



UNIVERSIDAD de VALLADOLID



ESCUELA de INGENIERÍAS  
INDUSTRIALES

## **INGENIERÍA TÉCNICA INDUSTRIAL, ESPECIALIDAD ELECTRÓNICA INDUSTRIAL**

### **PROYECTO FIN DE CARRERA**

# **MODELADO Y ANIMACIÓN FACIAL DE UN AVATAR VIRTUAL REALISTA MEDIANTE FORMAS CLAVE**

Autores:

**Saldaña Burgos, Irene**

**Del Río Blanco, Cristina**

Departamento: **Ingeniería de Sistemas y Automática**

Tutores:

**Gómez García-Bermejo, Jaime**

**Zalama Casanova, Eduardo**

**JULIO – 2013**



## **1. Introducción**

El ser humano está continuamente interactuando con los objetos que le rodean y crea expectativas sobre cómo estos deben comportarse basándose en pasadas experiencias con estos objetos u otros similares. Cuando los seres humanos y las máquinas interactúan lo hacen a través de un medio específico que se define como Interacción Hombre-Máquina, en inglés Human- Machine Interaction (HMI). Los sistemas HMI pueden entenderse como dos procesos de información (humano y equipo) que intentan comunicarse entre sí transmitiéndose mutuamente información, órdenes y datos a través de un software especializado.

El rápido desarrollo y la disminución de los costes de las tecnologías de hardware y software durante los últimos años junto con el desarrollo de nuevas técnicas como la síntesis del habla y el procesamiento de audio y visión por ordenador, ha originado la concepción de nuevas formas de interacción HMI que implican diferentes técnicas de entrada como el habla o la voz, la visión por computador, etc. De esta forma, los usuarios son capaces de interactuar con las máquinas de la misma forma en que se comunican con otras personas para expresar emociones, estados de ánimo, actitudes o su grado de atención. Así aparece un nuevo campo de investigación que centra su trabajo en el área de Interfaces Conversacionales de Usuario y es por eso por lo que ha ido creciendo el interés por las interfaces antropomorfas y expresivas.

Los avatares son una técnica empleada que posibilitan la comunicación hombre – máquina de una forma más natural y familiar para el ser humano como si de un ser real se tratara, facilitando sus labores y haciéndolas más comprensibles. De esto se deduce que estos sistemas antropomórficos, que pretenden una comunicación amigable entre usuario y máquina, han de incorporar las capacidades fisonómicas y sensoriales propias del rostro humano. Estos cumplen una serie de funciones en el sistema, como la de asistencia o guía a través de aplicaciones o elemento de navegación al brindar ayuda en la búsqueda de información, hasta poder mantener una conversación sobre cualquier tema. También, existen asistentes virtuales con fines pedagógicos, y además se pueden emplear para el tratamiento de terapias de trastornos y enfermedades mentales o entrenamiento de habilidades sociales, al poder responder de forma realista y socialmente con él.

De todo esto se puede deducir que el empleo de los avatares puede ser muy beneficioso y su uso es cada vez más común. Además, los avances realizados en las técnicas empleadas en la generación de gráficos y animaciones han hecho posible que se puedan emplear imágenes para incrementar la apariencia humana, obteniendo mayor grado de veracidad y consiguiendo que los entornos virtuales puedan ser más reales y sociales, consiguiendo que el usuario sienta una interacción más auténtica, y tenga la sensación de encontrarse dentro del entorno generado por el ordenador y responder socialmente a este.

El principal objetivo de este proyecto es el desarrollo de nuevas formas de interacción hombre-máquina que incluyan modos, estilos y canales de comunicación propios de la comunicación humana. A tal fin se ha generado un Avatar Virtual realista que posee una interfaz natural con apariencia antropomórfica y capaz de generar expresiones faciales complejas operando en tiempo real. Esto es así ya que la comunicación no verbal es un aspecto muy importante en la interacción social para expresar las emociones de los interlocutores. Cuando uno interactúa con otra persona, las expresiones emocionales de ambos interlocutores son observadas e interpretadas durante la conversación. Así, el interlocutor expresa su estado de ánimo o sus intenciones a los otros, influyendo también en la reacción de los demás hacia él.

## **2. Desarrollo**

Resulta obvio que el diseño de un avatar que presente similitudes con el rostro humano requiera de un cierto conocimiento tanto de la anatomía facial del ser humano como de los aspectos psicológicos asociados a la expresividad del rostro.

Para generar correctamente un Avatar Virtual con capacidad gestual y alto grado de realismo, que opere en tiempo real, es necesario resolver tres problemas principales:

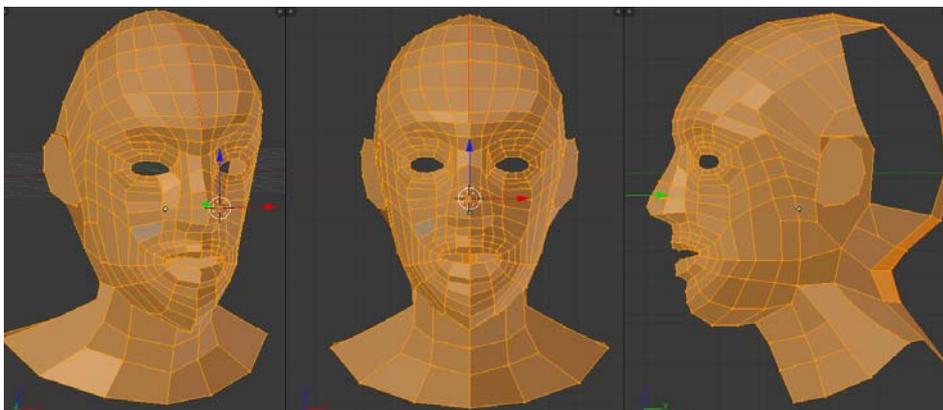
- Generar el modelo tridimensional de partida alcanzando un equilibrio entre dos componentes fundamentales: el nivel de detalle del rostro generado y las posibilidades computacionales del sistema que se dispone.
- Elegir el método más conveniente para la animación, que dote al sistema del mayor realismo posible y buscando la convergencia entre las características propias de la animación 3D de personajes y los fundamentos de la anatomía humana.
- Crear correctamente las animaciones y la aplicación de control del Avatar, siempre aunando de forma eficiente las necesidades y limitaciones de los recursos computacionales y humanos, y capaz de operar en tiempo real.

De entre todos los programas de modelado geométrico y animación existentes, el software elegido para el desarrollo íntegro del modelo ha sido la herramienta Blender, debido a las múltiples cualidades que ofrece. Este software de libre distribución integra un amplio rango de herramientas para la creación de contenidos 3D, desde el modelado hasta el desarrollo de gráficos en movimiento, animaciones y aplicaciones de objetos 3D. Es multiplataforma y compatible con la mayoría de los sistemas operativos. Además, cuenta con una arquitectura 3D de alta calidad, permitiendo un rápido y eficiente desarrollo y un tamaño pequeño de ejecutable para una fácil distribución.

## **2.1. Modelado del Avatar**

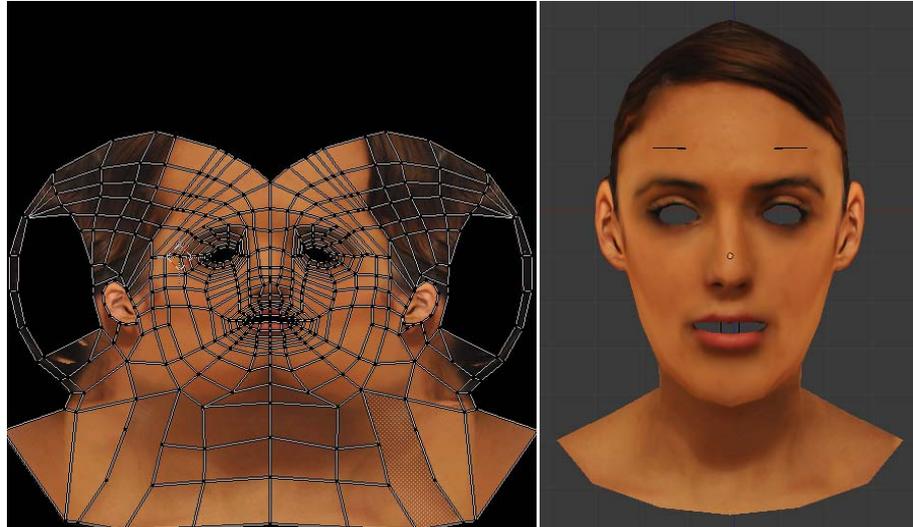
El modelado gráfico de personajes realistas 3D comprende la obtención de datos tridimensionales de un rostro real y la creación de una superficie a la que mediante técnicas de texturizado se la dota de información cromática logrando así un rostro realista.

Al modelar un objeto, hay que tener en cuenta la forma física y la apariencia visual que éste tiene. La forma física de un objeto comprende puntos unidos por líneas para formar superficies o polígonos y la forma que se atribuye a estos polígonos representa la forma del objeto a crear. Los rasgos faciales de un asistente virtual vienen dados por los objetos geométricos que lo conforman. Parece lógico que la forma que adopten estos objetos geométricos esté basada en la estructura del rostro humano. El principal objeto que define la apariencia facial del avatar es la estructura geométrica de la cabeza, por lo que la forma que se atribuya a la superficie poligonal debe ser afín a esta. El nivel de detalle viene determinado por el número de polígonos empleados en la representación ya que, obviamente, a mayor número de polígonos más fiel será la representación de las superficies con relieves curvos y pronunciados, como en la nariz y la región de la oreja. En aplicaciones de tiempo real se debe alcanzar un compromiso entre el nivel de detalle deseado y la velocidad de la animación final generada. A mayor número de polígonos mayor coste computacional empleado en su animación, y mayor espacio es necesario para almacenar el modelo. Uno de los puntos clave al construir u obtener un objeto 3D es lograr una superficie que sea fácilmente deformable entre varias poses. La superficie debe poder contraerse y expandirse de forma similar a la piel del rostro humano por lo que es conveniente que la disposición de los puntos y las aristas de los polígonos se realice teniendo en cuenta las deformaciones que sufrirá la superficie poligonal durante la animación.



*Figura 1 – Superficie del modelo 3D*

Pero al modelar el objeto es tan importante definir la estructura como aportar las características necesarias para otorgarle una apariencia visual realista, es decir, el texturizado. Para que el modelo tenga un aspecto realista, este debe estar dotado de una textura que recree la apariencia humana: el color y las cualidades de la piel, rasgos que caractericen al personaje como pecas, arrugas, vello o imperfecciones y añadir sombras en aquellas zonas que cambien o donde se producen pliegues al gesticular.



*Figura 2 – Resultado del texturizado del Avatar*

Los datos que constituyen el modelo 3D pueden obtenerse modelando la superficie a representar de forma manual, pero hay ciertos detalles como son la cavidad bucal, los ojos o el pelo que se agregan al modelo y se procesan de manera aislada al proceso de modelado de la superficie genérica del modelo.

Además, para generar un avatar virtual animado con alto grado de realismo es tan importante conseguir un óptimo resultado del modelado del avatar como generar un entorno natural que lo englobe y para ello entran en juego las luces, las cámaras y la generación de una escena que dote de mayor realismo al conjunto.



*Figura 3 – Resultado final del modelado y del modelado en escena*

Este proyecto está basado en proyectos anteriores, en los cuales se obtuvo el modelo tridimensional del rostro de la persona a través de un escáner 3D. Sin embargo, en este proyecto se ha obtenido de forma únicamente manual tomando de soporte dos imágenes con la vista frontal y perfil del rostro a modelar y a partir de las cuales se obtuvieron las características necesarias para generar el modelo, se extrajo información

sobre las coordenadas tridimensionales del rostro, las características cromáticas necesarias para la textura y la información necesaria para añadir aquellos elementos que generan el modelo, como los ojos, los dientes o la cavidad bucal. Realizar el modelado empleando una técnica exclusivamente manual supone una gran ventaja, hablamos de que el escáner 3D crea un modelo con una enorme cantidad de superficies poligonales, que aumenta la capacidad computacional y dificulta la creación de las animaciones del modelo. Por el contrario, en este modelo las superficies poligonales están dispuestas de una forma más simplificada, reduciendo su número y mejorando así la capacidad de deformación del modelo, lo que desemboca en unas animaciones más realistas y un gasto computacional menor.

## **2.2. Animación de un Avatar Virtual realista**

El proceso de animación consiste en adecuar la superficie poligonal modelada anteriormente mediante distintas técnicas que engloban la generación del movimiento del modelo, la composición de las animaciones y la generación de la aplicación que controla y emula el comportamiento muscular real.

Para la animación del modelo se han empleado las Formas Clave o *Blend Shapes*, las cuales generan animaciones mediante la deformación de la malla base buscando el equilibrio entre el empleo mínimo de Formas Clave sin que por ello la animación carezca de expresividad. Esta combinación se puede realizar mediante funciones de interpolación especificando el movimiento entre una posición origen y otra final dentro de un intervalo normalizado y calculando las posiciones intermedias por las que los vértices pasan. Las Formas Clave otorgan flexibilidad y expresividad al modelo, lo que apunta a que es necesario la combinación de varias Formas Clave para generar animaciones complejas. Sin embargo, el usuario debe fijar cada uno de ellos, por lo que conviene emplear el menor número de Formas Clave.

Debido a la riqueza de movimientos que el rostro puede generar, es necesario un sistema que sirva de estándar fundamentado en la anatomía del rostro y que describa la actuación de los músculos implicados para formar distintas expresiones en función de la región donde se realice el movimiento, considerando solamente aquellos músculos involucrados y siempre en función de lo que resulta distinguible al ojo humano, es decir, el Sistema de Codificación Facial, o Facial Action Coding System (FACS). Las bases del modelo de FACS establecen que las expresiones se dividen en acciones de pequeños grupos de músculos denominadas Unidades de Acción (AUs o Action Units), las cuales se relacionan anatómicamente con la contracción de un conjunto específico de músculos faciales, además de varias categorías de movimientos y posiciones de la cabeza y los ojos.

De tal forma, se consiguen reproducir distintos gestos y expresiones complejas como pueden ser las seis emociones consideradas universales al ser reconocibles por cualquier persona de cualquier tipo y condición en cualquier parte del mundo: alegría, tristeza, miedo, enfado, disgusto y sorpresa.

Para la animación del modelo generado se han determinado las Unidades de Acción más relevantes activadas en cada una de las seis emociones y mediante la combinación de estas se componen las expresiones teniendo en cuenta tanto la posición como la acción de los músculos que intervienen en la activación de cada Unidad de Acción implicada. Se ha empleado un total de 32 Formas Clave acorde a las diferentes Unidades de Acción presentes en las expresiones logrando que estas sean fácilmente reconocibles e incrementando el realismo del modelo al generar las expresiones y la sensación de naturalidad al interactuar con el Avatar.

## **2.3. Composición de las animaciones y generación de la aplicación de control del Avatar**

Una vez creados todos los componentes que intervienen en la generación de movimiento de los proyectos de contenido 3D y mediante la mezcla sincronizada de las diferentes Formas Clave correspondientes a las AUs que intervienen en cada expresión compleja, se construyen las animaciones. Las animaciones se componen por la sucesión de puntos de control que determinan la trayectoria de la curva de

interpolación que especifica el movimiento del objeto y por tanto de su animación, y estos puntos de control corresponden a las formas clave implicadas en las expresiones. Se ha generado un total de 7 animaciones, cada una correspondiente a cada emoción y una adicional como presentación del Avatar.

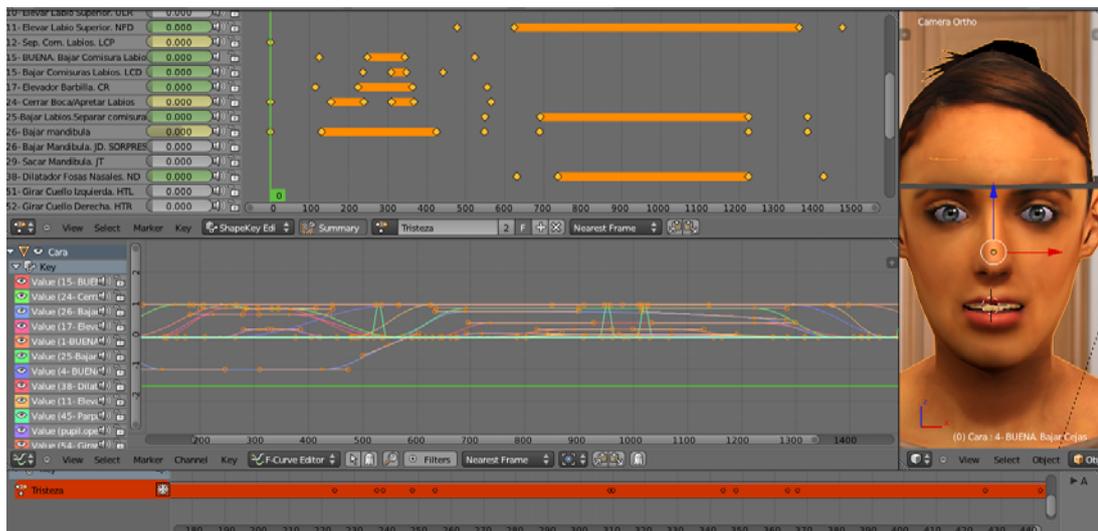


Figura 4 – Visualización del panel de animaciones con los editores y las Formas Claves

Para generar la aplicación interactiva que controla las animaciones del Avatar, se ha empleado el editor de lógica de Blender que permite agregar los sensores, actuadores y controladores precisos y que proporciona una interacción profunda con la simulación. Su funcionalidad puede ser extendida a través de secuencias de comandos de Python. Así, se asocia una tecla para el control de la animación de cada expresión compleja de forma que al pulsar cada tecla se reproduce la animación correspondiente, es decir, un sensor que al activarse envía un pulso al actuador vinculado al controlador de dicha animación, que maneja las acciones inducidas en la armadura y muestra las animaciones generadas de cada expresión compleja.

Empleando el motor de juego de Blender, se genera la lógica del mismo, supervisando un bucle de juego, que procesa la lógica, el sonido, la física y la representación de simulaciones en orden secuencial y se desarrolla en C++. Este permite que una aplicación pueda ser ejecutada sin necesidad de tener que cargar el sistema Blender o tener el programa instalado en el sistema. Para ello, automáticamente al exportar el archivo como reproductor independiente, se generan un ejecutable del archivo y una serie de bibliotecas automáticamente con los complementos, lo que hace que los juegos y aplicaciones interactivas puedan ser distribuidos a otros usuarios sin requerir que posean un conocimiento detallado de Blender y sin la posibilidad de modificación no autorizada.

### **3. Evaluación del Avatar**

Una vez desarrollado el avatar y preparado para su integración, se evaluaron las reacciones de los usuarios al interactuar con él y se comprobó si es capaz de generar emociones fácilmente identificables. Para ello se realizó un estudio para determinar si los observadores del avatar eran capaces de reconocer las emociones básicas que éste genera. La evaluación se centró en dos aspectos fundamentales: la apariencia del avatar tanto estática como dinámicamente y en la capacidad de los sujetos de identificar las expresiones que el avatar producía en ambos casos.

Los resultados extraídos del estudio indican que las expresiones que genera el avatar son identificadas fácilmente y el empleo de las animaciones resulta favorable para su reconocimiento y por lo tanto útiles para la percepción de emociones en el avatar y el empleo del mismo en la interacción hombre-máquina.

Además, se muestra en general una buena aceptación del avatar por parte de los usuarios y la apariencia y expresividad del mismo se han valorado positivamente ya que genera expresiones mediante movimientos naturales y basados en el comportamiento humano.

#### **4. Conclusiones**

Tras la realización del presente proyecto y la evaluación del Avatar Virtual desarrollado, se pueden deducir las siguientes conclusiones:

Se han conseguido los objetivos de generar un avatar virtual con alto grado de realismo tanto en apariencia como en comportamiento, capaz de generar expresiones emocionales de manera natural y fácilmente identificables.

El modelado manual permite simplificar esta etapa del proceso obteniendo un modelo de baja poligonalización alcanzando el equilibrio entre simplicidad y definición, mejorando el tiempo de producción y la calidad final del modelo, reduciendo el número de las superficies poligonales en comparación a otras técnicas de adquisición de datos como los escáneres digitales o digitalizadores. Esto desemboca en un gasto computacional menor y una generación de animaciones más realistas. Aun así, el modelado de objetos 3D mediante técnicas manuales resulta una tarea minuciosa que precisa un procesado y una dedicación mayor.

Se ha desarrollado un modelo de animación facial mediante Formas Clave, fundamentados en la fisionomía muscular real del rostro humano y el sistema de codificación facial o FACS, intentando que la aplicación tenga un elevado nivel de realismo. Las Formas Clave permiten el almacenamiento de formas relativas a partir de una malla base y generan animaciones de una forma sencilla pasando de una forma base a otra mediante la deformación de la malla. La cantidad de detalles en cada expresión puede variar, ya que las animaciones resultantes se generan mediante combinación de las formas clave. Mediante el empleo de esta técnica se obtienen unos resultados óptimos aunque requiere mucha especificación manual y generar animaciones completas puede requerir mucho tiempo.

El avatar resulta adecuado para su uso como asistente virtual o interfaz hombre-máquina, ya que produce expresiones mediante movimientos naturales y basados en el comportamiento humano y es capaz de generar emociones fácilmente identificables. Se podrían incluir otros sistemas perceptivos en los que ya se ha estado trabajando, como por ejemplo el reconocimiento de identidades de los usuarios, un sistema de captura de movimientos, reconocimiento facial, reconocimiento por voz o sistemas de síntesis de voz. De esta manera incluso se podría llegar a mantener una conversación humano – avatar muy realista. Estos sistemas mejorarían considerablemente la interfaz desarrollada, y seguramente implicarían un mayor grado de aceptación por parte de los usuarios del sistema.

Respecto a proyectos anteriores, se ha reducido el número de vértices y caras del modelo, que implica un mejor manejo de éste. El tamaño del archivo y de la aplicación también es menor, lo que conlleva que la aplicación podrá ser ejecutada en cualquier equipo ya que el consumo computacional es menor.

A la vista de las prestaciones conseguidas por el avatar, cabe comentar que podría ser un buen recurso en situaciones en las que es necesaria una interacción entre una persona y una máquina, posibilitando una comunicación bastante más amigable y sencilla en comparación con las existentes. De esta manera, podría utilizarse para proporcionar información a los usuarios, implementarse en internet como sistema de ayuda en páginas web o emplearse en el campo de la domótica.

Cabe añadir que en la animación 3D hay que dominar numerosas disciplinas: artísticas, biológicas, matemáticas, robóticas, etc. Si bien es verdad que la herramienta de modelado elegida ha permitido la abstracción de cálculos matemáticos complejos y de programación. Así mismo, se ha podido comprobar que el estudio de la animación del rostro humano es un trabajo interesante pero a la vez extremadamente complejo. Además, no todas las personas muestran la misma apariencia facial para especificar un estado de ánimo concreto.



UNIVERSIDAD de VALLADOLID



ESCUELA de INGENIERÍAS  
INDUSTRIALES

**INGENIERÍA TÉCNICA INDUSTRIAL, ESPECIALIDAD  
ELECTRÓNICA INDUSTRIAL**

**PROYECTO FIN DE CARRERA**

**MODELADO Y ANIMACIÓN FACIAL DE UN  
AVATAR VIRTUAL REALISTA MEDIANTE FORMAS  
CLAVE**

Autores:

**Saldaña Burgos, Irene**

**Del Río Blanco, Cristina**

Departamento: **Ingeniería de Sistemas y Automática**

Tutores:

Gómez García-Bermejo, Jaime

Zalama Casanova, Eduardo

**JULIO – 2013**



# Índice

<b>Capítulo 1</b>	– Introducción	
1.1.	Marco del proyecto	1
1.2.	Objetivos	2
1.3.	Organización de la memoria	3
<b>Capítulo 2</b>	– El interés de los interfaces virtuales	
2.1.	Introducción	5
2.2.	Las interfaces virtuales	6
2.2.1.	Tipos de interfaces virtuales	6
2.2.2.	La realidad virtual	8
2.2.2.1.	Características básicas de la realidad virtual	8
2.2.2.2.	Aplicaciones de la realidad virtual	9
2.2.3.	Los asistentes virtuales	10
2.3.	Expresividad en interfaces de realidad virtual	11
2.3.1.	El papel de las expresiones emocionales en la interacción social	11
2.3.2.	La expresión emocional en el ámbito médico	12
2.3.2.1.	Empleo de avatares virtuales en terapias	14
<b>Capítulo 3</b>	– Fundamentos del modelado y animación facial de un avatar realista	
3.1.	Introducción	17
3.2.	Herramientas de modelado gráfico	18
3.2.1.	Principales programas de modelado gráfico	18
3.2.2.	Blender	19
3.3.	Modelado y animación facial	20
3.4.	Características básicas del modelado facial	21
3.5.	Métodos de representación del modelo 3D facial	22
3.5.1.	Superficies poligonales	22
3.5.2.	Superficies paramétricas	23
3.5.3.	Subdivisión de superficies	24
3.5.4.	Modelos volumétricos	25
3.6.	Técnicas de modelado facial	26
3.6.1.	Modelado mediante la deformación de un modelo genérico	27
3.6.2.	Modelado mediante la combinación de modelos base	29
3.7.	Técnicas de animación	29
3.7.1.	Interpolación	30
3.7.2.	Parametrización	31
3.7.3.	Técnicas basadas en la simulación	32
3.7.3.1.	Modelos basados en la física muscular	32
3.7.3.1.1.	Técnicas que emplean modelos masa-muelle	32
3.7.3.1.2.	Técnicas que emplean modelos vectoriales	32
3.7.3.1.3.	Técnicas que emplean capas de superficies elásticas (Layered spring mesh)	37
3.7.4.	Performance driven	37
3.7.5.	Deformación geométrica	38
3.7.6.	Técnicas basadas en la combinación de formas o blend shapes	40

<b>Capítulo 4</b>	– Anatomía y expresiones del rostro	
4.1.	Introducción	43
4.2.	La anatomía del rostro humano.	44
4.3.	Las expresiones faciales	47
4.4.	Facial Action Coding System	49
4.4.1.	Las unidades de acción.	50
4.4.2.	Intensidad de las unidades de acción	53
4.4.3.	Evolución temporal de las unidades de acción	54
4.4.4.	Valencia de las expresiones.	54
4.5.	Expresiones complejas	55
4.5.1.	Descripción de las expresiones básicas universales	55
4.5.2.	EMFACS	57
4.6.	Elección de las unidades de acción a representar.	58
<b>Capítulo 5</b>	- Creación de un avatar animado	
5.1.	Introducción	61
5.2.	El proceso de modelado de un avatar realista	62
5.3.	Adquisición de datos a partir de un rostro real	62
5.4.	Ajuste del modelo	63
5.4.1.	Modelado de la superficie	63
5.4.2.	Ajuste de superficie genérica a modelo	66
5.5.	Texturizado de la superficie	68
5.5.1.	Tipos de mapeado UV	69
5.5.1.1.	Técnicas de proyección simples	69
5.5.1.2.	Técnicas de proyección complejas:	69
5.5.2.	Desenvolver empleando costuras	70
5.5.3.	Desenvoltura de la superficie y aplicación de texturas UV	71
5.6.	Modelado de los ojos y la boca	74
5.7.	Agrupación y emparentado de objetos	76
5.8.	Creación de la escena	77
5.8.1.	Iluminación de la escena	78
5.8.1.1.	Descripción de las fuentes lumínicas en Blender	79
5.8.1.1.1.	Luz solar (Sun Light)	79
5.8.1.1.2.	Luz Hemi	80
5.8.1.1.3.	Luz lámpara	80
5.8.1.1.4.	Luz Focal	81
5.8.1.2.	La iluminación con Blender y las sombras	81
5.8.1.2.1.	Sombras con el buffer	81
5.8.1.2.2.	Sombras mediante el trazado de rayos	82
5.8.1.2.3.	Otro tipo de sombras: La luz Volumétrica	82
5.8.1.3.	La incidencia de la luz en los objetos	83
5.8.1.3.1.	Sombreadores de difusión de luz	83
5.8.1.3.2.	Reflexión especular de la luz	83
5.8.2.	Renderizado de la escena. Las cámaras	85
5.9.	Animación del modelo	86
5.9.1.	Las formas clave	86
5.10.	Composición de las animaciones	90
5.10.1.	El editor de Curvas IPO	91
5.10.1.1.	Edición de fotogramas clave mediante las Curvas IPO	92
5.10.1.2.	Modificaciones y ajustes de las Curvas IPO	93
5.10.2.	El editor de acción	95
5.10.3.	El editor NLA	97
5.11.	Game Engine. Motor de juego	98
5.11.1.	Confección y generación de la aplicación mediante el motor de juego	101

<b>Capítulo 6</b>	- Resultados	
6.1.	Introducción	105
6.2.	Unidades de acción, AUs	105
6.3.	Expresiones complejas	117
6.4.	Resultados comparativos	120
6.5.	Resultados numéricos	123
6.6.	Conclusiones	124
<b>Capítulo 7</b>	- Evaluación del avatar	
7.1.	Introducción	125
7.2.	Método	126
7.2.1.	Diseño	126
7.2.2.	Participantes	128
7.2.3.	Materiales y Aparatos	129
7.2.4.	Procedimiento	129
7.3.	Análisis y resultados	130
7.4.	Discusión	132
7.5.	Conclusiones	135
<b>Capítulo 8</b>	- Estudio económico	
8.1.	Introducción	137
8.2.	Recursos empleados	137
8.3.	Costes directos	138
8.3.1.	Costes del personal	138
8.3.2.	Costes de amortización de equipos y programas	139
8.3.3.	Costes derivados de otros materiales	140
8.3.4.	Costes directos totales	140
8.4.	Costes indirectos	141
8.5.	Costes totales	141
<b>Capítulo 9</b>	- Conclusiones y líneas futuras	
9.1.	Conclusiones	143
9.2.	Líneas futuras	144
<b>Referencias</b>		147
<b>Anexo</b>	- Contenido del CD y utilización	155



# Índice de figuras

Figura 2.1 - Ejemplos de avatares virtuales, el segundo utilizado en la web de la empresa RENFE	11
Figura 3.1 - Fragmento de un modelo generado mediante superficies poligonales	23
Figura 3.2 - Fragmento de un modelo generado mediante curvas NURBS	24
Figura 3.3 - Resultado de subdivisión de superficies de 3 niveles mediante el algoritmo Catmull-Clark	25
Figura 3.4 - Modelado mediante superficies volumétricas	25
Figura 3.5 - Proceso de escaneo láser 3D	26
Figura 3.6 - Modelado mediante la deformación de un modelo genérico. (a)(b) Anotación de características faciales importantes. (c) Deformación del modelo genérico alineándola posición de las características faciales anteriores	27
Figura 3.7 - Efecto de un vector muscular en una red tridimensional	33
Figura 3.8 - Modelado del efecto de un vector muscular en el sector $V1$ Pr Ps. $R_s$ y $R_f$ representan el desplazamiento de $f'$ a lo largo del vector $V1$ $V2$	34
Figura 3.9 - Efecto de interpolación cosenoidal variando el factor de elasticidad	35
Figura 3.10 - Efecto en los límites de acción del músculo variándola constante de elasticidad $K$	35
Figura 3.11 - Curvas de aplicación en las funciones cosenoidales que definen el campo de acción del vector muscular	36
Figura 3.12 - Efecto de un vector muscular en el plano $x$ y relativo a (a), (b) y (c) respectivamente	36
Figura 3.13 - Prisma tricapa del modelo de Lee	37
Figura 3.14 - Bone, problema cinemático y ángulos de cada unión	38
Figura 3.15 - Ejemplo de esqueleto más complejo	39
Figura 3.16 - Rigging como proceso intermedio en la animación	39
Figura 3.17 - Proceso del modelado empleando Blend Shapes	40
Figura 4.1 - Estructura de la piel	44
Figura 4.2 - Estructura del cráneo	45
Figura 4.3 - Estructura muscular	46
Figura 4.4 - Imágenes de la investigación de Duchenne	47
Figura 4.5 - Las seis expresiones universales. De izquierda a derecha y de arriba abajo: disgusto, tristeza, enfado, alegría, miedo y sorpresa	48
Figura 4.6 - Diagrama en dos dimensiones de Russell y Fernández-Dols	49
Figura 4.7 - Músculos que intervienen en las unidades de acción superiores	50
Figura 4.8 - Grados de intensidad de las AUs	53
Figura 4.9 - Ejemplo de expresión con las correspondientes AUs y grados de intensidad	53
Figura 4.10 - Las seis expresiones universales	55
Figura 5.1 - Imágenes genéricas usadas para la construcción del modelo	63
Figura 5.2 - Proceso de creación del modelo 3D	65
Figura 5.3 - Modelo 3D de la cara	65
Figura 5.4 - Resultado final del modelo 3D	66
Figura 5.5 - Imágenes del modelo real	66
Figura 5.6 - Ajuste del modelo genérico al real	67

Figura 5.7	- Modelo 3D ajustado	67
Figura 5.8	- Modelo desenvuelto mediante una cremallera	68
Figura 5.9	- Efecto de la costura al desenvolver el modelo	70
Figura 5.10	- Uso de imagen en cuadrícula para el proceso de unwrap	71
Figura 5.11	- Resultado de texturizar usando sólo la imagen frontal (a) o lateral	73
(b)		
Figura 5.12	- Resultado de texturizar usando ambas imágenes	74
Figura 5.13	- Estructura del ojo	74
Figura 5.14	- Resultado final del ojo	75
Figura 5.15	- Modelado de los dientes superiores	75
Figura 5.16	- Modelado de los dientes inferiores	75
Figura 5.17	- Modelado de la lengua	75
Figura 5.18	- Modelado del interior de la boca	76
Figura 5.19	- Modelado del pelo	76
Figura 5.20	- Resultado final del modelado	77
Figura 5.21	- Modelado de la estancia de la escena	77
Figura 5.22	- Modelado de la indumentaria.	78
Figura 5.23	- Principio básico de motor de renderizado de Blender	78
Figura 5.24	- Esquema de la luz solar y su efecto	80
Figura 5.25	- Esquema de luz hemi y su efecto	80
Figura 5.26	- Esquema de luz focal	81
Figura 5.27	- Posición de la Luz Hemi en la escena	82
Figura 5.28	- Posición de la cámara y vista de la escena desde ésta	85
Figura 5.29	- Vistas de la escena desde la cámara empleando diferentes	85
perspectivas		
Figura 5.30	- Resultado final del modelado en escena	86
Figura 5.31	- Ejemplo de empleo de formas clave	87
Figura 5.32	- Editor de Curvas IPO	90
Figura 5.33	- Editor de acciones	90
Figura 5.34	- Editor NLA	91
Figura 5.35	- Editor de Curvas IPO Modo Curva	92
Figura 5.36	- Editor de Curvas IPO Modo Formas Clave	93
Figura 5.37	- Editor de Curvas IPO Modo Edición	93
Figura 5.38	- Ajuste horizontal de formas Clave en el Editor de Curvas IPO	94
Figura 5.39	- Ajuste de formas Clave en el Editor de Curvas IPO	94
Figura 5.40	- Interpoladores de Curvas ISO. Constante, Lineal y Bezier,	94
respectivamente.		
Figura 5.41	- Extensión lineal, extrapolación, Cíclico y Cíclico Extrapolado,	95
respectivamente.		
Figura 5.42	- Editor de Acción	96
Figura 5.43	- Visualización del panel de animaciones con los diferentes editores.	98
Figura 5.44	- Detalle de los sensores empleados para controlar las animaciones.	101
Figura 5.45	- Esquema de las conexiones en el Motor de Juego	102
Figura 6.1	- Imagen ejemplo de la AU1	106
Figura 6.2	- Posición de reposo (izquierda) y posición final de la AU1 (derecha)	106
Figura 6.3	- Imagen ejemplo de la AU2	106
Figura 6.4	- Posición de reposo (izquierda) y posición final de la AU2 (derecha)	107
Figura 6.5	- Imagen ejemplo de la AU4	107
Figura 6.6	- Posición de reposo (izquierda) y posición final de la AU4 (derecha)	107
Figura 6.7	- Imagen ejemplo de la AU5	108
Figura 6.8	- Posición de reposo (izquierda) y posición final de la AU5 (derecha)	108
Figura 6.9	- Imagen ejemplo de la AU6	108
Figura 6.10	- Posición de reposo (izquierda) y posición final de la AU6 (derecha)	108
Figura 6.11	- Imagen ejemplo de la AU7	109

Figura 6.12 - <i>Posición de reposo (izquierda) y posición final de la AU7 (derecha)</i>	109
Figura 6.13 - <i>Imagen ejemplo de la AU9</i>	109
Figura 6.14 - <i>Posición de reposo (izquierda) y posición final de la AU9 (derecha)</i>	109
Figura 6.15 - <i>Imagen ejemplo de la AU10</i>	110
Figura 6.16 - <i>Posición de reposo (izquierda) y posición final de la AU10 (derecha)</i>	110
Figura 6.17 - <i>Imagen ejemplo de la AU11</i>	110
Figura 6.18 - <i>Posición de reposo (izquierda) y posición final de la AU11 (derecha)</i>	110
Figura 6.19 - <i>Imagen ejemplo de la AU12</i>	111
Figura 6.20 - <i>Posición de reposo (izquierda) y posición final de la AU13 (derecha)</i>	111
Figura 6.21 - <i>Imagen ejemplo de la AU15</i>	111
Figura 6.22 - <i>Posición de reposo (izquierda) y posición final de la AU15 (derecha)</i>	111
Figura 6.23 - <i>Imagen ejemplo de la AU17</i>	112
Figura 6.24 - <i>Posición de reposo (izquierda) y posición final de la AU17 (derecha)</i>	112
Figura 6.25 - <i>Imagen ejemplo de la AU24</i>	112
Figura 6.26 - <i>Posición de reposo (izquierda) y posición final de la AU24 (derecha)</i>	112
Figura 6.27 - <i>Imagen ejemplo de la AU25</i>	113
Figura 6.28 - <i>Posición de reposo (izquierda) y posición final de la AU25 (derecha)</i>	113
Figura 6.29 - <i>Imagen ejemplo de la AU26</i>	113
Figura 6.30 - <i>Posición de reposo (izquierda) y posición final de la AU26 (derecha)</i>	113
Figura 6.31 - <i>Imagen ejemplo de la AU29</i>	114
Figura 6.32 - <i>Posición de reposo (izquierda) y posición final de la AU29 (derecha)</i>	114
Figura 6.33 - <i>Imagen ejemplo de la AU38</i>	114
Figura 6.34 - <i>Posición de reposo (izquierda) y posición final de la AU38 (derecha)</i>	114
Figura 6.35 - <i>Imagen ejemplo de la AU45</i>	115
Figura 6.36 - <i>Posición de reposo (izquierda) y posición final de la AU45 (derecha)</i>	115
Figura 6.37 - <i>Imagen ejemplo de la AU51, AU52, AU53 y AU54, respectivamente.</i>	115
Figura 6.38 - <i>Posición final de las AU51 (a), AU52 (b), AU53 (c) y AU54 (d).</i>	116
Figura 6.39 - <i>Imagen ejemplo de la AU61, AU62, AU63 y AU64, respectivamente</i>	116
Figura 6.40 - <i>Posición final de las AU61 (a), AU62 (b), AU63 (c) y AU64 (d).</i>	117
Figura 6.41 - <i>Comparación resultado animación e imagen real de alegría</i>	120
Figura 6.42 - <i>Comparación resultado animación e imagen real de tristeza</i>	121
Figura 6.43 - <i>Comparación resultado animación e imagen real de enfado</i>	121
Figura 6.44 - <i>Comparación resultado animación e imagen real de miedo</i>	122
Figura 6.45 - <i>Comparación resultado animación e imagen real de sorpresa</i>	122
Figura 6.46 - <i>Comparación resultado animación e imagen real de disgusto</i>	123



# Índice de tablas

Tabla 4.1	- Unidades de acción (AUs) definidas en el FACS	51
Tabla 4.2	- Descriptores de Watson y Tellegen	54
Tabla 4.3	- Algunas expresiones y AUs implicadas en ellas	58
Tabla 4.4	- AUs más significativas	59
Tabla 4.5	- AUs añadidas relativas al movimiento de ojos y cuello	60
Tabla 5.1	- Listado y descripción de las formas clave	88
Tabla 5.2	- Listado y descripción de los sensores	99
Tabla 5.3	- Listado y descripción de los controladores	100
Tabla 5.4	- Listado y descripción de los actuadores	100
Tabla 6.1	- AUs que intervienen en la expresión de alegría	118
Tabla 6.2	- AUs que intervienen en la expresión de tristeza	118
Tabla 6.3	- AUs que intervienen en la expresión de enfado	118
Tabla 6.4	- AUs que intervienen en la expresión de miedo	119
Tabla 6.5	- AUs que intervienen en la expresión de sorpresa	119
Tabla 6.6	- AUs que intervienen en la expresión de disgusto	119
Tabla 6.7	- Consumo y recursos de la aplicación	123
Tabla 6.8	- Estadísticas del modelo	124
Tabla 7.1	- Cuestionario	127
Tabla 7.2	- Distribución de participantes según sus características individuales	129
Tabla 7.3	- Preguntas realizadas para la evaluación, con los resultados del Alfa de Cronbach	130
Tabla 7.4	- Resultados significativos del ANOVA. Se muestran los factores principales y de interacción estadísticamente significativos, junto con el nivel de significación y los grados de libertad del análisis.	131
Tabla 7.5	- % reconocimiento de imágenes estáticas	132
Tabla 7.6	- % reconocimiento de animaciones dinámicas	132
Tabla 7.7	- Conclusiones más importantes del estudio	134
Tabla 8.1	- Coste anual del personal	138
Tabla 8.2	- Días efectivos por año	138
Tabla 8.3	- Distribución temporal de trabajo	139
Tabla 8.4	- Amortización del material	139
Tabla 8.5	- Costes indirectos	141
Tabla 8.6	- Costes totales	141



# Capítulo 1

## *Introducción*

### **1.1. Marco del proyecto**

El proyecto que aquí se expone está ubicado en el ámbito de la interacción hombre-máquina, en inglés Human – Machine Interaction (HMI). Los sistemas de HMI permiten al ser humano interactuar con máquinas y procesos mediante software especializado. No es tan sólo la mera interacción entre un usuario y la interfaz de un ordenador, sus aplicaciones van más allá. Las técnicas de HMI se aplican en entornos tan complejos como la cabina de un avión, sistemas de telecomunicaciones multimedia, e incluso en intervenciones quirúrgicas. Además actualmente, y con el rápido desarrollo de las tecnologías hardware y software, disponemos de múltiples maneras de interacción usuario – máquina: cámaras, sensores ópticos y micrófonos, entre otros.

Sin lugar a duda, uno de los componentes fundamentales en la comunicación entre personas es la capacidad gestual del rostro. Y es por eso por lo que en los últimos años ha ido creciendo el interés por las interfaces antropomorfas y expresivas, que posibilitan la comunicación hombre – máquina de una forma más natural y familiar para el ser humano. De esto se deduce que estos sistemas antropomórficos, que pretenden una comunicación amigable entre usuario y máquina, han de incorporar las capacidades fisonómicas y sensoriales propias del rostro humano.

Las cabezas virtuales son muy adecuadas para la interacción entre la componente mecánica y humana, sirven como complemento visual y dotan de un mayor realismo a la componente mecánica. Para generar correctamente una cabeza virtual con capacidad gestual, es necesario resolver tres problemas importantes:

- Generar el modelo tridimensional de partida.
- Elegir el método más conveniente para la animación, que dote al sistema del mayor realismo posible.
- Crear correctamente las animaciones, siempre aunando de forma eficiente las necesidades y limitaciones de los recursos computacionales y humanos.

Aunque sabemos que el campo de la animación se ha venido desarrollando extensamente en el mundo cinematográfico con fines de entretenimiento, la animación tridimensional de personajes realistas es de más reciente creación, debido principalmente a las limitaciones en cuanto a herramientas de procesado gráfico se refiere. No es hasta principios de los 70, con los trabajos de Frederic I. Parke, cuando surge el estudio concreto de la generación de modelos realistas y la animación facial.

Los sistemas de animación facial, por tanto, ayudan a mejorar la cada vez más profunda interacción hombre – máquina, además de esclarecer el difícil comportamiento del rostro humano tanto a nivel fisonómico como su modo de respuesta emocional ante el entorno que lo rodea.

## **1.2. Objetivos**

En este proyecto se pretende desarrollar un sistema virtual de animación facial humana tridimensional con capacidad gestual. Se buscará generar un modelo de animación que presente el mayor realismo facial posible, teniendo en cuenta las barreras y limitaciones computacionales, así como la complejidad propia de la simulación de un rostro humano. Como objetivo final se espera conseguir un sistema de animación gestual realista, que pueda ser controlado externamente, y que exprese satisfactoriamente las 6 emociones básicas del ser humano, a saber: alegría, tristeza, sorpresa, enfado, miedo y disgusto.

El presente proyecto está basado en proyectos anteriores, en los cuales se obtuvo el modelo tridimensional del rostro de la persona a través de un escáner 3D. Sin embargo, en este proyecto se plantea la obtención del modelo a partir de simples fotografías, lo cual proporciona una gran ventaja. Hablamos de que el escáner 3D crea un modelo con una enorme cantidad de superficies poligonales, que aumenta la capacidad computacional y hace más difícil la creación de las animaciones del modelo. Por el contrario, en el modelo que se propone las superficies poligonales estarán dispuestas de la forma que nos interesa, reduciendo su número y mejorando así la capacidad de deformación del modelo, lo que desembocará en unas animaciones más realistas.

Para llegar a cumplir el objetivo final, hay que conseguir unos objetivos más específicos, que son:

- Estudio de las muchas técnicas y metodologías de modelado y animación facial existentes, eligiendo la que más conviene a nuestro proyecto.
- Comprensión de la anatomía facial humana y la estructura de su sistema muscular, así como del FACS (sistema de codificación de acciones faciales). De esta manera conoceremos el efecto que los músculos producen en el rostro en cada expresión humana, para lograr una animación lo más realista posible.
- Búsqueda comparativa de las distintas herramientas de modelado gráfico y de animación que existen en el mercado. Elección, si procede, de aquellas que puedan resultar útiles.
- Estudio de las distintas técnicas de cinemática que podemos usar en la animación y elección de la más adecuada.
- Implementación de la animación a partir de los resultados obtenidos, siempre buscando el mayor equilibrio posible entre la parte fisonómica y la apariencia visual.
- Implementación de la aplicación para su control.
- Análisis de resultados y extracción de conclusiones respecto al objetivo final considerado.

- Propuesta de líneas futuras de trabajo, que permitan mejoras en el desarrollo de posteriores proyectos.

### **1.3. Organización de la memoria**

La memoria que aquí se expone está dividida en 8 capítulos y un último apartado de referencias. Los contenidos de cada uno de ellos son los siguientes:

- CAPITULO 1 – *Introducción*

Este capítulo presenta la temática del proyecto y sus objetivos generales.

- CAPITULO 2 – *El interés de los interfaces virtuales*

Se realiza un estudio sobre la importancia de los interfaces y asistentes virtuales y sus tipos. Se explica también el papel de las expresiones emocionales en la vida social, así como en el caso de individuos con daños neurales, y centrándonos algo más en el caso especial de la esquizofrenia.

- CAPITULO 3 – *Fundamentos del modelado y animación facial de un avatar realista*

En este apartado se exponen las distintas herramientas y técnicas del modelado gráfico y de animación, y los métodos de representación del modelo 3D.

- CAPITULO 4 – *Anatomía y expresiones del rostro*

Estudio de la anatomía del rostro humano, del sistema de codificación de acciones faciales (FACS) y de las unidades de acción.

- CAPITULO 5 – *Creación de un avatar animado*

En este capítulo se explica más detalladamente todo el proceso de creación del avatar, desde el proceso de modelado de la superficie hasta la generación de las animaciones, pasando por el ajuste del modelo genérico al real y el texturizado de éste.

- CAPITULO 6 – *Resultados*

Elección de las unidades de acción necesarias para la representación de las expresiones complejas y conclusiones.

- CAPITULO 7 – *Evaluación del avatar*

Una vez expuesto todo lo anterior, se explica el diseño del proyecto elegido, los materiales y herramientas utilizados y se realiza un estudio sobre la viabilidad del proyecto.

- CAPITULO 8 – *Estudio económico*

Cálculo de los costes directos, indirectos y totales que supone la realización del proyecto

- CAPITULO 9 – *Conclusiones y líneas futuras*

Por último, se exponen las conclusiones obtenidas y las posibles líneas de trabajo futuras que permitan mejorar la animación desarrollada.



# Capítulo 2

## *El interés de los interfaces virtuales*

### **2.1. Introducción**

En los últimos años está creciendo el uso de aparatos electrónicos, como ordenadores, móviles, tablets o PDAs, ya que además de facilitar ciertas tareas en la vida laboral también cumplen funciones de entretenimiento o aprendizaje. Este crecimiento no sólo se debe a que el precio de estos productos está bajando, sino a que están mejorando continuamente y su uso resulta cada vez más sencillo y atractivo. Sin embargo, sigue existiendo la necesidad de conseguir una comunicación más personal y natural, basada en la comunicación real entre humanos. Así, aparece un nuevo campo de investigación que centra su trabajo en el área de Interfaces Conversacionales de Usuario, incluyendo los Asistentes Virtuales. Estos asistentes, denominados también Avatares, son personajes virtuales cuyo objetivo principal es hacer la comunicación entre el usuario y la máquina más natural e interactiva. Estos cumplen una serie de funciones en el sistema, como la de asistencia o guía a través de aplicaciones o simplemente de presentar la información [Ca04].

La principal ventaja de utilizar avatares en las interfaces, es la ilusión generada al usuario de la existencia de un personaje con el cual puede interactuar como si de una persona real se tratase, lo que no sólo ayuda en aspectos técnicos de la interacción, sino también en los psicológicos. Esta ilusión se consigue dando al avatar la posibilidad de expresar emociones, realizar gestos faciales y corporales.

El uso de avatares está extendido a todo tipo de entornos, por ello la definición de su comportamiento no puede estar limitada a expertos programadores, así surge la necesidad de definir las tareas de animaciones de personajes virtuales [Ca04].

## **2.2. Las interfaces virtuales**

Lewis y Rieman [CI93] definen las interfaces hombre - computadora como: Las interfaces básicas de usuario son aquellas que incluyen cosas como menús, ventanas, teclado, ratón, los "beeps" y algunos otros sonidos que la computadora hace, en general, todos aquellos canales por los cuales se permite la comunicación entre el hombre y la computadora. Las máquinas y herramientas extienden nuestras capacidades permitiéndonos hacer cosas que sin ellas no podríamos, pero a su vez necesitamos algo que nos conecte con ellas, algo que las haga accesible a nuestro cuerpo, es decir, la interface. La idea fundamental en el concepto de interfaz es el de mediación entre el hombre y la máquina. La interfaz es lo que "media", facilita la comunicación o interacción entre dos sistemas de diferente naturaleza.

Cuando la herramienta en cuestión es un sistema informático, se tiene que hacer distinción entre la parte física y la parte virtual de la interface. El usuario dirige el funcionamiento de los ordenadores mediante instrucciones, denominadas genéricamente como entradas, las cuales se introducen mediante diversos dispositivos y se convierten en señales electrónicas que pueden ser procesadas. A continuación, estas señales electrónicas son coordinadas y controladas por la unidad de proceso central y el soporte lógico. Una vez ejecutadas las instrucciones indicadas por el usuario, se comunican los resultados mediante señales electrónicas que se transmiten a dispositivos de salida. La finalidad de los dispositivos de entrada/salida es la de conectar el cuerpo al ordenador para hacer accesible los datos, pero sin embargo, los ordenadores son una meta-herramienta, en el sentido que permite ejecutar otra herramienta (la aplicación). La herramienta física manipula la herramienta lógica, y esta, manipula información. Es decir, la herramienta informática se debe manifestar de alguna forma a través de los dispositivos de entrada/salida. Esta manifestación, es la parte virtual de la interface, también denominada interface de usuario. Así distinguimos la interface física, como los dispositivos de entrada/salida, y la interface virtual, la cara visible de la herramienta lógica representada por estos dispositivos de entrada/salida, es decir, está constituida por las representaciones que son mediatizadas a través de los dispositivos de entrada/salida.

### **2.2.1. Tipos de interfaces virtuales**

Los diferentes tipos de interacción se refieren a los diferentes estilos de comunicación entre una persona y una máquina basada en una plataforma tecnológica, que [Fo90] define como el elemento físico de intercambio de datos de entrada/salida entre una máquina y un ordenador para realizar tareas.

Según [Fe07] se distinguen los siguientes estilos de interacción:

- Lenguajes de línea de comandos: la interacción se realiza tecleando los comandos en la computadora que muestra un mensaje para expresar que está lista para aceptar comandos. Este estilo de interacción proporciona una forma de expresar instrucciones directamente a la computadora mediante teclas de función, caracteres, abreviaturas o comandos.
- Los interfaces de líneas de comandos: ofrecen acceso directo a la funcionalidad del sistema y pueden combinarse para aplicar un conjunto de herramientas sobre los datos. La desventaja de este método es que los comandos de texto en general son difíciles de aprender y usar, ya que emplean comandos trópicos de sintaxis estricta que el usuario debe conocer previamente. Este tipo de interfaces son más adecuadas para usuarios expertos.

- Menús: Los menús son un conjunto de opciones mostradas en pantalla donde la selección y ejecución de una o más de ellas origina un cambio en el estado de la interfaz [Pree94]. Al contrario que los sistemas regidos por línea de comandos, los menús presentan la ventaja de que los usuarios no tienen que recordar el comando que desean ejecutar, solo reconocerlo. Dentro de los menús se distinguen los siguientes tipos [Sh92]:
  - Menús pull-down
  - Menús pop-up
  - Menús jerárquicos

El empleo de menús hace decrecer el número de errores cometidos por el usuario y se reduce el tiempo necesario para realizar tareas sencillas, pero en el caso de tareas complejas, el empleo de menús incrementa el tiempo necesario para encontrar todas las opciones necesarias.

- Manipulación directa: Muestran un conjunto de objetos en pantalla y un repertorio de posibles manipulaciones que el usuario puede realizar sobre estos. Este tipo de interfaz es muy popular, al incluir manipulaciones análogas a acciones cotidianas (apuntar, arrastrar, mover objetos) lo que hacen que los errores de sintaxis se reduzcan y aumente la velocidad de realización de las tareas [Sh97].
- Formularios: Es el estilo más simple de interacción [Sh92]. En él, se pide al usuario que responda una serie de cuestiones con un formato predeterminado. El único tipo de interacción es proveer información útil como dato de entrada a las aplicaciones.
- Lenguaje natural: En este tipo de interacción los usuarios emplean el lenguaje natural como medio de comunicación con la máquina. Este tipo de interacción es muy recurrida cuando los usuarios no tienen acceso a dispositivos de entrada o teclados, o tienen experiencia limitada. Pero el principal problema radica en que los usuarios deben aprender qué frases reconoce la máquina y que los comandos de voz deben ser fácilmente reconocidos, siendo a veces muy estricto por la pronunciación o la entonación. Por ello a veces se recurre al empleo de lenguaje no verbal. Los gestos pueden ser empleados para crear frases fluidas y permite expresar conceptos importantes de forma clara, apropiada y natural. El empleo de lenguaje no verbal como lenguaje natural cada vez está más extendida ya que permite una comunicación más natural que empleando únicamente comandos de voz [Bux90].
- Diálogos de pregunta/respuesta y consulta: Se trata de un mecanismo simple para proveer información en un dominio específico. Se realiza al usuario una serie de cuestiones, de forma que es guiado paso a paso a lo largo del proceso de interacción. Estos interfaces son fáciles de aprender y usar, pero tiene muchas limitaciones en cuanto a sus capacidades. Se suele emplear para recopilar información del usuario y almacenarlo en una base de datos.
- Interfaces WIMP: los interfaces WIMP (Windows, Icons, Menus and Pointers/ Ventanas, Iconos, Menús y Punteros) son los interfaces más populares en entornos interactivos. Las ventanas son áreas de pantalla que se comportan como terminales independientes. Los iconos, son pequeñas imágenes empleadas para representar una ventana, un archivo o cualquier otro tipo de objeto. Este sirve de interfaz para apuntar, seleccionar, arrastrar objetos en la ventana que pueden moverse, editarse, explorarse o ejecutarse para que se adapten mejor al punto de vista del usuario. Este tipo de elementos emplean métodos gráficos y de visualización para proporcionar un interfaz más accesible que aquellos basados en líneas de comandos, proporcionando al usuario un entorno sencillo de trabajo. Sus principales ventajas son su aplicación general a todo tipo de tareas, convertir las funcionalidades en herramientas explícitas y permitir el soporte y ayuda inmediatos.

