University of Cagliari

PhD School of Mathematics and Computer Science

XXVII CYCLE

Course in Computer Science

New HCI techniques

for better living through technology

INF/01

Coordinator:
Prof. G. Michele Pinna

Author:
Fabio Sorrentino

Supervisor:
Prof. Riccardo Scateni

2014 - 2015

Acknowledgments

I would like to thank all the professors and researchers that I met in these years.

First of all, I would like to express my deepest gratitude to my supervisor Riccardo Scateni, whose patience, competence, table tennis abilities, care, enthusiasm and immense knowledge have been a constant during my PhD studies.

I would also like to thank Davide Spano, for his patience, competence, insightful comments, and considerable encouragements.

Unique bat-thanks go to my bat-colleagues, for creating a bat-amazing bat-work environment which I am bat-proud to join every bat-day.

I am grateful to Marco Winckler and Fabio Pittarello for patiently and deeply reviewing my thesis, providing interesting comments and useful suggestions.

Fabio Sorrentino gratefully acknowledges Sardinia Regional Government for the financial support of his PhD scholarship (P.O.R. Sardegna F.S.E. Operational Programme of the Autonomous Region of Sardinia, European Social Fund 2007 - 2013 Axis IV Human Resources, Objective l.3, Line of Activity l.3.1).

People think that computer science is the art of geniuses
but the actual reality is the opposite,
just many people doing things that build on each other,
like a wall of mini stones.

Donald Knuth

Sommario

All'interno della comunità scientifica che si occupa di Interazione Uomo-Macchina, molti lavori hanno come obiettivo quello di studiare l'efficacia delle nuove tecnologie per migliorare lo stile di vita e, diversamente da altri campi, questi devono essere perseguiti con un approccio multi-disciplinare.

La tecnologia è in continua evoluzione e questo porta a migliorare le nostre vite sotto diversi aspetti legati ad esempio all'educazione, alla salute e alle comunicazioni. Questo perché il fine ultimo della tecnologia è proprio quello di rendere le attività quotidiane più semplici.

In questa dissertazione si approfondiscono tre aspetti principali: il primo è l'apprendimento mediante l'utilizzo di nuove tecnologie, il secondo è il miglioramento delle attività quotidiane utilizzando dispositivi innovativi mentre il terzo riguarda l'utilizzo dei dispositivi mobili in combinazione con algoritmi e tecniche di elaborazione di immagini e informatica grafica.

Nella prima parte dell'elaborato viene descritto lo stato dell'arte e i lavori correlati che sono stati necessari per implementare i tool realizzati sia in ambiente mobile che desktop.

Nei progetti esposti proponiamo l'utilizzo di diverse tecnologie con il fine ultimo di migliorare l'esperienza utente. Per raggiungere questo scopo abbiamo comparato queste soluzioni, che fanno uso di strumenti legati a tecniche di realtà aumentata e realtà virtuale, confrontandole con i metodi tradizionali. In alcuni esperimenti abbiamo applicato tecniche di gamification, prendendo spunto da quelle utilizzate in ambiente mobile, per dimostrare come l'utente possa essere motivato e intrattenuto durante, ad esempio, la propria routine di allenamento.

Sono stati progettati e implementati diversi sistemi integrati atti a migliorare l'apprendimento dei bambini, per rendere più piacevole l'esperienza di shopping, per offrire nuovi servizi legati al turismo e per migliorare le attività legate al fitness e al benessere in generale.

Nel capitolo finale della tesi saranno discussi i risultati ottenuti inserendoli all'interno del contesto più generale del miglioramento delle attività quotidiane usando le nuove tecnologie, evidenziando le nozioni apprese da ogni progetto analizzato e al contempo proponendo possibili sviluppi futuri.

Abstract

In the Human Computer Interaction community, researchers work on many projects that investigate the efficacy of new technologies for better living, but unlike other research fields, these researchers must have an approach that is typically multi-disciplinary.

Technology is always developing thus improving our lives in many ways like education, health and communication. This due to the fact that it is supposed to make life easier.

This dissertation explores three main aspects: the first is learning with new technologies, the second is the improvement of real life by using innovative devices while the third is the usage of mobile devices in combination with image processing algorithms and computer graphics techniques.

We firstly describe the progress on the state of the art and related work that have been necessary to implement such tools on commodity hardware and deploy them in both mobile and desktop settings.

We propose the usage of different technologies in different settings, comparing these solutions for enhancing the interaction experience by introducing virtual/augmented reality tools for supporting this kind of activities. We also applied well-known gamification techniques coming from different mobile applications for demonstrating how users can be entertained and motivated in their working out.

We describe our design and prototype of several integrated systems created to improve the educational process, to enhance the shopping experience, to provide new experiences for travellers and even to improve fitness and wellness activities.

Finally, we discuss our findings and frame them in the broader context of better living thanks to technology, drawing the lessons learnt from each work while also proposing relative future work.

Contents

Chapter 1

Introduction and Motivation

In this dissertation we focused over three main aspects: the first is learning with new technologies [3], the second is the improvement of real life by using innovative devices [4] while the third is the usage of mobile devices in combination with image processing algorithms and computer graphics techniques [5].

Through the works presented in this thesis we have tried to address three main questions: *how does technology facilitate learning? how can technology advances affect our life? is the mobile the future of every thing?*

The main goal of the first aspect is to define some of the key issues in children and adult learning thus presenting several developed tools in order to demonstrate how to face such kind of challenges. In fact, we applied different technologies in different learning settings. One the most representative work presented in the learning section is Speaky Notes [3.1]. It is a web authoring system that allows teachers to create a mobile application supporting children in learning a new language in a more pleasant and entertaining way by using augmented reality.

With SoSphere [3.2], we compared two different solutions for enhancing the interaction experience of a planetarium application, both replicable at a reasonable cost. The first version is based on a simple multi-touch paradigm, while the second one exploits a full-body interaction together with a projection on geodetic sphere. We detail the technical implementation of both versions and, in addition, we discuss the results of user-study that compared the two modalities.

The test results highlight a tradeoff between the control and the users' involvement in the virtual environment.

With RIFTart [3.3], we introduce a VR tool for supporting the teaching and studying of Art History. Using RIFTart the teachers can configure virtual museum rooms, with artwork models inside, thus enhancing them with multimodal annotation. The environment supports both the teachers during the lesson and the students during rehearsal. The application, implemented completely by using Web technologies, can be visualised on large screens and head mounted displays. The user test results advance the understanding of the VR effects on classroom usage. We demonstrate that VR increases the motivation of high-school students towards studying Art History and we provide an in-depth analysis of the factors that contribute to this result.

We then applied similar technology settings for improving real life with new technologies. In RiftAbike [4.1], we apply well-known gamification techniques coming from different mobile applications for demonstrating how users can be entertained while working out with an exercise bike. We evaluate the effectiveness of such techniques through a user study, deriving a set of guidelines for creating what we call fitmersive games (immersive games for fitness).

In Interactive Shops [4.4] we present a proof-of-concept of an integrated system for enhancing the shopping experience in a shoes' shop from inside and outside. The system uses two modules:(i) one internal to the shop using an interactive totem;(ii) one external to the shop based on the customer's mobile device and the interactive external shop surface. We describe the technical architecture of the two modules and two different scenarios of the user experience.

The work WoBo (World in a Box) [4.2] aims to provide a new experience for travellers, allowing them to visit distant or hardly reachable places through the exploitation of consumer cameras and a head mounted display. The experience consists in watching a 360-degrees video with 3D audio in a dedicated cabin. The user can select videos shot in different places, which have been created with six consumer cameras. We describe the proposed experience, the hardware and the software used for a first prototype.

The SmartMirror project [4.3] aims at creating a seamless interactive support for fitness and wellness activities in touristic resorts. The overall idea is to evaluate the current physical state of the user through a technology-enhanced mirror.

We describe the state of the art technologies for building a smart mirror prototype.

In addition, we compare different parameters for evaluating the user's physical state, considering the user's impact, the contact requirements and their cost and we depict the planned setup and evaluation setting for the Virtuoso project.

Finally, given the importance of mobile devices and their utility as a learning tool, we developed four projects that implement well known algorithms that are often used on desktop environment. In these works the main challenge is the constraint of such devices due to their computational power.

Each work has its own distinct characteristics, but some of them are similar in terms of used techniques and approaches. Two of these projects, Chromagram [5.1] and Click and Share [5.3], exploit two algorithms that make use of image processing techniques. This kind of algorithms are computationally expensive, especially on mobile devices and we here propose their implementation and their evaluation in terms of computational times. The third and the fourth project aim to offer geo-located information, through augmented reality contents, are both thought especially for tourists. In fact, VisitAR [5.4] allows users to have a more immersive sightseeing experience, providing tourist information in a more entertaining way. Similarly, AR-Garden [5.2] is a navigation system for limited extensions inside urban areas which permits to wander around and gives access to related information using augmented reality.

For each tool we describe the lesson learnt from the design of a system and the techniques we identified for reaching an interface that is thought to be easy to use, fairly intuitive and self-explanatory. This characteristics are necessary for those tools that are designed for educators, parents and in general people that have no programming skills.

1.1 Pros and cons of technology for learning

The learning topic is often correlated to young learners, this is due to the fact that children are the bigger part of the totality of learners. It is not surprising that the major part of research publications aim to solve and improve the teaching dynamics oriented to preschoolers and school-aged children. In this dissertation, we focus on these categories but we also considered a wide variety of other typologies of learners/users.

New technologies are becoming the best friends of today's students, not only for the young ones but even for aged learners. In fact, in the everyday life, we find several examples and demonstrations of how this kind of technologies helps learners in extremely different fields.

Children are naturally curious and they tend to explore and discover their environment without any conscious effort. We can affirm that they learn from everything they do, for better or for worse. Like other human behaviours, if this explorations bring positive mental and emotional states, like happiness or enjoyment, children want to learn more and more. Due to this fact, children's educators and parents have a crucial role in learning. They have to give the right amount of support and encouragement, especially in the early period of childhood. In order to better analyse the user typologies we can define four age-related development periods:

- Newborn (ages 0-5 weeks)
- Infant (ages 5 weeks - 1 year)
- Toddler (ages 1-3 years)
- **Preschooler** (ages 3-5 years)
- **School-aged child** (ages 5-12 years)
- Adolescent (ages 13-19 years)

In most of our learning-oriented project, our user target is represented by the fourth and the fifth group. In fact, in developmental psychology and developmental biology, we find an interesting definition of the "children critical or sensitive period" that refers to several overlapping periods of development where a child is sensitive to a particular stimulus or type of interaction. Children become skilled at numerous activities without formal instruction. In literature, some researchers differentiate between the adjectives "critical" and "sensitive", but in our dissertation we do not consider them differently. Once a sensitive period is passed, the skill must be formally taught, thus taking a great deal of effort to learn.

Several studies have shown clear evidence for sensitive period effects in the language acquisition [79]. This effects refer to both first and second language acquisition [80].

Even if new technologies offer a good chance to improve the way people learn, we must keep in mind that new devices also bring cons, not only pros. These days, the children development includes technologies such as smartphones, tablets and computers. Researchers try to figure out which of the available gadgets are appropriate for each stage of childhood. Gwenn Schurgin O'Keeffe M.D., tech expert of the American Academy of Paediatrics (AAP), says that there is not really a *right* age to allow kids to dip a toe into the digital world [125].

It is clear that children are surrounded by technologies all the time, especially when nearby adults. Due to the fact that one of the social milestones for children, especially for toddlers, is imitating the behaviour of others it is easy to understand that they are influenced by their parents' habits with this kind of electronic devices.

Regarding these new technologies, such as smartphones and tablets, and their owners' age range, we can affirm that year after year, it is common to see kids using their parents' smartphones or even their own personal mobile devices. The same way children learn to speak in the first years of their lives they also learn to interact with these mobile devices while still in the childhood. At the beginning, they learn how to unlock the phone, then they tap randomly over the screen surface, and finally they start to explore the device features. It is impressive how fast children learn to use this kind of devices, and after a while they understand the correlation between icons and applications. Due to this it is clear that in primary school we will see skilled smartphone users that use a tablet to improve their learning. Having such skilled users makes it possible to use these devices for learning.

All that glitters is not gold.

Interacting with electronic devices may take away time of interacting with other children and people in general. Usually, interactive games for children allow them to only interact with the screen. Instead, it is important for children to have face to face interaction with other people in order to develop empathy and build social skills in general. Besides, at the moment, no educational software can compete in teaching tasks. Again, children touch everything, this is the way they learn about the surrounding environment. Take the sand for example. Sand is very well-suited for the need to explore and for the imagination inherent in children.

No existing software can simulate it for the time being. On the contrary, such kind of devices offer too much stimulation. It can be a trouble due to the fact that children are naturally exposed to an huge sensory stimulation. These devices can overstimulate their sensory system.

Another potential negative aspect is that kids who become too *tech-dependant* have trouble in focusing and paying attention to less high-tech educational mediums. Last but not least, children need to get moving. Physical activity has a major role in their growth, and unfortunately, computers can reduce such kind of activities with all the known negative consequences.

Even if Education is usually related to children, there are a lot of cases where the students are grown people [176] and the topics are taught by using technology-based approaches. Recent statistics show us that the number of learners who are interested by these needs are a considerable number. Thanks to the average life expectancy increasing in the past years, the world labor force is increasing worldwide too. This is especially true for workers that are 65 years old or older [37]. A recent article of an American press (USA Today) noted that *"Americans 50 and older are staying employed longer than at any time since such demographics were tracked"*. This trend seems to continue in the next years, in fact by 2018, a 25% of the American workforce will be aged between 55 and older individuals [164]. The same situation exists and will increase in the European Union. This leads to another mechanism, in fact these workers are postponing retirement and sometimes they engage new employments thus needing a training for their new jobs [108]. Such kind of new employees get their training through technology-based methods. These methods include interactive video systems, computer-based training, computer-assisted training, e-learning, computer-based simulations, virtual reality training and more. In several works, researchers have shown that, in certain instructional interventions, there is a different impact on learning results depending on participant age.

1.2 How technology helps us our daily lives

One of the main fields that leads to technological innovation is probably the military one. Leaving out the consideration if war is, or is not, morally acceptable, war creates a higher demand for technological innovation rather than the one under conditions of peace. A cliché tells that “No one wins in war”, but it is clear that much of the technology we have today is a direct side effect of wars.

New technologies in the Military field help to significantly improve training effectiveness while reducing its own cost. Some of them are Service-Oriented Architectures (SOA), cloud computing, data analytics and some gaming approaches. All the advances in simulation programs, gaming and other technologies have improved the training of military forces significantly. Last but not least, advances in Computer Graphics are making these simulators, year after year, more realistic [105]. In these simulators, aircraft pilots preparing for deployment to a specific war zone can fly simulated missions over the same geography they will encounter in the field. It is clear that by using these technologies it is possible to save a good amount of money since a fighter aircraft costs around a hundred million dollars, without even taking into account the pilot’s life, which is priceless.

Again, one of last products in the progress of military technologies are the combat drones. These vehicles allow military forces to deploy weapons in war remaining thousand of miles away from the center of the battlefield. This small and accurate vehicles require very small components, trustable and reliable wireless connections, precision sensors and, often, an HD vision system. All these requirements push forward the progress of this kind of electronic components, these parts will be integrated into different devices like tablets, smartphones and others. Actually, these improvements in technologies are a good chance to increase the capabilities of mobile devices and consequently they offer students more powerful tools for their education.

Some people think that technology is killing us. Killing friendships, marriages and even good habits. And it is quite clear that there can be a negative side resulting from inappropriate use of technology, and this negative side can have serious consequences. To make the best use of technology, teachers and especially parents must also recognise these aspects in order to avoid them for their children.

But it is also clear that technology makes our lives easier. Today’s students have tremendous opportunities to learn and to connect by using technology itself thus empowering their capabilities.

They can, for example, use search engines or educational sites to better understand the concepts learnt at school. But, these improvements in our daily lives are tangible for adults too. Think about the online calendars that help people coordinate their schedules and think about how hard it was to organise them without devices such as tablets or smartphones. Talking about travel, before there were modern vehicles, it took weeks or months for people to travel. Now it takes hours. People can now pay bills and bank online or can communicate with doctors to save time and money.

Technology gives us new experiences, like talking to people all across the world, or even seeing their face. Some people can not travel because they are sick or not mobile. With technology, you can see these people, not in person but you can talk and see their face.

Last but not least, we think that the use of technology can improve the quality of our personal experience, with respect to wellness and wellbeing. In fact, by using new devices that are oriented to the wellness analysis, such as activity trackers and smart watches, we can collect a lot of data that can be aggregated in a way that is very useful for the creation of refinement of wellbeing strategies.

1.3 Mobile is (already) the future of everything

As reported by the International Data Corporation (IDC), desktop sales have been in a free fall for years due to the fact that users have shifted to laptops, tablets and mobile devices in general. The only category that survives this trend is the mini-desktop one. But how far away are we from their death?

Apart from companies that were born to be mobile-oriented like Twitter or Instagram, even the companies that do not have a mobile strategy are implementing it because they could be out of the game in the coming days. Mobile affects anyone who is selling something, take restaurants as an example, the mobile is the mean through which users decide where to have lunch and dinner. And we can take thousands of other examples. Think about banks, television, social networks, business meetings, presentations, travels, photography, tourism.

If, before the mobile era people were enthusiastic to see that a company offered a dedicated mobile application, nowadays they want it, and if they can not do something on their mobile phone they are frustrated and ready to skip to a better offerer.

As companies should do, research has to work to understand users' mobile habits. This, in order to follow the trend thus evaluating and testing old tasks on such innovative technologies. We need to see users using their smartphone as they work and live.

Chapter 2

Related work

In this dissertation we focused over three main aspects: the first is learning with new technologies, the second is the improvement of real life by using innovative devices while the third is the usage of mobile devices in combination with image processing algorithms and computer graphics techniques.

It is clear that these aspects cover an huge amount of fields, where there is also a large number of works that are useful for the final goal. For this reason, in this section we have highlighted the used material by following the same thesis schema. For each one of the three aspects we analyse the major contributions that have influenced our work.

2.1 Learning

In some detailed works presented in this thesis, we focused on how playing can improve the way children learn [65]. Due to the known fact that playing is fun for (almost) every child, it is clear that the combination of playing and learning could give us better results in term of children satisfaction and educational success.

Playing is fun for children and it represents one of the way they actually learn [143] [130] [27]. Through play, children learn about themselves, their environment, their family and the whole world around. As they play, children learn to solve problems and enhance their creativity actually developing skills and personalities. Play in early childhood is the best foundation for success in school. Language develops as children play and interact with other people. Learning to cooperate and play by the rules are important interpersonal lifelong skills, all fostered by playing with other children. Positive play experiences develop positive emotional well-being.

Play also helps the child master the environment. Again, when children feel secure, safe, successful and capable, they acquire important components of positive emotional health.

Research [44] has proven that parental participation in the education process plays a major role in their scholastic capabilities and general development. Based on this observation, it is important to use an approach that encourages parents to help their children in doing their homework. Adults must provide an environment and resources that interest children by stimulating their curiosity. This fosters their imagination, provokes questions and inspires children to explore.

As explained in [65] *"we need to find a balance between safe and challenging environments for children. If there is too little challenge, children will find new ways to take risks, often in unsafe and destructive ways. Too much challenge and too little supervision can result in children feeling frustrated, unsafe and afraid."*. Due to this, the adult's role in children education is really important. They have to control and organise children's time, and, at the same time, provide opportunities for free and creative activities. Actually, finding the right balance between organised and free play is the main challenge for educators and parents.

In particular, the role of parents is a topic of critical importance in children's education [72]. In fact, many works in literature analyse the influence of parents in children education processes and this kind of works are mostly presented in psychological fields.

Thanks to the works shown in [126], researchers have studied the positive effects of parent involvement on children when schools and parents continuously support and encourage the children's learning and development. In fact, according to Henderson and Berla [68], the most accurate predictor of a student's achievements in school is not income or social status but the extent to which that student's family is able to:

- Create a home environment that encourages learning;
- Express high (but not unrealistic) expectations for their children's achievement and future careers;
- Become involved in their children's education at school and in the community.

This and other studies show that parent involvement activities, that are effectively planned and well implemented, result in substantial benefits to children, parents, educators and the school too.

One of the main topics that are involved in our works, and that are correlated to learning, is certainly the augmented reality. Augmented reality, abbreviated as AR, can be defined as a type of virtual reality where digitally rendered 2D/3D objects generated by the computer are overimposed onto the real world. Indeed, the final result is the combination of real scene enriched with the virtual one. The main goal of these techniques is to give users the illusion of perceiving both scenes as a single one, without noticing where the real scene ends and the virtual one starts. Milgram [113] defined a continuum of real to virtual environments, in which AR is one part of the general area of mixed reality. In both augmented virtuality and virtual reality, the surrounding environment is virtual, while in AR the surrounding one is real.

Even if augmented reality has been gaining worldwide significance since the nineties, it finds its origins in the fifties. Augmented reality is marked as a hot topic during these years. Existing devices that use it, such as smartphones and tablets, or that are being developed, like special glasses, allow the enrichment of our real environment with virtual information. This information is presented in a way that makes it much more fun, useful and accessible. This approach, if applied to the education field, means the creation of new incredible scenarios which would otherwise be impossible. For example, using augmented reality, we can hold something as small as an atom or as big as a star in our hands. In 2010, augmented reality reached number four in Time magazine's 10 Tech Trends list [50] and then BBC reported that it *"looks set to become a serious commercial tool"*. If augmented reality techniques offer new learning opportunities, it is clear that they also create new challenges. Augmented reality should be considered as a concept rather than a technology. Another aspect that motivates us to pursue our goals is the impact of augmented systems on students' enthusiasm; in fact, several studies have noticed the improvement of students' motivation towards learning with this approach [39] [107].

In [139] we can find an AR approach used for geometry teaching while in [49] a comparison between two 3D astronomical tangible models is shown. The first one is built using physical models while the second one uses an augmented reality approach. These tangible models show augmented views of the celestial bodies and support the students' investigation by using spatial visual guides. Built through an exploratory experiment with 39 primary class students, the authors investigated the potential of the augmented scenario comparing the learning results with those achieved with the traditional physical model. The results show that this AR learning environment enhances astronomical learning significantly. Moreover the model allows users to concentrate on their tasks also enhancing the task controllability.

Among augmented reality applications that are focused on children learning, we have analysed ZooBurst [185] that is a *"digital storytelling tool that lets anyone easily create his or her own 3D pop-up books. As an educational tool, ZooBurst gives students new ways to tell stories, deliver presentations, write reports and express complex ideas"*.

This tool is mentioned as one of those with the most potential for language teaching and learning case studies in the Salmon and colleagues' study [145] where an evaluative framework has established a potentially useful baseline for making decisions about the use of augmented reality applications in the classroom. Even if ZooBurst has not been specifically designed for language teaching and learning, it is a flexible tool that could be used for this purpose. This project highlights how children can learn through playing, that is our own way of thinking. Cascales and colleagues in [24] report an experiment where a sample of 36 preschoolers, having the same curriculum, were split in two groups (with/without an AR approach): even if the learning outcomes do not present significant statistical difference, they affirm that all participants considered the use of augmented reality techniques a good approach in the teaching/learning process.

In their conclusions we can find some interesting points, in particular they say that *"augmented reality also promotes communication skills, promoting all kinds of interactions in the classroom between teachers and teachers, teacher and students, students and students, students and families, families and families"*, therefore considering it a learning tool fostering social aggregation.

These augmented reality techniques and methodologies have been really useful for "Speaky Notes" [3.1], a web authoring system that makes it possible to create a mobile application that supports children in learning a new language in a more pleasant and entertaining way by using augmented reality. This application allows pupils to improve their speaking skills by turning language acquisition into a game under the supervision of both teachers and parents. Our contribution is focused on understanding how digital technology can facilitate learning while keeping in mind that it is a wide and interdisciplinary issue.

The potential of an augmented reality approach in education is clear, and it is easy to understand that such software may offer many exciting features to children in their learning process. But, even if these aspects are interesting for the community, little research has been done into the teaching and learning field by using such kind of methodologies as explained in [145].

During the development of our projects we used both augmented reality and 3D avatars. Their combination has been tested, for example, in the "SuperAvatar" project [3.4] where we used EPA (Embodied Pedagogical Agents) for presenting touristic and cultural information to children.

These embodied pedagogical agents, also knows as Guideboots, have been embedded in learning applications targeted for different types of users, ranging from the elderly [128]] to different categories of young people, such as college students [91] or children [64], [55]. They can have different roles in the software, such as mentors, learning companions or instructors. Their effectiveness is related to the social nature of the interaction they support, since they might be able to establish an affective link with the user [81].

Therefore, their visual appearance is crucial for their communication success. Different studies in literature provided contributions to this topic: here we focus on the children-related work.

Haake and Gulz [64] conducted a user study on 90 school children between 12 and 15 years old. They established three dimensions related to the selection of a particular EPA: the static visual appearance (detailed or simplified), its role (instructor or learning companion) and the communicative style (task or task-and-relation oriented). The analysis of the collected data shows a preference for a more stylised appearance of the EPA, in particular for female students. Similar results were found by Girard and Johnson [55] with a study on 86 primary school children, aged between 7 and 11. The children had to choose between an instructor and a learning companion EPA, simplified or detailed in appearance and humanoid or smiley shaped. In general, children preferred learning companions over instructors, simplified characters for both roles, humanoid-shaped for instructors and stylised for learning companions. According to these results, we selected familiar comic-book superheroes for the EPA role to be included in the AR visiting guide, for increasing the children's affection towards the virtual guide.

The idea of exploiting agents for enhancing the visiting experience has been investigated in different works in literature, in particular in museum settings. Most of them exploit these agents in stationary positions. For instance, in [92] the authors describe their experience with "Max", a conversational agent as guide for a computer museum in Paderborn. The case study demonstrated that the visitors communicated with the agent as if it was a human being. In [162], two conversational agents provided information to the visitors answering questions at the Museum of Science in Boston. Other solutions such as [69] distributed the agents on the whole museum, providing robots as physical counterparts of the virtual agents.

The visual appearance of the robots and the museum contents are enhanced by the guide software through head-mounted displays. The robot allows taking the interaction initiative moving towards a particular artwork in order to be followed by the visitor. Our approach exploits the agents in an ubiquitous setting, deploying the guide on the user's personal device.

Even if the augmented reality offers new interaction opportunities, especially in ubiquitous settings, it is not the only approach that offers such kind of innovative interactions. In fact, in our works we also analysed the multitouch and the full-body interactions. This interaction techniques have been evaluated in the "SoSphere" project, where we exploit different low-cost sensing devices in order to provide engaging interaction experiences, trying to put into practice the ubiquitous computing vision, where the user is expected to interact naturally with the technology without even realising its mediation.

Many investigators used the sky and space exploration to provide examples of immersive systems. A description of how virtual environments can be exploited for this kind of tasks can be found in [123] which describes, among the other settings, how such a kind of environment is exploited by the NASA. Another relevant example is the work in [153], where the authors exploited magnetic sensors in order to support the user while pointing or searching for real stars. In addition, they also exploited the Wiimote controller for guiding the recognition of the constellation shapes.

In [15], the authors exploited a spherical display for creating a 360-degrees space for visualising content for multiple users minimising the occlusion. The authors exploited such display for collaborative settings. We used a larger spherical display, and we implemented a free-hand interaction paradigm with such screen. In addition, we projected the image on the concave surface rather that the convex one. Francese and colleagues [51] measured the presence and the immersion of different 3D gesture interaction techniques (namely remote-based and full-body). They concluded that the perceived immersion of such an interaction technique is high and that the users pass quickly from a novice to an expert style of interaction. We created a setting that exploits the same interaction technique, adding a more realistic projection of the sky map, in order to increase the user's immersion feeling even further.

So far, we focused on the interaction modalities, but we also need to understand how to present the content to users. This content could be made through an augmented reality or a virtual approach.

This understanding has been necessary, for example, in the “RIFTart” project, a virtual reality tool for supporting the teaching and studying of Art History. By using RIFTart the teachers can configure virtual museum rooms, with artwork models inside, and enhance them with multimodal annotations. The environment supports both the teachers during the lesson and the students during rehearsal.

The recent development in consumer hardware lowers the cost barrier for adopting immersive Virtual Reality (VR) solutions, which could be an option for classroom use in the near future. By using “RIFTart” the teachers can configure virtual museum rooms, with artwork models inside, and enhance them with multimodal annotation. The environment supports both the teachers during the lesson and the students during rehearsal.

Training and learning were two of the most important applications of virtual reality (VR) since the first introduction of technologies supporting the creation of virtual environments. Already in the 1980’s, VR was used to replicate dangerous or safety-critical settings (e.g., airplane cockpits, space exploration etc.), or simulated contexts impossible to sense directly (e.g., cell evolutions, atomic reactions etc.) [67, 179]. At the time the hardware was really expensive and its cost was worth only if compensated by other relevant aspects like, for instance, the safety of air traffic. As the personal computers expanded their computational power, desktop-based solutions have been providing VR environments at a reasonable cost for many years, increasing the learner’s engagement even if providing a less immersive experience [40].

Nowadays, the technology evolution has led to the creation and commercialisation of different consumer-level devices allowing to create immersive experiences at a reasonable cost, and we foresee that the availability of such hardware will increase in the next future. For instance, the Oculus Rift [124] represents the first customer-level VR head mounted display (HMD) for gaming. Other mobile-based HMDs are currently under development, (e.g., the Samsung Gear VR [146]), while very cheap solutions for transforming mobile phones into HMDs [59] already exist.

Technology is ready for employing immersive VR experiences in classrooms, in the near future, for teaching more subjects than those covered in the past by VR environments.

There is an extensive research literature on virtual reality application to teaching and learning. Merchant et al. [110] provide a review on the effectiveness of VR-based instruction in elementary, middle and high school.

They classify the virtual reality environments in three categories:

- simulations, which are interactive digital environments imitating real-life processes and situations;
- games, which are special simulation environments which include goals, achievements and levels to be reached by following narrative plots;
- virtual worlds, which exploit the illusion to be in a 3D space, the ability to interact with 3D objects, the avatar representation of the learner and communication with other users inside the world.

Virtual reality has been adopted in different areas for teaching, especially when it allows the action of inexpert people that may cause danger or may raise ethical issues. For instance, in [41] the authors exploited the virtual reality for medical education, taking advantage of the reduced risks and costs, together with the possibility to instruct students from a distance.

The research community studied such advantages, analysing the effectiveness of web-based multi-user virtual environments from a pedagogical point of view [31], and comparing them against 2D alternatives [38]. Most of the studies focused more on desktop-based virtual environments with respect to more expensive settings like cave automatic virtual environments (CAVEs) or head mounted displays (HMDs).

The already mentioned availability of hardware supporting immersive experiences led to a new generation of learning environments, exploiting the increased fidelity perceived by users [12]. Both the topic and the target audience vary: veterinary anatomy [172], architecture and building engineering [173] for university students, biology for K-8 students [96] and even subway evacuation procedures for a larger audience [152]. We are interested to analyse the effect of adopting an immersive setting not only on the perceived fidelity of the environment, but also on the impact on student's motivation towards a specific topic. The application of virtual reality in the art and cultural heritage field has different purposes, e.g. the acquisition for preserving [103] or restoring [61] artworks; the reconstruction of a 3D scene from a painting or fresco [23] and more. There are different examples of educational VR-based applications that foster informal learning, especially in museums, where providing a playful interactive experience is crucial for attracting people, especially children. For instance, already in 2000 at the Foundation of the Hellenic World in Athens was possible to take a virtual guided tour in both Olympia and the ancient Miletus [53]. More recently, Kennedy et al. [90] reconstructed the St. Andrews Cathedral, which can be virtually visited with the Oculus Rift.

In [160], the authors survey such applications describing different implementation settings and technologies. They define a virtual learning museum as a specific type of virtual museum, which presents contents in a context and interest dependent way, in order to motivate a real visit by stimulating the curiosity on contents that better fit the user's interest.

In the RIFTArt project, we focus on a different type of virtual reality application, which provides material for a formal lesson on Art History and cultural heritage. Exploiting virtual reality for creating this kind of material has been under investigation by the research community considering again the high hardware cost for a classroom set-up. The lesson learnt in this study has been useful in two other projects that aim to improve real life by using new technologies. In the "Rift-A-bike", as deeply explained in its section [4.1], we used the Oculus Rift head mounted display for immersive fitness activities. The same device has been used in "WoBo" (World in a Box) [4.2], where we provide immersive and realistic virtual journeys.

2.2 New technologies in everyday life

Technology has definitely improved our lives, this is because of the various aspects of it: computers, internet, phones and more. Sometimes, technology leads people to be too much dependent on technology, but the key factor that has made us dependent is ourselves. In fact, the key is how, and how much, we use technology. Thanks to modern communication opportunities, we can talk with almost everyone almost everywhere, with a single click.

During the last years, we developed several social experiments that aim to improve real life aspects by offering new and innovative opportunities. Think about the fitness activities, it is clear that exercising outdoors may improve energy levels thus decreasing the stress level. But it is also clear that not everybody can do outdoor activities. For this reason in the "Rift-A-bike" project we apply well-known gamification techniques coming from different mobile applications (i.e. e-learning, lifestyle, fitness, etc.) for demonstrating how users can be entertained while working out indoor with an exercise bike. In this work we analysed two main aspects: how new technologies help people while doing fitness activities and which gamification techniques are suitable for this kind of activities. The introduction of devices for recognising the user's gestures (such as the Wiimote or the MS Kinect) marked a milestone in the videogame roadmap signing the birth certificate of the active video games [16]. Even if they cannot be considered a real substitute for a proper physical exercise [122], they have a positive impact on different physical abilities [114].

At the same time, the increasing number of people that own a smartphone shifted the focus in the fitness applications market towards mobiles. Many of them adopted the gamification in their design. For instance, in "Zombies, Run!" [184] the user's goal is to save people from zombies bringing different supplies and rescuing survivors. In order to complete these actions in the game, they have to run in the real world. Besides the single-player objective, commercial applications exploited also social networks for stimulating competition among friends. In this category, one of the most famous applications is Fitocracy [48], which logs the user's physical activity awarding points, badges and so on. Users are ranked according to such rewards, and this both stimulates the competition and keeps the motivation high.

Other mobile applications combine dedicated hardware (such as wristbands, chest belts etc.) to the standard smartphone sensors, in order to have a fine-grained measurement of the user's physical parameters. For instance, the application Nike+ [121] provides a user profile where the data coming from different devices is stored and, in a similar way with respect to Fitocracy, allows the comparison with other people in the community.

Different kind of technologies have been analysed in the "SmartMirror" project, which aims at creating a seamless interactive support for fitness and wellness activities in touristic resorts through a smart mirror. We reviewed the existing technologies that may be used for building interactive mirrors. Thanks to the recent introduction of new materials and the overall ICT developements, it is now possible to adopt different solutions for building interactive mirrors, transforming them from simple reflecting surfaces to more interactive devices. We can group them in three sets:

- multimedia players (e.g. music or videos);
- interactive home controllers;
- augmented reality devices.

We discuss their advantages and disadvantages in general, and in particular for the "SmartMirror" project.

The multimedia players category contains devices equipped with a touch screen, or TVs enhanced with external full body tracking devices (e.g. Microsoft Kinect). However, most of them only support entertainment tasks, such as playing music or videos.

- **Mirror 2.0** [62] combines the advantages of a smartphone and a mirror. It contains an LCD display, positioned behind the glass. The user controls its functionalities through gestures and vocal commands.

- **Smart Washbasin** [149] displays different information on a washbasin mirror such as mails, weather forecast, water temperature and pressure, the calendar and the user's weight measured through a built-in scale in the base portion. It is possible to control it without touching the screen surface, since it is equipped with proximity sensors able to track the hands position and motion.
- **NEOD Framed Mirror TV** [120] is a standard LCD screen (up to 50 inches), covered by a mirror, specifically designed for the screen. The screen only provides TV functionalities, and it does not provide more interactive features.

The interactive home controllers, instead, allow to control different home appliances through the interface displayed on the mirror. The following is the list of devices belonging to this category.

- The **Smart Mirror for home environment** [6] allows to control all the smart devices at home. It relays on face recognition for authenticating the user and displays personalised information (news, mail, messages etc.). The system exploits a touch screen monitor and two webcams, one for the face recognition and one for the home surveillance.
- The **Multi Display in Black Mirror** [165] by Toshiba is a prototype that combines the functionalities of a tablet together with the reflecting surface of a mirror. It provides two configurations taking into account two different home environments: the bathroom and the kitchen.

