

Universitat de les Illes Balears

Departament de Ciències Matemàtiques i Informàtica

Tesis doctoral

EXPRESIÓN DE EMOCIONES DE ALEGRÍA
PARA PERSONAJES VIRTUALES
MEDIANTE LA RISA Y LA SONRISA

Miquel Mascaró Oliver

Dirigida por Dr. Francisco J. Serón Arbeloa y Dr. Francisco J. Perales López

Palma de Mallorca, Marzo 2014

*A Andreu Mascaró i a Maria Oliver
in memoriam*

- En el nostre autòmat, benvolgut doctor, s'hi distingeixen quatre parts. La primera, a l'interior, és el que jo anomeno sistema vivent, i comprèn l'equilibri, la capacitat de moviment, la veu, la flexió de les extremitats, les diferents reaccions als estímuls bàsics, etcètera. En una paraula, allò que, si un cavall en tingués, podríem denominar l'ànima.

*Sebastià Alzamora
Crim de sang 2012*

Agradecimientos

En primer lugar, mi agradecimiento al Dr. Francisco José Serón Arbeloa y al Dr. Francisco José Perales Lopez, directores de esta Tesis, por su guía, ayuda y apoyo durante todo este tiempo.

A todos mis compañeros del Departamento de Ciencias Matemáticas e Informática y en especial a Ramon Mas, Pere Palmer, Miquel Mascaró Portells, Antoni Jaume, Diana Arellano, Biel Moyà, Silvia Ramis y a todas las personas que han pasado por el laboratorio de la Unidad de Gráficos y Visión que de una manera u otra me han acompañado en este camino. Así como a la Dirección del Departamento, Ricardo Alberich y Xavi Varona por su ayuda y colaboración.

Agradecimiento especial a Cristina Manresa por prestarse a sesiones de grabación y ayudarme a la realización de la base de datos de expresiones de risa. También gracias a Victoria Pericàs, Patricia Morales, Juana Maria Peralta, Jana Iencova, Zuzana Durková y Toni Pons de la Escola d'Art Dramàtic por su colaboración profesional y desinteresada en la elaboración de la base de datos.

A Francisca Negre y Barbara de Benito por su ayuda con el test de validación.

A Josep Hernández y Lluís Casals por sus ayudas y consejos con Autodesk Maya.

Y en especial a Francisca Mestre y Maria Mascaró por estar ahí en todo momento.

Resum

L'animació facial és un dels tònics encara no resoltos tant en el camp de la interacció home màquina com en el de la informàtica gràfica. Les expressions d'alegria associades a riure i somriure són pel seu significat i importància, part fonamental d'aquests camps. En aquesta tesi es fa una aproximació a la representació dels diferents tipus de riure en animació facial alhora que es presenta un nou mètode capaç de reproduir tots aquests tipus. El mètode es valida mitjançant la recreació de seqüències cinematogràfiques i mitjançant la utilització de bases de dades d'expressions facials genèriques i específiques de somriure. Addicionalment es crea una base de dades pròpia que recull els diferents tipus de rialles classificats i generats en aquest treball. D'acord a aquesta base de dades pròpia es generen les expressions més representatives de cadascuna de les rialles i somriures considerades en l'estudi.

Resumen

La animación facial es uno de los tópicos todavía no resueltos tanto en el campo de la interacción hombre máquina como en el de la informática gráfica. Las expresiones de alegría asociadas a risa y sonrisa son por su significado e importancia, parte fundamental de estos campos. En esta tesis se hace una aproximación a la representación de los diferentes tipos de risa en animación facial a la vez que se presenta un nuevo método capaz de reproducir todos estos tipos. El método se valida mediante la recreación de secuencias cinematográficas y mediante la utilización de bases de datos de expresiones faciales genéricas y específicas de sonrisa. Adicionalmente se crea una base de datos propia que recopila los diferentes tipos de risas clasificados y generados en este trabajo. De acuerdo a esta base de datos propia se generan las expresiones más representativas de cada una de las risas y sonrisas consideradas en el estudio.

Abstract

Nowadays, facial animation is one of the most relevant research topics still unresolved both in the field of human machine interaction and in the computer graphics. Expressions of joy associated with laughter and smiling are a key part of these fields mainly due to its meaning and importance. In this thesis an approach to the representation of different types of laughter in facial animation is done while a new method to reproduce all these types is proposed. The method is validated by recreating movie sequences and using databases of generic and specific facial smile expressions. Additionally, a proprietary database that lists the different types of classified and generated laughs in this work is created. According to this proprietary database the most representative of every smile expression considered in the study is generated.

Índice de contenido

Capítulo 1 Introducción.....	1
1.1 Motivación.....	1
1.2 Objetivos.....	2
1.3 Estructura del documento.....	3
Capítulo 2 Trabajo relacionado con la representación de la risa y la sonrisa ...	5
2.1 Estudio de la risa en el campo de la psicología.....	5
2.2 Aproximaciones artísticas a la representación de la risa: Representación de la cara humana.....	8
2.3 Estudio de la risa desde el punto de vista fisiológico.....	12
2.4 Representación de la risa en personajes digitales: Introducción a la animación facial.....	16
2.5 El proceso de la animación facial: Rigging facial.....	18
2.5.1 Morfología del personaje.....	20
2.5.2 Descripción del comportamiento.....	21
2.5.3 Entrada de datos.....	21
2.5.4 Interfaz de usuario.....	22
2.5.5 Deformaciones de la geometría.....	23
2.6 Captura de movimiento de las características faciales.....	25
2.7 Definición de los tipos de risa.....	28
2.7.1 Descripción de las expresiones de risa.....	31
Capítulo 3 Reproducción estática de las diferentes expresiones de risa.....	39
3.1 Construcción de un rig facial para la descripción de los diferentes tipos de risa.....	39
3.2 Reproducción de las 13 expresiones de risa.....	47
3.3 Evaluación de las expresiones generadas.....	52
Capítulo 4 Método para la reproducción animada de las diferentes expresiones de risa.....	57
4.1 Introducción.....	57
4.2 Adquisición de los datos para la simulación de las expresiones de risa.....	58
Capítulo 5 Cálculo de los valores animados de los controladores de expresión facial.....	61
5.1 Parametrización facial.....	61

5.2 El controlador SmileFrown.....	66
5.3 El controlador Twich.....	67
5.4 El controlador Jaw.....	69
5.5 El controlador Neck.....	70
5.6 El controlador Blink.....	70
5.7 Los controladores UpsetSad, BrowRaise y Shocked.....	71
5.8 El controlador Sneer y el M_F_U_MouthRolls.....	72
5.9 Otros controladores.....	73
5.10 Uso de los valores de los controladores como objetivo de otras configuraciones de rig facial.....	73
Capítulo 6 Aplicación del método.....	77
6.1 Generación de expresiones de risa a partir de secuencias de vídeo.....	77
6.1.1 Aplicación a las expresiones de risa procedentes de secuencias cinematográficas.....	77
6.1.2 Aplicación a las expresiones de bases de datos públicas específicas de análisis de la risa.....	81
6.2 Creación de una base de datos específica de los diferentes tipos de sonrisa.....	85
6.3 Generación de las expresiones faciales a partir de la base de datos propia.....	86
Capítulo 7 Modelado de expresiones más representativas en las que interviene la AU12.....	93
7.1 Introducción.....	93
7.2 Comparación de las curvas de animación con el algoritmo FastDTW.....	95
7.3 Detección y eliminación de outliers.....	99
7.4 Generación de las secuencias animadas a partir de los datos del FastDTW.....	102
7.5 Ajustes generales sobre las curvas de animación.....	105
7.6 Estudio particularizado de las curvas de cada expresión.....	107
Capítulo 8 Validación de las expresiones generadas a partir de la base de datos.....	113
8.1 Metodología aplicada.....	113
8.1.1 Test de validación.....	114
8.2 Resultados obtenidos.....	115
8.2.1 Valoración de la calidad del movimiento y de la expresión.....	115
Capítulo 9 Conclusiones y trabajo futuro.....	119
Apéndice 1. Tablas con los resultados del algoritmo FastDTW.....	123
Apéndice 2. Test de validación de resultados.....	141
Referencias.....	143

Índice de figuras

Figura 1: Representación de parte de la musculatura facial.....	12
Figura 2: Interfaz con control 2D en ventana 3D de Alexander et al.....	22
Figura 3: Interfaz con controladores 3D de Plantec.....	23
Figura 4: Las diferentes expresiones de la alegría según la clasificación de Gary Faigin: (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada (Original de Faigin 1990).....	31
Figura 5: Gráfico de la hipótesis del "Valle misterioso" (Original de Mori 1982).....	40
Figura 6: Reproducción de algunas de las AUs del FACS: (1) AU 1: Inner brow raiser, (2) AU 2: Outer brow raiser, (3) AU 4: Brow lowerer, (4) AU 5: Upper lid raiser, (5) AU 6: Cheek raiser, AU 7: Lid tightener, (6) AU 9: Nose wrinkler, (7) AU 10: Upper lid raiser, (8) AU 11: Nasolabial deepener, AU 20: Lip stretchner, AU 24: Lip pressor, (9) AU 12: Lip corner puller, (10) AU 13: Cheek puffer, (11) AU 14: Dimpler, (12) AU 15: Lip corner depressor, AU 17: Chin raiser.....	41
Figura 7: Personaje primitivo con malla de superficie de subdivisión y personaje naturalista con malla poligonal (edge loop en amarillo).....	41
Figura 8: Coherencia topológica (edge loop).....	42
Figura 9: Interfaz del rig, según Digital tutors.....	44
Figura 10: Acción de los controladores del rig sobre el modelo primitivo....	45
Figura 11: Deformación basada en huesos para el controlador de la mandíbula	46
Figura 12: Geometrías para el blenshape del personaje primitivo.....	46
Figura 13: Ejemplos cinematográficos de las diferentes expresiones de alegría: (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa	

embarazosa (12) falsa sonrisa y (13) falsa carcajada.....	48
Figura 14: Reproducción de las diferentes expresiones de alegría: (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada.....	50
Figura 15: Reproducción de las diferentes expresiones de alegría: (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada.....	51
Figura 16: Emparejamiento de las expresiones de alegría (dibujo de Faigin, personaje primitivo y fotograma): (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada.....	53
Figura 17: Expresiones de alegría relacionadas en un mismo contexto.....	55
Figura 18: Puntos de la malla de Face Tracker.....	59
Figura 19: Resultado del Face Tracker de Saragih aplicado a la secuencia de video antes citada.....	59
Figura 20: Importación de los puntos de Face Tracker aplicados a una secuencia de la base de datos Cohn-Kanade en el entorno de Autodesk Maya	62
Figura 21: Curva de animación del controlador SmileFrown para la misma secuencia con los valores de ϵ mínimo y máximo.....	64
Figura 22: Curva de animación para el controlador SmileFrown con el mismo factor de escala para dos personajes diferentes.....	65
Figura 23: Activación del smileFrownControl.....	66
Figura 24: Ajuste del factor de escala para la reproducción de la AU15 con el controlador SmileFrown.....	67
Figura 25: Cálculo del twitchControl para el ojo derecho.....	67
Figura 26: Activación del twitchControl.....	68
Figura 27: Activación del jawControl.....	69
Figura 28: Activación del blinkControl.....	71
Figura 29: Activación del shockedControl.....	72
Figura 30: Asignación de curvas de animación a un rig diferente.....	74

Figura 31: Rig de FaceWare Technologies en el que el manipulador de apertura de la boca se activa con valores de rotación en grados de un ángulo	74
Figura 32: Secuencia con extracción de datos errónea.....	79
Figura 33: Ejemplos de asignación de valores a los controladores del rig a partir de datos de secuencias cinematográficas. (1) y (2) sonrisa melancólica, (3) y (4) sonrisa astuta, (5) y (6) sonrisa congraciante y (7) y (8) sonrisa boca cerrada.....	81
Figura 34: Ejemplos de la adquisición de la animación a partir de secuencias de la base de datos Uva-Nemo.....	83
Figura 35: Ejemplos de la adquisición de la animación a partir de secuencias de la base de datos MAHNOB.....	84
Figura 36: Reproducción de las diferentes expresiones de alegría: (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada.....	86
Figura 37: Valores de los principales controladores del rig para las diferentes expresiones de el sujeto s07.....	88
Figura 38: Valores de los principales controladores del rig para las diferentes expresiones de el sujeto s07.....	89
Figura 39: Valores de los principales controladores del rig para las diferentes expresiones de el sujeto s07.....	90
Figura 40: Valores de los principales controladores del rig para las diferentes expresiones de el sujeto s07.....	91
Figura 41: Asignación de los datos del sujeto 7 para las expresiones de risa clamorosa, sonrisa melancólica y sonrisa astuta.....	91
Figura 42: Sonrisa clamorosa ejecutada por los diferentes sujetos del experimento.....	94
Figura 43: Valores del controlador smileFrown para la expresión "sonrisa clamorosa" de los 7 sujetos.....	94
Figura 44: Matriz de coste con la distancia mínima de alineamiento para el controlador smileFrown de la expresión "risa clamorosa" de los sujetos s04 (X) y s06 (Y).....	96
Figura 45: Matriz de coste con la distancia mínima de alineamiento para el controlador smileFrown de la expresión "risa clamorosa" de los sujetos s04 (X) y s03 (Y).....	97
Figura 46: Detalle del error en el seguimiento.....	99
Figura 47: Box plot para la representación de los valores del controlador	

SmileFrown de las 13 expresiones de risa.....	100
Figura 48: Box plot de los datos del controlador SmileFrown para la expresión 7 (Sonrisa ávida).....	101
Figura 49: Reproducción de las expresiones de risa más representativas: (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada.....	104
Figura 50: Proceso de generación de la expresión facial a partir de la base de datos de sonrisa.....	104
Figura 51: Simplificación de curvas.....	105
Figura 52: Imágenes que ilustran las acciones AU6 y AU12 en el FACS.....	106

Índice de tablas

Tabla 1: Acciones de los músculos faciales (a partir de Spaltehoz).....	13
Tabla 2: Relación de los músculos faciales que intervienen en la risa según los diferentes autores (Ruch 2001).....	15
Tabla 3: Relación de los controladores del rig con las AU que activan.....	45
Tabla 4: Músculos activados en las diferentes expresiones.....	49
Tabla 5: Relación de los parámetros faciales extraídos de los puntos de tracker con los manipuladores del rig.....	63
Tabla 6: Tabla de distancias de alineamiento para el controlador smileFrown de la expresión "risa clamorosa".....	97
Tabla 7: Tabla con los valores de mínima distancia de warping para el controlador SmileFrown.....	98
Tabla 8: Tabla con los valores de mínima distancia de warping para el controlador Jaw.....	98
Tabla 9: Tabla con los valores de mínima distancia de warping para el controlador SmileFrown una vez eliminados los outliers de la muestra.....	102
Tabla 10: Tabla con los valores de mínima distancia de warping para el controlador Jaw una vez eliminados los outliers de la muestra.....	102
Tabla 11: Sujetos más representativos para cada controlador de las diferentes expresiones de sonrisa.....	103
Tabla 12: Resultados de la prueba de valoración de las expresiones generadas.....	116
Tabla 13: Controlador: SmileFrown / Expresión: Carjada extrema.....	123
Tabla 14: Controlador: SmileFrown / Expresión: Carcajada.....	123
Tabla 15: Controlador: SmileFrown / Expresión: Risa clamorosa.....	124
Tabla 16: Controlador: SmileFrown / Expresión: Sonrisa boca cerrada.....	124
Tabla 17: Controlador: SmileFrown / Expresión: Risa sofocada.....	124
Tabla 18: Controlador: SmileFrown / Expresión: Sonrisa melancólica.....	125

Tabla 19: Controlador: SmileFrown / Expresión: Sonrisa Ávida.....	125
Tabla 20: Controlador: SmileFrown / Expresión: Sonrisa Congraciante.....	125
Tabla 21: Controlador: SmileFrown / Expresión: Sonrisa astuta.....	126
Tabla 22: Controlador: SmileFrown / Expresión: Sonrisa desleal.....	126
Tabla 23: Controlador: SmileFrown / Expresión: Sonrisa embarazosa.....	126
Tabla 24: Controlador: SmileFrown / Expresión: Falsa sonrisa.....	127
Tabla 25: Controlador: SmileFrown / Expresión: Falsa carcajada.....	127
Tabla 26: Controlador: Jaw / Expresión: Carcajada extrema.....	127
Tabla 27: Controlador: Jaw / Expresión: Carcajada.....	128
Tabla 28: Controlador: Jaw / Expresión: Risa clamorosa.....	128
Tabla 29: Controlador: Jaw / Expresión: Sonrisa boca cerrada.....	128
Tabla 30: Controlador: Jaw / Expresión: Risa sofocada.....	129
Tabla 31: Controlador: Jaw / Expresión: Sonrisa melancólica.....	129
Tabla 32: Controlador: Jaw / Expresión: Sonrisa ávida.....	129
Tabla 33: Controlador: Jaw / Expresión: Sonrisa congraciante.....	130
Tabla 34: Controlador: Jaw / Expresión: Sonrisa astuta.....	130
Tabla 35: Controlador: Jaw / Expresión: Sonrisa desleal.....	130
Tabla 36: Controlador: Jaw / Expresión: Sonrisa embarazosa.....	131
Tabla 37: Controlador: Jaw / Expresión: Falsa sonrisa.....	131
Tabla 38: Controlador: Jaw / Expresión: Falsa carcajada.....	131
Tabla 39: Controlador: UpsetSad y Shocked / Expresión: Carcajada extrema	132
Tabla 40: Controlador: UpsetSad y Shocked / Expresión: Carcajada.....	132
Tabla 41: Controlador: UpsetSad y Shocked / Expresión: Sonrisa clamorosa	132
Tabla 42: Controlador: UpsetSad y Shocked / Expresión: Sonrisa boca cerrada	133
Tabla 43: Controlador: UpsetSad y Shocked / Expresión: Risa sofocada.....	133
Tabla 44: Controlador: UpsetSad y Shocked / Expresión: Sonrisa melancólica	133
Tabla 45: Controlador: UpsetSad y Shocked / Expresión: Sonrisa ávida.....	134
Tabla 46: Controlador: UpsetSad y Shocked / Expresión: Sonrisa congraciante	134

Tabla 47: Controlador: UpsetSad y Shocked / Expresión: Sonrisa astuta.....	134
Tabla 48: Controlador: UpsetSad y Shocked / Expresión: Sonrisa desleal..	135
Tabla 49: Controlador: UpsetSad y Shocked / Expresión: Sonrisa embarazosa	135
Tabla 50: Controlador: UpsetSad y Shocked / Expresión: Falsa sonrisa	135
Tabla 51: Controlador: UpsetSad y Shocked / Expresión: Falsa carcajada..	136
Tabla 52: Controlador: LeftBlink / Expresión: Carcajada extrema.....	136
Tabla 53: Controlador: LeftBlink / Expresión: Carcajada.....	136
Tabla 54: Controlador: LeftBlink / Expresión: Sonrisa clamorosa.....	137
Tabla 55: Controlador: LeftBlink / Expresión: Sonrisa boca cerrada.....	137
Tabla 56: Controlador: LeftBlink / Expresión: Risa sofocada.....	137
Tabla 57: Controlador: LeftBlink / Expresión: Sonrisa melancólica.....	138
Tabla 58: Controlador: LeftBlink / Expresión: Sonrisa ávida.....	138
Tabla 59: Controlador: LeftBlink / Expresión: Sonrisa congraciante.....	138
Tabla 60: Controlador: LeftBlink / Expresión: Sonrisa astuta.....	139
Tabla 61: Controlador: LeftBlink / Expresión: Sonrisa desleal.....	139
Tabla 62: Controlador: LeftBlink / Expresión: Sonrisa embarazosa.....	139
Tabla 63: Controlador: LeftBlink / Expresión: Falsa sonrisa.....	140
Tabla 64: Controlador: LeftBlink / Expresión: Falsa carcajada.....	140

Capítulo 1 Introducción

En este capítulo se presentan de forma concisa las ideas y objetivos principales del trabajo de investigación realizado. En el apartado "Motivación" se pretende dar una visión inicial del problema. En "Objetivos" se presentan los objetivos planteados y finalmente el apartado "Estructura" introduce al lector el esquema de este documento.

1.1 Motivación

La simulación de expresiones realistas en actores virtuales es uno de los campos de mayor proyección en la animación de personajes 3D. La calidad de estas expresiones es fundamental, debido a la enorme empatía que genera el rostro humano, el espectador es capaz de interpretar los mensajes que puede generar la cara de un personaje animado y por consiguiente, el espectador también es capaz de evaluar la fidelidad de las expresiones faciales reproducidas.

Los avances tecnológicos en la última década han permitido desarrollar una gran diversidad de productos de todo tipo, culturales, de entretenimiento, comerciales, asistenciales, etc. que incluyen personajes sintéticos con sus correspondientes expresiones faciales. Aunque en diferentes aplicaciones se requieran diferentes tipos de necesidades los sistemas de bajo coste que ofrezcan calidades realistas aceptables tienen un enorme atractivo.

La animación facial presenta muchas dificultades, restricciones que tienen que ver con el tiempo, el coste y la complejidad. Tradicionalmente muchos de los procesos de producción de animación facial se realizan de forma manual y de forma individual para cada personaje. Incluso en procesos automatizados, como es la animación por captura de movimiento, es habitual encontrar procesos asociados que se llevan a cabo de forma manual, como por ejemplo el limpiado y ajuste de los datos de captura. De manera que sistemas que automaticen parte de este trabajo y a la vez puedan generalizarse para

diferentes personajes, son de mucho interés.

Desde el punto de vista fisiológico, la complejidad de las expresiones hace plausible el intento de estudiar su representación mediante la división del problema según los diferentes tipos de expresiones.

De todas las expresiones faciales, la risa es la que ofrece un mayor número de matices, tanto emotivos como visuales. La risa es la más compleja de todas las expresiones existentes. Hay más sutilidad y sombras emocionales en una sonrisa que en ninguna otra expresión. La sonrisa puede tener elementos de otras expresiones como la tristeza y la furia creando rostros de interesante ambigüedad y complejidad. La risa es capaz de ofrecernos un potente significado incluso cuando apenas se empieza a dibujar en el rostro. Detrás de una sonrisa se pueden enviar multitud de mensajes emotivos, de desafío, menosprecio, lujuria etc. Al mismo tiempo que una sonrisa puede expresar una ambigüedad desasosegante extremadamente difícil de interpretar como lo atestiguan retratos de Leonardo, Vermeer o Rembrandt.

Conseguir por tanto un sistema que sea capaz de realizar una representación realista de los diferentes tipos de risa y sonrisa, de forma sistemática y automática, que sea generalizable a diferentes personajes sintéticos y que a la vez sea integrable en el proceso de producción de la creación 3D son las motivaciones que nos han llevado a desarrollar este trabajo.

1.2 Objetivos

El objetivo principal del presente estudio es ofrecer un sistema para la representación realista de expresiones faciales en donde aparecen la risa y la sonrisa, esto incluye todas las expresiones que tradicionalmente se asocian a estados emocionales de alegría a las que se tienen que sumar todas las expresiones que involucran la acción del músculo cigomático mayor que eleva y estira hacia atrás las comisuras de los labios y que pueden referirse a ciertas expresiones de tristeza o de ira.

Este sistema de representación de risas pretende ser una herramienta que pueda integrarse con facilidad en la manera de trabajar de los animadores de personajes virtuales en un entorno 3D.

Objetivos secundarios derivados directamente de los dos objetivos anteriores son por una parte la realización de una clasificación precisa de los diferentes tipos de risa o sonrisa que hasta ahora han hecho los diferentes autores que han tratado el tema, esto abarca campos diversos que van desde la psicología

y la medicina a terrenos artísticos y a las soluciones que se proponen en dibujo, pintura y escultura. Por otra parte, será necesario también, hacer un estado del arte en la representación facial de lo que se ha venido realizando en el mundo de la animación 3D, estudiando las diferentes técnicas que los profesionales usan hoy en día en las producciones cinematográficas, televisivas, de videojuegos etc.

El problema de la representación realista de determinadas expresiones faciales por personajes virtuales involucra dos aspectos que formaran parte de los objetivos del presente trabajo. De un lado está el sistema que ha de permitir la simulación precisa de la actividad muscular y que deformará la malla geométrica que represente la cara del personaje. Y por otra parte está el sistema que permitirá establecer las claves de animación distribuidas de forma correcta en el tiempo de manera que la actividad muscular se lleve a cabo de forma coherente en la línea de tiempo de forma que la acción sea válida y creíble para cualquier espectador.

Otro objetivo importante del estudio a la vez que una validación del sistema es la aplicación del mismo a las diferentes bases de datos actuales que contienen información sobre expresiones faciales en general y expresiones de risa en particular. Así mismo se desea que nuestro sistema sea capaz de reproducir expresiones calcadas a las que puedan realizar actores en secuencias cinematográficas. Por lo que será necesario que nuestro trabajo incluya el uso y clasificación de esta información.

1.3 Estructura del documento

Esta tesis se estructura de la siguiente manera:

- **Capítulo 2** describe el estado del arte de la representación de la risa y la sonrisa en los campos de la psicología, la representación artística, la descripción fisiológica y la creación digital. Se presenta el concepto de *rigging* desde el punto de vista del proceso de producción. Se introducen los avances en captura de movimiento facial y se estudian las diferentes clasificaciones de la risa cubriendo las aproximaciones de la psicología y de la representación artística.
- **Capítulo 3** describe la reproducción estática de las diferentes expresiones de risa. En primer lugar se detalla la construcción de un rig facial aplicado a dos personajes distintos. Se hace una búsqueda de ejemplos reales de las expresiones de risa. Se reproducen los diferentes tipos de risa atendiendo a

las descripciones morfológicas y a la actividad muscular de cada una de ellas. Se evalúa el resultado de dichas expresiones, constatando la problemática del reconocimiento de las expresiones de risa a partir de cuadros sueltos.

- [Capítulo 4](#) describe la reproducción animada de las diferentes expresiones de risa. Introduce los requerimientos del método. Detalla el sistema de captura del movimiento. Describe el sistema de cálculo de los valores de animación de los controladores basado en reglas temporales entre los puntos faciales. Y comprueba la versatilidad del método aplicándolo a otra configuración distinta de *rig* facial.
- [Capítulo 5](#) expone el método de generación de las curvas de animación de los controladores según los datos extraídos del algoritmo de seguimiento de las características faciales.
- [Capítulo 6](#) aplicación del método a secuencias cinematográficas y a bases de datos públicas de expresiones de risa. Describe la creación de una base de datos propia con sujetos interpretando las diferentes expresiones de risa y aplica la captura de datos y la generación de expresiones automática a los datos de la base de datos.
- [Capítulo 7](#) generación de animaciones representativas para cada una de las 13 expresiones de risa y sonrisa consideradas en este estudio. A partir de la base de datos de las diferentes expresiones de risa generamos las curvas de animación para los controladores faciales.
- [Capítulo 8](#) validación de las animaciones más representativas mediante la aplicación de una encuesta donde se analiza la calidad del movimiento y la representación de las expresiones faciales.
- [Capítulo 9](#) se exponen las conclusiones y el trabajo futuro.
- [Apéndice 1](#) listado de las tablas con las distancias de alineamiento entre las curvas de animación de los controladores para cada expresión facial analizada.
- [Apendice 2](#) muestra del test de validación
- [Referencias](#) lista de todas las referencias bibliográficas citadas en el estudio.

Capítulo 2 Trabajo relacionado con la representación de la risa y la sonrisa

En este capítulo se realiza un repaso sobre el estado del arte del problema planteado. Debido a la transversalidad del tema se presentan aproximaciones pertenecientes a diferentes campos. En un primer apartado se introduce la risa desde la psicología, se continúa con los estudios anatómicos y fisiológicos y finalmente se recogen los trabajos referidos a la representación de la risa en personajes digitales. Describimos el proceso de rigging facial como parte fundamental del problema planteado. Hacemos una relación de los trabajos realizados en captura del movimiento de características faciales y se termina con una clasificación exhaustiva de los distintos tipos de risa sobre los que se basará nuestra aportación.

2.1 Estudio de la risa en el campo de la psicología

Según Bridges[1], la primera expresión facial que aparece en el ser humano es el llanto siendo la risa la segunda. Habitualmente la risa suele aparecer, en el bebé, durante el segundo mes de vida; al mismo tiempo en que este toma conciencia del entorno y aprende a usar los brazos y las piernas. Quedando a esta temprana edad, bien definidos ya, los dos estados emocionales más importantes, placer o alegría y dolor. La risa constituye una forma de comunicación innata heredada de los primates e íntimamente relacionada con el lenguaje. Mientras que la forma del llanto se irá redefiniendo conforme vaya creciendo el individuo, la risa que se presenta un bebé de dos meses no difiere significativamente de la de un niño de ocho años. Esto es: una vez, el individuo empieza a reír, siempre lo hace de la misma manera. Pudiendo convertirse, la risa, en un elemento característico del individuo.

Hess et al.[2] afirman que no sólo la risa es la expresión que se mantiene más intacta desde la infancia a la edad madura sino que además es la expresión más reconocida y apreciada en todas las sociedades humanas y por lo tanto la que más puede trascender las barreras del lenguaje y la cultura.

Hay numerosos estudios que se refieren al estudio de las expresiones faciales y su relación con las emociones humanas, y que por tanto tratan de la risa y la sonrisa como parte fundamental de estas.

De todas maneras no son muchos los artículos de investigación que específicamente hablan de la risa, algunos autores como Willibald Ruch[3] echan en falta un uso consistente del termino, incluso a nivel enciclopédico, y hacen constar que a menudo la sonrisa es confundida con la risa aunque el elemento vocal sea inexistente. Básicamente los estudios de investigación usualmente se centran en uno o dos sistemas de risa más que en la integración de todos componentes involucrados.