Resulta la forma más rápida y efectiva de intercambio de información que cualquier otro método.

- **Realidad virtual:** Consiste en “la simulación de medios ambientes y de los mecanismos sensoriales del hombre por computadora, de tal manera que se busca proporcionar al usuario la sensación de inmersión y la capacidad de interacción con medios ambientales artificiales (por ejemplo, a través de gestos naturales)” [Pree94]. Además, nuevos estilos de interacción han surgido con este método: dispositivos basados en la entrada/salida de datos mediante habla, visión (por ejemplo mediante gestos), interfaces de audio, táctil señales biofísicas... que brindan una nueva generación de interfaces que no se basen en comandos. [Rau03].

### **2.2.2. La realidad virtual**

Particularmente, estamos interesados en el empleo de la realidad virtual como medio de interacción entre los usuarios y las máquinas. La realidad virtual, permite a una persona interactuar con un ambiente simulado por medio de la tecnología informática. La experiencia de la interacción no se limita únicamente a sensaciones visuales, también incluye otros sentidos como el oído o el tacto empleando tecnologías especialmente diseñadas para ello.

Un sistema, para poder ser considerado de realidad virtual, debe ser capaz de generar digitalmente un entorno tridimensional en que el usuario se sienta presente y en el cual pueda interactuar intuitivamente y en “tiempo real” con los objetos que encuentre dentro de él. Los objetos virtuales deben ser tridimensionales, poseer propiedades propias, tales como fricción y gravedad y mantener una posición y orientación en el ambiente virtual independiente del punto de vista del usuario. El usuario deber tener libertad para moverse y actuar dentro del entorno sintético de un modo natural. De tal forma que la sensación de presencia será mayor cuanto más sean los canales sensoriales estimulados.

#### **2.2.2.1. Características básicas de la realidad virtual**

Se pueden distinguir tres fases o estados de la realidad virtual:

- 1) **Pasivo:** Son entornos inmersivos no interactivos. Es un entorno virtual en el cual podemos ver y oír y quizás sentir lo que sucede. El entorno puede moverse lo que da sensación de movimiento pero no es posible controlar el movimiento. En sentido estricto se trata de una pseudo-realidad virtual.
- 2) **Exploratorio:** Son sistemas que permiten desplazarse por un entorno virtual para explorarlo lo que supone un salto cualitativo en cuanto a funcionalidad. Es el estadio habitual de los paseos arquitectónicos y de las obras de arte virtuales.
- 3) **Interactivo:** Un sistema virtual interactivo permite experimentar y explorar el entorno y, además, modificarlo. A su vez dentro de un sistema de realidad virtual podemos distinguir diferentes niveles de interactividad. También es importante tener en cuenta las características de las interfaces de comunicación entre el usuario y el sistema. Cuanto menos intrusivos y más intuitivos sean los medios utilizados, mayores serán las posibilidades de acción del usuario dentro del entorno virtual.

Básicamente podemos distinguir entre tres tipos de realidad virtual agrupándolos en función de sus principales características:

- 1) Sistemas de sobremesa: se trata de sistemas no inmersivos que presentan el entorno digital en la pantalla de un ordenador. El usuario puede interactuar y desplazarse por él. Son plataformas adecuadas para el diseño industrial y otras aplicaciones que requieran sistemas avanzados de visualización 3D.
- 2) Sistemas proyectivos: Se trata de sistemas que intentan proporcionar la sensación de inmersión mediante la proyección de imágenes del mundo virtual en las paredes de un espacio cerrado dentro del cual se encuentra el usuario. El usuario controla sus movimientos en el entorno inmaterial y en algunos casos puede también interactuar con los objetos que encuentra en él, mediante el uso de un interfaz adecuado. Este tipo de sistema se adapta bien a las aplicaciones multiusuario. Los simuladores de vuelo y otros simuladores de conducción utilizan sistemas proyectivos basados en conceptos similares a los descritos. Los vehículos suelen incluir plataformas móviles para simular el movimiento físico. Son muy utilizados en presentaciones de arte virtual, demostraciones comerciales y aplicaciones educativas, estas plataformas se sitúan a medio camino entre los sistemas proyectivos tipo CAVE y los equipos de sobremesa.
- 3) Sistemas inmersivos: El objetivo es conseguir que el usuario tenga la sensación de encontrarse dentro del entorno generado por el ordenador. Para esto el equipo utilizado debe estar equipado de dispositivos capaces de estimular el mayor número de sentidos posibles. Este sistema, al que se conoce como realidad aumentada, se puede considerar un híbrido entre la experiencia material y la simulación digital. El uso de la realidad aumentada ofrece perspectivas prometedoras para aplicaciones médicas y para todas aquellas actividades que requieran simultáneamente la manipulación de dispositivos complejos y el acceso a datos e informaciones complementarias.

### **2.2.2.2. Aplicaciones de la realidad virtual**

En la mayoría de simulaciones realizadas en entornos visuales, basta con emplear displays 3D y dispositivos de sonido estéreo para provocar al usuario mediante imágenes y sonidos, la sensación de inmersión dentro del espacio virtual. No obstante, además de provocar en el usuario esta sensación de inmersión, debemos proporcionarle la posibilidad de interactuar con el medio virtual, pudiendo establecer entre el usuario y el entorno virtual una transferencia bidireccional y en tiempo real de información. Para ello, a veces se recurre a otros dispositivos en función de las características que deseamos controlar en nuestra aplicación, como interfaces tipo háptico, dispositivos Desktop, que permiten la interacción puntual con el objeto virtual a través de un terminal, lápiz, joystick o dedo virtual, o guantes.

El abanico de aplicaciones de la realidad virtual es muy amplio, y cada día su uso está más extendido. Aunque en rasgos generales, sus principales aplicaciones en los diversos campos son las siguientes:

- Medicina: Simuladores quirúrgicos para entrenamiento médico, aplicaciones en cirugía, tratamiento de trastornos de diferentes índoles: ansiedad de tipo fóbico (miedo a volar y a las alturas, fobia a insectos...), trastorno obsesivo compulsivo, desórdenes de déficit de atención [Ri00] [Ch02], entrenamiento de habilidades sociales en personas psicóticas, investigaciones del comportamiento interpersonal de pacientes con enfermedades mentales, tratamiento de trastornos alimenticios, evaluación cognitiva de los pacientes

[Ku03], simulación de actividades de la vida diaria para el tratamiento de pacientes con ‘ictus’ [Le88], etc.

- **Educacional:** Puede resultar particularmente útil en formación de los alumnos en disciplinas básicas, apoyada en imágenes 3D interactivas, como recurso didáctico para motivar y atraer la atención de los estudiantes, etc. Además, también se puede emplear en la formación de personal y entrenamiento de profesionales de diferentes campos.
- **Entretenimiento:** El desarrollo de la realidad virtual en este campo es sumamente reconocido, y está en continua evolución incorporando nuevas técnicas que permiten al usuario una interactividad e inmersión mayor al ser capaz de sentir y manipular objetos virtuales, etc.
- **Industria:** La realidad virtual se puede emplear en la planificación de una planta industrial, su remodelación o diseño, reproducción de equipamientos y procesos involucrados con el fin de no dañar los equipos en ocasiones muy costosos, o como tele operación de maquinaria en entornos hostiles o centros de información de instalaciones industriales.
- **Comercio:** Mostrar a los clientes aquellos productos que serían demasiado caro enseñar de otra manera o simplemente no están construidos porque se realizan a medida. Además, hoy en día las empresas integran estas tecnologías con fines promocionales, de marketing, o de asistencia virtual en páginas web.
- **Artes gráficas:** Exhibiciones virtuales de arte, museos, simuladores de entornos protegidos, etc.

### **2.2.3. Los asistentes virtuales**

Como hemos visto anteriormente, la realidad virtual se puede emplear como simulador de entornos, objetos o como recurso en educación o en empresas (se puede ver un ejemplo en la *Figura 2.1*). Pero además, según [Bu94], la realidad virtual consiste en la utilización de dispositivos que hacen que el usuario tenga la misma sensación que si estuviera interactuando realmente con un objeto físico determinado. Un aspecto de la realidad virtual como elemento de interacción, es la creación de asistentes virtuales o avatares. Un asistente virtual o avatar, es un personaje conversacional generado como programa informático capaz de reconocer, al menos de forma básica un lenguaje natural que simula una conversación para dar información y ofrecer un servicio mediante la voz o texto para los usuarios. Un asistente virtual incorpora comprensión de lenguaje natural, control de diálogo, conocimiento del dominio (el tema que debe tratar con el usuario) y un aspecto visual (fotografías o animaciones 3D) que en los más avanzados cambian o simulan estados de ánimo de acuerdo al contenido del diálogo. Los métodos de interacción son de texto a texto, texto a voz, de voz a texto o de voz a voz. Está formado por dos componentes independientes: la interfaz de usuario (IU), que es el aspecto que tendrá, es decir, una imagen 3D con emociones, y la base de conocimientos.



Figura 2.1 – Ejemplos de avatares virtuales, el segundo utilizado en la web de la empresa RENFE

Por lo que un avatar es un ser animado 3D que forma parte de la interfaz de un sistema o aplicación, que puede ser la representación de una persona real, o bien representar un ser imaginario con el que el usuario puede interactuar como si de un ser real se tratara, manteniendo conversaciones con él y expresando sus emociones o respondiendo de forma lógica y coherente de acuerdo al usuario.

Los avatares son una técnica empleada para que el usuario interactúe con máquinas como si de un ser real se tratara facilitando sus labores y haciéndolas más comprensibles. Pero no es su única aplicación, un avatar puede emplearse como asistente virtual, ayudando a los usuarios, indicando los pasos que tienen que dar para realizar sus tareas, o como elemento de navegación o ayuda en páginas web. También, existen asistentes virtuales con fines pedagógicos para entornos virtuales, que ayudan a los alumnos a utilizar y reparar una maquinaria virtual, como es Steve (Soar Training Expert for Virtual Environments) [Joh97]. Además se emplean para el tratamiento de terapias, trastornos, enfermedades o entrenamiento de habilidades sociales, al poder responder de forma realista y socialmente con él. Asimismo, también se usan para ayudar a las personas a comunicarse con otros, cumpliendo el rol de representación de usuarios. Este tipo de avatares se pueden encontrar en videojuegos como representación de los usuarios, o se emplean en foros de internet como Second Life, donde los avatares pueden interactuar con otras personas a través de sus personajes. Como se puede observar, la aplicación de los avatares como método de interacción ya sea como entrenamiento de habilidades sociales como para interactuar con otros usuarios es cada vez más extensa.

De todo esto se puede deducir que el empleo de los avatares puede ser muy beneficioso y su uso es cada vez más común. Además, los avances realizados en las técnicas empleadas en la generación de gráficos y animaciones han hecho posible que se puedan emplear imágenes para incrementar la apariencia humana, obteniendo mayor grado de veracidad y consiguiendo que los entornos virtuales puedan ser más reales y sociales, consiguiendo que el usuario sienta una interacción más auténtica, y tenga la sensación de encontrarse dentro del entorno generado por el ordenador y responder socialmente a este.

## ***2.3. Expresividad en interfaces de realidad virtual***

### ***2.3.1. El papel de las expresiones emocionales en la interacción social***

Un aspecto importante en la interacción social es la comunicación no verbal empleada para expresar las emociones de los interlocutores. Cuando uno interactúa con otra persona, las expresiones emocionales de ambos interlocutores son observadas e interpretadas durante el flujo

de la conversación [St96]. Así, el interlocutor expresa su estado de ánimo o sus intenciones a los otros, influyendo también en la reacción de los demás hacia nosotros.

La expresión de las emociones durante la interacción social, forma parte del sistema de comunicación de los seres humanos y ha evolucionado a lo largo de los años. Darwin fue el primero en poner en manifiesto la importancia de las emociones en su libro “La expresión de las emociones en animales y en el hombre”, donde señalaba que las emociones cumplían dos funciones: (1) facilitar la adaptación del organismo al medio, y su supervivencia al reaccionar de manera adecuada ante situaciones de emergencia, y (2) servir como medio de comunicación de sus intenciones a otros animales mediante la expresión de la conducta emocional.

Es importante la capacidad de interpretación de las distintas señales emocionales del interlocutor, que según [Ne76] se realiza de forma innata para un conjunto básico de emociones, evolucionando a medida que aumenta la interacción social, hasta permitir detectar expresiones más sutiles (resignación, frustración, aburrimiento, etc.) [St96].

La inteligencia emocional puede definirse entonces como la capacidad de reconocer sentimientos propios y ajenos, así como la habilidad para manejarlos y emplearlos como fuente de motivación personal [Go95]. Se cree que la inteligencia emocional se encuentra separada de los procesos intelectuales, ya que se expresan mediante actividad cerebral localizada en distintas regiones del cerebro. Así, el centro de inteligencia emocional se sitúa en la región del neocortex, mientras que los centros emocionales se hallan en el subcortex, en la región inferior del cerebro [Le96]. Según [Go95], para la realización de tareas en entornos sociales y colaborativos, la inteligencia emocional es mucho más importante que los procesos cognitivos. Sin embargo, aunque el sistema cognitivo y emocional de las personas tienen arquitecturas separadas, se cree que están relacionadas de forma intrínseca, y no pueden actuar aisladamente de forma correcta [Le96], [Go95].

Las emociones, además, están directamente relacionadas con la empatía. La empatía puede definirse como la habilidad de sentir, percibir y conceptualizar cómo experimentan el mundo otras personas. Reeves y Nass [Re96] fueron los primeros en discutir que tanto la inteligencia emocional como la empatía hacen que los humanos comprendan y acepten el comportamiento de otros seres vivos, y que apliquen el mismo modelo social para explicar, comprender y predecir el comportamiento de seres inertes. Las personas tienden a emplear modelos sociales para describir la forma de actuar de dispositivos (coches, computadoras, etc.) cuyo funcionamiento les resulta complicado o desconocido, de forma que el comportamiento de dichos dispositivos se ajuste al modelo de interacción con el entorno al que están acostumbrados. Es por eso que desde hace algunos años se han aplicado los modelos de interacción social en humanos al interactuar con máquinas. Hoy en día, el uso de estas técnicas cada vez es más extenso, no sólo para interactuar con máquinas complejas, sino que cada vez es más frecuente ver el empleo de avatares en internet, ya sea como asistente virtual o mecanismo de ayuda de empresas, juegos o perfiles en redes.

De todo lo anterior se desprende que las expresiones emocionales juegan, por tanto, un papel fundamental tanto en las relaciones sociales entre los seres humanos como en la interacción de los usuarios con los dispositivos que les rodean.

### ***2.3.2. La expresión emocional en el ámbito médico***

Las funciones sociales dependen en gran medida del reconocimiento de las emociones comunicadas por otras personas [Li98]. Estudios indican que los pacientes que sufren

enfermedades o trastornos mentales tales como la esquizofrenia, autismo, fobias o desórdenes de déficit de atención, se ven afectados en la capacidad de percibir las expresiones faciales de las emociones [Gur02] [Sch06] [Koh03] [Bru05] [Hoo02].

La cognición social se define como la habilidad del ser humano para percibir las intenciones y emociones de los otros, inferir en lo que piensen, comprender e interpretar las intenciones del otro y las normas que gobiernan las interacciones sociales [Ado01] [Bro90]. En individuos con trastornos mentales de este tipo, ha sido ampliamente reconocida su dificultad de funcionamiento social debido al deterioro de la cognición social. Respecto a la percepción social, [Pe02] estudios refieren que los pacientes desestiman el contexto social cuando procesan los estímulos sociales. Asimismo, [Nu84] comentan que presentarían problemas para captar la información que requiere un razonamiento abstracto. En este sentido, [Zi10] señalan que el deterioro de la cognición social en estos casos se debe a deficiencias en varios mecanismos, incluida la capacidad para pensar analíticamente, y para procesar la información relativa a emociones y señales. Por otro lado, [Le01] señalan las dificultades que los pacientes presentan en relación con el procesamiento emocional y con cuatro factores fundamentales: la abstracción, la familiaridad, la complejidad de la situación y el procesamiento semántico. Según esto, se puede concluir que el déficit en la cognición social podría estar en el origen de las dificultades interpersonales y en el deterioro del funcionamiento psicosocial que se observa en los enfermos de este tipo.

En el caso de la esquizofrenia, además, la dificultad de reconocer las emociones faciales es una característica permanente desde el inicio clínico de la enfermedad. Se ha indicado que las personas que sufren esta enfermedad no perciben los rostros de manera integradora, sino como una suma de partes y el déficit radica en este análisis configuracional [Ro10]. Numerosos estudios demuestran que los pacientes tienen mayor dificultad a la hora de reconocer las emociones negativas, como ira o miedo, además de reconocer erróneamente y otorgar de valor emocional a las expresiones neutrales.

Entre las estructuras cerebrales implicadas en el reconocimiento de las emociones destacan, principalmente, la corteza tempor-occipital, en especial el giro fusiforme, las zonas órbito-frontal y parietal derecha, la amígdala y los ganglios basales, entre otras [Ro10]. Según [Ad02], el reconocimiento de las emociones faciales se consigue por medio de tres estrategias complementarias: la percepción, el análisis y la identificación, que consisten en la activación de zonas de la corteza motora y visual que tendrían el efecto de representar interiormente las posturas observadas y generar el estado emocional correspondiente al observado. También se ha postulado una relación entre el estado clínico y la capacidad de reconocer las emociones faciales. Los pacientes estabilizados, tienen un desempeño mejor que aquellos que se encuentran en fases agudas.

Debido a la relevancia de la cognición social, se han desarrollado programas de rehabilitación que se centran en la recuperación de estos aspectos cognitivos. Dichos programas se pueden clasificar principalmente en dos tipos:

- a) El primero incluye intervenciones focalizadas en mejorar el reconocimiento de emociones, como el *Training of Affect Recognition*, en el que los pacientes aprenden a identificar y diferenciar los principales signos faciales de las seis emociones básicas.
- b) Programas de intervención más complejos, como el *Emotion Management Training*, cuyos objetivos principales son la recuperación de los déficits en la percepción de emociones, el ajuste social del paciente o enseñar estrategias de afrontamiento eficaces.

También existen programas integrales en los que se mejora tanto el funcionamiento cognitivo como el social. [Ro10].

### **2.3.2.1. Empleo de avatares virtuales en terapias**

Con la realidad virtual, han emergido muchas aplicaciones médicas, en particular, debido a la capacidad de proporcionar un entorno realista e interactivo. Dichas aplicaciones existen para la evaluación cognitiva de los pacientes [Ku03], para el tratamiento de pacientes con ‘ictus’ donde se simulen actividades de la vida diaria [Le88], para la evaluación e investigación de las características humanas en entornos tridimensionales tales como memoria espacial [Om02], o también para el tratamiento de varios tipos de trastornos mentales, tales como fobias [Ro95] [Wi02] [Ja02] o desórdenes de déficit de atención [Ri00] [Ch02] así como su empleo en investigaciones del comportamiento interpersonal de pacientes con enfermedades mentales como la esquizofrenia o el autismo.

Los programas de rehabilitación cognitivos descritos en el punto anterior requieren personal altamente cualificado y consumen mucho tiempo de los pacientes [Dy09]. La realidad virtual (VR) podría proporcionar una terapia más eficiente para tratar las deficiencias sociales y alteraciones emocionales, ya que tiene el potencial de crear un entorno realista, tridimensional, y más social al poder incluir representaciones de personas y objetos con los que el usuario pueda interactuar de forma independiente [Dy09]. Avances en las técnicas empleadas en la generación de gráficos y animación han hecho posible que se puedan emplear imágenes para incrementar la apariencia humana en los entornos virtuales. [Je06]. Otra ventaja del empleo de VR es que es un recurso que evalúa de forma objetiva los síntomas. Hay informes que muestran que la capacidad de percepción social puede ser fácilmente evaluada durante la interacción con personajes virtuales [Ki07]. Por otra parte, la distorsión de la percepción de la realidad y las funciones de memoria se puede medir dentro de un entorno virtual [So08] [We08].

De cara al empleo de rostros virtuales en tratamientos, su uso posibilitaría crear de forma controlada expresiones faciales con una graduación de dificultad [Dy09]. Este enfoque abriría una manera sistemática en los programas de rehabilitación, en el que los pacientes fueran entrenados inicialmente en el reconocimiento de las emociones con características más marcadas, y una vez mejorado el reconocimiento de éstas, incorporar en las expresiones virtuales rasgos más sutiles y raros. La intensidad de los estímulos influye en las tasas de reconocimiento de los pacientes. [KO03] comparó el reconocimiento de las emociones en pacientes con esquizofrenia aplicando diferentes grados de intensidad, variando desde leve a extremo y demostró que los pacientes no fueron capaces de reconocer expresiones con mayor intensidad. Se ha demostrado que hay expresiones que son más difíciles de reconocer por los pacientes empleando rostros más virtuales que naturales, como las emociones negativas, o que tienden a sobre-interpretar y atribuir valor emocional a estímulos neutrales (enfado, tristeza y felicidad) [Ka03] [Su06]. Tampoco se suele reconocer bien la expresión de disgusto en las caras virtuales [Mo06] [Sp01]. Este efecto puede ser explicado por la dificultad para generar una unidad de acción específica (arrugar la nariz de arrugas) propia de la expresión de disgusto.

Pese a todo esto, diversos estudios han demostrado que los pacientes con enfermedades mentales muestran déficits similares en el reconocimiento de caras virtuales y naturales [Dy09]. El hecho de que los pacientes presenten déficits en el reconocimiento de las emociones virtuales indica que previamente tienen carencias en el reconocimiento de las características de las emociones básicas, ya que las expresiones virtuales sólo contienen características importantes y frecuentes; mientras que las caras naturales, contienen características poco frecuentes y de menor importancia que varían individualmente [Ko04].

Además, experimentos realizados en los que los pacientes interactuaban con el avatar, éstos se comportaban como si el avatar estuviera en frente suyo e interactuara con ellos. Diversos estudios han demostrado que las personas dicen sentir un cierto nivel de presencia en casi todos los entornos virtuales e incluso responder socialmente a estos [No03]. Estos resultados apoyan la

hipótesis de que las personas responden socialmente tanto a entornos reales como virtuales, [Je06]. Es evidente también que el comportamiento de los pacientes de cara a avatares virtuales es variable en función de sus síntomas y esto también ocurre al interactuar con personas reales. De todo esto se concluye que el empleo de técnicas de realidad virtual en pacientes con enfermedades mentales es posible y justifica el empleo de técnicas de entornos virtuales para el tratamiento o evaluación de las características de comportamiento así como las características cognitivas de los mismos. [Je06]

Los personajes virtuales representan una valiosa herramienta que amplía las opciones de investigación en este campo ya que se pueden aprovechar las ventajas que estas ofrecen, como la facilidad de control, la animación y el cambio de los parámetros. Además, los avatares empleados también pueden ser implementados en juegos o videos [Dy09]. Por consiguiente, pueden ser fácilmente integrados en escenas realistas. Estos suministrarían un escenario óptimo para el ejercicio de reconocimiento de emociones en un contexto social y podría proporcionar una herramienta para la investigación de los procesos neuronales que subyacen a deficiencias en comportamiento social de una manera más realista.



# Capítulo 3

## *Fundamentos del modelado y animación facial de un avatar realista*

### **3.1. Introducción**

En el presente capítulo se detallan las técnicas de modelado y animación facial 3D más comunes para la generación de avatares, siguiendo el esquema general de construcciones de animaciones 3D. En él, se hace un breve estudio de los diferentes software de modelado y simulación 2D y 3D más empleados, se describen los aspectos y las cuestiones a tener en cuenta para la elaboración de un avatar con alto grado de realismo, se detallan las técnicas empleadas más comunes y las características y las consideraciones para la correcta elección de éstas, buscando la convergencia entre características propias de la animación 3D de personajes y los fundamentos de la anatomía y fisiología humanas.

El objetivo final de la creación de un avatar o agente animado es conseguir un organismo lo más realista posible que opere en tiempo real. Debido a la complejidad que reside en la anatomía facial humana, no existe un sistema que genere las expresiones faciales de forma minuciosa en tiempo real, además el proceso de generación de expresiones faciales supone una extensa intervención humana o un pre procesado elevado. No obstante, se puede generar un avatar con alto grado de realismo capaz de generar expresiones verosímiles en un periodo de tiempo tolerable.

## 3.2. Herramientas de modelado gráfico

Un programa de gráficos 3D es el conjunto de aplicaciones que permiten la creación y manipulación de gráficos 3D por ordenador. Estas aplicaciones son usadas tanto para la creación de imágenes como en la animación por computadora. Estas herramientas informáticas contienen las herramientas de modelado, animación y renderizado en 3D necesarias para desarrollar gráficos en movimiento, juegos y animaciones de objetos 3D.

### 3.2.1. Principales programas de modelado gráfico

Se pueden encontrar diversos programas de modelado gráfico. A continuación, se expondrá un listado de los principales programas en el mercado. Dentro de éstos, se pueden encontrar tanto software gratuitos como de pago. Entre los siguientes programas, los más destacados son Blender, 3D Studio Max y Maya.

- *3D Studio Max*: Este programa desarrollado por Autodesk, es uno de los programas de animación 3D más utilizado debido a su arquitectura abierta, su baja curva de aprendizaje al tener una interfaz simple e intuitiva y sus potentes herramientas. Se emplea especialmente para el desarrollo de videojuegos, publicidad, televisión y cine, arquitectura o ingeniería. Su precio es a partir de 3900€.

- *Blender*: Este software de libre distribución contiene un amplio rango de herramientas de creación, renderizado y animación de contenidos 3D. Es multiplataforma y cuenta con una arquitectura 3D de alta calidad que permite un rápido y eficiente desarrollo y un pequeño tamaño de ejecutable para una fácil distribución. Se encuentra en continuo desarrollo, y cuenta con una comunidad global que resulta muy útil para los usuarios. Su principal desventaja es que presenta una interfaz poco intuitiva, pero sus potentes herramientas han hecho que se emplee incluso en la industria cinematográfica (*Spiderman 2*) o cortos de animación.

- *Cinema 4D*: Este software permite la creación de gráficos avanzados 3D, texturización y animación. Destaca por su alta velocidad de renderización, una interfaz altamente personalizable y flexible la cual es modulable, se pueden añadir módulos especializados independientes en función de las necesidades del proyecto y una curva de aprendizaje muy vertical. Se puede obtener a partir de 10.000 €.

- *K-3D*: Es un programa libre de modelado de 3D y animación. Sus principales características es que es compatible con los principales sistemas operativos, es sencillo al tener una interfaz bastante intuitiva y potente al poder crear animaciones en tiempo real orientado a objetos con soporte para plugins.

- *Lightwave 3D*: Programa de modelado gráfico, renderización y animación, se divide en dos subprogramas: Modelos y Layout. En Modeler, se modela un objeto siguiendo un modelo e capas donde cada una es una única malla. En Layout, se realiza la configuración del esqueleto de una malla o rigging, y se generan las animaciones y el renderizado. Ambos están ligados de forma que si se desea hacer una modificación en uno u otro, los cambios aparecerán automáticamente en el otro. Ofrece interfaces basadas en lenguaje C además de su lenguaje de programación propio. Su precio parte de 1400€.

- *Maya*: Surgió por la evolución de Power Animator y de la fusión de Alias y Wavefront, que finalmente fue absorbida por Autodesk. Maya ha sido por muchos el software preferido por grandes empresas de efectos visuales debido a su gran flexibilidad, lo abierto que es a software de

terceros ya que tiene soporte para otros lenguajes como C++, Python, MEL, y que su arquitectura interna se puede transformar debido a sus opciones altamente personalizables. Éste es el único software de 3D acreditado con un Oscar gracias al enorme impacto que ha tenido en la industria cinematográfica. Se ha empleado en películas como “Lord of the Rings”, “La momia”, “Ice Age”, “King Kong”, “Narnia”, “La guerra de los mundos” etc. Su precio es a partir de 3900€.

- *Strata*: Este programa de modelado, animación y renderizado de gráficos 3D posee una interfaz poco amigable, y aunque sus trabajos se pueden integrar en Adobe Photoshop o emplearse en plugins de diseño 3D, presenta limitaciones en la exportación de ficheros y en lo que respecta a los formatos soportados.

- A pesar de que se citen solo algunos programas de animación y modelado existentes hoy en día, dentro de sus potentes herramientas incluyen la creación de personajes humanos con alta capacidad de realismo tanto en aspecto como en dinamismo. Sin embargo, existen también herramientas especializadas en la modelación y animación de estos. A continuación, se mencionarán algunos de ellos.

- *MotionBuilder*: Este software especializado en la animación de modelos y personajes desarrollado por Autodesk, se utiliza para la producción virtual, captura de movimiento o animación mediante fotogramas clave. Su principal ventaja es que genera automáticamente los bones del modelo, y el usuario se tiene que simplemente configurar los movimientos. Esto puede suponer una limitación para los usuarios que deseen crear otras funcionalidades o animaciones fuera de las que ofrece el programa. También, MotionBuilder permite al usuario crear herramientas, botones y modificaciones mediante scripts programados en Python, y es compatible con los diversos programas de modelado 3D que ofrece la casa (3DMax o Maya).

- *FaceGen*: Este programa permite crear retratos tridimensionales de gran realismo o caricaturas en base a los parámetros que el usuario establezca dentro de sus controles pudiendo crear rostros de cualquier raza o grupo de edad pudiendo personalizar los rasgos e incluir detalles pero dentro de los parámetros del programa. Puede exportar a formato propio (FG) o a formatos de otros programas 3D como 3D Studio, Maya o Lightwave.

- *Lifemi*: Este software posee potentes herramientas para la creación de rostros y animaciones, e incluso la sincronización de los labios al habla. Su principal desventaja es que no permite el modelado de rostros independientes al mismo. Presenta posibilidades de integración con 3D Studio Max o con Maya.

### 3.2.2. *Blender*

Dentro de todos los programas de modelado geométrico y animación existentes hoy en día, el software elegido para este proyecto ha sido la herramienta *Blender*, debido a las múltiples cualidades que ofrece y cuyo potencial lo eleva a la categoría del resto de herramientas empleadas por las industrias de animación.

Blender es un programa que integra una serie de herramientas para la creación de un amplio rango de contenidos 3D. Puede ser usado para visualizaciones 3D, tanto de imágenes estáticas como de vídeos de alta calidad, mientras que la incorporación de un motor 3D en tiempo real permite la creación de contenido interactivo que puede ser reproducido independientemente.

Originalmente desarrollado por la compañía 'Not a Number' (NaN), Blender es ahora desarrollado como 'Software Libre', con el código fuente disponible bajo la licencia GNU GPL, es decir, se puede usar, copiar, modificar y distribuir sin ningún límite.

Algunas de las características más destacables de este software es que tiene un paquete de creación totalmente integrado, ofreciendo un amplio rango de herramientas para la creación de contenido 3D, desde el modelado, hasta creación de juegos, incluyendo mapeado UV, texturizado, rigging, weighting, animación, simulación de partículas y otros, scripting, renderizado, composición y post-producción. Además, es multiplataforma con una interfaz unificada y compatible con todas las versiones de *Windows*, *Mac OS X*, *Linux*, *OSX*, *FreeBSD*, *Solaris*, *Irix* y *Sun*, y otros sistemas operativos. También cuenta con una arquitectura 3D de alta calidad, permitiendo un rápido y eficiente desarrollo y un tamaño pequeño de ejecutable para una fácil distribución.

El que Blender sea disponible de forma gratuita ayudó mucho en su adopción. Hoy en día cuenta con una gran comunidad, estable y activa, donde comparten y mejoran sus proyectos, donde se puede encontrar tutoriales o ayuda para obtener una mejor comprensión de las funcionalidades de Blender, además de foros de ayuda y discusión, fuentes de documentación o modelados y animaciones. Las principales comunidades son la comunidad de desarrollo, centrada alrededor de la web de la *Fundación Blender*, <http://www.blender.org>, o la comunidad de usuarios centrada en un sitio independiente, <http://www.blenderartists.com>, donde comparten sus proyectos, realizan críticas para mejorarlos o se ayudan con el desarrollo de éstos. Pero estas páginas web no son los únicos recursos de Blender. La comunidad mundial ha creado una gran cantidad de sitios diferentes, en diferentes idiomas y especializadas en diferentes áreas donde actualizan constantemente los recursos para Blender. También se puede encontrar mucha información y ayuda en los canales de *Youtube* y videos tutoriales de los usuarios, o canales chat, <irc.freenode.net>, para discusiones generales sobre *Blender*, [#blenderchat](#), preguntar sobre el uso de *Blender*, [#blenderqa](#) o discusiones sobre temas relacionados con la creación de juegos con el motor de *Blender*, [#gameblender](#) además de chats para los desarrolladores, [#blendercoders](#), [#blenderpython](#), [#blenderwiki](#).

Aún así, la principal desventaja de esta herramienta es su interfaz, que es poco intuitiva. Por eso, antes de comenzar a modelar, interesa que los nuevos usuarios se familiaricen con ella y comprendan la manera en que funciona, aprendiendo a manejar también los controles y atajos del teclado y ratón. Una vez que el usuario se familiariza con la interfaz, encontrará que puede realizar su trabajo de forma rápida y productiva. Por otro lado, la interfaz de Blender ha cambiado sustanciosamente a medida que han aparecido nuevas versiones del programa, en ocasiones dificultando la comprensión de éste al emplear en la mayoría de los tutoriales entrados en materia de la generación de avatares virtuales realistas versiones anteriores donde los comandos o los menús del programa han cambiado.

### **3.3. Modelado y animación facial**

El diseño y animación de objetos virtuales ha sufrido una constante evolución desde los años 70 hasta nuestros días aumentando en sofisticación y realismo. Los avances que ha sufrido la técnica permite generar modelos de alta calidad y han supuesto un profundo impacto en el desarrollo de animaciones empleados en el cine, videojuegos, o en líneas de investigación de medicina o tecnología. Sin embargo, a pesar de los avances realizados, las técnicas empleadas en el modelado y animación de objetos virtuales requieren muchos recursos e implican una gran labor manual.

El modelado de objetos 3D consiste en diseñar y crear la forma de los objetos que se emplean en la escena. Al modelar un objeto, hay que tener en cuenta la forma física y la apariencia visual que éste tiene. La forma física de un objeto comprende puntos unidos por líneas para formar superficies o polígonos y la forma que se atribuye a estos polígonos representa la forma del objeto a crear. Pero al modelar el objeto es tan importante definir la estructura como aportar las características necesarias para otorgarle una apariencia visual realista, es decir, el texturizado. El texturizado permite recrear el color y las cualidades del objeto a modelar, dotarle de detalles a determinadas zonas o simular su material. El proceso de modelado puede incluir algunas acciones relacionadas con la preparación del modelo 3D para su posterior animación. A los objetos se les puede asignar un esqueleto o una estructura central que puede influir en la forma y movimiento del mismo. Esto ayuda al proceso de animación, en el cual el movimiento del esqueleto automáticamente afectará a las porciones correspondientes del modelo.

Los rasgos faciales de un asistente virtual vienen dados por los objetos geométricos que lo conforman. Parece lógico que la forma que adopten estos objetos geométricos esté basada en la estructura del rostro humano. El principal objeto que define la apariencia facial del avatar es la estructura geométrica de la cabeza, por lo que la forma que se atribuya a la superficie poligonal debe ser afín a esta. Además para que el modelo tenga un aspecto realista, este debe estar dotado de una textura que recree la apariencia humana: el color de la piel, rasgos que caractericen al personaje como pecas, arrugas, vello o imperfecciones y añadir sombras en aquellas zonas que cambien o donde se producen pliegues al gesticular.

### ***3.4. Características básicas del modelado facial***

Para lograr una animación facial con suficiente grado de realismo y que opere en tiempo real, es necesario alcanzar un equilibrio entre dos componentes fundamentales: el nivel de detalle del rostro generado y las posibilidades computacionales del sistema del que se dispone. Waters [Wa90] propuso un listado de atributos de los que debe disponer el modelo de representación facial:

- *Piel*: se pueden emplear para su representación modelos volumétricos o superficies 3D. En caso de trabajar únicamente con una representación 3D de la superficie del rostro, ésta debe construirse de forma que pueda modificarse para generar expresiones suficientemente convincentes. En la piel se manifiestan los movimientos y contracciones de los músculos faciales. En un modelo facial no es necesario recrear la simulación exacta de los músculos, relieves, pero sí que el modelo conste de parámetros dinámicos que emule estas características primarias.
- *Pelo*: existen numerosos estudios por parte de las compañías de animación para construir un modelo de cabello que responda dinámicamente a la variación de fuerzas externas como corrientes de aire, desplazamientos bruscos de la cabeza, etc.
- *Ojos*: son una parte importante a la hora de crear expresiones convincentes. La principal característica que deben presentar es un alto grado de dinamismo y de respuesta a las variaciones de luminosidad externa.
- *Orejas/nariz*: demandan una superficie detallada debido a las cavidades y múltiples relieves que presentan.
- *Labios*: su principal característica debe ser la flexibilidad, ya que deben amoldarse a los movimientos de la boca.

- *Dientes*: dado que los dientes se mueven de forma rígida siguiendo a la boca y la mandíbula, su modelado no resulta complejo, pero se deben tener en cuenta su el modelo facial incorpora sistema de voz.
- *Lengua*: importante sobre todo si el modelo facial incorpora un sistema de voz.
- *Detalles asimétricos*: por simplicidad, la mayoría del modelado facial se realiza empleando herramientas que permiten el duplicado simétrico de las características faciales. Sin embargo, ningún rostro real es perfectamente simétrico por lo que la incorporación de ciertas características disonantes en el mismo resulta conveniente para lograr un mayor grado de realismo.

### **3.5. Métodos de representación del modelo 3D facial**

Fundamentalmente, los datos que constituyen el modelo 3D pueden obtenerse modelando de forma manual la superficie a representar o por medio de la adquisición de datos a partir de caras reales por medio de escáneres digitales 3D, cámaras, sistemas de obtención de modelos volumétricos como el MRI y otros. Con independencia del método de obtención elegido, los métodos más comunes de representación de modelos 3D son las superficies poligonales, las superficies paramétricas y los modelos volumétricos.

También se puede generar un modelo 3D mediante técnicas de captura de movimiento. Existe una variedad de técnicas de captura de movimiento, como por ejemplo, registrarlos mediante el empleo de una cámara de video estándar y marcadores. Además se pueden emplear técnicas de visión artificial para realizar un seguimiento de las características faciales en la secuencia. Las características incluyen las cejas, los ojos, el centro de la pupila y el contorno interior y exterior de los labios, o marcas auxiliares en mejillas y mentón. Un modelo de seguimiento facial puede estar basado en una base de datos. Para obtener la base de datos, inicialmente se emplean varias imágenes de la secuencia de vídeo y se marca el contorno facial y se distribuyen los marcadores de forma uniforme a lo largo de cada contorno. A continuación, estos se ajustan para eliminar datos poco relevantes. La base de datos es un análisis de componentes principales en un vector que contiene información sobre la intensidad de los píxeles y el etiquetado de las coordenadas de cada uno. Para realizar el seguimiento del modelo, las imágenes se deforman para ajustar el movimiento en relación con el seguimiento realizado por la cámara. Cada frame es analizado en función de la intensidad del píxel y se compara con el resto de muestras acumuladas en el vector, obteniendo así las características faciales del modelo.

#### **3.5.1. Superficies poligonales**

La superficie del modelo se forma por medio de polígonos llamados también “caras”. Dichas caras están formadas por vértices interconectados entre sí por los bordes de los polígonos (*Figura 3.1*). Generalmente los polígonos incluyen información sobre su vector normal, la dirección a la que apunta el polígono (qué parte es “externa”), y la normal de sus vértices, que es información adicional sobre la curvatura de la superficie [He06].

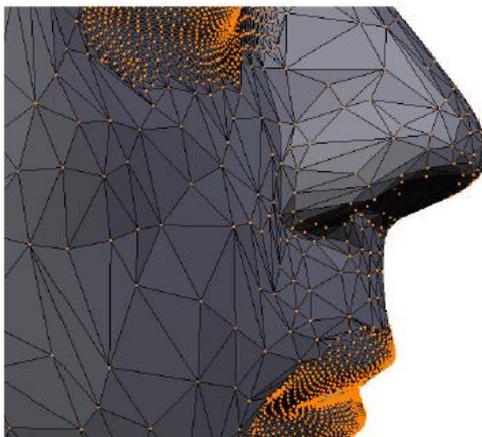


Figura 3.1 – Fragmento de un modelo generado mediante superficies poligonales

Uno de los puntos clave al construir u obtener un objeto 3D formado por polígonos es lograr una superficie que sea fácilmente deformable entre varias poses. La superficie debe poder contraerse /expandirse de forma similar a la piel del rostro humano por lo que es conveniente que la disposición de los puntos y las aristas de los polígonos se realice teniendo en cuenta las deformaciones que sufrirá la superficie poligonal durante la animación.

Los modelos construidos de forma manual generalmente buscan que las aristas de los distintos polígonos del modelo se sitúen formando líneas perpendiculares a la dirección de contracción de los distintos músculos faciales. En el caso de obtener modelos por medio de escáneres digitales 3D, los datos obtenidos no cumplen con éstas condiciones por lo que resulta más complejo conseguir animaciones que simulen de forma correcta la deformación de la piel.

El nivel de detalle en los modelos poligonales viene determinado por el número de polígonos empleados en la representación ya que, obviamente, a mayor número de polígonos más fiel será la representación de las superficies con relieves curvos y pronunciados, como en la nariz y la región de la oreja. En aplicaciones de tiempo real se debe alcanzar un compromiso entre el nivel de detalle deseado y la velocidad de la animación final generada. A mayor número de polígonos mayor coste computacional empleado en su animación, y mayor espacio es necesario para almacenar el modelo.

La ventaja del modelo poligonal estriba en su fácil representación e implementación de deformaciones, comparado con las superficies paramétricas o los modelos volumétricos. Como inconveniente presenta el hecho de que alcanzar un alto grado de detalle supone un elevado coste computacional.

### **3.5.2. Superficies paramétricas**

Las superficies paramétricas se construyen por medio de la combinación de curvas paramétricas, en las que es posible establecer su forma por medio de una serie de parámetros de control. Una de las curvas paramétricas comúnmente empleada en animación es del tipo NURBS (Non – Uniform Rational B-Spline) [Pi95]. Las superficies formadas por curvas NURBS son flexibles y pueden combinarse para formar superficies mayores. Al proceso de unión de las distintas superficies paramétricas se le denomina *trimming*. Se puede ver un ejemplo en la *Figura 3.2*.

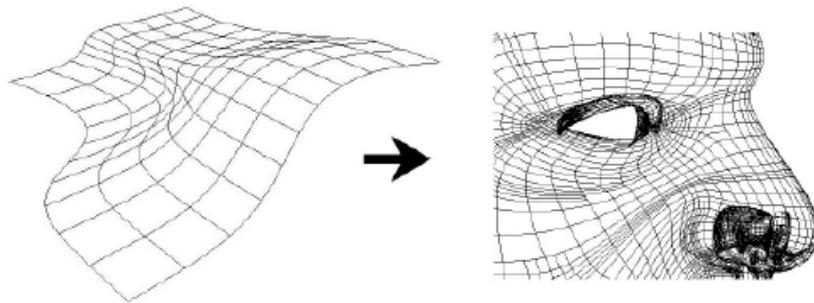


Figura 3.2 – Fragmento de un modelo generado mediante curvas NURBS

La definición de este tipo de curvas y planos, y los parámetros que los forman se muestra a continuación:

Curva NURBS	Plano NURBS
$\left\{ \begin{array}{l} p = \text{orden} \\ B = \text{función básica B-spline} \\ P_i = \text{puntos de control } P_0, \dots, P_n \\ w = \text{peso} \end{array} \right.$	$\left\{ \begin{array}{l} p, q = \text{orden} \\ B = \text{función básica B-spline} \\ P_i = \text{puntos de control } P_0, \dots, P_n \\ w = \text{peso} \end{array} \right.$
$C(t) = \frac{\sum_{i=0}^n B_{i,p}(t)w_i P_i}{\sum_{i=0}^n B_{i,p}(t)w_i}$	$S(u, v) = \frac{\sum_{i=0}^m \sum_{j=0}^n B_{i,p}(u)B_{j,q}(v)w_{i,j} P_{i,j}}{\sum_{i=0}^m \sum_{j=0}^n B_{i,p}(u)B_{j,q}(v)w_{i,j}}$

La principal ventaja de emplear este tipo de curvas reside en que permiten una mayor precisión y flexibilidad que las representaciones poligonales puesto que están basadas en métodos paramétricos. Además pueden ser renderizadas a una velocidad suficiente para producir animaciones en aplicaciones interactivas. Sin embargo, en ocasiones necesitan gran cantidad de uso de memoria. Otro de los inconvenientes de la representación paramétrica es que se emplean métodos poco intuitivos y, por tanto, de difícil construcción. Por último, destacar que las curvas paramétricas no son capaces de crear arrugas de forma realista, como las que se forman en la zona de las alas de la nariz o en las comisuras de los ojos y labios.

### 3.5.3. Subdivisión de superficies

Para conseguir un efecto de deformación adecuado en la superficie poligonal, se puede emplear un método de deformación basado en regiones (grupos de vértices), construidas de tal forma que simulen las inserciones musculares reales en la piel. La subdivisión de superficies permite añadir superficies suaves en los modelos poligonales manteniendo la posibilidad de crear arrugas de forma realista. Dicha subdivisión se basa en aproximaciones o interpolaciones. El método empleado varía en función del tipo de polígonos que formen la superficie: triángulos, cuadrados, etc.

En la siguiente imagen (Figura 3.3) se muestra el proceso de subdivisión empleando el algoritmo Catmull – Clark [Sy05] con tres niveles de subdivisión. Mediante éste algoritmo, se dividen los polígonos de la superficie triangular uniendo las mediatrices de los lados de cada

triángulo. La subdivisión de redes poligonales se emplea comúnmente para producir secuencias de vídeo a partir de modelos sencillos que requieran un bajo coste computacional durante su manipulación. Posteriormente, antes del renderizado, se procede a aplicar la subdivisión para mejorar la calidad del resultado final.

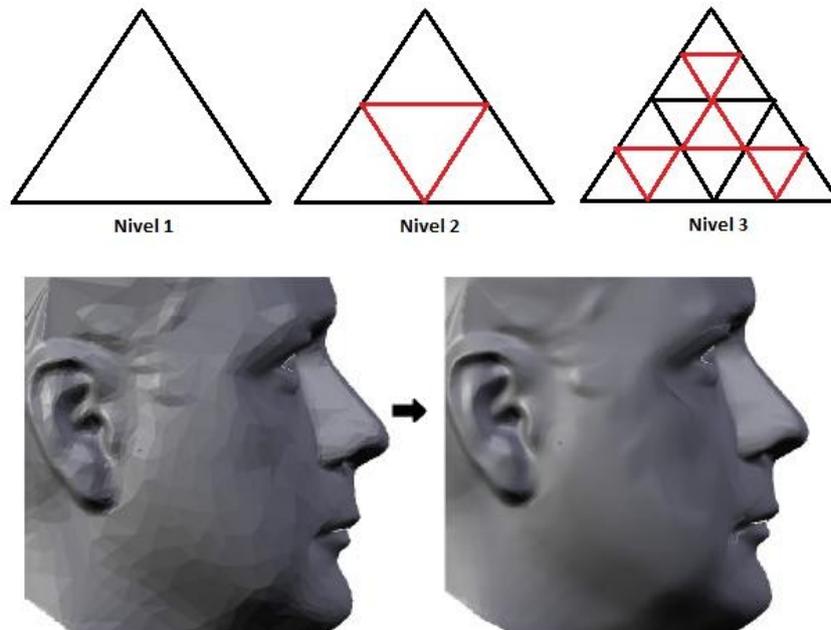


Figura 3.3 – Resultado de subdivisión de superficies de 3 niveles mediante el algoritmo Catmull-Clark

### 3.5.4. Modelos volumétricos

La construcción de modelos volumétricos se emplea en ciertas áreas médicas, en las que no sólo importa el aspecto exterior visual del modelo, sino también su parte interna: músculos, órganos etc. Estos modelos están formados por voxels, que son elementos volumétricos equivalentes a los pixels. Cada voxel está compuesto por cuatro vértices o más, y un atributo de color para el volumen que representa (Figura 3.4). Para su construcción, los datos se obtienen por medio de técnicas como MRI o Magnetic resonance imaging o CT, Computer Tomography. [Ma11]

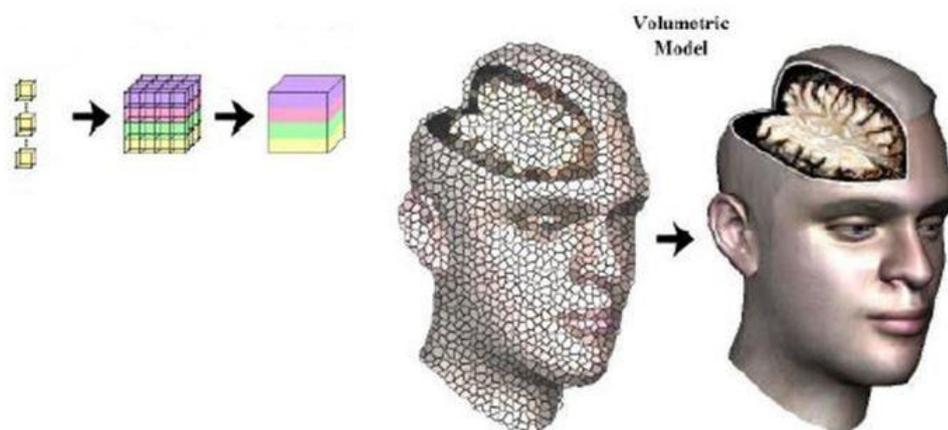


Figura 3.4 - Modelado mediante superficies volumétricas

Los modelos volumétricos son la forma más fiel de representar la anatomía humana, lo que los hace la única opción para aplicaciones médicas en las que la representación de la superficie externa es insuficiente. Sin embargo, estos modelos suponen una alta carga computacional lo que los hace poco adecuados para animación en tiempo real, aunque su uso se generaliza a medida que aumentan el potencial de los sistemas computacionales empleados. Uno de los primeros ejemplos de animación con modelos volumétricos basándose en métodos de elementos finitos puede encontrarse en [Zh98].

### 3.6. Técnicas de modelado facial

La tendencia actual en lo que a modelado facial se refiere es producir modelos con el mayor grado de realismo posible. Además, en caso de que se pretenda animar el modelo generado, es importante tener en cuenta las técnicas de animación que se van a emplear durante la construcción del modelo. De las formas de representación del modelo facial descritas en el apartado anterior, la más utilizada es la representación mediante superficies poligonales, dado que es la forma más básica y dichas superficies son soportadas de forma nativa por la mayoría del hardware y software de procesamiento gráfico.

La principal dificultad del modelado facial es capturar de forma específica la geometría 3D de un rostro concreto. Con el avance de los sistemas de captura de datos 3D, se ha generalizado el empleo de escáneres láser 3D o digitalizadores para la generación de rostros tridimensionales. Por medio del escaneo láser es posible adquirir información tridimensional y colorimétrica precisa del rostro escaneado. Sin embargo, los modelos obtenidos con estos procesos normalmente no son óptimos para la animación, dado que carecen de información de la estructura facial, el cálculo de ruido produce errores en el objeto y los puntos capturados están pobremente distribuidos y es necesario realizar un importante trabajo manual que permita adecuar el modelo para su animación. Además, muchos métodos de obtención de puntos producen modelos incompletos carecientes de pelo, ojos, orejas, etc.

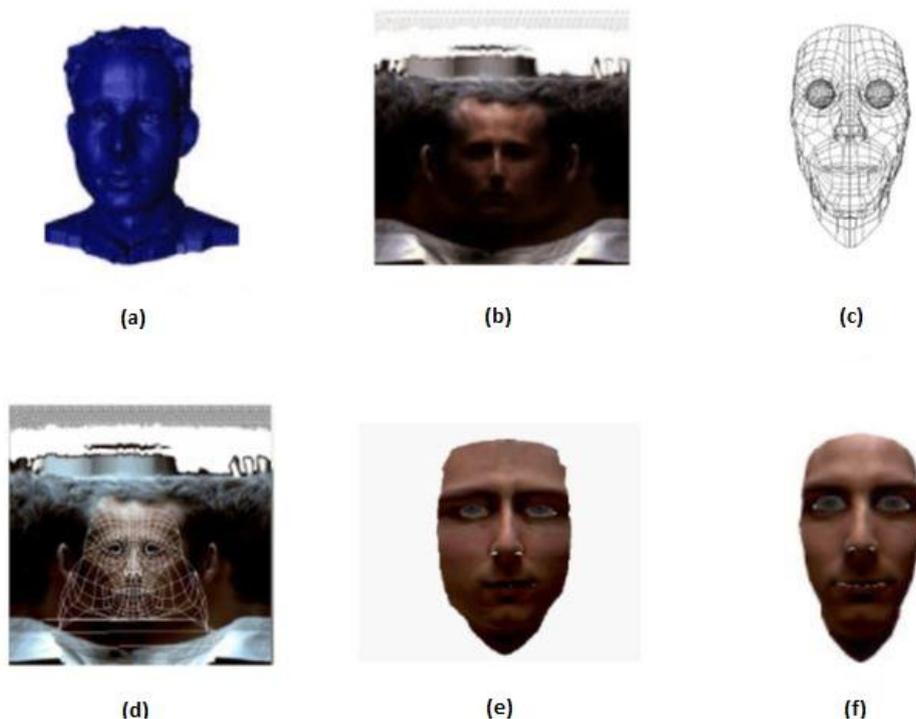


Figura 3.5 – Proceso de escaneo láser 3D.

En la *Figura 3.5* podemos observar el proceso de un escaneado láser 3D de un rostro real y apreciar su acabado:

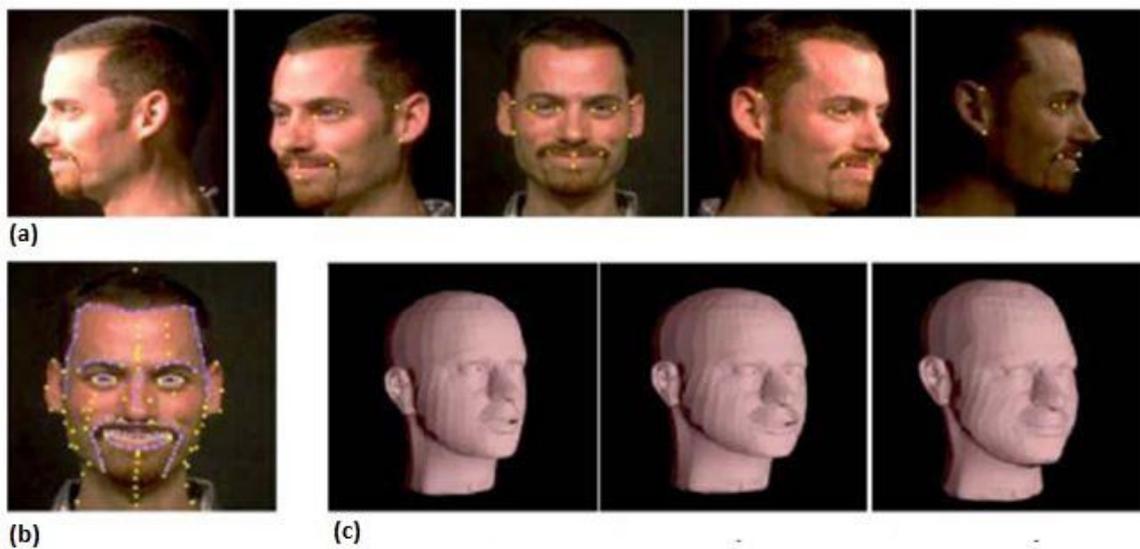
- a) Datos obtenidos mediante escaneo, contienen información detallada.
- b) Datos escaneados en reflectancia, contiene información del color.
- c) Red de puntos genérica, contiene información para la animación.
- d) Red de puntos proyectada en coordenadas cilíndricas para el ajuste.
- e) Red de puntos ajustada.
- f) Red de puntos antes del ajuste, mostrada en comparación con (e).