Finally, many mirror devices provide augmented reality features, which are especially useful for advertisement purposes in shop showcases. In addition, there are different attempts to use smart mirrors for supporting routine activities. In this category we can include, for instance, an application that monitors and guides a user while brushing teeth through the information displayed on the smart mirror. Finally, the so called "medical mirrors" are particularly interesting for our project, since they can measure different physical parameters such as heart and breath rate.

- The **Reveal** project [71], created in the New York Times research and development department, consists of an LCD Display covered by a mirror glass. The device exploits a Microsoft Kinect for tracking user's movements in real-time. It visualises different information on its surface (calendar, mail, news, online shopping websites, instant messenger and more). In addition, it responds to vocal commands. A peculiar feature is the medicine box scanner, which allows the user to buy medicines by recognising their packages.

- The **Cybertecture Mirror** [97] is a complete PC contained into a 37 inches mirror, equipped with a 32 inches LCD screen. Through a smartphone application, the user accesses different information overlaid on the reflected image. The interface allows to visualise instant messages, the calendar, the mailbox, and the weather forecast. In addition, it provides information on the user's physical state.

- The **Interactive Mirror** [129] by Panasonic seems to be an ordinary mirror: neither camera nor other sensors suggest the features of a smart object. Once the user sits down in front of it, the mirror displays an enlarged frame for her face, together with menus for accessing different functionalities. The system analyses the face hydration, wrinkles and other details in order to recommend products and treatments to take care of her skin (e.g. to make it softer etc.), to slow ageing and so on.

- The **Connected Store Demo** [43] by eBay and Rebecca Minkoff provides interactive experiences in both the store showcase and in the fitting room, equipped with a mirror surface overlaid by the store user interface. In the showcase setting, the user explores the different items in the store. Once she finds something interesting, she requests to try it in the fitting room. Once finished, the shopper prepares the fitting room with all the items. Inside the fitting room, the user exploits the mirror for looking for other items and/or providing feedback. In addition, she may select some of them for buying.

- The **Brushing Teeth Mirror** [117] displays the information collected by a smart brush about inflammations or infections of the teeth and gums.

- The **Medical Mirror** [136] combines computer vision and signal processing techniques for measuring the heart rate from the optical signal reflected of the face. The prototype consists of an LCD display with built-in camera and a two ways mirror fitted onto the frame.

These innovative technologies are also influencing the way shops sell their products. Such kind of new interaction opportunities are studied also from different perspectives than the computer science one. During the implementation of our prototype for a interactive shop, described in its [4.4] section, we analysed considerations that come from different scientific areas. In fact, several studies in sociology and related fields give sellers useful guidelines for creating attractive showcases. Of course, the purpose of them is to increase the number of customers attracted by the shop and, thus, keen to enter and browse the goods.

Moreover, these studies also analyse the buying behaviours of customers in correlation with different shop elements.

Bearden and colleagues [13] highlight that atmosphere, location, parking and salespeople's attitude are critical aspects affecting the store image while Petruzzellis and colleagues [131] explain that music is one of the most important elements of the retail atmosphere. Even the shop's entrance plays an important role by influencing shoppers' perception as explained by Bansal and colleagues [11]. Probably, the main aspect influencing the customers is the shop window rather than the entrance [141]. In fact this is the mean through which sellers can take their first step towards customers. In [99], the authors examine the consumers' attention towards sale signs placed outside the shops, and how these factors affect their intention to actually visit it.

All these aspects are useful to attract buyers into the shop, but it is therefore clear that, to be successful, a business has to satisfy the expectations of their customers. Taking this into consideration, the shop's interior has to be attractive and pleasant for them as well. Nowadays, technologies such as multi-touch displays are being installed in different kinds of stores like footwear and clothing stores since they are the brands more linked to new technologies, but it is not strange to also find them in jewellery or accessories ones, in order to make the shop more attractive for customers. Sometimes, these displays can be combined with other devices to offer new interaction experiences. For example we can take the Connected Store Demo [43] by eBay and Rebecca Minkoff where we see an interactive experience in both the store showcase and the fitting room. This room is equipped with a mirror surface overlaid by a user interface. In the showcase setting, the user explores the different items in the store. Once she finds something interesting, she requests to try them in the fitting room and the clerk prepares it with all the favourite items. This approach seems to be really attractive for customers, especially in clothing stores. For this reason we use a similar shop, a footwear store, as case study, working on both the shop's interior and exterior. As noticed in the introduction, a lot of works regarding the use of mobile devices as remote controllers for house appliances [58] or even vehicles [98] have been presented to the community. These works are not only related to industrial applications but they are related to domestic use too. Even if most of the works are focused on controlling home appliances, other works present innovative ways to control public smart objects. Usually, in this scenario a single user controls the system and the actions are visible to other people. An example of this interaction approach is still visible in Cagliari at the THotel where visitors can change the hotel tower's colour by tweeting a message writing the desired one [135].

2.3 Mobile world

Given the importance of mobile devices and their utility as learning tools, we developed four projects that implement well known algorithms that are often used on desktop environments. In these works the main challenge is the constraint of such devices due to their computational power. In “Chromagram” we have implemented an algorithm that allows users to create a chroma-key effect on their smartphone. The chroma key effect allows users to replace an entire area of an image which is identified by a specific colour with another image or video stream.

Due to the fact that we are working on mobile devices, we need an efficient, fast, cheap and easy-to-implement method for the real-time chroma key matting effects computation. We need to minimise the computational cost to face these issues and achieve good visual results too. It is clear that such effects can be achieved by using different techniques, and due to this we studied several keying (and matting) techniques to understand which one could be the best in a mobile setting.

As a result of our study, we can list the most used techniques:

- Two of the most widely used techniques are the **Bayesian** [33] and the **Poisson Matting** [161] that provide great results with random backgrounds and intricate boundaries.

- The **Luma Keying** technique [19] allows editors to create the matte by using brightness information (luminance) of the HLS representation of the image. By setting a threshold value for the luminance, that is independent from the color information of the image, we can create the binary matte for the image. This approach is useful if applied to bright foreground objects on dark background.

- The **Difference Keying** approach [177] provides the matte creating it from the subtraction of two photos, the first one is made up by using the background and foreground objects, while the second one contains only the background. This technique does not require a dedicated photo set, but only needs two photos (or videos) with the same background.

- The **Depth Keying-based** and [32, 63, 85, 86] techniques provides the matte through a special depth camera equipment that defines the depth of the scene objects, making everything that it is placed beyond a chosen depth threshold transparent.

- The **3D Keying** [133, 148] represents the image in a three dimensional space where every axis represents one of the primary colours: red, green, and blue. Then, the algorithm is able to separate the background from the foreground using two shapes.

All these matte techniques, that belong to the image processing field, are now easy to implement on mobile devices as well as other algorithms for face recognition. In fact, in “Click and Share”, we used these techniques in order to allow users to quickly and easily identify faces in pictures, by recognising people, and, thus, sharing pictures with them.

The problem of automatically recognising the faces contained in a photo can be split in three different phases [180]:

- the detection, which selects one or more portions of the photo containing a face;
- the feature extraction;
- the recognition.

The features extracted depend on the particular recognition algorithm. We can recall here the three main algorithms for the latter phase: *Eigenfaces* [166], *Fisherfaces* [14] and *Local Binary Pattern Histograms* [3]. Executing such algorithms on mobiles is becoming feasible, considering their increased computational power. In particular, the interaction opportunities provided by mobile devices in an ubiquitous setting are pushing for having the recognition computed directly on the smartphone. For instance, in [142], the authors discuss an automated and robust face verification system for securing mobile devices.

The authentication is also the focus of FIRME [106], which introduces a modular and flexible architecture for performing face and iris recognition, designed in particular for being executed on a mobile device. Apart from security applications, Vasquez et al. [170] propose a built-in module for face recognition in mobile devices, which works for both photos and videos. In order to demonstrate the possible benefits of the proposed module, they describe a proof-of-concept application that allows to quickly send a photo via email.

Working on mobile, we also faced issues related to the device sensors such as gyroscope, GPS and compass. In fact, in “AR-Garden” we created a prototype for a navigation tool for tourists. This tool makes use of augmented reality techniques in order to present touristic data to visitors.

During the prototype design process, we analysed similar projects that have common characteristics such as the technology used, the devices' specifications and the same user target. Most of them are general purpose augmented reality browsers that show the nearby point of interest giving their position and general information. Through the device camera, the reality augmentation allows to see the position of these point of interests but it does not make it possible to see the path necessary to reach them. This could be an annoying problem due to the fact that in metropolitan cities we have to find the point of interest within several blocks. In our work instead, we use the augmented reality to guide users through the shortest path thus allowing them to reach the point of interest easily.

It is clear that navigation systems may offer reliable information, as said in [118] Mulloni and colleagues say that the road towards this kind of applications is not easy due to a lack of precision of navigation systems. Nowadays, this drawback is a minor problem thanks to new mobile devices' sensors that are faster and more accurate. Their work demonstrates how users that have access to the information provided on a 2D map or via augmented reality, in most cases prefer the use of the 2D map because it gives greater precision and reliability compared to the use of the augmented reality one.

In order to have better results in terms of user experience and usability, augmented reality applications have to be more accurate. These improvements can be achieved by using techniques that are, for example, used in indoor navigation systems. In fact, these systems offer navigation functionalities in indoor environments by using both orientation sensors and phone camera, thus guiding users inside a building without the GPS information. These systems suit very well large buildings such as museum or art exhibition and good examples of such a kind of indoor navigation systems can be found in [151]. Actually, most of the works concerning the indoor navigation are thought for museum environments. Often these applications offer innovative ways for discovering the museum artworks. In [28], Chen et al. present a prototype that uses an augmented reality visualisation by using the phone camera in order to recognise hand gestures thus offering new interaction opportunities. In fact, by using this approach it is possible to observe and manipulate 3D digital replicas of artefacts in real-time without the traditional keyboard and mouse setting.

Similarly, in the "VisitAR" project as well we exploit AR techniques in order to present touristic information. For this reason we deeply analysed this topic figuring out its pros and cons. According to Lee et al. [100] AR is one of the thirteen key technologies used in Ubiquitous city.

Moreover, a consistent way of using augmented reality in Ubiquitous City concerns touristic applications. Augmented reality [9] allows users to feel a different perception of information and reality that surrounds them. In fact, augmented reality allows users to visualise contextual information of the real and virtual world directly from their point of view. The contemporary vision of both real and virtual world allows users to access a large amount of information, normally accessible by consulting different kinds of sources. In recent years augmented reality has seen a remarkable expansion of augmented reality navigators [175]. Gleue and Dähne [56] presented an overview of the necessary hardware for the development of the prototype used in "ARCHEOGUIDE" project. They used the system, as a first trial, at ancient Olympia in Greece. This work is important to understand the problems that early researchers in the field of tourism had to face. Their prototype consisted of an head mounted display, a laptop, a WLAN antenna, a GPS receiver, many cables and a backpack to carry out all this equipment. Nowadays all this hardware is integrated into mobile devices. This explains the main reason for using augmented reality in archaeological sites navigation with the use of these handled devices.

In fact, this augmented reality techniques for touristic purposes are really suitable for smartphone and tablet, especially due to the fact that tourists always have these devices in their pocket. However, it could be adapted to other mobile AR technologies like Optical SeeThrough head mounted display, or integrated with other features using new generation head up displays (HUDs), like the Google Glass.

Fritz et al. [52] described the development process of an interactive visualisation system, based on augmented reality Technologies, and its integration into a tourist application in the PRISMA project. The main idea was to combine touristic information on a binocular device with augmented reality techniques. This integration allows visitors to consult, with the same device, different sources of information normally available using different media.

The use of these technologies allows users to retrieve customised and interactive information about city monuments and historical buildings. Finally they suggest some interesting ideas that will be integrated into the PRISMA project such as seeing the surrounding area in different periods of the year, by combining the augmented reality with a virtual immersive enhancement and allow visitors to explore some places by offering a panoramic view.

Kounavis et al. [93] discuss the use of augmented reality applications for the touristic needs, and describe the evolution of the technology from pilot applications to nowadays mobile applications. First of all they present some examples of applications developed for tourism:

- Tuscany+, [167] the first AR application, developed specifically for the Tuscany region by "Fondazione Sistema Toscana", operates like a digital tourist guide. The information in Tuscany+ is both in Italian and English, and regards accommodation, dining, nightlife, and, of course, sightseeing: they are retrieved from Internet sources, such as the region's official portal, Wikipedia and Google Places.

- Basel [109] is another city provided with an AR tourist guide, as part of the project "Augmented Reality", accessible through the Layar browser. Users can retrieve valuable information on the city of Basel, in particular regarding its sites, museums, restaurants, hotels, events and shopping centres.

The authors then, list the main problems in the implementation of a tourism application. The major obstacle regards the lack of interoperability across mobile platforms. In addition augmented reality applications for tourism require an Internet connection through either Wi-Fi or 3G, but not all cities offer free Wi-Fi connection and 3G data roaming connection has a high cost for foreign visitors. Nevertheless they argue that the employment of augmented reality in tourism could help the tourist in the trip planning and to access interesting information, increasing their knowledge regarding a touristic attraction or a place.

From the presented works, we can conclude that up to now different types of augmented reality touristic applications have been designed. However none of these take into account the tourist experience as a whole, from the moment they arrive in the city and plan the sightseeing, to the moment they do the city tour and enjoy it. For this reason we developed an application that takes into account some of the touristic needs and we provided as much information as possible in a simple and intuitive way.

2.4 Evaluation methods

In this dissertation we used different validation techniques whose knowledge is useful to a full understanding of the experimental results. Due to this consideration, this section provides a brief description of these evaluation methods to the reader.

2.4.1 Demographic questionnaire

The demographic questionnaire is one of the main aspects of any survey. Its questions are designed to help researchers to determine which factors may influence the users' answers and opinions. The procedure of collecting demographic information will enable researchers to compare subgroups in order to see how responses vary between these groups. This questions are useful for collecting information on users' physical activity habits, experience level with the analysed environment and other data that characterises the sample test population such as gender and age.

2.4.2 SMEQ - Subjective Mental Effort Question

This questionnaire has been developed in 1985 by Zijlstra and Van Doorn [182]. Measuring cognitive workload involves assessing how much mental effort users expend whilst using a prototype or a ready to release system. The SMEQ is post-task single-item questionnaire with a rating scale from 0 to 150 with nine verbal labels ranging from *Not at all hard to do* (just above 0) to *Tremendously hard to do* (just above 110). Its aim is to measure the amount of mental effort that people feel they have invested in a specific task. This questionnaire allows researchers to measure the amount of effort users feel they have invested, and not the amount of effort they think the task may have demanded. In the first published version of the SMEQ, participants draw a line through the scale (which is 150 mm in length) to indicate the perceived mental effort of completing a task. In the online version published in 2009 and developed by Sauro and Dumas, participants use a slider control to indicate this value. The authors of the SMEQ claimed that it is a reliable and easy to use questionnaire, and that they placed the verbal labels by calibrating them psychometrically against tasks.

2.4.3 IMMS - Instructional Material Motivational Survey

The IMMS is used both as a pre-test and a post-test tool serving as either a motivational needs assessment prior to instruction or a measure of people's reactions to instructional materials afterward. The IMMS contains 36 questions, that are related to the ARCS model, in relation to the instructional materials they have just used in the user test. The ARCS model identifies four basic strategy factors for motivating the instruction process: *Attention* strategies for arousing and sustaining curiosity and interest of learners, *Relevance* strategies connected to learners' needs, interests and goals, *Confidence* strategies that help students develop a positive expectation for successful achievement and *Satisfaction* strategies that provide extrinsic and intrinsic reinforcement for effort.

The ARCS model provides a useful framework for both the design and improvement of the motivational quality of a range of informational entities and it also increases the chance that these entities will be used and enjoyed.

2.4.4 SUS - Software Usability Scale

The SUS [20], created in 1986 by John Brooke, provides a *quick and dirty*, reliable tool for measuring the usability of a website, a software or in general a developed system. It measures the overall usability of the tool being tested. In fact, SUS allows researchers to evaluate a wide variety of products and services, including hardware, software, mobile devices, websites and applications. It consists in a set of 10 questions providing a five scale answer. These steps vary from *Strongly agree* to *Strongly disagree.*

This method is a very easy scale to administer to users and can be used on small sample sizes with reliable results. Finally, it can effectively differentiate between usable and unusable systems. According to Brooke, participants should complete the SUS after having used the system under evaluation but before any debriefing or other discussion. Instructions to the participants should include asking them to record their immediate response to each item of the questionnaire rather than thinking too much about them.

Regarding the scoring system, for the odd-numbered items worded items the score contribution is the scale position minus 1. For negatively worded items (even numbers), the score contribution is 5 minus the scale position. To get the overall SUS score, we need to multiply the sum of the item score contributions by 2.5. Thus, overall SUS scores range from 0 to 100 in 2.5-point increments. Since the values are on a scale of 0-100, it is easy to interpret them as percentages but they are not. Usually, to better interpreting the results it is necessary to normalise the scores in order to produce a percentile ranking.

2.4.5 PACES - Physical Activity Enjoyment Scale

The original 18-item questionnaire was developed by Kendzierski and DeCarlo [89] for a college-age population, and was intended to be unidimensional, but further testing in other populations revealed problems with its factor structure. Motl and colleagues used a 16-item version, revised for adolescent girls, which has also been modified for use with younger children. An abbreviated 8-item version of the PACES has been used with adults of mixed ages and was found to be invariant across samples of adult runners and cyclists.

The PACES is a standard and promising questionnaire for evaluating the level of users' enjoyment during a certain physical activity. It has differentiated between the experience of exercising in pleasant versus unpleasant conditions and between modes of physical activity selected by participants versus modes selected by an investigator.

Chapter 3

Innovative learning

In the Human Computer Interaction community, researchers work on many projects that investigate the efficacy of new technologies for teaching and learning in both formal and informal settings. Unlike other research fields such as Computer Graphics or Geometry Processing, these researchers must have an approach that is typically multi-disciplinary. This approach involves experts in different fields such as computer science, behavioural science, design and several others depending on the topic.

Photo credits: Lucélia Riberio

Using technology as a learning tool, if done properly, can be an useful resource for young students. But, *are all technologies suitable for learning? Is there a limit to their use?* Even if new technologies can be *the cool thing* for young students, we have to keep in mind that sometimes new technologies are still unreliable for such important tasks.

Even if learning is not only a matter of entertainment or a way to have fun, as explained in [70], humor lightens learning and can enhance student's pleasure and willingness to learn. It is important to find the right balance between the rigorous learning and the learning with entertainment.

If this process does not occur properly, the learners will not be able to recall the acquired information or new skills. As Dror et al. explain in their work [42], the way to reach a good result is clear: The learning must fit the human cognition.

Dror et al. also say that *"The difficult and tricky challenge is how to translate this theoretical and academic research into practical ways to utilise technology so as to enhance learning."*.

We can create several tools and different approaches by changing the traditional learning methodologies. We need to evaluate these approaches, studying their effects in the middle and long term. Learning is not a new behaviour. It has existed for millions of years, it is one of the characteristics of the intelligence of the mankind. At the same time, advancements in technologies are a good chance to improve the way people learn, we need these innovative ways to engage students, to assist them and to improve their preparation in any field, from the language to the mathematical, historical or geographical ones.

The section [3.1] is a revised version of the paper [157].
The section [3.2] is a revised version of the paper [168].
The section [3.3] is a revised version of the paper [26].
The section [3.4] is a revised version of the paper [158].

3.1 Speaky Notes

Studying a foreign language is mandatory in most European countries. In fact, according to a 2012 report from **Eurostat**, pupils begin studying a foreign language as a school subject between the ages of 6 and 9. Even if English is one of the main studied foreign languages, UK and US students spend time and effort to learn a second language too. Despite the fact that English has become a global lingua franca over the past decades, according to recent statistics, only 5.6% of the world's total population speaks English as their primary language, while another 5-6% speaks English as second or third language. Nowadays, there is over 7000 languages and learning at least one foreign language will help you by opening up a world of new opportunities.

For example, it allows learners to have the ability to communicate and truly understand other cultures. This aspect is more important if we talk about the modern dependence between nations. Each nation needs competent persons able to bridge the gap between these different cultures also promoting world peace, international diplomacy and trade engagement. Learning a second language can also improve the first language itself. In fact, research shows that knowledge of other languages improves the overall linguistic abilities, enabling students to use their native language more effectively.

At the same time, learning a language involves a wide range of learning skills in general, thus enhancing one's ability to learn in different other areas. These studies show that children who have learned foreign languages in their childhood prove greater cognitive development in mental flexibility, creativity, problem-solving and reasoning in general.

Again, other pros are the appreciation and the enjoyment of several life-related aspects. Think about the masterpieces of the world's literature. Even if one day we will have a translation for all the spoken languages in the world of a specific work, we will never be able to fully appreciate its characteristics such as the metaphors and the word plays as well as its style and beauty. The same reasoning can be applied, *mutatis mutandis*, to theater, music and film.

Another activity strictly connected to speaking other languages is traveling. In most cases, English may be sufficient to explore new countries and cities. If tourists want to explore the *real* country they must know the official language or at least its basics.

In recent years, mobile devices have become very popular within young people. Thanks to developments in mobile technologies, these devices can now do much more than just voice calls and texts. We envision mobile devices as tools for improving the young users' lifestyle, especially for learning.

In this work we present a web authoring system that makes it possible to create a mobile application that supports children in learning a new language in a more pleasant and entertaining way by using Augmented Reality. This application allows pupils to improve their speaking skills by turning the language acquisition into a game under the supervision of both teachers and parents. Our contribution is focused on understanding how digital technology can facilitate learning while keeping in mind that it is a wide and interdisciplinary issue. Two of the most common questions among teachers of primary schools are: *How can I engage students? How can we get kids to do their homework?*

Research [44] has proven that parental participation in the education process plays a major role in their scholastic capabilities and general development. Based on this observation, it is important to use an approach that encourages parents to help their children in doing their homework. We think that new technologies, especially mobile devices, are the mean through which it is possible to improve the students' engagement and their learning process. Speaky Notes is part of a bigger project named LEAF that is a web system that allows users to create their own personal dictionaries (called e-dictionary). Each dataset record thus represents a single term object and contains the word syntax, its translation (in one or more languages), its phonetic transcription, its audio pronunciation and a representative image. Every dataset record is validated by linguistic experts and, thanks to its API, it is possible to use this collection in a fast and easy way. In fact, Speaky Notes is focused on children learning using new technologies, and its goal is to give teachers and parents a useful tool able to turn lessons and homework into more pleasant and fun activities.

Regarding these new technologies (e.g. smartphones and tablets) and their owners' age range, we can affirm that year after year, it is common to see kids using their parents' smartphones or even their own personal mobile devices. In the same way children learn to speak in the first years of their lives, they also learn to interact with these mobile devices while still in their childhood. At the beginning, they learn how to unlock the phone, then they tap randomly over the screen surface, and finally they start to explore the device features. It is impressive how fast children learn to use this kind of devices, and after a while they actually understand the correlation between icons and applications.

Due to this, it is clear that in primary school we will see skilled smartphone users that use a tablet to improve their learning. Having such skilled users makes it possible to use these devices for learning. In our case study, we focused on language learning, and, more precisely, our tool is oriented towards the acquisition of language vocabulary and pronunciation.

It is worth noting that children have a critical or sensitive period for language acquisition that lasts until puberty, and during these years, certain parts of the brain are more developed than others. These brain areas are responsible for the unconscious learning such as riding a bike or dancing and this happens for the language learning as well [7]. Thanks to this we can use this "golden period" to learn more than a single language, stimulating children with new approaches involving mobile devices and augmented reality.

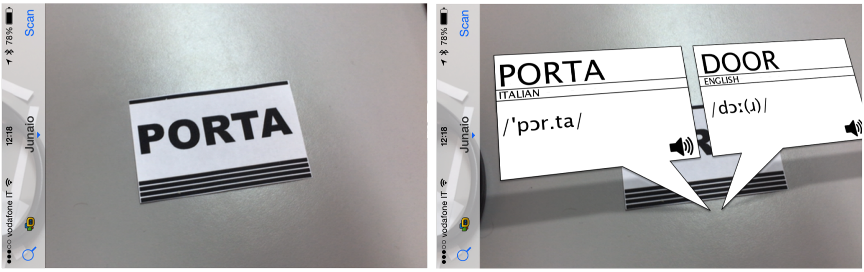


Figure 3.1: Speaky Notes as it appears on an iPhone 6 using the marker 'Porta' (that means Door in english) before (left) and after (right) the image recognition.

In detail, the presented system allows teachers and parents to easily create an application for mobile devices that uses AR. This tool becomes really useful in children's hands because they learn through playing. The main reason that motivates our case study is explained in [21], in fact, if we consider the English native speakers and second language speakers who use this language for their day-to-day needs, it is possible to cover around the 15% of the world's population.

Finally, this work tries to contribute to the improvement of children education pursuing the following objectives:

- Promote educational innovation by providing educators with a novel tool for teaching;
- Allow a self-regulated learning at home with the participation of parents;

- Analyse the possibilities that augmented reality techniques can have on children education.

Setup discussion

As we explained earlier, educators have a set of tracking cards and they have to place them over real objects, avoiding those that are placed in unreachable or dangerous areas. When all cards are well positioned the mobile device can be given to the child/pupil.

After that, the exploration game can start. Firstly, educators give the mobile device to children and they can assist them by giving some information or suggestions. For example they can ask them to find a specific item or by giving directions necessary to reach the next item (go right, go left and such).

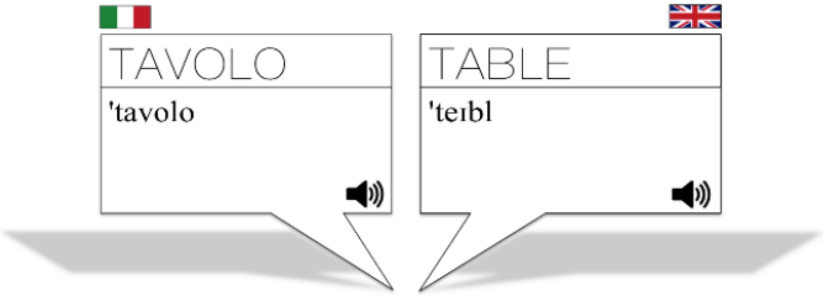


Figure 3.2: Augmented reality content that shows the same word (and phonetic transcription) in different languages and allows to listen to an audio file tapping over.

During the game, educators can even ask them to repeat the word or to say the correct spelling. This game could be conducted in a collaborative way, giving a single device to a group of pupils, thus fostering social interaction between them under the supervision of the educator. All the elements, including the tracking information, their behaviour and the interaction functionalities, are generated by our system.

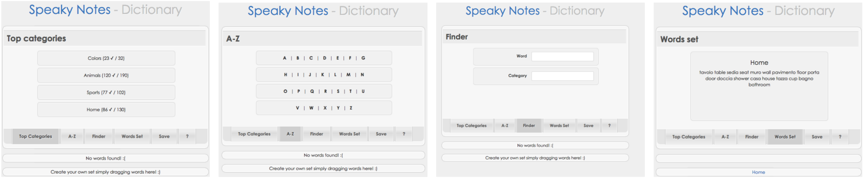
In addition, we can analyse a possible scenario of our approach: Bob is a primary school educator that teaches English in an Italian school. He usually trains his pupils with new technologies like tablets and touchable surfaces in general. He is preparing the next lesson that is about new English terms.

These words are connected to the house environment such as door, window, table, chair and more. He uses the Speaky Notes web system by choosing among the available categories thus creating his own set of words. Then he prints the PDF file and cuts the markers out. The day after Bob places them in the classroom explaining the game to his pupils. The school where Bob works has five mobile devices for each class, therefore he splits his pupils in five groups. He asks to the first group (Alessia, Carlo, Sara, Luca) to look for the *item with a flat top and one or more legs, providing a level surface for eating, writing, or working*, immediately the first group's pupils see the table and go towards it. Then, they frame the card with the smartphone and the relative information is shown over the card. Pupils can now interact with the device by tapping over the augmented content listening to the pronunciation of the word in both Italian and English. Bob asks Alessia and Paolo to slowly repeat the word in English, then he asks Luca and Carlo to spell the word. Finally he says a sentence to Sara along the lines of *The table is red* asking her to translate it in Italian. The game ends when all groups have framed all the placed cards.

It is clear that for some people learning a language requires more dedication than many other subjects even though it is not really difficult per se, but it certainly is time-consuming. Speaking a language fluently is made up of three main aspects: know the grammar rules; use a wide vocabulary; know the right pronunciation avoiding misspellings. Surely, there are a lot of correlated aspects that we have not mentioned, but in general these three ones are the most important.

We offer teachers and parents a tool for their children, making them able to improve their vocabulary and pronunciation of a new language, leaving the acquisition of the grammar rules to standard methodologies. Our approach does not aim to replace traditional learning methodologies, but would be a complementary tool for teachers to use during the training stage, oriented to the language vocabulary and its pronunciation. Children can easily interact with these devices, in fact they already read e-books on tablets and they do homework directly on them through some sort of web-based platforms.

This work is part of the LEAF project, which is a web system able to create personal dictionaries (called e-dictionary). For each word in the collection we can find its syntax, its phonetic transcription, its pronunciation (as an mp3 audio file) and a representative image. The system allows a categorisation of the terms; we can, for instance, have the home category that contains words like table, door and window.



No words found! :(

Create your own set simply dragging words here! :)

Home

tavolo table sedia seat muro wall pavimento floor porta door doccia shower casa house tazza cup bagno bathroom

Figure 3.3: Speaky Notes web interface. From left to right: top categories, first word letter search, term search, e-dictionary summary.

The entire system supports any language pairing, but in our case study we implemented the version for Italian students willing to learn English and also for the English students who want to learn Italian. In general we call the content relative to the native language *FL* (First Language), and *SL* (Second Language) the other one.

Our approach combines real world and virtual reality by using augmented reality. To achieve this we envision a card set that represents the words the educators want to teach in the lesson. Both LEAF and Speaky Notes are web-based software. We envision them as freely available systems where teachers and parents can create their own augmented reality experiences thus sharing their e-dictionaries. These experiences depend on which words they want to teach. Usually, children learn vocabulary terms day by day, and often by categories. Some examples could be categories like colours (red, blue, black), house items (table, chair, door, window) or school supplies (pen, pencil, marker, eraser). Teachers, browsing among the available categories in the system can select which of them are suitable for the lesson, and, if necessary, they can create a set of words combining more categories by simply removing or adding single terms.

The search of a word can be accomplished by using the category browser, by using the word's first letter or even by using a common search input box as shown in Fig. 3.3. Then, teachers can drag items into a specific box, building their own card set. After this step, the personal e-dictionary is used to automatically generate the Speaky Notes mobile application content just by clicking on the *Save* button. This process is fairly immediate and does not require any user interaction. Framing just one QR Code (Quick Response Code Fig. 3.4), the generated augmented reality application can be downloaded at a specific link through the Junaio (www.junaio.com) channel servers. In our prototype we used this approach and, thanks to the specifications available on the Junaio framework, it has been possible to programmatically create the entire channel starting from a root project that includes the required AREL (Augmented Reality Experience Language) libraries and some base content.

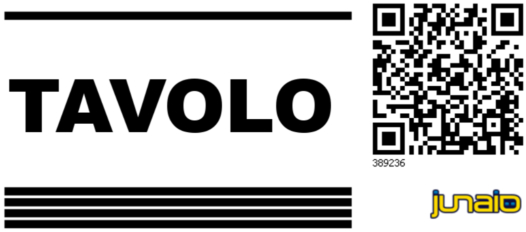


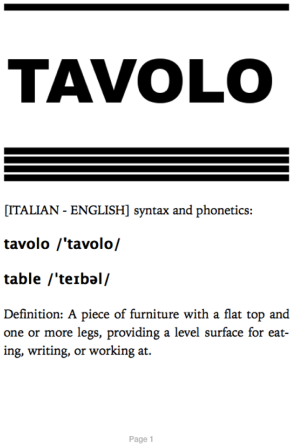
Figure 3.4: Card layout and QR Code needed to identify an AR channel.

Junaio is an AR browser with an augmented reality client application that runs on smartphones and tablets. Developers can create their standalone augmented experiences (called channels) allowing them to have a platform independent application.

After the content creation, the PDF containing the Speaky Notes card is given to teachers, as well as the QR Code. This document contains all the words saved in the e-dictionary. At this point the users have to print out and cut out the cards, then they can start the Speaky Notes application by framing the QR Code.

The card/marker layout is showed in Fig. 3.4. We envision each card as a piece of heavy paper similar in dimensions to those used for common card games (60mm x 90mm). Once the QR Code shown in Fig. 3.4 has been framed, and the application launched, the augmented reality experience can start. All the Speaky Notes cards are readable by smartphones too, since we use them as tracking images for the augmented reality experience. This way children can interact with the Speaky Notes just pointing at the cards with the mobile device and, thanks to augmented reality techniques, the word's information are over-imposed on the screen.

In fact, when the system generates the cards and the PDF file, it also retrieves the necessary content for the augmented reality experience through the LEAF API. This content consists in the word syntax, its phonetic transcription and an mp3 audio files that contains the pronunciation in both languages (FL and SL). In detail, the system creates all the virtual content that will be shown over the tracking marker such as syntax, phonetic transcription and language name as illustrated in Fig. 3.2. Pupils can tap on the virtual objects and listen to the audio file that reproduces the right pronunciation.



TAVOLO

[ITALIAN - ENGLISH] syntax and phonetics:

tavolo /'tavolo/

table /'teɪbəl/

Definition: A piece of furniture with a flat top and one or more legs, providing a level surface for eating, writing, or working at.

Page 1

Figure 3.5: Sample page of the generated eBook. In this case we can read the word in the native language (also used as marker for AR), in the second language, its phonetic transcription and definition.

Having fun with the eBooks

We have discussed how to use the AR application under the supervision of an educator at school, but we think that a lot of work can be done at home. For this reason our system generates a summary eBook that contains all the words chosen by the author/educator and their relative information.

Using an eBook reader (or a similar device) students can improve their study at home, rehearsing the words learnt at school. Each eBook page contains all the information about a single word and its marker image. The image lets students and parents to replicate the augmented game at home, transforming the student's house as a learning environment.

Conclusion

The main objective of Speaky Notes is to give educators a system that uses a dataset of validated words that allows the creation of personal dictionaries, easy to share among colleagues. Furthermore, we envision this system as a tool to engage students during the learning activity. Even if we use as case study the language acquisition process, the whole system could also be applied to other learning fields. The next step for the Speaky Notes tool is user testing. We are planning a first test with a school in Monza under the supervision of an expert in pedagogy in order to get feedback and suggestions.

At this moment the application is ready to use and the beta version will be available as open source. Advancements in technologies are a good chance to improve the way children learn, we need these innovative ways to engage students, to assist them and to improve their preparation in any field, from the language to the mathematical, historical or geographical ones.

In the education path, speaking a second language properly cannot be a secondary goal. Children have to be motivated and enthusiastic in language learning and we see our approach as a good way to give them the tool to reach these feelings.

3.2 SoSphere

The wide availability of low-cost sensing devices is opening the possibility to easily create different interaction settings, which exploit various techniques for a more natural interaction, especially in public and shared settings. In this section, we compared two different solutions for enhancing the interaction experience of a planetarium application, both replicable at a reasonable cost. The first version is based on a simple multi-touch paradigm, while the second one exploits a full-body interaction together with a projection on a geodetic sphere. We detail the technical implementation of both versions and, in addition, we discuss the results of a user-study that compared the two modalities, which highlights a tradeoff between the control and the users' involvement in the virtual environment.

In this work, we exploit different low-cost sensing devices in order to provide engaging interaction experiences, trying to put into practice the ubiquitous computing vision, where the user is expected to interact naturally with the technology without even realising its mediation.

Nowadays, the availability and the decreasing cost of such sensing devices allows for the creation of interactive spaces, especially in public and shared settings, even with limited resources. We describe how we created a more immersive and engaging version of an existing virtual planetarium software for desktop systems, transforming it into both a multi-touch and a gestural oriented application. In particular, the Kinect version exploits also a geodetic projection on a hemisphere, enhancing the realism of sky visualisation. Both settings are easily replicable for different applications and settings.

We report on a small-scale comparative evaluation of the two settings, which shows that, without an appreciable difference in the overall usability between them, the multi-touch version gives the users more confidence in controlling the application, while the Kinect version is reported closer to a real experience.

