3D Annotations in Immersive Environments

Francisco J. Pinzón L.

Universidad de los Andes
Departamento de Ingenería de Sistemas y Computación

Ecole Nationale Superieure d'Arts et Métiers. Paris Tech.
Institut Image

July 2011

USA.

0 Brief

The requirement of a 3D Annotation System for Immersive Environments with telecooperation (not concurrent) capabilities was a basic and unfulfilled need present at the Institute Image when I arrived. For this reason I was given the task of developing such system as my undergrad thesis. The use of the immersive facilities present at the Institute in the provided solution was a requirement, as well as keeping the compatibility with standard computers. The provided application succeed in the fulfillment of the original requirement. The following document describes this solution as well as the considerations taken in account for developing it.

Smoke whirls
After the passage of a train.
Young foliage.
Shiki Masaoka (1867-1902)

1 Introduction

For a long time now I've wanted to visit France. Actually, thinking about it, I had that desire before I started to study Systems and Computer Engineering at Universidad de los Andes. Someday I just asked Tiberio about the possibilities of making my thesis abroad. He gave me certain options and one of them was the Institut Image at Chalon-sur-Saône (in which I am writing this Introduction). Since then a few months passed before the bureaucratic nightmare started.

From the university demanding information they already had, and for which I had to pay a considerable amount of money; to the French Embassy and their Kafkesque process to give me a 5 minutes appointment after weeks of collecting innumerable documents, without taking in account my travel to Brazil to participate in Interactivos? 2010 BH, Christmas and New Year's eve, I thought I wouldn't be able to fly to France in time. After all it actually happened.

The subject of this work was given to me by Frédéric two weeks after my arrival at Chalon. In standard circumstances I prefer my projects to be product of my ideas and not someone else's, but the opportunity to come here for a whole semester easily took over that. Looking backwards, this project has been a great learning opportunity, both personal and professionally.

The instructions for developing my project were simple and precise: A system for making annotations on 3D models using the cave facilities present at the institute, however usable in a standard computer. The idea is to be able to comment collaboratively and in distant locations a 3D model. The annotations would be simply a plain text comment, with an author and a priority attached.

So, as you can see, rather than the traditional way to choosing a thesis I was given a determinate topic and a narrow and precise expected solution. For this reason, in this project, I will work on the possibilities, contexts in which it might be used, constraints and potential scopes of the given solution.

Some of the basic design choices of the solution were given by the facilities present at the Institute. Windows XP running in the cave computers, infrared tracking cameras, passive stereo projections and the Android powered tablet are some examples of these design choices. The software was developed using Microsoft Visual C++ for the software running in the servers and the Eclipse IDE for the tablet application; thus C++ and Java were used to build the solution. OpenSceneGraph and VRPN were also used.

Aca falta **Diseño e implementación, Resultados obtenidos. Estructura del documento (cómo leerlo y seguirlo)**

First of all I thank my mother and my sister for supporting me, I know it have been tough. I also thank Tiberio and Frédéric for supervising this work, and giving me this opportunity. Luisa for her friendship and love; David for his friendship and ¿?, and both for their support and our shared memories. Sadly I can't thank God because I don't believe in him, nevertheless I can thank Alan Turing for making this work possible.

2

General Description

2.1 Goals

The objective of the project is to create a 3D Annotation System for Immersive Environments without concurrent tele-cooperation capabilities and which also might be used in a standard computer. The projected is developed for fulfill the need of such system at the facilities of the Institut Image and at the COLIVRI Laboratory at the Universidad de los Andes; without being restricted to these two contexts.

The final solution must be capable of performing this task in models of a variety of 3D formats, associating with each annotation its type and urgency to be managed. On the other hand the system meet the ever-present needs of scalability, flexibility, performance and specially usability. The project must be developed using C++ and OpenSceneGraph and must run in Windows XP.

