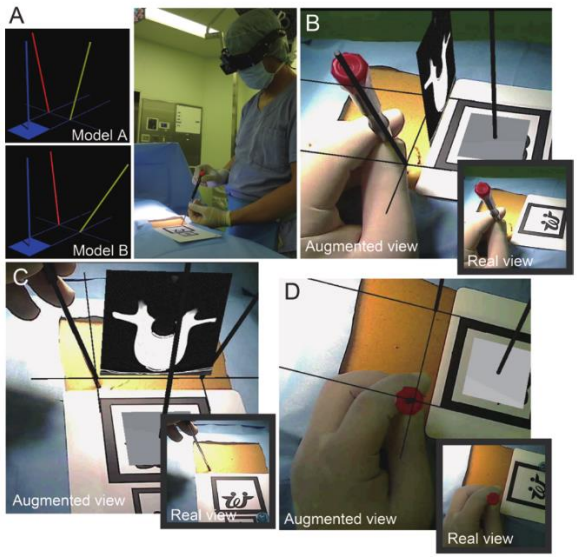


Fig. 2.16 *– Simulative models and AR-assisted needle insertion using the spine phantom according to Abe* [55]

It is clear that percutaneous screw positioning is a promising area for the application of AR. Wang et al. (2015) presented novel augmented reality-based navigation system for accurate insertion of percutaneous sacroiliac screws [57]. The study was a pilot one, but results were encouraging. Moreover, AR system presented included a head mounted display (Fig. 2.17). Other recent surveys about orthopaedic surgery are Londei et al. [58], Fischer et al. , Lee et al. [59].

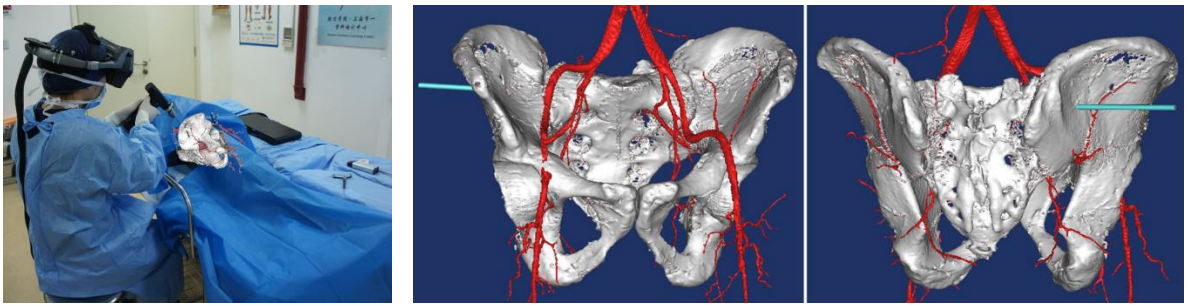


Fig. 2.17 *– The system developed by Wang et al.* [57].

2.4 AUGMENTED REALITY IN NEUROSURGERY

Neurosurgery is the first field where navigation was applied [24]. During the last decade, neuronavigation has become an essential neurosurgical tool for pursuing minimal invasiveness and maximal safety [60]. Unfortunately, ergonomics of such devices are still not optimal. Thus, AR found a fertile soil as a potential aid to overcome these limitations.

A very accurate review about AR in neurosurgery was published in 2016 by Meola et al. [61]. A total of 18 studies were included in their review. Six of the 18 studies reported neuro-oncological applications only, six reported neurovascular applications only, five reported both neuro-oncological and neurovascular applications, and one reported a neuro-oncological, neurovascular, non-neuro-oncological non-neurovascular application, the use of AR for external ventricular drainage placement. Authors suggest that current literature confirms that AR in neurosurgery is a reliable, versatile, and promising tool, although prospective randomized studies have not yet been published. However, this issue is certainly shared with many surveys on this topic among all medical specialties. According to this review, a practical ten-point multiparametric assessment for AR systems in neurosurgery is provided (Fig. 2.18).

Field of use	Options (examples)
1. Open neurosurgery	Macroscopic (allowed by all systems) Microscopic (only microscope and endoscope)
2. Endoscopy	Endoscope
3. Endovascular neurosurgery	X-ray fluoroscopy
AR system features	
4. Real data source	Microscope video camera (hand-held or head-held) Tablet-camera unaided view (projector; light field display and half mirror at 45°, LFD) Endoscope X-ray fluoroscopy
5. Tracking modality	Optical (broadly used for all the systems) Magnetic (used only by Doyle WK, 1996) None
6. Registration technique	Fiducial markers (applied to skin, teeth, skull) Skin surface registration Manual
7. Display type	Microscope oculars semitransparent head-up display (used only by Doyle WK, 1996 and questionable) External monitor (hand video camera, camera, microscope, endoscope, X-ray fluoroscopy) Tablet monitor LFD Projector
8. Perception location	Patient (microscope, tablet, projector, LFD) External monitor
AR scene parameters	
9. Virtual image source	CT MRI (and related techniques) Angiography
10. Visualization	Wireframe texture map surface mesh transparencies light field rendering standard navigation view

Fig. 2.18 *– A practical ten-point multiparametric assessment for AR systems in neurosurgery* [61].

A very prolific group about AR in neurosurgery is the one headed by Cabrilo, from the Neurosurgery Division of the Geneva University Medical Centre, (Switzerland), who successfully tested AR mainly for intracranial vascular anomalies in a large case series (Fig. 2.19) [62]–[64].

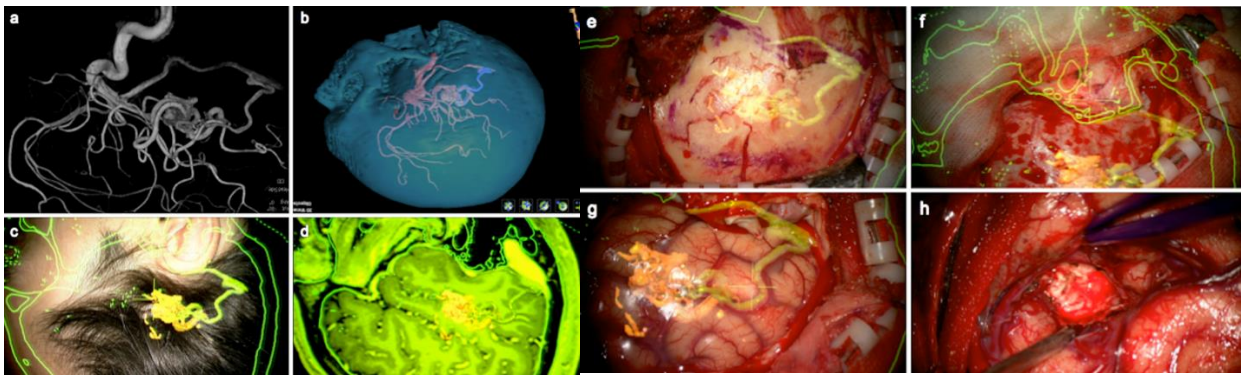


Fig. 2.19 *– An example of the display modality presented in the work by Cabrilo et al.* [63]

2.5 Augmented Reality in Other Surgical Sub-Specialties

Augmented reality has been successfully tested in many other subspecialty (some examples are displayed in Fig. 22.20) like urologic surgery [65]–[70], vascular surgery [71]–[74], paediatric surgery [75], [76], eye surgery [77], microsurgery [78], anaesthesiology [79] and, of course, head and neck and cranio-maxillofacial surgery which are the topic of the next paragraphs.

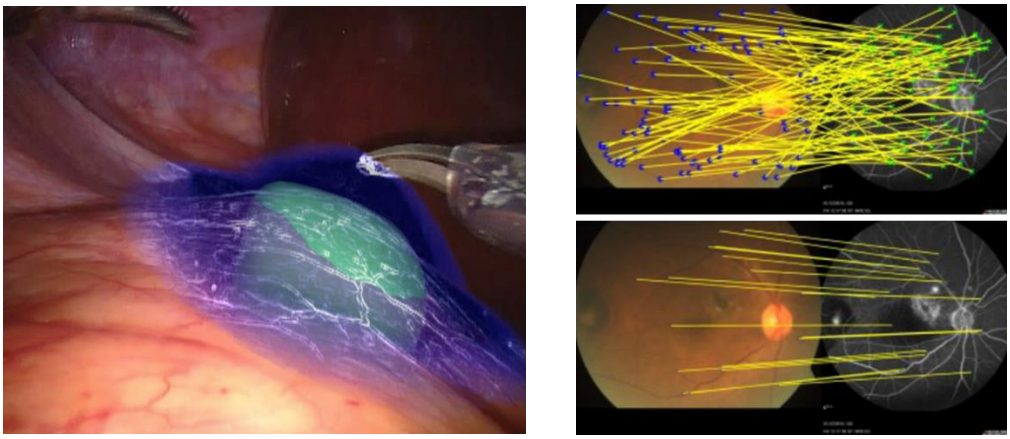


Fig. 2.20 *– Examples of AR application in other surgical subspecialties: (left) partial nephrectomy in urologic surgery* [65]*, (right) retinal treatments in eye surgery* [77].

2.6 Augmented Reality in ENT and Head & Neck Surgery

Head & Neck surgery is a broad field where new technologies have always had a preeminent role. AR is starting to provide some interesting future-oriented results, especially for robotic trans-oral surgery [80], [81], cervical surgery (like lymph node biopsy) [82], [83] and for endoscopic skull base surgery [84], [85]. This last topic is of preeminent interest, as long as endoscopic skull base anatomy is severely complex but grossly rigid, because every vascular, nervous, mucosal structure is firmly bonded to the surrounding bone. This features make skull base a perfect area where navigation in general and (even more) augmented reality could help in an effective way. The greatest issue is how to track the rigid but possibly bending endoscope. Great efforts have been put on this topic [84] with promising results (Fig. 20.21)

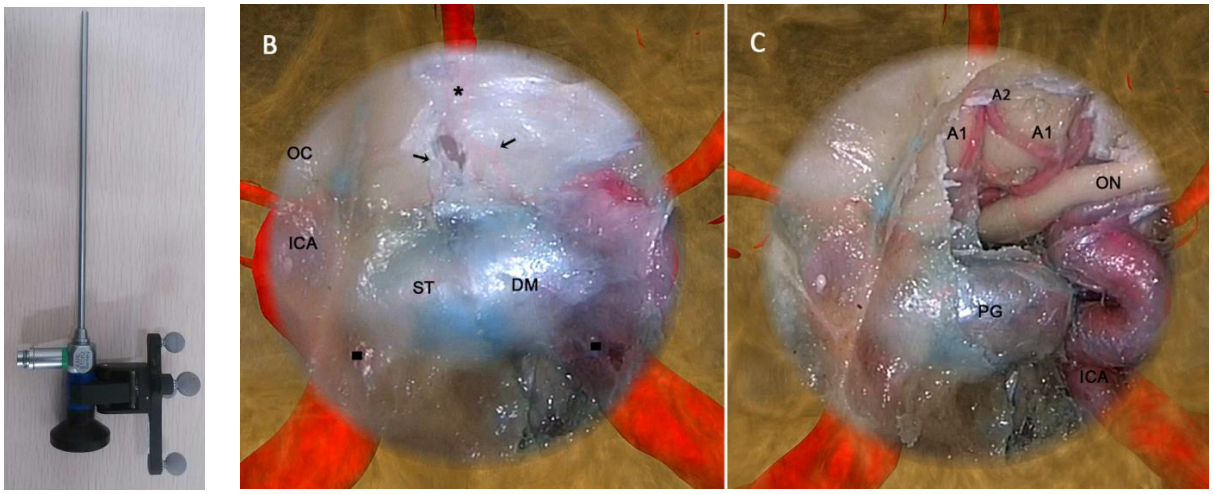


Fig. 2.21 *– Examples of an AR-based system providing augmented endoscopic view of the skull base anatomy* [82].

2.7 AUGMENTED REALITY IN CRANIO-MAXILLOFACIAL SURGERY AND DENTAL SURGERY

Finally, Cranio-maxillofacial (CMF) surgery is discussed, which is the area of research of this thesis. According to the geographical site, CMF surgery can share some or most topics and pathologies with Head & Neck surgery. Nevertheless, here CMF surgery is discoursed restricting his field to surgery of cranio-maxillofacial malformations and dysmorphosis (orthognathic surgery) (Fig. 2.22).

Fig. 2.22 *– Examples of facial dysmorphosis in the context of CMF surgery.*

Preoperative planning and navigation have been successfully applied to this topic. In particular, the group from the University of Bologna has published interesting results in the context of surgical simulation [25], [32] and navigation [33]–[35].

The software platforms that we helped to develop and tested have showed to be very useful in planning complex surgical bone movements and predicting the potential outcome (Fig. 2.23)

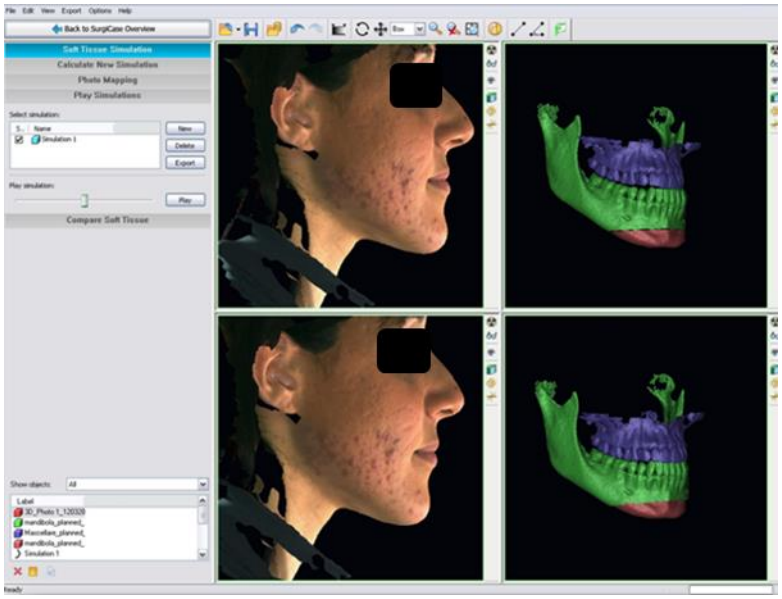


Fig. 2.23 *– Surgical planning with a dedicated software.*

Besides, the application of navigation to this surgery has been validated with significant results if compared to the standard procedure where computer-assisted surgery was not used (Fig. 2.24).

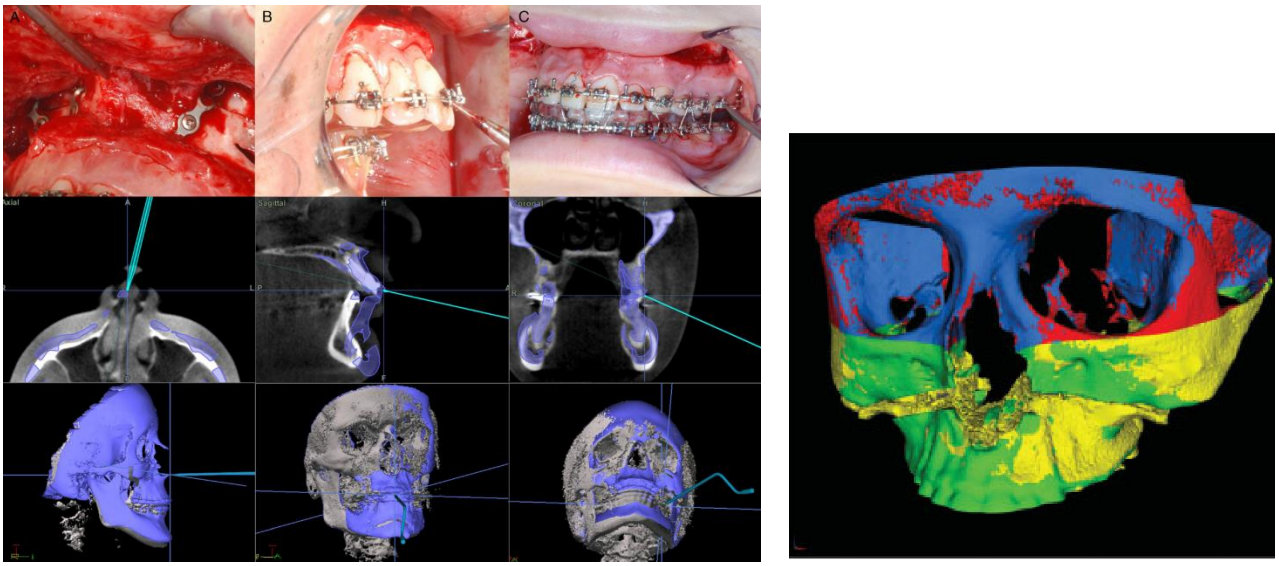


Fig. 2.24 *– A standard navigation system as an aid in orthognathic surgery.*

Thus, it is clear that, in the context of image-guided surgery, improvements based on AR may represent the next significant technological development in CMF surgery, because such approaches complement and integrate the current concepts of surgical navigation based on virtual reality.

In this field, the first attempt to introduce an AR-based navigation system is due to Mischkowski and Zinser [86], [87]. Both works, with an interval of six years, present a device based on AR for jaw bone repositioning in real clinical conditions. The device is a hand-held tablet in conjunction with a navigation system (Fig. 2.25).

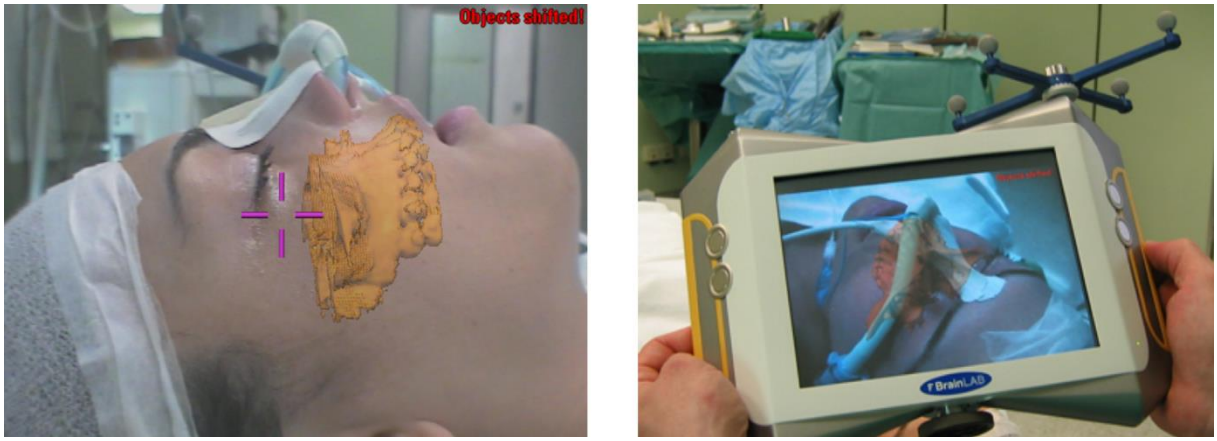


Fig. 2.25 *– AR-based navigation tool according to Mischkowski and Zinser* [86].

They present promising results in real clinical surgical conditions. Nevertheless, a great limitation of this system is represented by the need to hold the device while operating, which is not a problem for HMD tools.

A significant step forward, even if not tested in real clinical framework, has been made by Suenaga et al. [88], [89]. This group developed and evaluated, in vitro, an experimental AR system incorporating integral videography for imaging oral and maxillofacial regions. The three-dimensional augmented reality system (integral videography display, computed tomography, a position tracker and a computer) was used to generate a three-dimensional overlay that was projected on the surgical site via a half-silvered mirror (Fig. 22.26) [88]. This technology has the great benefit to leave surgeon's hands free even if the augmented content is not presented with a HMD. In a second survey [89], they improved the system and the algorithm, including a function for automatic registration based on teeth (Fig. 2.27).

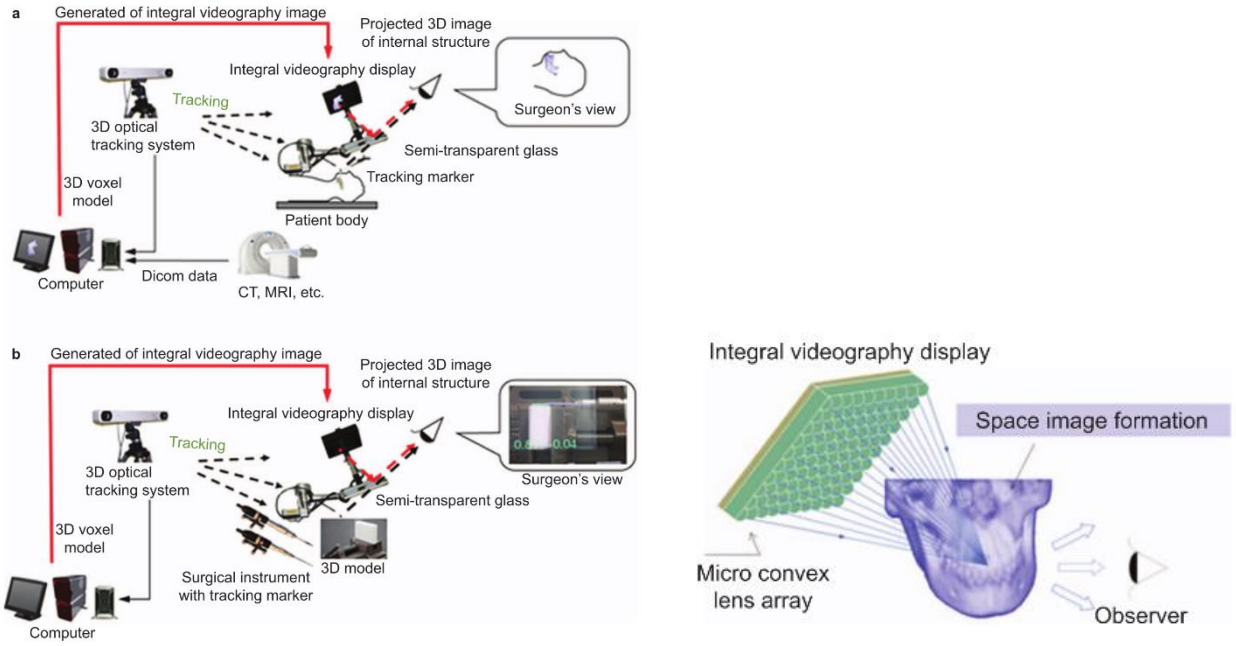


Fig. 2.26 *– The system built by Suenaga et al.; integral videography is a sort of holographic imaging* [88].

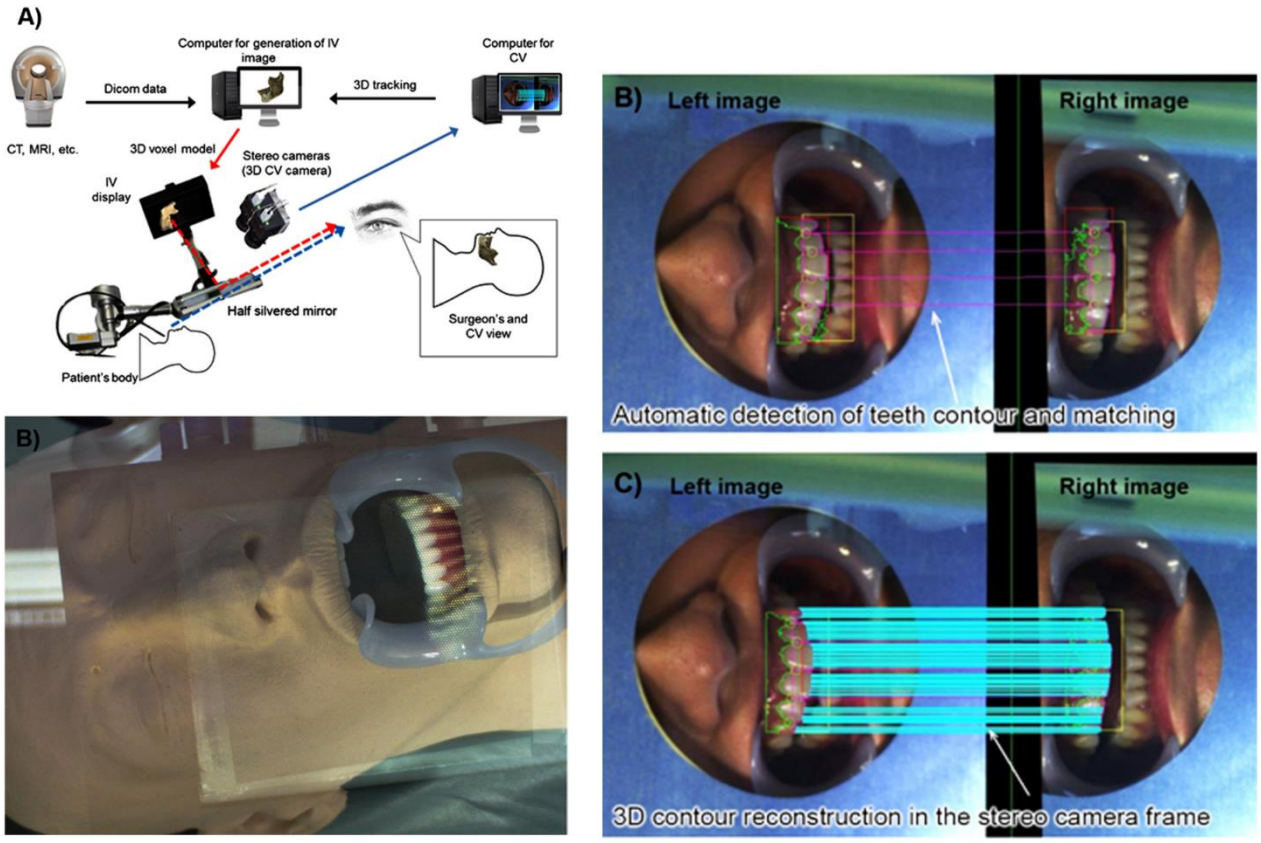


Fig. 2.27 *– (left) The improved system by Suenaga et al.; (right) the algorithm for automatic teeth detection* [89].

In more recent times, two interesting studies from Asia presented the use of different AR-systems for specific CMF procedures in real clinical surgical conditions.

Qu et al. presented their experience about the use of an augmented reality navigation system providing a surgical guide for the transfer of the osteotomy line in mandibular DO for patients with hemifacial macrosomia (Fig. 2.28) [90]. The survey is relevant and the presented tool is promising, even if it is not completely clear how the augmented content is displayed to the surgeon.

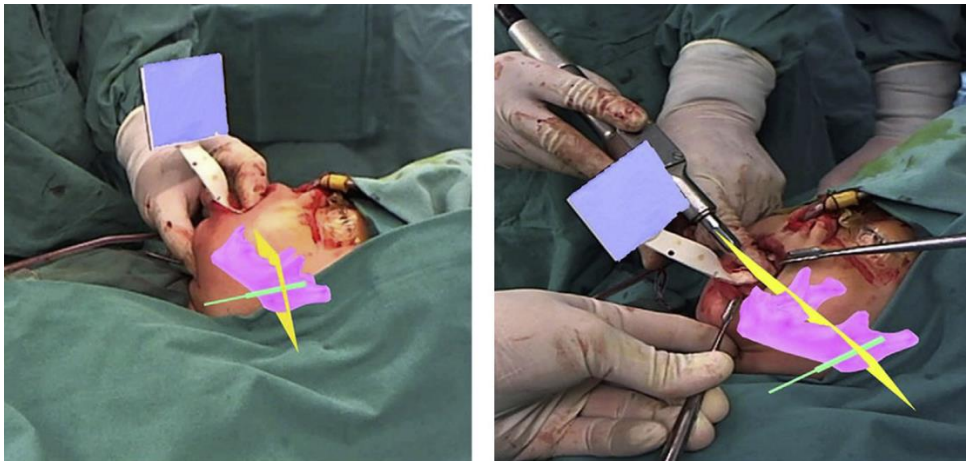


Fig. 2.28 *– A view of how the system by Qu et al. provides AR navigation* [90].

In the same year, Zhu et al. presented an analogue system for the treatment of orbital hypertelorism [91]. The AR-based device (Fig. 22.29) is similar to the one by Qu et al. and shows interesting results in a relatively large case series for this rare pathology (12 patients).

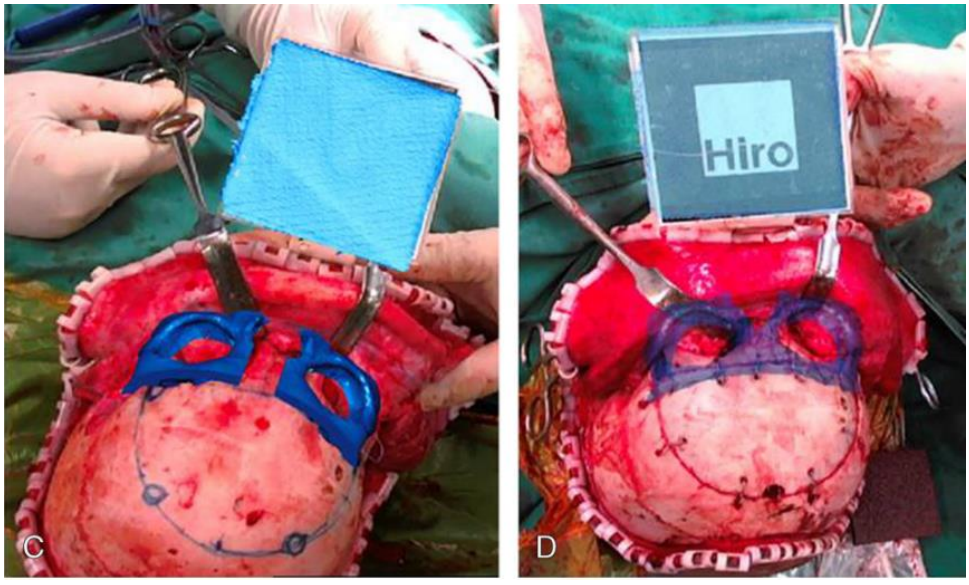


Fig. 2.29 *– A view of how the system by Zhu et al. delivers the AR information* [91].

Lastly, in 2016, Lin et al. presented another system for mandibular angle split osteotomy [92]. The system seems to be feasible, but it is tested just in vitro (Fig. 2.30).

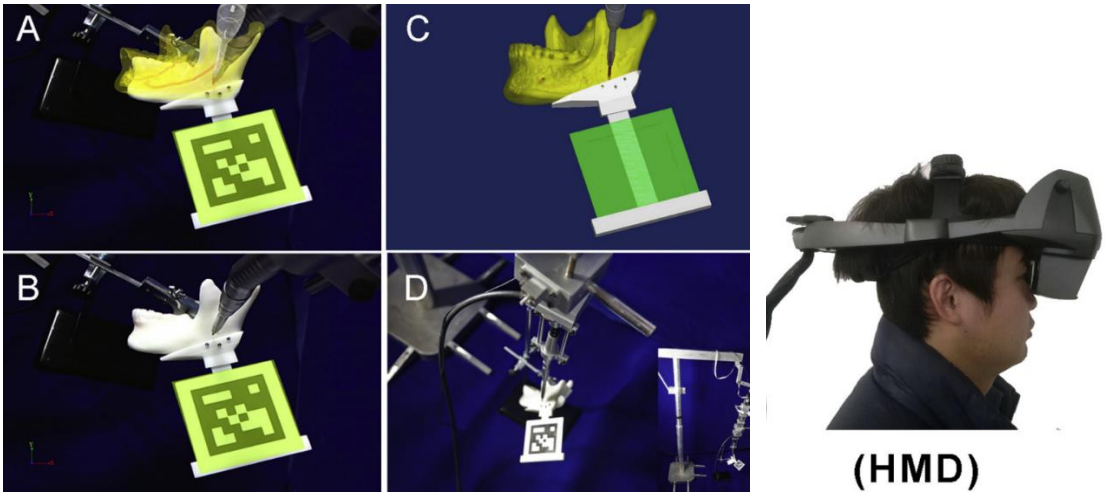


Fig. 2.30 *– The system by Lin et al. as an aid for mandibular angle split osteotomy* [92].

In the context of maxillofacial surgery, oral and dental surgery have not been forgotten by researchers aiming to AR-based tools. Katić et al. [93], Lin et al. [94] and Wang et al. [95] are some good examples on the topic. The main goal of these surveys is to develop a system as an aid for dental implant surgery. Yet, all studies available so far are restricted to experimental in vitro environment (Fig. 2.31).