Una aproximación a un rostro específico supone un arduo pre-procesado para preparar un prototipo con toda la información sobre estructura y animación necesarias. Sin embargo, durante los últimos años se han propuesto varias técnicas para la producción de modelos faciales de alta calidad con las que se reduzca el trabajo manual necesario para su obtención y ajuste. Dichas técnicas pueden dividirse en dos categorías:

- Obtención del modelo de un sujeto específico a partir de la deformación de un modelo genérico.
- Creación de modelos con las características deseadas a partir de la combinación de una colección de modelos que sirven como base.

### ***3.6.1. Modelado mediante la deformación de un modelo genérico***

Esta técnica de modelado genera modelos faciales de un individuo específico a partir de la deformación de un modelo genérico. Para ello, a partir de una imagen o una secuencia de vídeo, se extraen las posiciones de ciertas características faciales preseleccionadas del rostro tales como los bordes de los ojos o las posiciones de la nariz y boca. A continuación se genera un nuevo modelo a partir de la deformación de un modelo genérico buscando el alineamiento de las características faciales extraídas con las correspondientes del modelo genérico. Se puede apreciar el proceso en la *Figura 3.6*.



*Figura 3.6 – Modelado mediante la deformación de un modelo genérico. (a)(b) Anotación de características faciales importantes. (c) Deformación del modelo genérico alineándola posición de las características faciales anteriores.*

Existen diferentes técnicas de modelado en función de las características faciales empleadas, la forma en que éstas se obtienen y los distintos métodos para deformar el modelo tridimensional genérico. En [Pi98] se presenta un método de generación de expresiones faciales a partir de múltiples vistas de un mismo individuo. El modelo tridimensional genérico se deforma a partir de puntos característicos de las imágenes (esquinas de los ojos, la punta de la nariz...) por medio de interpolación de datos dispersos [Am02]. En [Le04b], la obtención de características se realiza a partir de una imagen lateral y otra frontal y se simplifica el proceso de correspondencia: primero por regiones en torno a cada una de los puntos característicos identificados en las imágenes para luego realizarlo de forma más fina para cada uno de los puntos considerados. El principal inconveniente de estos métodos de modelado es que requieren una gran cantidad de trabajo manual, por ejemplo para anotar todo el conjunto de puntos característicos en las imágenes de partida.

En Zhang et al. [Zh04] se presenta un algoritmo que permite adaptar un modelo funcional facial genérico provisto de un conjunto de pseudo-músculos que permiten posteriormente animar el rostro. En este método, se especifican un conjunto de marcas antropométricas sobre las imágenes 2D del modelo genérico y del modelo escaneado que se pretende adaptar y se proyectan a coordenadas 3D, adaptándolas a continuación primero a nivel global (posición, orientación) y luego a nivel local (vértices). Asimismo, la estructura muscular del modelo genérico se adapta de forma automática.

En [Kä03] se emplea un modelo genérico formado por múltiples capas que simulan la estructura anatómica facial incluyendo la piel, los músculos, el cráneo y una capa de tensores junto con otros componentes tales como los ojos, los dientes, la lengua, etc. y se emplean marcas antropométricas [Fa94] que permiten adaptar el modelo genérico a nuevos modelos 3D escaneados. Además, empleando medidas antropométricas para simular el crecimiento, el sistema es capaz de generar rostros animados de un mismo sujeto a diferentes edades.

En [Yi04] se crea una representación topográfica para transferir las características faciales de una imagen frontal del rostro a un modelo genérico. El modelo topográfico trata la imagen como un terreno, marcando cada uno de los pixels como marcas topográficas, incluyendo picos, crestas, valles, etc. Tras el análisis topográfico, el proceso de ajuste del modelo genérico se realiza mediante ecuaciones diferenciales de segundo orden que definen fuerzas internas y externas. Las fuerzas externas rigen la deformación del modelo a la forma objetivo, mientras que las fuerzas internas mantienen la forma del modelo durante la deformación.

En [Zh04] se emplean múltiples cámaras de vídeo para capturar los datos característicos mediante procesamiento estéreo que captura una secuencia de mapas de profundidad. A continuación se emplea un método de seguimiento y ajuste para adaptar el modelo genérico a los mapas de profundidad del primer frame mediante un método de optimización de Gauss-Newton y realizar un seguimiento a lo largo de la secuencia completa. La principal desventaja de este algoritmo radica en que requiere un elevado coste computacional y el empleo de equipamiento específico para capturar la geometría y el movimiento facial.

Las técnicas basadas en la deformación de un modelo genérico precisan de un único modelo de partida, por lo que son más adecuadas para aplicaciones en las que no se tenga acceso a grandes colecciones de modelos faciales. Además es posible utilizar estas técnicas en modelos faciales multicapa, modelando así piel, músculos y cráneo de forma simultánea. Sin embargo, requieren un proceso complejo de obtención de características faciales, lo que supone o bien una gran cantidad de trabajo manual, o el empleo de costosas técnicas de visión y equipamiento, como ocurre en [Zh04].

### 3.6.2. *Modelado mediante la combinación de modelos base*

Las técnicas descritas en esta sección se basan en la combinación de modelos faciales existentes para la creación de nuevos modelos. Dadas las características faciales requeridas, por ejemplo a partir de marcas en una imagen, se emplean métodos de optimización para encontrar los coeficientes adecuados para combinar la base de modelos. En [BI04] se emplea una base de modelos deformables para generar el modelo genérico. La técnica de modelos deformables se basa en la transformación de las formas y las texturas de una base de modelos 3D a un espacio de vectores. La forma y la textura de un nuevo modelo facial se representan como combinación lineal de dichos vectores transformados. En [HI03] se obtiene un modelo 3D de un rostro mediante modelos deformables y el seguimiento de un rostro en una secuencia de vídeo. En [VI05] se emplean también modelos de base para crear nuevos modelos faciales. De forma análoga al modelado mediante modelos deformables se realiza un pre-procesamiento de los datos de entrada para establecer correspondencias entre los vértices de los modelos de base. A continuación dichos modelos se organizan en forma de tensores, de manera que codifican las variaciones del modelo en términos de diferentes atributos tales como identidad, expresión, etc. permitiendo variaciones independientes para cada uno de dichos atributos. Empleando los datos de tensores, es posible crear un modelo facial arbitrario con la expresión deseada a partir de la combinación lineal de los modelos de base.

Las técnicas que emplean la combinación de un conjunto de modelos de base permiten generar modelos faciales a partir de una única imagen sin el requerimiento previo de la identificación de características faciales. Además permiten generar modelos 3D que incorporen expresiones faciales a partir de simples imágenes. Sin embargo, es necesario que los modelos faciales empleados como base sean registrados uno a uno para establecer una correspondencia entre vértices. Otro de los inconvenientes que presentan es la dificultad para extenderlos a modelos multicapa, ya que se necesitaría de un conjunto de modelos base multicapa con una correspondencia punto a punto establecida entre cada una de las capas consideradas.

### 3.7. *Técnicas de animación*

En la animación de objetos 3D, se emplea un instrumento fundamental denominado FRAME. Un frame es un fotograma, una imagen o un momento particular de la escena dentro de una sucesión de imágenes que componen una animación. En general, en un segundo hay 24 frames, pero esta cantidad puede variar en función de la animación deseada. La cantidad de frames por segundo también se denominan FPS (Frame per second) [VI02]. Existe un frame particular denominado Keyframe que es el frame desde el cual comienza un momento clave en la animación. Desde él se crean otros frames hasta llegar a un nuevo Keyframe que representa el momento final de la animación y el comienzo de otra. La creación de frames entre dos Keyframes se llama InBetweening y para llenar el espacio vacío entre estos se pueden emplear cualquiera de las técnicas descritas en los siguientes apartados.

Cuando se han generado los frames que componen la animación se pueden juntar y dentro de las diferentes secuencias generadas puede distinguirse:

- **Shot:** es una secuencia de frames rodeada por una sola cámara sin interrupciones y cuya duración es de pocos segundos.
- **Sequence:** es una sucesión de frames que siguen un mismo momento de la acción.
- **Scene:** Es una sucesión de Shots o Sequences, es decir, una acción continuada en un mismo lugar y en un momento preciso.

Si no se dispone de suficientes keyframes, la animación resultante no es fluida, sobre todo en modelos complejos [Vi02]. La línea de acción suele ser incorrecta y los objetos pueden chocar unos con otros.

### 3.7.1. Interpolación

La técnica de interpolación ofrece una idea intuitiva de la animación de objetos 3D. Típicamente, una función de interpolación especifica el movimiento suave entre dos posiciones clave extremas, dentro de un intervalo normalizado [Ma06]. Obviamente, para poder interpolarse entre dos posiciones clave la información contenida en ambas ha de ser de la misma naturaleza. Esta técnica consiste en tomar un vértice en un espacio 3D y transformar su posición a otra localización específica. Mediante una función de interpolación se especifica el movimiento entre una posición de origen y otra final dentro de un intervalo normalizado y se calculan los puntos intermedios por los que los vértices pasan [Ma06].

Se pueden emplear diferentes funciones de interpolación y dependiendo de la que se emplee, se pueden crear animaciones de diferente calidad con movimientos más o menos fluidos. La función de interpolación presenta como parámetros de entrada los valores contenidos en las posiciones clave, un coeficiente temporal (el cual determina la cuantía de variación con respecto a las posiciones inicial y final) y viene definida por una curva, [Vi02]. Los tipos de curvas más usados son las siguientes:

- Linear Spline: Es una interpolación creada por una recta que genera una animación con velocidad constante.
- Cardinian Spline: La animación depende de donde estén situados los keyframes y la curva tiene que pasar por todos ellos, generando una animación brusca.
- B-Spline: Con estas curvas, la animación generada será más suave ya que la curva no tiene que pasar por los keyframes, solo alrededor de estos.
- Bazier Curves: La curva es generada en función de los puntos de control, por lo que la animación tendrá una calidad mayor, aunque es la más complicada de controlar.

Como se puede observar, la interpolación emplea curvas para crear los frames entre dos keyframes, pero existen diferentes métodos de interpolación dependiendo del movimiento que se desee crear.

Existe también el método inverso, donde se efectúa el movimiento, y el ordenador calcula las articulaciones que se deben de mover y los ángulos de estas. También existe el método directo, en el que se elige la articulación y el tamaño del ángulo para mover la articulación. Este método requiere mucho trabajo manual y se emplea cuando los objetos tienen que realizar movimientos iguales y repetitivos.

Otro método muy utilizado es el MotionPath en el que se limita el movimiento del objeto en un recorrido fijo. El MotionPath crea un movimiento primario que se amplía con la creación de movimientos secundarios. Se emplea sobre todo en animaciones sencillas de objetos. [Vi02]

Para animaciones de movimientos secundarios como por ejemplo expresiones faciales, se emplea una técnica llamada Free-Form Shape. Es una técnica de animación punto a punto de la

superficie del objeto, es decir, toma el objeto en una posición y la copia de este en dicha posición donde se mueven los puntos que se desean modificar para generar la animación deseada. Con la interpolación entre estos dos keyframes el objeto pasa de la posición inicial a la posición deseada. Esta técnica requiere que el objeto sea el mismo, pero si se emplean objetos diferentes, se emplea el Morphing. Aunque se empleen diferentes objetos, para ser animados bien, tienen que tener el mismo número de puntos y un orden de correspondencia, es necesario crear una asociación entre los puntos ya que si no, no se pueden interpolar los dos objetos. [Vi02]

Por otro lado, si se desean crear animaciones particulares de la superficie del objeto como las olas del mar o el movimiento producido por el viento, se emplea Wave Function, el cual emplea funciones matemáticas y curvas planas o esféricas con las que se generan este tipo de animaciones.

La interpolación es un método rápido y genera fácilmente animaciones faciales simples. Con la interpolación se puede producir un pequeño grupo de animaciones a partir de pocas posiciones clave, pero su capacidad de generar un abanico mayor de configuraciones faciales está limitada por la información disponible en el modelo empleado. Es difícil producir combinaciones de movimientos independientes de forma eficaz. Además, esta técnica requiere mucha información para ser controlada de manera eficaz por un usuario y requiere de alto control y capacidad técnica y es preciso combinarla con otros procedimientos para lograr una animación más completa.

### **3.7.2. Parametrización**

Las técnicas de parametrización [Kö05] superan algunas de las limitaciones y restricciones de la interpolación. Como la interpolación tiene tan alto volumen de información, es común basarse en controles paramétricos para controlar grupos de vértices. En lugar de almacenar los datos para la deformación en varias poses clave, se emplean un conjunto de parámetros cada uno de ellos correspondientes a una región de vértices. La parametrización ideal permite obtener cualquier posible expresión mediante la combinación de los valores de parámetros independientes. A diferencia de las técnicas de interpolación, la parametrización permite un control de configuraciones específicas. La combinación de parámetros otorga un amplio conjunto de expresiones faciales con relativamente poco coste computacional.

La parametrización consta de dos etapas: la definición del conjunto de parámetros y el desarrollo de modelos en función de los parámetros definidos [Jun04]. El objetivo de la definición de parámetros es determinar cómo va a deformarse el rostro modelado. Se puede implementar de diversas formas: a partir de un análisis de la topología de la superficie facial, por las regiones que causan variaciones en la piel o acciones específicas del rostro. Sin embargo, la parametrización no produce expresiones realistas cuando dos parámetros conflictivos afectan a los mismos vértices. Por esta razón, las parametrizaciones se diseñan para que únicamente afecten a regiones faciales específicas. Otra limitación de la parametrización reside en que la elección de parámetros depende de la topología facial y una parametrización completa y genérica es inviable. Además, requiere un tedioso trabajo manual previo para especificar los valores de los parámetros y es posible que aún así se obtengan resultados poco realistas.

Un buen ejemplo de la parametrización es el estándar para animación en tiempo real MPEG-4, el cual define un conjunto de parámetros denominados “Facial Parameters Points”. Este estándar se complementa con lenguajes multimedia como SMIL o VRML para la ejecución de la animación.

### **3.7.3. Técnicas basadas en la simulación**

Este tipo de técnicas buscan crear expresiones faciales sintéticas a partir de la simulación de la acción muscular sobre un modelo facial. Esta técnica requiere de la definición previa de la funcionalidad y el emplazamiento, más o menos detallado, de las características físicas de las diferentes capas que componen el rostro (piel, tejido tisular, tejido muscular y cráneo).

#### **3.7.3.1. Modelos basados en la física muscular**

Las técnicas de animación basadas en la física muscular superan las limitaciones de correspondencia entre puntos de los modelos de interpolación y parametrización. Los modelos musculares intentan representar las características físicas de los músculos faciales además de las características de la piel y la estructura ósea humana. Es lo que se denomina descripción dinámica del movimiento. Los más tempranos esfuerzos hacia un modelo basado en músculos (muscle-based model) fueron los de Platt & Badler [P181].

La física muscular simula las propiedades visco-elásticas de la piel facial y los músculos para generar expresiones y animaciones. Las tecnologías dominantes en los modelos basados en la física muscular son los modelos masa-muelle (mass spring), representaciones vectoriales y capas de masas de muelles (layered spring mesh), los cuales pueden ser utilizados separadamente o combinados, dependiendo de la intensidad de la simulación.

##### **3.7.3.1.1. Técnicas que emplean modelos masa-muelle**

Los modelos masa-muelle (Mass-spring methods) propagan las fuerzas musculares en una malla elástica que modela la deformación de la piel. Los estudios realizados por Platt y Badler [P181] sobre el modelado de los músculos y la estructura anatómica del rostro humano, determinan que se pueden simular las expresiones faciales de una forma realista simulando las fuerzas que actúan en los músculos aplicadas en una malla elástica. Platt [P185] propone un modelado facial donde los músculos se representan como conjuntos de bloques interconectados por una red elástica donde cada conjunto viene definido en función de las regiones de la estructura facial generando una superficie formada por 38 de estos bloques.

##### **3.7.3.1.2. Técnicas que emplean modelos vectoriales**

Los modelos vectoriales deforman la superficie facial empleando campos que limitan el movimiento a las regiones de influencia.

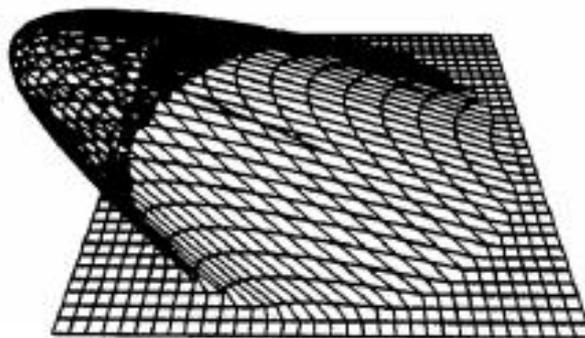
Un modelo muscular basado en las técnicas musculares vectoriales ampliamente aceptado y utilizado fue propuesto por Waters [Wa87]. Su modelo se basa en la mecánica muscular real. Los músculos se conectan a un extremo del cráneo y a la piel. La contracción y relajación de los músculos provoca cambios en la apariencia facial. En el modelo de Waters, se asigna a cada músculo simulado una región de la superficie que representa el rostro para modelar la acción de los músculos sobre la piel. La variación de estos músculos, siguiendo un conjunto de parámetros preestablecidos que simulan la mecánica muscular, provoca la variación de la región de la superficie del modelo asignada. En el modelo se definen tres tipos de músculos: lineal,

elipsoidal y por trozos (sheets). El modelo propuesto por Waters simula los músculos faciales que se contraen como una elipse simplificada que simula los músculos referentes a la región de la boca. Cada tipo de músculo está pensado para emular la actividad muscular facial en una región concreta del rostro, siguiendo el efecto de los distintos músculos faciales reales.

Los músculos lineales se implementan como vectores que se contraen y relajan, produciendo variaciones en la región de la superficie facial que tienen pre-asignada. El campo de acción de cada uno de los músculos viene definido por funciones cosenoidales y factores de disminución que producen una forma cónica al visualizarlo como un campo de intensidades. También se les conoce como músculos paralelos.

Se asigna a cada músculo simulado una serie de parámetros, como los extremos de desplazamiento y las zonas de máxima y mínima influencia. Esto puede parecer muy tedioso ya que solo se pueden medir los puntos de la superficie, y el rango de características varía mucho dependiendo de la zona, por ejemplo, los mayores desplazamientos se realizan en la zona de la boca. Además, la zona de influencia depende del grado de contracción. Summerfield [Su83] propuso establecer grados de libertad para los nodos alrededor de la boca.

Los fundamentos de los modelos vectoriales musculares determinan que un extremo del músculo tiene una inserción ósea y permanece estático en esta zona, mientras que el otro extremo es incrustado en el tejido de la piel, y cuando se acciona el músculo, su movimiento se ve reflejado en esta región, [Wa87]. De una forma simplificada, se puede tomar que la forma de la zona de influencia de los vectores musculares es circular, y su efecto se extiende a lo largo del radio de la red tridimensional como muestra la *Figura 3.7*. Los vectores musculares vienen definidos por su dirección y magnitud tanto en dos como tres dimensiones. La dirección del vector viene definida hacia el punto de unión, y su magnitud es máxima en el punto de inserción de la piel.



*Figura 3.7 - Efecto de un vector muscular en una red tridimensional*

En la *Figura 3.7* se muestra la forma en que el tejido adyacente donde se encuentra el vector muscular se ve afectado cuando este se contrae: se puede asumir que el desplazamiento será máximo en el punto de inserción de la piel, y en el punto de inserción ósea será nulo. Empleando un método de interpolación no lineal se puede representar la acción del músculo de una forma simple (*Figura 3.8*). El desplazamiento se transmite a lo largo de los puntos  $P_m$ ,  $P_n$  y  $V_1$ ,  $P_s$ . Aplicando los mismos principios se puede emplear el mismo modelo para una tercera dimensión, desplazando el punto  $p(x, y, z)$  a  $p'(x', y', z')$ .

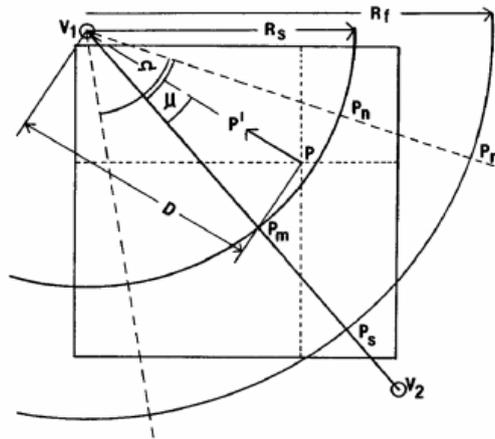


Figura 3.8 - Modelado del efecto de un vector muscular en el sector V1 Pr Ps. Rs y Rf representan el desplazamiento de f' a lo largo del vector V1 V2

En la Figura 3.8, V1 y V2 son dos puntos en un espacio bidimensional.

Rs representa el comienzo del radio de acción del vector, referenciado a V1.

Rf representa el máximo radio de acción del vector referenciado a V1.

Considerando cualquier punto P(x, y) de la malla y dentro de la zona V1 Pr Ps, es desplazado a lo largo del vector PV1 hasta P' (x', y') donde:

$$x' \in f(K, A, R.x)$$

$$y' \in f(K, A, R.y)$$

Siendo:

K la constante de elasticidad del músculo

Ω la zona máxima de influencia

D el vector V1 P

El desplazamiento angular, A se define como:

$$A = \cos \left( \frac{\mu}{\pi} \cdot \frac{\pi}{2} \right)$$

Donde μ es el ángulo entre V1 V2 y V1 P.

El factor de desplazamiento radial R se define como:

$$R = \cos \left( \left( 1 - \frac{D}{R_s} \right) \cdot \frac{\pi}{2} \right)$$

Donde si P se encuentra dentro de V1 Pn Pm, entonces:

$$R = \cos \left( \frac{(D - R_s)}{(R_f - R_s)} \cdot \frac{\pi}{2} \right)$$

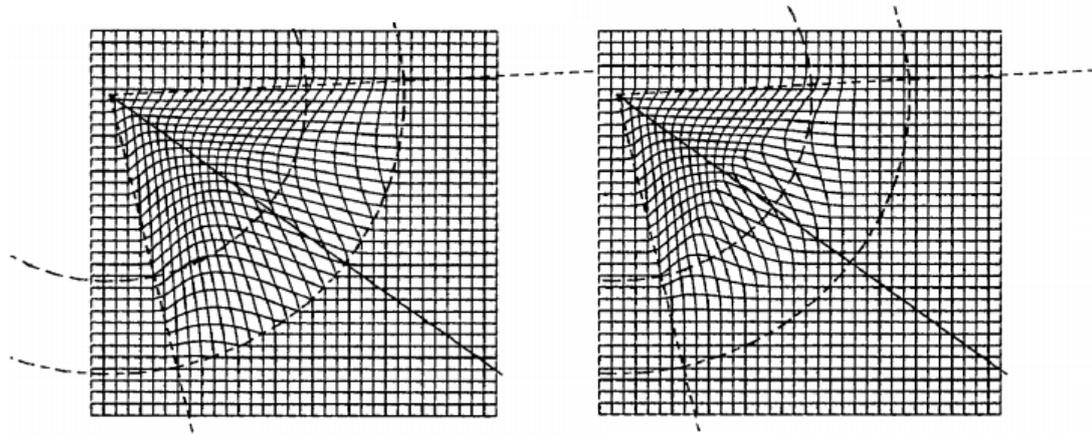


Figura 3.9 - Efecto de interpolación cosenoidal variando el factor de elasticidad

En la *Figura 3.9*, tanto en la de la izquierda como en la de la derecha se muestra una interpolación cosenoidal, donde todos los valores de las constantes permanecen invariables excepto el factor de elasticidad, que se ha incrementado para disminuir la elasticidad de la malla.

Los músculos que provocan contracciones en la piel se pueden describir como un punto en torno al cual se contrae la superficie formando una circunferencia. Se puede decir que este efecto ocurre de manera uniforme alrededor del punto de contracción, por lo que el desplazamiento angular ya no se requiere.

$$\begin{aligned} x' &\in f(K, R.x) \\ y' &\in f(K, R.y) \end{aligned}$$

Es evidente que los músculos no se comportan de manera tan regular, por lo tanto, se generan ondas elípticas a lo largo de los ejes longitudinales y verticales.

Los límites de acción del músculo se pueden determinar por la constante de elasticidad  $K$  (*Figura 3.10*), que representa el desplazamiento máximo del músculo, [Wa87]. El problema asociado a este modelo es que cada acción muscular es independiente, y el desplazamiento real de la superficie se determina por una sucesión de acciones musculares. Esto es más extremo cuando los nodos son desplazados fuera de la zona de influencia de los vectores musculares adyacentes, y existe el peligro de que se supere el grado de libertad de cualquier nodo. Para evitarlo, se puede definir el grado de libertad de cada nodo de la superficie poligonal en función de la posición facial de los músculos y el pre-procesado de la estructura.

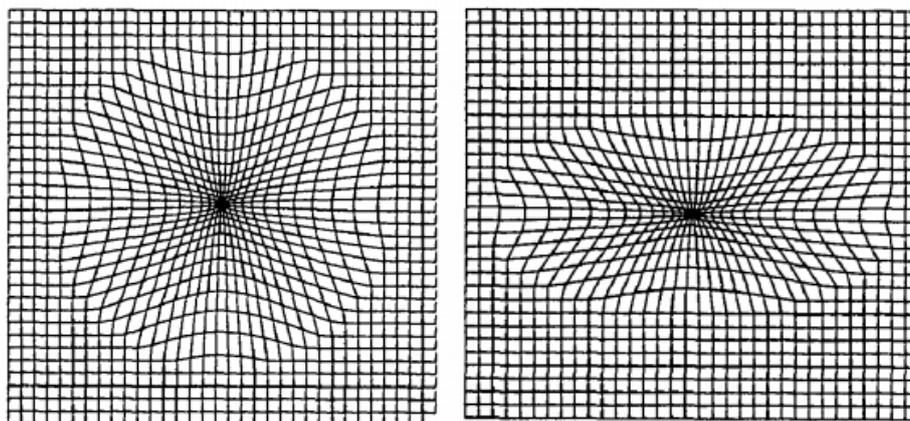
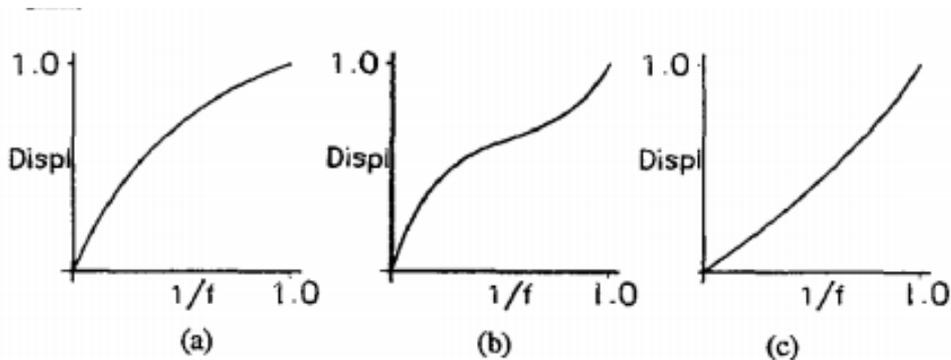
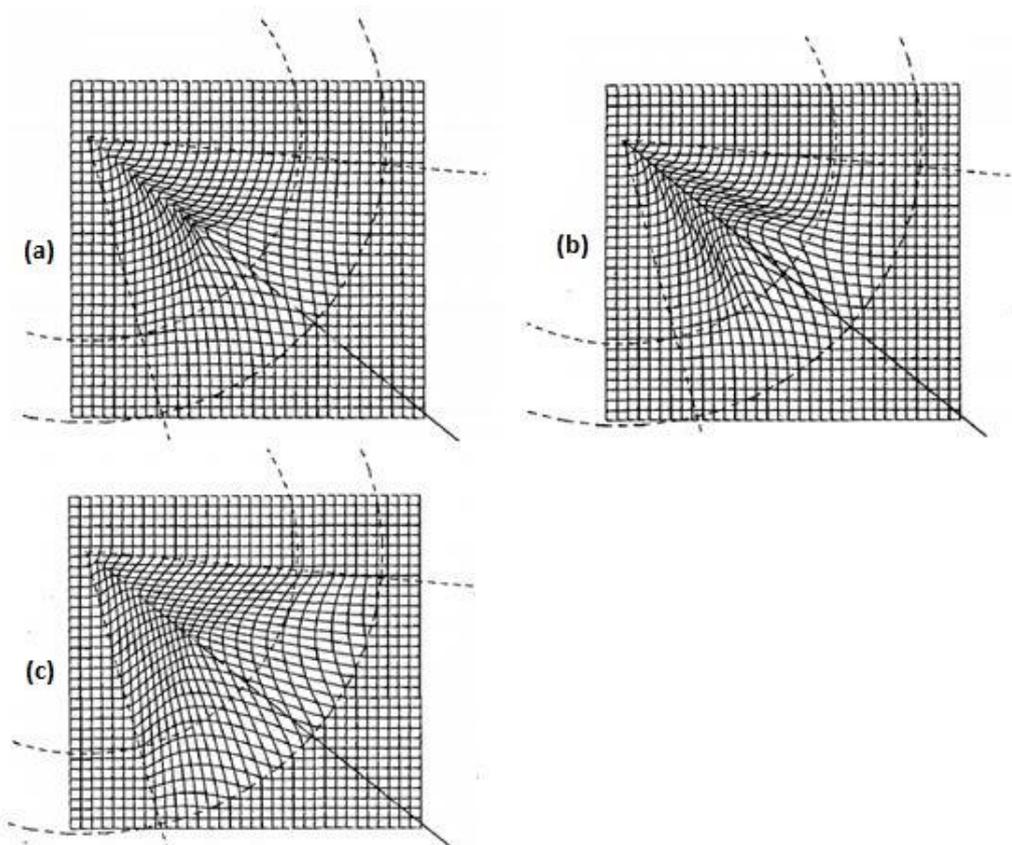


Figura 3.10 – Efecto en los límites de acción del músculo variándola constante de elasticidad  $K$ .

Se pueden emplear funciones cosenoidales para definir el campo de acción. El requisito para emplear funciones cosenoidales radica en que el desplazamiento entre el punto de unión entre la inserción en el hueso y la piel sea máximo. Se puede emplear cualquier curva mostrada en la *Figura 3.11*, cuyo efecto se muestra en las figuras (a), (b) y (c) de la *Figura 3.12*, respectivamente. [Wa87].



*Figura 3.11 - Curvas de aplicación en las funciones cosenoidales que definen el campo de acción del vector muscular*



*Figura 3.12 - Efecto de un vector muscular en el plano x y relativo a (a), (b) y (c) respectivamente.*

Uno de los modelos propuestos recientemente, emplea elementos finitos y lineales para determinar la acción del músculo, capturado por marcadores faciales [Ef06]. Lo interesante de esta aproximación es que el modelo puede ser modificado por fuerzas exteriores, como por ejemplo el impacto de un objeto sólido. Se puede tratar de incrementar el realismo incluyendo un esqueleto interior y con una deformación mejorada en los labios.

### 3.7.3.1.3. Técnicas que emplean capas de superficies elásticas (Layered spring mesh)

Las técnicas que emplean capas de superficies elásticas amplían las estructuras masa-muelle en tres superficies elásticas conectadas que modelan el comportamiento anatómico facial con mayor fidelidad que el resto de métodos.

Terzopoulos y Waters [Te90] propusieron un modelado que detalla la estructura anatómica y el movimiento dinámico del rostro humano. Las tres capas deformables de la red corresponden a la piel, el tejido adiposo y la capa muscular unida a los huesos. Los elementos elásticos conectan cada nodo de la malla y cada capa entre sí. Las fuerzas que actúan en los músculos se propagan a través de la red para generar la animación. Este modelo cuenta con un alto grado de realismo, sin embargo, esta técnica requiere una carga computacional muy grande. Un modelo simplificado reduce la carga computacional conservando el realismo que se obtiene con esta técnica [Mag94].

Lee et. Al. [Lee95] propone un modelo (Figura 3.13) basado en las propiedades físicas de la piel y los músculos de una forma más sencilla [Te90]. El modelo consta de tres componentes: una capa de tejido biológico con deformación no lineal, una capa muscular unido entre sí bajo la piel y una estructura ósea impenetrable bajo la capa muscular. La capa tisular se modela por medio de prismas que se dividen en cada una de las tres capas, epidermis, muscular, y superficie ósea. Los elementos de unión que conectan las tres capas simulan la elasticidad de la piel. Los elementos que simulan las fuerzas aplicadas por los músculos conectan la epidermis y la capa ósea. Este modelo ofrece la ventaja de que cuenta con un grado de realismo y fidelidad muy alto, sin embargo, requiere mucha carga computacional y un trabajo de pre-procesado muy extenso.

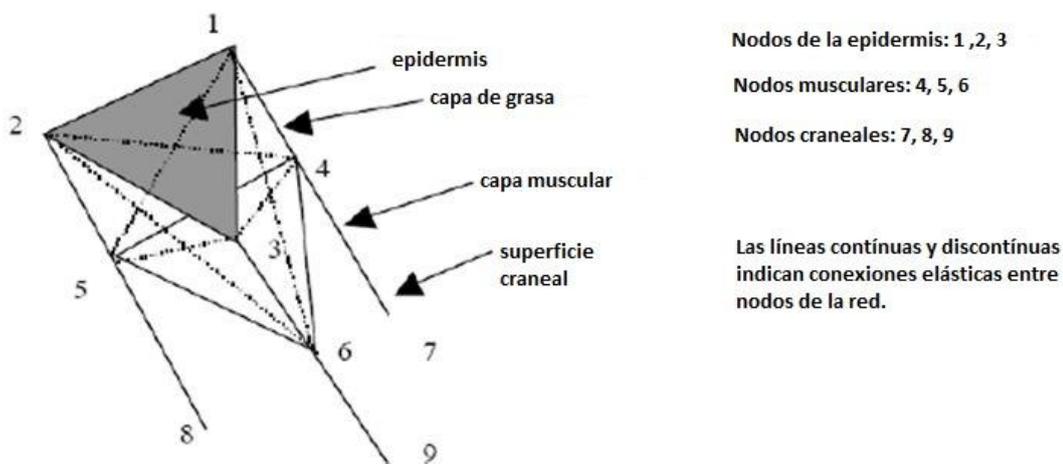


Figura 3.13 – Prisma tricapa del modelo de Lee.

### 3.7.4. Performance driven

Consiste en capturar los movimientos de personas reales (los performers) para emplearlos para animar personajes sintetizados [Vi02].

La transferencia de imágenes se realiza mediante la extracción de movimientos de un sujeto y su aplicación a una imagen objetivo. Dicha transferencia se basa en un modelo 3D

multilineal de rostros. Los modelos multilineales consisten en un conjunto de nubes de puntos que representan los distintos rostros junto con variaciones estimadas en términos de diferentes atributos como el tamaño, la identidad o la expresión. Idealmente, si la colección de rostros es lo suficientemente extensa, el modelado multilineal permite generar cualquier rostro con cualquier expresión o visema asociado.

En [Ks04] se muestra una técnica de captura del movimiento facial basada en el empleo de análisis de componentes principales (principal component analysis o PCA). Mediante PCA, se detecta un conjunto de movimientos faciales principales a partir de los datos capturados, los cuales proporcionan información sobre el movimiento general del rostro. Estos métodos requieren grandes cantidades de información. Asimismo, capturar dicha información significa utilizar una multitud de marcadores faciales. Aun así, esta técnica simplifica enormemente la fase de animación, además produce movimientos faciales más realistas ya que los movimientos incluyen visemas, expresiones faciales y posiciones del rostro y aporta la sutil interpretación de los actores, la expresividad de los cuerpos y el peso de las acciones.

### 3.7.5. Deformación geométrica

Consiste en utilizar un objeto para modificar otro más complejo pero con un control más simple e intuitivo que el keyframing o la parametrización.

Es posible implementar una animación facial mediante el empleo de esqueletos como estructuras que originen movimiento en la superficie del modelo representado. La estructura de esqueleto empleada en animación, denominado armadura, es jerárquica, y se forma posicionando estos elementos de naturaleza de translación o rotacional denominados uniones en los lugares deseados. Cuando uno de estos elementos rota, todos los elementos siguientes de la estructura jerárquica rotarán en la misma medida. Estas uniones son esencialmente elementos que poseen seis grados de libertad y se definen por medio de su posición y rotación. El término *bone* es empleado para definir el vector entre dos de estas uniones que representa su naturaleza jerárquica (Figura 3.14 y Figura 3.15).

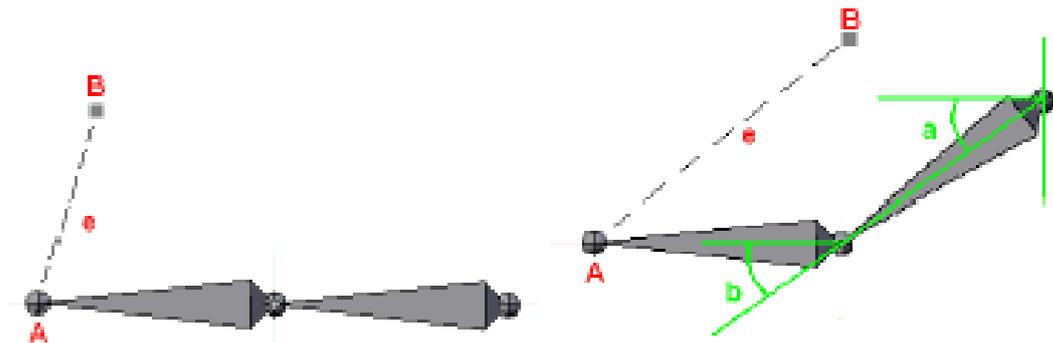


Figura 3.14 – Bone, problema cinemático y ángulos de cada unión

A cada una de las uniones se les puede asociar una parte de la superficie del modelo tridimensional poligonal, haciendo que el movimiento afecte únicamente a dicha región en función del valor de influencia especificado. El proceso de rotar las sucesivas uniones dependientes, con el objetivo de minimizar la distancia entre la última unión subordinada y la posición final deseada, es una operación de cinemática.

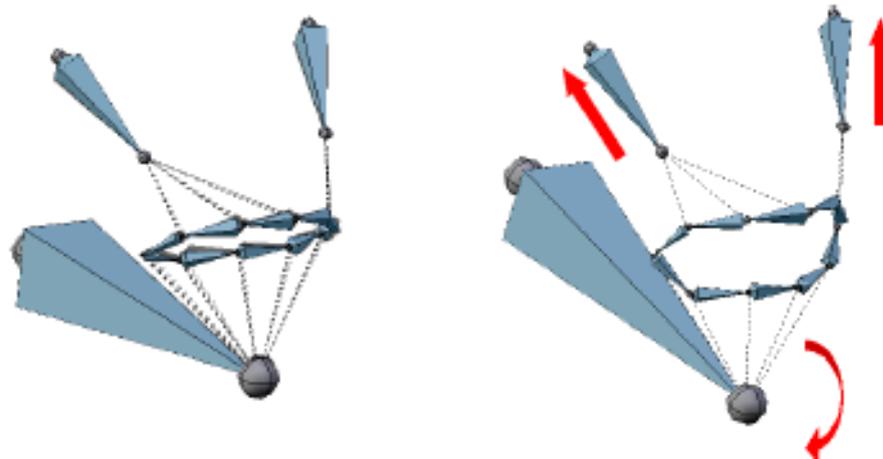


Figura 3.15 – Ejemplo de esqueleto más complejo

En la animación de objetos empleando esqueletos jerárquicos se define la posición final y empleando técnicas de cinemática inversa e interpolación, se calculan los estados intermedios entre la posición inicial y final del objeto. Para disminuir el coste computacional, los cálculos de rotaciones se realizan mediante cuaternios, y las interpolaciones entre distintas rotaciones y translaciones por medio de curvas de Bezier [Fo92], lo que permite suavizar los movimientos realizados.

El término “rig” se emplea para describir un conjunto de parámetros que controlan la rotación o translación de las uniones de la armadura, su posición, su región de influencia en la superficie poligonal, etc. Se puede considerar el rigging como la construcción y posicionamiento de los controles ligados a las uniones de la armadura que controlan y facilitan el manejo y el rango de movimientos que el objeto pueda hacer. El rigging (Figura 3.16) genera, a partir de los valores de los parámetros del rig, el posicionamiento de la geometría del modelo animado con respecto al sistema de referencia global, y lo transfiere al sistema de renderizado. [Ma11]



Figura 3.16 – Rigging como proceso intermedio en la animación

La deformación geométrica requiere un buen sistema de control por encima que controle los bones de manera automática, debido a la alta complejidad de controles y la alta cantidad de información que debe manejar.

### 3.7.6. Técnicas basadas en la combinación de formas o blend shapes

Los métodos basados en la combinación de formas generan expresiones faciales a partir de la combinación de un conjunto de formas de base. Esta combinación se puede realizar mediante interpolación lineal aplicada o bien a imágenes [Zh03] [Zh06], o bien a modelos faciales [Pi98] [Zh04] o a modelos deformables [B199].

El empleo de Blend Shapes genera animaciones de una forma sencilla pasando de una forma base a otra mediante la deformación de la malla. Cada BlendShape se puede modelar mediante diferentes métodos. La cantidad de detalles en cada expresión puede variar, ya que las caras resultantes se pueden combinar en alguna forma. Esta técnica requiere mucha especificación manual y generar una animación completa puede requerir mucho tiempo.

[Ch04] propone un método de animación facial en el que combina la interpolación mediante BlendShapes y captura de movimientos. Este método conserva la flexibilidad que se obtiene mediante el modelado empleando BlendShapes y las ventajas de captura de movimientos. Esta técnica consta de varias etapas: captura facial, descomposición facial y reorientación facial. La *Figura 3.17* muestra un diagrama del proceso descrito.

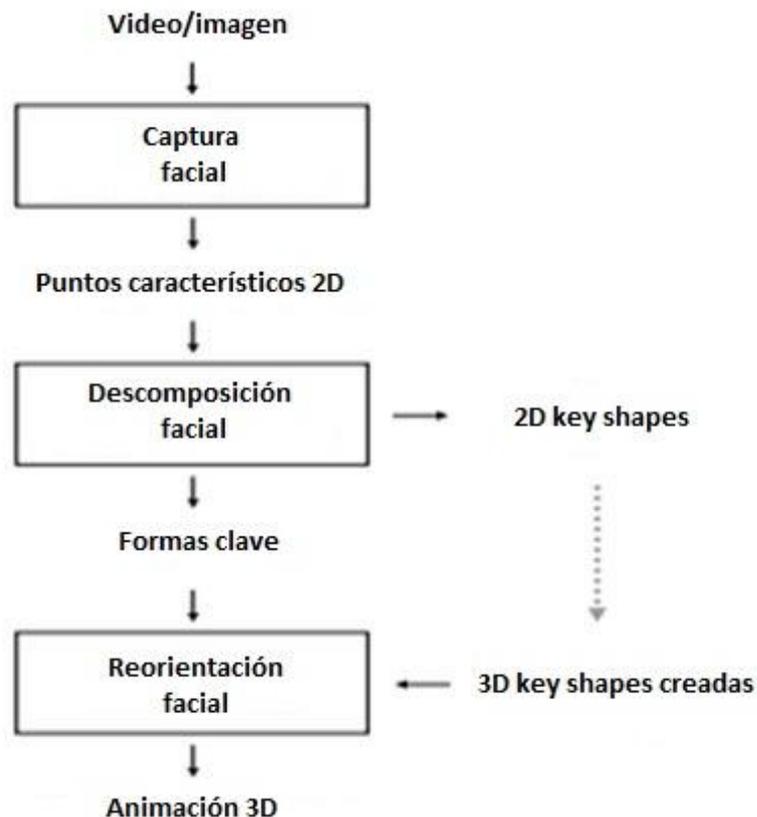


Figura 3.17 – Proceso del modelado empleando Blend Shapes

Para emplear las imágenes, es necesario extraer las posiciones geométricas de ciertos puntos característicos de las imágenes. A continuación es posible obtener una expresión facial realista mediante la combinación convexa de las expresiones faciales basadas en dichas posiciones. Dado que ésta técnica emplea imágenes de expresión de entrada de alta calidad, es posible generar expresiones faciales con alto nivel de detalle.

En la fase de captura facial, se realiza un seguimiento de las características faciales de un modelo empleando técnicas de visión por ordenador.

En la etapa de descomposición facial, los rasgos principales del rostro se descomponen en una combinación ponderada de las formas principales. Estas formas clave se seleccionan automáticamente entre los datos rastreados.

Muchas de las técnicas actuales de rastreo emplean PCA para encontrar las formas bases, pero la desventaja radica en que estas representaciones no se adaptan muy bien en la reorientación. En la reorientación facial se generan las formas claves para la animación facial del modelo de forma manual por el usuario. Estas figuras clave están basadas en la secuencia del video, así una cara sonriente se corresponde a una cara sonriente de la animación. Igualmente aunque la forma del modelo de la cara se modifique, las particularidades de la animación facial permanecen al poder transformar las regiones extraídas entre las poses clave.

La ventaja de esta técnica reside en que el proceso de captura facial y redistribución de puntos se dividen. Además el mapeado del movimiento entre la pose de origen y la destino se simplifica a transferir únicamente los shapes que componen el movimiento haciendo posible emplear la secuencia para otros modelos.

Escoger los KeyShapes apropiados es una parte importante de la fase de descomposición. Los KeyShapes otorgan flexibilidad y expresividad al modelo, lo que apunta a que se necesitarán varios KeyShapes. Sin embargo, el usuario debe fijar cada uno de ellos, por lo que conviene emplear el menor número de KeyShapes necesarios. Un método ideal sería aquel que encontrara el equilibrio entre el empleo mínimo de KeyShapes sin que por ello la animación carezca de expresividad.

El problema se puede expresar de la siguiente manera:

- Dada una secuencia de frames, cada una con su propia forma del rostro,  $Q=[Q_1, Q_2, \dots, Q_t]$  se quiere escoger los  $k$  shapes de  $Q$  que permitan expresar el resto de la secuencia como una combinación de estos KeyShapes siendo el error mínimo.
- Como encontrar el número óptimo de formas clave es un problema que aumenta con las combinaciones del número de estas, se proponen tres métodos para escoger las formas claves necesarias y se evalúan en función de los términos de sus ecuaciones.
- Cada uno está diseñado para escoger las formas claves que aparecen en los frames que componen la secuencia.
- Como el número de características faciales en una secuencia es relativamente grande, la dimensión del vector que contiene información sobre las características tendrá una dimensión mayor.
- Como cualquier problema de clasificación de patrones, una entrada de dimensiones menores hace que el problema sea más manejable. Por ello los tres métodos que emplean análisis de los principales componentes (PCA- Principle Component Analysis) reducen las dimensiones de los datos de entrada.

A continuación se evaluarán los diferentes métodos para la elección de los KeyShapes:

- Máxima difusión a través de los componentes principales: Este método recoge los datos de aquellos puntos que contienen mayor información al proyectarse sobre los ejes principales. Partiendo del primer eje principal con mayor valor propio, se toman aquellos frames con mayor y menor proyección en los ejes como formas clave. La base de este método es que para cada dirección de variación en la secuencia, se toman las poses extremas mínima y máxima que codifican esta variación. Otros Shapes se pueden establecer como la interpolación entre estos puntos extremos.
- Mediante agrupación de características/componentes: Cada forma de la cara se proyecta en un número menor de vectores propios para formar un vector de dimensiones menores. Estos son agrupados en k conjuntos empleando métodos euclídeos. El valor central de cada grupo se toma como un Keyshape. Con este método se garantiza cubrir el rango de movimientos posible. Los resultados de este método aparecen recogidos en vectores de dos y tres dimensiones.
- Convex hull: Cada forma de la cara se proyecta sobre un pequeño número de vectores propios produciendo un vector de dimensiones menores. Todas las formas que se encuentran en la zona convexa se toman como formas clave. Como cada otra forma tiene que estar en la envolvente convexa, se estima que el resto de las formas se pueden obtener mediante interpolación de las formas seleccionadas. Para mayores dimensiones que dos, cada forma de las secuencias se encuentran en la zona de convergencia. Dado que el uso de cada fotograma como una forma clave contradice el propósito de este método, los resultados se expresan en vectores de dos dimensiones.

# Capítulo 4

## *Anatomía y expresiones del rostro*

### **4.1. Introducción**

Resulta obvio que el diseño de un avatar que presente similitudes con el rostro humano requiera de un cierto conocimiento tanto de la anatomía facial del ser humano como de los aspectos funcionales y psicológicos asociados a la expresividad del rostro.

Debido a la complejidad que reside en la fisiología humana, se torna imposible desarrollar una animación que opere en tiempo real y que sea íntegramente fiel a la apariencia visual del movimiento facial. Para conseguir animaciones que simulen de forma fiel los movimientos faciales es necesario emplear un sistema que describa la actuación de los músculos faciales para formar distintas expresiones.

En el presente capítulo, se realiza un análisis de aquellos aspectos a tener en cuenta al realizar animaciones de los movimientos faciales de los rostros humanos. El rostro humano es la parte más expresiva del cuerpo humano y la región del cuerpo con mayor número de músculos permitiendo generar un sin fin de movimientos faciales. Debido a la riqueza de movimientos que el rostro puede generar, es necesario un sistema que sirva de estándar fundamentado en la anatomía del rostro y que describa la actuación de los músculos implicados para formar distintas expresiones en función de la región donde se realice el movimiento considerando solamente aquellos músculos involucrados y siempre en función de lo que resulta distinguible al ojo humano.

## 4.2. La anatomía del rostro humano.

Los tres principales componentes de la cabeza humana son la piel, la estructura ósea y la estructura muscular del rostro; aunque en la generación de expresiones, juegan un papel trascendental los principales órganos visibles: los ojos, los dientes y la lengua.

La piel (*Figura 4.1*) es un sistema complejo cuya función es proteger al organismo frente al medio que le rodea, regular la temperatura corporal y un medio de almacenamiento de agua, sangre y grasa. Desde el punto de vista de la apariencia, la piel del rostro está dividida en varias regiones: las sienes, las mejillas, la nariz, los párpados, los labios y el mentón. Dentro de cada unidad, la piel es consistente en color, textura, grosor y movilidad.

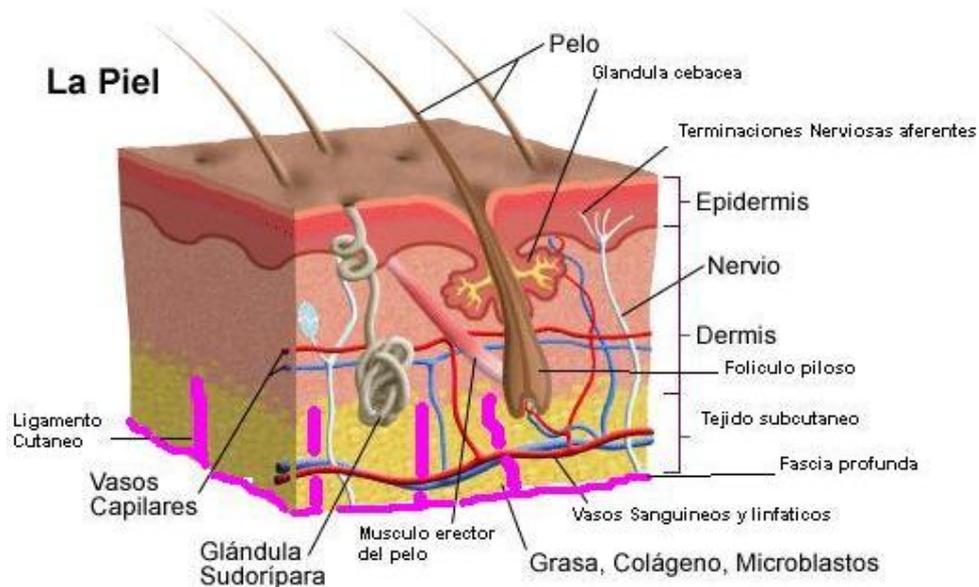


Figura 4.1 – Estructura de la piel

La cabeza ósea o calavera (*Figura 4.2*) es el conjunto de huesos que forman el esqueleto de la cabeza. Está formado por tejido vivo que se adapta a las fuerzas inducidas por los movimientos de la musculatura facial. Está compuesta por 28 huesos, de los cuales sólo la mandíbula es móvil. Su principal función es la de proteger al encéfalo y al resto de órganos sensoriales, además de contener el aparato de la masticación.

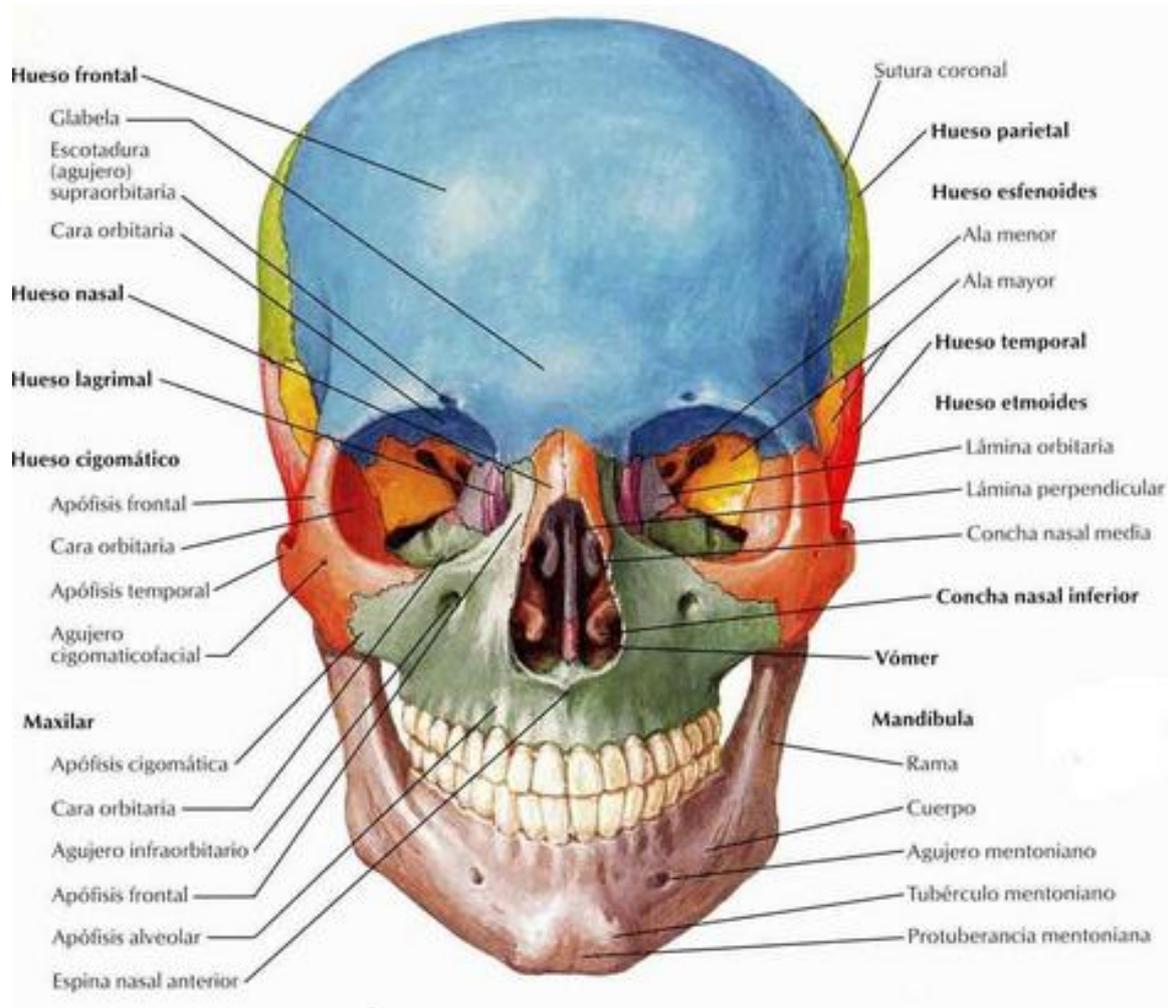


Figura 4.2 – Estructura del cráneo

Respecto a la estructura muscular (*Figura 4.3*), el rostro es la región del cuerpo con mayor número de músculos. La riqueza de movimientos faciales que el ser humano es capaz de realizar es posible gracias al amplio conjunto de músculos que se sitúan entre la piel y el cráneo (más de 50).