2.2 Prior Work

The most similar prior work I found while making the research for this work is the one presented by Kadobayaby et al. in [9]. It's similarity consists in the use of annotations in 3D immersive environments to deal with the inherent complexity of 3D models. The work is nevertheless, restricted by it's use of Croquet. The Croquet Project was a very promising project, now fallen into disregard and oblivion because of the sudden drop of its development. It also provides interesting ideas behind user defined paths inside the VR space that will be addressed later.

This work makes a good case behind the reasons of using a 3D annotation system to move beyond the "eye-candy" environment. It also provides a nice definition of annotation, as follows "Annotations may be thought of as author- attributable content placed within a Croquet scene in association with (or reference to) a particular element of that scene.". On the other hand it also provides a rather bizarre visualization of these annotations, "...we have developed basic conventions by which annotations can be created independent of the form of media. We allow objects to become annotation even if the object was not originally designed to be an annotation".

The work by *Sonnet et al.* [12] integrates a 3d probe with annotations, giving meaningful information and an exploded view of the 3D model depending of the position of the 3D probe. It uses an automated algorithm to select the best position to display the annotations given both the position of the probe and the centroid of the objet of interest. It also uses handy techniques to navigate the 3D space and isolate important parts of the 3D model. On the other hand it uses complex ways of annotating objects and rare shapes to work as containers of the annotations.

The work by *Duval et al.* [4] works the possibilities around collaborative explorations of 3D scientific data. The most remarkable parts of it concern the definition of annotations as 3D viewpoints rather than textual information freely located in the space as shown in [9] and [12]. This idea will be used later in the development of this project. It also uses a minimap to locate the users in the scene similar to the ones featured in the game industry.

The work on annotating and sketching on 3D models via a web interface developed by *Jung et al.* [8] draws interesting lines between the needs behind a post-it metaphor driven annotation system and a more free 3D sketching system. It also raises interesting points around the possible advantages of an asynchronous collaborative system and it's support for a more semantic history of the evolution of the model, versus a real-time voice based discussion system.

Other sources consulted did not work as prior work but as excellent references for a more specific part of the project. The work by *Gonçalves et al.* in their software Tag Around [6], concerned with correct tagging and metadata enriching of photos, proposes an interesting approach to usability tests focused in non conventional interaction. The work by *Burton et al.*[3] and by *Bernheim et al.*[2] presents a really insightful conclusion around collaborative annotation systems, which as summarized in [8] should "be unobtrusive but accessible, inform without overwhelming, separate higher and lower priority information for different actors at different times", even when their work concerns the annotation of ext documents the principles are extrapolable.

2.3 The Problem

The root of the problem is the inherent complexity of 3D models. The need to reduce these inherent complexity and expand the models through the use of textual information is the primary concern that gave light to this project. Textual information of a general kind would serve to cast clarity and allow a deeper understanding of the modeled scene. As said in [9] 'To be generally useful for collaborative research and learning, immersive virtual 3D spaces must include intuitive content creation and annotation tools'.

The uses of such technology are rather diverse. From enabling educational experiences not possible before, to reducing the time to market via reducing the time to design in a diverse range of manufacturing companies; passing through semantically augmenting architectonic heritage, the potential uses are innumerable.

These potential uses and its associated potential shareholders are still restricted to the access of a cave-like facility given it's high costs. On the other hand, the personal computer version could be used by anyone with access to a standard computer. However, this scenario provides a rather poor experience and similar solutions are already available in the market.

The context of the problem is nevertheless a rather specific one, the facilities of the Institut Image, which devotes most of its attention to educational and research ends. These facilities count with two immersive systems relevant to this work. The first one is the MoVe, a cave-like system with three faces and the floor projecting passive-stereo images and a IR tracking system with two cameras. The second immersive system is the Spidar, a single screen projected as well with passive-stereo images and 4 IR cameras.

As part of the problem definition the requirements of using the C++ programming language, along with OpenSceneGraph, running in various Windows XP machines, were given.

3

Design and Specifications

3.1 The Problem's Definition

Even thought in previous sections the problem to solve has been discussed considerably, this section focuses in provide a short and precise definition of it.