La risa ha sido objeto de estudios en el campo de la psicología, desde los trabajos pioneros de Duchenne[4] y Darwin[5], a aportaciones más recientes como las de Provine[6], Gruner[7], Morreall[8] y Ruch[3].

Los primeros trabajos, tanto el de Duchenne en 1862 como el de Darwin en 1872 abrieron un campo en la investigación psicológica de las expresiones faciales. En ambos se encuentra un capítulo completo relacionado con la risa.

Duchenne dio nombre a la sonrisa que lleva su mismo nombre, descrita como una sonrisa generada de forma involuntaria y por tanto genuina. Identificó los músculos involucrados en diferentes expresiones faciales, concretamente llamó al cigomático mayor y a la parte inferior del músculo orbicular de los párpados músculos de la alegría y la benevolencia. Estudió y fotografió la excitabilidad muscular facial a través de la estimulación con electrodos e identificó que la contracción de los mismos se produce de forma involuntaria de ahí que también se llame a la sonrisa de Duchenne, sonrisa verdadera. Argumentación basada en que mucha gente no puede contraer de forma voluntaria el músculo orbicular.

Por su parte, los estudios de Darwin, se centraron en la hipótesis de que las expresiones faciales son innatas, refiriendo experimentos sobre niños de corta edad. El razonamiento de Darwin se basaba en que los gestos faciales que expresan emociones básicas son muy parecidos en culturas muy diferentes, incluso en culturas que apenas han tenido contacto con el exterior. Darwin distingue la risa como respuesta refleja (por ejemplo, en las cosquillas), de la risa como respuesta emocional a una idea o hecho divertido. Señala la gran diferencia, en la frecuencia y espontaneidad, entre la risa infantil y la de los adultos, marcando ya la idea de un desarrollo evolutivo en la expresión de la alegría y sus posibles funciones según la edad. Describe con gran precisión la expresión de la risa en sus distintas intensidades, como una serie graduada que va desde la risa violenta y la moderada pasando por la sonrisa amplia, la sonrisa suave, hasta la simple expresión de buen humor. Señaló la importancia de los músculos cigomático mayor y orbicular, como anteriormente defendía Duchenne.

Aproximaciones más recientes estudian la evolución de la risa en el tiempo. Ruch define el concepto de ataque de risa, como aquel que englobaría a todo el evento, incluyendo los aspectos respiratorios, vocales y faciales de los músculos y del esqueleto. Idealmente la risa puede ser dividida en un punto de inicio que incluiría, por ejemplo, la parte facial antes de la vocal que en el caso de la risa explosiva sería muy corta; un punto de cima, en donde aparecen la vocalización o la exhalación forzada y que en el caso de una risa continuada pueden verse interrumpidas por inhalaciones; y un punto de compensación con una parte de post-vocalización continuada con una larga sonrisa que va desvaneciéndose. El período de vocalización está compuesto por los llamados ciclos de risa, en los que aparecen pulsos de risa con pausas intercaladas. Aunque pueden aparecer risas de solo uno o dos pulsos, como la risa en la exclamación o como la risa sofocada; habitualmente diferentes estudios como el de Rothänger[9] refieren un modo de cuatro pulsos en el ciclo de la risa. El máximo número de pulso viene limitado por el volumen pulmonar, pudiendo llegar a nueve o doce. En un episodio de risa se podrían encadenar varios ataques separados por inspiraciones llegando a un número mayor. Así mismo en cada pulso de risa se puede observar una segmentación acústica, manifestándose por una parte en el núcleo de la expresión vocal y por otra en la “h” aspirada que la precede.

Paul Ekman (Chicago DC 1934) encabeza la lista de psicólogos modernos que han estudiado las expresiones faciales y es considerado uno de los cien psicólogos más relevantes del siglo XX[10]. El trabajo de Ekman se fundamenta en una perspectiva evolutiva según la cual las expresiones faciales de las emociones no son determinadas culturalmente, sino que son más bien universales y tienen, por consiguiente, un origen biológico, tal como planteaba la hipótesis de Charles Darwin. Ekman clasificó las expresiones humanas en lo que llamo seis expresiones universales que expresan la ira, la repugnancia, el miedo, la alegría, la tristeza y la sorpresa[11]. Ekman y Friesen publicaron 1978 lo que denominaron Sistema de Codificación de la Acción Facial (FACS)[12], y que constituye el sistema más comúnmente usado a la hora de clasificar la acción tanto voluntaria como involuntaria de los músculos faciales. Este sistema ha sido citado aproximadamente por un millar de artículos de investigación y existe un software comercial que se ofrece a los investigadores del tema, siendo su última versión del año 2002.

El FACS se ha usado como estándar para categorizar la expresión de las emociones y ha sido usado tanto por psicólogos como por animadores. Con FACS se puede codificar cualquier expresión facial, descomponiéndola previamente en sus Unidades de Acción (AU). Estas unidades de acción pueden usarse en campos tan diferentes como el reconocimiento de expresiones básicas o como la programación de expresiones para agentes en entornos inteligentes. FACS define 32 AU como la contracción o relajación de

músculos. También define una serie de Descriptores de Acción que a diferencia de las AU no implican una acción muscular precisa. Entre otros describen algunos movimientos de la mandíbula, mejillas y labios.

Existen varios modelos implementados en aplicaciones gráficas que incluyen el sistema FACS para la generación de expresiones faciales como el CANDIDE[14], Artnatomy[15] o el más reciente FACe![16] entre otros. Todos ellos permiten la generación de expresiones artificialmente mediante la activación de determinadas AU. Cada aplicación desarrolla su propio interfaz para la manipulación de las unidades de acción. Por ejemplo cada acción tiene codificada una intensidad que varía desde la A (más baja) a la E (intensidad máxima). Así AU 1A es la activación más sutil para elevar las cejas desde su parte interior.

Entre otros trabajos que se acercan al tema de las expresiones faciales desde el punto de vista psicológico destaca el compendio que aparece en el libro “The Psychology of Facial Expression”[13] editada por James A. Russell y José Miguel Fernández Dolç donde se cuestionan los principios de Ekman y sus predecesores Tomkins e Izard. Estos presentan el significado de las expresiones faciales en función de las emociones mientras que las nuevas teorías que se desarrollan en este volumen abren un debate entre los que ven la expresión facial como un precursor o acompañante necesario de la experiencia emocional y los que ven a la expresión facial como un dispositivo comunicativo, expresivo y mimético remanente de un temprano lenguaje preverbal no necesariamente ligado a la presencia de emoción.

La relación entre las expresiones faciales y las emociones o estados emocionales constituyen por tanto una cuestión que queda fuera de nuestro estudio si bien tenemos en cuenta que un actor virtual con una expresión facial generada representa obligatoriamente un estado emocional y en buena medida este estado será el responsable que su presentación o actuación sea o no veraz y por tanto correcta.

2.2 Aproximaciones artísticas a la representación de la risa: Representación de la cara humana

En el campo del arte el estudio de la risa lo encontramos imbuido en el estudio general de la representación de los rasgos faciales y las emociones. Desde el punto de vista artístico la risa y la sonrisa como parte de las expresiones del ser humano han sido estudiadas tanto en aproximaciones realistas como en otras más figurativas.

La representación artística del cuerpo humano ha estado íntimamente ligada a estudios de anatomía por parte de los artistas. Durante el renacimiento se comenzaron a ilustrar mediante figuras las estructuras anatómicas observadas directamente en las disecciones. Primero fueron los artistas, pintores y escultores que acudían a los anfiteatros para ver lo que había debajo de la piel de los seres humanos, sobre todo los músculos superficiales[14]. Luego, algunos médicos como Ketham (1493) y Gersdorff dibujaron del natural el útero y el cerebro. Vesalio, culminó el proceso en su obra monumental “De humani corporis fábrica. Libri septem”, basada en abundantes disecciones humanas.

Miguel Ángel acudía frecuentemente a documentarse del natural, y lo mismo Alberto Durero, que llegó a hacer un dibujo señalando su bazo enfermo. Pero, entre los artistas del renacimiento, quien destaca por su interés por la anatomía es, sin duda, Leonardo da Vinci. En su estancia en Florencia (1500-1506) donde precisamente pintó su conocida obra “La Gioconda” (1503), hizo disecciones en el Hospital Santa María La Nueva, concretamente de un niño de dos años y de un viejo centenario, en el año 1513 se instaló en Roma y siguió practicando disecciones y haciendo gran número de apuntes con la intención, que no realizó, de publicar una obra sobre ellos. Al principio, siguió las enseñanzas de Aristóteles y, luego, las de Galeno, para al fin observar y dibujar por cuenta propia. En 1489 hizo una serie de dibujos del cráneo muy estimables; más tarde (hacia 1500) otra sobre el sistema cardiovascular y del sistema genitourinario femenino. Representó por vez primera, el corazón como un músculo. Describió también el sistema nervioso, el respiratorio y el digestivo. Semejante tarea, dio como resultado una colección de 779 dibujos de los que se conservan unos 600. Leonardo no fue un anatomista riguroso, en ocasiones, se inventaba trazos y suplía sus deficiencias en material humano incluyendo estructuras de animales o ideas tomadas de libros de las autoridades clásicas. De lo que no cabe duda es de que Leonardo da Vinci fue el inventor de los cortes anatómicos y de la representación de la figura humana en diferentes planos.

Leonardo dedicó a la anatomía del cráneo y los maxilares varias láminas. En una de ellas se puede observar el nervio maxilar superior saliendo por el agujero mayor de la base del cráneo y emergiendo al exterior por el agujero suborbitario. Al lado, una representación de la yugular y la carótida con sus ramificaciones faciales.

Aparte de lo que pudiera haber de curiosidad científica en la obra de Leonardo, no cabe duda de que el conocimiento anatómico estuvo al servicio de su obra pictórica y escultural. La boca y los dientes le sirven para dar carácter a sus personajes tallados en mármol o plasmados en lienzos y murales.

En la actualidad la mayoría de obras sobre aprendizaje de dibujo de la figura humana contienen referencias anatómicas en mayor o menor grado. Entre ellas cabe destacar la de George Brant Bridgman (1865 - 1943)[15], que fue profesor de anatomía para artistas durante 45 años en la Art Students League de Nueva York, pasando por sus clases algunos de los más afamados artistas e ilustradores del siglo XX y fue además, autor de varios libros que se han convertido en el estándar de anatomía para artistas. En su obra “Bridgman's Complete Guide to Drawing from Life” aparecen referencias a los músculos faciales y como afectan a las expresiones. Bridgman solo hace referencia a un reducido grupo de músculos, concretamente, cita el frontal, el corrugador, el orbicular de los ojos, el cigomático mayor y menor y el buccinador. No hace un estudio concreto de las diferentes expresiones pero precisamente se fija en la risa comentando que la acción de estiramiento hacia arriba de los labios se combina con la que presiona los ojos. También hace una muy breve referencia a la respiración entrecortada y a la agitación del diafragma.

Andrew Loomis fue uno de los alumnos de Bridgman y autor de una serie de obras de referencia en el campo de la ilustración realista que han influenciado a muchos ilustradores actuales. En su obra “Drawing the Head & Hands”[16] dedica un capítulo entero a justificar la importancia de la anatomía para los ilustradores. En su argumentación, cita que muchos artistas basan su trabajo en el estudio de fotografías para realizar las expresiones faciales pero comenta que las expresiones solo serán convincentes si reflejan las acciones de los músculos bajo la piel. Para ilustrar su afirmación elige también la expresión de la risa. Los aspectos en que se fija son, por una parte, en la forma de los labios, que atribuye al músculo cigomático, y por otra parte los pliegues que aparecen bajo los ojos y que cita como especialmente difíciles de representar en el caso del dibujo. Se fija también en las arrugas o patas de gallo que aparecen en los ojos y por último en que los dientes superiores aparecen más visibles que los inferiores. En la obra de Loomis aparecen representaciones de los músculos faciales pero no su nomenclatura, tampoco encontramos un intento más preciso de clasificación de diferentes tipos de risa, así como tampoco lo encontrábamos en Bridgman.

Hay que destacar también el trabajo Burne Hogarth en el campo de la divulgación artística de la anatomía humana. Hogarth es sobre todo conocido por su trabajo como autor de cómics pero en su faceta como profesor impartió cursos en varias instituciones norteamericanas como la School of Visual Arts de Nueva York, The Parsons School of Design, la Escuela Otis de Los Ángeles o el Centro de Arte de Pasadena. En su obra “Drawing The Human Head”[17] divide el estudio en una primera parte donde estudia las estructuras craneales y las formas de la cara, aquí Hogarth introduce un somero estudio de la musculatura facial, describiendo las funciones de cada músculo y su posición y forma. Describe la función de la anatomía desde un punto de vista

artístico, no como diseccionar y poner en exposición a los músculos sino como vía de analizar y evaluar las formas, distinguiendo las que tienen mayor o menor impacto visual de las que apenas son visibles. Esta consideración es perfectamente válida para nuestro propósito dado nuestro interés en reproducir fidedignamente las expresiones a nivel visual sin necesidad obligatoria de reproducir la mecánica de todos los músculos. En la segunda parte del libro Hogarth estudia el movimiento de la cabeza con el objeto de ejemplificar diferentes planos desde los que representarla. Resulta de especial interés el estudio que hace este autor sobre las arrugas y el envejecimiento de la cara humana, con una serie pormenorizada de imágenes de ejemplo.

En cualquier caso, uno de los estudios más completos desde el punto de vista artístico es el presentado por Gary Faigin en 1990 titulado “The Artist’s Complete Guide to Facial Expression”[16]. La obra se encuentra dividida en tres partes: en la primera se aborda la estructura de la cabeza, en ella se detallan las proporciones de los huesos de la calavera y la mandíbula, la mecánica ocular, la morfología de la nariz y la boca y aspectos relativos a la representación de la cabeza y la perspectiva; en la segunda parte se estudian los músculos de la expresión, aquí Faigin elabora una lista de once músculos principales y describe sus efectos separándolos por zonas donde actúan, distingue entre los músculos de la ceja y el ojo y los de la boca, no solo estudia las acciones musculares sino que también tiene en cuenta las arrugas que se producen en la piel por esas acciones; en la tercera parte se estudian cada una de las seis expresiones básicas universales enumeradas por Paul Ekman, es decir: tristeza, ira, alegría, miedo, asco y sorpresa, a estas expresiones les añade las producidas por lo que llama estados físicos, en las que incluye dolor, somnolencia, pasión y esfuerzo físico, que serían expresiones más ambiguas en términos de reconocimiento instantáneo. El trabajo de Faigin se ha convertido en objeto de referencia no solo para artistas plásticos convencionales sino que en la actualidad, miembros de otras disciplinas como animadores o artistas forenses encuentran en él información fundamental, documentada con numerosos ejemplos gráficos.

De nuestro estudio de la bibliografía sobre la representación artística de la risa y la sonrisa podemos concluir que si bien podemos echar en falta referencias específicas, queda patente la necesidad de pasar por estudios anatómicos y musculares. Otro aspecto que podemos destacar es la dificultad que existe a nivel bibliográfico para describir de una forma gráfica las diferentes acciones musculares que crean las expresiones faciales. La simulación de los músculos faciales en un modelo 3D que pretendemos realizar puede ayudar a entender de manera más comprensible estos aspectos.

2.3 Estudio de la risa desde el punto de vista fisiológico

Desde un punto de vista fisiológico la medición de la acción de la musculatura facial presenta diferentes problemas. Por ejemplo ha de tenerse en cuenta que en caso de risa intensa o carcajada intervienen más elementos que la propia estructura facial, movimientos de la cabeza y del tronco tiene que ser considerados para una correcta simulación.

Ruch[3] resume diferentes hipótesis y resultados de trabajos llevados a cabo por diferentes investigadores que cubren desde la observación en tiempo real, inferencias basadas en el conocimiento de la musculatura facial, el uso de electromiografía facial o el Sistema de Codificación de la Acción Facial de Ekman y Friesen.

La electromiografía facial EMG es la técnica que mide la actividad muscular mediante la detección y amplificación de los impulsos eléctricos minúsculos que se generan por la contracción de las fibras musculares. Con estas diferentes técnicas expuestas se confirma efectivamente que los dos músculos que conforman la Sonrisa de Duchenne, es decir el cigomático mayor y el orbicular de los párpados, están involucrados en la generación de la risa y la sonrisa, las transiciones entre la risa y la sonrisa así como en el inicio y el fin de un ataque de risa.

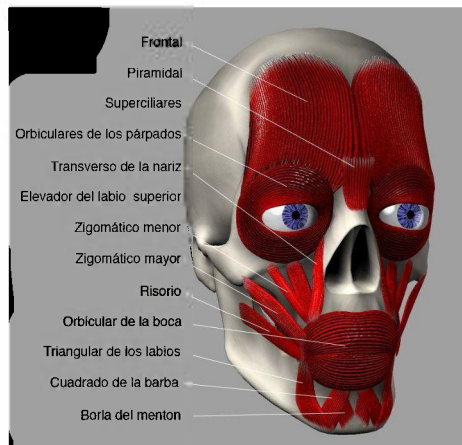


Figura 1: Representación de parte de la musculatura facial.

En el atlas de anatomía humana de Spalteholz[18] encontramos una descripción precisa de todos los músculos faciales. En la descripción se incluye el nombre, origen, inserción, acción e inervación. Una simplificación del sistema muscular facial se puede observar en la figura 1, en la que se

observa que se han eliminado los músculos que no intervienen en la generación de expresiones como el músculo auricular temporal, la aponeurosis epicránea, la aponeurosis parotideometérica, el músculo cutáneo del cuello y el músculo esternocleidomastoideo; tampoco está representado el músculo masetero responsable de la aducción de la mandíbula inferior y finalmente otros músculos internos situados bajo los superficiales como los superciliares, el bucinador, o el nasal también han sido simplificados. Estos últimos ejercen su efecto en zonas de la cara que comparten el efecto de otros músculos. En la tabla 1 se relacionan los músculos y sus acciones cuando estos son activados.

Músculo	Acción
Frontal, parte mediana	Eleva las cejas (la parte interior). Surcos horizontales en la frente (centro)
Frontal, parte lateral	Eleva las cejas (la parte exterior). Surcos horizontales en la frente (laterales)
Superciliares (corrugador)	Junta y baja las cejas. Arrugas verticales en el entrecejo
Piramidal (procerus, depresor)	Baja las cejas (la parte media). Puede arrugar la piel sobre el puente de la nariz
Orbiculares (parte orbital)	Estira la piel alrededor del ojo. Patas de gallo
Orbiculares (parte de los párpados)	Estira la piel alrededor del ojo, la parte inferior aumenta
Zigomático mayor	Tira de las comisuras de labios hacia arriba y atrás
Zigomático menor	Hunde el surco naso-labial
Elevador del labio superior	Eleva el labio superior y las aletas de las fosas nasales, se enseñan los dientes
Elevador del labio superior, aletas nasales	Eleva la parte central del labio superior y las aletas de las fosas nasales
Nasal	Dilata las fosas nasales
Buccinador	Comprime la mejilla formando un hoyuelo
Risorio	Estira las comisuras labiales a los lados
Orbicular de los labios	Aprieta, hace que sobresalgan o inviertan los labios
Triangular de los labios, depresor ángulo boca	Tira hacia abajo las comisuras de los labios
Mentoniano, borla del mentón	Eleva el mentón y el labio inferior sobresale
Masetero	Realiza la aducción de la mandíbula inferior
Canino, elevador del ángulo de la boca	Tira de la esquina del labio e infla las mejillas

Tabla 1: Acciones de los músculos faciales (a partir de Spaltehoz)

Cuando empieza la sonrisa, el ojo empieza a contraerse y a medida que aumenta de nivel, es decir, cuando pasa de risa a carcajada, por ejemplo, el ojo va contrayéndose cada vez más. Hay una relación directa entre el nivel de la risa y el grado de contracción del ojo; tanto es así que si no hay una correspondencia exacta entre ellos automáticamente se tiende a rechazar la, en el sentido de que es considerada como forzada o falsa. En este sentido se refieren los estudios llevados a cabo por Sumitsuji.

Durante la carcajada el ojo se estira debido a la contracción del músculo orbicular (orbicularis oculi); este músculo actúa también durante el llanto,

estirando también la piel que rodea a los ojos. La diferencia entre ambos es que en el caso del llanto, el músculo entero se contrae generando arrugas en las esquinas y entre los ojos, mientras que en la carcajada solo actúa la mitad inferior de la porción de los párpados del músculo, siendo el resultado el estiramiento del ojo desde la parte inferior.

Durante la risa actúa la misma parte del músculo de manera menos intensa, de manera que en este caso, el ojo permanece abierto al contrario que durante la carcajada. Cuando se produce la sonrisa la parte inferior del párpado se acorta, tensándose y deslizándose hacia arriba a través del ojo. De forma similar a lo que pasa con el labio superior, en ambos casos se tensan los extremos de la curva. Visualmente el párpado sube justo hasta la parte inferior de la pupila. Además de la forma del ojo en sí misma en la misma zona se han de tener en cuenta las arrugas que aparecen en las esquinas de los ojos, las llamadas “patas de gallo” y las bolsas que aparecen bajo el párpado inferior, ambas son comunes tanto a la risa como al llanto.

La acción del músculo orbicular de los ojos es diferente según la parte del músculo que actúe, así la parte orbital es la responsable de la aparición de las patas de gallo y la parte de los párpados la responsable de que aumente la bolsa del párpado inferior. Este motivo justifica que muchos autores consideren las dos partes del músculo por separado.

Durante la sonrisa la boca se abre y se muestra la parte superior de la dentadura. Desde un punto de vista visual se genera un fuerte contraste entre el blanco de los dientes y el negro del interior de la boca. Este contraste no se encuentra en ninguna otra expresión de manera tan manifiesta y es decisivo a la hora de identificarla. La potencia visual de la sonrisa unida a su fácil identificación queda patente en la masiva utilización de esta en publicidad por ejemplo. Esta acción es debida al músculo elevador del labio superior, éste también actúa sobre las aletas nasales ensanchándolas, de hecho este músculo es capaz de actuar sobre estas últimas sin mover el labio superior. En el estudio de Ruch se diferencian estas dos acciones como si de dos músculos se tratase.

Durante la sonrisa los labios se estiran y siguen la forma de la calavera. La forma no sufre modificaciones importantes respecto a la carcajada exceptuando los saltos provocados por la mandíbula en esta última. Estas formas son provocadas por la acción del músculo cigomático mayor, la labor principal de este músculo es precisamente la de activar la sonrisa en la boca. Conviene no confundir la función de este músculo con la del músculo risorio; este último junto con el platismo (músculo cutáneo del cuello) es el responsable del estiramiento hacia abajo de la boca que se produce en el llanto por ejemplo. Tampoco hay que confundirlo con el cigomático menor que es el responsable de hundir el surco naso-labial, esta acción aparece en la mayoría

de diferentes sonrisas pero no en todas.

Autores	Parte superior					Parte central					Parte inferior							
	Frontal med	Frontal lat	Superciliar	Piramidal	Orbicular orb.	Orbicular par.	Zigom. May.	Zigom. Men.	Elev. L. Sp. Dien	Elev. L. Sp. Alet	Nasal	Bucinador	Risorio	Orbicular lab.	Triangular lab.	Mentoniano	Masetero	Canino
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Hipótesis: Risa																		
Bloch et al. (1987)					x	x	x											x
Darwin (1872)					x	x	x		x	x?								
Dearborn (1897)			x	x		x	x	x	x	x	x		x			x		
Hecker (1873)								x		x	x							
Heller (1902)					x	x	x						x					
Hjortsjö (1970)	x	x			x	x	x	x		x		x	x					x
Piderit (1867)							x	x	x				x		x			
Poeck and Pilleri (1963)					x		x				x			x				
Raulin (1900)					x?		x	x						x				
Hipótesis: carcajada																		
Darwin (1872)			x		x	x	x		x	x?								
Heller (1902)	x	x	x	x	x	x	x			x			x					x
Piderit (1867)			x				x	x	x				x		x			
Resultados: Estudios FACS																		
Grammer (1990)									x									
Keltner and Bonanno (1997)					x		x											-
Ruch (1990, 1994)					x	x	x											-
Ruch (1997)					x		x											-
Resultados: Estudios EMG																		
Gallo and Palla (1995)																		x
Santibañez et al. (1986)														x				x
Sumitsuji (1967)	-	-	-		x		x		x				x	x	x	x		
Tanaka (1976)	-	-	-		x		x											x
Tanaka et al. (1991)	-	-	-		x		x		x				x					x
Risa sofocada																		
Tanaka (1976)	x	x			x		x											
Risa patológica																		
Tanaka et al. (1991)	x	x	x		x		x		x				x					x
Yamada et al. (1994)			x															

Tabla 2: Relación de los músculos faciales que intervienen en la risa según los diferentes autores (Ruch 2001)

Las mejillas se encuentran en el centro de los cambios cuando se produce la risa o la carcajada. Tanto el músculo orbicular del ojo como el cigomático mayor empujan a la mejilla cuando se contraen de forma individual. De manera que cuando actúan de forma conjunta las mejillas están sometidas a

doble presión. El redondeo de mejillas es una de las formas más características de la sonrisa que aparece fuertemente incluso en estadios muy iniciales de la sonrisa. A nivel visual el punto donde es más apreciable este redondeo coincide con la altura de las aletas de la nariz, sobresaliendo alrededor de los 45 grados respecto al frente de la cara.

Las mejillas son el área de la cara con mayor concentración de grasa subcutánea. La movilidad de la grasa es la que explica el redondeo de la mejilla. Las diferencias en la forma de las mejillas durante la sonrisa y la carcajada son casi inapreciables. Existe una diferencia sutil debida a que durante la carcajada la presión extra del músculo orbicular hace subir las aletas de la nariz y el tejido naso labial por lo que la mejilla gana prominencia. El músculo canino es el responsable de que se inflen las mejillas, mientras que el buccinador hace que aparezcan los hoyuelos característicos en las mejillas, presentes en muchas de las sonrisas estudiadas.

En la tabla 2 Ruch relaciona la inclusión o no de los diferentes músculos faciales en los estudios que los principales investigadores han realizado sobre la risa y la sonrisa. Concretamente en total se consideran 18 músculos faciales divididos en tres zonas de la cara: superior, media y baja.

2.4 Representación de la risa en personajes digitales: Introducción a la animación facial

En lo que se refiere al campo de la animación facial, tal y como describen Noh et al.[19] o Radovan et al.[20] se distinguen dos aproximaciones diferentes, las que se basan en deformaciones geométricas y las que se basan en la manipulación de imágenes. Si bien algunas técnicas combinan las dos metodologías.

La animación facial se ha beneficiado de los principales resultados de los avances en la animación corporal en general. El sistema basado en la simulación muscular comparte similares problemas de investigación, incluyendo la acción muscular y la deformación de los tejidos blandos. No obstante existen notables diferencias:

- El número de músculos que intervienen es mayor en el caso de la animación facial
- Las deformaciones de los tejidos blandos deben ser más realistas debido a la capacidad innata de reconocimiento de expresiones faciales que tiene el espectador.

- La articulación de la cabeza presenta mayor complejidad que la mayoría de las articulaciones del cuerpo debido al mayor grado de libertad que posee.
- De forma habitual las técnicas de fusión de formas (blendshape) se han usado en animación facial mientras que técnicas de animación por esqueleto se han aplicado a la animación del cuerpo.

Entre las aproximaciones basadas en la geometría se distinguen:

- Técnicas de interpolación de claves. En ellas los vértices 3D de los polígonos son trasladados a otra posición geométrica describiendo una posición clave para una expresión determinada. El ordenador calculará los cuadros intermedios interpolando los valores. Presenta inconvenientes relativos a la cantidad de claves que ha que generar para cada cara y la cantidad de vértices que hay que modificar. Los trabajos pioneros de Parke[21] apuntaron en esta dirección. Actualmente este método puro suele considerarse como obsoleto si bien las técnicas de interpolación se usan comúnmente en combinación de otras.
- Técnicas de parametrización: Con esta técnica cualquier cara puede expresarse mediante la combinación de una serie de parámetros independientes. Con lo que se permite el control de configuraciones faciales específicas. Existen inconvenientes debido intervención de parámetros que entran incompatibilidad. Iniciaron esta técnica Parke y Waters en sus trabajos en 1996[22]. Uno de los modelos paramétricos de uso público más usado, el Candide fue actualizado[23] para poder usar sistemas de codificación de expresiones. Igual que el método anterior, las técnicas de parametrización puras se consideran obsoletas en la actualidad.
- Aproximaciones basadas físicamente en la acción muscular: En este sentido abogaban los trabajos pioneros de Platt y Batler[24]. En estas aproximaciones se definen capas de mallas elásticas deformables según la acción de un número determinado de músculos que intervienen en las diferentes expresiones. Waters[25] estableció los tres diferentes tipos de músculos (lineal, fusiforme y esfinter) y la acción debida a estos.
- Aproximaciones basadas en la simulación de músculos: Usando técnicas de deformación como las de forma libre se simula la acción muscular. Con ellas es difícil la simulación de arrugas y pliegues de la cara. Estas técnicas empezaron a desarrollarse con los trabajos de Thomas W. Sederberg et al. [26] y más tarde Prem Kalra et al.[27]. Autores como Pasquariello y Pelachaud[28] dividen la cara en diversas zonas donde actúan los músculos según el estándar de animación MPEG-4 simulando arrugas y pliegues con técnicas de mapas de desplazamiento. Otra aproximación a este método es la de Bui[29] cuyo modelo también simula arrugas usando la acción de los músculos definidos por Waters.