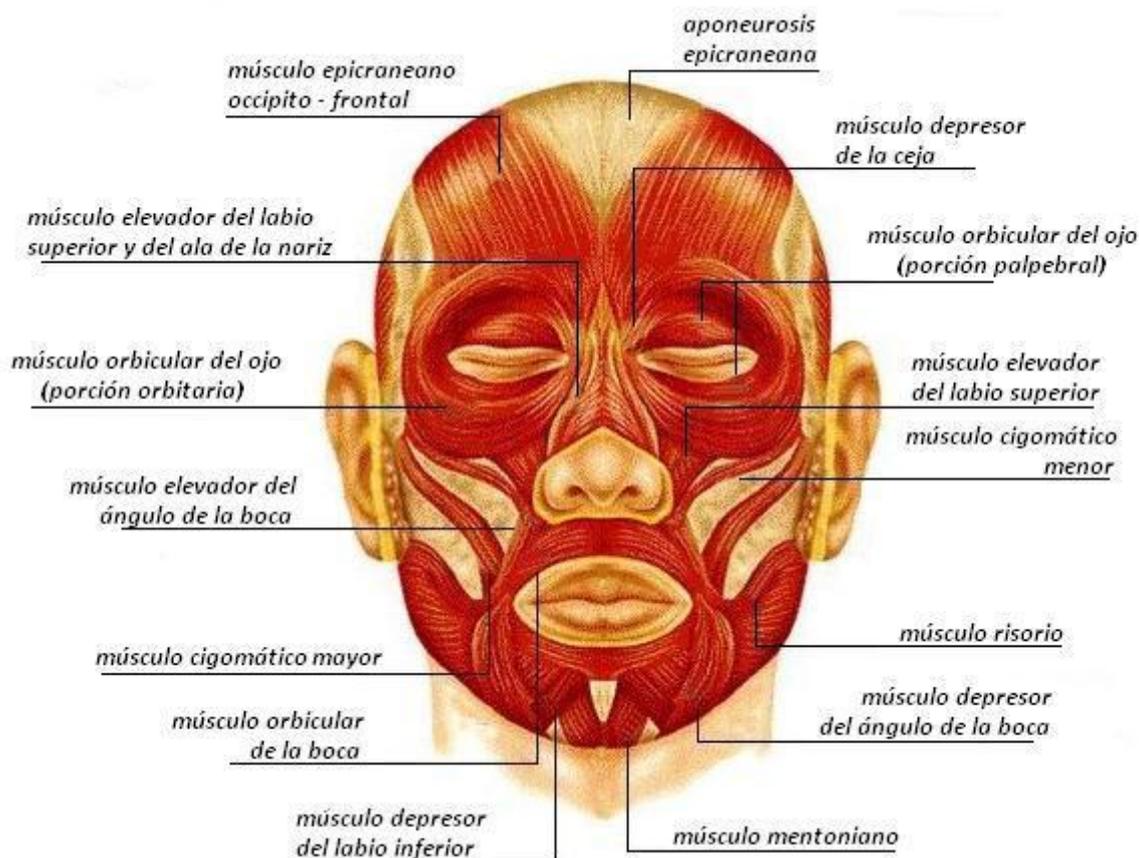


Figura 4.3 – Estructura muscular

Los músculos de expresión originan deformaciones en la piel al contraerse. La contracción generalmente origina arrugas en la piel, que se sitúan de forma perpendicular a la dirección de contracción de las fibras musculares.

Los músculos principales de expresión facial y sus efectos en la cara cuando se contraen son los siguientes:

- Cigomático mayor (Zygomatic Major): desplaza la comisura de los labios superior lateralmente. La parte inferior del surco naso labial se profundiza cuando se estira hacia arriba y hacia fuera
- Depresor del ángulo de la boca (Angular Depressor): tira de la comisura de la boca inferolateralmente. Profundiza la parte inferior del surco naso labial y lo empuja hacia abajo.
- Elevador del ángulo de la boca (Labi Nasii): eleva la comisura y el labio superior, empujando la parte interior de la mejilla hacia el ojo.
- Elevador del labio superior y del ala de la nariz (Inner Labii Nasii): atrae en dirección superior el ala de la nariz y el labio superior. Se eleva y profundiza el surco naso labial.
- Frontal interior (Frontalis inner): la parte media del músculo frontal eleva la parte interior de las cejas.
- Frontal exterior (Frontalis outer): la parte lateral del músculo frontal eleva la parte exterior de las cejas.
- Frontal (Frontalis major): eleva las cejas de forma que el arco formado es más prominente.
- Corrugador lateral (Lateral Corrugator): tira de la parte interna de las cejas.
- Corrugador de la ceja o superciliar (Corrugator Supercilli): junta las cejas.

El estudio de los principales músculos faciales permite conocer el resultado visual de las deformaciones más significativas que se produce en el rostro humano fruto de la contracción muscular.

### 4.3. Las expresiones faciales

El rostro es la parte más expresiva del cuerpo humano y el canal más importante en la comunicación no verbal. Cuando uno interactúa con otra persona, las expresiones emocionales de ambos interlocutores son observadas e interpretadas durante el flujo de la conversación [St96] expresando su estado de ánimo o sus intenciones hacia otros, ya que las emociones se usan para enfatizar o ayudar a alcanzar una meta como actos intencionales de comunicación.

La expresión de las emociones durante la interacción social, forma parte del sistema de comunicación de los seres humanos y se pueden considerar como impulsos que nos llevan a actuar, programas de reacción automática con los que nos ha dotado la evolución. El primero en demostrar la universalidad de las emociones fue Charles Darwin en 1872 en “The expression of emotions in man and animals”.

La investigación más importante sobre expresiones faciales fue desarrollada en 1862 por Duchenne (*Figura 4.4*), centrada en el estudio de los movimientos de los distintos músculos de la cara empleando electrodos en puntos de la superficie del rostro. De este modo, pudo clasificar los músculos en expresivos, inexpressivos o discordantemente expresivos [Con05]. A pesar de que existen discrepancias entre su clasificación y la de estudios recientes, fue el que definió esencialmente el campo.



Figura 4.4 – Imágenes de la investigación de Duchenne

Otro estudio destacable de este campo fue realizado por Fridlund en 1994 en su Programa de Expresión facial (P.E.F.) [Fri94] [Ru97]. Este programa consta de una serie de hipótesis, teorías y métodos, pero es difícil encontrar una aceptación compleja del conjunto, dado que cada autor presenta su propia visión sobre la materia. No obstante, la mayoría de ellos están de acuerdo en que la principal conclusión del programa es el carácter de la universalidad de las expresiones faciales al demostrar una emoción [Ru97], es decir:

En todos los grupos humanos ocurren los mismos patrones de movimientos faciales. Observadores de distintas sociedades atribuyen una misma emoción a patrones establecidos como universales. Esos patrones universales son por lo tanto manifestaciones de todas esas emociones

en todas las sociedades humanas. Por lo que las emociones pueden clasificarse en categorías discretas en un lenguaje común y universal [Ro09].

Las expresiones correspondientes a las principales emociones primarias son fácilmente distinguibles y categorizables en diferentes patrones de gestos realizados mediante el rostro. Paul Eckman y Wallace Friesen desarrollaron un estudio más riguroso y preciso sobre el estudio analítico de la expresión de la cara humana. Desarrollaron el Facial Action Coding System (FACS) que permitía medir con rigor científico todos los movimientos musculares de la cara. Dividieron las expresiones en acciones de pequeños grupos de músculos que denominaron Action Units (AUs).

Parte de su estudio fue comprobar científicamente si los gestos y las expresiones difieren con la cultura, y sus investigaciones concluyen que hay seis categorías que pueden considerarse universales, esto es, reconocibles por cualquier persona de cualquier tipo y condición en cualquier parte del mundo. [Ekm72] [Ek78] [Elf03] [Zhi07].

Estas categorías son: alegría, tristeza, enfado, miedo, disgusto y sorpresa, (*Figura 4.5*).



*Figura 4.5 – Las seis expresiones universales. De izquierda a derecha y de arriba abajo: disgusto, tristeza, enfado, alegría, miedo y sorpresa.*

Cada una de ellas tiene un amplio rango de expresiones con distinta intensidad y variaciones en los detalles de las mismas.

Gary Faigin consideró que existen expresiones no relacionadas con emociones que también están universalmente reconocidas: dolor, esfuerzo, somnolencia... [Fai90].

También, Russell y Fernández-Dols presentaron dos dimensiones para estudiar las emociones y su relación emocional con sentimientos afectivos (*Figura 4.6*). Con estas nuevas dimensiones, se establecen niveles de predominancia entre ellas ampliando la cantidad de emociones clasificables. [Ru97].

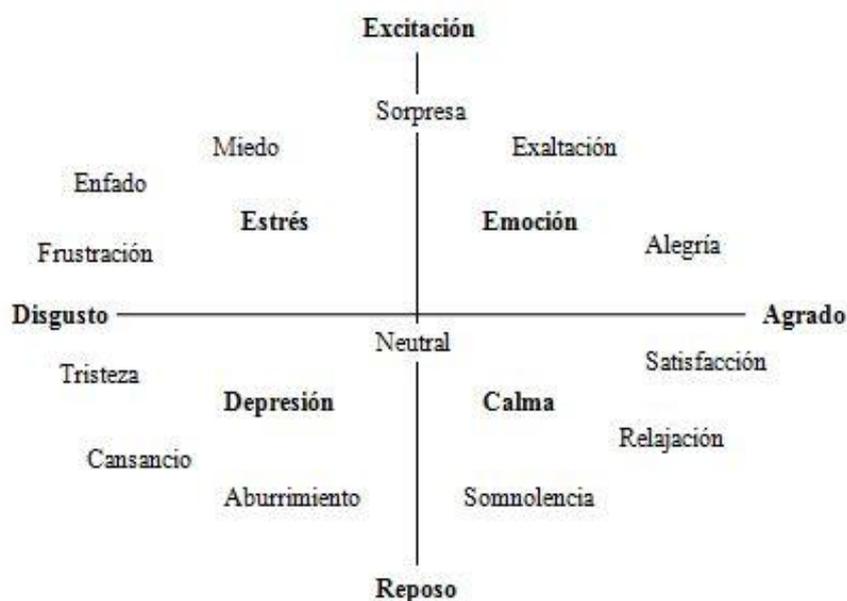


Figura 4.6 – Diagrama en dos dimensiones de Russell y Fernández-Dols.

#### 4.4. Facial Action Coding System

El Facial Action Code System o FACS propuesto por Ekman y Fiesen en 1978 [Ekm78], es el sistema más utilizado para el análisis del comportamiento facial e intenta responder a la ausencia de métricas para la clasificación de la emociones básicas con una serie de puntos a observar en el rostro. Se considera casi un estándar en la clasificación facial, estando presente en investigaciones tanto del área de psicología como de la animación 3D [Cu06] [Woj06].

Las unidades de medida del FACS son las unidades de acción (AUs o Action Units), las cuales se relacionan anatómicamente con la contracción de un conjunto específico de músculos faciales.

Es importante resaltar que aunque el modelo de FACS surge de la anatomía del rostro, no existe una correspondencia 1:1 entre los grupos de músculos y las AUs, esto se debe a que se agrupan un conjunto de músculos en una misma AU porque los cambios que generan los músculos individualmente de la apariencia no son distinguibles, o porque un mismo músculo puede dividirse en varias AUs al poder actuar de diferentes formas o contraerse en distintas regiones generando cambios de la apariencia visibles diferentes [Con09]. Un ejemplo claro de esto son los músculos frontales (*Figura 4.7*); la contracción de la zona media de ellos sólo levanta las esquinas interiores de las cejas (produciendo la AU 1), mientras que la contracción de la porción lateral de los frontales levanta las cejas desde su parte externa (produciendo AU 2).

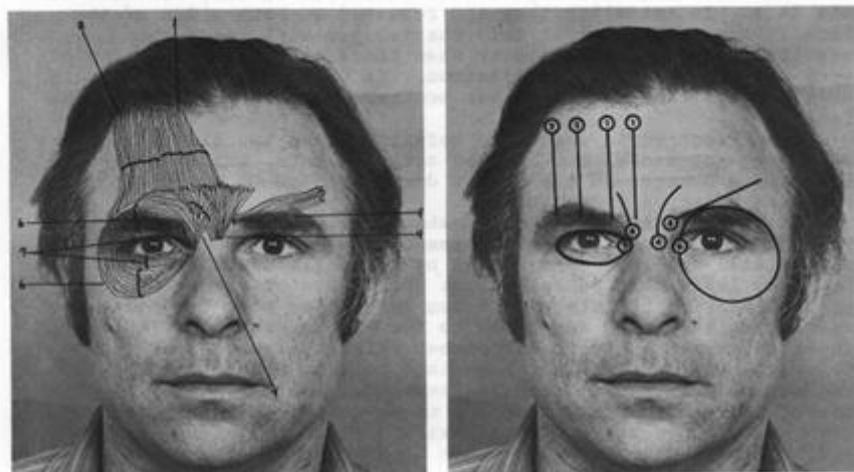


Figura 4.7 – Músculos que intervienen en las unidades de acción superiores

Las bases del modelo de FACS establecen que el comportamiento facial se descompone en 46 unidades de acción (AUs o Acción Units), cada una relacionada anatómicamente con la contracción de un conjunto específico de músculos faciales, además de varias categorías de movimientos y posiciones de la cabeza y los ojos. Aunque sólo se define un número pequeño de AUs, existen más de 7000 combinaciones de estas hasta el momento.

Una expresión facial observada se puede descomponer en las AUs que generan dicho movimiento y la intensidad, duración o simetría de cada una, por lo que el sistema de codificación facial es un método potente para la detección y medición de un gran número de expresiones faciales prácticamente por la observación de un pequeño conjunto de acciones musculares. Sin embargo, la apariencia exacta de cada unidad funcional varía de un individuo a otro, dependiendo de la estructura ósea, variaciones en la musculatura facial, depósitos de grasa, arrugas permanentes, características faciales, etc. por lo que éste sistema no tipifica al poder aparecer sutiles variaciones dependiendo de la persona.

Además, la formación de expertos y anotar todas las AUs es costoso y consume mucho tiempo, y la fiabilidad de la codificación humana inherentemente viene atenuada por la subjetividad del codificador humano. Por lo tanto, un sistema capaz de reconocer las AUs en tiempo real sin intervención humana resultaría útil en diversos campos de aplicación incluyendo herramientas automatizadas para la investigación del comportamiento, videoconferencia, interfaces hombre-máquina, reconstrucción 3D de rostros y animación entre otros.

#### 4.4.1. Las unidades de acción

Las unidades de medida del sistema de codificación facial son las unidades de acción o AUs. Se han definido 46 AUs, cada una relacionada con la contracción de un conjunto específico de músculos faciales. De estas 46 unidades de acción, 30 están relacionadas con las contracciones de músculos específicos del rostro: 12 para la parte superior de la cara y 18 para la inferior.

En el sistema de codificación facial se descartaron aquellas acciones musculares en las que su diferencia no se tipificara por no ser fiable. Además, en ocasiones los cambios de apariencia resultante de dos músculos distintos difieren mayormente en la intensidad de la acción,

no en la apariencia final. Para estos casos, se describe una única unidad de acción la cual puede ser producida por varios músculos diferentes.

Generalmente, es raro que aparezca una única AU aislada en la expresión facial. En cambio, frecuentemente se muestran algunos patrones de combinaciones de AUs en las expresiones de emociones. Así, aparecen varias AUs de forma simultáneamente en la mayoría de los casos. Las unidades de acción que aparecen de manera simultánea, pueden hacerlo por dos métodos diferentes: de forma aditiva, en la que se modifica la AU primitiva, y no aditiva, en la que la AU no es modificada, [Ek02]. El estudio de estas combinaciones concluye que los cambios de apariencia originados pueden no ser ni aditivos ni distintivos entre dos AUs, sino estar regidos por una relación de “dominio”. La unidad de acción dominante hace difícil la detección de la unidad subordinada.

Para lograr un consenso en la observación, se establecen ciertas reglas de forma que no se considera el movimiento de la unidad subordinada cuando existe una evidencia plausible de la existencia de dicha relación de dominio. El sistema no especifica aquellas combinaciones de unidades que no interactúan entre sí. Por ejemplo, tirar del extremo de los labios se lleva a cabo por un músculo que no afecta a aquellos responsables de la posición de las cejas. Por otro lado, existen un reducido número de AUs que no se pueden realizar de manera simultánea, como se describe en las normas definidas en FACS.

En la *Tabla 4.1* vienen recogidas y descritas las unidades de acción definidas en el sistema de codificación facial. Cada unidad de acción viene identificada por un número y un nombre indicando la región y movimiento descrito. La asignación de cada AU a un número es arbitraria, por ejemplo, números consecutivos de AUs no corresponden a una misma AU con cambio de apariencia similares.

<b><u>NÚMERO DE AU EN EL FACS</u></b>	<b><u>NOMBRE ORIGINAL</u></b>	<b><u>MÚSCULOS QUE INTERVIENEN</u></b>
<b>AU1</b>	<i>Inner Brow Raiser</i> / Levantar el interior de las cejas	Frontal (par medial)
<b>AU2</b>	<i>Outer Brow Raiser</i> / Levantar el exterior de las cejas	Frontal (par lateral)
<b>AU4</b>	<i>Brow Lowerer</i> / Bajar las cejas	Corrugador superciliar, Depresor superciliar, Próceros
<b>AU5</b>	<i>Upper Lid Raiser</i> / Levantar el párpado superior	Elevador del párpado superior
<b>AU6</b>	<i>Cheek Raiser</i> / Levantar las mejillas	Orbicular del ojo (par orbital)
<b>AU7</b>	<i>Lid Tightener</i> / Estrechar los párpados	Orbicular del ojo (par parpebral)
<b>AU9</b>	<i>Nose Wrinkler</i> / Arrugar la nariz	Elevador del labio superior
<b>AU10</b>	<i>Upper Lip Raiser</i> / Elevar el labio superior	Elevador del labio superior, alar de la nariz
<b>AU11</b>	<i>Nasolabial Furrow Deepener</i> / Elevar el labio superior	Elevador del labio superior
<b>AU12</b>	<i>Lip Corner Puller</i> / Separar las comisuras de los labios	Cigomático mayor
<b>AU13</b>	<i>Cheek Puffer</i> / Subida e inflamación de los carrillos	Triangular
<b>AU14</b>	<i>Dimpler</i> / Retraer los labios y estrechamiento de las comisuras	Buccinador
<b>AU15</b>	<i>Lip Corner Depressor</i> / Bajar las comisuras de los labios	Triangular

<b>AU16</b>	<i>Lower Lip Depressor</i> / Depresión del labio inferior	Depresor del labio inferior
<b>AU17</b>	<i>Chin Raiser</i> / Levantar la barbilla	Borla
<b>AU18</b>	<i>Lip Puckerer</i> / Contraer los labios adelantando y redondeando la boca (beso)	Incisivo del labio superior, incisivo del labio inferior
<b>AU20</b>	<i>Lip Stretcher</i> / Estirar horizontalmente los labios	Risorio
<b>AU22</b>	<i>Lip Funneler</i> / Labios en posición de embudo	Orbicular de la boca
<b>AU23</b>	<i>Lip Tightener</i> / Afinar los labios	Orbicular de la boca
<b>AU24</b>	<i>Lip Pressor</i> / Apretar los labios	Orbicular de la boca
<b>AU25</b>	<i>Lips Part</i> / Separar los labios	Orbicular de la boca, borla
<b>AU26</b>	<i>Jaw Drop</i> / Bajar la mandíbula	Masetero, temporal, Pterigoideo interno
<b>AU27</b>	<i>Mouth Stretches</i> / Apertura de la boca (bostezo)	Pterigoideo
<b>AU28</b>	<i>Lips Suck</i> / Succión de los labios hacia dentro	Orbicular de la boca
<b>AU29</b>	<i>Jaw Thrust</i> / Apretar los labios	Masetero, temporal
<b>AU38</b>	<i>Nostril Dilator</i> / Dilatar las fosas nasales	Alar de la nariz
<b>AU39</b>	<i>Nostril Compressor</i> / Compresión de las fosas nasales	Alar de la nariz
<b>AU41</b>	<i>Lids Droop</i> / Caída de los párpados	Orbicular del ojo (par parpebral)
<b>AU42</b>	<i>Eyes Slit</i> / Estrechamiento de la apertura de los párpados	Orbicular del ojo (par parpebral)
<b>AU43</b>	<i>Eyes Close</i> / Cerrar los ojos	Orbicular del ojo (par parpebral)
<b>AU44</b>	<i>Squint</i> / Estrechamiento acentuado de la apertura de los párpados	Orbicular del ojo (par parpebral)
<b>AU45</b>	<i>Blink</i> / Parpadeo	Orbicular del ojo (par parpebral)
<b>AU46</b>	<i>Wink</i> / Guiño	Orbicular del ojo (par parpebral)
<b>AU51</b>	<i>Head Turn Left</i> / Girar el cuello a la izquierda	Conjunto de músculos del cuello
<b>AU52</b>	<i>Head Turn Right</i> / Girar el cuello a la derecha	Conjunto de músculos del cuello
<b>AU53</b>	<i>Head Up</i> / Girar el cuello hacia arriba	Conjunto de músculos del cuello
<b>AU54</b>	<i>Head Down</i> / Girar el cuello hacia abajo	Conjunto de músculos del cuello
<b>AU55</b>	<i>Head Tilt Left</i> / Cabeza inclinada a la izquierda	Conjunto de músculos del cuello
<b>AU56</b>	<i>Head Tilt Right</i> / Cabeza inclinada a la derecha	Conjunto de músculos del cuello
<b>AU57</b>	<i>Head Forward</i> / Echar la cabeza hacia delante	Conjunto de músculos del cuello
<b>AU58</b>	<i>Head Back</i> / Echar la cabeza hacia atrás	Conjunto de músculos del cuello
<b>AU61</b>	<i>Eyes Turn Left</i> / Mirar hacia la izquierda	Músculos extraoculares
<b>AU62</b>	<i>Eyes Turn Right</i> / Mirar hacia la derecha	Músculos extraoculares

<b>AU63</b>	<i>Eyes Up / Mirar hacia arriba</i>	Músculos extraoculares
<b>AU64</b>	<i>Eyes Down / Mirar hacia abajo</i>	Músculos extraoculares

Tabla 4.1 – Unidades de acción (AUs) definidas en el FACS.

Respecto a esta tabla, actualmente es aceptado considerar como una única AU el grupo de AUs 25, 26 y 27, y de la misma forma las 41, 42 y 43, aplicando a cada una el criterio de intensidad, explicado en el siguiente punto.

#### 4.4.2. Intensidad de las unidades de acción

Las calificaciones de intensidad para un rango de expresión individual varía de uno a cuatro (1=bajo, 2=medio, 3=alto, 4=muy alto). Los valores más bajos están referidos a expresiones leves como una sonrisa en la que se elevan ligeramente las comisuras de la boca pero no muestra los dientes y se produce poco movimiento alrededor de los ojos. Los valores intermedios aparecen en expresiones más moderadas, como una sonrisa que roce la risa, con las cejas ligeramente levantadas y los labios separados y mostrando los dientes. Los valores más altos se reservan para aquellas expresiones muy intensas, por ejemplo, aquellas que evidentemente la persona está riendo con la boca completamente abierta y las cejas y mejillas elevadas.

El sistema de codificación facial emplea una serie de umbrales o niveles que establecen el grado de intensidad de las AUs. Cuando no existe evidencia de que se manifieste una AU específica, la cara se encuentra en un estado neutro. Cuando se manifiesta alguna AU, la intensidad de esta se califica en una escala de 5 puntos. Las letras A, B, C, D y E se refieren a la intensidad de la unidad de acción (Figura 4.8). Estas letras se escriben inmediatamente después del número de la AU aplicada para indicar qué parte de la variación total es debida directamente por la AU. En nivel más bajo corresponde a la letra A, y el más extremo a la E. Se puede ver un ejemplo en la Figura 4.9.

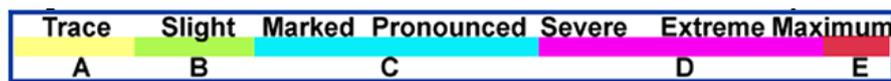


Figura 4.8 – Grados de intensidad de las AUs

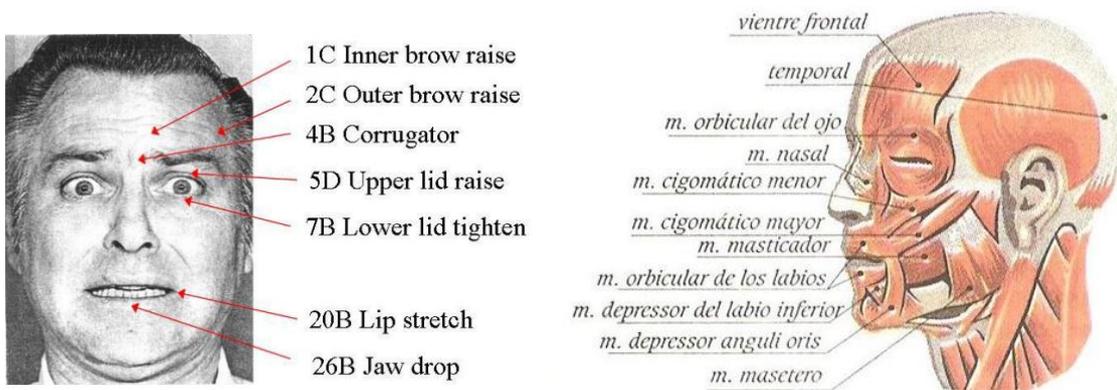


Figura 4.9 – Ejemplo de expresión con las correspondientes AUs y grados de intensidad

Aquellas expresiones con un ligero atisbo de microexpresión apenas imperceptible, se clasifican con la letra A. Expresiones donde se ve claramente una microexpresión, simplemente se ve marcada aunque no mucho, entran dentro de la categoría B. Las expresiones que se ven más pronunciadas que antes y las arrugas pueden provocar microexpresiones entran en el rango C. Por otro lado aquellas emociones muy intensas y las microexpresiones se ven de forma clara y casi sin llegar al extremo pertenecen al nivel D. Por último, las expresiones más extremas donde las expresiones aparecen claramente marcadas entran en la clasificación E.

Existe una correspondencia entre los 4 rangos de intensidad de las FACS y la escala de intensidad de las AUs que definen los umbrales de los niveles de intensidad.

### 4.4.3. Evolución temporal de las unidades de acción

Además del desarrollo dinámico de cada AU, se debe tener en cuenta la evolución temporal de éstas para reflejar de manera realista la evolución del rostro de un estado neutro a una muestra de emoción hasta el estado liberado.

Dicha concurrencia dinámica y semántica se puede modelar con una red dinámica de Bayesian (DBN). Dicho modelo es capaz de representar las relaciones entre diferentes AUs en una estructura jerárquica coherente y unificada, las regiones de incertidumbre en el reconocimiento de las AUs y las dependencias entre ellas, obteniendo así un modelado del dinamismo facial realista.

### 4.4.4. Valencia de las expresiones.

El sistema de codificación facial también consta de un sistema para determinar la valencia de la expresión. Watson y Tellegen establecieron estos descriptores como positivos y negativos. Se proporcionan como una guía para los codificadores en la determinación de la valencia de la expresión, y esta es consultada en caso de que existan dudas con la expresión. La *Tabla 4.2* muestra la clasificación de las expresiones como positivas o negativas.

DESCRIPTORES POSITIVOS	DESCRIPTORES NEGATIVOS
Happy / Feliz	Miserable / Miserable
Delighted / Encantado	Distressed / Angustiado
Glad / Contento	Annoyed / Molesto
Amused / Divertido	Jittery / Nervioso
Pleased / Satisfecho	Nervous / Nervioso
Content / Contento	Angry / Enfadao
Satisfied / Satisfecho	Gloomy / Melancólico
Calm / Calmado	Anxious / Ansioso
Serene / Sereno	Afraid / Asustado
Excited / Excitado	Tense / Tenso
Astonished / Asombrado	Alarmed / Alarmado
Cheerful / Alegre	Frustrated / Frustrado
Surprised / Sorprendido	Disgusted / Disgustado
Active / Activo	Depressed / Deprimido
Content / Contento	Hostile / Hostil

Tabla 4.2 – Descriptores de Watson y Tellegen

## 4.5. Expresiones complejas

### 4.5.1. Descripción de las expresiones básicas universales

Dada la universalidad descubierta por Ekman, es posible hacer un análisis y representar genéricamente estas emociones (Figura 4.10) [Conde09].

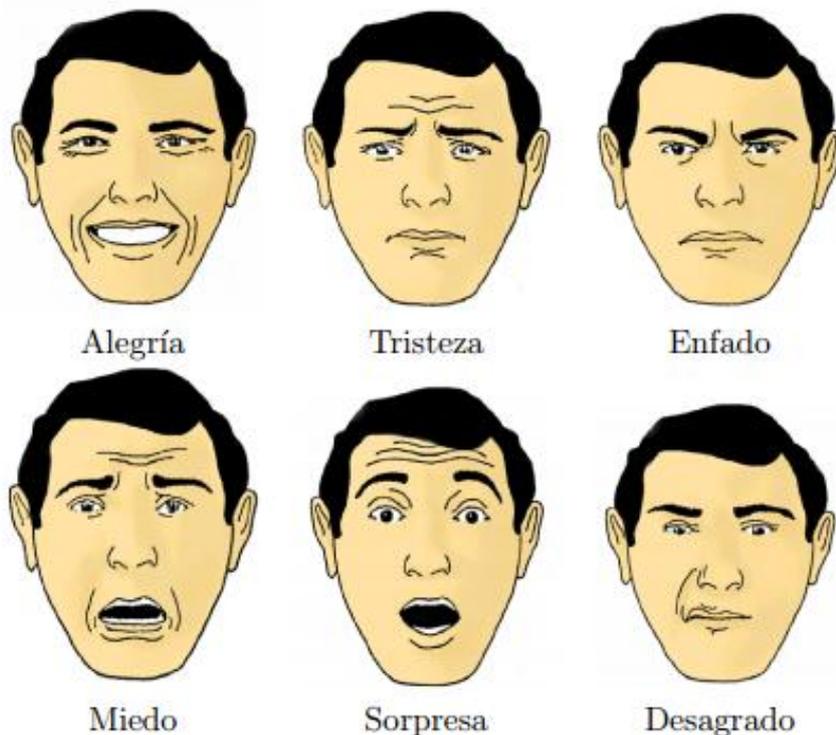


Figura 4.10 – Las seis expresiones universales

- **Alegría**

La alegría es una emoción que hace sentir bien al individuo y le induce a querer repetir la experiencia que lo produce. Puede ser causada por un estímulo agradable, por la atenuación del malestar o la consecución de un objetivo deseado.

La manifestación externa de esta emoción suele estar fuertemente influenciado por el entorno social que nos marca cuándo y dónde se puede expresar esta emoción. Puede producir un incremento en la flexibilidad cognitiva del individuo incrementando su creatividad y actitud ante los problemas.

En la expresión pura de alegría las cejas están relajadas. El párpado superior ligeramente bajado y el párpado inferior recto, siendo elevado por la mejilla. Los labios son finos y se encuentran presionando fuertemente contra el hueso. En la expresión de alegría, se forman también arrugas en los extremos de los ojos, o patas de gallo, hoyuelos, un pliegue bajo el párpado inferior y un pliegue profundo naso labial desde la nariz a la barbilla.

Algunas variaciones de la expresión de la alegría pueden ser risa desternillante, risa, sonrisa con la boca abierta, sonrisa ansiosa, sonrisa melancólica, sonrisa tímida, sonrisa perversa, falsa sonrisa y falsa risa.

- Tristeza

La tristeza es un sentimiento de sufrimiento. Está generado por la pérdida de algún ser u objeto apreciado o una situación favorable.

Se caracteriza por la falta de motivación que produce un decaimiento en la atención y concentración. Al contrario que en el resto de expresiones, la apariencia del rostro durante un sentimiento de tristeza es menos distintivo cuanto mayor intensidad tiene la emoción.

En la expresión pura de tristeza, la zona interior de las cejas se curva hacia arriba. La piel y los tejidos blandos debajo de la ceja se agolpan sobre el párpado superior. Los ojos se cierran ligeramente a consecuencia de la presión hacia debajo de los tejidos sobre el párpado y también por el movimiento hacia arriba del párpado inferior. La boca se encuentra relajada. Las arrugas asociadas a la tristeza incluyen pliegues horizontales en la frente, arrugas verticales entre las cejas, pliegues oblicuos sobre el párpado superior, y un pliegue en forma de sonrisa bajo el labio inferior.

Algunas variaciones de la expresión de la tristeza son lloro con boca abierta o cerrada, tristeza reprimida, casi lloro e infelicidad. Estas variaciones pueden conllevar cejas completamente bajadas, ojos fuertemente cerrados, una protuberancia en la barbilla y un pliegue naso labial muy pronunciado.

- Enfado

El enfado se produce como resultado de situaciones que producen frustración y hostigamiento. Interrumpe los procesos cognitivos en curso, centrando la atención en el agente que la instiga. El enfado provoca que aumente la presión arterial y el enrojecimiento del rostro.

En la expresión pura de enfado, los extremos de las cejas se empujan hacia abajo y se juntan. El extremo inferior de la ceja se pone al mismo nivel que el párpado elevado. Los ojos están medio abiertos, pero la presión de la frente impide que se vea el blanco del ojo por encima del iris. La boca se encuentra cerrada con el labio superior ligeramente comprimido. Las arrugas en el enfado incluyen pliegues horizontales sobre los párpados superiores y líneas verticales entre las cejas.

Algunas variaciones posibles de enfado son irritación, descontento, rabia y tensión. Implican labios ligeramente presionados con protuberancia en la barbilla o la boca totalmente abierta con una mueca del labio superior con el labio inferior recto, mostrando ligeramente los dientes superiores e inferiores.

- Miedo.

El miedo es una señal emocional de advertencia ante una amenaza o peligro presente e inminente. Dicho peligro puede ser real o imaginado. Produce ansiedad, incertidumbre e inseguridad, por lo que el ser humano tiende a protegerse de esta emoción.

La expresión de miedo se caracteriza por estar las cejas elevadas y juntas. La parte inferior de las cejas se curvan hacia arriba. La boca puede estar ligeramente abierta y caída, estando estirada en ambos lados. Las arrugas asociadas pueden incluir pliegues horizontales en la frente, líneas verticales entre los ojos, hoyuelos sobre las cejas y pliegues oblicuos sobre los párpados superiores.

La expresión de miedo puede variar desde la preocupación al terror. En la preocupación, los labios se fruncen fuertemente y los bordes de los labios desaparecen. Se abulta la parte por debajo del labio inferior y sobre la barbilla. En la expresión de terror, tanto la boca como los ojos se encuentran totalmente abiertos. El labio superior se relaja mientras que el inferior se estira

hasta mostrar los dientes inferiores. El pliegue naso labial se vuelve recto y poco profundo y aparecen pliegues en forma de paréntesis a los lados del labio inferior.

- Sorpresa.

Es la emoción más breve y generalmente no prevalece en el tiempo, a menos que el elemento que origina la sorpresa revele nuevos elementos sorprendentes.

En la sorpresa, las cejas se elevan rectas tanto como es posible. Los párpados superiores se abren al máximo con los inferiores relajados, La boca se abre totalmente formando un óvalo. Se forman pliegues horizontales en la frente.

La sensación de sorpresa incrementa la sensibilidad de los órganos sensoriales permitiendo una evaluación rápida e inconsciente de la situación

- Disgusto.

Se trata de una sensación de aversión hacia el elemento que origina la emoción. Es una emoción fuertemente influenciada por el entorno social y cultural.

En la expresión de disgusto, las cejas se relajan y los párpados están relajados o sólo ligeramente cerrados. El labio superior se eleva en una mueca, a menudo asimétrica. El labio inferior se relaja y el pliegue naso labial es más profundo a lo largo de la nariz.

Esta expresión varía desde el desdén al rechazo físico. En el desdén, los párpados pueden estar parcialmente cerrados con los ojos mirando hacia abajo. En el rechazo físico, las cejas se bajan especialmente en los extremos interiores. El labio superior se eleva en una mueca intensa mostrando los dientes superiores. Aparecen arrugas en las patas de gallo y por debajo del párpado inferior y desde el lagrimal hasta el puente de la nariz y una protuberancia en la barbilla.

#### **4.5.2. EMFACS**

Las expresiones complejas pueden codificarse como combinación de un grupo reducido de unidades de acción en las que se asocia a cada unidad de acción un nivel de intensidad.

El sistema EMFACS es un método para emplear el Sistema de codificación facial (FACS) en el que se analiza la combinación de acciones faciales que son relevantes en la generación de emociones.

EMFACS es un acrónimo de FACS Emotion. Ekman y Friesen desarrollaron el sistema EMFACS durante la década de los 80 como un método para reducir el tiempo de puntuación al analizar únicamente las señales de emoción en el rostro. El método EMFACS se basa en las FACS y requiere el conocimiento y capacidad de identificar la combinación de las unidades de acción que se generan en las distintas regiones del rostro implicadas en cualquier movimiento facial y analiza la intensidad de su acción y el grado de simetría del rostro en la emoción. El sistema EMFACS sin embargo no tiene en cuenta aquellas unidades de acción cuyo efecto no afecta a la interpretación de la emoción o no exista evidencia de su efecto en esta. El EMFACS no se debe aplicar en aquellos casos donde las expresiones emocionales aparecen mezcladas o aparecen enmascaradas o sutilmente.

Este método emplea los estándares de FACS, y como este, no tiene en cuenta todas las AUs involucradas, solo las que generan la expresión de emoción. También se debe tener en cuenta que el análisis de las expresiones es desde un punto de vista descriptivo, e identifica las

unidades de acción que se activan en una determinada expresión mediante la observación y analiza los mismos aspectos con los mismos patrones que en las FACS: la intensidad de la AU dentro de su banda de 5 niveles, su frecuencia, duración y valencia.

En la *Tabla 4.3* vienen recogidas algunas expresiones y las AUs implicadas en la generación de éstas.

Appraisal dimensions	Hot Anger	AUs	Elation	AUs
Suddenness	high	AU1b + AU2b + AU25	high/medium	AU1b + AU2b + AU25
Familiarity	low	AU4a + AU7		
Predictability	low	AU4b	low	
Intrinsic Pleasantness			high	AU6 + 12b
Concern relevance	high		high	
Outcome probability	very high		very high	
Expectation	low		low	
Conduciveness	obstruct	AU4b + AU7b + AU17b + AU23b	very high	AU6 + AU12d
Urgency	high	intensification, high tension	low	
Cause agent	other			
Cause motive	intent		chance/intent	
Control	high			
Power	high	AU17c + AU24		
Adjustment	high		medium	
External norms	low	AU10c	high	
Internal norms	low		high	

*Tabla 4.3 – Algunas expresiones y AUs implicadas en ellas*

#### **4.6. Elección de las unidades de acción a representar.**

El rostro humano es una de las regiones más complejas de la anatomía humana, tanto por el número de músculos que presenta como por la cantidad de combinaciones de movimientos que dichos músculos pueden realizar.

La integración de detalles y características anatómicas del rostro a un modelo animado resulta una tarea complicada, ya que a medida que el modelo es más complejo, es más difícil ajustar cada uno de los elementos introducidos para conseguir una animación óptima del modelo.

Actualmente no existe ningún sistema de animación facial que opere en tiempo real capaz de simular todos los movimientos de los músculos faciales, y en todas las propuestas se realizan simplificaciones que permiten reducir tanto el número de componentes faciales implicados como el comportamiento de éstos [Ma11]. El objetivo de este método es conseguir resultados realistas y visualmente efectivos, es decir, que proporcionen al usuario la misma cantidad de información visual que al interactuar un rostro real. Para ello se seleccionan aquellos músculos cuyo movimiento originen deformaciones fácilmente reconocibles por la mayoría de los usuarios.

Marc Fabri [Fa06] presentó en su tesis de avatares en entornos colaborativos una reducción del número de movimientos faciales que un avatar debe emular para conseguir mostrar las seis expresiones emocionales universales: disgusto, tristeza, enfado, alegría, miedo y sorpresa. Fabri fundamentó su estudio en los trabajos de Ekman relativos a las expresiones emocionales derivados del FACS [Ek75], en el que estudió el reconocimiento de las diferentes expresiones básicas de emoción y determinó las unidades de acción más relevantes activadas en cada emoción. Para cada una de las seis categorías de emoción, Fabri seleccionó para cada emoción dos expresiones significativas que presentaban una mayor tasa de reconocimiento por parte de los observadores, y a partir de ellas determinó las unidades de acción presentes en dichas expresiones. También minimizó el número de unidades de acción seleccionando únicamente aquellas que tienen mayor influencia en la percepción visual de las expresiones, reduciendo el número de AUs a 17, presentes en la *Tabla 4.4*.

<b>NÚMERO AU EN EL FACS</b>	<b>NOMBRE ORIGINAL</b>	<b>MÚSCULOS QUE INTERVIENEN</b>	<b>EXPRESIONES EN LAS QUE SE ACTIVA</b>
<b>AU1</b>	<i>Inner Brow Raiser</i> / Levantar el interior de las cejas	Frontal (par medial)	Tristeza, Sorpresa, Miedo
<b>AU2</b>	<i>Outer Brow Raiser</i> / Levantar el exterior de las cejas	Frontal (par lateral)	Sorpresa, Miedo
<b>AU4</b>	<i>Brow Lowerer</i> / Bajar las cejas	Corrugador superciliar, Depresor superciliar, Próceros	Tristeza, Enfado, Miedo, Disgusto
<b>AU5</b>	<i>Upper Lid Raiser</i> / Levantar el párpado superior	Elevador del párpado superior	Tristeza, Enfado, Miedo, Disgusto
<b>AU6</b>	<i>Cheek Raiser</i> / Levantar las mejillas	Orbicular del ojo (par orbital)	Alegría
<b>AU7</b>	<i>Lid Tightener</i> / Estrechar los párpados	Orbicular del ojo (par parpebral)	Enfado, Disgusto
<b>AU9</b>	<i>Nose Wrinkler</i> / Arrugar la nariz	Elevador del labio superior	Disgusto
<b>AU10</b>	<i>Upper Lip Raiser</i> / Elevar el labio superior	Elevador del labio superior, alar de la nariz	Miedo, Disgusto
<b>AU11</b>	<i>Nasolabial Furrow Deepener</i> / Elevar el labio superior	Elevador del labio superior	Tristeza
<b>AU12</b>	<i>Lip Corner Puller</i> / Separar las comisuras de los labios	Cigomático mayor	Alegría
<b>AU15</b>	<i>Lip Corner Depressor</i> / Bajar las comisuras de los labios	Triangular	Enfado
<b>AU17</b>	<i>Chin Raiser</i> / Levantar la barbilla	Borla	Enfado, Disgusto
<b>AU24</b>	<i>Lip Pressor</i> / Apretar los labios	Orbicular de la boca	Enfado
<b>AU25</b>	<i>Lips Part</i> / Separar los labios	Orbicular de la boca, borla	Alegría, Tristeza, Miedo
<b>AU26</b>	<i>Jaw Drop</i> / Bajar la mandíbula	Masetero, temporal, Pterigoideo interno	Alegría, Tristeza, Sorpresa, Miedo

<b>AU29</b>	<i>Jaw Thrust</i> / Apretar los labios	Masetero, temporal	Enfado
<b>AU38</b>	<i>Nostril Dilator</i> / Dilatar las fosas nasales	Alar de la nariz	Tristeza, Miedo

Tabla 4.4 – AUs más significativas

En su estudio, Fabri mostró una buena tasa de reconocimiento por parte de los usuarios de las seis expresiones emocionales básicas generadas con su avatar. Sin embargo no se tuvo en cuenta la posición de la cabeza o de los ojos al representar emociones, ya que a pesar de utilizar un modelo 3D animado, empleó imágenes estáticas del avatar para el reconocimiento. Es por este motivo por lo que se ha añadido a esas 17 unidades de acción ocho unidades de acción más a las consideradas por Fabri durante la generación de expresiones, relacionadas con el movimiento de los ojos y el cuello.

Dichas unidades de acción generalmente son obviadas en los desarrollos de interfaces antropomórficos, pero tal y como demuestran diferentes estudios [Ad03b], [Ad05], [He07], resultan determinantes para una correcta comunicación y percepción de las distintas expresiones emocionales. También se ha incluido la unidad de acción relacionada con el parpadeo de los ojos, ya que el parpadeo otorga al avatar una sensación de naturalidad aun cuando se encuentra en posición de reposo. Por tanto, al conjunto de unidades consideradas en [Fa06] como esenciales, se han añadido otras 9 unidades de acción más, presentes en la *tabla 4.5*.

Para la animación del modelo generado se tiene en cuenta tanto la posición como la acción de los músculos que intervienen en la activación de cada una de las unidades de acción. Además, con el modelo se pretende emular la acción muscular facial tanto desde el punto de vista del resultado visual externo, como desde el punto de vista anatómico y fisonómico.

<b>NÚMERO AU EN EL FACS</b>	<b>NOMBRE ORIGINAL</b>	<b>MÚSCULOS QUE INTERVIENEN</b>
<b>AU45</b>	<i>Blink</i> / Parpadeo	Orbicular del ojo (par parpebral)
<b>AU51</b>	<i>Head Turn Left</i> / Girar el cuello a la izquierda	Conjunto de músculos del cuello
<b>AU52</b>	<i>Head Turn Right</i> / Girar el cuello a la derecha	Conjunto de músculos del cuello
<b>AU53</b>	<i>Head Up</i> / Girar el cuello hacia arriba	Conjunto de músculos del cuello
<b>AU54</b>	<i>Head Down</i> / Girar el cuello hacia abajo	Conjunto de músculos del cuello
<b>AU61</b>	<i>Eyes Turn Left</i> / Mirar hacia la izquierda	Músculos extraoculares
<b>AU62</b>	<i>Eyes Turn Right</i> / Mirar hacia la derecha	Músculos extraoculares
<b>AU63</b>	<i>Eyes Up</i> / Mirar hacia arriba	Músculos extraoculares
<b>AU64</b>	<i>Eyes Down</i> / Mirar hacia abajo	Músculos extraoculares

Tabla 4.5 – AUs añadidas relativas al movimiento de ojos y cuello

Con estas unidades de acción adicionales se pretende que las expresiones sean más fácilmente reconocibles e incrementen el realismo al generar las expresiones y la sensación de naturalidad al interactuar con el avatar.

# Capítulo 5

## *Creación de un avatar animado*

### **5.1. Introducción**

En el presente capítulo se describirá con detalle todo el proceso seguido para la generación de un asistente virtual realista animado con alto grado de realismo, tanto en apariencia como en comportamiento.

Resulta obvio que el diseño de un avatar que presente similitudes con el rostro humano requiera de un cierto conocimiento tanto de la anatomía facial del ser humano como de los aspectos psicológicos asociados a la expresividad del rostro. Por ello a la hora de generar el avatar se ha buscado la convergencia entre características propias de la animación 3D de personajes y los fundamentos de la anatomía y fisiología humanas.

La construcción de un avatar consta de dos partes bien diferenciadas: el modelado, que consiste en la obtención de los objetos tridimensionales con los que se va a trabajar, es decir, la geometría y texturizado del rostro, y la animación, que define los movimientos que el objeto realizará, en este caso, los movimientos gestuales y emocionales que el rostro llevará a cabo.

El modelado gráfico de personajes realistas 3D comprende la obtención de datos tridimensionales de un rostro real y la creación de una superficie a la que mediante técnicas de pegado de texturas se dota a la superficie de información cromática logrando así un rostro realista [Ig04]. Al modelo obtenido se le añaden ciertos elementos que no es posible adquirir de forma directa: ojos, dientes o cavidad bucal.

El proceso de animación consiste en adecuar la superficie poligonal modelada anteriormente mediante una de las distintas técnicas de animación, de las que se hablará posteriormente, y poder controlar la animación y emular el comportamiento muscular real. Para ello será necesario definir los frames clave que formarán las distintas deformaciones del rostro en función de la acción muscular real y el sistema de codificación facial o FACS.

## **5.2. *El proceso de modelado de un avatar realista***

La primera etapa del proceso, consiste en la construcción de la escena en la que se va a trabajar, y el modelado o creación de los objetos tridimensionales que la forman.

Las escenas de animación constan de los objetos tridimensionales modelados, las fuentes de luz y las cámaras. Este proceso engloba la creación de una superficie 3D teniendo presente la anatomía facial del rostro humano y las relaciones morfológicas de los músculos implicados en la expresión facial, también, la obtención de datos tridimensionales de un rostro real para aportar a la superficie información cromática y dotar al avatar de un aspecto realista aplicando técnicas de texturizado, y añadir aquellos elementos que no se pueden adquirir de forma directa y requieren un procesado diferente, como son los ojos, los dientes o la cavidad bucal.

Los objetos tridimensionales son aquellos objetos que conforman la escena, las fuentes de luz se emplean para el renderizado de las imágenes, determinando el aspecto final de las texturas en las distintas superficies. Y las cámaras, permite especificar la región de la escena que será renderizada (clipping).

## **5.3. *Adquisición de datos a partir de un rostro real***

El primer paso para conseguir un buen modelado facial, es tener buenas referencias para modelar. Existen diferentes técnicas para obtener las coordenadas tridimensionales del rostro como escáneres digitales o digitalizadores, pero dichos procesos no son óptimos para la animación al carecer de información de la estructura facial, el cálculo de ruido produce errores en el objeto y los vértices están pobremente distribuidos. Además, muchos métodos de obtención de datos producen modelos incompletos, sin pelo, orejas, ojos...

También se puede capturar el modelo mediante técnicas de fotogrametría. Para obtener una buena referencia de los puntos, interesa captar numerosas fotos a partir de la vista frontal de la cara y realizar fotografías del modelo facial cada 45 grados. También hay que tener especial cuidado con la iluminación del lugar donde se adquieren los datos, se tiene que buscar un equilibrio entre que sea óptima para iluminar la escena, pero no genere brillos en el modelo, ya que pueden hacer que la apariencia del avatar pierda realismo.

Se optó por emplear esta técnica, ya que resulta una técnica viable al no suponer costes, y se consiguen extraer las características del rostro del modelo de forma sencilla con una calidad suficiente para su aplicación. Se empleó para ello una cámara Nikon D90, una cámara reflex digital semiprofesional de una única lente y autofocus, con 12.3 megapíxeles de resolución que cuenta con un sensor de imagen CMOS de formato DX y rendimiento de ruido reducido de ISO 200 a 3200.

A partir de diferentes poses de la modelo tanto de frente como de lateral o poses intermedias, se obtuvieron las características necesarias para generar el modelo, se extrajo información sobre las coordenadas tridimensionales del rostro, las características cromáticas necesarias para la textura y la información necesaria para añadir aquellos elementos que generan el modelo, como los ojos, los dientes o la cavidad bucal.

## 5.4. Ajuste del modelo

### 5.4.1. Modelado de la superficie

Para realizar el modelado de la superficie de un rostro, se tiene que tener presente la fisiología del rostro humano. Los tres principales componentes de la cabeza humana son la piel, la estructura ósea y la estructura muscular del rostro; aunque en la generación de expresiones, juegan un papel transcendental los principales órganos visibles: los ojos, los dientes y la lengua.

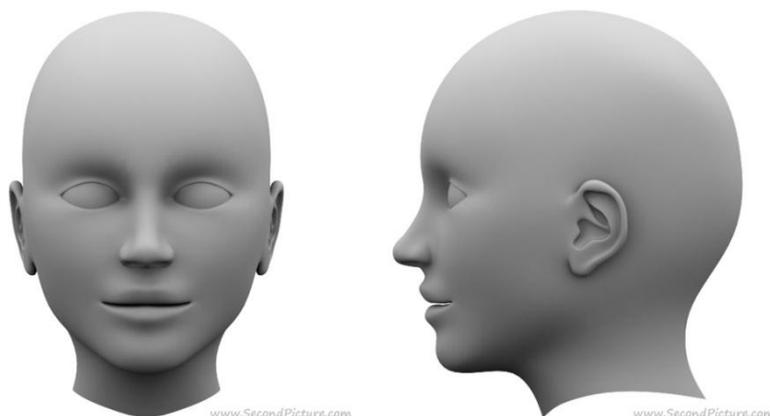
Un rasgo destacable de Blender es la comunidad global en torno a este, donde, como se indicó anteriormente, sus usuarios comparten sus trabajos o ponen a disposición del resto sus conocimientos sobre esta herramienta. Así, en internet se pueden encontrar modelos genéricos creados o tutoriales que muestran paso a paso cómo realizarlos.

Los datos que constituyen el modelo 3D pueden obtenerse modelando de forma manual la superficie a representar. Hoy en día, mejorando el tiempo de producción y la calidad final del modelo, se aconseja trabajar en un modelo de baja poligonalización y trabajar los detalles tales como surcos, arrugas, bultos, etc. Empleando este método se consigue que la superficie del modelo sea poco poligonalizada, esto resulta conveniente también al generar la animación ya que ésta resultará más sencilla de realizar y tendrá un menor gasto computacional.

El modelado manual permite simplificar esta etapa del proceso, y ofrece la libertad de generar la superficie a voluntad del usuario buscando un equilibrio entre simplicidad y definición, pero hay ciertos detalles como son la cavidad bucal, dientes u ojos que se agregan al modelo y se procesan de manera aislada a este proceso de modelado de la superficie genérica del modelo. Sin embargo, este método puede resultar muy tedioso y extenso sobre todo para los usuarios básicos de Blender. Además, la superficie que se obtiene puede ser inexacta o resultar muy tosca.

Para realizar la superficie es necesario tener en cuenta las proporciones antropométricas de la cara a modelar. En cuanto a los tipos de geometría a modelar, se incluyen vectores, polígonos, subdivisión de superficies y mallas poligonales.

Para ello, se construyó la malla mediante puntos utilizando de soporte dos imágenes base con las vistas 2D frontal y de perfil de un modelo genérico (*Figura 5.1*), y el objeto 3D se obtiene uniendo las operaciones de ajuste de cada vista 2D (*Figura 5.2*).



*Figura 5.1 – Imágenes genéricas usadas para la construcción del modelo*

Es posible personalizar el entorno de trabajo de Blender y agregar tantas ventanas como sean necesarias, además de poder emplear en cada ventana diferentes imágenes de fondo como soporte para el modelado (*Figura 5.3*).

El proceso de modelado de la superficie se realiza en el modo edición, en la ventana de Vista 3D. La geometría de una escena se construye a partir de uno o más objetos, como una malla plana de partida. Esta malla plana contiene cuatro vértices, cuatro aristas y una cara. En las mallas básicas, todo se construye a partir de tres estructuras básicas: vértices, aristas y caras. Estos pueden deformarse mediante transformaciones básicas como la rotación, translación o escalado seleccionando uno o más puntos de la malla.

Trabajando con la vista frontal y partiendo de un polígono plano, se ajusta su tamaño y posición y agregando nuevos vértices se generan nuevas caras a la superficie, creando así una malla siguiendo la forma del rostro. Los puntos que definen la cara se unen formando triángulos o cuadrados, cuantos más puntos, más polígonos y mayor definición, aunque una cantidad excesiva de vértices puede resultar inadecuada, ya que puede hacer que las animaciones generadas resulten poco realistas y dificultar esta etapa, pudiendo suponer una labor muy tediosa. Para evitar aumentar el gasto computacional, se optó por no modelar aquellas regiones que iban a permanecer ocultas o en estas regiones la malla estará constituida por el menor número de polígonos posibles, como la parte posterior de la cabeza y la nuca.