This project proposes a solution to Annotate 3D models in a immersive cave-like system using an android tablet to interact with the scene. The part of the software that will be running in the cave servers and is aimed to render the scene will be developed using C++ and OpenSceneGraph. The part of the software that runs in the tablet will be developed using Java.

The annotations must be saved using an XML archive, so it can be easily read by any other software. The annotations must show the information in the best displayable way, maximizing readability and relevance. This annotations must be textual comments with priority and author.

The solution also must be able to run in a standard PC running Windows XP.

3.2 Specifications

3.2.1 Functional and Non Functional Requirements

Functional Requirement 1 (FR1)
The user must be able to create annotations associated to the polygons of interest in the model.
Input:
The polygons of interest, the content of the annotation, its priority and its author.
Output:
The annotation is successfully linked to the model.

Functional Requirement 2 (FR2)
The user must be able to retrieve the annotations made in the model.
Input:
None.
Output:
The user can see a list of the annotations associated with de model.

Functional Requirement 3 (FR3)
The user must be able to save the annotations made in the model.
Input:
The annotations made by the user.
Output:
An archive containing the annotations.

Functional Requirement 4 (FR4)
The user must be able to load the annotations made in the model from a file.
Input:
The archive containing the annotations.
Output:
The user is now capable of seeing the annotations previously made.

Non Functional Requirement 1 (NFR1)
The application must work in the cave-like systems available at the facilities of the Institut Image.
Origin:
Context of the development of the project.
Level of Impact:
Sets the bases of the architecture of the system.

Non Functional Requirement 2 (NFR2)
The application work, as well in a standard PC running Windows XP.
Origin:
The need to use the system in other locations with different facilities.
Level of Impact:
Creates another branch of development inside the project.

Non Functional Requirement 3 (NFR3)
The part of the application running in the cave servers must be developed using C++ and OpenSceneGraph.
Origin:
These default technologies used for high demanding applications and the *de-facto* standard inside the laboratories related to the project.
Level of Impact:
Sets the technologies to be used in the development.

Non Functional Requirement 4 (NFR4)
The part of the application running in the cave servers must run in Windows XP..
Origin:
The OS of the cave is Windows XP.
Level of Impact:
Sets the environment of development. Given the use of OpenSceneGraph as the base layer the impact is not as deep as it might have been.

Non Functional Requirement 5 (NFR5)
The application must save the annotations in XML format.
Origin:
Asked by Frédéric, taking in account the advantages of the highly standardized format.
Level of Impact:
Creates another module in the software to save and load annotations from an XML file.

Non Functional Requirement 6 (NFR6)
The application will tour the user through all the annotations made to the model.
Origin:
Enhances usability.
Level of Impact:
Minimal, given that the user was already able to manually tour model searching the annotations.

Non Functional Requirement 7 (NFR7)
The application will make use of the VRPN technology to interchange information between devices.
Origin:
Enhances the modifiability of the system and is a standard practice in both laboratories.
Level of Impact:
Sets architectural constraints.

Non Functional Requirement 8 (NFR8)
The application will make use of an Android powered tablet as the front-end interaction device.
Origin:
Enhances the usability of the system and was proposed by Frédéric.
Level of Impact:
Creates a whole new branch of development but reduces it's complexity given the high-level API featured by the Android OS.

3.2.2 Acceptance Level

The minimal accepted solution would be the one who fulfills all the requirements (functional and nonfunctional) mentioned before, except the Non Functional Requirement 2. This is, a solution that runs and is fully functional in cave like systems but doesn't work in a standard PC running windows XP.

The optimal solution in the available time is then, the one described below through the requirements. The description of better global solutions will be reserved for the *Future Work* part of this document as they involve a significantly major development work, involve an external repository of the annotations and timeline functionalities upon the models and their textual information, that are beyond the reach of this project.

An incomplete solution would be developed in the case that the calendar designed for developing the solution falls short due to unforeseen serious technical difficulties during the implementation. In this disgraceful scenario the part of the project that would suffer the most, is without doubt the related to the reviewing of the annotations. This, because the first part that will be implemented is the one concerned with creating the annotations given it's need to afterwards see the annotations made.