Y en las basadas en imágenes encontramos:

- Aproximaciones basadas en imágenes: las técnicas basadas en imágenes se dividen en *morphing* entre imágenes, manipulación de textura, mezcla de imágenes y expresiones vasculares. Estas técnicas se iniciaron con los trabajos de Beier y Nelly[30]. Pighin et al.[31] combinaron técnicas de *morphing* 2D con transformaciones 3D para conseguir expresiones faciales realistas. Las técnicas de manipulación de textura fueron descritas inicialmente por Oka et al.[32] mostrando sistemas de texturización en tiempo real de una malla a partir de fotografías faciales. Las expresiones vasculares son usadas como complemento a otras técnicas, en concreto Kalra et al.[33] las usaron para ajustar el color de la piel para representar determinados estados emocionales. Especial mención en este campo se merecen los trabajos de Bregler et al.[34], Ezzat et al.[35] y Cossato et al.[36] en ellos se desarrollan técnicas para la simulación del habla.
- Métodos de captura del movimiento: Se basan en la captura del movimiento realizado por actores para después poder animar a los personajes sintéticos. Los primeros trabajos en este campo fueron desarrollados por Parke y Waters[22]. Kouadio et al.[37] presentaron un sistema para generar animación en tiempo real usando expresiones faciales premodeladas.
- Sistemas de codificación de expresiones: Han sido consideradas como la alternativa a los métodos de captura. Proporcionan una base de datos con las posibles expresiones faciales, de manera que el animador pueda generar un guión con las expresiones y el sistema pueda calcular la animación. El sistema más popular es el diseñado por Ekman et al.[12] junto con los modelos MPEG-4 de Doenges y Ostermann[38].

Es habitual encontrar combinación de técnicas de simulación muscular con técnicas basadas en imagen. Modelos como el de Sifakis[39] consiguen un importante nivel de realismo. En ellos se parte de una geometría de alta definición con capacidad para reproducir las acciones musculares para después traducir las capturas de movimiento a su correspondiente activación muscular. Los fonemas son sintetizados de forma separada y son usados para generar nuevas geometrías que permitirán la simulación del habla.

2.5 El proceso de la animación facial: *Rigging* facial

Recientemente el estudio de Orvalho[40] presenta un estado del arte en animación facial cubriendo todo el proceso de *rigging*. Este estudio es

especialmente interesante ya que aparte de introducir definiciones y especificaciones del concepto de *rigging* enfoca la problemática de la animación facial desde el punto de vista técnico y desde el punto de vista artístico poniendo un énfasis especial en el proceso de producción. Según Orvalho la principal dificultad consiste en la falta de una definición estándar de lo que es un *rig* facial y en la multitud de posibles configuraciones para la animación facial que usan los animadores. Define el *rigging* facial como el proceso de creación y posterior activación de una serie de controladores que se usarán para operar sobre un modelo 3D análogamente al funcionamiento de las cuerdas en una marioneta.

En función de las aplicaciones de la animación facial en la industria del entretenimiento éstas se pueden dividir en aplicaciones fuera de línea y aplicaciones interactivas en tiempo real. Orvalho muestra la evolución de la animación facial por décadas desde los setenta hasta nuestros días con los siguientes ejemplos, para la animación off-line: El modelo paramétrico de Parke de 1972, el primer corto de animación que usó un modelo facial paramétrico "Tony de Peltrie" de 1985, la primera película generada íntegramente por ordenador "Toy Story" de 1995, Gollum personaje realista generado por ordenador en un papel principal en la película "The Lord of the Rings: The Two Towers" en 2002 y la película "The Adventures of Tintin" de 2011 como ejemplo de los últimos avances en captura de movimiento aplicados a cine de animación. A su vez ilustra la animación en tiempo real con los ejemplos: "Mike the Talking Head" de 1988, la primera marioneta generada por ordenador, animada en tiempo real, el "Half Life" de 1998 que fue uno de los primeros juegos que contenía animación facial, el "Doom 3" de 2004 con un *rig* facial basado en huesos y "The Samaritan Demo" de 2011 consistente en una secuencia de demostración del motor de juego Unreal. Estos casos son un buen ejemplo gráfico que permite ver como han ido mejorando los resultados en animación facial desde sus inicios al presente.

Es a mediados de los 90 cuando emerge el concepto de *rigging* en animación facial y está asociado a personajes realizando acciones complejas, la técnica del *rigging* enlaza los procesos de modelado y animación con el procedimiento de la animación tradicional. Los avances en control del movimiento se basan en conceptos provenientes de la robótica y de las técnicas de animación de stop-motion. Se definen esqueletos como una serie de huesos unidos que permiten definir movimientos articulados. Estas estructuras están en un nivel abstracto y no forman parte de la imagen calculada, se encuentran dentro del modelo y están ligados a su estructura. Para el control de estas estructuras el animador usa técnicas de cinemática y de cinemática inversa. Las técnicas de *rig* se incrementan notablemente cuando se pasa a trabajar a dos niveles, a nivel de transformación y a nivel de deformación. El primero manipula el objeto y el segundo cambia la forma de las componentes de ese objeto, de manera que el

concepto de *rig* es un concepto que va evolucionando desde el control de objetos 3D simples a objetos más complejos como por ejemplo una cara. Objetivos recientes de los sistemas de *rig* es que los animadores puedan controlar el movimiento de una manera completamente intuitiva.

Un sistema de *rig* es una estructura capaz de realizar algún determinado movimiento que en el modelo 3D tendrá una determinada configuración que corresponderá con una determinada deformación. Siendo este el principal problema del *rigging*, es decir, la transformación de los datos del movimiento del controlador en datos de deformación de la geometría. Habitualmente la estructura del *rig* se encuentra dividida en tres niveles para hacer la conversión de datos: la interfaz de usuario que permite al artista el control de los datos de entrada, la estructura de transformación que define el origen y posición del modelo y los deformadores que modifican la geometría del modelo según la transformación.

El *rigging* se integra en el proceso de producción de creación 3D en función de si la aplicación es en sistemas *off-line* o en sistemas interactivos de tiempo real. En los sistemas *off-line* el proceso de *rigging* tiene lugar después de la modelización y es el que permitirá generar los datos de animación que a su vez configuraran el fichero de *render*. En los sistemas de tiempo real es necesario definir dos módulos de *rigging* uno para la animación y otro para el *render*, ya que la geometría de deformación y el *render* deben calcularse en tiempo de ejecución. En estos últimos sistemas puede haber variaciones en función de la aplicación, así por ejemplo en vídeo juegos es habitual el uso de secuencias pregrabadas que se ejecutan en función de la entrada o por ejemplo en aplicaciones interactivas donde las secuencias son calculadas en función de los datos suministrados por el usuario.

En el trabajo de Orvalho se incluye una tabla con las principales aportaciones al *rigging* facial desde 1996 a 2011. Cita, entre otras, las aproximaciones de Maraffi[41], Gorden[42], Ritchie[43], Miller[44], McKinley[45], Bredow[46], Chang[47], Pardew[48], O'Neill[49], Vilagrassa[50], Alexander[51], Osipa[52] y Arghinetti[53]. En esta tabla se reflejan las características de cada aportación en función de los diferentes criterios para clasificar los sistemas de *rigging* dependiendo de la morfología del personaje, la descripción que debe simular el *rig*, la entrada de datos al sistema, la interfaz del usuario y el tipo de deformación usado.

2.5.1 Morfología del personaje

Debido a la gran cantidad de tipos de caras y a la diversidad de estilos visuales, existen diferentes clasificaciones de tipos de caras. Aplicadas al entorno 3D, entre otros, Mclaughlin [54] distingue tres tipos de morfologías: naturalista, primitivo y abstracto. Donde naturalista se refiere a un personaje

que se mueve y comporta como una persona o una criatura del mundo real, primitivo es aquel que está asociado a una estética *cartoon* y abstracto se refiere a personajes cuya forma está compuesta por elementos que obedecen a movimientos físicos creíbles.

2.5.2 Descripción del comportamiento

Ya sea siguiendo descripciones realistas como las de Ekman y Friesen con FACS o según el estándar MPEG-4 para animación facial descrito por Pandzic y Forchheimer[55] o bien animaciones no realistas del estilo *cartoon* o animaciones que no siguen ninguna descripción estándar.

El MPEG-4 es un sistema estándar de parametrización para el control facial. Desde su introducción ha sido usado tanto en la industria del entretenimiento, como la medicina o la telecomunicación. Es independiente del modelo facial siempre que el modelo permita la anotación de acuerdo con el estándar. Se basa en tres tipos de datos faciales: parámetros de animación facial FAPs, parámetros de definición facial FDPs y tablas de interpolación FIT. Los FAPs son el mínimo conjunto de parámetros que el sistema establece para la animación facial esto incluye 84 puntos FP distribuidos en un modelo base en posición neutra para reproducir fácilmente los movimientos faciales, las expresiones, las emociones y el habla.

2.5.3 Entrada de datos

Entendiendo el *rig* como un sistema que realiza la conversión entre datos de movimiento y deformación de geometría, se proponen tres tipos de entrada de datos: por claves de animación, captura de movimiento y animación procedural.

El sistema de claves *keyframing* es el más popular de los sistemas de animación por ordenador, desde su introducción en los años 70 y más tarde con la aplicación de los principios de la animación tradicional a las estructuras con esqueletos propuestas por Lassetter[56]. Desde mediados de los 90 se generalizaron entornos de animación que usan curvas editables *spline* para el control de cada parámetro animable en el tiempo. Estos editores son comunes a todas las aplicaciones comerciales.

Los sistemas de captura de movimiento han cobrado gran relevancia en la última década debido sobre todo a los intereses de la industria. En términos generales estas técnicas están descritas en el estudio de Menache[57]. Por su importancia en nuestro estudio dedicaremos un punto posterior a ellos en este estado del arte.

Mediante la animación procedural se pretende introducir datos de movimiento describiendo la expresión o el humor del personaje, la mayoría de estos

métodos se basa en los estándares FACS y MPEG-4. Arya et al.[58] Describen el modelo de comportamiento para agentes sociales virtuales en función del conocimiento, personalidad y humor, que proporcionan parámetros para la deformación de la malla facial. Arellano et al.[59] determinan el estado emocional de un personaje virtual dependiendo de rasgos de su personalidad y de su experiencia emocional.

2.5.4 Interfaz de usuario

Para generar el movimiento que define la animación del personaje, el animador debe ver el sistema de *rigging* como una estructura que puede ser manipulada. La interacción con el animador puede establecerse con elementos hardware como dispositivos para la captura de movimiento y con elementos software como la interfaz del *rig*. Básicamente se distinguen dos tipos de interfaz los basados en ventana y los basados en vista 3D.

Los basados en ventana permiten la entrada de datos de diversas maneras: Villagrasa y Susín[50] presentan un sistema basado en FACS manipulable mediante deslizadores y Bredow et al.[60] definen el movimiento mediante una lista de atributos.

En los interfaces basados en vistas 3D es común usar objetos para la definición del movimiento, típicamente curvas que al ser desplazadas dentro de un entorno restringido activarán las deformaciones sobre la malla facial. Como es el caso del *rig* propuesto por Digital Tutors[61] o por Alexander et al.[62], en ambos se realizan movimientos restringidos en el plano 2D usando un control antropométrico. Otros autores como Komorowski et al.[63] o Plantec[64] presentan interfaces donde los manipuladores faciales son activados en un volumen 3D sobre la misma cara del personaje.

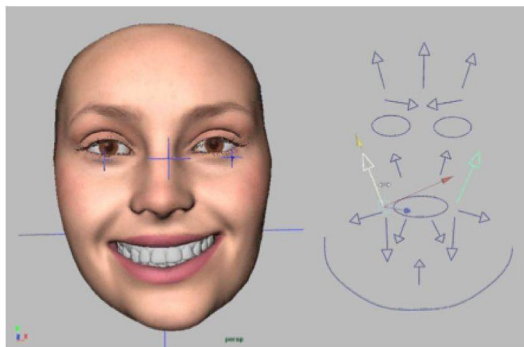


Figura 2: Interfaz con control 2D en ventana 3D de Alexander et al.

Recientemente se han introducido técnicas de bosquejo para activar las deformaciones faciales, como por ejemplo el método presentado por Miranda et al.[65].

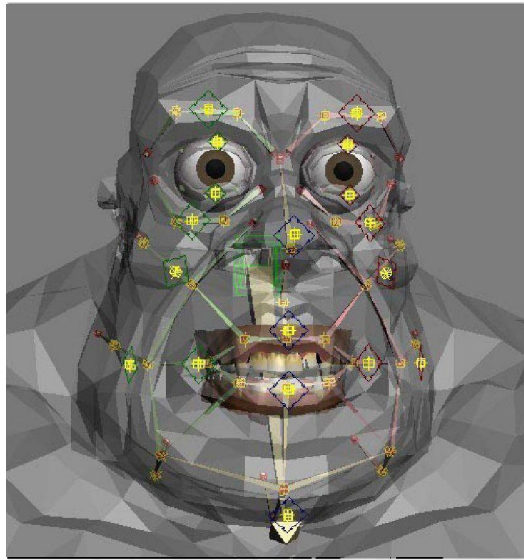


Figura 3: Interfaz con controladores 3D de Plantec

2.5.5 Deformaciones de la geometría

Básicamente la creación del *rig* se basa en técnicas de *blendshape* y en técnicas de esqueletos. Otros métodos como las deformaciones libres FFD o las basadas en física pueden añadirse para mejorar la precisión de los métodos anteriores.

El concepto de *blendshape* como interpolación geométrica fue introducido por Parke[21] en 1972 y popularizado por Bergeron y Lachapelle[66] en 1985 a partir de su corto de animación “Tony de Peltrie”. Este proceso consiste en esculpir formas faciales en diferentes mallas con la misma topología, la interpolación de estas mallas generará la animación del personaje. Por ejemplo, interpolando entre la malla con la boca abierta y la posición neutra generará el movimiento de abrir / cerrar. Es posible definir deformaciones localizadas en diferentes regiones faciales de manera que su combinación genere un amplio número de expresiones diferentes. La animación facial de un personaje puede estar completamente definida con este método, pero si la animación ha de ser realista el número de *blendshapes* puede dispararse, el personaje Gollum de “The Lord of the Rings: The Two Towers”[67] tenía 675 *blendshapes*, en lo que constituye un ejemplo de *rig* complejo de este tipo. Con un menor número de formas y combinando con otro tipo de animación es

posible lograr una animación realista. Con el uso de técnicas mixtas la dificultad para el animador aumenta en función de lo grande que sea la región a animar. La definición manual de las diferentes formas para la interpolación es un proceso costoso que requiere una cantidad importante de tiempo y además es un proceso que se repite para cada personaje animado. Osipa[52] introduce una manual para la generación de un *rig* facial usando la técnica de *blendshapes*. Existen diferentes trabajos enfocados en minimizar los inconvenientes que suponen la modelización de geometrías, el control y optimización del sistema de interpolación. Orvalho[68] presenta un algoritmo genérico para definir un *rig* y de forma automática calcular el cambio de objetivo de interpolación para poder aplicarlo a otros personajes. Li et al.[69] proponen un método específico para *rig* facial basado en *blendshapes* que crea de forma automática las geometrías faciales de interpolación a partir de un modelo genérico existente.

Otro tipo de sistemas de deformación son aquellos que implementan un sistema basado en la creación de un esqueleto facial articulado. Este tipo de control de la animación se encuentra implementado en todos los paquetes de creación 3D[41]. En general con este tipo de animación se generan movimientos más suaves que con el sistema de *blendshape*. En contraposición este sistema requiere una mayor planificación ya que los vértices de la malla geométrica solo se mueven en función de los huesos más cercanos. Para poder animar el modelo es necesario enlazarlo con la estructura de huesos, este proceso se denomina *skinning*. Existen dos tipos de *skinning* que son el suave y el rígido, en función de como se ejecuta la deformación sobre los vértices afectados. En este proceso es especialmente importante la tarea de definición de pesos que define el grado de influencia de cada vértice durante la deformación, de manera que cada hueso tiene su propio mapa de influencia que define su acción.

El *rigging* por deformación geométrica consiste en aplicar operadores simples que permiten modificar modelos complejos. Uno de los operadores, comúnmente usado es el llamado deformador de forma libre FFD, se trata de un sólido paramétrico que envuelve a la geometría de manera que las transformaciones aplicadas a sus vértices deforman el modelo sobre el que actúa. Este concepto se extiende para poder abarcar la animación facial en lo que se da en llamar deformador de forma libre racional RFFD que garantiza un mayor control sobre la geometría deformable. Una de estas aproximaciones es la de Noh et al.[70] que dividen el rostro en una serie de regiones y aplican un RFFD a cada una.

En las animaciones basadas en física se simulan las propiedades viscoelásticas de la piel, los músculos y los tejidos blandos. Reproduciendo el comportamiento de éstos para conseguir animaciones más realistas. Los principales métodos aplicados para este tipo de animación son los sistemas de

masa - muelles y los basados en elementos finitos. Como hemos introducidos en el punto 2.4 varios autores han hecho aportaciones según este método.

Rigs definidos por combinación de *blendshapes* y esqueletos. Teniendo en cuenta que el presupuesto y el tiempo son recursos limitados y los más importantes en la industria del entretenimiento, la importancia de la buena elección de un sistema de *rig* adecuado para la animación de los personajes es fundamental. Durante el proceso de creación del *rig* es habitual el tener que retocar o remodelar algunas partes del personaje que por ejemplo requieran más resolución, esto puede provocar que el *rig* deba empezarse de nuevo, debido a esto es necesario que el marco de trabajo sea lo más flexible posible. La configuración inicial de un *rig* basado en *blendshape* es sencilla y depende del número de mallas existentes, ahora bien, cambios en la topología provocarían que se tengan que rehacer las mallas. En el caso de *rig* basados en huesos son complejos a la hora de su definición pero son más robustos a cambios de topología. La combinación de ambos métodos es la más común en la industria[52], [61], de forma que se aprovecha la suavidad y la flexibilidad del sistema de esqueletos con la expresividad que proporciona el sistema de *blendshape*.

2.6 Captura de movimiento de las características faciales

Aunque un sistema de *rig* sea capaz de simular la actividad muscular facial de forma precisa y detallada, para conseguir una animación válida es necesario que el animador posea cierta destreza que le permita definir correctamente las claves de animación en la línea de tiempo.

Desde los inicios hasta el presente, una herramienta comúnmente usada por los artistas ha sido un espejo de mano que ha servido para reproducir las expresiones faciales[71].

En el caso concreto de la representación de la risa es fundamental el control del tiempo como parte de lo que Ruch y Ekman[3] denominan “laughter bout” que comprende el evento acústico juntamente con los elementos respiratorios, vocales y esqueleto-musculares de la cara. Durante la segmentación de la risa aparece además el llamado ciclo de risa formado por la repetición de sucesivos pulsos de risa separados por pausas.

El problema de la detección de rostros y de características faciales es un aspecto ampliamente estudiado en el área de visión por ordenador. Viola y Jones[72] describen un algoritmo basado en el aprendizaje automático, que después de un entrenamiento previo es capaz de procesar gran cantidad de

imágenes rápidamente y alcanzar niveles muy altos de detección. Existen variaciones de este método[73] que son usadas tanto para la detección de caras como para la detección de expresiones faciales. Castrillón[74] hace una comparación de los diferentes métodos basados en este algoritmo.

Cerezo et al.[75] aplican técnicas de visión para la detección de rostros para realizar un seguimiento de las características y de las expresiones faciales y aplican sus resultados a la animación del personaje virtual Maxine.

La animación facial basada en la captura de movimiento ha sido un tópico de referencia en la investigación durante los últimos años. Con gran número de sistemas de adquisición y procesamiento que comparten principios fundamentales y detalles específicos de aplicación. Una revisión completa de estos sistemas se encuentra en el trabajo de Phigin y Lewis[76].

Un equilibrio fundamental en todos estos sistemas es la relación entre la calidad de los datos adquiridos y la complejidad del sistema de adquisición. Se debe tener en cuenta que no todas las aplicaciones tienen las mismas necesidades de precisión, que varía desde los complejos avatares usados en la producción cinematográfica a avatares más simples no humanos o de juegos en tiempo real. A este nivel se usan técnicas basadas en marcadores debido básicamente a la robustez del método.

Para la digitalización realista de caras sin marcadores se usan escáneres 3D de tiempo real, capaces de detectar gran cantidad de detalles como arrugas y pliegues de la piel. Por regla general estos sistemas implican sensores altamente especializados y/o condiciones especiales de estudio, como iluminación. Huang et al.[77] presentan un sistema que combina un conjunto mínimo de captura de datos de un escáner 3D con un sistema de interpolación de *blendshapes* para conseguir reconstrucciones de caras con mucha fidelidad.

En el otro extremo se encuentran diferentes aproximaciones que usan simples cámaras para la adquisición de los datos faciales. Típicamente estos métodos, al trabajar en entornos con iluminación no controlada, se limitan a la detección de características faciales como ojos, cejas, pupilas o contorno interior y exterior de los labios.

Existen diferentes enfoques de este estilo, uno de ellos es el que presentan Patras y Pantic[78] basado en el filtrado de partículas para el seguimiento de características faciales, reconocimiento de eventos temporales y razonamiento en casos de incertidumbre. La idea del filtrado de partículas es tener una representación basada en partículas de la probabilidad a posteriori de un estado dadas todas las observaciones de una secuencia de imágenes. Pantras hace el seguimiento de 5 características faciales que se corresponden al contorno de la nariz y la boca. En un trabajo posterior[79] usando el mismo método hacen un reconocimiento de las AUs que aparecen en secuencias de

vídeo, para ello usan el seguimiento de 20 puntos faciales y hacen el reconocimiento de las acciones faciales de forma dinámica según una serie de reglas temporales definidas en función de la distancia existente entre los puntos considerados.

Otro enfoque es el basado en el ajuste de modelos deformables consistente en hacer el registro de un modelo paramétrico a una imagen de tal manera que sus puntos de interés se correspondan con los del modelo. Diferentes aproximaciones de estos métodos hacen una predicción de donde estarán los puntos de interés. En este campo es destacable el trabajo de Saragih et al. [80], la solución que ellos proponen se basa en el algoritmo de *mean-shift* con una regularización debida al conocimiento previo que se tiene del modelo, además de extensiones a la solución que tratan el problema de oclusiones parciales de la imagen así como la reducción de la complejidad computacional. Como resultado ofrecen la aplicación FaceTracker que es una aplicación de código abierto sobre la librería OpenCV. La entrada de la aplicación pueden ser archivos de vídeo o capturas de la cámara, no necesita ningún marcador, ni inicialización con ninguna posición determinada, ni ninguna iluminación especial. Funciona sobre Mac OS X y no tiene ningún requerimiento especial de hardware.

Recientemente aprovechando la tecnología de vídeo juegos han aparecido nuevos dispositivos como los sensores de la Nintendo Wii o del sistema Kinect de Microsoft. Existen aproximaciones de diferentes estudios que incorporan estas tecnologías como es el caso del Face Tracker [81] del SDK de Kinect. En este caso el resultado de salida de la aplicación es mayor, que la anterior, ya que aparte de una malla formada por 87 puntos 2D se tiene un posicionamiento de la cabeza 3D y una lista de las AUs realizadas. Todo ello con un nivel de error también bajo como en la opción anterior. Especialmente remarcable es el trabajo de Weise y otros [82] que ofrecen un sistema de animación facial de bajo coste. Presentan un nuevo algoritmo de seguimiento facial basado en una serie de animaciones faciales existentes que adaptan una malla genérica según los datos captados por la cámara. Su sistema no requiere ninguna iluminación especial, ni ningún tipo de marcador sobre el actor. Es necesario un proceso de entrenamiento de las diferentes acciones faciales para la adaptación probabilista de los datos pregrabados. Su solución se comercializa con el nombre de FaceShift y es integrable en las principales aplicaciones 3D.

2.7 Definición de los tipos de risa

En la literatura se encuentran diferentes clasificaciones de los tipos de risa, la mayoría de ellas provienen del campo de la psicología y su mayor interés estriba en dilucidar la autenticidad de la expresión.

Trabajos de Ruch y Ekman[3] apuntan a una clasificación de la risa en términos de las diferentes formas existentes para la generación de los sonidos de risa y la posible modulación de la emoción producida por la risa espontánea. De esta manera diferencian tres tipos: la risa espontánea, la risa artificial o falsa y la pronunciación de los sonidos de la risa, que pueden representarse fonéticamente como “ha ha ha”. Estos sistemas se diferencian en el grado del control voluntario y de forma inversa en el grado de emotividad.

En la risa espontánea se sigue un impulso, el impulso de reír sin restricciones. No hay ningún intento de suprimir la respuesta o ejercer cualquier control de su expresión; la persona que ríe se abandona a sí misma a la respuesta de su cuerpo. El aspecto involuntario también se puede apreciar en el hecho de que durante la risa, la auto-conciencia y auto-atención se ven disminuidas. Tratar de dirigir la atención durante un episodio de risa, hace que ésta se detenga o se reduzca significativamente. Las descripciones de la experiencia emocional se refieren principalmente a la conciencia del impulso de reír, en este sentido, la risa espontánea es claramente agradable. Es el tipo de risa que aparece en los recién nacidos[83].

El otro tipo de risa que distinguen Ruch y Ekman es la risa voluntaria. Durante la risa voluntaria se intenta reproducir el patrón de la risa espontánea. Un ejemplo cotidiano de este tipo de risa es el que se da cuando el sujeto intenta unirse a la risa de otros sin sentir ningún tipo de placer. Si bien la risa voluntaria puede controlar o pasar por encima a la risa espontánea no es posible de forma voluntaria generar risa emocional. De forma categórica Ruch y Ekman afirman que cualquier risa producida de forma voluntaria tendrá en mayor o menor grado una componente de falsedad. Desde el punto de vista del patrón respiratorio la risa espontánea y la voluntaria son muy parecidas, no obstante las diferencias entre ambas son muchas y están por definir, de hecho la cuestión de la detección de mentiras mediante el estudio del comportamiento fisiológico es uno de temas de las investigaciones de Ekman y su equipo[84] [85].

Finalmente otro tipo de risa puede producirse modulando el sonido de la risa, ya sea hablando o cantando. Igual que en la risa voluntaria, en esta, se pretende reproducir un patrón de sonido pero con más control sobre los resultados. La principal diferencia es que en esta última la fonación no se basa

en la respiración forzada sino en una dosificación correcta del aire que produce una menor resonancia traqueal, otra diferencia puede ser la carencia de entonación emocional que hace la expresión vocal sea más clara pudiendo llegar a coincidir con los fonemas del idioma hablado. La risa emocional o espontánea es independiente de la risa hablada pero puede darse el caso que se den al unísono al alargarse los pulsos de la risa[86].

Otras clasificaciones de tipos de risa tienen presentes características acústicas, como los trabajos llevados a cabo por Jo-Anne Bachorowski[87] que considera tipos de risa en función del tono, la modulación del aire como un bufido o un ronquido y si es producida por un hombre o una mujer. De forma genérica las clasificaciones dependientes de parámetros acústicos o que tengan que ver con la respiración son las que tendremos en menor consideración en nuestro estudio debido a que nuestro interés se basa sobre todo en el aspecto exterior de la expresión facial más que en su generación interna.

Messinger et al.[88] incluyen un experimento realizado con 30 bebés y que le permite hacer una clasificación entre sonrisa con elevación de mejillas, sonrisa con boca abierta y la que incluye las dos características.

Hess et al.[2] centrándose en las diferencias y similitudes culturales estudian las sonrisas y experimentan diferenciando 5 tipos dependiendo de la actividad del músculo cigomático mayor, así distinguen sonrisa débil, sonrisa media, con y sin arrugas en el contorno de ojos, sonrisa fuerte y sonrisa miserable.

En Faigin[16] encontramos una clasificación exhaustiva de diferentes tipos de expresiones de alegría. Esta clasificación se basa en la observación por separado de como actúan los elementos clave en la risa y la sonrisa. Usando métodos de observación ad hoc e ilustrándolos con dibujos propios, fotografías y reproducciones de obras clásicas, Faigin hace una primera clasificación según la intensidad que va de la **carcajada** o risa extrema, a la **sonrisa**.

A esta primera clasificación que iría de mayor a menor intensidad, le añade otras que llama sonrisas complejas que son el resultado de combinar a la acción de los músculos cigomático y orbicular de los párpados diferentes posiciones de las cejas correspondientes a las formas que se definen en las expresiones universales de Ekman. De esta manera distingue:

- La **risa astuta** como resultado de la combinación con las cejas bajadas y el ceño fruncido, es decir como aparecen las cejas en la expresión de enfado.
- La **risa ávida**, combinando con las cejas elevadas, como en la expresión de sorpresa.
- La **risa congraciante**, combinando con las cejas temerosas (elevadas y

ceño fruncido), posicionamiento de las cejas como en la expresión de miedo.

- La **risa melancólica**, combinando con las cejas correspondientes a la expresión de tristeza.

A estas les añade otras dónde es patente el ejercicio voluntario, de esta manera aparecen:

- La **risa sofocada**, producida al reprimir la risa espontánea. A la acción de los músculos cigomático y orbicular de los párpados, se les añade la del orbicular de la boca, que estira los labios, la del triangular y la del mentoniano, que estiran las esquinas y bajan los labios respectivamente. Además de la acción del cigomático menor y del labio superior, que afectan al hundimiento de surco naso-labial y al ensanchamiento de las fosas nasales.
- La **risa embarazosa**, igual que la anterior se trata de un intento de represión de la emoción, mientras que en la anterior se reprime la risa o la carcajada en este caso la emoción es menos intensa, es la represión de la satisfacción. Se caracteriza por tener los párpados completamente cerrados o con la mirada dirigida hacia abajo. Las cejas aparecen relajadas o levemente hacia arriba.
- **Sonrisa desleal**, en esta modalidad la sonrisa se combina con una posición de los párpados en la que las pupilas quedan en parte cubiertas, es la posición que se puede asociar a estados de somnolencia e intoxicación. La acción del músculo frontal es la que permite elevar las cejas y hacer que los párpados no caigan por completo. Cuando se combinan con la sonrisa el resultado sugiere intoxicación placentera (alcohólica por ejemplo), deseo sexual o actitud satírica. En este tipo de expresión pueden producirse asimetrías tanto en las cejas como en los labios.

