**Mécanique
& Industries**

CLS: A low cost visualization environment for the train industry

Diego Gutiérrez[a], Francisco J. Serón, Juan A. Magallón, Emilio J. Sobreviela
and José A. Gutiérrez

Instituto de Investigación en Ingeniería de Aragón, Grupo de Informática Gráfica Avanzada, Centro Politécnico Superior,
Universidad de Zaragoza, C/ María de Luna, 1, 50018 Zaragoza, Spain

Abstract – We present here a visualization environment for the train industry, where engineers, designers or executives can discuss and analyze visual, aesthetics and ergonomic issues of a train model before it is built. Instead of building a full-size real model, which is not too practical at all regarding design modifications, a virtual model is built in the digital realm, thus taking advantage of the flexibility it offers while keeping costs at approximately one third of the real model. The environment is based on a low cost, PC-based, CAVE-like architecture, (which we have named CLS, or CAVE-Like System) and combines static and dynamic computer generated imagery, both with and without stereoscopy for 3D visualization, as well as Augmented Virtuality techniques for the integration of the train with its environment. The system has already been tested and used by the Spanish companies Renfe and CAF in the design of the new Civia train.

Key words: Simulator / CAVE / industrial design / stereoscopy / virtual prototyping

Résumé – CLS : Un environnement de visualisation bon marché pour l'industrie ferroviaire. Nous présentons ici un environnement de visualisation pour l'industrie du train, où les ingénieurs, les dessinateurs ou les cadres peuvent discuter et analyser les aspects visuels, esthétiques et ergonomiques d'un modèle de train avant qu'il soit construit. Au lieu de fabriquer un modèle normal, en vraie grandeur, ce qui n'est pas trop pratique en ce qui concerne aux modifications de conception, un modèle virtuel est construit dans le domaine numérique, tirant ainsi profit de la flexibilité que ceci offre et coûtant pourtant à peu près un tiers du prix d'un vrai modèle. L'environnement est basé sur une architecture à bas coût, basée sur PC, de type CAVE, (que nous avons appelée CLS « CAVE-Like System », ou système de type CAVE) et qui combine des représentations dynamiques et statiques générées par ordinateur, avec ou sans stéréoscopie pour la visualisation tridimensionnelle, aussi bien que des techniques de Virtualité Augmentée pour l'intégration du train avec son environnement. Le système a été déjà mis à l'épreuve et employé par les compagnies espagnoles Renfe et CAF dans la conception du nouveau train « Civia ».

Mots clés : Simulateur / CAVE / dessin industriel / stéréoscopie / modélisation virtuelle

1 Introduction

Simulation is a way of working that allows the study of a physical system by substituting it for another, more suitable to observation or measure.

Different interests can be addressed by means of simulation. An astronaut can learn to pilot a space shuttle, a government can predict the evolution of the population in its country, a meteorologist can predict storms or a car company can preview its brand new model before it is actually built. These different approaches require different layers of complexity in the simulation: the astronaut trainer, for instance, will most likely need to be able to simulate, up to a certain degree, the dynamics of the system. On the other hand, the car company might not need such realism if only the visual appearance is to be discussed by means of the simulation. But, once the general shape is approved, it might then need to undergo a simulated wind tunnel test, which will require a new layer of complexity in the simulation. Thus, motivation guides the simulation itself.

Several simulations will work better by using Virtual Reality techniques [1–3]. These techniques create immersive environments that make the interaction between man and simulator as indistinguishable from reality as

---

[a] Corresponding author: `diegog@unizar.es`

possible. Most of the times this implies the concept of real time, although as we have seen its necessity will be driven by the goal of the simulation.

The CAVE (Cave Automatic Virtual Environment) is a virtual reality and/or scientific visualization system. Its motivation was rooted in the SIGGRAPH'92 Showcase. The CAVE succeeded in attracting serious collaborators, thus establishing itself as a consolidated virtual reality display [4].

Two of the goals that inspired the CAVE were the search for higher-resolution color imagery, including good surround vision without geometric distortion, and the need to show data and teach others in a reasonable way in artificial worlds.

