Permeating the Architectural Past in 4D: an Augmented Reality Interactive System

Gabriel Riboldi Anderson Maciel
Laboratório de Tecnologias Digitais – Universidade de Caxias do Sul (UCS)
Campus 8, Rod. RS 122, km 69, s/n CEP 95010-550 – Caxias do Sul – RS – Brazil
{griboldi, amaciel5}@ucs.br

Abstract

Important buildings and urban sites of the recent past were not adequately documented and have been forgotten for a long time. In this context, new experiences with the 3D modeling and simulation of such spaces in VR are becoming very impactful as documentation tools. However, these virtual spaces are not accessible to the general public as visualization tools are not available. The purpose of this work, then, is to create an interaction environment in augmented reality to explore historical areas in such a way that ancient versions of buildings can be visualized and explored directly in the real space where the buildings were in the past or their current versions are situated today. Users handling a mobile display device, as a tablet PC, walk around the real site, and as they point the display towards a neighboring building, they can see how it was in the past, which allows a travel in time, offering a fourth dimension to the experience. The results obtained demonstrate the potential of augmented reality applications for the dissemination of historical heritage.

1. Introduction

Augmented reality (AR) integrates concepts and techniques from areas such as computer vision, computer graphics and image processing. It amplifies the view a user has from the real world by means of the combination of 3D virtual objects with real world objects. Moreover, in this process, the user keeps a feeling of presence in the real world, in opposition to virtual reality (VR) where the user is fully immersed in the virtual environment and loses contact with the real environment.

In this work, we apply AR as a tool to amplify the current vision of the historical heritage in the city of *Caxias do Sul*, providing a view to the architectural past coupled to the real physical space.

Figure 1: Historical heritage can be better known and preserved with the assistance of an AR system.

Important buildings from the later years of the 19th century and the first decades of the 20th century were not adequately documented, and this matter have been forgotten for a long time by historiography. In this direction, experiences with the simulation of architectural and urban spaces in AR is innovative and can be very impactful. Thus, a totally new attention is now directed to the reconstruction and simulation of historical environments aiming at the education in cultural heritage, the construction of historical documentation and the preservation of the historical heritage. Architecture researchers are obtaining excellent results in the reconstruction of such environments, especially in modeling specific historical buildings in 3D [4]. Detailed models of these buildings are now available, but they are still inaccessible to the general public as the necessary tools to visualize and explore such virtual historic spaces are not available.

In face of this problem, the general goal of the present work is creating an augmented reality environment in which ancient versions of buildings can be visualized and explored directly in the real space where the buildings were in the past or their current versions are situated today. More specifically, we focus on providing AR tools to navigate in a virtual environment of the central area of the city of *Caxias do Sul*. Users handling a mobile display device, as a tablet PC, walk around the real *Dante Alighieri* Square, and as they point the display towards a neighboring building, 3D models depicting ancient constructions are superimposed and they can see how the building was, for example, in the year of 1920 (Figure 1). The date can be chosen interactively, which allows a travel in time, offering a fourth dimension to the experience.

In the development of the graphical environment, we used the augmented reality library irrAR, an integration of the ARToolkit and the game engine irrLicht. For display we used a mobile device display with good processing power (tablet PC), and a good webcam (PS3Eye) for video tracking the user position in relation to the real environment.

A study about these techniques and elements, as well as the system implementation, are described in the remaining of this paper. Section 2 presents some basic concepts of virtual and augmented reality and overviews the related work in the area of virtual historical heritage. Section 3 details the system implementation including how the 3D models of buildings used in this work have been acquired, and section 4 presents tests and the preliminary results obtained so far. Finally, in section 5 we discuss the results, present conclusions and ideas for future work.

