POLITECNICO DI MILANO
MECHANICAL DEPARTMENT
DOCTORAL PROGRAMME IN MECHANICAL ENGINEERING

LOW-COST AUGMENTED REALITY FOR INDUSTRIAL PROBLEMS

Doctoral Dissertation of:
Guido Maria Re

Supervisor:
Prof. Umberto Cugini

Tutor:
Prof. Federico Cheli

The Chair of the Doctoral Program:
Prof. Bianca Maria Colosimo

2013 – XXV Cycle

Acknowledgements

I would like to express my gratitude to my supervisor, Professor Umberto Cugini for his support during my research. Furthermore, I would also like to thank Professor Monica Bordegoni for the suggestions that she gave me in these years.

Research also means team-work, for this reason I express all of my appreciation to all of the guys of KAEMaRT group whom I met during these years. Thanks to all of them for their advices, in particular Giandomenico, Samuele, Michele, Paolo, Marco, Chicco, Mario e Francesco. Thank you also for our "Spaghetti Time".

In the last year of my PhD student period, I went abroad for some months as a visiting scholar at the Iowa State University. I would like to thank the University that hosted me and, in particular, Professor James Oliver for the precious hints, our weekly meetings and the new approach to do research that he has indirectly shown me.

A special thanks to my family that has supported me during these years. I know that it is not still clear for them what a PhD means, but at the same time they are proud of this personal big accomplishment.

In these years I also had some (few) time to hang out with friends. I am grateful for all of them, who now are located in several countries and continents, and the time we have spent together.

Finally, a special thanks to the girl that suddenly crashed in my life and who is giving me peacefulness everyday: thank you, Cindu.

Abstract

MANY technological supports have been proposed and integrated in industrial activities so far. Their purpose is to increase the working performances, in order to speed up the accomplishing time and decrease the costs of an operation. Among them, Augmented Reality (AR) has been proposed, which is an emerging technology that allows to visualize and to interact with virtual objects in the real world. The exploitation of AR in the industrial working field can provide several advantages because it is able to display digital information directly in the working environment. These pieces of information can be data, instructions or alerts that are represented in a textual or graphical manner, otherwise virtual objects in general.

Currently, the majority of the applications for AR in Industry have never left the laboratories where they have been developed. Their purpose was just to check the feasibility of AR in Industry and their use in a real industrial context is indeed difficult. Moreover, the technologies used during the development of the applications were limited to fulfill the industrial demands or too much expensive.

In this Thesis we propose the use of the technology embedded in low-cost devices, coming from the mass market, as substitute or completion for traditional technologies used for AR. Since these devices are cheap, their use for an AR application is easy to afford for industrial companies and consequently their integration in industrial activities should be an acceptable investment. These low-cost devices are also good candidates for AR purposes because they usually embed robust, cutting-edge technologies and they are easy to find.

Unfortunately, we cannot directly integrate these technologies in applications for Industry because they have been designed for different purposes. Thus, we made these technologies working in the new industrial context. This step has been carried out by modifying the algorithms to manage the technologies. For the modification, we took into account the different needs that are coming from the user and the environment wherein he is working.

Once the technology is modified to operate in the industrial environment, we integrate it in some application to support specific activities. In particular, we focus on the phases of Product Design, Manufacturing, Maintenance and Inspection.

We carried out testing sessions to evaluate the performances of the integrated technology and performed some user test to assess the usability of the final application. From these tests, we can state that it is possible to substitute standard AR technologies by integrating low-cost ones. In fact, these last ones can achieve good performances that are acceptable for AR purposes, and the final applications are usable by the user.

Sommario

FINO ad ora, molte tecnologie sono state proposte e integrate in attività industriali come supporto. Il loro scopo è di aumentare il rendimento lavorativo, in modo da velocizzare il tempo di esecuzione di un compito e diminuirne i costi. Tra queste, è stata proposta la Realtà Aumentata (AR), una tecnologia emergente che permette di visualizzare e interagire con oggetti virtuali nel mondo reale. Lutilizzo della AR nel campo industriale può fornire vari vantaggi poiché ci permette di rappresentare informazioni digitali direttamente sul posto di lavoro. Queste informazioni possono essere dati, istruzioni, allarmi, i quali sono rappresentati in forma testuale o grafica, oppure oggetti virtuali in generale.

Attualmente, la maggior parte delle applicazioni AR per lIndustria non hanno mai lasciato i laboratori in cui sono state sviluppate. Il loro scopo era soltanto di verificare la fattibilità della AR nellIndustria tanto che il loro utilizzo in un reale contesto industriale risulta difficoltoso. Oltretutto, le tecnologie usate nello sviluppo delle applicazioni sono limitate per soddisfare le richieste da parte dellIndustria oppure troppo costose.

In questa Tesi, proponiamo luso delle tecnologie integrate in dispositivi a basso costo provenienti dal mercato di massa come sostituto o complemento alle tradizionali applicazioni usate finora in AR. Poiché questi dispositivi sono economici, il loro utilizzo per applicazioni AR è facilmente sostenibile per società industriali e di conseguenza la loro integrazione in attività industriali dovrebbe essere un investimento accettabile. Questi dispositivi del mercato di massa sono inoltre buoni candidati per scopi AR perché solitamente sono robusti, dotati di tecnologie allavanguardia e sono semplici da reperire.

Purtroppo, non possiamo integrare queste tecnologie direttamente in applicazioni per lIndustria poiché sono state progettate per scopi differenti. Per questo motivo, abbiamo dovuto rendere queste tecnologie funzionanti nel contesto industriale. Ciò è stato eseguito modificando gli algoritmi che gestiscono le tecnologie. Per questa modifica, abbiamo considerato le varie necessità dellutente e dellambiente, dove esso lavora.

Una volta che la tecnologia è stata modificata per lavorare nel campo industriale, abbiamo sviluppato alcune applicazioni a supporto di specifiche attività. Nello specifico, abbiamo focalizzato la nostra attenzione sulle fasi della Progettazione, della Fabbricazione, della Manutenzione e Ispezione.

Abbiamo condotto delle sessioni di test in modo da valutare le caratteristiche della tecnologia integrata e abbiamo portato a termine alcuni test con utenti per verificare l'usabilità dellapplicazione finale. Da questi test, possiamo affermare che è possibile sostituire le usuali tecnologie per AR andando a integrare tecnologie a basso costo. Infatti, queste ultime possono raggiungere buone prestazioni le quali sono accettabili per scopi AR e le applicazioni finali sono usabili dallutente.

Contents

Part I

Overview

CHAPTER 1

Introduction

The use of electronic devices and information technology as support for industrial phases are nowadays very diffuse. Their integration within an industrial context is typically made to increase the efficiency and the effectiveness of a process. Actually, companies have to satisfy some demands from the global market, which pushes them to make goods with a high ratio between quality and price and from the final users, which are harder and harder to please. Therefore, companies continuously improve the way of making products, in order to cut time and costs of fulfillment and they resort to production tools that are more versatile, in order to manage requests with high variability. The support provided by technological solution in this scenario is usually to relieve the work of the operators by lightening some tasks.

The introduction of PCs in several industrial activities led to the use of technologies like Virtual Reality (VR) in the industrial context. The general concept of VR refers to a synthetic environment generated by computer. This environment is totally digitalized and only the user, who experiments and interacts with, is real. The creation of synthetic worlds by means of this technology enables to imitate the real world or to overcome the physical laws, providing to the user an environment with different physics and properties.

VR can be used in the industrial field to provide information about the product before it is physically made. The digital representation of the product, usually called *virtual prototype*, can be used to carry out some analysis during development phases. In this way, there is a minor need of tests on a real physical prototype. In addition, a virtual prototype is more versatile than a physical one because it can be modified by specific software in a quick way. The possibility to easily modify the virtual object let to create also variations of the product to test. In conclusion, VR makes the development of a product more effective because some industrial phases, usually carried out by using physical prototypes, can be by-passed and performed in a digital environment.

One of the limitations of VR technologies is that the virtual object is represented in a synthetic world. Thus, the creation of the virtual environment is mandatory in order to perform some analysis on the object.

1.1 Augmented Reality

There is a technology that is able to contextualize digital information in a real environment. This technology is called Augmented Reality (AR) and allows the user to see and interact with virtual contents in the real world.

AR can be seen as a technology strictly related to VR and the reality at the same time. Even if it seems that there is no relation between VR and Reality, they are bounded and can share contents (real and virtual) among them. According to Milgram [106, 107], it is possible to link them together by means of the Reality - Virtuality continuum, shown in Figure 1.1, wherein real world and VR are the two extremes. Mixed Reality stands in between the two extremes of this continuum because it shares something from the real world and from the synthetic in the same environment.

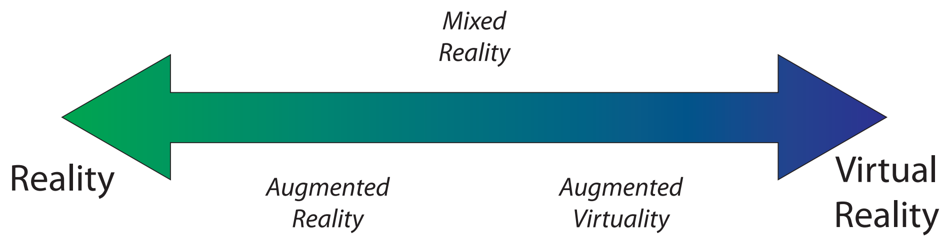


Figure 1.1: *Milgram's Reality-Virtuality Continuum.*

According to the amount of reality and virtual content present in the environment, it is possible to classify Mixed Reality in two categories: Augmented Reality, if the content is mainly real, and Augmented Virtuality, if the environment is mainly virtual. Thus, Augmented Reality can be defined as augmenting the natural feedback and stimuli to a user by means of simulated cues [108]. In this way, the user is still able to perceive the real world although it is integrated by something that is generated by computer.

Even though AR considers the integration of additional virtual information to the human senses in general, it is usually related on *vision*. This is due most likely because sight is the most developed of human senses and, for this reason, it is the predominant and suitable way to convey information. Indeed, human experience, knowledge and information have been historically communicated in words and pictures, which are related to the sense of sight. For this reason, AR will usually be related to visual augmentation in this dissertation.

The use of AR technology in the industrial field turns out to be more intuitive than VR because the digital information is directly placed in the working environment. This provided information can be of different typology. They can be data, which are represented in a graphic or textual manner, coming from sensors or analysis, instructions about the task to perform or about an alert, notations or simply virtual prototypes in a real context.

The integration of AR in an industrial context can relieve some working phases. Actually, pieces of information, which are usually provided by mean of paper supports or displayed on a monitor, can be directly placed in the working field or on top of the object to deal with. Since the information is contextualized, it should be easier for the user to understand it. Moreover, in case of analysis of a virtual prototype, its simpler comprehension makes possible to analyze the object even to people with different backgrounds, enabling the communication between people from divers fields. The greater versatility of AR despite VR and traditional approaches can lead to a major efficiency for the user in the working environment and consequently an increase of the efficiency of industrial phases.

1.2 Context

Some simple applications that use AR have already been developed in fields like entertainment and marketing. On the other hand, the development of AR industrial applications is still in an embryonic phase [67]. One of the reasons of the minor applicability of AR in industrial contexts is due to strict industrial demands. AR applications for Industry should be very easy to use, provide robust technologies and cheap. Unfortunately, many AR technologies are usually limited for being used in an industrial environment as their performances are not enough for working properly in this field and also because they are not handy or they are just too expensive if they want to fulfill the industrial demands.

Today, new technological devices addressed to everyday life purposes are constantly coming out on market. These devices usually offer technology embedded for a small price. They are usually produced in large quantity, so their cost is low, thanks to an economical strategy of mass productions. Examples of these low-cost devices are many video-game tools and communication systems as smart phones.

Even though these low-cost technological devices have been designed for a specific purpose, they can be adapted to be used in an industrial context. In this way, they could substitute the AR technologies proposed so far.

The technology embedded in these low-cost devices is usually described by the following characteristics:

Good performance. The technical features of these devices are usually very good because they are designed to be reactive in real-time and to ensure the satisfaction of the users as much as possible. For this reasons they are usually equipped by advanced or even at state-of-art technologies.

Robustness. The device is designed to work properly in order to always satisfy the customer requirements. Thus, the device is put through some tests to check its functioning before coming out on market.

Availability. They are widely distributed and easy to get in a small time. Moreover, the technologies available in this way are ready-to-use, because they have been integrated in a system already and nothing should be developed from scratch.

Low Price. We can get good performances from these technologies for a lower price than many high performance technologies usually experimented in AR. From an industrial point of view, this makes them more preferable because the investment is easy to afford. Moreover, the ratio between performances and price is often better that the one of traditional technologies proposed so far.

These characteristics make these low-cost technologies good candidates to be integrated in AR application for Industry.

The majority of the previous works regarding AR as tool for industrial contexts make use mainly of technologies already available. The purpose of these works is usually to show the potentiality of AR. It is not a problem if the subsequent application is often limited or difficult to use because these AR applications never leave the lab where they have been developed. There is a lack of research to take into account other technologies, in particular the low-cost ones, and, moreover, there is no evidence that their use is feasible in a real working environment.

1.3 Objective

The aim of this research is to investigate the feasibility of using low-cost technological devices for AR purposes in the industrial field. These devices come from the mass market and they could offer some benefits if they are integrated in some industrial phases. Actually, they could be a valid substitute and/or they could solve some open issues of the traditional technologies proposed for AR purposes so far.

Unfortunately, the use of these low-cost devices in an AR support for industrial activities is not straightforward and it could entail some issues. The first issue is *how* to integrate this kind of technology in Industry. In fact, their initial purpose is not addressed to any industrial phase and they should be adapted to work in these contexts. For this reason, we have to face with its integration, by managing the technology for a new purpose. Moreover, we have to consider that every industrial context has its own particular requests that the integration of these low-cost devices should satisfy. Once the low-cost technology is modified to work in an industrial environment, the second problem is related to the *limitations* that its use could involve. In fact, the technology could show poor performances if used for other purposes. Therefore, we have to compare their characteristics to other well-known technologies. If the technology turns out to be usable for AR purposes, we can use it for an AR application. Finally, we have to check the functioning of the application in *real working conditions*. For this reason, we have to perform some tests with users in order to assess its usability.

Summing up, the objectives of this Thesis can be expressed by some *research questions*. These questions are connected together so that only a positive response of a previous one allows us to access to the next one. The research questions are the following:

- is it possible to integrate low-cost technologies from mass market for AR purposes in the industrial context?
- do the performances of these technologies fulfill the industrial demands?
- is the AR application that relies on low-cost technologies usable in an industrial context?

In this work, we try to answer to these questions by taking into account some industrial activities as study cases. In particular, we justify the answers by some testing sessions.

1.4 Contribution

The thesis focuses on some industrial phases and it proposes some new low-cost approaches to overcome existing limitations that are present in technologies proposed for AR so far. In particular, this work is related to some phases of product development: Product Design, Manufacturing, Maintenance and Inspection. For each of these phases, we take into account some specific activities for the AR support, considering the present limitations. Consequently, some of these low-cost devices is analyzed and integrated. In order to do that, new algorithms to use the embedded technologies in the industrial context are developed and evaluated.

Research in Product Development focuses on interaction with virtual models in an immersive system for design evaluation and in a system for planning and evaluating spaces. In the former case, a wireless device coming from the entertainment field is integrated in an immersive system to provide interaction with virtual objects. The system allows the user to manage the virtual objects by means of using both gestures and the device itself as a pointer. In order to do that, we developed some new algorithms to extend the features of the device. In particular, we exploited some embedded technology to detect its position in the space. In the latter case, the aim of the work is to deal with virtual objects in environments for space planning purposes. Since the environments are usually empty and wide in these cases, the user should structure the room in order to have a functioning AR system, which is a time consuming task. Thus, the system proposed relies on a hybrid tracking approach to avoid the time-consuming set up task and at the same time, to provide tracking in wide environments. The proposed hybrid tracking approach integrates a cheap and commercial mobile robot, which has been designed for household purposes. A new control system for the robot is developed in order to manage it as a mobile reference support for tracking.

We analyzed the AR as assistance during manual assembly tasks in fields of manufacturing and maintenance. In particular, we evaluated an AR system to help the operator when he manages some circuit boards by comparing it to the common operating manual. The proposed system directly superimposes information about the manual task to accomplish directly on the object by means of AR. In order to do that, a Natural Features tracking approach is developed to recognize the correct circuit board and detect its position in the space. Natural Features tracking is a technology that has been already used in some simple marketing and entertainment application, wherein it usually works with pictures. In this work we adapted it for industrial purposes, extending its characteristics to deal with an orthographic picture of the circuit board.

Furthermore, we developed a versatile AR visualization system for the inspection of objects state. The system can be adapted to work for different typologies of inspection and the data regards to the object can be broadcasted on devices located remotely. The application for AR visualization allows the user to visualize the data and it is addressed to work with mobile devices. In this way, it is possible to perform the inspection task by means of common devices, like tablets or smart phone. We validated the AR solution during the inspection of specimen in a fatigue-loading test.

Each solution that we developed by low-cost devices has been evaluated by comparing it with well-known technologies. Once we were sure of the correct functioning of our proposed solution for AR purposes, we integrated it in an AR application for Industry. Finally, we evaluated the effectiveness of each final application by means of users tests.

By means of these evaluations, it turns out that the integration of low-cost devices from mass market in industrial phases is feasible. The technology embedded in these devices allows us to reach technological performances that fulfill the industrial demands and it is also cheaper than other AR technologies proposed so far. Finally, these low-cost technologies can open the doors to new solutions for technical issues in AR, so that they allow us to extend or improve of the possible AR technologies at disposal.

1.5 Structure of the Thesis

The Thesis is structured as follow:

Chapter 2. It is a general overview of the AR technologies.

Chapter 3. We examine the background of AR solutions that have been developed for industrial applications so far.

Chapter 4. We analyze the industrial context, its needs and we describe the approach adopted in this Thesis.

Chapter 5. We describe a new point-based tracking system for a wireless game device.

Chapter 6. We describe a hybrid tracking system for wide environments grounds on a mobile robot.

Chapter 7. We report a natural features tracking system for circuit boards that it is able to work with only on an orthographic picture of the board.

Chapter 8. We evaluate the wireless device tracking system of Chapter 5 for product evaluation purposes in an immersive AR environment.

Chapter 9. We integrate the hybrid tracking system of Chapter 6 in an AR system for space planning.

Chapter 10. We developed an interactive AR system to support manual tasks with circuit boards and it grounds on the tracking solution of Chapter 7.

Chapter 11. We report the development of a system for inspection purposes.

Chapter 12. We draw the final conclusion and we show the further developments of this work.

CHAPTER 2

Augmented Reality and Related Technologies

Augmented Reality (AR) is a technology allows the user to experience virtual information mixed with the perception coming from the real world. It is one of the emerging technologies that are growing up since '90s and many application fields have been investigated or they are under research. Actually, this technology was born in the research field of Computer Science, but it has been rapidly spread in several application contexts, such as military, medical, industrial and marketing. The results obtained by its use are promising for its integration into everyone's life and their activities in the future.

Since this dissertation regards to using Augmented Reality in the field of Product Development, this chapter reports on the overview about this technology. Moreover, the main related technologies, which can be useful in an industrial context, have been here examined.

2.1 AR today

The first AR application has been developed by Sutherland in 1968 [148], but only from the '90s AR became an interesting research field, thanks to some progresses in hardware technologies and algorithms to manage and communicate data with devices. According to the last Gartner Hype Cycle, which provides a graphic representation of the maturity and adoption of technologies and applications, in Figure 2.1, AR is potentially relevant. Nowadays, this technology is still in a phase of overstated expectations and it has been forecasted that it will be bring into an effective utilization in 5-10 years.

However, today, it is possible to have experience of AR in every day life, especially related to the sense of vision. Actually, we are constantly bombarded on TV by digital information merged with reality, for instance in weather forecast or during a live sport events (Figure 2.2(a)). Even with other media, like newspapers and magazines, it is possible to have an AR experience. By placing the magazine in front of a webcam, we can see additional content represented on the screen of a computer, as in Figure 2.2(c). Moreover, the arrival of Smart Phones and Tablets led to mobile computing and consequently to Mobile AR. Thanks to this new kind of devices, the number of AR application greatly increased [72], especially in games and navigation supports as shown in Figure 2.2(b).

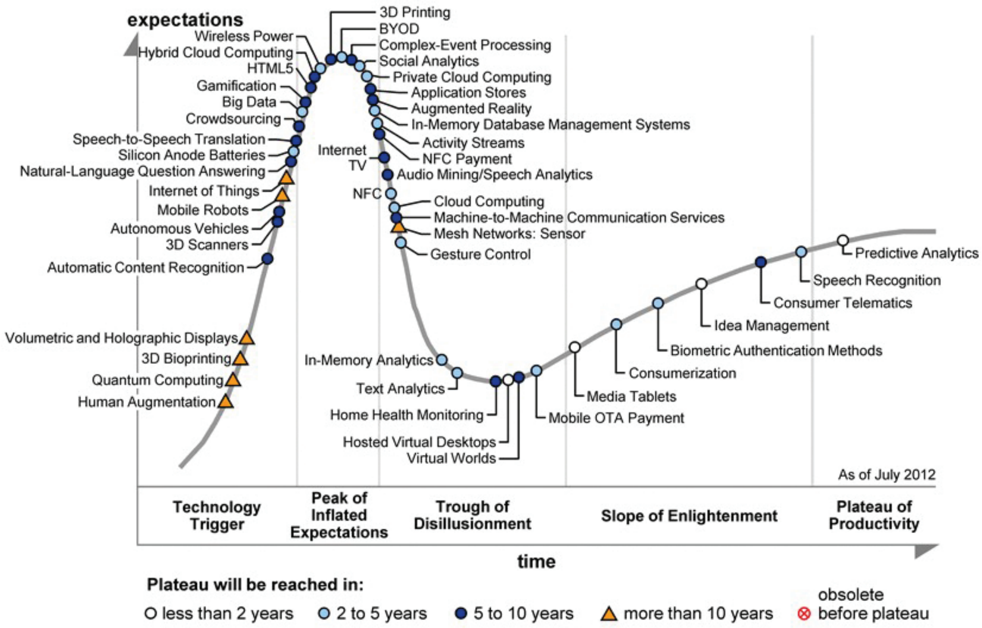
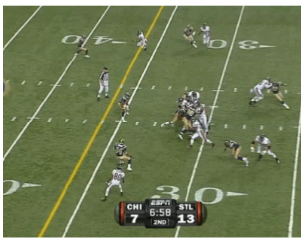
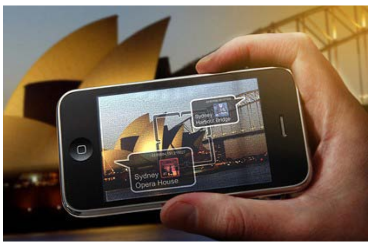


Figure 2.1: *Gartner Hype-Cycle of the current developing technologies as of July 2012.*



(a)



(b)

(c)

Figure 2.2: *Examples of AR in common life. In many TV programs, AR is exploited to convey additional information, for instance it can be used in sport events (a) to help the public to better understand the game. Additional information are achievable by means of mobile phones (b) and on magazines as well (c).*

2.1.1 The AR Community

Nowadays, the AR community is continuously expanding for what concern both research and the development of final applications for users. Currently, many conferences related to different interests have tracks about this topic. In particular, the first main important events have been workshops related with other ACM conferences as the Symposium on User Interface Software and Technology (UIST) and International Symposium on Wearable Computing (ISWC). These have been IWAR'98, IWAR '99, ISMR '99, ISMR '01, ISAR '00 and ISAR '01. Since 2002, ISMAR, the International Symposium of Mixed and Augmented Reality, is the main international conference about AR and the reference point about emerging AR technologies and social aspects of this growing community. Finally, Augmented Reality Event (ARE) is the main important annual event dedicated to AR for developers, industry and marketing since 2011.

2.2 Characteristics of an AR system

Any AR application should satisfy some necessary characteristics in order to provide an augmented experience for the user. According to [36], these are the characteristics:

- Combine real and virtual worlds in order to let the user to perceive them at the same time. In order to do that, sensors have to capture some pieces of information from the real environment and digitalize them. Then, the system merges it with virtual data.
- Give impression of coherence between the two words. Therefore, the system has to know specific characteristic of the environment, such as the object that are present in the space, or the position of the user (in particular his head and hands). This information can be provided to the system or it has to be estimated by sensors. Finally, these characteristics have to be taken into account during the synthesis of the Augmented Environment.
- Work in real time. The reality is continuously changing and the system should adapt to the augmented user experience. For this reason, AR needs to update the data coming from sensors and merge the data more quickly than the user perceptible capability.

These characteristics have to be supported by technologies that should work together in order to provide an AR output. The collaboration between these technologies is possible by a core software that manage them. As it is possible to see in Figure 2.3, the three main branches for technologies are:

Display technologies. They are the AR output and their purpose is to represent virtual objects in the real world;

Tracking technologies. They are systems that acquire data regarding to the environment, in order to understand the world and for the registration of object in it;

Interaction technologies. They are systems that grab the user actions in order to have coherency between the two worlds in real time.

In the next sessions, these technologies are described. For each of them, an overview is provided and pros and cons are discussed. Finally, the most important software platforms to manage these technologies are mentioned.

2.3 Display Technologies

An electronic device that is able to represent information is called display. In general, a display could work for at least one of the human senses. In the every day life we use these devices, such as headphones, that convey information by sound, TV screens, by images and sound, or mobile phones that combine sight, hearing and touch. As previously mentioned, only visual displays for AR are described in this dissertation, even if there are some other devices that are able to convey information by the other senses. Just to mention, haptic systems [42] are designed to provide the sense of touch and aural displays regards

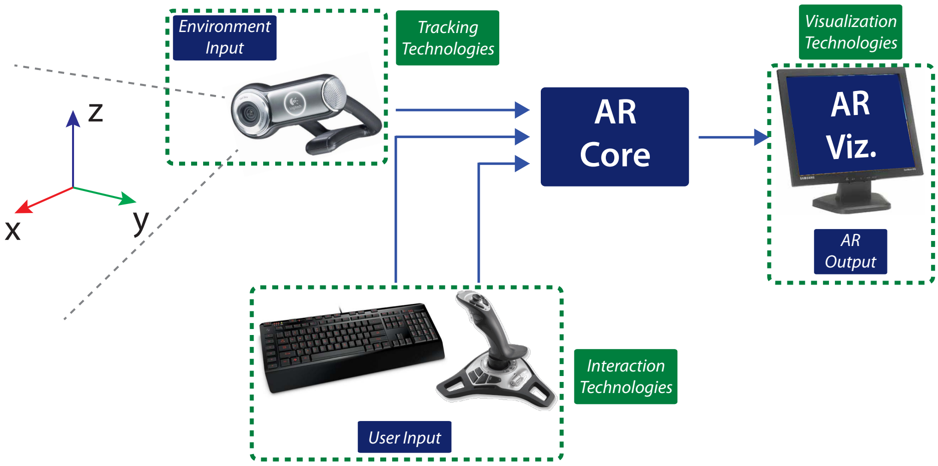


Figure 2.3: *The common framework of an AR system.*

sound, but their contribute in AR is still limited. Finally, there are some attempts of augment the sense of taste and smell by means of particular displays [112], but these technologies are still in the embryonic stage and have never left the lab.

Therefore, the adopted definition of display for AR in this dissertation is, according to Bimber,

> *an image-forming systems that apply a set of optical, electronic and mechanical components to generate images somewhere on the optical path in-between the observers eyes and the physical object to be augmented [46].*

In this section, AR display technologies are described and divided in a taxonomy by the location of the display regards to the user. In fact, the device can be located attached to the user body or placed in the environment and far from him. According to this taxonomy, it is possible to classify the devices in three groups:

- head-attached, if worn by the user,
- hand held display
- spatially aligned, when the device is not attached to the user.

Figure 2.4 shows these three categories.

Since the task of an AR visual display is to represent images, other two secondary features has been considered in the taxonomy. The first one regards the surface to represent the image. In fact, the majority of devices create images on a planar surface, like a computer screen, but there are other solutions where the image is produced on curved or generic surfaces. Finally, the last feature to consider is how the device mixes the real world with the virtual part. In fact the display has to manage two image sources: one related to the real world, the other one related to the virtual one. These fusion technologies are

video mixing The image is generated by the combination of a live video stream and graphics generated by the computer. This mixing is performed by a computer that shows the result on a screen.

optical combination A computer generates the graphic part for the augmentation and it makes visible the virtual content within the users field of view.

Each of the present technologies for visual augmentation involves some advantages and drawbacks that have to be taken into account for the final application. For this reason, a description of the devices developed is provided in the following sections.

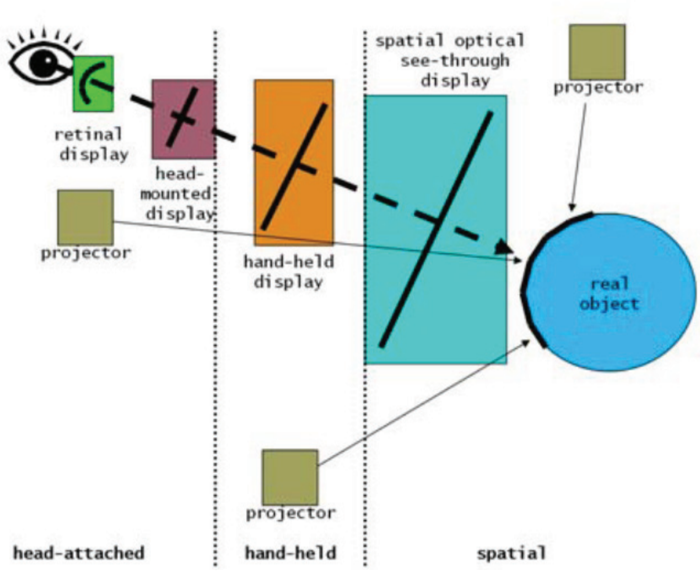


Figure 2.4: *Image generation for Augmented Reality displays [46].*

2.3.1 Head attached display

Head attached display is a visualization system for immersive AR applications that the user wears it on his head. Usually, this solution gives the user much freedom, since he can move free and in a natural way into the environment and his hands are free of any particular visualization devices. Moreover, this kind of display usually conveys an AR representation for each eye, providing stereo visualization as well.

According to the kind of surface where represent the image, there are three different kind of displays:

- Head-mounted displays;
- Retinal displays;
- Head mounted projectors.

Head Mounted displays

Head Mounted Displays (HMDs) are a technology inherited by Virtual Reality. According to Azuma [36], two kinds of HMD are present and they differ between them to the technique for superimposing the virtual part on the real world: *video see through* and *optical see through*. Video See Through Head Mounted display (VST-HMD) uses the video mixing technique in order to combine real world with the virtual one. As in Figure 2.5, the system captures the environment by mean of a camera mounted on the device, it elaborates the graphic part and merge everything in an image, which represented on a display close to the user eye.

Optical See Through Head Mounted Display (OST-HMD) makes use of the optical combination technique to generate the augmented view. Usually, this is performed by half-transparent mirrors or transparent displays, as shown in Figure 2.6. In this way, the user is able to directly see the real world with his eyes, while the digital information are overlaid.

Unfortunately, see-trough technology entails several drawbacks. According to [136] and [46] the main problems are described to follow.

Resolution The necessity to use miniaturized components in order to create a device that is light and not cumbersome is negative for the final image quality. In case of OST-HMD, only the virtual part is affected by resolution issues, while for VST-HMD, both real and virtual world are affected by the camera and display resolution.

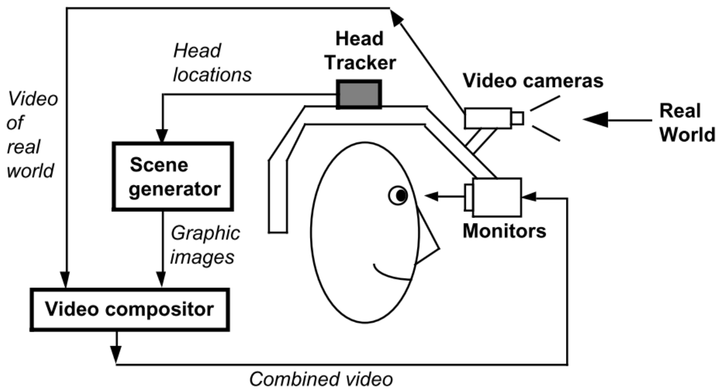


Figure 2.5: *Scheme of a generic VST-HMD.*

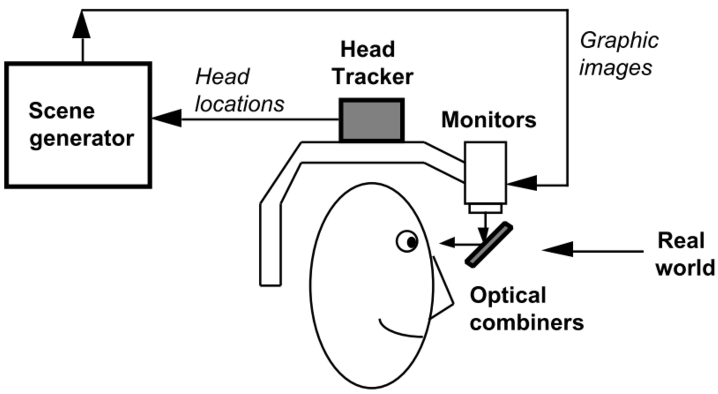


Figure 2.6: *Scheme of a generic OST-HMD.*

Field of View The Field of View (FOV) is usually narrower than the human one. In particular, the FOV is limited to the frontal view, while the peripheral view is never taken into account. This is due to a limitation in the optics and in the technique to convey images.

Perception Users that wear see-through HMD complain of perception problems. In OST-HMD, the virtual objects are always partially transparent due to the solution used to merge virtual and real world. For this reason the augmented part is often perceived as a ghost and it is possible to partially see through it. According to the first issue, OST suffers to a problem related to depth, in particular the occlusion problems between real and virtual objects.In fact, when a virtual object occludes the view of a real object, the user is still able to see the real object behind. This problem does not happen in case of VST, since the two worlds are merged by means of video mixing. Another problem related to depth is the eye accommodation to objects to different distances. In fact, human eyes adapt the focus according to the object distance, while, in case of VST, the camera focus is usually fixed. Thus, when the view is close, the user is not able to have a proper augmentation. In case of OST, instead, a proper eye accommodation is possible only for real objects. Moreover, since the virtual part is located on a different plane regards to the real world, the eyes is forced to see only one plane sharp at time.

View point matching Both the camera and the eye point of view cannot be the same in a VST-HMD, carrying to a shift in the view that it is easy to perceive by the user.

All these issues lead to a discomfort for the user, which he is not able to use an HMD for a long period. Fortunately research in the field of see-through HDM is going to overcome these problems by introducing new solutions. For example a big limitation was the weight and encumbrance of the devices since few years ago. Now there are solutions like [33] that has the same size of a normal pair of sunglasses. Finally, there are some experimental works to improve the perception on OST by using additional LCD display to block the light and increase the coherency of the virtual objects in the real world [91].

Retinal Display

Retinal display projects images directly on the retina of the user eye by mean of a low voltage laser [157]. This is a technology in its initial phase, but promising for the future. In fact, this solution provides a high detailed image with a large field of view, high contrast and high brightness. However, it suffers of some limitation, first of all the cost, since the mass production is not still available, the development of one display involves particular components and critical tolerances that require expensive fabrication techniques. Moreover this display supports for only one eye (no stereoscopic vision) and it does not still provide any adjustment for eye accommodation.

Head mounted projective display

The head mounted projective displays (HMPDs) reflect images directly in the user eyes by mean of a retro-reflective material placed in the environment and a couple of projectors mounted on the user head [83]. The first HMPD has been patented by Fergason in 1997 [62]. Compared with see-through HMDs, the major advantages of this technology are the lager FOV, an easier correction of distortion and parallax problems. However, these displays are not still applicable outside the lab. The main shortcomings are the encumbrance, the limited resolution and the need of a particular retro-reflective surface to work.

2.3.2 Hand Held display

Hand held displays provides a cheap solution for mobile AR visualization, compared to HMD. The biggest advantage of this kind of display is given by the wide diffusion in every day life, such as mobile phones, tablet PC, personal digital assistant (PDA) and ultra mobile PCs (UMPCs). Moreover, since the user holds the device in his hand, they are less intrusive than HMD, and it allows collaboration between users in an easier way [158]. However, the representation of images is at arm distance and it entails the impossibility of immersive experience.

Since the majority of hand held devices is already equipped by a screen and a camera, the common AR visualization is provided by the video mixing technique. However, solutions based on projection are present. In fact, the arrival of light mobile projectors led to hand held video projectors like the AR Flashlight [129].

The necessity of having a small and light device, which is possible to hold in hand, involves some limitations to this technology. These shortcomings are

- screen size, that limits the FOV of the augmented environment,
- low video quality, due to integrated and miniaturized cameras,
- limited performances regards computational power that require to develop optimized algorithms,
- necessity for the user to hold them by hands.

2.3.3 Spatial display

Spatial display is the last category and it is related to visualization systems that are not attached to the user. In fact, they are placed statically into the environment. This last kind of displays is suitable for large exhibitions, where several people are involved and limited interaction is required. According to the typology of merging and the surface where the image is represented, three main kinds of spatial displays can be identified:

- screen-based video see through display
- optical see through display
- projective display

In this section the main features of these three kinds of display are described.

Screen-based optical see-through displays

Screen-based optical see-through display is the AR visualization technology with the most cost efficient approach, since only a camera and a computer are needed. The computer joins the frames acquired by the camera with the augmented content and represent the result on the screen.

Since the AR experience is related to the screen, the two main drawbacks of this technology are the grade of immersion and the interaction. As a matter of fact, the dimension of the screen limits the FOV of the user, entailing a low level of immersion. Finally, since the screen is not related to the user, but fixed in the environment, this solution offers few possibility of interaction.

Spatial Optical see through display

Spatial optical see-through display allow the user to see the virtual part aligned in the real environment by using solution as projection on semi transparent mirrors, semi-transparent screens or holograms. This kind of display enjoys the same advantages and disadvantages of OST-HMD described in [136] compared to Video see-through technologies. Thus, it is possible to have an easy eye accommodation, but occlusion problem is still an open issue. Finally, the working volume for the AR visualization is limited so the user perceives an un-natural cropping of the AR part when he moves away from it.

Projective display

Projective displays do not deal with any particular surface to represent the augmented part, but they project directly on the physical object in the environment. This technology turns out to be very effective when it works on the visual properties of physical objects, like by projecting a new color or a new texture on its surface [130]. Furthermore, it is possible to visualize virtual 3D object in the environment, but it requires a more complex system. In fact, it is necessary to use more than one projector and to synchronize them in order to have a stereo image. Unfortunately, this modality is feasible only for one user, because it is dependent by his point of view.

Since this visualization technology is based on projection of light, a problem that turns out is the generation of shadows on the real objects. In addition, the projectors have only one plane of focus to a precise distance, so that the projection on objects to a different distance is blurred. Lastly, the field of view and resolution is limited to the size and distance of the object; they can be increased by using more than a projector, which increase the complexity and the management of the entire system as well.

2.4 Tracking Technologies

In AR, tracking is the technology that allows finding the relative position and orientation between two objects in the space. This is performed by means of algorithms that detect the relative pose by data from particular devices in the environment. If the user is able to move along, tracking is used to estimate where he is looking at in the space. Thus, the two objects under discussion are the user point of view and the fixed coordinate reference system in the environment. The estimation of the user point of view is necessary for the graphic algorithms of the AR application in order to correctly render the virtual part and consequently to have a proper alignment between the real world and the augmented content. On the other hand, the position can be also between the fixed reference system and a mobile object in the environment. In this way, tracking allows the system to use real objects for the interaction with virtual objects, as discussed in the Session 2.5.

The relative position and orientation between two objects is usually described by means of a matrix M, called roto-translation matrix. This matrix is a homogeneous transformation from a reference system to another one.The mathematical representation of a roto-translation matrix for a transformation in the space is usually described in a four by four matrix

$$M = \begin{bmatrix} R \mid T \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{12} & t_x \\ r_{11} & r_{12} & r_{12} & t_x \\ r_{11} & r_{12} & r_{12} & t_x \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{2.1}$$

In the Equation 2.1, R corresponds to the rotational components of the matrix while T represents the translation vector in the space.