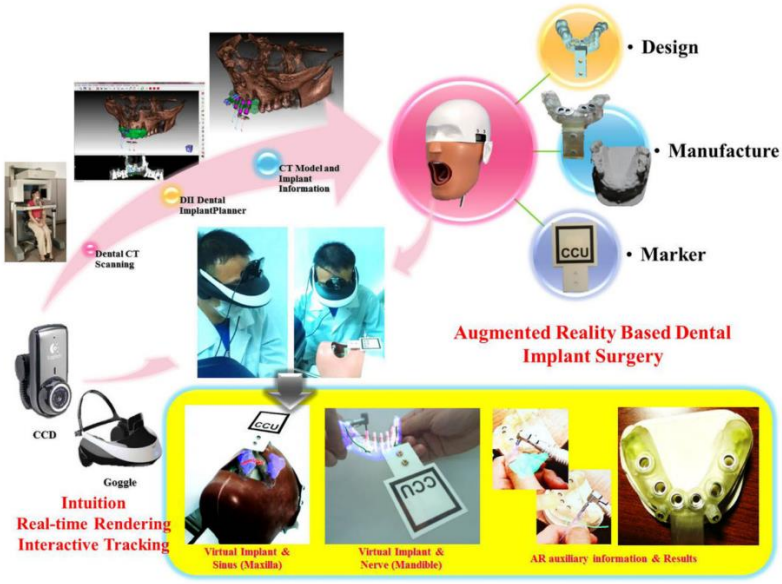


Fig. 2.31 *– An example of AR-based system for dental implant surgery according to Lin et al.* [94].

3 CHAPTER 3: DEVELOPING AN AUGMENTED REALITY DEVICE

As mentioned before in Chapter 1, this survey has been conducted on the basis of a close cooperation with the EndoCAS Center, headed by Prof. Mauro Ferrari, a preeminent clinical bioengineering laboratory of the University of Pisa, located in Pisa (Italy) in the context of the Cisanello University Hospital (Fig. 3.1). During the last decade, this research laboratory has been working on Augmented Reality applied to surgery, especially general surgery, spine surgery, neurosurgery and CMF surgery [17], [40], [61], [96]–[98]. This study results from the efforts of our two groups (EndoCAS Center and the Maxillofacial Surgery Department of the University of Bologna) towards an AR-based device for CMF surgery. The researchers who have mainly been working (and still are) on the presented study are Vincenzo Ferrari (head-engineer of the EndoCAS Center), Fabrizio Cutolo (as PhD student and main researcher involved in the technical aspects and improvements of the device) and Cinzia Freschi (as experienced engineer on this topic).

Fig. 3.1 *– Courtesy of EndoCAS Lab, University of Pisa (Italy).*

In particular, most of the technical content that follows, is borrowed from Fabrizio Cutolo's PhD thesis, as long as it pairs – like a "twin" study - with the clinical maxillofacial survey [17].

The aim of the technical job done by EndoCAS engineers was to develop a ***wearable AR-based device*** as an aid to open surgery. Despite the fact of being involved mainly in the clinical developing and testing of the device, we have had the opportunity to contribute also on the technical construction of the tool, as the following chapters will better clarify.

3.1 THE VIEW: HOW TO DISPLAY AR TO THE SURGEON

In 2012 Kersten-Oertel et al. [36] proposed a taxonomy focused on the mixed reality visualization systems in IGS and defined the required three major components. Subsequently, the authors demonstrated the efficacy of the proposed taxonomy by classifying 17 state-of-the-art research papers in the field of AR-based Image-Guided Surgery (IGS) systems [99]. The acronym for the taxonomy (DVV) derives from its three key components: *Data type*, *Visualization Processing* and View (Fig. 3.2). Having this taxonomy in mind, while classifying and assessing the value of a new AR system for IGS, the researcher must focus their attention on the specific surgical scenario in which the visualization system aims to be integrated. The surgical scenario affects each of the three DVV factors, i.e. the type of Data that should be displayed at a specific surgical step, the Visualization Processing technique implemented to provide the best pictorial representation of the augmented scene and how and where the output of the Visualization Processing should be presented to the end-user (namely, the View). Kersten-Oertel also tabulated all the subclasses, the main attributes, and the common instances of the DVV components of an AR-based IGS system (Fig. 3.3).

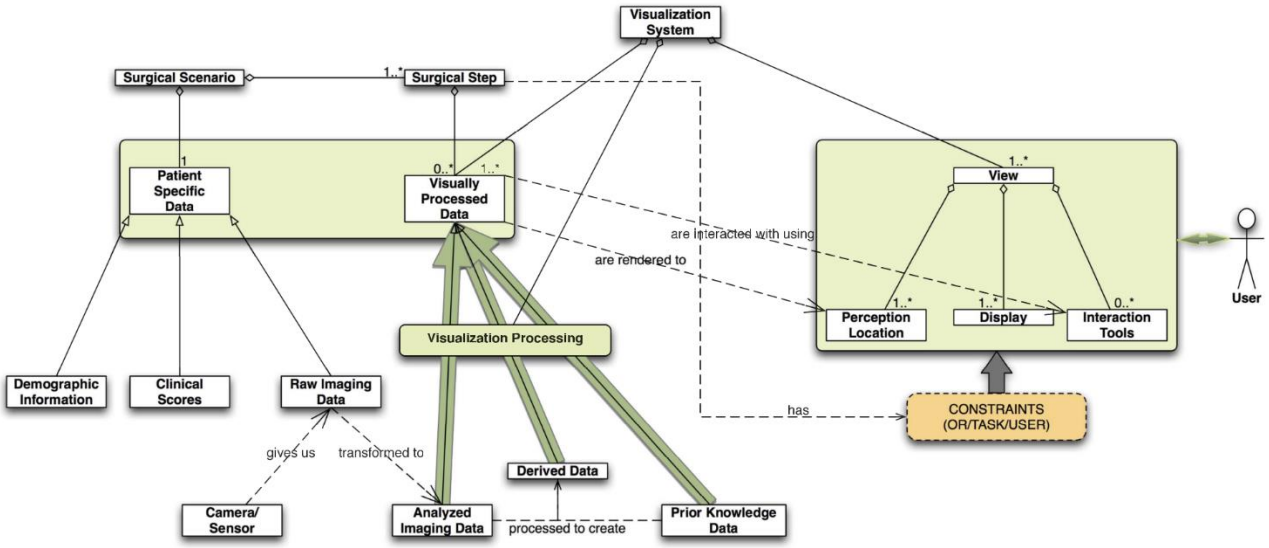


Fig. 3.2 *– Kersten-Oertel's taxonomy in AR visualization: the scheme outlinesnthe relations among the DVV factors (i.e. Data, Visualization Pro- cessing and View) and their classes and subclasses* [36].

FACTOR/CLASS	MAIN ATTRIBUTES	COMMON INSTANCES
Data	Dimensionality	0D, 1D, 2D, 2.5D, 3D, 4D
Raw Imaging Data	Acquisition sensor	CT, MRI, fMRI, EEG, angiography, mammography, X-rays, ultrasound, video endoscope, microscope
	Acquisition time	Pre-op, intra-op
Visually Processed Data	Semantic	Strategic, operational, anatomical
	Data object	Real, virtual
Analyzed Imaging Data	Data primitive	Point(s), line, contour, plane, surface, wireframe/mesh, volume
Derived Data	Underlying process	Labels, uncertainty, measurements (e.g. tumour volume, distance between regions of interest, etc.).
Prior Knowledge Data	Underlying model	Labels, atlases, system uncertainty, surgery roadmaps, tool models
Visualization Processing	Algorithm	Adding depth cues, using transparency, illustrative rendering, photorealistic rendering, non-photorealistic rendering, importance-driven volume rendering, saliency methods (outlining, highlighting, etc.), filters (colour, edge enhancement, etc.), animation, lighting models
View	Mixed reality	Augmented reality, virtual reality
Perception Location	Location	Patient, display device, surgical tool, real environment
Display	Device	*2D Technology:* Monitor/screen, touchscreen, wall, laser, microscope *3D Binocular Stereoscopic:* HMD, polarized glasses, silvered mirror, shutter glasses, colour glasses, polarized glasses, augmented microscope *3D Auto-stereoscopic:* multi-view lenticular displays, integral videography, holography
Interaction Tools	Hardware interaction tools	Mouse, keyboard, data glove, haptic device, tangible object (e.g. surgical tool)
	Virtual interaction tools	Change object properties (colour, transparency, edge thickness, visibility (on/off), etc.), change view or pose (rotate, zoom, focus window, etc.), transfer/windowing functions, volume cutting, voxel peeling, clipping planes

Fig. 3.3 *– Kersten-Oertel's common instances of the DVV visualization taxonomy* [36].

In Chapter 2, we grossly experienced how AR has been implemented in surgery in the last years. Historically, the first AR-based systems in surgical navigation have been implemented starting from commonly used devices [20]. Augmented operating microscopes were proposed in neurosurgery [63], [100]–[102]. In these systems, AR was displayed by inserting the virtual content directly into the optical trail of the real image, hence by inserting a beam splitter into the microscope optics. Other solutions featured the use of spatial monitors and video-based tracking modalities and were used in neurosurgery [103], maxillofacial surgery [86], [87], [104], laparoscopic general surgery [46], [48], [49], [105]. Another category of AR systems is based on the use of half-transparent screens in conjunction with display technologies providing monoscopic, stereoscopic or autostereoscopic parallax. Blackwell et al [106] introduced a prototype of AR window, providing a stereoscopic vision of the virtual content by means of a pair of shutter glasses. Stetten et al. [107] have proposed a simple and interesting optical see-through hand-held half-silvered mirror that overlays ultrasound scans aligned with the scanned area.

The major limitations of the optical see-through paradigm implemented in standard AR windows are due to the intrinsic mismatch between the 3D representation of the real world and the nature of the virtual content, rendered as a 2D image. To cope with these issues, alternative and promising

approaches, based on the integral imaging technology, were proposed [108]. Integral imaging displays use a set of 2D elemental images from different perspectives to generate a full parallax 3D visualization. Therefore, with integral imaging-based displays, a proper 3D overlay between virtual content and real scene can be obtained. The integral imaging paradigm thus is able to provide the user with an self-centred viewpoint and a full parallax augmented view in a limited viewing zone (imposed by the integral imaging display).

In laparoscopy, and generally in endoscopic surgery, the part of the setting where the attention of the surgeon is focused during the surgical task is a monitor. Thus, in such procedures, the surgeon operates watching endoscopic video images reproduced on the spatial display unit. Therefore, the virtual information is usually merged with the real-time video frames grabbed by the endoscope and presented on a monitor. These systems were also tested in robotic surgery [46], [47], [109], [110].

Otherwise, wearable AR systems particularly offer the most ergonomic solution in those medical tasks manually performed under user's direct vision (open surgery, introduction of biopsy needle, palpation, etc.). Wearable AR systems based on HMDs essentially provide the user with an self-centred viewpoint and they do not limit their freedom of movement around the patient. Standard HMDs provide both binocular parallax and motion parallax and smoothly augment the user's perception of the surgical scene throughout the specific surgical procedure. In these HMDs, the see-through capability is accomplished either through the above-mentioned video see-through paradigm or through the optical see-through paradigm.

3.2 Optical see-through vs Video see-through Head Mounted Displays

Generally, Head-Mounted Displays for Augmented Reality fall into two classes according to the *see-through paradigm* they implement: *video see- through* or *optical see-through*. In optical see-through systems, the user's direct view of the real world is augmented with the projection of virtual information on a bean combiner and then into the user's line of sight [111] (Figg. 3.4 – 3.5). Differently, in video see-through systems the virtual content is merged with camera images captured by two external cameras rigidly fixed on the visor (Figg. 3.6 – 3.7).

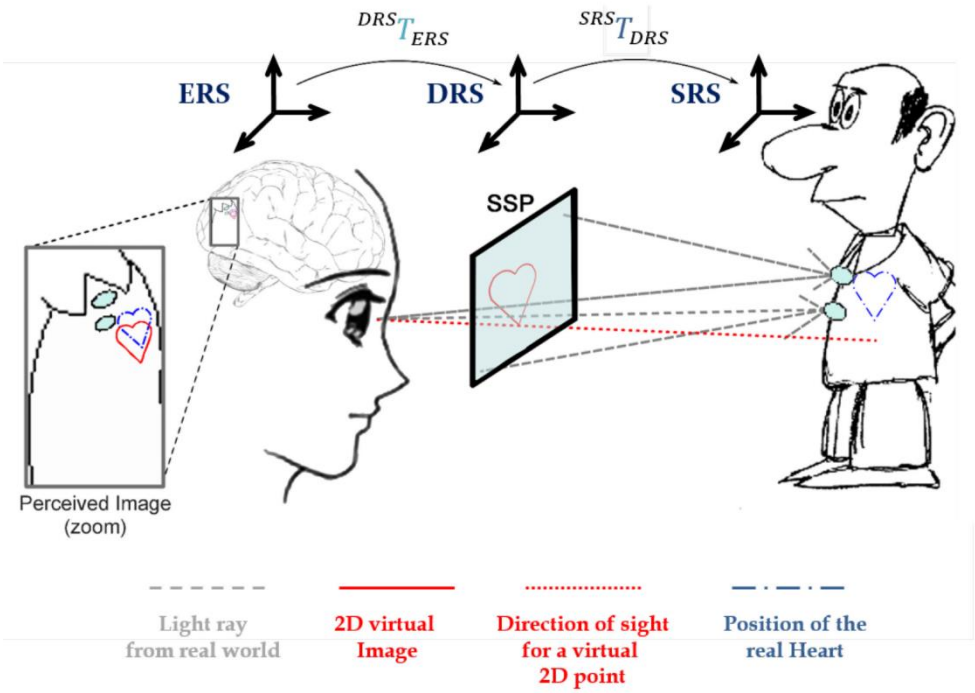


Fig. 3.4 *– Optical see-through paradigm: a virtual 2D image (heart) is rendered on the semi-transparent surface of projection (SSP). The user perceives the light rays coming from the real world merged with the 2D virtual image. Accurate geometric registration is difficult to achieve. Frequent calibrations are required to estimate eye's pose in relation to the display (${}^{DRS}T_{ERS}$). Moreover, the eye's projective model should be known (courtesy of EndoCAS).*

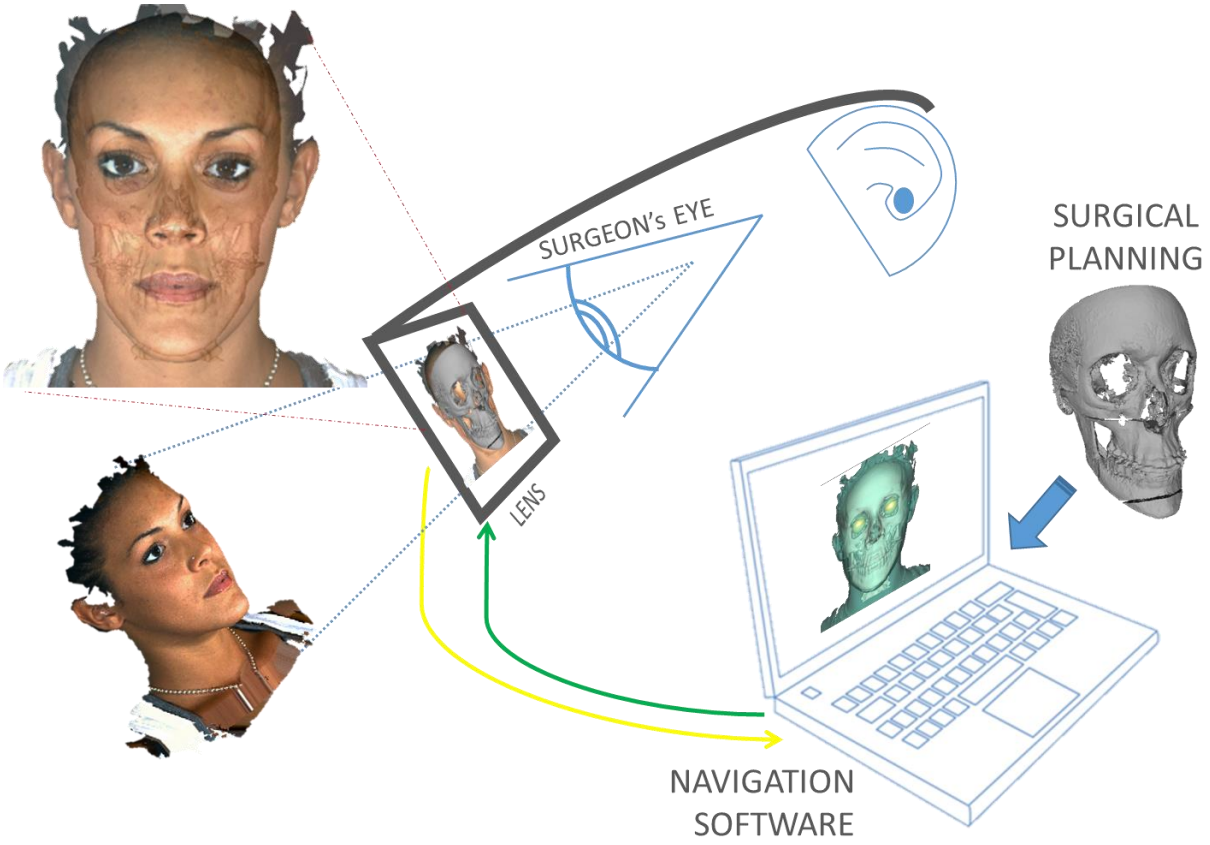


Fig. 3.5 *– The design of an AR-based Optical see-through system for CMF surgery.*

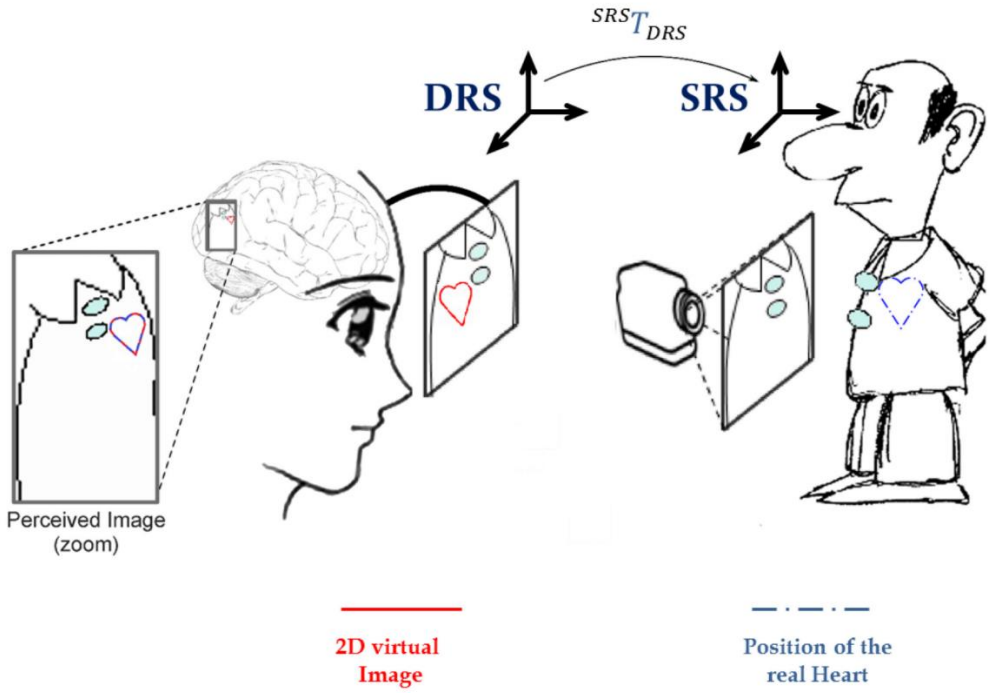


Fig. 3.6 *– Video see-through paradigm: the virtual 2D image (heart) is mixed on an image frame of the real world grabbed by the external camera. Accurate geometric registration solely relies on accurately determining* ${}^{SRS}T_{DRS}$ *(tracking) (courtesy of EndoCAS).*

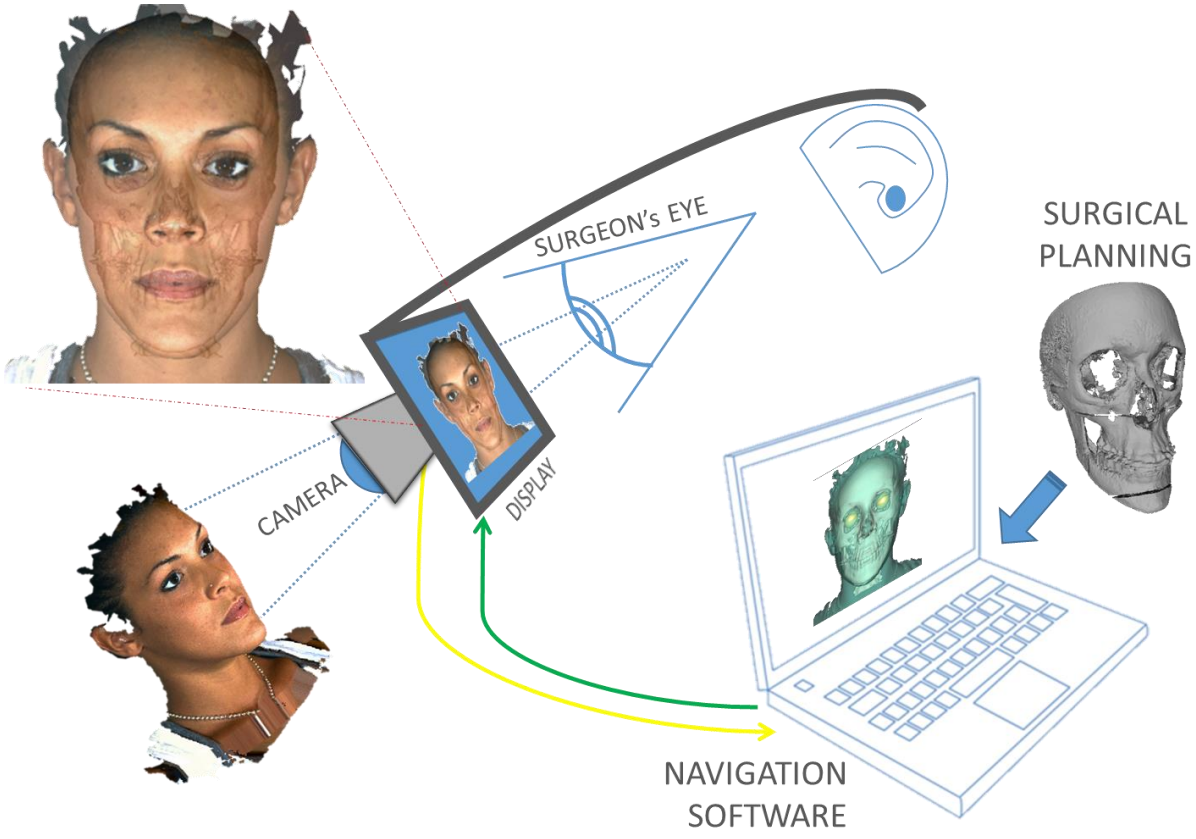


Fig. 3.7 *– The design of an AR-based Video see-through system for CMF surgery.*

Thus, optical-see through systems could seem to deliver a better experience of augmented vision, as long as they provide the user with a natural view of the real world with full resolution. On the other hand, the camera-mediated view typical of the video see-through paradigm could be considered drastically affecting the quality of the visual perception and experience of the real world. The fundamental optical see-through paradigm of HMDs is being implemented in many consumer-oriented devices available on the market or forthcoming (Google Glass, Microsoft HoloLens, Epson Moverio, Lumus optical, Meta 2). A straightforward implementation of the optical see-through paradigm includes the employment of a half-silvered mirror or beam combiner to merge real view and virtual content. The user's own view is augmented by rendering the virtual content on a two-dimensional (2D) micro display and by sending it to the beam combiner. Lenses can be placed between the beam combiner and the display to focus the virtual 2D image so that it appears at a comfortable viewing distance on a semi-transparent surface of projection (SSP) [112]. As an alternative, the use of high-precision waveguide technologies allows the removal of the bulky optical engine placed in front of the eyes [113]. The optical see-through paradigm is particularly suitable for augmenting the reality by means of simple virtual elements (models, icons or text labels) but shortcomings remain both from a technical and a perceptual standpoint, especially in case of virtual contents of greater complexity. Yet, the degree of adoption and diffusion of optical see-through systems has slowed down over the years due to technological and perceptual limitations. For example, the presence of a small augmentable field of view, the perceptual conflicts between the 3D real world and the 2D virtual image on the SSP, and the need for frequent recalibrations of the HMD for yielding accurate spatial registration.

Indeed, the optical see-through HMDs force the user to accommodate their eye for focusing all the virtual objects on the SSP placed at a fixed distance. On the other hand, the focus distance of each physical object in the 3D world depends on its relative distance from the observer and may dynamically vary over time. Even if an accurate geometric registration of virtual objects to the real scene is attained on the 2D SSP plane, the user may not be able to view both the virtual and real content in focus at the same time. This aspect is particularly relevant in applications devoted to surgical navigation, because it reduces the user's capacity to interact with the real surgical field while maintaining the virtual aid in focus.

The second major shortcoming of the standard optical see-through HMDs is related with the geometric registration required to obtain a geometrically consistent augmentation of the reality,

which is an essential prerequisite for being considered as reliable surgical guidance system [114]–[117].

Differently, the video mixing technology that underpins the video see-through paradigm, once integrated with monocular or binocular HMDs, can offer high geometric coherence between virtual and real content. In these systems, a user-specific calibration routine is not necessary, and this is the major advantage of the video versus the optical see-through approach. In video see-through systems, real scene and virtual information can be synchronized, whereas in optical see-through devices there is an intrinsic lag between the immediate perception of the real scene and the inclusion of the virtual elements. Further, from a perceptual viewpoint, in video see-through systems the visual experience of both the real and virtual content is unambiguously controllable by computer graphics, with everything on focus at the same apparent distance from the user.

Additionally, video see-through systems are much more suited than optical see-through systems, to rendering occlusions between real and virtual elements or to implementing complex Visualization Processing modalities that are able to perceptually compensate for the loss of the unobstructed real-world view. This aspect is of particular importance in IGS applications, wherever our goal is of trying to mimic the perceptive efficiency of the human visual system to allow a smoother interaction with the augmented visual information [118].

Therefore, at the current technological level, the use of video see-through systems is more straightforward, at least for those IGS applications in which we can tolerate slight delays between the capture of the real scene by the cameras and the presentation of the captured frames on the displays of the visor [17].

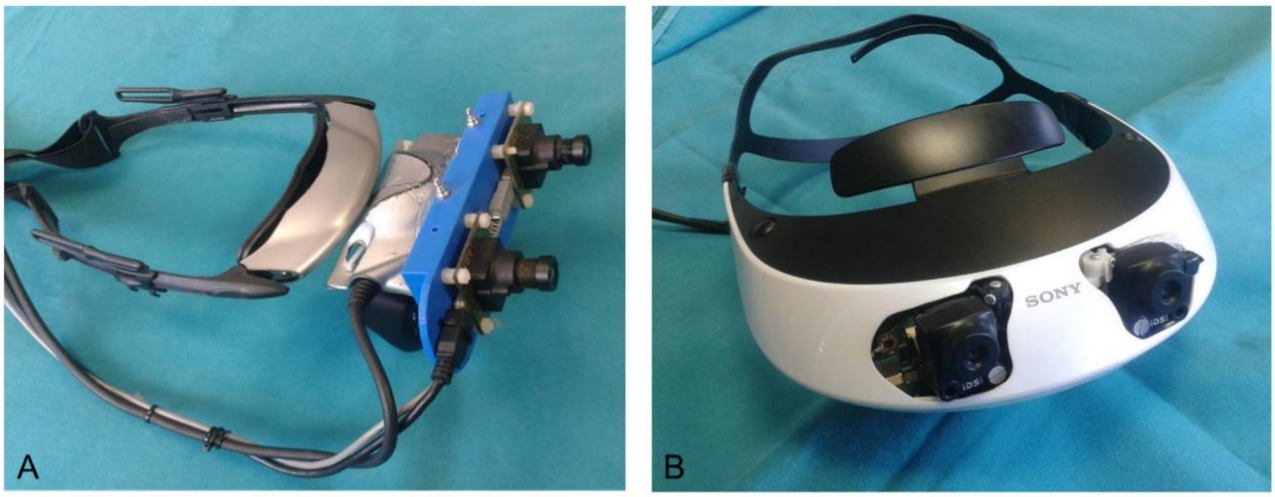
3.3 Developed Stereoscopic Video See-Through HMDs

The two custom-made stereoscopic video see-through HMDs built by EndoCAS (Pisa, Italy) and employed in this work comprise the following two major components: a commercial 3D visor and a pair of external USB cameras.

1. The first model of AR visor (Fig. 3.8A) was based on a commercial 3D visor by eMagin (eMagin Z800) provided with dual OLED panels and featuring a diagonal field of view (FoV) of 40°. The two external cameras were two USB cameras by IDS (uEye UI-1646LE) equipped with a 1.3 Megapixel

Aptina CMOS sensor (pixel pitch of 3.6 µm) that achieve a frame rate of 25 fps in freerun mode at full resolution (i.e. 1280x1024). A plastic frame (ABS) was built through rapid prototyping and fixed to the 3D visor as support for the two external USB cameras. In such a way, the two cameras could be aligned with the user's eyes as to provide a quasi-orthoscopic view of the augmented scene mediated by the visor (in a video see-through fashion).

2. The second model of AR visor (Fig. 3.8B) was based on a commercial 3D visor by Sony (Sony HMZ-T2) provided with dual 720p OLED panels and a horizontal FoV of 45°. The two external cameras were 2 USB cameras by IDS (uEye XS) equipped with a 5 Megapixel Aptina CMOS sensor (pixel size of 1.4 µm) that achieve a frame rate of 15 fps at 1280x720 resolution. Also here, the two external cameras were mounted on the visor with an anthropometric interaxial distance (≈7cm) and aligned with the user's eyes as to provide a quasi-orthoscopic view of the scene.



	eMagin Z800 & IDS uEye UI-1646LE	Sony Hmz-T2 & IDS uEye XS
Optimal Resolution	640x512	1280x720 / 640x480
Refresh Rate (fps)	25	15 / 30
Field of View (FoV)	40 deg diagonal	45 deg horizontal

Fig. 3.8 *– A) First embodiment of the video see-through HMDs from a 3D visor by eMagin and B) second embodiment assembled on a 3D visor by Sony.*

An ad-hoc application able to run with both the systems was developed. The software framework is able to load and display every AR scene. The system performs Video See-Through AR with commercial equipment ranging in a wide and constantly up-to-date list of devices.

The application is easily and highly manageable through configuration file. This allows the user to adapt the software framework for several different applicative scenarios and purposes, in terms of tracking methods, Visualization Processing modalities, and View/Display technology. Software architecture is easily extensible.

3.4 VIDEO MARKER-BASED REGISTRATION FOR STEREOSCOPIC AR

In AR-based IGS systems the goal of tracking is to determining the position and orientation (pose) of surgical tools and of patient's anatomical structures in relation to the reference display, so to obtain an optimally registered augmented scene [38]. In AR HMDs, the real-time accurate alignment between the real scene and the virtual content is usually performed by means of optical or electromagnetic external trackers [86]. Nevertheless, surgical navigation systems based on external infrared trackers have the main disadvantages of introducing unwanted line-of-sight restrictions into the OR and of adding technical difficulty to the surgical workflow. Other tracking modalities are based on more complex surface-based tracking algorithms [119]. As an alternative to optical tracking, electromagnetic tracking systems are particularly appropriate for tracking hidden structures [98] but their accuracy and reliability are heavily affected by the presence of ferromagnetic and/or conductive materials [120]. Finally, standard video-based tracking methods, featuring the use of large template-based markers [90], [91], [121] whose pose can be determined through machine vision routines, are hardly suited for use in a surgical setting because they limit surgeon's line-of-sight, given their planar structure, and/or they may obstruct the visibility of the operating field.

In 2009, Ferrari et al. already presented an early version of the HMD for AR applications. The distinctive feature of that AR HMD was that the pair of external cameras served both to capture the real scene and to act as stereo tracker [96]. Indeed, the lack of an external tracker is a key characteristic for allowing the smooth and profitable integration of AR systems into the surgical workflow.