2. Background

2.1. Virtual and augmented reality

The term virtual reality VR is defined by Ivan Sutherland as: "... a computer simulated world in which a user can navigate in real time as if they were moving in a real three-dimensional space" [11]. The basis of VR arose in the 1970's in flight simulators, where a synthetic simulated world was created based on a real environment [6]. To allow the feeling of presence in such virtual worlds, VR integrates sophisticated devices as datagloves [13], stereo glasses, biofeedback sensors, etc.

Later, Paul Milgram introduced the virtuality continuum, a continuous line upon which the real world is placed in the left side and the virtual world in the right side. Between them, the continuum is generically defined as mixed reality, and represents an environment in which the objects from the real and the virtual worlds are presented together in the same context [9]. Yet, in the continuum, mixed reality is subdivided into: augmented reality and augmented virtuality. Augmented Reality (AR) is a well known term, many times used in place of mixed reality. AR can also be defined as an improvement of the real world with computer generated elements, or a system which supplements the real world with virtually generated objects coexisting in the same space [1].

Augmented reality requires capturing the real world to the computer system (computer vision). This is usually made in AR using a video camera or a webcam. The basic requirements in the construction of an AR system are tracking, registration and display [2].

In tracking, the first step is capturing a video frame. The frame is then analyzed and patterns are recognized, which can be markers (tags) or pre-established information (like contours and edges). When markers are used, the approach is called marker-based. Artificial patterns are placed along the real environment (fiducial markers) for camera pose estimation. On the other hand, markerless approaches use patterns of the real scene captured by the camera (colors, lines, etc.).

When any of these patterns are found, the system registers their positions and calculates the transformation of the pattern on the frame of the camera.

Then, the three-dimensional virtual object is associated to the pattern – according to the estimated pose (position and orientation) registered for the pattern – and can be displayed at that emplacement. Tracking, registration and display procedures are repeated continuously during the video capture.

2.2. Virtual historical heritage

Visualization and interaction technology for immersive environments are in constant development. Complex and realistic virtual worlds can now be handled thanks to significative advances in computer graphics and simulation. Since VR has become possible, it has been assisting professionals in several areas of spatial exploration, as architecture. Historic reconstruction of monuments and buildings of the past and which are fading with time is now possible with 3D modeling software and VR tools for navigation. Moreover, augmented reality is there to make possible the experience of studying and understanding such environments which no longer exist right in the place they once were.

Research is being conducted all around the world which is producing reconstituted models, scenario animations, ancient environments, historical objects and buildings [5] [8]. All this media content generate a technological cosmos which favors new developments in VR and AR.

The ARCHEOGUIDE project [12] is one of the pioneers to combine AR and historical heritage. The project used GPS and compass to define the pose of a user in the Olympia site in Greece, and render historical and touristic information to users. The project, nevertheless, relies on heavy hi-tech equipment which despite its high cost is not accurate and depends on a considerably complex infrastructure.

The project ARTHUR (Augmented Round Table for Architecture and Urban Planning) of Fraunhofer Institut consists in a collaborative round table in AR which replaces drawings, plans and maquettes for building planing [3]. Without physical models, architects and clients modify their computer generated virtual models collaboratively around a common table (see Figure 2a). Gestures with fingers or a stylus can be used for object manipulation implementing a magic-wand approach. One inconvenient requirement of the project is that every user must wear a head mounted display (HMD). This can drastically limit the widespread use of the final application.

The University of Sao Paulo (USP), developed an immersive VR environment in which the user can take a walk through a virtual model of the *Sé* square in the year of 1911 (see Figure 2b). The enormous urban changes perpetrated since the 1880's motivated the choice of that specific site. As the population quadruplicated in about 10 years, a number of structural changes in the city remodeled the traditional city center. A number of references, including iconographic, cartographic and documental sources have been used for 3D reconstruction with Maya (Autodesk). The reconstructed model has been then visualized in a CAVE environment for a virtual walkthrough in the ancient urban space [14].