As originally conceived, the CAVE is a theater of roughly $3 \text{ m} \times 3 \text{ m} \times 3 \text{ m}$, made up of four rear-projection screens (three walls plus the floor). Projectors throw full-color imagery onto the screens. Stereographics' LCD stereo shutter glasses are used to separate the alternate fields that go to each eye. Four Silicon Graphics high-end workstations create the imagery, plus a fifth for serial communications to input devices and synchronization. The users' head and hand are tracked with devices equipped with electromagnetic sensors. The projectors' optics are folded by mirrors, thus reducing the space necessary for the setup. For a more complete description about stereoscopy and other related issues, the reader can refer to [5–7] or [8].

It is important to know that the terms "CAVE" or "cave" cannot be used to refer to generic technology, since the University of Illinois Board of Trustees registered the trademark, and subsequently licensed it to Fakespace Systems exclusively and worldwide. Given this legal frame, we will refer to our system as a CAVE-Like System, or CLS.

2 Definition of the problem

The process of building a train is by no means simple. Usually, the contractor will make a call for tenders and, based on engineering aspects, awards the project to one of the proposals. Once the project is signed, the phase of designing the train begins. This might include a third party, concentrating only on this task. The process involves a lot of work, including conceptual designs, color palettes, sketches... The information goes back and forth between the parties involved, until the contractor is satisfied with the results.

This design process is way more important than what most people might think. Engineering considerations set aside, the visual aspect plays a key role. How the public will accept the train depends of course on functionality issues such as speed or accessibility, but also on considerations relating comfort (are the seats too packed together? Are the corridors wide enough?), ergonomics (is the monitor visible from this specific place? Is the luggage area accessible?) and pure aesthetics (does this color scheme work? Do the shapes seem to belong to one unified design?).

To address most of these not-so-technical issues, the contract includes as a requirement the construction of a full-size model of only one of the cars that make up the train, usually the most representative, due to budget limitations. More often than not, the contractor finds out that what worked as a sketch on paper does not work when built full-size in 3D, but changing it is costly in terms of time and money. Even though this full-size model approach allows one person to actually get into the car, walk through it, sit down to test the seats etc. (which cannot be done with the digital model), it imposes serious restrictions:

– Size and mobility. The full-size model is obviously hard to transport from one place to another, thus limiting the physical space where it can be shown. Only one car is built, and it is a fixed model with no movement capabilities at all.
– Versatility. After examining the full-size model, changes in shapes, lighting or appearance can only be realized by rebuilding whatever parts are involved in the change, then reinstalling them in place. This is not only expensive, but involves heavy time delays as well. Consider the trouble it means to change a critical element such as the seat, for instance. Or a change in the color of the floor, that would involve fixing a new carpet to the whole car.
– Economy. These full-size models tend to be very expensive, which affects the number of changes can be done within a reasonable budget.
– Life. Once used and approved, the full-size model is usually destroyed for storage reasons.

As an alternative to this classical approach, the authors proposed the construction of a CAVE-Like System (CLS). The CLS has been designed to overcome all these shortcomings and change the methodology of the design process, turning it into a cheaper and more flexible one. More precisely, by building a digital model, the restrictions mentioned above are solved, while extra features are added:

– Mobility. The proposed CLS can be dismantled and shipped to be shown in different places or cities, in different frames (fairs, congresses, train stations...) which helps the dissemination of the product. Also, scaled-down versions can be given away in CD or DVD formats for yet deeper and wider dissemination.
– Versatility. All kinds of changes in shapes, lighting or appearance are feasible with the digital model, and much less costly in terms of time and money. Pre-planned changes and different options are only a mouse click away. For instance, the user can compare between five different color schemes, three types of seats and six floors to see which combination works best, add people, change from natural daylight to artificial night lighting... On top of that, the digital model can simulate movement through rendered animations.
– Economy. Once the initial hardware investment has been made, the digital approach is one order of magnitude cheaper than building the physical model.

The first model, including the hardware investment reusable for subsequent models, has roughly the same price as one full-size model.