On the other hand, another part of the project that could present non-optimal quality is the one concerned with the user experience. The project is thought to be a clean, minimal but satisfying solution to the problem it aims to solve, nevertheless a huge effort is being invested in improving the user experience and the quality attributes related to it.

Taking in account the ideas presented in the famous book by the creators of the successful company 37signals [5] and the Apple Development Guidelines [1], a nice user experience involves reducing the span of functional requirements to the minimum necessary to successfully solve the problem, while focusing in less evident goods of the software. In the particular case of the project, the rendering power, the net transmission and the responsivity of the front-end (the Android powered tablet) are the most relevant factors in providing a satisfactory user experience. Is yet to be seen how easy or hard is to tune them rightly.

3.3 Restrictions

The most significant restrictions at the time of developing the project is, with complete certainty, both the inexperience of the author with the technologies used and the time to deliver it. There wasn't any other technological restriction which could affect the project. Given the nature of the project there weren't any economic, environmental, social, political, ethical, legal, health or security related restrictions [1].

On the other hand, a significant restriction to broad the contexts, in which the application is a relevant solution to a similar problem, is the amount of money needed to build similar immersive facilities to the ones present at the Institut Image or the COLIVRI laboratory at the Universidad de los Andes.

[1] This remark is made taking into account the ABET directives for such documents.

4

Design Process

4.1 Information Gathering

The information sources used to develop the design and the architectural choices where basically the online databases provided by the Universidad de los Andes library site. Being the most used and consulted the ACM Digital Library and IEEE Computer Society Digital Library online libraries . Given the precise instructions to develop the project there wasn't a particular need to search information about the origin of the problem to solve, it was just a need present at the Institute Image, as discussed in previous sections of this document.

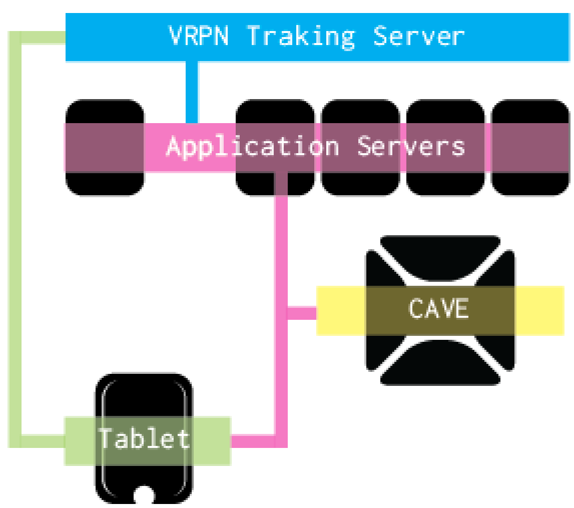
The reviosion of the bibliography and references used in articles describing similar projects [9], [12], [4] and [8] pointed other useful sources of information; even when they weren't so closely related to the project's subject. These references were used basically to evaluate potential design alternatives, taking in account the reduced number of iterations possible to develop the application, such as [2], [7]. Other sources of information that influenced the work were known by the author before, such as [1], [5] and [10].

4.2 Design Alternatives

The goal of this section is to explain the design alternatives and the reasons behind the selection of the alternative that was ultimately implemented. Even in the case that there weren't alternatives at all, like in the system architecture, this section will still explain the pros and cons of the default selection.

4.2.1 System Architecture

The system running the application is a three-layered system, as can bee seen in the image below . The top level, the VRPN Tracking servers receives tracking information from the tablet VRPN device, which design will be discussed later, and from the IR tracking cameras following the flat passive markers attached to the back of the tablet.