A major outcome from the project was to allow visualizing urban regions which cannot be conveniently understood along the temporal axis due to long intervention processes. The proposal has become part of the urban iconography, and is now subjected to criticism, appropriations and new developments. VR can then stimulate studies and discussions about the urban history and culture. It can also be part of the collection of existing documents, impacting in the management and availability of information for both researchers and inhabitants of the city.

In another approach, the project *Brasília 3D* [10], depicts the capital city in an electronic maquette showing details of its monuments and the architecture with textual description (see Figure 2c). Users can go for a virtual ride by car or helicopter choosing the spots to be visited. The goal is to encourage education, regional development and civic and cultural tourism using three-dimensional immersion and interactivity.

These projects are examples of AR and VR applications in the urban space. However, none of them is totally available to the public because they require complex equipment, and none of them can be explored in a 1 to 1 scale with the real urban space they are representing. Yet, the fast development in computer graphics and the availability of new mobile devices – with a constant decrease of cost and increase of usability – creates a fertile ground for the development of new applications in the area of virtual historical heritage.

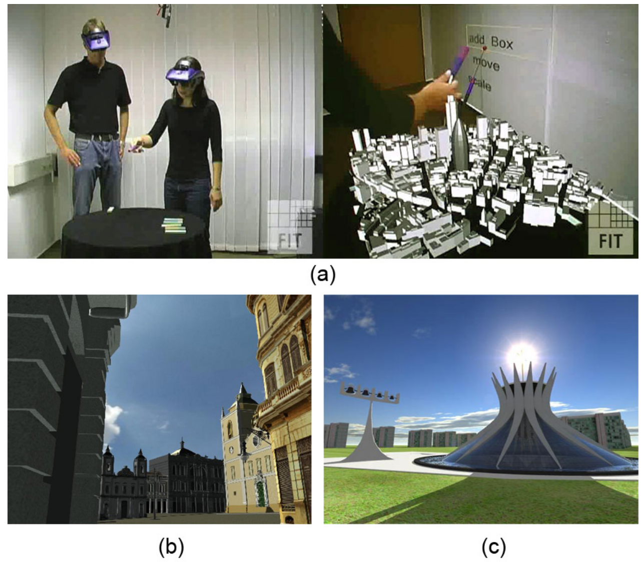


Figure 2: Previous work on virtual urban spaces: a) Augmented Round Table for Architecture and Urban Planning [3]; b) VR model of the *Sé* square in the year of 1911 [14]; c) *Brasília 3D* [10].

In the next section, we describe our approach to popularize the access to historical heritage using AR.

3. Design and implementation of the historical heritage augmented reality system

In the system implementation we used C++ with the game engine irrLicht. The libraries ARToolkit and irrAR are also integrated, as described in the subsections below. As for the 3D models, they have been prepared with 3D Studio Max and imported as a irrLicht scene.

3.1. Third party software

3.1.1 ARToolkit

In this work we used the ARToolkit library and fiducial markers (tags) as auxiliary elements to capture positions and prove our concept. The use of tags has demonstrated great practical results and is fast enough for real time applications. Tags are images with easy to extract visual features. They are printed in paper and do not need any electronic device associated. The most used resource based on tags for AR is the ARToolkit. It is popular because it offers 3D tracking, demands low computational cost and is opensource [16], incentivizing users to modify it according to their applications. The typical use of ARToolkit is in positioning and orienting virtual elements within a real world scenario, becoming an interaction mean between users and applications. A simple design for an AR application with the toolkit, which is nicely described in their website, consists of the following:

- start video configuration, reading markers database and camera parameters;
- capture a video frame;
- detect and identify markers on the video;
- calculate the marker transformation in relation to the camera;
- draw virtual objects using the pose defined by this transformation.

With ARToolkit, the visualization of a virtual object occurs when the associated marker appears on the video camera field of view. The display can be made in an indirect fashion (non-immersive) or direct (immersive). Monitors are often used as indirect visualization devices. They play the role of mirrors "reflecting" the real world augmented with real elements, placing the AR environment out of the user's action space. On the other hand, direct visualization devices, as HMDs, mix the two realities directly in the user's action space [7].

