A multiplatform architecture for 3D graphics

María del Puy Carretero, Amalia Ortiz, David Oyarzun, Isabel Torre, María Teresa Linaza - Vicomtech
Alejandro García-Alonso - University of the Basque Country

Our research focuses on enabling users to interact with 3D graphics and interactive applications with the same content, appearance and interaction paradigm, on different platforms (Internet, Television, and Mobile Devices). The main contribution of this paper is the design and implementation of a platform-independent architecture for 3D graphics and content, validated in the entertainment and tourism sectors. In the case of entertainment, people enjoy communicating with others through virtual characters, also called avatars, e.g. in chat rooms. 3D avatars are being used increasingly and are gaining interest because they are often more visually appealing. As people can choose among different communications platforms, they should also be able to use a virtual character on these platforms to interact with other people at any time and in any place. The interaction of tourists with 3D graphics occurs in two stages of a trip and on two main platforms: the PC when planning their trip on the Internet, and a mobile device while at their destination. Tourist offices invest substantially in virtually recreating some of their most prominent tourism resources, therefore these models should be reusable on several platforms. The general architecture presented in this paper has been validated in several projects that are presented as case studies.

Introduction

Nowadays, people are accustomed to using different interactive platforms in their daily lives. Digital TV is consolidating as the main interaction platform at home, while laptops and desktop computers are still the most important working tools in the office. It must also be mentioned the growing impact of mobile devices in both private and working environments. Nowadays, there are more mobile devices in the world than PCs and TVs combined [1]. Multiplatform applications are becoming more and more frequent in daily life.

Therefore, this paper focuses on enabling people to interact with 3D graphics and interactive applications with the same content, appearance and interaction paradigm in several platforms (Internet, television) and using different devices (PCs, PDAs, mobile phones or televisions).

This paper is organized as follows. The first section describes currently available architectures for different platform applications. Then, the proposed architecture is presented as well as the two sectors (entertainment and tourism) where the architecture will be validated. While in the former 3D contents are mainly virtual characters (or avatars) that represent people, the contents in the latter are tourist attractions such as buildings, sculptures, monuments, etc. Finally, some conclusions of this work are withdrawn.

Related Work

As Herring [2] stated, communication environments incorporating audio, video, and 3D graphics are increasingly available. For instance, in the entertainment sector, people enjoy communicating with others through avatars in virtual worlds such as Second Life (www.secondlife.com) and Imvu (www.imvu.com). These virtual environments allow users represented by a virtual character to interact and communicate with other users. In general, these applications are limited to PC and Web environments. However, the Internet is not the only medium used for entertainment, as other devices such as televisions or mobile phones are commonly used during leisure time for entertainment. For this reason, there are several approaches to multiplatform interaction with avatars and 3D content.

The use of avatars and 3D content on the Internet connected PC has increased in recent years. Several prototype applications allow users to chat, or virtually interact, with other users. For example, BodyChat [3] uses embodied avatars to mimic face-to-face human communication. In [4], Ma *et al.* present a chat application based on emotion estimation from the text that the user types. Moreover, an embodied avatar reproduces the text with emotion estimation. Flat3D [5] is a 3D virtual world that provides creative activities. Furthermore, there are also some 3D virtual worlds on the Web. For example, Second Life (www.secondlife.com) is a 3D virtual world with more than 4.5 million users. Outerworlds (www.outerworlds.com) is a virtual reality 3D chat application where people can talk with their friends, play interactive games, build a virtual house, etc. Worlds.com (www.worlds.net) has a 3D chat where avatars move through virtual rooms and chat with other users.

For television, several applications including avatars and 3D content have been developed. For instance, MARILYN [6] is a prototype for a virtual human avatar for intelligent interaction with digital TV and is focused on business television. An interactive TV show based on avatars is described in [7]. The show consists of a TV quiz in which viewers are represented by avatars, which simulate their behaviour, allowing them to play from their homes. In [8], the authors show a meta-model aimed at specifying an avatar's organisation from four points of view: structural, functional, contextual and normative. The model has been applied to TV game shows in which avatars are based on agents. These avatars are directly controlled by their corresponding viewers. With AmigoTV [9] people can connect with their friends watching the same TV programme. Each user is represented by an avatar and they can change the emotion of the avatar depending on their mood. AmigoTV allows viewers to communicate by voice.