– Life. The digital information lasts virtually forever.
– Extension. Using Augmented Virtuality techniques [9], the model can be inserted in real environments, or ergonomics can be studied placing people and objects in the desired places, providing at the same time a greater sense of realism to the viewer.
– Quality. The digital model is a perfect medium for presentations. From a commercial and publicity point of view, it also offers far more possibilities and visual effects than a presentation done with a real model.

Considering all the cited pros and cons of the digital, virtual model approach, it makes sense to think it is worth the effort. The decision to build the digital model was taken by Construcciones y Auxiliar de Ferrocarriles (CAF), Renfe Cercanías [10] and the Grupo de Informática Gráfica Avanzada (GIGA; Advanced Computer Graphics Group) of the University of Zaragoza. The proposed methodology changes the old full-size model paradigm, providing the CLS as a working system since the early stages of the design. The parties involved can propose a design, see it in a full-size stereoscopic display, change it or tweak it if necessary, go back to the CLS to test it and so on, until the cycle converges into the final design. Needless to say, all the changes done in the digital realm are a lot cheaper and can be done faster than with the real model. In addition to this, the digital model can be seen in different daylight conditions, at night, inside a tunnel, integrated with different real environments, with real people etc.

Computer generated imagery has therefore been generated, both static and animation, for regular and stereoscopic viewing. The following paragraphs describe the system specification as well as certain decisions taken during the implementation.

3 System specification

A low-cost CLS was chosen as output format [11]. Four to six people can fit in comfortably, although there is only one ideal point of view from which perspective and stereoscopy look perfect, since it is the point of view for which both have been calculated [12].

Given the nature of the project, the notion of real time was turned down in favor of high quality imagery. That means that the images have all been prerendered, instead of using some graphic library such as OpenGL [13, 14] to calculate and display them in real time. A commercial modeling and animation package was used to produce the geometry, shaders and camera paths. The chosen rendering algorithm was ray tracing, favored over more advanced Global Illumination methods such as photon mapping [15] due to the lesser time needed to generate one frame with the former approach. Global Illumination effects were cheated for a more soft, natural look than what off-the-shelf ray tracing usually offers.

3.1 Hardware

Hardware-wise, the system is made up of the following elements:

– flat panel screens for rear projection (two screens of 4×3 m, one of 3×3 m) assembled with wedge frame, t-bars and light baffles. Additional characteristics are gain 1.0 and 180 degrees of viewing angle. The washable, flame retardant screens do not depolarize the light projected on them to allow for stereoscopic images.
– LCD projectors (XGA panel, 1800 ANSI lumen, 1024×768 native panel resolution) to project images and videos onto the screens.
– fixed short focus lenses with throw ratio: 0.9:1. They are used to reduce the distance between the projector and the screen, instead of folding the projectors' optics using heavy and fragile mirrors.
– Polarizing filters placed in front of the projector lenses and polarizing glasses for the audience for stereoscopic views.
– Two mid-sized PC's (P-III @ 800 Mhz, 256 Mb RAM) running W2000. One is equipped with two MPEG-2 decoder cards, each one offering four channels of simultaneous MPEG-2 decoding. The two cards combined assure that the six videos projected during certain moments of the presentation (two for each eye times three screens) run in sync, keeping both the stereoscopic effect and the continuity between them. The other one will send static SVGA imagery, so it must have six SVGA output channels (different combinations of regular SVGA cards assure that).

The screens are supported by a frame made up of four beams assembled together, forming a rigid rectangle; the cantilevering wedge screen framing system allows each image to be rear projected completely to the outer edge of the seam – which minimizes the apparent gap between all adjacent images. The screens material is flexible, thus allowing for easy transportation by just rolling them up carefully. To assemble it, the screens are unrolled over the frame, fastened to it and tightened with a towrope. Once the screens are completely assembled, their frames are fixed to one another in their final position, so that no additional structure is needed for the screens to stand in the upright position.

3.2 Viewer–centered perspective

Our aim is to have the users see exactly the same things, at 1:1 scale, that they would see if they were standing inside the real train. We must make the projection planes in our perspective match the real CLS walls, and the projection focus should be exactly at the viewers' position. Since our CLS renders off-line in order to achieve the required photorealism in the images using low-end hardware, we must choose an "ideal user" fixed viewer position inside it. The obvious choice is to have this ideal user standing at the center of the floor, and looking at a

point located at the center of the front screen, with the same high as the viewer's eyes (which are also fixed to a mean standard value of 1.65 m).