3.1.2 irrLicht

irrLicht is an open source game engine implemented in C++. Game engines are computer programs and sets of libraries generally used to simplify the development of games and other real time graphics applications. Basic functionalities of irrLicht as a game engine are a graphics package for rendering, a physics package for collision detection and physics simulation, support to sound generation, support to artificial intelligent behavior and so on.

irrLicht supports a large variety of formats as 3ds, MD2, Maya, obj, Quake 3. Light sources, virtual camera and 3D models are organized in a hierarchy of scene nodes, which can be controlled by predetermined animation routines or manually by the user/programmer.

3.1.3 irrAR

irrAR is a library for AR development which integrates the irrLicht engine with the ARToolkit. irrAr increases the potential of AR applications allying the ease of scene managing provided by the engine with the possibility to independently assign scene nodes to markers.

3.2. The graphic environment

The methodology for creating the graphic environment started with our colleagues architects. The first step was a documental and iconographic research to build a chronological evolution hypothesis of every element around the square. From the hypotheses, they defined phases, comprehending the period starting with the first edification in a specific spot, passing through changes in elements of the composition, until demolition and new edification. After that, modeling of each phase follows the procedures defined in Figure 3: (a) volumetric simplification, (b) formal interpretation, (c) measuring, (d) general 3D model, (e) detailed 3D model and (f) filling the gaps. Modeling software as Blender, 3D Studio Max e Sketchup have been used [4].

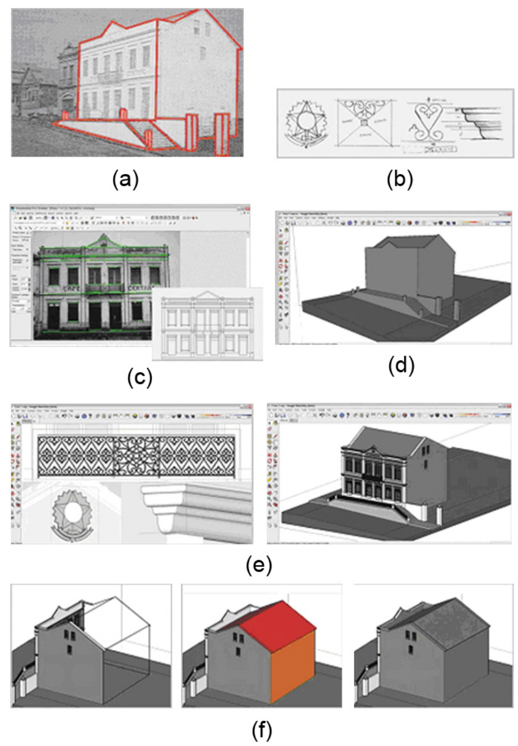


Figure 3: Modeling procedures.

For this AR project we had to rework the models. First, it was convenient to convert all model elements into triangle meshes. Then, we reduced the number of triangles for efficiency, but always taking care not to lose important geometric details. We used optimizer functions of 3D Studio to accomplish this task and then exported the mesh files in a format we could load with the irrLicht engine.

The models corresponding to four construction phases were loaded into the engine environment. We named them: year 1899, year 1920, year 1932 and year 1947. Anticipating that the system would have to display only one of the phases at each time, we worked on a transition strategy based on transparency with the alpha channel. In such a way, we would be able to provide a smooth transition between phases by interpolating the alpha channels of the previous and next phase to be displayed from 255 to 0 and 0 to 255 respectively, simulating a fade in/out. With irrLicht this is made using a texture to define a transparent material to each mesh which alpha value can be controlled.

3.3. Tracking and pose determination

An ARToolkit marker (tag) and a webcam (PS3Eye) were used for motion tracking. Tracking consists in detecting the presence of the marker and calculating its pose in relation to the webcam. The video processing by computer vision establishes the relation between marker and camera coordinates. Such relation is later used to compose the transformation between the virtual camera and the models for AR rendering.