Setup discussion

In this section, we discuss the two different interaction settings we created for an interaction-enhanced version of a planetarium virtual environment. We used "Stellarium" [73], which is able to show a 3D view of stars and planets according to the current time and the user position in the world. The application lets the user browse the sky by using the mouse and the keyboard for moving the current point of view or selecting planets and stars.

The basic version of the application is controlled by using a mouse: moving the position of the pointer moves the current view, the left button can be pressed for rotating or panning the view, while the right one is used for scaling. However, this standard scheme of scene exploration can be easily enhanced for offering a more engaging environment. In order to support a more natural interactive scheme, we have designed an interactive model where users perform their tasks in a three dimensional space or touching directly the sky representation.

In both versions, we exploited the same technical solution for implementing the communication between the two different gesture recognition sensors and the existing Stellarium application, which is based on the TUIO [84] network protocol. The scheme is shown in Figure 3.6: on the one hand, we created an intermediate layer between the devices and the application that generates TUIO events according to the current state of the device tracking; on the other hand, we extended the Stellarium code in order to be able to translate such events into application commands, like rotating or panning the view or moving the current camera position.

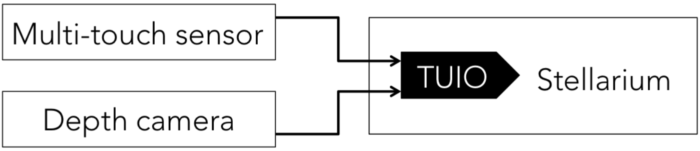


Figure 3.6: Connections among Stellarium and the gesture sensing devices.

Multi-touch version

We implemented the multi-touch version of the application exploiting the improved FTIR [66] table proposed in [74]. The resulting interaction setting is shown in Figure 3.7: the sky visualisation is projected onto the multi-touch screen. The user controls a rotation hinged on the barycentre of the scene by touching the scene and moving the finger in different directions. In addition, it is possible to resize the scene by touching the surface with two fingers: moving them apart from each other enlarges the scene, while moving the finger towards each other shrinks the scene proportionally.

With respect to the desktop version, the multi-touch interface has the obvious advantage of allowing the user to interact directly with the sky projection, without mapping the mouse movement into the scene. In this work, we consider such setting as the baseline for a low-cost solution for creating an application deployment that can be exploited in a shared public setting.

Figure 3.7: The multi-touch planetarium application.

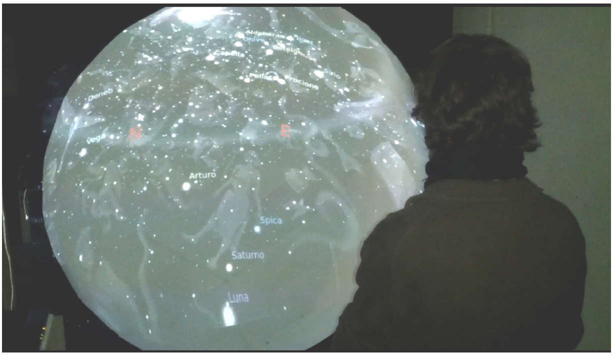


Figure 3.8: Full-body version of the planetarium application projected on a geodetic sphere.

Full-body version

We enhanced the virtual planetarium experience with a different environment, which mimics the visualisation of the sky ceiling through the projection of the sky map on a hemispherical surface.

In order to keep the cost of the implementation low, we have built the entire surface using paper and glue. Following the scheme in Figure 3.9 we built a hemispheric mesh of kraft paper triangles. After that, using thin wood planks we built the shell holding the surface. The shell length and height are equal to the diameter of the hemisphere, which is about two meters.

After securing the hemisphere to the shell, we painted the surface using a common white wall paint and finally we placed a black cover between the squared shell and the hemisphere borders which has a circular section. Once we completed the construction of the hemisphere, we used a short throw ratio projector for displaying the sky map on the curve surface. In this case, the projection exploits an orthographic filter that is provided with Stellarium. The resulting setting is shown in Figure 3.8.

The interaction with this kind of screen is based on a free-hand paradigm, using a Microsoft Kinect sensor. Such gestural interaction allows the user to navigate the sky map. In order to mitigate the Midas touch problem [174], we exploited a grab gesture that has to be performed by the user in order to engage the interaction with the planetarium application. Therefore, in order to start the interaction with the application, the user has to close at least one hand, which resembles the act of grasping a real object.

We initialise the hands position detection with the skeleton tracking algorithm provided by the NITE framework [137], which also detects the hands in the depth image. Then, we incrementally track such positions in the subsequent images, recognising hand open/closed shapes by estimating the local surface areas in the depth images (represented in red in Figure 3.10).

The hand region is compared with its convex-hull area: if the hand is closed, the two silhouettes are nearly coincident, and such ratio is close to 1; otherwise, when the hand is opened, the two silhouettes will consistently differ. In this way we are able to tell when the user is moving for interacting with the application.

The user can interact with the sky visualisation by rotating the scene around its barycentre and enlarging or reducing it in order to get the desired level of detail.

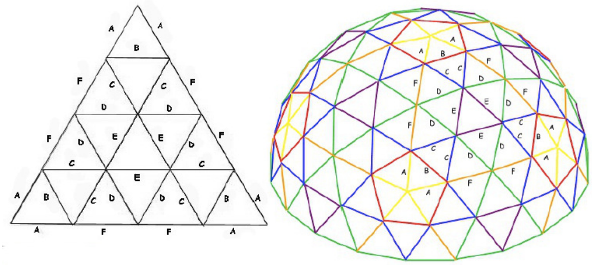


Figure 3.9: Geodetic sphere scheme.

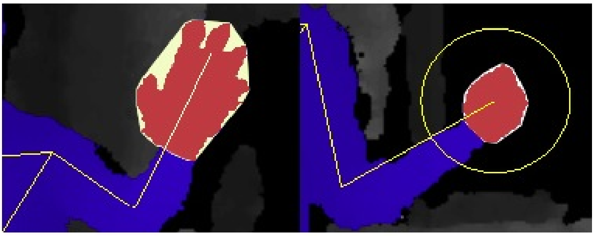


Figure 3.10: Hand open/closed recognition.

We support the interaction through two gestures:

- rotation is supported by simply performing an on-air grab with either the right or the left hand. Keeping the hand closed the user changes the hand position, and the view is rotated accordingly;
- the zoom is supported through a two hand gesture, closing both hands. The user can either enlarge or reduce the view size by respectively moving the hands apart or moving them towards each other.

After building the projection screen and implementing the interaction gestures, we had our interaction scenario ready: the user was able to see the projection of the planetarium image on a spherical screen and to interact with the application by simply standing in front of the hemisphere and performing free-hand gestures.

Such setting is engaging not only for a single user, but also for a group of people: one of them controls the visualisation while the others look at the projected sky map.

Evaluation

We conducted a small-scale user test in order to compare the two interfaces for the interactive planetarium, which gave us some hints on the differences encountered by users in the different platforms (namely multi-touch and Kinect).

The aim of the proposed test is threefold. First, we wanted to evaluate the overall perceived difficulty while executing different tasks with the planetarium interface in both settings. Second, we wanted to determine whether the overall usability of the application was affected by the change of platform. Finally, we investigated if there were differences in the factors that affect the user's presence perception, according to [102].

The test was organised as follows: after completing a small demographic questionnaire, the users had to complete the following tasks by using both versions of the application:

1. Starting from a visualisation where Saturn occupies the whole screen, the user had to go back to a point of view where it is possible to see the Earth.
2. The user had to complete a 360° horizontal rotation of the view (on the Z axis).
3. The user had to find Jupiter and visualise the name of its satellites.
4. The user had to change point of view in order to see at the center of the screen one constellations (selected by the user, but declared at the beginning of the task)
5. Starting from the a view of Saturn with a minimum zoom factor, the user had to enlarge it until it occupied the whole available space.

We alternated the starting version in order to minimise the carry-over effect. After the completion of each task, the users answered the Subject Mental Effort Question (SMEQ) [183] in order to evaluate the perceived difficulty. After completing the whole task set, the user had to fill two different questionnaires: the first one is the Software Usability Scale (SUS) [20] that evaluates the overall usability of the application, while the second one is the Presence Questionnaire [102], which measures different aspects of the user's presence perception. After completing both questionnaires, the user repeated the experiment with the other version (multi-touch if starting with the Kinect and vice-versa).

Thirteen users participated to the test, 9 males and 4 females, aged between 21 and 26 ($\bar{x} = 23.3, s = 1.8$), 5 had a high school, 5 a bachelor and 3 a master degree. The users were more proficient with multi-touch applications ($\bar{x} = 5.54, s = 1.6$ in a 1-7 Likert scale), if compared with the Kinect one ($\bar{x} = 4.3, s = 2.1$).

For the post-task evaluation, we report in Table 3.1 the upper bound ($\rho = 0.05$) of the perceived user effort.

According to [20], the perceived effort for all tasks is between 11 (minimum) and 25 (maximum), labelled respectively *“Not very hard to do”* and *“A bit hard to do”*. It is possible to notice that the multi-touch version required less effort for T1 and T2, while the Kinect version performed better for T4 and T5.

Task	**Multi-touch**	**Kinect**
T1	8.25	15.50
T2	8.67	13.80
T3	7.40	10.00
T4	18.76	11.96
T5	17.21	10.00

Table 3.1: Perceived task difficulty upper bounds ($\rho = 0.05$) .

The SUS post-study questionnaire did not reveal any difference in the overall usability of the two versions. The score was $\bar{x} = 74.04, s = 11.67$ for multi-touch and $\bar{x} = 70.97, \sigma = 11.57$ for the Kinect version. Therefore we can conclude that the perceived usability of the two versions is about the same.

For the presence post-study questionnaire, we disaggregated the scores of the different answers (1-7 Likert scale) according to the following factors [102]: Control Factors (CF), Sensitivity Factors (SF), Distraction Factors (DF) and Reality Factors (RF). Obviously, given the small number of participants, it is not possible to generalise the quantitative results. However, we want to point out here a qualitative tendency that explains the different perception of the effort for the different tasks.

In Table 3.2 we report the questionnaire results. The multi-touch version performed slightly better for the CF and the DF, while it was slightly worst for SF and RF. This means that the users had more difficulties with the Kinect version when a fine-grained control of the planetarium positioning was required (T1 and T2). However, the users were more involved from a sensory point of view, and they found more real the Kinect experience. Indeed, the more exploratory tasks had a higher rating with the Kinect version (T4 and T5).

From these results we can conclude that, given a comparable overall usability and cost of the two settings, it is better to select the multi-touch environment for a more fine-grained control, while if we want to increase the sensory and realism perception for the user (according to definitions in [102]), it is better to select the full-body version.

Factor	**Multi-touch**	**Kinect**
CF	$\bar{x} = 4.96, s = 0.82$	$\bar{x} = 4.77, s = 0.22$
DF	$\bar{x} = 5.19, s = 0.49$	$\bar{x} = 4.92, s = 2.12$
SF	$\bar{x} = 5.15, s = 1.2$	$\bar{x} = 5.5, s = 0.22$
RF	$\bar{x} = 3.62, s = 1.85$	$\bar{x} = 4.04, s = 1.63$

Table 3.2: Disaggregated results of the post-study presence questionnaire.

Conclusion

In this section, we described the creation of a low-cost setting for an immersive virtual planetarium experience. Starting from an existing software for desktop platforms, we created both a multi-touch and Kinect version of the application. The Kinect versions employs a geodetic display that provides the user with a more accurate representation of the sky. We performed a small-scale user study in order to investigate the perceived difficulty in performing different tasks with the two settings, which was low for both versions. In addition, the post-test questionnaires did not highlight any significant difference in the overall usability between the multi-touch and the full-body interaction. However, we found a difference in the perceived control of the application (which was higher in the multi-touch version) and in the perceived realism of the experience (which was higher in the Kinect version).

3.3 RIFTart

In this section, we describe the setup of RIFTart, a virtual reality environment for teaching Art History. It is thought for creating teaching materials to compare two or more artworks, putting them in the same virtual room. The teacher prepares different multimedia contents (e.g., audio or text descriptions), for highlighting and describing different aspects of the artwork. The application, implemented completely by using Web technologies, can be visualised on large screens and head mounted displays. The user test results advance the understanding of the VR effects on classroom usage. We demonstrate that VR increases the motivation of high-school students towards studying Art History and we provide an in-depth analysis of the factors that contribute to this result. This material may be used both during the lesson and also for individual study. The students can explore the artwork by using both wide shared displays and HMDs, thus replicating the visit in a (virtual) museum room.

In the following sections we describe both the RIFTart supported features and the user experience provided and the implementation's technical details, which are useful for researchers and practitioners that would like to create similar experiences through web-based solutions.

In addition we report on a user test, which provides insights on the adoption of immersive VR as teaching material. We measured the students' motivation in learning a particular Art History topic through the Instructional Material Motivation Survey instrument (IMMS) [88], comparing the immersive VR against a projected shared display (which is the current standard in Italian classrooms) on two different groups of high school students. The results show that the VR setting increases their motivation. We analyse in detail the factors that ground such difference.

Setup discussion

The exhibition curators accurately select the artwork position inside a museum, in order to ease the interpretation of sculptures and paintings following a reasonable path inside a specific period of time or the life of an artist. Similarly, teachers try to follow a logical path in their explanations, in order to highlight the main characteristics of genre, similarities and differences in execution techniques etc. However, even if comparing two sculptures in a specific place is easy, for example the Canova's work "Amor and Psyche" and the "Venus de Milo" at the Louvre museum, it is hard to do the same thing comparing artworks located in different places, especially for those artists that had a great impact and worked in different cities and countries.

Figure 3.11: RIFTart interface.

For instance, consider Michelangelo's "David" and the "Moses": the two masterpieces are located in different cities (the former in Florence and the latter in Rome), and they have very different sizes (410 cm vs 235 cm). The obvious solution to such physical problems is using photos of the whole sculpture and details for showing the students the artwork characteristics.

Considering the state of 3D scanning techniques and the advances in the simplification and manipulation of such large datasets, our idea is to complement such teaching material with a VR environment, where the teacher can position two or more 3D models of a sculpture in a virtual museum room. Such material would be available during the lesson, in order to support the teacher while explaining the concepts. In addition, it can be provided to students for autonomous study. This is the main idea of the prototype tool we discuss here, named RIFTart: empowering teachers with the possibility to create VR environments as teaching material. The possibility to explore the environment through the Oculus Rift HMD has a positive impact on the student's motivation. In the next two sections, we discuss first the interaction supported by the prototype and then its technical implementation.

Interface

Figure 3.11 shows the main RIFTart interface: the user sees a virtual museum room where it is possible to explore two sculptures. In our example, we have two Michelangelo's masterpieces: the "David" (1501-1504) and the "Youthful Captive" (1530-1534).

The 3D models were provided by the Stanford Digital Michelangelo project [103]. The user can freely move inside the room by changing the position of the camera in all directions, enabling her to admire every sculpture detail. Obviously, the tool supports different levels of scene configuration. First of all, it is possible to load different models, which can be selected according to the lesson topic. In addition, the teacher can control their position inside the room and their orientation. Moreover, it is possible to set a scale factor for the different models, in order to allow the comparison of sculptures with a relevant difference in size. Such feature breaks the environment fidelity with respect to the real counterparts, creating experiences that can create misconceptions in students. Therefore, teachers should use such feature carefully.

The virtual museum room was designed in order to minimize its interference with the inner content. It consists of four plain walls, a single wooden floor and a plain ceiling, whose dimensions depend on the size of the artwork models. It is possible to configure the position of several spotlights directing them towards the sculpture model. The tool proposes a default light configuration that can be modified by teachers, in order to enhance the virtual visit experience or for highlighting relevant details of the artwork.

Once the teacher has configured the environment, RIFTart allows users to visit the virtual room, moving the user's viewpoint with the standard keyboard and mouse coordinated control for first person video games.

The environment allows associating multimedia annotations to the whole sculpture or to parts of it. In this way, the teacher augments the artwork visualisation through a different contents, useful for e.g. individual lesson rehearsal. As explained in [101], a 3D-annotation is an annotation whose target is a 3D object and which is contextualised with regard to the object itself. In our application, notes may be either audio or textual. Textual notes can be easily read from different points of view, since their orientation is based on the user's point of view, as shown in figure 3.12. In the same way it is possible to activate audio notes, that play pre-recorded audio descriptions, focusing the user's attention on a specific detail with a cone-shaped pin. We do not include videos directly in the visualisation, but they may be linked through text annotations.

The annotations are associated to keyboard buttons, the list of associations is available pressing the L button. When the user activates an annotation, the tool provided automatically moves her position in the scene in order to visualise the artwork detail.

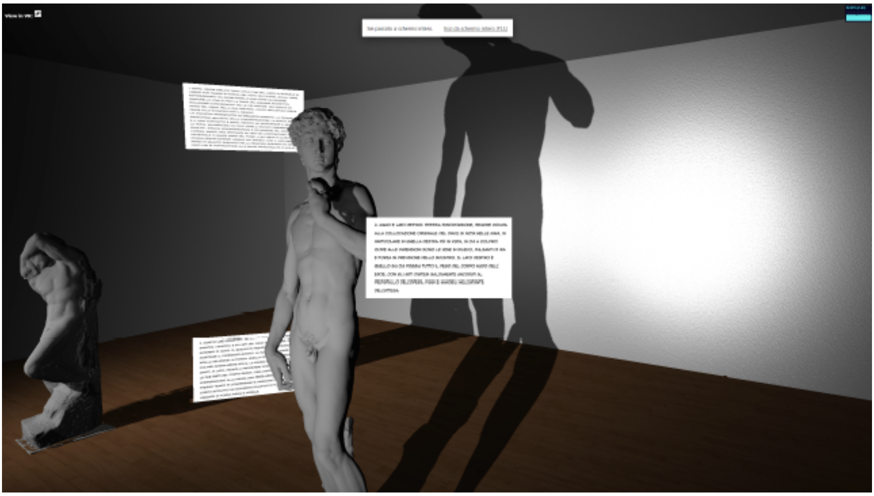


Figure 3.12: Textual annotations on artworks.

RIFTart allows users to explore the virtual environment through two different types of displays. The first one provides a monocular view on the 3D scene, suitable for normal displays. Such visualisation is useful for the currently adopted technology setting in Italian classrooms, which are currently provided with an interactive multimedia whiteboard (LIM): a wide projected screen where it is possible to interact through touch gestures or drawing on the surface through special pens. With this kind of display configuration, RIFTart supports the teacher in showing the students the artworks and commenting on particular details. The viewpoint is controlled by the teacher through a remote, a keyboard or through multitouch gestures. The teacher may use audio or video annotations if needed.

The second visualisation option allows to explore the artworks with a consumer VR HMD, such as the Oculus Rift or Google Cardboard. In this case, the tool provides a stereoscopic view on the environment, increasing the sense of immersion in the virtual environment and the depth perception. In this case, the user has two points of view on the scene, one for the left eye and one for the right eye, as shown in figure 3.14. The user controls only the position in the room through the keyboard, while she freely moves and rotates her head in order to change the looking direction, exploiting the inertial sensors in the HMD.

As we better detail in the following section, the tool implementation is web-based. Therefore, it is possible to support a scenario where each student wears an HMD and the teacher coordinates the lesson content.

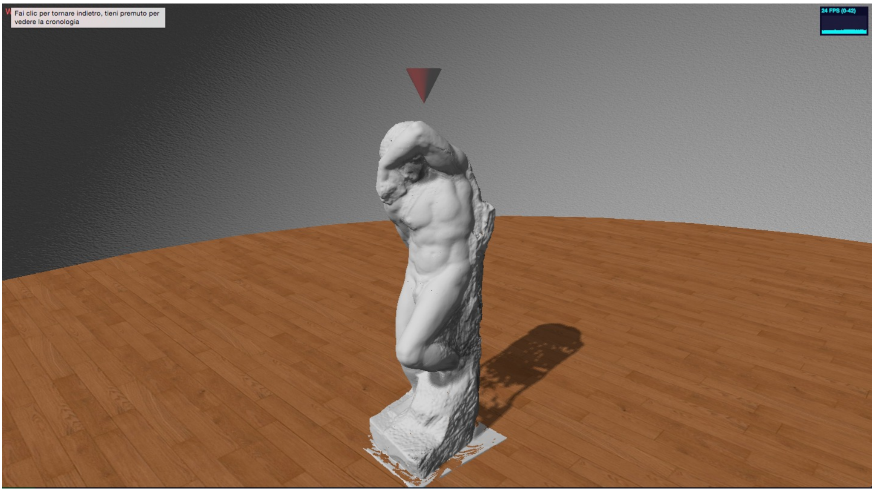


Figure 3.13: Audio annotations on artworks. A pointer focuses the user's attention on the part described by the audio recording.

However, considering the current technology status and cost, such configuration is not realistic even if technically possible. The HMD visualisation may be currently used for autonomous study or lesson rehearsal. However, considering that it is already possible to use high-end smartphones as VR viewers, we suppose that this scenario will be realistic in the near future, either using school equipment (e.g. setting-up a laboratory) or directly exploiting student's personal devices.

Implementation

Considering that the application provides teaching material to students, it is quite obvious that it should be accessible from different devices and in different places (e.g. at school, at home, etc.). As happens for other multimedia contents, web-based implementations provide the flexibility for supporting different users and devices. In fact, if offers the possibility to include the 3D visualisation inside other web contents, such as e.g. the school website information, or a museum description. In this way, it would be possible to reuse the environment not only for teaching purposes, but also for providing general or additional information on artworks.

It also offers the opportunity of updating contents without reinstalling any application, which is important for supporting teachers in creating their own contents (e.g. through an authoring environment), with a quick cycle.

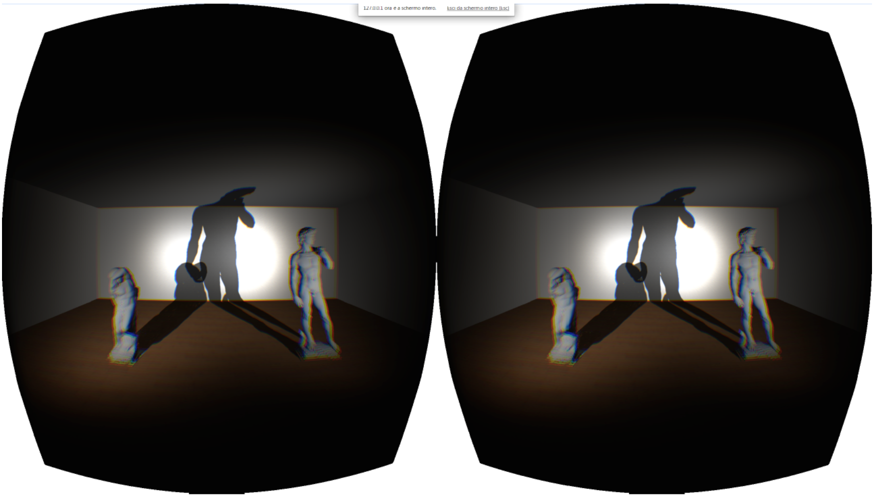


Figure 3.14: Stereoscopic view on artworks.

Therefore, we chose to implement a web application that uses WebGL [95] (Web Graphics Library) for creating the virtual environment. WebGL is a low level JavaScript API for rendering interactive both 2D and 3D computer graphics, based on OpenGL ES 2.0 [94].

In order to avoid using low-level drawing functions, we created the scene using the `Three.js` [78] library, which provides loaders for different 3D model formats, geometries, materials, lights, cameras etc. Using WebGL and Three.js guarantees the compatibility with the most important browsers (Chrome, Firefox, Opera, Internet Explorer, Safari) in both their desktop and mobile versions. We support the single camera display (the LIM mode) with a perspective camera, with a 45° field of view. In order to move inside the world, the user changes the position of the camera or the sight direction with the keyboard and mouse. The support for the HMD requires two components. The first is a renderer that creates the image for each eye. A VR effect decorator wraps the usual 3D scene renderer object, and it is provided by Three.js. Starting from the current position of a normal perspective camera, the decorator shifts its position to the left and to the right for simulating a separate camera for each eye, and then it renders the corresponding images splitting the screen as shown in figure 3.14.

The second component is required in order to change the camera orientation in the scene according to the user's head movements.

In order to do this, the browser must read the HMD sensor data, whose access is not currently provided by any desktop browser out of the box. In order to use the Oculus Rift with a desktop browser there are two alternatives: the first is a special build of the Chromium browser (named Chromium VR) and the second is the development Firefox nightly build. Considering mobile devices, Chrome for Android supports natively the Google Cardboard.

In order to enable the VR visualisation, the application checks that the client is a VR compatible browser, and, if it is not, the user can only run the monoscopic version. If the browser supports VR, it checks if there is an available HMD on desktop computer or if the device is compatible with Google Cardboard on mobile. If so, the stereoscopic visualisation can be simply activated by pressing a button, which is otherwise invisible.

Even if the web-based visualisation guarantees the flexibility in accessing the tool through different devices and operating systems, the differences in the hardware configuration have a high impact on the user's experience. The 3D rendering is obviously more resource consuming with respect to regular web pages and, especially when RIFTart is used with HMDs, the tool should update the scene with a high framerate. However, the number of the 3D models in the scene, their resolution, the number and the type light type may degrade the rendering performance especially on mobile devices.

Therefore, the tool is able to render the scene in two modes:

- In the *High Fidelity Mode*, RIFTart configures the scene using more realistic but computationally expensive elements, such as higher resolution models, more than one light source and so on.
- In the *Fast Rendering Mode*, RIFTart includes the models version with the lowest resolution (if provided), only one directional light following the camera direction and a spherical version of the museum room for avoiding unpleasant light reflections.

Figure 3.11 shows the rendering in the high fidelity mode, while figure 3.15 shows the same scene rendered in fast mode.

Evaluation

In order to assess the prototype effectiveness as a teaching support, we decided to evaluate its effects on students motivation, which has been defined as *that which explains what goals people choose to pursue and how actively or intensely they pursue them* [88, p.1]. Different studies related the student's motivation and performance [88, 163].

Figure 3.15: Virtual room rendering in fast mode.

In this work, we replicated the study in [150], which exploits the Instructional Material Motivation Survey Instrument (IMMS) [88] for evaluating the motivation according to four different factors, namely *Attention*, *Relevance*, *Confidence*, and *Satisfaction* (ARCS) [87].

The main difference with respect to the work by Di Serio et al. [150] is the setting type: while they analysed the Augmented Reality effect on motivation, here we analyse the effect of Virtual Reality (VR).

Method

Through the following user test, we aim at testing the hypothesis that the immersive VR setting described in this project is able to increase the motivation of students. Therefore, we compared our setting against the technology environment currently available in Italian high schools: the LIM, a projection-based widescreen with multitouch capabilities connected with a PC.

An expert on Art History prepared a lesson on Michelangelo's sculpting technique, to be supported with RIFTart (we provided only the technical support for creating the material). She inserted in the environment an example of his early years as an artist, the well-known David, and an example of his mature phase, the Youthful Captive.

For each sculpture, she included a general description, and a more detailed explanation of three different artwork parts: the head, the left and right arm for the David, the head, the right arm and the right leg for the Youthful. This information has been recorded in advance, reading texts provided by the Art History expert. A set of keyboard buttons activated the playback of the different audio descriptions.

In order to evaluate the effects of using a VR environment, we exploited the monocular and the binocular view provided by the RIFTart prototype respectively for the LIM (not immersive) and the Oculus Rift version. Since the information provided is the same in both version, with this setting we are able to run the experiment controlling two conditions: VR versus LIM.

The test was hosted in a high school, the Liceo Filippo Figari in Sassari, Italy, which has a specific programme in Arts. We selected two different classes: all the students in the first attended the lesson taught with the LIM (and they represented our control group), while the second class attended the Oculus supported lesson. At the end of each lesson, we requested the participants to fill a questionnaire in three parts: demographic information, the IMMS [88], which allows to evaluate the overall motivation of the students through a set of 36 questions, grouped according to the four factors that lead the human motivation [87] for learning:

- **Attention**: Capturing the interest of learners; stimulating the curiosity to learn.
- **Relevance**: Meeting the personal needs/ goals of the learner to effect a positive attitude.
- **Confidence**: Helping the learners believe/ feel that they will succeed and control their success.
- **Satisfaction**: Reinforcing accomplishment with rewards (internal and external).

The questionnaire contains 36 questions, each one in a 1 to 5 Likert scale. Reversing the ratings for questions with a negative formulation and aggregating the points, it is possible to analyse each one of the ARCS factor individually, since each question is associated with one of them.

Considering that we ran the test during school time, we did not have the time to request students to complete some exercises at the end of the lesson. Therefore, we removed the exercises related questions from the questionnaire (namely question 5, 19 and 27). Two of the questions we omitted are associated to *Satisfaction*, while one is associated to *Confidence*.

In summary, the maximum aggregate score in the modified version is 165 (instead of 180). The maximum scores for the four factor are: 60 for the *Attention*, 45 for the *Relevance*, 40 (instead of 45) for the *Confidence* and 20 (instead of 30) for the *Satisfaction*.

Besides evaluating the students motivation, we included five additional questions for evaluating a set of qualitative aspects more related to the tool usability, such as aesthetics, usefulness, enjoyableness, simplicity and will to reuse the environment. These additional questions complete the evaluation since the IMMS questionnaire is focused on learning rather than usability.

Results

Twenty-three people participated to the user test. Twelve tested the LIM ($G1$), while the remaining tested the Oculus version ($G2$). The LIM group was one year older than the Oculus group ($\bar{x}_{G1} = 18.58, s_{G1} = 1.38, \bar{x}_{G2} = 17.27, s_{G2} = 1.01$). The LIM group had moderate experience with 3D environments and 3 of them already used HMDs. In G2 none of the participants already used HMDs, and they have less experience with 3D environments if compared to G1.

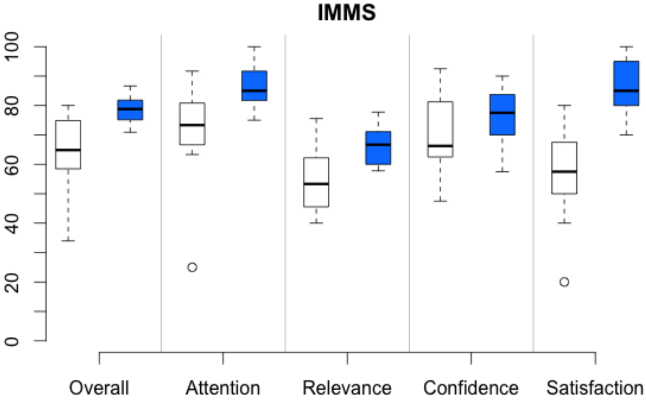


Figure 3.16: IMMS aggregated scores (normalised by the maximum value for each category). The white boxes represent the results for G1 (LIM), while blue boxes represent the results for G2 (Oculus).

Question	**G1** **(LIM)**	**G2** **(Oculus)**	**95% c.i.**	**p**
Motivation	$\bar{x} = 106.7$ $s = 20.53$	$\bar{x} = 129.6$ $s = 8.30$	$[9.23; 30.70]$	.002
Attention	$\bar{x} = 45.27$ $s = 10.18$	$\bar{x} = 52.00$ $s = 5.11$	$[2.01; 11.44]$	.013
Relevance	$\bar{x} = 24.42$ $s = 5.05$	$\bar{x} = 30.00$ $s = 3.06$	$[1.96; 9.21]$	.004
Confidence	$\bar{x} = 28.17$ $s = 5.34$	$\bar{x} = 30.36$ $s = 4.36$	$[-2.02; 6.41]$	.291
Satisfaction	$\bar{x} = 11.33$ $s = 3.26$	$\bar{x} = 17.27$ $s = 1.95$	$[3.61; 8.27]$	.000

Table 3.3: IMMS aggregated scores for the whole test (Motivation) and for each one of the ARCS factors. We highlighted in bold the ones with significant differences. For each factor we report the mean value ($\bar{x}$), the standard deviation (s), the 95% confidence interval around the mean (95% c.i.) and the p value.

Table 3.3 shows the aggregated scores means given by $G1$ and $G2$ for the whole IMMS questionnaire (indicated in table 3.3 as Motivation) and for each one of the ARCS factors. In order to compare the two groups, we used the Student's t-test for independent samples ($\alpha = .05$). We ensured that the measures for each group are normally distributed running a Shapiro-Welch test. We had to reject the normality hypothesis only for the G1 *Attention* data, we fixed the problem simply excluding an outlier value.

The t-test highlighted a significantly higher mean score for the Oculus version for the overall motivation, ranging between 9 and 30 points.

Such result confirms our hypothesis: the motivation of the students is higher if we use VR for presenting the teaching material, with respect to the technology setting currently employed in Italian high schools. We can analyse more in detail the results for the ARCS factors, which are depicted in figure 3.16. We found a significant higher difference for three out of four factors.

The difference for *Confidence* factor, even if the mean is higher in the G2 group, does not allow us to conclude that the VR or the LIM setting has an impact on the students' expectancy of success or control feeling on the subject.

In summary, we can draw the following conclusions from the IMMS result analysis:

- Using a VR setting has a positive impact on the students' interest in the lesson topic (*Attention* factor). Since the lesson contents were the same in the two versions (the 3D models or the audio descriptions), the advantage can be explained only in terms of the technology setting.
- The VR setting led to a higher satisfaction for students attending the lesson (the *Satisfaction*). This point requires more investigation: the satisfaction may be explained with the sense of novelty deriving from using a new technology for the first time, but we cannot exclude that this feeling may decrease in the long term.
- The VR setting increases the feeling that the lesson material fits the student's need (*Relevance*). The stereoscopic visualisation of sculptures provided by the Oculus Rift allows the students to better appreciate the details of an artwork, therefore they have a sensation closer to a real museum visit. For instance, the students that tried the Oculus version were much more impressed by the smoothing difference between the David and the Youthful Captive with respect to the other ones.

The second part of the questionnaire included five questions evaluating five qualitative aspects of RIFTart: aesthetics, usefulness, enjoyableness, simplicity and their wish to reuse the application. All questions requested a 1 to 7 Likert scale rating. As shown in figure 3.17, the ratings for the Oculus version are higher for all aspects considered.

More in detail, table 3.4 shows the results of the means comparison through a t-test ($\alpha = 0.05$). The differences are all significant, however the confidence interval for the aesthetics in the worst case may be not practically relevant. For all the other aspects, the students consistently prefer the Oculus version.

Conclusion

In this section we introduced RIFTart, a tool supporting teaching and studying Art History through Virtual Reality. With RIFTart, teachers can exploit 3D models for describing and comparing different artworks.

Question	**G1** **(LIM)**	**G2** **(Oculus)**	**95% c.i.**	**p**
Aesthetics	$\bar{x} = 3.83$	$\bar{x} = 5.11$	$[0.10; 2.46]$	.035
	$s = 1.26$	$s = 1.27$		
Userfullness	$\bar{x} = 5.08$	$\bar{x} = 6.67$	$[0.33; 2.84]$	.017
	$s = 1.93$	$s = 0.50$		
Entertainment	$\bar{x} = 4.00$	$\bar{x} = 6.67$	$[1.28; 4.05]$	.001
	$s = 2.13$	$s = 0.50$		
Simplicity	$\bar{x} = 4.08$	$\bar{x} = 6.00$	$[0.67; 3.16]$	.005
	$s = 1.78$	$s = 0.87$		
Reuse	$\bar{x} = 4.58$	$\bar{x} = 6.67$	$[0.56; 3.60]$	.011
	$s = 2.35$	$s = 0.50$		

Table 3.4: Qualitative questionnaire results comparison. For each aspect we report the mean value ($\bar{x}$), the standard deviation (s), the 95% confidence interval around the mean (95% c.i.) and the p value.

In addition, they can enhance them through multimedia annotations. The same material can be used by students for individual learning. RIFTart has been developed completely with web-based technologies, in order to be accessible from different devices. We discussed the technical solutions, their advantages and disadvantages. The prototype can be used on wide shared screens, such as LIMs (the multimedia whiteboard currently employed in Italian classrooms), but also on recent consumer level head mounted displays, such as the Oculus Rift or the Google Cardboard.

Considering that such solutions will be more and more available in the future, we foresee the possibility to equip schools laboratories with immersive VR, or to directly exploit student's mobiles as cheap HMDs.

We evaluated the impact of immersive VR on high-school students motivation through a user test in two classes of the Filippo Figari High School in Sassari. The results show that the immersive VR increases the students' motivation in studying the lesson topic, in particular increasing their attention, satisfaction and the perceived relevance of the teaching material.