Moreover, small spherical markers placed onto the surgical framework do not heavily affect the line-of-sight and they can be conveniently placed on the patient's anatomy with a reduced spatial impact on the surgical workflow. Still with the objective of increasing system usability, the minimum set of markers ensuring a finite number of solutions to the camera pose estimation problem was chosen, i.e. three.

One of the tasks conducted in EndoCAS Center by Fabrizio Cutolo was to develop an algorithm to correctly register patient's anatomy on the basis of small spherical markers and external cameras stereoscopic video-grabbing serving as tracker [17].

Figure 3.9 shows schematically how the stereoscopic video see-through mechanism is implemented for visualisation and registration, while Figure 3.10 shows the algorithm sequence.

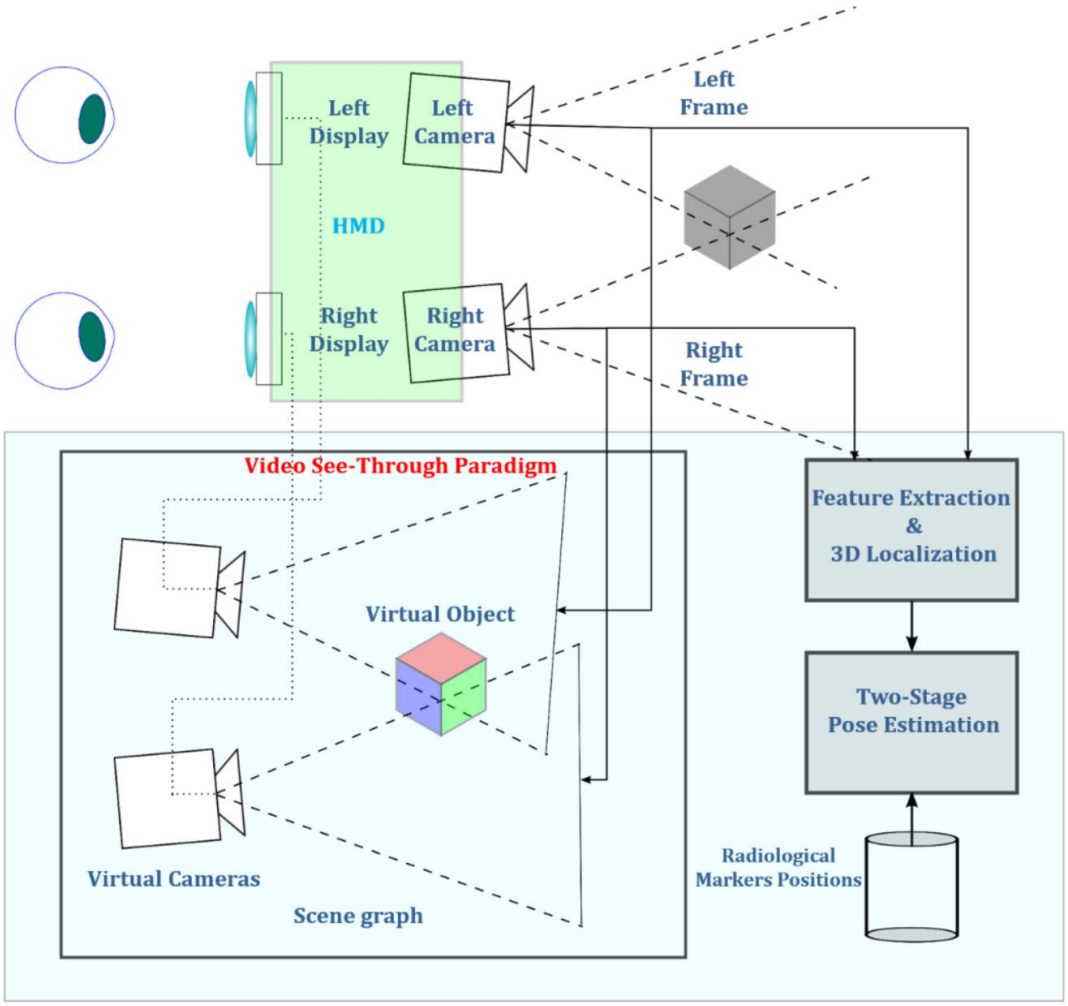


Fig. 3.9 *– Scheme representing the stereoscopic video see-through mechanism (courtesy of EndoCAS).*

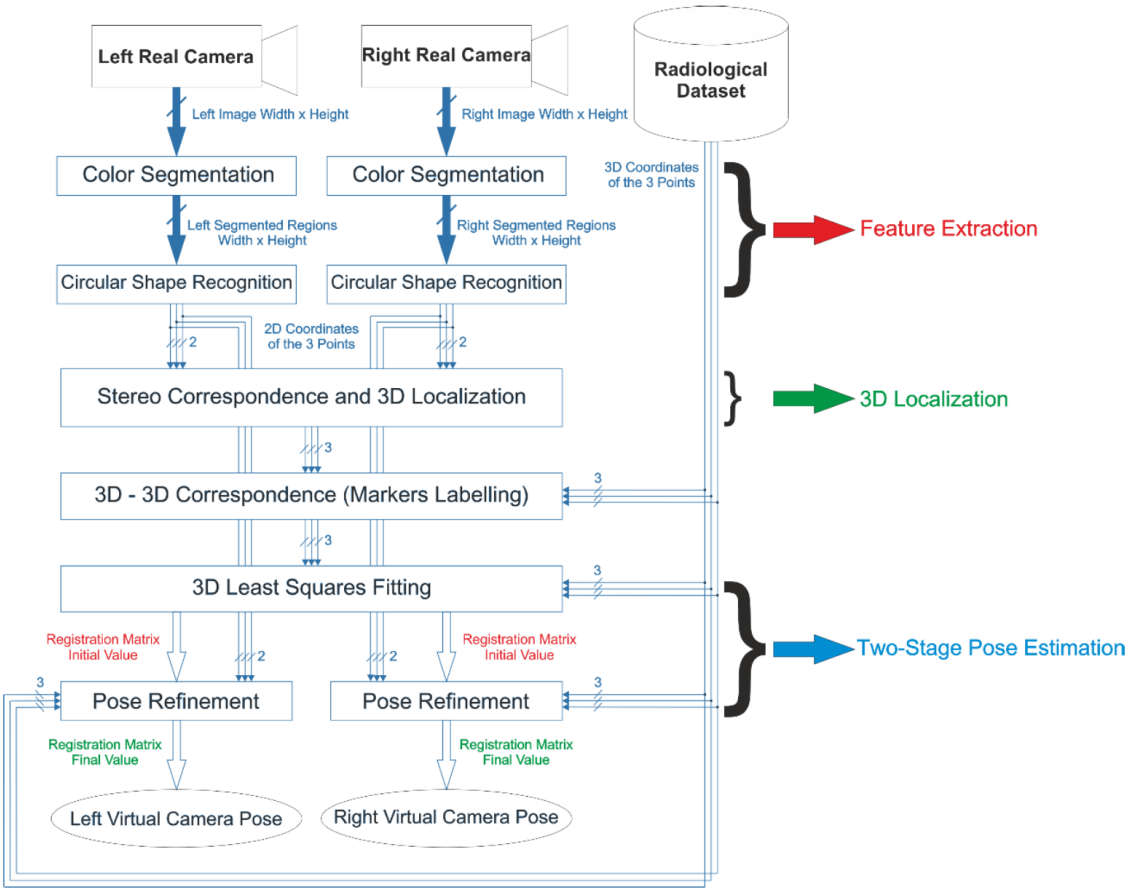


Fig. 3.10 *– Localisation and Registration Algorithm (courtesy of EndoCAS).*

This AR mechanism was implemented in software libraries built in C++ on the top of the multipurpose EndoCAS Navigator Platform modules [122]. The management of the virtual 3D scene was carried out through the open-source software framework OpenSG 1.8 (www.opensg.org), while regarding the machine vision routines, needed for implementing the video-based tracking method, Halcon 7.1 software library by MVTec® was adopted. The whole application was implemented to be compatible with several 3D displays (working either with side-by-side or alternate frames) and with all the cameras whose DirectShow drivers by Microsoft® are available.

In a video see-through system, to achieve an accurate and robust fusion between reality and virtuality we must render the virtual scene so that the following three conditions are satisfied:

1. The virtual cameras projection models ≈ to the real ones
2. The relative pose between the two virtual cameras of the stereo rig ≈ to the real one
3. The pose of the virtual anatomies/surgical tools ≈ to the real ones

To first condition implies that the virtual cameras viewing frustums must be modelled on the real ones in terms of image size, focus length and centre of projection (intrinsic calibration). At the same time, the second condition entails that the relative pose between the two virtual cameras of the stereo rig must be set equal to the pose between the two real cameras (extrinsic calibration).

Lastly, the pose of the virtual elements in the virtual scene must be set equal to the real poses between real anatomies/tools and the physical camera. This latest condition is satisfied by applying the video marker-based registration method.

The poses of the two cameras relative to the anatomy and vice-versa are determined by tracking passive coloured markers constrained to the surgical scene in predefined positions. The proposed video-based tracking solution relies on the stereo localization of three monochromatic red markers and it is robust to inconsistent lighting conditions. 3D coordinates of the markers in the left camera reference system (CRS) are retrieved by applying stereo 3D Localization routines on pairs of conjugate projections of the markers' centroids onto the image planes of the two cameras. Image coordinates of the centroids of the markers are determined by a feature extraction task, performed through Colour Segmentation and Circular Shape Recognition.

Schematically, the algorithm could be summarized as follows, as a Machine Vision methods applied on the grabbed images (Fig. 3.11 – 3.12):

A. *Colour Segmentation* → HS thresholding (Select the red-coloured regions)

B. *Shape-based Recognition* → Select the Circles

C. *Markers labelling* → Sequence of stereoscopic triangulations and geometrically-based comparisons

D. *Point Based Registration* → Least Squares Fitting through SVD

E. *Single Camera pose refinement* → Minimization back-projection error

On the basis of the testing conducted by Fabrizio Cutolo, this algorithm was found to be robust and accurate [17]. Thus, the clinical implementation of the presented method was considered possible and potentially promising. The following sections will clarify the efforts of the research team in developing an ergonomic and feasible device for CMF surgery (with potential benefits for any surgical application).

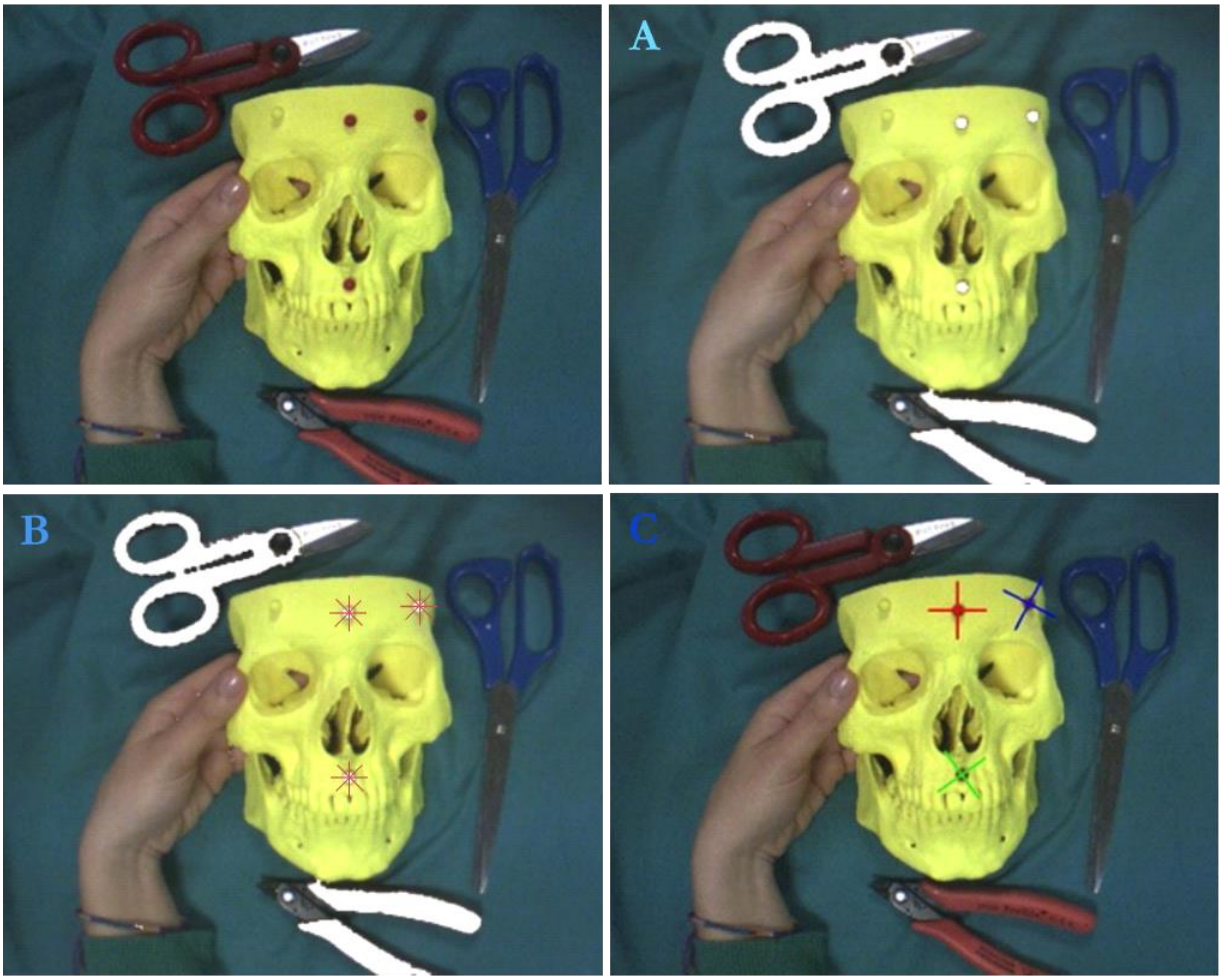


Fig. 3.11 *– A, Colour Segmentation; B, Circular Shape Recognition; C, Markers Labelling (courtesy of EndoCAS).*

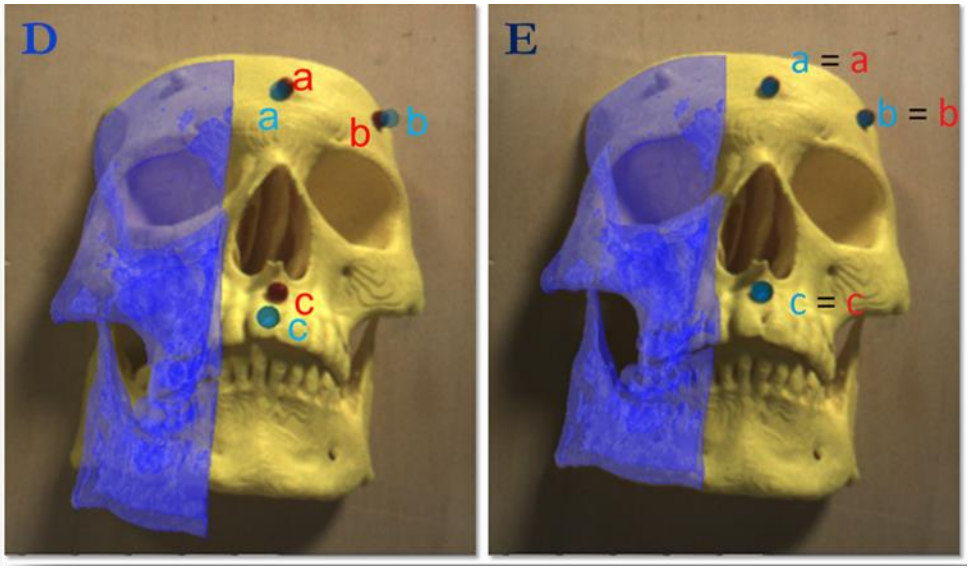


Fig. 3.12 *–3D-3D Correspondence Definition (D → E) and Point-Based Registration (courtesy of EndoCAS).*

4 CHAPTER 4: HUMAN-PNP - ERGONOMIC AR INTERACTION PARADIGM FOR MANUAL PLACEMENT OF RIGID BODIES

4.1 SPECIFIC CONTRIBUTION OF THIS SECTION

In the understanding of how humans perceive depth, the mechanisms behind the mutual interaction between physiological and psychological cues are still disputed. This question is of great importance in designing 3D Visors and it is even more important if the aim is to mimic the perceptive efficiency of the human visual system within augmented reality based surgical navigation systems. On this issue, unreliable modalities of AR Visualization can bring cognitive overload and perceptual conflicts causing misinterpretation and hindering clinical decision-making [123]. In this Chapter, a way to facilitate the profitable introduction of AR-based navigation systems into the surgical workflow is presented. The solution is based on the definition of an ergonomic interaction paradigm designed for aiding the manual placement of rigid bodies in space. The key idea behind this AR Visualization modality is that the minimization of the reprojection residuals, on the image plane, between a set of corresponding real and virtual feature points can direct the accurate placement, by successive manual adjustments, of a non-tracked rigid body relative to a tracked reference one. This AR interaction paradigm drew its inspiration from the general problem of estimating camera pose from a set of n-correspondences, known in computer vision community as the *perspective-n-point problem*. This work was presented at the 10th International Workshop, AE-CAI (Augmented Environments for Computer-Assisted Interventions) 2015, held in conjunction with MICCAI 2015 [124].

4.2 VISUALIZATION PROCESSING MODALITIES IN IGS

In the context of AR-based surgical navigation systems, several Visualization Processing techniques have been adopted to allow a more immersive viewing experience for the surgeon and a more precise definition of the spatial relationships between real scene and Visually Processed Data along the three dimensions. The human visual system exploits several physiological and psychological cues to deal with the ill-posed inverse problem of understanding a 3D scene from one retinal image. However, even when interacting with the real world, monocular and binocular cues are not always

sufficient to infer, with extreme accuracy, the spatial relationships between virtual and real objects. Therefore, a full comprehension of the mechanisms that underpin depth perception are still debated and therefore it becomes even more difficult trying to reproduce its working within an augmented scene [125].

In their general survey on AR, Azuma et al. [12] rightly claimed that «The basic goal of an AR system is to enhance the user's perception of and interaction with the real world through supplementing the real world with 3D virtual objects that appear to coexist in the same space as the real world». In the context of AR-based surgical navigation systems, researchers thus have tried to improve the perceptive efficiency of the human visual system in several ways. Some have modelled and contextually rendered the virtual content in a photo-realistic manner. Other researchers have adopted pixel-wise transparency maps and "virtual windows" [126] to recreate occlusions and motion parallax cues. Some of the proposed techniques for enhancing depth perception have included high-fidelity texturing [127] or colour coding methods [128], whereas others comprised lighting and shading cues and/or were based on the adoption of an interactive "virtual mirror" [129], [130]. Alternatively, depth perception can be improved by relying on standard stereopsis and two-view displays or on more complex full parallax multi-view displays. In any case, to the best of our knowledge, previously there have been no AR Visualization Processing techniques that have provided the surgeon with useful information able to significantly improve the postoperative outcome of specific surgical tasks.

Many surgical procedures in the field of orthopaedic surgery or maxillofacial surgery, generally involve the task of reducing displacements or correcting abnormalities between rigid anatomical structures, i.e. bones, based on pre-operative planning. The direct tracking of all the rigid anatomies involved in the procedure certainly would provide a measure of the six-degrees-of-freedom displacements between each couple of them and it would aid the correct performance of the surgical task, yet it is not always feasible for technical and logistic reasons. In case of single object tracking, the pointer of a standard surgical navigator can be employed by the surgeon to compare the final positions of clearly detectable reference points, over the repositioned anatomy with those of their counterparts from the surgical planning. Nevertheless, this approach does not allow the assessment of all of the six-degrees-of-freedom at the same time, and so the surgeon has to iteratively move the body and check the position of at least three points through the pointer of the navigator. This cycle can be tedious, since the verification is performed after any adjustment of the

pose of the object. Further, despite it is trivial to check the position of one point, same cannot be said if our goal is to align more points (i.e. at least three).

AR seems the optimal solution to aid this kind of surgical tasks because potentially it can show all the required information in a consistent and immediate fashion. Yet, the traditional AR interaction technique, featuring the superimposition of a semi-transparent virtual replica of the rigid anatomy in a position and orientation (pose) defined during planning, did not prove to be very effective in aiding the correct performance of those procedures as we demonstrated in the experiments that will be outlined in Chapter 5. To this end, it appears undoubtedly more beneficial and intuitive for the surgeon to deal with task-oriented Visualization techniques, more than with complex reproductions of the virtual anatomies through photorealistic rendering, transparencies and/or virtual windows.

4.3 METHODS

4.3.1 Perspective-n-Point Problem.

The task of estimating the pose of a camera with respect to a scene object given its intrinsic parameters and a set of n world-to-image point correspondences is known as the Perspective-n-Point (PnP) problem in computer vision and exterior orientation or space resection problem in photogrammetry.

This inverse problem concerns many fields of applications (structure from motion, robotics, augmented reality, etc.) and it was first formally introduced in the computer vision community by Fishler and Bolles in 1981 [131], albeit already used in the photogrammetry community before then.

According to Fishler and Bolles the PnP problem can be defined as follows (distance-based definition):

> *«Given the relative spatial locations of n control points* $P_i,\ i = 1, \dots n$*, and given the angle to every pair of these points from an additional point called the centre of perspective* C*, find the lengths* $D_i = |CP_i|$ *of the line segments joining* C *to each of the control points».*

The constraint equations are:

$$D_i^2 + D_j^2 - 2D_iD_j \cos\theta_{ij} = d_{ij}^2, \qquad i \neq j$$

Where $D_i = |CP_i|, D_j = |CP_j|$ are the unknown variables, $\theta_{ij} = \widehat{P_iCP_j}$ and $d_{ij} = |P_iP_j|$ are the known entries (Fig. 4.1). In computer vision θ_{ij} are determined finding the correspondences between world-to-image points and knowing the intrinsic camera parameters, while d_{ij} are established by the control points.

Following this definition, once each distance D_i is computed, the position of the points $P_i$can be expressed in the CRS. Therefore, being the position of each $P_i$in the SRS known, the problem of estimating camera pose with respect to the SRS is reduced to a standard absolute orientation problem whose solution can be found in closed-form fashion through quaternions [132] or singular value decomposition (SVD) [133].

The same problem is also known under the transformation-based definition [134] which can be formalized as:

$$\lambda_i \hat{p}_i = [K|0] \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \hat{P}_i \,, i = 1, \dots n$$

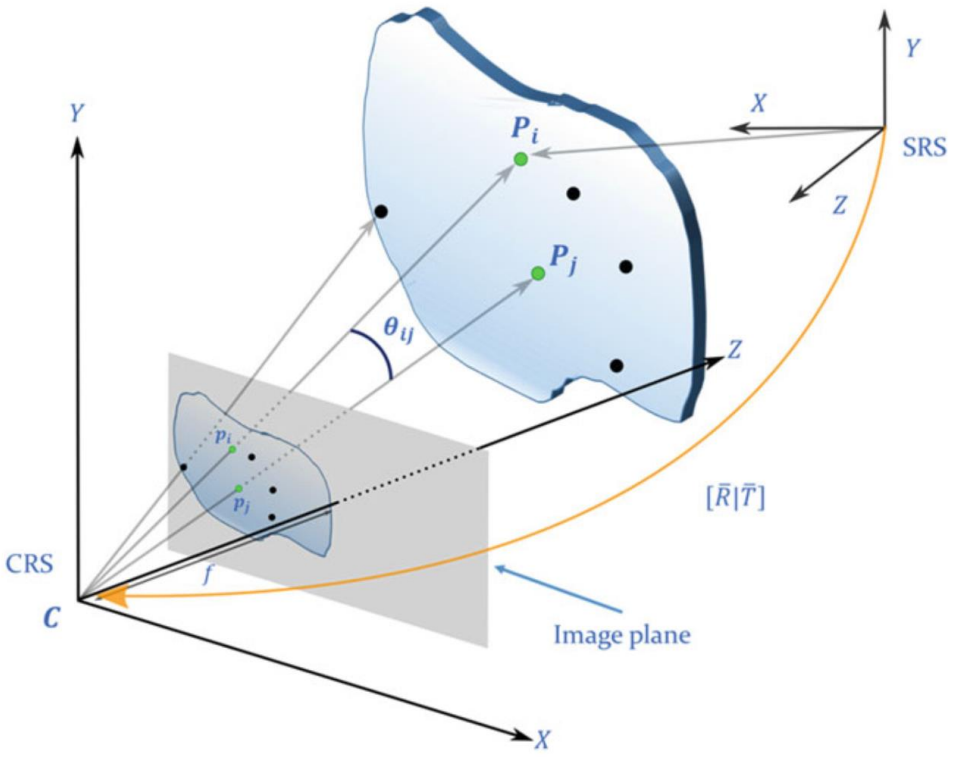


Fig. 4.1 *– Geometry of the PnP problem.*

Where the scene and image points $\hat{P}_i$and $\hat{p}_i$ are represented in homogeneous coordinates and the equation is up to a scale factor λ_i. Hence, according to this definition, the PnP problem aims at determining the pose (in terms of a rotation matrix R and a translation vector T) given a set of n world-to-image correspondences and known the intrinsic camera parameters encapsulated by the matrix K.

The PnP problem has been extensively studied by several groups, which have proposed different iterative, closed-form for solving it. Closed-form methods [135]–[141], directly provide an estimation of the camera pose but they are usually less accurate and more susceptible to noise than iterative methods. Iterative non-linear optimization methods solve the PnP problem by iteratively minimizing a cost function generally related to the geometric (reprojection residuals) or algebraic error but they need a good initial guess and yield only one solution at a time [142]–[144]. A useful overview of the state-of-the-art methods can be found in [140].

In terms of geometric reprojection residual, the non-linear cost function can be formulated as the sum of the squared measurement errors (d_i):

$$\overline{\mathrm{R}}|\overline{\mathrm{T}} = \arg min \sum_{i=1}^{n} d(p_i\, \hat{p}_i)^2 = \arg min \sum_{i=1}^{n} \| p_i - \hat{p}_i\big(K, \hat{R}, \hat{T}, P_i\big) \|^2$$

Where p_i are the measured image points, and $\hat{p}_i$ are the calculated projections of the corresponding control points as a function of $K, \hat{R}, \hat{T}$.

The other important research direction on the PnP problem is the study of the multi-solution phenomenon of the PnP problem [145], principally when n=3 (P3P) [146], being 3 the smallest subset of control points that yields a finite number of solutions. P3P problem yields at most four solutions which can be disambiguated using a fourth point, and it is the most studied case since it can be used as first step to reduce the complexity of the computation of a PnP problem, e.g. in a RANSAC scheme by removing the outliers.

4.3.2 AR Video-Based Camera Registration

Regardless of the method adopted for solving the PnP problem, an immediate application of the PnP problem is to locate the pose of a calibrated camera with respect to an object, given the 3D

position of a set of n control points rigidly constrained to the object and the 2D position of their correspondent projections onto the image plane.

For a correct registration of computer-generated elements to the real scene in AR-based surgical navigation systems, the image formation process of the virtual camera must perfectly mimic the real camera one. In mostly all the AR applications the estimation of the intrinsic camera parameters is the result of an off-line calibration process whereas the extrinsic camera parameters are determined online, e.g. solving a PnP problem in real-time. This video-based camera registration method suggested us the implementation of an ergonomic AR interaction paradigm for positioning and orienting a non-tracked rigid object in space.

4.3.3 Human-PnP.

Many surgical procedures in the field of orthopaedic surgery or maxillofacial surgery, involve the task of manually placing rigid anatomies on the basis of preoperative planning. In that case, let us assume that we can rely on a robust and accurate registration of the surgical planning onto the real scene, by means of the tracking of at least one rigid body (e.g. the head). The six-degrees-of-freedom pose of an additional and non-tracked rigid anatomy in relation to the SRS, can be retrieved by physically placing it as to minimize the geometric distance, on the image plane, between a set of real and virtual feature points. For brevity, from now on, we shall refer to these structures as "tracked anatomy" for the former and "non-tracked anatomy" for the latter, while the proposed method will be referred to as the *human-perspective-n-point problem (hPnP)*.

From a theoretical standpoint, our method draws its inspiration and physically mimics the paradigm on which the PnP problem is formulated. As mentioned in the previous section, the main goal of the PnP problem is to infer useful information on the real 3D scene, based on 2D observations of it. In an AR application, this spatial information is used to geometrically register the virtual elements onto the real scene. Thus, as a general rule and regardless of the method adopted for solving the PnP problem, a robust and accurate registration should minimize in the image plane the geometric reprojection residuals between measured and estimated projections. Similarly, the goal of our hPnP interaction paradigm is to achieve the desired placement of a non-tracked anatomy by manually minimizing the reprojection residuals between correct/planned projections $\bar{p}_i$ of virtual landmarks, and observed projections $\hat{\bar{p}}_i$ of real landmarks. The correct/planned projections $\bar{p}_i$ are rendered on

the image plane according to the real-time estimation of the camera pose $[\overline{\mathrm{R}}, \overline{\mathrm{T}}]$ relative to the tracked anatomy reference system (SRS) and assuming the intrinsic camera parameters, encapsulated by matrix K, are determined offline, e.g. through the Zhang's method [147]. The position of each virtual landmark P_i in the SRS is established during surgical planning.

The real projections $\hat{\bar{p}}_i$ are associated with the pose, encapsulated by $\left[\hat{\bar{R}}, \hat{\bar{T}}\right]$, between viewing camera and non-tracked anatomy reference frame (ARS): this resulting pose varies according to the manual placement of the rigid body relative to the camera:

$$\widetilde{\mathrm{R}}|\widetilde{\mathrm{T}} = \arg min \sum_{i=1}^{n} d(p_i\, \hat{\bar{p}}_i)^2 = \arg min \sum_{i=1}^{n} \| \bar{p}_i(K, \bar{R}, \bar{T}, P_i) - \hat{\bar{p}}_i\left(K, \hat{\bar{R}}, \hat{\bar{T}}, P_i\right) \|^2$$

In this way, we wish to obtain $\left[\widetilde{\mathrm{R}}|\widetilde{\mathrm{T}}\right] \approx [\overline{\mathrm{R}}|\overline{\mathrm{T}}]$ (Fig. 4.2), namely we seek to positioning and orienting the ARS as coincident with the planned and registered SRS (non-tracked anatomy reference frame ≈ planning reference frame).

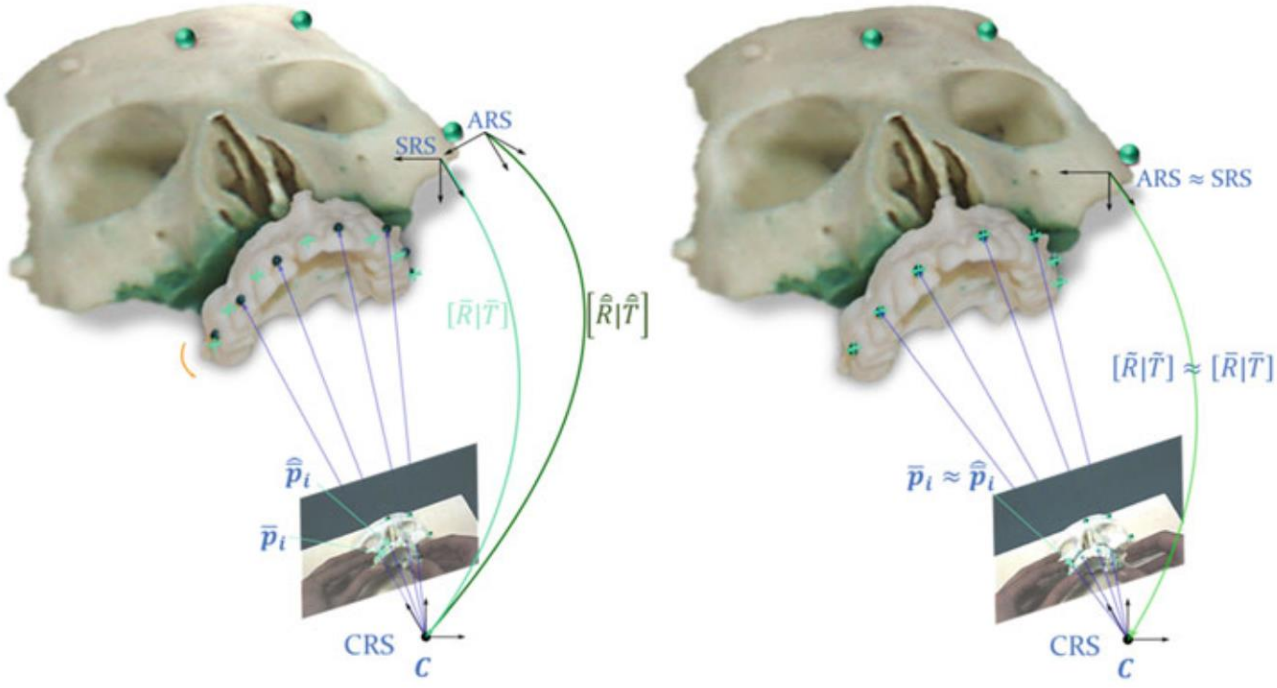


Fig. 4.2 *– Geometry of the hPnP: minimizing the reprojection residual between registered projections $\bar{p}_i$ and real projections $\hat{\bar{p}}_i$ is sufficient to aid the accurate placement of a rigid body (the maxilla in the image) in space.*

To implement this strategy, we add simple virtual elements (e.g. virtual asterisks, crosses, etc.) to the virtual scene during the surgical planning: one element for each of the clearly detectable physical landmarks on the rigid body. The landmarks may consist of a series of distinguishable feature points over the surface of the anatomy or rigidly constrained to it. Under such AR guidance, the user moves the non-tracked rigid body up to obtain a perfect overlapping between real and virtual landmarks, hence manually minimizing the reprojection residuals on the image plane: $\bar{p}_i \approx \hat{\bar{p}}_i \forall i$ (Fig. 4.3). The theoretical assumptions underpinning the PnP problem ensure that if $\bar{p}_i \approx \hat{\bar{p}}_i \forall i$, the non-tracked anatomy is placed in the correct pose as planned in SRS.