The VRPN server centralizes and retransmit the information from all the interaction devices provided to interact with the immersive system in which is used, in this case the tablet and the passive stereo glasses. The use of

VRPN was a must for the project, as shown by the Non Functional Requirements in the previous chapter. VRPN stands for *Virtual Reality Peripheral Network* and is "... a set of classes within a library and a set of servers that implement a device-independent, network-transparent interface between application programs and the set of physical devices (trackers, buttons, etc.) used in a virtual reality (VR) system" as defined by its creators in [11]. The use of VRPN is a development standar at the Institute Image as well as at the COLIVRI laboratory because of the flexibility provided by an abstraction level that separates front-end devices from the application input.

The second layer is composed by the Application Servers in charge of rendering the scene to project it in the Cave. One of these application servers render the scene from the tablet point of view, and stream the scene in real time, to allow the camera metaphor that will be explained in the User Interaction section of this chapter. Is important to note, that there isn't communication between those servers; each one, renders the scene from a particular perspective calculating the adequate frustrum for it's designated screen. Each of the Application Servers renders the images for both eyes of it's screen, but each image is projected by a different projector in order to work as a passive 3D system.

4.2.2 User Interaction

Front-end Device

Application Workflow

User Interface

4.2.3 Software

Cave

Tablet

5

Implementation

5.1 Implementation Review

5.2 Expected Results

6 Validation

6.1 Methods

6.2 Validation of Results

6.3 Discussion

6.4 Future Work

Bibliography

[1] Apple Inc. Apple development gidelines.

[2] Bernheim Brush A.J., Bargeron D., G. A. C. J. Robust annotation positioning in digital documents. In *Proceedings of the SIG-CHI Conference on Human Factors in Computing Systems* (2001), pp. 285–292.

[3] Burton N, Kilgour A C, T. H. A case study in the use of vrml 2.0 for marketing a product. In *Proceedings of From Desk-Top to Web-Top: Virtual Environments on the Internet, World Wide Web and Networks* (1997), pp. 1–24.

[4] Duval, T., Fleury, C., Nouailhas, B., and Aguerreche, L. Collaborative exploration of 3d scientific data. In *Proceedings of the 2008 ACM symposium on Virtual reality software and technology* (New York, NY, USA, 2008), VRST '08, ACM, pp. 303–304.

[5] Fried, J., Hansson, H., and Linderman, M. *Getting Real: the Smarter, Faster, Easier Way to Build a Successful Web Application.* 37signals, 2009.

[6] Gonçalves, D., Jesus, R., Grangeiro, F., Romão, T., and Correia, N. Tag around: a 3d gesture game for image annotation. In *Proceedings of the 2008 International Conference on Advances in Computer Entertainment Technology* (New York, NY, USA, 2008), ACE '08, ACM, pp. 259–262.

[7] Hong, L., Chi, E. H., and Card, S. K. Annotating 3d electronic books. In *CHI '05 extended abstracts on Human factors in computing systems* (New York, NY, USA, 2005), CHI EA '05, ACM, pp. 1463–1466.

[8] Jung, T., Gross, M. D., and Do, E. Y.-L. Annotating and sketching on 3d web models. In *Proceedings of the 7th international conference on Intelligent user interfaces* (New York, NY, USA, 2002), IUI '02, ACM, pp. 95–102.

[9] Kadobayashi, R., Lombardi, J., McCahill, M. P., Stearns, H., Tanaka, K., and Kay, A. Annotation authoring in collaborative 3d virtual environments. In *Proceedings of the 2005 international conference on Augmented tele-existence* (New York, NY, USA, 2005), ICAT '05, ACM, pp. 255–256.

[10] Lidwell, W., Holden, K., and Butler, J. *Universal principles of design: 100 ways to enhance usability, influence perception, increase appeal, make better design decisions, and teach through design*. Rockport, 2003.

[11] Russell M. Taylor II, Thomas C. Hudson, A. S. H. W. J. J. A. T. H. Vrpn: A device-independent, network-transparent vr peripheral system. *VRST* (2001).

[12] Sonnet, H., Carpendale, S., and Strothotte, T. Integrating expanding annotations with a 3d explosion probe. In *Proceedings of the working conference on Advanced visual interfaces* (New York, NY, USA, 2004), AVI '04, ACM, pp. 63–70.