Tracking is considered for AR one of the fundamental enabling technologies along with visualization. For this reason, research about this technology for AR is the most active [174]. Different tracking approaches are discussed in this section. In particular, the proposed technologies up to now are divided in three main groups of techniques: sensor based, vision based and hybrid tracking.

The main feature of a tracking system for AR is its functioning in *real time* [135]. The system must react in a synchronous way with what is happening in the environment and provide data about the pose in a short time. This time should be smaller than the user perception threshold, in order to have as much temporal coherence between virtual and real world as possible.

Other two important factors for tracking are *accuracy* and *resolution* [135]. The first one is related to the absolute error of the measurement of position and orientation. On the other hand, resolution is related to the noise that is present in the data provided by the sensor. Finally, it is worth noting that these two features should be considered both during static measurement and dynamic events. Error and noise result as an instable visualization of the virtual content [167].

2.4.1 Sensor Based Tracking

Sensor based tracking regards all the techniques that rely on sensors. Most of these techniques have already been fully developed and they are not still interesting in the research field. In fact, the majority of sensor based tracking has been explored until 90's and no big innovations have been carried out during the last years.

Since there are many physical phenomena in the world, many sensors are available today in order to detect them. For this reason, diverse sensor-based systems have been proposed for tracking in AR. According to the physical phenomenon, these techniques are:

Time of flight The position and orientation of a moving target are estimated by its distances to some devices placed in the environment. Each distance is achieved by calculating the time of flight, which is the time of propagation of signals from an emitter to a receiver. The emitter for this technology is usually on the target and the receivers are the devices in the room. Example of TOF techniques are by means of acoustic (usually ultrasonic) signals and GPS signals.

Inertial The principle of inertia is exploited to estimate the pose. In particular, the orientation is calculated by means of gyroscopes, while the position by a double integration of the data coming from tri-axial accelerometers.

Mechanical The mobile target is bound up with the reference by mechanical links and its pose is calculated by the angles among the links. The angles are usually measured by encoders or potentiometers.

Magnetic The pose of the mobile target is calculated starting from a magnetic field. In case of an electro-magnetic field generated by a coil, the orientation and distance of the sensor is calculated by flux produced by the combination of three coils perpendicular to each other. In case of magnetic field of the earth, instead, the orientation can be achieved by using a magnetometer.

All of the mentioned technologies have metrological features that enable them for AR tracking purposes. They are constituted by very small, compact and light devices, except the mechanical link technique. A common problem is the working range that it is usually limited, except for GPS and magnetometers that for this reason are mostly used for outdoor applications. Even inertial techniques are suitable for wide environments because theoretically they do not need any external reference point to work. However, they suffer from error that increases during time and they should be calibrated often. Moreover TOF and magnetic sensors can be affected by objects in the space. In fact, occluding objects can block the signal between emitter and receiver, while a ferromagnetic element into the working space can distort the magnetic field perceived by a sensor. Finally, air affects the magnetic field and the sound propagation, since it changes its properties with temperature and humidity.

2.4.2 Vision Based

Vision-based tracking technique provides the pose by processing images coming from the live video stream of a camera. This technique is now the most used in AR applications. Actually, many AR systems are already equipped with a camera because they perform video mixing to merge the real world and the virtual one. Thus a camera can be exploited for two different purposes: the augmentation of the visualization and tracking. Moreover, vision-based techniques are not invasive and not cumbersome solutions, because only a camera is needed. Finally, since many of these tracking algorithms are sufficiently precise even with low quality cameras, this solution does not require expensive devices and it can work with cheap cameras, such as webcams.

All the vision based tracking algorithms rely on three steps to provide the pose [100], as shown in Figure 2.7. The first one is the image processing, in which some computer vision operations are applied in order to extract some important information directly from the image. These pieces of information are then analyzed and matched with some known features previously stored, of which the system knows the position in the space. Finally, the correctly matched features are used to estimate the roto-translation matrix in the pose estimation step.

Most of the available vision-based tracking technique can be divided in two classes, mainly according to how they can get features from the image. The former is called *fiducial-based*, which detects features in the image by the recognition of particular object called *marker* in the environment. The latter is tracking by *natural features*, which works without any additional support for the image processing and it relies on features of the environment.

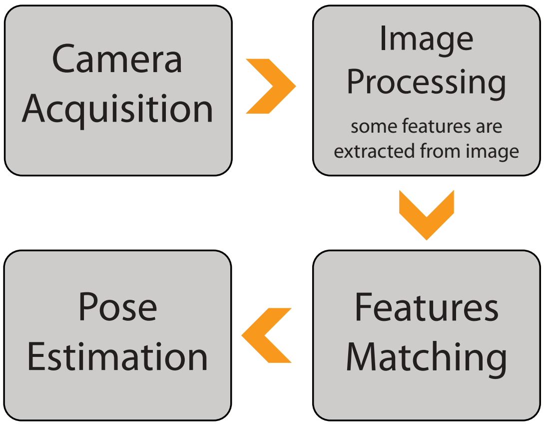


Figure 2.7: *The main step for tracking with a vision-based system.*

Once the coordinates of the features in the image are detected, they are matched with the related point in the space. The point in the image plane are described as the vector

$$a = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}, \tag{2.2}$$

where u and v are the coordinates in the image plane. The point in the space is represented by

$$A = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \tag{2.3}$$

where x, y, z are the three coordinates of the point in the space.

The final step for vision based tracking algorithms is the pose estimation, which means calculate the roto-traslation matrix M that satisfies for each feature find in the image the relation

$$a_i = K \cdot M \cdot A_i \qquad where \quad i = 1, \dots N, \tag{2.4}$$

where K is the calibration matrix [77] and N the number of points involved.

Fiducial-based Tracking

The use of fiducials, called also *markers*, has been the first approach to have a reliable and robust tracking for AR purposes. In fact, the addition of marker in the scene provides features in the images that are simple to detect.According to [100], there are two kinds of fiducial: *point* and *planar fiducial*.

Point fiducials are usually constituted by a pattern of spherical or circular elements so that they create geometry univocally recognizable. These elements are connected together in a fixed way, as in Figure 2.8. Since spheres and circular elements are represented respectively as circles and ellipses in the image plane, the identification of the centroids is feasible even with sub-pixel accuracy. Then, the pose estimation is straightforward.

The issue in this kind of tracking approach lies in the detection of the coordinates of the spheres or circles in the plane image. The common solution is to provide high contrast between the marker itself and the image background, in order to find easily the blobs corresponding to the fiducial. The blobs are detected by thresholding the image and then by applying some morphological operation on the binarized image. Afterwards, it is possible to detect the centroids of the blobs. Hoff et al. proposed a marker

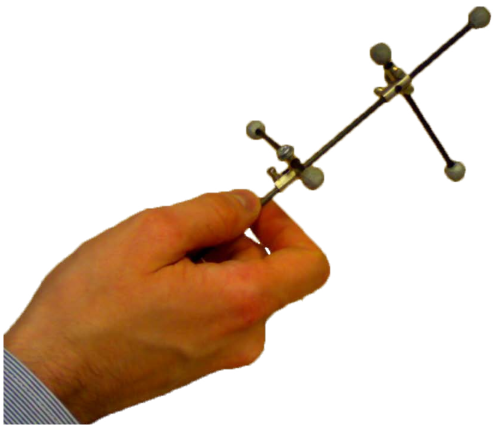


Figure 2.8: *A point fiducial marker.*

composed by two concentric circles one black and one white [82], called Concentric Contrasting Circle (CCC). This solution can be extended to use markers codified by using colors [146] or by particular black and white pattern that is the combination of CCC and DataMatrix [110]. Another very reliable solution is constituted by markers that are visible in the infra-red spectrum. These markers can be passive because they reflects light from a source, or active if they are the source of the infra-red light, like LEDs [111]. Thus, the detection of the marker is easy because it is possible to eliminate the background from the image by using IR-filters. Moreover, the detection is not disturbed by the illumination of the environment by visible light. Example of this kind of solution are some commercial products of Vicon® [31] and Advanced Real Time Tracking GmbH [5].

The second solution, *planar fiducial*, relies on black and white markers, as in Figure 2.9. Instead of finding the centroids of circles, this solution recognizes the four vertices of the square in the image by means of fitting straight lines around the borders of the marker (Figure 2.10). Then, the pattern inside the marker is analyzed in order to find the correct correspondences between the vertex a_i in the image and its position A_i in the space. Finally, the pose is estimated by the four vertices, minimizing the re-projection error.

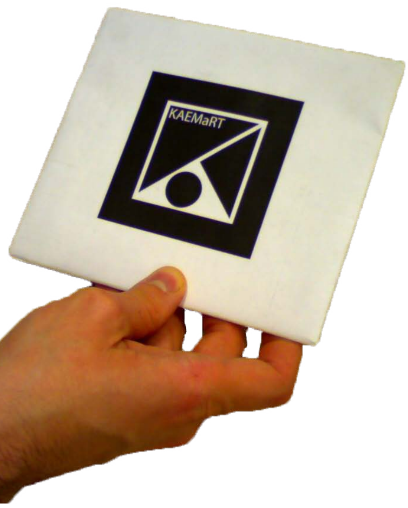


Figure 2.9: *A planar fiducial marker.*

The first attempt of this technique is [97], where blue and red squares inside the black marker are

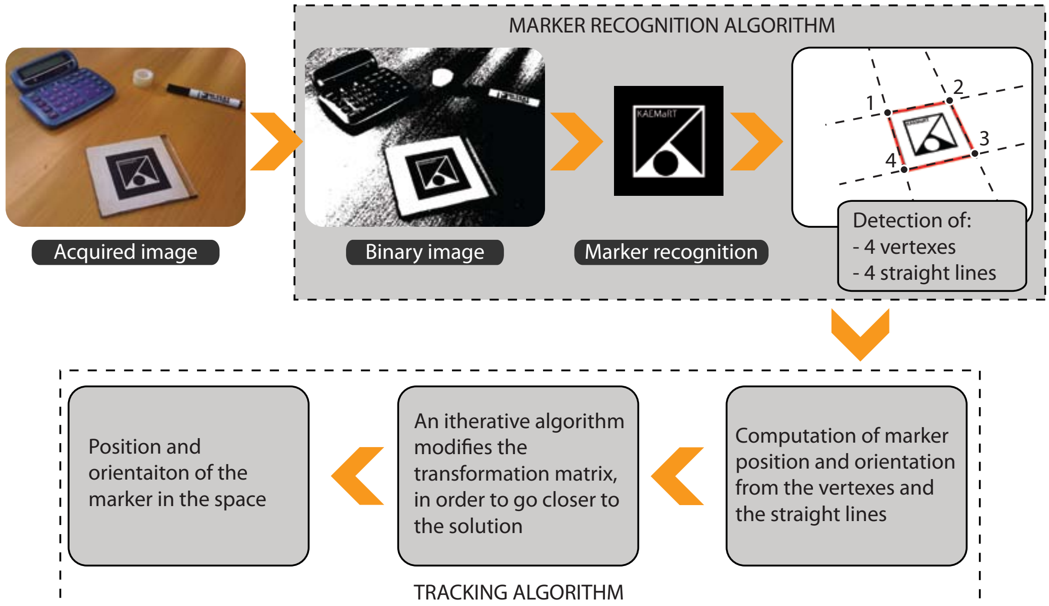


Figure 2.10: *General scheme of a marker-based tracking system.*

used to identify the marker. The identification of the marker is then improved by Rekimoto [134] by introducing a black and white pattern inside the marker. This last approach becomes popular by the introduction of ARToolkit [9, 88] that proved to be an effective solution. Moreover, ARToolkit library turned out to be robust to light conditions and its algorithms are not particularly intensive for CPUs. For this last reason, ARToolKit allowed AR applications by means of planar fiducials tracking to work with higher refresh rate than the previous solutions. In particular, real time applications (more than 20Hz refresh rate) have been possible. Finally, the distribution of ARToolkit since 2004 as public open source code made it one of the most important and famous library for AR tracking [174].

The detection and recognition of the pattern inside the planar marker can fail. In fact, since the pattern detection in ARToolkit is based on a template matching algorithm, sometimes the identification of the marker in the image is difficult due the image processing or false detections occur. In order to have a minor probability of misidentification, [123] proposes the creation of patterns by using Direct Cosine Transform (DCT), while [63] created codified digital pattern in order to have 2002 different IDs that the system is able to identify with a very low percentage of error. The use of digital codified markers has been optimized in order to be minimally onerous for the CPU in the library called ARToolKit Plus [162]. In this work, 4096 different markers can be detected and a new precise algorithm for the pose estimation, Robust Planar Pose Tracking (RPP), is developed [141]. Finally, the library has been optimized to work in real time also on mobile devices.

Natural Features Tracking

The use of marker is very effective because it provides a stable tracking solution, but requires engineering the working environment. The insertion of additional object in the space could entail some issues, since structuring the environment could be difficult to carry out, like in big or complex environments or in outdoor applications or since markers could be disturbing for the user. Thus, other solutions that do not deal with markers have been proposed. These solutions are able to detect some features in the image grabbed by the camera without any particular addition in the environment. The features under discussion can be points, lines, edges or textures.

The approach usually adopted for natural features tracking is by using a reference model. The model is knew a priori and it is recognized in the image. The model can be a 3D object, which is described by a CAD model [59]. As it is possible to see in Figure 2.11, the system detects lines and edges in the image and the pose is estimated by fitting the features with a digital model of the tracked object.

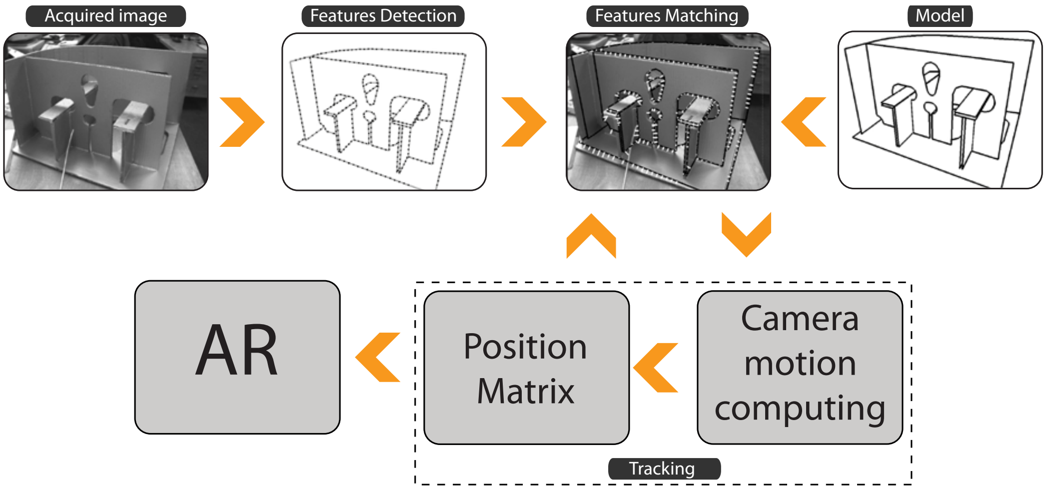


Figure 2.11: *General scheme of a model-based tracking system.*

Another kind of model can be an picture. The features extracted by this picture are local patterns differ from their immediate neighborhood [153] and they are usually associated with a change of image properties like intensity and color. These local features can be mono-dimensional, such as points, or described by small image patches. In order to find the pose, the extracted featured are then matched with some reference features. These ones have been previously chosen by an off line process that involves one or more images and then they are stored in a database [100].

It is worth nothing that finding correspondences between features in the image and the reference features is not a trivial task. In fact, some measurements are taken from features in order to compare them to the ones stored in the database. Typically, these measurements are taken from a region centred on the feature and collected in a descriptor. Since the extraction of features is performed by images from a video stream, the same feature can be taken from a different orientation, position and under different lighting conditions. Hence, the information saved in the descriptor has to be insensitive to rotations, light, scaling factors and noise.

Scale Invariant Feature Transform, SIFT, [103] and SURF [40, 41], Speed Up Robust Features are two local feature descriptor that worth mentioning. Unfortunately, even if they are very reliable, it is not possible to use them for real time application due to their high computational cost. They can achieve real time performances by parallelizing the image processing on GPGPU [56] or by restricting the algorithms effectiveness in favour of computational speed.

Real time performances are also achieved in [101] by means of feature classification. In fact, the matching of features is treated as a classification of a patch regards a class of views of the patch previously generated during an offline process. An example of application using this technique is called BazAR [11]. This approach is then further developed by the FERN classification, which uses binary features for the classification [125]. Finally Wagner et. At. developed a SIFT and FERN version for real time detection and tracking on mobile phones under some restriction and strong modification of the original algorithms [159, 161].

Finally, it is possible to track the camera pose even if a reference model is not present a priori. Actually, by means of a technique from robotics, Simultaneous Localization and Mapping (SLAM), the system is able to detect the camera point of view by features in the environment and, at the same time, increase the knowledge of the environment by mapping new recognized features. Klein et Al. developed PTAM [93], which exploit this method for a small AR workspace applications, and subsequently the method was extended on mobile phones [94] and on wide environments by using multiple maps [51].

2.4.3 Hybrid Tracking

Hybrid solutions require two or more different kinds of sensors work together in order to provide a more robust tracking. In this way, a hybrid solution can work even if one of the technologies involved fails because the other sensors are still able to get useful data for the pose estimation.

The main research branch for hybrid AR tracking is involved to increase vision-tracking systems performances by coupling them with inertial sensors. Actually, the vision-based approach is not able to estimate the camera pose when the system does not get useful information from the image, like in case of motion blur due to a fast movement. On the other hand, inertial sensors work even during fast movements, but they are prone to drift error if used for a long time. The join of both of the technologies is more robust because it uses vision for normal situations and at the same time it provides useful data to remove the drift to the inertial sensors. When rapid movements occurs, instead, the system takes advantage of the inertial sensors to estimate its position. Thus, [92] proposes a OST-HMD where the pose tracking is estimated by a camera and a gyroscope.

2.5 Interaction

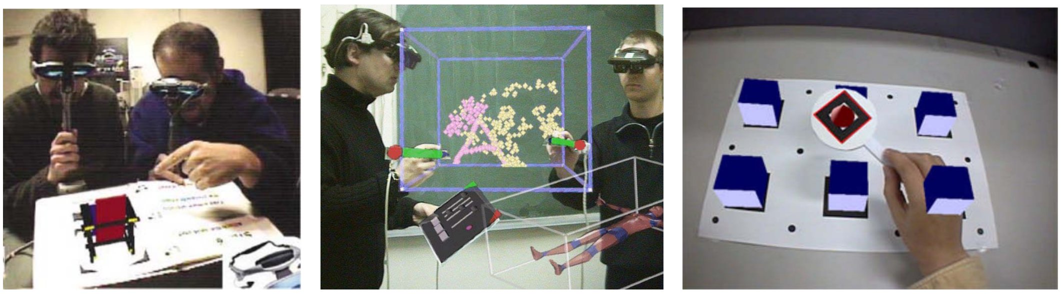
Since in every AR application the virtual content is located within the real environment, an AR system should afford the user to interact with both real and virtual objects. Research in the field of human-computer interaction is working to provide solutions to face with the virtual content in a way as simple, direct and natural as with real objects.

The common computer interfaces are often inadequate for AR purposes. The traditional interaction methodologies, like WIMP (windows, icons, menus, pointing) approach, are not always valid for AR because limiting the experience or not effective. Moreover, normal desktop interaction systems are usually designed to work on a 2D space, while in AR the user works in a 3D environment. In fact, the user does not deal with a pointer on a plane in AR, but with something more complex that usually requires managing up to six degrees of freedom (three for the translation and three for the orientation). Often, mouse and keyboard are consequently unacceptable in many AR application, since they are reductive for AR interaction and they are often encumbrance for the user.

2.5.1 Tangible Augmented Reality

The first solution for AR interaction is to manage and modify the virtual content by means of the physical manipulation of real objects This approach is inherit from the concept of Tangible User Interface (TUI), wherein objects of everyday life are the input and output devices for computer interfaces [86], and it is called Tangible Augmented Reality (TAR). The point of force of this solution is that the used objects are real, like a pen for example. Thus, they have an elevated affordance, since the user knows their physical limits and how to use them. Consequently the user has not to be trained to use the interaction device or he is able to learn it easily.

A first example is the Magic Book [45], where the TAR is a simple illustrated book, as shown in Figure 2.12(a). The pages of the book are provided with some markers so that the user can switch to different virtual contents by simply turning them. The augmentation is performed directly on the book, allowing the user to experience a sort of 3D pop up book in AR. A similar project is the Universal Media Book [74], where an interactive exploration of the human anatomy is performed by a book and a projective system. In the Studierstube project [139, 150], a pen is proposed for interaction as the natural extension in the space of the mouse cursor for a desktop application. As it is possible to see in Figure 2.12(b), the pen and a pad are tracked in the space by a magnetic system and they constitute the Personal Interaction Panel (PIP) [149]. PIP allows the interaction with the virtual content by means of 2D elements, like buttons and sliders, and 3D elements, which are new widgets proposed in the work. A paddle, instead, is used in [89] to move and manage digital objects on a table (Figure 2.12(c)). Some interactive metaphors have been designed to expand the interaction by the paddle. In particular, it is possible to pick the object, change it by shaking and then release it by tilting the paddle or delete objects by pushing them.



(a) (b) (c)

Figure 2.12: *Some solutions for Tangible Augmented Reality: (a) the Magic Book, (b) the Personal Interaction Panel in Studierstube and (c) interaction paddle.*

2.5.2 Finger Interaction

Another solution is the direct use of user hands to interact with the augmented environment. Actually, hands are the main means of interaction with objects in real life, so AR interfaces should allow us to perform free hand interaction with virtual objects. Moreover, this would enable a natural and intuitive interaction with both real and virtual objects at the same time. This kind of interaction requires systems and algorithms to track the finger position.

A first approach is to track the position of the fingers by attaching sensors directly to the user hand. Sometimes, this solution provides haptic feedback as well. A multi-fingered haptic interface device for AR is proposed in [163], where the user can manipulate a virtual object, which is visible in his hands, by means of a string-based haptic system. Another solution comes from VR applications, where gloves are used. These gloves are equipped with sensors to estimate the position of hand [147] and sometimes with some devices to provide haptic feedback. Zimmerman et. Al. released the first commercial glove in the 80s and it was called DataGlove [175]. DataGlove detects the position of the hand and its orientation by means of a magnetic tracking and the fingers flexion is perceived by ten light conductor stripes. However, even if this solution is effective and widely accepted in VR, turns out to be uncomfortable and cumbersome for many AR applications.

Another solution is the use of vision system to track the position of the hand in order to provide lighten gloves for the user. [58] uses punctual markers and a kinematic model of the finger to have interaction with the index finger for a chess game in AR. FingARtips [47] extends the interaction to a multi-fingered input device by placing some planar markers on a glove. The application is also able to recognize some gestures, such as grabbing, releasing, pointing, pressing and navigating and it is proposed for an urban planning system in AR. Finally, many applications avoid gloves by tracking only the position of the fingers by means of putting some colors easy to recognize directly on the fingertips, like in the SixthSense project of MIT [109]. In this project a mobile projective AR is proposed and the interaction is feasible by tracking four fingertips (two for each hands) by colored markers, as in Figure 2.13.

Vision-based techniques can also let the user to interact without the use of any devices and gloves. An augmented desk is designed in [137] [120], where the user can freely manipulate a virtual desktop. In this project, an infrared camera, which has been set up to the skin temperature, is used to find the user hands in the working environment and a template matching to detect the fingertips. [98] replaces the common planar marker to the user hand itself by means of a system that recognizes fingertips by adaptive color segmentation and geometry analysis. Wang et Al. [164] propose a 3D metaphor to interact with a virtual assembly by both hands. The pose estimation is performed by matching the silhouettes of the hands captured by two webcams to a database of pre-computed configurations. The use of depth sensors for interaction brought to new natural solution in the last years. OmniTouch [75] is a wearable system addressed to depict virtual information and objects on every day surfaces by means of a portable projector. A cheap depth sensor, coupled with an RGB camera, allows the user to naturally interact with virtual objects in Holodesk project [81] and MirageTable [43]. The former displays images by means of

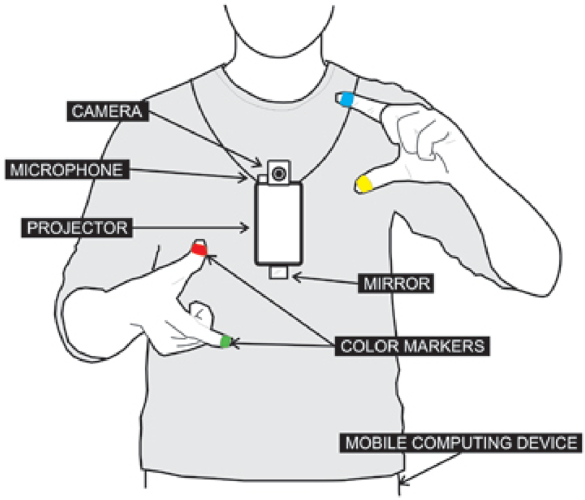


Figure 2.13: *The SixthSense Project from MIT. This one is a mobile projective AR system and the user can interact with the augmented part by means of his fingers.*

an optical see trough mirror and allows the user to grasp the objects. The latter uses a projective system to represent the virtual objects and allow the user to capture real objects in real time, in order to and have a *digital copy* of them for further interactions.

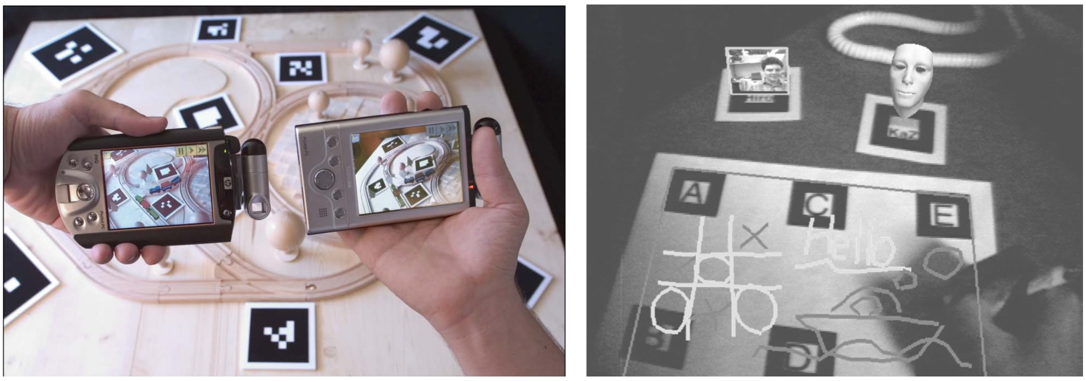
2.5.3 Collaborative Interaction

A weak point for main desktop applications is the difficulty of performing collaborative work. In the case of AR applications, users are located in the same place and, besides sharing the same environment and objects within, they can share the virtual content as well. In fact, TAR interaction methods can allow a collaborative co-located AR experience by using real objects as TUIs. Studierstube [150] is one of the first AR frameworks that supported collaborative work.

The advantages of computer-supported collaborative work (CSCW) in AR are present in [80] where they demonstrate, by means of a tennis game on mobile phones, that collaboration between users in an AR environment is more effective when the users can see each other. Another example of collaborative AR application is the Invisible Train [158] in which users can visualize a virtual train on a track and collaborate to manage the speed of two trains and the track switches (Figure 2.14(a)). The purpose of the game is to avoid collision between trains as long as possible.

AR collaboration is feasible between users located remotely, in order to enhance the telepresence. As shown in Figure 2.14(b), Billingurst et Al. propose a conferencing system that allows the user to visualize the people, which are involved into the meeting and located in other places, on planar markers that he can freely place in the environment [88]. Moreover the user shares a virtual working space where it is possible to put annotations. A more recent research extends the remote collaboration by sharing real and virtual objects in a common space by projective AR [87]. In this project, a test with some kids turns out that it is possible to have a natural remote interaction (playing games in this case) and it is possible to easily overcome the problem of sharing real objects.

Finally, the collaborative interaction in AR is even possible on multi-scales levels, in which users work from different point of views. In MagicBook [45], users can share the visualization of the book in AR but also each of them can visualize the virtual scene in immersive VR. In this way, the user that watches the scene in AR can see the others in the virtual world as animated avatars. A god-like interaction metaphor is developed by Stafford et Al. to provide directional information for an outdoor application [145]. Thus, an operator, which is remotely located, can use his hand to point locations or everyday objects as landmarks that the user can see in the AR outdoor application.



(a) (b)

Figure 2.14: *Examples of collaborative AR: (a) the Invisible Train and (b) a remote AR collaboration using ARToolkit.*

2.6 Software platforms

The technologies discussed need some software supports in order to be used in AR applications. Thus, several AR software platforms have been developed to facilitate the development of specific AR solutions.

These platforms have been usually created to manage only one technology. Actually, most of them are addressed only to the visualization, tracking or interaction. Only few of them deal with more technologies; Studierstube [30] was the first one able to do it. Finally, these platforms distinguish themselves from the others by the typology of distribution, since they can be Open Source, free SDK or with fee.

By means of these platforms, the user can develop the application by a high-level programming language. Usually, the programming is performed by a code language, but sometimes can be done by graphic methods (as Virtools [1]) as well.

The visualization, and the AR rendering in particular, ground on the OpenGL API [24]. Thanks to OpenSceneGraph [25], the programming level has been raised and it is possible to have an easy way to manage the visualization. OsgART [27], finally, takes advantages of OpenSceneGraph capability, focusing on AR rendering.

However, the majority of platforms is addressed to tracking. The one that is also considered the most important is ARtoolKit [9], which has been developed for marker-based tracking. This library has been written in C language and then it has been ported or improved to other languages. Examples of other marker-based tracking libraries are:

- FLARToolKit [17], for web applications
- FLARManager [16], for Flash Actionscript language
- SLARToolkit [29], for Silverlight
- NyARToolkit [23], for multiple languages (Java, C#, Android . . .)
- ARTag [6], in C++
- ARToolkit Plus [8], in C++
- Alvar [2], in C++, but with a plug-in for Virtools
- ARMES [4], for tracking of circular markers

For what concerns Natural Features Tracking, open source solutions for model based tracking are Bazar [11] and OpenTL [26], which is not directly addressed to AR, but tracking in general. Instead, ARToolKit NTF [7] from ARToolworks, Vuforia [32] from Qualcomm (only for mobile) and Metaio SDK [22] are solutions with fee. A more detailed list of the tracking libraries developed can be found at [21].

Lastly, there are some suites that directly manage visualization and tracking and for this reason they allow the development of AR application in a simple way, such as Build AR [12] from HITLab, DFusion from Total Immersion [14] and Layar [20], to augmented contend to print media.

2.7 Discussion

In this Chapter, the main features for an AR system have been described. An overview about visualization, tracking and interaction technologies has been carried out and the main software platforms for the development of an AR application has been taken into account.

Since many technologies supporting AR have been proposed, the developer should evaluate their pros and cons in order to choose the most suitable for the purposes of his final application.

According to the visualization technologies described in this Chapter, HMD systems provide the most immersive AR experience. However, they cannot be used for a long time since they are usually uncomfortable and because of their weight, which burdens the user head. Moreover, they still suffer from limited FOV, distortion and parallax problems. Other solutions, such as projective and spatial AR, release the user from wearing devices on his head and they usually provide a better visualization resolution and quality. However, they restrict the interaction and mobility of the user. Finally, the high computing performances of mobile devices, as tablets and smart phones, open the gates to the Mobile AR, but it results limited by the size of the displays and the interactive capability that it is usually limited to a surface or buttons.

Nowadays, several tracking technologies, which grounds on many different physical principles, have been proposed. Unfortunately, the best tracking system does not exist [167] and we have to choose the one more suitable for a particular AR purpose than the others, according to their pros and cons. Accuracy, resolution, latency time and cost are the most important factors for tracking technology. These factors should be under an acceptable threshold that varies according to the purpose and typology of the AR application. For example, using of AR in the medical field requires high features for the system, which could also entail high costs, while in video-games real-time, performances are often preferred to accuracy and the cost for development and deployment of the tracker should be affordable for a large scale production.

Since vision-based tracking technique is usually cheap, it is the most used solutions for pose estimation so far. In particular, the most robust are the marker-based techniques because the algorithms are well known and have been optimized in the last decade as well. For this reason, they usually provide a tracking with a good accuracy and a very low computational cost, which makes this solution feasible in real time for any kind of device today. However, this approach relies on marker and system fails whenever marker is not visible by the camera, like during rapid movements and occlusion. The problem can be partially solved by multiple markers or nested markers [152]. Natural Features Tracking allows the system to avoid partial occlusions and sometimes to work in wide environments. However, it usually needs an initialization process or a reference model, which its creation is usually a time consuming task, to work.

Interaction techniques proposed in AR application are usually influenced by visualization and tracking technologies adopted. Actually, different kinds of displays enables some interactive metaphors regards others. As said before, hands interaction in the space is used in immersive visualization, while simple finger interaction on a surface is possible in AR with mobile phones or tablets. The interactive purpose depends to the tracking system as well. Sensor-based tracking systems are very effective, but often uncomfortable. Thus, vision-based techniques have been proposed, but tracking of fingers and TUIs is still a challenge because of occlusion and the complexity of the object to track.

In the next chapter, some AR projects related to the industrial world are shown and their pros and cons regards to technologies adopted are discussed.

CHAPTER 3

Related Works in Industrial Augmented Reality

Augmented Reality has been introduced in the industrial field since '90s. The purpose of AR in this field is to provide additional information, numeric data, virtual objects or indications directly into the real environment where the industrial process takes place. This technology is addressed to anybody who is involved into an industrial process; operators, technicians in charge of management and maintenance of machines, engineers and managers. Since AR in an industrial context allows the user to see more than what reality can offer to him, this technology could change some paradigms in the industrial context. Thus, several investigations [113, 115] have been carried out on the field of AR in order to propose some new methods to perform tasks during an industrial process. AR should be able to ease the work of the user, by increasing the performances, improving the quality or reducing time and cost of an industrial process.

The first work on AR in Industry has been proposed by Mizell and Caudell in 1992 where AR is proposed as support of the operator during a wiring activity in the aircraft industrial context [52]. ARVIKA [69, 169] was the first founded consortium for the investigation of AR in the industrial field. This project involved several universities and companies in Germany, in particular from the automotive division [10].

The proposal of AR is meant to develop specific solutions to improve particular industrial aspects. Since industry is a very complex working environment, every industrial sector and product phase has its own features. Therefore, every industrial process has particular requirements and the applications have to be specialized to solve or support a specific challenge. Since it is not possible to propose a single AR killer application for industry, it is worth to divide them into different industrial areas. As suggested by [67], we divide the industrial areas that use AR according to the product and its own life-cycle such as:

Product Design. The creative process that leads to the conceptualization of ideas or needs into a physical object.

Manufacturing. The industrial process that creates a new product by starting from one or some goods.

Commissioning. The process after manufacturing where the final product is checked.

Inspection and Maintenance. The analysis of the condition of a product during its life.

Decommissioning and Revamping. The last process that a product is subjected to after its utilization life.

In the next sessions, we discuss the use of AR in these five areas and we provide some examples.

3.1 Product Design

The design of a new product is usually an iterative process that involves people with different skills and backgrounds. By starting from conceptualized ideas, the design is improved and modified as it goes along. However, it is often possible that some previous planning choices to be changed and consequently the design process to be readjusted. These changes in the design are usually due to factors related to costs, appearance of the final product, manufacturing necessities or ergonomics.

The design review, which is the process wherein changes are discussed and proposed, entails the collaboration of different people without the same background and knowledge. For this reason, the decisions are usually made by using one or more prototypes of final product. These prototypes can be real or virtual. Unfortunately, physical prototypes show some limitations; for example, they do not allow variants of shape and material, nor support interactive behavior. The production of the physical prototype is indeed costly and time consuming. Virtual objects allow the people involved into the design process to evaluate the product without making a real prototype and also easily propose various alternatives at an early stage.

The use of AR allows us to improve the perception of the final product, still in a virtual representation, by contextualizing it in the real world. Thus, the representation of a virtual object and the evaluation of its weak points turns out to be easier even for people without any technical background. Moreover, it is also possible to directly work on the virtual object in order to carry out some modifications in real time. In this way, the need of real prototype to evaluate the design is reduced by the use of virtual ones directly in the real world, shortening time and cutting cost to achieve a final product.

3.1.1 Evaluation

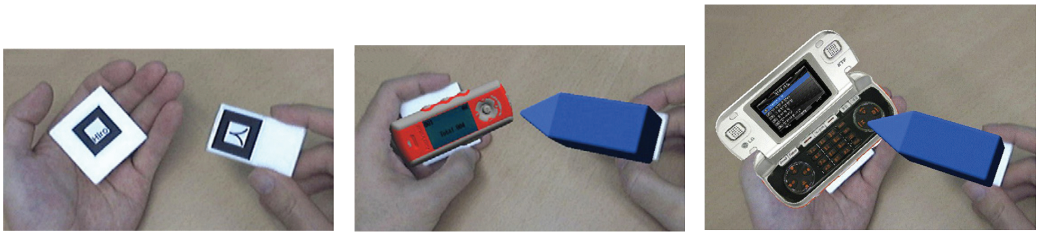
Since the primary function of AR is the visualization of virtual objects in the real environment, AR technology has been exploited to evaluate product design by representing virtual information and/or the virtual object directly into the real environment.

Aesthetic

Aesthetic evaluation has been performed in the field of automotive by representing a virtual model of a car or some of its parts. Fata Morgana project [95] proposes AR for the visualization of a virtual car in a show room. For this reason, an immersive visualization system has been used and the behavior of designer during aesthetic evaluation of a car has been analyzed. Fründ et Al. complete the visualization of a partial prototype of a car by superimposing the remaining part with virtual objects [70].

Functional Evaluation

TAR is used by Park et Al. to evaluate the functional behavior of electronic products [126] . As shown in Figure 3.1, the user, by means of a sort of stylus tracked by ARToolkit, interacts with the virtual representation of the object, which simulates the behavior of the real product through a state machine. A user study conducted with 10 testers shows better interaction performances with immersive stereo AR than the screen based one. Sidharta et Al. [143], instead, proposes TAR for a collaborative environment for the evaluation of mechanical products. In this work, the design review is performed by means of a HMDs with users located remotely. Regenbrecht et. Al. propose a solution for face-to-face collaboration during design review that is called MagicMeeting [131]. In this project, every user wears a VST - HMD and they share common virtual objects by means of some Personal Digital Assistants (PDAs). Moreover, it is possible to explore the virtual objects by means TUIs; for instance, the user can change the light setting, cut the object by clipping planes or add annotations.



(a) (b) (c)

Figure 3.1: *Interactive evaluation of electronic products in AR. In this work, a virtual stylus is used to interact with the virtual product.*

Data Analysis

Besides placing the virtual object in the real world, information about numerical analyses can also be visualized. Uva et Al. superimpose 3D model on top of the technical drawing and the user can see the results of FEM analysis on it [154]. Moreover, they implemented a framework in order to manage more than one user, allowing them to place labels on the virtual object and discuss about the changes by chat, as in Figure 3.2. Noelle, in a work within ARVIKA, proposes to superimpose the result of a FEM analysis directly on the real crashed object [117]. By means of a stereo augmentation, he is able to notice difference between the experimental results and numerical ones. In [104], the CFD analysis of an air diffuser in a room is visualized directly into the room itself by means of an HMD. Moreover, the authors propose the use of a tracked glove to interact with the representation of a CFD data.

Figure 3.2: *Example of a collaborative AR system for design evaluation. The system is able to project the virtual object on the draft paper and also show FEM analysis.*

Ergonomy

Finally, AR allows us to carry out ergonomic analysis by means of real or virtual humans. Balcisoy et Al. [38] exploit virtual humans in the real environment to perform ergonomic evaluation with real and virtual objects. Real users, instead, are used to evaluate the interior of a car in [119]. A reconfigurable driving seat is described in [48], wherein user can interact with the buttons and knobs on the car dashboard and he can also perceive a haptic feedback by them.