Finally, more and more applications for mobile devices are becoming available. For example, The SenseMS Application [10] allows the user to send an MMS including text, audio, photo and a 2D avatar expressing an emotion. ExMS [11] is another application for sending messages with 2D avatar animations that express emotion. LiveMail [12] presents a new way of communicating using 3D face models created from images taken by phone cameras. In [13], a scalable avatar for conversational user interfaces is available. Its function is to adapt the appearance of the avatar to the device; using, for example, a 2D avatar for small devices such as mobile phones. Y. Mochizui *et al*. [14] have developed a Virtual TV Phone application for conversation with an anthropomorphic agent, which is suitable for consumer communication equipment, especially mobile phones. The receiver of the call can select and display this agent instead of the caller.

With regard to the tourism sector, some recent attempts have been made at exploring 3D graphics for mobile tourist guides. Rakkolainen *et al* [15] have proposed a system that combines a 2D map of an area with a 3D representation of what users are currently seeing in the physical world, studying the effects of 3D graphics on navigation and finding ones way in an urban environment. They concluded that 3D models help users to recognise landmarks and find routes in cities more easily than traditional 2D maps. Unfortunately, the prototype was implemented on a laptop, not on a PDA. 3D city models for route guidance have also been tested by Kulju *et al* [16] who obtained similar results, but highlighted the need for detailed modelling of buildings and additional route information, such as street names. Unfortunately, their prototype used only predefined animations and picture sequences, not interactive 3D worlds. The LAMP3D [17] system provides tourists with a 3D view of the environment they are exploring, synchronised with the physical world through the use of GPS positioning information. The user can easily obtain information on existing objects in the real world, selecting directly them in the VRML world displayed on their device. In addition to supporting navigation via a positioning system, such as GPS, LAMP3D allows the user to navigate through the 3D environment using the PDA stylus, or even watch a virtual tour previously recorded by another user.

In this section it has been shown that the use of virtual characters to represent users and 3D content in different environments is more popular than ever. Nevertheless, at present there is no application that allows the use of the same avatar in different platforms.

The background material considered in the previous section shows that 3D rendering is disperse and in a wide variety of data structures and display engines. 3D-diversity has occupied research and industrial fields, so identifying the core elements and communication channels allows us to evolve in what is a rapidly changing environment. As a consequence our goal has been to minimise modules and communication in the high level architecture that this paper presents.

Architecture

This section describes the global architecture of the system. Our proposal follows an abstraction of the key components of the display system. A client/server architecture was implemented to access 3D graphics and interactive applications with the same content, appearance and interaction paradigm, in any place and at any time, in a multiplatform environment.

The proposed scheme is based on the experience gained during several projects. For instance, the SASTEK project (Figure 1) aimed at helping elderly Alzheimer's sufferers in there every day life. The project proposed the use of an avatar on a PDA, as a tool to explain the way the application works and to remind the user of the activities in their daily routine.

Figure 1 *A PDA based avatar application*

This concept has been extended in the European project I2HOME (Figure 2). In this case, the avatar appears on TV in order to inform the elderly user about their medication, or to warn them when something is wrong at home (for example, if the oven has been left on) in their digital home.

Figure 2 *An avatar appears on TV to inform to user about their schedule or to warn about a situation in the digital home.*

Finally, in the ELEIN project (Figure 3), the 3D avatar plays the role of a virtual teacher in order to help students on the Internet.

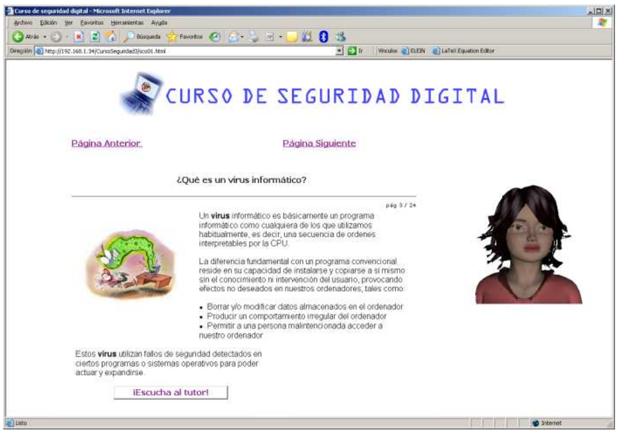

Figure 3 *An avatar on the Internet playing the role of virtual teacher.*

In these projects, the avatars have different appearances due to the fact that the requirements differed. However, using the proposed architecture we can use the same appearance in each environment (mobile device, TV or the Internet) in order to interact with the same avatar.