We placed the marker at a fixed position in the real environment in such a way that it does not occlude important parts of the video like important detail of the real buildings.

3.4. Calibration

The first time a marker is placed and every time it is relocated in the real world, the relation between the marker and the virtual world must be accordingly re-established. We implemented this through a manual calibration mode. In this mode, the user can use scroll bars on the touch screen to change scale and translation parameters by simply sliding with the fingers. Calibration parameters are saved and do not need to be recalculated provided that the marker is not moved.

3.5. Navigation

Since calibration is done, the user can navigate the 3D space by moving the display (tablet PC) around the real environment. The only requirement is that the marker is kept in the range of the camera adapted to the tablet. Then, the position and orientation of the display, stored in a 4x4 matrix by ARToolkit, is constantly updated by the system using the fixed marker as a reference. While the position is updated, new frames are rendered in real time and exhibited together with the background video, as described in section 3.6 below.

3.6. Rendering and shading

After a phase (year) is selected and the camera parameters are calculated from the marker, the scene can be rendered. The next step then is to configure the shading parameters. Shading is important because the virtual scene must match the more exactly possible the visual aspect of the real scene for a successful fusion in AR. In the current implementation, we take into account the date and time of the day to simulate the sunlight reflecting on the virtual buildings.

The position of a light source in the environment is defined through a system of spherical coordinates which allows us to estimate the position of the sun in relation to the scene. In such system, the cartesian coordinates (x, y, z) of any point can be computed from the spherical coordinates (r, Φ, Θ) using equation 1. r is the radius of the sphere and can be seen as the distance between the scene and the sun. Φ is the elevation angle an can be seen as the latitude angle for the sun. Θ is the azimuth angle, which in turn, can be seen as the longitude of the sun.

$$\begin{cases} x = r \sin\Phi \cos\Theta \\ y = r \sin\Phi \sin\Theta \\ z = r \cos\Phi \end{cases} \tag{1}$$

However, some changes have been necessary to adequate the cartesian axes to the reference system of our models and also to simplify the implementation, using one angle to the time of the day and the other to the seasons of the year. The system currently in use is the one of equation 2. Finally, we read the date and time of the operating system and use linear interpolation to define the appropriate angles. Of course, we had to take into account the geographic coordinates of the city of *Caxias do Sul* to determine the angle ranges for every season. Such coordinates are $29^o10'4"S$ $51^o10'44"W$. For the time of the day we used a simplification in this proof of concept, supposing sunrise at $6am$ and sunset at $6pm$.

$$\begin{cases} sun_x = -r \cos\Theta \\ sun_y = -r \cos\Phi \sin\Theta \\ sun_z = -r \sin\Theta \end{cases} \tag{2}$$

Using this approach, the system renders a scene predominantly lit by a sun rising from the east and setting to the west, which matches the real scene any time of the day. Other light components as ambient light are also set for realism. Ambient light simulates the indirect light received by faces not turned to the sun (reflected by other elements as other buildings and the atmosphere). Projected shadows are not used in this implementation.

3.7. Interaction with the mobile display

Along with 3D navigation and display, the tablet PC is also used to navigate in a fourth dimension: time. Using the touch screen the user can slide their finger on a timeline slider bar placed vertically on the left of the display area. In practice, the system starts showing the scenario as it was in the year 1899. To explore the local architecture in other years, the user rolls the slider with a finger. The system then switches smoothly to the scenario of the year 1920, for example, using a fade-in/out effect (Figure 4). As a year is chosen, additional descriptive and historical information can also be displayed on the text area at the bottom of the screen.

Figure 4: Interaction through the navigation bar placed in the left side of the touch screen. The user changes the year with a slide of a finger.

4. Results

In this section we describe tests performed with the system and show some of the results obtained so far. As already stated, our case study has been the urban space around the central square *Dante Alighieri* in the city of *Caxias do Sul*.