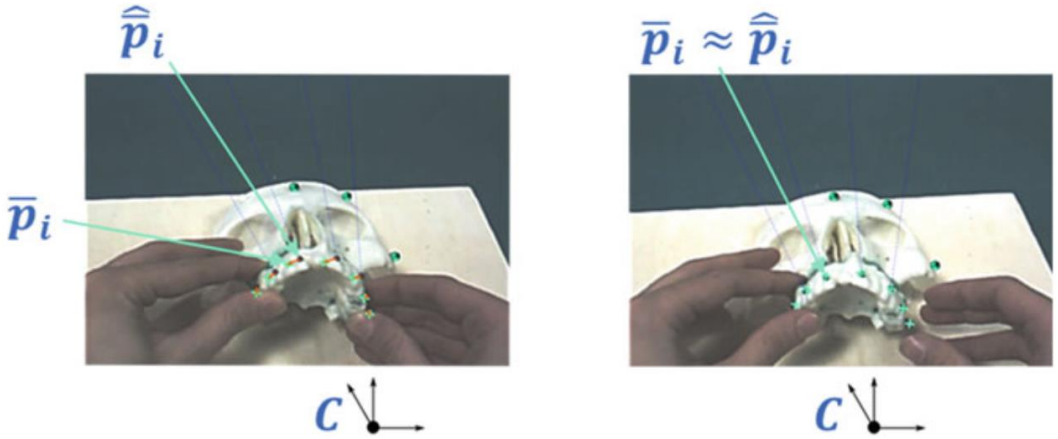


Fig. 4.3 *– Detail of the image plane with the minimization of the reprojection residuals. Here the virtual information consists of a cyan-colored asterisk for each physical landmark clearly detectable over the maxilla, e.g. coloured landmarks fixed on the brackets of the orthodontic appliance (Color figure online).*

4.4 DISCUSSION

A novel and ergonomic AR interaction paradigm has been introduced, that aims at obtaining the accurate placement of a rigid body in space without the need for multiple objects tracking. From a theoretical standpoint, our method draws its inspiration and physically mimics the paradigm on which the PnP problem in computer vision is formulated. This approach, termed *hPnP*, could be of help in those tasks, also not specifically surgical, where the AR guide aims at aiding the placement of a rigid body in space.

The key-principle behind this interaction paradigm can be exploited in many different AR-based navigation systems: it can be integrated with different end-products of the Visualization process in

terms of display technology and Perception Location and/or it could be realized in conjunction with various tracking modalities.

To improve robustness and applicability of the proposed AR interaction paradigm in a real clinical scenario, redundancy in choosing the set of landmarks must be granted. Further, the presence of line-of-sight occlusions caused by soft-tissues, surgeon's hands or surgical instrumentation may be restricted by conveniently selecting the position of the landmarks in relation to the surgical field.

Lastly, it is important to note that the chosen AR interaction paradigm was not bound to the particular video-based tracking technique, neither to the use of a specific wearable stereoscopic system. The user can enhance the accuracy in object placement by checking consistency of real and virtual landmarks from different viewpoints. In this regard, the ergonomics of the proposed method may benefit from the adoption of a wearable AR system. Moreover, the choice of such instance of Visualization data was, in that work, empirically inspired by the authors' endeavour of defining a modality that were ergonomic for the surgeon and that provided the smallest perceived parallax error.

5 CHAPTER 5: AR-BASED VIDEO SEE-THROUGH HMD AS AN AID FOR SURGICAL MAXILLARY REPOSITIONING

In recent years, the discipline of Cranio-Maxillofacial (CMF) Surgery has experienced an extraordinary rate of technological innovation. This is because the complex 3D anatomy of the face, together with the need for surgical precision and the increasing number of requests for morphological surgery, have resulted in surgeons demanding advanced technological assistance. Thus, surgical planning software platforms and navigation systems are today widely used by maxillofacial surgeons [25], [26], [34], [35], [148]–[150].

However, the complexity of such surgery and the extended operative times required, have compromised the widespread implementation of such devices. Moreover, the necessary equipment is expensive [151]. For these reasons, the technology demands both methodological and economic rationalisation.

Bearing these facts in mind, in this study the first version of HMD was used to elaborate a useful strategy for delivery of AR information to the CMF surgeon. As previously stated for the survey presented in Chapter 4, this study is the result of a collaboration between our Department, the Maxillofacial Surgery Unit of the S. Orsola-Malpighi University Hospital of Bologna, and EndoCAS Laboratory of the University of Pisa. The aim of the present study was to validate the accuracy of the video see-through HMD when used as an aid during surgery on facial bones, specifically repositioning of the maxilla after LeFort 1 osteotomy, a standard procedure in orthognathic surgery. The potential for the wider application of the HMD in maxillofacial surgery was also explored [152].

5.1 THE SELECTED SURGICAL PROCEDURE: LEFORT1 OSTEOTOMY

Orthognathic Surgery is that field of CMF Surgery aiming to treat dentofacial deformities, thus to normalize facial disharmony and/or asymmetry with a specific focus on dental occlusion (Fig. 5.1). The term *dentofacial deformity* refers to significant deviations from normal proportions of the maxillo-mandibular complex that negatively affect facial harmony and the relationship of the teeth within each arch and the relationship of the arches with one another (occlusion) [153].

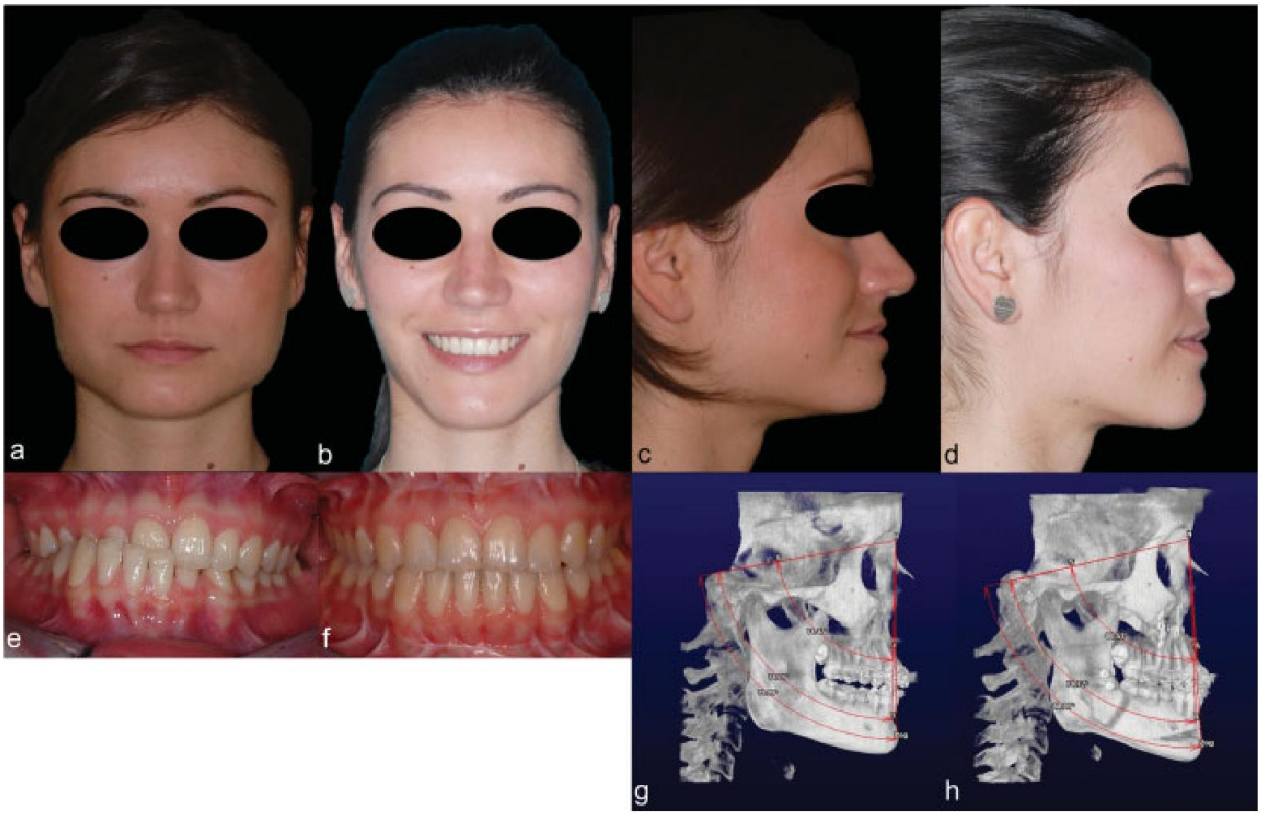


Fig. 5.1 *– An example of treatment of a dentofacial deformity* [35]

It is not the purpose of the present thesis to give and extended description of this area of CMF surgery. Thus, this subsection will be limited to the essential information needed to understand what kind of surgical procedure has been selected as a test and why.

To achieve the above described aim of orthognathic surgery, the surgeon can perform selective osteotomies of the facial bones, namely the maxilla, the mandible and its border (gonial angles, lower borders and the chin). The standard osteotomy to mobilize and reposition the maxilla is called LeFort1 osteotomy, according to the surgeon who first described this osteotomy line for trauma cases. It represents a complete detachment of the upper maxilla just above the palate (i.e., the nasal floor), potentially including also the lower part of the zygomatic buttresses for better volume enhancement of the malar regions (Fig. 5.2). This surgical manoeuvre is usually performed with the aid of a surgical occlusal wafer that reproduce the planned maxillary position in relation to the original mandibular dental arch. Then, maxilla is fixed in its planned position using titanium mini-plates and screws (Fig. 5.2).

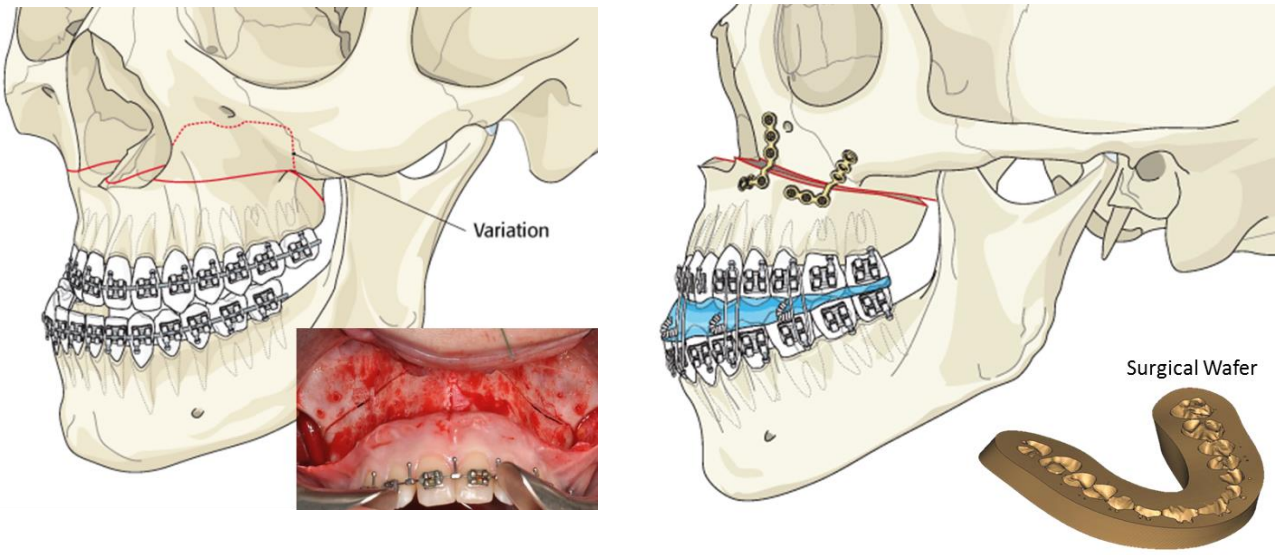


Fig. 5.2 *– LeFort1 osteotomy line (left) and surgical occlusal wafer(right) (Modified from AO Foundation online library).*

Nowadays, this procedure is planned with the aid of surgical simulation software, that allow the surgeon to three-dimensionally measure facial skeletal dimensions and bony parts mutual relationships (cephalometry) and to simulate the surgical procedure in a 3D environment (Fig. 5.3) [25], [26]. That is particularly useful for asymmetric faces.

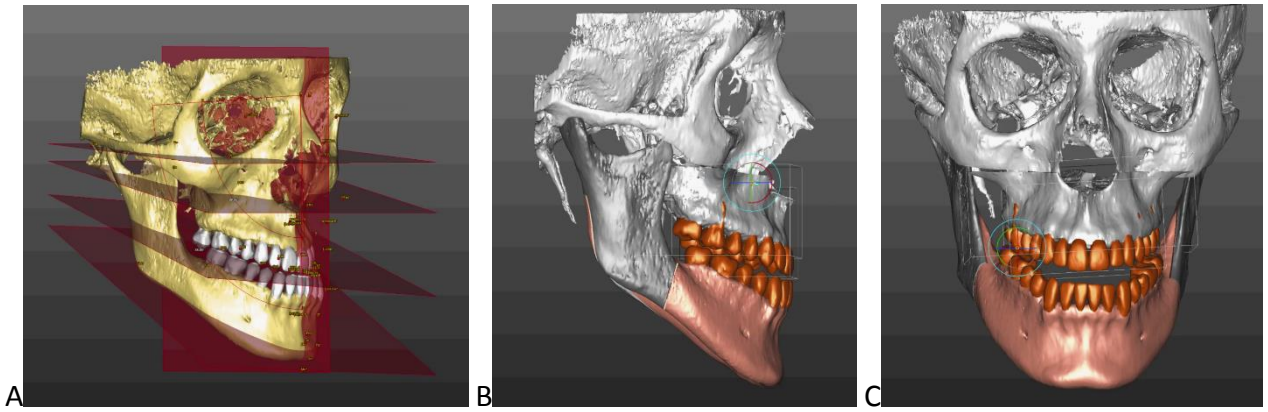


Fig. 5.3 *– A: 3D cephalometry; B,C: 3D rendered image of a LeFort1 osteotomy, lateral and frontal aspects (obtained by a computed surgical simulation software, Simplant O&O, Dentsply Sirona, York, USA).*

LeFort 1 osteotomy was chosen as a test procedure because it is a standard technique, very well-known and repeatable [153]. Moreover, it is a one-piece bone segment repositioning in the context of CMF surgery, properly exemplifying any other bone segment repositioning procedure for the cranio-maxillofacial area.

5.2 Materials and Methods

5.2.1 Implementation of the Video-Based Tracking Method

Here is a brief description of the specific implementation of the video-based tracking method extensively described in the previous chapters. The online estimation of the pose of the stereo camera pair (Camera reference system, abbreviated to CRS) relative to the reference system of the surgical planning (SRS), is the result of a marker-based video registration method. Such method relies on the localization of three physical markers (red spheres) rigidly constrained to the head phantom and whose position in the virtual scene (SRS) is recorded during planning. The key characteristic of the implemented method for registering the preoperative planning to the live views of the surgical scene (i.e. the patient phantom) is that it is not based on the adoption of a cumbersome external localizer. Standard surgical navigation systems, featuring the use of external infrared trackers, may in fact introduce unwanted line-of-sight constraints into the operating room as well as add error-prone technical complexity to the surgical workflow. The proposed video- based algorithm proved to be an ergonomic and functional implementation of the video see-through paradigm (Fig. 5.4 – 5.5).

5.2.2 In vitro Set-Up

The tests were conducted on a replica of a cadaveric human skull. The real skull underwent CT scanning and the DICOM files were segmented using a semi-automatic segmentation tool integrated into the ITK-Snap open-source platform [122].

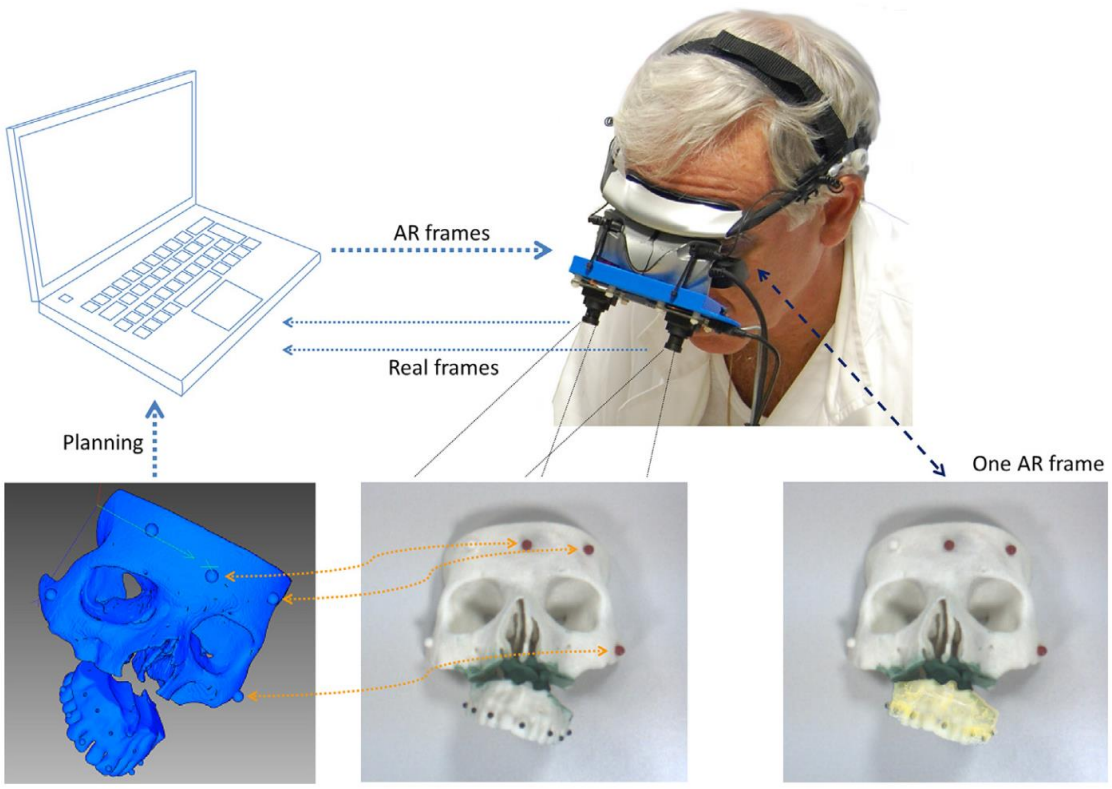


Fig. 5.4 *– The first implementation set-up.*

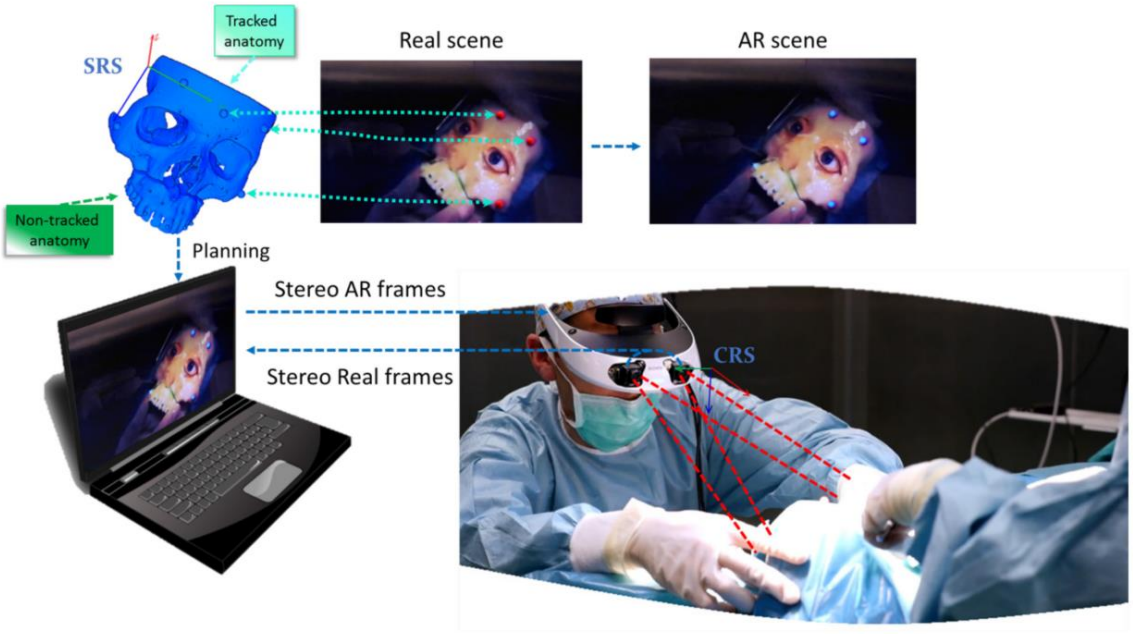


Fig. 5.5 *– The latter implementation set-up.*

Manual segmentation refinement was performed to obtain detailed information on small anatomical structures (e.g. the foramen rotundum, foramen spinosum, lamina cribriformis, and hypoglossal canal). The 3D virtual model distinguished pneumatised bones very well. In particular, the nasal cavities and the paranasal sinuses were computer-generated in minute detail.

The virtual model of the skull was cut along the LeFort 1 osteotomy line. The two resulting virtual objects (the upper skull and maxilla) were exported as STL files and replicated in ABS using a 3D printer (Stratasys Elite; Edina, MN). As already stated, LeFort 1 osteotomy and repositioning of the upper maxilla were chosen as test procedures featuring the principal features of maxillofacial surgery. Thus, the technique involves surgery on facial bones; the approach is a form of semi-buried surgery performed under real clinical conditions; the technique involves complex 3D movements of a rigid object in space; and the technique is often performed in clinical practice worldwide. It was chosen to perform and represent this manoeuvre without the aid of surgical occlusal wafers, namely wafer-less, to better understand the value of the AR device as a sole tool aiding the procedure.

Before printing, three 6-mm-diameter spheres were inserted into the virtual model as reference markers for the video-based tracking method described in the previous chapters. Further, three reference holes were drilled into the vestibular cortical bone, over the teeth (anterior in the pre-maxillary region; posterior left and posterior right in the respective molar regions). These holes were used as references to evaluate the position of the maxilla. Thus, each hole was designed to receive the tip of the tester probe used for validation (Fig. 5.6), to guarantee unique selection of each reference point.

The upper skull was fixed on a wooden holder. The maxillary piece was connected to the upper skull with plasticine (this material is highly malleable but rigid when shaped). This construction served as a fixing device for the maxilla once the planned position was attained, yet allowed the maxilla to be manually adjusted in space.

To evaluate the accuracy of our system, a traditional navigation platform (the eNlite Navigation System running iNtellect Cranial Navigation Software version 1.0; Stryker, Freiburg, Germany) featuring an active infrared localiser was used. The entire setup is shown in Figure 5.6, which identifies the tracking and pointing instruments of the commercial navigation system.

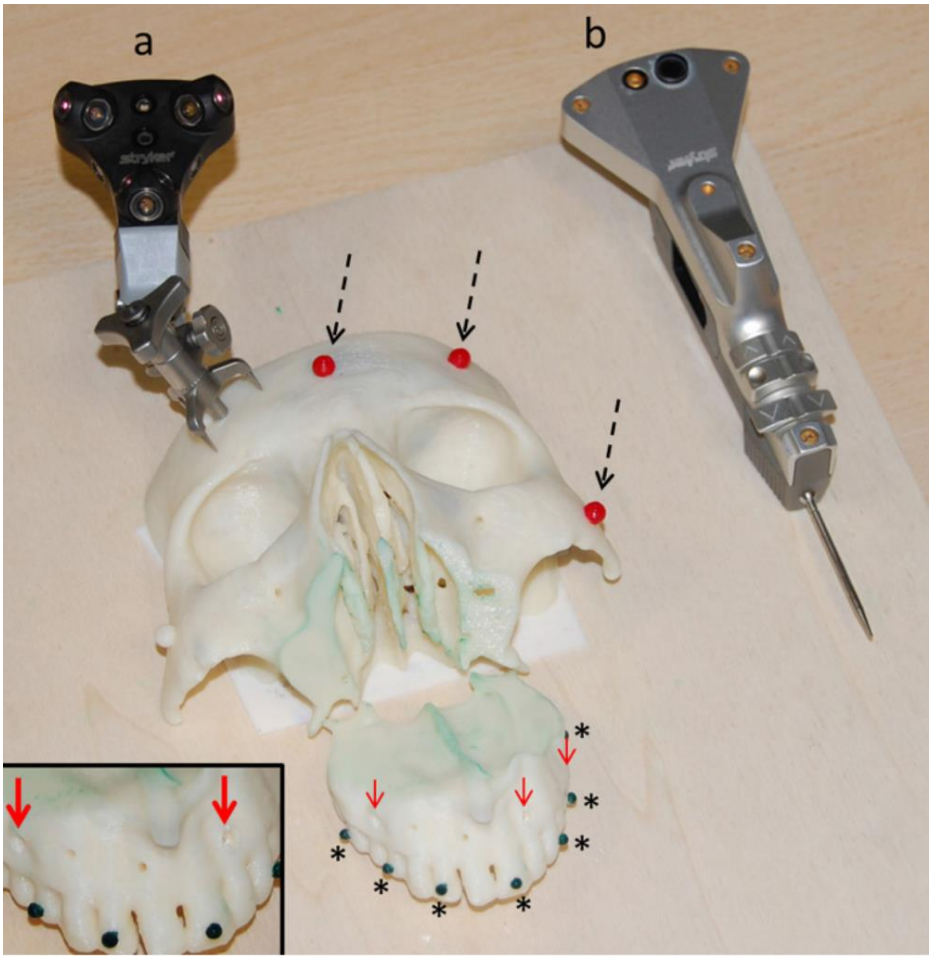


Fig. 5.6 *– The Experimental setup: A physical replica of the human skull is fixed onto a wooden holder, and the three coloured spheres on the model (the black dashed arrows) are used as reference markers for the video-based tracking method employed. The coloured brackets on the teeth (the black asterisks) are the reference markers for the hPnP-like AR interaction paradigm; three holes on the maxilla (the red arrows) were used to evaluate accuracy with the aid of an external navigation system. The tracker of the navigation system is fixed onto the model in (a). In (b), the pointer of the navigation system, used to assess the position of reference holes, is shown beside the model.*

5.2.3 AR Visualization: hPnP Approach

A preliminary assessment was conducted to evaluate the ergonomics of the device, actual usability in a surgical environment, and (in particular) the best method of displaying the virtual content. One surgeon and three engineers collaborated in this work. Tests were conducted using different Visualization Processing modalities in an attempt to define a modality that was optimally comfortable and that offered the smallest perceived parallax error. The traditional AR interaction technique, featuring the superimposition of a semi-transparent virtual replica of the maxilla, as

dictated by the surgical planning, did not prove to be very effective in aiding the surgeon in manually repositioning the upper maxilla. This was mostly due to the surgeon's limited perception of the relative distances of objects within the AR scene owing to the presence of unnatural occlusions between the real and the virtual maxilla. Conversely, a more ergonomic form of visualization consisted in the use of an interaction paradigm which did actualize the previously described hPnP approach (Chapter 4). In this approach, physical landmarks onto the maxilla were designated as reference markers for the AR Visualization modality. The physical landmarks correspond to coloured landmarks fixed on the brackets of the orthodontic appliance usually applied prior to this kind of interventions. The repositioning of the maxilla is therefore assisted by visually aligning small virtual asterisks, drawn in positions defined during planning (i.e. in the scene reference system SRS), with the corresponding real landmarks (Fig. 5.7).

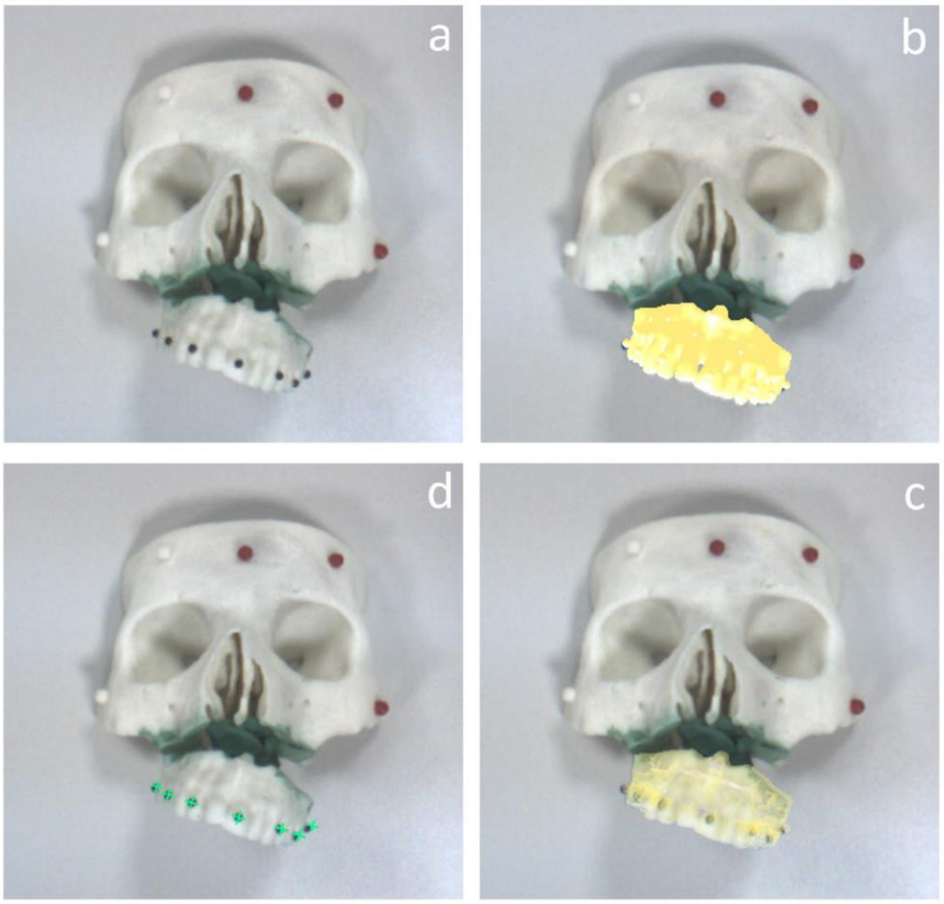


Fig. 5.7 *– Different Visualization Processing modalities: (a) Real video frame; (b, c) Traditional approach, presenting the virtual model on real camera frames. Using the approaches of (b) and (c), it was not possible to completely perceive the spatial relation- ships between the real and virtual world. (d) A more ergonomic form of Visualization, that actualizes the hPnP interaction paradigm. The virtual information consists of a green asterisk for each coloured landmark on the maxilla.*

5.2.4 Accuracy Evaluation Testing

Virtual Surgical Planning. Using Maya (Autodesk; Toronto, Canada), the virtual maxilla was moved in space as dictated by three surgical planning of increasing complexity.