3.1.2 Modification, drawing and planning

The most important aspect that makes a virtual object preferable than a real prototype is its own versatility. A digital object is indeed easier and faster to modify than a physical one and does not involve any physical effort to be moved because it is weightless. Thus, it is possible to switch quickly from the evaluation of a version to another one or change some properties like color, texture or shape.

Modification

[127] proposes a reconfigurable AR prototype by means of some assembling body parts. The user combines the real parts in the configuration that he desires and then he can visualize the augmented objects on top of the assembly. Moreover, it is possible to change the color of the object and interact with it by a fingertip tracking.

The use of projective systems can be useful to modify the appearance of a real object. [130] makes use of two projectors to superimpose new colors and textures on some real objects. Verlinden et Al., instead, use a projector and a turntable support to visualize colors and textures on a rapid prototype object [156]. This approach has been recently developed by Extend3D in a marketed product [15].

Drawing

AR can be use to directly design a product directly in the real environment. Construct3D is a first attempt to represent primitives in the real world by using Studierstube [90]. Spacedesign is another solution for design that grounds on Studierstube [65]. In this project, an immersive system for concept creation and design review is developed, wherein the user can design free form and interact with digital objects by means of a large table display, a Personal Interaction Panel and HMD that are tracked by a point-based system. A cheaper solution is proposed by Cheok at Al. by using ARToolKit for tracking [53]. In this work, the user is able to draw with his index finger thanks to a planar marker placed on his hand. Another immersive solution based on markers is [142] that expands the design capability in AR by proposing advanced tool for drawing with a pen, such as the snap to grid in planar surfaces, and the multi-user collaboration. This last work has then been further improved by replacing the tracked pen with gloves in [116].

Usually designer likes working with real materials to design the shape of a final object and then it is digitalized for the following design steps. In [105], a TAR solution that creates the virtual object by modifying a real one is developed by sculpturing a block of foam with a cutting tool. Both of the foam and the cutting tool are tracked by a magnetic system. Thus, the virtual object is created in real time, without the need of a reverse engineering to digitalize the model. Information about how to perform a cut, texture and color are superimposed to the real object by a projector.

Planning

Product design means also design of space both for living and industrial purposes. Siltanen et Al. develop a marker-based system for the furniture layout in a house [144]. Once the user has taken pictures of the environment, he can superimpose virtual objects on top of them, controlling the position and orientation of the object and adjusting the light setting for the virtual content according to the real one (Figure 3.3). On the inside of the ARVIKA project, a similar system has been developed for the factory planning of an automotive industry [128]. A multi-marker tracking system from Metaio is here used to plan the disposition of machines and evaluate possible interference with the existing ones.

3.2 Manufacturing

Nowadays, the human support during a production process is fundamental, even if the majority of manufacturing systems are automated. Actually, technicians are needed to program machines and supervise

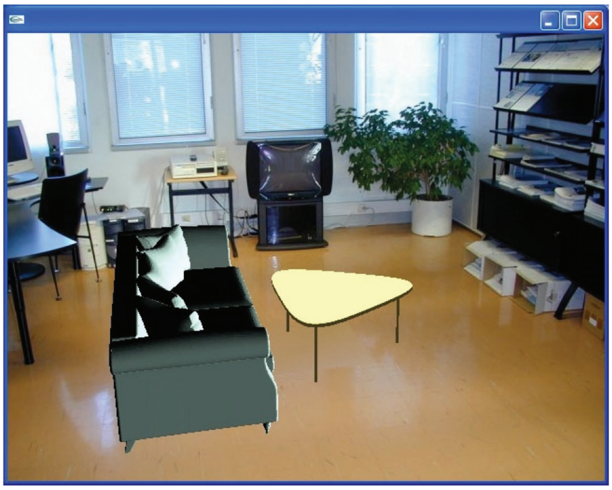


Figure 3.3: *Example of using AR for layout planning for a room.*

the work done by robots. If the process is not highly automated, trained or expert operators are required to perform assembly tasks.

AR can be a valid substitute of VR to simulate a productive process, to help the user to program machines and to provide useful information to the operator. In these cases, AR could be used as training for the operator or proposed as a supporting guide during the assembly. In these cases AR should superimpose some information or command about how to proceed with the assembly.

3.2.1 Automated Manufacturing

Robots can substitute humans in many repetitive and tedious manufacturing activities by performing tasks quickly and efficiently. Moreover, robots are useful for operations in dangerous environments. In order to work, robots should be programmed or controlled at distance. AR is used to interact with robots in an intuitive way both by programming and by controlling them.

Robot Programming

Robot activity planning is usually performed by using a programming language or, sometimes, by recording some configurations that the robot will get while functioning. This activity, done by users, has to be verified before preparing the robot for use. For this reason, the behaviour planned for the robot is often simulated in VR in order to avoid issues in the real environment, such as potential collision and damages. By using AR, the programming and simulation of the robot behaviour can be performed in the real environment.

Interactive system can be used to design the robot path. [172] and [133] use a point-based tracking system to draw the welding trajectories of a robot on an object. by means of a projective AR system, the welding lines are visible on the part and modifiable in real-time. [121] shows the optimized trajectories on a common display. The system describes the robot path with a parameterized line, which is estimated by a set of trajectories given by the user in a Bayesian Neural Network. Moreover, the system automatically checks possible collision during the robot work and rectify the trajectory.

Telerobotics

Robot motion at distance is not always a trivial task. The operator has to detect the position of the robot in the scene by one or more exocentric cameras. [54] simplify the motion of an end-effector, which is controlled by two joysticks, by augmenting the visualization of cameras. The virtual coloured axes, in particular, are represented in the views and are mapped to the joysticks movements. For what concerns mobile robots, the project TouchMe exploits a touch screen to move a robot and its arm in a room [78]. The user interacts with the screen, and he sees an AR preview of the position that the robot is going to reach.

3.2.2 Manual Assembly

The automated production line is not always convenient for companies, in particular for specific products that are not frequently produced or with a very small and specific market. For this reason, there are some assembly tasks that are still performed manually by people. In order to carry out manual assembly tasks, the operator requires training or a guide support, which are usually supplied by physical instruction manuals.

The use of traditional manuals can be overtaken representing assembly instruction and indications directly on the object by means of AR.

Assembly Guide

The first AR assembly proposal is by Caudel and Mizell for wiring activities into an aircraft [52]. Computer indicates the position of pegs which is used to assemble the wire by a see-trough display. AR support during a complicated assembly phase is described in [132], where an immersive marker-based system is used to visualize the steps to insert a door-lock into a car door. ARVIKA consortium proposes a system to support welding in the automotive field [60]. The project, called Intelligent Welding Gun, consists of an AR display mounted on a welding gun that guides the operator to the several wielding points of a car frame (Figure 3.4). The system estimates the pose of the gun by tracking point-based markers and it indicates the next welding stud by means of a compass metaphor on the display.

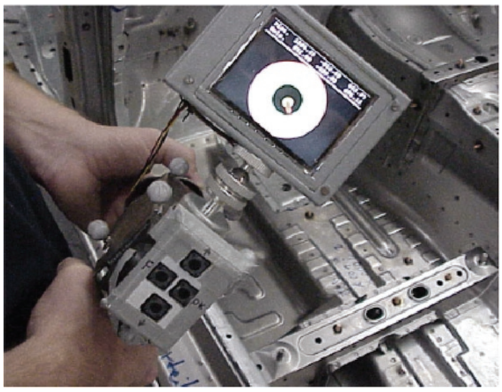


Figure 3.4: *The Intelligent Welding Gun proposed in the ARVIKA consortium.*

Assembly instructions are superimposed by means of a Virtual Interaction Panel in [171]. The user can interact with the virtual data using a tracked pen. [173], instead, developed an authoring tool to create assembly instructions for a marker-based AR assembly system. The use of mobile phones for assembly is explored by [44], where AR information is superimposed on static images taken by the phone. The authors sustain mobile phones are a versatile solution for AR assembly, which is always available and not intrusive.

The effectiveness of AR in assembly tasks has already been demonstrated with simple tests. A comparative test, carried out by Baird and Barfield with several visualization modalities for assembly instruction, shows that immersive AR guide is the most effective solution when dealing with electronic components [37]. In another comparative evaluation, Tang et. Al. measured the user workload during an assembly test with Lego® [151]. NASA Task Load Index [76] is used for the measurement. The results of a between subject test proved users that dealt with AR made less errors and showed a lower workload compared to others that worked with traditional assembly modalities.

Assembly with virtual objects

An AR system can also simulate the assembly of some components by means of virtual objects, instead of real ones. This solution can be useful to evaluate the feasibility of an assembly and the interaction among its parts or as a system to train the operators in a controlled environment.

In the AugmenTable project [155], a screen-based system allows the user to visualize and interact with virtual components of an electronic assembly. A stereoscopic tracking system detects the position of the user fingers in order to get a natural interaction with the virtual objects. [122] proposes a similar solution, but using an immersive visualization system. Finally, a system for arch welding training is proposed in [96]. In this project, the user visualizes the parts to weld in immersive visualization, a haptic system simulates the electrode and the sound of welding is synthesized.

3.3 Commissioning

Once a product is completed, it is checked during the commissioning phase. This process verifies if components and functionalities of a product meet the requirements. For what concerns small objects or goods with a wide market, commissioning is usually performed after the production line. In case of big products, for instance a plant, technicians do the evaluation. AR allows us to superimpose the original CAD model on the physical result achieved and to find discrepancies.

The project ARBA proposes a point-based tracking system to track the position of a visualization system in a factory [140]. In particular, the visualization system is composed by a display that is mounted on a trolley. Appel et Al. limits the discrepancy check on images taken in the factory [114]. By means of markers placed in the environment and several calibrated images, the system is able to correctly represent the technical drawing of the plant on the floor. In [68], non calibrated images are aligned to the 3D CAD model of the plant. Then, the system developed allows the user to visualize the virtual model of the plant superimposed on top of the images and navigate between them to compare the designed models and the final built items.

3.4 Inspection and Maintenance

In order to have a proper working condition of a product, its status should be periodically checked during its life. Moreover, maintenance operations, for instance replacing some damaged components, should be carried out. AR could be a good replacement to the common supervising approach and traditional repairing tasks. It could also let the user to visualize information about the status of the object directly on it, to visualize data about how to change a part of the product or have a remote support.

3.4.1 Data Visualization

Data signals coming from sensors placed on the object are usually represented on a display in a numeric or graphic fashion. By means of AR, it is possible to visualize these data in the working space of the product or directly on the product itself. Clothier et. Al. supervise a bridge with several sensor disposed on it [55]. The stress are then represented in a three-dimensional graphic visualization and superimposed on the real bridge, as in Figure 3.5. In [73], a grid of sensors on a wall captured the humidity distribution and the AR visualization is performed by a wireless communication system and mobile phones.

3.4.2 Maintenance task

One of the first suggestions of use of AR has been for maintenance tasks. In 1993, KARMA project (Knowledge-based Augmented Reality for Maintenance Assistance) allows the user to visualize graphical information for some simple operation on a laser printer, for instance replacing the toner cartridge and refilling the paper drawer [61]. The pieces of information are visualized by an OST-HMD and are coherently placed on top of the object and changed according to the position of the user or the task by an ultrasonic tracking system. Henderson and Feiner propose a comparative analysis for maintenance of

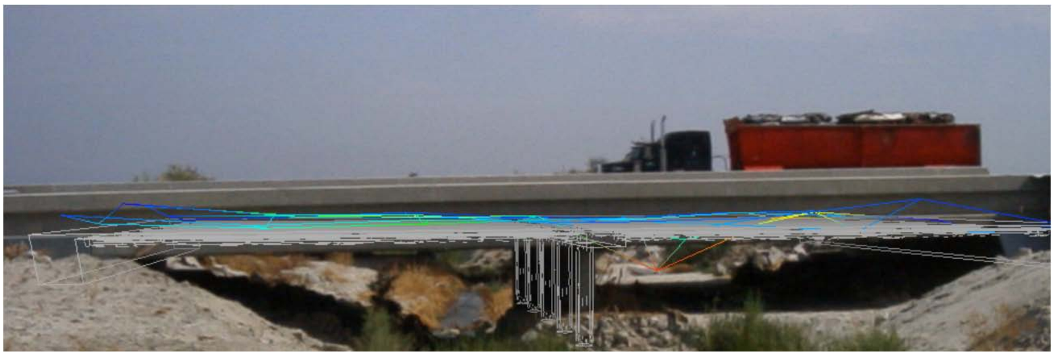


Figure 3.5: *Stress representation of a bridge in AR.*

a military vehicle [79]. The comparison is between three different visualization modalities: an immersive AR system with indications placed in the space, an HMD with only textual instructions and a LCD monitor. The result of the with-in subject test shows AR involves a reduction of time to accomplish the task.

3.5 Decommissioning and Revamping

When the utilization life of a product ended, it should be destroyed, disposed or recycled. This process turns out to be very expensive in case of big industrial plant. Moreover, it could be a dangerous activity due to possible high hazards during the dealing with components of the industrial structure. AR can be used as navigation tool, alert system for the operator and as decisional support for this phase by simulating how to dismantle or recycle big plants.

Ishii et Al. propose a multi-marker solution to track the position of the operator in a disused nuclear plant [84]. The AR system supports the user during the navigation into the site by means of a mobile device and it provides useful information about dangerous areas of the power plant for the user. In [85], the system is then extended by superimposing CAD models on images taken from the plant. This solution is exploited as evaluation support for the dismantling procedure in high-risk environment. Finally, Zokai et Al. developed a system for revamping wherein an existing component is substituted by a virtual one [176]. As it is possible to see in Figure 3.6, the simulation grounds on some computer vision algorithms, which are able to delete a real object from the scene and replacing it by a new virtual version.

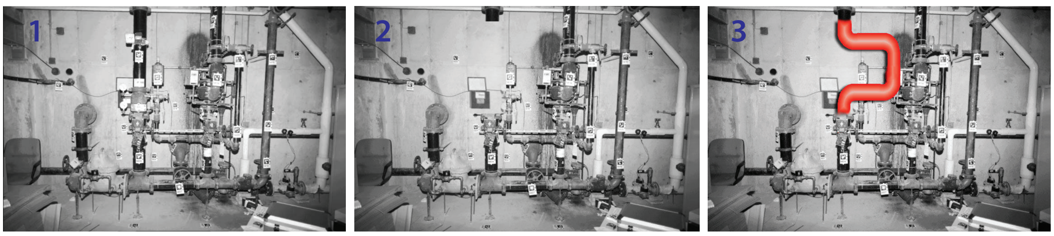


Figure 3.6: *The revamping process in [176]. From left to right, a picture of an industrial plant is taken, a pipe is removed from the image and then replaced with a virtual one*

3.6 Discussion

In this chapter, we presented an overview of AR applications for Industry. We analyzed the proposed AR solutions by dividing them into different groups according to the most important phases of the lifecycle of an industrial product. It turns out that each application involves several technologies, according to its purposes. The technologies used for an application is usually different to the others.

It is worth noting that the majority of the AR applications proposed so far has been developed to demonstrate the feasibility and the benefit of AR into an industrial context. These ones have not been further implemented because of technological issues. In fact, as also shown in the next Chapter, the industrial demands are very strict so that AR technologies rarely fulfill. However, it comes out AR is a valid support for industrial phases, since it entails a greater versatility and could reduce time and cost of a process.

CHAPTER 4

Research Context

The majority of AR applications that are distributed to final users are usually related to entertainment, marketing or digital events. These applications are generally simple and they allow the user to carry out only simple limited tasks. In fact, they are usually developed to show the real world some information, simple digital objects, animations and advertisement without any particular requirements or they have been simply developed with the purpose of impressing.

Many AR applications have been developed for the industrial field, but the majority of them remained in research labs [67]. This issue is likely due to the technologies used for developing the AR application, which does not fit well the industrial demands. Since Industry is a particular environment, technological limitations can bound the development of AR applications in this field. Otherwise, good solutions can be provided by high costs.

In this chapter, the industrial needs that an application should fulfill and the problems regarding to technologies for AR are analyzed. Afterwards, the research context of this dissertation is described by pointing out the issues taken into account and showing the method applied to solve them.

4.1 Industrial Requests

During the development of an AR application, the technologies used should abide by some specific requests from the final user. In particular in the industrial context, these requests become more pressing [115, 165]. Differently to application for entertainment, in which user can quit every time he gets bored, the operator in an industrial context should have support by AR. This means that AR applications should be designed by seriously taking into account ergonomics and usability aspects because often they are going to work continuously for a long period of time. Besides that, since the AR purpose is to positively change the way on how the operator performs tasks, it should simplify the operators work by making him easy to adapt to the new methodology.

For the mentioned requests, in order to design and build a usable AR application for Industry, researchers and developers have to take into account the following requests [113]:

Reliability. An AR application must provide a robust functioning and offer a minimum level of operation. As a matter of fact, the application cannot afford to stop providing the support because it could entail a drop in the performances in the overall process. Hence, the application should base on technologies that are stable in its entire operability field. The technological features of the system must be known beforehand, by evaluating the resolution, accuracy and real-time performance. Moreover, technologies could be paired together in order to have an alternative solution when one is not working properly.

Usability. Since the user works in an industrial context, he has particular demands. AR should provide some benefits in the user work, decreasing time and efforts to accomplish a task. For these reasons, the system has to be user-friendly. The user should be comfortable and feel safe when he interacts with the application and able to learn, set-up and customize it. Another aspect to take into account is AR replaces a methodology that was integrated in the working system and known by the user. Therefore, the integration of an AR application in an industrial process should provide continuity with the previous method. In this way, it is easier to accept the new approach and consequently learn how to use it.

Scalability. The application has to be easy to execute and distribute on a real industrial scenario. For these reasons, issues related to the installation, set-up of the application and maintenance of the technology must be considered.

Cost. The deployment of an AR application in Industry must be cheap. The cost should be justified regards to the financial expenses to support its deployment and to the economical saving compared to the replaced approach. For this reason, applications that rely on inexpensive development and technologies are preferred.

4.2 The Context

According to the reasons previously described, the development of AR solutions for industrial purposes is usually a challenging task. Most of the actual technologies do not fulfill the requests, in particular for what concerns reliability and costs. Nowadays, there are some reliable technologies that could be used in an industrial context. Unfortunately, even though they are robust tools for AR, they are expensive and often a technical background is required to use them.

Another solution is to adopt consumer technologies, as it has been experimented in this work. They are sometimes cutting-edge technologies, which are also robust, but they are not addressed to work in an industrial context. Since these technologies are developed for everyday purposes, they could not be reliable in a particular environment. Their limitations could turn out when used for industrial activities or it is impossible to use them as they are. For this reason, we should adapt their initial purpose in order to import them in an industrial context. The adaption phase of these technologies means changing the way we handle them, by a modification or a development of managing algorithms.

However, even though we integrate low-cost technologies in an industrial context by developing new managing algorithms, we are not sure about their performances. The capability of these technologies could not be adequate to fulfill the industrial requests. For this reason, we have to evaluate the technical characteristics achieved by the integration. In order to do that, we have to perform a comparison between the new proposed system and others that are well known in the AR field. Once we make sure of their performances match the industrial requests, we can use the adapted technologies to develop the final application.

4.3 Approach adopted

The approach that we used to explore low-cost devices for industrial AR can be divided in some steps, as it is possible to see in Figure 4.1. The first step is to evaluate if it is possible to use AR as support for a specific industrial activity. For this reason, we analyze the problems to tackle in order to have an AR support in the industrial activity; in particular these problems are related to the specific user needs during his work and the issues related to the industrial environment. In this first part, we also take into account the existing limitations of the current technologies, according to the state of art for AR.

After the feasibility analysis, we develop an *alternative technological approach* in order to overcome the limitations and consequently fulfill the industrial demands. For the development of this alternative solution, we take into account the integration of low-cost technological devices. These devices are modified for the new industrial purposes and some new algorithms to manage them are implemented. Afterwards, the new approach is evaluated in order to check its features by some comparative tests.

We use the new approach by means of the low-cost technology to develop the AR application support for the specific industrial activity. In particular, we integrate it into the final application. Then, the effectiveness of the application is validated by some tests. We perform tests in real industrial condition, or simulating the industrial environment, and we involve users to check or measure the performances.

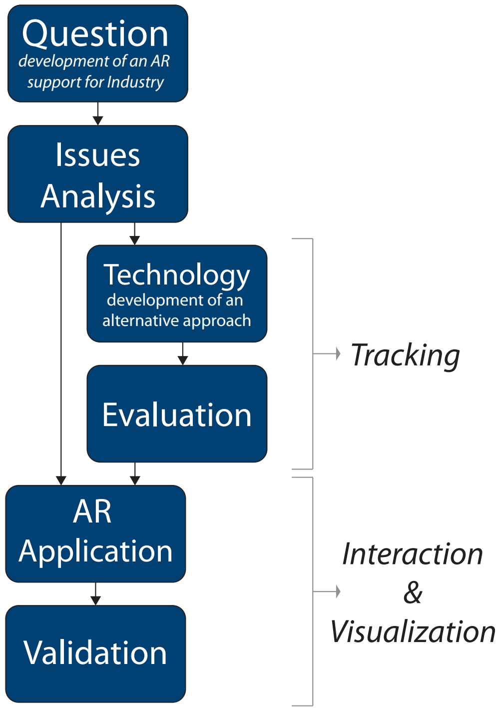


Figure 4.1: *Scheme of the approach adopted to integrate low-cost technologies in AR supports for Industry.*

The alternative solutions developed by low-cost devices in particular concerns new approaches for tracking. Actually, we consider the tracking process as fundamental for each AR application. If tracking is not robust enough for the purposes of the application, some drawbacks could affect visualization and interaction in the augmented environment. Consequently, a loss in usability occurs. We describe the three new approaches for tracking that we developed in Part II of this Thesis. The development of the applications, instead, focuses more on the visualization and interaction and it is reported in Part III. We developed these applications by exploiting the tracking approach previously described.

Part II

Tracking Approaches

CHAPTER 5

Point-Based Tracking for a Wireless Device

In this Chapter we describe the implementation and evaluation of a tracking system for a low-cost device. This tracking solution is then used in Chapter 8 for interacting with digital objects. The proposed tracking system relies on a the wireless control of the Nintendo Wii Console, the WiiMote.

The WiiMote is a very interesting tool that includes technologies in a single low cost device. Since its appearance on the market, many researchers have tried to exploit the WiiMote functionalities for different purposes [99, 138]. Similarly, we have decided to make use of its functionalities in order to make it a 3D input device dedicated to AR environments.

The motivations that led us to use the WiiMote are mainly related to the availability of different technologies integrated into one single and cheap device. The common use of the WiiMote is for navigating the interface on the TV screen, which is 2D, and as an interaction tool based on the accelerations imposed by the user. Instead, we aimed at using the WiiMote as a 6DOF device. To this purpose, we have developed an algorithm that allows us to detect the position and the orientation of the WiiMote by using its own sensors. The accuracy and precision of the 3D pointer was evaluated through a comparative testing session.

5.1 System Description

The WiiMote is the unofficial nickname of the innovative controller of the Nintendo's Wii game console. It is a device similar to a common remote control but that contains other technologies that allow acquiring information about its status.
As it is possible to see in Figure 5.1, the WiiMote is equipped with:

- 12 buttons,
- a triaxial accelerometer,
- an IR sensor,
- a speaker,
- a rumble.

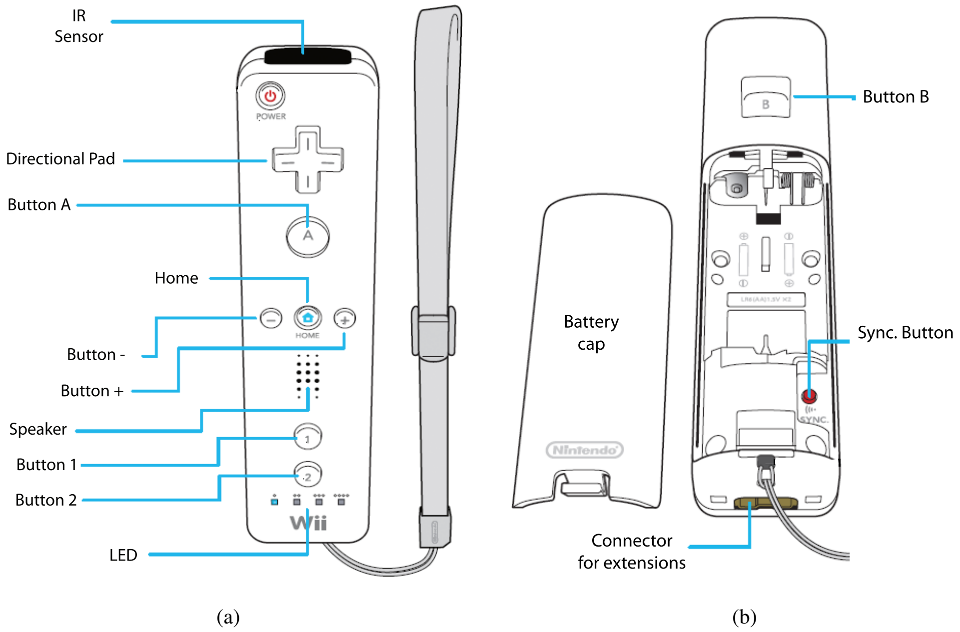


Figure 5.1: *The Nintendo WiiMote. The majority of buttons are located on the front part of the device (a). On the back of WiiMote, only the B button is present (b).*

The installed accelerometer is the ADXL330 tri-axial linear MEMS produced by Analog Devices [3] that allow us to measure an acceleration up to 3g. The IR sensor consists of a gray scale CMOS sensor with a resolution of 1024×768 pixel and a depth of 4bit for the luminous intensity.

All data coming from these sensors are transmitted through a Bluetooth connection. In order to limit the number of data transmitted, the WiiMote integrates a chip that recognizes the IR sources through the CMOS and detects the coordinates of their centroid. This chip, produced by the PixArt Imaging [28], integrates the PixArt Multi-Object Tracking engine that is able to manage up to four different IR emitters.

Unfortunately, Nintendo did not provide an official WiiMote Software Development Kit (SDK), but many open source libraries exist and allow managing this device by PC through the Bluetooth connection. Specifically, we used the WiiYourself! library [34] that is a powerful C++ open source library that enabled us to integrate the WiiMote in the final AR application.

5.2 Tracking

The tracking of an object in the space requires that all the six degrees of freedom have to be evaluated: the three spatial coordinates (X, Y, Z) and the three angles (yaw, pitch, roll). Normally, for tracking the WiiMote it is used an IR source, called sensor bar, wherein four IR LEDs are positioned in a row: two on the left side and two on the right side of it. The integrated IR sensor acquires the relative positions and dimensions of these LEDs, and the WiiMote transmits these data to the game console, which calculates the relative position and orientation of the WiiMote with respect to the sensor bar. So, the WiiMote can be used as a 2D mouse for navigating into the game console menu: just moving the device the user operates a change in the position of the cursor on the monitor display.

We cannot calculate the 6DOF of the WiiMote by knowing only the distance between the four IR emitters of the sensor bar because they are arranged on a straight line. Even using the accelerometer to evaluate the orientation of the device by means of detecting the gravity vector, it is not possible to evaluate the yaw angle.

In order to compute the spatial position and orientation of the WiiMote, we elaborated an alternative method, which it is described as following. This method works by using only data from the IR sensor.

5.2.1 Modification of ARTooolkit

In these years, many libraries, based on marker tracking, have been already developed with the aim of providing a simple and efficient way to develop AR applications. For this reason, algorithms already implemented has been taken into account to track the WiiMote. In particular, ARToolKit has been used for this work.

As already shown in Figure 2.10 at Section 2.4.2, ARToolkit provides the roto-translation matrix as output by means of information concerning the recognized marker and the 2D position of its four vertices in the image. However, since the WiiMote IR sensors are not able to detect this kind of marker, a new marker and a new recognition method is proposed here.

WiiMote processes every frame to directly detect the planar coordinates of the IR emitters centroids and it is able to recognize at most four of them, in order to send their coordinates to the PC by Bluetooth. Thus, we developed a new kind of marker consisting of four IR LEDs, which lay on a plane to form a square. These four LEDs were assembled on the same plane of an ARToolKit marker so that the two centers and the two reference systems coincide, as it is possible to see in Figure 5.2.

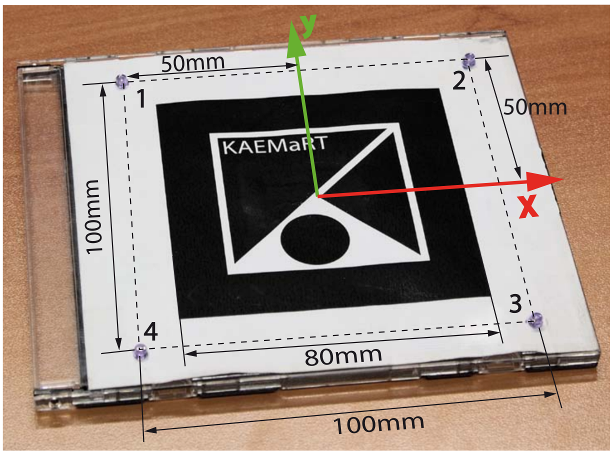


Figure 5.2: *The H-marker. The IR-LEDs are arranged to form a square, which is aligned with the ARToolKit's marker. The sizes of the two markers are superimposed on the picture.*

In this way, the same support provides both the pattern for camera tracking and the WiiMote. Since the marker can be used by two tracking systems, it has been called Hybrid Marker (H-Marker).

5.2.2 Labeling Process

ARToolkit usually perform the labeling process by means of the pattern inside the marker. Once the vertices of the square are detected by computer vision methods, they are arranged clockwise and labeled. The first point is always the one that corresponds to the angle at top left of the pattern. In case of H-Marker, instead, the four IR LEDs are arranged in a precise position to form a square, but the coordinates of the points, transmitted to the PC by the WiiMote, are not labeled.

Since WiiMote is not able to detect any pattern, an algorithm to order and label the points has been developed by means of geometrical considerations. An initialization procedure is proposed to solve the symmetrical ambiguities of the marker. This procedure involves placing the WiiMote in order to enable the IR sensor to frame the first LED at top left, as it is possible to see in Figure 5.3.

The labeling process starts when the user presses the button A on the WiiMote and continues during the tracking phase by comparing two consecutive frames. The algorithm recognizes a point in the new frame as the one located at the minimum distance from the equivalent point labeled in the previous frame.

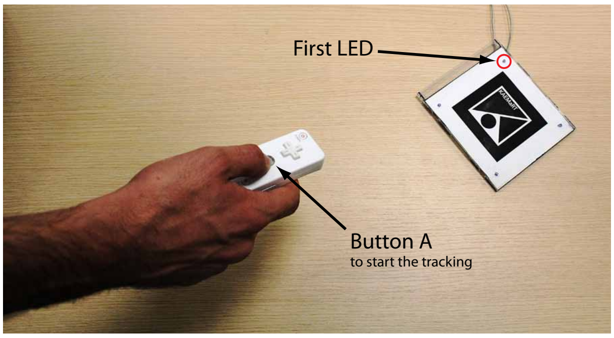


Figure 5.3: *The initialization procedure. Every time the user points to the H-Marker and pushes the A button, the system sorts and labels each of the four LEDs so that the first one is the one shown in the picture.*

5.2.3 Pose Estimation

Once the four coordinates of the LEDs have been ordered, these data are given to ARToolKit algorithms as input to estimate the pose. From this point, the system interprets the WiiMote marker as a common ARToolKit's marker and it is able to calculate the roto-translation matrix.

5.3 Evaluation of tracking performance

ARToolKit is a library that supplies a tracking system with good precision and accuracy for AR environments, it has not been evaluated if this library is suitable for WiiMote tracking and can provide correct tracking data. For this reason, we performed a comparison between the WiiMote tracking data and a camera tracking data.

We organized the testing sessions to compare two tracking data set: the first coming from the video camera and the second coming from the WiiMote. The peculiarity of these testing sessions was the correct repeatability of the test that we obtained by using a robotic arm. The robotic arm enabled us to impose to the marker the same trajectories, in every testing session, by allowing us to statistically assess the errors. To do this, we mounted our H-marker as it was an end effector of the robotic arm while the video camera and the WiiMote were arranged on a fixed tripod set in front of the robotic arm, as shown in Figure 5.4.

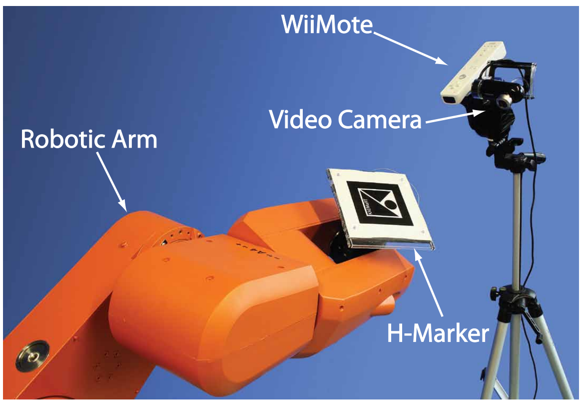


Figure 5.4: *The measuring system. The H-Marker is positioned on the end-effector of a robot. Both of two tracking systems (WiiMote and Camera) are placed on a tripod.*

5.3.1 Calibration

These comparative test sessions involved two different devices that were the WiiMote and a USB digital video camera with a resolution of 640×480 @ 30fps.

In order to obtain comparable values from the video camera and the WiiMote, before executing the acquisition phase, it was important to execute the calibrations of the two devices. In fact the two systems needed to detect the position of the marker with the same unit: we did it with a calibrating phase, where the intrinsic parameters of the two devices were detected.

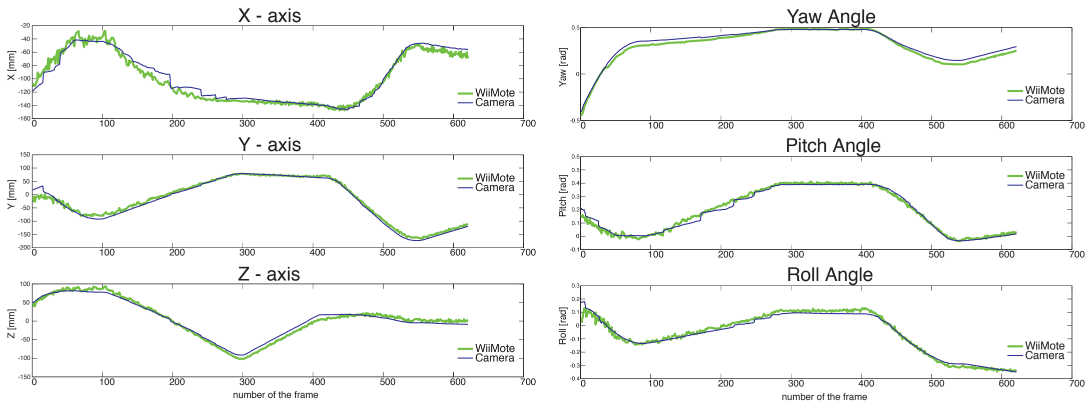
The video camera was calibrated using a tool developed for MATLAB [13]. This utility calculated the intrinsic parameters analyzing a planar checkerboard, which was framed by the video camera.

Instead, a different calibration has been made for the WiiMote, since it does not work with planar checkerboards. The calibration has been made through another code utility developed in MATLAB, which uses some considerations about linear projective geometry [77]. The algorithm is simpler than the previous one because neglects the lens distortion.

5.3.2 Testing session: accuracy and precision

The testing sessions consisted of moving the H-marker through two different paths (A,B), while the video camera and WiiMote tracked its location. Every path was repeated 15 times, in order to validate the acquired data statistically. These paths were created by imposing the robotic arm to move through seven points in the space. During the testing sessions, our tracking algorithm recorded the tracking data from the video camera and from the WiiMote synchronously, in order to simplify the data comparison. As described before, the video camera and the WiiMote were located on a fixed tripod and the distance between the center of the two sensors was about $5\,cm$. This offset has been deleted involving the first data acquired as reference system. After this procedure, the robotic arm moved the H-marker in the space according the two paths with two different speeds: $3.5\,cm/s$ and $7\,cm/s$.

The comparative graphs (figures 5.5(a) and 5.5(b)) show the data of a single testing session by presenting separated components of position and orientation.

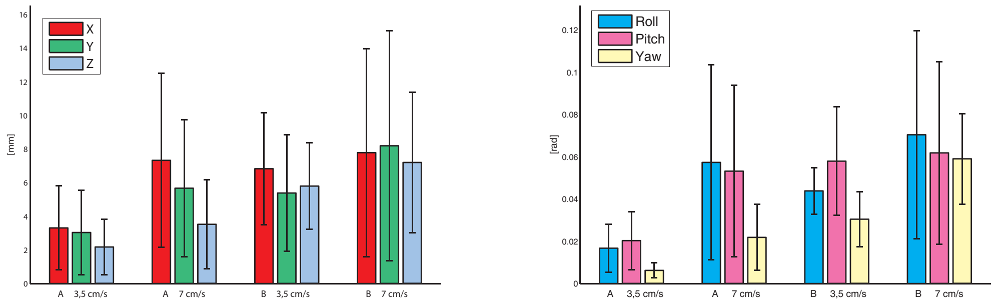


(a) (b)

Figure 5.5: *Comparative graphs for the position (a) and orientation (b) during a tracking test. Although there is some jitters in the trajectory of the WiiMote (in green), the tracking is similar to the one obtained by the camera (blue).*

After the execution of all testing sessions, we elaborated the data and estimated the average error values and the standard deviation of every session by analyzing the gap between the trajectories acquired by WiiMote and those acquired by the video camera, as shown in figures 5.6(a) and 5.6(b).

Finally we calculated the global error values, by estimating the accuracy and the precision of the position and orientation, as summarized in Table 5.1.



(a) (b)

Figure 5.6: *Average error of the WiiMote tracking regards to camera tracking data. The analysis has been carried out at different velocities for both the position values (a) and orientation (b).*

Table 5.1: *Global errors of WiiMote.*

Position

	X	**Y**	**Z**
Accuracy [mm]	6,31	5,57	4,67
Precision [mm]	4,30	4,22	2,76

Orientation

	Roll	**Pitch**	**Yaw**
Accuracy [rad]	$47,07 \cdot 10^{-3}$	$48,31 \cdot 10^{-3}$	$29,38 \cdot 10^{-3}$
Precision [rad]	$29,44 \cdot 10^{-3}$	$30,77 \cdot 10^{-3}$	$13,40 \cdot 10^{-3}$

5.3.3 Discussion

The tracking system proposed is less precise than the other marker-based tracking system used for the comparison. This is due to the precision of WiiMote to find centroids, since represents the position of the centroids as integer values. On the other hand, the standard ARToolKit algorithm elaborates the image in order to get the position of the four vetices with a sub-pixel approximation, describing them with float values. For this reason, the trajectories executed by the WiiMote are more uneven and the position is a little bit less stable than the ones obtained by a common camera tracking, as it is possible to see in the Figure 5.5.

However, the error obtained by the proposed tracking with WiiMote is limited and it is acceptable for AR purposes.

5.4 Conclusion

In this Chapter we described the development and the validation of a tracking system. The solution employs a commercial system for game console, the WiiMote, and extends the features of this device in order to detect its position in the space. For this purpose, we developed a point-based marker system, which is partially based on ARToolKit algorithm. The marker is composed by four IR-LEDs, which are disposed to form a square and are captured by the WiiMote. The system is able to order and label the position of the four IR emitters and then estimate the pose of the device by using the ARToolkit library.

The proposed solution has been evaluated by comparing the tracking with the one obtained by a standard ARToolKit system, composed by a camera and a black and white marker. Even though the system described is less precise, is acceptable for an AR interactive system.

CHAPTER 6

Hybrid Tracking for wide environments

This chapter describes a hybrid approach for tracking in wide environments, which is then used for space planning activities in Chapter 9. The purpose of this work is to provide a solution to estimate the camera pose in a empty and wide indoor space. Today, there are some technologies that can provide this kind of tracking, but they are usually expensive or they require long time to set up or for initialization. For example, a good (and cheap) solution is by means of multiple markers, which are placed in the environment. The system works if the camera is able to frame at least one marker. Unfortunately, the structuring of the environment is required. Baratoff et. Al. propose a solution to speed up the set-up process [39], but it is still a time-consuming task.

