



## II JORNADAS DE INNOVACIÓN DOCENTE, TECNOLOGÍAS DE LA INFORMACIÓN Y DE LA COMUNICACIÓN E INVESTIGACIÓN EDUCATIVA EN LA UNIVERSIDAD DE ZARAGOZA 2008

### **PERSONAJES VIRTUALES:**

#### **Una nueva herramienta para la enseñanza de Informática Gráfica**

Sandra Baldassarri

*Profesora Colaboradora, Departamento de Informática e Ingeniería de Sistemas  
Centro Politécnico Superior*

Eva Cerezo

*Profesora Contratada Doctora, Departamento de Informática e Ingeniería de Sistemas  
Escuela Universitaria Ingeniería Técnica Industrial*

Francisco J. Seron

*Profesor Catedrático de Universidad, Departamento de Informática e Ingeniería de Sistemas  
Centro Politécnico Superior*

#### **Síntesis:**

Este artículo presenta el uso de personajes virtuales para mejorar el aprendizaje de los estudiantes de las asignaturas de Informática Gráfica, en la carrera de Ingeniería Informática. Para este propósito se ha desarrollado una potente plataforma llamada Maxine para la creación y manipulación de entornos 3D. En ella, los personajes virtuales interactúan a través de texto, voz, expresiones faciales, movimientos corporales, etc. animando a los estudiantes a ser conscientes de sus progresos, transmiten entusiasmo, y hacen el aprendizaje más divertido... En este trabajo se presentan los resultados de la utilización de esta herramienta para la exposición de ciertos temas que suelen ser particularmente difíciles de aprender por parte de los alumnos.

#### **Palabras clave**

Educación, presentadores virtuales, profesores virtuales, motor gráfico.

### **1. INTRODUCCIÓN**

Actualmente mayoría de las investigaciones acerca de interfaces sociales están enfocadas al diseño de los llamados agentes corpóreos conversacionales (ECAs, Embodied Conversational Agents) [1]. Estos agentes se hacen visibles en el interfaz como actores animados que representan seres humanos. Algunas veces, consisten simplemente en caras animadas que hablan, mostrando en ocasiones expresiones faciales y, cuando se usa síntesis de voz, aportando sincronización labial. Otras veces tienen una representación gráfica en 3D, con movimientos corporales complejos, etc...

Existen diversos ejemplos de presentadores virtuales: el hombre del tiempo virtual de Noma y Badler [2, 3], el periodista de TV de Thalmann y Kalra [4], el presentador

del proyecto “Personalized Plan-based Presenter (PPP)” de André et al. [5], el avatar 3D de Baus et al. [6] o, más recientemente el trabajo sobre presentadores virtuales para reuniones virtuales de Nijlot et al. [7], directamente relacionado con el sistema presentado en este trabajo. Todos estos agentes constituyen una prometedora opción de interfaz desarrollada siguiendo un estilo de comunicación e interacción con el cual los seres humanos se encuentran familiarizados [8]. Actualmente esta tecnología y estos modelos se están utilizando en gran variedad de aplicaciones, como por ejemplo como guías en vistas a museos o entornos virtuales [9, 10] o en aplicaciones domóticas [11].

En este trabajo nos centramos en el uso de los agentes corpóreos virtuales para la realización de presentaciones virtuales, y su utilización en la enseñanza de las asignaturas de Informática Gráfica.

La estructura del artículo es la siguiente: en la Sección 2 se presenta la plataforma de animación desarrollada, con sus características básicas y arquitectura. En la Sección 3 se explica en detalle el sistema de presentaciones desarrollado (MaxinePPT). La Sección 4 muestra los resultados obtenidos al utilizar esta herramienta, y, finalmente, en la Sección 5 se presentan las conclusiones.