• Maxilla 6 mm forward;

• Maxilla 5 mm forward and 1 mm downward;

• Maxilla 6 mm forward, 1 mm downward, and with 15° roll and 10° pitch.

Each planning was stored as an STL file to be loaded by the software that manages the AR (Fig. 5.8).

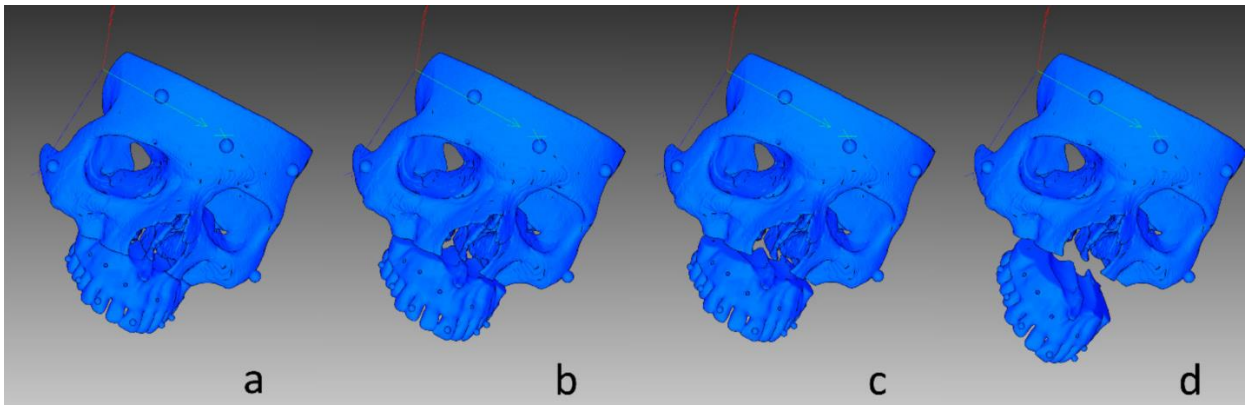


Fig. 5.8 *– The virtual maxilla (a) was moved in space as dictated by three surgical planning of increasing complexity: (b) 6 mm forward; (c) 5 mm forward and 1 mm downward; (d) 6 mm forward, 1 mm downward, with 15° roll and 10° pitch.*

Tests. Three maxillofacial surgeons, three trainees in maxillofacial surgery, and three engineers were involved in the testing; we evaluated interobserver variability. Hence, the three groups included appropriate representatives of users with different levels of surgical skill (from unskilled engineers to highly skilled surgeons). After only one warm-up session, during which each subject familiarized with the HMD and with the interaction paradigm, the subject was asked to manually reposition the maxillary segment, using the AR guide. Each subject for each of the three virtual plans repeated the procedure; the maximum test duration was 15 min. After completion of each test, the position of the maxillary segment was confirmed using the navigation system described below.

Accuracy Measurement. The CT scan of the skull was imported into the navigation system as a DICOM file and the three plans, defined in the CT reference system (SRS), were loaded into the navigation system as STL files (Fig. 5.9). The tracker of the navigation system was fixed on the model of the skull and the registration process featured a point-based procedure (using defined anatomical points) with subsequent surface refinement; the target registration error was 0.3 mm. After each trial session, the navigation system probe was inserted into each of the three reference holes on the maxilla and the probe tip positions were saved (Fig. 5.10). For each subject, the linear distances between the real positions of the reference holes (measured using the navigation system) and the expected positions (defined during surgical planning).

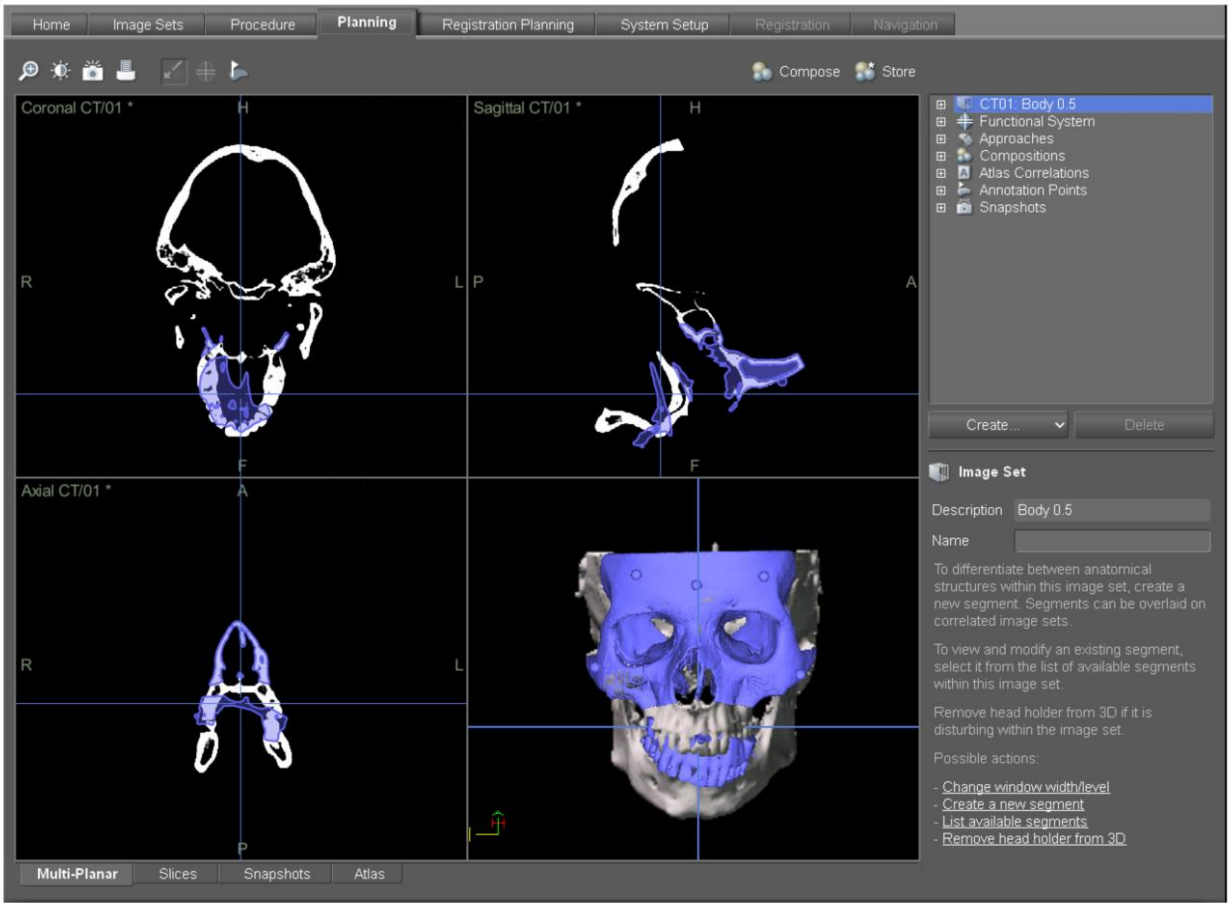


Fig. 5.9 *– A screenshot of the navigator. The blue planning scenario is loaded together with the original CT scan*

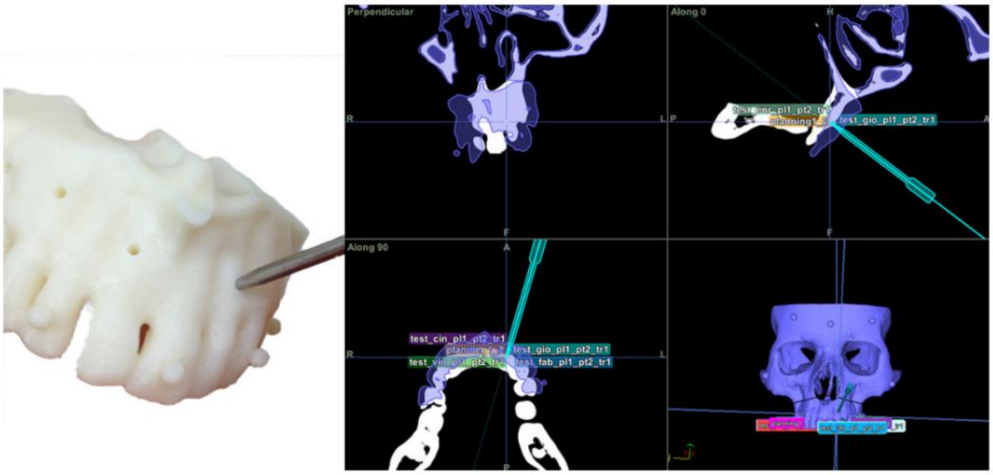


Fig. 5.10 *– The accuracy evaluation process is shown in detail. On the left, the pointer slides into a reference hole of the maxilla (the hole termed "anterior one"); on the right, the navigation system shows where the tip of the pointer is actually located (compared with the planned location). The coordinates of the real position are recorded and used to estimate errors in linear measurements.*

5.3 RESULTS

The linear results are shown in Table 5.1. The mean error was 1.70 ±0.51 mm. The axial errors were 0.89 ± 0.54 mm on the sagittal axis, 0.60 ± 0.20 mm on the frontal axis, and 1.06 ± 0.40 mm on the cranio-caudal axis. The simplest planning was associated with a slightly lower mean error (1.58 ± 0.37 mm) than the more complex plans (medium:1.82 ± 0.71 mm; difficult:1.70 ±0.45 mm). The mean error for the anterior reference point was lower (1.33 ± 0.58 mm) than those for the posterior right (1.72 ± 0.24 mm) and posterior left (2.05 ± 0.47 mm) points.

	Plan 1	Plan 2	Plan 3	Mean
Target 1	1.71 mm (±0.24)	1.80 mm (±0.18)	1.63 mm (±0.34)	1.72 mm (±0.24)
Target 2	1.07 mm (±0.17)	1.47 mm (±0.12)	1.45 mm (±0.45)	1.33 mm (±0.58)
Target 3	1.96 mm (±0.32)	2.18 mm (±0.69)	2.02 mm (±0.49)	2.05 mm (±0.47)
Mean	1.58 mm (±0.37)	1.82 mm (±0.71)	1.70 mm (±0.45)	**1.70 mm (±0.51)**

Table 5.1

Angular errors were also computed, according to Fig. 5.11 model. Mean pitch was 3.13°±1.89°, mean roll was 1.99°±0.95° and mean yaw was 3.25°±2.26°. Angular errors were computer also according to the plan, i.e. the complexity of the simulated procedure: these results are summarized in Table 5.2.

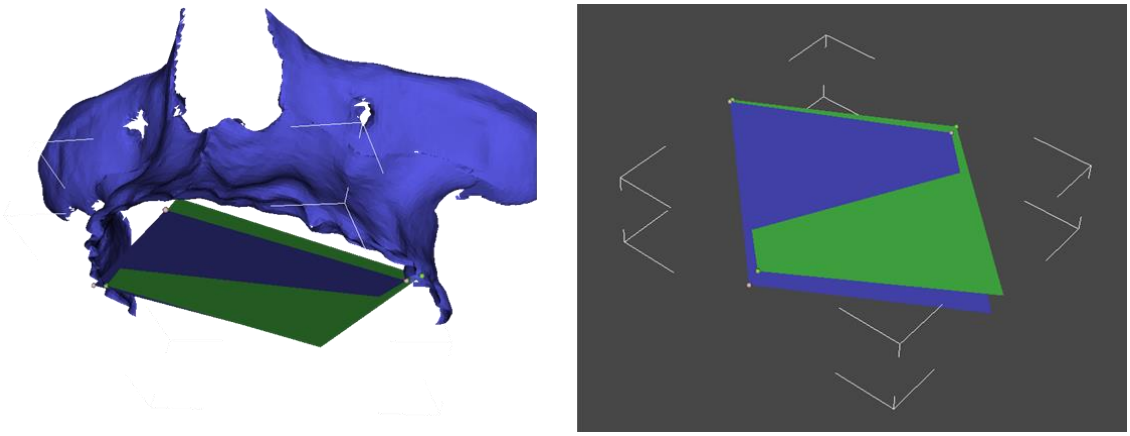


Fig. 5.11 *– Angular errors were computer according to the angular discrepancies between the plane built on the three planned points and the plane built on the three achieved points*

	Plan 1	**Plan 2**	**Plan 3**
PITCH	3.12°±2.68°	2.89°±0.80°	3.40°±1.89°
ROLL	2.27°±1.22°	1.46°±0.57°	2.24°±0.81°
YAW	3.25°±3.29°	3.02°±0.82°	3.48°±2.24°

Table 5.2

Moreover, no significant difference was noted among operators, despite variation in surgical experience (Fig. 5.12). Feedback from surgeons was acceptable; all procedures were completed within 15 min and the tool was found to be both comfortable and usable.

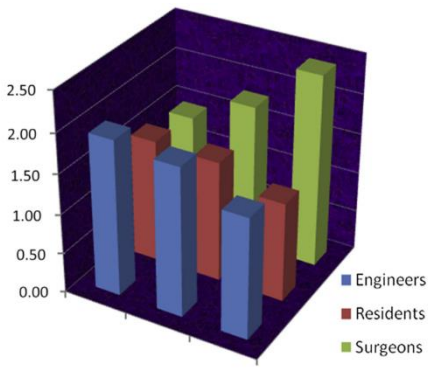


Fig. 5.12 *– Mean errors in mm (over three trials and three reference holes) for each of the nine participants. No difference between engineers and physicians is evident.*

5.4 DISCUSSION

This represents the first study testing an AR-based HMD device for CMF surgery. The efficacy of the wearable AR system as surgical navigator was validated in combination with the ergonomic AR interaction paradigm presented in Chapter 4. The system proved to be an effective surgical aid when used to assist in wafer-less maxillary repositioning.

The positive results were obtained without the tracking of the maxilla but just relying on the chosen AR interaction paradigm: the overlapping on the image plane between virtual feature points and real landmarks, visible over the non-tracked anatomy, proved to be sufficient to assist the accurate repositioning of the maxilla. Further, the video-based tracking solution obviates any need for an external localiser, because the stereo camera pair acts as frame-grabber and as localiser.

The obtained results suggested that the AR HMD can be both comfortable and functional, permitting a surgeon to maintain their natural operative posture during surgery performed at different angles, without losing the 3D relationship between the real scene and that afforded by virtual planning. This is of particular importance. During this typology of surgery, surgeons in fact are frequently asked to change their line of view to control the 3D position of the maxilla from all angles. In the same manner, with the proposed AR interaction paradigm, the user can enhance the accuracy in object placement by checking consistency of real and virtual landmarks from different viewpoints. In this regard, the ergonomics of the proposed method may benefit from the adoption of a wearable AR system even if it is important to note that the chosen Visualization modality was not bound to the particular video-based tracking solution adopted in the study, neither to the use of a specific wearable stereoscopic system.

From the clinical point of view, the device was certainly promising. Results suggest that the device would be accurate when used to assist in wafer-less maxillary repositioning during the LeFort 1 orthognathic procedure. Linear and angular errors were considered clinically acceptable according to current good surgical practice, especially for a wafer-less procedure.

Further, data suggest that the method can be extended to aid the performance of many surgical procedures on the facial skeleton, but more tests are needed to improve user's experience in terms of quality of information displayed. Also, in vivo testing should be performed to assess system accuracy under real clinical conditions.

6 REFERENCES

[1] B. Furht, *Handbook of Augmented Reality*. 2011.

[2] P. Milgram and F. Kishino, “Taxonomy of mixed reality visual displays,” *IEICE Trans. Inf. Syst.*, vol. E77-D, no. 12, pp. 1321–1329, 1994.

[3] S. Mann, “Fundamental issues in mediated reality, WearComp, and camera-based augmented reality,” *Fundam. Wearable Comput. Augment. Reality, Lawrence Erlbaum Assoc. Inc*, pp. 295–328, 2001.

[4] S. Mann, “Mediated Reality with implementations for everyday life,” *Presence Teleoperators Virtual Environ.*, pp. 1–13, 2002.

[5] J. Smart, J. Cascio, and J. Paffendorf, “Metaverse Roadmap: Pathways to the 3D Web,” *Metaverse a cross-industry public foresight Proj.*, 2007.

[6] N. Stephenson, “Snow crash,” *Futures*, vol. 26. pp. 798–800, 1994.

[7] M. Billinghurst, A. Clark, and G. Lee, “A Survey of Augmented Reality Augmented Reality,” *Found. Trends Human-Computer Interact.*, vol. 8, no. 2–3, pp. 73–272, 2015.

[8] I. E. Sutherland, “The ultimate display,” *Proc. Congr. Int. Fed. Inf. Process.*, vol. 21, no. 3, pp. 506–508, 1996.

[9] T. P. Caudell and D. W. Mizell, “Augmented reality: an application of heads-up display technology to\nmanual manufacturing processes,” *Proc. Twenty-Fifth Hawaii Int. Conf. Syst. Sci.*, vol. ii, 1992.

[10] N. Navab, J. Traub, T. Sielhorst, M. Feuerstein, and C. Bichlmeier, “Action- and workflow-driven augmented reality for computer-aided medical procedures,” *IEEE Comput. Graph. Appl.*, vol. 27, no. 5, pp. 10–14, 2007.

[11] R. Azuma, “A survey of augmented reality,” *Presence Teleoperators Virtual Environ.*, vol. 6, no. 4, pp. 355–385, 1997.

[12] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre, “Recent advances in augmented reality,” *IEEE Comput. Graph. Appl.*, vol. 21, no. 6, pp. 34–47, 2001.

[13] O. Bimber and R. Raskar, “Modern approaches to augmented reality,” *ACM SIGGRAPH 2006 Courses*, p. 1, 2006.

[14] O. Bimber and R. Raskar, *Spatial Augmented Reality Merging Real and Virtual Worlds*, vol. 6. 2005.

[15] Gartner, “Newsroom Gartner ’ s 2015 Hype Cycle for Emerging Technologies Identifies the Computing Innovations That Organizations Should Monitor,” *Gartner*, no. Stage 5, pp. 2–3, 2015.

[16] M. Steinert and L. Leifer, “Scrutinizing Gartner’s Hype Cycle Approach,” in *PICMET 2010 Proceedings*, 2010, no. October, pp. 1–13.

[17] F. Cutolo, “Wearable Stereoscopic Augmented Reality System for Medical Procedures,” University of Pisa, 2015.

[18] G. A. Moore, *Crossing the Chasm: Marketing and Selling High-Tech Products to Mainstream Customers*, vol. Rev. 1999.

[19] E. M. Rogers, *Diffusion of innovations*. 1995.

[20] D. W. Roberts, J. W. Strohbehn, J. F. Hatch, W. Murray, and H. Kettenberger, “A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope.,” *J. Neurosurg.*, vol. 65, no. 4, pp. 545–549, 1986.

[21] C. B. Wilson, “Adoption of new surgical technology.,” *BMJ*, vol. 332, no. 7533, pp. 112–114, 2006.

[22] T. M. Peters, “Image-guidance for surgical procedures.,” *Phys. Med. Biol.*, vol. 51, no. 14, pp. R505–R540, 2006.

[23] T. M. Peters, “Image-guided surgery: from X-rays to virtual reality.,” *Comput. Methods Biomech. Biomed. Engin.*, vol. 4, no. 1, pp. 27–57, 2000.

[24] T. Peters and K. Cleary, *Image-guided interventions: Technology and applications*, vol. 53, no. 9. 2008.

[25] C. Marchetti, A. Bianchi, M. Bassi, R. Gori, C. Lamberti, and A. Sarti, “Mathematical modeling and numerical simulation in maxillo-facial virtual surgery (VISU).,” *J. Craniofac. Surg.*, vol. 17, no. 4, pp. 661–7; discussion 668, Jul. 2006.

[26] A. Bianchi, L. Muyldermans, M. Di Martino, L. Lancellotti, S. Amadori, A. Sarti, and C. Marchetti, “Facial soft tissue esthetic predictions: validation in craniomaxillofacial surgery with cone beam computed tomography data.,” *J. Oral Maxillofac. Surg.*, vol. 68, no. 7, pp. 1471–9, Jul. 2010.

[27] L. Adams, W. Krybus, D. Meyer-Ebrecht, R. Rueger, J. M. Gilsbach, R. Moesges, and G. Schloendorff, “Computer-Assisted Surgery,” *IEEE Comput. Graph. Appl.*, vol. 10, no. 3, pp. 43–51, 1990.

[28] M. Hofer, G. Strauss, K. Koulechov, M. Fischer, T. Neumuth, C. Trantakis, W. Korb, A. Dietz, and T. Lüth, “Computer Assisted ENT Surgery,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 1, no. S1, pp. 311–323, 2006.

[29] J. L. P. De Engenharia and U. De, “Computer-Assisted Surgery,” *Vis. Comput. Med.*, vol. 39, no. Suppl 1, pp. 593–623, 2014.

[30] A. Hodgson, “Computer-assisted orthopedic surgery,” in *Image-Guided Interventions: Technology and Applications*, 2008, pp. 333–386.

[31] P. S. Young, S. W. Bell, and A. Mahendra, “The evolving role of computer-assisted navigation in musculoskeletal oncology,” *Bone Jt. J.*, vol. 97-B, no. 2, pp. 258–264, 2015.

[32] A. Bianchi, L. Muyldermans, M. Di Martino, L. Lancellotti, S. Amadori, A. Sarti, and C. Marchetti, “Facial Soft Tissue Esthetic Predictions: Validation in Craniomaxillofacial Surgery With Cone Beam Computed Tomography Data,” *J. Oral Maxillofac. Surg.*, vol. 68, no. 7, pp. 1471–1479, 2010.

[33] S. Mazzoni, G. Badiali, L. Lancellotti, L. Babbi, A. Bianchi, and C. Marchetti, “Simulation-guided navigation: a new approach to improve intraoperative three-dimensional reproducibility during orthognathic surgery.,” *J. Craniofac. Surg.*, vol. 21, no. 6, pp. 1698–1705, 2010.

[34] A. Bianchi, G. Badiali, L. Piersanti, and C. Marchetti, “Computer-Assisted Piezoelectric Surgery,” *J. Craniofac. Surg.*, vol. 26, no. 3, pp. 867–872, 2015.

[35] G. Badiali, A. Roncari, A. Bianchi, F. Taddei, C. Marchetti, and E. Schileo, “Navigation in Orthognathic Surgery: 3D Accuracy,” *Facial Plast. Surg.*, vol. 31, no. 5, pp. 463–473, 2015.

[36] M. Kersten-Oertel, P. Jannin, and D. L. Collins, “DVV: Towards a Taxonomy for Mixed Reality Visualization in Image Guided Surgery,” *Med. Imaging Augment. Real.*, vol. 6326, no. 2, pp. 334–343, 2010.

[37] C. Bichlmeier, O. Ben, S. M. Heining, A. Ahmadi, and N. Navab, “Stepping into the operating theater: ARAV - Augmented Reality Aided Vertebroplasty,” in *Proceedings - 7th IEEE International Symposium*

on Mixed and Augmented Reality 2008, ISMAR 2008, 2008, pp. 165–166.

[38] T. Sielhorst, M. Feuerstein, and N. Navab, “Advanced medical displays: A literature review of augmented reality,” *IEEE/OSA J. Disp. Technol.*, vol. 4, no. 4, pp. 451–467, 2008.

[39] S. Nicolau, L. Soler, D. Mutter, and J. Marescaux, “Augmented reality in laparoscopic surgical oncology,” *Surg Oncol*, vol. 20, no. 3, pp. 189–201, 2011.

[40] V. Ferrari, M. Carbone, C. Cappelli, L. Boni, F. Melfi, M. Ferrari, F. Mosca, and A. Pietrabissa, “Value of multidetector computed tomography image segmentation for preoperative planning in general surgery.,” *Surg. Endosc.*, vol. 26, no. 3, pp. 616–26, Mar. 2012.

[41] T. T. Nguyen, H. Jung, and D. Y. Lee, “Markerless tracking for augmented reality for image-guided Endoscopic Retrograde Cholangiopancreatography.,” *Conf. Proc. IEEE Eng. Med. Biol. Soc.*, vol. 2013, pp. 7364–7, Jul. 2013.

[42] E. Marzano, T. Piardi, L. Soler, M. Diana, D. Mutter, J. Marescaux, and P. Pessaux, “Augmented reality-guided artery-first pancreatico-duodenectomy.,” *J. Gastrointest. Surg.*, vol. 17, no. 11, pp. 1980–3, Nov. 2013.

[43] D. Katić, A. L. Wekerle, J. Görtler, P. Spengler, S. Bodenstedt, S. R??hl, S. Suwelack, H. G. Kenngott, M. Wagner, B. P. M??ller-Stich, R. Dillmann, and S. Speidel, “Context-aware Augmented Reality in laparoscopic surgery,” *Comput. Med. Imaging Graph.*, vol. 37, no. 2, pp. 174–182, 2013.

[44] F. López-Mir, V. Naranjo, J. J. Fuertes, M. Alcañiz, J. Bueno, and E. Pareja, “Design and validation of an augmented reality system for laparoscopic surgery in a real environment.,” *Biomed Res. Int.*, vol. 2013, p. 758491, Jan. 2013.

[45] S. Onda, T. Okamoto, M. Kanehira, F. Suzuki, R. Ito, S. Fujioka, N. Suzuki, A. Hattori, and K. Yanaga, “Identification of inferior pancreaticoduodenal artery during pancreaticoduodenectomy using augmented reality-based navigation system,” *J. Hepatobiliary. Pancreat. Sci.*, vol. 21, no. 4, pp. 281–287, 2014.

[46] P. Pessaux, M. Diana, L. Soler, T. Piardi, D. Mutter, and J. Marescaux, “Robotic duodenopancreatectomy assisted with augmented reality and real-time fluorescence guidance,” *Surg. Endosc. Other Interv. Tech.*, vol. 28, no. 8, pp. 2493–2498, 2014.

[47] P. Pessaux, M. Diana, L. Soler, T. Piardi, D. Mutter, and J. Marescaux, “Towards cybernetic surgery: robotic and augmented reality-assisted liver segmentectomy,” *Langenbeck’s Arch. Surg.*, vol. 400, no. 3, pp. 381–385, 2014.

[48] T. Okamoto, S. Onda, K. Yanaga, N. Suzuki, and A. Hattori, “Clinical application of navigation surgery using augmented reality in the abdominal field,” *Surg. Today*, vol. 45, no. 4, pp. 397–406, 2015.

[49] T. Okamoto, S. Onda, J. Yasuda, K. Yanaga, N. Suzuki, and A. Hattori, “Navigation surgery using an augmented reality for pancreatectomy,” *Dig. Surg.*, vol. 32, no. 2, pp. 117–123, 2015.

[50] P. Edgcumbe, P. Pratt, G. Z. Yang, C. Nguan, and R. Rohling, “Pico Lantern: Surface reconstruction and augmented reality in laparoscopic surgery using a pick-up laser projector,” *Med. Image Anal.*, vol. 25, no. 1, pp. 95–102, 2015.

[51] J. Hallet, L. Soler, M. Diana, D. Mutter, T. F. Baumert, F. Habersetzer, J. Marescaux, and P. Pessaux, “Trans-thoracic minimally invasive liver resection guided by augmented reality,” *J. Am. Coll. Surg.*, vol. 220, no. 5, pp. e55–e60, 2015.

[52] N. Haouchine, S. Cotin, I. Peterlik, J. Dequidt, E. Kerrien, M. Berger, N. Haouchine, S. Cotin, I. Peterlik, J. Dequidt, M. Sanz-lopez, N. Haouchine, S. Cotin, I. Peterlik, J. Dequidt, M. S. Lopez, E. Kerrien, and M. Berger, “Impact of Soft Tissue Heterogeneity on Augmented Reality for Liver Surgery,” *IEEE*

Trans. Vis. Comput. Graph., vol. 21, no. 5, pp. 584–597, 2015.

[53] Y. Koreeda, Y. Kobayashi, S. Ieiri, Y. Nishio, K. Kawamura, S. Obata, R. Souzaki, M. Hashizume, and M. G. Fujie, "Virtually transparent surgical instruments in endoscopic surgery with augmentation of obscured regions.," *Int. J. Comput. Assist. Radiol. Surg.*, 2016.

[54] E. Wild, D. Teber, D. Schmid, T. Simpfendörfer, M. Müller, A.-C. Baranski, H. Kenngott, K. Kopka, and L. Maier-Hein, "Robust augmented reality guidance with fluorescent markers in laparoscopic surgery," *Int. J. Comput. Assist. Radiol. Surg.*, 2016.

[55] Y. Abe, S. Sato, K. Kato, T. Hyakumachi, Y. Yanagibashi, M. Ito, and K. Abumi, "A novel 3D guidance system using augmented reality for percutaneous vertebroplasty: technical note," *J. Neurosurg. Spine*, vol. 19, no. 4, pp. 492–501, 2013.

[56] J. R. Wu, M. L. Wang, K. C. Liu, M. H. Hu, and P. Y. Lee, "Real-time advanced spinal surgery via visible patient model and augmented reality system," *Comput. Methods Programs Biomed.*, vol. 113, no. 3, pp. 869–881, 2014.

[57] H. Wang, F. Wang, A. P. Y. Leong, L. Xu, X. Chen, and Q. Wang, "Precision insertion of percutaneous sacroiliac screws using a novel augmented reality-based navigation system: a pilot study," *Int. Orthop.*, 2015.

[58] R. Londei, M. Esposito, B. Diotte, S. Weidert, E. Euler, P. Thaller, N. Navab, and P. Fallavollita, "Intra-operative augmented reality in distal locking," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, no. 9, pp. 1395–1403, 2015.

[59] S. C. Lee, B. Fuerst, J. Fotouhi, M. Fischer, G. Osgood, and N. Navab, "Calibration of RGBD camera and cone-beam CT for 3D intra-operative mixed reality visualization," *Int. J. Comput. Assist. Radiol. Surg.*, pp. 1–9, 2016.

[60] D. Inoue, B. Cho, M. Mori, Y. Kikkawa, T. Amano, A. Nakamizo, K. Yoshimoto, M. Mizoguchi, M. Tomikawa, J. Hong, M. Hashizume, and T. Sasaki, "Preliminary study on the clinical application of augmented reality neuronavigation," *J Neurol Surg A Cent Eur Neurosurg*, vol. 74, no. 2, pp. 71–76, 2013.

[61] A. Meola, F. Cutolo, M. Carbone, F. Cagnazzo, M. Ferrari, and V. Ferrari, "Augmented reality in neurosurgery: a systematic review," *Neurosurg. Rev.*, 2016.

[62] I. Cabrilo, K. Schaller, and P. Bijlenga, "Augmented Reality-Assisted Bypass Surgery: Embracing Minimal Invasiveness," *World Neurosurg.*, vol. 83, no. 4, pp. 596–602, 2014.

[63] I. Cabrilo, P. Bijlenga, and K. Schaller, "Augmented reality in the surgery of cerebral arteriovenous malformations: Technique assessment and considerations," *Acta Neurochir. (Wien).*, vol. 156, no. 9, pp. 1769–1774, 2014.

[64] I. Cabrilo, P. Bijlenga, and K. Schaller, "Augmented reality in the surgery of cerebral aneurysms: A technical report," *Neurosurgery*, vol. 10, no. 2, pp. 252–261, 2014.

[65] G. P. Moustris, S. C. Hiridis, K. M. Deliparaschos, and K. M. Konstantinidis, "Evolution of autnomous and semi-autnomous robotic surgical systems: a review of the literature," *Int. J. Med. Robot.*, vol. 7, no. April, pp. 375–392, 2011.

[66] M. Müller, M.-C. Rassweiler, J. Klein, A. Seitel, M. Gondan, M. Baumhauer, D. Teber, J. J. Rassweiler, H.-P. Meinzer, and L. Maier-Hein, "Mobile augmented reality for computer-assisted percutaneous nephrolithotomy.," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 8, no. 4, pp. 663–75, Jul. 2013.

[67] T. Hakky, D. R. Martinez, L. I. Lipshultz, P. E. Spiess, and R. E. Carrion, "Mp23-11 Augmented Reality Assisted Urologic Surgery (Araus): a Surgical Training Tool," *J. Urol.*, vol. 193, no. 4, p. e271, 2015.