Multiplatform architecture

As previously mentioned, a client/server architecture (Figure 4) was implemented to access 3D graphics and interactive applications through different platforms.

The server is responsible for providing the 3D interactive content to the clients. Content is described in files using a markup language based on XML and stored in the *3D object DB*. These files make reference to the 3D objects and the textures which form the 3D graphic scene. In addition, materials and the position of objects are included.

The content will be near identical for all platforms. However, the visualisation will depend on the device. For example, a special format (*.m3g) is needed to load 3D objects embedded in J2ME applications on mobile devices, even though PCs can load 3D objects in the *.obj format. Thus, the *3D Object Manager* will select the appropriate format for each platform.

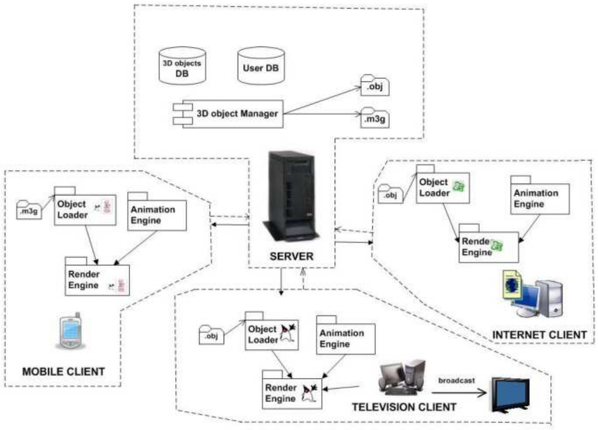

Figure 4 *System Architecture*.

On the other hand, each platform has specific hardware and display features. Therefore, a common architecture based on three modules is used for each of them: *the Object Loader*, *the Animation Engine* and *the Render Engine.*

Firstly, the *Object Loader* has specific functions for each content and hardware combination. These functions import the 3D interactive object from the server (3D object Manager) using the most appropriated libraries or functions for each specific rendering engine. Models are stored and managed by the server.

Likewise, the *Animation Engine* is responsible for generating the 3D animation, when necessary. Current technologies make it possible to run the animation engine on mobile phones. However, the animation can be slow due to current limitations of mobile phones. This performance is expected to improve in the immediate future, as some current mobile phones already integrate hardware accelerated graphics chips (e.g. NVidia Tegra), which allow a better integration of the animation engine and a smoother running application.

Finally, the *Render Engine* is built using the most appropriate libraries for the user interface on each platform. At times, Internet clients are reluctant to download and install additional plug-ins. Thus, the Internet scenario uses the Anfy3D library, which is integrated into the application. This client is oriented towards desktop browsing. In the mobile environment, clients use the Mobile 3D Graphics API (M3G), which is the 3D graphics library used by J2ME. It must be mentioned that J2ME is considered as a *de facto* standard for mobile applications. Therefore, it has been selected as the base environment for such applications. Finally, there are no such usability restrictions in the television environment, such as those previously mentioned; a more powerful PC can be used, mitigating the need to install additional software libraries. Therefore, Java3D has been selected for the render engine, as it is more powerful than Anfy3D.

Validation

The proposed architecture was tested in two different sectors: entertainment and tourism. Both sectors try to enable people to interact with the same interactive applications based on 3D graphics on any platform.

It is needless to say that the architecture should be separated from the quality of the models, as it provides the same 3D content to different platforms. However, not all of the 3D content can be loaded on all platforms, as there is 3D content which cannot be loaded even on a typical PC. Therefore, the capabilities of each platform should be taken into account, as it is obvious that current mobile technologies are less powerful than PC technologies. The larger the number of polygons a 3D model has, the more difficult it will be to load and animate it on a mobile device. Thus, a temporary solution has been designed in order to validate the proposed architecture.

It is clear that the models should be adapted to the power of the platform and therefore the PC model should have the highest level of detail. It is also clear that smaller details are unnoticeable on mobile devices with smaller screens. Moreover, as previously explained, the format of the model has to be changed; minor model details can be discarded in order to adapt the models to the mobile devices' capabilities. For example, the body of the avatar could be omitted, because the facial details are more important. Thus, the *3D Object Manager (server)* will decide which models are sent to each client.