This means our CLS ideal user has to be looking at the front wall, while the side walls are used to fill the field of view, i.e., to bee seen through the corner of the eye; if the user rotates his head to look directly to, say, the left wall, he will not see a real stereo view.

3.3 Video display and slide show

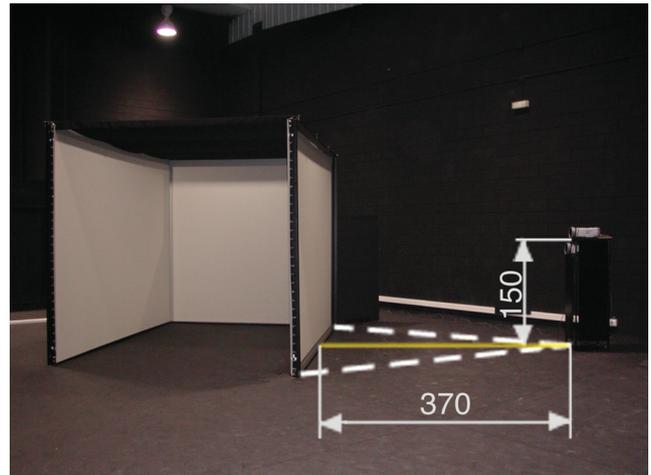We faced two main issues in the process of designing the video display subsystem:

- The ability to deliver six (three stereo pairs) video streams in real time from a single computer.
- The need to play them synchronized at frame level.

The first point above leads us to the need of special hardware add-on cards to play the six streams. These cards also offer the possibility to link the streams on a time or frame basis for synchronization. The control software is based on the SDK provided with the aforementioned cards, that gives some low level functions to program the cards the way it is needed. This software offers access to functions like synchro-start, state querying, etc. We had to program the high-level flow logic of the system and the error control and recovery. GUI programming was done with the GTK+ for Windows library, that offers both an efficient and complete toolkit.

On the other hand, the slide show program is a conceptually simple tool: it just takes six images and displays them in six monitors. The main problems came from the state of the technology in commodity-off-the-shelf hardware in PCs.

One of the conditions of the design was that the show could not be stored in memory, so the slides must be loaded on real time. For this purpose we designed a caching-prefetching system that loads from disk the next frame the user is going to see after displaying the current one. This lets the computer unusable while heavily reading from disk (and perhaps decompressing JPEG or other format), but this time is spent by the users either discussing about what is on screen in that moment or just paying attention to the specific details shown in the picture. We wanted the presentation to include some nice effects, like fading between slides. Hardware acceleration is defective (in the best case) under Windows in multi-monitor conditions (remember you only have one AGP slot, so many of the additional cards were on a 33 MHz PCI bus). Just the process of dumping raw images to the cards works roughly at 5 fps. In addition we needed to generate the intermediate frames for the transitions in real time, and an operation like weight-averaging two frames for a simple fade is very time consuming. This process was added also in the prefetching code.

The amount of interaction with the program is limited. The system is built to be controlled from inside the CAVE, so the remote control has to be radio-based (i.e., non optical). This kind of controls are usually mice or

Fig. 1. Dimensions of the system in centimeters.

small trackballs with 2 or 3 buttons, and are not very comfortable to use (try turning the trackball with the thumb while you press a button...). So no complex operations were designed. Just step forward and backwards, and the possibility to add a popup menu of basic functions (rewind, fast-forward) on the third button. This limited amount of interaction, though, proved to be more than enough for the desired type of presentations.

Before the presentation, the audience is given polarized glasses. During the presentation, they are prompted (by means of a static virtual actor) to put them on or take them off, depending on whether or not the images are stereoscopic.

Figure 1 shows roughly the dimensions of the system described.

Figure 2 shows a few pictures of the actual CLS built. For simplicity, it was decided not to have a floor or a ceiling, at least in this first prototype. The ceiling, though, has been covered with a black canvas, not to let any light out. Similarly, the room the CLS is in has been painted in matt black to absorb any light filtering in or bounced off the screens.