[68] A. Chowriappa, S. J. Raza, A. Fazili, E. Field, C. Malito, D. Samarasekera, Y. Shi, K. Ahmed, G. Wilding, J. Kaouk, D. D. Eun, A. Ghazi, J. O. Peabody, T. Kesavadas, J. L. Mohler, and K. A. Guru, "Augmented-reality-based skills training for robot-assisted urethrovesical anastomosis: A multi-institutional randomised controlled trial," *BJU Int.*, vol. 115, no. 2, pp. 336–345, 2015.

[69] A. J. Hung, S. H. Shah, L. Dalag, D. Shin, and I. S. Gill, "Development and Validation of a Novel Robotic Procedure Specific Simulation Platform: Partial Nephrectomy," *J. Urol.*, vol. 194, no. 2, pp. 520–526, 2015.

[70] G. Gandaglia, P. Schatteman, G. De Naeyer, F. D'Hondt, and A. Mottrie, "Novel Technologies in Urologic Surgery: a Rapidly Changing Scenario," *Curr. Urol. Rep.*, vol. 17, no. 3, 2016.

[71] Y. Jeon, S. Choi, and H. Kim, "Evaluation of a simplified augmented reality device for ultrasound-guided vascular access in a vascular phantom," *J. Clin. Anesth.*, vol. 26, no. 6, pp. 485–489, 2014.

[72] M. Gao, Y. Chen, Q. Liu, C. Huang, Z. Li, and D. Zhang, "Three-Dimensional Path Planning and Guidance of Leg Vascular Based on Improved Ant Colony Algorithm in Augmented Reality," *J. Med. Syst.*, vol. 39, no. 11, p. 133, 2015.

[73] D. G. Armstrong, T. M. Rankin, N. A. Giovinco, J. L. Mills, and Y. Matsuoka, "A heads-up display for diabetic limb salvage surgery: a view through the google looking glass," *J. Diabetes Sci. Technol.*, vol. 8, no. 5, pp. 951–956, 2014.

[74] I. Cheng, R. Shen, R. Moreau, V. Brizzi, N. Rossol, and A. Basu, "An augmented reality framework for optimization of computer assisted navigation in endovascular surgery," *Conf. Proc. ... Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.*, vol. 2014, pp. 5647–5650, 2014.

[75] R. Souzaki, S. Ieiri, M. Uemura, K. Ohuchida, M. Tomikawa, Y. Kinoshita, Y. Koga, A. Suminoe, K. Kohashi, Y. Oda, T. Hara, M. Hashizume, and T. Taguchi, "An augmented reality navigation system for pediatric oncologic surgery based on preoperative CT and MRI images," *J. Pediatr. Surg.*, vol. 48, no. 12, pp. 2479–2483, 2013.

[76] J. Marescaux and M. Diana, "Next step in minimally invasive surgery: Hybrid image-guided surgery," *J. Pediatr. Surg.*, vol. 50, no. 1, pp. 30–36, 2015.

[77] E. P. Ong, J. A. Lee, J. Cheng, H. Lee, G. Xu, A. Laude, S. Teoh, T. H. Lim, D. W. K. Wong, and J. Liu, "An Augmented Reality Assistance Platform for Eye Laser Surgery *," pp. 4326–4329, 2015.

[78] S. Nishimoto, M. Tonooka, K. Fujita, Y. Sotsuka, T. Fujiwara, K. Kawai, and M. Kakibuchi, "An augmented reality system in lymphatico-venous anastomosis surgery," *J. Surg. Case Reports*, vol. 2016, no. 5, p. rjw047, 2016.

[79] H. A. D. Ashab, V. A. Lessoway, S. Khallaghi, A. Cheng, R. Rohling, and P. Abolmaesumi, "An augmented reality system for epidural anesthesia (AREA): Prepuncture identification of vertebrae," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 9, pp. 2636–2644, 2013.

[80] W. P. Liu, M. Azizian, J. Sorger, R. H. Taylor, B. K. Reilly, K. Cleary, and D. Preciado, "Cadaveric feasibility study of da Vinci Si-assisted cochlear implant with augmented visual navigation for otologic surgery.," *JAMA Otolaryngol. Head Neck Surg.*, vol. 140, no. 3, pp. 208–14, 2014.

[81] W. P. Liu, J. D. Richmon, J. M. Sorger, M. Azizian, and R. H. Taylor, "Augmented reality and cone beam CT guidance for transoral robotic surgery," *J. Robot. Surg.*, vol. 9, no. 3, pp. 223–233, 2015.

[82] A. C. Profeta, C. Schilling, and M. McGurk, "Augmented reality visualization in head and neck surgery: An overview of recent findings in sentinel node biopsy and future perspectives," *Br. J. Oral Maxillofac. Surg.*, no. 2015, pp. 10–12, 2016.

[83] J. D'Agostino, J. Wall, L. Soler, M. Vix, Q.-Y. Duh, and J. Marescaux, "Virtual Neck Exploration for

Parathyroid Adenomas," *JAMA Surg.*, vol. 148, no. 3, p. 232, 2013.

[84] L. Li, J. Yang, Y. Chu, W. Wu, J. Xue, P. Liang, and L. Chen, "A Novel Augmented Reality Navigation System for Endoscopic Sinus and Skull Base Surgery: A Feasibility Study," *PLoS One*, vol. 11, no. 1, pp. 1–17, 2016.

[85] A. Citardi, M.J., Agbetoba, A., Bigcas, J.-L., Luong, "Augmented reality for endoscopic sinus surgery with surgical navigation: a cadaver study," *Int. Forum Allergy Rhinol.*, vol. 00, no. 0, pp. 1–6, 2016.

[86] R. a Mischkowski, M. J. Zinser, A. C. Kübler, B. Krug, U. Seifert, and J. E. Zöller, "Application of an augmented reality tool for maxillary positioning in orthognathic surgery - a feasibility study.," *J. Craniomaxillofac. Surg.*, vol. 34, no. 8, pp. 478–83, Dec. 2006.

[87] M. J. Zinser, R. A. Mischkowski, T. Dreiseidler, O. C. Thamm, D. Rothamel, and J. E. Z??ller, "Computer-assisted orthognathic surgery: Waferless maxillary positioning, versatility, and accuracy of an image-guided visualisation display," *Br. J. Oral Maxillofac. Surg.*, vol. 51, no. 8, pp. 827–833, 2013.

[88] H. Suenaga, H. Hoang Tran, H. Liao, K. Masamune, T. Dohi, K. Hoshi, Y. Mori, and T. Takato, "Real-time in situ three-dimensional integral videography and surgical navigation using augmented reality: a pilot study.," *Int. J. Oral Sci.*, May 2013.

[89] H. Suenaga, H. H. Tran, H. Liao, K. Masamune, T. Dohi, K. Hoshi, and T. Takato, "Vision-based markerless registration using stereo vision and an augmented reality surgical navigation system: a pilot study," *BMC Med. Imaging*, vol. 15, no. 1, p. 51, 2015.

[90] M. Qu, Y. Hou, Y. Xu, C. Shen, M. Zhu, L. Xie, H. Wang, Y. Zhang, and G. Chai, "Precise positioning of an intraoral distractor using augmented reality in patients with hemifacial microsomia," *J. Cranio-Maxillofacial Surg.*, vol. 43, no. 1, pp. 106–112, 2015.

[91] M. Zhu, G. Chai, L. Lin, Y. Xin, A. Tan, M. Bogari, Y. Zhang, and Q. Li, "Effectiveness of a Novel Augmented Reality-Based Navigation System in Treatment of Orbital Hypertelorism.," *Ann. Plast. Surg.*, vol. 00, no. 00, pp. 1–7, 2015.

[92] L. Lin, Y. Shi, A. Tan, M. Bogari, M. Zhu, Y. Xin, H. Xu, Y. Zhang, L. Xie, and G. Chai, "Mandibular angle split osteotomy based on a novel augmented reality navigation using specialized robot-assisted arms—A feasibility study," *J. Cranio-Maxillofacial Surg.*, vol. 44, no. 2, pp. 215–223, 2016.

[93] D. Katić, P. Spengler, S. Bodenstedt, G. Castrillon-Oberndorfer, R. Seeberger, J. Hoffmann, R. Dillmann, and S. Speidel, "A system for context-aware intraoperative augmented reality in dental implant surgery," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, no. 1, pp. 101–108, 2015.

[94] Y.-K. Lin, H.-T. Yau, I.-C. Wang, C. Zheng, and K.-H. Chung, "A Novel Dental Implant Guided Surgery Based on Integration of Surgical Template and Augmented Reality.," *Clin. Implant Dent. Relat. Res.*, Jul. 2013.

[95] J. Wang, H. Suenaga, K. Hoshi, L. Yang, E. Kobayashi, I. Sakuma, and H. Liao, "Augmented reality navigation with automatic marker-free image registration using 3-d image overlay for dental surgery," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 4, pp. 1295–1304, 2014.

[96] V. Ferrari, G. Megali, E. Troia, A. Pietrabissa, and F. Mosca, "A 3-D mixed-reality system for stereoscopic visualization of medical dataset.," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 11, pp. 2627–33, Nov. 2009.

[97] A. Moglia, V. Ferrari, L. Morelli, M. Ferrari, F. Mosca, and A. Cuschieri, "A Systematic Review of Virtual Reality Simulators for Robot-assisted Surgery," *Eur. Urol.*, 2015.

[98] V. Ferrari, R. M. Viglialoro, P. Nicoli, F. Cutolo, S. Condino, M. Carbone, M. Siesto, and M. Ferrari,

"Augmented reality visualization of deformable tubular structures for surgical simulation.," *Int. J. Med. Robot.*, vol. 12, no. 2, pp. 231–40, Jun. 2016.

[99] M. Kersten-Oertel, P. Jannin, and D. L. Collins, "The state of the art of visualization in mixed reality image guided surgery," *Comput. Med. Imaging Graph.*, vol. 37, no. 2, pp. 98–112, 2013.

[100] P. J. Edwards, D. J. Hawkes, D. L. Hill, D. Jewell, R. Spink, a Strong, and M. Gleeson, "Augmentation of reality using an operating microscope for otolaryngology and neurosurgical guidance.," *J. Image Guid. Surg.*, vol. 1, no. 3, pp. 172–178, 1995.

[101] P. J. Edwards, A. P. King, C. R. Maurer, D. A. De Cunha, D. J. Hawkes, D. L. G. Hill, R. P. Gaston, M. R. Fenlon, A. Jusczyzck, A. J. Strong, C. L. Chandler, and M. J. Gleeson, "Design and evaluation of a system for microscope-assisted guided interventions (MAGI)," *IEEE Trans. Med. Imaging*, vol. 19, no. 11, pp. 1082–1093, 2000.

[102] A. P. King, P. J. Edwards, C. R. Maurer Jr., D. A. de Cunha, D. J. Hawkes, D. L. Hill, R. P. Gaston, M. R. Fenlon, A. J. Strong, C. L. Chandler, A. Richards, and M. J. Gleeson, "A system for microscope-assisted guided interventions," *Stereotact Funct Neurosurg*, vol. 72, no. 2–4, pp. 107–111, 1999.

[103] W. E. L. Grimson, G. J. Ettinger, S. J. White, T. Lozano-Pérez, W. M. Wells, and R. Kikinis, "An automatic registration method for frameless stereotaxy, image guided surgery, and enhanced reality visualization," *IEEE Trans. Med. Imaging*, vol. 15, no. 2, pp. 129–140, 1996.

[104] M. J. Zinser, R. A. Mischkowski, T. Dreiseidler, O. C. Thamm, D. Rothamel, and J. E. Zöller, "Computer-assisted orthognathic surgery: waferless maxillary positioning, versatility, and accuracy of an image-guided visualisation display.," *Br. J. Oral Maxillofac. Surg.*, vol. 51, no. 8, pp. 827–33, Dec. 2013.

[105] S. Bernhardt, S. A. Nicolau, V. Agnus, L. Soler, C. Doignon, and J. Marescaux, "Automatic localization of endoscope in intraoperative CT image: A simple approach to augmented reality guidance in laparoscopic surgery," *Med. Image Anal.*, vol. 30, pp. 130–143, 2016.

[106] M. Blackwell, C. Nikou, a M. DiGioia, and T. Kanade, "An image overlay system for medical data visualization.," *Med. Image Anal.*, vol. 4, no. 1, pp. 67–72, 2000.

[107] G. Stetten, V. Chib, D. Hildebrand, and J. Bursee, "Real time tomographic reflection: Phantoms for calibration and biopsy," in *Proceedings - IEEE and ACM International Symposium on Augmented Reality, ISAR 2001*, 2001, pp. 11–19.

[108] H. Liao, N. Hata, S. Nakajima, M. Iwahara, I. Sakuma, and T. Dohi, "Surgical navigation by autostereoscopic image overlay of integral videography," *IEEE Trans. Inf. Technol. Biomed.*, vol. 8, no. 2, pp. 114–121, 2004.

[109] B. P. M. Yeung and P. W. Y. Chiu, "Application of robotics in gastrointestinal endoscopy: A review," *World J. Gastroenterol.*, vol. 22, no. 5, pp. 1811–1825, 2016.

[110] N. C. Buchs, F. Volonte, F. Pugin, C. Toso, M. Fusaglia, K. Gavaghan, P. E. Majno, M. Peterhans, S. Weber, and P. Morel, "Augmented environments for the targeting of hepatic lesions during image-guided robotic liver surgery," *J. Surg. Res.*, vol. 184, no. 2, pp. 825–831, 2013.

[111] J. P. Rolland and H. Fuchs, "Optical Versus Video See-Through Head-Mounted Displays in Medical Visualization," *Presence Teleoperators Virtual Environ.*, vol. 9, pp. 287–309, 2000.

[112] N. S. Holliman, N. A. Dodgson, G. E. Favalora, and L. Pockett, "Three-dimensional displays: A review and applications analysis," *IEEE Trans. Broadcast.*, vol. 57, no. 2 PART 2, pp. 362–371, 2011.

[113] H. Mukawa, K. Akutsu, I. Matsumura, S. Nakano, T. Yoshida, M. Kuwahara, K. Aiki, and M. Ogawa, "8.4: Distinguished Paper : A Full Color Eyewear Display using Holographic Planar Waveguides," *SID*

08 Dig., pp. 89–92, 2008.

[114] A. Plopski, Y. Itoh, C. Nitschke, K. Kiyokawa, G. Klinker, and H. Takemura, "Corneal-Imaging Calibration for Optical See-Through Head-Mounted Displays," *IEEE Trans. Vis. Comput. Graph.*, vol. 21, no. 4, pp. 481–490, 2015.

[115] F. Kellner, B. Bolte, G. Bruder, U. Rautenberg, F. Steinicke, M. Lappe, and R. Koch, "Geometric calibration of head-mounted displays and its effects on distance estimation," *IEEE Trans. Vis. Comput. Graph.*, vol. 18, no. 4, pp. 589–596, 2012.

[116] S. J. Gilson, A. W. Fitzgibbon, and A. Glennerster, "Spatial calibration of an optical see-through head-mounted display," *J. Neurosci. Methods*, vol. 173, no. 1, pp. 140–146, 2008.

[117] Y. Genc, M. Tuceryan, and N. Navab, "Practical solutions for calibration of optical see-through devices," in *Proceedings - International Symposium on Mixed and Augmented Reality, ISMAR 2002*, 2002, pp. 169–175.

[118] K. Abhari, J. S. H. Baxter, E. C. S. Chen, A. R. Khan, T. M. Peters, S. De Ribaupierre, and R. Eagleson, "Training for planning tumour resection: Augmented reality and human factors," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 6, pp. 1466–1477, 2015.

[119] N. Haouchine, J. Dequidt, M.-O. Berger, and S. Cotin, "Deformation-based augmented reality for hepatic surgery.," *Stud. Health Technol. Inform.*, vol. 184, pp. 182–8, Jan. 2013.

[120] A. M. Franz, T. Haidegger, W. Birkfellner, K. Cleary, T. M. Peters, and L. Maier-Hein, "Electromagnetic tracking in medicine -A review of technology, validation, and applications," *IEEE Trans. Med. Imaging*, vol. 33, no. 8, pp. 1702–1725, 2014.

[121] K. Hirokazu and M. Billinghurst, "Marker tracking and HMD calibration for a video-based augmented reality conferencing system," *Proc. 2nd IEEE ACM Int. Work. Augment. Real.*, pp. 85–94, 1999.

[122] G. Megali, V. Ferrari, C. Freschi, B. Morabito, F. Cavallo, G. Turini, E. Troia, C. Cappelli, A. Pietrabissa, O. Tonet, A. Cuschieri, P. Dario, and F. Mosca, "EndoCAS navigator platform: A common platform for computer and robotic assistance in minimally invasive surgery," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 4, no. 3, pp. 242–251, 2008.

[123] C. Ware, *Information visualization: perception for design*. 2004.

[124] F. Cutolo, G. Badiali, and V. Ferrari, "Human-PnP: Ergonomic AR Interaction Paradigm for Manual Placement of Rigid Bodies," in *Augmented Environments for Computer-Assisted Interventions*, vol. 1, 2015, pp. 104–113.

[125] T. Sielhorst, C. Bichlmeier, S. M. Heining, and N. Navab, "Depth perception--a major issue in medical AR: evaluation study by twenty surgeons," *Med Image Comput Comput Assist Interv*, vol. 9, no. Pt 1, pp. 364–372, 2006.

[126] C. Bichlmeier, F. Wimmer, S. M. Heining, and N. Navab, "Contextual anatomic mimesis: Hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality," in *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR*, 2007.

[127] N. Haouchine, J. Dequidt, M. O. Berger, and S. Cotin, "Single view augmentation of 3D elastic objects," in *ISMAR 2014 - IEEE International Symposium on Mixed and Augmented Reality - Science and Technology 2014, Proceedings*, 2014, pp. 229–236.

[128] M. Kersten-Oertel, S. J. S. Chen, and D. L. Collins, "An evaluation of depth enhancing perceptual cues for vascular volume visualization in neurosurgery," *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 3, pp. 391–403, 2014.

[129] C. Bichlmeier, S. M. Heining, M. Feuerstein, and N. Navab, “The Virtual Mirror: A New Interaction Paradigm for Augmented Reality Environments,” *IEEE Trans. Med. Imaging*, vol. 28, no. 9, pp. 1498–1510, 2009.

[130] C. Bichlmeier, E. Euler, T. Blum, and N. Navab, “Evaluation of the virtual mirror as a navigational aid for augmented reality driven minimally invasive procedures,” in *9th IEEE International Symposium on Mixed and Augmented Reality 2010: Science and Technology, ISMAR 2010 - Proceedings*, 2010, pp. 91–97.

[131] M. a. Fischler and R. C. Bolles, “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography,” *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[132] B. K. P. Horn, “Closed-form solution of absolute orientation using unit quaternions,” *J. Opt. Soc. Am. A*, vol. 4, no. 4, p. 629, 1987.

[133] K. S. Arun, T. S. Huang, and S. D. Blostein, “Least-Squares Fitting of Two 3-D Point Sets.,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, no. 5, pp. 698–700, 1987.

[134] Y. Wu, “PnP problem revisited,” *J. Math. Imaging Vis.*, vol. 24, no. 1, pp. 131–141, 2006.

[135] B. M. Haralick, C. N. Lee, K. Ottenberg, and M. Nölle, “Review and analysis of solutions of the three point perspective pose estimation problem,” *Int. J. Comput. Vis.*, vol. 13, no. 3, pp. 331–356, 1994.

[136] L. Quan and Z. Lan, “Linear N-point camera pose determination,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 8, pp. 774–780, 1999.

[137] P. D. Fiore, “Efficient linear solution of exterior orientation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 140–148, 2001.

[138] X. S. Gao, X. R. Hou, J. Tang, and H. F. Cheng, “Complete solution classification for the perspective-three-point problem,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 8, pp. 930–943, 2003.

[139] A. Ansar and K. Daniilidis, “Linear pose estimation from points or lines,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 5, pp. 578–589, 2003.

[140] V. Lepetit, F. Moreno-Noguer, and P. Fua, “EPnP: An accurate O(n) solution to the PnP problem,” *Int. J. Comput. Vis.*, vol. 81, no. 2, pp. 155–166, 2009.

[141] J. A. Hesch and S. I. Roumeliotis, “A Direct Least-Squares (DLS) method for PnP,” in *Proceedings of the IEEE International Conference on Computer Vision*, 2011, pp. 383–390.

[142] R. M. Haralick, H. Joo, C. N. Lee, X. Zhuang, V. G. Vaidya, and M. Bae Kim, “Pose Estimation from Corresponding Point Data,” *IEEE Trans. Syst. Man Cybern.*, vol. 19, no. 6, pp. 1426–1446, 1989.

[143] D. G. Lowe, “Fitting Parameterized Three-Dimensional Models to Images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 5, pp. 441–450, 1991.

[144] C. P. Lu, G. D. Hager, and E. Mjolsness, “Fast and globally convergent pose estimation from video images,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 6, pp. 610–622, 2000.

[145] Y. Guo, “A note on the number of solutions of the coplanar P4P problem,” in *2012 12th International Conference on Control, Automation, Robotics and Vision, ICARCV 2012*, 2012, pp. 1413–1418.

[146] T. Wang, Y. Wang, and C. Yao, “Some discussion on the conditions of the unique solution of P3P problem,” in *2006 IEEE International Conference on Mechatronics and Automation, ICMA 2006*, 2006, vol. 2006, pp. 205–210.

[147] Z. Zhang, “A flexible new technique for camera calibration,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 11, pp. 1330–1334, 2000.

[148] S. Mazzoni, G. Badiali, L. Lancellotti, L. Babbi, A. Bianchi, and C. Marchetti, “Simulation-guided navigation: a new approach to improve intraoperative three-dimensional reproducibility during orthognathic surgery.,” *J. Craniofac. Surg.*, vol. 21, no. 6, pp. 1698–1705, 2010.

[149] S. Mazzoni, A. Bianchi, G. Schiariti, G. Badiali, and C. Marchetti, “Computer-aided design and computer-aided manufacturing cutting guides and customized titanium plates are useful in upper maxilla waferless repositioning,” *J. Oral Maxillofac. Surg.*, vol. 73, no. 4, pp. 701–707, 2015.

[150] M. Zinser, H. Sailer, and L. Ritter, “A Paradigm Shift in Orthognathic Surgery? A Comparison of Navigation, Cad/Cam Splints and ‘Classic’ Intermaxillary Splints to Surgical Transfer of Virtual,” *… Oral Maxillofac. …*, 2013.

[151] G. Turchetti, E. Spadoni, and E. E. Geisler, “Health technology assessment. Evaluation of biomedical innovative technologies.,” *IEEE Eng. Med. Biol. Mag.*, vol. 29, no. 3, pp. 70–6, Jan. 2010.

[152] G. Badiali, V. Ferrari, F. Cutolo, C. Freschi, D. Caramella, A. Bianchi, and C. Marchetti, “Augmented reality as an aid in maxillofacial surgery: validation of a wearable system allowing maxillary repositioning.,” *J. Craniomaxillofac. Surg.*, vol. 42, no. 8, pp. 1970–6, Dec. 2014.

[153] J. C. Posnick, *Orthognathic Surgery: Principles and Practice*. Elsevier, 2015.

Human-P*n*P: Ergonomic AR Interaction Paradigm for Manual Placement of Rigid Bodies

Fabrizio Cutolo[1(✉)], Giovanni Badiali[3], and Vincenzo Ferrari[1,2]

[1] Department of Translational Research and New Technologies in Medicine and Surgery, EndoCAS Center, University of Pisa, Pisa, Italy
fabrizio.cutolo@endocas.org
[2] Information Engineering Department, University of Pisa, Pisa, Italy
[3] School in Surgical Sciences, University of Bologna, Bologna, Italy

Abstract. The human perception of the three-dimensional world is influenced by the mutual integration of physiological and psychological depth cues, whose complexity is still an unresolved issue per se. Even more so if we wish to mimic the perceptive efficiency of the human visual system within augmented reality (AR) based surgical navigation systems. In this work we present a novel and ergonomic AR interaction paradigm that aids the manual placement of a non-tracked rigid body in space by manually minimizing the reprojection residuals between a set of corresponding virtual and real feature points. Our paradigm draws its inspiration from the general problem of estimating camera pose from a set of n-correspondences, i.e. perspective-*n*-point problem. In a recent work, positive results were achieved in terms of geometric error by applying the proposed strategy on the validation of a wearable AR system to aid manual maxillary repositioning.

Keywords: Augmented reality and visualization · Computer assisted intervention · Interventional imaging

1 Introduction

In the context of image-guided surgery (IGS), augmented reality (AR) technology represents a promising integration between navigational surgery and virtual planning.

In 2012 Kersten-Oertel et al. [1] proposed a taxonomy of mixed reality visualization systems in IGS and defined the three major components based on which they then presented a systematic overview of the trends and solutions adopted in the field [2]. The acronym for the taxonomy (DVV) derives from its three key components: Data type, Visualization Processing and View. According to the taxonomy, for classifying and assessing the efficacy of a new AR system for IGS, we must focus our attention on the particular surgical scenario in which the visualization system aim to be integrated. The surgical scenario affects each of the three DVV factors, namely the type of data that should be displayed at a specific surgical step, the visualization processing technique implemented to provide the best pictorial representation of the augmented scene

C.A. Linte et al. (Eds.): AE-CAI 2015, LNCS 9365, pp. 50–60, 2015.
DOI: 10.1007/978-3-319-24601-7_6

and how and where the output of the visualization processing should be presented to the end-user.

Several visualization processing techniques have been adopted to allow a more immersive viewing experience for the surgeon and a more precise definition of the spatial relationships between real scene and visually processed data along the three dimensions. The human visual system exploits several physiological and psychological cues to deal with the ill-posed inverse problem of understanding a three-dimensional scene from one retinal image. However, monocular and binocular cues are not always sufficient to infer the spatial relationships between objects in the three-dimensional scene. Therefore, a full comprehension of the mechanisms underpinning depth perception is not a completely resolved issue per se in a real scene and it results even more complex within an augmented scene [3]. In this regard, among the suggested visualization processing techniques, researchers have tried to improve the perceptive efficiency by modeling and contextually rendering the virtual content in a photo-realistic manner, and/or by using pixel-wise transparency maps and "virtual windows" [4] to recreate occlusions and motion parallax cues. Some of the proposed techniques for enhancing depth perception comprise high-fidelity texturing [5] or colour coding methods, whereas others consist in lighting and shading cues and/or on the adoption of an interactive "virtual mirror" [6, 7]. Alternatively, depth perception can be improved by relying on standard stereopsis and two-view displays or on more complex full parallax multi-view displays. In any case, to the best of our knowledge, hitherto there are no visualization processing techniques that provide the user with useful information able to improve the postoperative outcome for those specific surgical tasks that involve the accurate manual placement of rigid anatomies in space.

Many surgical procedures in the field of orthopedic surgery or maxillofacial surgery, involve the task of reducing displacements or correcting abnormalities between rigid anatomical structures, i.e. bones, on the basis of a pre-operative planning. The direct tracking of all the rigid anatomies involved in the procedure would yield a measure of the six-degrees-of-freedom displacements between them and it would aid the correct performance of the surgical task, yet it is not always feasible for technical and logistic reasons. In case of single object tracking, the pointer of a standard surgical navigator can be employed by the surgeon to compare the final positions of clearly detectable reference points, over the repositioned anatomy, with those of their counterparts from the surgical planning. Nevertheless, this approach does not allow the assessment of all of the six-degrees-of-freedom at the same time.

AR seems the optimal solution to aid this kind of surgical tasks. Yet, the traditional AR interaction technique featuring the superimposition of a semi-transparent virtual replica of the rigid anatomy in a position and orientation (pose) defined during planning, is not very effective in aiding the surgeon in the correct performance of those procedure. In this regard, it is more beneficial and intuitive for the surgeon to deal with task-oriented visualization techniques, more than with complex reproductions of the virtual anatomies through photorealistic rendering, transparencies and/or virtual windows.

The goal of this work is to present a novel and ergonomic AR interaction paradigm based on a simple visualization processing technique that aims at aiding the accurate manual placement of a non-tracked rigid object in space. Our strategy relies on the tracking of a single object in the scene (e.g. the patient's head), namely on the real-time

estimation of the geometric relation between a scene reference system (SRS) and the camera reference system (CRS), e.g. performed by means of a video based registration approach. In this scenario, the AR guide aids the surgeon in placing other non-tracked rigid bodies (e.g. bones fragments) at a planned pose relative to the CRS. Our paradigm draws its inspiration from the general problem of estimating camera pose from a set of n-correspondences, i.e. perspective-n-point problem. The key idea is that manually minimizing the distance, in the image plane, between a set of corresponding real and virtual feature points is sufficient to aid the accurate placement of a non-tracked rigid body in space.

2 Methods

Perspective-*n*-Point Problem. The task of estimating the pose of a camera with respect to a scene object given its intrinsic parameters and a set of n world-to-image point correspondences is known as the Perspective-*n*-Point (P*n*P) problem in computer vision and exterior orientation or space resection problem in photogrammetry.

This inverse problem concerns many fields of applications (structure from motion, robotics, augmented reality, etc.) and it was first formally introduced in the computer vision community by Fishler and Bolles in 1981 [8], albeit already used in the photogrammetry community before then. According to Fishler and Bolles the P*n*P problem can be defined as follows (distance-based definition):

> *Given the relative spatial locations of n control points* $\boldsymbol{P_i}, \boldsymbol{i} = 1, \ldots n$*, and given the angle to every pair of these points from an additional point called the center of perspective* $\boldsymbol{C}$*, find the lengths* $D_i = |\boldsymbol{CP_i}|$ *of the line segments joining* $\boldsymbol{C}$ *to each of the control points.*

The constraint equations are:

$$D_i^2 + D_j^2 - 2D_iD_j \cos\theta_{ij} = d_{ij}^2, i \neq j \tag{1}$$

Where $D_i = |\boldsymbol{CP_i}|$, $D_j = \left|\boldsymbol{CP_j}\right|$ are the unknown variables, $\theta_{ij} = \widehat{\boldsymbol{P_iCP_j}}$ and $d_{ij} = \left|\boldsymbol{P_iP_j}\right|$ are the known entries (Fig. 1). In computer vision θ_{ij} are determined finding the correspondences between world-to-image points and knowing the intrinsic camera parameters, while d_{ij} are established by the control points.

Following this definition, once each distance D_i is computed, the position of the points $\boldsymbol{P_i}$ can be expressed in the CRS. Therefore, being the position of each $\boldsymbol{P_i}$ in the SRS known, the problem of estimating camera pose with respect to the SRS is reduced to a standard absolute orientation problem whose solution can be found in closed-form fashion through quaternions [9] or singular value decomposition (SVD) [10].

The same problem is also known under the transformation-based definition [11] which can be formalized as:

$$\boldsymbol{\lambda_i \widehat{p}_i} = [\mathrm{K}|0] \begin{bmatrix} \mathrm{R} & \mathrm{T} \\ 0 & 1 \end{bmatrix} \widehat{\boldsymbol{P}}_i, i = 1, \ldots n \tag{2}$$

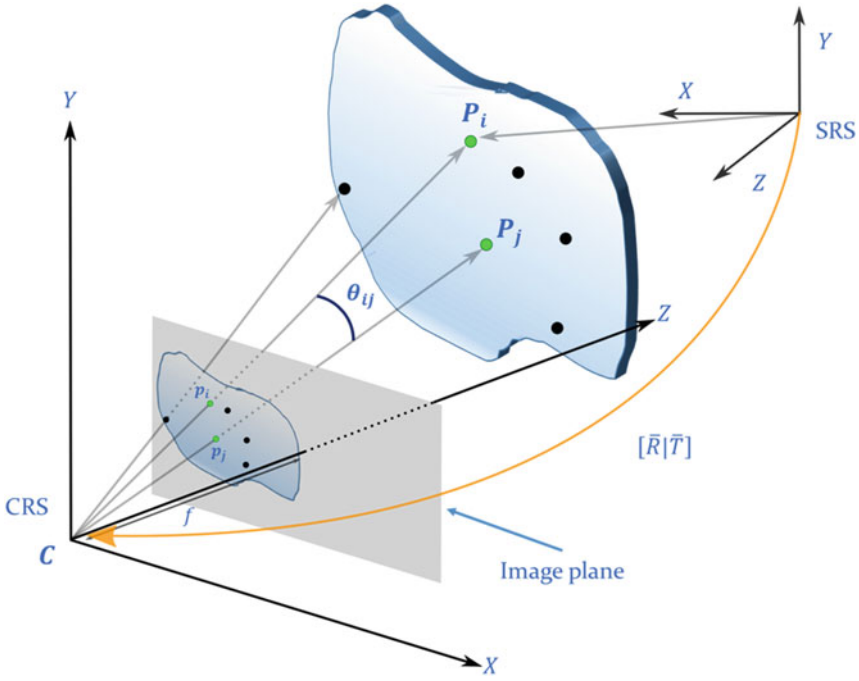


Fig. 1. Geometry of the P*n*P problem.