Entertainment Sector

Regarding the entertainment sector, a multi-media virtual character which can be used in the more popular entertainment environments has been implemented. The avatar is a virtual representation of the human user who wants to be present in remote locations using their desired medium. The prototype was validated on three platforms: Internet (desktop browser), mobile phones (SMS) and Television.

A markup language based on XML was used to define the appearance of the avatar. This markup language called ACML (Avatar Configuration Markup Language) describes the 3D objects that form the avatar (head, hair, eyes, body) and the materials and textures required. These ACML files are stored in the 3D object database (server).

In the Internet scenario, people can configure the physical appearance of their avatars. This appearance is saved in (or loaded from) an ACML file which refers to the 3D objects in *.obj format. Moreover, it can be used to chat in an avatar chat application. Users must be connected to contact other users. If a user is not on line, their avatar (photo) will appear in an email. As in other chats, the user can write a text, although in this case, the reader can also see the avatar reproducing this text, which has been synthesised. In such a way, users have the feeling of talking to a "real" person. In addition, users can select realistic emotions for the avatar when reproducing the text such as happy, sad, angry, etc.

As in the Internet scenario, the user can edit the physical appearance of their virtual character using the mobile device or use this avatar on the Internet. In this case, *.m3g files will be used. Figure 5 shows an avatar created on the Internet which has been loaded on a mobile phone

Figure 5 *An avatar created on the Internet and loaded on a mobile phone.*

Users can send messages with the avatar attached to other users. Two ways of sending personalised messages have been developed (Figure 6). The first method sends a message that appears together with an animation of the sender's avatar, including the emotion selected by the sender. In this case, the body of the avatar is omitted in order to speed up the animation. The second method sends an MMS with the text and an image of the user's avatar. While the former is oriented towards mobile devices that are capable of rendering 3D graphics, the latter is applied to those which cannot.

Figure 6 *Message with an animated 3D avatar vs. MMS with the avatar's caption.*

Finally, in the TV scenario, avatars represent users who send an SMS to a television programme. Currently, some TV programmes display text messages sent by viewers. Using the proposed application, an avatar can reproduce the message using two different methods: text displayed on the screen with a static avatar ("photo"), or an animated avatar reading the text (*.obj files). The first method is more appropriate as it prevents the avatar from interrupting the TV programme's host and guests. In such a way, the avatar's image and text can appear on the screen combined with the programme, so that viewers at home can read the messages and see a picture of the sender's avatar. Some messages may be read by the avatar when appropriate. TV hosts may break in the programme in order to display the avatar. One important issue with SMSes is that senders often shorten words. Current text-to-speech synthesisers cannot understand abbreviated messages. Therefore, as a preliminary solution, a tool that TV employees can use to edit SMS message content has been implemented.

Tourism sector

Regarding the tourism sector, the main goal of the architecture is to allow tourists to access tourist content in different stages of their journey: when planning their visits and while at the destination. Therefore, two different platforms have been used for validation: the Internet and mobile devices.

In the Internet scenario, users can search information to plan their holidays or their visit to unknown places. They can make 3D simulated guided visits; see a map of the tourist attractions, read interesting information, etc. With all of this information, users will have an idea of what they should expect to experience when they visit.

It is obvious that when people are at their destination, they cannot access the tourist information through their desktop personal computers. However, they can download a tourist guide [18] to their mobile devices and enjoy different kinds of information. This guide is created by content providers that take advantage of our platform in order to include 3D objects in the guides.

Using our architecture, users can see what amounts to the same content and information on any device.

Figure 7 *Example of a tourist attraction on a mobile phone*

An XML file was used to define the tourist scene that is to be loaded. Although this markup language is still under development, an XML file which references the 3D objects, materials, textures and positions has been used.

On the Internet, the XML file which is imported references *.obj files and can also reference several 3D objects which describe the tourist scene in detail. On the mobile phone, the 3D objects (*.m3g) that are of lesser importance can be omitted. For example, an XML which refers to tourist attractions, buildings, parks and other street furniture which are close to it can be loaded in full on the Internet. However on the mobile phone, it is only the tourist attraction data is loaded as shown in Figure 7.