## **2. NUESTRA PLATAFORMA DE ANIMACIÓN**

Maxine [12] es un motor gráfico para la creación y gestión de entornos virtuales tridimensionales controlado por medio de *scripts*. En Maxine es posible cargar modelos, animaciones, texturas, sonidos, etc. en tiempo real, a medida que se necesitan en la presentación virtual. Aunque es una plataforma muy genérica, está especialmente orientada a trabajar con actores sintéticos en escenarios virtuales. Una de las principales ventajas de Maxine es que permite la especificación de contenidos a través de un editor, pero también es posible la comunicación directa con herramientas de bajo nivel. La plataforma está íntegramente programada en C++ y construida sobre un amplio conjunto de librerías de código abierto. A continuación se describen las características generales y funcionalidades de la plataforma, y, posteriormente, la arquitectura en la que se ha basado su desarrollo.

### **2.1 Características generales**

Maxine permite la creación, organización, manipulación y animación de diferentes elementos que constituyen un escenario virtual, a través de un “grafo de escena”. En un grafo de escena se pueden representar los siguientes elementos:

- Imágenes, que se muestran y posicionan como rectángulos orientados en el espacio. Es posible cargar los principales formatos gráficos (bmp, gif, jpeg, pic, png, rgb, tga, tiff) y se puede representar canal alpha en aquellos que lo soporten.
- Textos. El texto se trabaja de forma similar a las imágenes, pero, además, es posible utilizar diferentes fuentes, especificar el tamaño y color del texto.
- Primitivas geométricas simples. Existe un conjunto de primitivas geométricas básicas como esferas, cubos, prismas, o, incluso, los ejes de coordenadas.
- Modelos geométricos complejos. Permite importar una serie de formatos de ficheros de geometría de uso extendido como 3ds, flt, lwo, md2, obj, osg, DirectX, etc. De este modo es posible crear escenarios virtuales complejos con gran nivel de detalle.

- Luces simples. Básicamente luces omnidireccionales y focos.
- Sonido tridimensional y ambiental. Las fuentes de sonido se caracterizan por una posición en el espacio virtual y con atenuación debida a la distancia, trabajando así sonido tridimensional. También permite sonido ambiental, es decir, independiente del punto de vista del observador.
- Personajes animados. Se utiliza el formato de la librería de animación Cal3D [13], pero los personajes se pueden generar mediante herramientas de modelado y animación 3D comerciales. El sistema proporciona diferentes tipos de animaciones, incluyendo animaciones secundarias para incrementar el realismo y la expresividad de los actores.
- Actores animados. Son una especialización de los anteriores y están provistos de síntesis de voz y animación facial con sincronización labial, siguiendo el estándar VHML y la especificación MPEG-4 para la animación facial.
- Voces sintéticas. Las voces se pueden asociar a un elemento. Se pueden configurar y controlar ciertas características de audio como la selección de una voz concreta, el control de volumen, la velocidad del habla, inserción de pausas, el tono, la enfatización de las palabras, especificación de la pronunciación, etc

Maxine también permite manipular elementos auxiliares como:

- Cámaras. Fijan el punto de vista del usuario.
- Animadores simples. Permiten añadir a cualquier elemento la capacidad de desplazarse por la escena y realizar movimientos rotatorios, con determinados parámetros, a largo del tiempo.
- Animadores complejos de conjuntos de elementos. Este tipo de animadores pueden llevar a cabo movimientos más elaborados, como el seguimiento de trayectorias definidas por puntos o la reorganización y animación simultánea de conjuntos de objetos.
- Grupos. Son nodos que permiten agrupar un conjunto de elementos, llamados hijos, y tratarlos como uno sólo.

## 2.2 Arquitectura de Maxine

En la Figura 1 se presenta la arquitectura general de Maxine, con sus bloques más importantes y los elementos utilizados para la interacción con el usuario. El usuario debe proveer al sistema con los ficheros que contienen la definición del entorno virtual (modelos geométricos, imágenes, texturas, sonidos, etc). La carga y la manipulación de las escenas se realizan a través de *scripts*, por medio de ficheros o a través de comandos por consola.

En la Figura 2, en cambio, se detallan las librerías utilizadas en cada uno de los bloques.