Where the scene and image points $\hat{\boldsymbol{P}}_{\boldsymbol{i}}$ and $\hat{\boldsymbol{p}}_{\boldsymbol{i}}$ are represented in homogeneous coordinates and the equation is up to a scale factor $\boldsymbol{\lambda_i}$. Hence, according to this definition, the P*n*P problem aims at determining the pose (in terms of a rotation matrix R and a translation vector T) given a set of n world-to-image correspondences and known the intrinsic camera parameters encapsulated by the matrix K.

The P*n*P problem has been extensively studied by several groups, which have proposed different iterative, closed-form for solving it.

Closed-form methods [12–18], directly provide an estimation of the camera pose but they are usually less accurate and more susceptible to noise than iterative methods. Iterative non-linear optimization methods solve the P*n*P problem by iteratively minimizing a cost function generally related to the geometric (reprojection residuals) or algebraic error but they need a good initial guess and yield only one solution at a time [19–21]. A useful overview of the state-of-the-art methods can be found in [17] and in [22].

In terms of geometric reprojection residual, the non-linear cost function can be formulated as the sum of the squared measurement errors (d_i):

$$\begin{aligned} \bar{\mathrm{R}}|\bar{\mathrm{T}} &= \arg\min \sum_{i=1}^{n} d(\boldsymbol{p_i}, \hat{\boldsymbol{p}}_i)^2 \\ &= \arg\min \sum_{i=1}^{n} \left\| \boldsymbol{p_i} - \hat{\boldsymbol{p}}_i(\mathrm{K}, \hat{\mathrm{R}}, \hat{\mathrm{T}}, \boldsymbol{P_i}) \right\|^2 \end{aligned} \tag{3}$$

Where $\boldsymbol{p_i}$ are the measured image points, and $\hat{\boldsymbol{p}}_{\boldsymbol{i}}$ are the calculated projections of the corresponding control points as a function of $\mathrm{K}, \hat{\mathrm{R}}, \hat{\mathrm{T}}$.

The other important research direction on the P*n*P problem is the study of the multi-solution phenomenon of the P*n*P problem [23], principally when $n = 3$ (P3P) [24, 25], being three the smallest subset of control points that yields a finite number of solutions. P3P problem yields at most four solutions which can be disambiguated using a fourth point, and it is the most studied case since it can be used as first step to reduce the complexity of the computation of a P*n*P problem, e.g. in a RANSAC scheme by removing the outliers.

AR Video-Based Camera Registration. Regardless of the method adopted for solving the P*n*P problem, an immediate application of the P*n*P problem is to locate the pose of a calibrated camera with respect to an object, given the 3D position of a set of n control points rigidly constrained to the object and the 2D position of their correspondent projections onto the image plane.

For a correct registration of computer-generated elements to the real scene in AR-based surgical navigation systems, the image formation process of the virtual camera must perfectly mimic the real camera one. In mostly all the AR applications the estimation of the intrinsic camera parameters is the result of an off-line calibration process whereas the extrinsic camera parameters are determined online, e.g. solving a P*n*P problem in real-time. This video-based camera registration method suggested us the implementation of an ergonomic AR interaction paradigm for positioning and orienting a non-tracked rigid object in space.

Human-P*n*P. As written in the introduction, many surgical procedures in the field of orthopedic surgery or maxillofacial surgery, involve the task of manually placing rigid anatomies on the basis of preoperative planning. In that case, let us assume that we can rely on a robust and accurate registration of the surgical planning onto the real scene, by means of the tracking of at least one rigid body (e.g. the head). The six-degrees-of-freedom pose of an additional and non-tracked rigid anatomy in relation to the SRS, can be retrieved by physically placing it as to minimize the geometric distance, on the image plane, between a set of real and virtual feature points. For brevity, from now on, we shall refer to these structures as "tracked anatomy" for the former and "non-tracked anatomy" for the latter, while the proposed method will be referred to as the human-perspective-*n*-point problem (hP*n*P).

From a theoretical standpoint, our method draws its inspiration and physically mimics the paradigm on which the P*n*P problem is formulated. As mentioned in the previous section, the main goal of the P*n*P problem is to infer useful information on the real 3D scene, based on 2D observations of it. In an AR application, this spatial information is used to geometrically register the virtual elements onto the real scene. Thus, as a general rule and regardless of the method adopted for solving the P*n*P problem, a robust and accurate registration should minimize in the image plane the geometric reprojection residuals between measured and estimated projections (see Eq. 3). Similarly, the goal of our hP*n*P interaction paradigm is to achieve the desired placement of a non-tracked anatomy by manually minimizing the reprojection residuals

between correct/planned projections $\bar{\boldsymbol{p}}_{\boldsymbol{i}}$ of virtual landmarks, and observed projections $\hat{\bar{\boldsymbol{p}}}_{\boldsymbol{i}}$ of real landmarks.

The correct/planned projections $\bar{\boldsymbol{p}}_{\boldsymbol{i}}$ are rendered on the image plane according to the real-time estimation of the camera pose $[\bar{\mathrm{R}}, \bar{\mathrm{T}}]$ relative to the tracked anatomy reference system (SRS) and assuming the intrinsic camera parameters, encapsulated by matrix K, are determined offline, e.g. through the Zhang's method [26]. The position of each virtual landmark $\boldsymbol{P}_{\boldsymbol{i}}$ in the SRS is established during surgical planning.

The real projections $\hat{\bar{\boldsymbol{p}}}_{\boldsymbol{i}}$ are associated with the pose, encapsulated by $[\hat{\bar{\mathrm{R}}}, \hat{\bar{\mathrm{T}}}]$, between viewing camera and non-tracked anatomy reference frame (ARS): this resulting pose varies according to the manual placement of the rigid body relative to the camera:

$$\begin{aligned} \tilde{\mathrm{R}}|\tilde{\mathrm{T}} &= \arg\min \sum_{i=1}^{n} d\left(\bar{\boldsymbol{p}}_{\boldsymbol{i}}, \hat{\bar{\boldsymbol{p}}}_{\boldsymbol{i}}\right)^2 \\ &= \arg\min \sum_{i=1}^{n} \left\| \bar{\boldsymbol{p}}_{\boldsymbol{i}}(\mathrm{K}, \bar{\mathrm{R}}, \bar{\mathrm{T}}, \boldsymbol{P}_{\boldsymbol{i}}) - \hat{\bar{\boldsymbol{p}}}_{\boldsymbol{i}}(\mathrm{K}, \hat{\bar{\mathrm{R}}}, \hat{\bar{\mathrm{T}}}, \boldsymbol{P}_{\boldsymbol{i}}) \right\|^2 \end{aligned} \tag{4}$$

In this way, we wish to obtain $\left[\tilde{\mathrm{R}}|\tilde{\mathrm{T}}\right] \approx [\bar{\mathrm{R}}|\bar{\mathrm{T}}]$ (see Fig. 2), namely we seek to positioning and orienting the ARS as coincident with the planned and registered SRS (non-tracked anatomy reference frame ≈ planning reference frame).

To implement this strategy, we add simple virtual elements (e.g. virtual asterisks, crosses, etc.) to the virtual scene during the surgical planning: one element for each of the clearly detectable physical landmarks on the rigid body. The landmarks may consist of a series of distinguishable feature points over the surface of the anatomy or rigidly

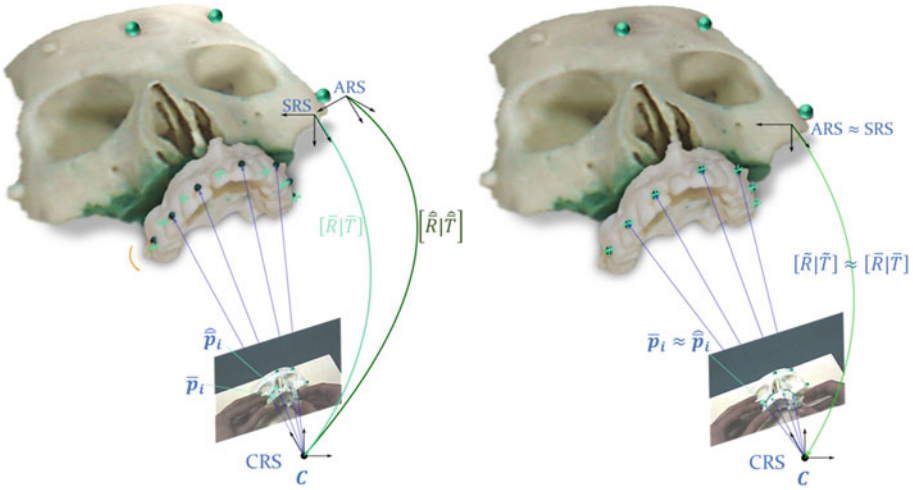


Fig. 2. Geometry of the hP*n*P: minimizing the reprojection residual between registered projections $\bar{p}_i$ and real projections $\hat{\bar{p}}$ is sufficient to aid the accurate placement of a rigid body (the maxilla in the image) in space.

constrained to it. Under such AR guidance, the user moves the non-tracked rigid body up to obtain a perfect overlapping between real and virtual landmarks, hence manually minimizing the reprojection residuals on the image plane: $\bar{\boldsymbol{p}}_{\boldsymbol{i}} \approx \hat{\bar{\boldsymbol{p}}}_{\boldsymbol{i}} \forall \boldsymbol{i}$ (Fig. 3). The theoretical assumptions underpinning the P*n*P problem ensure that if $\bar{\boldsymbol{p}}_{\boldsymbol{i}} \approx \hat{\bar{\boldsymbol{p}}}_{\boldsymbol{i}} \forall \boldsymbol{i}$, the non-tracked anatomy is placed in the correct pose as planned in SRS.

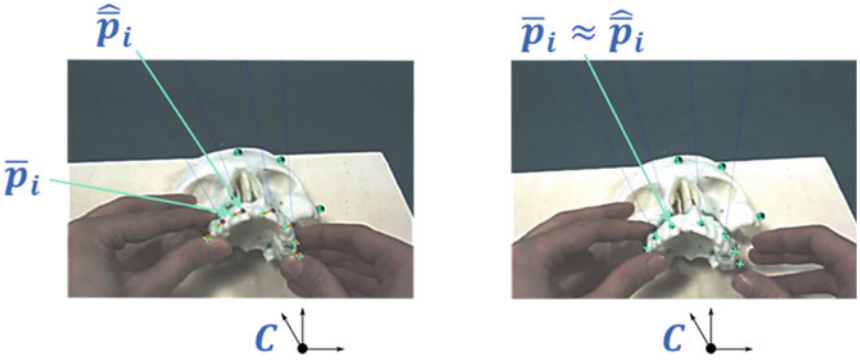


Fig. 3. Detail of the image plane with the minimization of the reprojection residuals. Here the virtual information consists of a cyan-colored asterisk for each physical landmark clearly detectable over the maxilla, e.g. coloured landmarks fixed on the brackets of the orthodontic appliance (Color figure online).

3 Results

In a recent work [27], the described strategy was applied in the validation of a wearable AR system to aid maxillary repositioning. AR system consisted of a stereoscopic video see-through head mounted display equipped with two external USB cameras placed in a quasi-orthoscopic position [28, 29]. The video see-through paradigm of the system is implemented as follows (Fig. 4): real-world views are grabbed by a pair of calibrated external cameras; the captured frames, after compensation of the radial distortion, are screened as backgrounds of the virtual scene onto the corresponding display; the virtual elements, defined during planning, are added to the real scene and observed by a pair of virtual cameras whose processes of image formation mimic those of the real cameras in terms of intrinsic and extrinsic camera parameters. Zhang's method is used to calibrate the two cameras. The estimation of the extrinsic parameters, allowing the real-time registration of the virtual elements to real scene, is achieved through a marker-based video-registration method [29].

In the study, manual repositioning of the upper maxilla following LeFort 1 osteotomy was chosen as test procedure. The test was conducted on a CT-scanned/3D-printed replica of a cadaveric human skull. The planned pose of the maxilla, as defined during preoperative planning, acts as a guide for the surgeon during the intervention performed in-vitro. The traditional AR interaction technique, featuring the superimposition of a semi-transparent virtual replica of the maxilla, as dictated by the surgical planning, did

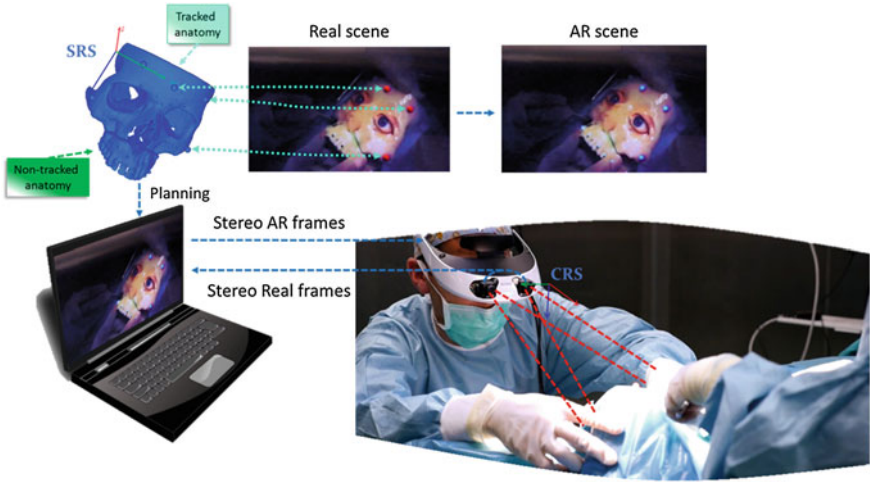


Fig. 4. Video see-through paradigm of the stereoscopic head mounted display used to aid maxillary repositioning.

not prove to be very effective in aiding the surgeon in manually repositioning the upper maxilla. This was mostly due to the surgeon's limited perception of the relative distances of objects within the AR scene owing to the presence of unnatural occlusions between the real and the virtual maxilla. Conversely, a more ergonomic form of visualization consisted in the use of an interaction paradigm which actualizes the above described hP*n*P approach: physical landmarks onto the maxilla and corresponding to coloured landmarks fixed on the brackets of the orthodontic appliance usually applied prior to this kind of interventions, were designated as reference markers for the AR view modality. The repositioning of the maxilla is assisted by visually aligning small virtual asterisks, drawn in positions defined during planning (relative to the SRS), with the corresponding real landmarks.

The upper surface of the maxilla (corresponding to the post-osteotomy surface) was covered with highly malleable plasticine so to be fixed to the upper skull once the surgeon performed the repositioning. The surgical accuracy was validated with the aid of an optical navigation system that recorded the coordinates of three reference points on the non-tracked maxilla after repositioning. Six surgeons and three unskilled engineers were involved in the testing, each of whom was asked to manually reposition the maxilla as dictated by three surgical plannings of variable complexity. Results in terms of linear distances between the real positions of the reference holes and the expected positions (defined during planning) were very promising: mean error was 1.70 ± 0.51 mm. The axial errors were 0.89 ± 0.54 mm on the sagittal axis, 0.60 ± 0.20 mm on the frontal axis, and 1.06 ± 0.40 mm on the cranio-caudal axis. Such results were obtained without the tracking of the maxilla but just relying on the ergonomics of the chosen AR interaction paradigm: the overlapping on the image plane between virtual feature points and real landmarks, visible over the non-tracked anatomy, proved to be sufficient to aid the accurate repositioning of the maxilla.

4 Discussion

It is important to note that the chosen AR interaction paradigm was not bound to the particular video-based tracking modality exploited in the cited study, neither to the use of a specific wearable stereoscopic system. Howbeit, the user can enhance the accuracy in object placement by checking consistency of real and virtual landmarks from different viewpoints. In this regard, the ergonomics of the proposed method may benefit from the adoption of a wearable AR system. Moreover, the choice of such instance of visualization data was, in that work, empirically inspired by the authors' endeavor of defining a modality that were ergonomic for the surgeon and that provided the smallest perceived parallax error: no further discussion was held on the theoretical hypotheses behind such interaction paradigm which are here discussed for the first time.

5 Conclusion

In this work, we proposed a novel and ergonomic AR interaction paradigm that aims at obtaining the accurate placement of a rigid body in space without the need for multiple objects tracking and/or complex visual representations of the virtual guide. From a theoretical standpoint, our method draws its inspiration and physically mimics the paradigm on which the P*n*P problem in computer vision is formulated. This approach, represented by the acronym hP*n*P, could be of help in those tasks, also not specifically surgical, where the AR guide aims at aiding the placement of a rigid body in space. The key-principle behind this interaction paradigm can be exploited in many different AR-based navigation systems: it can be integrated with different end-products of the visualization process in terms of display technology and perception location and/or it could be realized in conjunction with various tracking modalities.

To increase robustness and applicability of the proposed AR interaction paradigm in a real clinical scenario, redundancy in choosing the set of landmarks must be granted. Further, the presence of line-of-sight occlusions caused by soft-tissues, surgeon's hands or surgical instrumentation may be restricted by conveniently selecting the position of the landmarks in relation to the surgical field.

Acknowledgments. This work was funded by the Italian Ministry of Health grant SThARS (Surgical training in identification and isolation of deformable tubular structures with hybrid Augmented Reality Simulation, 6/11/2014–5/11/2017). Grant "Ricerca finalizzata e Giovani Ricercatori 2011–2012" Young Researchers – Italian Ministry of Health.

References

1. Kersten-Oertel, M., Jannin, P., Collins, D.L.: DVV: towards a taxonomy for mixed reality visualization in image guided surgery. In: Liao, H., Edwards, P.J., Pan, X., Fan, Y., Yang, G.-Z. (eds.) MIAR 2010. LNCS, vol. 6326, pp. 334–343. Springer, Heidelberg (2010)
2. Kersten-Oertel, M., Jannin, P., Collins, D.L.: The state of the art of visualization in mixed reality image guided surgery. Comput. Med. Imaging Graph. **37**, 98–112 (2013)

3. Sielhorst, T., Bichlmeier, C., Heining, S.M., Navab, N.: Depth perception - a major issue in medical AR: evaluation study by twenty surgeons. In: Larsen, R., Nielsen, M., Sporring, J. (eds.) MICCAI 2006. LNCS, vol. 4190, pp. 364–372. Springer, Heidelberg (2006)
4. Bichlmeier, C., Wimme, F., Heining, S.M., Navab, N.: Contextual anatomic mimesis hybrid in-situ visualization method for improving multi-sensory depth perception in medical augmented reality. In: 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, ISMAR 2007, pp. 129–138 (2007)
5. Haouchine, N., Dequidt, J., Berger, M.O., Cotin, S.: Single view augmentation of 3D elastic objects. In: 2014 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 229–236 (2014)
6. Bichlmeier, C., Heining, S.M., Feuerstein, M., Navab, N.: The virtual mirror: a new interaction paradigm for augmented reality environments. IEEE Trans. Med. Imaging **28**, 1498–1510 (2009)
7. Bichlmeier, C., Euler, E., Blum, T., Navab, N.: Evaluation of the virtual mirror as a navigational aid for augmented reality driven minimally invasive procedures. In: 2010 9th IEEE International Symposium on Mixed and Augmented Reality (ISMAR), pp. 91–97 (2010)
8. Fischler, M.A., Bolles, R.C.: Random sample consensus - a paradigm for model-fitting with applications to image-analysis and automated cartography. Commun. ACM **24**, 381–395 (1981)
9. Horn, B.K.P.: Closed-form solution of absolute orientation using unit quaternions. J. Opt. Soc. Am. A-Opt. Image Sci. Vis. **4**, 629–642 (1987)
10. Arun, K.S., Huang, T.S., Blostein, S.D.: Least-squares fitting of 2 3-D point sets. IEEE Trans. Pattern Anal. Mach. Intell. **9**, 699–700 (1987)
11. Wu, Y.H., Hu, Z.Y.: PnP problem revisited. J. Math. Imaging Vis. **24**, 131–141 (2006)
12. Haralick, R.M., Lee, C.N., Ottenberg, K., Nolle, M.: Review and analysis of solutions of the 3-point perspective pose estimation problem. Int. J. Comput. Vis. **13**, 331–356 (1994)
13. Quan, L., Lan, Z.D.: Linear N-point camera pose determination. IEEE Trans. Pattern Anal. Mach. Intell. **21**, 774–780 (1999)
14. Fiore, P.D.: Efficient linear solution of exterior orientation. IEEE Trans. Pattern Anal. Mach. Intell. **23**, 140–148 (2001)
15. Gao, X.S., Hou, X.R., Tang, J.L., Cheng, H.F.: Complete solution classification for the Perspective-Three-Point problem. IEEE Trans. Pattern Anal. Mach. Intell. **25**, 930–943 (2003)
16. Ansar, A., Daniilidis, K.: Linear pose estimation from points or lines. IEEE Trans. Pattern Anal. Mach. Intell. **25**, 578–589 (2003)
17. Lepetit, V., Moreno-Noguer, F., Fua, P.: EPnP: an accurate O(n) solution to the PnP problem. Int. J. Comput. Vis. **81**, 155–166 (2009)
18. Hesch, J.A., Roumeliotis, S.I.: A direct least-squares (DLS) method for PnP. In: 2011 IEEE International Conference on Computer Vision (ICCV), pp. 383–390 (2011)
19. Haralick, R.M., Joo, H., Lee, C.N., Zhuang, X.H., Vaidya, V.G., Kim, M.B.: Pose estimation from corresponding point data. IEEE Trans. Syst. Man Cybern. **19**, 1426–1446 (1989)
20. Lowe, D.G.: Fitting parameterized 3-Dimensional models to images. IEEE Trans. Pattern Anal. Mach. Intell. **13**, 441–450 (1991)
21. Lu, C.P., Hager, G.D., Mjolsness, E.: Fast and globally convergent pose estimation from video images. IEEE Trans. Pattern Anal. Mach. Intell. **22**, 610–622 (2000)
22. Garro, V., Crosilla, F., Fusiello, A.: Solving the PnP problem with anisotropic orthogonal procrustes analysis. In: Second Joint 3dim/3dpvt Conference: 3d Imaging, Modeling, Processing, Visualization & Transmission (3dimpvt 2012), pp. 262–269 (2012)

23. Hu, Z.Y., Wu, F.C.: A note on the number of solutions of the noncoplanar P4P problem. IEEE Trans. Pattern Anal. Mach. Intell. **24**, 550–555 (2002)
24. Zhang, C.X., Hu, Z.Y.: Why is the danger cylinder dangerous in the P3P problem. Zidonghua Xuebao/Acta Automatica Sinica **32**, 504–511 (2006)
25. Wang, T., Wang, Y.C., Yao, C.: Some discussion on the conditions of the unique solution of P3P problem. IEEE ICMA 2006: Proceeding of the 2006 IEEE International Conference on Mechatronics and Automation, vols. 1–3, Proceedings, pp. 205–210 (2006)
26. Zhang, Z.Y.: A flexible new technique for camera calibration. IEEE Trans. Pattern Anal. Mach. Intell. **22**, 1330–1334 (2000)
27. Badiali, G., Ferrari, V., Cutolo, F., Freschi, C., Caramella, D., Bianchi, A., Marchetti, C.: Augmented reality as an aid in maxillofacial surgery: validation of a wearable system allowing maxillary repositioning. J. Craniomaxillofac. Surg. **42**, 1970–1976 (2014)
28. Ferrari, V., Megali, G., Troia, E., Pietrabissa, A., Mosca, F.: A 3-D mixed-reality system for stereoscopic visualization of medical dataset. IEEE Trans. Biomed. Eng. **56**, 2627–2633 (2009)
29. Cutolo, F., Parchi, P.D., Ferrari, V.: Video see through AR head-mounted display for medical procedures. In: 2014 IEEE International Symposium on Mixed and augmented reality (ISMAR), pp. 393–396 (2014)

Journal of Cranio-Maxillo-Facial Surgery 42 (2014) 1970–1976

Contents lists available at ScienceDirect

Journal of Cranio-Maxillo-Facial Surgery

journal homepage: www.jcmfs.com

Augmented reality as an aid in maxillofacial surgery: Validation of a wearable system allowing maxillary repositioning

Giovanni Badiali [a, *], Vincenzo Ferrari [b], Fabrizio Cutolo [b], Cinzia Freschi [b], Davide Caramella [b, c], Alberto Bianchi [d], Claudio Marchetti [d]

[a] PhD School in Surgical Sciences (Head: Prof. Andrea Stella, MD), University of Bologna, Italy
[b] EndoCAS Center (Head: Prof. Mauro Ferrari, MD), University of Pisa, Italy
[c] Department of Radiology, Santa Chiara Hospital (Head: Prof. Davide Caramella, MD), University of Pisa, Italy
[d] Oral and Maxillofacial Surgery Unit, S.Orsola-Malpighi University Hospital (Head: Prof. Claudio Marchetti, MD, DMD), University of Bologna, Italy

ARTICLE INFO

Article history:
Paper received 22 March 2014
Accepted 1 September 2014
Available online 11 September 2014

Keywords:
Augmented reality
Computer-assisted surgery
Image-guided surgery
Maxillofacial orthognathic surgery
Maxillofacial abnormalities

ABSTRACT

Aim: We present a newly designed, localiser-free, head-mounted system featuring augmented reality as an aid to maxillofacial bone surgery, and assess the potential utility of the device by conducting a feasibility study and validation.
Methods: Our head-mounted wearable system facilitating augmented surgery was developed as a stand-alone, video-based, see-through device in which the visual features were adapted to facilitate maxillo-facial bone surgery. We implement a strategy designed to present augmented reality information to the operating surgeon. LeFort1 osteotomy was chosen as the test procedure. The system is designed to exhibit virtual planning overlaying the details of a real patient. We implemented a method allowing performance of waferless, augmented-reality assisted bone repositioning. In vitro testing was conducted on a physical replica of a human skull, and the augmented reality system was used to perform LeFort1 maxillary repositioning. Surgical accuracy was measured with the aid of an optical navigation system that recorded the coordinates of three reference points (located in anterior, posterior right, and posterior left positions) on the repositioned maxilla. The outcomes were compared with those expected to be achievable in a three-dimensional environment. Data were derived using three levels of surgical planning, of increasing complexity, and for nine different operators with varying levels of surgical skill.
Results: The mean error was 1.70 ± 0.51 mm. The axial errors were 0.89 ± 0.54 mm on the sagittal axis, 0.60 ± 0.20 mm on the frontal axis, and 1.06 ± 0.40 mm on the craniocaudal axis. The simplest plan was associated with a slightly lower mean error (1.58 ± 0.37 mm) compared with the more complex plans (medium: 1.82 ± 0.71 mm; difficult: 1.70 ± 0.45 mm). The mean error for the anterior reference point was lower (1.33 ± 0.58 mm) than those for both the posterior right (1.72 ± 0.24 mm) and posterior left points (2.05 ± 0.47 mm). No significant difference in terms of error was noticed among operators, despite variations in surgical experience. Feedback from surgeons was acceptable; all tests were completed within 15 min and the tool was considered to be both comfortable and usable in practice.
Conclusion: We used a new localiser-free, head-mounted, wearable, stereoscopic, video see-through display to develop a useful strategy affording surgeons access to augmented reality information. Our device appears to be accurate when used to assist in waferless maxillary repositioning. Our results suggest that the method can potentially be extended for use with many surgical procedures on the facial skeleton. Further, our positive results suggest that it would be appropriate to proceed to in vivo testing to assess surgical accuracy under real clinical conditions.

1. Introduction

Augmented reality (AR) is an innovative technology allowing merger of data from the real environment with virtual information.

* Corresponding author. Oral and Maxillofacial Surgery Unit, S.Orsola-Malpighi University Hospital, Viale Massarenti 9, 40138 Bologna, Italy. Tel.: +39 051 6363415; fax: +39 051 6363641.
E-mail addresses: giovanni.badiali2@unibo.it, giovanni.badiali@gmail.com (G. Badiali).

http://dx.doi.org/10.1016/j.jcms.2014.09.001
1010-5182/

The virtual data may be simply informative (such as textual or numerical values relevant to what is under observation) or may consist of three-dimensional virtual objects inserted within the real environment in spatially defined positions.

In the context of image-guided surgery, improvements based on AR may represent the next significant technological development in the field, because such approaches complement and integrate the concepts of surgical navigation based on virtual reality. AR provides a surgeon with a direct perception of how virtual content, generally obtained via medical imaging, is located within an actual scene (Ferrari et al., 2009; Freschi et al., 2009). This is particularly valuable in the context of head-and-neck surgery, in which the extreme anatomical complexity has encouraged the development of several innovative devices. However, the sophistication of such surgery and the longer operative times required have compromised the widespread implementation of such devices. Moreover, the necessary equipment is expensive (Hupp, 2013; Turchetti et al., 2010). For these reasons, the technology demands both methodological and economic rationalisation.

In recent years, tools (or defined applications) employing AR have been designed and tested in the context of several surgical and medical disciplines, including maxillofacial surgery (Marmulla et al., 2005b, 2005a; Mischkowski et al., 2006; Zinser et al., 2013b), dentistry (Bruellmann et al., 2013), ENT surgery (Caversaccio et al., 2008; Nakamoto et al., 2012), neurosurgery (Inoue et al., 2013; Mahvash and Besharati Tabrizi, 2013) and general surgery (Kowalczuk et al., 2012; Azagury et al., 2012; Marzano et al., 2013). The user experiences an AR view presented with the aid of various technical modalities, such as a traditional display, a tablet display, or a wearable display (Freschi et al., 2009; Mischkowski et al., 2006; Mezzana et al., 2011; Shenai et al., 2011; Gavaghan et al., 2012; Deng et al., 2013; Suenaga et al., 2013). Nevertheless, as is true of many emerging technologies, no standard method by which AR technology could/should be transferred to clinical practice has yet been developed (Dixon et al., 2013).

Bearing these facts in mind, we used a new localiser-free, head-mounted, stereoscopic, video see-through display to develop a useful strategy for delivery of AR information to the surgeon. Our study is the result of collaboration between the EndoCAS Laboratory of the University of Pisa (Italy) and the Maxillofacial Surgery Unit of the S. Orsola-Malpighi University Hospital of Bologna (Italy).

For brevity, the system will be termed the “wearable augmented reality for medicine” (WARM) device. The aim of the present study was to describe our new tool and to validate the accuracy thereof when used as an aid during surgery on facial bones. We also explore the potential for its wider application in maxillofacial surgery in general.

2. Materials and methods

2.1. The WARM device

The device (Fig. 1) is based on a lightweight, stereoscopic head-mounted display (HMD) that is widely available; this is the Z800 instrument of eMagin (Bellevue, WA, USA). A support placed in front of the HMD holds two USB SXGA cameras (uEye UI-1646LE; IDS, Obersulm, Germany) and a 1/3” image sensor placed precisely in front of the user's eyes. Two optics (mounted on either camera) ensure an anthropometric field of view. Augmented reality is provided by software that runs on conventional personal computers (Ferrari et al., 2009). Alignment between the real and virtual world is achieved in the absence of an external tracking system, via processing of video frames grabbed by the cameras. In particular, a machine vision algorithm is used to superimpose the virtual content onto real data provided by the cameras, with subpixel accuracy, using small coloured spheres that do not compromise the surgeon's view of the real scenario (Fig. 2).

2.2. In vitro setup

The test was conducted on a replica of a cadaveric human skull. The real skull underwent CT scanning and the DICOM files were segmented using a semi-automatic segmentation tool integrated into the ITK-Snap open-source platform (Ferrari et al., 2012). Manual segmentation refinement (using a touch screen) was performed to obtain detailed information on small anatomical structures (e.g. the foramen rotundum, foramen spinosum, lamina cribriformis, and hypoglossal canal). The 3D virtual model distinguished pneumatised bones very well. In particular, the nasal cavities and the paranasal sinuses were computer-generated in minute detail.