Finalmente la clasificación se completa con las llamadas risas y sonrisas falsas. Hay una serie de características que identifican este tipo de expresiones que tienen en común la falta de alegría aunque intenten expresarla.

- **Falsa sonrisa**: En la boca de esta expresión interviene la acción del músculo risorio que tira hacia abajo del labio inferior, con lo que la boca presenta una forma más cuadrada y sobresalen más los dientes inferiores que en la sonrisa verdadera.
- **Falsa carcajada**: En esta expresión se repite la acción del músculo risorio de la anterior y además se diferencia de la verdadera en la parte de los ojos. Aquí estos no aparecen contraídos o no lo suficientemente contraídos. Es la acción del músculo orbicular de los ojos. En menor grado se observa también una diferencia en lo que afecta a las mejillas. En la carcajada real

la tensión debida la cigomático mayor es máxima y las mejillas aparecen completamente redondeadas mientras que en las falsa es menor y por tanto la forma de las mejillas es más sutil.

En total Faigin distingue y describe 13 tipos de expresiones referidas a la expresión del estado de alegría. Esta clasificación ofrece un espectro amplio y detallado sobre el que vamos a desarrollar nuestro estudio. Ver figura 4.

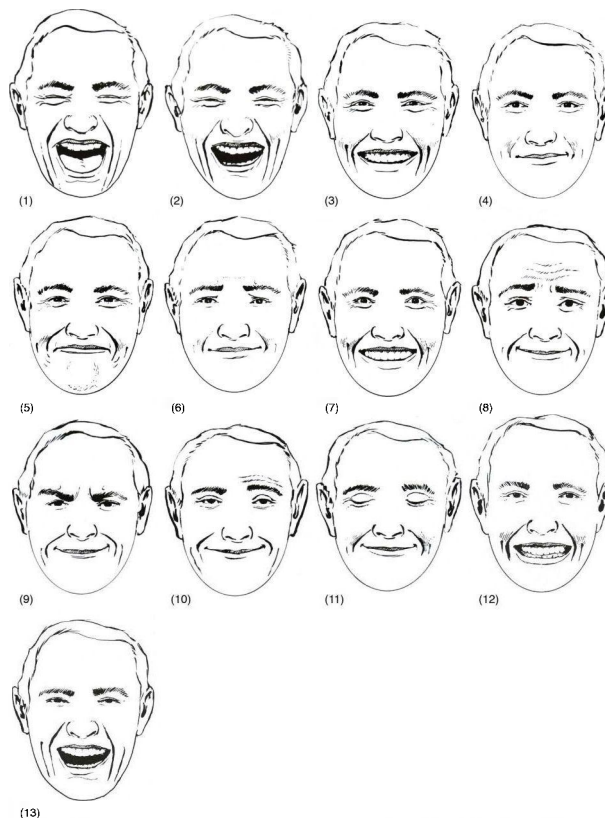


Figura 4: Las diferentes expresiones de la alegría según la clasificación de Gary Faigin: (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada (Original de Faigin 1990)

2.7.1 Descripción de las expresiones de risa

Carcajada extrema.

- Cejas: Se presentan relajadas. El músculo frontal no actúa.

- Ojos: Cerrados por compresión del músculo orbicular. Cuanto más intensa es la carcajada más fina es la línea de las pestañas.
- Boca: Completamente abierta y agrandada. Las esquinas están elevadas. El labio superior completamente elevado, el inferior en la mitad con los costados casi verticales. Se enseñan los dientes superiores y pueden mostrarse las puntas de los inferiores.
- Arrugas de expresión: Patas de gallos y pliegues en el párpado inferior, arrugas en forma de estrella en las esquinas de los ojos, barbilla robusta.
- Similitud con otras expresiones: Similar al bostezo exceptuando que en este las mejillas no están llenas y los ojos están activos.

Carcajada

- Cejas: Relajadas.
- Ojos: Cerrados o muy ligeramente abiertos. Las pestañas se encuentran en una línea. El párpado inferior se ve lleno.
- Boca: Completamente abierta. Esquinas elevadas. El labio superior elevado enseñando los dientes superiores. El labio inferior en la mitad. Mejillas angulosas. De los dientes inferiores se muestran solo las puntas.
- Arrugas de expresión: Arrugas bajo los ojos en forma de sonrisa. Arrugas verticales desde la nariz a la barbilla. Aplanamiento de la barbilla.
- Similitud con otras expresiones: Similar al llanto y a la falsa carcajada. Las diferencias con el primero estriban en que en éste el músculo corrugador contrae las cejas, el músculo orbicular de los ojos se contrae completamente haciendo aparecer las arrugas características. Solo se enseñan las puntas de los dientes superiores mientras que los inferiores están abajo enseñándose las esquinas de la dentadura. Respecto a la falsa carcajada la diferencia principal se encuentra en que los ojos están más abiertos, las mejillas menos llenas y también hay menos arrugas.

Sonrisa clamorosa.

- Cejas: Relajadas.
- Ojos: En forma de media luna. Párpado superior ligeramente bajado mientras que el inferior sube contrayéndose con tensión.
- Boca: Abierta con las esquinas hacia atrás en dirección a las orejas. El labio superior recto enseñando la dentadura superior, el labio inferior a la mitad y las mejillas hinchadas.
- Arrugas de expresión: Arrugas bajo los ojos en forma de sonrisa. Arrugas

desde la nariz a la barbilla. Hoyuelos.

- Similitud con otras expresiones: Similitudes con la falsa sonrisa. Diferenciándose de ésta en que no hay una relación directa entre el grado de la sonrisa y el grado de oclusión de los ojos. En la sonrisa falsa el ojo está más abierto, el párpado inferior no es tan fino y las mejillas no están tan llenas.

Sonrisa boca cerrada.

- Cejas: Relajadas.
- Ojos: Crispados. Párpado superior ligeramente bajado mientras que el inferior sube contrayéndose con tensión.
- Boca: Agrandada con las esquinas estiradas en dirección a las orejas. Labios en forma de V. Labios afinados, los dos se mueven arriba presionando contra la calavera.
- Arrugas de expresión: Arrugas bajo los ojos en forma de sonrisa. Hundimiento del surco naso-labial. Hoyuelos.
- Similitud con otras expresiones: Similitudes con la falsa sonrisa. Diferenciándose de esta en que no hay una relación directa entre el grado de la sonrisa y el grado de oclusión de los ojos. En la sonrisa falsa el ojo está más abierto, el párpado inferior no es tan fino y las mejillas no están tan llenas.

Risa sofocada.

- Cejas: Relajadas.
- Ojos: Forma de media luna. Párpado superior ligeramente bajado mientras que el inferior sube contrayéndose con tensión.
- Boca: Ensanchada, tensada y cerrada. Los labios sonríen pero las mejillas están bajadas. Pueden estar un poco hacia delante. La línea entre los labios, recta y estirada.
- Arrugas de expresión: Arrugas bajo los ojos en forma de sonrisa. El surco naso-labial se contrae alrededor de la boca. Hoyuelos. El mentón se arruga.
- Similitud con otras expresiones: Esfuerzo. En esta expresión los ojos están apretados sin llenar el pliegue del párpado inferior y las mejillas están planas en vez de llenas. La boca es más fina y menos ensanchada.

Sonrisa melancólica

- Cejas: Hacia el centro de la cara. Inclinas hacia delante o ligeramente

enroscadas.

- Ojos: Ligeramente estrechados. El párpado superior ligeramente bajado mientras que el inferior permanece en tensión y alto.
- Boca: Ensanchada con las esquinas hacia atrás en dirección a las orejas. La línea labial estirada y en forma de V. Los labios suavizados y finos se mueven hacia delante presionando contra la calavera.
- Arrugas de expresión: leves patas de gallo y pliegues en forma de sonrisa bajo los ojos. El pliegue naso-labial se contrae alrededor de la boca. Aparecen hoyuelos en las mejillas. Arrugas horizontales en la frente sobre las cejas. Arrugas oblicuas sobre los párpados.
- Similitud con otras expresiones: Tristeza. En ésta la boca no está curvada hacia arriba. El pliegue naso-labial se acerca a la nariz y se arruga el mentón. En cualquier caso la sonrisa melancólica se interpreta más próxima a la tristeza que a la alegría debido a la fuerza de los ojos.

Sonrisa ávida

- Cejas: Pueden estar elevadas.
- Ojos: En forma de media luna. El párpado superior está extra elevado, puede llegar a enseñar el blanco sobre el iris mientras que el inferior permanece en tensión y alto.
- Boca: Ensanchada con las esquinas hacia atrás en dirección a las orejas. El labio superior está recto enseñando la dentadura superior. El labio inferior recto en el medio. Las mejillas llenas.
- Arrugas de expresión: Patas de gallo y pliegues bajo los ojos en forma de sonrisa. El pliegue naso-labial se contrae alrededor de la boca. Aparecen hoyuelos en las mejillas.
- Similitud con otras expresiones: Falsa sonrisa. En la sonrisa forzada el párpado superior y las cejas suelen estar levantadas pero no coinciden con el relleno de las mejillas y párpado inferior.

Sonrisa congraciante

- Cejas: Similar a como están en la expresión de miedo. Esto es, elevadas conjuntamente. Inclinas hacia delante aproximadamente un tercio de la ceja. Cejo ligeramente fruncido. En el caso de la sonrisa congraciante aparece una elevación extra en todo el conjunto. Así como un ligero fruncimiento extra.
- Ojos: Forma de media luna. El párpado superior se eleva ligeramente mientras que el inferior permanece en tensión y alto.

- **Boca:** Ensanchada con las esquinas hacia atrás en dirección a las orejas. La línea labial estirada y en forma de V. Los labios suavizados y finos se mueven hacia delante presionando contra la calavera.
- **Arrugas de expresión:** Arrugas horizontales en mitad de la frente. Pliegues verticales en mitad de las cejas. Patas de gallos y pliegues en forma de sonrisa bajo los ojos. Pliegue naso-labial hundido. Hoyuelos en las mejillas.
- **Similitud con otras expresiones:** No es similar a ninguna otra expresión.

Sonrisa astuta.

- **Cejas:** Las esquinas están inclinadas hacia abajo hacia el centro de la cara. El costado inferior de la ceja cae a nivel del párpado superior.
- **Ojos:** Estrechados por la presión de las cejas desde arriba y desde abajo por la presión del párpado inferior.
- **Boca:** Ensanchada con las esquinas hacia atrás en dirección a las orejas. La línea labial estirada y en forma de V. Los labios suavizados y finos se mueven hacia delante presionando contra la calavera.
- **Arrugas de expresión:** Líneas verticales entre las cejas, hoyuelos encima. Pliegues en forma de sonrisa bajo los ojos. Pliegue naso-labial hundido. Hoyuelos en las mejillas.
- **Similitud con otras expresiones:** No es similar a ninguna otra expresión.

Sonrisa desleal, corrupta.

- **Cejas:** Pueden variar. Pueden estar levantadas similares a haciendo un esfuerzo para prevenir la somnolencia o pueden estar levantadas de forma asimétrica.
- **Ojos:** El párpado superior cae sobre el iris cubriendo parte de la pupila. El párpado inferior puede estar estirado. Puede salirse de convergencia.
- **Boca:** Ensanchada con las esquinas hacia atrás en dirección a las orejas. La línea labial estirada y en forma de V. Los labios suavizados y finos se mueven hacia delante presionando contra la calavera.
- **Arrugas de expresión:** Arrugas horizontales a través de parte de la frente. Patas de gallo y pliegues en forma de sonrisa bajo los ojos. Pliegue naso-labial hundido. Hoyuelos en las mejillas.
- **Similitud con otras expresiones:** No es similar a ninguna otra expresión.

Sonrisa embarazosa

- **Cejas:** Relajadas.

- Ojos: Pueden estar casi cerrados o mirando abajo. Boca: Ensanchada con las esquinas hacia atrás en dirección a las orejas. La línea labial estirada y en forma de V. Los labios suavizados y finos se mueven hacia delante presionando contra la calavera.
- Arrugas de expresión: Patas de gallo y pliegues en forma de sonrisa bajo los ojos. Pliegue naso-labial hundido. Hoyuelos en las mejillas.
- Similitud con otras expresiones: No es similar a ninguna otra expresión.

Falsa sonrisa

- Cejas: Normalmente están relajadas pero pueden presentarse elevadas como en la sonrisa ávida.
- Ojos: Ligeramente estrechados pero no lo suficiente para coincidir con la sonrisa.
- Boca: Abierta con las esquinas estiradas hacia atrás. No asciende hacia las orejas como en la sonrisa real. El labio superior está recto enseñando la dentadura superior. El labio inferior tiene la sección central recta y alargada, enseñando la dentadura inferior y la esquina inferior de la boca. La forma de la boca es más cuadrada que en la sonrisa relajada.
- Arrugas de expresión: Patas de gallo y ligeros pliegues en forma de sonrisa bajo los ojos.
- Similitud con otras expresiones: Sonrisa clamorosa. En la sonrisa natural los párpados están más estrechos y el labio inferior y las mejillas más llenos. La forma de la boca es más angular en la parte media inferior, no tan cuadrada. La dentadura inferior no se muestra.

Falsa carcajada.

- Cejas: Relajadas.
- Ojos: Párpados estrechados, bajando el párpado superior y subiendo el inferior. El iris permanece visible mirando hacia adelante. También pueden aparecer cerrados de forma relajada con la forma curvada de las pestañas y algunas arrugas.
- Boca: Ampliamente abierta. Esquinas altas. Labio superior recto y labio inferior recto en el centro y anguloso en los extremos. Se muestra la dentadura superior y solo las puntas de la inferior.
- Arrugas de expresión: Ligeras patas de gallo y pliegues poco profundos bajo los ojos que pueden ser ausentes. Pliegues de la nariz a la barbilla. Hoyuelo en la barbilla.

- Similitud con otras expresiones: Carcajada. Cuando la carcajada es sincera el ojo está casi completamente escondido. Los párpados se contraen con fuerza, las mejillas presentan forma de bola y las patas de gallo son profundas.

Capítulo 3 Reproducción estática de las diferentes expresiones de risa

En este capítulo se expone el proceso completo de la generación de expresiones faciales de risa. El primer apartado se refiere a la generación del sistema de rig, en el se incluye la construcción de la geometría del avatar, la definición de la interfaz y la definición del esqueleto facial y de las deformaciones que permitirán la animación. En el segundo apartado se generan manualmente las 13 expresiones de risa que constituyen nuestro estudio. Finalmente en un último apartado se hace una evaluación de estas expresiones.

3.1 Construcción de un rig facial para la descripción de los diferentes tipos de risa

Siguiendo los puntos que hemos desarrollado en el punto 2.5 nuestro sistema de rig debe considerar los siguientes **desafíos**:

Diversidad de caras: En este sentido pretendemos reproducir de forma realista las diferentes sonrisas, de manera que las variaciones de forma que adopten nuestros modelos geométricos se deben corresponder a las deformaciones debidas a la acción muscular que se produce en la realidad. Debemos considerar también que los personajes fantásticos o de dibujos animados puedan reproducir estas expresiones de alegría, con las sutilezas que les diferencia, por lo que los personajes deberán tener una estructura básicamente antropomórfica pero no tienen porque ser realistas. Por estas razones hemos optado por la creación de dos personajes uno realista y otro no que nos servirán para observar la fiabilidad de nuestras representaciones.

Inconsistencia de los movimientos faciales: En 1970 Mori[89] introdujo la teoría del "Valle misterioso" según la cual tanto los movimientos antinaturales como la apariencia antinatural de un gesto estático de un agente virtual son fácilmente reconocidos debido a la sensibilidad que posee el ser humano

referida a la interpretación facial. De manera que si bien es cierto que con la apariencia antropomórfica se logra una inmediata aceptación del personaje, existe un cierto umbral en el cual la apariencia "casi humana" provoca un rechazo. Esta teoría se resume según el gráfico de la figura 5.

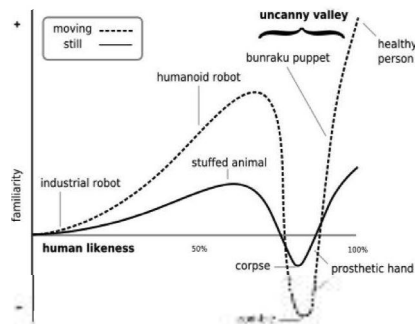


Figura 5: Gráfico de la hipótesis del "Valle misterioso" (Original de Mori 1982)

A la hora de elegir un buen sistema de *rig* es necesario poder asegurar no solo la representación realista de movimientos y gestos sino que a priori que el sistema deberá ser capaz de poder representar todas las necesidades que tendrá el animador, esto incluyen los movimientos realistas más los movimientos propios del personaje fantástico. Teniendo en cuenta todo lo expuesto en las referencias del capítulo anterior, diremos que para garantizar la realización de movimientos precisos y realistas, es necesario que con el *rig* se puedan simular todas las acciones musculares detalladas en el FACS y que aplicadas según las descripciones de Faigin den lugar a las 13 expresiones que buscamos reproducir. Efectivamente, como se observa en la figura 6 el *rig* que hemos construido es capaz de reproducir las AUs descritas por Ekman y Friesen. En la imagen se contraponen en la parte izquierda las fotos originales que describen las AUs y en la parte derecha aparece una ventana 3D que a su vez está dividida en dos partes, en una aparece la interfaz del *rig* que explicaremos más adelante y en la otra nuestro personaje primitivo (o no realista) ejecutando la misma AU.

Falta de norma: No existe una norma estándar para definir el sistema de *rig*. En este sentido nuestro sistema tendrá como objetivo principal su uso por parte de los artistas, es decir que pretende ser lo suficientemente intuitivo para que cualquier animador sea capaz de reproducir las expresiones. Por ejemplo mediante la definición de un sistema, que de forma interactiva, deforme la malla de la manera equivalente a lo que haría la acción muscular en la realidad.

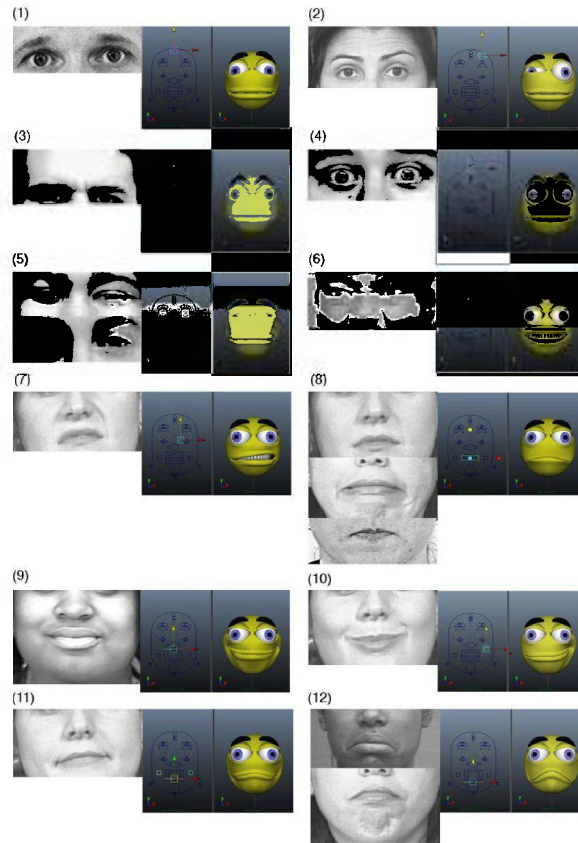


Figura 6: Reproducción de algunas de las AUs del FACS: (1) AU 1: Inner brow raiser, (2) AU 2: Outer brow raiser, (3) AU 4: Brow lowerer, (4) AU 5: Upper lid raiser, (5) AU 6: Cheek raiser, AU 7: Lid tightener, (6) AU 9: Nose wrinkler, (7) AU 10: Upper lid raiser, (8) AU 11: Nasolabial deepener, AU 20: Lip stretchner, AU 24: Lip pressor, (9) AU 12: Lip corner puller, (10) AU 13: Cheek puffer, (11) AU 14: Dimpler, (12) AU 15: Lip corner depressor, AU 17: Chin raiser

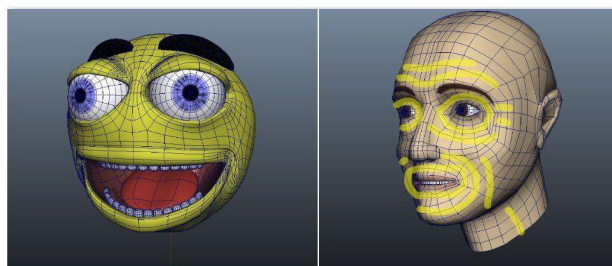


Figura 7: Personaje primitivo con malla de superficie de subdivisión y personaje naturalista con malla poligonal (*edge loop* en amarillo)

A la hora de elegir la **morfología** de los modelos para representar las risas se ha optado por la construcción de dos personajes que según las definiciones McLaughlin[54] se corresponderían a uno primitivo, asociado a la estética de dibujo animado y otro naturalista, es decir que se parece y comporta como una persona en el mundo real. En lo referente a la elección del tipo de geometría para la construcción de los modelos se ha optado por el uso de superficies de subdivisión para el primero y una malla poligonal para el segundo, con el objetivo de ejemplificar diferentes técnicas. Para el caso de la malla poligonal se han tenido en cuenta las técnicas de ciclo de bordes[90] (*edge loop*) que garantizan la coherencia topológica de la geometría y optimizan las deformaciones faciales. Ver figura 7. Esta técnica es independiente del tipo de malla geométrica que se usa y trata de optimizar la distribución de la malla para optimizar la deformación basada en criterios anatómicos. La geometría se construye siguiendo las líneas musculares, esto ayuda a mantener su ubicación en la cara y proporcionar una animación más realista ya que la deformación seguirá esta línea. En la figura 8 se observa el detalle de la comisura de los labios de dos versiones de la malla del personaje primitivo, en la primera la solución aplicada es con superficies de subdivisión y en la segunda tenemos una malla poligonal con *edge loop*.

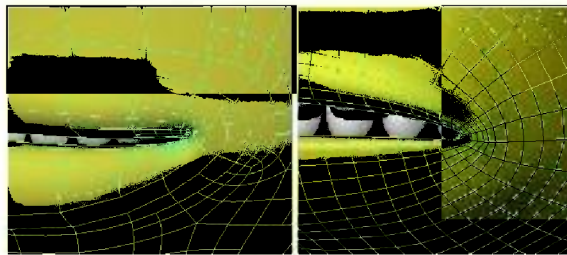


Figura 8: Coherencia topológica (*edge loop*)

Aunque sea un tema que depende en gran manera de la apariencia de la cara que queramos modelar, en el caso del modelo primitivo se ha optado por el uso de superficies subdivisión de Catmull-Clark[91]. Que de una parte ofrece la estandarización de uso de los polígonos y por otra la reducción de información de la geometría propia de las NURBS. Además de estar implementadas en la mayoría de programas de creación 3D. La ventaja que supone tener una malla de puntos más reducida que la malla poligonal, será apreciable sobre todo, a la hora de generar las geometrías que represen la acción muscular, facilitando la manipulación de puntos y evitando la aparición de errores, como los de cruce de aristas que se producen en mallas muy densas.

A la hora de diseñar un personaje primitivo se han tenido en cuenta además otras consideraciones: Se ha decidido elegir un personaje inspirado en el

diseño del emoticón smiley[92]. El *smiley* se ha convertido en un elemento fundamental de la cultura de Internet, con GIFs animados y otras representaciones gráficas, así como el omnipresente emoticón textual :-) o :). De ahí surgieron los demás emoticonos. Aunque la estructura facial del *smiley* pueda presentar algunos inconvenientes inherentes a su propio diseño, es una estructura facial que permite la simulación de todas las actividades faciales de una representación realista con la libertad de trabajar con un modelo de caricatura. Entre los problemas de diseño que puede presentar cabe citar: la desproporción entre la boca y la zona de las mejillas, esto puede provocar que el hinchamiento de la mejilla que se produce en determinados tipos de sonrisa y que sirve como un elemento diferenciador entre una expresión falsa de otra que no lo es, sea poco apreciable; el gran tamaño de los ojos y boca pueden originar problemas sobre todo en lo que refiere a la deformación de la geometría de la piel que los rodea, así para cerrar los párpados, los escasos puntos que rodean al ojo han de bastar para cubrir la esfera ocular, a su vez, por su diseño, la boca podrá abrirse mucho más que una boca realista sin que ello deba provocar errores por intersección con las geometrías que representan el interior; otro problema propio del diseño es la falta de nariz, aunque esta como tal no interviene en las expresiones faciales, el estiramiento del músculo elevador del labio superior que interviene sobre todo en las risas más marcadas provoca un ensanchamiento de las aletas nasales que obviamente no estará representado en nuestro modelo.

Animación. Nuestra intención es desarrollar la simulación de la risa dentro del entorno de una aplicación de creación 3D, de manera que las diferentes técnicas de animación que ofrece el paquete se podrán aplicar a nuestro objetivo. En concreto la técnica clásica de animación por claves *keyframe* definidas en los manipuladores del *rig* será la que definirá el movimiento del personaje. Además se aprovecharán todas las posibilidades que ofrecen este tipo de aplicaciones como uso de texturas, iluminación y cálculo de imágenes.

En la actualidad las aplicaciones comerciales de producción 3D se comunican con los sistemas de MOCAP con *pluggins* y convenciones para compartir archivos. De forma generalizada los datos generados por un dispositivo de captura de movimiento pueden verse como una secuencia de valores o claves que evoluciona en el tiempo. Este es el concepto de curva de animación que se implementa en la totalidad de las aplicaciones de animación.

Interfaz del usuario. Entre las diferentes opciones para definir la interfaz del *rig* se ha descartado el sistema basado en ventana de parámetros; básicamente por razones de uso, asociar un nombre, con un botón de deslizamiento o con un valor numérico, a un movimiento puede ser más ambiguo que la asociación entre un manipulador gráfico al movimiento cuyo efecto será similar al deseado en el modelo. El sistema elegido es la interfaz basada en una ventana 3D con manipuladores 2D, similar a las soluciones propuestas por Alexander

et al.[62] o Digital tutors[61]. Con este último se tiene un control exhaustivo de la deformación de la malla geométrica para la simulación de los músculos faciales a la vez que una herramienta intuitiva que permite realizar de forma sencilla todas las acciones descritas en el FACS. En la figura 9 se observan todos los controladores del rig facial con el nombre que reciben que nos da una idea de su función.

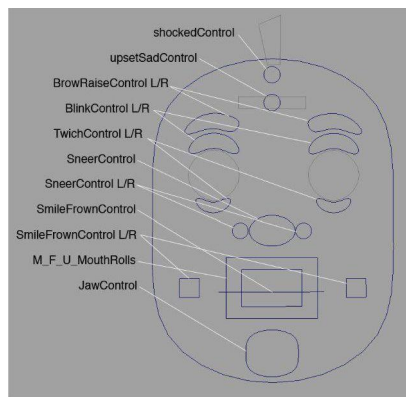


Figura 9: Interfaz del rig, según Digital tutors

Sobre estos controladores el animador aplicará transformaciones lineales sencillas de translación, rotación o escalado que serán las que tendrán asociadas las deformaciones pertinentes sobre el modelo. En la figura 10 se observan ejemplos de la activación de algunos de estos controladores, en la primera fila una translación negativa en el eje Y del *JawControl* abre la boca del modelo y una translación negativa en el eje Y de los dos *BlinkControl* cierra los ojos, en la segunda fila una translación positiva en el eje Y del *BrowRaiseControl* eleva una ceja a la vez que abre mucho un ojo y entrecierra el otro y una translación positiva en el eje Y del *ShockedControl* eleva las dos cejas a la vez que abre mucho los ojos, en la tercera fila una translación positiva en Y del *SmileFrownControl* estira las comisuras de los labios hacia arriba y una translación negativa del mismo controlador tira de las comisuras hacia abajo, en la última fila una translación positiva en el eje Y del *UpsetSadControl* tira de las cejas hacia arriba y una negativa del mismo controlador las tira hacia abajo y frunce el ceño.

Combinando estos controladores de forma conveniente se pueden reproducir las unidades de acción descritas en el FACS y por tanto la reproducción de cualquier expresión facial. Concretamente en la tabla 3 observamos las acciones faciales AUs que activan los controladores de nuestro sistema. En la tabla se han obviado las simetrías a derecha e izquierda que provocan las mismas AUs.

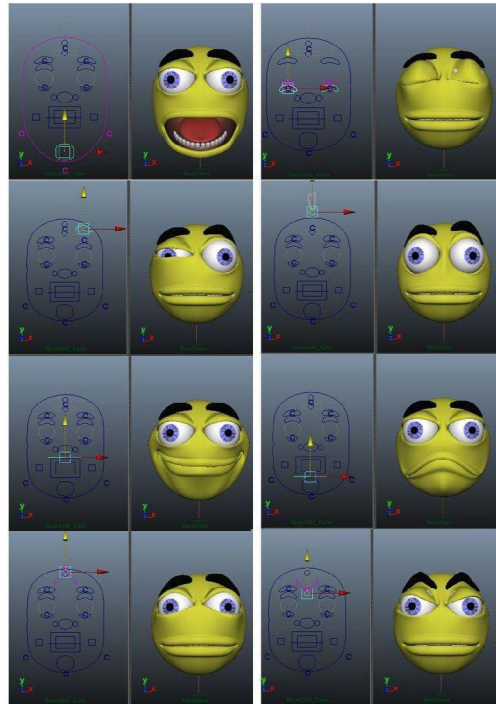


Figura 10: Acción de los controladores del rig sobre el modelo primitivo

Controlador del rig	AU
upsetSad	AU 1: Inner Brow Raiser, AU 4: Brow Lowerer, AU 9: Nose Wrinkler, AU 11: Nasolabial Deepener
shocked	AU 2: Outer Brow Raiser, AU 5: Upper Lid Raiser
twich	AU 6: Cheek Raiser, AU 7: Lid Tightener
sneer	AU 10: Upper Lip Raiser, AU 11: Nasolabial Deepener
smileFrown	AU 6: Cheek Raiser, AU 12: Lip Corner Puller, AU 13: Cheek Puffer AU 14: Dimpler, AU 15: Lip Corner Depressor, AU 17: Chin Raiser
jaw	AU 24: Lip Pressor, AU 25: Lips part, AU 26: Jaw Drop, AU 27: Mouth Stretch
blink	AU 41: Lid droop, AU 42: Slit, AU 42: Slit, AU 43: Eyes Closed, AU 44: Squint, AU 45: Blink, AU 46: Wink
neck	AU 55: Head Tilt Left, AU 56: Head Tilt Right

Tabla 3: Relación de los controladores del rig con las AU que activan

Deformadores. Para definir las deformaciones de las mallas de los personajes nuestro *rig* usa la combinación de técnicas basadas en hueso y técnicas basadas en interpolación geométrica *blendshape* [61], con las primeras se controlará el movimiento de la mandíbula (JawControl) y con las segundas todos los cambios en la malla que representa la acción muscular facial restante. En concreto se definen las interpolaciones geométricas entre el modelo base y 25 nuevas geometrías creadas aplicando deformaciones locales de forma manual. Estas geometrías se dividen en dos grupos: el primero de 11 geometrías para la simulación de la parte superior de la cara y que afectaría de forma visual a los ojos y a las cejas y un segundo con 14 geometrías cuyo efecto visual se centra en la boca y las mejillas. Ver imágenes 11 y 12.