The hybrid system that we describe here uses a mobile robot for tracking purposes. The role of the mobile robot is to co-work with the user for extending the working space of the marker-based tracking technique. The robot follows the user's movements during the exploration in the AR environment and updates the position of a fiducial marker, which is fixed on it. The robot is automatically controlled through the device used to visualize the AR scene.

The use of a mobile robot in AR environment has been investigated by several authors for different purposes, as also described in Section 3.2.1. Other works describe how AR techniques can be used to automatic control a mobile robot [118] even able to manage a robot swarm [64]. However, in none of these works the use of the robot tracking, meant as a support of AR tracking, has been investigated.

6.1 System Description

The hybrid tracking system consists of hardware and software components that have been developed and integrated with the aim of extending the working space of the AR marker-based tracking approach. In fact, the approach allows the user to exploit only two markers to track a working area instead of using a standard multi-marker solution. This section describes both the components, which have been used, and the global functioning.

The hybrid tracking system is mainly set up by three modules, as shown in Figure 6.1:

1. Server, which is the module for managing the wireless communications through the devices involved into the system.

2. Robot Client, which integrates a mobile robot, a fiducial marker and a wireless communication module;

3. AR Client, which is the module used for the visualization and interaction with the AR environment.

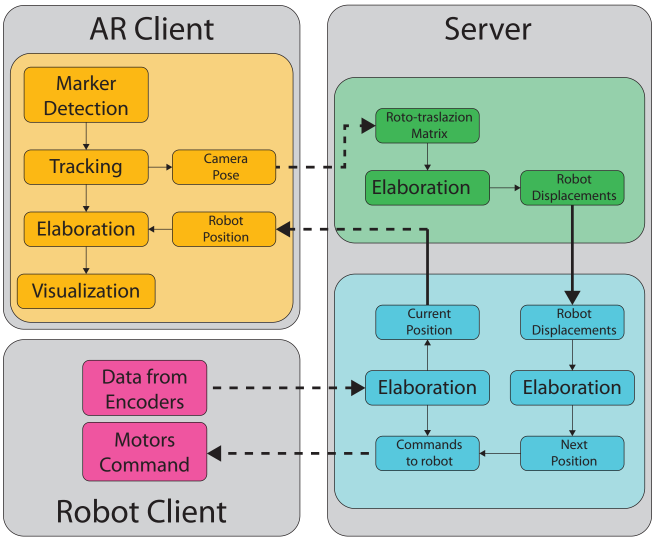


Figure 6.1: *Architecture of the hybrid tracking system. The arrows represent the data flow within the modules. The dashed lines indicate the data flow managed through wireless communication.*

6.1.1 Server

The Server is the module that manages the wireless communication between the AR Client and the Robot Client and elaborates the data coming from them. The Server gets the tracking data from the AR Client and the position of the robot from the Robot Client. By means of these data, it calculates whether and how much the robot has to move and it consequently sends the motion commands to the Robot Client. In this way, the camera always frames the robot correctly.

Moreover, the Server Client estimates the position of the robot in the environment. As explained in the following sections, this is a fundamental task for obtaining the camera tracking within the space. These position data are sent to AR Client for a correct registration between the real world and the virtual contents.

6.1.2 Robot Client

The Robot Client runs on iRobot Roomba 560 [19]. This device is a commercial mobile robot, which is normally used as automatic cleaning machine. The characteristics of this robot are that it can be programmed by using third-part programs and it can be controlled thanks to the I/O serial port mounted on its top. Then, the internal and external status of the robot can be continuously monitored.

The movements of the robot are managed by a differential drive, which controls two driving wheels, while a mechanical bumper and an infrared sensor enable the robot to avoid obstacles and to safely move within the room. In particular, two magnetic encoders, which are placed on the two driving wheels, provide the odometric data that are used to estimate the position of the robot in the environment.

In our system, the robot is controlled at a constant speed of $150mm/s$ by a continuous checking of the status of the two encoders. The RS-232 cable has been replaced by using wireless transmission that is carried through two Xbee modules [35]. One of these modules has been mounted on the robot, while the other one has been placed on the computer that runs the Server (see Figure 6.1).

The Robot Client cooperates with the user in defining the camera pose in the AR environment. For this reason, the Robot Client continuously sends the values coming from the two magnetic encoders to the Server and, as shown in Figure 6.2, a fiducial marker has been mounted on top of the robot.

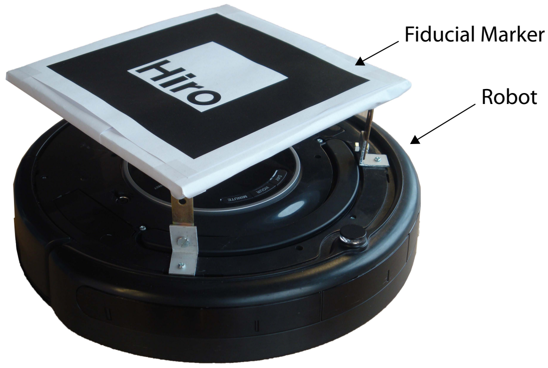


Figure 6.2: *The Robot Client. A marker is present on top of iRobot Roomba.*

6.1.3 AR Client

The module called AR Client enables the user to see and interact with the AR environment and the Robot Client. The hardware of this module mainly consists of a laptop and an USB camera. The software application, which runs on the AR Client, concurrently manages the following activities:

- markers detection,
- communication with the Server,
- camera pose estimation,
- rendering of the augmented environment.

The markers detection module enables the AR Client to calculate the position of the two markers in the scene. These two markers are the Robot Client marker in the scene and a Reference Marker, which is used to define a fixed reference system in the environment and correct the robot drift. The detection has been implemented by using the ARToolKit Plus [162] open-source library. By means of wireless TCP/IP connection, the AR Client sends the positions of the two markers to the Server module, which can manage the robot position in this way. As shown in Figure 6.1, the AR Client asynchronously receives the position of the robot from the Server and, finally, calculates the camera pose according to the absolute reference system, as explained in the next session.

It is worth noting that the AR Client is able to estimate the camera pose according to the absolute reference system, although the reference marker is not visible. The AR Client, in fact, combines the marker tracking with the robot odometry to evaluate its point of view.

6.2 The hybrid tracking system

In the proposed tracking system, the AR Client is able to calculate the position of the camera in the space in two different ways, as shown in Figure 6.3. In our approach, a Reference Marker is used to define the absolute reference system and can be located everywhere inside the AR scene. If the Reference Marker is visible, the camera tracking works in a standard way. However, the hybrid system is able to estimate the camera pose even if only the marker on Robot Client is visible, by means of a hybrid approach. In fact, once an absolute reference system is defined by the Reference Marker, it is possible to combine two measurements to reach the same purpose. In particular, the two measurements are the position and orientation of the robot in the absolute reference system with the pose of the camera obtained by tracking

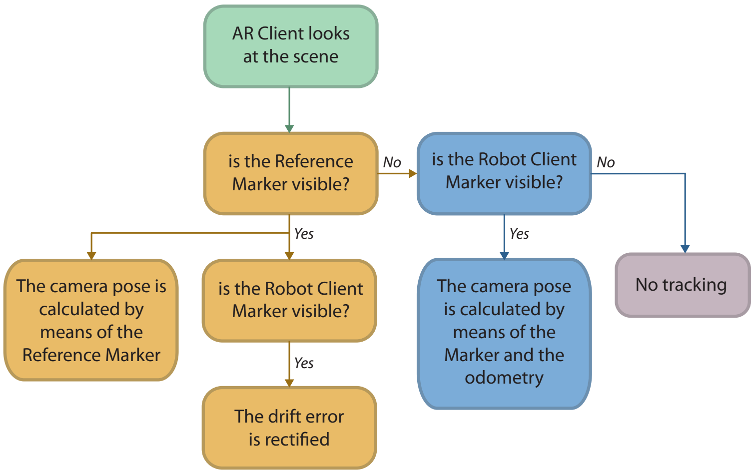


Figure 6.3: *The workflow of the hybrid tracking system. The proposed solution works even if th reference marker is not visible by means of a robotic support.*

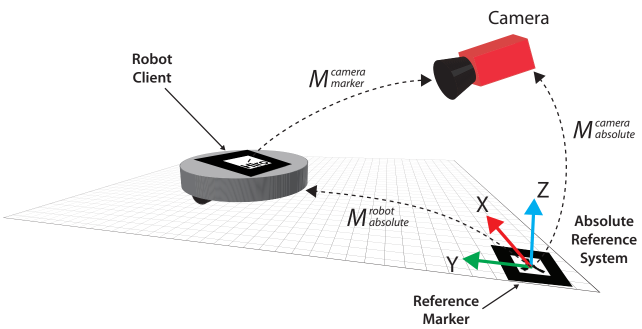


Figure 6.4: *Schema of the absolute camera pose estimation. It is possible to calculate the camera position and orientation by combining the position of the robot and the tracking of the marker on top of it.*

Robot Client, as shown in Figure 6.4. This second modality gives us the possibility of working with a mobile marker in a large space, thus avoiding the use of multiple markers or a structured environment.

The integration of two tracking technique (vision and odometric) carries out some issues thought. In fact, Robot Client calculates its own location by an incremental method and its reference system coincides with the position in which it has been activated. The initial point could not be overlapped with the absolute reference of the camera tracking. Moreover, the position could be affected by error due to drift. For these reasons, it is necessary to carry out a calibration of the system, when it is activated, to estimate the initial position of the robot relative to an absolute reference system, and correct the drift error during the use of the tracking system.

6.2.1 Automatic robot displacement

Since the main purpose of Robot Client is to provide a mobile support for the marker, we developed two different modalities for its handling. The first one is the manual modality, since the user can control the robot position my means of a GUI.

The second modality is totally automatic and managed by the Server. In this way, the robot is moved so that the marker results to be always centred by the camera. AR Client sends the tracking data to the Server that calculates the new position that the robot has to reach with a frequency of 1Hz. As shown in Fig. 6.5, the position that the Robot Client has to reach has been calculated as the intersection between the plane in which the marker moves and the straight line given by the optical axis of the camera.

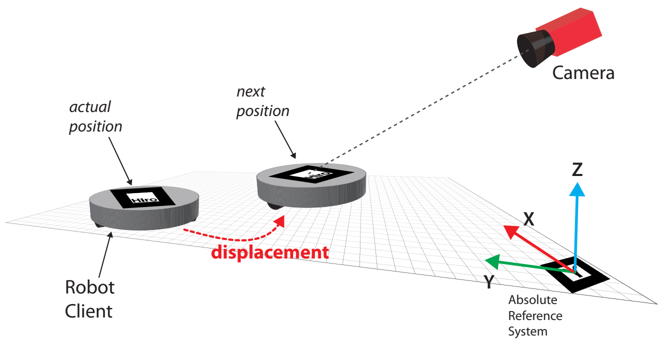


Figure 6.5: *Automatic robot positioning. The new position for Robot Client is calculated in order to be at the center of camera field of view.*

6.2.2 Calibration of Robot Client initial position

Robot Client identifies its initial location with the origin of its own reference system, so it is not possible to directly exploit its position for the hybrid tracking. Actually, it is not necessarily true that the robot reference system overlaps with the absolute one. For this reason, a calibration of the Robot Client reference system is fundamental. This allows the hybrid system to detect the rigid transformation to switch from the reference system of the Robot Client to the absolute reference system given by the Reference Marker.

The calibration consists in finding the offset between the two reference systems, which is represented by a roto-translation matrix. For this reason, both markers have to be tracked by the camera at the same time. The Server automatically performs this task every times the hybrid system is activated. The roto-translation between two different reference systems consists in calculating a matrix with six degrees of freedom (DoF); three degrees for the spatial position and three for the rotations. However, if the Reference Marker is placed on the floor, the two markers are on two parallel planes, consequently some DoF are constant and only three of them have to be found. As shown in Figure 6.6, the DoF under discussion are the translation on the X axis, the one on the Y axis and the rotation around the Z axis, which have been defined as X_0, Y_0 and θ respectively. By means of these values, it is possible to derive the initial position of the robot in the absolute reference system, as shown in the following equation,

$$M_{absolute}^{robot}{}^{(init)} = \begin{bmatrix} cos\theta & -sin\theta & 0 & X_0 \\ sin\theta & cos\theta & 0 & Y_0 \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6.1}$$

where Z_0 is the height of the Robot Clients marker to the floor, which is constant.

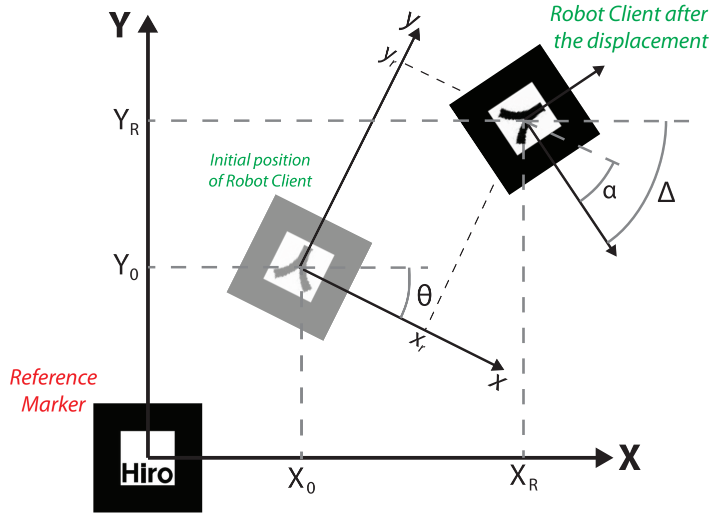


Figure 6.6: *Position and orientation of robot client in the absolute reference system.*

The data coming from the camera tracking have been exploited to calculate these three values. The AR Client sends the roto-translation matrices of the two markers to the Server, which deals with estimating the rigid transformation. Given the roto-translation matrices of the camera relative to the two markers as $M_{absolute}^{camera}$, if relative to the reference marker, and M_{robot}^{camera}, if relative to the robot one, it is possible to estimate the position of the robot in the absolute reference system by the following equation:

$$M_{absolute}^{robot}{}^{(init)} = M_{robot}^{camera} \cdot inv\,(M_{absolute}^{camera}) \tag{6.2}$$

and then calculate the values of the three degrees of freedom.

6.2.3 Camera Tracking by using the marker on robot

Every time the position of the robot changes, the robot immediately sends the encoders data to the Server. These data are consequently combined with the values in the Eq. 6.1 by the Server in order to estimate where the Robot Client is located, according to the absolute reference system. Figure 6.6 shows the coordinates of the robot in the absolute system, X_R and Y_R, and the orientation by the angle Δ. By assuming x_r and y_r as the robot coordinates position in plane and α the orientation, according to its odometric sensors, the absolute position and orientation is given by

$$\begin{bmatrix} X_R \\ Y_R \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \end{bmatrix} + \begin{bmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{bmatrix} \cdot \begin{bmatrix} x_r \\ y_r \end{bmatrix} \tag{6.3}$$

$$\Delta = \theta + \alpha \tag{6.4}$$

Consequently, the position of the robot in the absolute reference system is given by the following equation

$$M_{absolute}^{robot} = \begin{bmatrix} cos\Delta & -sin\Delta & 0 & X_R \\ sin\Delta & cos\Delta & 0 & Y_R \\ 0 & 0 & 1 & Z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6.5}$$

Finally, as shown in Figure 6.4, the estimation of the camera pose, according to the absolute reference system ($M_{absolute}^{camera}$), can be obtained by a composition of two matrices:

$$M_{absolute}^{camera} = M_{absolute}^{robot} \cdot M_{robot}^{camera}, \tag{6.6}$$

where M_{robot}^{camera} is the roto-translation matrix of the camera, which is calculated according to the marker mounted on the Robot Client.

6.2.4 Automatic drift compensation

As previously mentioned, the robot encoders are influenced by drift errors. This tendency has been taken into account. The drift error increases during the use of the Robot Client and it is directly associated to the usage time of the robot. For this reason, the real position and orientation of the robot in its coordinate system could not coincide with the one obtained by the encoders. The real measure can be modelled as a sum of two components:

$$\begin{bmatrix} x_r \\ y_r \end{bmatrix} = \begin{bmatrix} \tilde{x}_r \\ \tilde{y}_r \end{bmatrix} + \begin{bmatrix} e_x \\ e_y \end{bmatrix} \tag{6.7}$$

$$\alpha = \tilde{\alpha} + e_\alpha \tag{6.8}$$

where $\tilde{x}_r$, $\tilde{y}_r$ and $\tilde{\alpha}$ are the data given by the encoders, while e_x, e_y and e_α are the unknown errors given by the odometry. If the error becomes significant, the proposed hybrid tracking method becomes less effective.

The reference marker, which is located in the scene, allows us also to compensate the drift error. The Server automatically corrects the position of the Robot Client when the reference maker and the Robot Client are framed together. In fact, once the initial calibration has been carried out, it is possible to find the coordinates of the robot in its reference system by using data from the camera tracking. Hence, it is possible to estimate the drift error by combining the Eq. 6.3, Eq. 6.8, 6.7 and 6.8, as shown in the following equation

$$\begin{bmatrix} e_x \\ e_y \end{bmatrix} = \begin{bmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{bmatrix} \cdot \begin{bmatrix} X_R - X_0 \\ Y_R - Y_0 \end{bmatrix} - \begin{bmatrix} \tilde{x}_r \\ \tilde{y}_r \end{bmatrix} \tag{6.9}$$

$$e_\alpha = \Delta - \theta - \tilde{\alpha} \tag{6.10}$$

Once the drift is estimated, every new data from Robot Client are rectified.

6.3 Evaluation of tracking performance

In order to overcome the issues arising when the Reference Marker is not visible by the camera, we integrated two different tracking techniques for estimating the camera position in the absolute reference system. One of these two tracking techniques is based on the odometry, while the other one on computer vision and they are combined to estimate the camera pose. Both are affected by several errors. Therefore, the camera pose in the absolute reference system is even more affected by errors.

For this reason a test has been performed to assess the accuracy of our hybrid tracking method when the automatic re-calibration is not feasible. The new hybrid approach has been compared with a more precise tracking system that we consider our gold standard measurement system. In this case the comparative measurement system exploited is the Vicon System [31]. This is a vision based tracking system that works with spherical markers and infra-red light. For the testing session, we have recorded the tracking data from our system and from the Vicon one at the same time and afterwards we have compared the results.

6.3.1 Test Setup

Vicon and the hybrid tracking system had to estimate the position of the Robot Client and the AR Client camera point of view in the same tracking area. For this reason, an IR-marker set, which is necessary for Vicon to work, has been placed on the camera and another one has been mounted on the robot. This has allowed the Vicon System to track the two objects as well. The six Vicon cameras, model M2, have been positioned along the walls of the room used for the evaluation. Once the Vicon system was calibrated, it was able to cover a working area of $3m \times 3m$, with a precision of $1mm$ for the translation and about $0.5°$ for rotations. Thanks to this configuration, the two markers were visible at least by three cameras in the working space.

6.3.2 Testing session

In order to evaluate the accuracy of the developed system, we invited a user to use it. During the test, the user moved freely the AR Client within the Vicon tracking volume by framing the marker through the camera. For this test, any particular trajectory was requested. The user had just to continuously change the camera point of view, forcing the robot to always move in order to be visible by the camera. The Robot Client adapted its position in order to be always at the centre of the screen. For every new frame, the positions of the robot and the camera have been recorded in two log files, one for the Vicon data and another one for the hybrid system data. Afterwards, the tracking data has been analyzed, in order to estimate the error of the position of the robot, compared to Vicon, and the presence of a drift error.

Table 6.1: *Mean values and standard deviation error of Robot Client position and orientation during the test.*

	Error	Std. Deviation
X position [mm]	1.850	8.607
Y position [mm]	2.412	7.946
Rotation [deg]	3.163	1.506

6.3.3 Discussion

The hybrid system is quite precise to estimate its position, even if the position of the robot is not fixed by the automatic correction of the drift. In the Table 6.1, in fact, it is shown that the error made by the Robot Client during the test is limited. The hybrid tracking system works very well, as it is possible to notice in the Figure 6.7 as well. Also the camera position calculated by the AR Client is correct, even if it is less precise than the position provided by the Vicon, as it is possible to see in Table 6.2. In Figure 6.8 the trend of the camera position is shown. In this case the data deriving from the Z-axis of the camera tracking has not been shown, since the camera is on a trolley at a fixed distance to the ground.

Even if the data are affected by some errors and drift, the hybrid tracking method results to be effective. Indeed, the test carried out simulates an outer condition that hardly occurs during the use of the application. In the test, the user continuously changed the camera point of view, forcing the robot to always move so has to remain visible by the camera. During the common use, instead, once the user has

Table 6.2: *Mean values and standard deviation error of camera pose during the test.*

	Error	Std. Deviation
X position [mm]	5.535	7.281
Y position [mm]	5.845	7.986
Rotation [deg]	3.419	2.892

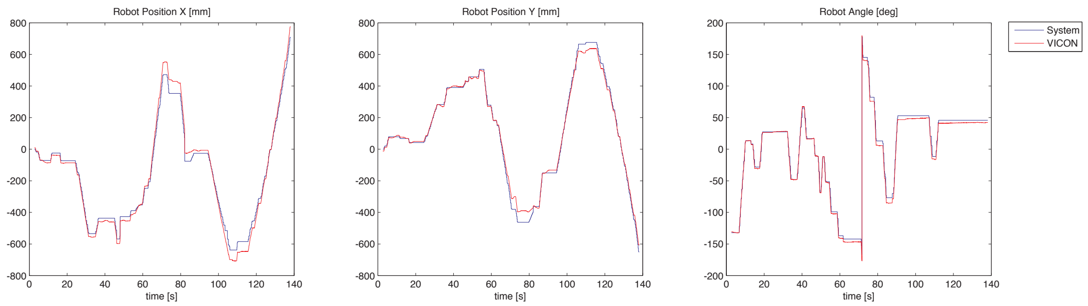


Figure 6.7: *Graph of Robot Client position and rotation during the test with the user. The blue lines correspond to the data estimated by means of encoders, the red ones correspond to data by Vicon System.*

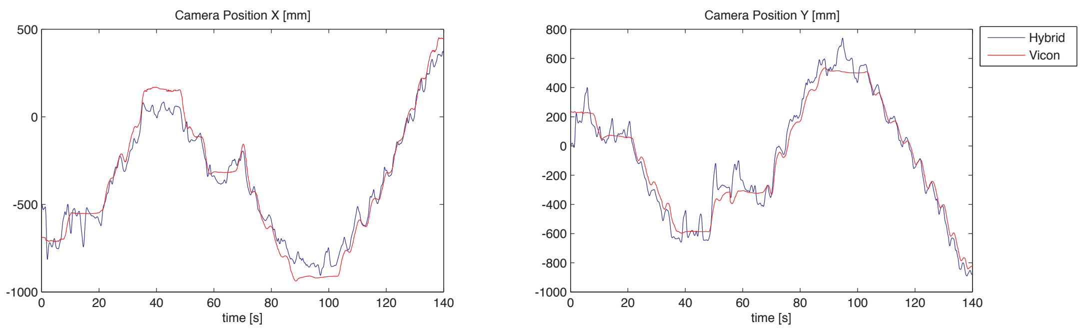


Figure 6.8: *Graphs of the camera position in X and Y during the test with user. The blue lines correspond to the data estimated by means of our hybrid system, the red ones correspond to data calculated by Vicon System.*

chosen the camera point of view and the robot is right placed, it is not necessary any new movement. For this reason we can state that this hybrid tracking approach can be used for the camera pose estimation in the environment.

6.4 Conclusion

In this Chapter we described the development and validation of a hybrid tracking system. The system uses a marker-based tracking system and a mobile robot to perform tracking in wide environments without structuring the environment or look to expensive solutions. Actually, we estimated the camera point of view by means of a combination of marker-based tracking and the data coming from the robot odometry. The position of the robot can be managed manually by the user or automatically by the system, which commands the robot to move so that it is always framed by the camera. Moreover, we reported the algorithms developed to obtain a quick initialization of the system and an auto-calibration of the robot position.

We evaluated the system by comparing it with a precise tracking solution, which is usually exploited to track in wide environments. Even though the proposed system is not very precise, it provides a good tracking for AR purposes.

CHAPTER 7

Natural Features Tracking for Circuit Boards

We describe in this Chapter the detection and estimation of circuit boards pose for AR purposes. The challenge of this work regards to the capacity of the tracking system to work without any additional device in the environment, such as markers or RFID. The system proposed works only by using vision and drew inspiration from tracking solutions already explored in other fields. We imported the tracking system into an industrial context and we further improved it to deal with circuit boards. Finally, we evaluate the tracking system in order to use it in the application described at Chapter 10.

7.1 System Description

The circuit board is an object that we can know a priori. For this reason, a tracking system based on natural features and a model of the circuit board can be used. The *reference model* is generated during an offline process, usually by means of some pictures. Through this phase, some unique features are detected in the images, their position in the environment regards to a fixed reference system is calculated and then used to describe the reference model. During tracking, we use the reference model to

- recognize the circuit board in the frame
- estimate the camera pose relative to the circuit board.

In the recognition phase, the system establishes if the circuit board that corresponds to the model is present in the framed environment. For this reason, the system has to analyze the new frame and extract some features that it can relate with the ones in the model. Every correct relation between features from the frame and the reference model is called *matching*.

As already described in the Equation 2.4 at page 19, the pose estimation corresponds to find the roto-translation matrix. In this case, the data about the position of the features in the space have already been elaborated during the offline process of the model.

7.1.1 Planar Assumption

The tracking system described in this Chapter considers the circuit boards as planar objects. As first approximation, we can approximate circuit boards to planar, because they are usually stretched out along two dimensions. In this way, we can exploit an image-based approach for the Natural Features tracking systems. Thus, the reference model relies only on a *orthographic picture* of the circuit board and 2D features that the system is able to get from the picture.

Tracking systems that deal with planar images have already been used in simple applications in the field of entertainment and marketing, as shown in Figure 7.1. In this work, we import this approach in the industrial field and we extend it to deal with circuit boards.

(a) (b)

Figure 7.1: *Examples of image-based natural features tracking for AR. On the left, the tracking system relies on a picture card to perform the augmentation. On the right, a marketing AR application from LEGO®; the system recognizes the box and it superimpose the 3D model of the construction on top of it.*

7.1.2 The Ferns Classifier

The biggest problem for a Natural Features tracker is the matching between the keypoints in the reference image and the ones detected in the video stream. One way to establish correspondences is by using local feature descriptors, like SIFT [103] and SURF [41]. Once the local descriptor for a keypoint in the image has been calculated, it is possible to compare it to the ones in the reference image. The matching is between the two most similar local descriptors. The problem of this approach is due to the computational time to calculate the descriptors and find the right matchings that are not feasible in real time.

In order to work in real time, we treated here the matching problem as a classification one [102]. This is a solution that can be used on a modern computer without any particular adaption. A matching approach based on classification is fasted than using a local descriptor, thanks to the offline process to train the classifier.

In this work, we chose to use a Fern Classifier [124, 125]. During the offline phase, the patches around each keypoint are used to synthesize the view of the patch from other point of view by mean of affine transformations. In this way, it is possible to recognize the patch even tough big prospective distortion of the patches occurs. Then, a class made of all the possible appearances of the image patch is associated to the keypoint of the reference image. Each patch is described by many binary features, each of them depending on the intensities of pair of pixels in the patch. During the offline process, the classes are used to train the classifier.

7.1.3 Real time performances

After the training phase, the Fern Classifier can be used to find the correspondences between the keypoints in the reference image and the one in the image stream. As it is possible to see in Figure 7.2, keypoints are located in the image and the patches around each of them are extracted. For each patch, the binary features are calculated. Afterwards, the trained classifier finds the best class of patches that corresponds to the patch extracted to the image. In this way, the classifier finds the matching between a keypoint in the reference image and the one in the stream image. Since the features are calculated as binary data, the comparison is very fast and makes the classification process suitable for real time applications.

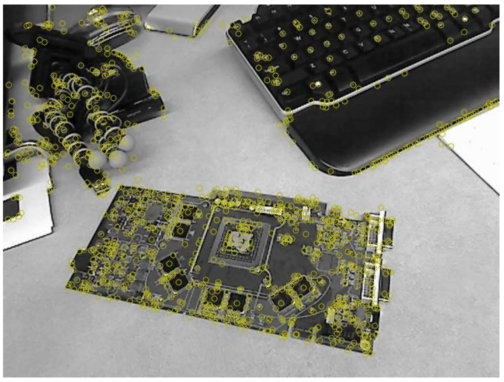


Figure 7.2: *Keypoints extracted from an image. The yellow circles indicate keypoints the system found in the image.*

7.2 Development of Circuit Boards tracker

We improved the natural features approach for tracking in order to reach good performances when working with circuit boards. In particular, we optimized the use of a Fern classifier for this purpose.As it is possible to see in Figure 7.3, the common matching process has been modified in order to check the correct correspondences and to focus the image processing algorithms only on the area of the image where the circuit board is present.

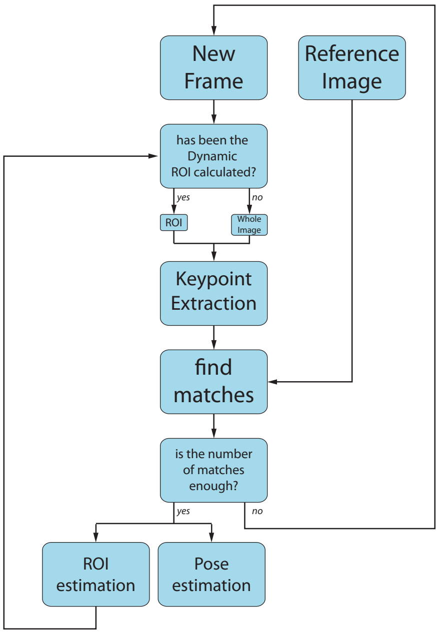


Figure 7.3: *The circuit board tracking workflow.*

7.2.1 Wrong Matching removing

The Fern Classifier sometimes fails and provides some wrong matched keypoint. These ones could affect the estimation of the final pose. For this reason, we developed an algorithm to remove the wrong correspondences after the classification task.

The algorithm to remove the wrong keypoint grounds on the use of homograpyes. Since the circuit board and the reference image can be assumed as a plane in the space, it is possible to find the homography H to transform the plane of the reference image to the one of the circuit board in the current frame. In this case the homografy has been calculated by using a RANSAC approach [66].

The center of each feature detected in the stream image lies on a plane and it is described as the point a, which coordinates are

$$a = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, \tag{7.1}$$

while the correspondent one in the reference image is

$$b = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}. \tag{7.2}$$

By means of the homography H, it is possible to transform the point b, in the reference image, to the correspondent point $\hat{a}$ on the plane in the space.

$$\hat{a} = H \cdot b \tag{7.3}$$

In case of correct matching between the feature in the reference image and the image stream, a and $\hat{a}$ should coincide, or at least lie close to each other if some error occurs. According to a tolerance $toll$, the match is considered good if

$$\|a - \hat{a}\| \leq toll \tag{7.4}$$

otherwise it is discarded. We fixed the tolerance to 20 pixel. This is a quite high value of tolerance, but we chose it high in order to take into account that the circuit board is not completely planar and the presence of possible distortions in the image. Two frames represented in Figure 7.4 regards the matching process between the reference image and the camera stream.

7.2.2 Dynamic Region of Interest

Every time the system analyses a new frame, it should deal with the entire image to find the circuit board. However, the position of the circuit board in two consecutive frames is similar, so it is not necessary to work with the entire image. For this reason we developed an algorithm that focuses only on a region of interest, that corresponds to an area of the image wherein the circuit board was present in the previous frame. The feature detection becomes faster, because it is only on a portion of image and it does not have to deal with the data of the entire image. Since the position of the circuit board keeps on changing, this area of image is not still and it has been called Dynamic Region of Interest.

As it is possible to see in the Figure 7.5, the system calculates the region of interest starting from the coordinates of the vertices of the circuit board. The area of the region of interest is described by a rectangle that inscribes the four vertices. Finally we decided to increase the size of the rectangle in order to cover the board in the next frame with much more probability.

In case of failure of recognition of the circuit board in the image, the system cannot use the dynamic ROI and automatically switch to the feature detection task on the entire image.

7.2.3 Pose estimation

Since the pose estimation grounds on the Equations 2.4 , it is necessary to know the position in the space of the features in the reference image. As it is possible to see in Figure 7.6, we can have a linear

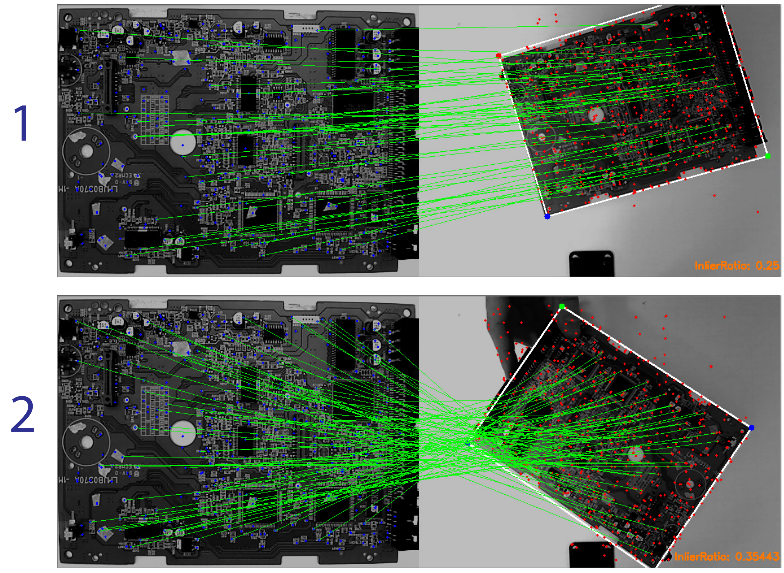


Figure 7.4: *Example of matching process in two frames. On the left of each picture the reference image is represented: the blue small dots corresponds to the reference keypoints. On the right the current image captured by the camera: the red dots are the detected keypoints. The green lines represent a good match between a keypoint in the reference image and its corresondent one in the captured image.*

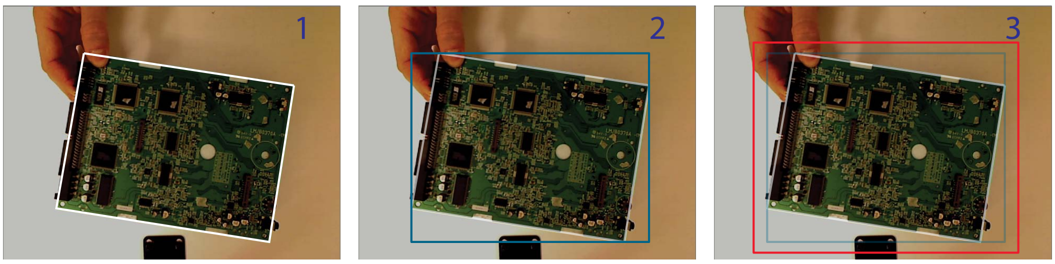


Figure 7.5: *The three steps of the creation of a dynamic region of interest.*

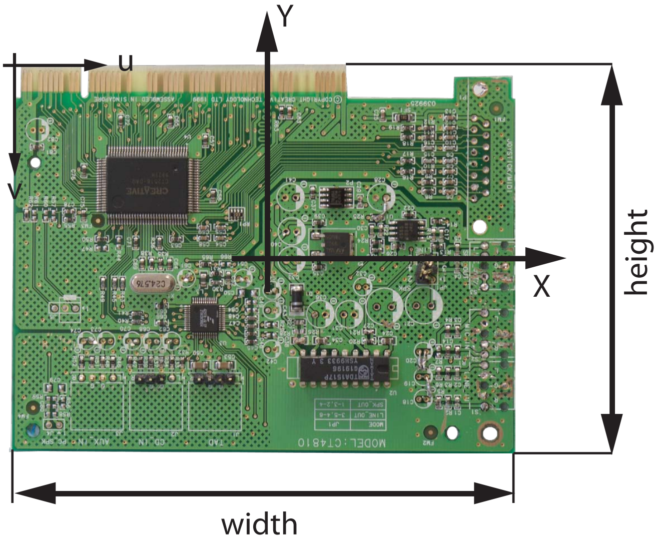


Figure 7.6: *The assessment of the position of a keypoint in the space by means of the reference image is quite straightforward. Actually, we can perform it by a linear transformation between the image plane coordinates* (u, v) *and the spatial coorinates* (x, y *and* $z = 0$)*.*

relation between the position of the feature in the image reference system (in pixels) and the fixed one (in millimeters) because the reference image is orthographic.
We easily correlate the position in pixel with the one in millimeters by the conversion factor

$$c = \begin{cases} \frac{width_{mm}}{width_{px}} & \text{or} \\ \frac{height_{mm}}{height_{px}} \end{cases} \tag{7.5}$$

where $width_{mm}$ and $height_{mm}$ are the real dimensions of the circuit board in millimeters, while $width_{px}$ and $height_{px}$ are the equivalent ones in pixel. Consequently, the linear transformation between one reference system to another one can be described by the following Equation

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = c \cdot \begin{bmatrix} 1 & 0 & -\frac{width_{px}}{2} \\ 0 & -1 & \frac{height_{px}}{2} \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{7.6}$$

It is worth noting that z is null because of the assumption of planarity of the circuit board. This operation is performed only once when the system loads the data of the reference model.

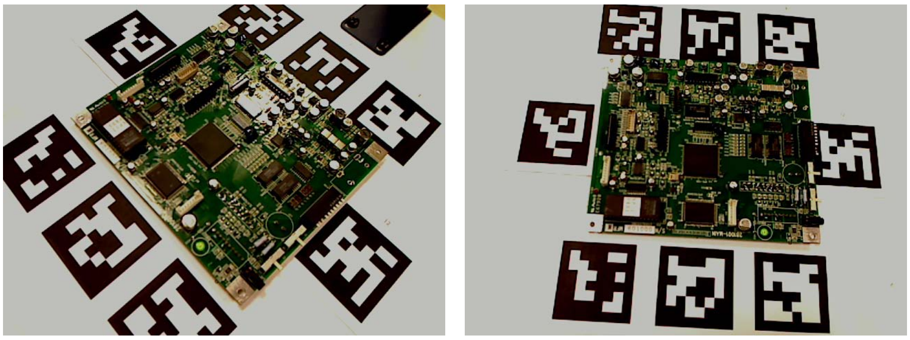
7.3 Evaluation

We carried out an evaluation session in order to assess if the tracking performances of the proposed tracking system are feasible for AR purposes. Thus, we compared the tracking data coming from our tracking system with the one achievable from a marker-based tracking solution.

7.3.1 Test setup

We placed a circuit board on a planar surface and we arranged some ARToolKit markers around it. The markers are grouped together in order to form a multi-marker system. In this manner, it is possible to track the position of the reference system even if some of markers are not visible by the camera. Moreover, a multi-marker system turns out to be robust than using a single one; the pose is estimated by means of more data, so it is less prone to error due to noise.

The size of the circuit board taken into account is $120mm \times 95mm$, while the markers side is $50mm$. Figure 7.7 shows how we disposed the circuit board and the markers.



(a) (b)

Figure 7.7: *Setup for the evaluation of the circuit board tracking.*

The application that detects the pose works by means of recorded videos. For each frame, it calculates the pose according to the markers reference system and the circuit board. The data about pose are then stored in a file. We used a calibrated camera Logitech Quick Cam Pro 9000 to record the videos.

7.3.2 Test session

We recorded some videos in order to check the tracking performances of our proposed solution for circuit boards. For each video, we moved the camera in order to have frames from different point of views. In order to detect the offset between the markers reference system and the circuit board one, we also recorded a video where the camera is fixed on a tripod and it statically frames markers and circuit board. Afterwards, we estimated the camera pose by means of the software application previously described. Figure 7.8 shows the position and orientation of the camera during one of the videos. The results are in the same reference system because we removed the offset.