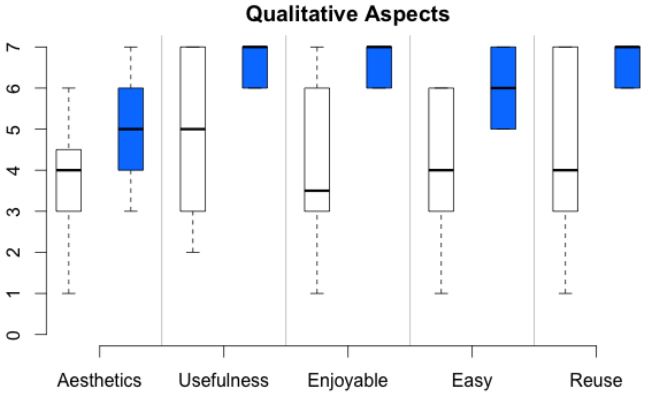


Figure 3.17: Summary of the qualitative question scores. The white boxes represent the results for G1 (LIM), while blue boxes represent the results for G2 (Oculus).

In future work, we aim to provide teachers with a proper authoring environment, a sort of Power Point for VR content, in order to better support them in the creation of the teaching material, and to evaluate both its usability and effectiveness with teachers. In addition, we would like to enhance the evaluation on two ways: on the one hand we would like to measure effect of the immersive VR visualisation in a collaborative lesson setting, where all students are provided with HMDs; on the other hand it would be interesting to perform a long term study on both motivation and students' learning outcome in classes using immersive VR settings.

3.4 SuperAvatar

When tourists are wandering around in a town or city they do not know, it is normal to use a guide to make up their mind on what is more interesting in the surrounding. If the guide is an interactive one on a mobile device, they can also consult multimedia material and listen to audio descriptions. But what if the user is a child that is much more prone to get distracted? In this work, we propose the presentation of touristic and cultural information to children through an augmented-reality approach. In order to keep the attention-span of the young users high we make use of a virtual tourist guide, appearing as a comic book superhero. To get a realistic representation of the avatar we exploit a technique for fast simulating talking heads, which is portable on mobile devices. The technique is based on pre-loading a set of meshes representing different phonemes and switching among them in order to simulate animation, without the need to compute the update for an entire face model [5.6]. In addition, we describe our first Android prototype, which shows the effectiveness of the approach for increasing children's learning by using as case study the city of Cagliari.

The wide availability of smartphones and tablets and their continuous development allows the combination of real-world perception and multimedia contents directly on a personal device, by providing directions or describing and enhancing the surrounding environment with text, images or other media.

This is specifically important if we consider travellers that can be provided with touristic and cultural contents through an augmented reality approach. The combination of real world and virtual contents has a particular relevance for conveying information to a young audience, who may benefit from a playful and enjoyable presentation. In order to stimulate the children's curiosity and to attract their attention, we propose to enhance city visits through a virtual guide with the appearance of a comic book superhero (e.g. Batman), which describes the different points of interest (POI) through ad-hoc multimedia contents. The superhero guide can be activated by downloading a mobile application on the visitor's device.

Following recent studies on Embodied Pedagogical Agent (EPA), which show a preference for cartoonish and stylised character representations for different age ranges [64], [55] for learning companions, we chose a superhero as virtual guide. Before we explain our idea for exploiting the Embodied Personal Agents for enhancing city visiting, we will describe an application scenario, highlighting the requirements and describing the envisioned application usage during the visit itself.

After that, we introduce the guide prototype implementation while the technical details of the avatar animation could be deepened in the section of the main work [5.6].

Sara and Paolo, together with their son Marco are an Italian family from Milan, and they are visiting Cagliari in October. They have already been in Sardinia for vacation, but last time they went to the seaside since it was August. This time, they want to discover the history of the city. Sara and Paolo would like to find a way to entertain Marco, who is 8 years old, in order not to bore him during the visit. When they arrive at the first monument, called "Bastione di San Remy", Paolo reads the information panel, searching for information on its history. Oddly, he sees that the panel contains a picture of Batman, and a text that invites to frame it with a smartphone camera. Intrigued by the strange message, Paolo uses his smartphone and finds out that there is a virtual city guide which can be installed on his own smartphone, simply by framing the Batman poster with the camera. Once the application has been installed, he sees that a virtual Batman appears on the screen. Paolo calls Marco and shows the strange application to his son. Marco is excited to see his favourite superhero on the mobile phone, and starts listening to him describing the monument. After 20 minutes, the family moves towards another monument, the "Torre di San Pancrazio". When they reach its location, Batman appears again and Marco listens to history of that monument. Sara and Paolo found how to enjoy the city and entertain Marco at the same time.

Setup discussion

Exploiting the user's personal device in order to visualise the EPA poses a set of technical problems. Having updated contents is crucial for the success of a touristic application. Therefore, information providers such as municipalities, itinerary curators or event organisers must be able to create contents easily. Most of the times they want to produce different descriptive texts. In addition, since cities have many points of interest that can be described during the visit and it is not possible to establish a priori the users' path, contents should be downloaded by the virtual guide on-demand.

The solution we adopted is to keep the content that have to be presented by EPAs in a text format, which can be easily and quickly downloaded, and to exploit avatar animation together with text-to-speech (TTS) for simulating the talking-head on the mobile device. The problem has been solved by using an improved version of our THAL-k [5.6] for Android. For the vocal rendering we used the TTS engine set as default by users on their Android device. We implemented it with the Google TTS [60] and IVONA [76].

Figure 3.18: Batman and Joker describing a point of interest.

Besides the mouth animation, the THAL-k allows to define the look-at point for the avatar. We exploited such capability for engaging young users more, looking at her or him while talking, or looking in a different direction for suggesting where to look for finding the current subject in the real world.

In order to use a superhero as EPA, we needed to build the appropriate 3D model of the character.

Figure 3.18 shows two EPA configurations, representing Batman and Joker, but many more can be created changing the default textures or adding new ones and setting the relative parameters.

We used FaceGen [46] for the model creation, a software that allows users to easily customise the appearance of a human face. In order to create Batman, we used a basic adult male face and we added a supplementary 3D model for the well-known character's mask. Joker, on the other hand, required some work on face texture for simulating the make-up, together with the selection of his hairstyle and colour. It is clear that an easy work is sufficient in order to customise the avatar and to add supplementary models (e.g. the mask) for obtaining different EPAs. Instead, the animation is completely handled by our THAL-k library, just by providing the text to be read.

In order to build a first city guide prototype, we considered Cagliari as case study, provided the availability of tourist information previously acquired through a related work on visiting the city [111]. The main role of the talking head is to describe and explain this information like a real touristic guide, by using a language and an appearance that result more interesting for a young audience.

The talking head animation is launched when the user is nearby monuments and points of interest (POIs). Every POI is associated with its GPS coordinates (longitude, latitude, altitude), a title and a description. In addition, it is possible to add other multimedia content, such as images and videos. The application catches the user's position via GPS and, if there is a POI nearby, the application shows the EPA and the content presentation starts. Figure 3.18 shows the two sample EPAs describing the "Bastione of San Remy", one of the most famous fortifications in Cagliari. Apart from the textual description, the push notification also provides information about the additional multimedia content availability, which can be opened directly from the guide application.

Users can easily install the application, using a QR-code printed on the monument description panel. In addition, if the GPS signal is not available and it is impossible to determine the user position, the QRCode can be used as a fallback localisation mechanism, in order to start the talking animation.

To attract the children's attention and to invite them to use the application, a super-hero poster is displayed nearby the POI information panel. When the child frames it with the mobile phone, application matches the poster and the super hero wakes up, greeting the child and starting to explain the monument history. This matching is achieved by an image-processing algorithm that tracks and estimates the camera position in real time by correlating the real-world coordinate system and the 2D image taken from the camera thus allowing the visualisation of 3D objects by over-imposing them in the exact position over the printed poster.

Conclusion

In this work we described the design and a prototype implementation of a mobile guide for city visits for a children audience. In order to attract their attention, a comic-book superhero avatar presents the information. This choice is grounded over recent studies about the Embodied Pedagogical Agents, which demonstrates the preference for cartoonish and stylised characters as learning companions.

We discussed our solution to the problem of animating the superhero while talking, with a low consumption of resources, suitable for mobile devices. In addition, we detailed how we created the guide prototype. In future work, we will refine the guide implementation and we will perform a user study with real children. Our hypothesis is that the proposed approach may increase the children learning during the city visit, if compared to reading information on a printed or electronic guide.

Chapter 4

Improving Real Life with New Technologies

Most people alive today grew up without the Internet network or mobile devices, such as smartphones and tablets equipped with apps for everything. Technology is always developing thus improving our lives in many ways like education, health and communication. This is due to the fact that it is supposed to make life easier.

Photo credits: Sergey Galyonkin

This happens for several and really different works and tasks, such as practising for a sport or doing shopping. Think about the mirrors in the changing rooms, what if they were able to suggest ways to dress? Or if they were able to check your physical state? This can improve your life. Again, what is better, to stimulate the increasing practice of indoor physical activity at home or at the gym, or using Immersive Virtual Reality (IVR) devices? Our bet is that they will be part of the next generation of fitness apparel.

Technology is entertainment too. It offers more opportunities to have great time through innovative systems, take for example the time machine.

Unfortunately, we still cannot travel through time and space as we wish. But we can simulate it, we have done such kind of experiment in WoBo (World in a Box), reinterpreting the concept of travelling, offering users the possibility of visiting distant or hardly reachable places, simply by wearing a head-mounted display inside an appropriate cabin.

The section [4.2] is a revised version of the paper [47].
The section [4.3] is a revised version of the paper [144].
The section [4.4] is a revised version of the paper [159].

4.1 Rift a Bike

What is better, to stimulate the increasing practice of indoor physical activity at home or at the gym, or using Immersive Virtual Reality (IVR) devices? Our bet is that they will be part of the next generation of fitness apparel. The decreasing hardware cost makes it affordable to pair IVR devices with treadmills and exercise bikes. In this work, we apply well-known gamification techniques coming from different mobile applications (i.e. e-learning, lifestyle, fitness, etc.) for demonstrating how users can be entertained while working out with an exercise bike. We evaluate the effectiveness of such techniques through a user study, deriving a set of guidelines for creating what we call fitmersive games (immersive games for fitness). The gamification [181] approach consists of including video games features in applications outside the playing context, in order to increase the users' engagement in different activities. In particular, the gamification has been applied in contexts related to the physical and psychological status of the user, such as learning, habit changing, keeping a good lifestyle and fitness. Considering that usually such activities envision long-term goals, gamified applications run on personal devices, in particular mobiles. Aerobic physical exercises, such as running or cycling, can be particularly tedious when carried out indoors. Nowadays, consumer-level immersive Virtual Reality (IVR) devices (e.g. the Oculus-Rift [124]) provide an affordable solution for creating IVR settings at home or at the gym, which can be used for entertaining people while training.

In this work, we evaluate the effects of gamification techniques in IVR environments for fitness that, from now on, we call *Fitmersive Games*. We first describe both the hardware setup and the design of a bike fitmersive game (Rift-a-bike), and how we applied the different gamification techniques. After that, we report on a user study, where we quantitatively measured the gamification effect on the users' physical activity enjoyment. In addition, we collected qualitative feedback on the different gamification features, according to three dimensions (usefulness, motivation and fun) together with an overall evaluation of the game experience. Finally, we derive a set of guidelines from the study results, which we will use for selecting the different techniques to apply in other fitmersive games.

Setup discussion

We created a prototypal implementation of a fitmersive game, called Rift-a-bike, whose goal is to have a bicycle ride in a digital city. The user focuses on keeping the appropriate speed for the training phase.

Figure 4.1: Rift-a-bike hardware setup.

The hardware setting only includes low cost devices and it is easily replicable at home or at gym. We used a normal exercise bike, which we instrumented with a Raspberry PI [140] for sensing the user's cycling speed. In order to explore the environment, the user wears an Oculus Rift head mounted display (HMD), which provides a stereoscopic view on the virtual world and a set of earphones for the audio. In addition, we used a Kinect for Windows v2 [112], for tracking the user's movement. This ensures a higher presence perception in the virtual environment, since we are able to let the avatar replicate the user's movement. Figure 4.2 shows the user's avatar in the game environment, while figure 4.1 shows the hardware configuration for playing the game.

From the software point of view, we created the virtual city using CityEngine [45], a procedural city environment generator. We imported the city model in Unity 3D [169] and we built the bike tour simulation. We added the gamification features on top of this environment. In figure 4.2 we included a set of screenshots of the different game features. In the rest of this section we describe how we applied a set of gamification techniques in the virtual environment.

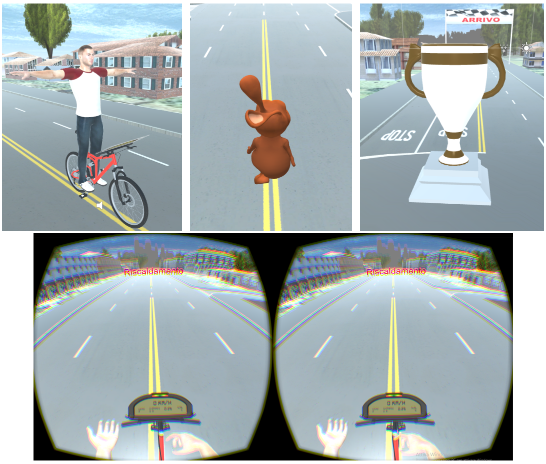


Figure 4.2: Rift-a-bike fitmersive game environment. From left to right: the user's avatar, the 'follow the rabbit' challenge, a session final trophy (badge) and the stereoscopic view for the Oculus Rift.

Levels

In a videogame each level should set goals for the users which are difficult but reachable. At the same time it should make them challenging but not frustrating for the average user. Throughout the different levels, players gain experience and follow the game plot. In Rift-a-bike, we included three different levels during the virtual city bike tour, each one corresponding to the phases of a physical work-out. In the *Warm-up*, the user pedals at a medium-low speed; in the *Exercise*, the user pedals at the speed required for the training; in the *Cool-down*, the speed gradually decreases leading to the end of the exercise.

From the users experience point of view, the partition helps them in gaining confidence with the system before starting the real exercise, and entertains them while showing the different phases in a playful way. From the physical exercise point of view, the first phase prepares the user to the physical strain, while the last one relaxes the body and guides it to the resting phase in a gradual and controlled manner [34]. It is possible for the user to set-up the duration of the different phases autonomously (e.g., 5 minutes for warm-up, 15 for the exercise and 5 for the cool-down, 25 minutes in total).

Points

The points are virtual rewards that stimulate the users to perform a set of actions. In addition, they provide also a quantitative measure of their performance, according to different metrics.

For instance, the user gradually gains Experience Points (XP) completing quests, or she increases her Skill Points (SP) by completing a certain action, or she is rewarded with Karma Points (KP) by helping other players.

We included a skill point system in Rift-a-bike, which rewards the user if she is able to maintain the correct speed. The user gains points collecting the coins positioned by the game engine along the path. The number of coins depends on the user's speed: the closer the pace is to the correct one, the more coins the user will collect.

Challenges

The challenges (or missions) represent particular tasks or quests that the user has to complete for gaining experience. For instance, at the very beginning of a game the users are requested to complete very simple challenges, such as breaking a box or jumping an obstacle, in order to become familiar with the game mechanics. Later in the game, challenges are related with the development of the plot (e.g., finding a treasure), or they may be optional (and more difficult) parts of the game, which usually attract the more eager and skilled players.

In Rift-a-bike, the whole training exercise is represented as a challenge: the user's objective is to follow a rabbit, trying to keep-up with him as long as possible (see figure 4.2). In this way, the player gains a precise objective during the training. This strategy has a twofold goal: it distracts the user from fatigue while motivating her in continuing the exercise. In addition, the rabbit is an implicit feedback mechanism: the user is able to evaluate her performance by simply considering the distance from the rabbit.

Badges and Prizes

Badges are special rewards for different achievements in the game. For instance, when the user finds 100 diamonds, she may be rewarded with a treasure finder badge, which acknowledges her ability in that particular activity. Badges increase the reputation of the player in the community and also her self-esteem. In addition, they provide additional motivation for playing.

It is usual in a videogame that a single achievement may be split in different sub-activities that the user should complete to receive a prize. When she collects all the prizes, the achievement is unlocked. Consider, for instance, a car-driving game: the professional driver badge may be unlocked when the user wins all the championships at amateur level.

In Rift-a-bike we included both badges and prizes. The badges are represented through a set of trophies that the user receives at the end of the training (see figure 4.2).

Their value is related to the overall evaluation of the training session. In addition, the user collects three different prizes during the session, one for each phase. Their value is related to the user's performance in the considered phase. If the user gets the best prize in each phase, she will get the most valuable trophy.

Evaluation

We have carried out a user test for evaluating two different aspects of our fitmersive game prototype. The first objective is to establish whether the users enjoy more the physical activity with gamification elements while interacting with IVR environments. The second one is to provide a qualitative assessment of the different gamification techniques used in our prototype, according to three dimensions: usefulness, fun and motivation.

The test consists of two 10-minutes sessions at the exercise bike, in which we vary the environment condition between two nominal values: Virtual Environment (VE) and Game (G). In both conditions we used the Oculus Rift sensor. In the VE condition, the user pedals in the virtual environment without any gamification feature. She is still able to observe the virtual city with the HMD and the avatar replicates her movements. In the G condition, the user pedals in the same virtual environment, enhanced with the gamification techniques. Each user carried out the test in both conditions, half of them started from the VE condition, while the other half started from the G condition.

Before starting the two sessions, the users completed a demographic questionnaire, useful for collecting information on their physical activity habits, experience level with virtual environment and other data that characterises the sample population. After each session, we requested the user to fill the Physical Activity Enjoyment Scale (PACES) [89], a standard questionnaire for evaluating the level of users' enjoyment during the physical activity.

At the end of the test, the users filled a qualitative questionnaire for evaluating the overall usability of the environment and the different gamification techniques according to their contribution to the usefulness, fun and motivation.

Twenty-one users participated to the test, 19 males and 3 females. The average age is 23 years old ($\bar{x} = 23.18$, $s = 3.0$, $min = 16$, $max = 29$), with different levels of education: junior high school (1), high school (13), bachelor (4) and master (3). Only three users exercise once a day, the majority of users performs physical activity once a week (10), the others are uniformly distributed among the lower levels of activity: 4 once a month, 3 once a year and 2 less than once a year.

The experience with virtual environments varies among the users: 3 uses them more than once a day, 2 once a day, 3 once a week, 6 once a month, 6 once a year and 1 never used them before. In addition, most of them already tried an HMD (18), while for three of them it was the first time.

The PACES test consists of 18 questions in a 1 to 7 Likert scale. Figure 4.3 summarises the results for all questions, showing the ratings for the G condition are consistently higher with respect to the VE condition. In particular, the mean aggregated for the VE condition is $\bar{x}_{VE} = 91.76$, $s_{VE} = 15.52$, which is already high, but for the G condition is even higher ($\bar{x}_G = 106.90$, $s_G = 10.53$). We compared the two means with a paired t-test ($\alpha = 0.05$) and we found a significant difference ($p = .001$), whose 95% confidence interval is $[6.83; 23.45]$. Therefore, we can conclude that our users perceived a difference between the two conditions and that using the gamification techniques positively affects the physical activity enjoyment in the IVR setting, that was the same for both condition. In the post-test questionnaire we asked the users to evaluate the gamification features according to three different aspects: usefulness, fun and motivation using a 1 to 7 Likert scale. Figure 4.4 summarises the questionnaire results. First of all, we noticed that the *Challenge* received consistently the higher score, and such difference is significant with the *Prize* feature for all aspects (respectively $p = .002$, $p = .001$, and $p = .002$).

Considering the usefulness aspect, we found a significant difference also between the *Levels* and *Prizes* ($p = .01$), while between *Points* and *Levels* there is a practical significance ($p = .06$). The feature that has the most relevant impact on the users' motivation is the *Challenge*, followed by *Points* and *Badges.* The impact of *Levels* and *Prizes* is less relevant.

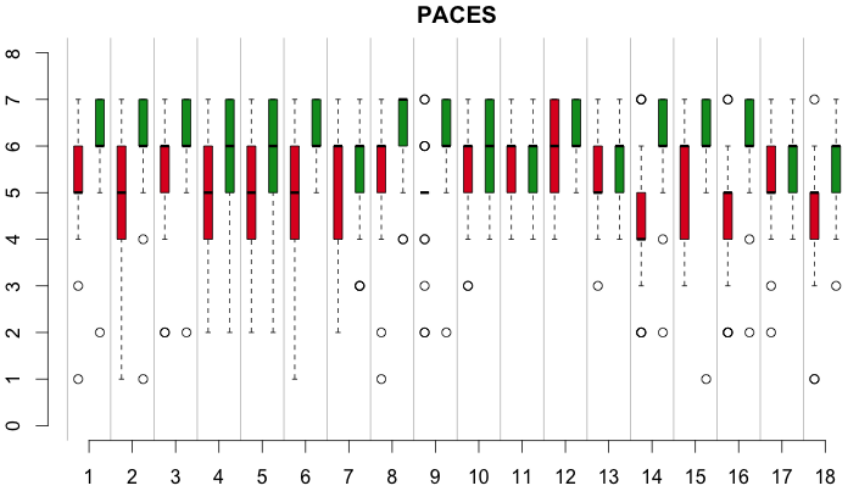


Figure 4.3: Paces questionnaire results. Red boxes are related to the Virtual Environment (VE) condition, while green ones to the Game (G) condition. We reversed the scores for questions with a negative connotation.

The general usability questions scores indicates a good overall satisfaction for our users, and their will to reuse the environment again. We included also open-ended questions in the last part of the questionnaire, in order to have insights on the motivation behind the scores. The majority of the user explicitly cited the rabbit challenge as the most entertaining and useful feature of the game, since it both motivates the exercise and provides a good feedback on the user's performance. The users appreciated also the immersion in the environment, and the possibility to change the viewpoint moving the head.

The negative aspects are related to a well-known side effect of HMDs: the motion sickness. Six of them have had a slight nausea, especially in bends at high speeds. This is an open issue for HMDs in general and for Oculus Rift applications in particular. However, we should try to mitigate the problem also while designing the city paths. Other minor problems are related to a general improvement of the graphics quality, and some synchronisation problems between the movement of the user and the avatar with the Kinect sensor.

We can summarise the lesson learnt from the user test results and from the open-ended questions, providing some guidelines for applying gamification features in fitmersive games.

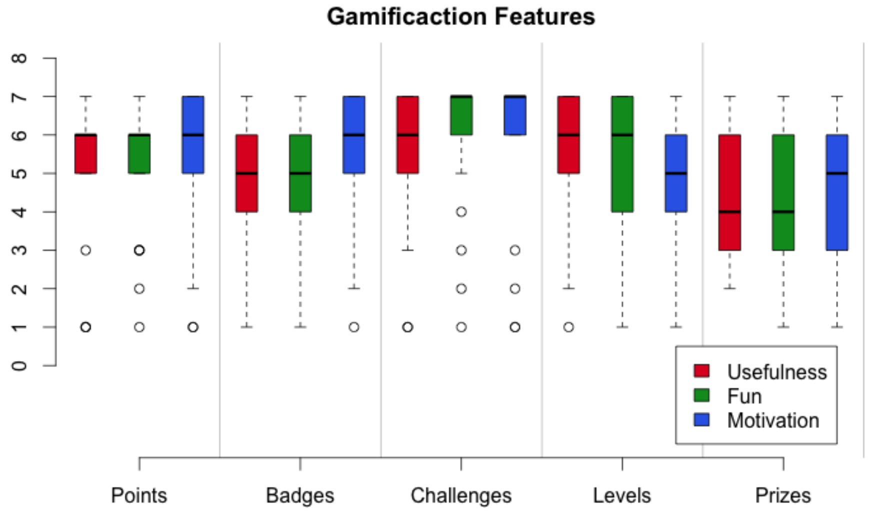


Figure 4.4: Qualitative assessment of the gamification features included in Rift-a-bike. Red boxes represent the usefulness rating for each feature, the green boxes the fun scores and the blue boxes the motivation scores.

Challenges are the most relevant gamification feature in fitmersive games: they can be used for both motivating users and for helping them in completing the workout correctly. They are also important for the overall experience, since the main objective of the user is not any longer a boring exercise, but a challenge inside a game, which provides a higher motivation for completing the activity.

Levels guide the user inside the physical training system and helps them in maintaining a consistent progress plan. In addition, they support the user in understanding how difficult will be the current task. Completing a level provides them a sensation of fulfilment, which facilitates their permanence in the fitmersive game.

Collecting **Points** provide an immediate goal for the users and, even if they are virtual rewards, they positively affect the motivation and the fun during the exercise, distracting the users from the physical strain. We suggest to select a point system able to maximise the exercise objective, such as the regularity of speed for the endurance, the precision of the movement for a particular muscle training and so on.

Badges provide an indication on the overall player status. Considering that we evaluated a single-player game, their impact was limited. However, we recognise that including a social network support in our prototype would have lead to different results. Therefore, we suggest to give priority to other features in single-player settings.

Including **Prizes** as intermediate rewards does not seem to have a particular influence on users with respect to the other gamification features. On the other hand, they do not degrade the user experience. Our suggestion is to include them only if needed, but to give priority to all the other features first.

Conclusion

In this work we discussed the application of gamification techniques in Immersive Virtual Reality environment for supporting physical exercise. We describe the implementation of Rift-a-bike, a fitmersive game that enhances the exercise bike experience with challenges, levels, points, badges and prizes. Through the results of a user test, we show that the gamification elements increase the user's enjoyment during the physical activity and we provide a set of guidelines for applying gamification features for fitmersive games. In future work, we will provide multiplayer and social support in Rift-a-bike, in order to stimulate the competition among the members of a social network. In addition, we will analyse the gamification effect in the long term.

4.2 WoBo

WoBo (World in a Box) aims to provide a new experience for travellers, allowing them to visit distant or hardly reachable places through the exploitation of consumer cameras and a head mounted display. The experience consists in watching a 360-degrees video with 3D audio in a dedicated cabin. The user can select videos shot in different places, which have been created with six consumer cameras. We describe the proposed experience, the hardware and the software used for a first prototype.

In many science fiction movies, one of the most desired possibilities is breaking the limit of physical distances for reaching far places, envisioning high speed vehicles, tele-transportation or telepresence mechanisms. With WoBo (World in a Box), we propose a new way of interpreting the concept of travelling, able to provide the user with the possibility of visiting distant or hardly reachable places, simply by wearing a head-mounted display inside an appropriate cabin. In this way, the users have the impression to be carried away in a new reality, and it will be as if the whole world could fit in the cabin. We describe in this work the experience, the technical set-up of a first prototype and the results of a preliminary evaluation, showing an encouraging acceptance by end-users. World in a Box (WoBo) provides immersive and realistic virtual journeys using the latest VR technologies, allowing a fast and cheap world (virtual) travelling. Both virtual and real world destinations are reachable by entering in a special box that we called WoBoX.

We envision the installation of several WoBoXes around the world, all connected to each other. For instance, it would be possible for the user to step into one located in New York and watch what is happening around the WoBoX in Seoul. In addition, it would be possible to have WoBoXes also on beautiful but hardly reachable places, such as mountain tops (e.g. the Everest), deserts etc. Each WoBoX is both able to transmit and receive signals to and from other boxes. Therefore, a WoBoX consists of two parts: the first one is a 360 degrees video recording equipment, for creating an immersive and realistic visualisation of the area around the WoBoX. The second part is a cabin where the user can, wearing a head mounted display and a set of earphones, experience the travel.

Setup discussion

We acquire a single 360-degree video stitching six different ones captured from six cameras. We position two cameras along each one of the three orthogonal axes in a 3D space, looking at opposite directions.

We created a small mount for securing the camera position while recording, following the schema in Figure 4.5. In our idea, such cube should be positioned on the top of each WoBoX, in order to capture in real-time what is happening around a specific place.

We used six GoPro Hero3+ in our prototype implementation. For having a complete coverage of the space around the cube mount, we shoot 4:3 videos having a resolution of 1440p at 30 fps. The 16:9 ratio was not an option, since it does not cover the entire space around the cube, considering the two horizontal cuts of the sensor image.

The six recorded videos are merged into a single panoramic video that can be mapped over a sphere as texture. This merging operation is done using a third party software (Kolor Autopano video) that uses a stitching algorithm to merge and create the final 360 degrees video.

The algorithm uses a feature-based approach for the video stitching, identifying the different points where it is possible to combine the frames coming from the different cameras. More precisely, the software exploits the SIFT features [104] for calculating the salient points and computing the matchings between two camera frames.

In the final WoBoX setting, we envision an almost real time stitching and broadcast of the 360-degrees video, trying to create a fast version of the algorithm that exploits the particular configuration of the mount cube. We consider acceptable also a delay of a few seconds between the video shooting and the user visualisation, provided the user has no means for modifying the surrounding environment and, consequently, for noticing such delay during the interaction.

In order to record a three-dimensional sound of the scene we need to use the six GoPros built-in microphones using the same concept of the binaural recording method. In our work we can not adopt the exact binaural method due to the fact that the user can move her head around. To overcome this issue we have decided to record the surrounding sounds and then replay them accordingly to the user's head position in the 3D space. In this way we are able to give a 3D stereo sound sensation to the user of actually being in the place where the sound have been recorded. Such recording method is intended for replay using exclusively headphones. The virtual travel through one WoBoX consists in watching the 360-degrees video. In order to provide a seamless and immersive experience, we used a head mounted display for the video visualisation.

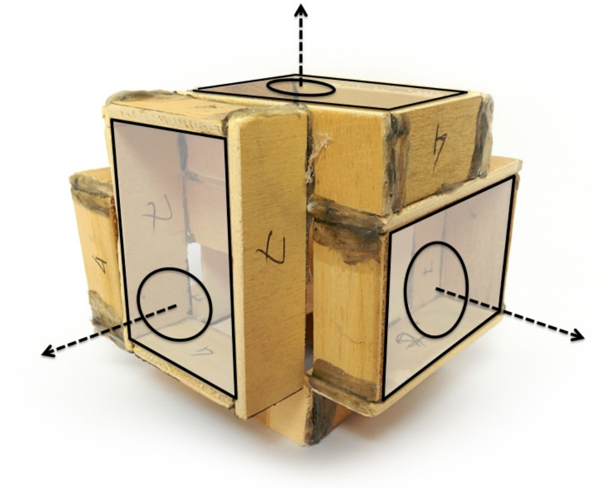


Figure 4.5: Mount prototype

In particular, the prototype implementation exploits the Oculus Rift[1] head-mounted display, which is able to provide a stereoscopic visualisation and to update the point of view according to the user's head movements.

The prototype application for watching the 3D video has been developed using Unity 3D. In order to set-up the environment, we inserted into the world a high-quality model of a sphere. The 360-degrees video is mapped as a texture on the sphere, while the world camera, corresponding to the user's head, is positioned in its center. The world has one light source, positioned again in the center of the sphere. We reversed its normals for viewing the video texture from the inside. The configuration of the 3D world is shown in Figure 4.6. As explained in the audio acquisition section, we record six sound sources using the six GoPros built-in microphones.

In the Unity framework we can add all these sounds positioning them in the 3D scene and, in addition, we can stream two different audio tracks: one for the right ear and the other one for the left ear.

---

[1] `http://www.oculus.com/`

Figure 4.6: 3D Scene with video texturing

Evaluation

In order to exemplify the travelling experience, we built a first prototype of a WoBoX, with the aim to obtain something close to telephone booths, but without interfering with the user movements. The prototype consists of a wooden box, whose dimensions are 1.5m x 2m x 2m. To simulate the availability of different WoBoXes in different places, we recorded three videos in different sea places in Sardinia, Italy. Once a user enters in the WoBoX, she can select one of the available destinations. Then, she can start the virtual visit. The prototype is equipped with an Oculus Rift and high quality headphones, connected to a desktop computer running the Unity 3D application. Since it was not possible to avoid using a wired connection for the Oculus Rift, we tried to reduce the interference with the user's movements fixing the wire at the top of the cabin.

We conducted a small-scale user study in order to evaluate the proposed user experience. The study was organised as follows: after completing a small demographic questionnaire, the users watched for a minute a seascape video in the WoBoX, with the head mounted display and 3D audio. After that, the user filled a Presence questionnaire as in [102].

We removed the questionnaire items related to the Control Factors (see [102] for a description of the factors and their mapping on the questions) since they were not meaningful considering that the users can control only the camera point of view, but they are not able to modify the virtual world configuration. The item ratings exploit a 1 to 7 Likert scale.

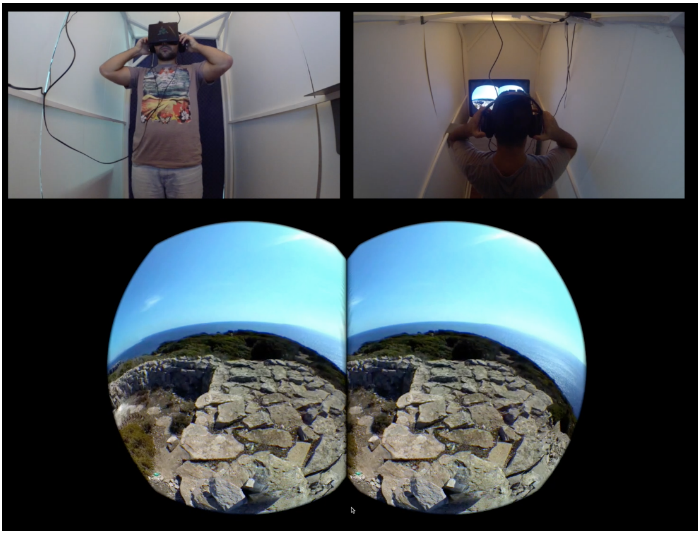


Figure 4.7: A user interacting with the WoBo prototype

Twelve users participated to the test, 9 males and 3 females, aged between 22 and 29 years old ($\bar{x} = 23.33$, $\tilde{x} = 23$, $s = 2.46$) and the test results were encouraging. First of all, the sensory engagement was highly rated considering the overall experience (q4: $\bar{x} = 5.67$, $s = 0.71$), the visual (q5 $\bar{x} = 5.83$, $s = 0.71$) and audio experience (q6: $\bar{x} = 5.83$, $s = 1.19$). In addition, the user did not find inconsistencies in the information coming from the different senses (q11: $\bar{x} = 2.42$, $s = 1.73$, 1 means no inconsistencies) and they were not confused after the session (q22: $\bar{x} = 2.33$, $s = 1.67$).

The debriefing session after the test highlighted a major usability problem for the prototype represented by the wires (for both the Oculus Rift and the headphones) that, physically interfering with the exploration, degrades the feeling of being into another place.

We plan in the future to exploit devices that embed a smartphone or a tablet inside the head mounted display (such as the Samsung Gear VR) in order to avoid using wires.

Conclusion

In this work we introduced WoBo, a new experience for travellers which aims to connect different locations reachable immediately for a virtual journey. A WoBoX is able to record a 360 video of the surrounding environment, which can be viewed in any other WoBoX in the world. We described the entire setting for a first prototype and the results of a preliminary evaluation of the proposed experience, demonstrating a high level of sensory engagement.

4.3 Smart Mirror

In this section we introduce the Virtuoso project, which aims to create a seamless interactive support for fitness and wellness activities in touristic resorts. The overall idea is to evaluate the current physical state of the user through a technology-enhanced mirror. In 2.2, we described the state of the art technologies for building a smart mirror prototype. In addition, we compare different parameters for evaluating the user's physical state, considering the user's impact, the contact requirements and their cost. Finally we depict the planned setup and evaluation setting for the Virtuoso project.

During our vacation time, especially in the summer, we would like to interrupt our daily routine and to regain the energies we lost during the whole year. At the beginning of the summer, many people do not like the image reflected in the mirror, and it is common to start the vacation with the intention of spending some time in physical activities. Mirrors do not lie... But what if mirrors were able to suggest solutions rather than cause complaints?

This is the objective of the Virtuoso project: to create a seamless interactive support for fitness and wellness activities in touristic resorts. The overall idea is to evaluate the current physical state of a resort guest through a mirror metaphor. Then, the system will try to help her in enhancing it during the stay, suggesting different activities available in the resort. Both the analysis and the suggestions should be provided in a playful way. In this work, we introduce the project requirements and we provide a survey on the existing enabling technologies for creating the mirror metaphor and to acquire data from the user. It can be helpful for researchers and practitioners interested in creating a similar environment.