4 Results

The simulation was asked for by Construcciones y Auxiliar de Ferrocarriles (CAF), with Renfe, the biggest train company in Spain, as the final client. More than twenty minutes of computer generated imagery made the cut for the final version of the simulation. As it has already been stated, this imagery included both static images and animation, with and without stereoscopy. In addition to pure synthetic images, Augmented Virtuality techniques were also used to, for instance, have the train come to a stop in a real train station carefully photographed in advance for that purpose. The images were sent to the projectors either in SVGA format or S-Video format. SVGA was used for static imagery, with transitions and fades calculated in real time by one of the PC's.
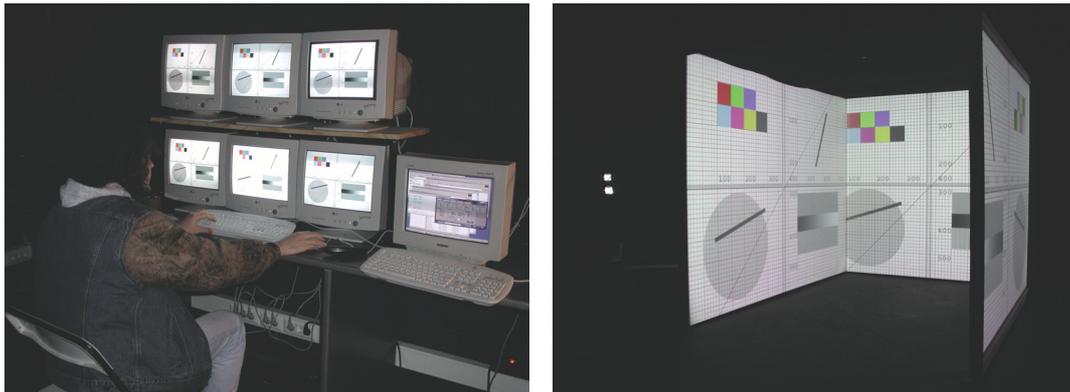
Fig. 2. The CLS system.

Those images were rendered at a resolution of 1024×768. Images were limited to that resolution by the projectors' native panel. Animations were sent in S-Video, at a resolution of 720×576. This time, the limitation was imposed by the graphic cards used to decode six MPEG-2 channels simultaneously. Perfect synchronization of those cards was required in order to keep continuity in the images projected onto the different screens, and to maintain the stereoscopy. This was achieved by writing another program, that would use the software libraries of the own cards.

The simulation was controlled by one script in the case of static imagery. A wireless mouse was used to advance or go back. An automatic version of the script that required no mouse interaction was also written but never used, since one of the goals of the simulation was to be able to stop at any given point and discuss as much as desired over the selected image. The animations were run by the graphic cards' software, since not too much interaction was required.

After initial tests, it was decided to start the presentation with the S-Video animations, including the train coming to a stop in the real station and the surrounding stereoscopic walks inside the train. Then we would turn to SVGA, thus presenting higher-resolution images. Doing it the other way around would have mean a decline in the quality, which initial tests showed to be unpleasant.

A synthetic hostess guides the users throughout the presentation, explaining what parts are going to be seen, and prompting them to put on the stereoscopic glasses or take them off. A small map of the train is shown on one of the side screens when necessary, to help the users locate themselves.

Animations were rendered on a render farm made of six double-processor stations with 1 Gb of RAM, running Windows 2000. For static imagery eight additional machines were used whenever the user would log off.

Table 1 shows some of the most relevant data of this simulation. Figure 3 shows the simulation running on the CLS, as well as several of the images produced.