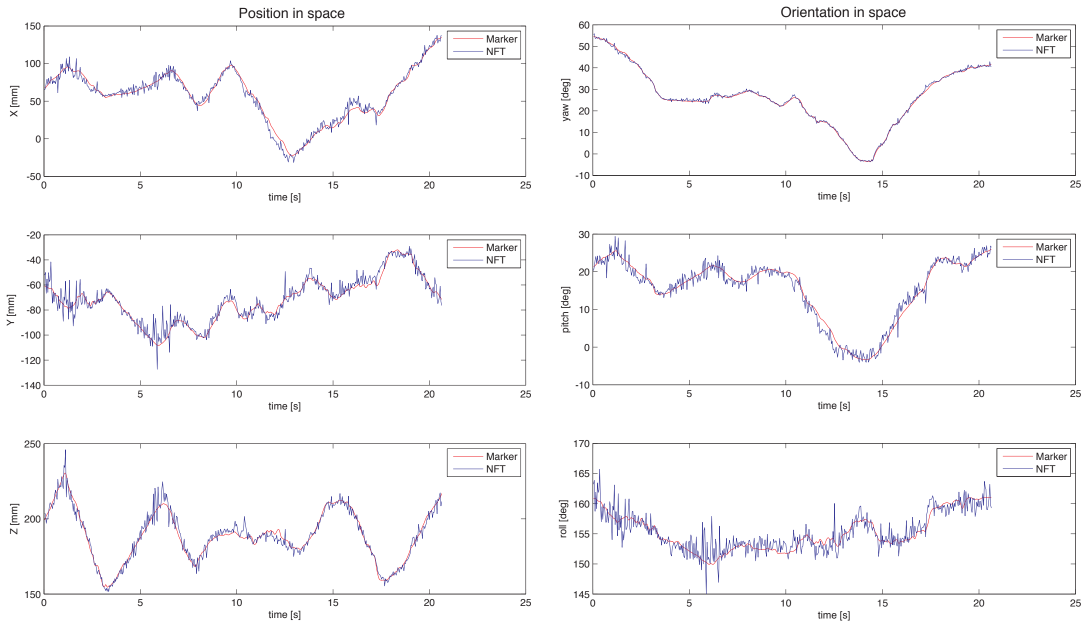
7.3.3 Discussion

As it is shown in Figure 7.8, the data about tracking by natural features are comparable with the ones obtained by the marker-based tracking. The higher variability during the estimation with our proposed solution entails some jittering, but it is still acceptable. The higher variability issue is probably due to the assumption of planarity of the circuit boards. Actually, the keypoints are not on the same plane because of some components in relief on the circuit board. Thus, the data used for tracking have some error since the beginning.

If the marker-based system is considered as the basis of comparison, we can calculate the error during tracking the circuit board. Table 7.1 reports the average error and the standard deviation of the tracking error for each of the six DoF. Once more time, we can verify that the tracking error of our solution is suitable for AR purposes.

7.4 Conclusion

In this chapter we described the development and evaluation of a tracking system that works with circuit boards. The developed approach grounds on natural features tracking techniques wherein the reference model is an orthographic image. This solution has already been used in other fields and we imported it here to work with circuit boards.



(a) (b)

Figure 7.8: *Graphs of the camera position (a) and orientation (b) during one of the evaluation tests. The blue lines correspond to the data estimated by means of our circuit board tracking system, the red ones correspond to data calculated by the multi-marker one.*

Table 7.1: *Average value and standard deviation of the errors for the six DoF of the circuit board tracking obtained during the evaluation test.*

	Error	Std. Deviation
X [mm]	4.559	3.884
Y [mm]	3.856	3.900
Z [mm]	2.727	2.595
yaw [deg]	0.442	0.382
pitch [deg]	1.183	1.007
roll [deg]	1.152	1.068

We evaluated the proposed solution by comparing its tracking performances with the one obtained by a multi-marker ARToolKit system. It turned out that our tracking solution is acceptable for an AR purposes.

Part III

Interactive Applications

CHAPTER 8

Interactive system for product evaluation in immersive environments

During a Design Review phase wherein AR is involved, the use of an HMD is a good solution for the designer to visualize the virtual object in the environment [95]. Actually, designer should be able to see a three-dimensional representation of a *virtual prototype* (VP) from different points and interact with the virtual content in the more natural way as possible. A tracked VST-HMD increases the feeling of immersion of the user that can move in the AR environment and see the virtual objects into the real world without the need to hold any visualization device in his hands.

The user's hands are not constrained and they are free to interact with the AR environment. Unfortunately, at time of developing of this system, only few solutions provided a natural hand-free interaction, but they were limited [47, 58]. The most common solution is interaction by means of input devices such as joysticks or data gloves coupled with an external tracking system for interacting with the 3D environment. Many of them are uncomfortable or bulky for a natural experience, most of the time because of wires for communication of data and power. For this reason, a good solution for 3D interaction would be a wireless device simple and easy to use.

In this Chapter an alternative solution for interaction in immersive AR has been chosen by using a device from a not-industrial field and adapting it for Design Review purposes. This device is the Nintendo WiiMote, which is from video game field. As already described in Section 5.1, WiiMote is a cheap hardware that is equipped with several technologies embedded; all of these resources are used here for managing virtual objects.

8.1 System Description

The system is made up of some hardware components that are managed by a software, which is the AR application for Design Review. The components of this system are:

- VST-HMD for the immersive stereoscopic visualization
- WiiMote for interaction with digital objects
- Application software for managing visualization and interaction

Since WiiMote has been described in Section 5.1 already, only the other two components are reported as following.

8.1.1 VST-HMD

At the moment of developing of this system, commercial VST-HMD were rare. For this reason, a custom one, which is oriented to Design Review purposes, is used here [49]. This HMD solves the issue related to the correct visualization of 3D virtual objects closely and far away by controlling the convergence angle of the video cameras mechanically by using two micro servos.
The device is composed by:

- Two cameras to acquire the scene. Their resolution is 640×480 pixels and they work with a refresh rate of $30Hz$.
- Two monocular OLED display. Their resolution is 800×600 pixel and the field of view is 32 degrees.
- Two analog servos.

As it is possible to see in Figure 8.1, all components were arranged on a light safety helmet so that the weight of the device is about $850g$. A specific frame provides all of the degrees of freedom for a correct registration, according to different Inter Pupillar distance (IPD) and head size. The development of the helmet is described in detail in [49, 50].

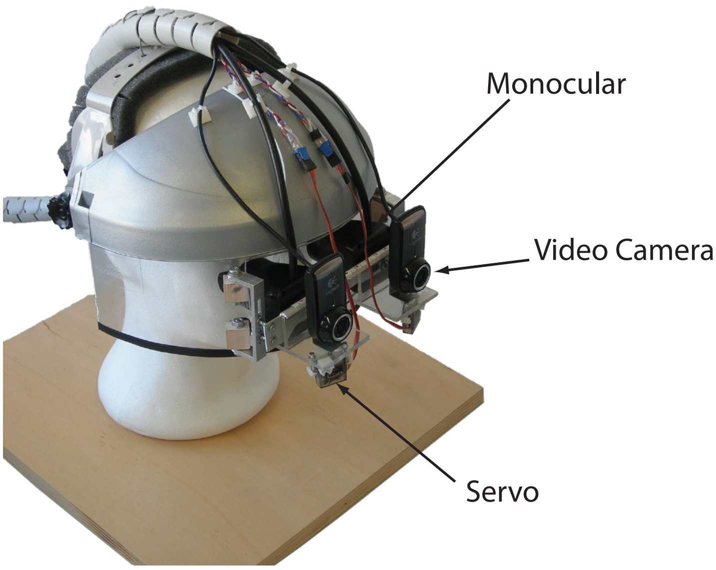


Figure 8.1: *The VST-HMD used for the immersive visualization. The helmet is equipped with a camera, a monocular display and a servo for each eye.*

8.1.2 Application Software

The developed software deals with visualization and interaction with virtual objects. The two cameras frame the environment and the system detects their pose by means of ARToolKit. The virtual object is then displayed correctly in the environment. In case of close range view, the two servos are managed in order to adapt the visualization. All the data, passed wirelessly by WiiMote, are then used to find the position of the device and for interactive purposes by managing the virtual object.

8.2 Interaction metaphors for Design Review

We used the majority of equipment embedded in the WiiMote to interact with the virtual object. Since there are several technologies involved, even the interactive metaphors are various. Thus, it is possible to have solutions to accomplish the same interactive task, but in different ways. The redundancy of interaction allows the user to exploit the best suitable for each occasion.

8.2.1 Interaction by gestures

The first developed interactive metaphors are based on the hand movement recognition and are activated by pressing one of the four buttons of the WiiMote directional pad. Effectively, the user modifies a spatial feature of the virtual object by pressing one of the four buttons and rotating the WiiMote along the X or Y axis (Figure 8.2). In this way, the system make use of the value of the yaw or roll angle detected by accelerometers to manage the object. As it is also summarized in Table 8.1, we used the buttons LEFT, RIGHT and UP to move the object in the space, if the WiiMote is rotated on the Y axis, or to change its orientation, if the WiiMote is rotated on the X axis. Instead, it is possible to change the size of virtual object by means of pressing the button DOWN and rotating the WiiMote on its Y axis. We designed this metaphor by starting from the concept of HI-FI knob because it allows increasing or decreasing the virtual object size as the HI-FI knob does with the volume. Besides, the user can reset all of the virtual object parameters (position, orientation, size) by pressing the button HOME.

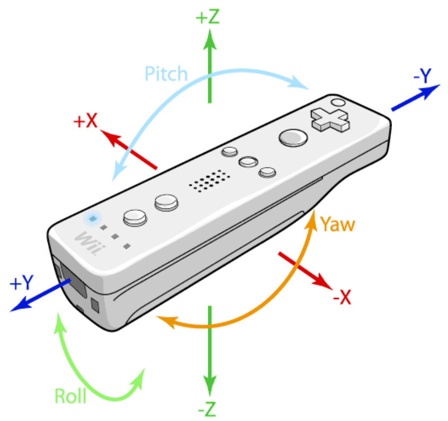


Figure 8.2: *The WiiMote angles and axes.*

Table 8.1: *The interaction metaphors. For each button pressed, the system modifies the position or the orientation of the virtual object along one axis.*

Button pressed	Rotation axis	Interaction
LEFT	X	Rotate around X axis
LEFT	Y	Translate along X axis
RIGHT	X	Rotate around Z axis
RIGHT	Y	Translate along Z axis
UP	X	Rotate around Y axis
UP	Y	Translate along Y axis
DOWN	Y	Change size
HOME		Reset

8.2.2 Interaction by tracking

The WiiMote tracking in the space by means of a IR-marker described in Chapter 5 enables other interaction metaphors. As a matter of fact, the tracking functionality for the WiiMote is used to developed a 3D pointing device with the metaphor of virtual pointer. The virtual pointer is a 3D object that helps the user to see more precisely where he is pointing at. As shown in Figure 8.3, the virtual pointer is represented as a 3D arrow at front of WiiMote and it follows the position and orientation imposed by the user. In this way, the perception of the placement of the WiiMote in the AR space is increased by the vision of this arrow as an extension of it.

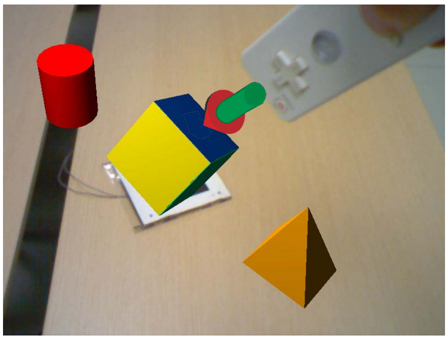


Figure 8.3: *The virtual pointer.*

This feature is useful for an AR application for design review because allows the user to indicate or mark a specific area of the 3D virtual object. In this case, it is possible to place a mark in the space by pressing the button ONE.

The tracking functionality is used for a "grabbing" metaphor with the WiiMote as well. The user points the object and, by pressing the button B on the WiiMote, the virtual object is hooked to the controller. Then, the object follows the movements of the user hand. When the button is released, the object remains in the new position.

8.3 Evaluation

The functionalities of the AR system, described in the previous sections, have been subsequently evaluated through a testing session where 10 users were involved for trying our AR system by performing some specific tasks. As it is possible to see in Figure 8.4, the user wears the HMD and interact in the augmented environment by means of the WiiMote. All the users were people skilled in the use of AR environments because, the selection of expert users sample, allowed us to identify many problems associated with the design of our user interfaces [18] .

8.3.1 Description of the test

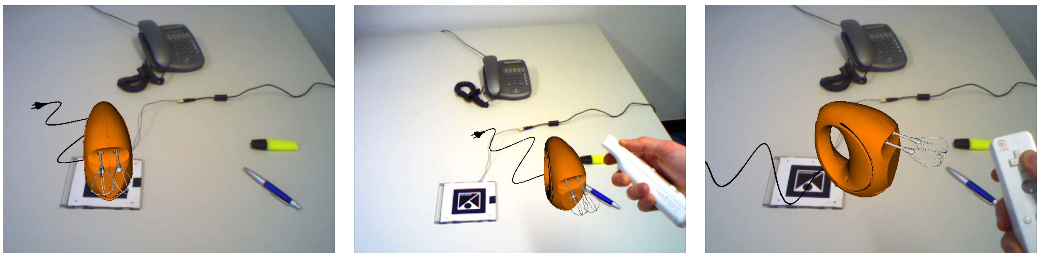
The scenario was the typical design review one where an expert has to evaluate the shape of an industrial product. For this test session we made a VP of a cooking mixer and the user have to interact with it. The expert have to see the model from different points of view and he has to use the metaphors developed to explore the object. The task to accomplish were:

Task 1: handling the VP through the interaction metaphors;

Task 2: moving the VP by means of the WiiMote tracking;

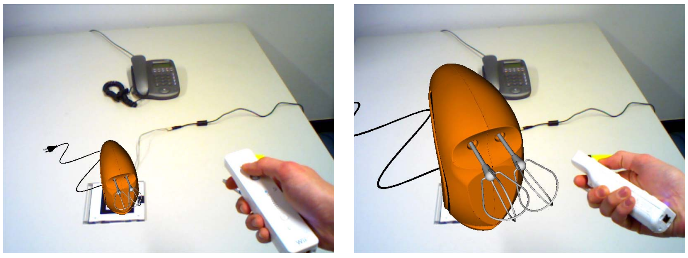
Task 3: using the virtual pointer to set a mark on a specific part of the VP.

Figure 8.4: *One of the expert user during the test.*



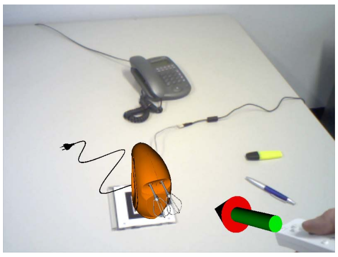
(a) (b) (c)

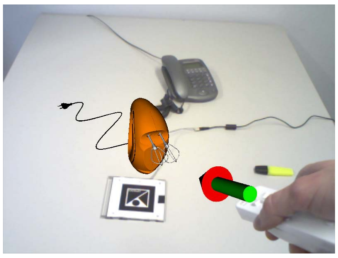
Figure 8.5: *Screen shoot during the testing session. The Virtual Object in the Real Scene (a) and change of position (b) and orientation (c) of the model.*



(a) (b)

Figure 8.6: *Resizing of the virtual object (b).*


(a)


(b)

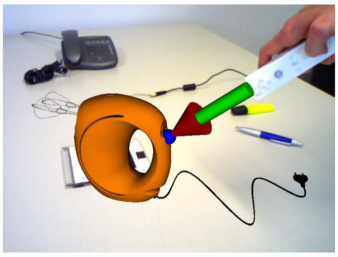

(c)

Figure 8.7: *Virtual Pointer for interaction with the virtual object (a). The interaction metaphor is used to move the virtual object (b) an to place a mark on the Virtual Prototype (c).*

Before starting each testing session, we described briefly the functioning of our system to the users and then, we explained the tasks that they had to perform. Afterwards, they had to execute the tasks. In Figure 8.5, 8.6 and 8.7 , some screenshots during the execution of the tasks are presented.

After the test execution, we invited the users to answer to some questions about AR visualization and interaction with WiiMote. The questionnaire that we used is reported in the Appendix A.1. They had to give a score from 1 (bad) to 6 (good) for each question and some comments about the AR system as well. The questions and the average results of the answers are summarized in Table 8.2.

Table 8.2: *Users' answers and average score.*

Visualization	
A1. Video quality	4.8
A2. Is the representation of the VP realistic?	4.7
A3. Is the VP integrated in the real scene?	4.4

Interaction	
B1. How easy to move the VP by means of the WiiMote?	4.4
B2. Were you able to move the VP where you wanted?	3.1
B3. Was easy to use the interaction metaphors?	5.5
B4. Were you able to learn the interaction metaphors rapidly?	4.8
B5. How much is useful the virtual pointer in this environment?	5.6
B6. Were you able to mark the VP where you wanted?	4.4

8.3.2 Discussion

The results in Table 8.2 and the comments given by the users are good and indicate a positive evaluation of the AR system for design review.

Visualization

All of the users saw the scene in the correct way by perceiving the depth sensation both of the real and the virtual objects. The analysis of the answers to the questionnaire highlights that the VST-HMD gives the possibility to see the virtual and the real objects at the same time with good quality (question A1). Moreover, the representation of the VP appears realistic enough and well integrated with the real environment (questions A2, A3). The unique relevant issue pointed out by the users was about the little field of view of the displays. A more detailed evaluation of the visualization part is described in [50].

Interaction

The results of the questionnaire demonstrate the users were able to use the WiiMote to interact with the virtual object. The users confirm that the proposed interaction metaphors by gestures are easy to learn and to use (question B3). Obviously the answer value of the question B4 reveals that it required some time in order to learn the metaphors correctly. Even the interaction by means of tracking are simple and intuitive because this type of interaction is similar to the natural movement made manually when handling a real object. However, the question B2, instead, reveals that the problems due to the small working field of IR tracking does not let the user to move the VP far away from the marker. Since the interaction metaphors by gestures received better judgement than the one by tracking, they are considered easier to manage by the users. This is due to the tracking that is a little bit unsteady, as previously reported in Section 5.3. For this reason the users suggested to manage the object by the tracked WiiMote and use the interaction metaphors as a fine tuning of position and orientation. Finally, the users appreciated the virtual pointer and considered it very useful in this AR environments (question B5). They confirmed that the virtual pointer improves the immersion of their hand in interacting with the VP. Every user was able to mark a specific part of VP and considered that the execution of this task did not require particular ability.

8.4 Conclusion

An immersive AR system for Design Review purposes needs not-traditional devices to interact with a Virtual Prototype. In this chapter, the possibility of using a device for entertainment purposes in a industrial application is shown. Actually, a WiiMote has been integrated into the AR system and used for interaction purposes.

By exploiting the technologies embedded within WiiMote, some interactive metaphors has beed developed. One of these metaphors has been made by means of the tracking system developed in Chapter 5.

Afterwards, the interactive system has been evaluated by expert users. In order to do that, a test wherein the users had to interact with a virtual cooking mixer has been performed. The positive results obtained by the test demonstrate a good integration of WiiMote in the Design Review System.

CHAPTER 9

Interactive System for Space Planning

The Space Planning (SP) is a design process regarding to disposition of objects in a space. For example, objects at issue are furniture, if the environment is a house, or machines, if the SP is carried out in a factory. Space Planning is not a simple process and often it involves more than a single category of specialists (interior designers, architects, ergonomists, etc.), who have to collaborate together to define an optimal furniture layout in terms of aesthetics, ergonomics and functionality.

One of the main challenges of this process is to show the proposed planning in a realistic manner. The AR approach, instead, enables specialists to evaluate the planning directly into the real scene by making the process more effective and efficient. In this Chapter, we propose an interactive system for space planning my means of AR and we evaluate it by some tests in different kind of environments.

9.1 Space Planning Activity

The definition of an optimal furniture layout implies taking into account several aspects, for instance ergonomics, aesthetics and safety, according to the space that has to be arranged. Skilled interior designers follow guidelines to optimizing the space layout. Some of these guidelines are regulated while others have been elaborated, over the years, starting from the personal experience and by common sense.The usual methodology to obtain a good furniture layout is divided in two phases:

Analysis phase, which is a pre-design process. During this phase, data are collected, analyzed and summarized. These data are very heterogeneous since they include information about users needs, structural constraints, furniture functionalities, etc.

Synthesis phase, which is the out-and-out design process. In this phase the space planner elaborates several solutions, according to the results of the analysis phase. These different solutions can be done by using tools that enable a quickly arranging and rearranging of the position of the furniture.

9.1.1 Computer Aided Tools

Nowadays, the SP process is supported by Computer-Aided tools that allow the easily rearranging of virtual furniture in a virtual environment. Simplest tools, which are based on bi-dimensional drawings, can be useful to assess some of the criteria, since it is possible to simply evaluate the encumbrance, the position and the orientation of each piece of the furniture with respect to the dimensions of the room. However, this simplified representation of the furniture does not allow correctly evaluating other criteria such as the aesthetic impact of the furniture layout.

3D modeling and rendering techniques, instead, give the possibility to improve the virtual representation both of furniture and the surrounding environment, where the furniture has to be arranged. The result of this virtual simulation is very effective and the visualization, of different furniture layouts, reaches high level of realism. However, making this kind of virtual simulation is a time-consuming process and requires specific skills. Furthermore, while the 3D models of furniture can be imported by pre-existent databases, often, the surrounding environment has to be modeled from scratch.

9.1.2 AR to support Space Planning

Technologies related to AR allows the space planner to overcome the issue of modeling the environment, since virtual furniture can be rearranged directly in the real one. In addition, the use of the real environment increases the comprehension of the final result as well. In fact, if the furniture arrangement is displayed in the real world, no abstraction activities are required to imagine the final results. Consequently, the better understanding of the final result improves the discussion with whom have no skills in space planning and in 3D modeling (e.g. the final customer).

The need to correctly visualize virtual furniture in the real room implies using a system able to track the whole working area. A cheap solution consists of marker-based tracking systems, where fiducial markers are located within the real scene. This approach has been exploited by several authors. In [144] a tracking approach that integrates still images with a single marker is described. However, the user is not able to physically move from a position in the space to another one, but he can only switch from different pictures, limiting his capacity to evaluate a spatial layout. Another approach is based on the use of several markers placed inside the room: each marker is associated to a different piece of furniture like in [57] and [170]. Unfortunately, even if this kind of solution is effective for AR visualization, it does not allow the user to remotely manage the furniture position, which is an important feature to conduct space-planning activities. Finally, some authors use multiple markers to estimate the camera pose [71] and associate the position of each piece of furniture according to a reference system, which is defined by combining the reference systems provided by each marker. However, in this case the user has to fix and calibrate several markers and this tasks require a lot of time in order to be accomplished.

Since the marker-based tracking is a cheap solution for AR applications but shows some drawbacks due to setup and initialization, we use here the tracking system described in Chapter 6 for this space-planning application. Actually, this tracking approach allow us to speed up the initialization process and it is possible to track in a wide area by moving the robot in the environment.

9.2 System Description

The system used for this AR application is based on hardware already described in Section 6.1 and it is shown in Figure 9.1. In particular, we describe in this part the AR Client and its Graphic User Interface (GUI), which we expressly developed for space planning purposes. The AR Client is settled on a trolley, in order to relieve strain on the user during the use.

9.2.1 Estimation of the camera pose

The hybrid tracking system grounds on the two markers reported in Section 6.2. The Reference Marker is placed on the ground, while the one on the Robot Client is at $170mm$ from the ground. A Kalman filter [166] has been implemented to relieve tracking errors that entail jitters. In this way, the virtual objects are more stable during the visualization.

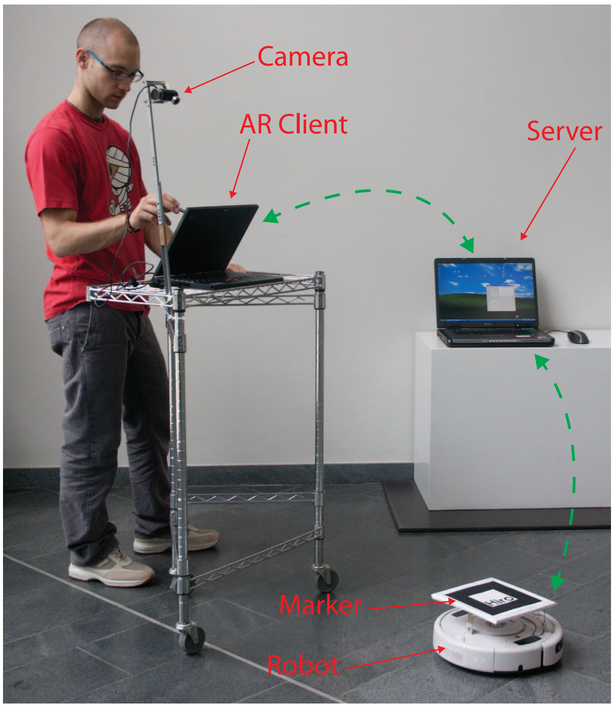


Figure 9.1: *Hardware components of the hybrid tracking system. The dashed green lines represent the communication between the devices.*

Robot Interaction

The user can also interact with the robot in order to change the camera point of view. In fact, it is possible to manage the position of Robot Client in the working space in order to cover a wide AR volume. The interaction with the Robot Client is performed by a GUI that allows the user to manage the robot in two different modalities, automatic or manual, as shown in Figure 9.2.

In the automatic case, the user changes the camera point of view and the system automatically calculates the robot displacement as deeply described in Section 6.2.1. Subsequently, the system sends commands to Robot Client in order to move the robot and let the marker visible by the camera. In the second case, the user decides where to place the robot in the environment using the buttons on the GUI. This second modality could be useful in order to have a complete control of the robot, as complex environments for example. Actually, if the room is full of corners or obstacles, a manual control of the robot should be better.

9.2.2 AR Objects

The AR client developed for space planning has to manage several virtual objects. Actually, these objects are not only virtual pieces of furniture or other big objects, but also the lights in the environment and the drawing of the plans previously designed.

Models

All the models that the designer can use are stored in the memory and it is possible to visualize them by a GUI, which is shown in Figure 9.3. The designer can switch trough them, select the one he is interested to and place it in the augmented environment. By means of the GUI, it is also possible to move independently the objects added in the environment, by means of the interactive solutions described in detail in the next section, and delete them. Finally, it is possible to save a configuration of objects placed in the space or load a layout previously saved, in order to easily evaluate different solutions.

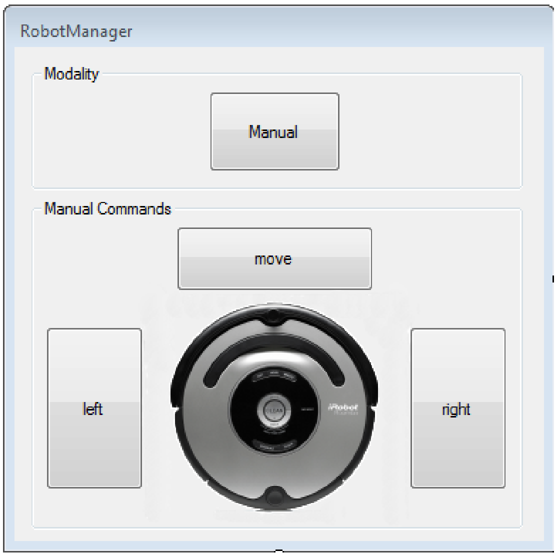


Figure 9.2: *The GUI to manage Robot Client. The user can select an automatic or manual modality to control the robot. In the second case, he can move it through the buttons in the GUI.*

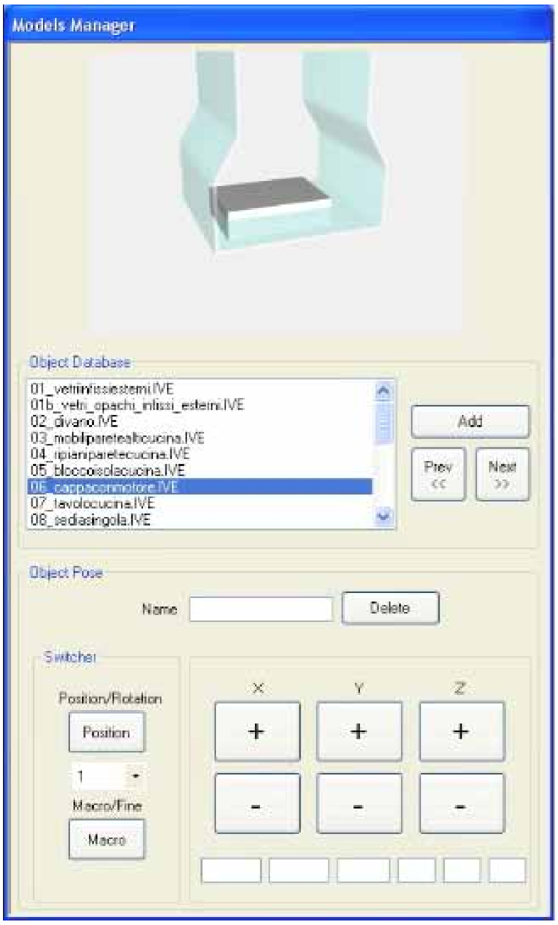


Figure 9.3: *The Model Manager. On top of the window, the user can see the digital object selected in the list.*

Plans

The application allows the user to visualize the plan of the environment, previously drawn by the designer, directly into the augmented scene. Once the plan is selected and loaded, it is superimposed to the floor, as shown in Figure 9.4. The representation of the plan is relative to the Reference Marker and it is fixed.

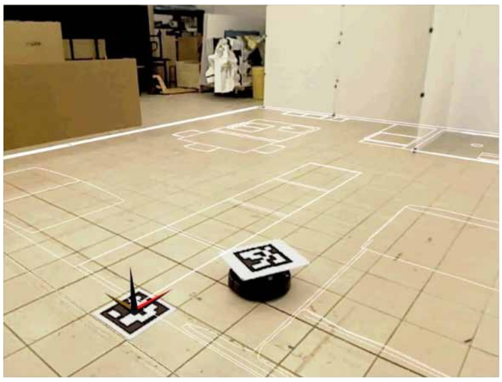


Figure 9.4: *Representation of the design plan. Once the system is able to track the camera point of view, it is possible to visualize the design plan on the floor.*

The main purpose of using them in our AR application is to check the correct design. Actually it can be used as a reference to locate the virtual object in the space, in this way the user can visualize the virtual objects accordingly to the designed plan and check if they correctly fit into the environment.

Lights

Even if the virtual models are placed in the real environment, they are not affected by the real illumination. For this reason adding virtual lights in the space is necessary to increase the coherency between real and digital and the feeling of realism in general. In the developed application it is possible to load different configurations of light through the GUI. In this way, the designer is able to evaluate also if the illumination can impact the final layout.

9.3 AR Space Planning

The space planning activity in the AR environment consists of the interaction with the digital objects previously described. The user visualizes the environment to augment though the AR Client display and he performs the design of space by means of AR Clients GUI. The space planning activity in the augmented environment is schematized in Figure 9.5.

Once he decided the desired piece of furniture in the database, he can place it in the environment. The virtual object is visualized in front of the camera point of view, to a distance of 1.80 meters. Afterwards, the user has to correctly place the object in the desired location; this task is performed by some interactive solutions that are described as following. If the virtual object does not match the designer's expectation, it is possible to remove it from the augmented environment. Once the designer finishes with one object, he can get through the next one. Finally, the designer can save the layout during the design process. He can store different configuration and quickly switch from different solution. In this way, it is possible to show the results to customers or keep on working on a previous space plan.

9.3.1 Objects Highlighting

When more than one virtual object is placed in the environment, a way to recognize the object, which user is working with, is due. Thus, once the user select the object by clicking it in the scene, the

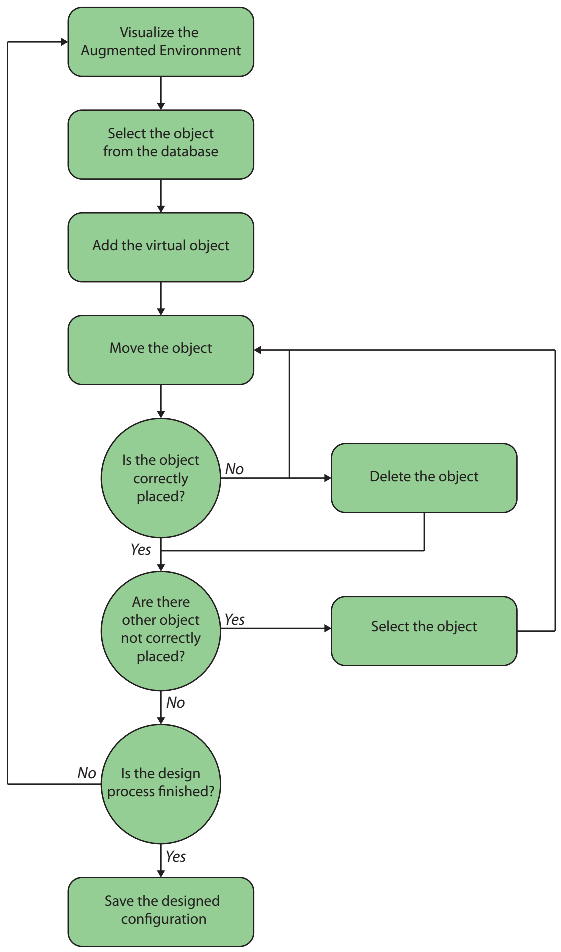
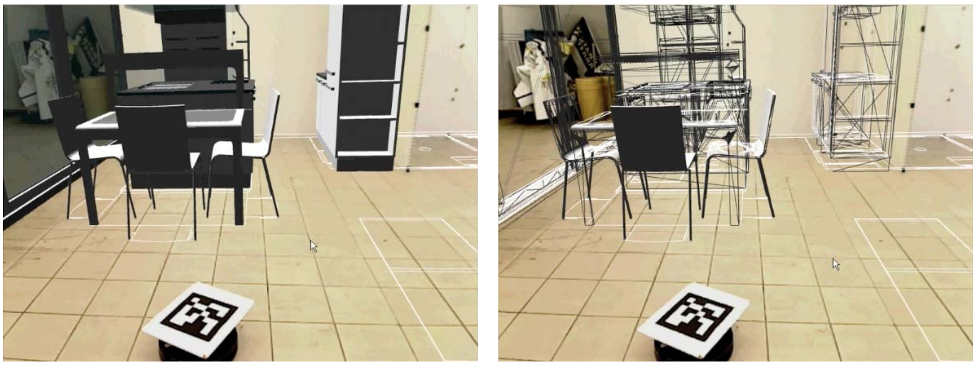


Figure 9.5: *The space planning workflow.*

Table 9.1: *Interaction modalities features. The main features of the proposed interaction modalities are evaluated in a qualitative way. The evaluation is described by the symbols* plus *(+),* minus *(−) or* neutral *(o) according if the feature is good, bad or not relevant.*

Interaction Type	Precision	Quickness	Usability
Buttons	+	-	o
Moving Trolley	-	+	+

system highlights it. As it is possible to see in Figure 9.6, the other objects, which are not selected, are represented in wireframe mode.



(a) (b)

Figure 9.6: *Object highlighting. If the user does not select any particular object, all the virtual pieces of furniture are correctly visible (a). When an object is selected (b), all the others are visualized in wireframe mode.*

9.3.2 Interactive solutions to move objects

We developed two interaction modalities to modify the position of virtual objects. Both of them has pros and cons. The interaction modalities are performed by mean of:

- buttons,
- moving the trolley.

Their main features are summarized in Table 9.1.

Buttons

The first typology of interaction that has been developed grounds on the buttons of the GUI. As it is possible to see also in Figure 9.3, six buttons (three for each axis) allow the user to modify the position and orientation along the three reference axis. The numeric values of position or orientation are increased or reduced according to an incremental value that can vary from $1mm$ to $1000mm$ for position and from $1°$ to $10°$ for the rotation.

This is an interaction solution easy to understand that allows the user to control very precisely the position of the virtual object. However, since the user has to deal with several buttons, this is not a fast way of interaction.

Moving the trolley

We developed another interactive approach to move the virtual object by using the trolley itself, since the AR Client is placed on it. Once the object is selected, it follows the movements of the trolley, which are performed by the user.

This is a very intuitive kind of interaction, since it grounds on the interactive metaphor of the trolley to carry objects. Moreover, the trolley movement enables the user to perform very quick placements of objects. Unfortunately, it is not possible to carefully control the position of the trolley, consequently the final position of the virtual object could not be very precise.

9.4 Evaluation

The process of space planning can be applied to different kind of environments, according to their final purposes. For this reason, the proposed system has been evaluated in three space-planning contexts

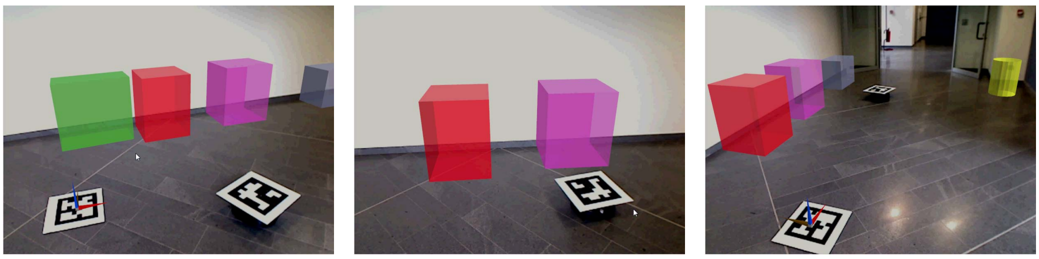
- Industry,
- Interior Design,
- Exibition,

which are described as following.

Industrial Context

The AR space planning system can be used to evaluate the encumbrances of machines in a factory. A virtual model of each machine can be placed in the environment and the final layout can be visualized before buying all industrial equipment.

The virtual model of the object to place is not available sometimes, but is it still possible to use the system to evaluate the encumbrances. Actually, it is possible to visualize a parallelepiped with the same size of the real object, as in Figure 9.7.



(a) (b) (c)

Figure 9.7: *Example of using the system for encumbrances evaluation. It is possible to represent some simple geometries, like a parallelepiped, to represent the volume of an object in the environment.*

Interior Design

In case of interior design, the system is a valid support for the aesthetic and functional evaluation of the layout of furniture in a room or house. In this way, it is possible to see in the real environment what the architect designed on the map. Moreover, the position of pieces of furniture can be rectified. Figure 9.8 shows a testing session performed at the Interior Design Lab of Politecnico di Milano, wherein an entire apartment has been evaluated by the system. As it is possible to notice, the interior designers dealt with a map of the apartment on the floor and they inserted the furnishings in their right location.

(a) (b) (c)

Figure 9.8: *Example of using the system for interior design.*

Exhibition

During the preparation of an exhibition, the organizers usually deal with works of art, which are very precious. Moving these objects is a very careful task because they are sometimes fragile and often expensive. As it is possible to see in Figure 9.9, the system has been used at Fondazione Michelangelo Pistoletto to design an hypothetical exhibition and evaluate the final result by means of virtual models of the statues. Thus, it is also possible to modify the position of the digital model or try some options without any risk for the real statues.

(a) (b) (c)

Figure 9.9: *Example of using the system as support for exhibition planning.*

9.5 Conclusion

In this Chapter, we described an AR system to support design of space. The system allows the user to visualize and manage the layout of virtual objects in a wide empty environment. It relies on the tracking method described in Chapter 6 so that it is possible to track the camera point of view in a wide environment. Moreover, we developed a GUI to visualize and interact within the Augmented Environment, in particular two interaction metaphors to deal with virtual objects.

The system has been validated in varying contexts of use, showing the versatility in working fields which are different to each others.

CHAPTER 10

Interactive AR Support for Manual Tasks

In a factory, many activities are still carried out manually by an operator. In particular, manual operations are required in case of tasks that are not frequent, complex or if they require some problem-solving skills that machines cannot do. Manual operations can be found in assembly phases and in maintenance. In the former one, manual operations are required during not highly automated tasks, like specific products that are not frequently produced or with a very small and specific market, which cannot justify the investment of a production line. In the latter case, there is a high variability of situation to solve when an object is damaged, for this reason operators skills and knowledge are fundamental to repair it.