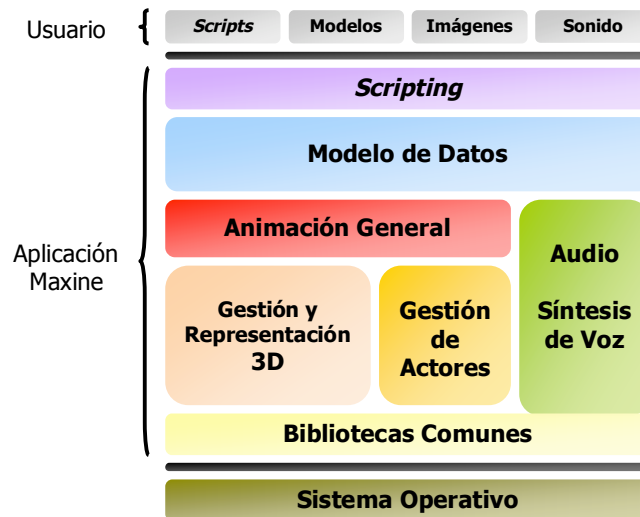


Fig 1. Arquitectura general de Maxine

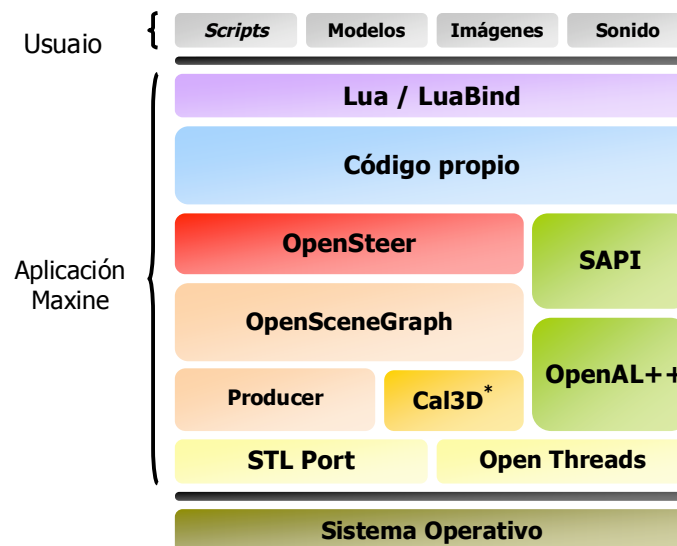


Fig 2. Detalle de las librerías utilizadas.

- OpenSceneGraph [14]: constituye el núcleo de Maxine y se encarga de manipular los grafos de escenas y las representaciones gráficas del entorno virtual. La visualización se realiza utilizando la librería OpenGL.
- Cal3D [13]: esta librería permite realizar la animación local de los personajes virtuales. Se basa en animación esquelética, que produce las deformaciones de la malla. Ha sido necesario realizar modificaciones en esta librería para poder adaptarla a las necesidades de Maxine.
- OpenAL++ [15]: permite trabajar con el sonido ambiente y 3D.
- SAPI [16]: realiza la síntesis de voz y genera la información necesaria para realizar la sincronización labial. Es de libre distribución, pero pertenece a Microsoft.
- OpenSteer [17]: permite enriquecer la navegación de los elementos de la escena añadiendo comportamiento de grupos de elementos.

- Lua y LuaBind [18]: son una parte esencial del sistema ya que hacen posible la creación y manipulación de objetos a través del lenguaje de *scripting* Lua.
- Producer [14]: esta librería gestiona la visualización 3D y el manejo de cámaras.
- STL port [19]: librería estándar C++ multiplataforma
- OpenThreads [20]: librería estándar con herramientas de gestión de *threads* y programación básica.