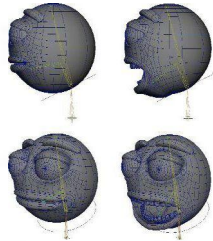


Figura 11: Deformación basada en huesos para el controlador de la mandíbula

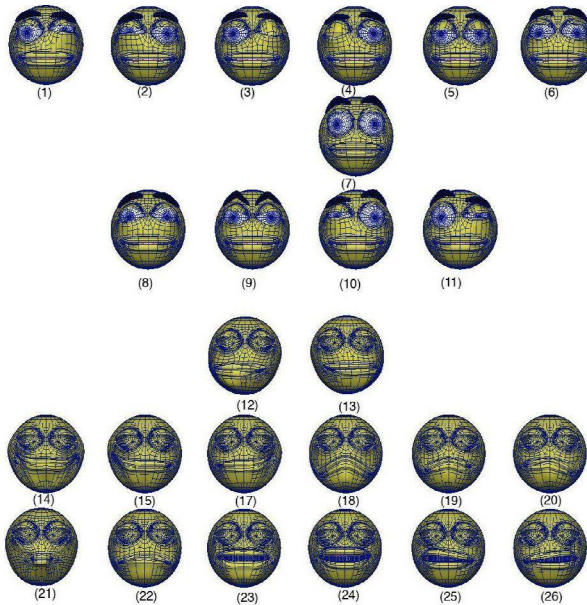


Figura 12: Geometrías para el *blendshape* del personaje primitivo

Cada una de estas geometrías será definida como objetivo de la deformación de *blendshape* desde la geometría que representa la expresión neutra. Una vez definida esta deformación se establecerá la relación entre la transformación lineal y su controlador correspondiente. De manera que cuando el animador actúe sobre el controlador se ejecute la deformación previamente definida. De esta manera tenemos a la geometría 1 le corresponde el controlador *LeftTwich*, a la 2 el *RightTwich*, a la 3 el *LeftBlink*, a la 4 el *RightBlink*, la 5 y la 6 están asociadas al controlador que gestiona el punto de vista y se usan para mirar abajo y arriba respectivamente, a la 7 el *Shocked*, la 8 y la 9 están asociadas al *UpsetSad* siendo la primera para valores positivos y la segunda para los negativos, a la 10 el *LeftBrowRaise*, a la 11 el *RightBrowRaise*, la 14 y la 18 están asociadas al controlador *SmileFrown* la primera se activa con translaciones positivas en Y, y la segunda con negativas, la 15 y la 19 activa el *LeftSmileFrown*, la 17 y la 20 el *RightSmileFrown*, gestionando de forma similar valores positivos y negativos, a la 24 le corresponde el *Sneer*, a la 25 el *LeftSneer*, a la 26 el *RightSneer* y finalmente la 12, 13, 21, 22 y 23 están asociadas al controlador *M_F_U_MouthRolls*, este controlador tiene diferentes funcionalidades: Translaciones a la derecha e izquierda simulan movimientos de la mandíbula en este mismo sentido, translaciones positivas en Y simulan el estiramiento de los labios que se producen cuando se pronuncia el sonido F, escalados en X encogen los labios de forma similar a cuando se pronuncia el sonido U y escalados en Y hacen que se presionen los labios de forma similar a cuando se pronuncia la M.

3.2 Reproducción de las 13 expresiones de risa

Para la reproducción de las risas según la clasificación de Faigin con el *rig* construido se han seguido varios caminos: por una parte la simulación visual ad hoc de la ilustración de la figura 4, y por otra la reproducción de la actividad muscular descrita por Spaltehoz[18] en la representación de cada una de las sonrisas. A parte de la guía que representan las ilustraciones de Faigin se han buscado ejemplos cinematográficos representativos de cada una de las expresiones.

Las expresiones de actores reales se han escogido a partir de una colaboración con muchocine.net, web independiente de crítica de cine en la que sus colaboradores elaboraron diferentes listas con las películas que representan las secuencias con risas más representativas del cine. Los expertos consultados elegían de entre las películas de todos los tiempos aquellas que de alguna manera, según su criterio, tuviesen algún aspecto que las hiciese

representativas respecto a la risa de alguno de sus personajes. Han sido encuestados 9 expertos, cada uno de los cuales ha elaborado su propia lista.

Como resultado de la encuesta se ha confeccionado una lista con 28 títulos de películas. Algunos de los resultados han sido despreciados debido a que involucraban a personajes no realistas, o con deformaciones que dificultaban o hacían imposible la aplicación correcta del seguimiento de puntos faciales; entre ellas están las secuencias de las risas del Joker en “Batman” de Tim Burton o las de Gary Oldman en “Drácula” de Francis Ford Coppola.

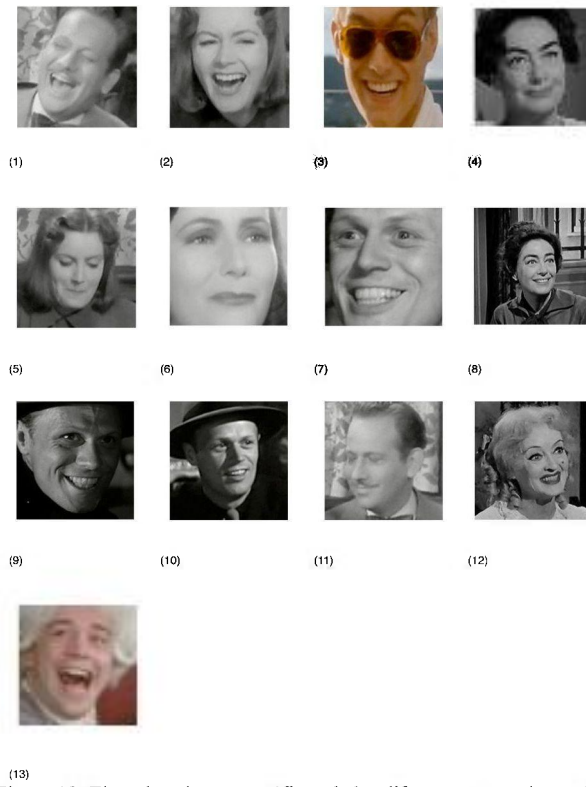


Figura 13: Ejemplos cinematográficos de las diferentes expresiones de alegría: (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada

Dependiendo del contexto de las diferentes escenas y atendiendo a la descripción morfológica y muscular que hace Faigin se han encontrado los ejemplos representativos de las 13 expresiones que se observan en la figura 13 y que se corresponden a las películas con mayor número de votos de la

encuesta que son: “Ninotchka”[93] de Ernst Lubitsch (imágenes 1, 2, 5, 6 y 11), “The talented Mr Ripley”[94] de Antony Minghella (imagen 3), “What Ever Happened to Baby Jane”[95] de Robert Aldrich (imágenes 4, 8 y 12), “Kiss Of The Death”[96] de Henry Hathaway (imágenes 7, 9 y 10) y “Amadeus”[97] de Milos Formann (imagen 13).

A partir de la información obtenida del trabajo de Ruch respecto a los músculos que actúan en la acción de la sonrisa y de la descripción precisa de Spalteholz tenemos una guía de como se va a deformar la malla geométrica de nuestro modelo para simular dicha acción.

En la obra de Faigin encontramos representaciones gráficas de las diferentes acciones musculares que definen cada expresión facial considerada. Así mismo se incluyen indicaciones por medio de flechas, de como se produce el movimiento durante la acción que se describe. Esta información gráfica será usada como guía a la hora de definir las deformaciones de nuestra geometría que después serán usadas como *blendshape* para la simulación de la acción muscular.

Con estas tres referencias anteriores, se propone la tabla 4. Esta tabla aporta la descripción de la actividad muscular para cada una de las expresiones de risa frente a la que proporciona Ruch en donde la clasificación se hace según los músculos considerados por los autores que han estudiado el tema. En la tabla la x implica la acción del músculo, el · indica una breve tensión muscular, la ? denota que puede aparecer en determinados individuos aunque no es generalizable y el ° que nuestros modelos no contemplan la acción de este músculo ya sea porque nuestro modelo primitivo no realiza estas acciones al carecer de nariz o porque la actividad del músculo masetero se simula con la acción del hueso de la mandíbula.

	Parte superior			Parte central						Parte inferior										
	Frontal med.	Frontal lat.	Superciliar	Pyramidal	Orbicular orb.	Orbicular par.	Zigom. May.	Zigom. Men.	Elev. L. Sp.	Dien	Elev. L. Sp.	Alet	Nasal	Budialar	Risorio	Orbicular lab.	Triangular lab.	Mentonlar	Masetero	Canino
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Carcajada extrema	·	·	·	·	x	x	x	x	x	x	°	°	°	°	x	°	x	°	x	°
Carcajada	·	·	·	·	x	x	x	x	x	x	°	°	°	°	x	°	x	°	x	°
Sonrisa clamorosa	·	·	·	·	x	x	x	x	x	x	°	°	°	°	x	°	x	°	x	°
Sonrisa					·	x	x							°	°	°	°	°	°	°
Risa sofocada					x	x	x							°	°	°	°	°	°	°
Sonrisa melancólica	x	x	·	·	·	·	·	x	x					°	°	°	°	°	°	°
Sonrisa ávida	·	x			x	x	x	x	x	x				°	°	°	°	°	°	°
Sonrisa congraciante	x	·	·	·	x	·	·	x	x					°	°	°	°	°	°	°
Sonrisa astuta		x	x	x	x	x	x	x	·	·				°	°	°	°	°	°	°
Sonrisa desleal	x	x			x	·	x	x						°	°	°	°	°	°	°
Sonrisa embarazosa					x	x?	x	x						°	°	°	°	°	°	°
Falsa sonrisa	·	x?			·	·	x	x	x	x				°	°	°	°	°	°	°
Falsa carcajada					·	·	x	x	x	x				°	°	°	°	°	°	°

Tabla 4: Músculos activados en las diferentes expresiones

A nivel muscular en nuestros modelos se realizarán las siguientes simplificaciones:

- No se diferenciará entre las acciones del músculo superciliar (3) y las del piramidal (4), el primero se encuentra en la cara profunda de la piel de la parte interna de la ceja y ocasiona la producción de un pliegue vertical de la piel, por fuera de la raíz de la nariz, mientras que el segundo tiene su origen en los huesos nasales y termina en la cara profunda de la piel entre las dos cejas, su acción provoca la aparición de arrugas en la raíz nasal. Estas acciones musculares aparecen sobre todo en la sonrisa astuta y las asociaremos a una única deformación que se activará juntamente a la acción del músculo orbicular de los párpados al tirar las cejas hacia adentro y abajo.

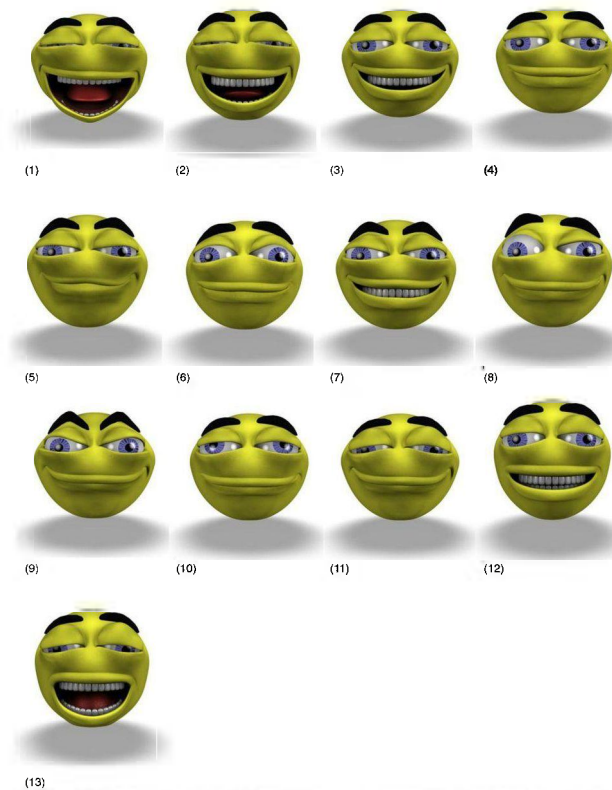
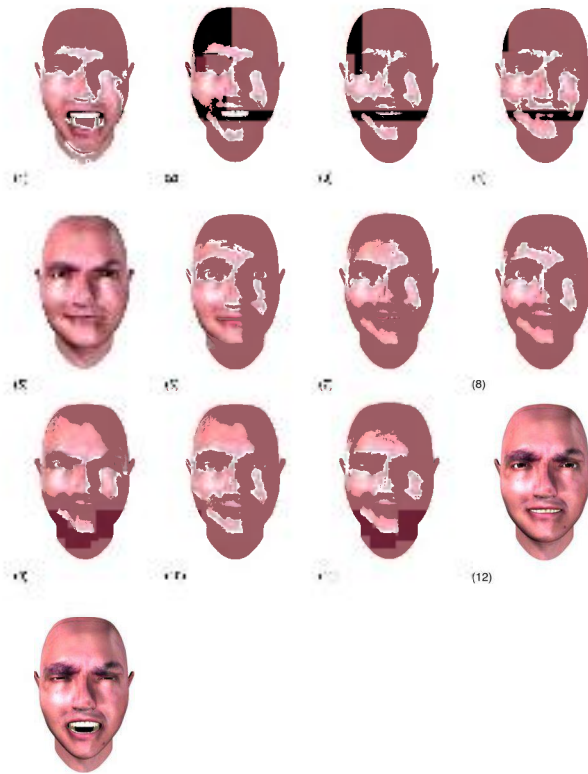


Figura 14: Reproducción de las diferentes expresiones de alegría: (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada

- No habrá diferencia tampoco entre el elevador del labio superior que enseña los dientes (9) y el elevador del labio superior que dilata las aletas nasales (10). Tampoco se considerará el músculo nasal (11), cuya acción es también la de dilatar las fosas nasales. De hecho en la descripción que hace Spaltheholz de estos músculos distingue en primer lugar al elevador superficial común del ala de la nariz y del labio superior y en segundo al elevador profundo y la acción de estos dos es la misma que la del cigomático menor y consiste en la elevación del ala nasal y el surco nasolabial. Nuestros modelos contemplaran una única deformación diferenciando la acción del cigomático menor de la parte derecha e izquierda de la cara.



(13)
 Figura 15: Reproducción de las diferentes expresiones de alegría: (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada

- Y por último la acción del bucinador (12) se simplifica y no se reproducen los hoyuelos en las mejillas, ya que ello implicaría necesariamente una

subdivisión mayor de la malla que podría incidir en la coherencia topológica argumentada anteriormente.

Respecto a la tabla anterior, se observa además la coincidencia total entre la carcajada y la carcajada extrema y que en todas las expresiones aparece la acción de los músculos cigomático mayor (7) y con diferentes intensidades, la parte baja del orbicular de los párpados (6); a los que Duchenne llamó músculos de la alegría o la benevolencia.

Con estas consideraciones en mente se ha manipulado la interfaz de nuestro *rig* creando las 13 expresiones que se observan en la figura 14 aplicadas al personaje primitivo y en la figura 15 aplicadas al personaje naturalista.

3.3 Evaluación de las expresiones generadas

Se ha realizado un experimento que demuestra que la generación de imágenes estáticas es insuficiente para la representación de determinadas expresiones. A la vez que se ha buscado cual de las distintas expresiones es más reconocible.

La evaluación de la calidad de las imágenes generadas se ha realizado mediante una encuesta. Se han llevado a cabo dos tests en los que se emparejaban expresiones faciales de alegría, en el primero se relacionaban las representaciones con los dibujos originales de Faigin con la reproducción de dichas expresiones por parte de nuestro modelo primitivo, mientras que en el segundo se tenían que relacionar las 13 expresiones de alegría que representaban sendos fotogramas también con nuestro modelo primitivo. Los fotogramas habían sido escogidos de películas representativas de la cinematografía según se ha explicado en el punto anterior. Se pretendía que los sujetos encuestados emparejasen las imágenes que se ven en la figura 16. En general los fotogramas no se corresponden ni a primeros planos de la cara ni a encuadres perfectamente centrados y rectos, en este sentido se ha procedido a un recorte y escalado de la imagen.

Las pruebas de test consistían en un formulario colgado en la web <http://dmi2.uib.es/~cinemiquel/> en donde los usuarios de forma anónima relacionaban parejas de imágenes mediante un menú numérico desplegable. Los encuestados fueron llamados a la participación por correo electrónico masivo.

De la primera prueba se obtuvo un muestra de 38 respuestas a las 13 preguntas de emparejamiento de imágenes. Exceptuando la primera, correspondiente a la carcajada extrema que se ofrecía como ejemplo se ha

obtenido una media de correlación correcta del 37,94%, con una desviación estándar de 7,87. Siendo la falsa carcajada la pareja más reconocida seguida de la sonrisa sofocada, la sonrisa clamorosa, la sonrisa melancólica, la sonrisa desleal y la carcajada; con valores superiores a la mediana. En el otro lado la mayor cantidad de correlaciones incorrectas se deben a la confusión entre la sonrisa congraciante y melancólica de nuestro modelo que se confunden con la sonrisa desleal de Faigin, seguidas de la sonrisa desleal y la clamorosa que se confunden respectivamente con la congraciante y la ávida de Faigin.

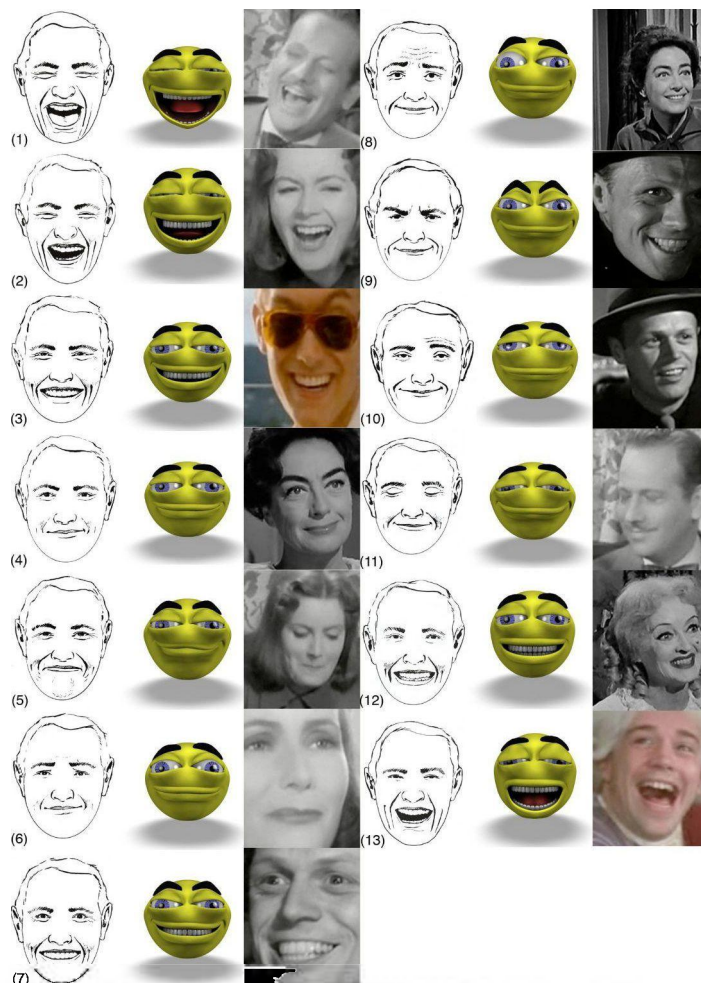


Figura 16: Emparejamiento de las expresiones de alegría (dibujo de Faigin, personaje primitivo y fotograma): (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada

En la segunda prueba donde se relacionaban fotografías de actores interpretando secuencias con las 13 sonrisas y nuestro modelo se obtuvo una muestra de 29 respuestas. Exceptuando también la primera que se correspondía al ejemplo se ha obtenido solo un 18,20% de correlaciones correctas con una desviación de 3,73. Las únicas parejas reconocidas se corresponden a las expresiones de carcajada, sonrisa melancólica y sonrisa con la boca cerrada. En dónde se han registrado mayores errores ha sido en la sonrisa desleal de nuestro modelo que ha sido confundida con el fotograma correspondiente a la sonrisa embarazosa, le siguen la falsa sonrisa confundida con la sonrisa ávida, la complaciente con la melancólica y finalmente la astuta con el fotograma representando a la sonrisa sofocada.

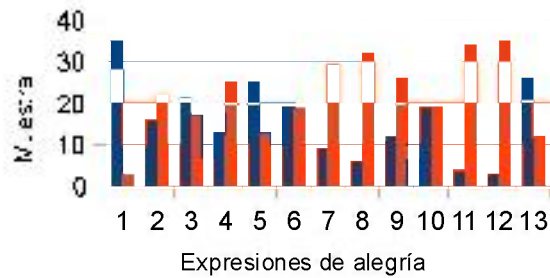


Gráfico 1: Test de emparejamiento del modelo primitivo con las representaciones de Faigin

En los gráficos 1 y 2 se observan las correlaciones correctas (en azul) y las incorrectas (en rojo) de ambos tests, en eje de las X se representan las 13 preguntas correspondientes a las expresiones faciales de alegría, dispuestas en el mismo orden con que aparecen en las figuras 4 y 14.

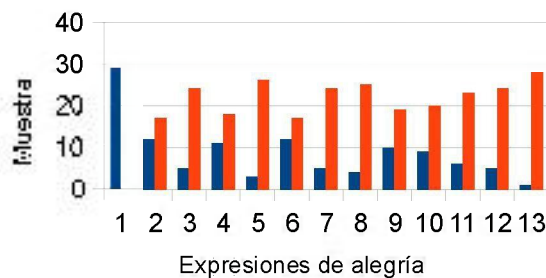


Gráfico 2: Test de emparejamiento del modelo primitivo con las fotografías de actores

Se concluye rápidamente en que la tarea de relacionar imágenes estáticas basadas en la reproducción de la actividad muscular según guías visuales a la activación de los músculos implicados en la expresión resulta ineficaz. Estos datos ponen de manifiesto una serie de problemáticas:

- Diferentes sujetos ejecutan la misma expresión con sutiles diferencias que hacen que un test de equivalencia basado en una imagen sea menos válido, estas diferencias ya las apunta Faigin en su obra, y se reflejan también en nuestras imágenes de prueba, por ejemplo en el caso de la sonrisa sofocada puede aparecer oclusión ocular, en la sonrisa astuta el elevador del labio superior puede hacer que los dientes sean visibles, las asimetrías de expresiones como la sonrisa congraciante y la desleal pueden darse en menor o mayor grado etc.
- Para algunas expresiones es fundamental verlas contextualizadas en la escena, expresiones como la sonrisa melancólica, la ávida, la congraciante, la desleal y la embarazosa se comprenden adecuadamente cuando son visualizadas en su contexto.



Figura 17: Expresiones de alegría relacionadas en un mismo contexto

Por ejemplo la secuencia de imágenes que se muestra en la figura 17 pertenecen a la misma escena de la película *Ninotchka* a la que antes nos hemos referido, en ella el personaje que interpreta Melvin Douglas acaba de caerse de la silla, el personaje de Greta Garbo reprime su sonrisa (5), en el contraplano Melvin se muestra avergonzado (11 sonrisa embarazosa), en el siguiente Garbo rompe en carcajada (2) y finalmente Douglas le corresponde con la carcajada extrema (1).

El análisis del contexto en el que se produce una expresión facial de alegría cae fuera del objetivo de este estudio aunque se pone de manifiesto la necesidad de evaluar la expresión en el tiempo. Todas las expresiones faciales implican una evolución temporal, aunque el análisis de la imagen en el zenit de la expresión sea fundamental la secuencia de imágenes previas son también de vital importancia. Dado que la risa es un fenómeno temporal, en este experimento vemos que no se pueden correlacionar bien imágenes estáticas. Lo que abriría el paso a la necesidad de la captura automática que introducimos en el siguiente capítulo.

Capítulo 4 Método para la reproducción animada de las diferentes expresiones de risa

En este capítulo se presenta un método para la reproducción automática de las expresiones de risa. En el primer apartado introducimos las necesidades del sistema. En el segundo introducimos el método de captura de las características faciales.

4.1 Introducción

Nuestro interés está en reproducir los diferentes tipos de risa teniendo en cuenta su desarrollo en el tiempo, esto implica la reproducción de una secuencia animada que se iniciaría en la expresión neutra y terminará en una expresión en donde se refleje el valor máximo de cada expresión. Una simulación realista es más compleja que la simple activación de la deformación en una determinada clave de tiempo y la asignación de un determinado número de cuadros. La aceleración / frenado del movimiento y la temporalización son algunas de las características básicas que un animador con una cierta experiencia deberá aplicar para conseguir un efecto creíble y realista.

Con el objetivo de facilitar esta tarea al animador o automatizarla para el caso de un animador inexperto proponemos un sistema que integra la captura de movimiento y sea capaz de generar las expresiones de nuestro interés.

Nuestro sistema debe ser capaz de sintetizar diferentes sonrisas a partir de secuencias de imágenes para después poder transferirlas a diferentes avatares sintéticos. Nuestro interés está en ofrecer un sistema rápido, sencillo y de bajo coste e integrable en los entornos de animación que permita ajustes a los animadores. Descartaremos, por tanto, de entrada, usar marcadores sobre los actores, iluminaciones específicas y escáneres 3D. Elegiremos un sistema que nos permita usar secuencias de vídeo grabadas con cámaras estándar o a

partir de simples secuencias de vídeo por diferentes razones: nos interesa la posibilidad de reproducir expresiones más o menos míticas ejecutadas por actores que no tienen por que estar en activo, desde el punto de vista de la metodología cinematográfica es preferible grabar las expresiones faciales en el mismo set donde se graba la acción del actor que no repetir solo la acción facial delante de un dispositivo especial y finalmente no debemos obviar que el animador no tiene por que ser buen actor, de manera que para nuestros intereses no consideraremos métodos que incluyan el tiempo real de forma que tenga que ser el animador el que realice la expresión.

El flujo de trabajo que proponemos consiste en elegir la expresión de un conjunto de secuencias de vídeo, de esa expresión extraer las curvas de animación que son importadas a diferentes controladores del *rig* del avatar y finalmente ajuste a esas curvas por parte del animador para afinar el movimiento.

4.2 Adquisición de los datos para la simulación de las expresiones de risa.

A la hora de decidir el sistema de seguimiento facial automático a usar, se ha optado por el FaceTracker de Saragih[98]. Las razones de dicha elección son que la aplicación es de código abierto, que no requiere ningún hardware especial para la toma de información y que de forma versátil se puede aplicar tanto a datos capturados desde la cámara como a secuencias de vídeo grabadas de forma específica o procedentes de bases de datos ya existentes o extraídas de cualquier película.

Como resultado de la ejecución de FaceTracker sobre una secuencia de imágenes se obtienen los valores en el plano XY de los 66 puntos de la malla (ver fig. 18) para cada cuadro de la secuencia. La aplicación es versátil y acepta ficheros de vídeo de diferente formato y resolución. La figura 19 pertenece a la ejecución de la aplicación sobre un plano de la película Ninotchka mencionada en el capítulo anterior, en ella se observan los puntos resultantes, en rojo los del óvalo de la cara y nariz y en amarillo los de las cejas, ojos y boca.

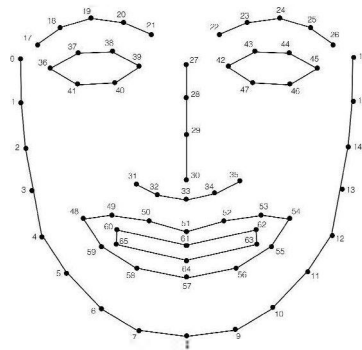


Figura 18: Puntos de la malla de Face Tracker



Figura 19: Resultado del Face Tracker de Saragih aplicado a la secuencia de vídeo antes citada

La aplicación nos devuelve la posición de los 66 puntos en el plano imagen para cada cuadro del plano. Con estos valores XY de cada punto se pueden generar 66 ficheros de animación de un entorno 3D que nos mostrarían la evolución en el tiempo de cada punto recogido en la aplicación de seguimiento. De estos datos extraeremos la información que será asignada a los controladores del *rig* y permitirá simular el movimiento original.

Capítulo 5 Cálculo de los valores animados de los controladores de expresión facial

En este capítulo se expone el método aplicado para la obtención de la animación de las expresiones faciales. Se detallan para cada controlador del rig las diferentes incidencias y particularidades encontradas.