The virtual model of the skull was cut along the LeFort 1 osteotomy line. The two resulting virtual objects (the upper skull and maxilla) were exported as STL files and replicated in ABS using a 3D printer (Stratasys Elite; Eden Prairie, MN, USA). LeFort 1 osteotomy and repositioning of the upper maxilla were chosen as test procedures featuring the principal features of maxillofacial

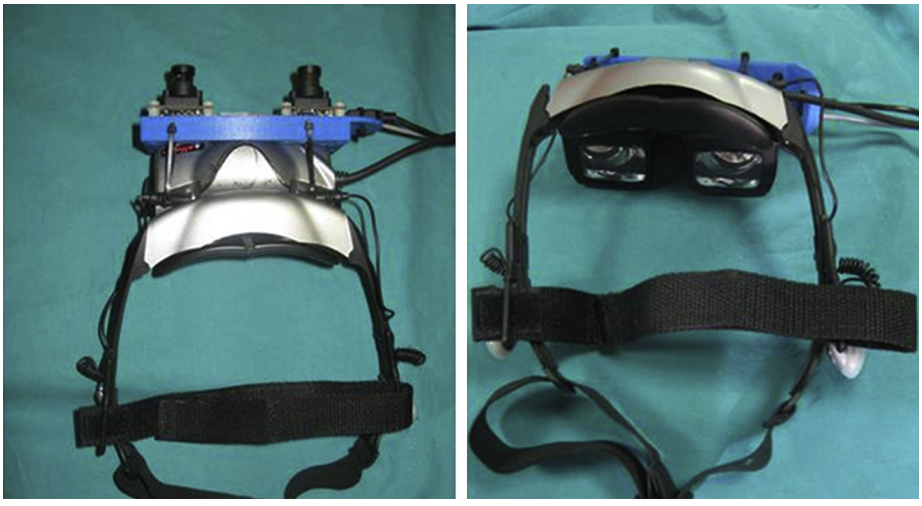


Fig. 1. The WARM system features the mounting of two external cameras on top of a commercial 3D visor.

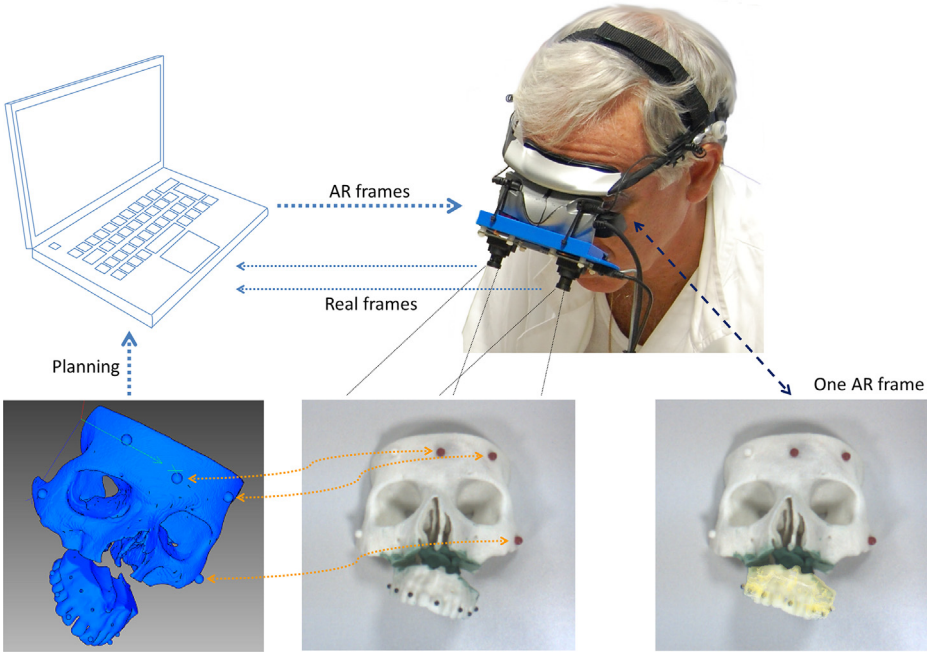


Fig. 2. The WARM system. The two external cameras acquire real video frames. Our software application merges the virtual 3D model derived during surgical planning with real data from the camera frames and sends the result to the two internal monitors. Alignment between real and virtual information is obtained by calculating the positions of coloured markers relative to camera data, with respect to their known positions (recorded during planning), using detailed pre-operative CT images.

surgery. Thus, the technique involves surgery on facial bones; the approach is a form of "semi-buried" surgery when performed under real clinical conditions; the technique involves complex three-dimensional movements of a rigid object in space; and the technique is often performed in clinical practice worldwide.

Before printing, three 6 mm-diameter balls were inserted into the virtual model as marker references for the WARM device. Further, three reference holes were drilled into the vestibular cortical bone, over the teeth (*anterior* in the premaxillary region; *posterior left* and *posterior right* in the respective molar regions). The holes were used as references to evaluate the position of the maxilla. Thus, each hole was designed to receive the tip of the tester probe used for validation (see below), to guarantee unique selection of each reference point.

The upper skull was fixed on a wooden holder. The maxillary piece was connected to the upper skull with plasticine (this material is highly malleable but rigid when shaped). This construction served as a fixing device for the maxilla once the planned position was attained, yet allowed the maxilla to be manually adjusted in space.

To evaluate the accuracy of our system, we used a traditional navigation platform (the eNlite Navigation System running iNtellect Cranial Navigation Software version 1.0; Stryker, Freiburg, Germany) featuring an active infrared localiser. Our setup is shown in Fig. 3, which identifies the tracking and pointing instruments of the navigation system.

2.3. AR visualisation: ergonomic evaluation

A preliminary assessment was conducted to evaluate the ergonomics of the device, actual usability in a surgical environment, and (in particular) the best method of displaying the virtual content. One surgeon (GB) and three engineers (VF, FC, and CF) collaborated in this work. Tests were conducted using different display modalities in an attempt to define a modality that was optimally comfortable and that had the smallest perceived parallax error. We commenced with the display modality most frequently adopted in similar work (Mischkowski et al., 2006; Suenaga et al., 2013); thus, a rendered virtual reality was superimposed on the real camera frames. We found that this display modality, although allowing us to change the transparency settings, did not satisfactorily establish

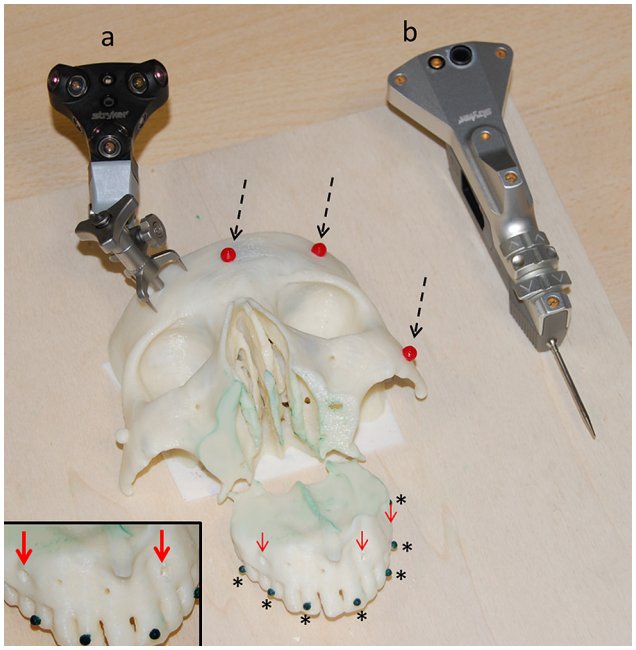


Fig. 3. Our setup: A physical replica of the human skull is fixed onto a wooden holder, and the three coloured spheres on the model (the black dashed arrows) ensure alignment between the real and virtual world in the absence of any external tracking system. We used a machine-based vision algorithm. The coloured brackets on the teeth (the black asterisks) are the reference markers for the AR display modality; three of the six holes on the maxilla (the red arrows) were used to evaluate accuracy with the aid of an external navigation system. The tracker of the navigation system is fixed onto the model in (a). In (b), the pointer of the navigation system, used to assess the position of reference holes, is shown beside the model.

the relative positions of the real and the virtual (planned) maxilla, particularly in terms of depth, and was thus unable to aid in correct performance of the surgical task.

The display modality that we finally selected is shown in Fig. 4. The virtual information consists of green asterisks drawn in positions defined during planning. For each virtual asterisk, a coloured landmark was fixed on the maxilla. Use of this display modality allowed us to study how to move the maxilla to replicate planning, and also if a planned position had been attained with high precision. Coloured landmarks can be fixed (for example) on the brackets of an orthodontic appliance, as shown, or on a CAD/CAM splint or guide.

2.4. Accuracy evaluation testing

2.4.1. Virtual surgical planning

Using Maya (Autodesk; Toronto, Canada), the virtual maxilla was moved in space as dictated by three surgical plans of increasing complexity (Fig. 5).

1) Maxilla 6 mm forward.
2) Maxilla 5 mm forward and 1 mm downward.
3) Maxilla 6 mm forward, 1 mm downward, and with 15° roll and 10° pitch.

Each plan was saved as an STL file.

2.4.2. Test

Three maxillofacial surgeons (AB, GB, and LP); three trainees in maxillofacial surgery (SA, EB, and FR); and three engineers (VF, FC, and CF) were involved in the testing; we evaluated interobserver variability. Hence, the three groups included appropriate representatives of users with different levels of surgical skill (from unskilled engineers to highly skilled surgeons). After only one warm-up session, during which each subject was trained to use the WARM device, the subject was asked to manually reposition the maxillary segment, using guidance afforded by the device. The procedure was repeated by each subject for each of the three virtual plans; the maximum test duration was 15 min. After completion of each test, the position of the maxillary segment was confirmed using the navigation system described in the following paragraph.

2.4.3. Accuracy measurement

The CT scan of the skull was imported into the navigation system as a DICOM file and the three plans, defined in the CT reference system, were loaded into the navigation system as STL files (Fig. 6). The tracker of the navigation system was fixed on the model of the skull and the registration process featured a point-based procedure (using defined anatomical points) with subsequent surface refinement; the target registration error was 0.3 mm. After each trial session, the navigation system probe was inserted into each of the three reference holes on the maxilla and the probe tip positions were saved (Fig. 7). We next determined, for each subject, the linear distances between the real positions of the reference holes

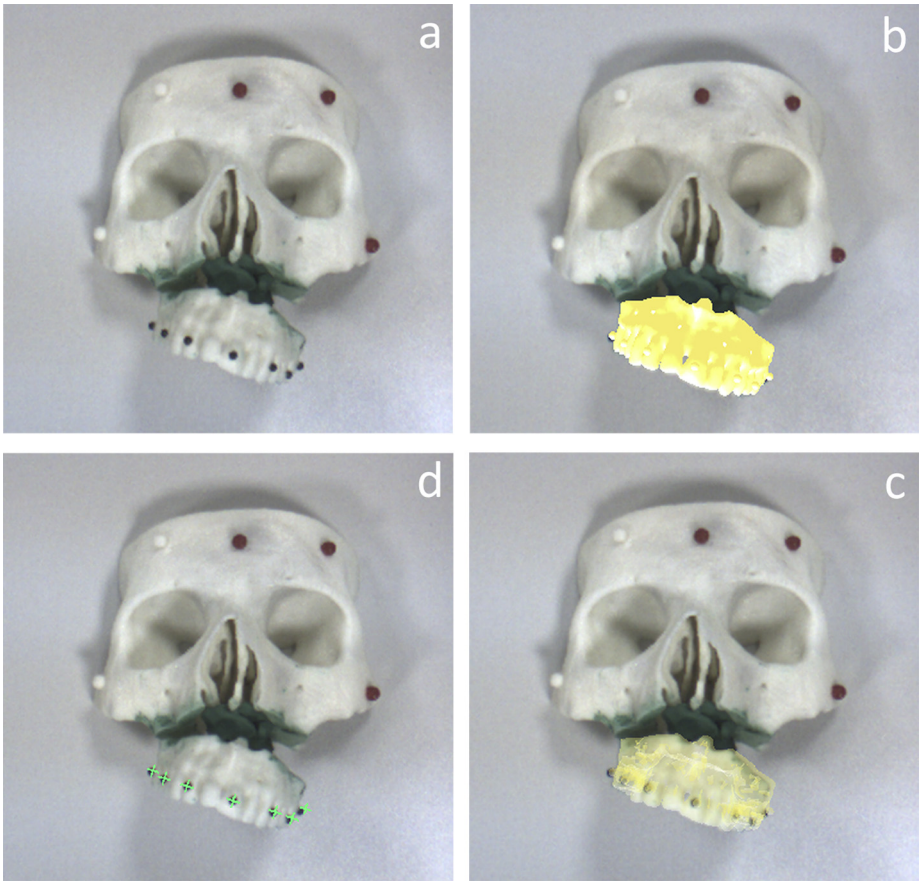


Fig. 4. Different approaches to presentation of AR information: (a) A real video frame; (b, c) A traditional approach, presenting the virtual model on real camera frames. Using the approaches of (b) and (c), it was not possible to completely perceive the relationship between the real and virtual world. (d) A more ergonomic form of visualisation, ultimately selected by us to permit the subject to determine if the real maxilla was positioned correctly. The virtual information consists of a green asterisk for each coloured landmark on the maxilla.

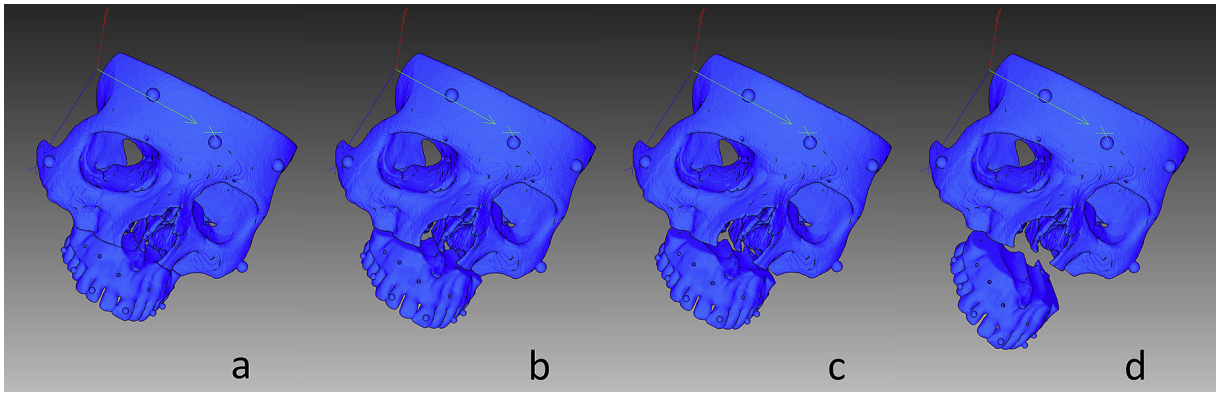


Fig. 5. The virtual maxilla (a) was moved in space as dictated by three surgical plans of increasing complexity: (b) 6 mm forward; (c) 5 mm forward and 1 mm downward; (d) 6 mm forward, 1 mm downward, with 15° roll and 10° pitch.

(measured using the navigation system) and the expected positions (defined during planning).

2.5. Statistical analysis

The linear distances between the expected and real positions were computed with the aid of MatLab (Mathworks Inc.; Natick, MA, USA) to obtain descriptive statistics.

3. Results

The results are shown in Table 1. The mean error was 1.70 ± 0.51 mm. The axial errors were 0.89 ± 0.54 mm on the sagittal axis, 0.60 ± 0.20 mm on the frontal axis, and 1.06 ± 0.40 mm on the craniocaudal axis. The simplest plan was associated with a slightly lower mean error (1.58 ± 0.37 mm) than the more complex plans (medium: 1.82 ± 0.71 mm; difficult: 1.70 ± 0.45 mm). The mean error for the anterior reference point was lower (1.33 ± 0.58 mm) than those for the posterior right (1.72 ± 0.24 mm) and posterior left (2.05 ± 0.47 mm) points. No significant difference was noted among operators, despite variation in surgical experience (Fig. 8). Feedback from surgeons was acceptable; all procedures were completed within 15 min and the tool was found to be both comfortable and usable.

4. Discussion

In recent years, the discipline of maxillofacial surgery has undergone a remarkable rate of technological innovation. This is because the complex three-dimensional anatomy of the face,

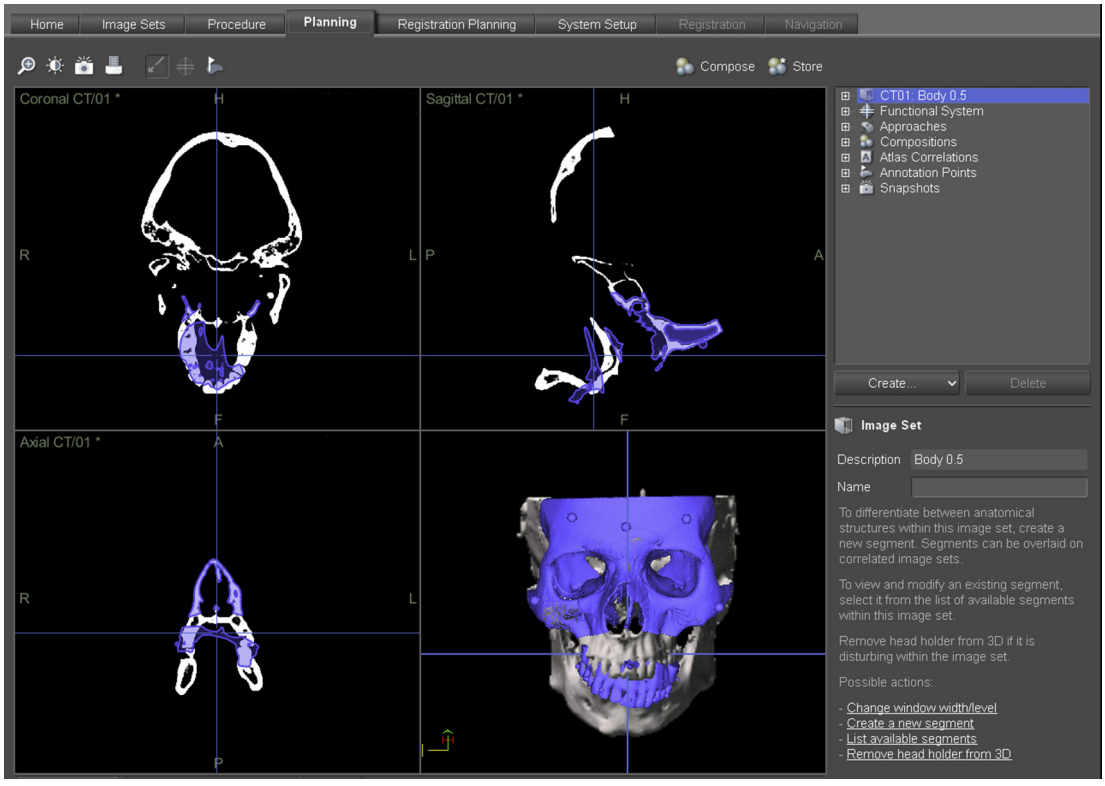


Fig. 6. A screenshot of the navigator. The blue planning scenario is loaded together with the original CT scan.

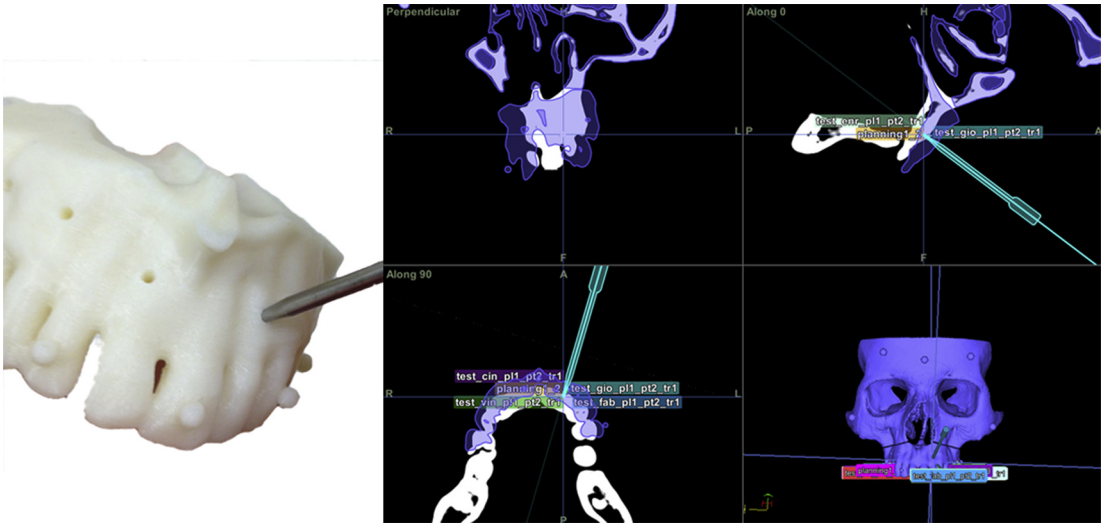


Fig. 7. The accuracy evaluation process is shown in detail. On the left, the pointer slides into a reference hole of the maxilla (the hole termed "anterior one"); on the right, the navigation system shows where the tip of the pointer is actually located (compared with the planned location). The coordinates of the real position are recorded and used to estimate errors in linear measurements.

together with the need for surgical precision and the increasing number of requests for morphological surgery, have resulted in surgeons demanding advanced technological assistance. Thus, virtual planning software and navigation systems are today widely used by maxillofacial surgeons (Mazzoni et al., 2010; Zinser et al., 2013a). However, substantial room for improvement remains. The accuracy afforded by the technology must increase, as must the usability of devices in real clinical practice.

AR represents an important step toward the practical integration of several ground-breaking technologies. AR fuses navigational surgery and virtual planning with the real surgical field. AR can be displayed on a traditional monitor, or directly in front of the eyes of a surgeon who uses a wearable system such as WARM.

Our results suggest that wearable AR is both comfortable and functional, permitting a surgeon to maintain their natural operative posture during surgery performed at different angles, without losing the three-dimensional relationship between the real scene and that afforded by virtual planning. This is of particular importance. We found that surgeons frequently change their line of view during an operation to control the three-dimensional position of the maxilla from all angles. Further, the use of a stereoscopic device obviates any need for an external localiser, because the device can serve as both a frame-grabber and a localiser.

Our system has other significant features; these are the registration and tracking modalities. Indeed, WARM does not require an external infrared camera or an electromagnetic field generator (unlike standard navigation systems), but uses visible light. The head-mounted cameras grab the scene and use frames both to track the patient's skull and to realise the AR environment. In our laboratory setup, three coloured (red) spheres were placed on the skull surface to simplify the experimental conditions, but, in the clinic, a skull-mounted tracker with coloured spheres could be used. This would obviously require that a patient-specific registration process be conducted.

In terms of validation, our results suggest that the device affords an average accuracy of 1.70 ± 0.51 mm, which is good in the context of maxillofacial surgery. This result is even more significant because waferless surgery was planned. Considering the axial error components, the lowest error (0.60 ± 0.20 mm) was measured along the frontal axis, the next-largest error (0.89 ± 0.54 mm) along the sagittal axis, and the greatest error (1.06 ± 0.40 mm) along the vertical axis. Thus, use of the device is associated with very small errors (below 1 mm) in terms of frontal and sagittal malposition of the maxilla; this is very good compared with orthognathic surgical standards. Further, even the error on the vertical plane (around 1 mm) is excellent, because the vertical dimension remains the most complex in terms of intraoperative control (Song and Baek, 2009). Such errors are not discernible when a patient is evaluated

Table 1
Errors for each target and plan for all operators, and the relative means (overall mean in bold).

	Plan 1	Plan 2	Plan 3	Mean
Target 1	1.71 mm (±0.24)	1.80 mm (±0.18)	1.63 mm (±0.34)	1.72 mm (±0.24)
Target 2	1.07 mm (±0.17)	1.47 mm (±0.12)	1.45 mm (±0.45)	1.33 mm (±0.58)
Target 3	1.96 mm (±0.32)	2.18 mm (±0.69)	2.02 mm (±0.49)	2.05 mm (±0.47)
Mean	1.58 mm (±0.37)	1.82 mm (±0.71)	1.70 mm (±0.45)	**1.70 mm (±0.51)**

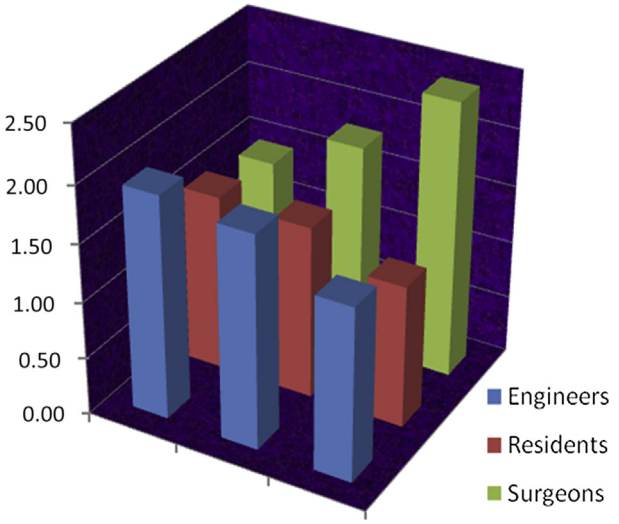


Fig. 8. Mean errors in mm (over three trials and three reference holes) for each of the nine participants. No difference between engineers and physicians is evident.

after intervention, and surgery can thus be considered as having been performed optimally.

No significant difference in errors was evident when the three planning modes were compared. The simplest plan was associated with error values slightly lower, on average, than the others; this is quite understandable. This suggests that our method could be extended to aid in the performance of any orthognathic procedure on the maxilla, regardless of the complexity of the required movements.

Average errors measured to the anterior reference hole were lower than those to the posterior hole. This is probably because the position of the anterior reference hole is the only one that can be controlled from every viewpoint.

Another interesting result is the non-dependency of accuracy on user experience; all of the experienced surgeons, trainees, and (even) engineers obtained comparable results. All test procedures were completed within 15 min after a single 15 min warm-up session.

The use of small virtual asterisks, corresponding to coloured landmarks fixed on the brackets of the orthodontic appliance or onto splints, turns out to be an efficient way to present AR guidance to the surgeon. Our device is simple and easy to use, and shows promise for assisting in maxillofacial orthognathic procedures.

5. Conclusion

We used a new, localiser-free, head-mounted, stereoscopic, video see-through display to develop a useful strategy affording the surgeon access to AR information. Our results suggest that the WARM device would be accurate when used to assist in waferless maxillary repositioning during the LeFort 1 orthognathic procedure. Further, our data suggest that the method can be extended to aid the performance of many surgical procedures on the facial skeleton. Also, in vivo testing should be performed to assess system accuracy under real clinical conditions.

Conflict of interest statement

This work was partially supported for EndoCAS by Opera Project (Advanced OPERAting room). Tuscany Regional Funds: PAR FAS 2007–2013 Azione 1.1 P.I.R. 1.1.B.

The authors have no financial interest or personal relationships with other people or organizations that may have inappropriately influenced the work presented here.

References

Azagury DE, Ryou M, Shaikh SN, San José Estépar R, Lengyel BI, Jagadeesan J, et al: Real-time computed tomography-based augmented reality for natural orifice transluminal endoscopic surgery navigation. Br J Surg 99: 1246–1253, 2012

Bruellmann DD, Tjaden H, Schwanecke U, Barth P: An optimized video system for augmented reality in endodontics: a feasibility study. Clin Oral Investig 17: 441–448, 2013

Caversaccio M, Giraldez JG, Thoranaghatte R, Zheng G, Eggli P, Nolte L, et al: Augmented reality endoscopic system (ARES): preliminary results. Rhinology 46: 156–158, 2008 Special report (technical note)

Deng W, Li F, Wang M, Song Z: Easy-to-use augmented reality neuronavigation using a wireless tablet PC. Stereotact Funct Neurosurg 92: 17–24, 2013

Dixon BJ, Daly MJ, Chan H, Vescan AD, Witterick IJ, Irish JC: Surgeons blinded by enhanced navigation: the effect of augmented reality on attention. Surg Endosc 27: 454–461, 2013

Ferrari V, Megali G, Troia E, Pietrabissa A, Mosca F: A 3-D mixed-reality system for stereoscopic visualization of medical dataset. IEEE Trans Biomed Eng 56: 2627–2633, 2009

Ferrari V, Carbone M, Cappelli C, Boni L, Melfi F, Ferrari M, et al: Value of multi-detector computed tomography image segmentation for preoperative planning in general surgery. Surg Endosc 26: 616–626, 2012

Freschi C, Troia E, Ferrari V, Megali G, Pietrabissa A, Mosca F: Ultrasound guided robotic biopsy using augmented reality and human-robot cooperative control. Conf Proc IEEE Eng Med Biol Soc 2009: 5110–5113, 2009

Gavaghan K, Oliveira-Santos T, Peterhans M, Reyes M, Kim H, Anderegg S, et al: Evaluation of a portable image overlay projector for the visualisation of surgical navigation data: phantom studies. Int J Comput Assist Radiol Surg 7: 547–556, 2012

Hupp JR: Advanced technology—hammers looking for nails. J Oral Maxillofac Surg 71: 465–466, 2013

Inoue D, Cho B, Mori M, Kikkawa Y, Amano T, Nakamizo A, et al: Preliminary study on the clinical application of augmented reality neuronavigation. J Neurol Surg A Cent Eur Neurosurg 74: 71–76, 2013

Kowalczuk J, Meyer A, Carlson J, Psota ET, Buettner S, Pérez LC, et al: Real-time three-dimensional soft tissue reconstruction for laparoscopic surgery. Surg Endosc 26: 3413–3417, 2012

Mahvash M, Besharati Tabrizi L: A novel augmented reality system of image projection for image-guided neurosurgery. Acta Neurochir (Wien) 155: 943–947, 2013

Marmulla R, Hoppe H, Mühling J, Eggers G: An augmented reality system for image-guided surgery. Int J Oral Maxillofac Surg 34: 594–596, 2005a

Marmulla R, Hoppe H, Mühling J, Hassfeld S: New augmented reality concepts for craniofacial surgical procedures. Plast Reconstr Surg 115: 1124–1128, 2005b

Marzano E, Piardi T, Soler L, Diana M, Mutter D, Marescaux J, et al: Augmented reality-guided artery-first pancreatico-duodenectomy. J Gastrointest Surg 17: 1980–1983, 2013

Mazzoni S, Badiali G, Lancellotti L, Babbi L, Bianchi A, Marchetti C: Simulation-guided navigation: a new approach to improve intraoperative three-dimensional reproducibility during orthognathic surgery. J Craniofac Surg 21: 1698–1705, 2010

Mezzana P, Scarinci F, Marabottini N: Augmented reality in oculoplastic surgery: first iPhone application. Plast Reconstr Surg 127: 57e–58e, 2011

Mischkowski RA, Zinser MJ, Kübler AC, Krug B, Seifert U, Zöller JE: Application of an augmented reality tool for maxillary positioning in orthognathic surgery – a feasibility study. J Craniomaxillofac Surg 34: 478–483, 2006

Nakamoto M, Ukimura O, Faber K, Gill IS: Current progress on augmented reality visualization in endoscopic surgery. Curr Opin Urol 22: 121–126, 2012

Shenai MB, Dillavou M, Shum C, Ross D, Tubbs RS, Shih A, et al: Interactive presence and augmented reality (VIPAR) for remote surgical assistance. Neurosurgery 68: 200–207, 2011

Song K-G, Baek S-H: Comparison of the accuracy of the three-dimensional virtual method and the conventional manual method for model surgery and intermediate wafer fabrication. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 107: 13–21, 2009

Suenaga H, Hoang Tran H, Liao H, Masamune K, Dohi T, Hoshi K, et al: Real-time in situ three-dimensional integral videography and surgical navigation using augmented reality: a pilot study. Int J Oral Sci 5: 98–102, 2013

Turchetti G, Spadoni E, Geisler EE: Health technology assessment. Evaluation of biomedical innovative technologies. IEEE Eng Med Biol Mag 29: 70–76, 2010

Zinser MJ, Mischkowski RA, Dreiseidler T, Thamm OC, Rothamel D, Zöller JE: Computer-assisted orthognathic surgery: waferless maxillary positioning, versatility, and accuracy of an image-guided visualisation display. Br J Oral Maxillofac Surg 51: 827–833, 2013a

Zinser M, Sailer H, Ritter L: A paradigm shift in orthognathic surgery? A comparison of navigation, cad/cam splints and “classic” intermaxillary splints to surgical transfer of virtual orthognathic planning. Oral Maxillofac 71: 2151.e1–2151.e21, 2013b