Aprovechando las herramientas que Blender ofrece y la simetría del rostro, se modeló la mitad del rostro y mediante herramientas de duplicado simétrico se proyectó la otra mitad, simplificando la labor y evitando así cometer errores al tener los vértices a la misma distancia del eje y en ambos lados del rostro.

Para completar el modelo, se ajustó la vista de perfil otorgando al modelo profundidad y, por consiguiente, obteniendo un objeto 3D con la superficie del rostro que más adelante se ajustará al rostro de la modelo (*Figura 5.4*).

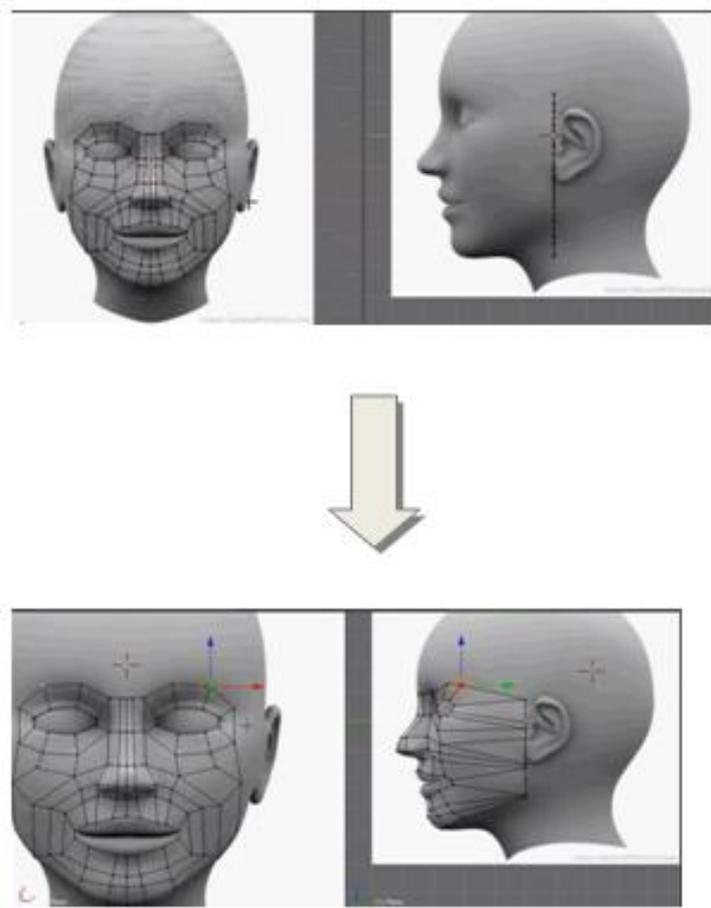


Figura 5.2 – Proceso de creación del modelo 3D

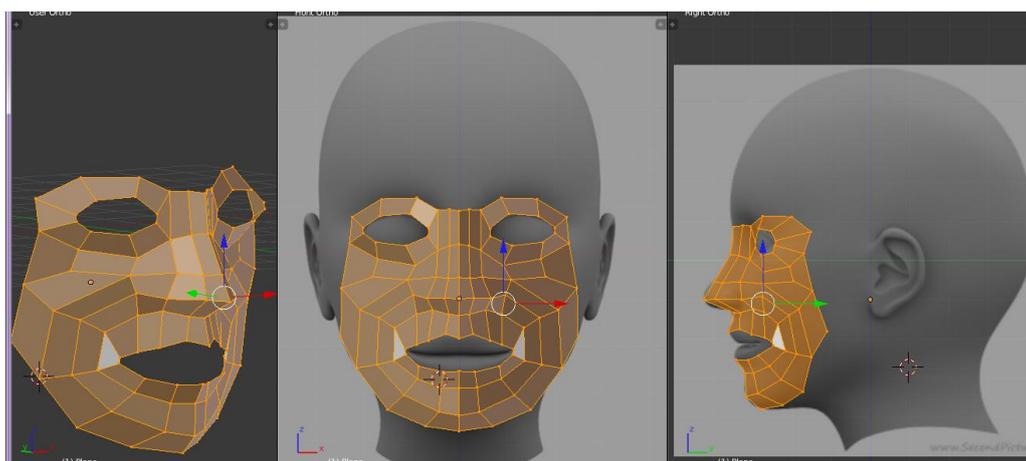


Figura 5.3 – Modelo 3D de la cara

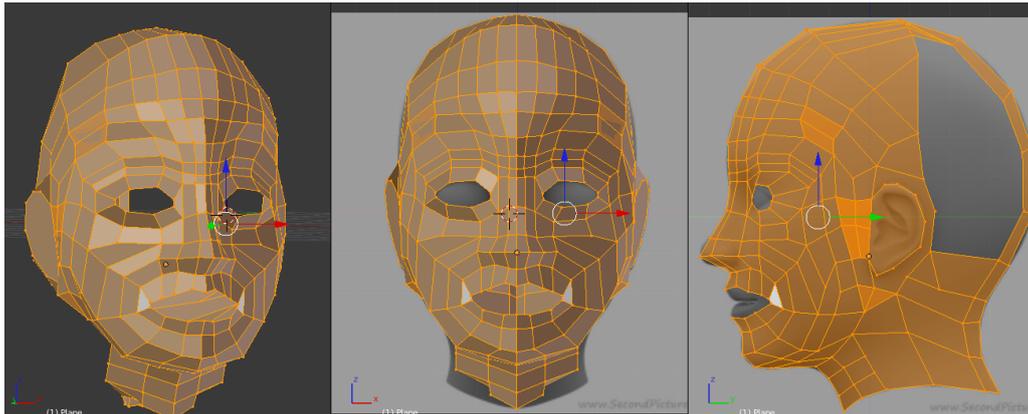


Figura 5.4 – Resultado final del modelo 3D

### 5.4.2. Ajuste de superficie genérica a modelo

Partiendo del modelo genérico, a continuación se procederá a ajustar la superficie al rostro de la modelo. Para ello, igual que se generó el modelo base, se toma de soporte dos imágenes con la vista frontal y perfil del rostro a modelar (Figura 5.5), y se ajusta el tamaño y posición de los vértices siguiendo la forma del rostro.

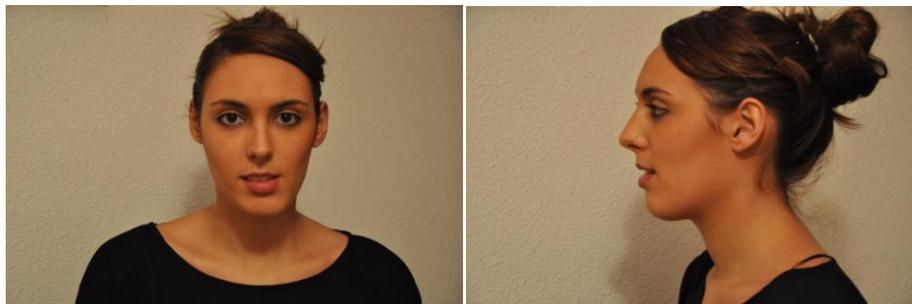


Figura 5.5 – Imágenes del modelo real

El ajuste de la superficie al modelo debe realizarse minuciosamente (Figura 5.6), ya que a esta superficie posteriormente se le agregará la textura dotando al modelo de color y realismo, y las animaciones se generarán modificando dicha superficie.

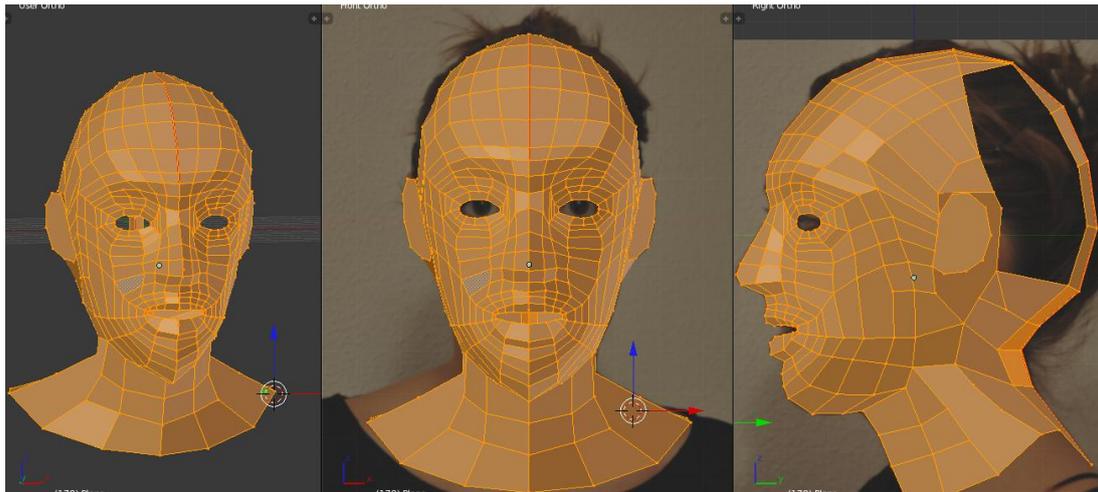


Figura 5.6 – Ajuste del modelo genérico al real

Con referencia a la superficie realizada, se ha intentado simplificar el número de vértices o polígonos que la conforman, para dotar de mayor realismo al modelo y, a la hora de hacer las animaciones, evitar emplear vértices innecesarios que podrían dificultar la animación y que incrementaría el coste computacional.

Aun así, el número de vértices de la malla es mayor en aquellas zonas que requieren mayor detalle al ser zonas más determinantes a la hora de generar expresiones (Figura 5.7), como puede ser ojos, cejas o boca. También, es importante la disposición de los vértices en la malla, ya que debe estar en consonancia con el movimiento de la región donde se encuentra dicho punto. Por ejemplo, en las zonas de la boca y los ojos éstos están distribuidos formando una elipse de modo que el movimiento que se forma en esas zonas se propague a lo largo de los puntos adyacentes, recreando el movimiento real de los músculos del rostro.

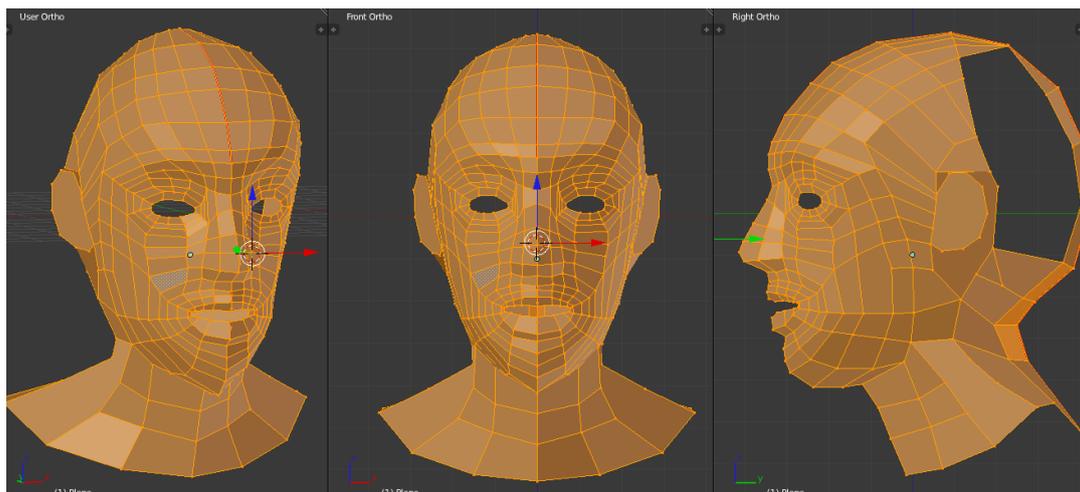
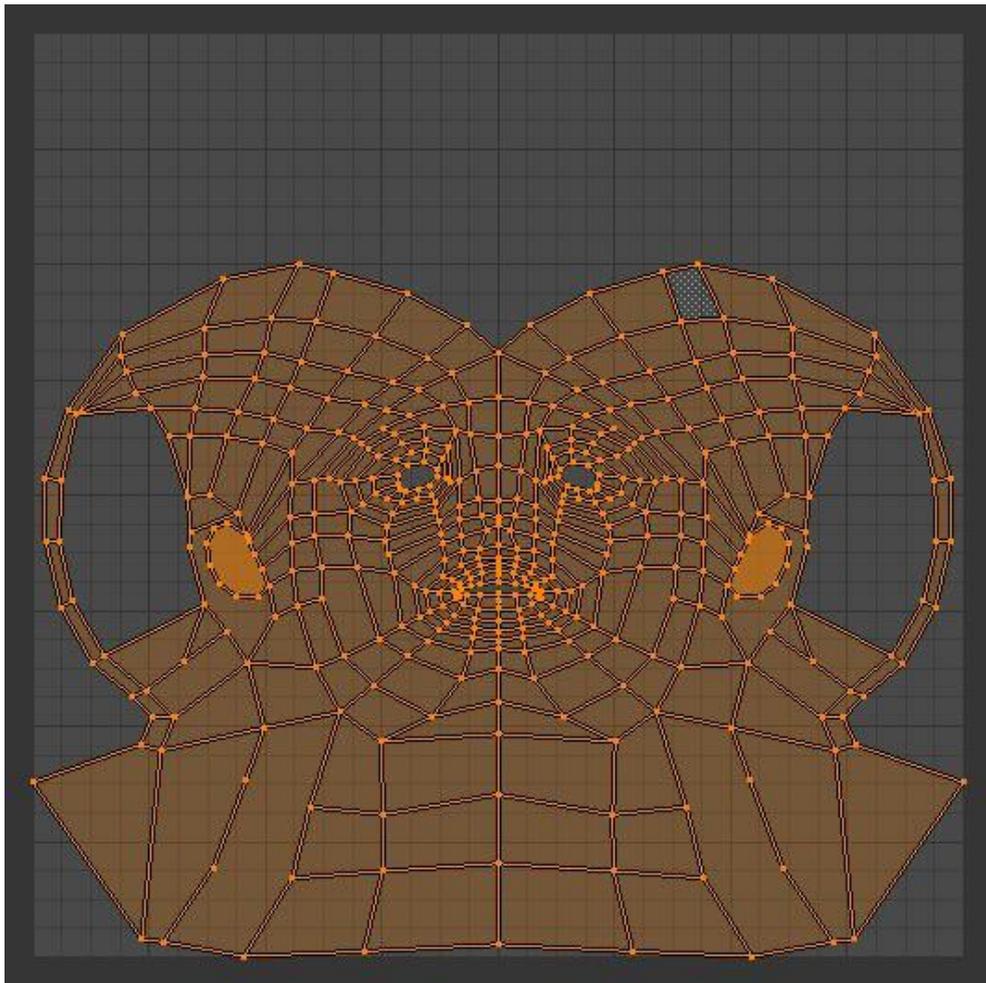


Figura 5.7 – Modelo 3D ajustado

## 5.5. Texturizado de la superficie

La fase de texturizado es tan importante como la de modelado, sobre todo si lo que se busca es realismo.

Para añadir información colorimétrica al modelo, se realiza el mapeo de imágenes sobre las superficies del modelo, donde cada polígono puede tener una imagen o una región de una imagen asignada. Esto se realiza por medio del mapeo UV, donde se toma la malla tridimensional (X,Y,Z) y se desenvuelve en una imagen plana bidimensional asignando a cada vértice de la malla un par de coordenadas 2D (U,V) que definen la región que es mapeada y a la que se asigna una región de color (*Figura 5.8*).



*Figura 5.8 – Modelo desenvuelto mediante una cremallera*

La operación de la generación de estos mapas UV también se denomina “desenvolver”, o unwrap en inglés, ya que es como si la malla se desplegara en un plano 2D. Al desenvolver una superficie en una textura UV en el editor de imágenes, a cada cara de la malla se le asigna automáticamente unas coordenadas UV que definen cómo se proyecta una imagen o una textura sobre la superficie.

El proceso de unwrap es sencillo, pero hay infinitas opciones disponibles, cada una con un efecto diferente del proceso. El proceso a seguir será el siguiente:

- 1) Marcar costuras si fuera necesario.
- 2) Seleccionar todos los componentes de la malla.
- 3) Seleccionar un método de mapeado UV de los que ofrece el menú de “Mapeado UV”.
- 4) Ajustar la configuración de unwrap.
- 5) Agregar una imagen de prueba para comprobar si se producen distorsiones.
- 6) Ajustar UVs en el editor UV.

### **5.5.1. Tipos de mapeado UV**

Para obtener modelos simples 3D, Blender tiene un conjunto de algoritmos automáticos de desenrollar que se pueden aplicar fácilmente. Los métodos de proyección más simples emplean algoritmos que interpolan la posición de los puntos del espacio 3D hacia los puntos/ejes a través de la superficie 2D. Los métodos más avanzados se pueden emplear con modelos más complejos, y tienen usos más específicos. Todas las técnicas descritas a continuación vienen recogidas en el menú “Mapeado UV” que se muestra en el Visor 3D seleccionando el objeto en Modo edición y cuyo efecto se muestra en el Editor de Imágenes/UV.

#### **5.5.1.1. Técnicas de proyección simples**

Existen cuatro tipos de proyecciones simples:

- Proyección cúbica  
El mapeado de la malla se realiza sobre las caras de un cubo que se despliega.
- Proyección esférica  
La proyección se realiza sobre una esfera. Esta técnica es solamente útil en objetos con formas esféricas, como ojos, planetas, etc.
- Proyección cilíndrica  
En este caso el objeto se desenrolla sobre una superficie cilíndrica.
- Proyección desde vista  
Toma la vista actual 3D del objeto y la proyecta como se muestra.

#### **5.5.1.2. Técnicas de proyección complejas:**

Y cuatro tipos de proyecciones más complejas:

- Desenrollar  
Resulta útil para las formas orgánicas. Desenrolla la malla en una superficie plana y suaviza la malla a lo largo de las costuras.
- Proyección UV inteligente  
Desenrolla la superficie en regiones basándose en un umbral de ángulos.
- Empaquetar mapa de luz  
Divide la superficie en función de las caras y las distribuye a lo largo de la cuadrícula UV.
- Seguir cuadrángulos activos  
Realiza un seguimiento de los cuadrángulos activos de la superficie mediante un bucle continuo.

### 5.5.2. *Desenvolver empleando costuras*

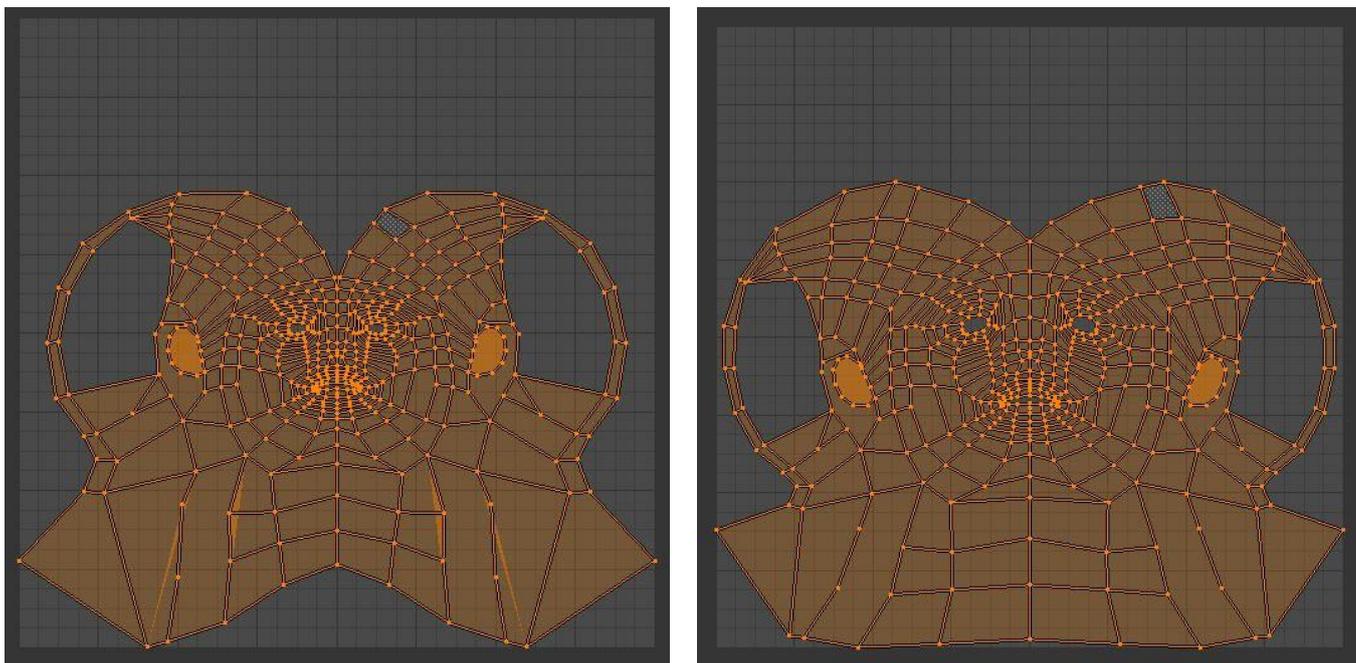
En muchos casos, las técnicas de desenvolver simples generarán un buen mapeado UV de la superficie del objeto, pero los modelos 3D más complejos requieren de un procesamiento previo como el empleo de costuras para guiar la asignación del mapeado UV y conseguir una proyección uniforme del modelo.

Las costuras limitan y guían el proceso de desenvoltura de la malla a lo largo de éstas. Hay que tener especial cuidado si se emplean un gran número de ellas, ya que a mayor costuras, menor capacidad de la malla de desenvolverse a lo largo del espacio UV.

El proceso a seguir será el siguiente:

- Crear las costuras. Las costuras se marcan desde la ventana Vista 3D en modo edición seleccionando las aristas donde se generarán las costuras empleando el comando Mark Seam.
- Desenvolver la superficie y proyectarla en el plano UV.
- Ajustar las costuras hasta obtener la superficie deseada en el plano UV.
- Ajuste manual UV.

Para el modelo, se empleó una costura que recorría desde la mitad de la frente hasta la nuca. El efecto del empleo de las costuras en el modelo se muestra en la *Figura 5.9*.



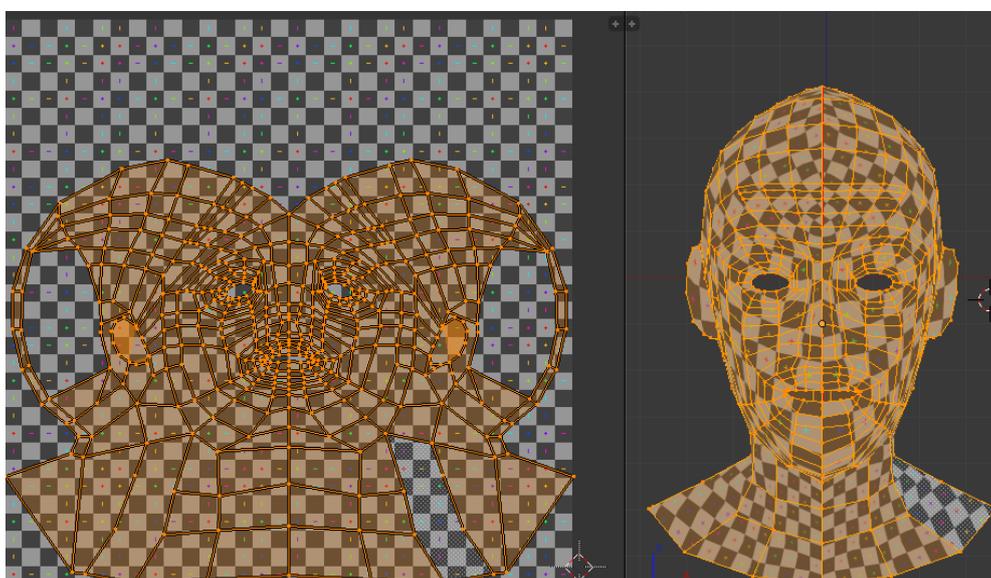
*Figura 5.9 – Efecto de la costura al desenvolver el modelo*

### 5.5.3. *Desenvoltura de la superficie y aplicación de texturas UV*

Desde la ventana Vista 3D en modo edición, se seleccionan todas las caras del objeto y se desenvuelve la superficie mediante el comando UV Unwrap.

Para poder ajustar la malla, Blender permite visualizar la distribución de las texturas sobre las regiones de la superficie y asignar texturas directamente a las caras de la malla. La vista 3D muestra el objeto que se está texturizando configurando esta pantalla en modo textura, y los cambios realizados se mostrarán en la ventana de editor de Imágenes/UV y viceversa.

Se puede también emplear una imagen de prueba de cuadrícula de verificación UV para comprobar la disposición de la textura en la superficie y ajustar la textura y modificarla si existen distorsiones en las caras, como se muestra en la *Figura 5.10*.



*Figura 5.10 – Uso de imagen en cuadrícula para el proceso de unwrap*

Los objetos creados en Blender se definen tanto por la forma que adopta su estructura, como por la apariencia física del objeto. Las características colorímetras y los efectos ópticos que poseen se determinan mediante los materiales y las texturas.

Un material define las propiedades ópticas de un objeto, su color y las cualidades del elemento que conforma el objeto, como el brillo, la refractancia, la transparencia o difracción de la luz sobre la superficie, etc. Las texturas modifican los materiales de diversas formas, como alterando sus colores.

Se pueden editar y cargar imágenes e incluso realizar animaciones al asignar texturas UV a los objetos sin asignarles a estos materiales. En la ventana Vista 3D en modo textura, se mostrará el objeto con la textura, sin embargo, cuando se renderiza la escena, el objeto aparecerá en color gris por defecto, o en negro si no se carga ninguna imagen para la textura. Para que al renderizar y animar la escena la apariencia del objeto se muestre correctamente al emplear una imagen como textura es necesario definir previamente un material cuya apariencia esté definida por esta. Para ello, se debe crear un material para el objeto y determinar la textura UV que se emplea en el objeto al renderizar.

Un material se crea añadiendo un nuevo material en el panel de Materiales del menú de propiedades. En el panel de Texturas, se añade una nueva textura y se define como una imagen. A continuación, se carga la imagen que se desea usar. En la sección de Mapping o Mapeo, hay que configurar como coordenadas de textura la opción UV y en Map o Mapa la capa UV que se quiere emplear.

La textura de un objeto se asigna mediante uno o más mapas UV. Las imágenes empleadas para definir la textura del avatar, son las vistas frontales y laterales de la modelo. Se generaron dos mapeados UV de la malla empleando la proyección desde vista de los dos perfiles y se ajustó la distribución de esta en el espacio UV haciendo corresponder de forma correcta la malla y las características del modelo. Además, se realizó un tercer mapeado desarrollando la superficie y se definió un material para almacenar los datos de la apariencia del objeto. Una vez definidos los mapas UV de la superficie, el modelo estará preparado para comenzar el proceso de texturizado.

Blender incorpora una herramienta específica para editar texturas UV e imágenes de una forma rápida y sencilla llamado “Pintar texturas”. Aun así si se desea, se pueden emplear programas de procesamiento de imágenes externos e importar la imagen a Blender. Una vez que se activa “Pintar texturas”, el ratón se convierte en un pincel y se muestran los ajustes en la pestaña de propiedades para definir el efecto que se desea crear con él. A medida que se arrastra el ratón por la superficie de la malla, se pinta y genera automáticamente la textura del objeto.

Blender tiene un conjunto de pinceles preestablecidos configurados para crear diferentes efectos al generar la textura. Con los pinceles el usuario puede añadir color a la textura pintándola, suavizarla o difuminarla, o generar la textura extrayendo los colores de una imagen especificada clonándola. Cuando se utiliza el pincel de clonación, se muestra la imagen de referencia de la cual se extraen los colores que conforman la textura. A medida que el usuario arrastra el ratón por la superficie del objeto en la ventana de visor 3D, se genera la textura del objeto a su paso, y los cambios que generan los pinceles en la textura se muestran automáticamente en la ventana de Editor de imágenes/UV, y viceversa. Sin embargo, la textura modificada no se guarda automáticamente, sino que se debe guardar como una imagen manualmente en el Editor de Imágenes.

También se puede editar la imagen empleando un programa de editor de imágenes como Photoshop siempre y cuando el archivo .blend se encuentre activo, se reedite y se vuelva a guardar memorizando los últimos cambios también en el Editor UV/Imagen. También se debe volver a guardar el archivo si se han asignado nuevas regiones a una imagen, y la vista 3D se actualizará con la última asignación realizada del objeto y la imagen. Si se mueve la imagen del directorio actual, Blender no la encontrará y se deberá reemplazar con el nuevo directorio origen. Por este motivo, Blender incorpora la opción de empaquetar la imagen, lo cual significa que la imagen se asociará al archivo .blend de manera intrínseca, sin tener en cuenta el directorio en el que se encuentra dicha imagen, o incluso si se dispone o no de ella.

Como se puede observar en la *Figura 5.11*, al clonar la textura empleando las imágenes frontal o lateral sobre el mapa UV de dichas vistas, la textura se ajusta perfectamente a la imagen, sin embargo, la apariencia es pésima al observar el modelo desde otra perspectiva.



Figura 5.11 – Resultado de texturizar usando sólo la imagen frontal (a) o lateral (b).

La textura del modelo se generará, pues, mediante una combinación de ambas imágenes (Figura 5.11). Combinar ambas imágenes en una misma textura puede resultar una ardua tarea especialmente en aquellas regiones donde las dos imágenes se mezclan, ya que las sombras que se generan debido a la iluminación producen que una misma región del rostro tenga diferente tonalidad. Por lo que el procesado en dichas zonas será más minucioso para conseguir que este contraste sea nulo y no se produzcan variaciones de la tonalidad del rostro y lograr que la apariencia del modelo sea realista.

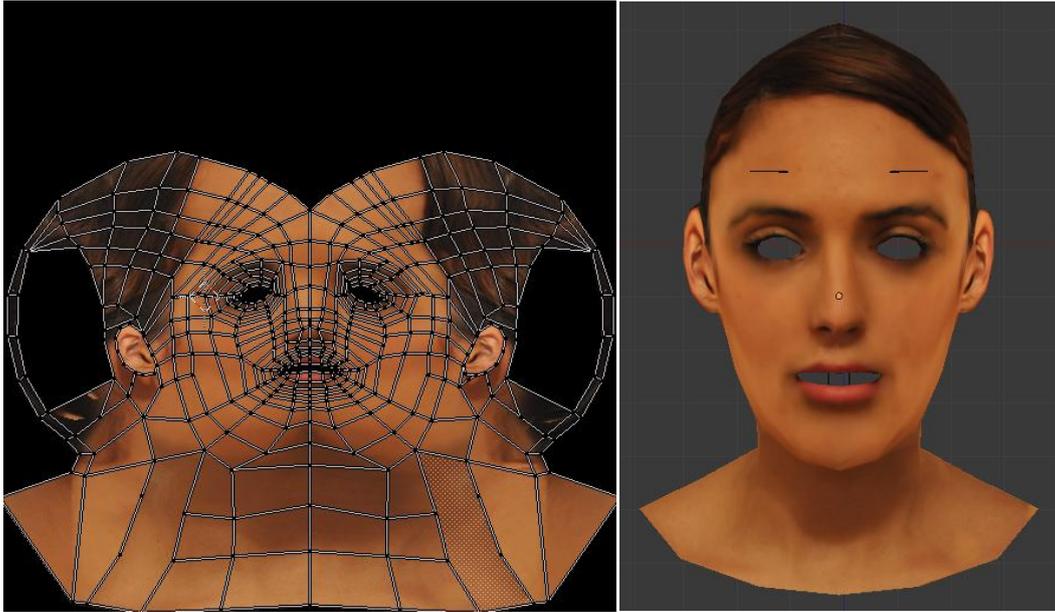


Figura 5.12 – Resultado de texturizar usando ambas imágenes.

### 5.6. Modelado de los ojos y la boca

Los datos que constituyen el modelo 3D pueden obtenerse modelando la superficie a representar de forma manual, pero hay ciertos detalles como son la cavidad bucal u ojos que se agregan al modelo y se procesan de manera aislada al proceso de modelado de la superficie genérica del modelo.

Para lograr unos ojos de aspecto realista y que requieran bajo coste computacional, se emplea una serie de técnicas en la construcción del modelo poligonal que permiten que el ojo interactúe con la luz aplicada a la escena de forma similar al ojo real.

El modelo consta de cuatro superficies como se puede observar en la *Figura 5.13*, donde la superficie azul situada más a la izquierda es la córnea. Su forma y textura hace que incluso una fuente de luz débil situada lejos en la escena aparezca reflejada sobre ella. La siguiente superficie corresponde al iris, que se modela mediante una superficie cóncava que origina que aparezca, tras el proceso de renderizado, un haz amplio de luz especular en el lado contrario siguiendo la dirección de la lámpara. Esto “esconde” la luz refractada por la superficie que representa la córnea dando la impresión de que el interior del ojo recoge luz y creando una sensación de profundidad. El iris se modela como una superficie circular oscura y el globo ocular como la superficie abierta de una esfera achatada.

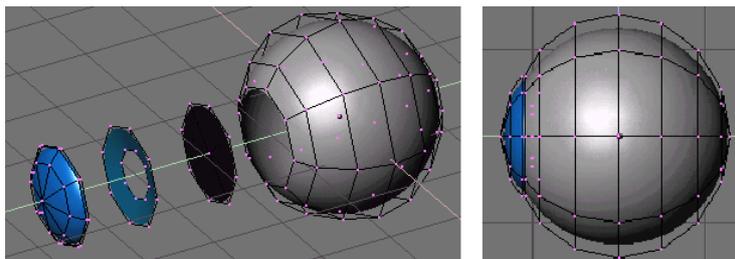
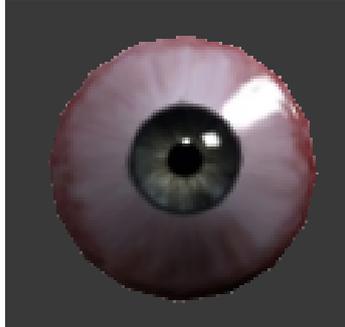


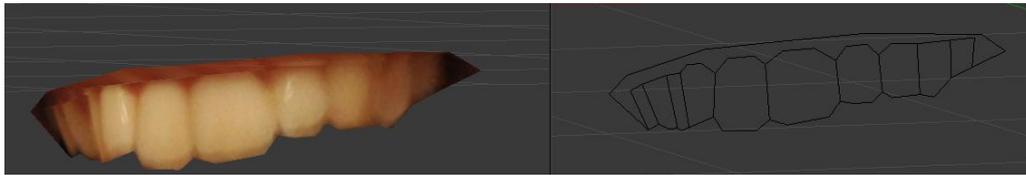
Figura 5.13 – Estructura del ojo.

La apariencia de los ojos se obtuvo mapeando una textura empleando una imagen (*Figura 5.14*). De esta manera, se independiza la información geométrica y colorimétrica en el modelo, pudiendo emplear diferentes imágenes para el texturizado de los ojos sin necesidad de modificar la geometría del modelo y conservando así los efectos creados para simular la reflexión de la luz en los ojos.

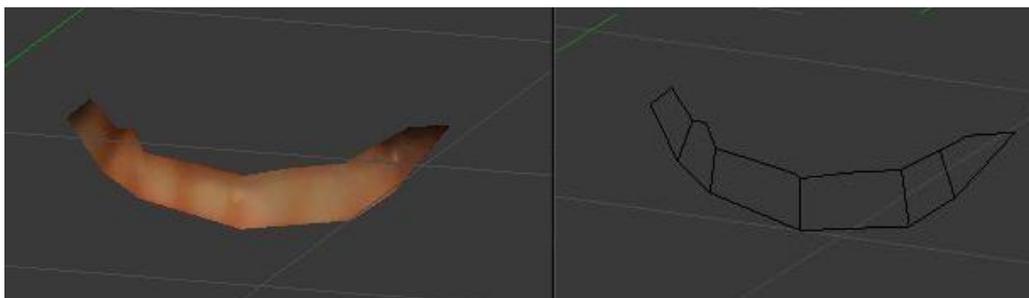


*Figura 5.14 –Resultado final del ojo.*

La cavidad bucal está constituida por los dientes superiores e inferiores, la lengua, y una superficie semiesférica que simula el interior de la boca (*Figuras 5.15, 5.16, 5.17 y 5.18*). Estas estructuras no requieren un modelado exhaustivo, ya que son estructuras rígidas que permanecen ocultas durante gran parte de la animación. Los dientes se modelaron mediante un plano curvado con pocos vértices, al igual que la lengua y la textura de ambos objetos se generó clonando las imágenes.



*Figura 5.15 –Modelado de los dientes superiores.*



*Figura 5.16 –Modelado de los dientes inferiores.*



*Figura 5.17 –Modelado de la lengua.*

El interior de la boca se ha elaborado simplemente para darle una apariencia realista al modelo empleando una semiesfera y pintándola de un tono oscuro.

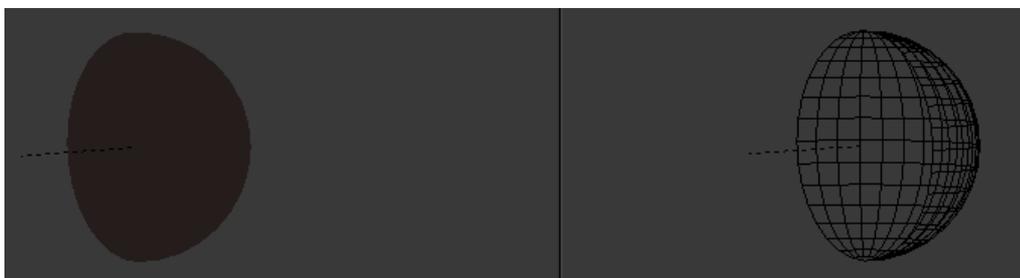


Figura 5.18 –Modelado del interior de la boca.

También, para lograr una apariencia más realista del modelo se ha añadido más cabello en forma de moño al modelo (Figura 5.19). Se trata de una estructura que permanecerá rígida durante la animación y no requiere de mucho detalle. Para ello se construyó una malla compuesta por pocos vértices, y la textura se obtuvo empleando las imágenes de la modelo.

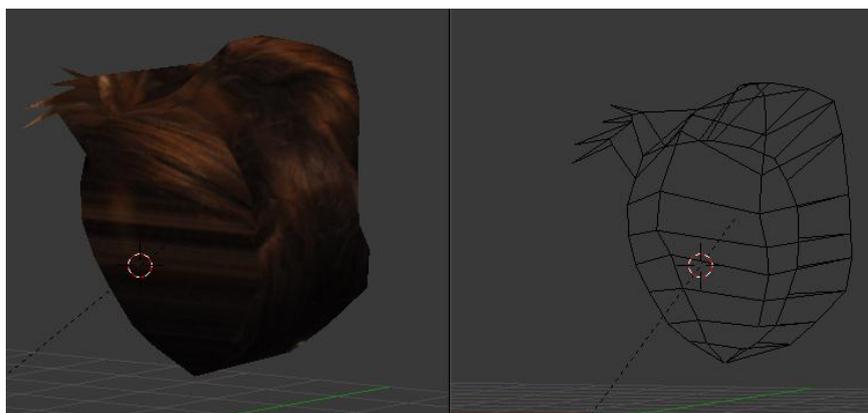


Figura 5.19 –Modelado del pelo.

### 5.7. Agrupación y emparentado de objetos

En una escena pueden existir muchos objetos: luces, cámara, varios objetos que conforman el mismo objeto, etc. Blender permite mantener todos los objetos organizados agrupando un conjunto de objetos en uno mismo. Cuando se modela un objeto complejo, se puede optar por modelar todo el conjunto o modelar las diferentes partes que forman el objeto por separados. Sin embargo, todas estas partes pueden estar unidas entre sí. En este caso, se debe designar un objeto como el objeto padre del resto de objetos hijos. De esta manera, los movimientos y rotaciones que realice el objeto padre afectará al movimiento de los objetos hijo, pero no al revés.

En la escena del modelo, existen varios objetos además de las luces y la cámara: el rostro, los ojos, el pelo, los dientes, la lengua y el interior de la boca. Para que todo el modelo se mueva en conjunto, se han agrupado todos los objetos en un mismo objeto. La Figura 5.20 muestra el resultado final.



*Figura 5.20 –Resultado final del modelado.*

### **5.8. Creación de la escena**

Para generar un asistente virtual realista animado con alto grado de realismo es tan importante conseguir un óptimo resultado del modelado del avatar como generar un entorno realista que lo englobe y para ello entran en juego las luces, las cámaras y la generación de una escena que dote de mayor realismo a la escena. Es por ello que se ha añadido a la escena una estancia para dotar al modelo de un entorno más realista (*Figura 5.21*).



*Figura 5.21 –Modelado de la estancia de la escena*

Además, para que el efecto de la escena completa fuera más realista, se le añadió al modelo ropa siguiendo el mismo proceso descrito para la creación de aquellos objetos que precisaban de un procesamiento manual como la cavidad bucal o los ojos descritos en el apartado anterior. (Figura 5.22)



Figura 5.22 –Modelado de la indumentaria.

### 5.8.1. Iluminación de la escena

La iluminación es un tema muy importante para la generación de la escena, tanto como el modelado, los materiales y las texturas. Una escena que cuente con un modelado y texturizado con gran precisión podría menoscabarse si no cuenta con un óptimo esquema de iluminación, al igual que un modelo simple puede llegar a ser muy realista si esta hábilmente iluminado.

Antes de exponer las diferentes fuentes de iluminación que ofrece Blender, se debe entender cómo interactúan las luces simuladas y las superficies con el motor de renderizado de Blender, y cómo los parámetros del material controlan éstas interacciones. La imagen renderizada que se crea con Blender es una proyección de la escena en una superficie imaginaria llamada plano de proyección (viewing plane). El plano de proyección es análogo a la película en una cámara tradicional, o a la córnea en un ojo humano, excepto que recibe luz simulada en lugar de luz real. Para renderizar una imagen de una escena primero debe determinar que luz de la escena está llegando a cada punto del plano de proyección. La mejor manera de responder a esta pregunta es seguir una línea recta (el rayo de la luz simulada) desde dicho punto en el plano de proyección y el punto focal (la posición de la cámara) hasta que impacte sobre una superficie representable en la escena, en este punto podremos determinar que luz debería impactar ese punto como muestra el esquema de la Figura 5.23. Las propiedades de la superficie y el ángulo de incidencia de la luz indican la cantidad de luz debería ser reflejada por el ángulo de visión incidente.

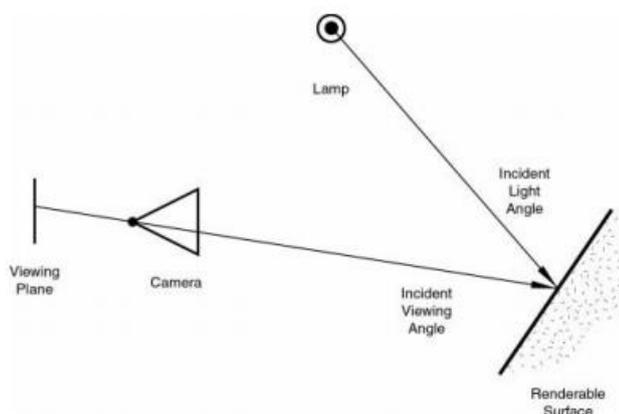


Figura 5.23 – Principio básico de motor de renderizado de Blender

Cuando un rayo de luz impacta sobre cualquier punto de una superficie, se dan lugar dos tipos básicos de fenómenos: difusión y reflexión especular. La difusión y la reflexión especular se distinguen la una de la otra principalmente por la relación entre el ángulo de luz incidente y el ángulo de la luz reflejada.

Blender cuenta con diferentes tipos de fuentes de luz los cuales generan diferentes efectos en la escena virtual. En el mundo real, hay una sola luz presente, ésta se ve modificada por saltos y se vuelve a irradiar en toda la escena por los objetos en los que rebota, haciendo sombras ligeras y zonas no completamente oscuras, pero iluminadas parcialmente.

Las físicas de las ondas lumínicas son simuladas en Blender por los renders de trazado de rayos (Ray Tracing) y pueden ser simuladas activando el motor "Radiosity" (chapter\_radiosity). Sin embargo, el trazado de rayos y la radiosidad son procesos lentos y suponen una gran carga computacional. Se puede conseguir un renderizado más rápido con su renderizado interno de escaneo de líneas (Scanline). Este tipo de motor de renderizado es mucho más rápido debido a que no intenta simular el comportamiento real de la luz, asumiendo muchas hipótesis que lo simplifican.

Blender proporciona cuatro tipos de luces:

- Sun Light- Fuente de luz Solar
- Hemi Light- Fuente de luz por hemisferio o área.
- Lamp Light- Fuente de luz focal
- Spot Light- Fuente de luz puntual

Al incorporar a la escena cualquier luz de las mencionadas, se pueden modificar los parámetros que las definen consiguiendo así diferentes efectos lumínicos en la escena.

### **5.8.1.1. Descripción de las fuentes lumínicas en Blender**

#### **5.8.1.1.1. Luz Solar (Sun Light)**

El tipo de luz más sencillo es probablemente la luz solar (Sun light)(Luz solar). Una luz solar tiene una intensidad constante que viene de una dirección dada. En la vista 3D la luz solar se representa por un punto amarillo metido en un círculo, el cual se pone de color púrpura cuando se selecciona, además de una línea de puntos. Esta línea indica la dirección de los rayos del sol (*Figura 5.24*)

La luz viene de una dirección constante, tiene una intensidad uniforme y no arroja sombras. Esto último es un punto muy importante de comprender en Blender: ninguna luz, excepto las del tipo "Spot", hacen que los objetos arrojen sombras a la escena. La razón para esto reside en la implementación de la luz en un renderizado por líneas y se tratará brevemente con las luces Puntuales (Spot). Por último, es importante darse cuenta que en tanto que la luz solar se define por su energía, color y dirección, la ubicación real en sí de la luz solar no es importante.

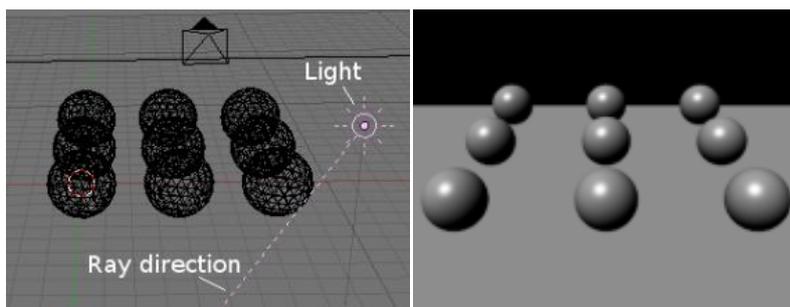


Figura 5.24 – Esquema de la luz solar y su efecto

### 5.8.1.1.2. Luz Hemi

La luz Hemi es un tipo de foco lumínico muy particular ya que está diseñada para simular la luz proveniente de un cielo muy nublado o de un cielo uniforme. Es decir, este tipo de luz se proyecta, uniformemente, por un hemisferio que rodea la escena como muestra la *Figura 5.25*.

La configuración de esta luz básicamente recuerda a la de la luz solar. Su localización no es importante, mientras que su orientación sí que lo es. La línea de puntos representa la dirección en la que se radia la máxima energía, que es la normal al plano que define el corte del hemisferio, apuntando hacia el lado oscuro.

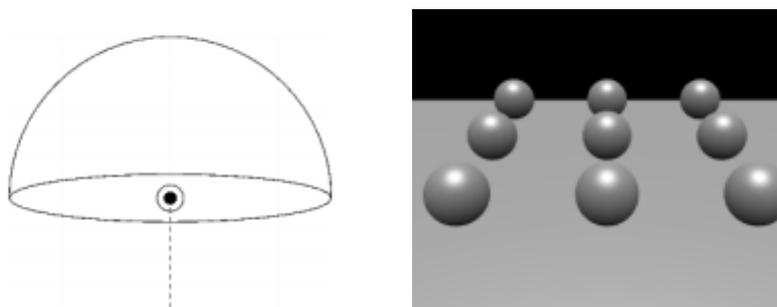


Figura 5.25 – Esquema de luz hemi y su efecto

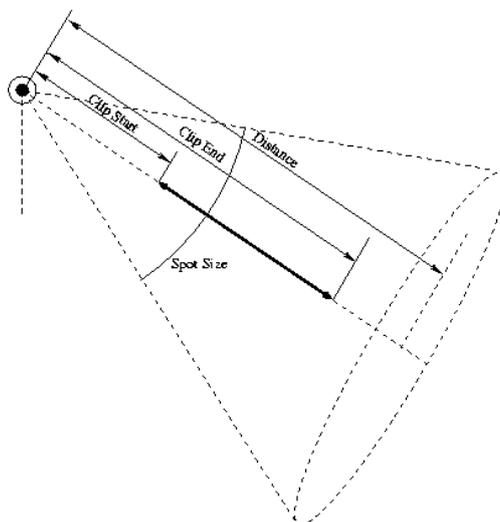
### 5.8.1.1.3. Luz lámpara

La luz de tipo lámpara es una luz de tipo omni-direccional, es decir, es un punto sin dimensiones que irradia la misma cantidad de luz en todas las direcciones. En Blender se representa por un punto amarillo encerrado en un círculo. La dirección de los rayos de luz en la superficie de un objeto es dada por la línea que unes el punto de la fuente de luz y el punto en la superficie del objeto. Es más, la intensidad de la luz decrece acorde a un radio dado de la distancia del foco.

Los objetos que están más cerca que ese punto, reciben más luz, mientras que los que están más lejos reciben menos luz. El usuario puede configurar el radio de acción de la luz además del tipo de atenuación de esta, que puede variar desde una atenuación lineal como la usada por defecto en Blender, a un ratio de atenuación cuadrático en distancia, o conseguir que la luz arrojada por la fuente resulta confinada en forma esférica en lugar de dispersarse hacia el infinito con su radio de atenuación.

#### 5.8.1.1.4. Luz focal

La luz focal es la más compleja de las luces de Blender y es de las más usadas gracias al hecho de ser la única que puede arrojar sombras. Una luz focal es un rayo con forma de cono que parte de la ubicación de la fuente luminosa, la cual es el vértice del cono, en una dirección determinada como muestra el esquema de la *Figura 5.26*.



*Figura 5.26 – Esquema de luz focal*

Entre las diferentes opciones que se pueden configurar para este tipo de lámparas, cabe destacar las opciones de Sombras. Este panel activa/desactiva la proyección de sombras para este foco. Se puede conseguir que el foco sólo proyecte la sombra y no la luz mediante la opción “Only Shadow”. Las luces focales habitualmente proyectan un cono de luz de sección circular. Hay casos en los que sería útil una sección rectangular, y tener una pirámide de luz en lugar de un cono mediante la opción “Square”. También, se puede conseguir que el foco proyecte un halo como si los rayos de luz atravesaran un medio nublado mediante la opción “Halo”.

### 5.8.1.2. La iluminación con Blender y las sombras

#### 5.8.1.2.1. Sombras con el Buffer

Los esquemas de iluminación analizados hasta ahora producen en los objetos áreas más o menos luminosas, pero no proyectan sombras, y una escena sin las sombras apropiadas pierde profundidad y realismo. Por otro lado, un cálculo apropiado de sombras requiere un trazado de rayos completo (y lento). En un renderizador como Blender, las sombras son calculadas usando el shadow buffer, lo que implica que una 'imagen', es iluminada por un foco de luz y al renderizarse se almacena la distancia de cada punto al foco. Cualquier punto de la imagen renderizada que este más lejos que cualquiera de estos puntos, se considera como sombra. El shadow buffer almacena estos datos. Para mantener el algoritmo compacto, eficiente y rápido el shadow buffer tiene un tamaño que es fijado desde el comienzo y que en Blender puede ser desde 512x512 hasta 5120x5120. Cuanto mayor sea este valor, más preciso será.

### 5.8.1.2.2. Sombras mediante el trazador de Rayos

A diferencia de las sombras producidas con Buffer, las sombras producidas por raytracing se obtienen proyectando rayos desde una fuente lumínica uniformemente y en todas direcciones. Se registra aquellos pixeles de la imagen final que fueron alcanzados por un rayo de luz y los que no, y entonces estos últimos son oscurecidos por una sombra.

Los tipos de luces que Blender ofrece proyectan rayos de manera distinta. Por ejemplo, una luz Spot genera los rayos en las direcciones dentro de un cono, mientras que una luz tipo Sol (Sun) emite los rayos como si viniesen desde un punto en el infinito por lo cual todos esos rayos son paralelos y tienen la dirección de la luz en cuestión.

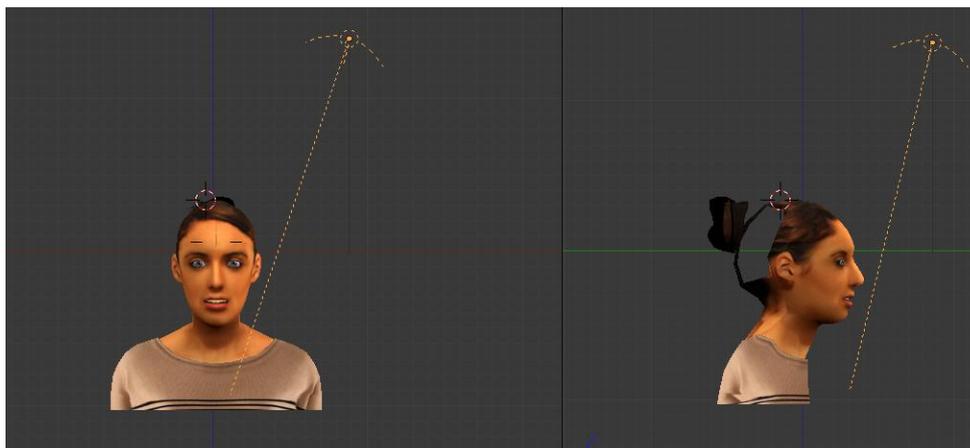
Debido a que cada luz emite una gran cantidad de rayos en la escena, las sombras por raytracing tienden a ser mucho más lentas que las calculadas usando un Buffer, pero los límites de las áreas iluminadas/oscuras son mucho más cortantes y definidos.

### 5.8.1.2.3. Otro tipo de sombras: La luz Volumétrica

La luz volumétrica consigue el efecto que se puede observar con el aire brumoso, cuando los rayos de luz se hacen visibles debido a la dispersión que ocurre debido a niebla, bruma, partículas de polvo etc.

Si se usa cuidadosamente puede añadir mucho realismo a una escena o conseguir que la escena sea pésima. También con este tipo de luces, el usuario puede conseguir que la fuente lumínica proyecte un halo de luz y el objeto proyecte este tipo de sombras.

Tras este estudio de las diferentes fuentes de iluminación que ofrece Blender, se optó por emplear una Luz tipo “Hemi” ajustando su Energía a 2.6, es decir, su intensidad. Con este tipo de luz se consigue una iluminación óptima de la escena, y generar sombras en la superficie del objeto, para la simulación de arrugas al generar las diferentes expresiones en las animaciones, como muestra la *Figura 5.27*.



*Figura 5.27 – Posición de la Luz Hemi en la escena*

### **5.8.1.3. La incidencia de la luz en los objetos**

Cuando un rayo de luz impacta sobre cualquier punto de una superficie, se dan lugar dos tipos básicos de fenómenos: difusión y reflexión especular. La difusión y la reflexión especular se distinguen la una de la otra principalmente por la relación entre el ángulo de luz incidente y el ángulo de la luz reflejada.

#### **5.8.1.3.1. Sombreadores de difusión de luz**

El impacto de la luz sobre una superficie y la posterior reirradiación por fenómeno de difusión puede ser muy dispersa. Esto significa que la cámara verá la misma cantidad de luz desde ese punto de superficie sin importar cual sea su ángulo de visión incidente.

Por supuesto, la cantidad de luz que impacta una superficie depende del ángulo de luz incidente. Blender tiene implementado tres fórmulas matemáticas diferentes para calcular la difusión de la luz sobre los objetos, y el fenómeno de difusión y reflexión especular.