In order to carry out manual tasks, the operator requires training or a guide support. This second option is usually preferred in case of objects that the operator rarely deals with, and it usually consists of an operating handbook.

AR Support

This kind of manual operation can be improved by means of AR, which integrates information about the task to perform into the working space of the user. The performance of AR as support for assembly and maintenance tasks has already been demonstrated in [61, 132, 173], but only with simple objects or tracking solutions that cannot be used in an industrial context. In this Chapter we evaluate the performance of AR as support of assembly and maintenance tasks, in particular, focusing on the industrial context related to circuit boards.

Tracking

In order to give digital information directly on the circuit boards, a tracking solution is necessary to recognize which circuit board the user is working with and its pose in the environment. Unfortunately, it is not possible to deal with markers. In fact, markers embedded on the board would entail a modification of the design and production of the board. Therefore, a tracking solution that grounds on Natural Features is exploited. This solution is the one described in Chapter 7.

10.1 System Description

The system for the support of manual task is oriented to long usage in an industrial context. Thus, we discarded the use of an immersive visualization because the HMDs available so far cannot be used for a long time. Actually, when the user wears one of them, he gets tired quickly so that it is not possible to use immersive systems for an entire working day. Therefore, the system adopted is constituted by a monitor-based AR solution.

10.1.1 Structure

The structure of the system is a workbench with a camera and a video display, as shown in Figure 10.1. The camera is a Logitech Pro 9000 and it is fixed on a trestle in order to frame only a portion of the working area. The monitor is set to a resolution of 1024×768 and it is built in the workbench. The operator can see the augmented information trough the monitor, which has been arranged to relieve strain for the operator during the use.

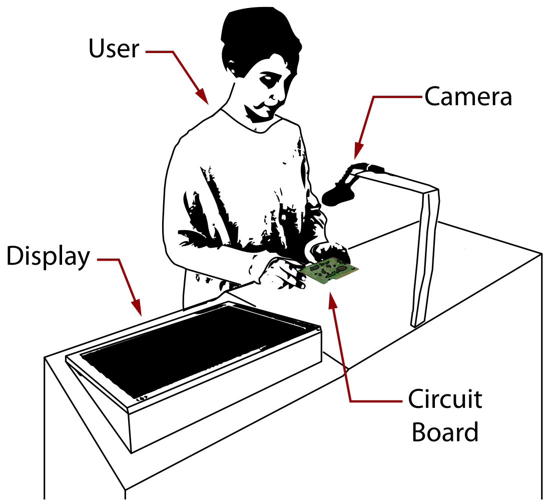


Figure 10.1: *Representation of the system. The system is mainly composed by a camera to frame a portion of the workbench and a monitor to represent the AR instructions.*

10.1.2 Software

The software developed is able to load the data related to several classifiers, which have been created previously offline, and tracks the circuit boards. The system allows the operator to switch from one circuit board to another one.

By means of the tracking data, the software is able to superimpose virtual objects on top of the circuit board, as shown in Figure 10.2. The objects under discussion are

- virtual arrows, which point at the area of the circuit board to work on
- circles, which mark the position of the hole wherein insert the electronic component
- text messages, which are related to the indication to accomplish the task.

These virtual objects are easily managed by means of some configuration files.

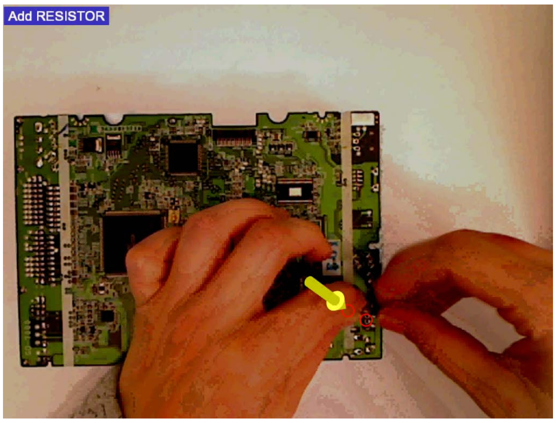


Figure 10.2: *The system developed is able to track the circuit board and consequently superimpose instruction on top of it. Here the system indicates where the components should be inserted.*

10.2 Evaluation

We demonstrated the benefit of AR as interactive guide for manual tasks by means of an evaluation phase in which the work with system proposed is compared to the one supported by an operating handbook.

10.2.1 Hypothesis

During a traditional assembly or repairing manual task, the first task that the user should do is to find the point to work on. Thus, the operator should find some common reference points among the physical object and its representation in the operating handbook. Once the operator has a mental map of the object, he usually adjusts the position of the real object as the orientation of the one represented in the manual. In case of AR support, instead, the mental demand to comprehend the object is not necessary because the proposed system allows the user to see information on top of the object, which are always correctly oriented for any position of the circuit board. For this reason, when the user deals with AR, he does not need any reference point and the instructions should be more intuitive.

Another problem during a work with circuit boards is that some tasks can be more difficult than others. Actually, some areas of the circuit board that the user has to focus on can be complex, due to many holes that create some symmetries or patterns. Since the user has to accomplish the task, a standard solution bases on using reference points and mental structures. In AR this problem is lightened because the system directly indicates on the object the hole to deal with.

An operating handbook delivers only information and it is not able to give a reply about the status of the work. By means of computer vision tools it is possible to provide some feedbacks and in this system are represented in a graphic way. In this evaluation test, the feedback of the task is not provided for all of tasks.

According to the assumption previously described, the hypotheses that the we want to check by the test are:

H1 : AR support for manual tasks reduces the total amount of time to accomplish the task compared to a traditional support;

H2 : AR support reduces the time to execute difficult operations compared to the traditional support

H3 : the mapping process that the operator should perform when dealing with the handbook is a time consuming phase

H4 : the feedback, given by the AR system, about the correctness of the operation is useful to increase the performances of the whole task.

10.2.2 Description of the test

We conducted a test between subjects. We involved the user to follow some instruction to manually deal with some components on circuit boards. According to the information provided, the user should:

- pick up the correct circuit board from a group of several circuit board
- insert LEDs, resistors, capacitors and cables into the correct hole(s) shown by the support

We proposed three methods to support the user work:

- AR Information about the current task are directly superimposed on the circuit board and displayed on the monitor.
- Traditional Operating manual. On each page the task to perform is represented by a picture.
- Traditional Flipped over. Like the previous one, but some images are sometimes turned by 180 degrees.

Appendix B reports the manual used for the test. It is worth nothing that in the AR support case, the user receives an immediate feedback if the circuit board to deal with is the correct one or not. Finally, the last method is to analyze if breaking the correspondences object-image is an issue for the user.

Working area

The working area for the evaluation test is represented in Figure 10.3. The user works standing in front of the workbench. All circuit boards and the component boxes to use during the test are placed on the left of the participant. Since the test does not require any particular skills and the user does not have to deal with welder or any other particular tool, every circuit board was glued on a base of Styrofoam. This solution enables the user to insert the components into the holes and fix them in a easy manner.

On the right side, the support for the manual operations was placed, which could be the AR based display or the paper handbook. The camera is used only in case of AR support.

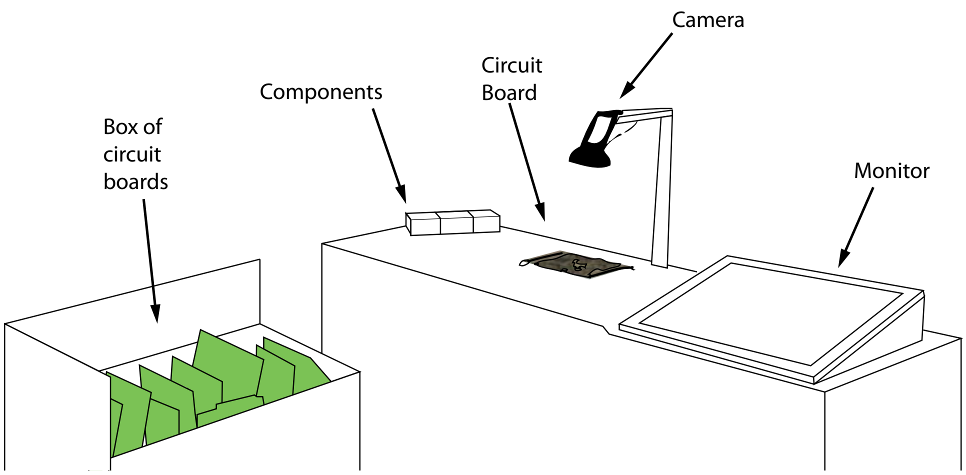


Figure 10.3: *Representation of the working area for the evaluation test.*

Procedure

The participants involved in the test were 60 (82.5% men, 17.5% women, age 26.15 ± 5.26), equally distributed on the three methods to support manual tasks (20 with AR, 20 with the operating manual and 20 with the manual flipped over). Before the main test, each participant has time to get familiar with the task, by completing some operations on a trial circuit board. When participants felt comfortable with the trial board, they can proceed to the main test.

The main test deals with 5 circuit boards, which has to be chosen in a group of 11 different ones. The selection of the circuit board was shown by the support used for the test. In case of AR support, the participants have a direct feedback if the circuit board was the right one by superimposing a blue square, with written "Found!", on top of the object as it is possible to see in Figure 10.4.

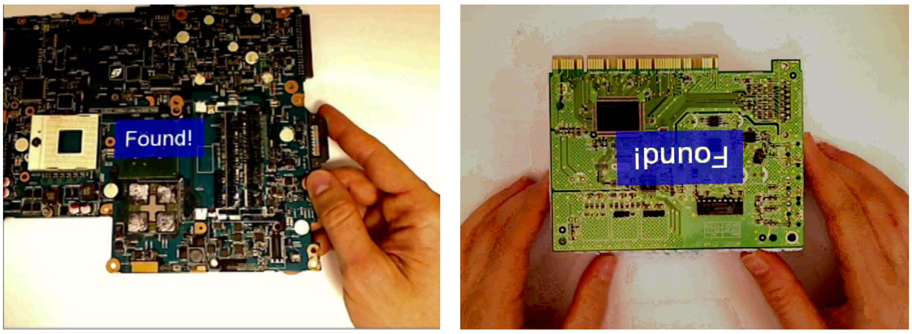


Figure 10.4: *The system developed place a blue square with written "Found!" on top of the circuit board. This is a useful feedback when the user has to chose the correct circuit board.*

We asked each Participant to accomplish each single operation being careful, in order to do not make mistakes, and being quick at the same time. The total number of operation to perform is 34, distributed on the 5 circuit boards. The participants were able to switch to the next operation by voice commands in case of AR support and by turning the page of the handbook in the other two cases.

Classification of difficulty

In order to measure the increase of performances during each single operation on the circuit board, we classified every operation to a value of difficulty, according to the Table 10.1.

Table 10.1: *Value of classification for each task.*

Difficulty Value	Description
0	Find the circuit board / Insert a connector
1	Insert a component into one hole, which is isolated
2	Insert a component into two holes, which are isolated to the other holes on the board
3	Insert a component into one hole, which is in a pattern
4	Insert a component into two hole, which is in a pattern
5	Insert a component into two hole, which is in a complex pattern (more than 5 holes close together)

Measurements

For each test, we took into account two types of measurements: time to accomplish each operation and the cumulative number of errors made during the test. The former one is measured as the time between to consecutive operations. Thus, we took the time between two vocal commands, in case of AR support, and between two consecutive turnings over the pages, in case of the operating manual. The latter one takes into account the errors that the participant could make, such as picking and using the wrong component for an operation or inserting the component in the wrong hole.

10.2.3 Analysis of the results

We analyzed the data regards to time to accomplish each operation and mistakes done by the participants by means of statistical tests. The purpose of the tests is to assess if there is significantly difference between the results with the three methods.

We conducted the tests by the analysis of variance (ANOVA). This analysis assumes that data are distributed on a Gaussian (normality condition) and that the variance is equally distributed (homoscedasticity of data). Since the data were not normally distributed, we had to transform them. We found out that the logarithm of the data satisfied the request of normality.

Finally it is worth nothing that the mistakes carried out by the participants are very few and distributed equally in the three methods, as it is possible to see in Table 10.2. For this reason, we discarded them from the analysis, since it is not possible to perform any kind of useful test on them. This issue is due to the test itself, which can be considered simple and feasible for all of the participants without any particular effort.

Table 10.2: *Average number of errors for each assembly methodology.*

	Errors		
	AR	Trad.	Flip.
Mean	0.50	0.60	0.35
Std	0.76	0.88	0.59

Time taken to accomplish the task

The first analysis that we performed is about the time taken by the participants to complete the whole test. As it is possible to see in Figure 10.5 and in Table 10.3 , the mean time to accomplish the test by means of AR support than the other two. In particular, the time is reduced of almost 20%.

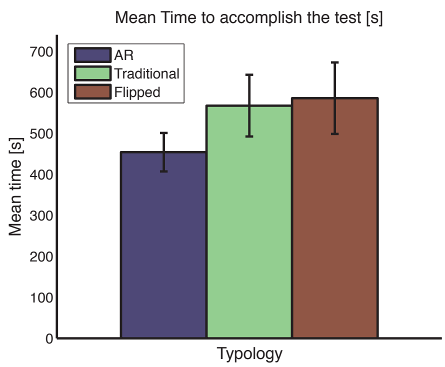


Figure 10.5: *Bar graph of the mean time taken to accomplish the test according to the three different supports.*

Table 10.3: *Mean time taken to accomplish the test according to the three different supports.*

Time for the whole test [s]			
	AR	Trad.	Flip.
Mean	454.01	567.49	585.78
Std	46.99	75.33	87.18

We discovered by a one way ANOVA, which factor is on 3 levels (AR, Traditional and Flipped Over), the means are significantly different among the levels ($F = 21.9748$, $p = 8.4225e^{-8}$). Moreover, we verified that the mean time of the tests performed in AR is different to the other two by a t-test on the means.

Time to find a circuit board

Since in the AR support the participants had a visual feedback of the correct circuit boards, we analyzed the time taken to find the circuit boards. Figure 10.6 and Table 10.4 show the mean time used to find one of the 5 circuit boards involved in the test is lower in AR support.

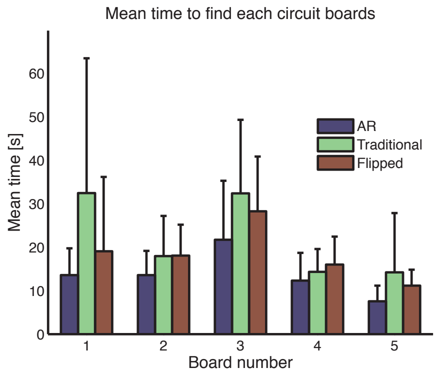


Figure 10.6: *Bar graph of the mean time to find a circuit board.*

Table 10.4: *Mean time taken to find a circuit board.*

Board	AR	Trad.	Flip.
1	13.58 ± 6.17	32.47 ± 31.03	19.07 ± 17.12
2	13.58 ± 5.58	17.96 ± 9.26	18.09 ± 7.09
3	21.72 ± 13.59	32.40 ± 16.92	28.28 ± 12.60
4	12.31 ± 6.4	14.32 ± 5.27	16.01 ± 6.44
5	7.55 ± 3.61	14.23 ± 13.64	11.15 ± 3.68

We verified this by a one way ANOVA, analyzing the total time during a test to find the circuit boards. The factor of ANOVA is the same of the previous analysis. It turns out that there is significantly difference between the means ($F = 14.1631$, $p = 1.0153e^{-5}$). Finally, verified by a t-test that the mean with the AR support is different to the other two.

Time to accomplish each operation, according to the difficulty

The third analysis that we carried out aims at verifying the performances of the participants with the three methods during manual task with various difficulties. As it is possible to see in Figure 10.7 and

Table 10.5 , the AR support leads to lower execution times for each level of difficulty. It is worth nothing that there are no values for level 3 of difficulty. In fact, we did not find any operation for this level of classification.

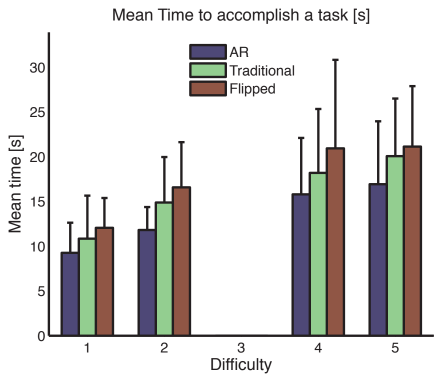


Figure 10.7: *Bar graph of the mean time to accomplish a task, according to the difficulty.*

Table 10.5: *Mean time to accomplish a task, according to the difficulty.*

Difficulty	AR	Trad.	Flip.
1	9.26 ± 3.36	10.83 ± 4.81	12.04 ± 3.33
2	11.80 ± 2.56	14.87 ± 5.08	16.56 ± 5.05
3	$NaN \pm NaN$	$NaN \pm NaN$	$NaN \pm NaN$
4	15.78 ± 6.30	18.19 ± 7.14	20.93 ± 9.89
5	16.92 ± 7.02	20.06 ± 6.43	21.12 ± 6.76

The greater performance of AR support regards the other two is then confirmed by a two ways ANOVA. The two factors taken into account are the kind of support (AR, Traditional and Flipped Over), and levels of difficulty of the operations. The results indicate that both the first factor ($F = 72.3751$, $p = 0$) and the second one ($F = 315.1767$, $p = 0$) are significant for the test, while the interaction between them does not describe the variance ($F = 0.8333$, $p = 0.5440$). Finally, by analyzing the three methods for assembly, we can state that

difficulty 1 means significantly different between them

difficulty 2 means significantly different between AR and the other two supports

difficulty 4 means significantly different between AR and the Flipped Over support

difficulty 5 means significantly different between them

10.2.4 Discussion

Thanks to the results obtained during the tests, we noticed that using AR as support for manual operations increase the performances of the user. In particular, AR support reduces the working time: the hypothesis H1 is supported.

AR is also useful for complex operations, since we demonstrated that time to accomplish a single operation, even a difficult one, is less than the other two methodologies. Actually, the user finds the indication directly on top of the real circuit board, at the point wherein he should work. In this manner, it is required a minor the mental work. The hypothesis H2 is supported.

All the results point out that dealing with the flipped over operation handbook extends the working time. Indeed, the user has to control the references between the circuit board and its image in the manual.

This mapping process is time-consuming and decreases the performances. AR support is the fastest one, since data are directly superimposed on the object and they are always coherent with it. The hypothesis H3 is supported.

To conclude, the results shows that finding the correct circuit board is faster by means of AR. Actually, the user has an immediate confirm about the circuit board by a visual feedback from the system. The hypothesis H4 is supported.

10.3 Conclusion

In this Chapter we described and evaluated a system to support manual tasks on circuit boards by means of AR. The system exploits the tracking approach reported in Chapter 7, which works with only an orthographic image of the circuit board. In this way, we can have a system that can provide information for assembly or maintenance purposes on top of the board.

Finally, we evaluated the system comparing it with traditional approaches to support manual operations. From the results of some statistical analysis, we noticed an increase of the working performances when the user deals with the AR support.

CHAPTER 11

Interactive AR Framework for Inspection purposes

Sometimes, it is useful to know the condition of an object during its working life, in order to forecast damages or plan its maintenance. Therefore, some sensors are placed on the object to acquire data regarding its status. A system usually samples the data from sensors and records them. Finally, the user can look trough these data in real time or after the recording of samples, if the visualization requires some mathematical tools for the analysis.

The visualization of the data is usually performed by representing them in a textual or graphic fashion on a monitor. Besides data from sensors, it is possible to apply some mathematical tools to the data in order to estimate information about the object state and them allows the user to visualize them.

In this Chapter we describe the development and evaluation of a framework for inspection purposes on an object by improving the visualization of these data trough AR. In this manner, the user can see the object and, at the same time, contextualized data from sensors. The system is oriented to work with more visualization devices at the same time, in particular, mobile devices. The proposed solution, indeed, needs devices equipped with a camera and a wireless connection to provide augmentation. Technological devices as tablets and smart phones satisfy this requirements nowadays, they are socially accepted and available for a small price. Thus, they are good candidates for inspection purposes.

Mobile AR for visualization of sensor data has been reported in [73, 168] already, but they have just evaluated the feasibility of this technology. In this Chapter, we describe a framework adaptable to different typologies of inspection purposes. The system is also able to represent information about some characteristics of the object by analysing data from sensors. Finally, we report an initial evaluation of its use in a real industrial context.

11.1 System Description

We developed the system by dividing it in modules, as illustrated in Figure 11.1. Every module has a specific task and we established a data exchange among them in order to finally provide an augmented visualization of the data from sensors on devices located remotely. Moreover, it is possible to modify every single module without altering the entire system. In this way, the system is versatile and can be easily customized for different kinds of inspection purposes.

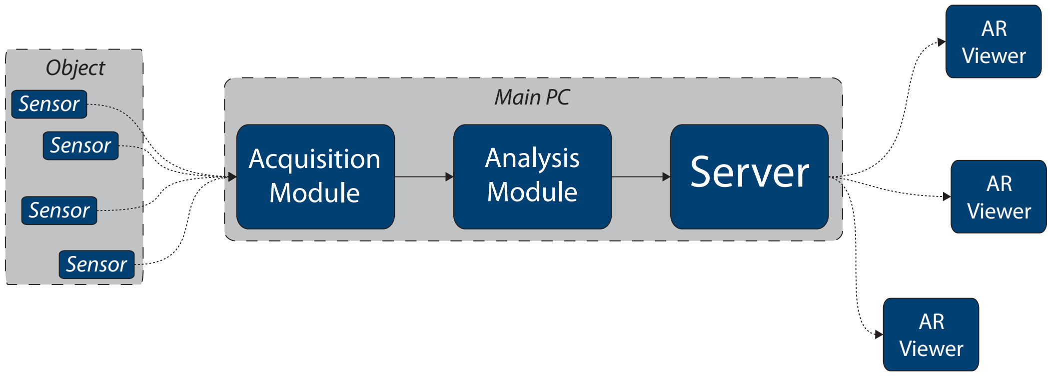


Figure 11.1: *The system framework. Data from some sensors placed on the object are acquired by the Acquisition Module and then analyzed by a dedicated module in order to estimate the object state. All the data are then taken by the Server and distributed to all the AR Viewers connected.*

11.1.1 Acquisition module

Sensors are placed on the object to inspect and they send signals to the Acquisition Module. The task of the module is to acquire the signals and to make them available for the next visualization and analysis. For this reason it has to convert the acquired data in measured values and write all the useful data in a text file.

11.1.2 Analysis module

The measurements from the Acquisition Module are then picked up by the Analysis Module to evaluate the state of the object. We developed this module by using Matlab and it executes the following tasks:

- get the measured values from the Acquisition Module
- analysis of the measured values
- estimation of state of the object
- communication of data to the Server

We developed the tasks regarding to analysis and estimation as two separated functions. In fact, we should change the functions according to the object under inspection, the kind of inspection and the sensors configuration. The analysis function is an analytical or numerical model that describes the state of the object by the data from sensors. The output data are saved in a text file, in order to be accessible to the next module.

11.1.3 Server

The task of the server regards taking the data about measurement and about the state of the object and making them available for remote AR application. The communication of data to remote applications is performed by a TCP/IP socket connection in a local wifi network. The server sends all the data to any connected client; in this way, the users can be located on different PC and perform inspection in AR.

The server reads the data, which has been saved by the Data Analysis Module in the text files, every a period of time; the timing can be selected by the user. After reading data, the Server makes them available on the network.

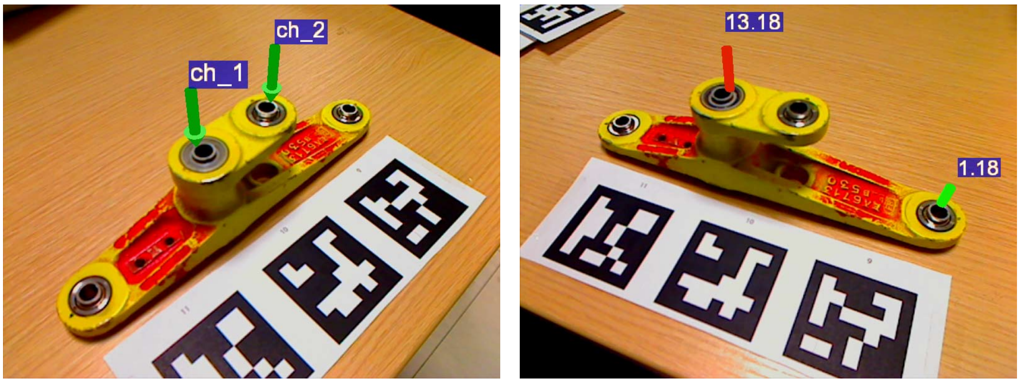
11.1.4 AR Interface

The AR Interface is composed by a client, which receives data from the server, and by a GUI, which allows the user to visualize the data in AR by a marker-based tracking system. The tracking system used in this work is ARToolKit Plus [160]. Two configuration files allows the user to easily modify the interface according to the object and inspection purposes. In fact, one configuration file describes the position of the sensors on the object, while the other one the position of the markers. Thus, every time we launch the AR Client, it reads the two files and it dynamically creates the virtual objects to visualize.

The objects, represented by the AR Interface, change dynamically according to the data from the Server. These objects are:

- 3D arrows, which point at the position of the sensors. On top of it the name of the sensors is displayed (Figure 11.2(a))
- 3D bars, which show the values acquired by sensors (Figure 11.2(b)). Each bar is placed on the corresponding sensor and it changes dimension and color according to the value measured. The measured value is also displayed on top of the bar in a textual way.
- critical points of the objects, according to its state, which are represented by a red area on the object.

It is worth nothing that all the objects are modified in real time as soon as the AR interface receives data from server.



(a) (b)

Figure 11.2: *The system allows us to place virtual information on top of the real object. These pieces of information are (a) arrows, indicating the name and position of the sensors, and (b) the value from the sensor, represented in a graphic and textual manner. The object in the images is used only for testing and it is without sensors.*

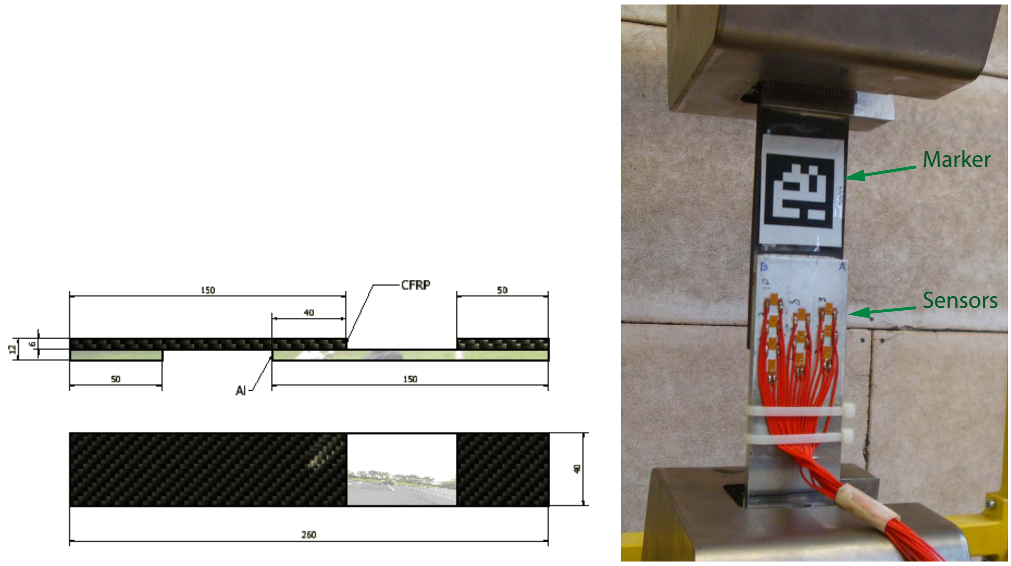
The GUI allows the user to visualize only some of the digital objects in the environment by selecting or deselecting them in a window.In this way, it is possible to visualize only the most important data. Otherwise, in case the data are too close to each other and the visual representation is not clear, it is possible to visualize only some of them at once.

11.2 Evaluation

We carried out an evaluation study with some expert users in order to check the achievability of the framework. The evaluation consists to the data visualization during a fatigue test of a mechanical component. During a fatigue test, the object under inspection is subjected to dynamic loads.

The object

The object taken into account for the fatigue test is the specimen in Figure 11.3(a). This one is composed by an aluminum bar and a carbon-fibre bar, which are glued together. The most critical part of this object is the overlap area of the two bars, so that a crack could start here. For this reason we applied 10 sensors on this area; the sensors are strain gauges from HBM series 1-LY13-0.6/120. As shown in Figure 11.3(b), we applied a marker on the specimen in order to perform tracking.



(a) (b)

Figure 11.3: *The object used for the system evaluation is a hybrid bar specimen, which dimensions are represented in (a). We placed 10 sensors and a marker on it (b).*

11.2.1 Setup Description

Acquisition Module

We acquired the data from the strain gauges through two control units HBM Spider 8 and by the software HBM Catman. The machine that applied the loads is a servo-hydraulic system MTS Landmark with a loading cell of $100kN$. During the test, the machine applied traction forces from $0kN$ to $2kN$ by steps of $0.5kN$. For each step the acquisition module took the values from the sensors with a frequency of $10Hz$ for 1 second.

Analysis Module

The analysis module takes the data acquired and it estimates the position of the crack in the specimen. We modelled the position of the crack by some numerical simulations the specimen under loading. From these analysis we discovered that the position of the crack p has a linear correspondence with the position of minimum strain in the specimen,

$$p = min(X_{strain}) + 1.9758. \tag{11.1}$$

We entered the Equation 11.1 in a Matlab function. In Figure 11.4 the interface of this module is presented.

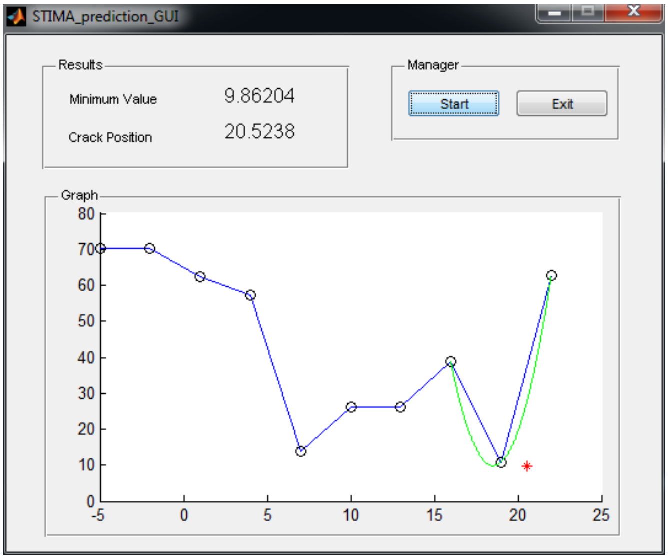


Figure 11.4: *The MATLAB interface. The interface represents the values from the sensors in a graphic manner and it estimates the crack position, which is displayed as a red dot.*

AR Visualization Module

The AR visualization module represents the data from Server on the positions described by the configuration file. Besides these kinds of visualization, we add the chance to visualize the data in a graph as well. As it is possible to see in Figure 11.5, the data from sensors are represented by white points. A red stripe passes trough the points and it is described by an interpolating spline function.

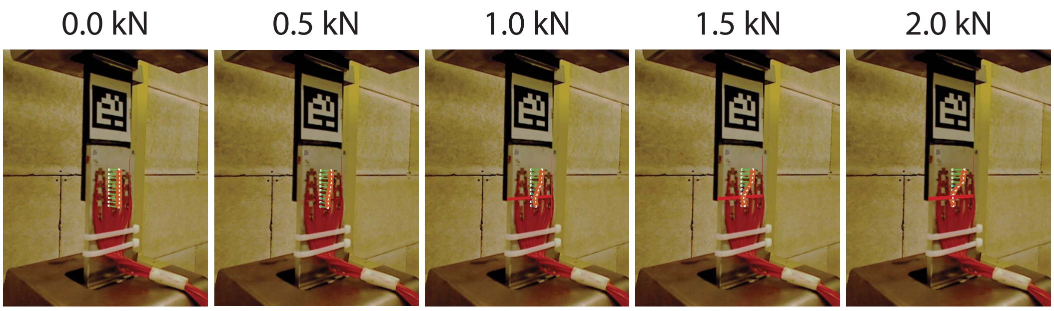


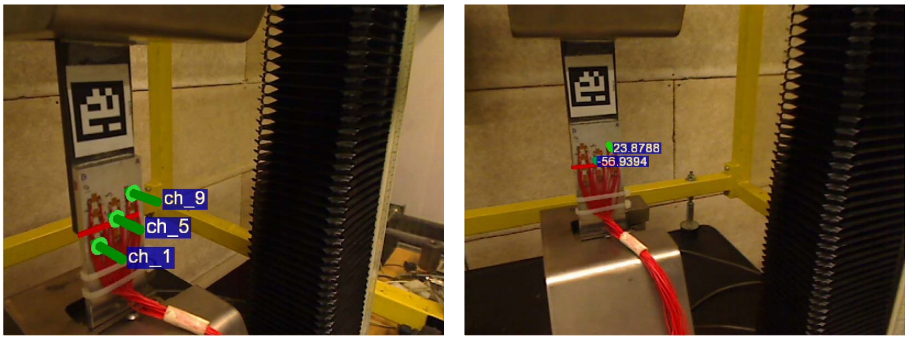
Figure 11.5: *For this specific evaluation we decided to represent a graph about data from sensors on top of the specimen as well. The sequence of pictures is at different loadings, in order to show the trend of the data during the traction of the specimen.*

11.2.2 Description of the Test

We installed the visualization system on three different devices, in order to check the functioning in different scenarios. The devices taken into account are:

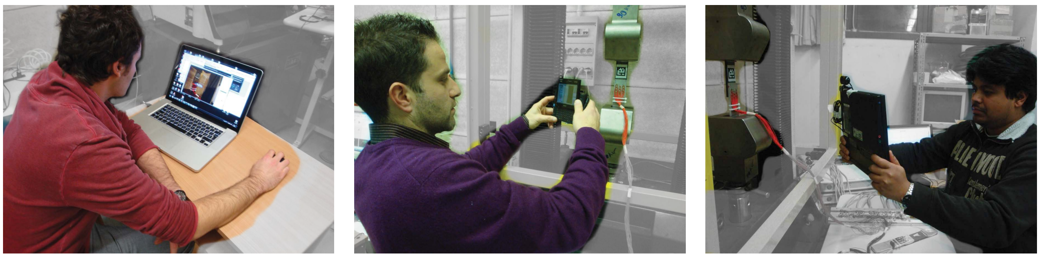
- a laptop, in order to simulate the visualization system from a fixed workspace. The camera is mounted close to the specimen.
- a tablet PC and a Ultra Mobile PC, in order to simulate the scenario of occasional monitoring of the object state. The user carries these devices with him and he observes the object from different point of view.

We involved four users to test the system. We asked them to monitor the status of the specimen by switching through the different visualization modalities, as shown in Figure 11.6. The user dealt with all the three devices (Figure 11.7) and, in case of mobile ones, they were free to move in the environment around the specimen.



(a) (b)

Figure 11.6: *The representation of the virtual objects on the specimen during the evaluation phase. On the left, the arrows indicates the position of some sensors (the number 1, 5 and 9 in this case), while on the right the system shows the data from them. The red stripe on the specimen, which it is possible to see in both the images, corresponds to the crack position.*



(a) (b) (c)

Figure 11.7: *Users during the evaluation with different visualization device: (a) laptop, (b) UMPC and (c) tablet PC.*

In order to get some feedback from the users experience, we asked them to fill a questionnaire, which is also in Appendix A.2. The results of the questionnaire are in Table 11.1.

11.2.3 Analysis of results

The positive feedbacks from the test with expert users show the effectiveness of AR for inspection purposes. Regarding to visualization, we obtained positive results for both AR monitor and AR mobile devices. Only in question 2 we can notice a difference between the two visualization approaches. Actually, the environment understanding is difficult in case of AR Monitor, because the user looks through a fixed point of view. A mobile device, instead, allows the user to better understand the environment because he can move around. One problem the users noticed is in the crack visualization, since they consider important to have more information about it, such as its dimension.

The AR monitor turns out to be better for what concern interaction. In question 1, AR monitor allows the user to sit down so that he is in a position that relieves stress; in mobile case, he holds the device and has to move along. Moreover, as it is possible to notice from question 3, the mobile devices are

Table 11.1: *Users' answers and average score.*

	Question	AR Monitor	AR Mobile
Visualization			
1.	The system allows me to correctly see the object under inspection	5.50	5.75
2.	The system allows me to comprehend the environment where the object is placed	3.50	5.50
3.	The system allows me to understand the position of the sensors on the object	5.00	5.25
4.	The system gives me useful data from the sensors	5.50	5.25
5.	The graph gives me a clear idea about the state of the object	5.50	5.50
6.	The element that identifies the crack gives me a clear idea about the object state	5.50	5.50
Interaction			
1.	The system is comfortable	5.50	4.25
2.	The system allows me to easily control the visualization of the data from the system	5.25	4.75
3.	The system works in real-time	5.75	3.75

not completely reactive in real-time. Actually, the mobile devices used for the test are quite old and for this reason they have limited performances. They work at $10Hz$, which is the lower limit for a fluid interactive interface.

11.3 Conclusion

In this Chapter we described a versatile system that allows us to inspect objects equipped with sensors by means of AR. The system is composed by modules, each of them can be modified according to the typology of inspection to perform. Moreover, the system is addressed to work also with mobile systems, which are more and more present on the consumer market.

We performed a pilot study with some users in order to evaluate the system. The study consist of inspecting a specimen of composite material during a fatigue test. By the results of the study we can state that the system is effective and usable by not AR expert users.

Part IV

Conclusion

CHAPTER 12

Conclusion

The aim of the thesis has regarded to the analysis of low-cost technologies integration in AR supports for Industry. These technologies are embedded in devices that are available on the consumer market and used for daily life purposes. In this dissertation, they have been exploited to solve some technological requirements or replace some expensive solutions for using AR within an industrial context.

As first step, we analyzed the technologies for AR that are available nowadays in Chapter 1. We classified them as Visualization, Tracking and Interaction technologies and we evaluated pros and cons of the most important of them that have been proposed so far. Then, we analyzed the background of AR used in the industrial field in Chapter 2. We carried out the analysis by dividing the performed researches in groups, according to the industrial phase to which each research belongs. The groups correspond to the five main phases of the lifecycle of a product: Product Design, Manufacturing, Commissioning, Inspection & Maintenance and Decommissioning. From this analysis it turns out that the majority of these works have been done only to verify the feasibility of AR in Industry and they have never been brought to being produced and commercialized.

The main problem of deployment of an AR application for the industrial fields is due to the technologies used. As we described in Chapter 3, the industrial context usually expects some strict demands, so that AR technologies proposed so far are sometimes limited or too expensive if exploited to satisfy all the industrial requests. Thus, we propose to integrate low-cost technologies, which come from the consumer field, to replace the standard ones. These kinds of low-cost technologies are actually robust, already integrated in a device, easily available and cheap. For this reason they could be good candidates for AR in Industry, since they could overcome some issues of actual technological solutions.

However, we should adapt these technologies for our AR reasons, because many of them are addressed to the everyday life use. Therefore, we developed some new algorithms in order to manage them for the new purposes. The new algorithms mainly focus on exploiting the technologies for tracking, in particular to lighten some problems about structuring the working environment or objects to deal with and facilitating the initialization of the AR application.

The development of these algorithms is reported in Part II of this dissertation. In practice, the technologies used are

- a wireless hand-held device for video gaming (Chapter 5); we implemented a solution to track its position in space by infra-red emitters using the sensors which it is equipped with.

- a mobile robot designed for households (Chapter 6); we exploited its mobility to create a hybrid tracking solution for wide environments. The robot has been used to increase the working area of a marker-based tracking system. Thus, it is possible to estimate the camera position without structuring the environment.