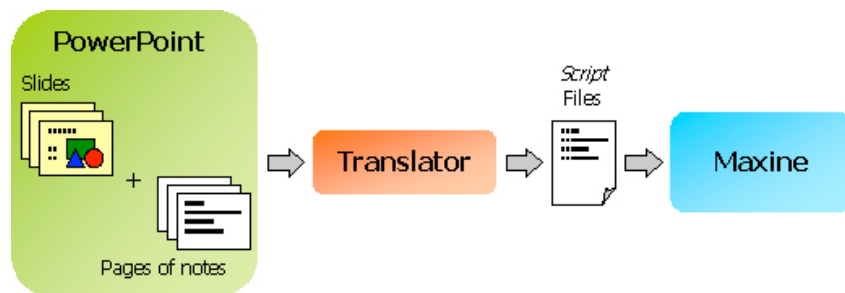
Se ha realizado un esfuerzo considerable en la integración y comunicación de todas estas librerías. El uso de estas librerías ha facilitado enormemente el desarrollo de la aplicación, pero lamentablemente, en algunos casos la documentación era prácticamente inexistente, y se tuvieron que realizar modificaciones importantes sobre el código para adaptar las funcionalidades.

### 3 EL SISTEMA DE PRESENTACIONES VIRTUALES

El sistema general de animación descrito previamente se ha utilizado en una aplicación específica que permite realizar presentaciones de información al estilo “PowerPoint” por medio de personajes virtuales. Dicha aplicación se puede ejecutar en tiempo real y la interacción entre el presentador virtual y el usuario se realiza por medio de texto, reconocimiento de la voz y síntesis de voz, imágenes a través de webcam, o incluso, detectando la posición del usuario por medio de un sistema de posicionamiento.

#### 3.1 Descripción general

MaxinePPT es un conjunto de aplicaciones que permite utilizar la información contenida en un fichero PowerPoint para realizar una presentación utilizando un actor virtual 3D (ver Figura 3).



**Fig. 3.** Visión global de la aplicación MaxinePPT.

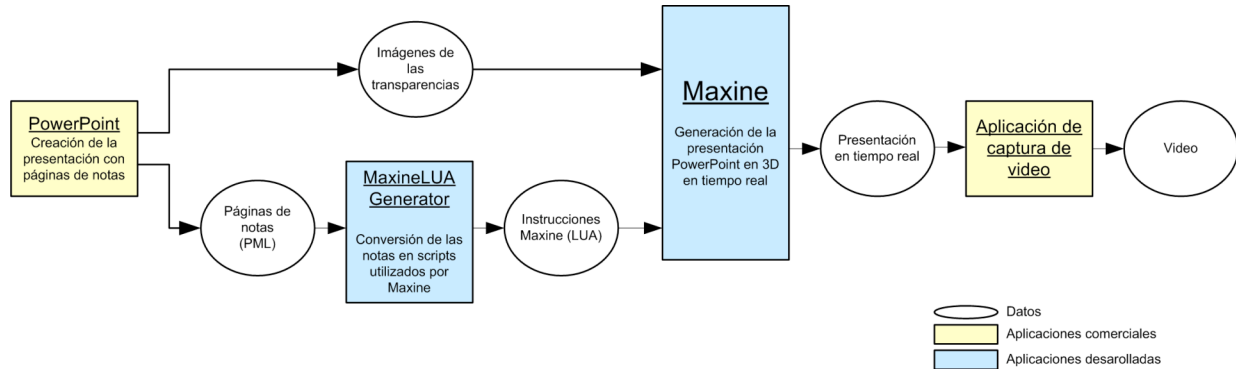
Las características más importantes de la aplicación son las siguientes:

- Permite crear y realizar presentaciones virtuales en un escenario virtual 3D, desde un fichero PowerPoint clásico al que se agrega información adicional.
- Se utiliza un lenguaje tipo XML, denominado PML, que controla los aspectos de la presentación virtual; sus instrucciones se agregan en las páginas de notas de cada transparencia de PowerPoint y se utilizan, por ejemplo, para determinar el texto que va a “leer” el presentador virtual.
- Una vez que se crea la presentación, ésta se presenta automáticamente (no se necesita ninguna intervención del usuario/creador) y de forma verbal, tanto en español como en inglés, utilizando un presentador virtual creado por el usuario, o el avatar por defecto incluido en el sistema.

- El presentador virtual no lee la transparencia, sino que dice lo que se le ha indicado previamente en la página de notas.