4.1. Qualitative evaluation

For all tests during development and for demonstrations in the lab we used a maquette of the most important building in our scenario: the 110 years old Saint Theresa Cathedral. The maquette has been constructed in a $40 : 1$ scale and a tag was placed on the front wall for calibration and tests.

In calibration mode we used the sliders through the touch screen to control scale and translation parameters in such a way that a direct spatial correspondence has been established between the maquette and the 3D model. The parameters are then saved and we could start navigation.

Figure 5 depict the visual results. The interaction was smooth, without flickering and very intuitive for users with experience in 3D modeling tools and games. The tag has always been detected, even at sharp angles, being lost only when the lighting conditions were very poor.

A second test has been performed in the actual site of the Cathedral. A tripod was used to place the tag at a convenient fixed position between the user and the building. New calibration was required due to the different scale and position of the tag in relation to the scene. Here we noticed that, as manual calibration can take some time, the 2 kg weight of the tablet PC becomes tiresome. While this situation is noticed during calibration, the weight does not bother the user in general navigation. Although automatic calibration could solve this problem by using two tags or detailed measurement of the relative tags positions, we believe that as calibration is a one time procedure, the current manual approach is generally convenient.

The tests around the real cathedral are illustrated in Figure 6. The user could easily navigate and interact to see how construction details changed along time. It was also possible to analyze neighboring buildings, as the Canonical House, situated in the right side of the Cathedral.

4.2. Quantitative evaluation

This quantitative evaluation aims at verifying the mobile hardware capability to process all tasks required in real time.

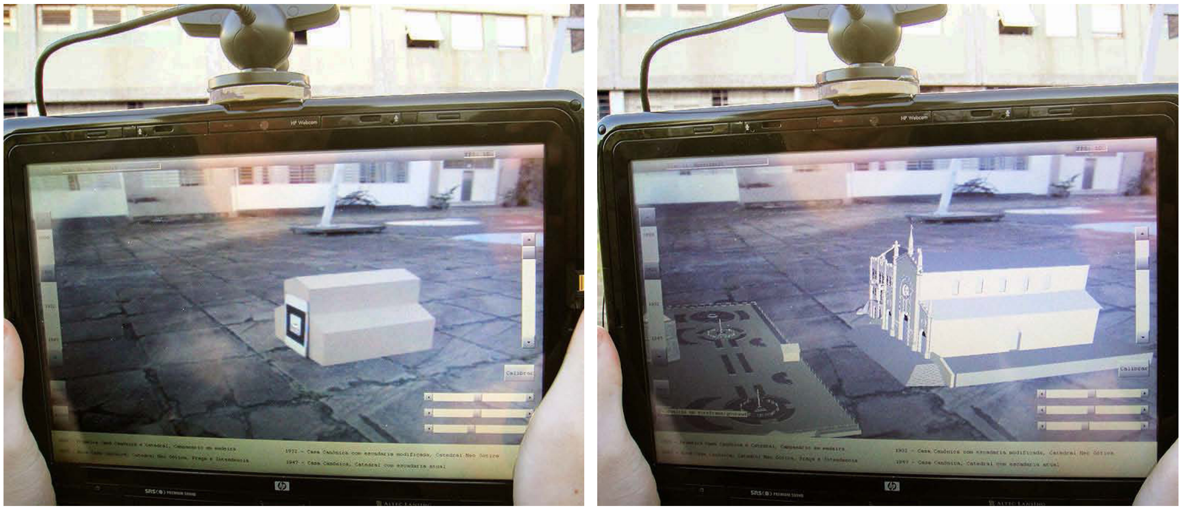
All tests used the full 4D model. This means that all four 3D scenes, one corresponding to each year and version of the buildings, are loaded and rendered. While the selected scene is presented with the alpha channel set to 255, the remaining three scenes are also rendered with alpha channel set to 0. This makes a total of 235,040 triangles, a value clearly above the necessary regarding the scene complexity, which was used in purpose to press the hardware to the maximum.