We would also like to mention the fact that the legal patent process for the system has already been started by the authors.
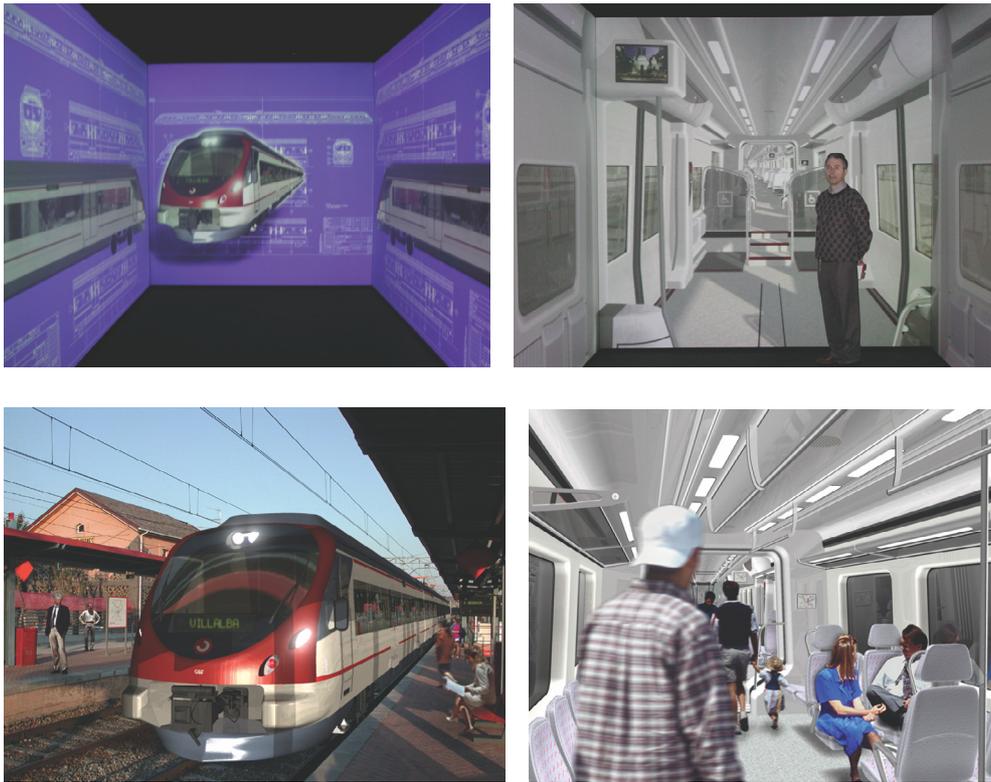
5 Conclusions

A train simulator entirely developed by the Grupo de Informática Gráfica Avanzada of the University of Zaragoza has been presented. The simulator, based on a low-cost CLS architecture, is fully functional, and has been used in several simulations already with CAF and Renfe staff with satisfactory results.

The possibility of not using a real model for the design process of a train has been proved. Lots of variations can be presented sharing a common design in a cheap, fast way. Images with different floors, seats, doors, windows or color schemes can be rendered to be viewed and compared full-size, even side to side if necessary. Elements can be hidden or shown in isolation, full-size or blown up full screen for better study.

By adding people to the images, the real perception of the train can be seen. An empty train interior tends to look dull, since usually the color schemes used are meant that way. As the Renfe management puts it, *"people are the color of the train"*. That fact was proved by showing empty images first, then adding different configurations of people sitting, staring out the window, walking down the corridors. . .

Lots of changes in the design of the train were induced by the images produced. The seats, for instance, which design had been approved based on their orthographic views, were found not to look so good from a natural, passenger's point of view, and so they were sent back to the drafting table. Also, a slightly darker hue was chosen when early render tests showed poor contrast with the train walls and ceiling. Displays on the outside were made bigger for better reading, after the Augmented Virtuality test of the train coming to a stop in a real station showed poor readability. Monitors were raised a little bit higher, after visibility tests were conducted placing virtual people standing up next to them. Several seats were removed after animatics showed they took up too much space. The original color chosen for the floor carpet was changed after illumination tests, waste baskets redesigned, speakers made a bit smaller. . . These are just some of the changes that could be easily done in the digital realm, but would have cost a lot more to do with a real model. Not only

Fig. 3. Simulation running on the CLS and several of the images produced.

Table 1. Data of the simulation.