- a planar image tracking system (Chapter 7); we extended the features of this tracking approach in order to work also with circuit boards. In this way it is possible to recognize and track the position of a board without the addition of any marker on it.

The developed algorithms have been also evaluated by comparing their characteristics to other well-known tracking approaches, in order to validate their use for AR purposes.

Finally, we reports some AR supports for industrial activities, which we created by using low-cost technologies, in Part III. Some of these applications exploit the solution proposed in Part II. In particular, we developed four applications to support specific activities:

- an interactive system for Design Review. We used the wireless device in Chapter 5 to interact with virtual prototypes during the evaluation phase of the design of a new product.

- an interactive system for space planning. We used the hybrid tracking solution described in Chapter 6, which take advantage of the mobile robot, in order to have the possibility to interact with virtual objects in wide environments.

- an interactive support for manual operations on circuit boards. We represent some pieces of information directly on top of circuit boards by means of the tracking system in Chapter 7. The AR information could be addressed to manual activities for both assembling and maintenance.

- a modular system for objects inspection. We developed a framework that allows us to represent data from sensors and the object state on all the visualization devices connected to a wireless network. In particular, the solution is addressed to common mobile devices, such as tablets and smart phones. We evaluated all the applications by means of some tests with users, in order to check their usability.

12.1 Lesson learned

By the work carried out and reported in this thesis, we can state that low cost technologies can be a valid alternative to technologies traditionally used for AR applications addressed to Industry. In particular, we noticed the following features:

Feasibility. We can adapt technologies addressed for everyday purposes in order to use them in industrial contexts. However, we need new algorithms to manage them.

Performances. Many consumer technologies offer good characteristics that make them acceptable for industrial purposes. Actually we demonstrated through some evaluation tests that the performances are less good than technologies specifically addressed to AR purposes, but still usable.

Limitation Reduction. These low-cost technologies can open the doors to new solutions for technical issues in AR. Their integration can bring useful instruments for further AR applications. As result, we have the extension of the possible AR technologies at disposal.

Usability. The users did not show any problem to use the applications. This means that the technology is still functioning and usable even if it has been converted to a new purpose. Moreover, the technological devices introduced are usable also because they come from everyday life. The users already know them and consequently it is easy for him to understand how to use the technology.

12.2 Further development

The consumer market keeps on proposing new technological products so that we can have more and more innovative devices in short time. This growth is a good signal for their use for AR supports for Industry because they can fill the gap between them and the existing reference technologies for AR. In particular, new technologies offer the possibility of new activities for the user. For this reason they can overcome some AR issues.

At this moment, the most interesting technologies to import in AR are likely related with video-games and mobile devices. As result we are having more and more hidden and accessible technologies, which are intuitive and natural.

For this reason, the next steps that we should carry out in the industrial context are the developing of technologies that take more into account the user and the environment. In the former case, we should provide solutions for AR as natural as possible, while in the latter one we should deeply analyze the restriction given by the environment.

The most recent technologies to consider for an integration in industrial application in a short period of time are related to the new generation of cheap cameras and displays. Nowadays it is possible to find high-resolution webcams that it is possible to connect to any computer. Their use could be useful for tracking, because of the sensors that can provide more information for the pose estimation. However, an adaption of the current tracking algorithms in order to deal in real-time with big images is mandatory. Another kind of interesting cameras are the so-called depth cameras, which are able to provide also information about the depth of objects from them and are already in use for video-game purposes. They could bring some advantages in tracking and interaction. Actually, we can get more information by them about the environment and about the user, without any device on him. Finally, high resolution displays are currently available everywhere, even for mobile devices. They can be coupled with high resolution cameras in order to provide a high quality AR visualization with many details.

Part V

Appendix

APPENDIX *A*

Questionnaires

A.1 Virtual Product Evaluation questionnaire

In this Section we present the questionnaire used during the user test for the AR system for Product Evaluation described in the Chapter 8.

Utente n°: ………………………………

Per ognuna delle affermazioni di seguito, seleziona il cerchio secondo te più adeguato.

Visualizzazione

1. La qualità della visualizzazione è soddisfacente.

Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d'accordo

2. La rappresentazione del prototipo virtuale è realistica.

Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d'accordo

3. Il prototipo virtuale è integrato nella scena reale.

Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d'accordo

Interazione

1. Lo spostamento the prototipo virtuale per mezzo del WiiMote è semplice.

   Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d’accordo

2. Il sistema mi permette di muovere il prototipo virtuale dove voglio.

   Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d’accordo

3. Le metafore d’interazione sono facili da usare.

   Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d’accordo

4. Sono stato in grado di imparare le metafore d’interazione rapidamente.

   Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d’accordo

5. Il puntatore virtuale è semplice da usare nell’ambiente.

   Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d’accordo

6. Il sistema permette di marcare il prototipo virtuale ovunque voglia.

   Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d’accordo

A.2 AR Inspection questionnaire

In this Section we present the questionnaire used during the user test for the AR Inspection system described in the Chapter 11.

Utente n°: ………………………………

Sistema di Ispezione:

Monitor	AR Monitor	AR Mobile

Per ognuna delle affermazioni di seguito, seleziona il cerchio secondo te più adeguato.

Visualizzazione

1. Il sistema mi permette di visualizzare correttamente l'oggetto sotto ispezione.

Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d'accordo

2. Il sistema mi permette di comprendere l'ambiente in cui l'oggetto è posto.

Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d'accordo

3. Il sistema mi permette di capire la posizione dei sensori sull'oggetto.

Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d'accordo

4. Il sistema mi fornisce chiare informazioni dai sensori.

Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d'accordo

5. Il grafico mi permette di avere una chiara idea sullo stato dell'oggetto.

Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d'accordo

6. L'elemento di identificazione dello stato della cricca mi permette di avere una chiara idea sullo stato di degrado dell'oggetto.

Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d'accordo

Interazione

1. Il sistema in generale è confortevole.

Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d'accordo

2. Il sistema mi permette di controllare con facilità la visualizzazione dei dati del sistema.

Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d'accordo

3. Il sistema è reattivo in tempo reale.

Sono pienamente in disaccordo ○—○—○—○—○—○ Sono pienamente d'accordo

APPENDIX *B*

User Manual

B.1 User Circuit Board Operating Manual

In this Section we show the operating manual that we used during the user test for the AR guide for manual assembly tasks. The test is also described in Chapter 10

Circuit Board Assembly Test

User Manual

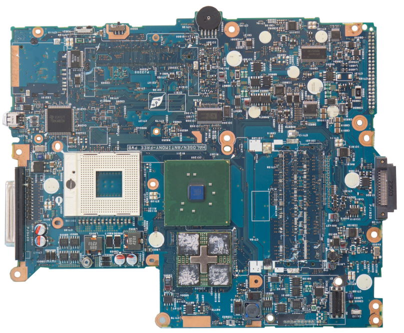
Guido Maria Re
October 2012

1

Test....

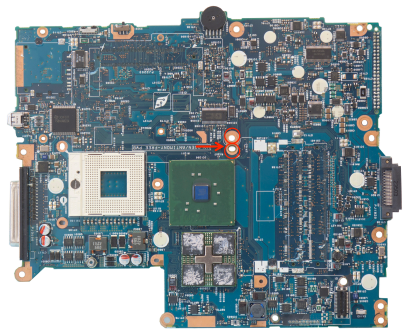
0

Take the following circuit board



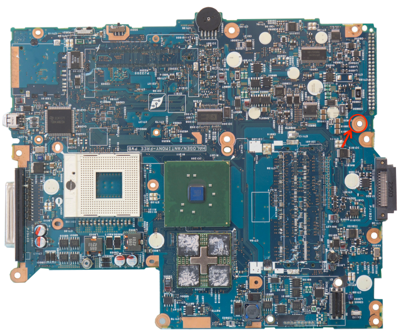
1

Insert the RESISTOR in the holes shown in the picture



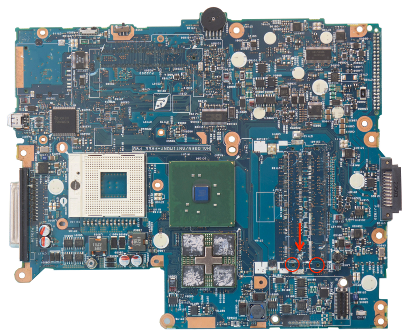
2

Insert the LED in the hole shown in the picture



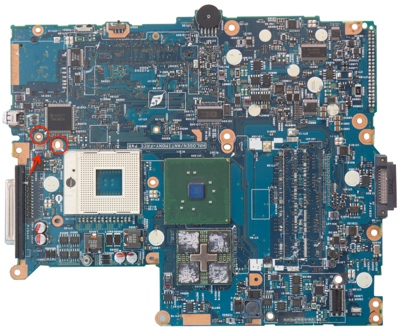
3

Insert the CAPACITOR in the holes shown in the picture



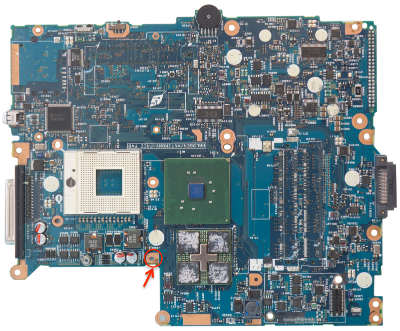
4

Insert the RESISTOR in the holes shown in the picture



5

Insert the LED in the hole shown in the picture

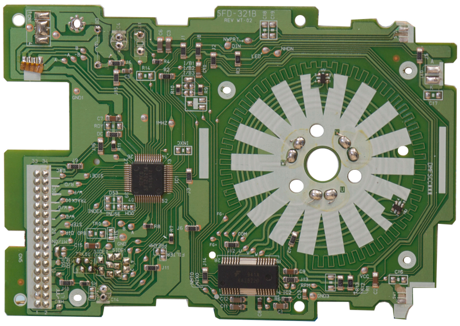


6

Finished

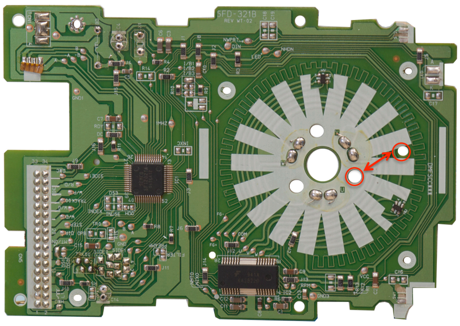
7

Take the following circuit board



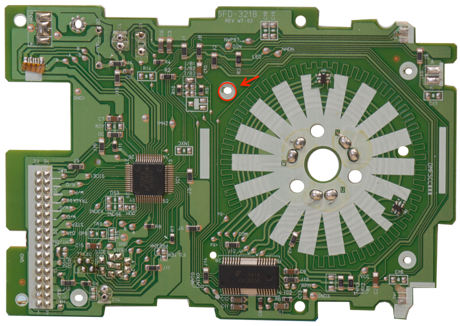
8

Insert the RESISTOR in the holes shown in the picture



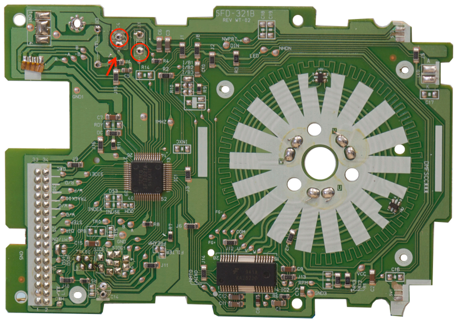
9

Insert the LED in the hole shown in the picture



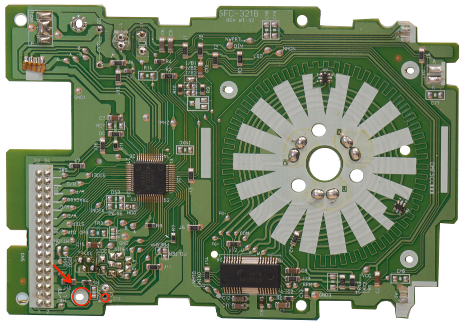
10

Insert the RESISTOR in the holes shown in the picture



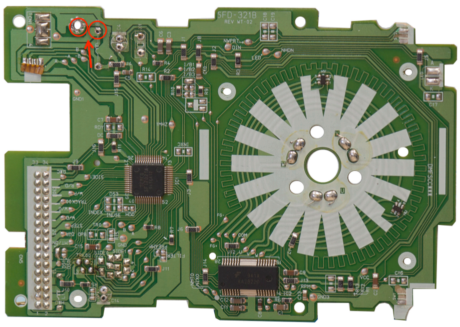
11

Insert the CAPACITOR in the holes shown in the picture



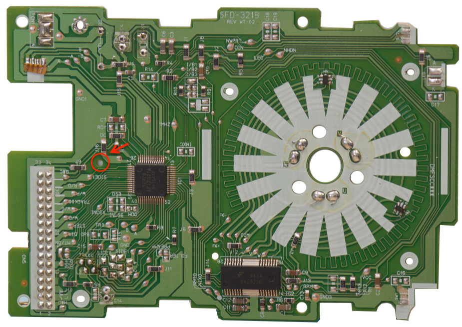
12

Insert the CAPACITOR in the holes shown in the picture



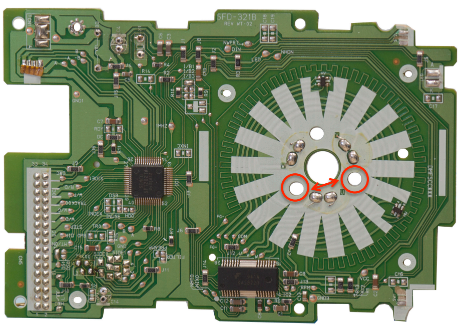
13

Insert the LED in the hole shown in the picture



14

Insert the RESISTOR in the holes shown in the picture

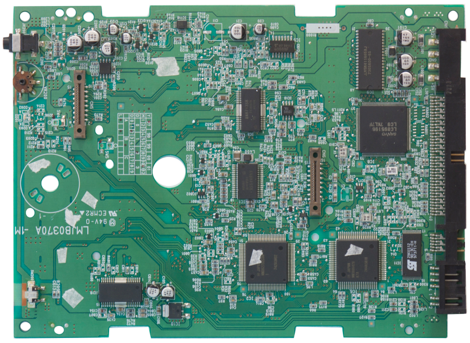


15

Finished

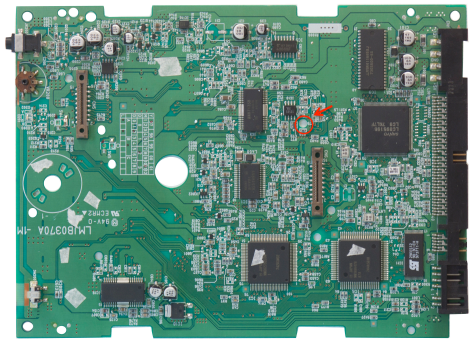
16

Take the following circuit board



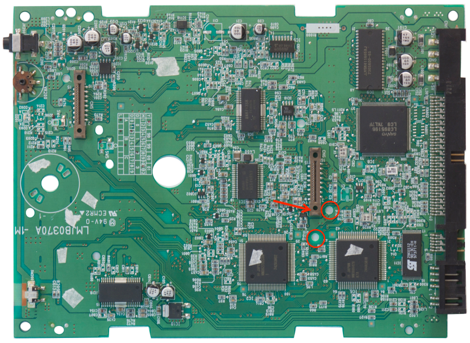
17

Insert the LED in the hole shown in the picture



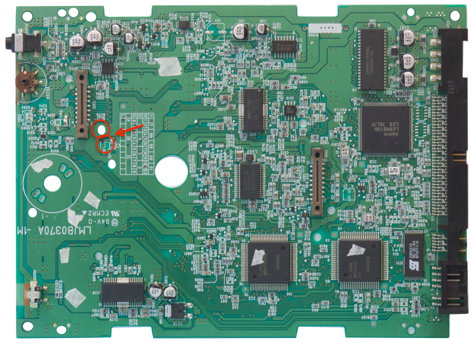
18

Insert the CAPACITOR in the holes shown in the picture



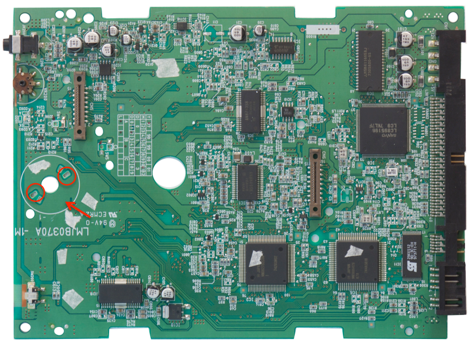
19

Insert the CAPACITOR in the holes shown in the picture



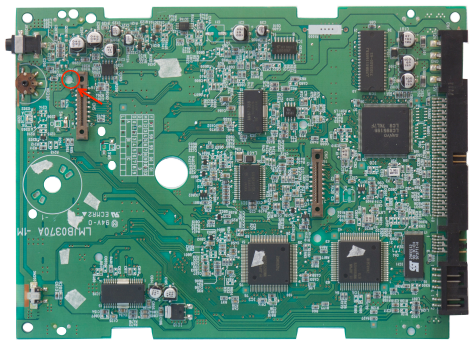
20

Insert the RESISTOR in the holes shown in the picture



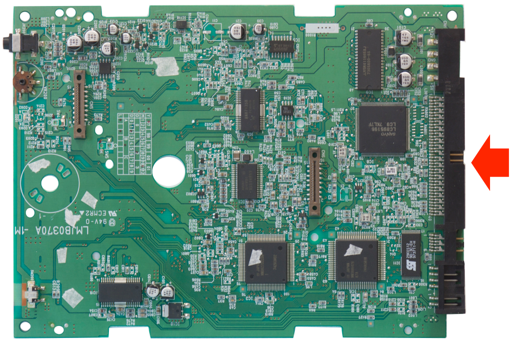
21

Insert the LED in the hole shown in the picture



22

Insert the CONNECTOR

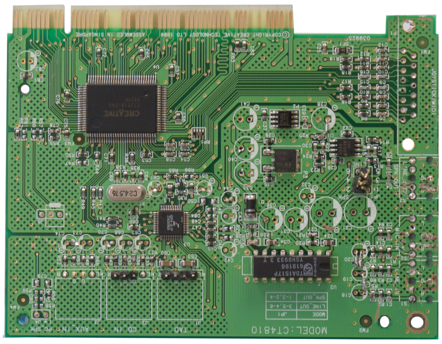


23

Finished

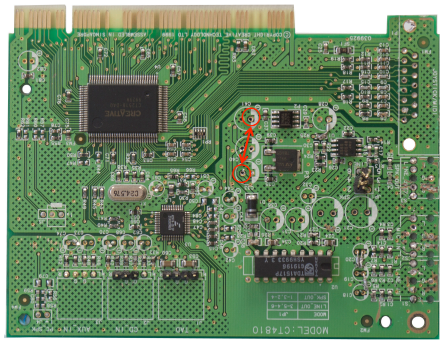
24

Take the following circuit board



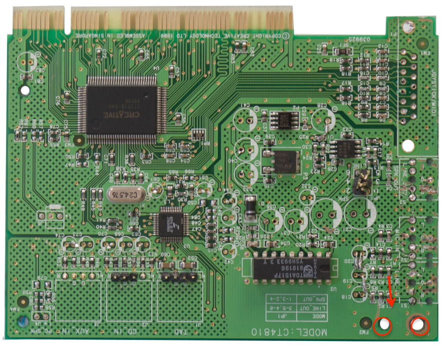
25

Insert the RESISTOR in the holes shown in the picture



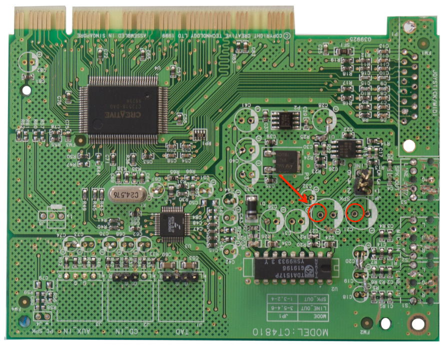
26

Insert the RESISTOR in the holes shown in the picture



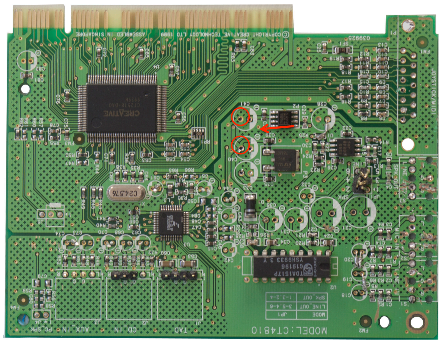
27

Insert the CAPACITOR in the holes shown in the picture



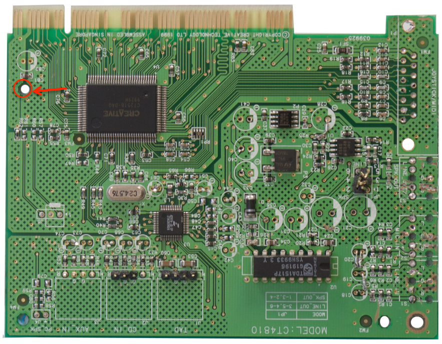
28

Insert the CAPACITOR in the holes shown in the picture



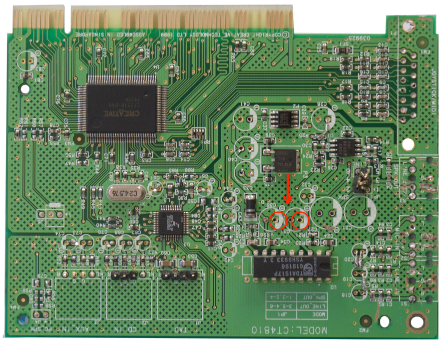
29

Insert the LED in the hole shown in the picture



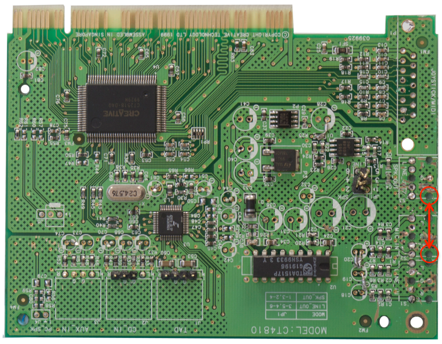
30

Insert the CAPACITOR in the holes shown in the picture



31

Insert the RESISTOR in the holes shown in the picture

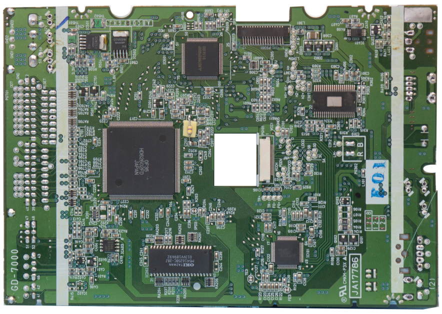


32

Finished

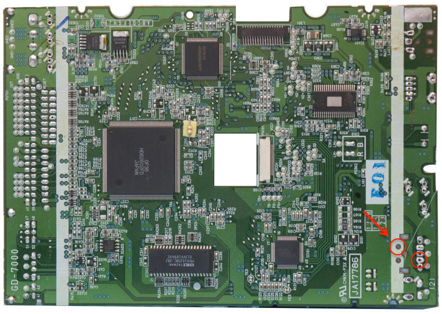
33

Take the following circuit board



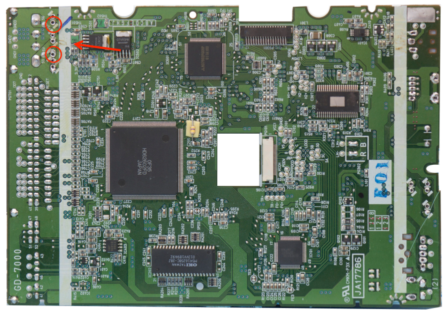
34

Insert the RESISTOR in the holes shown in the picture



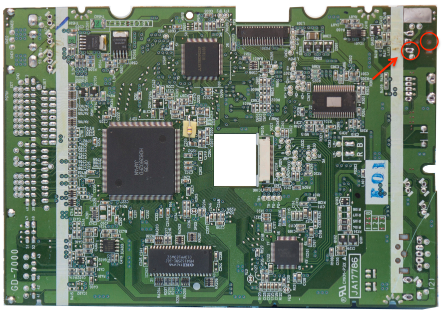
35

Insert the RESISTOR in the holes shown in the picture



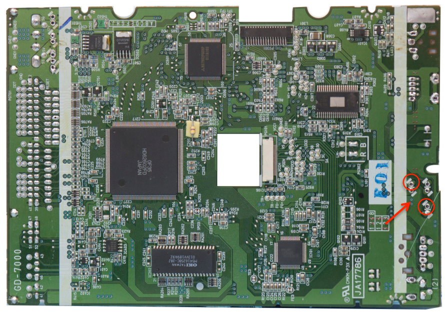
36

Insert the RESISTOR in the holes shown in the picture



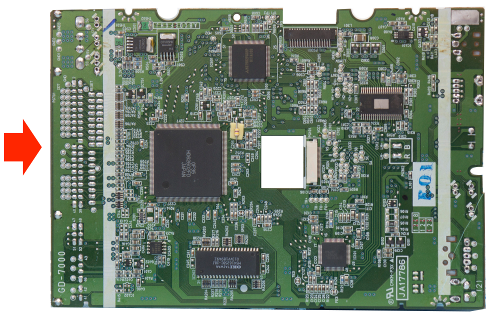
37

Insert the RESISTOR in the holes shown in the picture



38

Insert the CONNECTOR

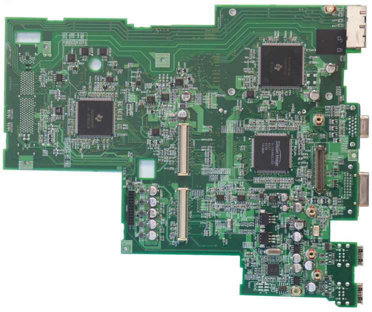


39

Finished

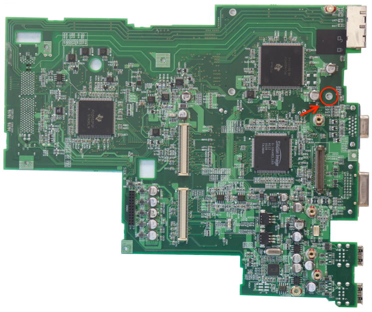
40

Take the following circuit board



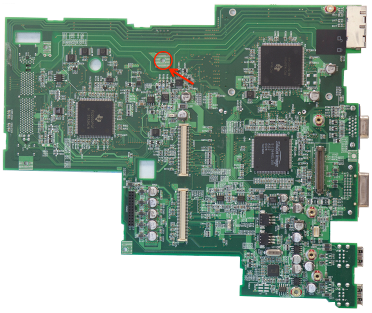
41

Insert the CAPACITOR in the hole shown in the picture



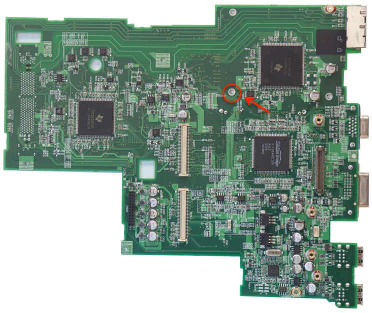
42

Insert the LED in the hole shown in the picture



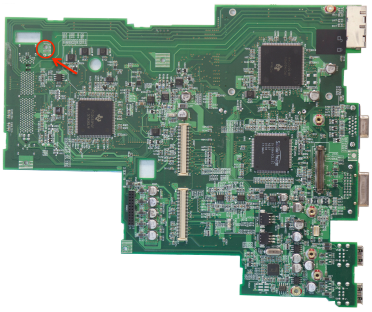
43

Insert the CAPACITOR in the hole shown in the picture



44

Insert the LED in the hole shown in the picture



45

Finished

46

B.2 User Circuit Board Flipped Over Operating Manual

In this Section we show the randomly flipped over operating manual that we used during the user test for the AR guide for manual assembly tasks. The test is also described in Chapter 10

Circuit Board Assembly Test

User Manual

Flipping Mode

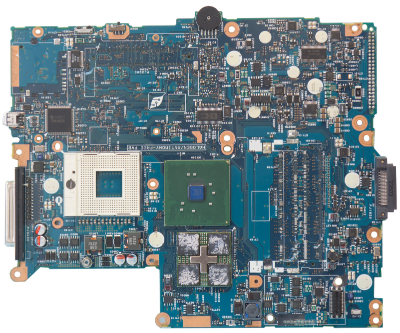
Guido Maria Re
October 2012

1

Test….