From a technological point of view, creating such environment poses a set of challenges, both technical and related to the user experience. The first challenge is how to design and realise the smart mirror and its related apparels. Different companies provide off-the-shelf smart mirrors whose functions are related to controlling and supporting smart home environments, multimedia controllers and augmented reality devices. In 2.2, we provided a detailed review of the existing devices in all three categories, discussing their advantages and disadvantages in general, and in particular for the Virtuoso project.

The second challenge for our project is how to support the evaluation of the user's physical state. Such evaluation does not require the same precision if compared to a real diagnosis, which is out of our scope.

In addition, since a resort is a relaxed environment where users expect to enjoy themselves, it is not feasible for Virtuoso to use intrusive equipments or methodologies for acquiring data. Therefore, in this work we analyse different measurements that can be quickly performed on a user. For each of them, we detail the procedure for acquiring the data, assessing its cost both money-wise (i.e., the equipment to be used) and user-wise (i.e., how bothering it can be). At the end of the vacation the mirror should profile our user again, if the activities had their effect, the resort guest can bring home a valuable souvenir: an improved physical shape. Or at least some good suggestions!

The main aim of the Virtuoso project is to provide a set of tools, easy to use, that does not necessarily require to be used by specialised health personnel (medical or paramedical), to assist tourist operators. The typical user of the system will be the fitness and/or wellness coach. This kind of professional is a figure that is more and more diffused also in contexts of vacation resort. The activity of the wellness coach is to identify the tourists' need and guide them in choosing and practicing fitness activities offered by the resort. The main tool we want to develop will be a computer-based application for the general assessment of the psychological and physical conditions of the customer. We aim, in other words, to measure his or her well-being with some simple and compact indicators. The application used by the center will have a companion mobile app, for end users, that allows to browse the information produced in the initial evaluation session, to log the activities during the vacation, to socially share it, and to bring back home the lesson learnt and continue to practice the good customs.

To reach this goal we will study and develop methods and techniques for the analysis of the general health of an individual, based on heterogeneous data acquired from non-invasive and non-intrusive sensors like, for example, cameras and platforms with sensors, contactless or minimally contact, such as multifunction armbands. The basic idea behind the whole project is the ability to assess the physical and psychological condition of an individual and give him or her good advices on how to improve, through an analysis of diagnostic signs collected using different techniques and technologies.

Physical State Acquisition

In this section we report different physiological parameters that can be evaluated through different sensors. We summarise their definition and the measurement procedure. After that, we compare them in table 4.1, according to three dimensions:

1. The *Cost* of the equipment for measuring the parameter. We considered three cost levels (*Low*, *Medium* and *High*).
2. Whether the sensors used for measuring the parameter require some *Contact* or not.
3. How annoying the procedure is or whether or not it requires uncommon actions for a user (*Impact*). We consider the project's audience: people monitoring their physical status for wellness purposes and not for a real diagnosis (they are spending their holidays in a resort). We consider three possible degrees: *High* if the procedure takes a long time or it requires difficult actions, *Low* if the procedure is short or very easy, and *Medium* for values in-between.

For instance, acquiring user's weight through a pressure sensitive board (e.g., the Wii Balance Board) and her height through a laser tool, makes it possible to calculate the Body Mass Index (BMI). Combining such data with a volumetric representation of the user's body allows to analyse the body weight distribution without any direct contact with the user. So the BMI has a low cost in terms of intrusiveness and a reasonable cost in terms of hardware. Such measurements together with other information provide the evaluation data and could drive the activity during the stay. The following is the list of parameters we considered:

1. **Metabolic Balance of Fat and Glucose**. Breath acetone is a parameter for detecting the correct exploitation of all energetic substrates [54]. Indeed, some metabolic diseases cause a predominant use of the fat substrate instead of glucose. This increases the production of acetone which is expelled through the lungs. The presence of acetone in the breath may indicate infectious diseases or diabetes. It can be measured through colorimetric reactions on a disposable transparent support with a reactive substance. We can detect the colour change with a camera. Such measure is *intrusive* and needs a *contact* with the user.
2. **Glucose metabolism**. There is a correlation between the saliva and the blood glucose level. A variation of the value from the reference levels reveals an altered glucose metabolism, caused by organ or hormonal diseases. It can be measured again through colorimetric reactions (*intrusive* and *contact*).
3. **Circulation of blood**. An irregular body heat distribution indicates blood circulation problems such as e.g. the venous stasis. Therefore, we can obtain such distribution through an infrared camera and analysing the thermal image [10]. The technique is both *unintrusive* and *contactless*.

4. The **Body Mass Index** (BMI) is the ratio of a person's weight and height squared. This biometric parameter indicates whether the weight correlates well with a height. In general, a person having a BMI greater than 25 is overweight, greater than 30 is obese, while below 18.5 is underweight. For calculating this index it is sufficient to measure the weight through a pressure board (e.g. the Wii Balance Board [35]) and the height with a laser sensor (*unintrusive* and *contactless*).

5. **Fat distribution**. It is possible to use the ratio between waist and hips circumference for identifying an excessive visceral adipose tissue mass accumulation, which is strongly correlated with cardio vascular diseases. The circumferences can be measured through laser and infrared sensors [5] (*unintrusive* and *contactless*).

6. Heart and breath functionalities can be evaluated through the **oxygen-haemoglobin saturation** levels (indicating a good lung ventilation and blood) and the heart rate. We can measure these parameters using a pulse oximeter, which requires a light pressure on a finger (*unintrusive*, and *contact*).

7. **Stress level**. People react to specific visual stimuli with different face movements, which can be evaluated for establishing the stress level. Therefore, it is possible to create a test where the user looks to a sequence of pictures and a software analyses the face movements [178]. The method is *unintrusive* and *contactless*.

8. **Muscular tone** evaluation. The muscular strength allows to evaluate the physiological condition of the active lean mass. We consider the strength of the quadriceps and of the dominant hand and forehand muscles, since they represent better the general condition of the entire muscle mass. We can measure the quadriceps strength with a dynamometer set in a fixed position while we measure the hand strength measuring the grip force [2]. The method is *unintrusive*, but it requires some *contact*.

9. **Body mass distribution**. The intracellular and extracellular water volume allows to evaluate the lean mass percentage (through a simple ratio). We can measure such volumes through the bio-impedance analysis: the user stands on impedance board and grabs two handles to allow the electricity flow from hands to feet [77]. The method is *unintrusive* but it requires some *contact*.

10. **Hearing** can be evaluated with an audiometric curve: a device emits a sound scale, ranging from 15 to 20000 Hz. The volume is variable between 0 and 100 dB. The user listens to the sounds and signals at which pitch s/he starts hearing [75]. The method is both *unintrusive* and *contactless*.
11. **Night vision**. We evaluate the functionality of the rods (the black-/white and shape photoreceptors) testing the adaptability of the eye to night vision, asking the user to recognise the objects in a set of images in low light conditions. We evaluate the recognition and detail degree of their image descriptions. The method is *unintrusive* and *contactless*.
12. **Blood pressure**. It provides information about the circulation of blood and also on the stress level. We can measure it with a sphygmomanometer applied at the user's wrist. The method is unintrusive and requires some contact.

Id	Parameter	Impact	Contact	Cost
1	Fat and Glucose Balance	High	Yes	Low
2	Glucose Metabolism	High	Yes	Low
3	Circulation of blood	Low	No	High
4	Body Mass Index	Low	No	Low
5	Fat distribution	Low	No	Low
6	Oxygen-haemoglobin saturation	Low	Yes	Medium
7	Stress level	Low	No	Low
8	Muscular tone	Medium	Yes	Medium
9	Body mass distribution	Low	No	High
10	Hearing	Medium	No	Low
11	Night Vision	Low	No	Low
12	Blood pressure	Low	Yes	Low

Table 4.1: Physical state parameters summary.

Setup discussion

The ideal setup of the system developed in the Virtuoso project is as following. Imagine a tourist arriving at his or her chosen resort, in summer, in Sardinia, to enjoy a relaxing week or fortnight of vacation. Most probably, he or she does not like to undergo a set of medical exams to figure out what is his or her state of health. But, what if, he or she will enter wearing a swimsuit, a funny room with a board on the ground, and a Microsoft Kinect mounted on a rotating arm capable of turning the board around, and a mirror interacting with him or her asking to answer simple questions just waiving a hand?

This will be fun, subtracting just five or ten minutes to the beach, and will be the basic set of information to compute a profile containing age, height, weight, BMI, body volume, fat vs muscle ratio, and some more parameters that, put together in a model, will tell which kind of activity is best suit and, why not, the best restaurant you can find in the resort for your needs. This is the long term goal of the project, scheduled for the summer of 2016.

In order to put these objectives into practice, we plan to create our mirror device prototype, starting from existing consumer hardware (Kinect 2.0 and a wide LCD screen). Considering the physical state parameters, we plan to include those having a low or medium cost, without any contact and with a low user impact in the measurement. With respect to the evaluation of the prototype, we plan to validate the Virtuoso results in a holiday resort scenario. Being this setting one of the most important for wellness in both research and industrial effort, we expect to collect important insights on the effectiveness and the acceptance of the approach. We plan to perform both technical test and evaluations with end users. On the technical level, the evaluation should address two points. The first one is assessing the reliability of the physical insights collected through our prototype with ground-truth data. In order to do this, we will evaluate a group of people whose physical state has been already assessed with proper diagnosis techniques and we will compare the results. The second point is related to the deployment of the entire solution in a real resort setting, considering hardware, software and the staff training.

Last but not least, we plan to evaluate the user overall usability of the approach through a long-term study. We will deploy the system in a well-known resort in Sardinia, and we will collect both qualitative and quantitative data. In particular, we plan to apply the SUM model [147] for combining different usability metrics into a single score.

In order to collect the data, considering that it would be annoying for resort guests to complete questionnaires, we plan to instrument the software for tracking task completion, time, and errors. With respect to the post-task satisfaction ratings, we will provide a playful interface for collecting the answer, e.g. punching the rating number as physical exercise. Such evaluation method is convenient since it would be possible to automatise the data collection and to perform analysis at both a global and a task level.

Conclusion

In this work, we reported on the state of the art technologies for building smart mirrors and for sensing the user's physical status. We described smart mirrors devices and prototypes working as multimedia players, interactive home controllers and augmented reality devices. Moreover, we described and compared different parameters for evaluating the user's physical state, considering whether they require annoying or uncommon actions, the contact between the sensors and the user and the procedure cost. Finally, we introduced the scope and the objectives of the Virtuoso project, which aims to create a seamless interactive support for fitness and wellness activities in touristic resorts. We explained how, in future work, we are planning to create the project setup and to evaluate its results.

4.4 Interactive shops

In this section we present a proof-of-concept of an integrated system for enhancing the shopping experience in a shoes' shop. The system uses two modules: (i) one internal to the shop using an interactive totem; (ii) one external to the shop based on the customer's mobile device and the interactive external shop surface. We describe the technical architecture of the two modules and two different scenarios of the user experience.

The increasing spread of smart mobile devices generated a great deal of interest over the possible usage of mobile devices as remote controllers. There are a lot of studies in the home automation field, for instance, on the possibility to allow users to remotely control house appliances like doors, lights, TVs, music boxes and more [4]. Of course, the mobile devices have much more features than this, but their availability and versatility (they can use dedicated apps or connect to the appliance using a browser) make them a really good choice as controllers. At the beginning of this millennium, infrared ports for wireless data transfer were commonplace on many mobile devices and they have been used as remote controllers to manage TVs and similar home appliances [1].

Wi-Fi and Bluetooth (BT) connections are now used, instead, to achieve this kind of interaction, thanks to the fact that they are more versatile and radio-based, thus not influenced by obstructions. More recently, some controlling functionalities make use of new protocols like NFC (Near Field Communication) or RFID (Radio-Frequency IDentification). It is common to use programmable boards in order to connect mobile devices and appliances since, except for the most recent ones, they do not have any native wireless communication capabilities.

Figure 4.8: The iTotem installation inside the shop.

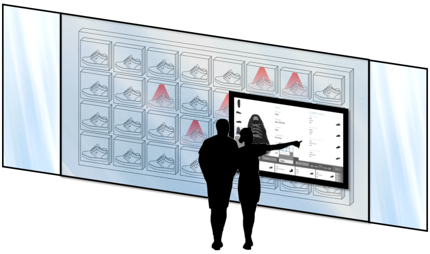


Figure 4.9: The customers interacting with the showcase from outside the shop.

Our main idea was to apply the knowledge built in this field to enhance the shopping experience of the customer having a smart mobile device with the goal of making the shopping easier and funnier. At the same time we develop the proof-of-concept for sellers that give them more options and tools. This system allows sellers to attract new customers by offering them new ways to discover the shop's products, improving their shopping experience.

We present a prototype for a footwear store based on this approach. In this setting we envision two main interaction areas: inside the shop using a dedicated area (see Fig. 4.8) and outside the shop through the showcase using a smartphone (see Fig. 4.9). The former allows users to get all the information about a chosen pair of shoes by placing it on a dedicated platform (that we named iTotem, see Fig. 4.8). The latter takes place outside the shop, allowing users to make simple and fast searches in the shop's catalogue using their mobile device (or a touchable surface) by interacting with the products placed in the showcase (see Fig. 4.9).

In the last years, many commercial brands are creating dynamic showcases that show videos, change the lights' colours or even move the exposed products. In some shops, users can interact with the environment by exploring the catalogue through multi-touch projected surfaces or displays, but they can not change the shop's appearance.

This is our main contribution: we want to give users/customers a new way of interaction, even when the shop is closed, by allowing them to explore the shop's catalogue and change the store's appearance.

Setup discussion

We describe here the guidelines of the system development. The graphical user interface is divided in five areas (see Fig. 4.12 for label coding):

A block In this section we display all the images and videos of the chosen shoe. We used some photos taken from different points of view and a 360 degree footage when available.

B block This section contains the user selected item.

C block Here we show all the shoe's details, like brand, colours, category, and whether the shoe is for males or females.

D block Here the user can scroll other shoes that are related to the selected one through this list view.

E block In this area there is a pop-up tab that allows users to search other products by looking in the shop's catalogue choosing by category, brand, size, colour, gender and a price interval. Once the user sets one or more filters the results are listed in the same tab and, by clicking over one of them, the page's content will refer to the new selected pair of shoes.

When nobody interacts with the system it remains in idle mode and awaits for an RFID reading or a touch event. In fact, the iTotem (see Fig. 4.8, Fig. 4.10) is equipped with an RFID reader that allows the users to get more information about the shoe by placing it on the platform. Every shoe is tagged with an RFID tag and each tag ID links to the shoes' details saved in the database. In stand-by mode it shows a sort of screen-saver explaining how to start the system (i.e., place a shoe over the iTotem). As soon as this happens, the system shows the main view with the five areas as explained earlier.

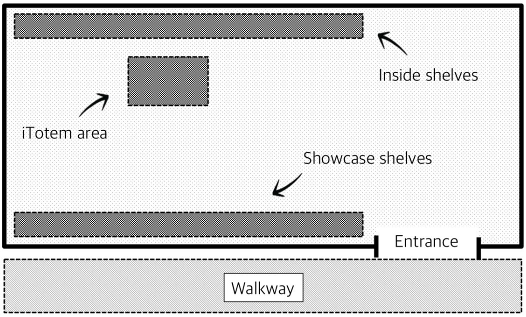
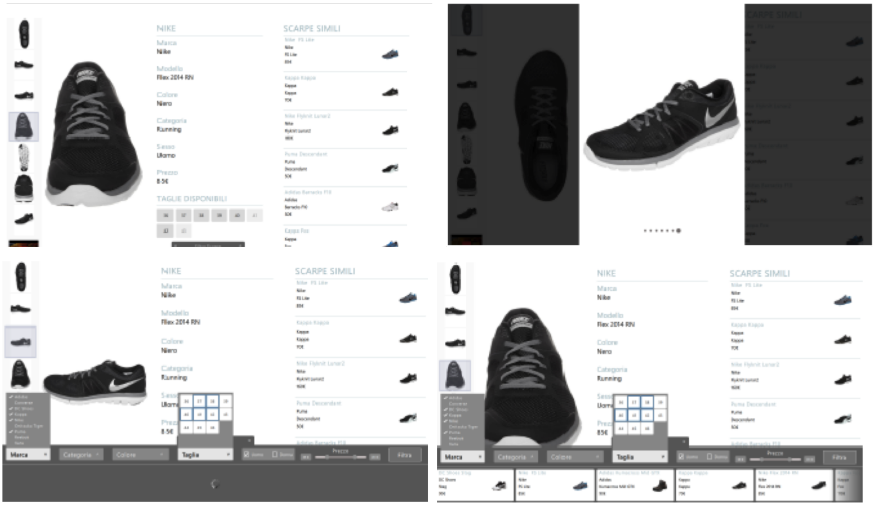
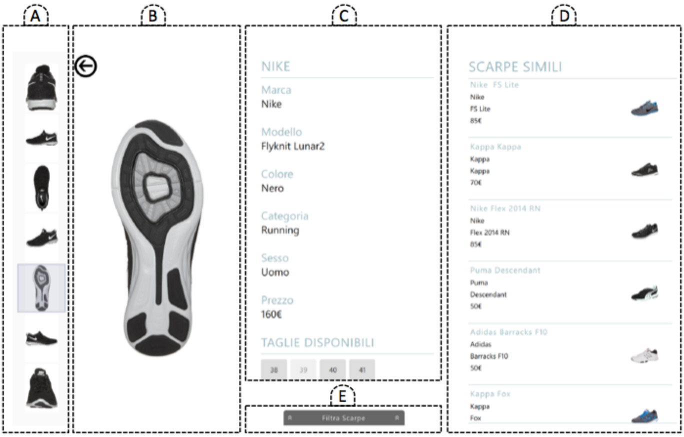


Figure 4.10: Possible plan of the ideal store where the system is installed.



(a) Shoe's information (b) Selected image (c) Pop-up filter tab (d) Filtered results

Figure 4.11: A few screenshots of the running application.



NIKE
Marca
Nike
Modello
Flyknit Lunar2
Colore
Nero
Categoria
Running
Sesso
Uomo
Prezzo
160€
TAGLIE DISPONIBILI
38 39 40 41

SCARPE SIMILI
Nike FS Lite
Nike
FS Lite
85€
Kappa Kappa
Kappa
Kappa
70€
Nike Flex 2014 RN
Nike
Flex 2014 RN
85€
Puma Descendant
Puma
Descendant
50€
Adidas Barracks F10
Adidas
Barracks F10
50€
Kappa Fox
Kappa
Fox

Figure 4.12: The layout of the UI of the interactive screen.

First scenario: inside the box

Alice is looking for a new pair of shoes and she is walking in a shopping street when she sees a footwear shop. She enters the store and looks for the exposed models and immediately sees a display that informs her to take a shoe and place it on the platform (the iTotem) in front of the screen.

Alice takes a sneaker and places it on the iTotem. The shoe is now recognised by the RFID reader and the display starts to show the shoe's information as explained in Fig. 4.11(a).

Alice can now watch the shoe's images and she is able to horizontally scroll among them as showed in Fig. 4.11(b). She can also watch information such as the price, available sizes, materials and even take a look at other similar shoes (see D-Block in Fig. 4.12). In the same moment Alice sees that the shoes' shelves change their appearance. The lighting system reflects the user's interaction, and, due to Alice's selection, only the shoes she is interested in are lightened by the spotlights.

In this way Alice immediately sees which shoes are the ones listed in the "related shoes" section on the display (D-Block in Fig. 4.12). She now wants to look for cheaper shoes; to do this she uses the filters' tab by tapping over it as explained in Fig. 4.11(c), then she selects a price range between 30 and 60 euros and her number. After pressing the filter button both the screen and the shelves change their appearance. The display shows the result of her search as showed in Fig. 4.11(d) while the shelves spotlights are illuminating only the shoes that correspond to the applied filter (a mockup example of this result is in Fig. 4.8). Alice finds the right pair of shoes, and she can finally try them on.

Second scenario: outside the box

Bob is looking for a new cheap pair of shoes but unfortunately his favourite store is still closed. He is looking at the showcase where all the shoes are exposed when he sees a QRCode (Quick Response Code) placed on the shop's window. This code allows customers to connect their smartphones to the interactive showcase. He sees all the exposed shoes in the smartphone's screen, in this way he can retrieve information such as the price, available sizes, available colours and more. He can also use a simple filter in the same way as discussed in the first scenario.

Once he chooses a set of filters, the shelves' lighting system changes the spotlights' states, by illuminating only the shoes that correspond to the selected filter (a mockup example of this result is in Fig. 4.9). Now Bob knows that some interesting shoes cost just under his budget price and are available in the store. When he will be able to visit the store he will ask directly for those shoes, thus saving some time. On the contrary, if no shoes are illuminated he will look in another store, always saving his time. In case the touchable window, where the same interface discussed in the first scenario, is not available, the user can use its own smartphone.

The system is divided in four main components: (i) the RFID manager; (ii) the Arduino controller; (iii) the Database agent; (iv) the GUI controller.

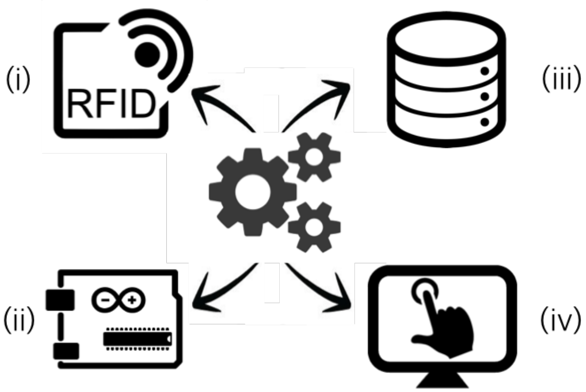


Figure 4.13: The architecture of the whole system.

All these components are controlled by the core component that can communicate with each single part that is hidden to the others (see Fig. 4.13). We made the whole system using different frameworks and programming languages. The main application has been developed using QT, a cross-platform application framework supporting different programming languages such as Java, Python, C and PHP. It includes a useful API that allows developers to build GUIs and manage external tools such as SQL databases, XML parsers, thread managers and more. We also used QtQuick, a collection of technologies designed to help developers for the creation of intuitive, modern, and fluid user interfaces working well especially on mobile devices. In order to make the system work we set up a server equipped with Apache, MySQL and the PHP engine.

The iTotem allows the users to get more information about the shoe by placing it on the platform by using an RFID reader that requires every shoe to be tagged with an RFID tag.

Each ID links to the shoes' details saved in the database. These details are displayed over a touchable surface (in our prototype we used a 24-inch multi-touch display) and, through it, users can perform the tasks described earlier. To identify the shoes equipped with an RFID tag, we used an *ID Innovations* branded reader (ID-12LA) connected through a USB module (SEN-09963) of the *Sparkfun Electronics* that is able to detect a tag from a distance lower than 12 cm (4.7 inch); the tags used are 125 kHz passive cards.

The USB module is directly connected to the PC where the main application runs (in our prototype this PC corresponds to the main server), and it works as an input device, while through another USB connection the application sends commands to the micro-controller board that manages the shelf's lighting system.

For our projects we used the Arduino board that is an open-source physical computing platform based on a simple micro-controller board, that came with a development environment for writing software for it. In detail, the light system is made up of 32 LEDs that are disposed as a 4x8 matrix and through the Arduino it is possible to turn on/off each single LED. This prototype represents a shelf of shoes and each LED simulates a spotlight. As previously said, in our prototype we used a 24-inch multi-touch display (DELL P2714T) that supports up to ten touch points. Even though in a real scenario this should be replaced with a multi-touch wall, thanks to the application responsive design, it runs well in both scenarios.

Conclusion

In this work, we described a proof-of-concept of an integrated system for enhancing the shopping experience in a shoes' store offering customers two different interactions inside and outside the shop. We are currently studying an approach that allows a multi-user usage of the system since, at the moment, it can be controlled by a single user decreasing the customers' engagement. We are also planning a real test with an Italian brand in order to understand the customers' feelings and to evaluate the user's experience when interacting with our proposed system. Nowadays, as more shopping moves online, clothing and footwear shops are trying to catch customers' attention by improving their buying experience with new technologies that are more interactive like smart mirrors or touchable displays. We think that these smart devices could change the way customers shop, and our contribution make a step forward in this direction.

Chapter 5

Having fun with mobile devices

Given the importance of mobile devices and their utility as learning tools, we developed four projects that implement well known algorithms that are often used on desktop environment. In these works the main challenge is the constraint of such devices due to their computational power.

Photo credits: Japanexperterna

Each work has its own distinct characteristics, but some of them are similar in terms of used techniques and approaches. Two of these projects, Chromagram and Click and Share, exploit two algorithms that make use of image processing techniques. This kind of algorithms are computationally expensive, especially on mobile devices and we here propose their implementation and their evaluation in terms of computational times. The third and the fourth projects aim to offer geo-located information, through augmented reality contents. They are both thought especially for tourists.

In fact, VisitAR allows users to have a more immersive sightseeing experience by providing tourist information in a more entertaining way.

Similarly, AR-Garden is a navigation system for limited extensions inside urban areas which permits to wander around and gives access to related information by using augmented reality techniques.

The section [5.1] is a revised version of the paper [36].
The section [5.2] is a revised version of the paper [29].
The section [5.3] is a revised version of the paper [25].
The section [5.4] is a revised version of the paper [111].
The section [5.5] is a revised version of the paper [17].
The section [5.6] is a revised version of the papers [155] and [156].

5.1 Chromagram

Chromagram is a mobile application prototype that allows users to apply a chroma key-based effect to the video stream coming from the device camera. This prototype has been developed on the Android platform but the same algorithm can be also implemented on a different mobile operating system. We discuss the algorithm used to achieve the Chroma Key effect focusing on the computational performance and on the quality of the final result.

The chroma key effect is a well known technique widely used by television and motion picture industries for image composition. In detail, this technique allows users to replace an entire area of an image (or a set of frames) which are identified by a specific colour with another image or video stream. This substitutive content can be easily selected from the device gallery, and it can be a video or an image.

Figure 5.1: Use of the blue of the sky as key colour.

Nowadays the Keying effect plays a very important role in the film industry field, and it is also widely used by amateur video-makers and photographers during the post-processing image composition to generate visual effects (see Fig. 5.1 for a simple example).

The main drawback of this technique is that it has a high computational cost, however this negative aspect is not really significant thanks to the increasing of the processing capabilities of both mobile devices and computers in general.

In the last years, the improvement of mobile devices cameras offers to users the possibility to take HD pictures and videos and we think that such kind of effect should be available among the built-in effects in the OS features (i.e. black and white effect, sepia effect and such).

Due to well known mobile devices constraints, we need an efficient, easy-to-implement, fast and cheap method for the real-time chroma key matting effect computation. We need to minimise the computational cost whilst achieving a good visual result too.

In this study we try to understand how much the Android platform, and in general a mobile setting, is suitable for the real-time image processing. It is clear that this effect can be achieved by using different techniques, and due to this we studied several keying (and matting) techniques to understand which one of these could be the best to use in a mobile setting.

As we said in the related work section, the most used techniques are:

- **Bayesian** [33]
- **Poisson Matting** [161]
- **Luma Keying** technique [19]
- **Difference Keying** [177]
- **Depth Keying-based** [32, 63, 85, 86]
- The **3D Keying** [133, 148]

In this work, we chose the 3D Keying technique for three main reasons: the quality of the results, the good computational efficiency and, especially, its easy implementation. These aspects allow us to evaluate the performance of the tool on the Android platform with a real-time image processing test.

Application design and User Interface

We designed a very simple interface, allowing the use of all the implemented features in an easy way without additional texts or tips (see Fig. 5.2).

The Background Substitution effect can be done by simply touching the camera preview. Through this operation it is possibile to select the entire image section characterised by a key colour. Then, using the side menu buttons, we can choose the substitutive background photo, edit the algorithm threshold and smoothness or take a screenshot of the applied effect.

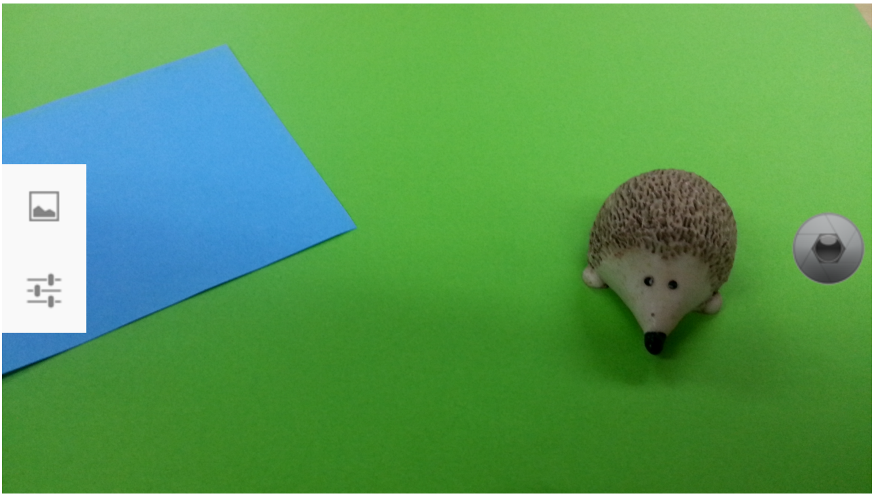


Figure 5.2: UI screenshot.

How it works

The basic idea to obtain the background substitution effect is to find and replace all pixels in the video stream that are characterised by a colour within a range. This can be achieved by using a key colour, previously chosen by the user, and applying a certain tolerance level (in Fig. 5.3 a single-coloured piece of fabric is placed in the scene for these purposes).

In order to implement the algorithm we used the OpenGL ES 2.0 graphic library. In fact, every frame obtained through the device camera will be computed and processed through the OpenGL fragment shaders thus obtaining the background substitution matting effect. Then it will be possible to render this frame on the screen using it as a texture of a plane. For every image texel of the frame obtained through the camera we can apply the matting effect through the OpenGL fragment shaders.

The algorithm used in this application makes use of the 3D keying technique as described in [148] where they represent all the pixels in a R^3 Euclidean space where the three dimensions represent each of the primary colours: red, green and blue.

Given a point k representing the chosen key colour, that will be replaced with an alternative background, by using two spheres S_1 and S_2 centred in k (with different radius), it is possible to define the transparency of every image pixel (or texel).

Figure 5.3: Using a green screen to achieve the final effect.

Every pixel p placed inside the inner sphere S_1 will be fully transparent while the pixels placed between S_1 and the outer sphere S_2 will have a proportional transparency to the distance from S_2: closer to the inner sphere S_1, higher transparency value; closer to the outer sphere S_2, lower transparency value (see an example in Fig. 5.4).

In details, for every pixel, to calculate the distance of this pixel p from the key colour k, we use the Euclidean distance:

$$distance = \sqrt{(p_x - k_x)^2 + (p_y - k_y)^2 + (p_z - k_z)^2}$$

After this calculation, if *distance* is lower than the radius r_1 of the inner sphere S_1, then the α opacity value of the pixel will be equal to 0.0 (fully transparent), else if the *distance* is higher than r_1 but lower than the radius r_2 of the outer sphere S_2, the α value will be calculated with the following formula:

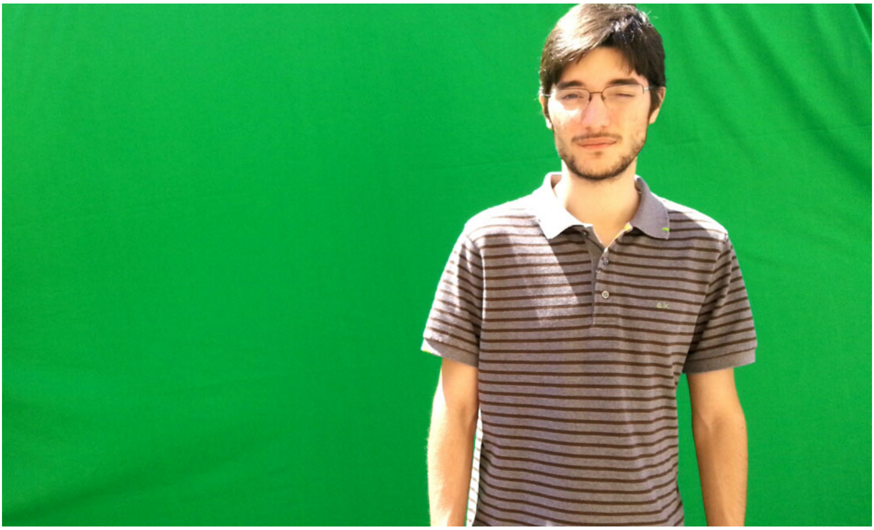
$$\alpha = (distance - r_1)/(r_2 - r_1)$$

The final pixel p_f value will be calculated as linear combination of original pixel p_o and substitutive image pixel p_s with α as the scalar value, as represented in the following expression:

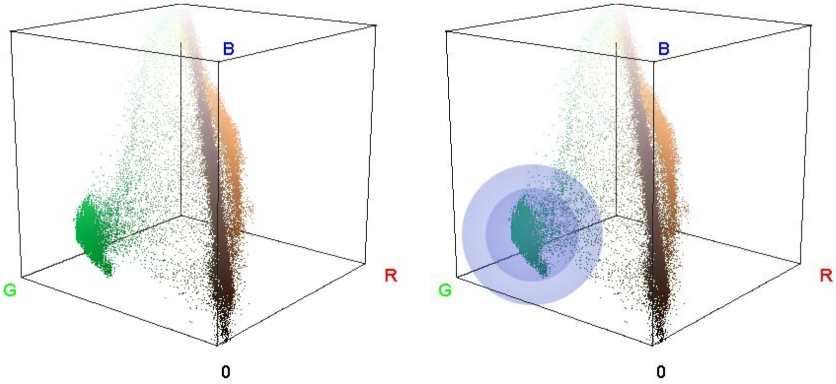
$$p_f = \alpha p_o + (1 - \alpha)p_s$$

So, the r_1 radius represents the threshold of the algorithm and the r_2 radius shows the smoothness.

As we said before, in this study we tested two different approaches.



(a) sim



(b) histo1

(c) histo2

Figure 5.4: An example of how the 3D Keying effect works: given an image (a), this technique represents all the pixels in the RGB space (b), separating the background from the foreground using two spheres (c).

The former by using the OpenCV and the latter by using the OpenGL library. We tested both approaches on the Android platform by using the same device obtaining very different performance results.

Figure 5.5: Color spill around the girl's hair.

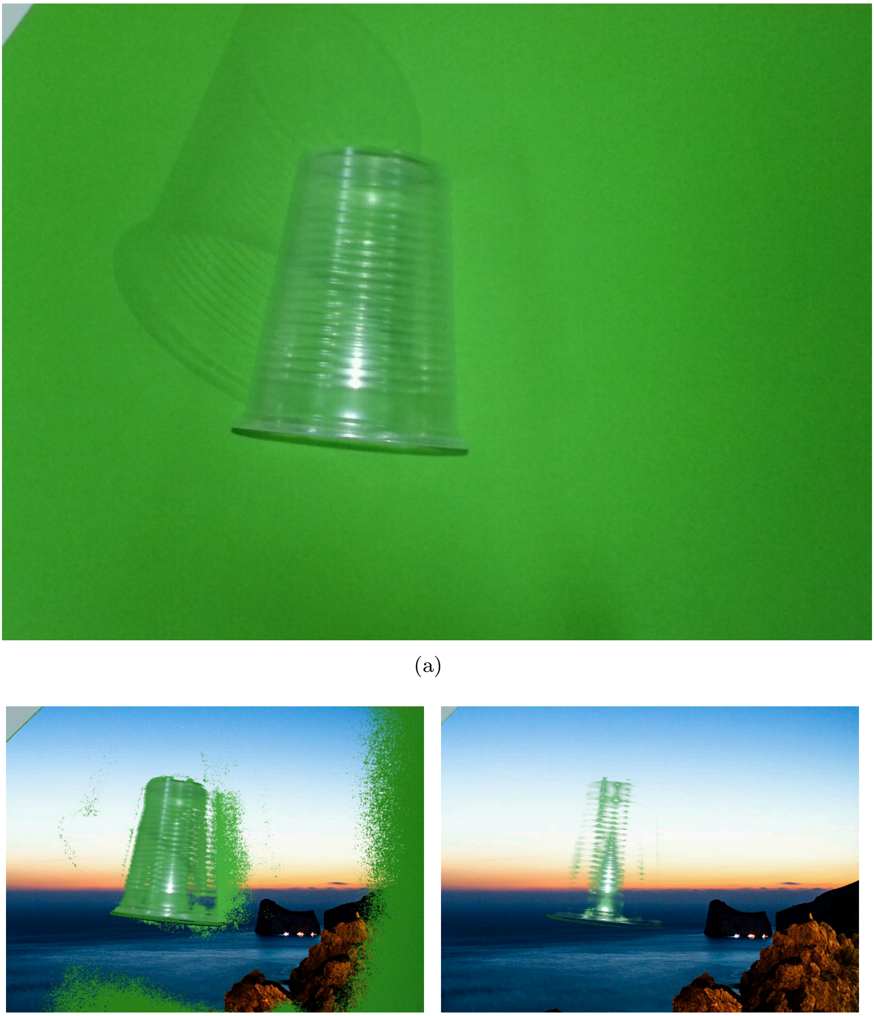
In Table 5.1 we reported the performances of both applications on several devices with different resolutions and processing power.