Las tres implementaciones de difusión, o sombreadores (shaders), usan dos o tres parámetros cada una. Los dos primeros parámetros son compartidos por todos los sombreadores de difusión y son los Colores de Difusión o simplemente color, del material, y la cantidad de energía de la luz incidente que es difundida en realidad.

Los sombreadores son los descritos a continuación:

- Lambert  
Es el sombreador difuso por defecto, y su efecto suele ser en general aceptable. Entre sus controles, se puede determinar la cantidad de luz disponible reflejada.
- Oren-Nayar  
Toma una especie de aproximación 'física' al fenómeno de la difusión. Además de los dos parámetros por defecto, tiene un tercero que tiene en cuenta la cantidad de rugosidad de la superficie.
- Toon  
Es un sombreador muy poco 'físico' ya que no está pensado para imitar la realidad, sino para producir renders de 'dibujos animados', con límites muy claros de luz-sombra y regiones uniformes de iluminadas en sombra. Aunque es relativamente simple, también requiere dos parámetros más que definen el tamaño de las áreas iluminadas y el suavizado de los límites de sombra.

#### **5.8.1.3.2. Reflexión especular de la luz**

Al contrario que la difusión, la reflexión especular es dependiente del punto de visión. De acuerdo con la ley de Snell, la luz que impacta sobre una superficie especular se verá reflejada con un ángulo espejado al ángulo de luz incidente, lo cual hace que el ángulo de visión cobre mucha importancia. La reflexión especular crea finos y brillantes reflejos, haciendo que la superficie parezca pulida.

En realidad, la Difusión y la reflexión especular son generadas por el mismo proceso exacto de dispersión de la luz. La difusión es dominante en una superficie que tiene una pequeña escala de rugosidad en ella con respecto a la longitud de onda, de forma que la luz se ve reflejada en muchas direcciones por cada pequeño fragmento de superficie, con cambios muy pequeños en el ángulo de la superficie. Por otro lado, la reflexión especular, predomina en una superficie que es suavizada, con respecto a la longitud de onda. Esto implica que la dispersión de los rayos de cada punto de la superficie serán direccionados en su mayoría en la misma dirección, más que al ser dispersada de forma difusa. Es simplemente un problema de la escala de detalle: si la rugosidad de la superficie es mucho menor que la longitud de onda de la luz incidente parecerá plano y actuará como un espejo.

Como la difusión, la reflexión especular tiene bastantes implementaciones diferentes, o sombreadores especulares. De nuevo, cada una de dichas implementaciones comparte dos parámetros comunes: el Color Especular y la energía de la especularidad, en un rango de [0,2]. Esto permite que se pueda emitir más energía con eficiencia como reflexión especular como si fuera energía incidente. Como resultado, un material tiene al menos dos colores diferentes, uno difuso y uno especular. El color especular suele ser blanco puro, pero pueden ser cambiados sus valores para conseguir efectos interesantes.

Los cuatro sombreadores especulares son:

- CookTorrance

Este tipo de sombreador especular resulta útil para aquellas superficies que simulan plásticos brillantes. Además de los dos parámetros estándar, este sombreador usa un tercero, dureza (hardness), que regula la anchura de los resaltes especulares. Cuanto menor dureza tenga, más anchos serán los reflejos.

- Phong

Éste es un algoritmo matemático diferente, usado para calcular los reflejos especulares. No es muy diferente de CookTor, y está configurado por los mismos tres parámetros.

Este tipo de sombreador especular resulta más conveniente emplearlo en aquellos objetos de superficies orgánicas y que simulen piel.

- Blinn

Éste es un sombreador especular más 'físico', pensado para combinarlo con el difuso OrenNayar. Es más físico debido a que añade un cuarto parámetro, un índice de refracción (IDR - IOR) a los tres anteriores. Este parámetro en realidad no se usa para calcular la refracción de los rayos (para ello se necesita un trazado de rayos), sino para calcular correctamente en la reflexión especular la intensidad y extensión por la ley de Snell. La dureza y los parámetros especulares dan más grados de libertad.

- Toon

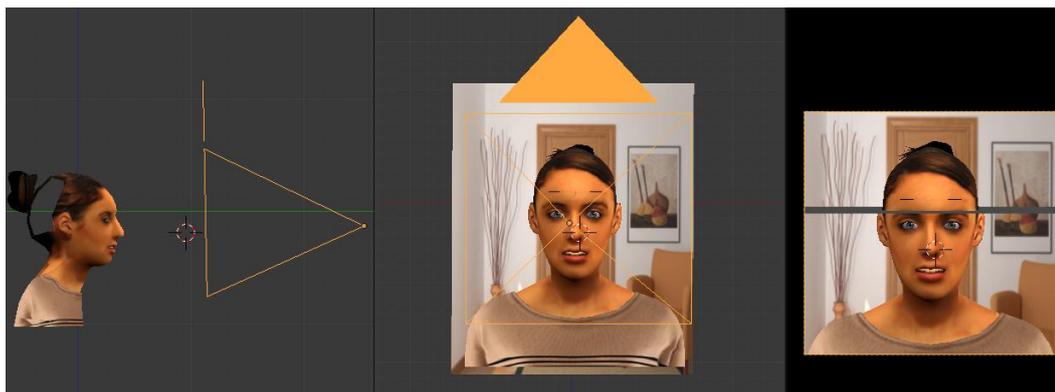
Este sombreador coincide con el de difusión Toon. Está diseñado para producir la agudeza y reflejos uniformes de los dibujos.

Para conseguir un buen efecto de la reflexión de la luz sobre la superficie, se estableció entre los parámetros del material una difusión tipo "Lambert" con una intensidad de 0.8, y una especularidad tipo Phong, ya que resulta la más conveniente para simular el efecto de la luz sobre la piel con una intensidad del 0.1 y 50 de Dureza.

## 5.8.2. Renderizado de la escena. Las cámaras

El renderizado es la fase en la que la escena 3D se crea finalmente y es el proceso de generar una imagen o vídeo partiendo de un modelo en 3D. Para renderizar la escena, es necesario tener presente el empleo de las cámaras y su posicionamiento, ya que estos objetos recopilan la información gráfica para generar las secuencias de animación.

La *Figura 5.28* muestra la posición de la cámara en la escena y la visualización de la escena desde ésta.



*Figura 5.28 – Posición de la cámara y vista de la escena desde ésta*

Es importante también determinar el tipo de cámara que se emplee en la escena, ya que puede conseguirse diferentes efectos en función de la perspectiva de ésta como muestra la *Figura 5.29* además de configurar las dimensiones de la lente, la resolución o el rango de fotogramas y la calidad de los píxeles.



*Figura 5.29 – Vistas de la escena desde la cámara empleando diferentes perspectivas*

Se ha implementado una cámara con un tipo de lente Ortogonal escalada en 14.4 y se ha escogido la cámara predefinida de Blender, aunque se puede escoger una cámara diferente entre los diferentes tipos de Cámaras comerciales que el programa ofrece.

La imagen es renderizada en función de las dimensiones definidas en el panel Format. Por defecto las dimensiones SizeX y SizeY son 320x256 pero se han modificado a 1150x1080. El ratio de proporción de los píxeles de la imagen entre las dimensiones X e Y se ha fijado en 1:1 dado que los píxeles de la pantalla son cuadrados, pero pueden ser modificados por ejemplo para el empleo de la televisión ya que estos píxeles no son cuadrados.

Se puede observar en la *Figura 5.30* la composición final de la escena.



*Figura 5.30 –Resultado final del modelado en escena.*

## **5.9. Animación del modelo**

En este apartado, se explica las técnicas de animación empleadas para el modelo, los BlendShapes o Formas Clave, cómo se han realizado las composiciones de las animaciones y las herramientas necesarias para ello, es decir, la planilla de tiempos y el editor de gráficas y se explican las herramientas empleadas para generar y exportar la aplicación como un ejecutable, el motor de juegos de Blender.

### **5.9.1. Las formas clave**

Las formas Clave permiten el almacenamiento de formas relativas a partir de una malla base. Otras aplicaciones 3D las denominan como ‘objetos de morph’. Las formas clave generan animaciones de una forma sencilla pasando de una forma base a otra mediante la deformación de la malla. Además, estas formas clave pueden ser mezcladas en base a un porcentaje con otras formas clave para conseguir el efecto deseado.

Tomando por ejemplo una cara humana, el usuario puede modelar una cara con una expresión neutral y tener claves de forma para la sonrisa, el ceño fruncido, el ojo izquierdo cerrado, el ojo derecho cerrado, etc. Entonces, las claves de forma tendrían que permitir la combinación de ellas. Así, la malla podría estar sonriendo con su ojo izquierdo cerrado completamente y su ojo derecho cerrado al 50%, como muestra la *Figura 5.31*.

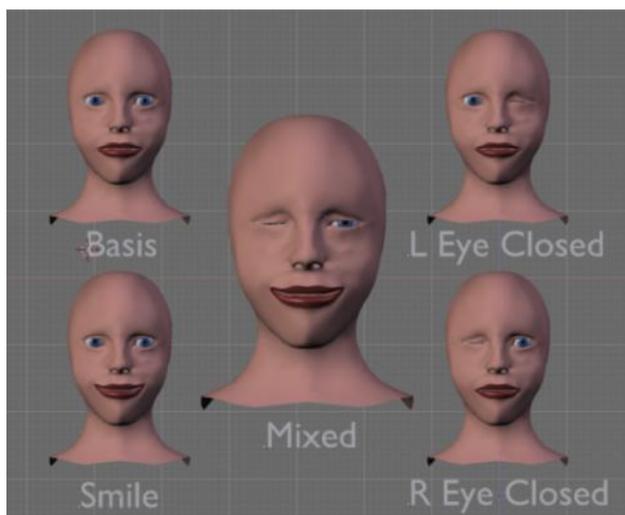


Figura 5.31 –Ejemplo de empleo de formas clave

Empleando una malla básica, el usuario habilita las formas clave insertando una forma clave. Este paso se hace seleccionando el objeto malla y creando una forma clave mediante el botón Add Shape Key del panel de Shapes (formas) en los botones del Modo Edición. El usuario puede identificar y renombrar la forma clave a placer además de modificar el rango de intensidad de la forma clave generada. Además, se pueden añadir tantas formas clave como el usuario requiera, e incluso insertar nuevas formas clave basadas en la forma seleccionada actualmente, por lo que se pueden crear formas clave tanto a partir de la forma base, en cuyo caso la nueva forma será una copia de ésta; o crear una nueva forma clave mientras otra forma alterada está seleccionada, donde la nueva forma será una copia de esa forma alterada.

Una vez se crea una forma clave, esta puede ser editada seleccionándola desde el panel Shapes. Desde el modo edición, el usuario puede alterar la posición, el escalado o la rotación de los diferentes vértices que componen el objeto que se está editando, y observar el efecto desde el panel de Modo Objeto variando la barra de control del rango de acción al emplear la forma clave. Pero añadir o borrar un vértice una vez que las formas clave se han añadido puede resultar problemático, ya que los cambios se propagan a las otras formas basándose en su posición en la forma actual y puede tener un efecto drástico en otras formas. Esto también ocurre al realizar modificaciones en la malla básica, añadir o quitar puntos o variar la posición de estos. Hay que tener especial cuidado en este aspecto, ya que además estos cambios pueden producir modificaciones indeseadas en la textura del objeto a medida que se incrementa el valor de acción de la forma clave aplicada si dichos puntos tienen posiciones extremas en alguno de sus ejes o se pueden generar sombras indeseadas en la malla.

El rango de acción de las formas clave viene limitado entre los valores 0 y 1 por defecto, pero el usuario puede modificar dichos valores si lo desea e incluso emplear parte de la escala negativa. Sin embargo, una parte importante de las Formas Clave es el uso de posiciones aditivas o extrapoladas. Por ejemplo, si la posición base para una cara es con la boca recta, y una clave se define para sonreír, entonces es posible que la aplicación negativa de la clave resulte en curvar la boca al contrario de una sonrisa. En el caso de extendamos la Curva IPO por encima de 1.0 se "extrapolará" esa clave, construyendo una sonrisa extrema

Hay tres métodos para trabajar con Formas Claves de Vértices:

1) El método 'animación cronológica'. Este método trabaja enteramente en Modo Edición, cronológicamente de posición en posición:

- Insertar Forma Clave. Especificar la referencia.

- Unos fotogramas después: Insertar Forma Clave. Editar la Malla para la segunda posición.
- Otros pocos fotogramas después: Insertar Forma Clave. Edita la Malla para la tercera posición.
- Continuar con el proceso anterior.

2) El método 'edición'.

- Primero se insertan todas las formas claves que se necesiten, a menos que ya se hayan creado empleando el método descrito anteriormente.
- Blender no está en Modo Edición.
- Deshabilitar el Modo Objeto seleccionando una Forma Clave. Cuando entra en Modo Edición, se pueden realizar las modificaciones necesarias en la Malla y a continuación salir del Modo Edición
- Seleccionar una Forma Clave. Entrar en Modo Edición, cambiar la Malla y salir del Modo Edición.
- Continuar el proceso anterior.

3) El método 'inserción'. En este método no importa si ya se han creado Claves o si estamos en Modo Edición.

- Se va hasta el fotograma en el que se quiere insertar la nueva Clave.
- Se inserta la Clave.
- Se va a un nuevo fotograma, se inserta la Clave.
- Continuar con el proceso anterior.

Mientras el usuario se encuentra en el Modo Edición, las Claves no se pueden cambiar. Si el usuario intenta hacerlo, aparecerá un aviso. Cada Clave se representa por una línea que se dibuja a cierta altura. La altura será de modo que la clave inserte la "Velocidad" IPO en el fotograma en el que se encuentra la Clave en el proceso de composición de las animaciones que se explicará más adelante.

Se ha empleado un total de 32 Formas clave las cuales vienen detalladas en la *Tabla 5.1*. con las cuales se generarán las composiciones de las animaciones.

<b>Nombre Forma Clave</b>	<b>Rango de Acción</b>	<b>Expresión en la que interviene</b>
<b>Basis (Básica)</b>	[-]	
<b>1- Levantar Interior Cejas. IBR</b>	[0, 1]	Tristeza, Miedo, Sorpresa
<b>2- Levantar Exterior Cejas. OBR</b>	[0, 1]	Alegría, Miedo, Sorpresa
<b>2 Levantar Exterior Cejas con arrugas. OBR ARRUGAS</b>	[0, 1]	
<b>4- Bajar Cejas. BL</b>	[0, 1]	Tristeza, Enfado, Miedo, Disgusto
<b>5- Levantar Parpado Superior. ULR</b>	[0, 1]	Enfado, Miedo, Sorpresa, Disgusto
<b>6- Levantar Mejillas. CR</b>	[0, 1]	Alegría,
<b>7- Estrechar Párpados. LT</b>	[0, 1]	Enfado, Sorpresa, Disgusto

<b>9- Arrugar Nariz. NW</b>	[0, 1]	Disgusto
<b>10- Elevar Labio Superior. ULR</b>	[0, 1]	Enfado, Miedo, Disgusto
<b>11- Elevar Labio Superior. NFD</b>	[0, 1]	Tristeza,
<b>12- Sep. Com. Labios. LCP</b>	[0, 1]	Alegría,
<b>15- Bajar Comisura Labios. LCD</b>	[0, 1]	Enfado
<b>15- Bajar Comisuras Labios mov. labios. LCD MOV LABIOS</b>	[0, 1]	Tristeza
<b>17- Elevador Barbilla. CR</b>	[0, 1]	Tristeza, Enfado, Disgusto
<b>24- Apretar Labios. LP</b>	[0, 1.3]	Alegría, Tristeza, Enfado, Miedo, Sorpresa, Disgusto
<b>25- Separar labios. LP</b>	[0, 1]	Alegría, Tristeza, Miedo
<b>26- Bajar mandíbula solo mand. JD SOLO MAND.</b>	[-1, 1]	Tristeza, Miedo
<b>26- Bajar Mandibula con boca. JD CON BOCA</b>	[0, 1]	Sorpresa
<b>29- Sacar Mandibula. JT</b>	[0, 1]	Enfado
<b>38- Dilatador Fosas Nasales. ND</b>	[0, 1]	Tristeza, Miedo
<b>45- Parpadeo. Blink</b>	[0, 1]	Alegría, Tristeza, Enfado, Miedo, Sorpresa, Disgusto
<b>51- Girar Cuello Izquierda. HTL</b>	[0, 1]	Miedo
<b>52- Girar Cuello Derecha. HTR</b>	[0, 1]	Miedo, Disgusto
<b>53- Girar Cuello Arriba. HU</b>	[0, 1]	Alegría, Miedo, Sorpresa, Disgusto
<b>54- Girar Cuello Abajo. HD</b>	[0, 1]	Tristeza, Enfado
<b>61-Mirar Izquierda. ETR</b>	[0, 1]	Miedo, Disgusto
<b>62-Mirar Derecha. ETL</b>	[0, 1]	Miedo
<b>63- Mirar Arriba. EU</b>	[0, 1]	Tristeza, Enfado, Sorpresa, Disgusto
<b>64- Mirar Abajo. ED</b>	[0, 1]	Alegría, Miedo, Sorpresa, Disgusto
<b>Contraer Pupilas. Pupil Closed</b>	[0, 1]	
<b>Dilatar Pupilas. Pupils Open.</b>	[0, 1]	Tristeza, Disgusto

Tabla 5.1 –Listado y descripción de las formas clave

### 5.10. Composición de las animaciones

La animación es uno de los pasos esenciales cuando se crean gráficos 3D por ordenador, y supone una tarea que requiere gran precisión y muchas veces supone un trabajo repetitivo. Una vez creados todos los componentes que intervienen en la generación de movimiento de los proyectos de contenido 3D y mediante el empleo conjunto de todos ellos, se generan las animaciones. Se han desarrollado muchas técnicas diferentes para acelerar y reducir la carga de trabajo de los animadores.

Para gestionar todas estas características, Blender cuenta con tres editores diferentes de Animación (tres tipos de ventanas):

- El editor de curvas IPO

Este es el editor de nivel más bajo, en el que emplea una curva para controlar cada propiedad animada / ajuste. Como se comentó anteriormente, dadas dos posiciones del objeto, se genera automáticamente una función de interpolación y mediante esta se especifica el movimiento entre una posición de origen y otra final dentro de un intervalo normalizado calculando los puntos intermedios por los que los vértices pasan. Mediante este editor (*Figura 5.32*), se puede configurar las diferentes curvas de interpolación creadas para las diferentes Formas Clave y controlar la velocidad, la duración o la intensidad de estas en la animación.

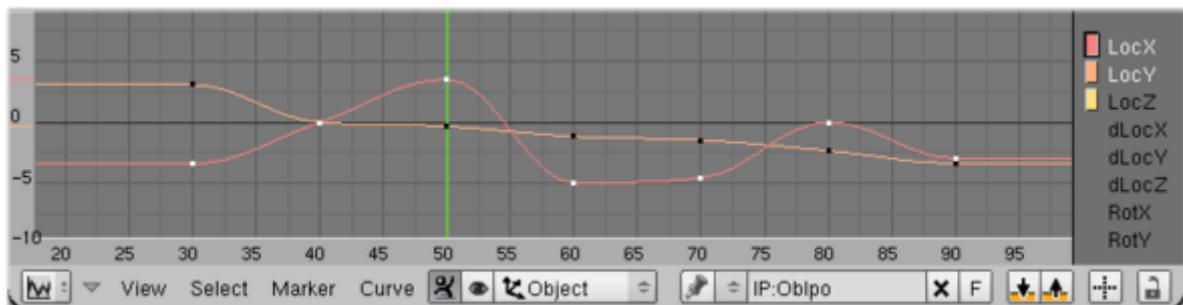


Figura 5.32 – Editor de Curvas IPO

- El editor de Acción

Este editor es algo similar al Editor de Curvas IPO, ofrece una visión menos precisa y más general de la animación del objeto, pero, a diferencia del Editor de Curvas IPO, permite editar al mismo tiempo múltiples Formas Clave o Bones. (*Figura 5.33*)

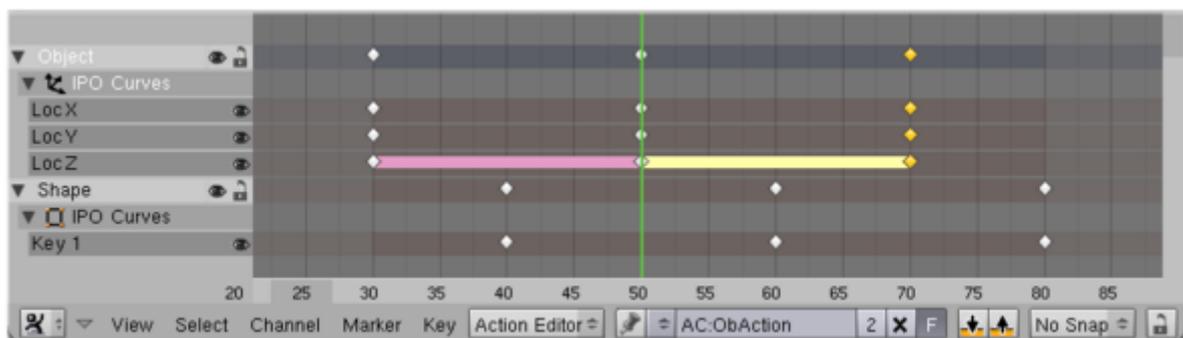


Figura 5.33 –Editor de acciones

- El editor NLA

Este es un editor de alto nivel cuyo funcionamiento está basado en los editores de video no lineales donde cada elemento representa una acción que puede ser movida, duplicada, acelerada o ralentizada a gusto del usuario. (Figura 5.34)

La idea detrás de este editor de alto nivel ha sido tomada de los editores de video no lineal: cada elemento representa una acción que se puede mover, duplicar, reducir / extracción (es decir, sujetar / lento), etc., a su gusto. Si las claves IPO son consideradas más bien caras de la malla, en una “Acción NLA” todos estos objetos se encuentran encapsulados y cuentan con un modificador.

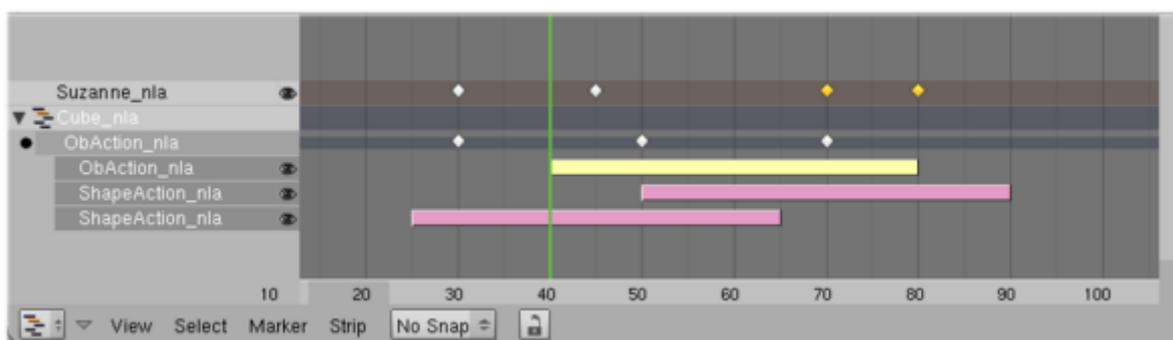


Figura 5.34 - Editor NLA

Es importante entender qué Editor de Animaciones le interesa más al usuario para generar su composición o qué rasgo desea destacar para escoger qué editor le resulta más cómodo y útil para trabajar.

También hay que tener presente que en el momento que el objeto contiene animación (la posición de un objeto, el color difuso de un material, la asignación de una textura, etc.) su curva IPO siempre anula toda la edición que podría realizar manualmente en las propiedades de animación a menos que determine esta edición como relevante en las curvas IPO insertando Formas Clave.

### 5.10.1. El editor de curvas IPO

Una curva IPO, es una curva que controla el valor de una propiedad, dependiente de los parámetros temporales asignándole el tiempo de duración de dicho valor.

Además de poder editar las animaciones mediante el Editor de Curvas IPO, también puede agregar, modificar o eliminar fotogramas clave en la línea de tiempo, en la vista 3D del objeto, en la ventana de botones o empleando los otros editores de animación y por último y no menos importante, grabar los movimientos de un objeto desde el motor de juego que trataremos posteriormente en el apartado 5.11.

En el área de trabajo, además de las opciones de edición se muestra en una de las ventanas las diferentes Curvas ISO gráficamente. El eje X, corresponde a la evolución de la Curva temporalmente. Este viene medido en fotogramas, o segundos, y es preciso para ello determinar el número de fotogramas o frames por segundo en las propiedades de la cámara. En el eje Y se representa el valor ponderado de la curva ISO, es decir, el valor del peso de la forma clave en un preciso instante. Además, se muestra un listado de las diferentes Curvas ISO existentes y por lo tanto las diferentes animaciones del objeto.

### 5.10.1.1. Edición de fotogramas clave mediante las Curvas IPO

Mediante el Editor de Curvas IPO se puede generar composiciones de animaciones empleando las formas clave. Las Curvas IPO se componen por la sucesión de puntos de control que determinan la trayectoria de la curva y por lo tanto de la animación y estos puntos de control corresponden a las formas clave generadas de los objetos. Para generar las Curvas IPO, es necesario insertar los fotogramas clave que componen las animaciones. Para ello resulta útil duplicar los fotogramas para obtener más puntos de control que compongan la curva y unir los elementos seleccionados.

A la hora de componer una animación, se debe tener presente la velocidad de reproducción de esta y la velocidad de un objeto viene dada por la pendiente de su curva en un gráfico de tiempo o posición. Para ello, se emplea LocX, LocY y LocZ, los cuales establecen los valores de los gradientes en un momento dado y por lo tanto, la velocidad en un determinado momento.

Para editar la trayectoria de la Curva IPO, se deben seleccionar las formas clave, ya sea seleccionando el punto central para seleccionar el punto de control conjunto o uno de los extremos para señalar sólo una empuñadura lateral. Se debe tener en cuenta que las curvas en modo constante o lineal de interpolación tienen puntos de control más simples, compuesto de un único vértice. Es posible realizar transformaciones de los puntos de control para determinar la trayectoria de las Curvas IPO o bien seleccionando cada uno individualmente o editándolos todos simultáneamente, pero se debe tener en cuenta que tendrán exactamente el mismo valor en el tiempo.

Mediante el menú marcador se editan las curvas IPO y se determinan las formas clave que las componen. Para ello, Blender dispone de tres métodos de visualización diferentes según las preferencias del usuario.

#### - Modo Curva

Se muestran las diferentes curvas creadas para cada animación donde los fotogramas clave se materializan por puntos blancos de las curvas seleccionadas y color negro para el resto. (Figura 5.35)

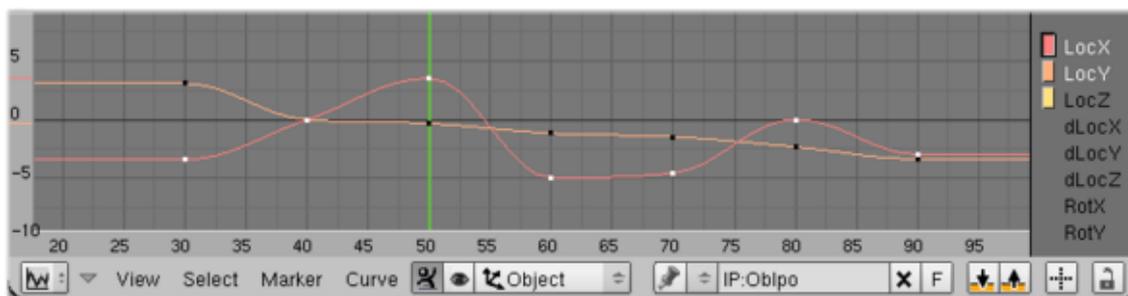


Figura 5.35 - Editor de Curvas IPO Modo Curva

#### - Modo de Fotogramas Clave

En este caso, los fotogramas clave se visualizan por una línea vertical amarilla y se puede editar la posición-tiempo de cada uno. En este modo de visualización, el usuario no tiene control directo sobre las curvas. (Figura 5.36)



Figura 5.36 - Editor de Curvas IPO Modo Formas Clave

- El modo edición

Aquí se puede editar cada fotograma clave de las curvas seleccionadas como controles de la Curva (mediante interpolación por defecto). (Figura 5.37)

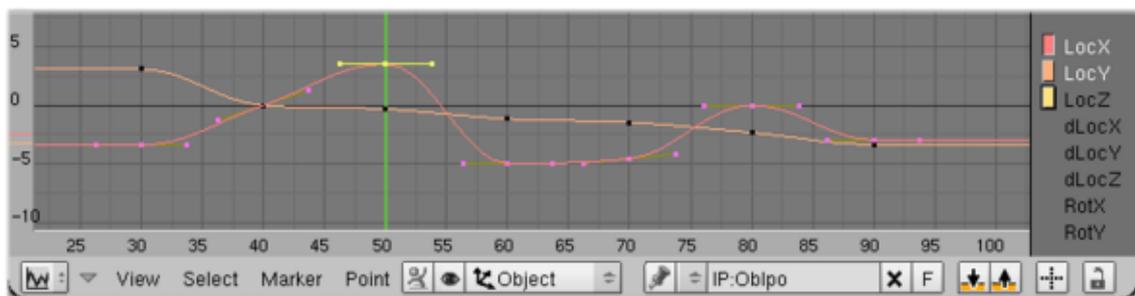


Figura 5.37 - Editor de Curvas IPO Modo Edición

### 5.10.1.2. Modificaciones y ajustes de las Curvas IPO

Seleccionando o activando los bloques de datos que se desea animar, se generan las Curvas IPO de cada objeto, y una vez agregados todas estas al entorno de trabajo, Una vez agregada una Curva IPO de un objeto en el panel de trabajo, comienza el proceso de edición de estas. Se pueden editar de una forma muy similar a la forma de trabajo de la “Vista 3D” ya que la mayoría de las opciones y comandos son comunes a los diferentes menús, pudiendo realizar transformaciones de las Curvas (traslación, rotación, escalado) o dividir la curva en varios fragmentos de una manera muy simple. Además, existen diferentes opciones complementarias que facilitan la tarea de animación, como el modo Espejo, o copiar o eliminar la curva editada.

Para definir el movimiento generado para el objeto mediante cada Curva IPO que componen la animación, el usuario debe ajustar los parámetros de estas para conseguir el efecto deseado. Para el ajuste existe una serie de herramientas que resultan muy útiles, como el ajuste “horizontal” mediante el cual todos los puntos de la curva seleccionados se colocan horizontalmente generando así posiciones intermedias en las animaciones. Se puede apreciar el efecto de este ajuste en la Figura 5.38.

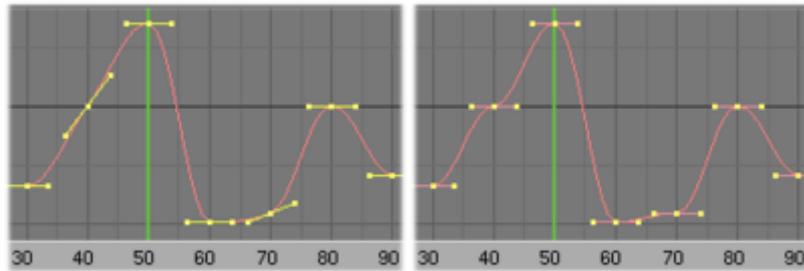


Figura 5.38 – Ajuste horizontal de formas Clave en el Editor de Curvas IPO

También se puede asignar a los puntos de control de la curva (Formas Clave) seleccionados de la curva el valor de la siguiente Forma Clave, produciendo así un desplazamiento de la curva como muestra la *Figura 5.39*.

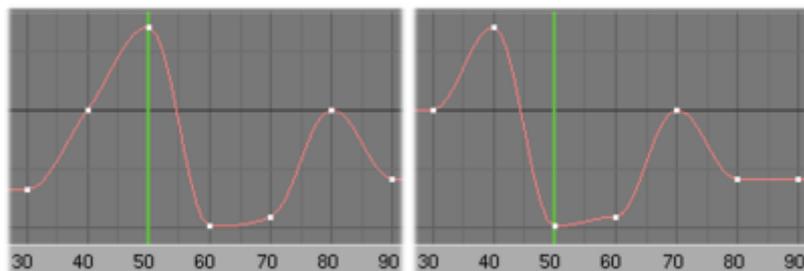


Figura 5.39 – Ajuste de formas Clave en el Editor de Curvas IPO

Otra opción importante en el ajuste de las Curvas IPO, es la limpieza y alisado mediante el cual se eliminan los puntos de control que están muy próximos entre sí en valor y tiempo además de poder establecer un umbral para eliminar aquellos puntos que resultan más convenientes.

Existen tres propiedades adicionales que controlan la interpolación y comportamiento de la extensión, mediante las cuales se puede modificar el tipo de función de interpolación empleada para generar el movimiento (*Figura 5.40*):

- Constante, donde no se realiza ningún tipo de interpolación.
- Lineal, en el que se crea un segmento recto entre fotogramas generando una línea quebrada.
- Bezier, que supone la interpolación más potente y útil y la opción por defecto, mediante la cual se consiguen curvas muy suavizadas y se obtienen animaciones suaves.

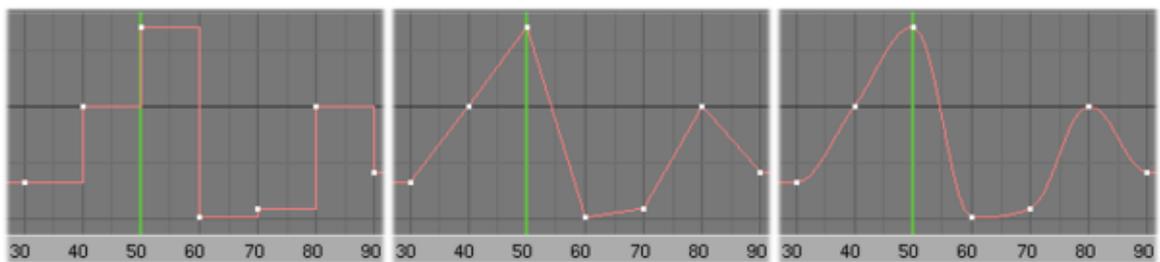


Figura 5.40 - Interpoladores de Curvas ISO. Constante, Lineal y Bezier, respectivamente.

Para la extensión de la curva, se pueden emplear cualquiera de las siguientes opciones que permite Blender (*Figura 5.41*):

- Constante: Las curvas antes y después de su primer fotograma clave tiene un valor constante

- Extrapolación: Los extremos de las curvas son rectas definidas por los dos primeros fotogramas clave y los dos últimos.
- Cíclico: La porción de la curva con formas clave se repite ‘ad infinitum’ antes y después de ella. Se debe tener presente al emplear esta opción que si los puntos de control de inicio y fin no tienen el mismo valor en el Eje Y, existe un vacío en la curva cada vez que pasa por este punto.
- Extrapolación cíclica: Es muy similar a la cíclica pero se genera una desviación vertical al pasar por el punto donde se encuentra la forma clave que se copia cíclicamente y si es necesario una desviación vertical a modo que no se produzcan esos vacíos.

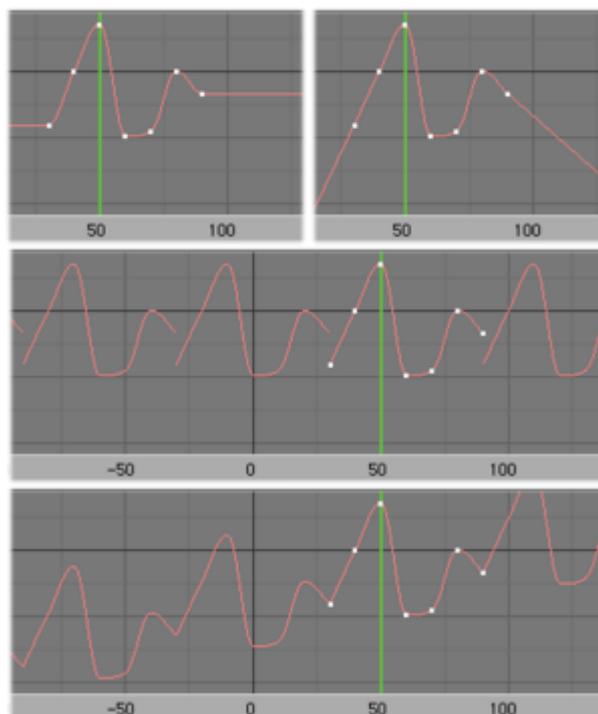


Figura 5.41 - Extensión lineal, extrapolación, Cíclico y Cíclico Extrapolado respectivamente.

### 5.10.2. El editor de acción

El editor de acción (Figura 5.42) permite ver y editar los bloques de datos IPO que componen las distintas animaciones. Este método de edición es más sofisticado que el Editor de Curvas IPO, pero no tan avanzado como el editor NLA y se suele emplear combinado con este último. Ofrece una visión más simplificada de los bloques de dato IPO y resulta interesante al generar animaciones de una forma más completa y sencilla que mediante el empleo del editor de las Curvas IPO y no tan extenso como el editor NLA. Además, permite modificar el tiempo de duración de las formas claves en bloque.

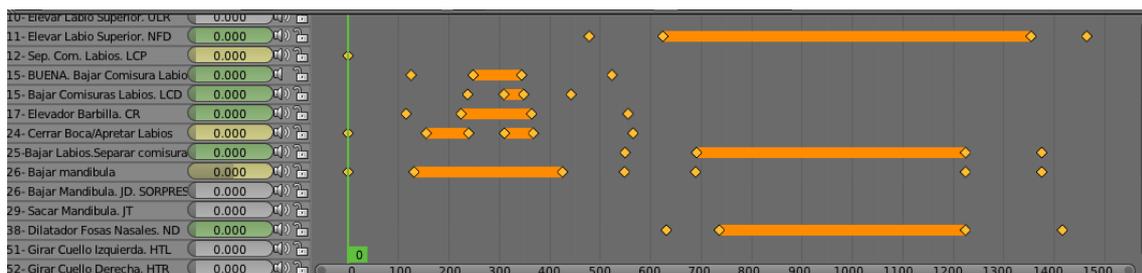


Figura 5.42 – Editor de Acción

La interfaz del Editor de Acciones es algo similar al Editor de Curvas Ipo, dividiéndose en tres áreas principales: el menú principal donde se muestran los principales controladores y editores entre otras herramientas, la planilla de tiempos, la cual es el área principal de trabajo y contiene los fotogramas clave para todos los canales seleccionados y el listado de los canales de las acciones, su jerarquía y el peso de cada acción. Al igual que en la ventana de visualización de las Curvas Ipo, se muestra una panel con la composición de cada animación y el peso de cada forma clave en la evolución temporal correspondiendo el eje X al tiempo, pero en este caso, el eje Y no tiene el mismo significado. Corresponde a la evolución de los canales de acción donde cada uno se muestra como una banda de color en horizontal y donde se incluyen los fotogramas clave representados como diamantes amarillos o gris dependiendo de si están seleccionados o no.

Una de las características clave de esta ventana es que permite visualizar de una forma muy clara el efecto y el peso de cada canal en la composición: cuando el valor de un determinado canal no cambia entre dos fotogramas clave vecinos, se dibuja una tira entre ellos. Si un fotograma o una secuencia de fotogramas hay menos de 8 rastros, se volverá rojo, si hay entre 8 y 16 rastros, el fondo será amarillo.

Esto es solo una respuesta visual que no significa que no se pueda reconstruir el movimiento de la cámara con menos de 8 rastros, únicamente el usuario debería prestar atención a estos fotogramas, lo cual no significa que la solución no vaya a ser precisa y comprobar si todos los posibles puntos de rastreo buenos se están rastreando allí correctamente. Más aún, rastrear más puntos de rastreo malos reducirá la precisión de la solución.

En el listado con todos los canales, se muestran los bloques de datos Ipo y se muestran todos los fotogramas clave que le definen. Al lado de cada forma clave, se muestra un indicador con el valor de cada fotograma clave empleado en la composición y permite cambiar el valor de fotogramas clave actuales o añadir otros nuevos. Se puede observar también los diferentes valores que adquieren las distintas formas clave que componen las animaciones a medida que avanza el cursor por la planilla de tiempos.

El ajuste de marcadores 3D es una etapa del cálculo de movimiento de cámara que mejora la solución, minimizando los errores de proyección que afecta a la posición de los marcadores 3D y las cámaras y aplica un método no lineal de mínimos cuadrados, obteniendo animaciones más precisas.

Otra herramienta que ofrece la planilla de tiempos, es la precarga en paralelo de fotogramas en el Editor de clips de película. La precarga es un operador que al invocarlo rellena el caché con fotogramas. Se cargarán tantos fotogramas desde el disco como quepan en el caché. Esto permite rellenar rápidamente el caché cuando se necesita rastrear algo, pero mantiene el ancho de banda de la CPU y el disco duro libres si se mantiene el Editor de clips de películas abierto, pero no se está interactuando con él. De esta manera además la memoria se encuentra completamente disponible para el clip en el que se está trabajando actualmente.

El editor de acciones tiene tres modos de trabajo disponibles:

- **Editor de acciones:** es el valor por defecto y más útil. Aquí se puede definir y controlar las acciones de las composiciones que se están configurando. Para ello, se emplean también los mismos comandos prácticamente que los empleados en las otras ventanas de Blender.
- **Editor de Formas Clave:** como el propio nombre del editor indica, está pensado para editar los fotogramas clave de la animación. Emplea el mismo bloque de datos que el editor de acciones.
- **Grease Pencil:** Está dedicado a la herramienta de lápiz graso de fotogramas clave, de forma que cada capa de lápiz existe una franja a lo largo de la cual se pueden seleccionar las formas clave, y por tanto fácilmente ser editados en la planilla de tiempos.

Como todo lo demás en Blender, las acciones son bloques de datos, y estos pueden ser agrupados en un único objeto los cuales se almacenan siempre en el archivo de Blender, incluso si no hay ningún objeto que lo utiliza. Esto se debe al hecho de que las acciones están diseñadas para ser empleadas en el NLA, donde varias acciones distintas pueden afectar a un mismo objeto. De esta manera se pueden emplear diferentes acciones si se requiere (y por lo tanto diferentes bloques de datos Ipo de la misma especie) para animar un mismo objeto. De esta manera, se garantiza que no se pierdan las acciones editadas previamente en una composición de animación al trabajar en una nueva.

El editor de acciones resulta una herramienta muy útil al componer las animaciones, ya que muestra de una forma muy simple la forma clave que se está editando, el valor y la evolución de esta en la planilla de tiempos. Además, simplemente añadiendo un valor a una forma clave automáticamente esta se inserta en la composición, aunque también resulta muy útil para añadir los fotogramas clave el duplicado de los mismos.

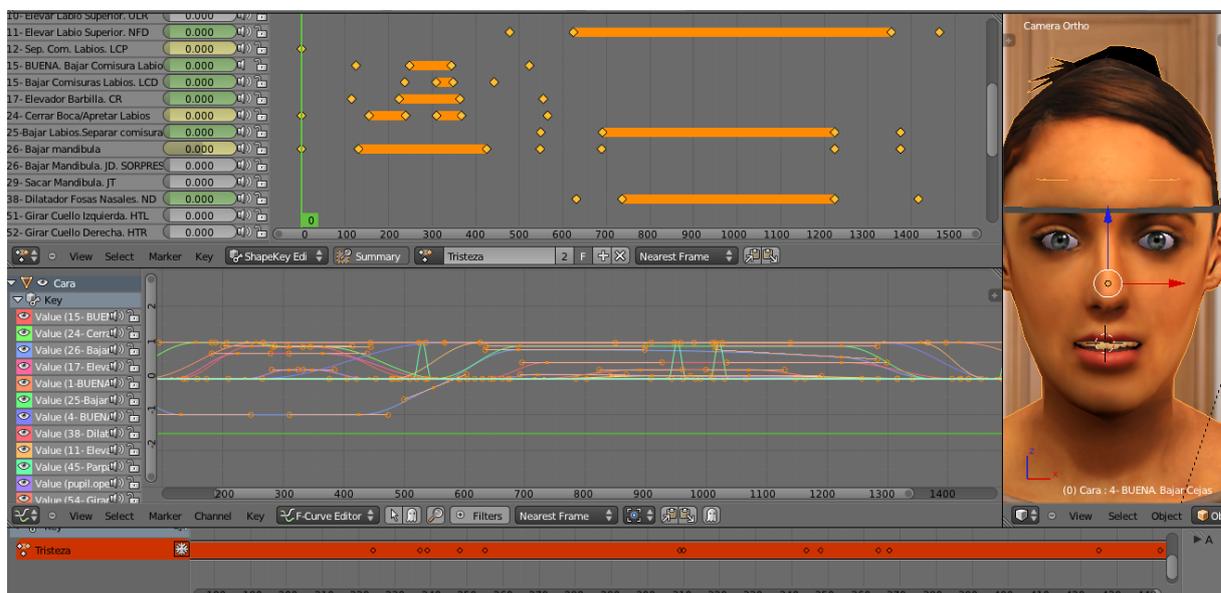
### **5.10.3. El editor NLA**

La ventana de Editor NLA permite realizar animaciones de una forma sencilla en Blender ya que su funcionamiento está basado en el funcionamiento de editores de vídeos externo. La animación mundial (NLA) se compone de fragmentos pequeños de acciones en las que cada una realiza una acción determinada. Estas acciones simples se pueden aplicar a diferentes objetos, ser reprogramadas, aceleradas o ralentizadas, repetidas varias veces, etc. para componer la animación compleja final.

Las pequeñas animaciones, se pueden componer empleando únicamente el Editor de Curvas IPO, sin embargo, interesa realizar aquellas animaciones más avanzadas y complejas mediante este editor ya que se puede descomponer en pequeños fragmentos con los que se trabaja de una manera más sencilla. Por ejemplo, si se generaran animaciones relacionadas con el movimiento de los labios en el síntesis del habla, conviene generarlas con este editor si se desea conseguir mucho realismo.

Aún así, el NLA requiere tiempo y práctica para entender y dominar las múltiples opciones que ofrece. Para generar las composiciones generadas, no se ha empleado este editor, pero resulta interesante mencionarlo en este documento porque resulta interesante conocerlo.

La *Figura 5.43* muestra del panel de animaciones donde se muestra de arriba abajo, el Editor de Planilla de Tiempos, el Editor de Curvas IPO y el Editor NLA respectivamente de la animación de “Tristeza”.



*Figura 5.43 - Visualización del panel de animaciones con los diferentes editores*

## 5.11. Game engine. Motor de Juego

Blender cuenta con su propio motor de juego el cual permite crear aplicaciones interactivas en 3D o simulaciones. La principal diferencia entre el motor del juego y el sistema convencional de Blender está en el proceso de renderizado. En el motor normal de Blender, las imágenes y animaciones se construyen fuera de línea, es decir, una vez prestados no pueden ser modificadas. A la inversa, el motor de juego de Blender hace escenas de forma continua en tiempo real, e incorpora funciones para la interacción del usuario durante el proceso de renderización.

El motor de juego de Blender está diseñado para proveer una poderosa herramienta para crear la lógica del juego a través de un entorno gráfico. Este supervisa un bucle de juego, que procesa la lógica, el sonido, la física y la representación de simulaciones en orden secuencial y está escrito en C++. El editor lógico proporciona una interacción profunda con la simulación, y su funcionalidad puede ser extendida a través de secuencias de comandos Python. Está diseñado para abstraer las características del motor complejos en una interfaz de usuario sencilla, que no requiere experiencia en programación.

El motor del juego puede simular el contenido dentro de Blender, sin embargo, también incluye la posibilidad de exportar tiempos de ejecución de binarios para Windows, Linux y MacOS.

El diseño, la construcción, la depuración y la ejecución de un juego emplea una amplia gama de funciones de Blender. El Editor Lógico es donde la lógica, propiedades y estados están configurados para controlar el comportamiento de los objetos en el juego. Los bloques representan funciones pre-programadas las cuales puedes ser ajustadas y combinadas para crear juegos y/o aplicaciones.

La lógica en el motor de juegos de Blender está ligada a los objetos. Los objetos están listados por sus nombres y aparecen en la ventana de lógica cuando son seleccionados. Además se puede señalar tanto un único objeto como varios. La lógica de un objeto es sólo visible cuando el objeto asociado con esta lógica es seleccionado.

El sistema está dividido en tres partes: sensores, controladores y actuadores. Los sensores perciben cuando suceden cosas como una colisión, la pulsación de una tecla o el movimiento del ratón. Los sensores están (conectados) a los controladores, quienes los evalúan y activan los actuadores.

- Conexiones

Las conexiones son la dirección del flujo de lógica entre los objetos. Las conexiones son dibujadas arrastrando desde un nodo de conexión hasta otro. Las conexiones solo pueden ser dibujadas desde los sensores hasta los controladores y desde los controladores hasta lo actuadores. Los conectores de salida (círculos negros, a la derecha de los sensores y controladores) pueden conectarse a múltiples conectores de entrada (a la izquierda de los controladores y actuadores). Además, los conectores de entrada pueden recibir múltiples conexiones también. Si se intenta conectar directamente sensores con actuadores, se creará automáticamente un controlador entre ellos. Si se desea lograr que un sensor se active luego de que un actuador se haya completado, se deberá usar el sensor Actuador.

- Sensores

Los sensores comienzan todas las acciones lógicas. El sensor indica cosas como la cercanía de un objeto, la pulsación de una tecla, eventos programados, etc. Cuando un sensor es activado, un pulso es enviado a todos los controladores enlazados. Estos proporcionan una salida cuando algo sucede, por ejemplo, un evento de disparo, como una colisión entre dos objetos, una tecla presionada en el teclado, o un temporizador para un evento con hora de apagarse. Cuando se activa un sensor, un pulso positivo se envía a todos los controladores que están vinculados a la misma. Los diferentes tipos de sensores que ofrece Blender, vienen reflejados en la *Tabla 5.2*.

<b>Actuador</b>	Detecta cuando un actuador particular recibe un pulso de activación.
<b>Siempre</b>	Da una señal de salida continua a intervalos regulares.
<b>Colisión</b>	Detecta colisiones entre objetos o materiales.
<b>Retrasar</b>	Los retrasos de salida por un número especificado de pasos lógicos.
<b>Palanca de mando</b>	Detecta el movimiento de los controles de palanca especificados.
<b>Teclado</b>	Detecta la entrada de teclado.
<b>Mensaje</b>	Detecta tanto mensajes de texto o valores de propiedad
<b>Ratón</b>	Detecta eventos autónomas.
<b>Cerca</b>	Detecta objetos que se mueven a una distancia específica de sí mismos.
<b>Propiedad</b>	Detecta los cambios en las propiedades de su objeto propietario.
<b>Radar</b>	Detección de objetos que se mueven dentro de una distancia específica de sí mismos, dentro de un ángulo de un eje.
<b>Azar</b>	Genera pulsos aleatorios.
<b>Rayo</b>	Tira un rayo en la dirección de un eje y detecta accesos.
<b>Toque</b>	Detecta cuando el objeto está en contacto con otro objeto.

*Tabla 5.2. –Listado y descripción de los sensores*

- Controladores

Los controladores manipulan la lógica, evalúan los pulsos de los sensores y envían pulsos a los actuadores en respuesta. Los diferentes tipos de controladores son los que vienen recogidos en la *Tabla 5.3*.

<b>Y</b>	Todos los sensores conectados deben estar activados para enviar un pulso positivo a los actuadores vinculados.
<b>O - O exclusivo</b>	Al menos uno de los sensores debe estar activado para enviar un pulso positivo.
<b>OeX</b>	Uno, y solo uno, de los sensores conectados debe estar activado.
<b>NoY</b>	Controlador 'Y' invertido.
<b>NoO</b>	Controlador 'O' invertido.
<b>NoOEx</b>	Controlador 'OEx' invertido.
<b>Expresión</b>	Escribe su propia expresión.
<b>Python</b>	Controla el sensor con un módulo o script de Python.

*Tabla 5.3. –Listado y descripción de los controladores*

- Actuadores

Los actuadores afectan los objetos o el juego de alguna manera. Los actuadores cambian movimiento, sonido, propiedades, objetos, etc. Estos cambios pueden provocar eventos en otros bloques lógicos. Los actuadores realizan acciones, como mover, crear objetos o reproducir un sonido. Los actuadores inician sus funciones cuando reciben un pulso positivo desde uno (o más) de sus controladores. Los bloques lógicos de todos los tipos de actuador pueden ser construidos y cambiados usando el Editor de lógica. La *Tabla 5.4* muestra los diferentes tipos de actuadores disponibles.

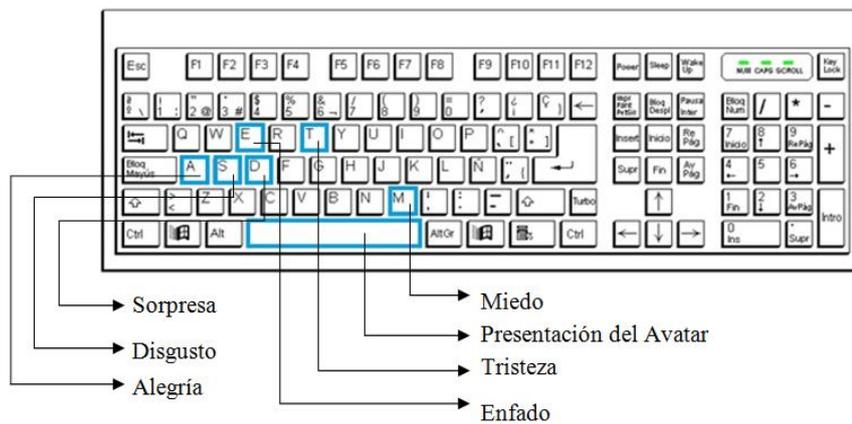
<b>Acción</b>	Maneja acciones inducidas. Esto sólo es visible si se ha seleccionado una armadura.
<b>Cámara</b>	Tiene opciones para seguir objetos sin problemas, sobre todo para los objetos de la cámara, pero cualquier objeto puede usar esto.
<b>Restricción</b>	Las restricciones se utilizan para limitar ubicación del objeto, la distancia, o la rotación. Estos son útiles para el control de la física del objeto en el juego.
<b>Editar Objeto</b>	Edita malla del objeto, añade objetos, o los destruye. También puede cambiar la malla de un objeto (y pronto también volver a crear la malla de colisión).
<b>Filtro 2D</b>	Filtros para efectos especiales como colores sepia o la falta de definición.
<b>Juego</b>	Maneja todo el juego y puede hacer cosas como reiniciar, dejar de fumar, cargar y guardar.
<b>Mensaje</b>	Envía mensajes, que pueden ser recibidas por otros objetos para activarlos.
<b>Movimiento</b>	Establece objeto en movimiento y / o rotación. Existen diferentes opciones, de "teletransportarse" para empujar físicamente girar objetos.
<b>Superior</b>	Puede fijar un padre al objeto, o un parent ella.
<b>Propiedad</b>	Manipula las propiedades del objeto, como asignar, añadir o copiar.
<b>Aleatorio</b>	Crea valores aleatorios que pueden ser almacenados en las propiedades.
<b>Escena</b>	Administrar las escenas en su archivo. Mezcla. Estos pueden ser usados como los niveles o de la interfaz de usuario y el fondo.
<b>Sonido</b>	Se utiliza para reproducir sonidos en el juego.
<b>Estado</b>	Cambios estados del objeto.
<b>Conducción</b>	Proporciona opciones de búsqueda de caminos para el objeto.
<b>Visibilidad</b>	Los cambios de visibilidad del objeto.

*Tabla 5.4. –Listado y descripción de los actuadores*

### 5.11.1. Confección y generación de la aplicación mediante el motor de juego

Para realizar la aplicación interactiva, se han empleado 7 sensores, 7 actuadores y 7 controladores, cada uno para cada animación generada para la aplicación.

Se ha empleado un total de 7 sensores del tipo “teclado” y en la *Figura 5.44* aparecen todos ellos recogidos con su correspondiente animación. Se ha escogido emplear la primera letra de cada expresión para el control de las animaciones de las mismas para que resultara más sencillo e intuitivo (A-Alegría, T- Tristeza, E- Enfado, S-Sorpresa, M-Miedo, D- Disgusto y adicionalmente la barra espaciadora como presentación del Avatar).