To impose strict constraints in computation power from the beginning, a modest PC was used during development. It is an Intel Pentium D 2.80 GHz, 1GB RAM and nVidia GeForce 7100 GS GPU.

With this development configuration we obtained a rate of 5 frames per second in the final display set to a resolution of 1280x1024 pixels. This is a poor frame rate, but it made us confident that we would reach real time with the final target device, a tablet PC with CPU AMD Turion X2 Ultra Dual-Core Mobile ZM-82 2.20GHz, 4GB RAM and ATI Radeon HD 3200 GPU. Indeed, the application run at 15 to 20 frames per second, which provides smooth interaction. Performance data are depicted in the chart of Figure 7.

5. Conclusion and Discussion

The system developed is a new tool for visualization and exploration of the urban historical heritage. Relying on computer vision techniques, graphics and game libraries,



(a) (b)

Figure 5: System being tested with a maquette of the cathedral building.

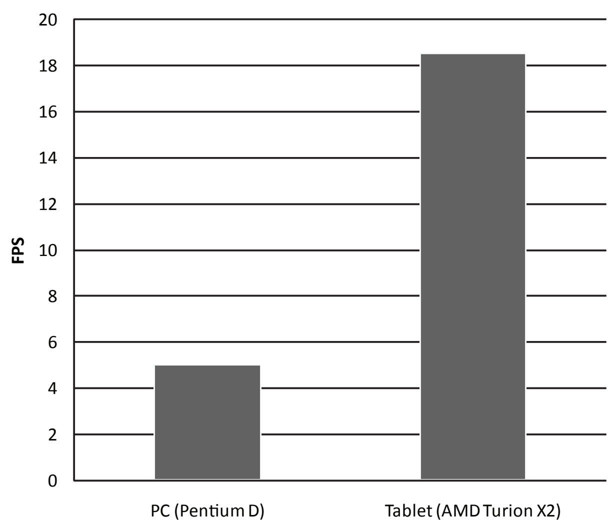


Figure 7: Performance measured with a modest PC used for development and with the tablet PC targeted for the final application. Both results have been obtained with a worst case scenario with over $230k$ triangles.

and new technology for visualization, we configured procedures and instruments to create an augmented reality environment accessible to the general public. AR techniques enabled real time interaction with 3D scenarios of the past, aggregating value to the already widely available models of ancient buildings in other areas as architecture and historiography.

This work is highly multidisciplinary. From the results obtained so far, we visualize a great potential for applications of VR and AR in the knowledge dissemination and preservation of the historical heritage, not only in *Caxias do Sul*, our case study. The results prove the concept, and if the system is further extended around the common goal of maintaining the architectural heritage present, it might become an important integration factor between the areas of historiography, architecture, VR and AR.

As immediate future work, we propose the use of markerless computer vision tracking for AR. This would give greater freedom to users as it does not require to spread a number of tags in the environment, which would end by generating visual pollution. We suggest the use of an edge based video technique, as edges are abundant in any edification. Processing edge detection algorithms in current mobile hardware in real time is still challenging, but we may expect that mobile hardware will increase processing power soon. Without worrying with the tags, the users could move more freely and stay in tune with the mobility trend of our time.

Other future work would be in experimenting new devices. Augmented reality glasses (see-through) combined with some sort of hand held interaction device as the *Wiimote* could increase the feeling of presence and be much less cumbersome systems than tablet PCs. This is also true when cell-phones can be used, and they are already in our pockets to be used anywhere in the world. In this case, while GPS and video tracking could precisely determine the user pose, the local 3D (or 4D) data must be available for download at any time, just as maps are available today with Google Maps. This would be a great tool and an innovative attraction for tourism and related areas.