### 3.2 Generación de una presentación 3D

La Figura 4 muestra el proceso general para crear una presentación virtual:



**Fig. 4.** Proceso general para generar una presentación 3D

A continuación se presenta una explicación más detallada de los pasos que se llevan a cabo.

**Generación de la presentación PowerPoint:** Básicamente, las presentaciones de MaxinePPT consisten en dos partes: las imágenes que se extraen de las transparencias PowerPoint y sus páginas de notas, que se utilizan para controlar la presentación. La página de notas contiene marcas PML (*Presentation Markup Language*) que pueden ser de dos tipos:

- Marcas de inicialización y configuración: permiten seleccionar un presentador virtual, configurar la voz, seleccionar el escenario, configurar la iluminación y especificar la disposición global y el movimiento de las transparencias.
- Marcas de control de la presentación: permiten controlar la visualización, el flujo, la voz y el sonido correspondiente a cada transparencia, y en particular, la elección de la cámara, las pausas en la presentación, los movimientos del presentador, el orden de las transparencias, o la reproducción del sonido.

En la Figura 5 se puede observar una transparencia inicial con su correspondiente página de notas.

**Conversión de PML a LUA:** Se realiza a través de “MaxineLUAGenerator”. Las páginas de notas escritas en PML se traducen en el lenguaje de *script* LUA para que Maxine las pueda interpretar.

```

<actor name="maxine">
<scene name="future">
<voice name="carmen" speed=40>
<lighths name="default">
<cameras name="default">

<backgroundSound name="back.wav" volume=120>
<pause value=2000>

<camera value=0>
<background_sound volume=50>
<text>
    Bienvenidos a la presentación de Maxine.
    <pause value=500> Hoy vamos a hablar de
mi.
</text>

```

**Fig. 5.** Transparencia inicial de una presentación y su página de notas

**Generación de presentaciones en tiempo real:** El motor de animación Maxine interpreta los resultados obtenidos en los pasos anteriores (las imágenes de las transparencias y las instrucciones LUA), y se realiza una presentación virtual en tiempo real.

Además de proyectar las presentaciones en una pantalla tradicional de PC, se pueden proyectar en un sistema de visualización avanzado, denominado CLS (Cave-Like System) [21] desarrollado por nuestro grupo y que constituye una alternativa de bajo coste al sistema inmersivo CAVE [22].

**Captura de vídeo:** Como la generación de presentaciones en tiempo real de alta calidad requiere la utilización de equipos potentes, es posible capturar la presentación en vídeo para luego verla en cualquier equipo sin necesidad de instalar la plataforma Maxine.

#### 4. RESULTADOS

El uso de MaxinePPT es muy simple. El usuario define una presentación de forma análoga a la realización de una película, a través de un *script* que especifica en cada momento lo que tiene que suceder. Los usuarios más avanzados pueden generar sus propios presentadores y sus propios escenarios.

La presentación generada por medio de esta aplicación se puede usar para realizar la exposición de un tema, mostrando los contenidos de las transparencias y utilizando el

personaje virtual para comunicarse con la audiencia a través de sus movimientos, su voz, sus expresiones... Sin embargo, también la audiencia puede interactuar con el avatar.

Las Figuras 6, 7 y 8 muestran a una actriz virtual “explicando”, en el primer caso, los contenidos de una transparencia inicial típica; en el segundo caso, mostrando una serie de transparencias; y, en último lugar, esperando preguntas por parte del auditorio. Todas las presentaciones se realizan en un escenario gestionado por el motor de animación Maxine.