*Figura 5.44 –Detalle de los sensores empleados para controlar las animaciones*

Los actuadores son todos de tipo “Y- And”, de esta manera, al activar el usuario el sensor, es decir, únicamente pulsando la tecla correspondiente a la expresión que desea animar, se envía un pulso positivo al actuador vinculado a dicha animación.

Por otro lado, los controladores elegidos son tipo “Acción”. Estos controladores manejan las acciones inducidas en la armadura, es decir, se mostrarán las animaciones generadas mediante el editor de lógica y la planilla de tiempos. Para ello, se tiene que determinar el tipo de acción a realizar es decir, “Reproducción”, y señalar entre las animaciones realizadas qué animación corresponde a dicho actuador. En cada controlador también se tiene que indicar la duración de las animaciones, el frame inicial y final de cada animación. Por defecto estos parámetros están a cero, y al generar la aplicación, al pulsar el actuador correspondiente no se realizará ninguna acción. Es importante reparar en este hecho, ya que en caso de no determinar estos parámetros ya que las animaciones pueden mostrarse incompletas. En la *Figura 5.45* se puede observar parcialmente el esquema de las conexiones realizados en el Motor de Juego.

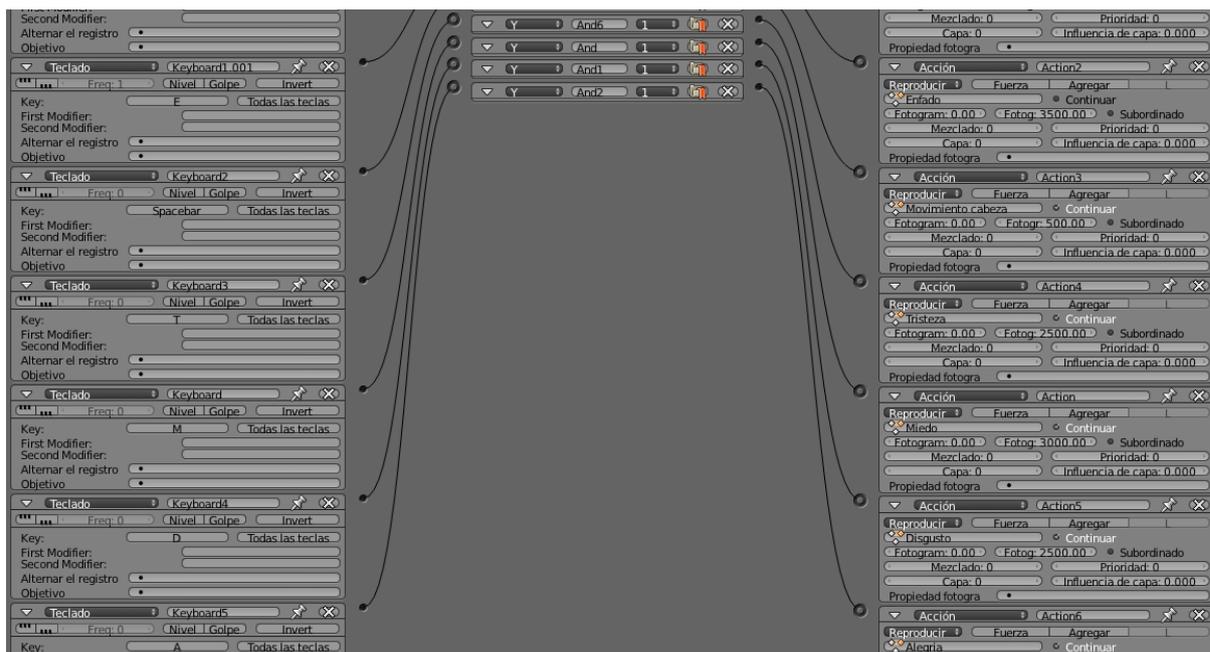


Figura 5.45 –Esquema de las conexiones en el Motor de Juego

Una vez confeccionada la aplicación y determinando las acciones que se van a realizar en esta, es momento de generarla. Blender permite que una aplicación pueda ser ejecutada sin necesidad de tener que cargar el sistema Blender o tener el programa instalado en el sistema. Para ello, automáticamente al exportar el archivo como reproductor independiente, se generan un ejecutable del archivo y una serie de bibliotecas automáticamente por los complementos. En general, estas son las bibliotecas generadas automáticamente por el programa, pero puede incluir alguna más en función de la versión del programa y el sistema operativo empleado.

- avcodec-51.dll
- avformat-52.dll
- avutil-49.dll
- libfaac-0.dll
- libfaad-0.dll
- libmp3lame-0.dll
- libx264-59.dll
- pthreadVC2.dll
- python25.dll
- SDL.dll
- swscale-0.dll
- vcomp90.dll
- xvidcore.dll
- zlib.dll

Esto permite que los juegos y aplicaciones interactivas puedan ser distribuidos a otros usuarios sin requerir que posean un conocimiento detallado de Blender y sin la posibilidad de modificación no autorizada.

Si se desea que todos los datos (texturas, sonidos, animaciones, etc.) sean guardados como parte del archivo .blend, se debe empaquetar el archivo previamente como archivo .blend, sino, es posible que se pueda perder parte de la información o detalles de la aplicación generada. También se debe tener en cuenta que el motor de juegos genera y guarda la aplicación en tiempo

en ejecución y esta debe ser precargada e inicializada antes de generar la aplicación para que tenga un correcto funcionamiento, es decir, se deben fijar todos los valores a los parámetros iniciales.

La aplicación será ejecutada mediante la ejecución del correspondiente archivo .exe e incluye las extensiones y bibliotecas generadas automáticamente y sin las cuales la aplicación no podrá ejecutarse, es decir, APLICACIÓN.exe



# Capítulo 6

## *Resultados*

### **6.1. Introducción**

Como ya se ha comentado anteriormente, la simulación de la acción muscular se ha llevado a cabo mediante blendshapes, correspondiendo cada uno a una AU, y siendo cada animación compleja el producto de distintas AUs.

En todo momento se trata de que la animación sea lo más realista posible y de reproducir de la forma más fielmente posible el comportamiento muscular. Sin embargo, se ha de tener en cuenta de que el método de animación empleado (blendshapes) simula la acción muscular en función de las posiciones inicial y final de los puntos, es decir, interpolando entre dichas posiciones.

### **6.2. Unidades de acción, AUs**

La disposición final del rostro propia de cada acción facial (AU) se ha realizado en función de la apariencia visual según la referencia presente en el FACS. A continuación se muestran los resultados obtenidos para dichas AUs:

- AU1: En el sistema de codificación de acciones faciales, FACS (capítulo 4), la unidad de acción 1 produce el levantamiento de la parte interior de las cejas, es decir, la región próxima al tabique nasal. Este movimiento se produce como consecuencia de la contracción voluntaria de la región central del músculo frontal.



Figura 6.1 –Imagen ejemplo de la AU1

La siguiente imagen muestra la posición de reposo y la posición final de la animación creada para esta AU. Podemos observar de esta manera la similitud entre la imagen ejemplo del FACS y la unidad de acción creada.



Figura 6.2 – Posición de reposo (izquierda) y posición final de la AU1 (derecha)

- AU2: De forma similar a la anterior, esta unidad de acción produce el levantamiento de la parte exterior de las cejas, y es consecuencia de la contracción voluntaria del músculo frontal en su parte externa.



Figura 6.3 –Imagen ejemplo de la AU2

Mostramos también las imágenes correspondientes a la posición de reposo y la posición final de esta unidad de acción, donde podemos compararla con la imagen ejemplo del FACS.



*Figura 6.4 – Posición de reposo (izquierda) y posición final de la AU2 (derecha)*

- AU4: Corresponde al movimiento de bajar las cejas, producido voluntariamente por los músculos corrugador y depresor superciliar. De la misma forma podemos comparar los resultados:



*Figura 6.5 –Imagen ejemplo de la AU4*



*Figura 6.6 – Posición de reposo (izquierda) y posición final de la AU4 (derecha)*

- AU5: Corresponde a la acción de levantar el párpado superior y es producida por el efecto del músculo elevador del párpado superior.



Figura 6.7 –Imagen ejemplo de la AU5



Figura 6.8 – Posición de reposo (izquierda) y posición final de la AU5 (derecha)

- AU6: Corresponde al movimiento de levantar las mejillas y se produce por la contracción del músculo orbicular del ojo.



Figura 6.9 –Imagen ejemplo de la AU6



Figura 6.10 – Posición de reposo (izquierda) y posición final de la AU6 (derecha)

- AU7: El movimiento corresponde al de estrechar los párpados y es producido por el mismo músculo que la unidad anterior, el orbicular del ojo.



Figura 6.11 –Imagen ejemplo de la AU7



Figura 6.12 – Posición de reposo (izquierda) y posición final de la AU7 (derecha)

- AU9: Producido por el músculo elevador de la nariz, produce el arrugamiento de la nariz.



Figura 6.13 –Imagen ejemplo de la AU9

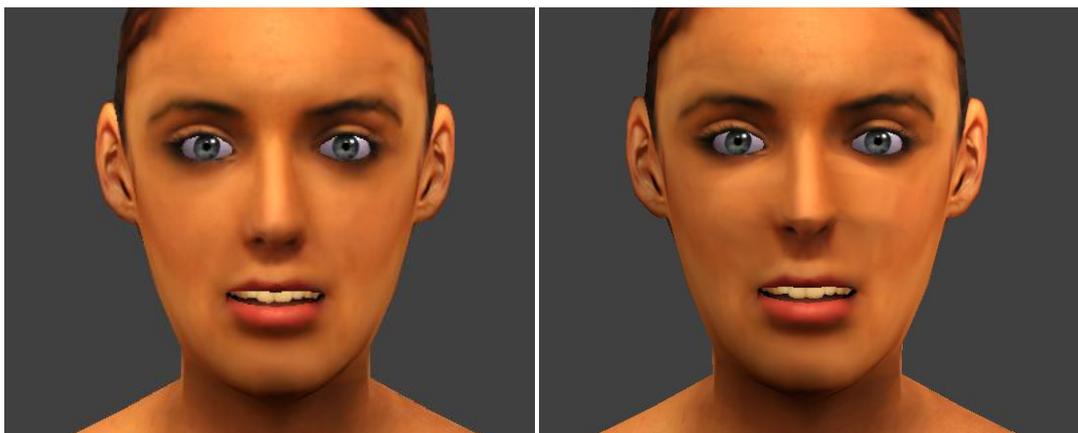


Figura 6.14 – Posición de reposo (izquierda) y posición final de la AU9 (derecha)

- AU10: Producido por los músculos alar y elevador de la nariz, produce la elevación del labio superior.



Figura 6.15 –Imagen ejemplo de la AU10



Figura 6.16 – Posición de reposo (izquierda) y posición final de la AU10 (derecha)

- AU11: Igualmente, es consecuencia del músculo elevador de la nariz y produce la elevación del labio superior con una acentuación del pliegue nasolabial.

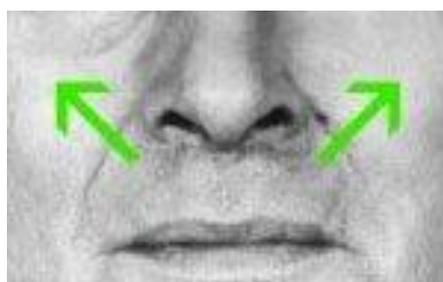


Figura 6.17 –Imagen ejemplo de la AU11

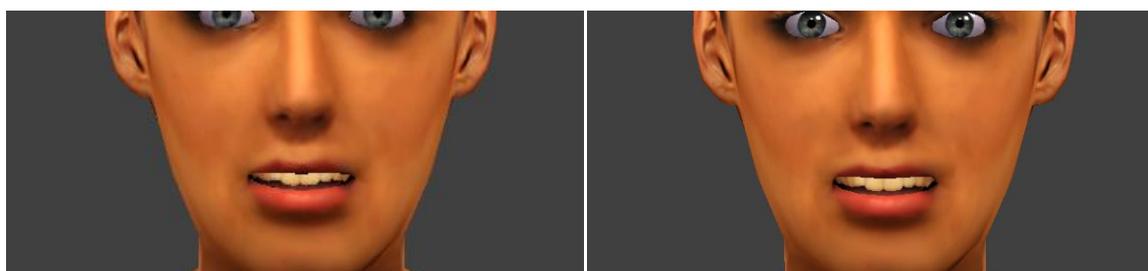


Figura 6.18 – Posición de reposo (izquierda) y posición final de la AU11 (derecha)

- AU12: Corresponde al movimiento de separación de las comisuras de la boca y es producido por el músculo cigomático mayor.



*Figura 6.19 –Imagen ejemplo de la AU12*



*Figura 6.20 – Posición de reposo (izquierda) y posición final de la AU13 (derecha)*

- AU15: Corresponde al movimiento de bajar las comisuras de la boca y es producido por el músculo triangular de la boca.



*Figura 6.21 –Imagen ejemplo de la AU15*



*Figura 6.22 – Posición de reposo (izquierda) y posición final de la AU15 (derecha)*

- AU17: Acción de levantar la barbilla, producido por la borla.



Figura 6.23 –Imagen ejemplo de la AU17

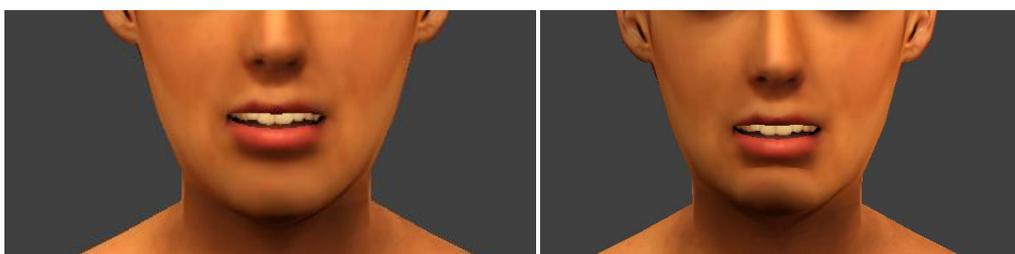


Figura 6.24 – Posición de reposo (izquierda) y posición final de la AU17 (derecha)

- AU24: Acción de apretar los labios, producido por el músculo orbicular de la boca.

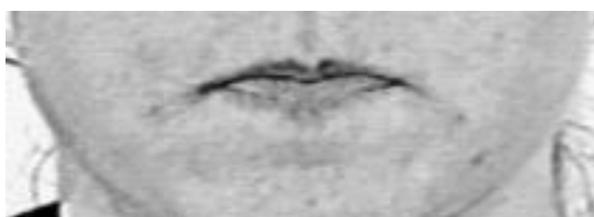


Figura 6.25 –Imagen ejemplo de la AU24

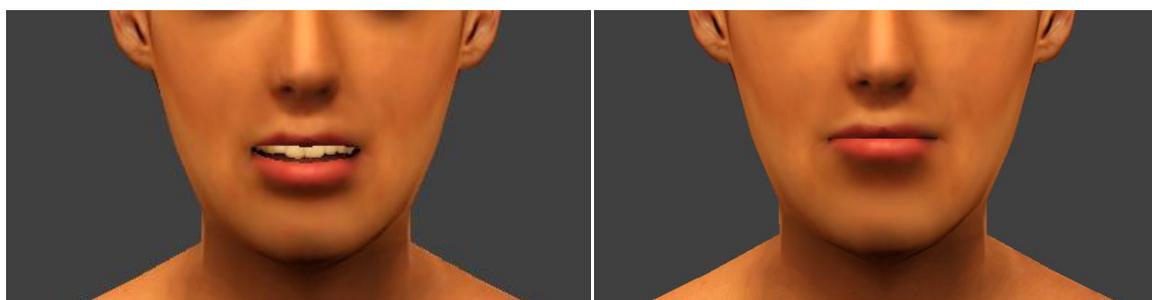


Figura 6.26 – Posición de reposo (izquierda) y posición final de la AU24 (derecha)

- AU25: Acción de separar los labios, producido por el orbicular de la boca.



Figura 6.27 –Imagen ejemplo de la AU25

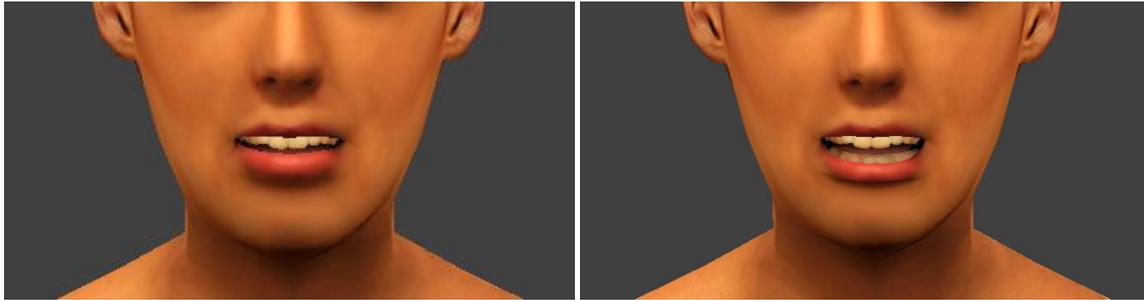


Figura 6.28 – Posición de reposo (izquierda) y posición final de la AU25 (derecha)

- AU26: Acción de bajar la mandíbula, producido por los músculos masetero, temporal y pterigoideo interno.



Figura 6.29 –Imagen ejemplo de la AU26

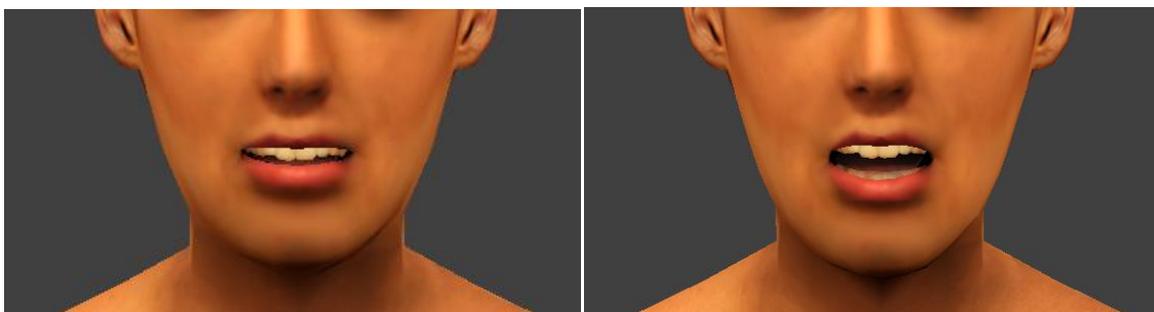


Figura 6.30 – Posición de reposo (izquierda) y posición final de la AU26 (derecha)

- AU29: Acción de apretar los labios y sacar la mandíbula inferior, producido por los músculos masetero y temporal.

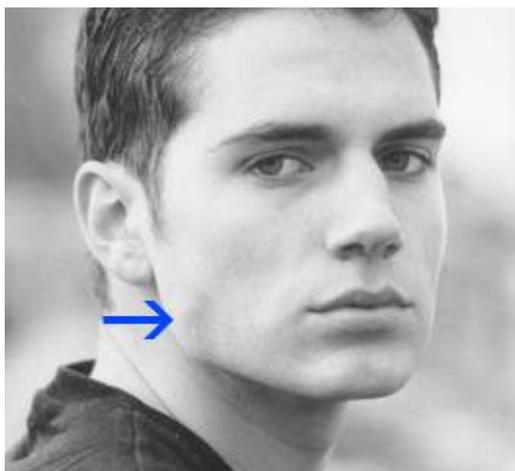


Figura 6.31 – Imagen ejemplo de la AU29

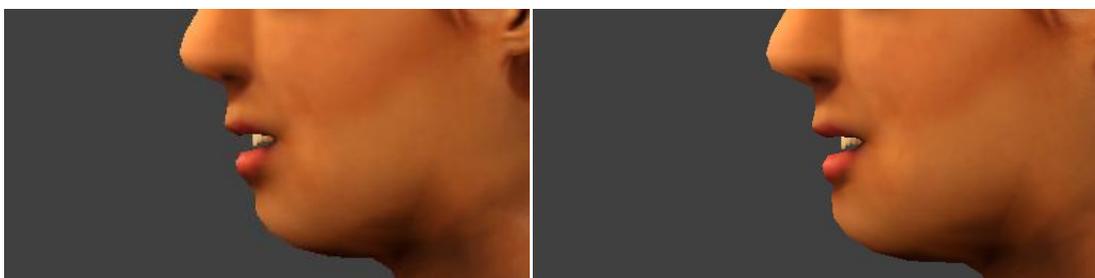


Figura 6.32 – Posición de reposo (izquierda) y posición final de la AU29 (derecha)

- AU38: dilatación de las fosas nasales producido por el alar de la nariz.



Figura 6.33 – Imagen ejemplo de la AU38

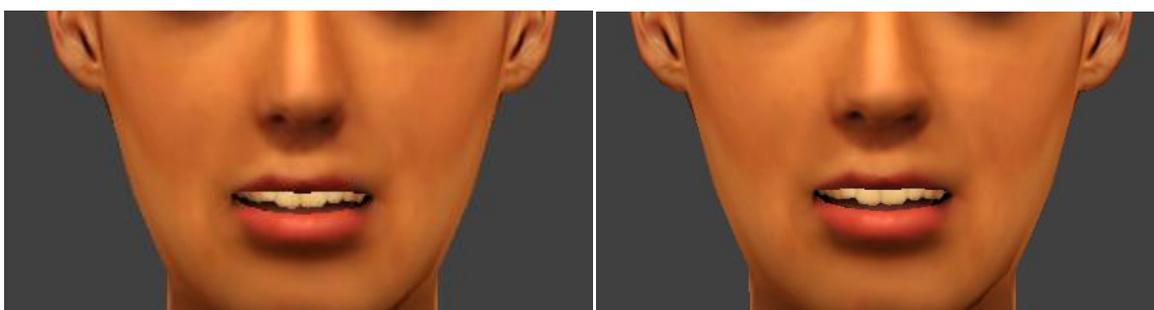


Figura 6.34 – Posición de reposo (izquierda) y posición final de la AU38 (derecha)

- AU45: Acción de parpadear, producido por el orbicular del ojo.



Figura 6.35 –Imagen ejemplo de la AU45



Figura 6.36 – Posición de reposo (izquierda) y posición final de la AU45 (derecha)

- AU51, AU52, AU53 y AU54: Giro del cuello a la izquierda, derecha, arriba y abajo, respectivamente. Producidos por los músculos del cuello.



Figura 6.37 –Imagen ejemplo de la AU51, AU52, AU53 y AU54, respectivamente.



Figura 6.38 – Posición final de las AU51 (a), AU52 (b), AU53 (c) y AU54 (d).

- AU61, AU62, AU63 y AU64: Mirada hacia la izquierda, derecha, hacia arriba y abajo, respectivamente. Producidos por los músculos extraoculares.



Figura 6.39 – Imagen ejemplo de la AU61, AU62, AU63 y AU64, respectivamente.



Figura 6.40 – Posición final de las AU61 (a), AU62 (b), AU63 (c) y AU64 (d).

### 6.3. Expresiones complejas

Como ya se comentó anteriormente, la generación de expresiones complejas se logra con la combinación de las distintas AUs elementales, cada una de ellas con un peso específico. Estos pesos son los que establecen la proximidad temporal (en número de frames) a las posiciones inicial y final generadas. De esta manera, una unidad de acción que no esté activa para una expresión concreta, se encontrará en el frame inicial del conjunto, pero le corresponderá un peso 0. De la misma forma, una unidad que tenga que activarse completamente para una expresión tendrá un peso 1.

A continuación indicaremos, para las seis expresiones complejas, las AUs que intervienen en cada una y su peso correspondiente:

- Alegría

AUs activadas	DESCRIPCIÓN	PESO ASOCIADO
AU2	Levantar el exterior de las cejas	60%
AU6	Levantar las mejillas	100%
AU12	Separar las comisuras	100%
AU24	Apretar los labios	90%
AU25	Separar los labios	40%
AU45	Parpadeo	95%
AU53	Girar el cuello hacia arriba	45%
AU64	Mirar hacia abajo	15%

Tabla 6.1 – AUs que intervienen en la expresión de alegría

- Tristeza

AUs activadas	DESCRIPCIÓN	PESO ASOCIADO
AU1	Levantar el interior de las cejas	100%
AU4	Bajar las cejas	100%
AU11	Elevar el labio superior	100%
AU15	Bajar las comisuras de los labios	90%
AU17	Levantar la barbilla	70%
AU24	Apretar los labios	100%
AU25	Separar los labios	90%
AU26	Bajar la mandíbula	80%
AU38	Dilatar las fosas nasales	45%
AU45	Parpadeo	100%
AU54	Girar el cuello hacia abajo	25%
AU63	Mirar hacia arriba	8%

Tabla 6.2 – AUs que intervienen en la expresión de tristeza

- Enfado

AUs activadas	DESCRIPCIÓN	PESO ASOCIADO
AU4	Bajar las cejas	100%
AU5	Levantar el párpado superior	100%
AU7	Estrechar los párpados	100%
AU10	Elevar el labio superior	100%
AU15	Bajar las comisuras de los labios	100%
AU17	Levantar la barbilla	90%
AU24	Apretar los labios	100%
AU29	Apretar los labios, sacar mandíbula	15%
AU45	Parpadeo	100%

AU54	Girar el cuello hacia abajo	45%
AU63	Mirar hacia arriba	20%

Tabla 6.3 – AUs que intervienen en la expresión de enfado

- Miedo

AUs activadas	DESCRIPCIÓN	PESO ASOCIADO
AU1	Levantar el interior de las cejas	100%
AU2	Levantar el exterior de las cejas	100%
AU4	Bajar las cejas	100%
AU5	Levantar el párpado superior	100%
AU10	Elevar el labio superior	100%
AU24	Apretar los labios	100%
AU25	Separar los labios	90%
AU26	Bajar la mandíbula	100%
AU38	Dilatar las fosas nasales	35%
AU45	Parpadeo	100%
AU51	Girar el cuello a la izquierda	40%
AU52	Girar el cuello a la derecha	35%
AU53	Girar el cuello hacia arriba	24%
AU61	Mirar hacia la izquierda	15%
AU62	Mirar hacia la derecha	15%
AU64	Mirar hacia abajo	85%

Tabla 6.4 – AUs que intervienen en la expresión de miedo

- Sorpresa

AUs activadas	DESCRIPCIÓN	PESO ASOCIADO
AU1	Levantar el interior de las cejas	100%
AU2	Levantar el exterior de las cejas	100%
AU5	Levantar el párpado superior	65%
AU7	Estrechar los párpados	50%
AU24	Apretar los labios	100%
AU26	Bajar la mandíbula	100%
AU45	Parpadeo	100%
AU53	Girar el cuello hacia arriba	28%
AU63	Mirar hacia arriba	8.5%
AU64	Mirar hacia abajo	6%

Tabla 6.5 – AUs que intervienen en la expresión de sorpresa

- Disgusto

AUs activadas	DESCRIPCIÓN	PESO ASOCIADO
AU4	Bajar las cejas	60%

AU5	Levantar el párpado superior	68%
AU7	Estrechar los párpados	100%
AU9	Arrugar la nariz	100%
AU10	Elevar el labio superior	100%
AU17	Levantar la barbilla	100%
AU24	Apretar los labios	100%
AU45	Parpadeo	100%
AU52	Girar el cuello a la derecha	20%
AU53	Girar el cuello hacia arriba	20%
AU61	Mirar hacia la izquierda	8.5%
AU63	Mirar hacia arriba	8.5%
AU64	Mirar hacia abajo	5.5%

Tabla 6.6 – AUs que intervienen en la expresión de disgusto

Hay que tener en cuenta, al observar estas tablas, que no se dan todas a la vez, si no que se dan a lo largo de la animación durante tiempos limitados, y que el peso especificado corresponde al punto de la animación donde más intensidad tiene la unidad de acción.

El resultado visual final de las expresiones puede verse en el apartado siguiente, donde se realiza una comparación con la foto real.

#### 6.4. Resultados comparativos

A continuación se realizará una comparación de las fotos reales y los resultados finales de cada expresión.

- Alegría

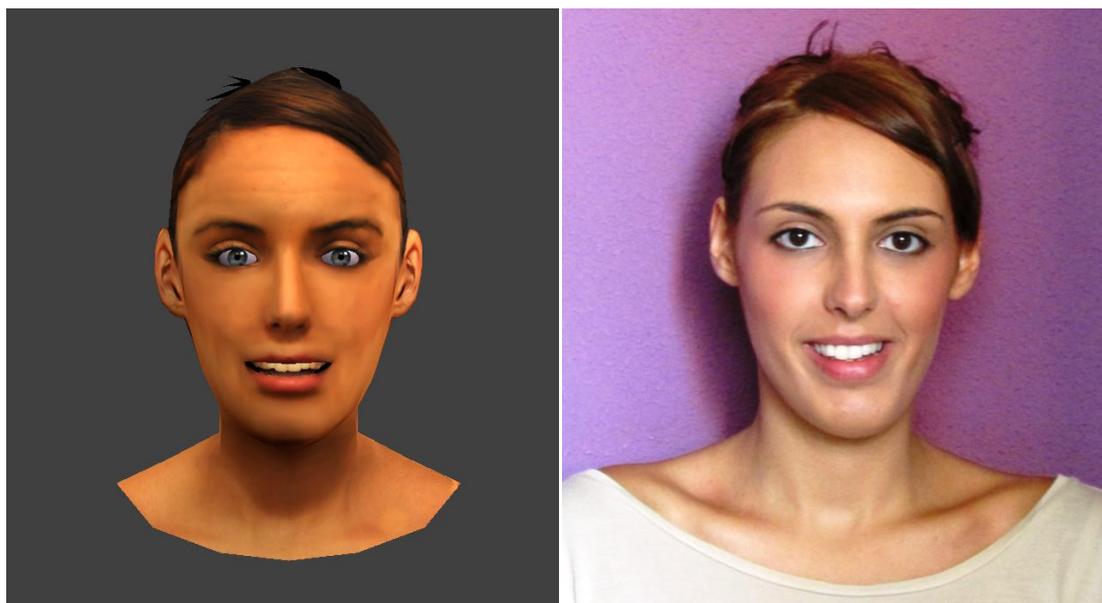
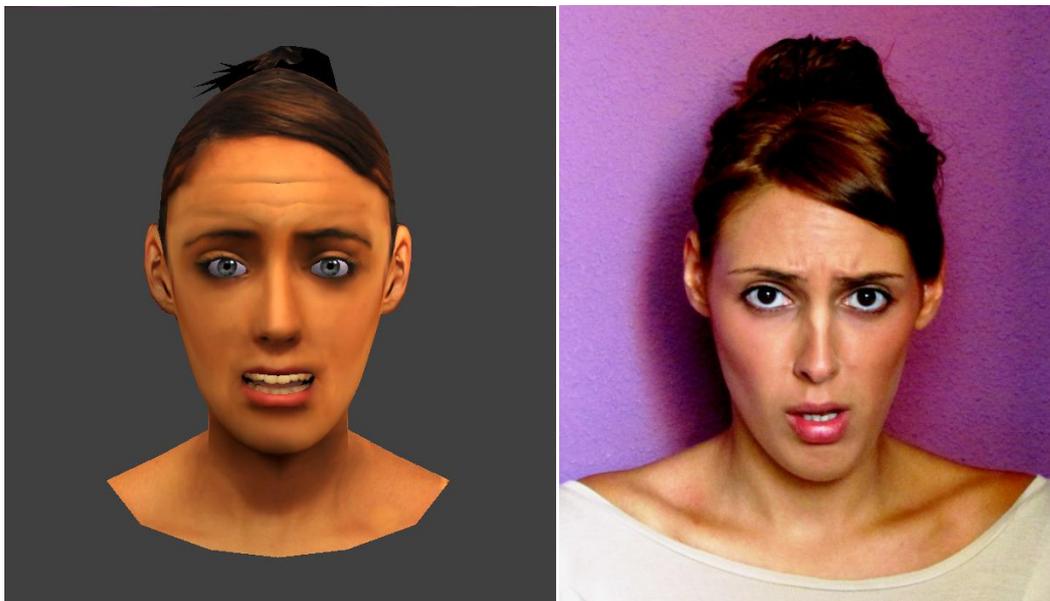


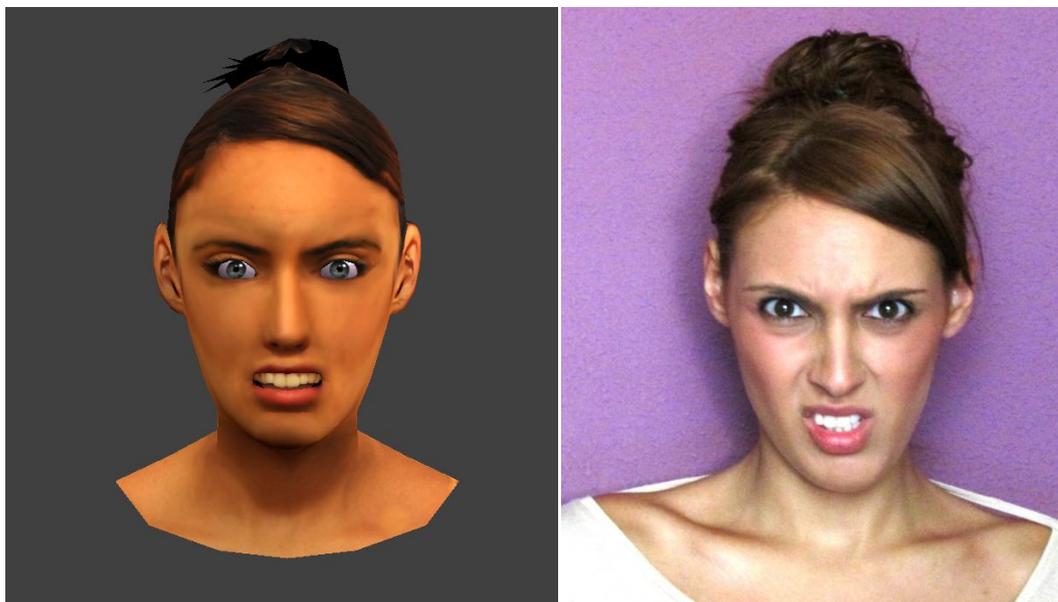
Figura 6.41 – Comparación resultado animación e imagen real de alegría.

- Tristeza



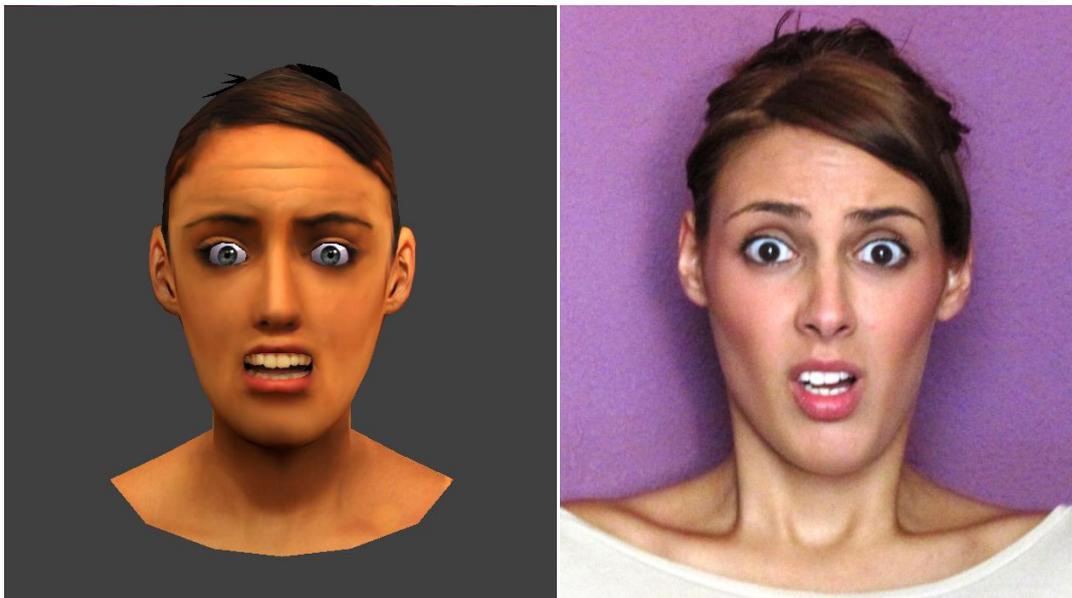
*Figura 6.42 – Comparación resultado animación e imagen real de tristeza.*

- Enfado



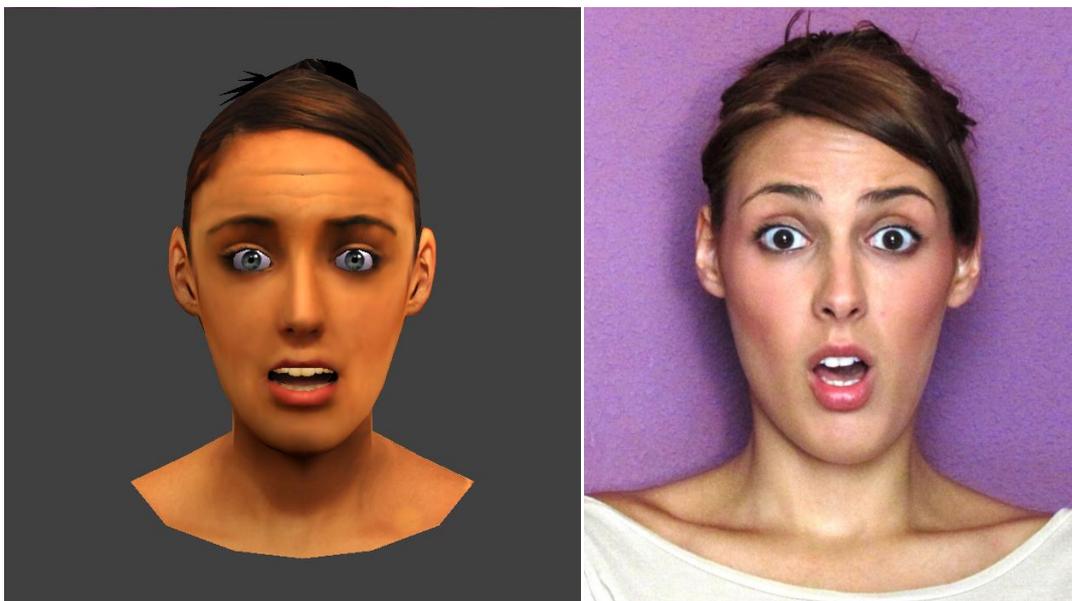
*Figura 6.43 – Comparación resultado animación e imagen real de enfado.*

- Miedo



*Figura 6.44 – Comparación resultado animación e imagen real de miedo.*

- Sorpresa



*Figura 6.45 – Comparación resultado animación e imagen real de sorpresa.*

- Disgusto

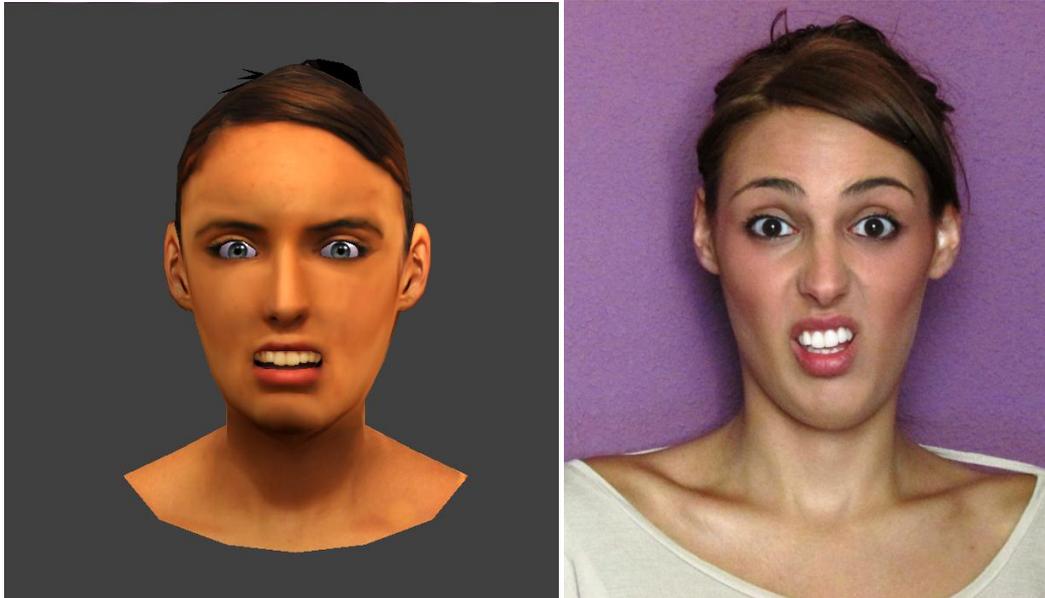


Figura 6.46 – Comparación resultado animación e imagen real de disgusto.

## 6.5. Resultados numéricos

Los resultados detallados a continuación (*Tabla 6.7*) se fundamentan en las pruebas realizadas para la aplicación final generada en dos ordenadores de sobremesa. El uso de recursos de la aplicación es un aspecto importante a tener en cuenta si se pretende que el sistema desarrollado pueda ser integrado en otra aplicación. Se ha tenido en cuenta también el consumo de RAM y CPU.

Recordemos, para comprender la *Tabla 6.7*, que cuando hablamos de frame nos referimos a una imagen o fotograma particular dentro de una sucesión de imágenes que conforman una animación. La continúa sucesión de estos fotogramas produce a la vista la sensación de movimiento. La frecuencia, FPS (frames per second), es el número de fotogramas por segundo que se necesitan para crear movimiento.

CONFIGURACIÓN	SISTEMA 1		SISTEMA 2	
<b>CPU</b>	Intel Core 2 Dúo Solo Processor SU3500 (1.4GHz, 800MHz FSB)		Intel Core i5 (5.67 GHz)	
<b>RAM</b>	4GB DDR3		4GB	
<b>GFX</b>	Intel GMA4500MHD. Up to 1244MB		Radeon	
<b>HDD</b>	320 GB HDD		500 GB	
<b>SO</b>	Windows 7		Windows 7	
MEDIA (FPS) RESOLUCIÓN	REPOSO	ANIMACIÓN	REPOSO	ANIMACIÓN
<b>640x480</b>	60.2 fps	59.9 fps	60.0 fps	60.1 fps
<b>800x600</b>	60.1 fps	59.6 fps	59.8 fps	59.9 fps
<b>1150x1080</b>	60.0 fps	59.6 fps	59.6 fps	59.4 fps

CONSUMO	REPOSO	ANIMACIÓN	REPOSO	ANIMACIÓN
CPU	0.17	19.6	0.15	18.8
RAM	150.23	197.2	137.47	184.76

Tabla 6.7 – Consumo y recursos de la aplicación

ESTADÍSTICAS DEL MODELO	
Vértices	1805
Caras	1555
UV	1555
Tamaño archivo .blend	16.1 MB
Tamaño ejecutable	29.3 MB

Tabla 6.8 – Estadísticas del modelo

Se va a realizar una breve comparación sobre las características de este proyecto y el precedente a éste:

- El número de vértices y de caras es inferior a la mitad que en el proyecto anterior, por lo que el gasto computacional es menor. Por el mismo motivo, el sistema se podrá implementar en un ordenador menos potente que el del proyecto anterior.
- Los tamaños tanto del archivo como del ejecutable son muy inferiores en este proyecto.

## 6.6. Conclusiones

En este capítulo se han mostrado los resultados, tanto visuales como estadísticos, de la aplicación creada.

Se ha podido comprobar fielmente la similitud con la realidad de las imágenes. Esto se debe principalmente al modelado manual en Blender y a las opciones que ofrece esta herramienta. Sin embargo, resulta muy complicado modelar la deformación de las regiones de la piel y de la formación de arrugas de forma totalmente correcta.

Otra característica de la aplicación es su corto tiempo de respuesta. Las expresiones se generan de forma inmediata tras pulsar en el teclado la tecla asociada correspondiente, por lo que la aplicación resultaría idónea para ser utilizada de forma interactiva.

Por otro lado, los recursos requeridos para la implementación de esta aplicación no son determinantes, lo cual es un valor añadido ya que es posible ejecutar la aplicación sin necesidad de capacidades computacionales grandes.

Por último, señalar que el método de animación usado simplifica el proceso de generación de expresiones, es decir, combinando las distintas unidades mínimas de acción originamos expresiones más complejas. Partiendo de esta idea, es posible ampliar el rango de expresiones solamente variando las unidades mínimas y su intensidad o peso.

# Capítulo 7

## *Evaluación del avatar*

### **7.1. Introducción**

Una vez que el avatar ha sido desarrollado y preparado para su integración, la siguiente etapa debe ser evaluar las reacciones de los usuarios al interactuar con él.

Cuando un avatar animado se diseña para interactuar con humanos, existen numerosos aspectos que influyen en la naturalidad con la que es capaz de comunicarse. Este hecho se acentúa aún más cuando el avatar presenta un rostro realista, ya que los humanos somos verdaderamente sensibles a las diferentes características del rostro, [Mo82].

Existe una gran cantidad de estudios que se centran en los aspectos no verbales de la interacción, en la expresividad y en la apariencia de los avatares. Por ejemplo, en [Go08] se estudia la hipótesis de que aquellos avatares con mayor grado de antropomorfismo obtienen una mayor respuesta social de los usuarios. En su experimento se presenta a los participantes un total de 12 avatares con diferentes grados de antropomorfismo demostrando que aquellos avatares más realistas son capaces de expresar un mejor comportamiento social desde el punto de vista del usuario. Griesser [Gr07] estudió los efectos de la variación de un conjunto de regiones faciales y de la combinación de dichas regiones en la expresividad de un avatar. Sus resultados muestran la importancia y la interrelación que las diferentes regiones faciales estudiadas tienen en la expresividad del avatar. Otros estudios que se centran en la expresividad de los avatares pueden verse en [Wa06c] y [Co04b], donde se evalúa la calidad de las animaciones generadas en función de la tasa de acierto que tienen los participantes al identificar expresiones complejas realizadas por diferentes avatares. En la evaluación de aspectos no verbales, existen estudios que se centran en aspectos mucho más específicos. En [Ta08], por ejemplo, se realiza un estudio del efecto que tiene en los participantes la tasa de parpadeo de diferentes avatares, de forma que un parpadeo similar al humano es muy positivo en la percepción de los usuarios hacia el avatar.

Según [Ca99], los humanos hacen un gran uso del canal visual en la interacción, y son capaces de extraer información de pequeñas señales obtenidas de la mirada de su interlocutor. En situaciones colaborativas, la mirada es un elemento fundamental del comportamiento que complementa diferentes actividades y roles.

Con el siguiente experimento realizado se ha tratado de evaluar el avatar completo, teniendo en cuenta las diferentes características otorgadas. La finalidad del estudio ha sido obtener conclusiones sobre la percepción de los usuarios cuando son enfrentados a un avatar animado que emplea un método de interacción basado en los fundamentos de la robótica social.

## **7.2. Método**

### **7.2.1. Diseño**

El experimento realizado ha sido del tipo entre sujetos o de muestras no relacionadas. Los participantes han observado el comportamiento del avatar identificando las expresiones que este genera, tanto estática como dinámicamente.

Uno de los problemas que puede presentarse cuando se diseña un experimento del tipo entre sujetos es que una diferencia obtenida en la variable dependiente o variable a estudio podría deberse al hecho mismo de emplearse distintos grupos de sujetos (con diferente edad, estado de ánimo, nivel de formación, etc.) y dichas diferencias puedan confundirse con la variable independiente de interés (es decir, la variable que se modifica para observar sus efectos en la o las variables dependientes). La distribución aleatoria de sujetos entre grupos o condiciones permite minimizar dicha influencia. De esta forma, en promedio, los grupos tienen características similares y puede manipularse la variable independiente (observar e identificar las emociones generadas por el avatar) sin la influencia de agentes derivados de la composición de los grupos a estudio.

Con la realización de un conjunto de test preliminares se ha identificado que los usuarios que conocen a la persona modelada tienen una primera reacción muy diferente que aquellas personas que desconocen totalmente a la persona que estaba en pantalla. Dicho efecto indica que podría existir cierta dependencia entre conocer a la persona cuyo rostro estaba representado por el avatar y la reacción que se tiene con respecto al mismo, lo que ha llevado a formular la siguiente hipótesis para el experimento: el haber mantenido cierta interacción previa con la persona real modelada en el avatar puede modificar la percepción y sensaciones del usuario que interactúa con él.

La principal finalidad del estudio consiste en determinar si los observadores del avatar son capaces de identificar su comportamiento y si son capaces de reconocer las emociones básicas que éste genera. La evaluación se ha centrado en dos aspectos fundamentales: la apariencia del avatar tanto estática como dinámicamente y en la capacidad de los sujetos de reconocer las expresiones que el avatar producía en ambos casos. Para ello se ha empleado un cuestionario en el que se solicitaba a los usuarios que identificaran las diferentes expresiones del avatar y siete preguntas referidas a su percepción de este.

Una práctica común en los estudios psicológicos es organizar las preguntas de un cuestionario en grupos para su análisis, también llamados constructores, de forma que cada grupo de preguntas esté relacionado con una característica que se quiere evaluar. De esta forma, se considera cada agrupación de preguntas o constructor como un único tema de estudio, y puede enfocarse la formulación de las preguntas de un constructor para que abarquen todos los aspectos relacionados con el concepto que representa. En el cuestionario realizado, las 7 preguntas se han repartido en 3 constructores que representan cada uno de los aspectos a evaluar: expresividad, expectativa de rendimiento y apariencia. En el caso de la expresividad, se han planteado tres preguntas que evalúan la movilidad completa del avatar al gesticular, y otra que evalúa la capacidad del avatar de realizar expresiones reconocibles por los usuarios. Para evaluar la apariencia, se ha planteado una cuestión referida a si les satisface la apariencia del mismo y para

evaluar la expectativa del rendimiento, una cuestión sobre si las animaciones mejoraban la interacción hombre-máquina.

La *Tabla 7.1* muestra las preguntas del cuestionario junto con los constructores creados y los grupos de preguntas que los evalúan. Mostraremos además el resultado del alfa de Cronbach, que será explicado y discutido en la sección de análisis. Las respuestas a dichas preguntas se han recogido en una escala de Likert de cinco niveles, entre —totalmente de acuerdo—totalmente en desacuerdo.

### **EVALUACIÓN DE AVATAR VIRTUAL REALISTA ANIMADO**

Sexo	Masculino	
	Femenino	
Edad	Menos de 20	
	Entre 20 y 40	
	Entre 40 y 60	
	Más de 60	
Conocimientos Informáticos	Nunca he usado un ordenador	
	Sólo para navegar por internet	
	Experiencia ofimática	
	Avanzado	
Conocimientos de la persona modelada	No conozco a esa persona	
	Le conozco de vista	
	Nos han presentado	
	Le conozco desde hace tiempo	

**A continuación indique qué expresión de las generadas por el avatar ha reconocido y en caso que no sea ninguna de las recogidas en la tabla, indique el número de la expresión y la emoción que ha identificado.**

RECONOCIMIENTO EMOCIONES								
APARIENCIA	Alegría	Tristeza	Miedo	Enfado	Sorpresa	Disgusto	Neutro	Otro
Expresión 1								
Expresión 2								
Expresión 3								
Expresión 4								
Expresión 5								
Expresión 6								
Expresión 7								
EXPRESIVIDAD	Alegría	Tristeza	Miedo	Enfado	Sorpresa	Disgusto	Neutro	Otro
Expresión 1								
Expresión 2								
Expresión 3								
Expresión 4								
Expresión 5								
Expresión 6								

**Valore los siguientes aspectos en una escala del 1 al 5, siendo:**  
**1- Muy poco 2- Poco 3- Regular 4- Aceptable 5-Muy aceptable**

REALISMO (Constructores)					
APARIENCIA Y EXPRESIVIDAD DEL AVATAR	1	2	3	4	5
Me ha gustado la apariencia del Avatar					
Me ha parecido una persona real					
Los movimientos del Avatar me han parecido naturales					
Me ha gustado cómo se movía el Avatar					
Me ha parecido que se movía a una velocidad adecuada					
En general he reconocido las expresiones que ha realizado					
Creo que el movimiento ayuda a mejorar la interacción hombre-computador					
Observaciones y comentarios:					

Tabla 7.1 – Cuestionario

Los cuestionarios además incluyen preguntas de tipo socio – cultural que piden a los participantes indicar su edad, sexo y conocimientos informáticos. Para la edad los participantes pueden seleccionar su respuesta en una escala de cuatro niveles: menos de 20, entre 20 y 40, entre 40 y 60 y más de 60. Los conocimientos informáticos también se han recogido en cuatro niveles: nunca ha usado un ordenador, sólo para navegar por internet, experiencia ofimática y conocimientos avanzados. En el caso del avatar animado, se ha incluido una pregunta para saber si los participantes habían tenido algún tipo de contacto previo con la persona modelada. Esta pregunta permite las siguientes respuestas: no, le conozco de vista, nos han presentado o le conozco desde hace tiempo.

### 7.2.2. Participantes

El avatar ha sido presentado a diferentes grupos de personas que acudían a la Escuela de Ingenierías Industriales de la Universidad de Valladolid, así como a otros colectivos como familiares o amigos.

Se ha realizado una primera encuesta a un grupo de 40 personas (GRUPO A) y se han estudiado los resultados de tales encuestas. Teniendo en cuenta las opiniones de dichas personas, se ha procedido a mejorar el avatar a partir de la realimentación recibida de este grupo. Posteriormente se ha vuelto a realizar la misma encuesta a otras 30 personas (GRUPO B), intentando que el grupo fuera lo más parecido al primero en cuanto a edad, sexo, etc.

En total se ha realizado el experimento sobre 70 sujetos. La distribución completa por edades, sexo, conocimiento informático y el haber interactuado con la persona modelada en el avatar 3D se muestran en la *Tabla 7.2* obtenidas de las respuestas del cuestionario.

	<b>POSIBLES RESPUESTAS</b>	<b>FRECUENCIA</b>	<b>PORCENTAJE</b>
<b>SEXO</b>	Masculino	37	52.94
	Femenino	33	47.06
<b>EDAD</b>	Menos de 20	6	8.82
	Entre 20 y 40	31	44.12
	Entre 40 y 60	31	44.12
	Más de 60	2	2.94
<b>CONOCIMIENTOS INFORMÁTICOS</b>	Nunca he usado un ordenador	2	2.94
	Sólo para navegar por internet	14	20.58
	Experiencia ofimática	27	38.24
	Avanzado	27	38.24
<b>CONOCIMIENTO DE LA PERSONA MODELADA</b>	No conozco a esa persona	46	65.71
	Le conozco de vista	8	11.43
	Nos han presentado	6	8.57
	La conozco desde hace tiempo	10	14.29

*Tabla 7.2 – Distribución de participantes según sus características individuales*

### **7.2.3. Materiales y Aparatos**

Para el experimento, se ha empleado un ordenador portátil convencional con un monitor de 15.6 pulgadas y la aplicación del Avatar desarrollada en la plataforma Blender descrita en el capítulo 5 y cuyo resultado final puede verse en la *Figura 5.21*. Para que resultara más sencillo pautar el orden del cuestionario y fuera el mismo en todos los casos, se ha generado un archivo PowerPoint en el que se muestran primero las imágenes y posteriormente videos de cada una de las animaciones.

### **7.2.4. Procedimiento**

El cuestionario fue presentado a diferentes grupos que acudieron durante un periodo de una semana a la escuela de Ingenierías Industriales, así como a colectivos cercanos como familiares y amigos. A aquellas personas que accedían a tomar parte del experimento se les mostraba inicialmente una serie de imágenes estáticas en las que se mostraba al avatar expresando las emociones básicas universales: alegría, tristeza, miedo, sorpresa, enfado, disgusto y neutro, y a continuación las animaciones compuestas en las que el avatar generaba dichas emociones. Posteriormente, el sujeto debía rellenar el cuestionario donde constaban dos tablas como las mostradas en la *Tabla 7.1* señalando la emoción que había identificado en cada caso, además de completar el cuestionario con las preguntas de percepción sobre la apariencia y expresividad del avatar.