Rendering time for the animations	576 hours
Animations generated	27.997 frames (18 min 40 s)
Animations selected for the simulation	17.275 frames (11 min 31 s)
Number of static images rendered including low-rez tests	5000 (roughly)
Number of static images selected for the simulation	207
Postproduction and generation of MPEG-2 files	360 hours
Textures	23.5 Mb, 217 files
Materials and shaders	3.140
3D objects	9.379, for a total of 194 models
Polygons	2.029.027

they were faster this way, but the flexibility of the system allowed CAF and Renfe to explore more creative paths than they usually would have, coming up with a far better product from an aesthetic and ergonomics point of view.

The modifications done in the digital model were classified by the authors during the production of this work as "short", "medium" or "long", according to how long it would take to implement them. This recorded implementation time started from the moment the client asked for a modification until he could see the final, complete image with the changes finished. Table 2 reflects an example of each one, along with the time it took to complete with the digital model and the estimated time (based on CAF's experience) needed with the real model.

A short modification usually involved a change in the visual characteristics of an object, such as the color of the floor carpet. A medium modification meant changes also

Table 2. Comparison of modification times.

Type	Example	Digital	Real
short	carpet color	15–30 min	7 hours
medium	TV shape	2 hours	6–7 days
long	seat shape	1.5 days	1 month

in the geometry of not-too-complex objects, such as the TV monitors. Long modifications meant the same kind of changes, but this time done on far more complex objects, such as the passengers' seats.

The digital model helped save money as well. The seat, for instance, was to be replaced only once at most with the real model, because the fabrication process of each model made it a very expensive object to modify once the initial prototype was built. All told, it was estimated that for each modification afforded in a real model, approximately seventy were realized with the digital one.

Running the simulation with people inside the CLS let us come up with two important conclusions as well. First one is the fact that, even though there is only one ideal point of view from which perspective and stereoscopy have been calculated, any mismatches or discontinuities perceived by people not standing exactly on the ideal point are subconsciously assumed, not causing any discomfort while viewing the images. Second conclusion is the tendency of people to look up front to the central screen. Unless the focal point of the image lies intentionally on one of the side screens, people use them just as peripheral vision, enough so they can feel the sense of immersion. This fact helps in the stereoscopic views, since they are calculated supposing the person is looking straight down the central screen.

References

[1] R. Hollands, The virtual reality home brewer's handbook, John Wiley & Sons, Inc., New York, 1996

[2] G. Burdea, Force and touch feedback for virtual reality, John Wiley & Sons, Inc., New York, 1996

[3] R. Kalawsky, The science of virtual reality and virtual environments, Addison-Wesley Publishing Company, New York, 1993

[4] C. Cruz-Neira, D.J. Sandin, T.A. DeFanti, Surround-screen projection-based Virtual Reality: the design and implementation of the CAVE. Computer Graphics Proceedings, Annual Conference Series, SIGGRAPH'93; 135–142

[5] L. Lipton, CrystalEyes handbook. StereoGraphics Coorporation, 1991

[6] S. Pastoor, 3D television: a survey of recent research results on subjective requirements, Signal Processing Image Communication, Vol. 4, 21–32, 1991

[7] A. Woods, T. Docherty, R. Koch, Image distortions in stereoscopic systems, in: SPIE, Vol. 1915. Stereoscopic Display and Applications IV, February 1993

[8] Stereographics developer's handbook. StereoGraphics Corporation, 1997

[9] P. Milgram, H. Takemura, A. Utsumi, F. Kishino, Augmented reality: a class of displays on the Reality-Virtuality continuum. SPIE, Vol. 2351, 282–292, Telemanipulator and Telepresence Technologies, 1994

[10] <`http://www.renfe.es`>

[11] <`http://www.sgi.com/products/appsdirectory.dir/apps/app_number859393.html`>

[12] D. McAllister, Stereo computer graphics and other true 3D technologies, Princeton University Press, 1993

[13] M. Woo, J. Neider, T. Davis, D. Schreiner, OpenGL programming guide, Addison-Wesley Developers Press, 1996

[14] R. Kempf, C. Frazier, OpenGL reference manual, Addison-Wesley, 1997

[15] H. Wann Jensen, Realistic image synthesis using photon mapping, A.K. Peters Ltd., 2001