5.1 Parametrización facial

Con el objetivo de verificar el método de extracción de las características faciales en el tiempo y su posterior asignación a los controladores del *rig* se ha usado la base de datos Cohn-Kanade[99]. Se pretende visualizar la expresión generada con el cálculo que se presenta en este punto junto con la expresión original de donde proceden los datos, de manera que se compruebe de forma rápida la fiabilidad de la expresión.

Esta base de datos, en su versión original del 2000, contiene secuencias de entre 9 y 60 cuadros de un total de 210 individuos de entre 18 a 50 años de edad, con representación de ambos sexos y de diferentes grupos raciales. La resolución de las imágenes es de 640x480 en niveles de gris y en color, la iluminación es relativamente uniforme y el encuadre es un primer plano frontal del sujeto. Todas las secuencias están convenientemente etiquetadas con las AUs que se activan. En la actualización CK+ del 2010 se amplía el número de secuencias y el de sujetos, además se incluyen las sonrisas producidas de forma involuntaria durante las sesiones de grabación.

Previamente a la generación de los valores para los controladores del *rig* se han visualizado los valores de los puntos de la malla de Face Tracker en el mismo entorno gráfico del *rig* y de los personajes virtuales. Como se aprecia en la figura 20. Para ello se ha creado una malla con curvas NURBS similar a la que aparece en la aplicación FaceTracker con 17 puntos de control

correspondientes a la curva que delimita el óvalo de la cara, dos curvas de 5 puntos de control para las dos cejas, dos curvas cerradas de 6 puntos de control para cada ojo, una curva de 4 y otra de 5 para la nariz, una cerrada de 12 para la boca y una última también cerrada con 6 puntos de control para el interior de la boca.

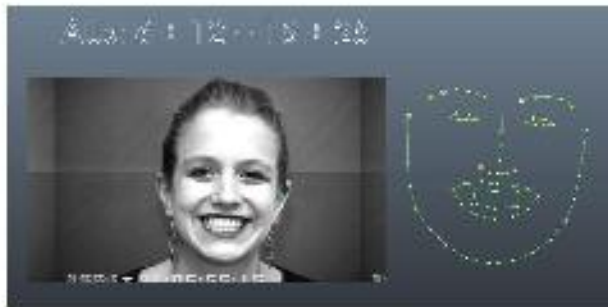


Figura 20: Importación de los puntos de Face Tracker aplicados a una secuencia de la base de datos Cohn-Kanade en el entorno de Autodesk Maya

La animación de estas curvas NURBS se ha creado asignando a cada punto de control los datos extraídos de la aplicación FaceTracker. Para ello ha sido necesaria una conversión al formato de fichero de animación propio del entorno 3D con el que estamos trabajando. En nuestro caso el entorno es el de la aplicación Autodesk Maya y el formato de los ficheros es el .anim cuya documentación se encuentra en su manual de usuario[100]. Igual que se han importado los datos de la aplicación de seguimiento también se importan las secuencias de vídeo dentro del mismo entorno. Este proceso permite detectar fácilmente posibles errores ocurridos durante la captura del movimiento e incluso subsanarlos manualmente editando las claves adecuadas. Este mismo control de errores servirá para comprobar la eficacia de la asignación de valores al *rig*, en la medida que el personaje virtual reproduzca los mismos gestos que se aprecian en vídeo y se descarten posibles errores producidos durante el *tracking*, y de forma que las curvas se correspondan con las imágenes del vídeo.

Para llevar a cabo la reproducción de las AUs por un personaje sintético es necesaria la activación de los correspondientes manipuladores del *rig*. Por ejemplo en la figura 20 donde se producen las AUs 6, 12, 15 y 25 tenemos que: para la AU 6 (Cheek Raiser) activar el Right & leftTwichControl, para la AU 12 (Lip Corner Puller) y la AU 15 (Lip Corner Depressor) activar el smileFrownCotrol y para la AU 25 (Lips part) activar jawControl. Según la interfaz definida en la figura 9.

Para calcular los valores de la animación de los controladores del *rig*

proponemos un método basado en reglas temporales similar al de Pantic y Patras[79] para la extracción de forma automática las acciones faciales en secuencias de imágenes. Nuestra aportación difiere de la anterior en que nuestro interés no está en reconocer que acción facial se ha producido, sino que a partir del movimiento detectado convertiremos las distancias entre puntos en valores adecuados para los controladores de *rig*. Para ello consideramos la tabla 5 en dónde se relacionan los parámetros faciales *FP* y los puntos del *tracker* que los definen con el controlador del *rig* que simula la acción facial.

Las transformaciones que se aplicarán a los manipuladores dependen de un factor de escalado ε de los datos de la secuencia de vídeo. Estos datos varían dependiendo de la resolución de la imagen y del tipo de plano. Por este motivo se toma la distancia entre los ojos *FP1* para aproximar este factor de escala. De todas maneras se ha de tener en cuenta de que si el personaje es de tipo primitivo o *cartoon* pueden interesar movimientos más exagerados, por lo que este factor de escala podrá ser ajustado manualmente según los intereses del animador.

FP1	Distancia entre los ojos	39 - 42	Factor de escala
FP2	Media apertura parpado inferior (derecha)	46 - 47	rightTwich
FP3	Distancia interna entre ceja y párpado (derecha)	22 - 27	upsetSad, shocked, rightBrownRaise
FP4	Media apertura parpado inferior (izquierda)	40 - 41	leftTwich
FP5	Distancia interna entre ceja y párpado (izquierda)	21 - 27	upsetSad, shocked, leftBrownRaise
FP6	Distancia apertura ojo (derecha)	22 - 42	leftBlink
FP7	Distancia apertura ojo (izquierda)	21 - 39	rightBlink
FP8	Ancho de la boca	48 - 54	smileFrown
FP9	Distancia comisura boca nariz (derecha)	33 - 54	rightSmileFrown
FP10	Distancia comisura boca nariz (izquierda)	33 - 48	leftSmileFrown
FP11	Altura de la boca	51 - 57	jaw
FP12	Vertical entre boca y nariz	51 - 33	sneer
FP13	Distancia comisura boca aleta nariz (derecha)	35 - 54	rightSneer
FP14	Distancia comisura boca aleta nariz (izquierda)	31 - 48	leftSneer
FP15	Vertical entre barbilla y nariz	8 - 33	jaw
FP16	Distancia entre ovalo cara	0 - 16	neck

Tabla 5: Relación de los parámetros faciales extraídos de los puntos de *tracker* con los manipuladores del *rig*

Los parámetros faciales del *FP2* al *FP7* son de los que se deduce la animación de la parte de cejas y ojos, esto incluye a los controladores *shockedControl*, *upsetSadControl*, *brownRaiseControl* (derecha e izquierda), *blinkControl* (derecha e izquierda) y *twichControl* (derecha e izquierda). Mientras que los que van del *FP8* a *FP15* son los correspondientes a la animación de la parte de la boca y las mejillas controlada desde la interfaz del *rig* por el

SmileFrownControl (derecha e izquierda), el sneerControl (derecha e izquierda) y el jawControl. Finalmente el $FP16$ es usado por el neckControl para representar torsiones de la cabeza. Excepto $FP2$ y $FP4$ los demás parámetros faciales se definen como las distancias entre los dos puntos que aparecen detallados en la tercera columna de la tabla 5. $FP2$ y $FP4$ se calculan como una media de distancias y se explican en 5.3.

Para asignar los valores que tomarán en el tiempo los distintos controladores consideraremos un cuadro de la secuencia en donde aparezca la expresión facial neutra, que en el caso de CK+ se corresponde a los primeros cuadros de la secuencia; y a partir de los parámetros faciales sacados del FaceTracker en cada cuadro tendremos la información del movimiento producido. De manera que para un determinado cuadro i el valor del controlador C vendrá dado por la expresión:

$$C_i = (FPx_i - FPx_0) * \varepsilon \quad (1)$$

Donde el subíndice 0 se refiere al cuadro que representa la expresión neutra, ε es el factor de ajuste a la escala del manipulador y FPx es el parámetro facial específico de cada controlador introducido en la tabla 5. El valor de ε depende en último término del animador y es el responsable de lo marcada o suave que pueda ser la expresión. De forma empírica podemos afirmar que valores correctos visualmente de ε pueden oscilar entre 0.1 y 0.3 para una imagen de resolución 640x490 correspondiente a un primer plano frontal del sujeto.

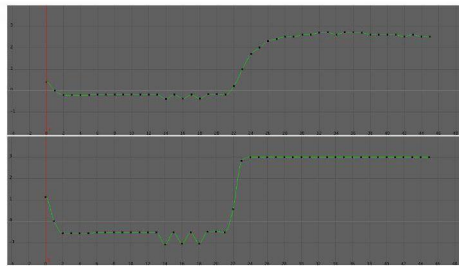


Figura 21: Curva de animación del controlador *SmileFrown* para la misma secuencia con los valores de ε mínimo y máximo

En la figura 21 observamos la curva de animación para el controlador *SmileFrown* correspondientes a la secuencia de vídeo de la base de datos Cohn-Kanade etiquetada con el nombre s055_005 que reproduce las AUs 6, 12, 15 y 25. En las abscisas se representan el número de cuadro de la secuencia y en las ordenadas la variación del manipulador. La curva superior de la imagen se corresponde a un valor de ε 0.1 mientras que la inferior corresponde a un valor de 0.3.

El controlador *SmileFrown* toma valores de translación en el eje Y que varían entre -3 y +3, siendo 0 el valor para la expresión neutra. Este rango de valores es fijado durante la creación del *rig* y es usado para definir la interacción del animador con este controlador. Mediante el factor de escala ε los valores recogidos por la aplicación de seguimiento serán convertidos a este rango.

En las dos curvas de la imagen 21 vemos como la expresión se activa en el cuadro 21. En la curva superior la activación es suave y nunca llega al valor máximo en la inferior con solo tres cuadros ya se llega al cenit de esta expresión. De forma general podemos afirmar que dependiendo de la estética del personaje animaciones más o menos acentuadas pueden ser aceptables visualmente, así a personajes primitivos o con estética *cartoon* les corresponderían valores de ε mayores mientras que para personajes más realistas o naturalistas sería lo contrario.

En la figura 22 se observa una curva de animación de *SmileFrown* generada con un valor ε de 0.28 que visualmente es válido para el personaje naturalista y para el primitivo.

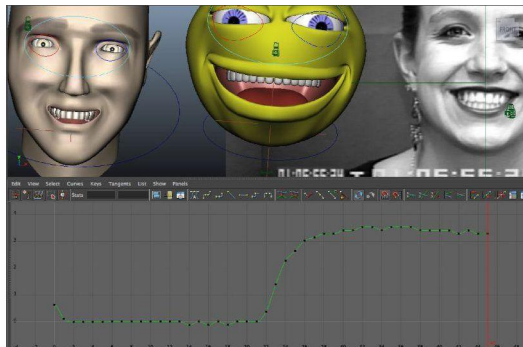


Figura 22: Curva de animación para el controlador *SmileFrown* con el mismo factor de escala para dos personajes diferentes

Dependiendo del rango de valores que tome el controlador y de las propiedades antropométricas de la expresión a simular la fórmula (1) será ajustada con valores múltiplo de ε .

Atendiendo a las particularidades de cada controlador tenemos:

5.2 El controlador *SmileFrown*

Como se ha dicho anteriormente este controlador toma valores en el eje Y, concretamente entre 3 y -3, valores positivos simulan la acción del músculo cigomático mayor cuando tensa y eleva las comisuras de los labios y valores negativos que simulan la acción del músculo triangular cuando tira de las comisuras hacia abajo. En base al *FP8* (ancho de la boca) calcularemos el valor de translación del controlador. Se activarán los controladores derecha / izquierda en el caso de que se aprecie diferencia en los valores *FP9* y *FP10* (distancias entre las comisuras de la boca y la nariz).

En la figura 23 se observa la activación del *smileFrownControl* en el cuadro 25 de la secuencia etiquetada con *s055_005* de la base de datos CK+, en la parte superior se ve el efecto sobre el modelo primitivo y el naturalista y en la parte inferior la curva simplificada de la animación del controlador en el tiempo y la imagen de referencia de donde se han extraído los datos con el FaceTracker. En el gráfico se observa como el controlador empieza a activarse a partir del cuadro 21 cuando se produce la sonrisa.



Figura 23: Activación del *smileFrownControl*

Por razones antropométricas, en la expresión (1) será necesario un escalado 2ε para los valores negativos responsables de generar la AU15, tales que $FP8_i < FP8_0$ de manera que tenemos:

$$SmileFrown_i = (FP8_i - FP8_0) * 2\varepsilon \quad (2)$$

que expresa el valor del controlador en el cuadro i en función del parámetro facial referido en la tabla 5.



Figura 24: Ajuste del factor de escala para la reproducción de la AU15 con el controlador *SmileFrown*

En la figura 24 se representa visualmente el ajuste del factor de escala introducido en (2). Efectivamente, como se ve en la imagen al cuadro 12 le corresponde el punto álgido de la reproducción de la AU15, en este cuadro el controlador tiene asignado el valor -1.5 que produce la imagen.

5.3 El controlador *Twich*

Este controlador toma valores positivos en el Y, concretamente entre los valores 0 y 4 definidos en el *rig*. Para el cálculo de los valores que tomará el controlador se consideran el *FP2* para el controlador del ojo derecho y el *FP4* para su equivalente ojo izquierdo. Para *FP2* consideramos los puntos de *tracker* 42, 45, 46 y 47 que se corresponden a la línea inferior del párpado. El parámetro facial se calcula como la media de las distancias de los puntos 46 y 47 a la recta que forman los puntos 42 y 45. Ver figura 25. De forma similar se calcula *FP4* para el ojo izquierdo.

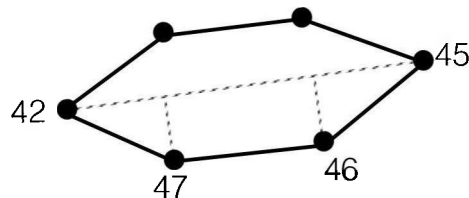


Figura 25: Cálculo del *twitchControl* para el ojo derecho

Dependiendo del tipo de imagen de entrada la aplicación de seguimiento facial puede producir un leve parpadeo en los puntos del contorno de los ojos. Este parpadeo que no se produce en la imagen real puede confundirse con el movimiento simulado por este controlador por lo que será necesario un proceso de filtrado de los resultados en el que se descartan cambios de sentido abrir / cerrar párpado inferior en entornos de tres cuadros. Nótese aquí que el parpadeo real que pueda realizar el sujeto va a reproducirse con el controlador *Blink* y no con el que nos ocupa. Como resultado se tiene una mitigación del problema de claves que representan temblores.

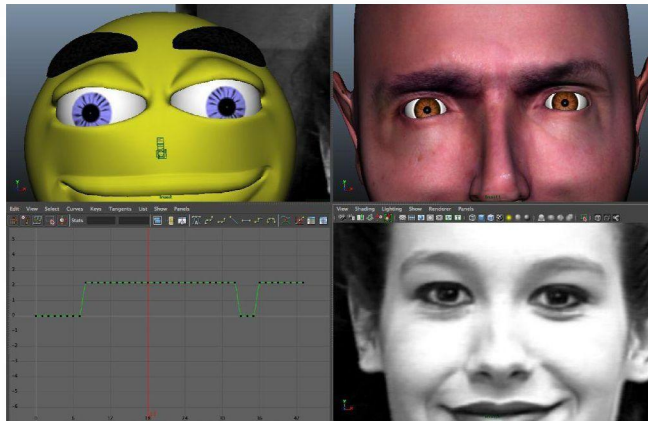


Figura 26: Activación del twitchControl

En la figura 26 se observa la activación del *twitchControl* en el cuadro 18 de la secuencia etiquetada con s074_005 de la base de datos de CK+, en la parte superior se ve el efecto sobre el modelo primitivo y naturalista y en la parte inferior la curva filtrada de la animación del controlador y la imagen de referencia de donde se han extraído los datos con el FaceTracker. En el gráfico se observa la curva suavizada con el proceso de filtrado que se activa durante la sonrisa. En la curva se observa la activación del controlador cuando empieza la sonrisa, se puede apreciar en la curva un valle entre los cuadros 33 y 35 correspondiente a un error de filtrado que produce una pérdida de presión en el párpado inferior. Estos errores podrán ser corregidos manualmente.

Para este controlador la expresión (1) de conversión de valores quedará:

$$rightTwich_i = -(FP2_i - FP2_0) * 3 \varepsilon \quad (3)$$

Donde FP2 es el parámetro facial involucrado (ver tabla 5) y 3ε es el factor de escala ajustado a valores correctos visualmente de forma empírica.

5.4 El controlador *Jaw*

Este controlador toma valores positivos y negativos en el plano XY, en el rango 5 y -5 en X y 2 y -5 en Y según la definición del *rig*. Los valores positivos en Y simulan la AU 24 que representa la acción del músculo orbicular de la boca provocando la presión de los labios y los negativos las AUs 25, 26 y 27. En este caso los parámetros faciales considerados son *FP11* (altura de la boca) y *FP15* (vertical entre la barbilla y la nariz) que tienen valores similares. Los valores en X se introducen para simular las asimetrías que se pueden generar con la mandíbula, notemos aquí que en las diferentes expresiones de sonrisa que se analizan en este estudio no aparecen asimetrías de este tipo.

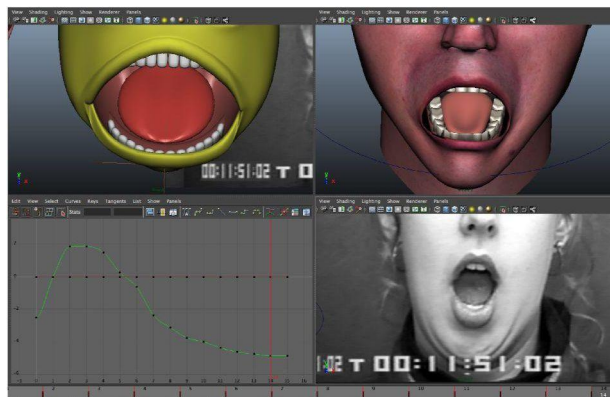


Figura 27: Activación del *jawControl*

Este controlador es el único que no está asociado a deformaciones de interpolación de geometrías sino que como se ha explicado en el punto 3.1 activa la deformación por hueso de la mandíbula. Nótese aquí que esta consideración no afecta en modo alguno al cálculo del valor del controlador.

En la figura 27 se observa la activación del *jawControl* en el cuadro 14 de la secuencia de la base de datos CK+ etiquetada con el nombre *s52_001* en la que aparecen entre otras las AU 25 y 27.

Para este controlador la expresión que nos da su valor en un cuadro *i* será:

$$Jaw_i = -(FP11_i - FP11_0) * 0.5 \varepsilon \quad (4)$$

Donde *FP11* es el parámetro facial involucrado (ver tabla 5) y 0.5ε es el factor de escala ajustado a valores correctos visualmente de forma empírica.

5.5 El controlador *Neck*

Si bien en la base de datos Cohn – Kanade todas las secuencias se corresponden a primeros planos frontales de diferentes individuos, se observa que un número significativo de ellos presentan ligeras rotaciones de la cabeza, que se simulan activando la rotación en el eje Z del *neckControl*, consideramos para ello el *FP16* (distancia entre el punto inicial y el final de la curva del óvalo de la cara) y extraemos el ángulo de rotación que se aplicará al controlador por trigonometría simple.

5.6 El controlador *Blink*

El *blinkControl* toma valores negativos en el eje Y, y simula la acción del músculo orbicular de los ojos cuando los cierra. Estos valores oscilan entre 0 (neutro) y -5 (completamente cerrado). En términos de AUs se refiere a las que van de la número 41 a la número 46. Para el cálculo de los controladores del ojo derecho e izquierdo se consideraran los parámetros *FP6* y *FP7* (distancia de apertura del ojo). Los errores en el seguimiento pueden hacer aparecer un efecto de temblor similar al que se ha descrito en el caso del controlador *Twitch* por lo que también es preciso un proceso de filtrado de estos errores. En este caso el filtro será diferente ya que no se pueden eliminar los parpadeos reales que aparezcan. Oscilaciones menores que un coeficiente determinado son despreciadas. Este coeficiente se calcula en base al factor de escala. La expresión para el cálculo del valor de este controlador quedará:

$$rightBlink_i = -(FP6_i - FP6_0) * 8 \varepsilon \quad (5)$$

Donde *FP6* es el parámetro facial involucrado para la parte derecha (ver tabla 5) y 8ε es el factor de escala ajustado a valores correctos visualmente de forma empírica.

En la figura 28 se observa la activación del *blinkControl* en el cuadro 1 de la secuencia etiquetada con s055_005 de la base de datos de CK+, en la parte superior se ve el efecto sobre el modelo primitivo y el naturalista y en la parte inferior la curva filtrada de la animación del controlador y la imagen de referencia. En el gráfico se observa la curva filtrada al inicio de la secuencia se observa la zona del parpadeo, mientras que durante la toma los ojos aparecen abiertos con valores cerca de cero, los valores positivos no se han

filtrado ya que el controlador no los contempla.

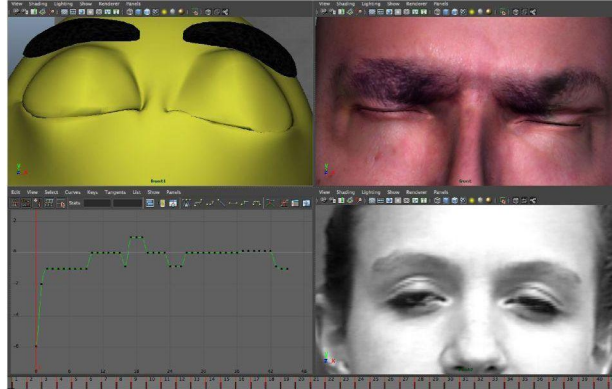


Figura 28: Activación del *blinkControl*

5.7 Los controladores *UpsetSad*, *BrowRaise* y *Shocked*

Estos controladores definen todo el movimiento que afecta a las cejas y parte superior de la cara. El *upsetSadControl* toma valores positivos y negativos en el eje Y, entre los valores 3 y -3. Los positivos colocan las cejas en la posición que se le atribuye a la expresión de la tristeza que se corresponde con las AUs 1 y 2 debidas a la acción del músculo frontal, tanto a su parte media como lateral. Los valores negativos son los responsables de hacer bajar las cejas en lo que habitualmente se asocia a la expresión de ira, se corresponde a la AU 4 debida a la actividad de los músculos corrugador y depresor superciliares. Para realizar el cálculo de este controlador se consideraran los parámetros *FP3* y *FP5* para la ceja derecha e izquierda respectivamente. En el caso de tener valores positivos simétricos se representa la AU1 y en el caso de que exista alguna asimetría se representa la AU2, si $FP3 > FP5$ se activa el controlador de la derecha y en caso contrario se activa el de la izquierda. Para las expresiones con asimetrías se usarán los controladores *rightBrowRaise* y *leftBrowRaise* que toman valores positivos en Y entre 0 y 3. Activar un controlador de este tipo implica la elevación de la ceja en cuestión, al tiempo que se abre más el ojo correspondiente mientras que la otra ceja baja y su ojo se cierra levemente.

Los parámetros *FP3* y *FP5* permiten hacer el seguimiento del movimiento de cejas para las expresiones anteriores pero también para la expresión de

sorpresa. Para la configuración de esta expresión se activa también la AU5 correspondiente a la acción del músculo elevador del párpado superior. El *shockedControl* puede ser activado por tanto con la misma curva de animación que el *upsetSadControl* quedando esta decisión en manos del animador así como los ajustes finos indicados con anterioridad. El *shockedControl* solo toma valores positivos entre el rango 3 (sorpresa máxima) y 0 (neutro). En la figura 29 se observa la activación del *shockedControl* con los datos extraídos a partir de los puntos mencionados de la secuencia s111_001 de la base de datos CK+.



Figura 29: Activación del shockedControl

Para estos controladores tenemos la expresión:

$$UpsetSad_i = -(FP3_i - FP3_0) * 4 \epsilon \quad (6)$$

Donde $FP3$ es el parámetro facial involucrado en expresiones sin asimetrías o con la expresión dominante en la parte derecha (ver tabla 5) y 4ϵ es el factor de escala ajustado a valores correctos visualmente de forma empírica.

5.8 El controlador *Sneer* y el *M_F_U_MouthRolls*

El *sneerControl* toma valores positivos en el eje Y para representar las AUs 10 y 11, y se controla con el parámetro $FP12$, con los parámetros auxiliares $FP13$ y $FP14$ se detectan posibles asimetrías activándose en su lugar los controladores de la derecha o la izquierda dependiendo de que se verifique que $FP13 > FP14$.

El *M_F_U_MouthRolls* toma valores positivos y negativos en el eje X para

representar variaciones de la AU26 en lo que es la acción del músculo masetero; toma valores positivos en el eje Y para simular la AU10 al tiempo que la AU16; valores de escalado en el eje Y simulan la AU25 o la AU28 mientras que en el eje X simulan la AU23. Los parámetros faciales involucrados en estos movimientos son los del contorno de la boca es decir de *FP8* a *FP14*.

Según las descripciones de las diferentes expresiones de risa que van a captarse de forma automática las acciones de estos controladores son mínimas y puntuales y en donde aparezcan servirán para matizar la expresión. El método que se presenta no pretende ser un método genérico para cualquier expresión, en este sentido no se considera ningún sistema para captar la forma de los labios, cuando ejercen presión por ejemplo, o situaciones en que se enseñe la dentadura.

5.9 Otros controladores

El estudio de la acción de los controladores correspondientes al punto de vista de los ojos del personaje así como el de los movimientos de la lengua, que si bien son comunes a la mayoría de *rigs* faciales no intervienen para las expresiones de risa que nos interesan. Pueden ser incorporados al sistema en futuras ampliaciones.

5.10 Uso de los valores de los controladores como objetivo de otras configuraciones de rig facial

Las mismas curvas de animación que se han generado para los controladores de nuestro *rig* pueden ser importadas a otras configuraciones de rig con resultados visualmente correctos. Por ejemplo en el caso del *rig* hecho para el proyecto Emily[62], de *FaceWare Technologies* que puede descargarse de forma gratuita desde su página web.

En la figura 30 se observa la asignación de curvas obtenidas a partir de la secuencia *s124_004* a un *rig* diferente de la aproximación de Digital Tutors que es el que hemos usado para construir nuestras curvas. En la imagen se observa la selección del controlador llamado *con browRaise 1*, en el entorno del rig de *FaceWare*, al cual se le ha asignado nuestra curva generada para

nuestro *upsetSadControl*, para generar la expresión que se observa en la imagen se ha asignado esta misma curva también a los controladores que elevan la otra ceja y la parte central de ambas, la curva del *jawControl* se ha asignado a su homóloga, mientras que la curva de nuestro *smileFrownControl* se ha asignado a sus depresores de las esquinas de los labios. El *rig* de Emily difiere radicalmente del que hemos usado, en la manera con que se controlan los párpados ya que en vez de usar la translación en el eje Y se ha optado por la creación de un atributo adhoc *OpenClose* que controla el movimiento, para este caso en concreto se ha tenido que modificar el fichero con los datos de la animación intercambiando el nombre del atributo animado para que coincidiese con el nombre *OpenClose* lo que permite la importación de las claves de animación sin problemas. Es interesante hacer notar como controles que el *rig* de Digital Tutors no contempla, como es el que permite elevar la barbilla se puede usar asignándole el movimiento de la mandíbula.

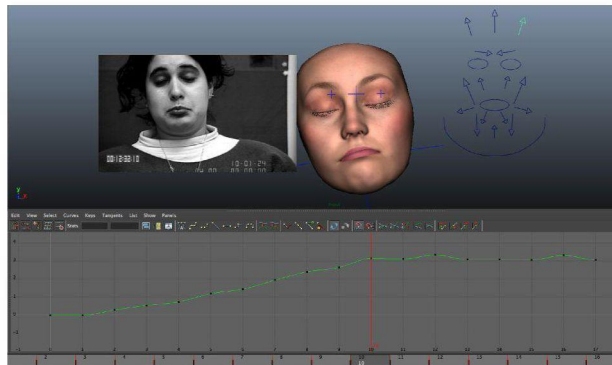


Figura 30: Asignación de curvas de animación a un *rig* diferente

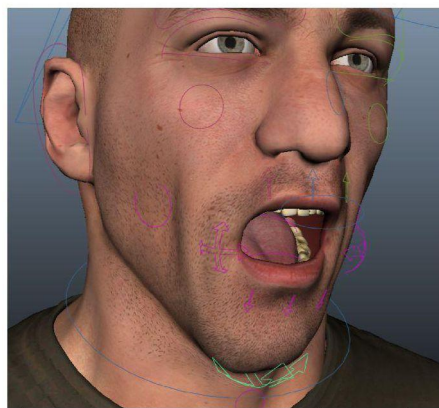


Figura 31: *Rig* de *FaceWare Technologies* en el que el manipulador de apertura de la boca se activa con valores de rotación en grados de un ángulo

Otros rigs de acceso libre consultados como los de *Creative Crash*[101] u otros de *FaceWare*[102] permiten de forma similar la asignación de los valores calculados como se ha expuesto. Si bien es preciso notar detalles como el controlador que permite la apertura de la boca en algunos rigs como el de FaceWare que ilustra la imagen son activados con valores de rotación de un ángulo, por lo que en este caso sería necesario la adecuación de nuestros valores mediante operaciones de trigonometría simple.

Capítulo 6 Aplicación del método

En este capítulo se prueba el funcionamiento del método generación de expresiones de risa presentado en el capítulo anterior. Para ello se generan expresiones a partir de videos correspondientes a secuencias cinematográficas representativas y a bases de datos públicas con expresiones de risa. Finalmente en un apartado de este capítulo elaboramos nuestra propia base de datos de expresiones que usaremos para la validación del método y para otras aportaciones de este estudio.