Device	**Resolution**	**OpenCV**	**OpenGL**
Sensation XE	800x480	15 fps	18 fps
Galaxy S3	1280x720	15 fps	30 fps
Galaxy S4	1920x1080	8 fps	30 fps

Table 5.1: Performances measured on Samsung Galaxy S3, Samsung Galaxy S4 and HTC Sensation XE.

Even if the test results seem promising, the chroma key matting effects have several issues regarding the scene illumination, the video compression and the camera quality. Other minor issues are related to the limitation of the algorithm used for the matting effects.

In particular, we might see some reflections on foreground objects that have the same value of the key colour (this effect is called color spill) caused by one of these issues or their combination: low smoothness value of the algorithm, bad scene illumination, excessive color reflection of the chroma key screen, and lastly, the materials that characterise the foreground objects (see Fig. 5.5).



(a)

(b)

(c)

Figure 5.6: An example of semitransparent object (a). With a small threshold, we preserve the glass details (b) but we still see some of the back key colour. Increasing the threshold (c), we lose glass details.

Implementing a spill removal algorithm is a good way to solve this problem, but this way decreases the application performance and it requires a good setup of the studio scene.

Another issue that we might encounter is the illumination difference between foreground objects and substitutive background, that produces unrealistic results of the final image composition.

A last noticeable issue is that we can lose small details of foreground objects, like hair for framed persons, or semitransparent objects made with materials like glass (see Fig. 5.6).

At the moment, to overcome this kind of issues thus obtaining good results, we have to correctly set up the shooting scene. It could be achieved by following some hints: use independent lights for foreground and background objects, illuminate the shooting scene as uniformly as possible, use foreground and background colours with a good contrast.

In the next future we plan to add new features, allowing the recording of video in addition to the images. We want also to improve the chroma key algorithm using a Bayesian or a Poisson matting filter that provides better visual results, solving some of the issues previously described. We are also working to decrease the complexity of the algorithm, preserving the computational performance results.

5.2 AR-Garden

In this section we describe a navigation system that allows visitors to discover cities and their urban areas easily and quickly. Its main goal it to find the shortest path to reach a given spot thus giving directions to the user saving time and energy. In detail, in this work we tried to set up a navigation system for small urban areas which permits to wander around them thus giving access to geo-located information by using augmented reality techniques.

Nowadays, the Google StreetView service offers the possibility to discover interesting places (indoors and outdoors) around the world by using Google Maps. It also offers a navigation system based on the same main service: Google Maps. Google, and many others, are still working on navigation systems for indoor environments. Even if there are a lot of available solutions for indoor and outdoor navigation systems, we think that, in the middle of them, there are other opportunities that require to be investigated.

These opportunities are given by the small urban areas that are too small for common navigation systems and too big for the indoor navigation ones. This areas are often places of touristic interest. Take as example the botanical gardens, the archeological sites or even the protected natural parks.

A Navigation system is usually based on the street maps given by the major mapping services such as Tele Atlas [8], Navteq [119] or OpenStreetMap [127]. Google StreetView offers its own navigation system but it also provides the possibility to *enter* inside the map and actually discover different city places that are both indoors and outdoors. Among these places we can not find the areas that are not accessible to the Google crew although they could be really interesting for tourists.

In these sites, visitors can discover the area thanks to flyers and guide books that usually include a map with the points of interest (POIs). It is not unusual that also e-guides (electronic guides) are available to tourists for their mobile devices. As we said before, the aim of this project is to offer a system as a result of a combination of two main features: the former is to give tourists an augmented reality navigation system for such kind of urban areas, and the latter, to give them an e-guide that, thanks to the GPS position, shows the POIs information when a tourist reaches one of them. All the necessary data are retrieved from a web server thus requiring a single download at the very first use of the application.

In tourism related applications, the augmented reality is often used to show information related to a city or a monument. There are basically two types of augmented reality:

- The first is the geo-localised AR, usually used in mobile devices that are equipped with a positioning system and a set of motion sensors; the surrounding environment is enriched with additional information through the screen device in correlation to the user's position;
- The second is based on augmented reality tags. It works on both mobile devices and desktop computers by detecting and recognising a specific marker (usually drawn with black and white geometric shapes); by recognising these markers it is possible to show digital content like photos, videos, 3D models even adding interactivity with them, thus improving the user experience.

In our case study we used the first approach by using the GPS and the compass sensors.

Application design and User Interface

As a proof of concept we created AR-Garden, considering the botanical garden of Cagliari as case study. After a very first configuration, the application allows users to navigate in the botanical garden sectors autonomously. The application provides information like sectors, plants species, points of interest, thus giving the directions necessary to reach them. Its main goal is to simplify the tourist visit keeping track of the user's position. As a common navigation tool, users have to choose a destination point and, starting from the user's current position, the application shows the path necessary to reach the chosen sector. One of the main challenges is to achieve results similar to those offered by the car navigation system but with the constraints that streets are smaller, POIs are closer to each other and the GPS error could be very high in correlation to the distances taken into consideration. This is really important in our application scenario, that is represented by the small urban area. In detail, we define as "small urban zone" those areas such as botanical gardens, parks or whatever part of the city that have these characteristics:

- are dedicated to pedestrians and/or cyclists;
- can be mapped on a 2D map;
- are smaller (even much smaller) than 20 square kilometres;
- have interesting spots for tourists;
- are located outdoors.

We designed the interface as minimal as possible. By using the starter screen it is possible to access all the application features. The main feature is the navigation one that allows the turn-by-turn navigation. In the lower part of the screen a 2D draggable map is shown; over it are displayed the user position, the final destination and the path that allows its reaching. While, in the upper part of the screen, the camera preview is combined with augmented reality 3D elements that give the directions for reach the destination point. The user can customise the size of both parts depending on her preferences (see Fig. 5.7).

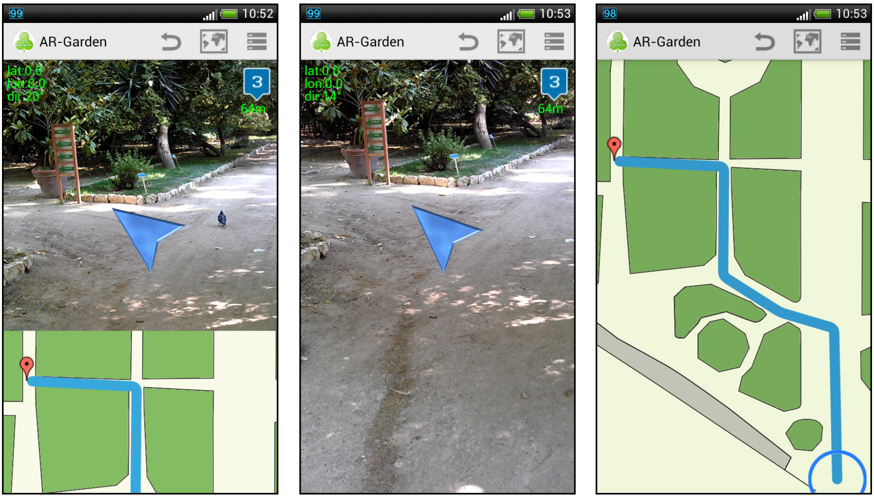


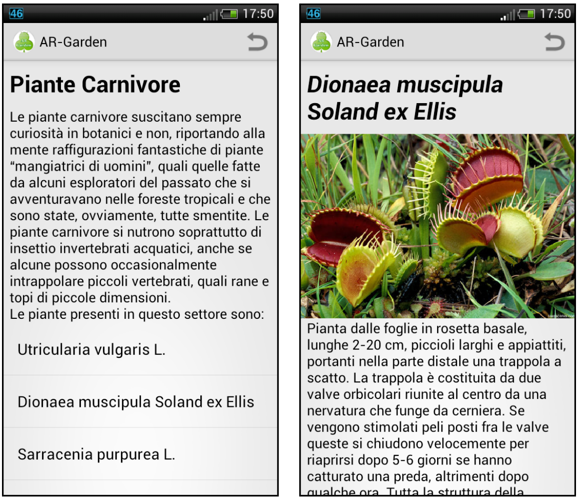
Figure 5.7: The three layouts for navigation: combined (left), AR only (center), and map only (right).

The other tabs allow users to see the sectors' information (in this case study, the plant information) (see Fig. 5.8).

The information of each plant regards the scientific name, the genre, the family, the habitat, the flowering period and a representative image.

How it works

As we said before, we used a 2D map overlaying on it the path to follow in order to reach the destination point. This map has an overall resolution of 2000×4205 pixels, that allows an optimal visualisation on devices with different screens sizes. The map viewer supports the common touch gestures such as drag and pinch-to-zoom. In order to find the best path we need to build a graph structure that represents the map, where each point of interest is represented by a node as well as the cross between two paths.



Piante Carnivore

Le piante carnivore suscitano sempre curiosità in botanici e non, riportando alla mente raffigurazioni fantastiche di piante “mangiatrici di uomini”, quali quelle fatte da alcuni esploratori del passato che si avventuravano nelle foreste tropicali e che sono state, ovviamente, tutte smentite. Le piante carnivore si nutrono soprattutto di insettio invertebrati acquatici, anche se alcune possono occasionalmente intrappolare piccoli vertebrati, quali rane e topi di piccole dimensioni.
Le piante presenti in questo settore sono:

Utricularia vulgaris L.

Dionaea muscipula Soland ex Ellis

Sarracenia purpurea L.

Dionaea muscipula Soland ex Ellis

Pianta dalle foglie in rosetta basale, lunghe 2-20 cm, piccioli larghi e appiattiti, portanti nella parte distale una trappola a scatto. La trappola è costituita da due valve orbicolari riunite al centro da una nervatura che funge da cerniera. Se vengono stimolati peli posti fra le valve queste si chiudono velocemente per riaprirsi dopo 5-6 giorni se hanno catturato una preda, altrimenti dopo qualche ora. Tutta la struttura della

Figure 5.8: The information on a single sector (left) and on a single plant (right).

Furthermore, the arcs between them represent the connections between the points. Thanks to the Dijkstra's shortest path algorithm, it is actually possible to highlight the best routes between two points. By using an appropriate geometric operation, the user's position can be plotted over the map. In our implementation we use the map as a Cartesian plane, transforming the previous (A point) and the current (B point) GPS position in XY coordinates on it. The distance between A and B and the angle between the straight line passing through the two points and the X-axis are computed every position update. To find the new position we calculate Δ_x and Δ_y and we add (or subtract) them appropriately to the current position. In particular $\Delta_x = \text{dist} \times \cos\alpha$ and $\Delta_y = \text{dist} \times \sin\alpha$, where *dist* is the distance between the two geographical positions transformed into pixels, and α is the angle between the straight line passing through the two points and the X-axis calculated using the following formula (where Lat=latitude, Lon=longitude):

$$\begin{aligned} dy &= \text{Lat}(B) - \text{Lat}(A) \\ dx &= \left(\cos\left(\frac{\pi}{180} \times \text{Lat}(A)\right) \times \text{Lon}(B) - \text{Lon}(A)\right) \\ \alpha &= \arctan(dy, dx) \times \frac{180}{\pi} \end{aligned}$$

The following step consists in recomputing the orientation of the arrow (that indicates the way the user should follow to reach the destination) based on the last position update. This is done by considering the current position, the target position and the user's orientation in correlation to the destination. We calculate the angular coefficient m of the straight line passing through the points A and B where A represents the current position and B is the destination point.

$$\begin{aligned}
\Delta_x &= x_A - x_B \\
\Delta_y &= y_A - y_B \\
m &= \frac{\Delta_x}{\Delta_y}
\end{aligned}$$

Δ_x or Δ_y equal to 0 means that the user is moving along one of the Cartesian axes otherwise we find the angle of rotation of the arrow calculating $\arctan(m)$.

The AR-Garden application has been tested in two steps: the former by emulating the user position and the latter by trying it directly on-site. To emulate the GPS signals we used an Android emulator by using a GPX file.

In the first simulation, the provided directions have been shown correctly as expected. On the other hand, testing it on site we obtained different and not reliable directions. The well known error in GPS coordinates is really significant and it causes wrong directions. The main problem is the signal strength, since the signal acquisition does not work very well in covered or partially covered areas. This leads to high errors in the position calculation during the visit of the area. In our case of study, the Botanical Garden of Cagliari, we have encountered this problem because it is almost totally covered by trees forming a barrier between satellites and the device. We have estimated that this error is around 20 meters in the worst cases. This implies that the navigation based on GPS coordinates is not enough to offer a good turn-by-turn navigation in such kind of areas.

5.3 Click and Share

Click and Share is an Android application for mobile devices that allows users to quickly and easily identify faces in pictures, by recognising people, and, thus, sharing pictures with them. Each identified person matches against a contact registered in the phone directory, and, if no match is found, the detected face can be used for the creation of a new contact.

In current popular culture Facebook is considered a wonderful and very powerful tool. As the name itself suggests its main feature is “to put oneself’s face”, not only metaphorically, but also in a proper meaning. This is one of the main reasons for its great diffusion and, as a consequence, people continuously use it for sharing close-up pictures with friends. This is also true for other social networks such as Twitter or Instagram. In this context, simply using the tools can be a factor of self-assessment, as several studies show (e.g., see [186]). That is why we wanted to simplify the task of sharing pictures with friends over social networks, using the possibilities given by powerful techniques of face recognition, now available also on mobile devices, aiming to enrich the user’s experience and satisfaction.

Application design and User Interface

We firstly describe the description of its practical usage through a scenario. Sara, Fabrizio, Davide, Giammy and Alessandro met each other at the University of Cagliari, since they attended the same Image Processing course. They became friends during the classes and they decided to visit Paris in the summer. Sara loves taking pictures, and she never misses the opportunity to save a souvenir on her mobile phone. When a photo is particularly beautiful, her friends always ask Sara to share photos with them: Fabrizio sometimes desires to have a picture posted on his Facebook wall immediately, Alessandro would like receiving photos via email since he does not use social networks for privacy issues, Giammy prefers using Instagram, while Davide only has a Twitter account. After coming back to Cagliari, they watch the photos remembering the places they visited and the fun they had during the trip. Again, Sara’s friends ask her to share other pictures that they missed in Paris.

The scenario describes two main contexts where sharing contents in a mobile setting is needed. In the first one, the user takes the photo, she looks at it and she decides that it is good enough to be shared with someone (e.g. through a social network) immediately after. The second one envisions a higher distance in time between the moment when the photo has been taken and the sharing action, which may be stimulated by other people or events.

The purpose of *Click and Share* is to provide a smartphone application that allows users to share photos quickly and easily, either if a picture was taken immediately before using the built-in camera or if it was stored into the smartphone's memory.

Our idea for speeding up the sharing task while maintaining a rich set of channels for delivering the content, is to associate the different photos to the contact information already contained in the mobile phone, by using a face recognition algorithm. In this way, the application is able to recognise which contacts are present in a given photo and to exploit the contact information for sharing (e.g. the email address, the Facebook user name etc.). This approach limits the need for further user input, which may be tedious and slow on mobile devices.

The main application screen presents two options: the first allows the user to take new pictures from the built-in camera, the other one allows to load an existing picture from the device storage.

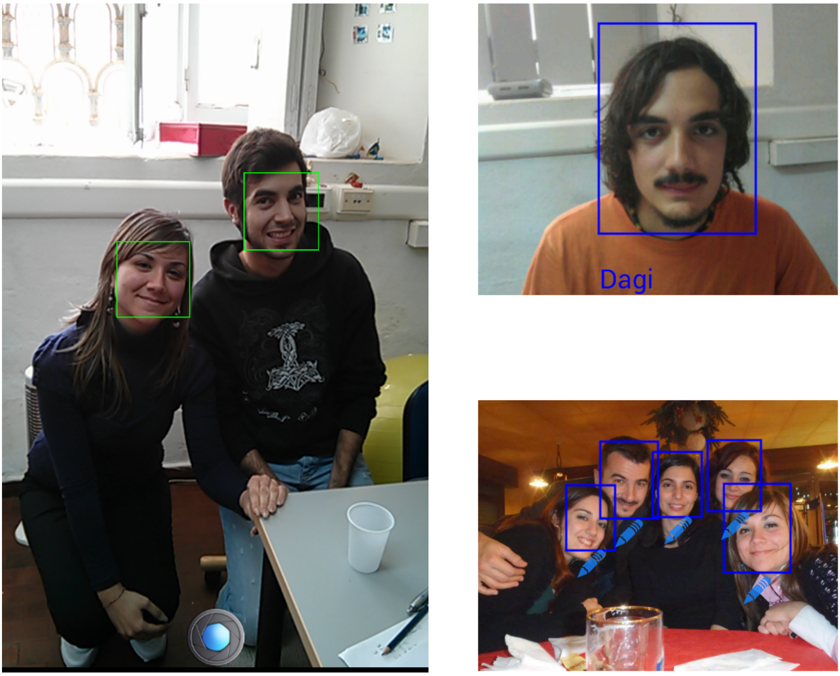
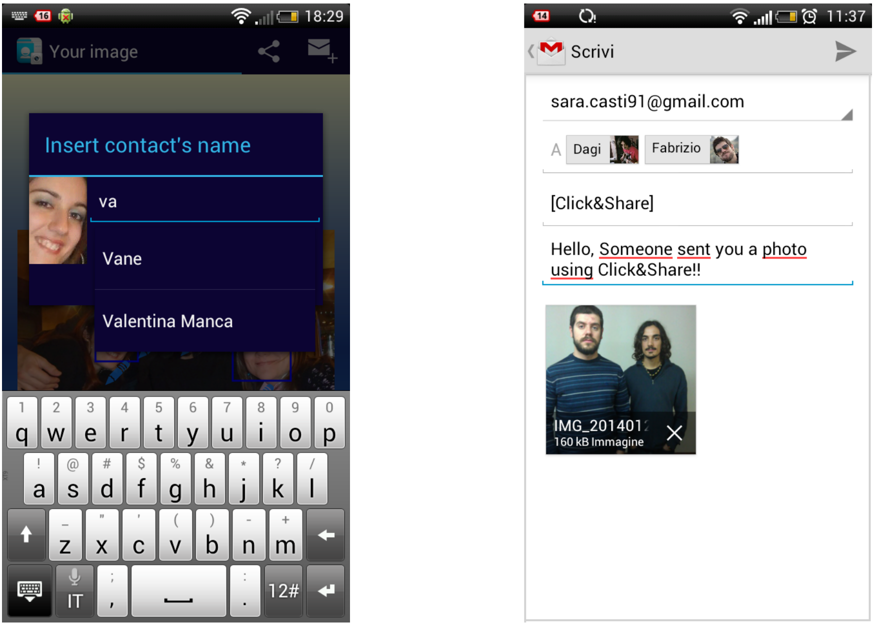


Figure 5.9: Face detection (left) and recognition (right).

If the user chooses to take a new photo, the application shows the camera preview, and a green-bordered square is shown in correspondence of all the faces detected in the photo in real-time (Figure 5.9, left). Tapping over the capture button, the image is saved on the smartphone storage, and the application executes the face recognition algorithm. In a similar way, if the user selects an existing picture from the gallery, the recognition process is started, when no face information is already associated to the selected photo. In both cases, if the algorithm is able to recognise one or more faces given the existing dataset, the application shows the contact's name below the face bounding box. Otherwise, the application shows a pencil icon, suggesting the need for the user's input. The face recognition UI is shown in the right side of Figure 5.9.



Insert contact's name

va

Vane

Valentina Manca

sara.casti91@gmail.com

Dagi Fabrizio

[Click&Share]

Hello, Someone sent you a photo using Click&Share!!

IMG_201401
160 kB Immagine

Figure 5.10: UI's for changing the identity of the recognised person (on the left) and for sharing a photo via email (on the right).

The user can manually specify the correspondence between the face and an existing contact either by tapping on the pencil icon, if the person was not recognised, or on the person's name suggested by the application, if the recognition was wrong. The interface allows the user to select an existing contact name from a filtered list, which shows the names that correspond to the user's input (see Figure 5.10, left part).

Otherwise, she can create a new contact if the person was not registered before. The selection of the correct person trains the face recognition algorithm, in order to improve precision for further pictures.

Once the picture is analysed and the faces are recognised the user can share it with her friends via mail or social-networks. In case the user decides to send the pictures via mail, the application fills the receiver field automatically for each recognised person, if they have an email address registered in the contacts' records. It is also possible to add or remove further recipients through the usual mail interface. In addition, the application generates a default template for the message and, obviously, attaches the picture to it. If no modification is needed, the user only has to press the send button to deliver the mail (Figure 5.10, right part).

In addition, it is possible to share photos through a social network. For instance, if the user selects Facebook, the application automatically creates a wall post with the photo and the tags related to the recognised people.

How it works

The face detection is performed in real-time while the user frames the scene with the camera or when an existing photo has been loaded. This operation has been implemented through a built-in Android OS feature, which returns the bounding-boxes for the faces contained in the photo. The face recognition phase is based on the Local Binary Pattern Histograms algorithm [3], which has been proved to be tolerant to the changes of illumination, computationally simple and to have a good recognition performance even with a small training set [115]. All these aspects are crucial for employing the algorithm on a mobile device for sharing contents. First of all, the computation should consume as little resources as possible, not only for providing a real-time feedback, but also for saving battery and to work even on low-end devices. The tolerance for different lighting conditions is important since people have different pictures of the same subjects in different places and situations. The good precision even with a small training set allows the user to benefit from the face recognition even with a low number of photos, limiting the learning phase that is needed at the beginning of the application usage.

Even if the number of subjects that may be present in a photo collection may be high, we can assume that the number of people that need a high recognition precision are only a small part of the entire set. Indeed, people tend to share information on social networks with a small subset of their declared contacts (e.g. friends on Facebook) [57].

Therefore, it is likely that the user has already produced or will produce an adequate number of photos for this subset. The recognition accuracy depends on several aspects like ambient lighting, subject pose, facial expression and, of course, camera sensor characteristics; unfortunately these factors can degrade the result in a substantial manner. In particular, it is possible that a face contained in a picture is not detected or recognised, or that an extremely different object (i.e. a bottle) can be recognised as a face.

Therefore, the application needs to support the user in changing the recognised person or to remove a detected bounding box if needed. In order to estimate the time required by a mobile device for recognising a person, we performed a preliminary evaluation of the algorithm performance on two different devices: a HTC Sensation XE (CPU Qualcomm MSM8260, 1.5 Ghz) and a Nexus 5 (CPU Snapdragon 800, 2.3 Ghz). The two models represent respectively a former high-end, now average, and a high-end smartphone.

The algorithm has been tested against the well-known AT&T Database of Faces [22], which contains 400 photos representing 40 people. We split the test in two parts: we first evaluated how long it takes to train the algorithm with the considered dataset, repeating the process six times. From the application point of view, this is performed only once when, for example, the user buys a new phone. For each measure we report the mean, the median value and the standard deviation. The HTC Sensation XE took about 7s ($\bar{x} = 7.0, \tilde{x} = 6.9, s = 1.4$), while for the Nexus 5 2.5s were sufficient ($\bar{x} = 2.4, \tilde{x} = 2.4, s = 0.1$). In the second part of the test, we recorded the time elapsed for recognizing a person with the leave-one-out technique. We repeated the experiment 10 times for each one of the 40 people. The HTC Sensation XE needed about 260 ms for completing the recognition ($\bar{x} = 257, \tilde{x} = 241, s = 55.8$), while the Nexus 5 needed about 150 ms ($\bar{x} = 152, \tilde{x} = 157, s = 13.2$).

We can conclude that the time needed for the recognition is acceptable for all the phone categories we considered. Therefore, the technique may be already considered for a broad set of mobile phones, without any relevant delay for the user. In order to evaluate the improvements introduced by the approach described in this work, we performed a user test. The aim of the study is threefold: we want to assess i) the efficiency, ii) the errors and iii) cognitive effort with respect to the existing approaches for creating and sharing photos through mobile devices. We considered the Facebook mobile application as the baseline for the comparison, which allows the users to take pictures, to recognise people into them and to share the photos.

After completing a small demographic questionnaire, the participants performed a set of four tasks with both mobile applications. We alternated the initial condition (Click & Share and Facebook) in order to minimise the carry-over effect. The tasks to be completed were the following:

1. Take a picture of two subjects and identify the person into the photo;
2. Share the photo with the identified subjects via email and Facebook;
3. Select an existing picture from the mobile phone gallery and identify the people in the photo;
4. Share the existing photo with the identified subjects via email and Facebook.

At the end of each task, the users answered the *Subjective Mental Effort Question* (SMEQ) [183] in order to evaluate the cognitive difficulty of the task, while the moderator recorded the completion time and the number of errors during the task.

Eighteen users participated to the test, 13 male and 5 female, aged between 19 and 54 ($\bar{x} = 26.17, \tilde{x} = 22.50, s = 9.7$), 12 with a high-school degree, 3 with a bachelor and 3 with a PhD. They were highly experienced in using mobile devices (all of them use them at least once everyday) and in particular they had good experience with the Android OS ($\bar{x} = 3.22, \tilde{x} = 3.00, s = 1.26$, in a $1 - 5$ Likert scale). Most of them (16) use social networks everyday, one once a week and one once a month, while 10 of them share information with other people through the mobile device once a day, 4 once every three days and 4 once a week.

Figure 5.11 shows the test results. For each metric, we compared the means with a paired t-test, and we obtained the results in Table 5.2. The completion time was significantly lower for T1, T2 and T4. The minimum difference for all of them ranges from 23 to 36 seconds, which is significant if compared with the mean of the completion time for each task (from 26 seconds for T4 to 44 of T2). For T3, we registered no significant difference. However, in the worst case users spend the same time with Click & Share and Facebook. Click & Share is less error-prone for all tasks considered. The difference in the worst case goes from 1 (T3 and T4) to 2 more errors per task, which is again significant considering a mean of about 0 or 1 error per task. The cognitive effort is significantly lower for Click & Share for T1. The difference, according to [183], goes from the *Not at all hard to do* to *Not very hard to do*. For T2 we registered a significant difference in mean, but no change of difficulty level. We have no evidence of any difference for T3 and T4.

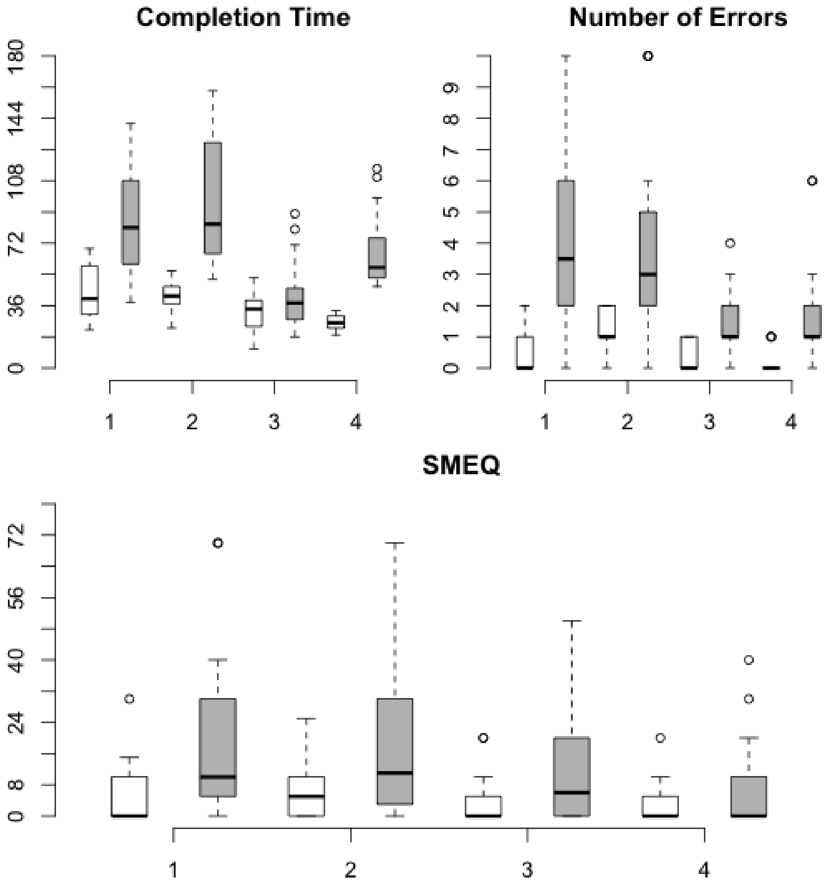


Figure 5.11: User study results for the completion time (in seconds), number of errors and the SMEQ. The white boxes represent Click and Share results, while the gray ones represent Facebook. The results are grouped by task.

Completion Time

Task	**Mean**	**95% c.i.**	**p-value**	**Sig.**
T1	44.3	$[23.3; 55.9]$	$8 \cdot 10^{-5}$	**Yes**
T2	41.5	$[36.4; 74.5]$	$1 \cdot 10^{-5}$	**Yes**
T3	34.0	$[-1.8; 20.00]$	0.096	No
T4	26.0	$[30.0; 50.6]$	$2 \cdot 10^{-7}$	**Yes**

Number of Errors

Task	**Mean**	**95% c.i.**	**p-value**	**Sig.**
T1	0.44	$[2.4; 5.2]$	$2 \cdot 10^{-4}$	**Yes**
T2	1.1	$[1.7; 4.6]$	$2 \cdot 10^{-4}$	**Yes**
T3	0.33	$[0.6; 1.7]$	$3 \cdot 10^{-4}$	**Yes**
T4	0.17	$[0.7; 2.4]$	$1 \cdot 10^{-4}$	**Yes**

SMEQ

Task	**Mean**	**95% c.i.**	**p-value**	**Sig.**
T1	5.3	$[7.0; 24.7]$	0.001	**Yes**
T2	7.22	$[4.2, 24.2]$	0.008	**Yes**
T3	3.6	$[-0.2; 12.2]$	0.06	No
T4	3.17	$[-0.5; 8.7]$	0.08	No

Table 5.2: Confidence intervals around the difference between the Facebook and Click & Share results (paired t-test). The mean is referred to the task performance with Click & Share.

In conclusion, the user test highlighted a difference between Click & Share and Facebook on both completion time and number of errors, while maintaining the same cognitive effort for reaching the user's goal. The prototype implementation performance shows that the approach is already feasible for being employed in mobile devices, while the user test shows an increase of the user performance from both the efficiency and the effectiveness point of view. In future work, we want to apply the same technique to a broader set of media and to allow the synchronisation of the data collected through different mobile devices.

5.4 VisitAR

In this section we present VisitAR, a mobile application that has two main goals: the first is to offer tourists a more immersive sightseeing experience, by providing tourist information in a more entertaining way; the second is to give them a tool for planning their visit. Both aspects are achieved by using augmented reality techniques. In fact, during the trip planning it is possible to frame the city map and see 3D content that is over-imposed on the map, while during the sightseeing, tourists can follow a specific route and, when they reach a city attraction, they gain access to several extra-contents like images, videos, cultural information and much more.

Application design and User Interface

In this work we used Junaio [83] that is an AR browser which has an augmented reality client application (freely available in the Apple App Store and in the Android Market), for smartphones and tablets. Our application is the result of a combination of two Junaio channels, the Glue channel (GC) and the location-based channel (LBC). The GC allows to see an interactive augmented reality map aimed to help the tourists in the sightseeing planning. On the traditional paper map, we over-imposed virtual balls, text and 3D models of the major attractions. These points of interest are presented in a way that allows users to freely customise the amount and type of attractions showed on the map.

The guided sightseeing was instead carried out using a LBC, which makes use of GPS and phone compass for the user tracking. The developed application interface has required a preliminary phase of study in order to create a simple and intuitive way to get information of the city. The customisation of the points of interest is made possible by using a filter among different categories such as monuments, restaurants, shops and more. Finally, when users reach a tourist attraction, they can see additional contents and have a more immersive experience in the place of interest. For example, they could see how the point of interest appears in different seasons or how it was before a particular event like a war or a natural disaster. This goal can be simply achieved by scanning a placed QR code by using the device directly in the tourist attraction. Users can benefit from the augmentation in different ways such as: a slideshow of pictures, a panoramic 360 degrees view or a 3D reconstruction layered over the device camera preview. To gain access to our application the user has to search it in the Junaio channels research page, or simply by scanning the associated QR code that is generated directly on Junaio web site. Both the GC and LBC we created, are dynamic channels.

The former uses the touch input while the latter uses the user's position.

How it works

In the GC, the tracking is performed through the use of either a picture (Natural Feature Tracking, NFT) or an ID marker. During the environment scanning by using the phone camera an image processing algorithm recognises the tracking pattern and estimates, in real time, the camera pose. In this way a correlation between the real-world's coordinate system and the augmented one (displayed on the phone screen) is made, making the registration of 3D objects (virtual world) possible in the exact position on the map (real world). One of the biggest challenges in the creation of this app was the choice of the tracking pattern.

In order to have a more stable tracking we decided to use a 40 mm ID marker placed on the top of the right side of the paper map. However with the use of the ID marker the user is forced to keep the marker entirely visible in the camera view, while, using natural feature targets, the tracking algorithm works even if it is partially visible. So if they want to focus the opposite corner on the map, they will be forced to move the device away from the map maintaining the marker visible.

Another important issue in the channel development was the 3D models rendering. In an initial stage we decided to locate the entire 3D city model on the map, in order to offer to the user a truly 3D map. However adding several 3D city models attractions as well as animations degrades the channel loading making it very slow. We overcame this issue by replacing the 3D model of the city with spheric elements, where each one represents a major attraction.

Above these spheric 3D models we place the attraction name, as showed in Figure 5.12. The color is necessary due to the occurrence of different types of attractions: museums/theatres (yellow), buildings/churches (red), natural attractions (green). These views offer new interaction opportunities. In fact, tapping on a spheric element, an animation starts: from the sphere, a sort of flower arises, and on its surface a 3D model of the specific attraction is shown (see Figure 5.12).

Then, tapping again on a 3D model, the user can visualise a popup window on the device screen with useful information about that place (web links, pictures, videos, comments from other visitors and much more). To return at the original configuration, the user has to simply tap on the 3D flower again to start the reverse animation.

Figure 5.12: The glue channel at work: 3D balls and name of attractions, placed on their position on the map (left); “flower” with the model of the Oulu Library on its top (right).

We decided to use a scale of 1:800, for example, for the Oulu Cathedral which has a dimension on the plant of about 40 m x 60 m, the related model turned out to have a dimension of 5 cm x 7,5 cm. Since we could not show all the models on the same plane, we chose to display the 3D flowers by using different lengths to fill all the available space above the map. Moreover the map remained uncovered, allowing the user to view in the same time both the models and the map.

Another advantage of this augmented map, compared to traditional maps, is that the user can easily customise the number of information.

In fact, even if we considered only three categories of attractions in our demo, the list can be easily extended, including everything that could help the tourist in the sightseeing planning (e.g. public transports, restaurants, coffee shops, etc.).

The Location Based channel

In the Junaio platform the POIs and their relative information can be displayed in three different ways: augmented reality view, map view and list view. The second and the third modality are fairly self explanatory. In augmented reality view, the POIs are displayed as virtual tags located in the surroundings of the user.

Furthermore the users can enable the view of the 3D models of several attraction thus easily recognising the point of interest. Our work was focused in the development of a very simple and user-friendly UI. Especially in the augmented reality field, the UI design is a very significant step of development. This consideration is even more important when working on the mobile environment, where the screen size can be very small. To identify the more useful categories we chose to take a look at the several touristic applications available in the App Store.