Numerical data

The avatar shown in Figure 5, the validation prototype, is composed of a set of objects. Table 1 shows the main characteristics of its geometric model. In addition, two textures complement the appearance of the avatar: 512x512 and 1524x512 pixels.

Object	Vertices	Edges	Faces
Head	514	1454	939
Hair	563	1150	589
other	628	1396	672
Total	1705	4300	2200

Table 1 *Number of vertices, edges and faces of the avatar used in the validation prototype.*

As we have explained, the models used both in the Internet based application and in the phone based application, are identical. The only difference is the format that stores the model, .obj for the Internet and M3G for the mobile phone.

The mobile phone used for the tests was a Sony Ericsson K750. Obviously, the Internet based application is able to animate much more complex models, but in order to keep the same appearance in both devices we have limited the model complexity to the size which the test phone animates correctly. Every few months the performance specifications of mobile phones improve. However, the proposed architecture will not need changes to support richer avatar models. Even if the model load module or the image generation module needs to be updated, the architecture can easily adapt to these new requirements.

Conclusions

We have defined a 3D interactive content management architecture for multiple devices. This was validated in the entertainment sector and with the most recent validation being in the tourism sector where satisfactory results were achieved. Using this architecture, the application has the following features that do not appear in related work.

The architecture allows applications from different devices to handle the identical (or near identical) 3D interactive content.

Although the clients load different 3D graphic models, for example *.obj in the Internet client and *.m3g in the Mobile client, the geometry of the model is the same, therefore the loaded avatars and other 3D models have similar appearances on different devices.

In the entertainment sector, users can configure the physical appearance of their 3D virtual characters, which will represent them in different media such as SMS, the Internet, or the television, and on different devices such as computers, mobile phones and the television.

In the tourism sector users can access tourist information, such as 3D objects that represent buildings or monuments, with the device that they prefer.

This architecture can easily be used by other sectors where 3D interactive content is needed.

References

[1] http://www.mobilemarketer.com/cms/news/software-technology/1959.html Revised November 2008

[2] S. C. Herring, "Slouching toward the ordinary: Current trends in computermediated communication". New Media & Society, 6(1), 26-36. 2004

[3] H. Vilhjálmsson, and J. Cassell, "Bodychat: Autonomous Communicative Behaviors in Avatars," International Conference on Autonomous Agents. Proceedings of the second international conference on Autonomous agents, Minneapolis, Minnesota, United States, pp.: 269-276, 1998.

[4] C. Ma, A. Osherenko, H. Prendinger, and M. Ishiizuka, "A Chat system based on emotion estimation from text and embodied conversational messengers," Active Media Technology, 2005. (ATM 2005) Proceedings of the 2005 International Conference, pp. 546-548. May 19-21, 2005

[5] H. Tobita, J. and Rekimoto, "Flat3D: a shared virtual 3D world grown by creative activities and communication through the network," Proceedings Computer Graphics International, pp. 476-479, 2004.

[6] S. Maad, ."MARILYN: A Novel Platform For Intelligent Interactive TV (IITV)," Proceedings of Human Computer Interaction International HCII2003 Conference, Crete, Greece, July 2003

[7] M. Khadraoui, B Hirsbruner, F. MeinKohn, D.J. Khadraoui, and M Courant,. "Interactive TV Show based on Avatars," 2005 Systems Communications (ICW'05, ICHSN'05, ICMCS'05, SENET'05) pp. 192-197, 2005.

[8] B. Gâteau, and D. Khadraoui, "Arbitration of Autonomous Multimedia Objects with a Multi-Agent System,". Information and Communication Technologies ICTTA '06, volume 2, pp 3007-3012, 2006

[9] T. Coppens, L., Trappeniers, and M. Godon, "AmigoTV: towards a social TV experience," Proceedings from the Second European Conference on Interactive Television "Enhancing the experience", University of Brighton, 2004.

[10] A.K. Amin, B.T.A, Kersten, O.A. Kulyk, P.H. Pelgrim, C.M Wang, and P. Markopoulos, "SenseMS: A User-centered Approach to Enrich the Messaging Experience for Teens by Non-verbal Means," Proceedings of the 7th international conference on Human computer interaction with mobile devices & services, pp. 161-166, Salzburg, Austria, 2005

[11] P. Persson, "ExMS: A field study of expressive mobile messaging with avatar-based animations," Proceeding of CHI pp 161-166, 2003 (CHI '03)

12] M. Mosmondor, T. Kosutic, and L. Pandzic, "LiveMail: Personalized avatars for mobile entertainment," Proceedings of. The Third International Conference on Mobile Systems, Applications, and Services, June 2005.