6.1 Generación de expresiones de risa a partir de secuencias de vídeo

Para la evaluación del funcionamiento del *rig* animado de forma automática aplicado al estudio de las expresiones que involucran risas y sonrisas se analizan diferentes tipos de secuencias de vídeo, en primer lugar secuencias cinematográficas procedentes de las películas a las que nos referíamos en el punto 3.2. y después se prueba sobre diferentes bases de datos de expresiones faciales.

6.1.1 Aplicación a las expresiones de risa procedentes de secuencias cinematográficas

Entre las más votadas se han buscado secuencias en las que se visualicen las expresiones de interés. Como característica común a todas las que se han analizado cabe destacar las siguientes **consideraciones**:

- A diferencia de las secuencias que forman la base de datos de Cohn-Kanade en donde todas las grabaciones se habían hecho manteniendo un primer plano similar, en las secuencias de las películas elegidas, como es natural, no hay ninguna **unificación de tomas** de los planos, se encuentran desde primeros planos a planos medios. Esta diferencia de escala se traduce en una variación proporcional de los valores resultantes que deberán ser

importados a los controladores del *rig*. Este problema se corrige usando el factor de escala introducido en el punto anterior o bien escalando los valores de la curva de animación resultante en el API 3D.

- Igual que los planos no están unificados tampoco lo está la posición del sujeto, en la muchas secuencias el sujeto no está **alineado frontalmente** con el plano de la cámara. La aplicación de seguimiento facial funciona correctamente incluso en posiciones de perfil de manera que la problemática se localiza en la interpretación de los datos de dicha aplicación. Como se ha explicado en el capítulo anterior, según la interfaz del *rig* se tienen entradas separadas para la parte izquierda y derecha de la cara, con el método que hemos expuesto es factible la asignación de curvas entre controladores, de manera que problemas debido a la oclusión de partes de la cara queda mitigado. Para la parte inferior de la cara el análisis de los puntos de las comisuras y de la parte central superior e inferior de los labios no produce errores apreciables mientras no haya oclusión de estos puntos.
- La falta de **iluminación normalizada** también es común en la serie de secuencias de estudio. Por tanto se tendrá que prescindir de las secuencias en que la iluminación sea demasiado débil de manera que haga que se produzcan errores en el seguimiento de las características faciales.
- Todas las secuencias de la base de datos Cohn-Kanade parten de una **expresión neutra** y al final de la secuencia se encuentra el punto máximo de la expresión realizada, cosa que evidentemente no ocurre en las secuencias cinematográficas. Añadiremos por tanto a nuestra aplicación la posibilidad de insertar el número de cuadro en el que la expresión sea la más parecida a la neutra despreciando aquellas secuencias en las que no podamos encontrar ninguna.
- En las secuencias de estudio no encontramos las expresiones completamente separadas del **diálogo** sino que en muchas ocasiones se producen de forma simultánea, esto puede dificultar la percepción de detalles de la expresión. Por lo que se elegirán o se cortarán las escenas convenientemente para eliminar diálogos.
- La problemática que se deduce de no tener las **secuencias etiquetadas** con la expresión realizada o con las AUs llevadas a término queda paliada por la interpretación del contexto de la escena.
- **Errores poco apreciables en el tracker** pueden dar lugar a extracción de valores incorrectos para la animación. Por ejemplo en la figura 32 se muestra un cuadro correspondiente a la película *Ninotchka*, en ella un personaje esboza una sonrisa con la boca cerrada partiendo de una expresión seria. Se ha asignado como cuadro neutro el que aparece en la

imagen (cuadro 1), en la imagen se ve como el *tracker* se ha ajustado correctamente a la cara pero si se observa con detalle se ve que para el cálculo del FP8 (anchura de la boca) toma como referencia la línea que define el bigote del personaje. Durante la realización de la expresión esta distancia apenas cambia por lo que el controlador no se activa.



Figura 32: Secuencia con extracción de datos errónea

- Durante la validación del método se han encontrado algunos **problemas inherentes al avatar** naturalista. En primer lugar se ha notado una falta de definición geométrica en la zona situada sobre las mejillas y en los párpados inferiores, esto provoca que en expresiones donde los controladores *blink* o *twitch* toman sus valores máximos, el efecto resultante simula una presión muscular mayor que la que ocurre realmente. En segundo lugar se ha detectado un problema de textura aplicada al modelo geométrico: en la fotografía original tomada como referencia el sujeto tenía el ceño fruncido mientras representaba la expresión neutra, de manera que cuando se activa el controlador *upsetSad* para valores positivos, es decir para la simulación de las cejas levantadas, el efecto visual queda mitigado al igual que queda ampliado en el caso inverso.

Para la selección de las escenas se han tenido en cuenta la descripción de las situaciones en que se producen, la morfología facial que hace Faigin y que además verifiquen que están tomadas en un primer plano o plano medio similar al frontal, sin diálogo, sin sombras u objetos que ocluyan parte de la cara, que en un cuadro determinado contengan la expresión neutra y que el *tracker* no dé errores. Todas estas restricciones hacen que la tarea de búsqueda de secuencias representativas de las 13 expresiones de risa sean una tarea ardua.

En http://dmi.uib.es/~miquel/webdedocencia/Video_Results_3_3.html se encuentran tres ejemplos de generación automática de expresiones a partir de

secuencias cinematográficas.

En el primer ejemplo se reproduce la sonrisa melancólica a partir de una secuencia de la película "*Ninotchka*", se trata de un primer plano de la protagonista en donde aparece su rostro bien iluminado y ligeramente ladeado a la derecha. El seguimiento de las características faciales no registra ningún error. Se han asignados los valores de animación generados al *rig* del modelo naturalista y del primitivo. Finalmente el vídeo de resultado muestra la secuencia renderizada con una iluminación similar a la de la secuencia cinematográfica, para esta última secuencia se han suavizado los valores de las curvas de animación con las utilidades que ofrece el API 3D.

En el segundo ejemplo se reproduce la sonrisa con la boca cerrada según un plano perteneciente a la película "*What Ever Happened to Baby Jane?*" en el que aparece un plano medio de una de las actrices protagonistas. El plano empieza con la sonrisa definida, por lo que se tendrá que buscar un cuadro donde aparezca la expresión neutra, además en esta secuencia la protagonista habla y en un momento dado se pone un cigarrillo en los labios con la consecuente oclusión de parte de la cara. Inicialmente el FaceTracker presenta problemas con esta secuencia ya que no es capaz de inicializar los valores de la malla de puntos coincidiendo con la cara, este problema se ha corregido escalando la imagen, convirtiendo el plano medio, en un primer plano de la cara. Los problemas derivados del habla y de la oclusión son apreciables. En el vídeo resultante se han realizado las mismas secuencias explicadas en el ejemplo anterior.

El tercer ejemplo reproduce la sonrisa astuta con los datos extraídos de una secuencia de la película "*Kiss of death*" en donde aparece un plano medio del antagonista de la película, aunque su cara aparece ladeada y hay sombras debido a la iluminación y a que lleva un sombrero no se producen errores en el seguimiento de características faciales. En el vídeo resultado vemos la animación aplicada a los dos modelos y finalmente la secuencia renderizada con el modelo naturalista.

En la figura 33 se observan diferentes imágenes correspondientes a los ejemplos, en la parte de la izquierda se aprecian cuadros correspondientes a la asignación de valores al *rig* del modelo naturalista, mientras que en la parte de la derecha se ven los mismos cuadros aplicados al modelo primitivo. En ellas se observa la similitud en las expresiones reales y las generadas por nuestros modelos a los cuales se les ha aplicado las curvas de animación resultantes del método de extracción de valores. En las imágenes 1 y 3 se observan los problemas comentados respecto al la textura del modelo naturalista y en la 1 además se aprecia el problema de la presión al cerrar los ojos. De forma general se han hecho una serie de ajustes manuales para unificar cuestiones de escala debido a los diferentes tipos de planos de las secuencias y suavizar las

curvas de animación simplificando el número de claves.

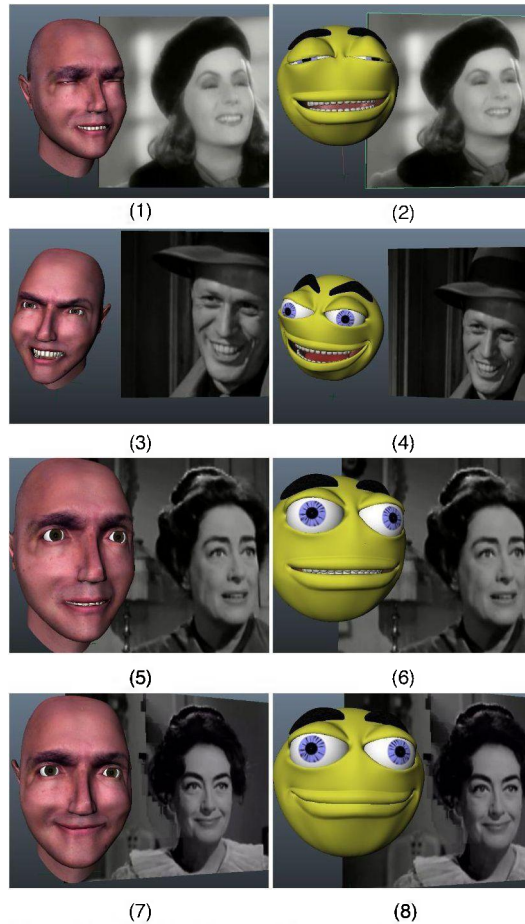


Figura 33: Ejemplos de asignación de valores a los controladores del *rig* a partir de datos de secuencias cinematográficas. (1) y (2) sonrisa melancólica, (3) y (4) sonrisa astuta, (5) y (6) sonrisa congraciante y (7) y (8) sonrisa boca cerrada

6.1.2 Aplicación a las expresiones de bases de datos públicas específicas de análisis de la risa

Con el objetivo de validar los resultados de síntesis de expresiones que involucran la acción del músculo cigomático mayor se ha acudido a bases de datos de acceso público que incluyan este tipo de expresiones.

Las bases de datos de expresiones faciales como la CK+, referenciada anteriormente, u otras como la MMI[103], aunque no sean precisamente

específicas para expresiones de sonrisa, incluyen algunas expresiones de este tipo, si bien las anotaciones correspondientes a ellas no diferencian tipos de risa, más allá de la sonrisa y la carcajada.

Recientemente han aparecido estudios donde se investiga sobre la detección y clasificación de tipos de risa. En concreto en la base de datos de la BBC Smile Dataset [104] se trata la diferencia entre la sonrisa espontánea y la voluntaria o posada, haciendo uso de una base de datos específica de sonrisas, que presenta los datos obtenidos a partir de un test de detección de sonrisas falsas en un conjunto de vídeos con 10 sonrisas espontáneas y 10 voluntarias, empezando y terminando con la expresión neutra; o como en Hoque et al. [105] donde a partir de varios experimentos se graba la reacción de 15 sujetos en diferentes situaciones frente a una pantalla de ordenador distinguiendo risas voluntarias e involuntarias correspondientes a estados de alegría y de frustración.

Para probar la efectividad del sistema de adquisición de datos para la animación facial de risas se ha accedido a las bases de datos públicas Uva-Nemo y MAHNOB.

La base de datos Uva-Nemo [106] se ha generado para estudiar la dinámica de la sonrisa espontánea respecto a la generada de forma voluntaria, esta formada por vídeos de alta resolución en un entorno con iluminación controlada y con la cámara situada a la misma distancia de los diferentes sujetos. Se trata de la mayor base de datos de este tipo que hay actualmente, contiene 1240 vídeos de sonrisas clasificadas en 597 espontáneas y 643 voluntarias, además de las clasificaciones por sexo y por edad que oscila entre los 8 y 76 años.

Otra base de datos específica para el estudio de la risa es la MAHNOB Laughter database [107], que ha sido usada para la diferenciación entre risa y locución y para la detección de la risa durante sesiones de discurso.

Esta base de datos contiene secuencias de vídeo, de audio y ficheros con sus anotaciones correspondientes; aparecen 22 sujetos que realizan risas espontáneas, risas voluntarias y locuciones narrando vivencias divertidas. Las secuencias han sido grabadas usando una cámara de vídeo estándar con su micrófono, un micrófono de corbata y una cámara termográfica. Las grabaciones se han realizado a una distancia constante y con un plano y una iluminación frontal o cercana a la frontal. De manera que la distorsión por sombras es mínima. Todas las secuencias se encuentran anotadas con la herramienta de anotación ELAN [108], usando etiquetas como risa, locución o risa-locución cuando aparecen conjuntamente.

De forma similar al procedimiento que hemos aplicado en el test de nuestro *rig* con la base de datos Cohn-Kanade hemos aplicado a algunas secuencias de Uva-Nemo y de MAHNOB el *FaceTracker* de Saragih y hemos exportado los

datos a ficheros .anim para su posterior asignación a los controladores del *rig*. Se han obtenido resultados satisfactorios en la reproducción de sonrisas simuladas y espontáneas tanto para nuestro personaje naturalista como para el primitivo.

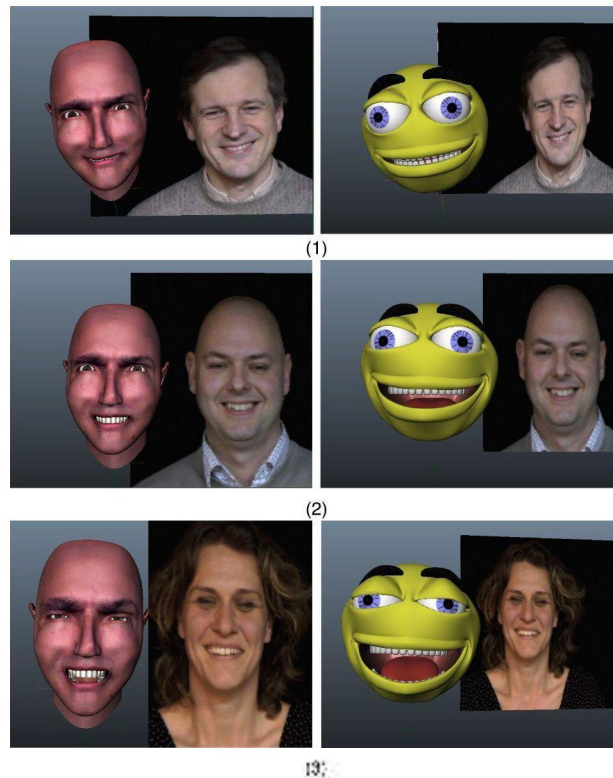
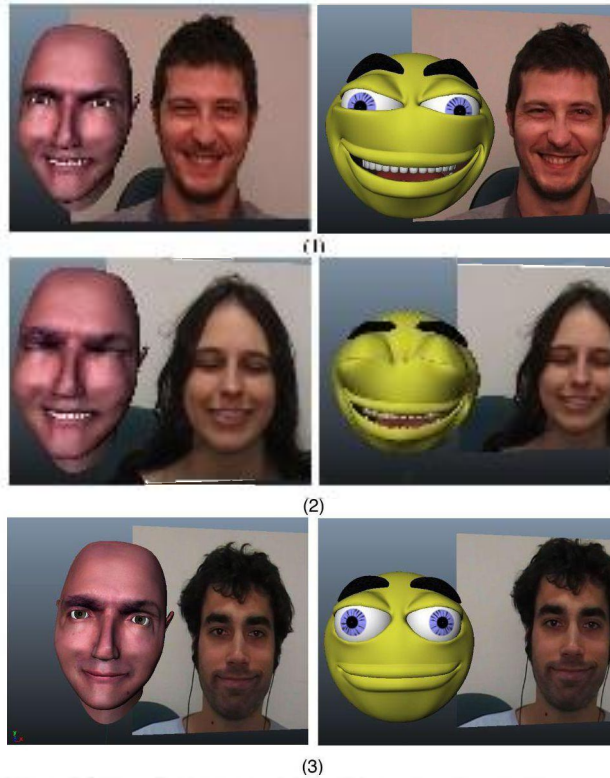


Figura 34: Ejemplos de la adquisición de la animación a partir de secuencias de la base de datos Uva-Nemo

En la figura 34 se observan diferentes cuadros correspondientes a secuencias (1) de risa espontánea y (2) y (3) de risa deliberada, asignadas al modelo naturalista y al primitivo, los dos pertenecen al mismo cuadro de la secuencia. Todas estas secuencias pertenecen a la base de datos Uva-Nemo y han sido elegidas al azar ya que al estar todas ellas grabadas en la misma situación no implican diferencia alguna que no sea proveniente del individuo y los tests realizados sobre el *tracker* ya garantizan esta independencia.

En la figura 35 se observan secuencias (1) y (3) anotadas como risa voluntaria y (2) como risa espontánea pertenecientes a la base de datos MAHNOB, igual que en la imagen anterior los dos personajes se encuentran en el mismo cuadro de la animación. En las 6 secuencias que se han reproducido se han repetido ajustes manuales del mismo nivel que en el caso de las secuencias

cinematográficas.



(3)
Figura 35: Ejemplos de la adquisición de la animación a partir de secuencias de la base de datos MAHNOB

Como se observa en las imágenes anteriores la reproducción de la expresión es satisfactoria y reconocible a simple vista, aún con las limitaciones y características de nuestros modelos.

En cuanto a la validación general para cualquier tipo de expresión de risa es notorio que estas bases de datos no tratan otros tipos de risa diferentes a las llamadas risas asociadas a expresiones de alegría. Tal y como se explica en el protocolo llevado a cabo a la hora de hacer la grabación, en los casos en que se pide a los sujetos que realicen sonrisas de forma voluntaria delante de la cámara no se les da ningún tipo de especificación respecto a la expresión que deben realizar para garantizar una mayor naturalidad de la expresión. De forma que respecto a la clasificación de Faigin que venimos usando encontramos ejemplos representativos de carcajada extrema, carcajada, sonrisa clamorosa, sonrisa boca cerrada, falsa sonrisa y falsa carcajada pero no de las demás.

6.2 Creación de una base de datos específica de los diferentes tipos de sonrisa

Con el objetivo de suplir la problemática expuesta en los puntos anteriores se ha desarrollado una base de datos específica que contemple exactamente los 13 tipos de risa descritos por Faigin.

La base de datos está formada por 91 vídeos en color RGB, con una resolución de 1280x720, con el códec Motion JPEG Video (MJPG), grabados con una cámara fotográfica Olympus EPI, en un entorno con la iluminación controlada con un foco y luz natural indirecta y con el objetivo a una distancia aproximada de 1,5 metros. Se ha usado un fondo de imagen del mismo color blanco en todas las tomas. Las tomas originales cuentan con el canal de audio estéreo grabado mediante el micrófono incorporado de la cámara con el códec PCM S16 LE (araw).

Para la realización de los vídeos se han buscado sujetos mediante una convocatoria para la participación voluntaria en el experimento. Se ha contado con la participación de 7 individuos, 6 de los cuales provenientes de la Escuela Superior de Arte Dramático de la Islas Baleares con experiencia en la interpretación y en el seguimiento de indicaciones para llevar a cabo expresiones faciales. Todos los sujetos han dado su consentimiento por escrito para el uso las imágenes grabadas con fines educativos y de investigación. Las grabaciones se han realizado en una habitación a la cual accedían de forma individual sin tener previamente conocimiento alguno de las interpretaciones que iban a realizar ni presión alguna de presencia de público. A cada uno de los sujetos, se les ha pedido que interpretasen las 13 expresiones de interés, para lo que se ha seguido un protocolo consistente, en primer lugar, en ofrecer la visualización gráfica de la expresión según la ilustración de Faigin y el fotograma cinematográfico correspondiente, después se les ha puesto en contexto describiendo situaciones donde se producen las expresiones como las requeridas partiendo siempre de la expresión neutra.

A posteriori las secuencias grabadas han sido editadas marcando los puntos de entrada y salida de forma que en la entrada se encontrase la expresión neutra y en la salida el punto más alto de la expresión realizada. Se ha eliminando el sonido y se han convertido a una resolución de 640x360 con el códec H24 - MPEG - A AVC.

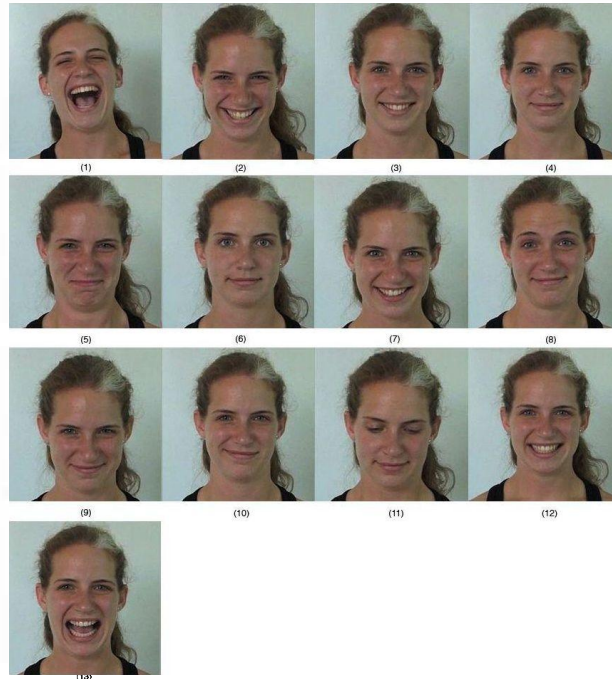


Figura 36: Reproducción de las diferentes expresiones de alegría: (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada

En la figura 36 se observan la reproducción de las 13 expresiones llevadas a cabo por la sujeto etiquetada como s07. A diferencia de la base de datos Uva-Nemo o MAHNOB en nuestro caso todas las expresiones son interpretadas, como se deduce del protocolo seguido. Solo en el caso de la sujeto s01 sin experiencia interpretativa se han usado expresiones reales grabadas fuera de las indicaciones. Con todos los sujetos en el caso concreto de la falsa sonrisa y la falsa carcajada se han dado indicaciones concretas referidas a la menor tensión existente en la parte inferior de los párpados y en la forma de la boca.

6.3 Generación de las expresiones faciales a partir de la base de datos propia

Se ha ejecutado el *Face Tracker* con los 91 vídeos de la base de datos

obteniéndose el mismo número de ficheros de animación de los 66 puntos del seguimiento facial. Por cada uno de estos ficheros la aplicación que implementa nuestro método ha creado un fichero con los valores de la animación de cada controlador del *rig* facial.

En las figuras 37, 38, 39 y 40 se observan la variación en el tiempo de los diferentes valores de algunos de los controladores del *rig* para cada una de las 13 expresiones interpretadas por el sujeto 07 que aparecen en la figura 36. En el eje de las abscisas se representan los cuadros del tiempo mientras que en las ordenadas aparecen la variación del valor del controlador según la magnitud definida por el *rig* facial. Los datos que se muestran aquí son los mismos que serán importados en el entorno de la aplicación 3D donde podrán afinarse en caso de que el animador lo crea conveniente. Después de su importación la curva de animación quedará formada con una clave de animación en cada cuadro, esto hace que estas curvas sean difíciles de usar, no tanto por la cantidad de información y peso de cálculo que se maneja sino más bien por la dificultad de edición. Este problema es habitual en los procesos de captura de movimiento, igual que la solución al uso, consistente en una simplificación de la curva con la correspondiente reducción del número de cuadros.

Entre las observaciones que se pueden hacer sobre la base de la visualización de los gráficos cabe destacar la curva que describe en las diferentes expresiones el controlador *smileFrown* que define la simulación del músculo cigomático mayor. Este controlador se mueve entre los valores -3 y 3 siendo los valores positivos los responsables de la AU6 Cheek raiser y la AU12 Lip corner puller. Esta representada por las curvas de color azul de la figura 38. En los gráficos se observa que para la expresión s07_09 sonrisa astuta y para la s07_11 sonrisa embarazosa aparecen valores negativos que se asociarían a las AU15 Lip corner depressor y AU17 Chin raiser y que serían acciones inapropiadas para la descripción de las dos expresiones a las que se refieren los datos. En ambos casos el error proviene de movimientos de rotación en eje X de la cabeza. Este tipo de errores justifica que las curvas sean ajustables posteriormente desde el entorno de la aplicación 3D.

En los gráficos también se observa la forma regular de las curvas correspondientes a los controladores del *twitch* y del *blink*, esto se ha conseguido con el proceso de filtrado mencionado anteriormente en el punto 5.3. En contraposición se observa que las curvas correspondientes al *upsetSad* presentan numerosos picos y que pueden generar temblores no deseados en la animación y que podrán ser eliminados con su correspondiente proceso de filtrado o bien a posteriori en el entorno 3D.

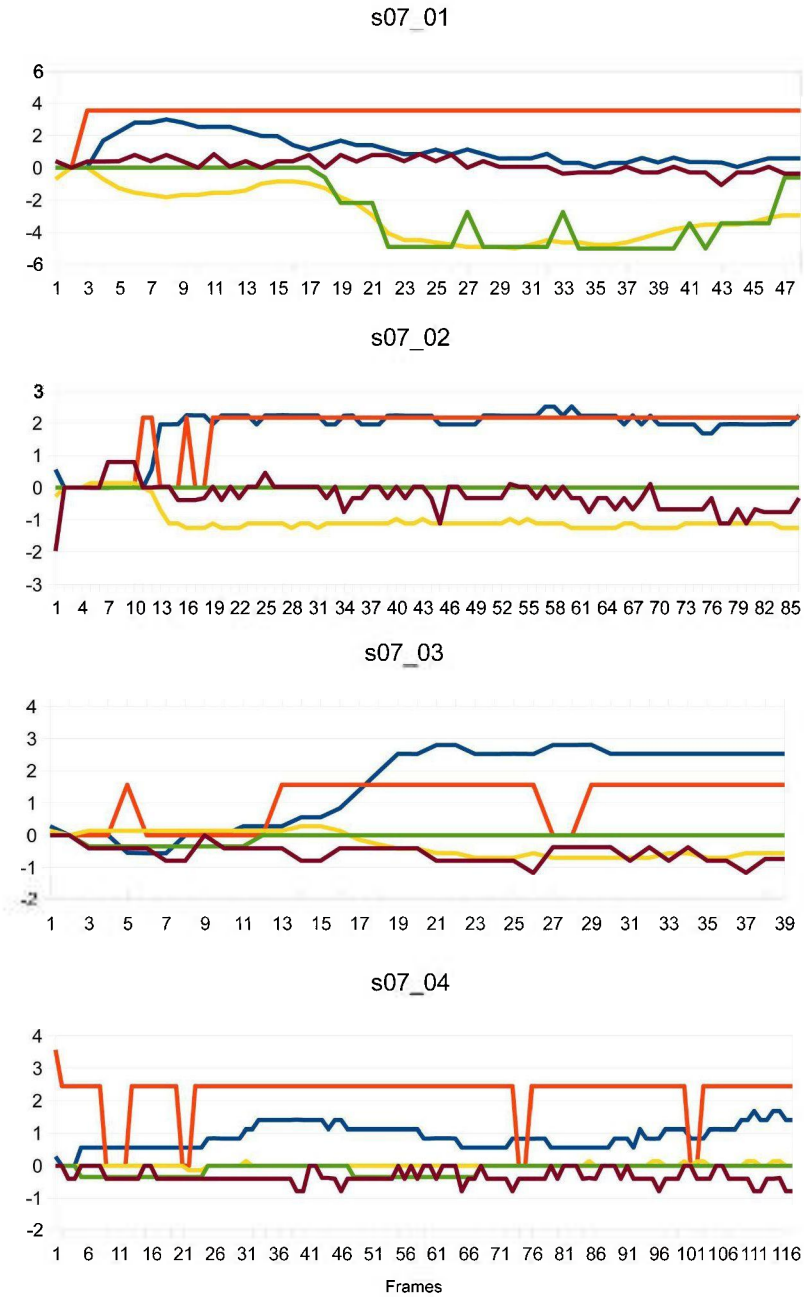


Figura 37: Valores de los principales controladores del *rig* para las diferentes expresiones de el sujeto s07

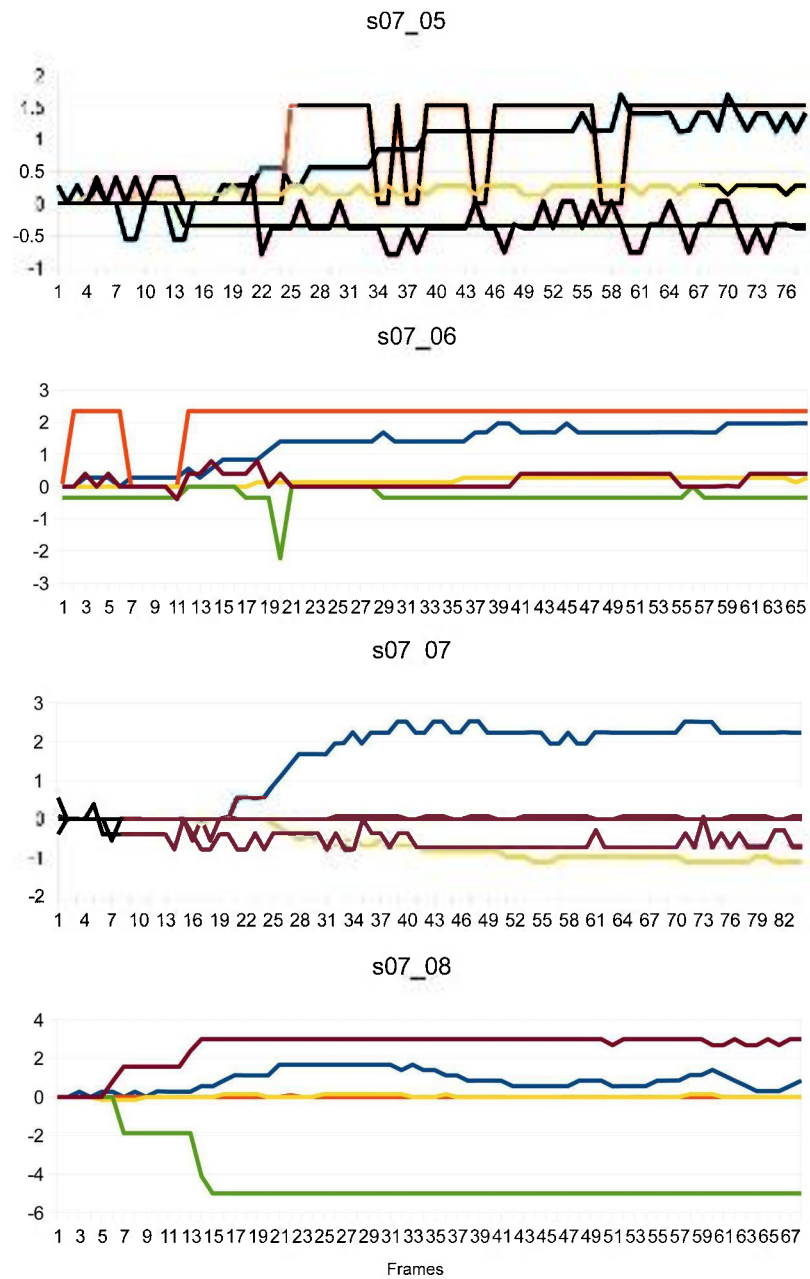


Figura 38: Valores de los principales controladores del *rig* para las diferentes expresiones de el sujeto s07