After this investigation we have identified seven main categories:

- Bus stops
- Museums and Churches
- Public offices
- Monuments
- Hotel / Bed and breakfast
- Food and Drink
- Shops

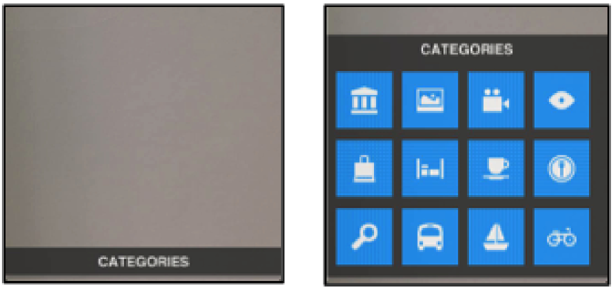


Figure 5.13: The popup-box in the initial appearance (left) and after a tap over the label “Categories” (right).

In augmented reality applications, the entire screen is used to display the camera preview overlaying 2D or 3D object enriching the scene with textual or visual information. Since the screen is not always large enough to show all available information we need a filter option. For this reason, we designed a popup box, showed in Figure 5.13, composed by several category switch-buttons in order to give to the user this functionality. For each category switcher, we have chosen a representative icon.

In order to give a feedback to the user, to understand which categories are filtered, a different opacity is applied to the category switch-buttons. In detail, 1.0 opacity value is applied to all categories POIs showed on the screen, vice versa, a 0.3 value is applied to the remaining POIs. Alternatively, we can apply a green color for the selected ones while a red colour for the others. At the same time, when the user switches to a category button the application makes the relative POIs visible and when the user switches the categories off the relative POIs are hidden.

This hide-and-show operation does not have a high computational cost because all POIs are downloaded and loaded at the starting time, thus 2D and 3D objects are pre-loaded and making them visible is a cost-free operation. These filter operations are done by using several functions written in Javascript (JS) and are performed using a JS library called JQuery [82].

The POIs

We have chosen as case study the city of Oulu in Finland. The entire 3D city model was used for the GC, while for the LBC we found out several POIs in order to create a sample (30 point of interest) to test our prototype. This points of interest include photos, videos and descriptions.

In addition to these POIs we added some 3D models of city monuments, the same used in the GC. In this way, using the live mode on the LBC the users can see the 3D monuments placed in the camera flow. Actually, the displayed monuments are the same that could be found in real life, but in the next future we will be possible to place old 3D models or new ones ad-hoc created. If a 3D artist reproduces old monuments it will be possible to add these models in the application contents in a trivial and easy way. Every POI contains several data: GPS related information (longitude, latitude, altitude), title, description, and eventually, connected images, videos and links. All POIs information are stored in a DB and are retrieved by using a simple PHP script. After this, all contents are sent to the client after the Junaio request. We can divide the POIs in three categories:

- Simple POIs: geo-located attractions like monuments, hotels, shops and so on;
- Image POIs: geo-located images shot in different ages or seasons, showed as layered-images or 360 degrees images;
- 3D POIs: geo-located 3D models of buildings built in different periods of the city life (e.g. before a natural event or a war).

Figure 5.14: From left to right: Live view (LV), LV+ over layered photo (OL) with opacity=0.3, LV+OL with opacity=0.7, LV+OL with opacity=1.0.

We must specify that for the same attraction it could be possible to have different kinds of POIs. For example, we can have a simple POI for a church with general information and related medias, and also have a 3D POI with the old model of the church (e.g. before a war) and again an image POI with a 360 degrees photo shoot in the same place but in another season.

Layered-images and 360 degrees panoramas

In the GC we focused on the 3D models of monuments, while in the LBC we focused on the POIs and their information. We have defined two different types of images: layered-images and 360 panoramas; both images are geolocated and the GPS coordinates indicate the place where the photo was taken. The first one (see Figure 5.14) is a simple image taken with a common camera in a different period and its display-interface shows the entire photo but the user can change the opacity of the image by using a slider controller. In this way, the user can see the live view and the old view at the same time.

The second one (see Figure 5.15) is a panoramic image taken with a third-party application (in our case Photosynth on iPhone 5 was used [132]), using a Junaio feature able to display spherical images, it is possible to show this kind of photos in specific places of the city.

Figure 5.15: Panoramic photo taken with Photosynth on iOS.

The main target of this type of images would be to show an immersive view in a different period of the year. The user during summer months can, for example, see how the city looks in winter.

We performed a semi-structured interview, in order to collect a qualitative feedback on our application and proposed interaction. After answering a set of pre-defined questions the users expressed their comments on the test application. The participants were 20 students (9 among them were girls) aged 20-30 years with different degrees and fields of study. To understand an early impression of the application, and the user interaction preferences, we asked them what they thought about the usefulness of the main idea of our Junaio channels and if they had more confidence with this kind of technologies or simple e-guides. Some questions were focused on the UI design (e.g., if the buttons and their behaviours were self-explanatory), and then we asked if they would use the application in their next trip.

After this early semi-structured interview, we can assert that the large majority of the interviewed users (about 80%) think that both channels are useful and very interesting, in detail the users preferred the LBC slightly more than the GC. Around 80% of the interviewed people would use the system in the next travel. Provided that, particularly interesting is that the majority of users (around 85%) think that a city with this touristic service acquires a positive judgment independently from other factors.

The users think that every city should have this kind of application, especially working with augmented reality techniques. Lastly, one person out of four affirmed that she is not glad to use the smartphone continuously during a city sightseeing, the same percentage thought that using the application only in specific places is more appropriate.

Due to these considerations, we can assert that the main idea discussed in this work is very attractive for the sample users.

In fact, they affirmed that if a city used this, or a similar application, it would acquire a visitor positive judgment in term of touristic oriented services.

Even if the Oulu city has been chosen as a test-case, the entire case of study could be applied to another city without particular adjustments. The main element for this kind of tourism applications is certainly the application content.

5.5 3D-ize U!

The goal of the work presented in this section is to build an application that allows the user to see the virtual three-dimensional representations of their friends and interact with them. The main challenge is to achieve results similar to those that a computer would produce, optimising the process to deal with the constraints of the technology used.

The result of this work is an application named **3D-ize U!** that allows users to build and display a 3D model of one of their friends simply by using a picture. In order to reach the goal, the application goes through several steps:

- **Taking** a picture, or choosing an existing one;
- **Detecting** a set of *feature points* in the picture;
- **Processing** the image using warping techniques in order to create a texture;
- **Applying** the texture onto a predefined 3D model;
- **Rendering** the 3D model and his texture;
- **Interacting** with the model;

Figure 5.16: Final results

In this work we used a warping algorithm in order to obtain the texture necessary for the 3D representation.

Warping is the set of proceedings focused on the transformation and deformation of shapes. This technique has been largely used in shape animation, modelling, analysis and matching (see for example [171] [116]). Basically, the image is deformed in order to distort the geometry of the objects depicted in the scene; it is a process mainly used to correct the distortion caused by the lenses, but has been recently used also for artistic photo editing. Using the warping technique, it is possible to modify an image aiming to align a set of image features to the same features on a different image, usually referred as the template image. It is also possible to calculate the amount of transformation needed to align the features: in this way it is possible to choose, among a pool of images, picking the one that is most similar to the template, to recognise people or objects that appear in damaged or distorted images, even if the object itself is disguised. Intuitively, a graphical object U consists of any processable entity in a computer graphics system: points, lines, surfaces, images and so on. In the particular case where the graphical object is an image I, the transformation $T : I \to R^n$ defined only for the points of I is named intrinsic transformation.

The easiest way to work on an image consists of defining a global transformation of I, that is, a transformation applied uniformly onto every pixel of the image; however, it's really hard to transform a specific object of the scene by means of a global image transformation. Instead, this goal can be achieved working locally on a subset of pixels. A good strategy to build local transformations is based on the idea of subdividing the object into smaller pieces, then defining a transformation for each piece separately and finally merging all the pieces together to obtain a new image. Working locally also has the advantage of approximating each transformation by using simple functions, obtaining good results and reducing the computational cost. Triangles and quadrilaterals has been widely used in computer graphics as blocks to build graphical objects. Dealing with a triangulated image or a subdivision made of a quadrilateral mesh is a rather diffused scenario and the techniques to compute the transformation from a polygon to another gain considerable importance (see [18]).

One of the challenges of the techniques based on locally defined transformations is the need for computing new attribute values for every point of the deformed object using only the values stored in the original one. Considering the particular case in which the affine transformation has been calculated between two triangles ABC and DEF, the transformation T on the vertices is defined as $T(A) = D$, $T(B) = E$ and $T(C) = F$. So it is possible to apply T to every point inside the triangle using barycentric coordinates.

A point p of a triangle ABC has barycentric coordinates $(\lambda_1, \lambda_2, \lambda_3)$, that is

$$p = \lambda_1 A + \lambda_2 B + \lambda_3 C$$

when $\lambda_i \geq 0$ and $\lambda_1 + \lambda_2 + \lambda_3 = 1$.

The transformation to obtain the position of point p in relation to the new reference triangle is the following:

$$\begin{aligned} q &= T(p) \\ &= T(\lambda_1 A + \lambda_2 B + \lambda_3 C) \\ &= \lambda_1 T(A) + \lambda_2 T(B) + \lambda_3 T(C) \\ &= \lambda_1 D + \lambda_2 E + \lambda_3 F \end{aligned}$$

After studying and analysing the mathematical model, the next section will explain how these techniques can be implemented in order to obtain the desired deformation.

Application Design and User Interface

The project has been divided into different phases: image acquisition, detection of feature points, image warping, texturing/rendering of the model and finally user interaction. The first step is the image acquisition. The user must provide an image to the application, and he can do this in two ways: by taking a picture “on-the-fly”, or by using an existing photo. Once the picture has been selected, a special module verifies the goodness of the image and notifies the users of the necessity of taking the picture again (or choosing a better one). In this specific case, a good image is an image that has good contrast, good lighting, a well-defined background and where the face of the person portrayed is clearly visible, making a full-automatic face detection step possible. To identify the face, the Android face detection feature has been used; thanks to this feature, it is possible to correctly locate the face, and center it into an ellipse that will be then used to cut the image.

For best results it is necessary that the user modifies the ellipse to contain the face. Through the use of finger gestures, the user can zoom in or out at will; once the size of the oval is adjusted it’s possible to move to the next phase, the detection of facial feature points. The feature points are the key element in the process.

A good tradeoff in terms of quality and quantity needs to be achieved, because on the one hand too few points produce unreliable results and poor accuracy, while on the other hand too many points make the application computationally heavier and less practical to use. As seen in [154], the most important points to be identified are represented by the eyes, the nose and the mouth. For best results, however, the eyebrows and the silhouette of the face should also be used. We can list the feature point used in the process (see Figure 5.17):

- 4 points for the left eye;
- 2 points for the left eyebrow;
- 4 points for the right eye;
- 2 points for the left eyebrow;
- 3 points for the nose;
- 4 points for the mouth;
- 12 points for the face silhouette;

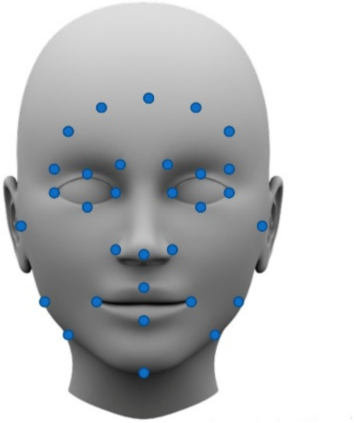


Figure 5.17: All feature points taken with the user interaction.

How it works

The 12 points of the face silhouette are automatically extracted from the ellipse placed over the image by the user. The acquisition of the remaining points is performed in a sequential manner, using a specific panel that tells the user which part of the face he has to pick.

The application shows a cursor made up of a vertical and a horizontal line intersecting the point touched. The user can scroll with the fingers in order to refine the selection and maximise the accuracy. Usually, the texture contains all the information to represent the entire surface of the 3D model; in this case the model is a head, so the texture should include information about the front, the sides and the back.

In this work the choice has been to acquire a single frontal picture, and as a result of this, the resulting texture is incomplete.

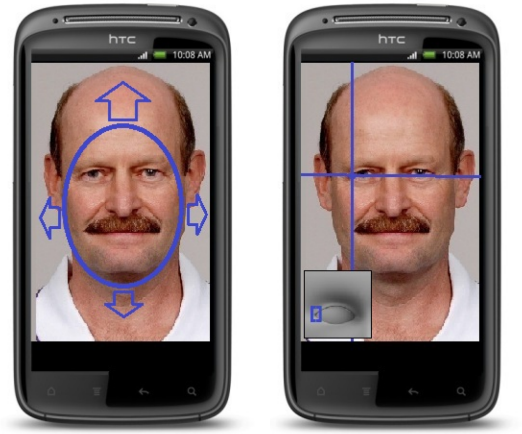


Figure 5.18: Oval fitting activity and point selection activity.

For each point the user has to press a dedicated button to confirm the coordinates indicated, that saves the current point and displays the image of the next point to be identified. This assisted acquisition allows to create a point-to-point correspondence map between the acquired picture and the final texture, and this will be particularly useful for the warping phase. At the end of this procedure every point is known and the process of warping that will generate the texture to apply to the 3D model will start.

After the acquisition of the feature points, the image has to be warped in order to fit a pre-existing template image. The goal of this step is to modify the original image in order to use it for texturing the three dimensional model of the head. Thanks to a 3D model deep analysis, the application knows both the coordinates of the texels associated with the feature points and the correspondences with the points of the 3D model. The information is then used to build the template that will serve as reference for the warping.

Since the system can not detect the position of feature points in an automatic way, the detection of these points is a user duty. As soon as the user locates the points, it becomes possible to warp the original picture in order to fit on the template. The warping operation is performed using the concept of barycentric coordinates ([134]). The first step of this process consists in carrying out the triangulation of the feature points indicated by the user. Since the position of the feature points in the acquired image is known, and also the corresponding template points are known, it is possible to triangulate the image to obtain its discrete subdivision, working locally on the pixels.

The triangulation is a classic *Delaunay triangulation* computed incrementally (implementation provided in [30]). This technique makes it possible to incrementally build the triangulation while the user picks the feature points; in this way, both triangulation and selection of feature points end at the same time. First, the algorithm builds a triangular bounding box such that every point of the triangulation, named as set P, falls within it.

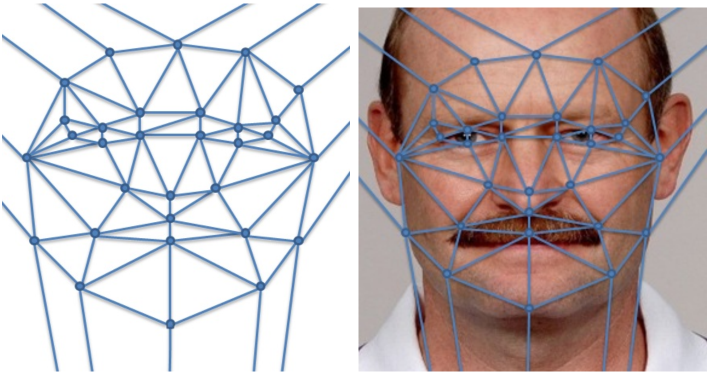


Figure 5.19: Delauney triangulation of feature points and its overlay over the texture.

For every point belonging to P, it is necessary to search for the triangle where it lies. This triangle is then split into three new triangles. This operation is trivial for the first point, when the only existing triangle is the bounding box. Subsequently, a test is performed over the newly created edges in order to avoid the presence of illegal edges. In case of illegal edges, the structure is updated to solve the problem. This is the incremental way to build the Delaunay triangulation of the feature points; the triangles are stored keeping track of the correspondences between the original triangles and the new ones.

The algorithm performs only one triangulation using the set of feature points defined by the user, then it builds a hash map containing the correspondences between those points and the ones located on the template image. In this way, for each triangle T of the image triangulation, there exists a triangle T' that has three vertices corresponding to the 3 feature points of T. The only difference is that these points have different spatial coordinates. To overcome this problem, it is possible to represent every pixel of the image by means of barycentric coordinates, with the assumption of knowing which triangle contains the specified pixel. Then, every pixel is processed locally, moving it from the old reference system (the triangle that contains it) to the new reference system (the corresponding triangle with moved vertices).

Once the new location of the pixel has been identified it is possible to assign his old color to it. Repeating this process for each pixel leads to the desired transformation. However, the image may require some other adjustment to avoid that some pixels may appear incorrectly. The reason is that, because of the discrete nature of pixels, by undergoing this transformation not every pixel of the final texture happens to be the destination of a pixel of the original image. Applying a smoothing filter that, for every not coloured pixel, gives it the average intensity of the 8 neighbouring pixels, overcomes this problem. Each pixel of the neighbourhood is taken into account to calculate the mean if and only if his intensity is not zero.

The choice of this technique has been made after finding out that achieving the image deformation by means of a global transformation was too time consuming, and the result could not be achieved in a reasonable amount of time.

Obtaining the texture from a single picture has the drawback that there is not enough information to texture the sides and the back of the head. In this work, the developers preferred to exclude the details of hair, which would require to be independently studied and that can be integrated in the application afterwards. For this reason, the model obtained has no hair and the choice to assign the color to the skin in the back of the head has been achieved using some kind of averaging of the available information.

Working on the morphology of the human face, it is possible to state that, except in some areas such as eyes, nose and mouth, the skin has the same regular pattern everywhere, excluding the area on the jaw where, especially in men, there may be a thick hair that hides the natural color of the skin. To complete the texture, gaining a realistic effect, the information on the pattern of the skin is taken directly from cheeks and forehead.

To improve this result and reduce the difference between the color of the facial skin and the synthetic skin, the color of the facial image is modified, applying a different enlightenment, similar to the one of the pattern just computed. In this way the color difference is almost not noticeable. The texture is then ready to be adapted to the model.

The Android operating system provides a subset of the OpenGL graphics library to render 3D objects. The only thing needed is a parser that read and build the model, while the visualisation part is carried by the OS. It is obvious to notice that the more complex the model, the slower the visualisation. See [138] for further details on OpenGL for mobile system.
Another constraint to consider is the number of vertices and faces that is representable by an array of short integers, limiting the complexity of the models displayable. Besides that, the number of vertices and faces has been limited to 1016 vertices and 1999 faces due to the excessive computational cost, and also because on small screens it is not necessary to have highly detailed objects to get a good result.

In the testing phase the model has been directly inserted in the application package as a RAW file, then the file is parsed to retrieve the coordinates of the vertices, information about faces, vertex normals and the texture map. Notice that it is possible to save the texture map on a file because this information never changes going from a texture to another. The model is now ready for the visualisation; after the model is rendered, the user can finally interact with it. Currently, the user can interact with the model like any 3D viewer: he can zoom, rotate and scale the displayed object. Furthermore, having different models with different facial expressions, it will be possible to replicate different expressions depending on user interaction.

Having some useful / funny / cool features is not enough to make a "good" application. A good application needs to perform these operations quickly. During several tests the time spent by various processes has always been monitored, to identify the steps in which the code was too computationally expensive. In detail, the application has been tested on a device with internal processor Qualcomm MSM7201A 528 Mhz and 192 Mb RAM. Since phase 1 and 2 are strictly dependent on the user, they are not taken into account, while phase 3 requires in average 53 seconds and phase 4 requires 43 seconds. Times are hardware dependent; brand new smartphones have 1.5Ghz dual core processors, and these operations can be performed in much less time. An interesting way to speed up the process is the use of a remote server for executing the third phase and fourth phase; this alternative will be better discussed in the following chapter.

The results of the work reflect the objectives set at the beginning of the development phase: developing an application for the growing mobile world, based on the Android platform, that allows the user to have a real time textured 3D model, where the texture can be obtained just by taking a picture of someone's face. Actually the application can:

- take a picture, or choose it from the SD card;
- recognise the presence of a face in the picture and draw a silhouette around it; eventually, the user can manually adjust it;
- help the user to identify the feature points, and pick them;
- warp the image so that it can be used as a texture for the 3D model;
- show the 3D model and let the user interact with it through touchscreen, as a classic 3D viewer.

One of the main problems solved during the development concerns the warping of the image. This deformation is necessary for our work to obtain a texture easily applicable to the 3D model, but the warping is based on mathematical functions that can be very complex, and the tradeoff between complexity and quality result is not fair. Despite the fact that the results are truly better using complex functions than using simple ones, the quality difference does not justify the gap in computation time. Hence, on mobile platforms, linear functions must be used in order to have the output back in a reasonable amount of time; in particular, the transformation used is based on the idea of barycentric coordinates, therefore a linear transformation.

Some minor issues, due to the Android platform, rose during the phase of the acquisition of the image. When the image is acquired, the device takes some time to store it either in the internal memory or in the SD card, delaying any work on that image until the storing process is finished.

In the future work, it would be interesting to make an application where people could try different hairstyles or hair color, or see themselves with a beard or a moustache. It would also be possible to use the images of two people to have an idea of the face of their son, or again, use the images of a father and his son to highlight the common features and see how the grown up child will appear, or even warp the faces to see how our friends would be if they were fatter, thinner, elder or younger.

5.6 3D Talking Head

In this section we present THAL-k, the first release of a library for mobile devices supporting the animation of 3D talking heads both for Android OS and iOS. The library is thought as a support for all the developers wishing to build applications on smartphones or tablets, and it uses 3D avatars to enhance their interaction functionalities.

Nowadays, the touch-less interaction (e.g., the usage of voice commands) on mobile devices is particularly attractive. The device can also possibly answer to our questions (e.g., Siri-Speech Interpretation and Recognition Interface, which according to Apple is “the intelligent personal assistant that helps you get things done just by asking”). Sometime, these voice commands are used in combination with 3D avatars that could represent a character, a digitalised real life person, a concept or an artificial entity. These virtual assistants are able to answer questions and perform tasks during conversational dialogs with users.

The use of talking avatars can improve the quality of the interaction by offering useful and pleasant feedback. Since avatars are static models, but the interaction requires dynamics, it is almost obliged to introduce avatars’ animations. Animation in desktop environment has been studied for decades. However transferring the same techniques to the mobile environment requires to reinterpret them due to the aforementioned constraints. This is especially true when we want to achieve real-time animations. The goal is more complicated when the animation of a 3D model is related to facial expressions. A real-time realistic animation has an heavier computational cost than pre-calculated animation. Due to this reason, in our work, we have analysed a possible methodology to simulate the speech of a 3D-avatar in the most fluid and realistic possible way.

How it works

We centred our approach on animating the eyes and the mouth of the avatar to simulate speech. Using a mid-res model (3000 faces and 1500 vertices), the animation would require the identification of some feature points to be used as anchors to realise the facial expression. The problems related to processing power are still present, since we need to move approximately 300 vertices per frame, with an heavy computational effort. We started studying the usage of two poses for each phoneme and calculating the shift for each vertex obtaining a list of pairs [vertex, shift]. These data are pre-computed and loaded into the app to allow the reading when necessary.

Figure 5.20: Talking head as it appears on Android smartphone.

With this approach, our model was poorly performant, the code heavy and not very robust. We, thus, decided to use warping to simulate the human speech, working on the texture of the 3D object. This time, the quality of the results was low since we used a static geometry of the model. Our final solution is described in the rest of the section.

Figure 5.21: Different openness degrees of the same phoneme.

We decided to split the model in 3D sub-meshes. THAL-k includes all the models needed for the animation. In detail, we have different meshes that represent the entire avatar: 1 head (with two textures: male and female), 2 hair meshes (long and short), 4 mouth meshes (tongue, throat, upper and lower teeth), 2 eye meshes (left and right), 50 lips meshes.

After this we edited the sub-meshes to represent the movements frame by frame. We then generated three meshes for the eyes (look-right, look-center, look-left), sharing the same texture and with the same number of vertices and connectivity, but differing in the spatial coordinates of each vertex and the associated normals.

To simulate the movement of the eyes we need to preload in memory the three meshes (about 50 points per mesh) and then work on the visibility of a single one of them; for labial animation we, instead, preload the meshes for each pronunciation, ten for each phoneme (about 400 faces and 200 vertices for mesh). We use only three phonemes: A, F and C. and showing them randomly we produce a realistic and language-independent speech.

The ten meshes for each phoneme, represent different degrees of openness, that is, nine different lip movements between neighbour meshes. The library allows developers to change in real-time the face texture, the length and colour of the hair and the color of the eyes. It is also possible to change the rendering mode (wireframe, texture and cube map) or move the head up, down, left and right even changing the looking direction.

Figure 5.22: Screenshot taken from the demo app.

Figure 5.23: Different view modalities: wireframe, flat, texture and skybox.

Chapter 6

Conclusion and Final Remarks

At the end of this dissertation we can draw the following conclusions and the lessons learnt from each work. As we said in the introduction, we focused over three main aspect: the first is learning with new technologies [3], the second is the improvement of real life by using innovative devices [4] while the third is the usage of mobile devices in combination with image processing algorithms and computer graphics techniques [5].

In all these works we have tried to address three main questions: *how does technology facilitate learning? how can technology advances affect our life? is the mobile the future of every thing?* In order to address these questions we have investigated a huge number of interesting issues that are related to really different aspects of living. In some of them the investigation has been followed by the implementation of prototypes functional for the the initial hypothesis evaluation. The overall conclusion is that technology is an essential component in everybody's life. This happens since the 1980s, when technology were used daily by the majority of the population. This happens nowadays too, in particular with Mobile technology. People's lives are made easier, and such kind of devices have become more than a necessity. Technology makes things in life easier, faster, and more efficient. Advancements in technologies are a good chance to improve the way children learn too, we need these innovative ways to engage students, to assist them and to improve their preparation in any field, from the language to the mathematical, historical or geographical ones. In the following paragraphs we draw the results and the conclusions given by each single project also providing direction and areas for future research.

Through Speaky Notes [3.1] we offer educators a system that uses a dataset of validated words that allows the creation of personal dictionaries, easy to share among colleagues. Furthermore, we envision this system as a tool to engage students during the learning activity. Even if we use as case study the language acquisition process, the whole system could also be applied to other learning fields. The next step for the Speaky Notes tool is the user testing. We are planning a first test under the supervision of an expert in pedagogy in order to get feedback and suggestions. At this moment the application is ready to use and the beta version will be available as open source. In the education path, speaking a second language properly cannot be a secondary goal. Children have to be motivated and enthusiastic in language learning and we see our approach as a good way to give them the tool to reach these feelings.

With SoSphere [3.2], we compared two different solutions for enhancing the interaction experience of a planetarium application, both replicable at a reasonable cost.

In this work, we described the creation of a low-cost setting for an immersive virtual planetarium experience. Starting from an existing software for desktop platforms, we created both a multi-touch and Kinect version of the application. The Kinect versions employs a geodetic display that provides the user with a more accurate representation of the sky. We performed a small-scale user study in order to investigate the perceived difficulty in performing different tasks with the two settings, which was low for both versions. In addition, the post-test questionnaires did not highlight any significant difference in the overall usability between the multi-touch and the full-body interaction. However, we found a difference in the perceived control of the application (which was higher in the multi-touch version) and in the perceived realism of the experience (which was higher in the Kinect version). In future work, we plan to exploit many other devices for controlling the application (e.g. the Leap Motion for a more precise hand tracking) and to combine different sensors to increase the application interaction capabilities.

With RIFTart [3.3], we propose a tool supporting teaching and studying Art History through Virtual Reality. With RIFTart, teachers can exploit 3D models for describing and comparing different artworks. In addition, they can enhance them through multimedia annotations. The same content can be used by students for individual learning. RIFTart has been developed completely with web-based technologies, in order to be accessible from different devices. We discussed the technical solutions, their advantages and disadvantages.

The prototype can be used on wide shared screens, such as LIMs (the multimedia whiteboard currently employed in Italian classrooms), but also on recent consumer level head mounted displays, such as the Oculus Rift or the Google Cardboard. Considering that such solutions will be more and more available in the future, we foresee the possibility to equip school laboratories with immersive VR, or to directly exploit student's mobiles as low-cost HMDs. We evaluated the impact of immersive VR on high-school students motivation through a user test in two classes of the Filippo Figari High School in Sassari.

The results show that the immersive VR increases the students' motivation in studying the lesson topic, in particular increasing their attention, satisfaction and the perceived relevance of the teaching material.

In future work, we aim to provide teachers with a proper authoring environment, a sort of Power Point for VR content, in order to better support them in the creation of the teaching material, and to evaluate both its usability and effectiveness with teachers. In addition, we would like to enhance the evaluation on two ways: on the one hand we would like to measure effect of the immersive VR visualisation in a collaborative lesson setting, where all students are provided with HMDs; on the other hand it would be interesting to perform a long term study on both motivation and students' learning outcome in classes using immersive VR settings.

In those works that concern the use of 3D avatar such as 3.4, 5.6 and 5.5 we faced the mobile devices' constraints by using necessary smart techniques in order to overcome well known computational issues. In detail, in 3.4 we described the design and a prototype implementation of a mobile guide for city visits for a children audience. In order to attract their attention, a comic-book superhero avatar presents the information. In future work, we will refine the guide implementation and we will perform a user study with real children. Our hypothesis is that the proposed approach may increase the children learning during the city visit, if compared to reading information on a printed or electronic guide. This work has been developed thanks to our previous works in this topic, especially the library presented in 5.6.

We then applied similar technology settings for improving real life with new technologies. In RiftAbike [4.1], we discussed the application of gamification techniques in Immersive Virtual Reality environment for supporting physical exercise.

We describe the implementation of RiftAbike, a fitmersive game that enhances the exercise bike experience with challenges, levels, points, badges and prizes. Through the results of a user test, we show that the gamification elements increase the user's enjoyment during the physical activity and we provide a set of guidelines for applying gamification features for fitmersive games. In future work, we will provide multiplayer and social support in Rift-a-bike, in order to stimulate the competition among the members of a social network. In addition, we will analyse the gamification effect in the long term.

In Interactive Shops [4.4], we described a proof-of-concept of an integrated system for enhancing the shopping experience in a shoes' store offering customers two different interactions inside and outside the shop. We are currently studying an approach that allows a multi-user usage of the system since, at the moment, it can be controlled by a single user decreasing the customers' engagement.

We are also planning a real test with an Italian brand in order to understand the customers' feelings and to evaluate the user's experience when interacting with our proposed system.

Nowadays, as more shopping moves online, clothing and footwear shops are trying to catch customers' attention by improving their buying experience with new technologies that are more interactive like smart mirrors or touchable displays. We think that these smart devices could change the way customers shop, and our contribution make a step forward in this direction.

In the work WoBo [4.2], we introduced a new experience for travellers which aims to connect different locations reachable immediately for a virtual journey. A WoBoX is able to record a 360-degree video of the surrounding environment, which can be viewed in any other WoBoX in the world.

We described the hardware and software setting for a first prototype and the results of a preliminary user evaluation of the proposed experience, demonstrating a high level of sensory engagement. Twelve users participated to the test and its results were encouraging. First of all, the sensory engagement was highly rated considering the overall experience, the visual and audio experience. In addition, the user did not find inconsistencies in the information coming from the different senses and they were not confused after the session.

The SmartMirror project [4.3] aims at creating a seamless interactive support for fitness and wellness activities in touristic resorts.

In this work, we also reported on the state of the art technologies for building smart mirrors and for sensing the user's physical status. We described smart mirrors devices and prototypes working as multimedia players, interactive home controllers and augmented reality devices. Moreover, we described and compared different parameters for evaluating the user's physical state, considering whether they require annoying or uncommon actions, the contact between the sensors and the user and the procedure cost. Finally, we introduced the scope and the objectives of the Virtuoso project, which aims to create a seamless interactive support for fitness and wellness activities in touristic resorts. We explained how, in future work, we are planning to create the project setup and to evaluate its results.

Finally, given the importance of mobile devices and their utility as learning tools, we developed four projects that implement well known algorithms that are often used on desktop environment. In these works the main challenge is the constraint of such devices due to their computational power.

In Chromagram [5.1], we tested two different approaches. The former by using the OpenCV and the latter by using the OpenGL library. We tested both approaches on the Android platform by using the same device obtaining very different performance results. In Table 5.1 we reported the performances of both applications on several devices with different resolutions and processing power. Even if the test results seem promising, the chroma key matting effects have several issues regarding the scene illumination, the video compression and the camera quality. Other minor issues are related to the limitation of the algorithm used for the matting effects. Implementing a spill removal algorithm is a good way to solve this problem, but this way decreases the application performance and it requires a good setup of the studio scene.

A last noticeable issue is that we can lose small details of foreground objects, like hair for framed persons, or semitransparent objects made with materials like glass. At the moment, to overcome this kind of issues thus obtaining good results, we have to correctly set up the shooting scene. In the next future we plan to add new features, allowing the recording of video in addition to the images.

We want also to improve the chroma key algorithm using a Bayesian or a Poisson matting filter that provides better visual results, solving some of the issues previously described. We are also working to decrease the complexity of the algorithm, preserving the computational performance results.

In Click and Share [5.3], we evaluated the algorithm used for the face recognition task by splitting it in two parts: we first evaluated how long it takes to train the algorithm with the considered dataset and then we recorded the time elapsed for recognising a person with the leave-one-out technique. We can affirm that the time needed for the recognition is acceptable for all the phone categories we considered.

Therefore, the technique may be already considered for a broad set of mobile phones, without any relevant delay for the user. In order to evaluate the improvements introduced by the approach described in this work, we performed a user test. The aim of the test is to assess the efficiency, the errors and the cognitive effort with respect to the existing approaches for creating and sharing photos through mobile devices. We considered the Facebook mobile application as the baseline for the comparison, which allows the users to take pictures, to recognise people into them and to share the photos.

Eighteen users participated to the test, 13 male and 5 female, aged between 19 and 54. They were highly experienced in using mobile devices and in particular they had good experience with the Android OS. In conclusion, the user test highlighted a difference between Click & Share and Facebook on both completion time and number of errors, while maintaining the same cognitive effort for reaching the user's goal. The prototype implementation performance shows that the approach is already feasible for being employed in mobile devices, while the user test shows an increase of the user performance from both the efficiency and the effectiveness point of view. In future work, we want to apply the same technique to a broader set of media and to allow the synchronisation of the data collected through different mobile devices.

In VisitAR [5.4], we performed a semi-structured interview, in order to collect a qualitative feedback on our application and proposed interaction. Some questions were focused on the UI design (e.g., if the buttons and their behaviours were self-explanatory), and then we asked if they would use the application in their next trip. After this early semi-structured interview, we can assert that the large majority of the interviewed users (about 80%) think that both channels are useful and very interesting. Around 80% of the interviewed people would use the system in the next travel. Provided that, particularly interesting is that the majority of users (around 85%) think that a city with this touristic service acquires a positive judgment independently from other factors. The users think that every city should have this kind of application, especially working with AR techniques.

Lastly, one person out of four affirmed that she is not glad to use the smartphone continuously during a city sightseeing, the same percentage thought that using the application only in specific places is more appropriate. Due to these considerations, we can assert that the main idea discussed in this work is very attractive for the sample users. Even if the Oulu city has been chosen as a test-case, the entire case of study could be applied to another city without particular adjustments. The main element for this kind of tourism applications is certainly the application content.

In AR-Garden [5.2], we tested the application in two steps: the former by emulating the user position and the latter by trying it directly on-site. To emulate the GPS signals we used an Android emulator by using a GPX file. In the first simulation, the provided directions have been shown correctly as expected. On the other hand, testing it on site we obtained different and not reliable directions. The well known error in GPS coordinates is really significant and it causes wrong directions. The main problem is the signal strength, since the signal acquisition does not work very well in covered or partially covered areas. This leads to high errors in the position calculation during the visit of the area. In our case of study, the Botanical Garden of Cagliari, we have encountered this problem because it is almost totally covered by trees forming a barrier between satellites and the device.

We have estimated that this error is around 20 meters in the worst cases. This implies that the navigation based on GPS coordinates is not enough to offer a good turn-by-turn navigation in such kind of areas.