[13] U. Berner, and T. Rieger, "A Scalable Avatar for Conversational User Interfaces," User Interfaces for all, Paris 2002.

[14] Y. Mochizuki, K. Senda, K., Orimoto, J., Tagawa, and S. Asahara, "Humanoid Agent for Communication and Entertainment Driven by Emotion and "KANSEI"," in Embedded System. Consumer Communications and Networking Conference, pp. 551-556, 2004

[15] J. Rakkolainen, T. Vainio, "A 3D City Info for mobile users," Computer and Graphics 25(4), pp. 619-625, 2001.

[16] M. Kulju, E. Kaasinen, "Route Guidance Using a 3D City Model on a Mobile Device," Workshop on Mobile Tourism Support, Mobile HCI2002, Pisa, Italy, 2002.

[17] C. Kray, J. Baus, "A survey of mobile guides," Workshop HCI on mobile guides at Mobile HCI, 2003.

[18] I. Torre, M. T. Linaza, O. Otaegi, A. García. "Self-creating 3D mobile tourist guides" Proceedings of TTRA Europe Conference 2008 - Competition in Tourism: Business and Destination perspectives 2008

Acknowledgements

One author was partially supported by the Spanish Ministry of Education and Science (MEC) grant TIN2006-14968-C02-01.

María del Puy Carretero: Studied Computer Science at the University of the Basque Country UPV (2004). She was an intern during the summer of 2002 and carried out her final year project at Vicomtech (July 2003-February 2004). Since October 2004 she has worked as a Scientific Collaborator at Vicomtech, within the area of Interaction for Education, Leisure and e-Inclusion. She is currently pursuing her doctoral studies and has research interests in Interaction with Virtual Characters on mobile devices and the Internet.

Amalia Ortiz: Is the head of the Interaction for Edutainment and eInclusion department at Vicomtech. Her research interests are related to animation, virtual characters and human computer interaction. She has a PhD from the Faculty of Informatics at the University of the Basque Country, in the research area of emotional interaction trough avatars.

David Oyarzun: Studied Computer Science in the University of the Basque Country (2002). From July 2001 to October 2001 he was an intern at LANIK S.A. modifying and improving the application which is the basis of the production process. From October 2002 to April 2003 he taught Artificial Intelligence. From December 2002 to March 2003 he worked for SINTEC developing biometric access control systems for networks. Currently he works in Vicomtech as a researcher in the Interaction for Education, Leisure and e-Inclusion area. He is a PhD research student in the area of avatar animation and interaction in virtual environments at the University of the Basque Country.

Isabel Torre: From 1999 to 2003 she studied Ingeniería Técnica Industrial especialidad en Electrónica (Technical Industrial Engineering specialising in Electronics) and between 2003 and 2006 she studied Ingeniería de Telecomunicación (Telecommunications Engineering) at the University of Deusto. From October 2006 to August 2007 she *held a scholarship at Vicomtech in the Tourism, Culture Heritage and Creativity department. Since then, she has been working in projects related to wireless technologies and mobile devices programming.*

Mª Teresa Linaza: received a six year degree in Electrical Engineering (1996) and a Ph.D. in Industrial Engineering (2001) from the University of Navarre, Spain. From September 1996 to December 2001, she was member of the R&D team in the Electronic and Communications Department of the Centro de Estudios e Investigaciones Técnicas de Gipuzkoa (CEIT) in San Sebastian, Spain. She has been a member of the Vicomtech R&D team since 2002, assuming the position of Head of the Tourism, Heritage and Creativity department in 2006. She also participate in teaching and training activities in the Faculty of Engineering of the University of Navarre from September 1996 to June 1998 and she has been an Associate Professor in the University College of Technical Industrial Engineering in Donostia (University of the Basque Country). She has also given lectures concerning Arts and New Technologies and is the author of several publications.

Alejandro García-Alonso: He is an Associate Professor of Computer Graphics at the University of Basque Country. His research interests include computer graphics, geometrical computation and haptic systems. He has a PhD in Mechanical Engineering from the University of Navarre.