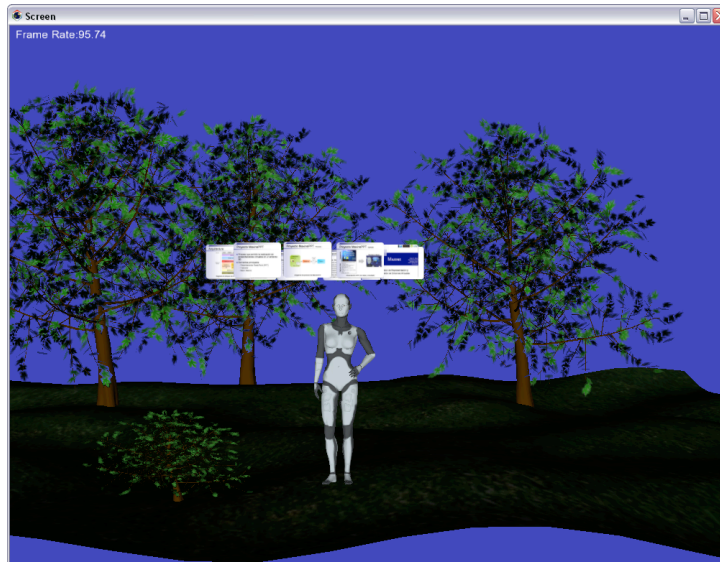


**Fig. 6.** Presentadora virtual “explicando” una transparencia con un índice típico



**Fig. 7.** Presentación de imágenes





**Fig. 8.** Presentadora virtual esperando preguntas de los usuarios.

El sistema desarrollado nos ha permitido plantear diferentes tipos de actividades educativas entre las que destacan las presentaciones virtuales, los tutoriales interactivos para aprender las funcionalidades de un motor gráfico.

En el primer caso, la herramienta ha sido utilizada para realizar presentaciones de temas concretos a los alumnos. El agente virtual, en este caso, actúa como un Experto que informa y guía a los alumnos en un tema específico. En particular se han realizado presentaciones de las actividades del Grupo de Informática Gráfica Avanzada y de las diferentes aplicaciones en el campo de la Informática Gráfica a los alumnos de Grado de Ingeniería Informática e Ingeniería Técnica en Diseño Industrial. Antes y después de estas presentaciones se han realizado encuestas entre los alumnos para comprobar el alcance de la herramienta desde el punto de vista educativo.

Para alumnos más avanzados, el sistema ha sido utilizado como un ejemplo de implementación que motor gráfico. Para ello se han realizado una serie de ejemplos interactivos que permitían mostrar de forma práctica aquellos temas en los cuales se había detectado más dificultades de comprensión por parte de los alumnos. Estos ejemplos permiten realizar modificaciones en tiempo real sobre el motor gráfico para reforzar las explicaciones teóricas del profesor. Por medio de esta herramienta se abordan los siguientes temas de las asignaturas de Informática Gráfica: representación paramétrica de puntos, curvas y superficies, aplicación de transformaciones geométricas sobre objetos, modificar luces y cámaras, modificar el tipo de proyección, cargar animaciones, crear actores simples, asociar voces a los actores, hacerlos hablar, etc.

## 5. CONCLUSIONES

En este trabajo se presenta el desarrollo de una herramienta de animación de código abierto, denominada Maxine, y desarrollada para la gestión de entornos y personajes virtuales. Su potencial y sus posibilidades de uso se presentaron mediante una aplicación específica: MaxinePPT. El motor de animación ha sido diseñado para cumplir los siguientes requerimientos:

- Movimiento natural y capacidad para realizar presentaciones.
- Generación de movimiento en tiempo real y sincronización con la voz.

- Capacidad para representar diferentes escenarios de presentación
- Capacidad para ampliar las posibles aplicaciones.

Esta herramienta se ha utilizado en diferentes asignaturas relacionadas con la Informática Gráfica, con una aceptación muy buena por parte del alumnado. Las encuestas realizadas han comprobado que la introducción de agentes virtuales ha motivado enormemente a los alumnos.

En cuanto al trabajo futuro, varias líneas de investigación permanecen abiertas, la mayor parte de ellas centradas en enriquecer la interacción emocional entre el presentador virtual y los usuarios o espectadores.

## **6. AGRADECIMIENTOS**

Este trabajo ha sido financiado parcialmente por la Dirección General de Investigación, N° TIN2007-63025 y por el Gobierno de Aragón a través del Convenio WALQA Ref. 2004/04/86 y el proyecto CTPP02/2006.