Se han realizado dos análisis sobre los datos obtenidos. En el primero se ha tratado de investigar qué características principales tenían algún tipo de influencia en las sensaciones que los participantes tuvieron al observar el avatar. En el segundo se ha tratado de comprobar si las

emociones generadas por el avatar son fácilmente identificables por los usuarios tanto estáticamente como en movimiento.

Como se describirá en la siguiente sección, se ha encontrado una diferencia evidente entre los participantes que conocen a la persona modelada y los que no han tenido ningún contacto previo con ella.

### 7.3. Análisis y resultados

Para determinar la consistencia de los constructores considerados, es decir, si los grupos de preguntas realmente expresan un mismo concepto, se ha realizado un test del Alfa de Cronbach. El Alfa de Cronbach es un índice ampliamente utilizado en psicometría y que sirve para comprobar la consistencia interna del instrumento que se está evaluando, es decir, si recopila información defectuosa y por tanto nos llevaría a conclusiones equivocadas o si se trata de un instrumento fiable que hace mediciones estables y consistentes. El alfa es un coeficiente de correlación al cuadrado que toma valores entre 0 y 1 y, a grandes rasgos, mide la homogeneidad de un conjunto de preguntas promediando todas las correlaciones entre todos los ítems para ver que, efectivamente, se parecen. La interpretación del Alfa de Cronbach supone que cuanto más se acerque el índice al extremo 1, mejor es la fiabilidad, considerando una fiabilidad correcta a partir de 0.80. Sin embargo, cuando se usa en experimentos psicológicos, un resultado superior a 0.7 se considera aceptable. Las condiciones para poder realizar el test del Alfa de Cronbach sobre un conjunto de preguntas son que:

- Debe estar formado por un conjunto de ítems (preguntas) que se combinan aditivamente para hallar una puntuación global (esto es, la puntuaciones se suman y dan un total que es el que se interpreta).
- Todos los ítems miden la característica deseada en la misma dirección. Es decir, los ítems de cada una de las escalas tienen el mismo sentido de respuesta (a mayor puntuación, más expresividad, por ejemplo).

Tal y como se muestra en la *Tabla 7.3*, las puntuaciones para los constructores fueron mayor que 0.7, excepto en el caso de la apariencia en el primer grupo. En el segundo grupo, y habiendo mejorado el avatar a partir de los resultados del primero, se observa una sensible mejora de las puntuaciones, indicando que los grupos de preguntas creados tienen consistencia interna y por tanto representan el mismo concepto.

CONSTRUCTOR	PREGUNTA	ALFA DE CRONBACH GRUPO A	ALFA DE CRONBACH GRUPO B
EXPRESIVIDAD	Los movimientos del avatar me han parecido naturales	0.74	0.8315
	Me ha gustado cómo se movía el avatar		
	Me ha parecido que se movía a una velocidad adecuada		
	En general he reconocido las expresiones que ha realizado		
APARIENCIA	Me ha gustado la apariencia del avatar	0.6	0.7527
	Me ha parecido una persona real		

*Tabla 7.3 – Preguntas realizadas para la evaluación, con los resultados del Alfa de Cronbach.*

Para determinar si alguno de los factores diferenciales a nivel individual, sexo, edad, conocimientos informáticos o el hecho de conocer a la persona modelada influía en los resultados obtenidos, se ha realizado un análisis ANOVA entre sujetos, tanto para cada uno de los constructores como para el reconocimiento de cada una de las emociones. El análisis ANOVA permite determinar si alguna variable de un conjunto de variables independientes tiene algún efecto sobre una variable dependiente (lo que se conoce como efectos principales) y además indica si dos o más variables independientes tienen algún tipo de interacción entre sí.

Los resultados relevantes del ANOVA se resumen en la *Tabla 7.4*. La columna fuentes de variación muestra únicamente las variables e interacciones entre variables que obtuvieron resultados significativos. Además se muestra los niveles de significación (también llamado p-valor). Un factor es significativo si su p-valor es menor que el nivel de significación considerado, en éste caso de 0.05. La última columna muestra los grados de libertad entre grupos. Para un efecto principal, el número de grados de libertad entre grupos es el número de niveles de la variable menos 1 (por ejemplo, para la variable sexo existen dos niveles: hombre y mujer, por lo que se tiene un grado de libertad). Para un efecto de interacción, el número de grados de libertad entre grupos viene dado por el producto de los grados de libertad de los efectos principales.

Tal y como se muestra en la *Tabla 7.4*, para los constructores de expresividad y apariencia para el avatar animado existen dos efectos principales significativos: la edad y conocer a la persona modelada. Respecto al reconocimiento de las emociones los efectos principales son el sexo y conocer a la persona. Para el resto de efectos de interacción mostrados, no existen suficientes grados de libertad para considerarlos significativos

	<b>FUENTE DE VARIACIÓN</b>	<b>Significación para un nivel 0.05</b>	<b>Grados de libertad</b>
<b>EXPRESIVIDAD</b>	Conocer * Edad * Informática	0.04	1
<b>APARIENCIA</b>	Conocer*Edad	0.039	1
<b>RECONOCIMIENTO DE EMOCIONES</b>	Sexo	0.015	1
	Conocer	0.029	3
	Sexo*Informática	0.049	1

*Tabla 7.4 –Resultados significativos del ANOVA. Se muestran los factores principales y de interacción estadísticamente significativos, junto con el nivel de significación y los grados de libertad del análisis.*

Sin embargo, un análisis en profundidad de los datos mostró ciertas tendencias en los diferentes grupos de participantes considerados: aquellas personas con mayor conocimiento informático dieron mayores puntuaciones al avatar y reconocieron mejor las emociones que expresaba. En el caso de la edad, los participantes más jóvenes lo valoraron de forma más positiva que aquellos de edad más avanzada. En el caso del sexo, las diferencias en las puntuaciones no fueron tan evidentes como en el resto de grupos, aunque se observa cómo las puntuaciones de los hombres son ligeramente superiores que las de las mujeres. Y, como ya hemos comentado, los participantes que conocían a la persona modelada dieron mayores puntuaciones y reconocieron mejor las expresiones dinámicas del avatar.

Por último, vamos a mostrar en la *Tabla 7.5* y *Tabla 7.6* el porcentaje de aciertos en la identificación de cada expresión tanto estática como dinámicamente, tratando la muestra global de personas.

<b>% RECONOCIMIENTO DE IMÁGENES ESTÁTICAS</b>																
	ALEGRÍA		TRISTEZA		MIEDO		ENFADO		DISGUSTO		SORPRESA		NEUTRA		OTRO	
ALEGRÍA	64.8	85	2.9	-					2.9	-			20.6	12.5	8.8	2.5
TRISTEZA			67.7	87.5					2.9				29.4	12.5		
MIEDO					70.7	85	2.9	2.5	20.6	2.5	2.9	7.5			2.9	-
ENFADO							70.6	90	26.5	10			2.9	-		
DISGUSTO					5.9	2.5	17.6	7.5	73.6	90			2.9	-		
SORPRESA											97.1	97.5			2.9	2.5
NEUTRO							5.9	5					88.2	95	2.9	-

Tabla 7.5 – % reconocimiento de imágenes estáticas.

<b>% RECONOCIMIENTO DE ANIMACIONES DINÁMICAS</b>																
	ALEGRÍA		TRISTEZA		MIEDO		ENFADO		DISGUSTO		SORPRESA		NEUTRA		OTRO	
ALEGRÍA	85.4	92.5	2.9	-	2.9	-							5.9	5.0	2.9	2.5
TRISTEZA			64.7	85.0	8.8	5.0					11.8	7.5	8.8	-	5.9	2.5
MIEDO					76.5	85.0	11.8	7.5	8.8	5.0	2.9	2.5				
ENFADO					8.8	10	53.0	72.5	35.3	15	2.9	2.5				
DISGUSTO					2.9	2.5	5.9	2.5	91.2	95						
SORPRESA											97.1	97.5			2.9	2.5

Tabla 7.6 – % reconocimiento de animaciones dinámicas.

## 7.4. Discusión

De un primer análisis global de resultados, puede observarse que el avatar animado tiene un rendimiento aceptable en los diferentes tópicos estudiados, por lo que puede considerarse ventajoso su uso.

Los sujetos identificaron por lo general correctamente las expresiones mejor en las animaciones que observando únicamente las imágenes, o los datos no variaban entre unas y otras, a excepción de la expresión de enfado, que en las animaciones esta expresión resultó más complicada de identificar. Esto puede ser debido a que los rasgos de la expresión son parecidos a la de disgusto, que es con la que más se ha confundido. Los resultados son: Alegría- 64.8% estático 85.4% dinámico, Tristeza- 67.7% estático 64.7% dinámico, Miedo 70.7% estático 76.5% dinámico, Enfado 70.6% estático 53% dinámico, Disgusto 73.6% estático 91.2% dinámico, Sorpresa 97.1% estático 97.1% dinámico. Además, la mayoría de los encuestados coincidieron en que las animaciones ayudaban a reconocer las expresiones de una manera más sencilla empleando las animaciones.

Como se puede observar, se obtuvieron mejores resultados en cuanto a la expresividad y las expectativas de rendimiento en el reconocimiento de animaciones dinámicas, debido a que la sincronización del movimiento proporciona una percepción más real del avatar. Por lo general, los resultados que muestra el primer estudio indican que el avatar realizaba emociones fácilmente identificables (Alegría- 85.4%, Tristeza- 67.7%, Miedo 76.5% dinámico, Enfado 53%, Disgusto 91.2%, Sorpresa 97.1%).

El conocer previamente a la persona modelada supuso un incremento en las puntuaciones en todos los apartados considerados, pero fundamentalmente en apariencia, lo que lleva a concluir que el avatar reproduce fielmente los gestos reales de la persona y que el haber mantenido contacto previo con ella es un factor clave en la percepción que tienen los usuarios del avatar, ya que entender los gestos de una persona resulta más fácil cuando se la conoce previamente y se tiene cierta familiaridad con sus gestos.

Aunque el análisis no mostró diferencias significativas en otros apartados, se observaron ciertas tendencias en las puntuaciones de los participantes en función de los parámetros socioculturales. Cuanto mayor es la experiencia que los participantes tienen en el manejo de ordenadores, las puntuaciones obtenidas por el avatar son más altas. Los participantes que emplean ordenadores para realizar sus tareas cotidianas parecen más acostumbrados al uso de nuevas tecnologías, y a aceptar la ayuda de este tipo de interfaces para la realización de tareas.

Existe además una relación inversamente proporcional entre la edad de los participantes y los resultados obtenidos. Este hecho puede deberse a que las personas más jóvenes están más acostumbrados a ver avatares y robots en videojuegos, películas, internet, etc., por lo que se sienten más cómodos al interactuar con ellos. En el caso del género, las diferencias entre hombres y mujeres no son tan destacables como en el resto de grupos considerados, aunque los participantes masculinos dieron puntuaciones ligeramente superiores.

Los participantes que habían tenido interacción previa con la persona modelada valoraron al avatar de forma más positiva en las preguntas relacionadas con la expresividad.

Por lo general, los encuestados coincidieron en que aunque las animaciones resultaran ventajosas para identificar las expresiones, consideraban que el avatar realizaba los movimientos muy lentos y aumentar la velocidad lo harían más dinámico y el comportamiento más realista.

Es por ello que se decidió realizar un segundo estudio modificando las animaciones que habían sido más difíciles de identificar, como Enfado (53%) o Tristeza (67.7%), además de realizar todas las animaciones más dinámicas aumentando la velocidad del avatar al mostrar expresión.

Los resultados de esta segunda encuesta muestran resultados mucho más favorables tanto estática como dinámicamente. A continuación se puede observar la mejora de esta segunda encuesta:

- Imágenes:	
Alegría	64.8% → 85.0%
Tristeza	67.7% → 87.5%
Miedo	70.7% → 85.0%
Enfado	70.6% → 90.0%
Disgusto	73.6% → 90.0%
Sorpresa	97.1% → 97.5%
Neutra	88.2% → 95.0%
- Animaciones	
Alegría	85.4% → 92.5%
Tristeza	64.7% → 85.0%
Miedo	76.5% → 85.0%
Enfado	53.0% → 72.5%
Disgusto	91.2% → 95.0%
Sorpresa	97.1% → 97.5%

Se puede observar también en las *Tablas 7.5 y 7.6* el porcentaje de resultados de las expresiones confundidas. Estos porcentajes también son menores en el caso de las animaciones que en las imágenes, demostrando una vez más cómo estas hacían más fácilmente identificables las expresiones. En líneas generales, alegría fue confundida con expresión neutra tanto en las imágenes (12.5%) como en las animaciones (5.0%), tristeza con expresión neutra en las imágenes y en las animaciones con sorpresa o miedo (5% y 7.5% respectivamente). Miedo por otro lado fue confundido con enfado, disgusto o sorpresa en las imágenes (2.5%, 2.5% y 7.5%) y con enfado, disgusto o sorpresa (7.5%, 5.0% o 2.5%). Enfado se confundió con disgusto (26.5%) estáticamente y con miedo, sorpresa o disgusto dinámicamente (10%, 15%, 2.5%), sorpresa sin embargo obtuvo resultados muy favorables en ambos casos y a la expresión neutra se la atribuyó enfado en las imágenes (5.9%).

Resulta relevante cómo determinadas expresiones son reconocidas erróneamente en un porcentaje mayor identificándolas con otras expresiones como es el caso de tristeza con sorpresa o miedo con disgusto o miedo con enfado y disgusto o enfado y disgusto, pero se puede observar que en el segundo cuestionario estos errores fueron en un porcentaje menor y las expresiones fueran acertadas correctamente en un porcentaje superior al 85% como media general.

Diversos estudios demuestran que las expresiones negativas son más difíciles de reconocer o se suelen sobre-interpretar y atribuir valor emocional a estímulos neutrales [Ka03][Su06] además de que virtualmente no se reconoce bien la expresión de disgusto y se suele mezclar con otras expresiones negativas como enfado o miedo, es por ello que los resultados se consideran ampliamente satisfactorios por el porcentaje en el que se han reconocido correctamente las expresiones.

Los resultados obtenidos para el constructor que representaba las expectativas de rendimiento que los usuarios tenían, dejó claro que el avatar virtual ayuda a la hora de identificar las emociones que muestra.

Los resultados relativos a si a los usuarios les agradaba la apariencia y el comportamiento del avatar mostraron buenos resultados, considerando que tenía una buena apariencia y comportamiento aceptable. Los participantes parecían no mostrar rechazo, posiblemente debido a su aspecto caricaturesco. También, los usuarios que conocían previamente a la persona modelada calificaron estas cuestiones con puntuaciones mayores.

En resumen, se extraen un conjunto de conclusiones del estudio realizado que se muestran en la *Tabla 7.8*, y que deben tenerse en cuenta dado que pueden servir para futuros diseños y parametrizaciones de agentes animados.

PROS	CONTRAS
Es capaz de lograr expresiones realistas y fácilmente reconocibles.	Las expresiones negativas se confunden en algunos casos.
Su velocidad hace que los usuarios aprecien mejor sus movimientos.	Su valoración depende de si los usuarios conocen a la persona que representa.
Las animaciones ayudan a la identificación de las expresiones y por tanto a la interacción hombre-máquina	

*Tabla 7.7 – Conclusiones más importantes del estudio.*

Cabe destacar que las conclusiones se extraen de este estudio particular, con las configuraciones consideradas, por lo que es posible que deban ser modificadas a la luz de estudios futuros con diferentes construcciones:

- *Las expresiones son identificadas fácilmente y el empleo de las animaciones resulta favorable para su reconocimiento, con lo que las animaciones resultan útiles para la percepción de emociones en el avatar.*

- *Las expresiones negativas son confundidas y a veces son difíciles de reconocer.* Este hecho debe tenerse en cuenta a la hora de animar al avatar e intentar conseguir que estas expresiones estén más marcadas para que sean más sencillas de reconocer.

- *La valoración del avatar animado está muy influenciada por el hecho de conocer a la persona modelada.* Este aspecto debe tenerse en cuenta a la hora de desarrollar avatares para interactuar con usuarios y, teniendo en cuenta las consideraciones éticas pertinentes, emplear avatares que representen a personas conocidas para lograr una mayor aceptación.

- En términos de la predisposición de los usuarios a que el movimiento ayuda a mejorar la interacción hombre-máquina y por lo tanto su predisposición a emplear agentes o avatares virtuales en las tareas se concluye que *los usuarios aceptan de buena manera un avatar animado al favorecer su percepción sobre el comportamiento de este.*

## **7.5. Conclusiones**

A partir del experimento realizado, se ha mostrado que el avatar resulta adecuado para el uso en interfaces hombre-máquina, ya que los resultados obtenidos para las facetas estudiadas muestran una aceptación positiva por parte de los usuarios.

Se han conseguido los objetivos de generar un avatar realista tanto en apariencia como en comportamiento y estos aspectos se han valorado positivamente. Las expresiones que genera el avatar son fácilmente identificables y estos resultados son mejores en las animaciones, además de que el avatar genere las expresiones mediante movimientos naturales y basados en el comportamiento humano.

Dichos resultados muestran además que los avatares animados podrían usarse en aplicaciones donde el presupuesto de que se dispone sea limitado. Además, obtienen muy buenos resultados en términos de expresividad y expectativa de rendimiento. En el caso de la expresividad, este hecho puede ser debido a la capacidad del avatar de mostrar expresiones más detalladas.

Es de destacar el efecto producido por el hecho de conocer a la persona modelada a la hora de interactuar con avatares animados. Conocer previamente al modelo ciertamente incrementó las puntuaciones otorgadas para cada uno de los constructores, pero fundamentalmente en los relacionados con la expresividad y la apariencia, lo cual lleva a concluir que los avatares que representan rostros de personas conocidas tienen un mayor grado de aceptación por parte de los usuarios.



# Capítulo 8

## *Estudio económico*

### **8.1. Introducción**

En el presente capítulo se realizará un estudio del coste económico que supone la realización del proyecto desarrollado.

Hasta este momento se ha estudiado la viabilidad de este proyecto desde un punto de vista técnico, es decir, si cumple o no el objetivo marcado desde el comienzo. Habiendo concluido este punto, queda pendiente comprobar la viabilidad económica del mismo.

Hay que apuntar que el coste del proyecto es debido a que se parte prácticamente de cero.

En los siguientes apartados se estudiarán por separado costes directos e indirectos y finalmente se mostrarán los costes totales del proyecto.

### **8.2. Recursos empleados**

A continuación se muestra un resumen de los recursos hardware y software empleados en el desarrollo de la aplicación. Es importante tener en cuenta únicamente la amortización del material durante el período de tiempo que ha sido utilizado en el proyecto, para calcular el coste real.

Software:

- Sistema operativo: Windows 7 Professional
- Software de modelado gráfico: Blender 2.63 y 2.66

Hardware:

- Dos ordenadores portátiles: Acer Aspire 5810T (4GB RAM) y HP Pavilion g6 (4GB RAM)

Material ofimático:

- Libros de consulta.
- Otros consumibles.

### 8.3. Costes directos

En cuanto a los costes directos, se evalúan:

- Coste del personal.
- Cortes amortizables de programas y equipos.
- Coste de materiales directos empleados.

#### 8.3.1. Costes del personal

La realización del presente proyecto ha sido llevada a cabo por dos ingenieras, encargadas del diseño y puesta a punto del sistema y aplicaciones correspondientes.

Se calcula el coste anual para un ingeniero para posteriormente adecuarlo al número de horas trabajadas por el mismo. Este coste anual incluye:

- Sueldo bruto anual, así como los posibles incentivos por su trabajo.
- Cotización a la Seguridad Social, que es un 35% del sueldo bruto.

Teniendo en cuenta esto, el coste anual del ingeniero será:

<b>COSTE ANUAL</b>	
Sueldo bruto más incentivos	35.000,00 €
Seguridad Social (35% sueldo bruto)	12.250,00 €
<b>Coste total</b>	<b>47.250,00 €</b>

Tabla 8.1 – Coste anual del personal

Se calcula a continuación una estimación de los días efectivos trabajados al año:

<b>DÍAS EFECTIVOS POR AÑO</b>	
Año medio	365,25 días
Sábados y Domingos	-104,36 días
Días de vacaciones efectivos	-20,00 días
Días festivos reconocidos	-15,00 días
Días perdidos estimados	-5,00 días
<b>Total días efectivos estimados</b>	<b>220,89 días</b>

Tabla 8.2 – Días efectivos por año

Conociendo ya el número total de días efectivos de trabajo, y que la jornada laboral es de 8 horas, obtenemos el total de **horas efectivas de trabajo**:

$$220,89 \text{ días/año} \times 8 \text{ horas/día} = \mathbf{1.767,12 \text{ horas/año}}$$

El **coste por hora** de un ingeniero se calcula como la división del sueldo anual entre las horas efectivas trabajadas al año:

$$\frac{\text{Coste}}{\text{hora}} = \frac{47.250 \frac{\text{€}}{\text{año}}}{1.767,12 \frac{\text{horas}}{\text{año}}} = 26,73 \text{ €/hora}$$

En la siguiente tabla (Tabla 8.3) se muestra una distribución temporal aproximada del trabajo empleado por ambos ingenieros para el desarrollo del proyecto:

<b>DISTRIBUCIÓN TEMPORAL DE TRABAJO</b>	
Formación y documentación	200 horas
Estudio del problema	250 horas
Desarrollo de la aplicación	300 horas
Puesta a punto del sistema	300 horas
Elaboración de la documentación	100 horas
<b>Total de horas empleadas</b>	<b>1150 horas</b>

Tabla 8.3 – Distribución temporal de trabajo

El **coste personal directo** es calculado como la multiplicación de las horas y el coste efectivo de una hora de trabajo del ingeniero:

$$1150 \text{ horas} \times 26.73 \text{ €/hora} = 30.739,50 \text{ €}$$

<b>COSTE PERSONAL DIRECTO</b>	<b>30.739,50 €</b>
-------------------------------	--------------------

### 8.3.2. Costes de amortización de equipos y programas

Para el cálculo de estos costes se debe realizar previamente la inversión total y calcular la amortización lineal correspondiente según los criterios aconsejados por la ley. En este apartado estudiaremos tanto los costes de la amortización del material de oficina como los costes de amortización de los equipos.

Se estima como tiempo de amortización un periodo de dos años, ya que es el considerado como vida útil del material informático empleado. De forma que al calcular el coste hay que multiplicar por un factor del 0.33 los precios mostrados.

<b>MATERIAL</b>	<b>IMPORTE (aprox.)</b>	<b>AMORTIZACIÓN N 33.3 %</b>
Sistema operativo Windows 7	140,00 € x 2 ord. = 280,00 €	93,25 €
Paquete ofimático Microsoft Office 2007	335,00 € x 2 ord. = 670,00 €	223,15 €
Ordenador HP Pavilion g6	600 €	200,00 €
Ordenador Acer Aspire 5810T	700 €	233,15 €
<b>Total material</b>	<b>2250,00 €</b>	<b>750,00 €</b>

Tabla 8.4 – Amortización del material

Hay que apuntar que los programas informáticos no incluidos anteriormente son de libre distribución y su coste, por tanto, es cero.

El coste final por hora de utilización del material es calculado mediante la división de la amortización anual entre el número de horas de uso en dichos equipos.

$$\frac{\text{Coste final}}{\text{hora}} = \frac{750 \frac{\text{€}}{\text{año}}}{1.767,12 \frac{\text{horas}}{\text{año}}} \approx \mathbf{0,43 \text{ €/hora}}$$

Se considera el tiempo de uso como el tiempo total necesario para la realización del proyecto, por ser necesario en las etapas de análisis, diseño, programación y documentación. Por tanto, el coste de amortización de material será:

$$575 \text{ horas} \times 0,43 \text{ €/hora} = 247,25 \text{ €}$$

<b>COSTE DE AMORTIZACIÓN DE MATERIAL DE OFICINA</b>	<b>247,25 €</b>
---	-----------------

### **8.3.3. Costes derivados de otros materiales**

Los costes recogidos a continuación se denominan consumibles, e incluyen, por ejemplo, libros de consulta, papel de impresora, fotocopias, cartuchos de tinta, etc.

Este tipo de material es necesario para la realización de los diferentes trabajos, tanto en la fase de desarrollo y edición, impresión de listados, manuales e informes, almacenamiento de programas y documentos, etc.

El coste total de este material es de 180 €.

<b>COSTE DE CONSUMIBLES</b>	<b>180,00 €</b>
-----------------------------	-----------------

### **8.3.4. Costes directos totales**

De todos los costes obtenidos anteriormente concluimos que los costes directos totales son los derivados de la suma de los costes de personal, amortización de material y de consumibles. Por tanto, será:

$$30.739,50 \text{ €} + 247,25 \text{ €} + 180,00 \text{ €} = 31.166,75 \text{ €}$$

<b>COSTES DIRECTOS</b>	<b>31.166,75 €</b>
------------------------	--------------------

#### 8.4. Costes indirectos

Los costes indirectos son los gastos producidos por la actividad requerida para la elaboración del proyecto y que no se pueden incluir en ninguno de los gastos directos.

<b>COSTES INDIRECTOS PARCIALES</b>	
Dirección y servicios administrativos	150,00 €
Consumo de electricidad	180,00 €
Consumo de telefonía	25,00 €
Consumo de desplazamiento	140,00 €
<b>Total gastos indirectos</b>	<b>495,00 €</b>

Tabla 8.5 – Costes indirectos

Por tanto, los costes indirectos totales ascienden a:

<b>COSTES INDIRECTOS</b>	<b>495,00 €</b>
--------------------------	-----------------

#### 8.5. Costes totales

Los costes totales son el resultado de sumar los gastos directos e indirectos, siendo el montante total para este proyecto:

<b>COSTES TOTALES</b>	
Costes directos	31.166,75 €
Costes indirectos	495,00 €
<b>Coste total del proyecto</b>	<b>31.661,75 €</b>

Tabla 8.6 – Costes totales

En conclusión, el coste total del proyecto asciende a la cantidad de:

<b>COSTES TOTALES DEL PROYECTO</b>	<b>31.661,75 €</b>
------------------------------------	--------------------



# Capítulo 9

## *Conclusiones y líneas futuras*

### **9.1. Conclusiones**

En este proyecto se ha realizado la implementación de un avatar virtual realista y expresivo, con capacidad gestual, para la interacción hombre – máquina. Se ha buscado generar un modelo de animación que presente el mayor realismo facial posible, teniendo en cuenta las barreras y limitaciones computacionales, así como la complejidad propia de la simulación de un rostro humano. El principal objetivo ha sido conseguir un sistema de animación gestual realista, que pueda ser controlado externamente, y que exprese satisfactoriamente las 6 emociones básicas del ser humano, a saber: alegría, tristeza, sorpresa, enfado, miedo y disgusto.

Para su desarrollo se ha creado un modelo facial en tres dimensiones a partir de dos fotografías, evitando así el uso de un escáner 3D usado en anteriores proyectos. De esta manera, el modelado nos permite mayor dinamismo a la hora de animar, ya que situamos los vértices de la forma que más nos conviene para su animación, reduciendo su número y mejorando así la capacidad de deformación del modelo, lo que desembocará en unas animaciones más realistas.

Con el objetivo de conocer el estado actual de este tipo de investigaciones, se ha realizado un estudio de los distintos métodos de modelado y animación 3D existentes. Y dentro de este estudio, se han diferenciado diferentes métodos de adquisición de datos, de modelado tridimensional, técnicas de animación y de simulación de las expresiones.

Una vez creado el modelo se ha procedido a su texturización y animación. Ésta última está basada en el sistema de codificación de acciones faciales o Facial Action Coding System (FACS) y llevada a cabo mediante el empleo de blend shapes o formas clave.

Con el objetivo de obtener un modelo con el mayor grado de realismo posible, se ha realizado el estudio del sistema de codificación de acciones faciales (FACS). Este sistema describe el efecto de la actividad muscular del rostro sobre su apariencia, fraccionando dicha actividad en unidades mínimas de acción o Action Units (AUs). Mediante la combinación de ellas, en mayor o menor grado de intensidad, se consiguen expresiones faciales complejas. Las seis expresiones

complejas universales que se han desarrollado son: alegría, tristeza, enfado, miedo, disgusto y sorpresa.

Respecto a la arquitectura necesaria en el proceso de animación, se han propuesto una serie de herramientas de modelado y animación, comparando las existentes en el mercado y optando finalmente por la herramienta de licencia GPL Blender.

De entre todos los métodos para la animación existentes, se ha optado por el empleo de blendshapes y no de esqueletos jerárquicos. Esta elección está basada en dos razones:

a) Los proyectos anteriores de los que parte este estudio están realizados mediante el empleo de esqueletos jerárquicos, e innovar en el uso de otra técnica como son los blendshapes permite compararlas.

b) Mediante el empleo de blendshapes nos aseguramos de que las animaciones sean lo más marcadas posible, y permite manejarlas más minuciosamente, ya que se modifican las posiciones punto por punto. Si bien es verdad que de esta manera el tiempo que conlleva es mayor, pero el resultado final es preferible.

Cabe resaltar, por último, los más que satisfactorios resultados obtenidos en la evaluación del avatar, que ponen de manifiesto la calidad del trabajo realizado (apariencia realista, expresiones reconocibles por el usuario, gastos económicos y computacionales muy asequibles, etc.). Podemos concluir, por tanto, que este proyecto resulta útil en la interacción hombre-máquina y que puede ser usado en otras aplicaciones, como se propone en el apartado siguiente.

Finalmente se ha podido comprobar que la animación 3D es un trabajo interesante pero a la vez de una complejidad considerable y para la que hay que dominar numerosas disciplinas: artísticas, biológicas, matemáticas, robóticas, etc. Si bien es verdad que la herramienta de modelado elegida ha permitido la abstracción de cálculos matemáticos complejos y de programación, y la aplicación de métodos de animación tridimensional. Además, como ya hemos comprobado, el estudio de la animación 3D es entendido de diversas maneras dependiendo de la cultura, nivel de conocimientos, edad, etc.

## **9.2. Líneas futuras**

Finalmente se han conseguido alcanzar satisfactoriamente los objetivos planteados al comienzo del proyecto, sin embargo cabe considerar algunos aspectos de mejora a través de las líneas de trabajo expuestas a continuación:

Nuestro sistema de animación se basa en la interpolación entre poses o formas clave. Sin embargo, es posible generar animaciones en tiempo real indicando únicamente a un motor gráfico las cuantías de movimiento de la estructura.

Una de las características del sistema empleado es la posibilidad de acceder a las distintas partes del modelo de forma independiente. Partiendo de esta propiedad, se podría tratar de automatizar el proceso de asignación de vértices y generación de movimiento, de manera que, partiendo de una configuración básica del modelo, se pudiera adaptar y emplear para cualquier modelo tridimensional de rostro que se introdujera al sistema.

Se podrían incluir otros sistemas perceptivos en los que ya se ha estado trabajando, como por ejemplo el reconocimiento de identidades de los usuarios, un sistema de captura de movimientos, reconocimiento facial, reconocimiento por voz o sistemas de síntesis de voz. De esta manera incluso se podría llegar a mantener una conversación humano – avatar muy realista. Estos sistemas mejorarían considerablemente la interfaz desarrollada, y seguramente implicarían un mayor grado de aceptación por parte de los usuarios del sistema.

A la vista de las prestaciones conseguidas por el avatar, cabe comentar que podría ser un buen recurso en situaciones en las que es necesaria una comunicación entre una persona y una máquina, posibilitando una comunicación más amigable y sencilla en comparación con las existentes. De esta manera, podría utilizarse para proporcionar información a una persona o a un grupo de gente, integrarse en cajeros automáticos para dar información al cliente, implementarse en internet como sistema de ayuda en algunas páginas, o usarse también en el campo de la domótica. Se podría integrar también en sistemas de control de accesos, video-vigilancia, en la enseñanza, etc.

Por último, se tiene en proyecto evaluar los diferentes aspectos de la interacción de este tipo de interfaces con personas con problemas en cognición social, como pueden ser los pacientes que sufren esquizofrenia, autismo u otros trastornos mentales, como ya hemos comentado anteriormente. Los programas de integración social que se han venido utilizando con este tipo de pacientes únicamente hacen uso de imágenes estáticas y no dinámicas. Por tanto, este aspecto se ha tratado poco en los últimos años, pero ha dado muy buenos resultados y es una buena línea de trabajo futura. El empleo de este tipo de avatares animados permitiría estudiar las ventajas que presenta trabajar aspectos de cognición social con un modelo que puede moverse y expresar emociones que el paciente puede reconocer.



# Referencias

[Ad01] Adolphs R. The neurobiology of social cognition. *Curr Opin Neurobiol.* 2001; 11:231—9.

[Ad02] Adolphs R. Neural Systems for Recognizing Emotion. *Current Opinion Neurobiological.* 2002; 12:1—9.

[Ad03b] Adams, R. B., Jr., Kleck, R. E., 2003. Perceived gaze direction and the processing of facial displays of emotion. *Psychological Science*, vol. 14(6), pp. 644–647.

[Ad05] Adams, R. B., Jr., Kleck, R. E., 2005. Effects of direct and averted gaze on the perception of facially communicated emotion. *Emotion*, 5(1), pp. 3–11. [Ag06] Agelfors, E., Beskow, J., Karlsson, I., Kewley, J., Salvi, G., &

[Am02] Amidror, I., 2002. Scattered data interpolation methods for electronic imaging systems: a survey. *Journal of Electronic Imaging*, Vol. 11 No. 2, pp. 157-176.

[bio] [http://www.biopus.com.ar/emiliano/catedras2006/inter1/apunte\\_interface/interface.html](http://www.biopus.com.ar/emiliano/catedras2006/inter1/apunte_interface/interface.html). Último acceso: Mayo 2013

[Bl04] Blanz, V., Scherbaum, K., Vetter, T., Seidel, H.-P., 2004. Exchanging faces in images. *Comput. Graph. Forum* 23(3), pp. 669–676.

[blog] <http://blog.pucp.edu.pe/media/541/20080620-Zolezzi-2007-2.pdf>. Último acceso: Mayo 2013

[Br05] Brune, M., 2005. Emotion recognition, ‘theory of mind’, and social behavior in schizophrenia. *Psychiatry Research* 133, 135–147.

[Bu90] Buxton, W., 1990. The Natural Language of Interaction: A Perspective on Non Verbal Dialogues. In Laurel, B. (Ed.). *The Art of Human-Computer Interface Design*, Reading, MA: Addison-Wesley. 405-416.

[Ca04] Carretero M.P, Oyarzun D., Aizpurua I., Ortiz A. Animación Facial y corporal de Avatares 3D a partir de la Edición e Interpretación de Lenguajes de Marcas. 2004. <http://www.sc.ehu.es/ccwgamoa/colaboradores/Tesis/ArticuloPuy-Amalia.pdf>

[Ca99] Cassell, J., Vilhjálmsón, H., 1999. Fully embodied conversational avatars: making communicative behaviors autonomous. *Autonomous Agents and Multi-Agent Systems* vol. 2, pp. 45–64.

[Ch02] Cho, B.H., Ku, J., Jang, D.P., et al. (2002). The effect of virtual reality cognitive training for attention enhancement. *CyberPsychology & Behavior* 5:129–137.

[Ch04] Chuang E., Bregler Ch., 2004. Performance Driven Facial Animations using Blendshape Interpolation. Computer Science Department, Stanford University.

- [Cl93] Clayton L. y Rieman J., Task-centered interface design: A practical introduction. 1993.
- [Co04b] Cosker, D., Paddock, S., Marshall, D., Rosin, P.L., Rushton, S., 2004. Towards perceptually realistic talking heads: models, methods and McGurk. In: Symposium on Applied Perception in Graphics and Visualization, pp. 151–157.
- [Con05] Conde López R. 2005. Mapeo facial de emociones sintéticas.
- [Cu06] Curio, C.; Breidt, M; Kleiner, M.; Vuong, Q.; Giese, M.; Bühlhoff, H., Semantic 2006. 3D Motion Retargeting for Facial Animation, Proc. 3th Symposium on Applied perception in graphics and visualization, pp.77-84. ISBN:1-59593-429-4
- [Dy09] Dyck M. , Winbeck M. , Leiberg S. , Chen Y., Mathiak K., 2009. Virtual faces as a tool to study emotion recognition deficits in schizophrenia. 0165-1781 –2009 Elsevier Ireland Ltd.
- [Ef06] Simulating speech with a physics-based facial muscle model. Eftychios Sifakis, Andrew Selle, Avram Robinson-Mosher, Ronald Fedkiw. Vienna : Symposium on Computer Animation archive, 2006.
- [Ek02] Ekman, P., Friesen, W.V., Hager, J.C., 2002. The Facial Action Coding System - Second edition. Weidenfeld & Nicolson, London, UK.
- [Ek75] Ekman, P., Friesen, W.V., 1975. Unmasking the Face. New Jersey, Prentice-Hall Inc.
- [Ekm72] Ekman, Paul., 1972. Universals and cultural differences in facial expressions of emotion, Journal of Personality and Social Psychology Vol.4 N°53 pp.712-717, 1972.
- [Ekm78] Ekman, P.; Friesen, W., 1978. Facial Action Coding System, Ed. Consulting Psychologists Press.
- [Elf03] Elfenbein, H.; Ambady, N., 2003. Universals And Cultural Differences In Recognizing Emotions, Current Directions In Psychological Science Vol.12 N°5 pp.159-164.
- [eps]  
[http://www.eps.uam.es/esp/alumnos/trabajos\\_fin\\_master/Rojas\\_Bello\\_Richard\\_Nolberto.pdf](http://www.eps.uam.es/esp/alumnos/trabajos_fin_master/Rojas_Bello_Richard_Nolberto.pdf).  
Último acceso: Mayo 2013
- [Fa06] Fabri, M., 2006. Emotionally Expressive Avatars for Collaborative Virtual Environments. Tesis doctoral, Leeds Metropolitan University, Noviembre 2006.
- [Fa94] Farkas, L., 1994. Anthropometry of the Head and Face, 2nd ed. Raven Press, New York.
- [Fai90] Faigin, G. 1990. The Artist's Complete Guide to Facial Expressions. Watson-Guptill.
- [Fe07] Fetaji M., Loskoska S., Fetaji B., 2007. Investigating human computer interaction issues in designing efficient virtual learning environments. <http://halshs.archives-ouvertes.fr/docs/00/19/00/68/PDF/313-324.pdf>

[Fo90] Foley J.D. , Dam van A., Feiner S.K. , Hughes J.F. , 1990. Computer Graphics: Principles and Practice, Addison-Wesley, Reading, Mass  
<http://lipas.uwasa.fi/~mj/hci/hci11.html>

[Fo92] Foley, J.D. et al.: Computer Graphics: Principles and Practice in C. Addison Wesley, 2nd ed.,

[Fri94] [56] Fridlund, A. J., Human facial expression: An evolutionary view, Ed. Academic Press, Inc, 1994.ISBN:0122676300

[Go08] Gong, L., 2008. How social is social responses to computers? The function of the degree of anthropomorphism in computer representations. Computers in Human Behaviour pp. 1494-1509.

[Go95] Goleman, D., 1995. Emotional intelligence. New York: Bantam Books.

[Gr07] Griessen, R.T., Cunningham, D.W., Wallraven, C., Bülthoft, H.H., 2007. Psychophysical investigation of facial expressions using computer animated faces. In Proceedings of the 4th Symposium on Applied perception in Graphics and Visualisation, pp. 11-18.

[Gu02] Gur, R.E., McGrath, C., Chan, R.M., Schroeder, L., Turner, T., Turetsky, B.I., Kohler, C., Alsop, D., Maldjian, J., Ragland, J.D., Gur, R.C., 2002b. An fMRI study of facial emotion processing in patients with schizophrenia. American Journal of Psychiatry 159, 1992–1999.

[He06] Herbelin, B., 2006. Modeling Virtual Environments, Medialogy department Aalborg, University Esbjerg. URL: <http://www.aau.dk/~bh/MVE/index.html>

[He07] Hess, U., Adams, R. B., Jr., Kleck, R. E., 2007. Looking at You or Looking Elsewhere: The Influence of Head Orientation on the Signal Value of Emotional Facial Expressions. Motivation and Emotion, vol 31 (2), pp. 137-144.

[Hi03] Hiwada, K., Maki, A., Nakashima, A., 2003. Mimicking video: real-time morphable 3D model fitting. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology, pp. 132–139. ACM SIGGRAPH, Osaka.

[Ho02] Hooker, C., Park, S., 2002. Emotion processing and its relationship to social functioning in schizophrenia patients. Psychiatry Research 112, 41–50.

[Ja02] Jang, D.P., Ku, J., Choi, Y.H., et al. 2002. The development of virtual reality therapy (VRT) system for the treatment of acrophobia and therapeutic case. IEEE Trans Inf Technol Biomed 6:213–217.

[Je06] Jeonghun Ku, Ph.D., Hee Jeong Jang, M.A., Kwang Uk Kim, M.S., Sung Hyouk park, M.D., Jae Jin Kim, M.D., Chan Hyung Kim, M.D., Ph.D., Sang Won Nam, Ph.D., In Young Kim, M.D., Ph.D., and Sun I. Kim, Ph.D., 2006 Pilot Study for Assessing the Behaviors of Patients with Schizophrenia towards a Virtual Avatar. Cyberpsychology & behavior Volume 9, Number 5, 2006 Mary Ann Liebert, Inc.

[Jo03] Joshi, P., Tien, W.C., Desbrun, M., Pighin, F., 2003. Learning controls for blend shape based realistic facial animation. In: Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pp. 187–192. EUROGRAPHICS, San Diego.

- [Jun04] Jun-yong N., 2004, A survey of Facial Modeling and Animation Techniques. Integrated Media Systems Center, University of Southern California.
- [Ka03] Kapur, S., 2003. Psychosis as a state of aberrant salience: a framework linking biology, phenomenology, and pharmacology in schizophrenia. *American Journal of Psychiatry* 160, 13–23.
- [Kä03] Kähler, K., Haber, J., Seidel, H.-P., 2003. Reanimating the dead: reconstruction of expressive faces from skull data. *ACM Trans. Graph.* 22(3), pp. 554–561.
- [Ki07] Kim, K., Kim, J.J., Kim, J., Park, D.E., Jang, H.J., Ku, J., Kim, C.H., Kim, I.Y., Kim, S.I., 2007. Characteristics of social perception assessed in schizophrenia using virtual reality. *Cyberpsychology and Behavior* 10, 215–219.
- [Ko03] Kohler, C.G., Turner, T.H., Bilker, W.B., Brensinger, C.M., Siegel, S.J., Kanesh, S.J., Gur, R.E., Gur, R.C., 2003. Facial emotion recognition in schizophrenia: intensity effects and error pattern. *American Journal of Psychiatry* 160, 1768–1774.
- [Ko04] Kohler, C.G., Turner, T., Stolar, N.M., Bilker, W.B., Brensinger, C.M., Gur, R.E., Gur, R.C., 2004. Differences in facial expressions of four universal emotions. *Psychiatry Research* 128, 235–244.
- [Kö05] Königsson, J., 2005. Facial Animation, Master's Thesis in Computing Science, Umea University, Department of Computing Science, SE-901 87 Umea, Sweden.
- [Ks04] Kshirsagar, S., Egges, A., Garchery, S., 2004. Expressive speech animation and facial communication. In: Magnenat-Thalmann, N., Thalmann, D. (eds.) *Handbook of Virtual Humans*, chap. 10. John Wiley & Sons, West Sussex, England.
- [Ku03] Ku, J., Cho, W., Kim, J.J., et al. (2003). A virtual environment for investigating schizophrenic patients' characteristics: assessment of cognitive and navigation ability. *CyberPsychology & Behavior* 6:397–404.
- [Le01] Leonhard C, Corrigan PW. Social Perception in schizophrenia. En: Corrigan PW, Penn DL, editors. *Social Cognition and schizophrenia*. Washington, DC: American Psychological Association; 2001. p. 73—96.
- [Le04b] Lee, W., Goto, T., Kshirsagar, S., Molet, T., 2004. Face cloning and face motion capture. In: Magnenat-Thalmann, N., Thalmann, D. (eds.) *Handbook of Virtual Humans*, chap. 2. John Wiley & Sons, West Sussex, England.
- [Le88] Lee, J.H., Ku, J., Cho, W., et al. A virtual reality system for the assessment and rehabilitation of the activities of daily living. *Cyberpsychology & Behavior* 6:383–388.
- [Le96] LeDoux, J., 1996. *The Emotional Brain*. New York: Simon&Schuster.
- [Lee95] Y. C. Lee, D. Terzopoulos, K. Waters. Realistic face modeling for animation. *Siggraph proceedings*, 1995, pp. 55-62
- [Li98] Liberman, R.P., Wallace, C.J., Blackwell, G., Kopelowicz, A., Vaccaro, J.V., Mintz, J., 1998. Skills training versus psychosocial occupational therapy for persons with persistent Schizophrenia. *American Journal of Psychiatry* 155, 1087–1091.

[Ma06] MacDorman, K. F., 2005. Androids as an experimental apparatus: Why is there an uncanny valley and can we exploit it?. CogSci-2005 Workshop: Toward Social Mechanisms of Android Science, pp.106-118.

[Ma11] Marcos Pablos, S. Avatar virtual realista con capacidad perceptiva orientado a la robótica social, Universidad de Valladolid

[Mag94] [135] Y. Wu, N. Magnenat-Thalmann, D. Thalmann, A Plastic-Visco-Elastic Model for Wrinkles in Facial Animation and Skin Aging, Proc. 2nd Pacific Conference on Computer Graphics and Applications, Pacific Graphics, 1994

[mic] <http://microexpression.wordpress.com/2010/03/05/lista-aus/#more-933>. Último acceso: Mayo 2013

[Mo06] Moser, E., Derntl, B., Robinson, S., Fink, B., Gur, R.C., Grammer, K., 2006. Amygdala activation at 3T in response to human and avatar facial expressions of emotions. Journal of Neuroscience Methods 161, 126–133.

[Mo82] Mori, M., 1982. The Buddha in the Robot. Charles E. Tuttle Co., ISBN 4333010020.

[Ne76] Neisser, U., 1976. Cognition and Reality. San Francisco, Freeman.

[No03] Nowak, K.K., & Biocca, F. (2003). The effect of the agency and anthropomorphism on user's sense of telepresence, copresence, and social presence in virtual environments. Presence: Teleoperators and Virtual Environments 12:481–494.

[Nu84] Nuechterlein KH, Dawson ME. Information processing and attentional functioning in the developmental course of schizophrenic disorders. Schizophrenia Bulletin. 1984;10: 160—203.

[Om02] Oman, C.M., Shebilske, W.L., Richards, J.T., et al. (2002). Three dimensional spatial memory and learning in real and virtual environments. Spat Cogn Comput 2:355–372.

[Pe02] Penn DL, Ritchie M, Francis J, Combs D, Martin J. Social perception in schizophrenia: the role of context. Psychiatry y Research. 2002;109:149—59.

[Pi95] Piegl, L., Tiller, W., 1995. The NURBS Book. Springer-Verlag.

[Pi98] Pighin, F., Hecker, J., Lischinski, D., Szeliski, R., Salesin, D.H., 1998. Synthesizing realistic facial expressions from photographs. In: Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, pp. 75–84. ACM SIGGRAPH, New York.

[PI81] Animating facial expressions. Platt, Stephen M. y Badler, Norman I. Dallas, Texas : International Conference on Computer Graphics and Interactive Techniques, 1981.

[PI81] S. Platt, N. Badler, Animating facial expression. Computer Graphics, 1981, vol. 15(3) pp. 245-252

[PI85] [105] S. M. Platt, A Structural Model of the Human Face, Ph.D. Thesis, University of Pennsylvania, 1985

---

[Pree94] Preece, J., 1994. *Human-Computer Interaction*. Pearson Education Limited: Essex, England.

[Rau03] Rauterberg M., 1990. "Interaction styles" Vol. 4, No. 6. (1990), pp. 72-82

[Re96] Reeves, B., Nass, C., 1996. *The Media Equation, How People Treat Computers, Television and New Media Like Real People and Places*. Cambridge University Press, Nueva York.

[red] <http://redesprotocolodecomunicacion.blogspot.com.es/2009/04/conceptos-de-interfaz-lewis-y-rieman.html>. Último acceso: Mayo 2013

[Ri00] Rizzo, A.A., Buckwalter, J.G., Humphrey, L., et al. (2000). The virtual classroom: a virtual environment for the assessment and rehabilitation of attention deficits. *CyberPsychology & Behavior* 3:483–499.

[Ro09] Rojas Bello, R. Identificación de características relevantes para el reconocimiento de emociones en el rostro. 2009. Programa Oficial de Posgrado en Ingeniería Informática y de Telecomunicaciones, Universidad Autónoma de Madrid

[Ro10] Rodríguez J.T., Acosta M., Rodríguez L., 2010. Teoría de la mente, reconocimiento facial y procesamiento emocional en la esquizofrenia. 1888-9891 –2010 SEP y SEPB. Elsevier España, S.L.

[Ro95] Rothbaum, B.O., Hodges, L.F., Kooper, R., et al. 1995. Effectiveness of computer-generated (virtual reality) graded exposure in the treatment of acrophobia. *American Journal of Psychiatry* 152:626–628.

[Ru97] Russell, James A.; Fernández-Dols, J.M.I., 1997. *The Psychology of Facial Expression*, Ed. Camb.U.P.

[Sc06] Schneider, F., Gur, R.C., Koch, K., Backes, V., Amunts, K., Shah, N.J., Bilker, W., Gur, R.E., Hab e U., 2006. Impairment in the specificity of emotion processing in schizophrenia. *American Journal of Psychiatry* 163, 442–447.

[Sh92] Shneiderman, B., 1992. *Designing the User Interface*, Addison-Wesley 1992.

[Sh97] Shneiderman B, Maes P., (1997) "Direct manipulation vs. interface agents interactions", Vol. 4, No. 6. (1997), pp. 42-61.

[So08] Sorkin, A., Weinshall, D., Peled, A., 2008. The distortion of reality perception in schizophrenia patients, as measured in virtual reality. *Studies in Health Technology and Informatics* 132, 475–480.

[Sp01] Spencer-Smith, J., Wild, H., Innes-Ker, A.H., Townsend, J., Duffy, C., Edwards, C., Ervin, K., Merritt, N., Paik, J.W., 2001. Making faces: creating three-dimensional parameterized models of facial expression. *Behavior Research Methods, Instruments & Computers* 33, 115–123.

[St96] Strongman, K.T., 1996. *The Psychology of Emotion. Theories of Emotion in Perspective* (cuarta edición), New York, Wiley & Sons.

[Su06] Surguladze, S., Russell, T., Kucharska-Pietura, K., Travis, M.J., Giampietro, V., David, A.S., Phillips, M.L., 2006. A reversal of the normal pattern of parahippocampal response to neutral and fearful faces is associated with reality distortion in schizophrenia. *Biological Psychiatry* 60, 423–431.

[Su83] Summerfield Q "Analysis, Synthesis and Perception of Visible Articalatory Movements", Academic Press Inc 1983

[Sy05] Symbolcraft, Catmull-Clark Subdivisión Surfaces, URL: <http://symbolcraft.com/graphics/subdivision>

[Ta08] Takashima, K., Omori, Y., Yoshimoto, Y., 2008. Effects of avatar's blinking animation on person impressions. *Proceedings of Graphics interface* vol. 322, pp. 169–176.

[Te90] D. Terzopoulos and K. Waters, Physically-based facial modeling, analysis, and animation. *J. of Visualization and Computer Animation*, March, 1990, vol. 1(4), pp. 73-80

[Vi02] Villagrasa Falip, S., Durán Castells J., Clares Gavilán, J. 2002. Animación facial: cómo (re)crear el perfecto rostro humano 3D, Barcelona, España.

[VI05] Vlastic, D., Brand, M., Pfister, H., Popović, J., 2005. Face transfer with multilinear models. *ACM Trans. Graph.* 24(3), pp. 426–433.

[Wa06c] Wallraven, C., Fischer, J., Cunningham, D.W., Bartz, D., Bülthoff, H.H., 2006. The evaluation of stylized facial expressions. In *Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization*, Boston, Massachusetts.

[Wa09] Wang, T. H., James, J. J. L., 2009. Facial expression recognition system based on rigid and non-rigid motion separation and 3D pose estimation. *Pattern Recognition* vol. 42 (5), pp. 962-977.

[Wa87] Waters K., 1987. A muscle model for Animating Three-Dimensional Facial Expression, Animation Research and Development, National Centre for Computer Aided Art and Design, Middlesex Polytechnic, England.

[We08] Weniger, G., Irle, E., 2008. Allocentric memory impaired and egocentric memory intact as assessed by virtual reality in recent-onset schizophrenia. *Schizophrenia Research* 101, 201–209.

[Wi02] Wiederhold, B.K., Jang, D.P., Gevirtz, R.G., et al. 2002. The treatment of fear of flying: a controlled study of imaginal and virtual reality graded exposure therapy. *IEEE Trans Inf Technol Biomed* 6: 218–223.

[wik] Sistema de codificación de acciones faciales (FACS). [http://en.wikipedia.org/wiki/Facial\\_Action\\_Coding\\_System](http://en.wikipedia.org/wiki/Facial_Action_Coding_System). Último acceso: Mayo 2013

[Woj06] Wojdel, A.; Rothkrantz, L., 2006. Intelligent system for semiautomatic facial animation.

[Yi04] Yin, L., Weiss, K., 2004. Generating 3D views of facial expressions from frontal face video based on topographic analysis. In: *Proceedings of the 12th Annual ACM International Conference on Multimedia*, pp. 360–363. ACM SIGMULTIMEDIA, New York.

[Zh03] Zhang, Q., Liu, Z., Guo, B., Shum, H., 2003. Geometry-driven photorealistic facial expression synthesis. In: Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, pp. 177–186. Eurographics, San Diego.

[Zh04] Zhang, Y., Sim, T., Tan, C.L., 2004. Rapid modeling of 3D faces for animation using an efficient adaptation algorithm. In: Proceedings of the 2nd International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia, pp. 173–181. ACM SIGGRAPH, Singapore.

[Zh06] Zhang, Q., Liu, Z., Guo, B., Terzopoulos, D., Shum, H., 2006. Geometry-driven photorealistic facial expression synthesis. *IEEE Trans. Vis. Comput. Graph.* 12(1), pp. 48–60.

[Zh98] Zhu, Q., 1998. 3d voxel-based muscle volume deformation by finite element method. Master's thesis, The State University of New York.

[Zhi07] Zhihong, Z.; Maja, P.; Roisman, G.; Huang, T. S., 2007. A Survey of Affect Recognition Methods: Audio, Visual and Spontaneous Expressions , *Procs. of the 9th international conference on Multimodal interfaces*, pp.126-133.

[Zi10] Ziv I, Leiser D, Levine J. Social cognition in schizophrenia: Cognitive and affective factors. *Cogn Neuropsychiatry*. 2010:1—21.

# Anexo

## *Contenido del CD y utilización*

El presente proyecto lleva adjunto un CD cuyo contenido es el siguiente:

- Copia en .pdf de esta memoria.
- Aplicación desarrollada.

La carpeta que contiene la aplicación incluye una gran cantidad de archivos y extensiones sin las cuales la aplicación no podría ejecutarse. Dicha aplicación lleva el nombre de APLICACIÓN.exe.

Una vez ejecutada la aplicación, la forma de controlarla es mediante las teclas asignadas a cada animación. De esta manera:

- Pulsando la tecla **A**, el avatar realizará la expresión de **alegría**.
- Pulsando la tecla **S**, el avatar realizará la expresión de **sorpresa**.
- Pulsando la tecla **E**, el avatar realizará la expresión de **enfado**.
- Pulsando la tecla **D**, el avatar realizará la expresión de **disgusto**.
- Pulsando la tecla **T**, el avatar realizará la expresión de **tristeza**.
- Pulsando la tecla **M**, el avatar realizará la expresión de **miedo**.
- Pulsando la barra espaciadora, el avatar moverá la cabeza a la izquierda, derecha, arriba y abajo.