(a) (b)

Figure 6: System being tested in the actual site of the *Dante Alighieri* Square. Photographs show the Saint Theresa Cathedral and a “Hiro” tag fixed with a tripod.

Acknowledgments

Thanks are due to the professor Jaqueline Viel Caberlon Pedone for the reconstructed 3D models. Thanks are also due to CNPq-Brazil for the support through processes 481268/2008-1, 481762/2008-6 and 302679/2009-0, and to Rio Grande do Sul State research foundation for the grant BIC/FAPERGS.

References

[1] R. Azuma. Overview of augmented reality. In *SIGGRAPH '04: ACM SIGGRAPH 2004 Course Notes*, page 26, New York, NY, USA, 2004. ACM.

[2] D. A. Bowman, E. Kruijff, J. Joseph J. LaViola, and I. Poupyrev. *3D User Interfaces Theory and Pratice*. Ed Addison-wesley, Boston, MA, 2005.

[3] W. Broll. The augmented round table - a new interface to urban planning and architectural design. In *INTERACT*, 2003.

[4] A. Cattani, A. Antoniazzi, J. V. C. Pedone, M. Stumpp, and R. Osmainschi. Criação de modelos virtuais para registro do patrimônio histórico: estudo de caso com softwares livres. In *SIGraDi 2008 - Proceedings of the 12th Iberoamerican Congress of Digital Graphics*, 2008.

[5] A. Foni and N. Magnenat-Thalmann. A virtual heritage case study: an approach to the revival of ancient sites through application of 3d real-time computer graphics. In *Proc. Workshop on Augmented Virtual Reality (AVIR'03)*. MIRALab, University of Geneva, September 2003.

[6] T. Furness. Super cockpit amplifies pilot's senses and actions. pages p.76–77, 1988.

[7] C. Kirner and R. Tori. Fundamentos de realidade aumentada. In *Fundamentos e Tecnologia de Realidade Virtual e Aumentada - Pré Simpósio SVR 2006*, pages 22–37, Belém, 2006. SBC.

[8] N. Magnenat-Thalmann and G. Papagiannakis. Virtual worlds and augmented reality in cultural heritage applications. In *Proc. Recording, Modeling and Visualization of Cultural Heritage, Baltsavias et al. (eds)*, pages 419–430. Taylor and Francis Group, January 2006.

[9] P. Milgram, H. Takemura, A. Utsumi, and F. Kishino. Augmented reality: A class of displays on the reality-virtuality continuum. In *Proceedings of the SPIE Conference on Telemanipulator and Telepresence Technologies*, volume 2351 of *Proceedings of SPIE*, pages 282–292, Boston, Massachusetts, USA, November 1995.

[10] A. Rangel. Brasília 3d, 2005. [Online; accessed 15-setembro-2009].

[11] I. E. Sutherland. *Seminal graphics: poineering efforts that shaped the field*, chapter A head-mounted three dimensional display, pages 295–302. ACM, New York, NY, USA, 1998.

[12] V. Vlahakis, J. Karigiannis, M. Tsotros, M. Gounaris, L. Almeida, D. Stricker, T. Gleue, I. T. Christou, R. Carlucci, and N. Ioannidis. Archeoguide: first results of an augmented reality, mobile computing system in cultural heritage sites. In *VAST'01: Proceedings of the 2001 conference on Virtual reality, archeology, and cultural heritage*, pages 131–140. ACM, 2001.

[13] T. G. Zimmerman, J. Lanier, C. Blanchard, S. Bryson, and Y. Harvill. A hand gesture interface device. In *CHI '87: Proceedings of the SIGCHI/GI conference on Human factors in computing systems and graphics interface*, pages 189–192, New York, NY, USA, 1987. ACM.

[14] M. K. Zuffo and R. de Deus Lopes. Ambientes de realidade virtual e realidade aumentada na preservação do patrimônio histórico. In *Computação gráfica: pesquisas e projetos rumo à Educação Patrimonial (workshop)*, São Paulo, 2008. DPH.