## Referencias bibliográficas y citas

1. Casell J., Sullivan J., Prevost S., Churchill E (eds). (2000): "Embodied Conversational Agents", Cambridge: MIT Press.
2. Noma T., Badler N. I. (1997). "A virtual human presenter"; Proc. IJCAI-97 Workshop on Animated Interface Agents, 45-51.
3. Noma T., Zhao L., Badler N. I. (2000) "Design of a Virtual Human Presenter"; IEEE Journal of Computer Graphics and Applications, Vol.20, No.4, July/August, pp. 79-85.
4. Magnenat-Thalmann N., Kalra P. (1995) "The Simulation of a Virtual TV Presenter", Proc. Pacific Graphics 95, World Scientific, Singapore, pp. 9-21.
5. André E., Rist T., Müller J. (1998) "Integrating reactive and scripted behaviours in a Life-Like presentation agent", Proceedings of the 2nd International Conference on Autonomous Agents, pp. 261-268.
6. Baus J., Butz A., Krüger A. (2000) "Incorporating a Virtual Presenter in a Resource Adaptive Navigational Help System"; Workshop notes of the Workshop on Guiding Users through Interactive Experiences, Paderborn, Germany.
7. Nijholt A., van Welbergen A., Zwiers J. (2005) "Introducing an embodied virtual presenter agent in a virtual meeting room"; Proc. of the 23rd IASTED International Multi-Conference Artificial Intelligence and Applications. Innsbruck, Austria, pp. 579-584.
8. Rist T., André E., Baldes S., Gebhard P., Klesen M., Kipp M., Rist P., Schmitt M. (2003) "A Review of the Development of Embodied Presentation Agents and Their Application Fields in Life-Like Characters: Tools, Affective Functions, and Applications"; Helmut Prendinger and Mitsuru Ishizuka (eds), Springer Series on Cognitive Technologies, pp. 377-404.
9. Yuan X., Chee Y.S. (2005) "Design and evaluation of Elva: an embodied tour guide in an interactive virtual art gallery"; Computer Animation and Virtual Worlds 16, pp.109-119.
10. Rocchi C., Stock O., Zancanaro M., Kruppa M., Krüger A. (2004) "The Museum Visit: Generating Seamless Personalized Presentations on Multiple Devices"; Proceedings of the 9<sup>th</sup> International Conference on Intelligent User Interfaces IUI'04, January 13-16, Madeira, Portugal.
11. Focken D., Stiefelhagen R. (2002) "Towards vision-based 3-D people tracking in a smart room"; Proceedings 4th IEEE International Conference on Multimodal Interfaces.
12. Baldassarri S., Cerezo E., Seron F. (2007) "An open source engine for embodied animated agents"; CEIG'07: XVII Congreso Español de Informática Gráfica, pp. 91-98.
13. Cal3D: Character Animation Library, <http://cal3d.sourceforge.net/>
14. Open Scene Graph, <http://www.openscenegraph.org/>
15. OpenAL++, <http://alpp.sourceforge.net/>
16. Long B. "Speech Synthesis & Speech Recognition Using SAPI 5.1", <http://www.blong.com/Conferences/DCon2002/Speech/SAPI51/SAPI51.htm>
17. OpenSteer, <http://opensteer.sourceforge.net/>
18. LuaBind, <http://www.rasterbar.com/products/luabind.html>
19. STLPort, <http://sourceforge.net/projects/stlport>
20. OpenThreads, <http://openthreads.sourceforge.net/>
21. Gutierrez D., Seron F.J., Magallon J.A., Sobreviela E.J., Gutierrez J.A. (2004) "CLS: A Low Cost Visualization Environment for the Train Industry"; Journal Mécanique & Industries 5: 139-145.
22. Cruz-Neira C., Sandin D.J. DeFanti T. A. (1993) "Surround-screen projection-based virtual reality: The design and implementation of the CAVE"; Computer Graphics (SIGGRAPH '93 Proceedings), vol. 27: 135-142.