Bibliography

[1] Mobile hci '02: Proceedings of the 4th international symposium on mobile human-computer interaction. F. Paternò, Ed., Springer-Verlag. 97

[2] Agre, J., Magness, J., Hull, S., Wright, K., Baxter, T., Patterson, R., and Stradel, L. Strength testing with a portable dynamometer: reliability for upper and lower extremities. *Archives of physical medicine and rehabilitation 68*, 7 (July 1987), 454—458. 93

[3] Ahonen, T., Hadid, A., and Pietikäinen, M. *Computer Vision - ECCV 2004: 8th European Conference on Computer Vision, Prague, Czech Republic, May 11-14, 2004. Proceedings, Part I.* Springer Berlin Heidelberg, Berlin, Heidelberg, 2004, ch. Face Recognition with Local Binary Patterns, pp. 469–481. 27, 123

[4] Al-Ali, A., and AL-Rousan, M. Java-based home automation system. *Consumer Electronics, IEEE Transactions on 50*, 2 (May 2004), 498–504. 97

[5] Anunciacao, P., Ribeiro, R., Pereira, M., and Comunian, M. Different measurements of waist circumference and sagittal abdominal diameter and their relationship with cardiometabolic risk factors in elderly men. *Journal of Human Nutrition and Dietetics 27*, 2 (2014), 162–167. 93

[6] Anwar Hossain, M., Atrey, P., and El Saddik, A. Smart mirror for ambient home environment. In *Intelligent Environments, 2007. IE 07. 3rd IET International Conference on* (Sept 2007), pp. 589–596. 23

[7] Asher, J. J., and García, R. The optimal age to learn a foreign language. *The Modern Language Journal 53*, 5 (1969), 334–341. 39

[8] Atlas, T. Digital maps. 115

[9] Azuma, R. T., et al. A survey of augmented reality. *Presence 6*, 4 (1997), 355–385. 29

[10] Bagavathiappan, S., Saravanan, T., Philip, J., Jayakumar, T., Raj, B., Karunanithi, R., Panicker, T. M., Korath, P., and Jagadeesan, K. Investigation of peripheral vascular disorders using thermal imaging. *The British Journal of Diabetes & Vascular Disease 8*, 2 (2008), 102–104. 92

[11] Bansal, H., Yammiyavar, P., and Anita, P. Y. *ICoRD'15 – Research into Design Across Boundaries Volume 1: Theory, Research Methodology, Aesthetics, Human Factors and Education.* Springer India, New Delhi, 2015, ch. A Study on Entrances and Foyers in Shopping Malls and Their Role in Influencing Perceptions, pp. 457–468. 25

[12] Bastiaens, T., Wood, L. C., and Reiners, T. New landscapes and new eyes: The role of virtual world design for supply chain education. *Ubiquitous Learning: An International Journal* (2014), 37–49. 20

[13] Bearden, W. O. Determinant attributes of store patronage-downtown versus outlying shopping centers. *Journal of Retailing 53*, 2 (1977), 15. 25

[14] Belhumeur, P., Hespanha, J., and Kriegman, D. Eigenfaces vs. fisherfaces: recognition using class specific linear projection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 19*, 7 (Jul 1997), 711–720. 27

[15] Benko, H., Wilson, A. D., and Balakrishnan, R. Sphere: multi-touch interactions on a spherical display. In *Proceedings of the 21st annual ACM symposium on User interface software and technology* (2008), UIST '08, ACM, pp. 77–86. 18

[16] Biddiss, E., and Irwin, J. Active video games to promote physical activity in children and youth: a systematic review. *Archives of pediatrics & adolescent medicine 164*, 7 (2010), 664–672. 21

[17] Boi, S., Sorrentino, F., Marras, S., and Scateni, R. 3D-ize U! A Real-time 3D Head-model Texture Generator for Android. In *Eurographics Italian Chapter Conference 2011* (2011), A. F. Abate, M. Nappi, and G. Tortora, Eds., The Eurographics Association. 106

[18] Botsch, M., Kobbelt, L., Pauly, M., Alliez, P., and Lévy, B. *Polygon Mesh Processing.* A. K. Peters Ltd., 2010. 138

[19] Brinkmann, R. *The Art and Science of Digital Compositing.* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1999. 26, 108

[20] Brooke, J. SUS: A "quick and dirty" usability scale. In *Usability Evaluation in Industry*, P. Jordan, B. Thomas, and B. Weerdmeester, Eds. Taylor & Francis, London, 1996, pp. 189–194. 32, 51

[21] Broughton, G., Brumfit, C., Pincas, A., and Wilde, R. D. *Teaching English as a foreign language.* Routledge, 2002. 39

[22] Cambridge, A. The database of faces, 2002. 124

[23] Carrozzino, M., Evangelista, C., Brondi, R., Tecchia, F., and Bergamasco, M. Virtual reconstruction of paintings as a tool for research and learning. *Journal of Cultural Heritage 15*, 3 (2014), 308 – 312. 20

[24] Cascales, A., Laguna, I., Pérez-López, D., Perona, P., and Contero, M. An experience on natural sciences augmented reality contents for preschoolers. In *Virtual, Augmented and Mixed Reality. Systems and Applications.* Springer, 2013, pp. 103–112. 16

[25] Casti, S., Sorrentino, F., Spano, L. D., and Scateni, R. Click and share: A face recognition tool for the mobile community. In *Image Processing (ICIP), 2014 IEEE International Conference on* (2014), IEEE, pp. 1952–1956. 106

[26] Casu, A., Spano, L. D., Sorrentino, F., and Scateni, R. Riftart: Bringing masterpieces in the classroom through immersive virtual reality. In *Smart Tools and Apps for Graphics-Eurographics Italian Chapter Conference* (2015). 36

[27] Chance, P. *Learning through Play. Pediatric Round Table: 3.* ERIC, 1979. 13

[28] Chen, C.-Y., Chang, B. R., and Huang, P.-S. Multimedia augmented reality information system for museum guidance. *Personal and Ubiquitous Computing 18*, 2 (2013), 315–322. 28

[29] Cherchi, G., Sorrentino, F., and Scateni, R. AR Turn-by-turn Navigation in Small Urban Areas and Information Browsing. In *Smart Tools and Apps for Graphics - Eurographics Italian Chapter Conference* (2014), A. Giachetti, Ed., The Eurographics Association. 106

[30] Chew, L. P. Voronoi/delauanay applet, 2007. `www.cs.cornell.edu/home/chew/chew.html`. 142

[31] Chittaro, L., and Ranon, R. Web3d technologies in learning, education and training: Motivations, issues, opportunities. *Computers & Education 49*, 1 (2007), 3 – 18. Web3D Technologies in Learning, Education and Training. 20

[32] Cho, J.-H., Ziegler, R., Gross, M., and Lee, K. H. Improving alpha matte with depth information. *IEICE Electronics Express 6*, 22 (2009), 1602–1607. 26, 108

[33] Chuang, Y.-Y., Curless, B., Salesin, D., and Szeliski, R. A bayesian approach to digital matting. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on* (2001), vol. 2, IEEE, pp. II–264–II–271 vol.2. 26, 108

[34] Clark, M. A., Lucett, S., Corn, R. J., et al. *NASM essentials of personal fitness training*. Lippincott Williams & Wilkins, 2008. 77

[35] Clark, R. A., Bryant, A. L., Pua, Y., McCrory, P., Bennell, K., and Hunt, M. Validity and reliability of the nintendo wii balance board for assessment of standing balance. *Gait & Posture 31*, 3 (2010), 307 – 310. 93

[36] Corda, F., Sorrentino, F., and Scateni, R. CHROMAGRAM: A Real-time Chroma Key Application for Mobile Devices. In *Smart Tools and Apps for Graphics - Eurographics Italian Chapter Conference* (2014), A. Giachetti, Ed., The Eurographics Association. 106

[37] Czaja, S. J., and Lee, C. C. The impact of aging on access to technology. *Universal Access in the Information Society 5*, 4 (2007), 341–349. 8

[38] Dalgarno, B., and Lee, M. J. What are the learning affordances of 3-d virtual environments? *British Journal of Educational Technology 41*, 1 (2010), 10–32. 20

[39] Di Serio, A., Ibáñez, M. B., and Kloos, C. D. Impact of an augmented reality system on students' motivation for a visual art course. *Computers & Education 68* (2013), 586–596. 15

[40] Dickey, M. D. Teaching in 3d: Pedagogical affordances and constraints of 3d virtual worlds for synchronous distance learning. *Distance education 24*, 1 (2003), 105–121. 19

[41] Dimitropoulos, K., Manitsaris, A., and Mavridis, I. Building virtual reality environments for distance education on the web: A case study in medical education. *International Journal of Social Sciences 2*, 1 (2008), 62–70. 20

[42] Dror, I. E. Technology enhanced learning: The good, the bad, and the ugly. *Pragmatics & Cognition 16*, 2 (2008), 215–223. 36

[43] Ebay. Rebecca Minkoff Connected Store Demo. `https://www.youtube.com/watch?v=6G3JIyG_GeY#t=10`. Accessed: 2015-01-26. 24, 25

[44] Emerson, L., Fear, J., Fox, S., and Sanders, E. *Parental engagement in learning and schooling: Lessons from research.* Family-School and Community Partnerships Bureau, 2012. 14, 38

[45] Esri. CityEngine software. `http://esri.com/software/cityengine`, 2015. Online; accessed 20-September-2015. 76

[46] FaceGen Inc. FaceGen Modeller 3.0 from Singular Inversions Inc. `http://www.FaceGen.com/`. Online; Accessed: 2014-01-18. 70

[47] Fibbi, S., Spano, L. D., Sorrentino, F., and Scateni, R. Wobo: Multisensorial travels through oculus rift. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems* (2015), ACM, pp. 299–302. 74

[48] Fitocracy. Fitocracy online game. `https://www.fitocracy.com/`, 2015. Online; accessed 20-September-2015. 22

[49] Fleck, S., and Simon, G. An augmented reality environment for astronomy learning in elementary grades: An exploratory study. In *Proceedings of the 25 ième conférence francophone on l'Interaction Homme-Machine* (2013), ACM, p. 14. 15

[50] Fletcher, D. 10 Tech Trends for 2010. `http://www.time.com/`. Online; accessed 04-May-2012. 15

[51] Francese, R., Passero, I., and Tortora, G. Wiimote and Kinect: gestural user interfaces add a natural third dimension to HCI. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (2012), AVI '12, ACM, pp. 116–123. 18

[52] Fritz, F., Susperregui, A., and Linaza, M. T. Enhancing cultural tourism experiences with augmented reality technologies. 6th International Symposium on Virtual Reality, Archaeology and Cultural Heritage (VAST). 29

[53] Gaitatzes, A., Christopoulos, D., Voulgari, A., and Roussou, M. Hellenic cultural heritage through immersive virtual archaeology. In *Proc. 6th International Conference on Virtual Systems and Multimedia, Ogaki, Japan* (2000), pp. 57–64. 20

[54] Galassetti, P. R., Novak, B., Nemet, D., Rose-Gottron, C., Cooper, D. M., Meinardi, S., Newcomb, R., Zaldivar, F., and Blake, D. R. Breath ethanol and acetone as indicators of serum glucose levels: an initial report. *Diabetes technology & therapeutics 7*, 1 (February 2005), 115—123. 92

[55] Girard, S., and Johnson, H. What do children favor as embodied pedagogical agents? In *Intelligent tutoring systems* (2010), Springer, pp. 307–316. 17, 68

[56] Gleue, T., and Dähne, P. Design and implementation of a mobile device for outdoor augmented reality in the archeoguide project. In *Proceedings of the 2001 Conference on Virtual Reality, Archeology, and Cultural Heritage* (New York, NY, USA, 2001), VAST '01, ACM, pp. 161–168. 29

[57] Golder, S. A., Wilkinson, D. M., and Huberman, B. A. *Communities and Technologies 2007: Proceedings of the Third Communities and Technologies Conference, Michigan State University 2007.* Springer London, London, 2007, ch. Rhythms of Social Interaction: Messaging Within a Massive Online Network, pp. 41–66. 123

[58] Gomes Sakamoto, S., Miranda, L. C., and Hornung, H. *Universal Access in Human-Computer Interaction. Aging and Assistive Environments: 8th International Conference, UAHCI 2014, Held as Part of HCI International 2014, Heraklion, Crete, Greece, June 22-27, 2014, Proceedings, Part III.* Springer International Publishing, Cham, 2014, ch. Home Control via Mobile Devices: State of the Art and HCI Challenges under the Perspective of Diversity, pp. 501–512. 25

[59] Google. Google Cardboard. `https://www.google.com/get/cardboard/`. Online; accessed 20-September-2015. 19

[60] Google Text-to-Speech. Google Text-to-Speech. `https://play.google.com`. Online; Accessed: 2014-01-18. 69

[61] Grün, A., Remondino, F., and Zhang, L. Photogrammetric reconstruction of the great buddha of bamiyan, afghanistan. *The Photogrammetric Record 19*, 107 (2004), 177–199. 20

[62] Grynkofi, R. Mirror 2.0. `http://bathroominnovation.com.au/finalists#Year2013`. Accessed: 2015-01-23. 22

[63] Gvili, R., Kaplan, A., Ofek, E., and Yahav, G. Depth keying. In *Electronic Imaging 2003* (2003), vol. 5006, International Society for Optics and Photonics, pp. 564–574. 26, 108

[64] Haake, M., and Gulz, A. A look at the roles of look & roles in embodied pedagogical agents-a user preference perspective. *International Journal of Artificial Intelligence in Education 19*, 1 (2009), 39. 17, 68

[65] Hamilton, N., and McFarlane, J. Children learn through play. *Putting Children First 14* (2005), 8–10. 13, 14

[66] Han, J. Y. Low-cost multi-touch sensing through frustrated total internal reflection. In *Proceedings of the 18th annual ACM symposium on User interface software and technology* (2005), UIST '05, ACM, pp. 115–118. 47

[67] Hawkins, D. G. Communication in the age of virtual reality. L. Erlbaum Associates Inc., Hillsdale, NJ, USA, 1995, ch. Virtual Reality and Passive Simulators: The Future of Fun, pp. 159–189. 19

[68] Henderson, A. T., and Berla, N. *A new generation of evidence: The family is critical to student achievement.* ERIC, 1994. 14

[69] Holz, T., Dragone, M., O'Hare, G. M., Martin, A., and Duffy, B. R. Mixed reality agents as museum guides. In *ABSHL'06: Agent-Based Systems for Human Learning, AAMAS 2006 Workshop* (2006), Citeseer. 17

[70] Honig, A. S. Research in review. *Young Children* (1983), 57. 35

[71] House, B., Lloyd, A., and Zimbalist, M. Reveal project. `http://brianhouse.net/works/reveal/`. Accessed: 2015-01-23. 23

[72] Hughes, M., Wikeley, F., and Nash, T. *Parents and Their Children's Schools.* Blackwells, Oxford, April 1994. 14

[73] Hughes, S. W. Stellarium–a valuable resource for teaching astronomy in the classroom and beyond. *Science Education News (SEN) 57*, 2 (2008), 83–86. 46

[74] Iacolina, S. A., Soro, A., and Scateni, R. Improving FTIR based multi-touch sensors with IR shadow tracking. In *Proceedings of the 3rd ACM SIGCHI symposium on Engineering interactive computing systems* (2011), EICS '11, ACM, pp. 241–246. 47

[75] ISO389. Acoustics-standard reference zero for the calibration of pure-tone air conduction audiometers, 1991. 94

[76] IVONAText-to-Speech. Text-to-Speech voices. `http://www.ivona.com`. Online; Accessed: 2014-01-18. 69

[77] Jaffrin, M. Y. Body composition determination by bioimpedance: an update. *Current opinion in clinical nutrition and metabolic care 12*, 5 (September 2009), 482—486. 93

[78] Javascript 3D. Threejs Library. `http://threejs.org/`. Online; accessed 20-September-2015. 59

[79] Johnson, J. S., and Newport, E. L. Critical period effects in second language learning: The influence of maturational state on the acquisition of english as a second language. *Cognitive Psychology 21*, 1 (1989), 60 – 99. 7

[80] Johnson, J. S., and Newport, E. L. Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. *Cognition 39*, 3 (1991), 215–258. 7

[81] Johnson, W. L. Interaction tactics for socially intelligent pedagogical agents. In *Proceedings of the 8th international conference on Intelligent user interfaces* (2003), ACM, pp. 251–253. 17

[82] JQuery. Jquery library. 133

[83] Junaio. Junaio ar browser. 129

[84] Kaltenbrunner, M., Bovermann, T., Bencina, R., and Costanza, E. TUIO: A protocol for table-top tangible user interfaces. In *Proc. of the The 6th International Workshop on Gesture in Human-Computer Interaction and Simulation* (2005). 47

[85] Kanade, T., Oda, K., Yoshida, A., Tanaka, M., and Kano, H. Video-rate z keying: A new method for merging images. Tech. rep., DTIC Document, 1995. 26, 108

[86] Kawakita, M., Iizuka, K., Aida, T., Kurita, T., and Kikuchi, H. Real-time three-dimensional video image composition by depth information. *IEICE Electronics Express 1*, 9 (2004), 237–242. 26, 108

[87] Keller, J. Development and use of the arcs model of instructional design. *Journal of instructional development 10*, 3 (1987), 2–10. 61, 62

[88] Keller, J. M. *Motivational design for learning and performance: The ARCS model approach.* Springer Science & Business Media, 2009. 54, 60, 61, 62

[89] Kendzierski, D., and DeCarlo, K. J. Physical Activity Enjoyment Scale: Two validation studies. *Journal of Sport & Exercise Psychology 13* (1991), 50–64. 32, 79

[90] Kennedy, S., Fawcett, R., Miller, A., Dow, L., Sweetman, R., Field, A., Campbell, A., Oliver, I., McCaffery, J., and Allison, C. Exploring canons amp; cathedrals with open virtual worlds: The recreation of st andrews cathedral, st andrews day, 1318. In *Digital Heritage International Congress (DigitalHeritage), 2013* (Oct 2013), vol. 2, pp. 273–280. 20

[91] Kim, Y., and Baylor, A. L. A social-cognitive framework for pedagogical agents as learning companions. *Educational Technology Research and Development 54*, 6 (2006), 569–596. 17

[92] Kopp, S., Gesellensetter, L., Krämer, N. C., and Wachsmuth, I. A conversational agent as museum guide–design and evaluation of a real-world application. In *Intelligent Virtual Agents* (2005), Springer, pp. 329–343. 17

[93] Kounavis, C. D., Kasimati, A. E., Zamani, E. D., and Giaglis, G. M. Enhancing the tourism experience through mobile augmented reality: Challenges and prospects. *International Journal of Engineering Business Management 4*, 10 (2012), 1–6. 30

[94] Kronos Group. OpenGL ES 2.0 Specification. `https://www.khronos.org/registry/gles/specs/2.0/es_full_spec_2.0.25.pdf`. Online; accessed 12-June-2015. 59

[95] Kronos Group. WebGL 2.0 Specification. `https://www.khronos.org/registry/webgl/specs/latest/2.0/`. Online; accessed 12-June-2015. 59

[96] Lartigue, J., Scoville, T., and Pham, M. Promoting k-8 learning using oculus rift: Employing virtual reality to increase learning outcomes in elementary biology. In *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2014* (New Orleans, LA, United States, October 2014), Association for the Advancement of Computing in Education (AACE), pp. 1100–1105. 20

[97] Law, J. Cybertecture Mirror. `http://www.jameslawcybertecture.com/index.php?section=Company`. Accessed: 2015-01-23. 24

[98] Lee, B.-H., hai An, S., and Shin, D.-R. A remote control service for osgi-based unmanned vehicle using smartphone in ubiquitous environment. In *Computational Intelligence, Communication Systems and Networks (CICSyN), 2011 Third International Conference on* (July 2011), pp. 158–163. 25

[99] Lee, H. K., Yoon, S.-Y., Lee, J. Y., Kim, H. B., Kwon, H. J., Kang, H. J., Hur, H. J., Lee, M.-A., and Jun, D. G. The effects of sale signs on consumer intentions to visit a store. *Journal of Global Fashion Marketing 6*, 1 (2015), 20–32. 25

[100] Lee, S. H., Han, J. H., Leem, Y. T., and Yigitcanlar, T. Towards ubiquitous city: Concept, planning, and experiences. *Igi Global* (2008), 148–169. 28

[101] Lenne, D., Thouvenin, I., and Aubry, S. Supporting design with 3d-annotations in a collaborative virtual environment. *Research in Engineering Design 20*, 3 (2009), 149–155. 56

[102] Lessiter, J., Freeman, J., Keogh, E., and Davidoff, J. A cross-media presence questionnaire: The itc-sense of presence inventory. *Presence 10*, 3 (June 2001), 282–297. 50, 51, 52, 87

[103] Levoy, M., Pulli, K., Curless, B., Rusinkiewicz, S., Koller, D., Pereira, L., Ginzton, M., Anderson, S., Davis, J., Ginsberg, J., Shade, J., and Fulk, D. The digital michelangelo project: 3d scanning of large statues. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (New York, NY, USA, 2000), SIGGRAPH '00, ACM Press/Addison-Wesley Publishing Co., pp. 131–144. 20, 56

[104] Lowe, D. G. Object recognition from local scale-invariant features. In *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on* (1999), vol. 2, Ieee, pp. 1150–1157 vol.2. 85

[105] Macedonia, M. Games, simulation, and the military education dilemma. In *Internet and the University: 2001 Forum* (2002), pp. 157–167. 9

[106] Marsico, M. D., Galdi, C., Nappi, M., and Riccio, D. Firme: Face and iris recognition for mobile engagement. *Image and Vision Computing 32*, 12 (2014), 1161 – 1172. 27

[107] Martín-Gutiérrez, J., and Contero, M. Improving academic performance and motivation in engineering education with augmented reality. In *HCI International 2011–Posters' Extended Abstracts*. Springer, 2011, pp. 509–513. 15

[108] Maurer, T. J., Shultz, K., and Adams, G. Employee development and training issues related to the aging workforce. *Aging and work in the 21st century* (2007), 163–178. 8

[109] mCRUMBS (2011) Basel Augmented Reality TouristGuide. `http://www.perey.com/AugmentedRealityForBasel/`. 30

[110] Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., and Davis, T. J. Effectiveness of virtual reality-based instruction on students' learning outcomes in k-12 and higher education: A meta-analysis. *Computers & Education 70*, 0 (2014), 29 – 40. 19

[111] Michele, G., Michele, D. D., and Fabio, S. Visitar: a mobile application for tourism using ar. In *SIGGRAPH Asia 2013 Symposium on Mobile Graphics and Interactive Applications* (2013), ACM, p. 103. 71, 106

[112] Microsoft. Kinect for Windows V2. `https://dev.windows.com/en-us/kinect`, 2015. Online; accessed 20-September-2015. 76

[113] Milgram, P., Takemura, H., Utsumi, A., and Kishino, F. Augmented reality: a class of displays on the reality-virtuality continuum, 1995. 15

[114] Miyachi, M., Yamamoto, K., Ohkawara, K., Tanaka, S., et al. METs in adults while playing active video games: a metabolic chamber study. *Medicine & Science in Sports & Exercise 42*, 6 (2010), 1149–53. 21

[115] Moore, S., and Bowden, R. Local binary patterns for multi-view facial expression recognition. *Computer Vision and Image Understanding 115*, 4 (2011), 541 – 558. 123

[116] Mullens, S., and Notley, S. Image morphing. `www.stephenmullens.co.uk`. 138

[117] Mullins, T. Brushing Teeth Mirror. `http://www.designboom.com/contest/view.php?contest_pk=36&item_pk=44258&p=1`. Accessed: 2015-01-23. 24

[118] Mulloni, A., Grubert, J., Seichter, H., Langlotz, T., Grasset, R., Reitmayr, G., and Schmalstieg, D. Experiences with the impact of tracking technology in mobile augmented reality evaluations. In *MobileHCI 2012 Workshop MobiVis* (2012). 28

[119] Navteq. Digital maps. 115

[120] NEOD. NEOD Framed Mirror TV. `http://www.neod.org/`. Accessed: 2015-01-23. 23

[121] Nike. Nike Plus. `https://secure-nikeplus.nike.com/plus/`, 2015. Online; accessed 20-September-2015. 22

[122] Nitz, J., Kuys, S., Isles, R., and Fu, S. Is the wii fit a new-generation tool for improving balance, health and well-being? a pilot study. *Climacteric 13*, 5 (2010), 487–491. 21

[123] Noor, A. K. Potential of virtual worlds for remote space exploration. *Advances in Engineering Software 41*, 4 (2010), 666 – 673. 18

[124] Oculus. Oculus Rift. `https://www.oculus.com/`, 2015. Online; accessed 20-September-2015. 19, 75

[125] O'Keeffe, G. S. *CyberSafe: Protecting and Empowering Kids in the Digital World of Texting, Gaming, and Social Media*, 1st ed. American Academy of Pediatrics, 2010. 7

[126] Olsen, G., and Fuller, M. The benefits of parent involvement: What research has to say. *Education. com* (2008). 14

[127] OpenStreetMap. Digital maps. 115

[128] Ortiz, A., del Puy Carretero, M., Oyarzun, D., Yanguas, J. J., Buiza, C., Gonzalez, M. F., and Etxeberria, I. Elderly users in ambient intelligence: Does an avatar improve the interaction? In *Universal access in ambient intelligence environments*. Springer, 2007, pp. 99–114. 17

[129] Panasonic. The Future Mirror: A Beautiful Innovation. `http://youtu.be/-2_kc9GQYIE`. Accessed: 2015-01-23. 24

[130] Pat, B. Developing an understanding of young children's learning through play: the place of observation, interaction and reflection. *British Educational Research Journal 32*, 2 (2006), 191–207. 13

[131] Petruzzellis, L., Chebat, J.-C., and Palumbo, A. *Ideas in Marketing: Finding the New and Polishing the Old: Proceedings of the 2013 Academy of Marketing Science (AMS) Annual Conference.* Springer International Publishing, Cham, 2015, ch. "Hey Dee-Jay Let's Play that Song and Keep me Shopping All Day Long". The Effect of Famous Background Music on Consumer Shopping Behavior., pp. 756–765. 25

[132] Photosynth. 360 photos. 134

[133] Photron. Chromakey effect. 27, 108

[134] Pickover, C. A. *The Mobius Strip.* Thunder's Mouth Press, 2006. 142

[135] Pintus, A., Carboni, D., and Piras, A. Paraimpu: A platform for a social web of things. In *Proceedings of the 21st International Conference on World Wide Web* (New York, NY, USA, 2012), WWW '12 Companion, ACM, pp. 401–404. 25

[136] Poh, M.-Z., McDuff, D., and Picard, R. A medical mirror for non-contact health monitoring. In *ACM SIGGRAPH 2011 Emerging Technologies* (New York, NY, USA, 2011), SIGGRAPH '11, ACM, pp. 2:1–2:1. 24

[137] PrimeSense. NITE Middleware. `http://www.primesense.com/solutions/nite-middleware/`. 49

[138] Pulli, K., Vaarala, J., Miettinen, V., Aarnio, T., and Roimela, K. *Mobile 3D Graphics: With OpenGL ES and M3G.* Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, November 2007. 144

[139] Purnama, J., Andrew, D., and Galinium, M. Geometry learning tool for elementary school using augmented reality. In *Industrial Automation, Information and Communications Technology (IAICT), 2014 International Conference on* (2014), IEEE, pp. 145–148. 15

[140] Raspberry PI Foundation. Raspberry PI. `https://www.raspberrypi.org/`, 2015. Online; accessed 20-September-2015. 76

[141] Reitberger, W., Meschtscherjakov, A., Mirlacher, T., Scherndl, T., Huber, H., and Tscheligi, M. A persuasive interactive mannequin for shop windows. In *Proceedings of the 4th International Conference on Persuasive Technology* (New York, NY, USA, 2009), Persuasive '09, ACM, pp. 4:1–4:8. 25

[142] Ren, J., Jiang, X., and Yuan, J. A complete and fully automated face verification system on mobile devices. *Pattern Recognition 46*, 1 (2013), 45 – 56. 27

[143] Roussou, M. Learning by doing and learning through play: An exploration of interactivity in virtual environments for children. *Comput. Entertain. 2*, 1 (Jan. 2004), 10–10. 13

[144] Saba, M., Scateni, R., Sorrentino, F., Spano, L. D., Colantonio, S., Giorgi, D., Magrini, M., Salvetti, O., Buonaccorsi, N., and Vitali, I. Smart mirror where i stand, who is the

leanest in the sand? In *Universal Access in Human-Computer Interaction. Access to Learning, Health and Well-Being*. Springer, 2015, pp. 364–373. 74

[145] Salmon, J., and Nyhan, J. Augmented reality potential and hype: towards an evaluative framework. *The Journal of Language Teaching and Learning 3*, 1 (2013), 54–68. 16

[146] Samsung. Samsung VR. `http://www.samsung.com/global/microsite/gearvr/gearvr_features.html`. Online; accessed 20-September-2015. 19

[147] Sauro, J., and Kindlund, E. A method to standardize usability metrics into a single score. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2005), CHI '05, ACM, pp. 401–409. 95

[148] Schultz, C. Digital keying methods. *University of Bremen Center for Computing Technologies, Tzi 4*, 2 (2006), 3. 27, 108, 109

[149] Seraku, C. Smart Washbasin. `http://smart-washbasin.seraku.co.jp/english/`. Accessed: 2015-01-23. 23

[150] Serio, A. D., Ibáñez, M. B., and Kloos, C. D. Impact of an augmented reality system on students' motivation for a visual art course. *Computers & Education 68*, 0 (2013), 586 – 596. 61

[151] Serra, A., Dessì, T., Carboni, D., Popescu, V., and Atzori, L. Inertial navigation systems for user-centric indoor applications. *Networked and Electronic Media Summit, Barcelona* (2010). 28

[152] Sharma, S., Jerripothula, S., Mackey, S., and Soumare, O. Immersive virtual reality environment of a subway evacuation on a cloud for disaster preparedness and response training. In *Computational Intelligence for Human-like Intelligence (CIHLI), 2014 IEEE Symposium on* (Dec 2014), pp. 1–6. 20

[153] Soga, M., Matsui, K., Takaseki, K., and Tokoi, K. Interactive Learning Environment for Astronomy with Finger Pointing and Augmented Reality. In *Advanced Learning Technologies, 2008. ICALT '08. Eighth IEEE International Conference on* (2008), pp. 542–543. 18

[154] Sohail, A. S. M., and Bhattacharya, P. *Signal Processing for Image Enhancement and Multimedia Processing*. Springer US, Boston, MA, 2008, ch. Detection of Facial Feature Points Using Anthropometric Face Model, pp. 189–200. 140

[155] Sorrentino, F., and Scateni, R. Talking heads on mobile devices. In *Proceedings of the 3rd Symposium on Facial Analysis and Animation* (2012), ACM, p. 8. 106

[156] Sorrentino, F., and Scateni, R. Thal-k: Talkinghead animation library. In *Smart Tools and Apps for Graphics-Eurographics Italian Chapter Conference* (2013). 106

[157] Sorrentino, F., Spano, L., and Scateni, R. Speaky notes learn languages with augmented reality. In *Interactive Mobile Communication Technologies and Learning (IMCL), 2015 International Conference on* (Nov 2015), pp. 146–150. 36

[158] Sorrentino, F., Spano, L., and Scateni, R. Superavatar children and mobile tourist guides become friends using superpowered avatars. In *Interactive Mobile Communication Technologies and Learning (IMCL), 2015 International Conference on* (Nov 2015), pp. 222–226. 36

[159] Sorrentino, F., Spano, L. D., and Scateni, R. Interactive shops: how the customer can deal with them both from inside and outside. In *Proceedings of the 11th Biannual Conference on Italian SIGCHI Chapter* (2015), ACM, pp. 78–81. 74

[160] Styliani, S., Fotis, L., Kostas, K., and Petros, P. Virtual museums, a survey and some issues for consideration. *Journal of Cultural Heritage 10*, 4 (2009), 520 – 528. 21

[161] Sun, J., Jia, J., Tang, C.-K., and Shum, H.-Y. Poisson matting. *ACM Trans. Graph. 23*, 3 (Aug. 2004), 315–321. 26, 108

[162] Swartout, W., Traum, D., Artstein, R., Noren, D., Debevec, P., Bronnenkant, K., Williams, J., Leuski, A., Narayanan, S., Piepol, D., et al. Ada and grace: Toward realistic and engaging virtual museum guides. In *Intelligent Virtual Agents* (2010), Springer, pp. 286–300. 17

[163] Theall, M. New directions for theory and research on teaching: A review of the past twenty years. *New directions for teaching and learning 1999*, 80 (1999), 29–52. 60

[164] Toossi, M. Employment outlook: 2008-18-labor force projections to 2018: Older workers staying more active. *Monthly Lab. Rev. 132* (2009), 30. 8

[165] Toshiba. Toshiba to Unveil Leading-edge Technologies at CES 2014. `http://www.toshiba.co.jp/about/press/2014_01/pr0702.htm`. Accessed: 2015-01-26. 23

[166] Turk, M., and Pentland, A. Face recognition using eigenfaces. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on* (Jun 1991), pp. 586–591. 27

[167] Tuscany+, F. S. T. . 3d software. 30

[168] Tuveri, E., Iacolina, S. A., Sorrentino, F., Spano, L. D., and Scateni, R. Controlling a planetarium software with a kinect or in a multi-touch table: a comparison. In *Proceedings of the Biannual Conference of the Italian Chapter of SIGCHI* (2013), ACM, p. 6. 36

[169] Unity. Unity3D. `https://unity3d.com/`, 2015. Online; accessed 20-September-2015. 76

[170] Vazquez-Fernandez, E., Garcia-Pardo, H., Gonzalez-Jimenez, D., and Perez-Freire, L. Built-in face recognition for smart photo sharing in mobile devices. In *Multimedia and Expo (ICME), 2011 IEEE International Conference on* (July 2011), pp. 1–4. 27

[171] Velho, L., Frery, A., and Gomes, J. *Image Processing for Computer Graphics and Vision.* Springer London, London, 2009, ch. Warping and Morphing, pp. 387–412. 138

[172] Viehdorfer, M., Nemanic, S., Mills, S., and Bailey, M. Virtual dog head: Using 3d models to teach complex veterinary anatomy. In *ACM SIGGRAPH 2014 Studio* (New York, NY, USA, 2014), SIGGRAPH '14, ACM, pp. 17:1–17:1. 20

[173] Villagrasa, S., Fonseca, D., and Durán, J. Teaching case: Applying gamification techniques and virtual reality for learning building engineering 3d arts. In *Proceedings of the Second International Conference on Technological Ecosystems for Enhancing Multiculturality* (New York, NY, USA, 2014), TEEM '14, ACM, pp. 171–177. 20

[174] Wigdor, D., and Wixon, D. *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*, 1st ed. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2011. 49

[175] Wikitude. Augmented reality sdk. 29

[176] Wolfson, N. E., Cavanagh, T. M., and Kraiger, K. Older adults and technology-based instruction: Optimizing learning outcomes and transfer. *Academy of Management Learning & Education 13*, 1 (2014), 26–44. 8

[177] WRIGHT, S. *Digital Compositing for Film and Video with Cdrom.* Butterworth-Heinemann, Newton, MA, USA, 2001. 26, 108

[178] YONG, C. Y., SUDIRMAN, R., AND CHEW, K. M. Facial expression monitoring system using pca-bayes classifier. In *Future Computer Sciences and Application (ICFCSA), 2011 International Conference on* (June 2011), IEEE, pp. 187–191. 93

[179] YOUNGBLUT, C. Educational uses of virtual reality technology. Tech. rep., DTIC Document, 1998. 19

[180] ZHAO, W., CHELLAPPA, R., PHILLIPS, P. J., AND ROSENFELD, A. Face recognition: A literature survey. *ACM Comput. Surv. 35*, 4 (Dec. 2003), 399–458. 27

[181] ZICHERMANN, G., AND CUNNINGHAM, C. *Gamification by design: Implementing game mechanics in web and mobile apps.* " O'Reilly Media, Inc.", 2011. 75

[182] ZIJLSTRA, R., AND VAN DOORN, L. The Construction of a Scale to Measure Subjective Effort. Tech. rep., Delft University of Technology, Department of Philosophy and Social Sciences, Delft, Netherlands, 1985. 31

[183] ZIJLSTRA, R., AND VAN DOORN, L. The Construction of a Scale to Measure Subjective Effort. Tech. rep., Delft University of Technology, Department of Philosophy and Social Sciences, Delft, Netherlands, 1985. 51, 125

[184] ZOMBIELINK. Zombies, Run! `https://zombiesrungame.com/`, 2015. Online; accessed 20-September-2015. 22

[185] ZOOBURST. ZooBurst. `http://www.zooburst.com`. Online; accessed 04-May-2012. 16

[186] ZYWICA, J., AND DANOWSKI, J. The faces of facebookers: Investigating social enhancement and social compensation hypotheses; predicting facebook^{TM} and offline popularity from sociability and self-esteem, and mapping the meanings of popularity with semantic networks. *Journal of Computer-Mediated Communication 14*, 1 (2008), 1–34. 120