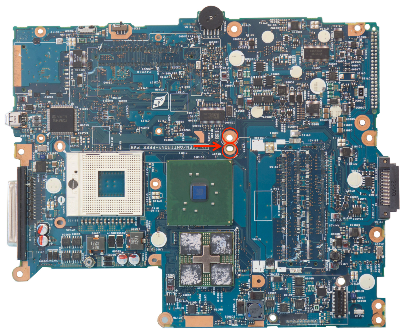
0

Take the following circuit board



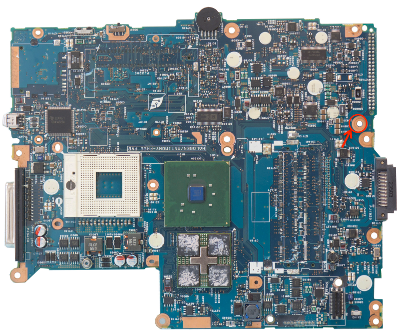
1

Insert the RESISTOR in the holes shown in the picture



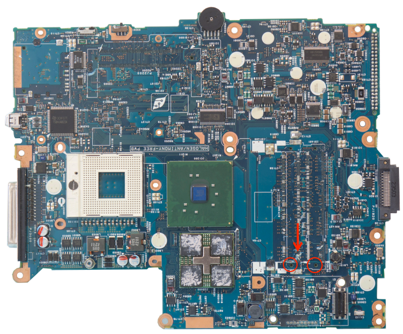
2

Insert the LED in the hole shown in the picture



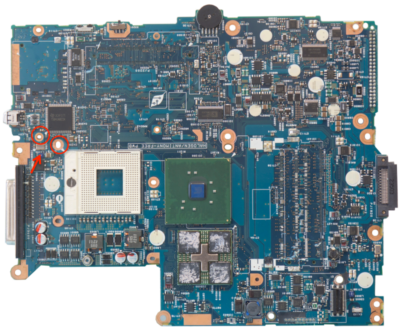
3

Insert the CAPACITOR in the holes shown in the picture



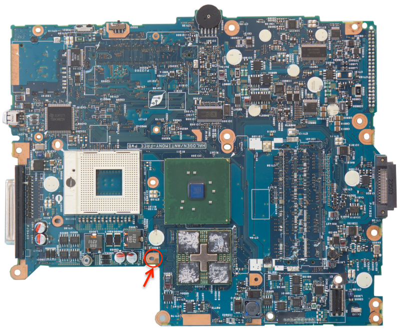
4

Insert the RESISTOR in the holes shown in the picture



5

Insert the LED in the hole shown in the picture

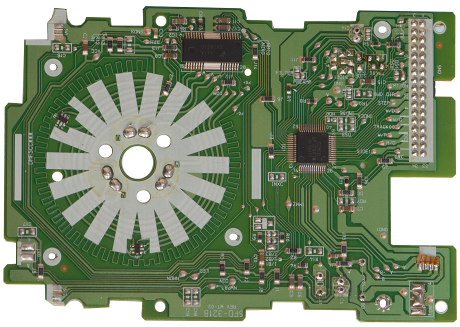


6

Finished

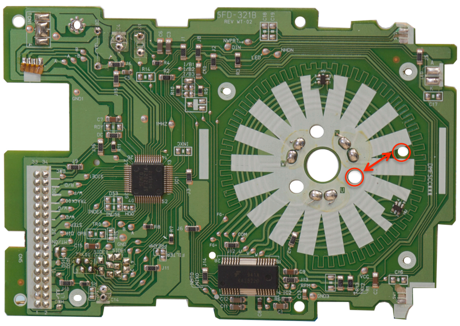
7

Take the following circuit board



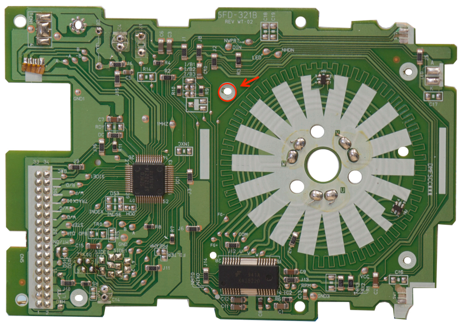
8

Insert the RESISTOR in the holes shown in the picture



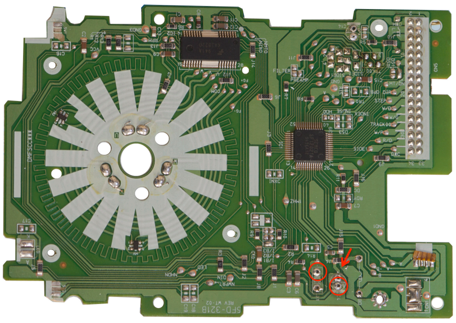
9

Insert the LED in the hole shown in the picture



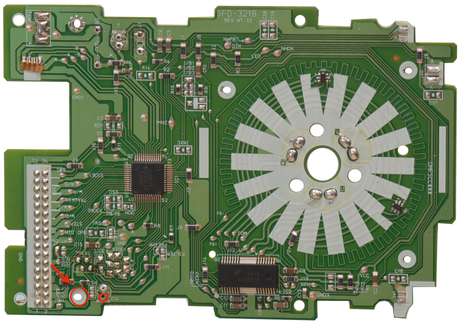
10

Insert the RESISTOR in the holes shown in the picture



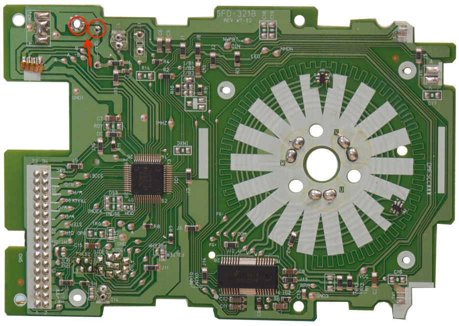
11

Insert the CAPACITOR in the holes shown in the picture



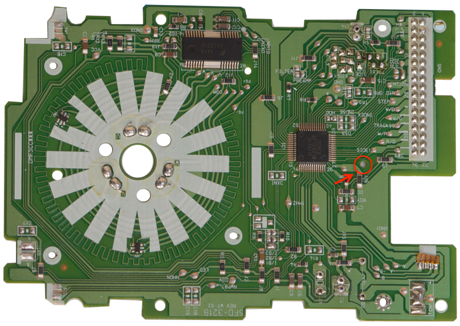
12

Insert the CAPACITOR in the holes shown in the picture



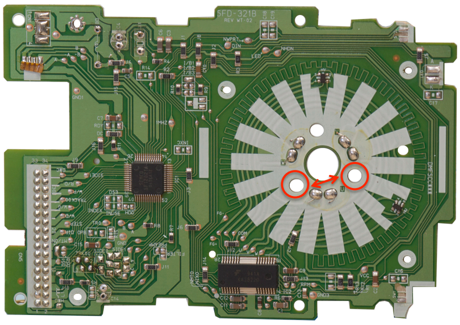
13

Insert the LED in the hole shown in the picture



14

Insert the RESISTOR in the holes shown in the picture

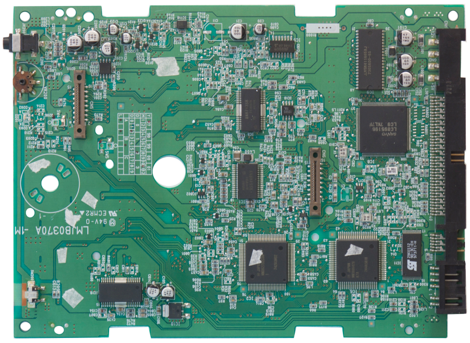


15

Finished

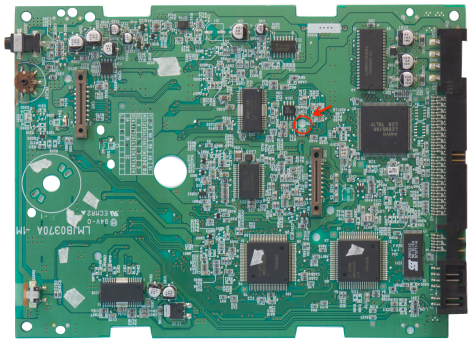
16

Take the following circuit board



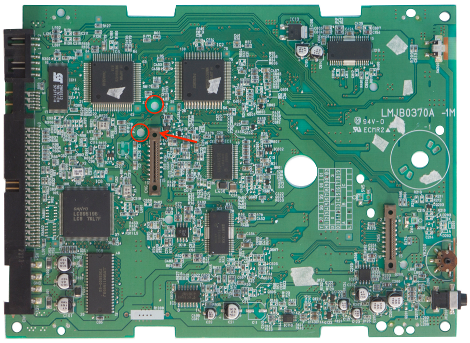
17

Insert the LED in the hole shown in the picture



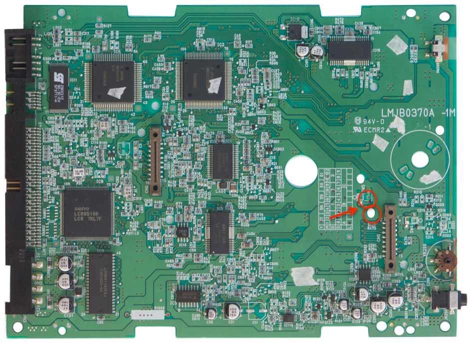
18

Insert the CAPACITOR in the holes shown in the picture



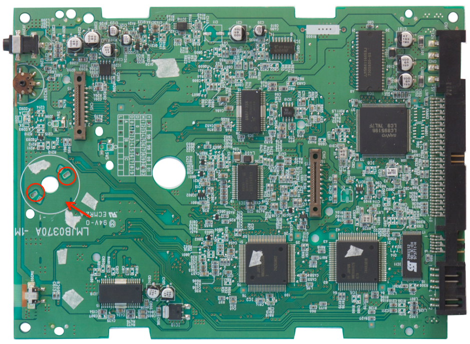
19

Insert the CAPACITOR in the holes shown in the picture



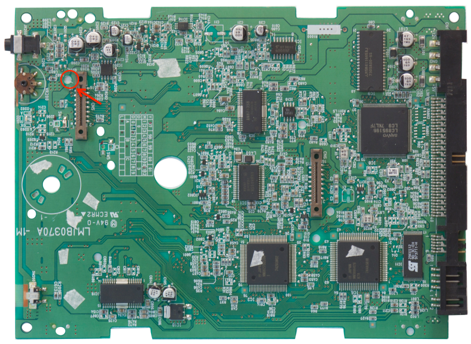
20

Insert the RESISTOR in the holes shown in the picture



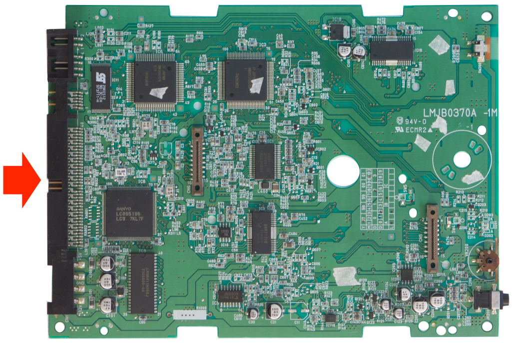
21

Insert the LED in the hole shown in the picture



22

Insert the CONNECTOR

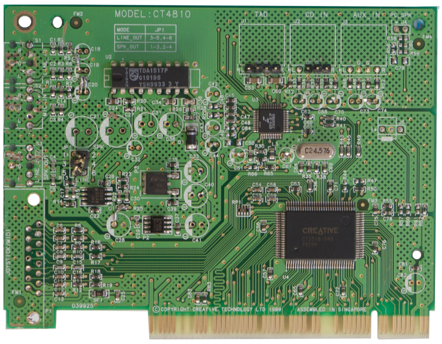


23

Finished

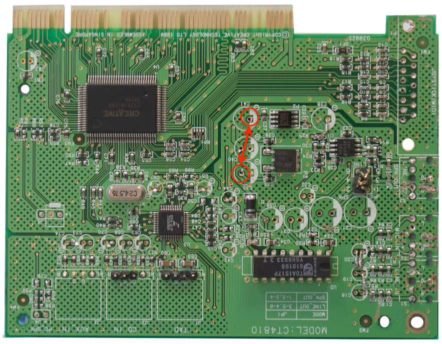
24

Take the following circuit board



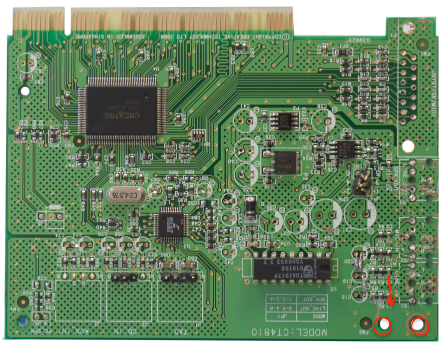
25

Insert the RESISTOR in the holes shown in the picture



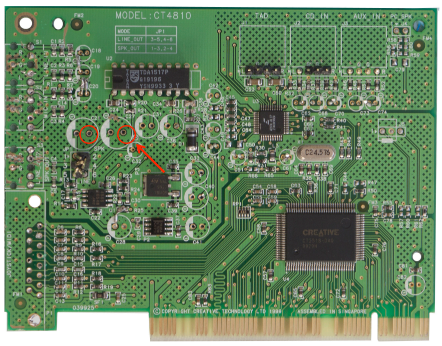
26

Insert the RESISTOR in the holes shown in the picture



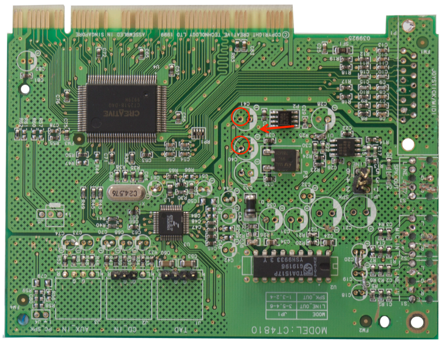
27

Insert the CAPACITOR in the holes shown in the picture



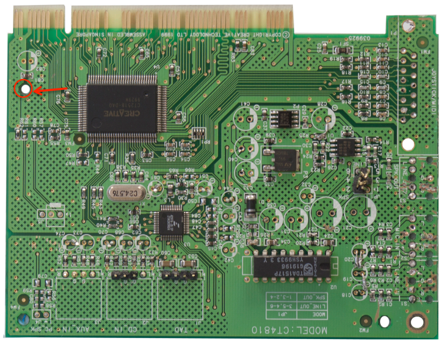
28

Insert the CAPACITOR in the holes shown in the picture



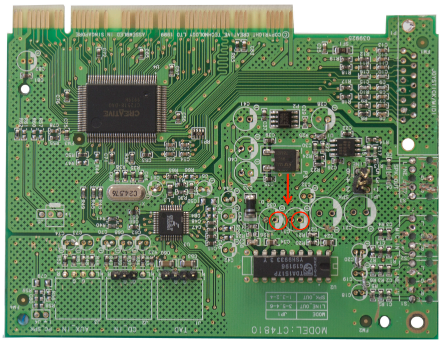
29

Insert the LED in the hole shown in the picture



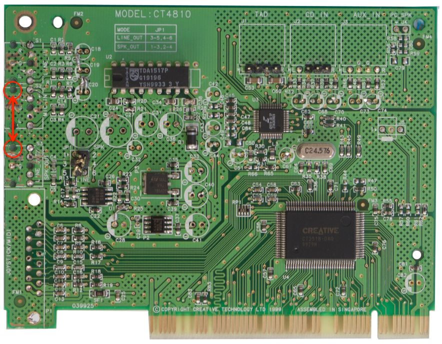
30

Insert the CAPACITOR in the holes shown in the picture



31

Insert the RESISTOR in the holes shown in the picture

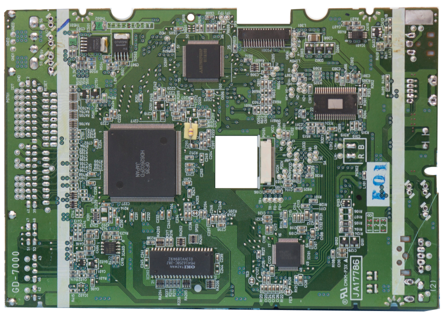


32

Finished

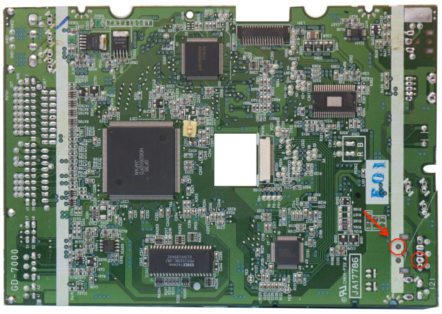
33

Take the following circuit board



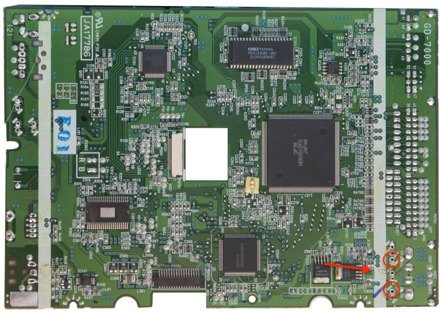
34

Insert the RESISTOR in the holes shown in the picture



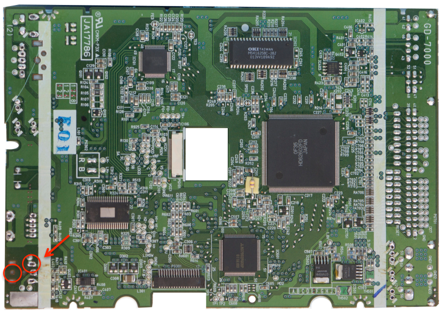
35

Insert the RESISTOR in the holes shown in the picture



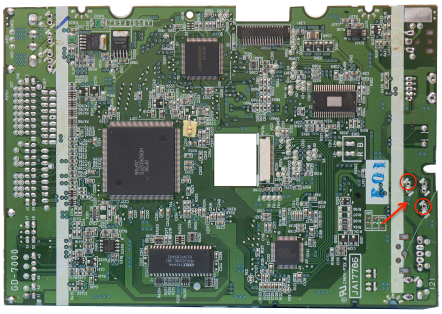
36

Insert the RESISTOR in the holes shown in the picture



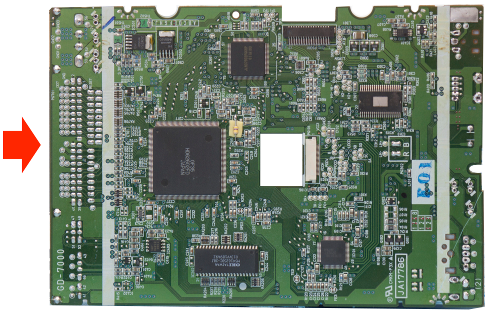
37

Insert the RESISTOR in the holes shown in the picture



38

Insert the CONNECTOR

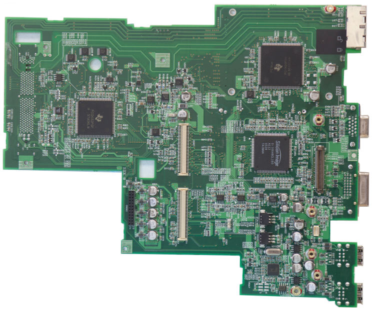


39

Finished

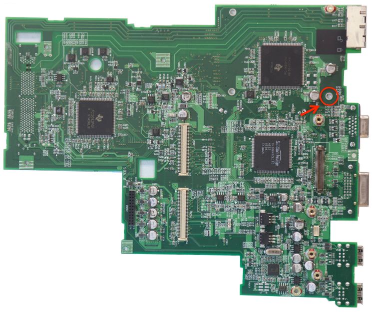
40

Take the following circuit board



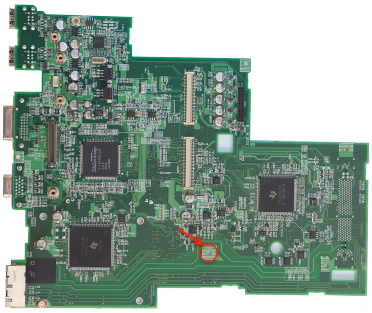
41

Insert the CAPACITOR in the hole shown in the picture



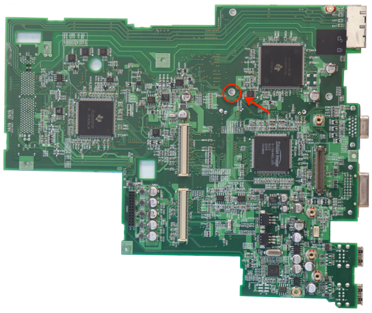
42

Insert the LED in the hole shown in the picture



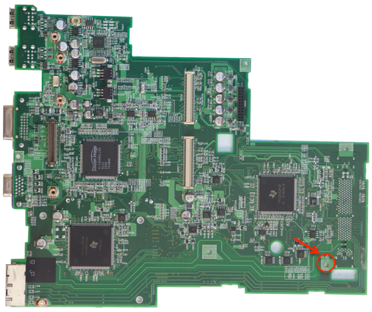
43

Insert the CAPACITOR in the hole shown in the picture



44

Insert the LED in the hole shown in the picture



45

Finished

46

List of Figures

List of Tables

Bibliography

[1] 3DVIA Virtools. www.virtools.com/. Last access: november 2012.

[2] Alvar. http://virtual.vtt.fi/virtual/proj2/multimedia/alvar/index.html. Last access: november 2012.

[3] Analog Devices. http://www.analog.com/en/index.html. Last access: November 2012.

[4] ARMES. http://www.armes-tech.com/. Last access: november 2012.

[5] ART Advanced Realtime Tracking. http://www.ar-tracking.com/. Last access: November 2012.

[6] ARTag. http://www.artag.net/. Last access: november 2012.

[7] ARToolkit NFT. http://www.artoolworks.com/support/library/ARToolKit_NFT. Last access: november 2012.

[8] ARToolkit Plus. http://handheldar.icg.tugraz.at/artoolkitplus.php. Last access: november 2012.

[9] ARToolKit website. http://www.hitl.washington.edu/artoolkit/. Last access: november 2012.

[10] ARVIKA Consortium. http://www.arvika.de/www/e/home/home.htm. Last access: December 2012.

[11] BazAR: A vision based fast detection library. http://cvlab.epfl.ch/software/bazar/. Last access: November 2012.

[12] Build AR. buildar.com/. Last access: december 2012.

[13] Camera Calibration Toolbox for Matlab. http://www.vision.caltech.edu/bouguetj/calib_doc/. Last access: November 2012.

[14] D'Fusion by Total Immersion. http://www.t-immersion.com/. Last access: november 2012.

[15] Extend3D. http://www.extend3d.de/en/. Last access: December 2012.

[16] FLARManager. http://words.transmote.com/wp/flarmanager/. Last access: november 2012.

[17] FLARToolkit. http://www.libspark.org/wiki/saqoosha/FLARToolKit/en. Last access: november 2012.

[18] How to conduct an euristic evalutation. http://www.useit.com/papers/heuristic/heuristic_evaluation.html. Last access: November 2012.

[19] iRobot Roomba. http://www.irobot.com/us/robots/home/roomba/Features.aspx. Last access: November 2012.

[20] Layar. www.layar.com/. Last access: november 2012.

[21] List of Augmented Reality SDKs. http://www.icg.tu-graz.ac.at/Members/gerhard/augmented-reality-sdks. Last access: december 2012.

[22] Metaio SDK. `http://www.metaio.com/products/sdk/`. Last access: november 2012.

[23] NyARToolkit. `http://nyatla.jp/nyartoolkit/wp/`. Last access: november 2012.

[24] OpenGL API. `http://en.wikipedia.org/wiki/OpenGL`. Last access: november 2012.

[25] OpenSceneGraph. `www.openscenegraph.org/`. Last access: november 2012.

[26] openTL. `http://www.opentl.org/`. Last access: november 2012.

[27] osgART. `www.osgart.org/`. Last access: november 2012.

[28] PixArt Imaging. `http://www.pixart.com.tw/index.asp`. Last access: November 2012.

[29] SLARToolkit. `slartoolkit.codeplex.com/`. Last access: november 2012.

[30] Studierstube. `www.studierstube.org/`. Last access: november 2012.

[31] Vicon Systemes. `vicon.com`. Last access: November 2012.

[32] Vuforia by Qualcomm. `http://www.qualcomm.com/solutions/augmented-reality`. Last access: november 2012.

[33] Vuzix AR Glasses. `http://www.vuzix.com/augmented-reality/index.html`. Last access: December 2012.

[34] WiiYourself! Library. `http://wiiyourself.gl.tter.org/`. Last access: November 2012.

[35] XBee wireless RF Module. `digi.com/xbee`. Last access: November 2012.

[36] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier, and B. MacIntyre. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, 21(6):34–47, 2001.

[37] K. M. Baird and W. Barfield. Evaluating the effectiveness of augmented reality displays for a manual assembly task. *Virtual Reality*, 4(4):250–259, Dec. 1999.

[38] S. Balcisoy, M. Kallmann, P. Fua, and D. Thalmann. A framework for rapid evaluation of prototypes with Augmented Reality. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 61–66, 2000.

[39] G. Baratoff, A. Neubeck, and H. Regenbrecht. Interactive multi-marker calibration for augmented reality applications. In *Proceedings. International Symposium on Mixed and Augmented Reality*, pages 107–116. IEEE Comput. Soc, 2002.

[40] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, 110(3):346–359, June 2008.

[41] H. Bay, T. Tuytelaars, and L. Van Gool. Surf: Speeded up robust features. In *European Conference on Computer Vision (ECCV)*, volume 3951, pages 404–417. Springer, 2006.

[42] M. Benali-Khoudja, M. Hafez, J.-M. Alexandre, and A. Kheddar. Tactile interfaces : a state-of-the-art survey. In *International. Symposium on Robotics*, volume 31, pages 1–9, 2004.

[43] H. Benko, R. Jota, A. D. Wilson, and V. Inesc-id. MirageTable : Freehand Interaction on a Projected Augmented Reality Tabletop. In *Conference on Human Factors in Computing Systems (CHI)*, pages 199–208, 2012.

[44] M. Billinghurst, M. Hakkarainen, and C. Woodward. Augmented assembly using a mobile phone. In *Proceedings of the 7th International Conference on Mobile and Ubiquitous Multimedia (MUM)*, pages 84–87. Ieee, Sept. 2008.

[45] M. Billinghurst, H. Kato, and I. Poupyrev. The MagicBook: a transitional AR interface. *Computers & Graphics*, 25(5):745–753, Oct. 2001.

[46] O. Bimber and R. Raskar. Modern approaches to augmented reality. In *ACM SIGGRAPH*, page 1, New York, New York, USA, 2006. ACM Press.

[47] V. Buchmann, S. Violich, M. Billinghurst, and A. Cockburn. FingARtips: gesture based direct manipulation in Augmented Reality. *Virtual Reality*, 1(212):212–221, 2004.

[48] G. Caruso. Mixed Reality System for Ergonomic Assessment of Driver s Seat. *The International Journal of Virtual Reality*, 10(2):69–79, 2011.

[49] G. Caruso and U. Cugini. Augmented Reality Video See-through HMD Oriented to Product Design Assessment. *Virtual and Mixed Reality*, pages 532–541, 2009.

[50] G. Caruso and G. M. Re. Interactive Augmented Reality System for product design review. In *IS&T/SPIE Electronic Imaging*, volume 39, pages 75250H–75250H, 2010.

[51] R. Castle, G. Klein, and D. W. Murray. Video-rate Localization in Multiple Maps for Wearable Augmented Reality. In *12th IEEE International Symposium on Wearable Computers (ISWC)*, pages 15–22, 2008.

[52] T. Caudell and D. Mizell. Augmented reality: an application of heads-up display technology to manual manufacturing processes. *Proceedings of the Twenty-Fifth Hawaii International Conference on System Sciences*, 2:659–669, 1992.

[53] A. D. Cheok, N. Weng, C. Edmund, and A. W. Eng. Inexpensive non-sensor based augmented reality modeling of curves and surfaces in physical space. In *Proceedings of the 1st International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 272–273, 2002.

[54] K. Chintamani, A. Cao, D. R. Ellis, and A. K. Pandya. Improved telemanipulator navigation during display-control misalignments using augmented reality cues. *IEEE Transaction on Systems, Man and Cybernetics, Part A: Systems and Humans*, 40(1):29–39, 2010.

[55] M. Clothier and M. Bailey. Augmented Reality Visualization Tool for Kings Stormwater Bridge. In *Visualization, Imaging, and Image Processing*, 2004.

[56] N. Cornelis and L. Van Gool. Fast Scale Invariant Feature Detection and Matching on Programmable Graphics Hardware. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 1–8, 2008.

[57] F. Doil, W. Schreiber, T. Alt, and C. Patron. Augmented reality for manufacturing planning. In *Proceedings of the workshop on Virtual environments*, pages 71–76, Zurich, Switzerland, 2003. ACM.

[58] K. Dorfmuller-Ulhaas and D. Schmalstieg. Finger tracking for interaction in augmented environments. In *4th International Symposium on Augmented Reality (ISAR)*, pages 55–64, 2001.

[59] T. Drummond and R. Cipolla. Real-time tracking of complex structures with on-line camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):932–946, 2002.

[60] F. Echtler, F. Sturm, K. Kindermann, and G. Klinker. The Intelligent Welding Gun : Augmented Reality for Experimental Vehicle Construction. In S. Ong and A. Nee, editors, *Virtual and Augmented Reality Applications in Manufacturing*, chapter 17, pages 1–27. Springer Verlag, 2003.

[61] S. Feiner, B. Macintyre, and D. Seligmann. Knowledge-based augmented reality. *Communications of the ACM*, 36(7):53–62, 1993.

[62] J. L. Fergason. Optical System for a head moounted display using a retro-reflector and method of displaying an image, 1997.

[63] M. Fiala. ARTag, a fiducial marker system using digital techniques. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 590–596, 2002.

[64] M. Fiala. A Robot Control and Augmented Reality Interface for Multiple Robots. In *Canadian Conference on Computer and Robot Vision*, pages 31–36. Ieee, May 2009.

[65] M. Fiorentino, R. de Amicis, G. Monno, and A. Stork. Spacedesign: A mixed reality workspace for aesthetic industrial design. In *International Symposium on Mixed and Augmented Reality (ISMAR)*, number figure 1, pages 86–318, 2002.

[66] M. A. Fischler and R. C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[67] P. Fite-Georgel. Is there a reality in Industrial Augmented Reality? In *10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 201–210. Ieee, Oct. 2011.

[68] P. Fite-Georgel, P. Schroeder, and N. Navab. Navigation Tools for Viewing Augmented CAD Models. *Ieee Computer Graphics And Applications*, 29(6):65–73, 2009.

[69] W. Friedrich, D. Jahn, and L. Schmidt. ARVIKA - Augmented Reality for Development, Production and Service. In *Proceedings of the ieee/acm International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 1–16, 2002.

[70] J. Fründ, J. Gausemeier, C. Matysczok, and R. Radkowski. Using Augmented Reality Technology to Support the Automobile Development. In *Computer Supported Cooperative Work in Design*, pages 289–298, 2005.

[71] R. Galantay, J. Torpus, and M. Engeli. living-room Interactive, Space-Oriented Augmented Reality. In *Proceedings of the 12th annual ACM international conference on Multimedia*, pages 64–71. ACM, 2004.

[72] M. Gervautz and D. Schmalstieg. Anywhere Interfaces Using Handheld Augmented Reality. *IEEE Computer*, 45(7):26–31, 2012.

[73] A.-s. Gunnarsson, M. Rauhala, A. Henrysson, and A. Ynnerman. Visualization of sensor data using mobile phone augmented reality. In *IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 233–234. Ieee, Oct. 2006.

[74] S. Gupta and C. Jaynes. The Universal Media Book : Tracking and Augmenting Moving Surfaces with Projected Information. In *Proceedings of the 5th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 177–180, 2006.

[75] C. Harrison, H. Benko, and A. D. Wilson. OmniTouch: Wearable Multitouch Interaction Everywhere. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology (UIST)*, pages 441–450, 2011.

[76] S. G. Hart and L. E. Staveland. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Human mental workload*, volume 1, pages 139–183, 1988.

[77] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[78] S. Hashimoto, A. Ishida, M. Inami, and T. Igarash. TouchMe: An Augmented Reality Based Remote Robot Manipulation. In *The 21st International Conference on Artificial Reality and Telexistence (ICAT)*, pages 1–6, 2011.

[79] S. Henderson and S. Feiner. Exploring the benefits of augmented reality documentation for maintenance and repair. *IEEE transactions on visualization and computer graphics*, 17(10):1355–68, Oct. 2011.

[80] A. Henrysson, M. Billinghurst, and M. Ollila. Face to Face Collaborative AR on Mobile Phones. In *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 80–89, 2005.

[81] O. Hilliges, D. Kim, S. Izadi, M. Weiss, and A. D. Wilson. HoloDesk: Direct 3D Interactions with a Situated See-Through Display. In *CHI Conference on Human Factors in Computing Systems*, pages 159–168, 2012.

[82] W. a. Hoff. Computer vision-based registration techniques for augmented reality. In *Proceedings of SPIE*, volume 2904, pages 538–548. Spie, 1996.

[83] H. Hua, a. Girardot, C. Gao, and J. P. Rolland. Engineering of head-mounted projective displays. *Applied optics*, 39(22):3814–24, Aug. 2000.

[84] H. Ishii, Z. Bian, H. Fujino, T. Sekiyama, T. Nakai, A. Okamoto, and I. HShimoda. Augmented Reality Applications for Nuclear Power Plant Maintenance Work. In *Internationa Symposium on ymbiotic Nuclear Power Systems for 21st Century*, pages 262–268, 2007.

[85] H. Ishii, H. Shimoda, T. Nakai, M. Izumi, Z. Bian, and Y. Morishita. Proposal and Evaluation of a Supporting Method for NPP Decommissioning Work by Augmented Reality. In *12th World Multi-Conference on Systemics, Cybernetics*, volume 6, pages 157–162, 2008.

[86] H. Ishii and B. Ullmer. Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms. In *Conference Proceedings of Human Factors in Computing Systems (CHI)*, number March, pages 234–241, 1997.

[87] S. Junuzovic, K. Inkpen, T. Blank, A. Gupta, and O. M. Way. IllumiShare: Sharing Any Surface. In *CHI Conference on Human Factors in Computing Systems*, pages 1919–1928, 2012.

[88] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *IEEE and ACM International Workshop on Augmented Reality (IWAR)*, pages 85–94, 1999.

[89] H. Kato, M. Billinghurst, I. Poupyrev, K. Imamoto, and K. Tachibana. Virtual object manipulation on a table-top AR environment. In *IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*, pages 111–119. Ieee, 2000.

[90] H. Kaufmann and D. Schmalstieg. Mathematics and geometry education with collaborative augmented reality. *Computers & Graphics*, 27(3):339–345, June 2003.

[91] K. Kiyokawa, Y. Kurata, and H. Ohno. An optical see-through display for mutual occlusion of real and virtual environments. In *Proceedings IEEE and ACM International Symposium on Augmented Reality (ISAR 2000)*, pages 60–67. Ieee, 2000.

[92] G. Klein and T. Drummond. Robust visual tracking for non-instrumental augmented reality. In *Proceedings of The Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 113–122. IEEE Comput. Soc, 2006.

[93] G. Klein and D. Murray. Parallel Tracking and Mapping for Small {AR} Workspaces. In *Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 225–234, Nara, Japan, 2007.

[94] G. Klein and D. Murray. Parallel Tracking and Mapping on a camera phone. In *8th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 83–86. Ieee, Oct. 2009.

[95] G. Klinker, A. Dutoit, M. Bauer, J. Bayer, V. Novak, and D. Matzke. Fata Morgana - A presentation system for product design. In *Proceedings of the International Symposium on Mixed and Augmented Reality*, pages 76–85. IEEE Comput. Soc, 2002.

[96] K. Kobayashi, S. Ishigame, and H. Kato. Simulator of Manual Metal Arc Welding with Haptic Display. In *International Conference on Artificial Reality and Telexistence (ICAT)*, pages 175–178, 2001.

[97] D. Koller, G. Klinker, E. Rose, D. Breen, R. Whitaker, and M. Tuceryan. Real-time vision-based camera tracking for augmented reality applications. In *Proceedings of the ACM symposium on Virtual reality software and technology*, pages 87–94. ACM, 1997.

[98] T. Lee and T. Hollerer. Handy AR: Markerless inspection of augmented reality objects using fingertip tracking. In *IEEE International Symposium on Wearable Computers (ISWC)*, pages 83–90. IEEE, 2007.

[99] T. Lee and T. Hollerer. Hybrid feature tracking and user interaction for markerless augmented reality. In *Virtual Reality Conference, 2008. VR'08. IEEE*, pages 145–152. Ieee, 2008.

[100] V. Lepetit and P. Fua. Monocular Model-Based 3D Tracking of Rigid Objects. *Foundations and Trends in Computer Graphics and Vision*, 1(1):1–89, 2005.

[101] V. Lepetit, P. Lagger, and P. Fua. Randomized trees for real-time keypoint recognition. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 775–781, 2005.

[102] V. Lepetit, J. Pilet, and P. Fua. Point matching as a classification problem for fast and robust object pose estimation. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 2, pages 244–250. Ieee, 2004.

[103] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, Nov. 2004.

[104] A. M. Malkawi and R. S. Srinivasan. A new paradigm for Human-Building Interaction: the use of CFD and Augmented Reality. *Automation in Construction*, 14(1):71–84, Jan. 2005.

[105] M. R. Marner and B. H. Thomas. Augmented Foam Sculpting for Capturing 3D Models. In *IEEE Symposium on 3D User Interfaces (3DUI)*, pages 63–70, 2010.

[106] P. Milgram and H. J. Colquhoun. A Taxonomy of Real and Virtual World Display Integration. In *Environments*, number 1999, chapter 1, pages 1–26. Springer, 1999.

[107] P. Milgram and F. Kishino. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information Systems*, E77-D(12), 1994.

[108] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented Reality: A class of displays on the reality-virtuality continuum. In *Proceedings of Telemanipulator and Telepresence Technologies*, volume 2351, pages 282–292, 1994.

[109] P. Mistry, P. Maes, and L. Chang. WUW - Wear Ur World - A Wearable Gestural Interface. In *CHI '09*, pages 4111–4116, 2009.

[110] L. Naimark and E. Foxlin. Circular Data Matrix Fiducial System and Robust Image Processing for a Wearable Vision-Inertial Self-Tracker. In *Proceedings International Symposium on Mixed and Augmented Reality*, number Ismar, pages 27–36, 2002.

[111] L. Naimark and E. Foxlin. Encoded LED System for Optical Trackers. In *Fourth IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 150–153, 2005.

[112] H. Nakamura and H. Miyashita. Development and evaluation of interactive system for synchronizing electric taste and visual content. In *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*, pages 517–520, New York, New York, USA, 2012. ACM Press.

[113] N. Navab. Developing Killer Apps for Industrial Augmented Reality. *IEEE Computer Graphics and Applications*, 24(3):16–20, 2004.

[114] N. Navab, B. Bascle, M. Appel, and E. Cubillo. Scene augmentation via the fusion of industrial drawings and uncalibrated images with a view to marker-less calibration. In *Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR)*, pages 125–133. IEEE Comput. Soc, 1999.

[115] A. Nee, S. Ong, G. Chryssolouris, and D. Mourtzis. Augmented reality applications in design and manufacturing. *CIRP Annals - Manufacturing Technology*, 61(2):657–679, Jan. 2012.

[116] L. X. Ng, S. W. Oon, S. K. Ong, and a. Y. C. Nee. GARDE: a gesture-based augmented reality design evaluation system. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 5(2):85–94, Mar. 2011.

[117] S. Noelle. Stereo augmentation of simulation results on a projection wall by combining two basic ARVIKA systems. *Proceedings. International Symposium on Mixed and Augmented Reality*, pages 271–322, 2002.

[118] K. Ohmori and K. Sakamoto. Automatic Mobile Robot Control and Indication Method Using Augmented Reality Technology. In H. Yang, R. Malaka, J. Hoshino, and J. Han, editors, *Entertainment Computing - ICEC 2010*, volume 6243 of *Lecture Notes in Computer Science*, pages 464–467. Springer Berlin / Heidelberg, 2010.

[119] T. Ohshima, T. Kuroki, H. Yamamoto, and H. Tamura. A mixed reality system with visual and tangible interaction capability: application to evaluating automobile interior design. *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.*, pages 284–285, 2006.

[120] K. Oka, Y. Sato, and H. Koike. Real-Time Fingertip Tracking and Gesture Recognition. *IEEE Computer Graphics and Applications*, 22(6):64–71, 2002.

[121] S. Ong, J. Chong, and a.Y.C. Nee. A novel AR-based robot programming and path planning methodology. *Robotics and Computer-Integrated Manufacturing*, 26(3):240–249, June 2010.

[122] S. Ong and Z. Wang. Augmented assembly technologies based on 3D bare-hand interaction. *CIRP Annals - Manufacturing Technology*, 60(1):1–4, Jan. 2011.

[123] C. B. Owen, F. Xiao, and P. Middlin. What is the best fiducial? In *The First IEEE International Workshop Agumented Reality Toolkit*, 2002.

[124] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua. Fast keypoint recognition using random ferns. *IEEE transactions on pattern analysis and machine intelligence*, 32(3):448–61, Mar. 2010.

[125] M. Ozuysal, P. Fua, and V. Lepetit. Fast keypoint recognition in ten lines of code. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 1–8, 2007.

[126] H. Park, H. Moon, and J. Lee. Tangible augmented prototyping of digital handheld products. *Computers in Industry*, 60(2):114–125, Feb. 2009.

[127] J. Park. Augmented Reality Based Re-formable Mock-Up for Design Evaluation. *2008 International Symposium on Ubiquitous Virtual Reality*, pages 17–20, July 2008.

[128] K. Pentenrieder, C. Bade, F. Doil, and P. Meier. Augmented Reality-based factory planning - an application tailored to industrial needs. *2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 1–9, Nov. 2007.

[129] R. Raskar and P. Beardsley. iLamps : Geometrically Aware and Self-Configuring Projectors. In *ACM SIGGRAPH*, pages 1–10, 2003.

[130] R. Raskar, G. Welch, K.-L. Low, and D. Bandyopadhyay. Shader lamps: Animating real objects with image-based illumination. In *roceedings of the 12th Eurographics Workshop on Rendering Techniques*, pages 89–102, 2001.

[131] H. Regenbrecht, M. Wagner, and G. Baratoff. MagicMeeting: A Collaborative Tangible Augmented Reality System. *Virtual Reality*, 6(3):151–166, Oct. 2002.

[132] D. Reiners, D. Stricker, G. Klinker, and M. Stefan. Augmented Reality for Construction Tasks : Doorlock Assembly. In *IEEE and ACM IWAR*, pages 31–46, 1998.

[133] G. Reinhart, U. Munzert, and W. Vogl. A programming system for robot-based remote-laser-welding with conventional optics. *CIRP Annals - Manufacturing Technology*, 57(1):37–40, Jan. 2008.

[134] J. Rekimoto. Matrix : A Realtime Object Identification and Registration Method for Augmented Reality. In *Third Asian Pacific Computer and Human Interaction (APCHI)*, pages 63–69, 1998.

[135] J. P. Rolland, Y. Baillot, and A. A. Goon. A survey of tracking technology for virtual environments. In Barfield and Caudell, editor, *Fundamentals of Wearable Computers and Augmented Reality*, chapter 3, pages 67–112. Mahwah, NJ, 2001.

[136] J. P. Rolland, R. L. Holloway, and H. Fuchs. A comparison of optical and video see-through head-mounted displays. *Telemanipulator and Telepresence Technologies*, 2351:293–307, 1994.

[137] Y. Sato, Y. Kobayashi, and H. Koike. Fast Tracking of Hands and Fingertips in Infrared Images for Augmented Desk Interface. In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, pages 462–467, 2000.

[138] T. Schl, B. Poppinga, N. Henze, and S. Boll. Gesture Recognition with a Wii Controller. *Gesture*, pages 18–21, 2008.

[139] D. Schmalstieg, A. L. Fuhrmann, G. Hesina, Z. Szalavari, L. M. Encarnaçao, M. Gervautz, and W. Purgathofer. The Studierstube Augmented Reality Project. *Presence*, 11(1):33–54, 2002.

[140] R. Schoenfelder and D. Schmalstieg. Augmented Reality for Industrial Building Acceptance. In *IEEE Virtual Reality Conference*, pages 83–90, 2008.

[141] G. Schweighofer and A. Pinz. Robust pose estimation from a planar target. *IEEE transactions on pattern analysis and machine intelligence (PAMI)*, 28(12):2024–30, Dec. 2006.

[142] Y. Shen, S. Ong, and a.Y.C. Nee. Product information visualization and augmentation in collaborative design. *Computer-Aided Design*, 40(9):963–974, Sept. 2008.

[143] R. Sidharta, J. Oliver, and A. Sannier. Augmented reality tangible interface for distributed design review. In *International Conference on Computer Graphics, Imaging and Visualisation*, pages 464–470, 2006.

[144] S. Siltanen and C. Woodward. Augmented interiors with digital camera images. In *ACM International Conference Proceeding Series*, volume 169, pages 33–36, 2006.

[145] A. Stafford, W. Piekarski, and B. Thomas. Implementation of god-like interaction techniques for supporting collaboration between outdoor AR and indoor tabletop users. In *IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 165–172. Ieee, Oct. 2006.

[146] A. State, G. Hirota, D. T. Chen, W. F. Garrett, and M. a. Livingston. Superior augmented reality registration by integrating landmark tracking and magnetic tracking. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques (SIGGRAPH)*, volume 30, pages 429–438, New York, New York, USA, 1996. ACM Press.

[147] D. J. Sturman and D. Zeltzer. A Survey of Glove-based Input. *IEEE Computer Graphics and Applications*, 14(1):30–39, 1994.

[148] I. E. Sutherland. A head-mounted three dimensional display. In *Proceedings of the fall joint computer conference*, pages 757–764, 1968.

[149] Z. Szalavári and M. Gervautz. The Personal Interaction Panel - a Two-Handed Interface for Augmented Reality. *Computer Graphics Forum*, 16(3):C335–C346, June 1997.

[150] Z. Szalavári, D. Schmalstieg, A. Fuhrmann, and M. Gervautz. Studierstube: An environment for collaboration in augmented reality. *Virtual Reality*, 3(1):37–48, Mar. 1998.

[151] A. Tang, C. Owen, F. Biocca, and W. Mou. Comparative effectiveness of augmented reality in object assembly. In *Proceedings of the conference on Human factors in computing systems (CHI)*, number 5, page 73, New York, New York, USA, 2003. ACM Press.

[152] K. Tateno, I. Kitahara, and Y. Ohta. A Nested Marker for Augmented Reality. In *IEEE Virtual Reality Conference (VR)*, pages 259–262, 2007.

[153] T. Tuytelaars and K. Mikolajczyk. Local Invariant Feature Detectors: A Survey. *Foundations and Trends in Computer Graphics and Vision*, 3(3):177–280, 2007.

[154] A. Uva, S. Cristiano, M. Fiorentino, and G. Monno. Distributed design review using tangible augmented technical drawings. *Computer-Aided Design*, 42(5):364–372, May 2010.

[155] M. VanWaardhuizen, J. Oliver, and J. Gimeno. Table Top Augmented Reality System for Conceptual Design and Prototyping. In *ASME World Conference on Innovative Virtual Reality (WINVR)*, pages 1–11, 2011.

[156] J. C. Verlinden, A. D. Smit, A. W. J. Peeters, and M. H. V. Gelderen. Development of a flexible augmented prototyping system. *Journal of WSCG*, 11(1), 2003.

[157] E. Viirre, H. Pryor, S. Nagata, and T. a. Furness. The virtual retinal display: a new technology for virtual reality and augmented vision in medicine. *Studies in health technology and informatics*, 50:252–7, Jan. 1998.

[158] D. Wagner, T. Pintaric, F. Ledermann, and D. Schmalstieg. Towards Massively Multi-User Augmented Reality on Handheld Devices. In *Pervasive*, volume 3468, pages 208–219, 2005.

[159] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg. Pose Tracking from Natural Features on Mobile Phones. In *7th IEEE and ACM International Symposium on Mixed and Augmented Reality*, pages 125–134, 2008.

[160] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg. Pose tracking from natural features on mobile phones. *IEEE International Symposium on Mixed and Augmented Reality*, pages 125–134, 2008.

[161] D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg. Real-time detection and tracking for augmented reality on mobile phones. *IEEE transactions on visualization and computer graphics*, 16(3):355–68, 2010.

[162] D. Wagner and D. Schmalstieg. Artoolkitplus for pose tracking on mobile devices. In *Proceedings of 12th Computer Vision Winter Workshop (CVWW)*, pages 139–146, 2007.

[163] S. Walairacht, K. Yamada, S. Hasegawa, Y. Koike, and M. Sato. 4 + 4 Fingers Manipulating Virtual Objects in Mixed-Reality Environment. *Interface*, 11(2):134–143, 1998.

[164] R. Y. Wang, S. Paris, and J. Popovic. 6D Hands : Markerless Hand Tracking for Computer Aided Design. In *ACM User Interface Software and Technology (UIST)*, pages 549–558, 2011.

[165] J. Weidenhausen, C. Knoepfle, and D. Stricker. Lessons learned on the way to industrial augmented reality applications, a retrospective on ARVIKA. *Computers & Graphics*, 27:887–891, 2003.

[166] G. Welch and G. Bishop. An Introduction to the Kalman Filter. In *SIGGRAPH, Course*, volume 8, pages 27599–3175, 2001.

[167] G. Welch and E. Foxlin. Motion Tracking: No Silver Bullet, but a Respectable. *IEEE Computer Graphics and Applications*, 22(6):24–38, 2002.

[168] S. White. Interaction with the Environment: Sensor Data Visualization in Outdoor Augmented Reality. In *International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 5–6, 2009.

[169] W. Wohlgemuth and G. Triebfürst. ARVIKA: augmented reality for development, production and service. In *Proceedings of DARE 2000 on Designing augmented reality environments*, pages 151–152, 2000.

[170] Y. Yasumuro, Y. Ishikawa, K. Minami, and M. Imura. Consistent Presentation of Interactive Virtual Objects in Real Space with 3D Markers. *Interactive Virtual Interior Design*, pages 10–12, 2003.

[171] M. L. Yuan, S. K. Ong, and A. Y. C. Nee. Assembly Guidance in Augmented Reality Environments Using a Virtual Interactive Tool. In *Singapore-MIT Alliance Symposium Symposium*, 2005.

[172] M. Zaeh and W. Vogl. Interactive laser-projection for programming industrial robots. In *IEEE/ACM International Symposium on Mixed and Augmented Reality*, pages 125–128. Ieee, Oct. 2006.

[173] J. Zauner, M. Haller, A. Brandl, and W. Hartman. Authoring of a Mixed Reality Assembly Instructor for Hierarchical Structures. In *The Second IEEE and ACM International Symposium on In Mixed and Augmented Reality*, pages 237–246, 2003.

[174] F. Zhou, H. B.-L. Duh, and M. Billinghurst. Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR. In *7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 193–202. Ieee, Sept. 2008.

[175] T. G. Zimmerman, J. Lanier, C. Blanchard, S. Bryson, and Y. Harvill. A hand gesture interface device. *ACM SIGCHI Bulletin*, 18(4):189–192, 1987.

[176] S. Zokai, J. Esteve, Y. Genc, and N. Navab. Multiview paraperspective projection model for diminished reality. In *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 217–226. IEEE Comput. Soc, 2003.