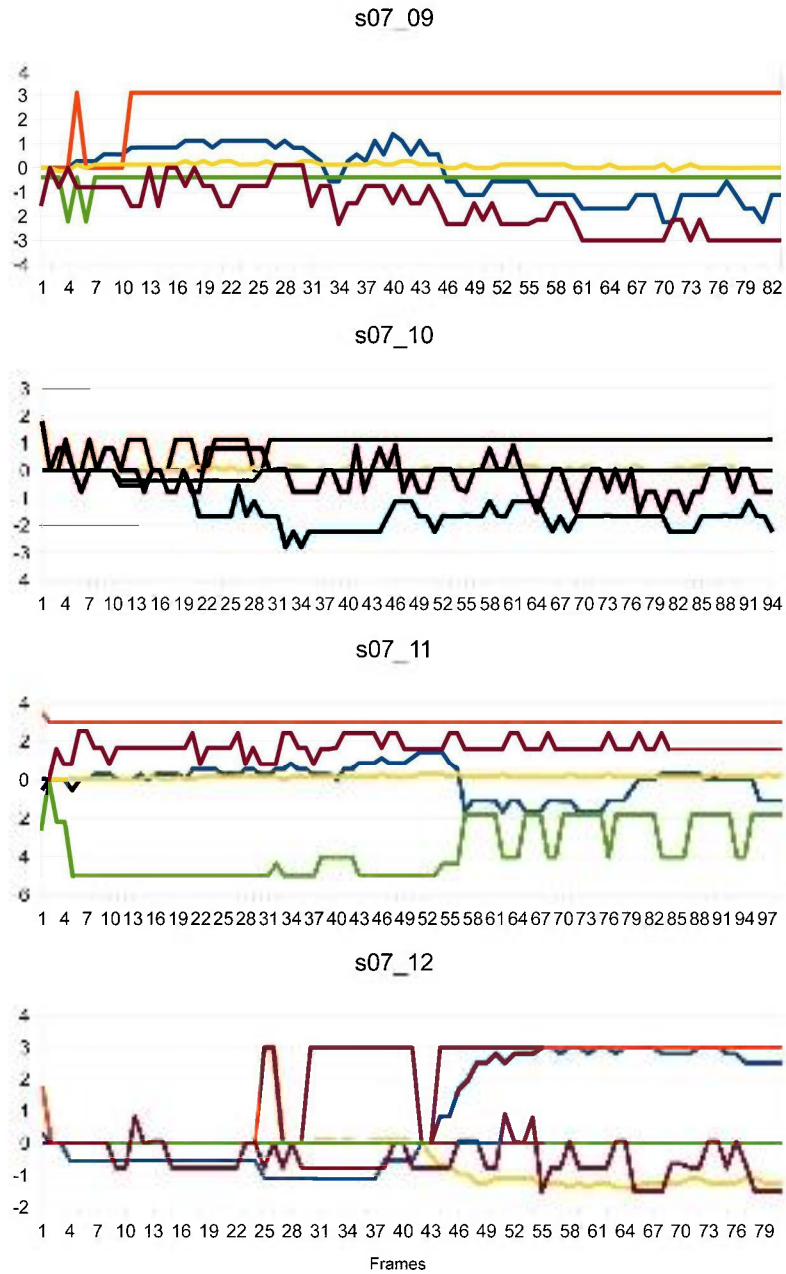


Figura 39: Valores de los principales controladores del *rig* para las diferentes expresiones de el sujeto s07

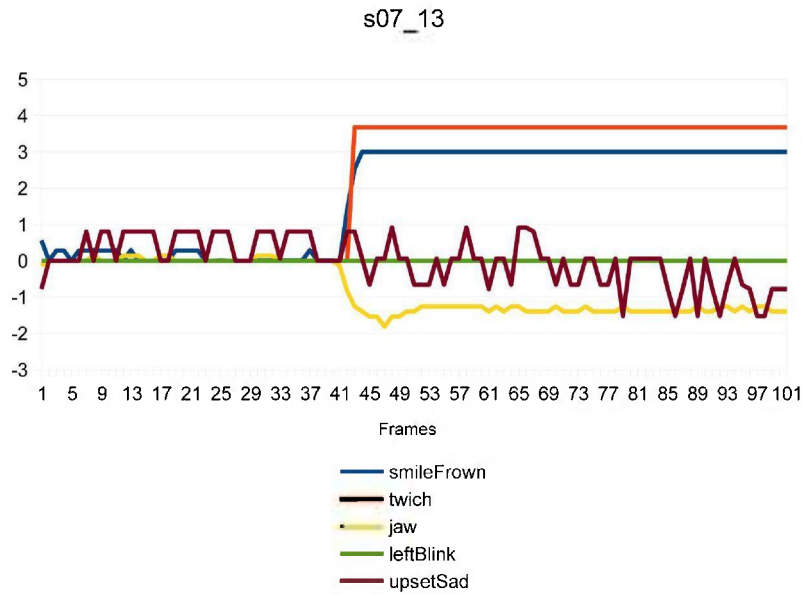


Figura 40: Valores de los principales controladores del *rig* para las diferentes expresiones de el sujeto s07

A partir de estos datos se genera la información de animación asignable al *rig* facial. En la imagen 41 se observa la asignación al modelo naturalista de la expresión risa clamorosa, sonrisa melancólica y sonrisa astuta. En la imagen aparece el modelo y el sujeto en el mismo cuadro de animación y en la parte inferior los gráficos de tiempo correspondientes al controlador *smileFrown* con los datos de la captura sin limpiar, es decir, con una clave de animación en cada cuadro de tiempo.



Figura 41: Asignación de los datos del sujeto 7 para las expresiones de risa clamorosa, sonrisa melancólica y sonrisa astuta

La visualización conjunta de las secuencias reales y la animadas permiten comprobar el grado de similitud entre ellas, no aparecen errores visuales en el seguimiento de las características faciales y por tanto las claves de animación se corresponden exactamente con la secuencia real, el ajuste de escala

empleado en todas las secuencias es el mismo y se corresponde a un valor de ϵ de 0.28 ajustado a la distancia de grabación y a la resolución de las imágenes de trabajo. Como se ha justificado en el capítulo anterior para este factor tenemos resultados correctos tanto para nuestro avatar naturalista como para el primitivo. Se han realizados ajustes manuales a las curvas de animación como la simplificación de curvas, eliminación de temblores y ajustes puntuales de escala.

Capítulo 7 Modelado de expresiones más representativas en las que interviene la AU12

En este capítulo se presenta un método para extraer de nuestra base de datos de risas las curvas de animación de los controladores del rig facial que mejor representan a cada una de las expresiones. Los datos de animación se comparan con el método Dynamic Time Warping y se extraen los resultados después de eliminar los sujetos outliers. Con los datos más representativos se generan las expresiones y se comparan con las que ejecutaban los sujetos de donde se ha extraído los datos.

7.1 Introducción

Analizando los datos obtenidos a partir de nuestra base datos intentaremos establecer curvas representativas para la animación de las 13 expresiones de risa consideradas en este estudio. El objetivo de este apartado es conseguir generar las animaciones de los controladores del *rig* facial eligiendo de entre las obtenidas de la base de datos de risas la que sea más representativa de todas ellas. Esto ha de permitir generar un sistema de animación procedural para la reproducción de los diferentes tipos de sonrisa.

Tomando por ejemplo el caso de estudio de la expresión de la "sonrisa clamorosa", grabada por los 7 sujetos de nuestro experimento, y analizando los datos correspondientes a la activación del controlador *smileFrown* se explica el método que se ha desarrollado.

En la imagen 42 se observa el último cuadro de la secuencia correspondiente al clímax de la expresión, junto con los puntos extraídos al aplicar el seguimiento facial, el número anotado al lado de la identificación del sujeto es el número de cuadros necesarios para llegar de la expresión neutra al punto culminante. A cada individuo de nuestra base de datos se le han dado las mismas especificaciones para reproducir cada expresión pero cómo es natural

cada uno la ejecuta con unas pequeñas variaciones, en concreto se observan diferencias en el tiempo para llegar al cenit de la expresión. En la imagen 43 se representan los valores en el tiempo del controlador *smileFrown* para esta expresión para cada uno de los sujetos (del s01 al s07) y en donde se puede apreciar este desplazamiento temporal.

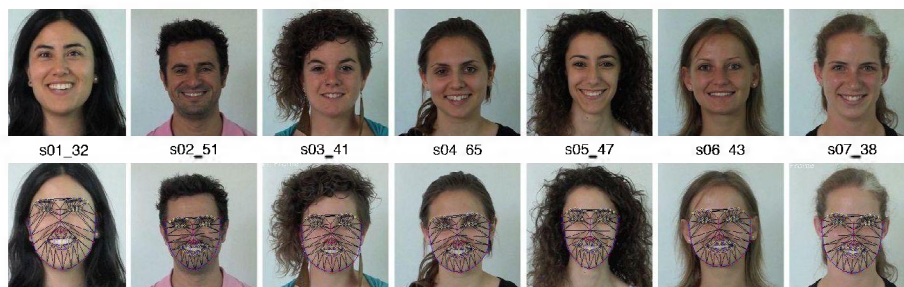


Figura 42: Sonrisa clamorosa ejecutada por los diferentes sujetos del experimento

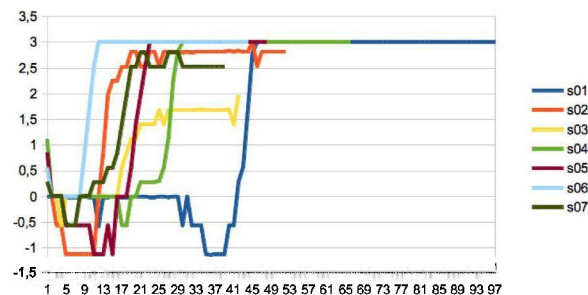


Figura 43: Valores del controlador *smileFrown* para la expresión "sonrisa clamorosa" de los 7 sujetos

Las secuencias de vídeo han sido cortadas convenientemente de manera que empiezan en un cuadro con la expresión neutra y en el siguiente empieza el movimiento de alguna parte de la cara. En el gráfico se observa como la curva correspondiente al sujeto s01 aparece claramente más desplazada que las demás, se aprecia también que los valores de las sujetos s01, s04, s05 y s06 llegan al punto máximo (valor 3 del controlador) mientras que los demás sujetos quedan por debajo de este umbral, especialmente claro es el caso del sujeto s03 cuya expresión es de una marcada baja intensidad.

En esta sección se plantea un método que permita escoger entre estas curvas la más representativa, o lo que es lo mismo, la curva que mejor defina los valores de animación de esta expresión.

7.2 Comparación de las curvas de animación con el algoritmo FastDTW

Para analizar secuencias de valores en el tiempo pueden utilizarse técnicas de Dynamic Time Warping (DTW) [109] que permiten el alineamiento óptimo entre dos series de tiempo mediante el estiramiento o encogimiento de una de ellas de forma no lineal a lo largo de su eje de tiempo. Estas deformaciones pueden ser usadas para encontrar regiones correspondientes entre dos series o bien para determinar la similitud entre dos series temporales. El DTW se ha convertido en un estándar para el estudio de series temporales con aplicaciones diversas en reconocimiento de patrones, en minería de datos, etc.

Para analizar dos series temporales se usa una medida de distancia, tanto para determinar su similitud como para su clasificación. El uso de la distancia euclidiana es efectivo y se consigue sumando el cuadrado de las distancias de cada punto n de una serie a cada punto n en la otra, el problema que aparece es que si dos series son idénticas pero una de ellas se encuentra ligeramente desplazada a un lado entonces la distancia euclidiana puede considerar que son muy diferentes entre sí. El DTW se introduce para superar esta limitación y dar una medida intuitiva de la distancia entre series de tiempo ignorando desplazamientos locales y globales en la línea de tiempo.

El problema se formula de la siguiente manera: Sean A y B dos series de datos de longitud M y N respectivamente

$$A = \{a_1, a_2, \dots, a_i, \dots, a_M\} \quad B = \{b_1, b_2, \dots, b_j, \dots, b_N\} \quad (7)$$

Llamando C a la función de alineamiento:

$$C = \{c(1), c(2), \dots, c(k), \dots, c(K)\} \quad (8)$$

Donde $c(k)$ es un par de punteros a los elementos a comparar:

$$c(k) = [i(k), j(k)] \quad (9)$$

Para $c(k)$ tenemos una función de coste:

$$d\{c(k)\} = \delta(a_{i(k)}, b_{j(k)}) \quad (10)$$

Que refleja la discrepancia entre los elementos comparados. Una función de coste típica es la distancia euclídea:

$$d\{c(k)\}=(a_{i(k)}-b_{j(k)})^2 \quad (11)$$

La función de alineamiento será la que minimice la función de coste total:

$$D(C)=\sum_{k=1}^K d\{c(k)\} \quad (12)$$

El problema de calcular la función de alineamiento se resuelve mediante técnicas de programación dinámica. El algoritmo DTW es de complejidad cuadrática lo que limita su uso a series de datos pequeñas. Salvador y Chan [109] presentan una aceleración de este algoritmo: el FastDTW con complejidad lineal que ofrece un enfoque multinivel que recursivamente proyecta una solución a partir de una resolución gruesa que va refinando.

En las figuras 44 y 45 se observa la matriz de coste resultado de aplicar el algoritmo FastDTW a los datos correspondientes al controlador *smileFrown* de las secuencias de la expresión "risa clamorosa" de los sujetos 4 y 6 en la figura 44 y los sujetos 3 y 4 en la figura 45.

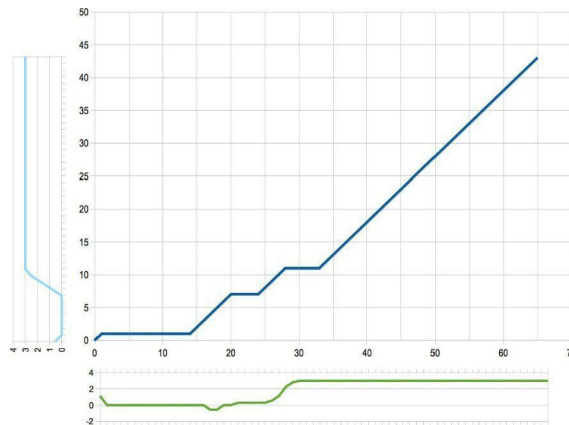


Figura 44: Matriz de coste con la distancia mínima de alineamiento para el controlador *smileFrown* de la expresión "risa clamorosa" de los sujetos s04 (X) y s06 (Y).

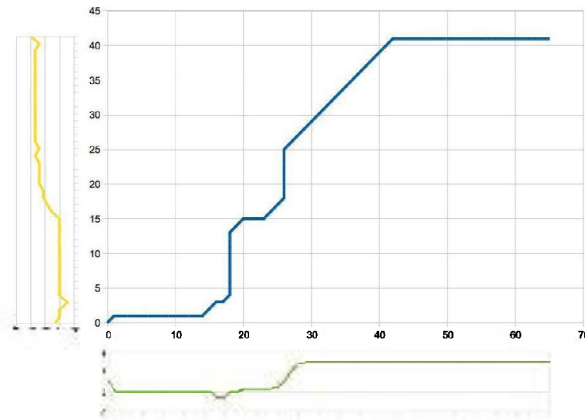


Figura 45: Matriz de coste con la distancia mínima de alineamiento para el controlador *smileFrown* de la expresión "risa clamorosa" de los sujetos s04 (X) y s03 (Y).

Los datos del sujeto s06 (en azul claro) están formados por 65 valores y los del sujeto s04 (en verde) por 43 y los del sujeto s03 (en amarillo) por 42, muestran como varía en el tiempo el controlador que simulará el efecto del músculo cigomático mayor. El resultado numérico que ofrece FastDTW para estos datos de entrada es una distancia de alineamiento (*warping*) de 4.29 cuando comparamos las curvas del s06 y s04 y de 48.25 cuando comparamos s03 y s04, estos valores nos dan una idea de la similitud entre las tres series. En concreto para el manipulador *smileFrown* de esta expresión las curvas del s04 y s06 son las que se encuentran a una distancia menor y las del s03 y s04 las que están a una distancia mayor.

	s01	s02	s03	s04	s05	s06	s07
s01		12.85	12.88	27.82	16.68	36.82	11.31
s02	12.85		42.15	10.90	8.66	16.17	9.16
s03	12.88	42.15		48.25	35.20	38.21	18.76
s04	27.82	10.90	48.25		4.78	4.30	13.42
s05	16.68	8.66	35.20	4.78		9.56	12.43
s06	36.82	16.17	38.21	4.30	9.56		13.66
s07	11.31	9.16	18.76	13.42	12.43	13.66	
PROMEDIO	19.73	16.65	32.57	18.24	14.55	19.79	13.12
DES. EST	10.32	12.79	13.82	17.00	10.87	14.32	3.21

Tabla 6: Tabla de distancias de alineamiento para el controlador *smileFrown* de la expresión "risa clamorosa"

Si aplicamos el FastDTW buscando la distancia de alineamiento entre todas las series comparándolas dos a dos tenemos una tabla como la 6. En la que comparamos por parejas las distancias entre todos los sujetos. Aplicando la media aritmética se deduce que la serie correspondiente al sujeto s07 es la que presenta en promedio menor distancia de *warping*, 13.12 con una desviación

estándar de 3.21 y por tanto la más representativa para la descripción del controlador smileFrown de la expresión "risa clamorosa".

Procediendo de manera similar podemos encontrar las curvas del controlador SmileFrown más representativas para la descripción de las 13 expresiones descritas por Faigin, como se ve en la tabla 7. En donde aparecen los sujetos cuyo promedio de distancias de *warping* mínimas respecto a los demás sujetos.

Expresión	Sujeto	Dist. Warping	Desv Est
Carcajada extrema	s02	69.87	51.35
Carcajada	s06	36.30	15.48
Risa clamorosa	s07	13.12	3.21
Sonrisa boca cerrada	s01	13.79	6.80
Risa sofocada	s03	46.16	13.27
Sonrisa melancólica	s01	27.83	16.99
Sonrisa ávida	s02	17.04	19.46
Sonrisa congraciante	s02	33.89	9.36
Sonrisa astuta	s04	47.16	52.99
Sonrisa desleal	s01	50.77	44.19
Sonrisa embarazosa	s06	57.86	28.50
Falsa sonrisa	s01	8.44	4.94
Falsa carcajada	s01	54.39	37.95

Tabla 7: Tabla con los valores de mínima distancia de warping para el controlador SmileFrown

De la misma forma se obtienen los valores representativos para el controlador del hueso de la mandíbula que se puede ver en la tabla 8, estos dos controladores son los principales a la hora de definir la animación de la boca.

Expresión	Sujeto	Dist. Warping	Desv Est
Carcajada extrema	s01	42.38	26.21
Carcajada	s04	18.60	12.19
Risa clamorosa	s05	4.89	1.76
Sonrisa boca cerrada	s06	2.14	1.14
Risa sofocada	s05	5.17	3.29
Sonrisa melancólica	s03	3.47	2.35
Sonrisa ávida	s07	12.39	7.68
Sonrisa congraciante	s07	9.27	7.05
Sonrisa astuta	s04	26.06	42.51
Sonrisa desleal	s02	10.04	9.52
Sonrisa embarazosa	s01	2.87	1.27
Falsa sonrisa	s02	11.97	13.46
Falsa carcajada	s05	18.77	12.45

Tabla 8: Tabla con los valores de mínima distancia de warping para el controlador Jaw

En las tablas 7 y 8 se evidencia que tanto para el controlador SmileFrown como para el Jaw las mayores distancias aparecen en el caso de la expresión **carcajada extrema**, en concreto 69.87 y 42.38 para cada controlador, con valores de error muy grandes, 51.35 y 26.21 lo que nos indicaría que esta

expresión es la menos generalizable y que es la más dependiente del sujeto. Existen diferentes factores que influyen en estos valores, por una parte se debe considerar que en esta expresión es en la que el ciclo de la risa tiene mayor presencia, Ruch [3] describe cómo el ciclo depende básicamente del volumen pulmonar, que obviamente depende del sujeto. Además debemos tener en cuenta que durante la realización de esta expresión es cuando los sujetos realizan mayor cantidad de movimiento, en concreto observamos inclinaciones del tronco y/o de la cabeza y cuello hacia adelante y atrás. En muchos casos estos movimientos se producen de forma brusca y repentina lo que origina que la aplicación de seguimiento produzca leves errores que observamos al reproducir la secuencia cuadro a cuadro.

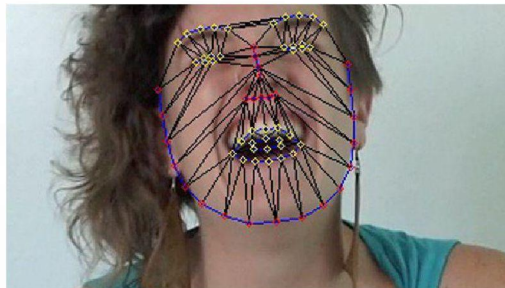


Figura 46: Detalle del error en el seguimiento

En la figura 46 se observa este error producido en la secuencia correspondiente a la grabación del sujeto s03, se aprecia a simple vista como los puntos del exterior de la boca no se corresponden a los de la imagen.

Analizando estas dos tablas se observa que determinadas expresiones presentan valores de dispersión elevados, en el caso de la tabla 6 se ve como el promedio de las distancias del sujeto 03 es mucho mayor que el de los demás, esto nos da pie a aplicar una búsqueda de los *outliers* de cada expresión para sacarlos de la muestra.

7.3 Detección y eliminación de *outliers*

Según la definición de Crubbs (1969), una observación periférica o atípica *outlier*, es la que parece desviarse marcadamente de otros miembros de la muestra en la que se produce. Los *outliers* son valores muy altos o muy bajos que pueden ser debidos a diferentes motivos, entre ellos la simple desviación

natural en poblaciones, y que introducen errores. Se pueden identificar los errores y eliminar su efecto contaminante en el conjunto de datos y purificar los datos para su procesamiento. Existen diferentes métodos de detección de *outliers*. Hodge y Austin[110] hacen un compendio de los principales métodos que son usados en la actualidad.

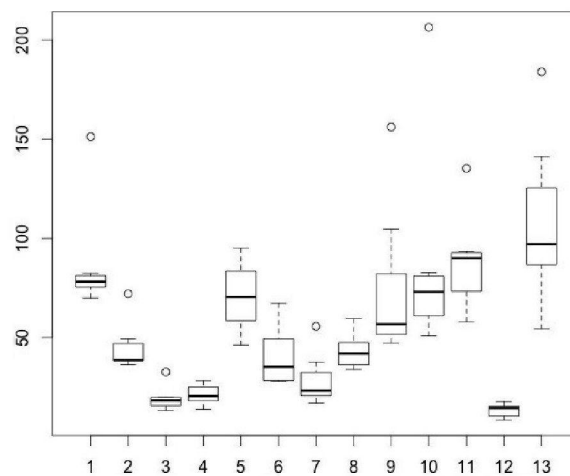


Figura 47: Box plot para la representación de los valores del controlador *SmileFrown* de las 13 expresiones de risa

Para nuestro estudio hemos escogido la técnica de detección de outliers estadísticos descrita por Laurikkala et al.[111] que se basa en utilizar diagramas de caja *boxplot* para identificar valores atípicos en los conjuntos de datos. Este sistema produce una representación gráfica, como se muestra en la figura 47 y permite que un auditor humano pueda identificar visualmente los puntos *outlier*. Este enfoque puede manejar valores reales, y atributos ordinales y categóricos (sin orden). Los diagramas de caja representan el extremo inferior, el primer cuartil, la mediana, el tercer cuartil y el extremo superior. Para datos univariantes, como los nuestros, se presenta un sencillo gráfico de 5 puntos. Los valores *outliers* son los puntos más allá de los extremos inferior y superior del box plot, y se representan por pequeños círculos como los de la figura.

En la figura se representa el diagrama de caja correspondiente al controlador *smileFrown* para las 13 expresiones de risa ordenadas en el mismo orden que en las figuras de capítulos precedentes que aparecen en el eje de las abscisas

mientras que en las ordenadas se representan los valores de las distancias de alineamiento. En el gráfico se observa que no aparece ningún *outlier* muy pequeño y que para las expresiones 1, 2, 3, 7, 9, 10, 11 y 13 tienen solo un *outlier* grande cada una.

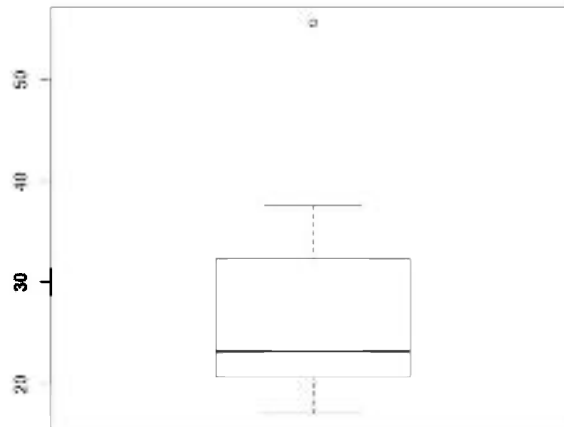


Figura 48: Box plot de los datos del controlador SmileFrown para la expresión 7 (Sonrisa ávida)

La expresión número 7 etiquetada con el nombre de sonrisa ávida, la podemos ver en detalle en la figura 48 en ella se representan los valores medios de las distancias de *warping* para los siete sujetos de la muestra, correspondiendo a los valores:

s01	s02	s03	s04	s05	s06	s07
20.38	17.04	20.83	55.61	27.07	23.13	37.54

En la gráfica se representa el primer cuartil 20.60, la mediana 23.13 y el tercer cuartil 32.30. Laurikkala et al. sugieren una heurística de 1,5 por el rango intercuartil para el cálculo de los valores extremos, este valor puede variar según el conjunto de datos. El valor 55.61 fuera del rango se representa con un círculo y es considerado un outlier.

Aplicando el proceso de detección de outliers y eliminándolos de las muestras las tablas con los sujetos más representativos para cada expresión quedan como las que se aprecian en 9 y 10 que sustituyen a las anteriores 7 y 8.

En el apéndice 1 se listan todas las tablas correspondientes a los controladores *SmileFrown*, *Jaw*, *LeftBlink* y *UpsetSad* cuyos valores coinciden con *Shocked*. Hay una tabla para cada expresión que contiene la distancia de alineamiento entre cada sujeto y los demás.

Expresión	Sujeto	Dist. Warping	Desv Est
Carcajada extrema	s02	51.53	27.8
Carcajada	s04	27.61	21.22
Risa clamorosa	s05	10.42	4.44
Sonrisa boca cerrada	s01	13.79	6.80
Risa sofocada	s03	46.16	13.27
Sonrisa melancólica	s01	27.83	16.99
Sonrisa ávida	s02	10.29	11.5
Sonrisa congraciante	s02	33.89	9.36
Sonrisa astuta	s04	25.67	6.89
Sonrisa desleal	s01	33.90	17.56
Sonrisa embarazosa	s06	49.87	23.17
Falsa sonrisa	s01	8.44	4.94
Falsa carcajada	s01	40.21	17.14

Tabla 9: Tabla con los valores de mínima distancia de *warping* para el controlador SmileFrown una vez eliminados los *outliers* de la muestra

Expresión	Sujeto	Dist. Warping	Desv Est
Carcajada extrema	s01	42.38	26.21
Carcajada	s04	18.60	12.19
Risa clamorosa	s05	4.89	1.76
Sonrisa boca cerrada	s06	2.14	1.14
Risa sofocada	s05	4.08	2.16
Sonrisa melancólica	s03	3.47	2.35
Sonrisa ávida	s02	9.92	4.59
Sonrisa congraciante	s07	9.27	7.05
Sonrisa astuta	s04	9.46	13.81
Sonrisa desleal	s02	10.04	9.52
Sonrisa embarazosa	s01	2.66	1.30
Falsa sonrisa	s02	7.33	8.09
Falsa carcajada	s05	14.71	8.36

Tabla 10: Tabla con los valores de mínima distancia de *warping* para el controlador Jaw una vez eliminados los *outliers* de la muestra

7.4 Generación de las secuencias animadas a partir de los datos del FastDTW

Siguiendo el mismo proceso que el llevado a cabo para la creación de las tablas anteriores, podemos generar tablas similares para todos los controladores del *rig* facial, como se observa en la tabla 11.

Expresión	CtrlJaw	CtrlSmileFrown	CtrlLeftBlink	CtrlLeftTwich	CtrlUpsetSad
	Sujeto	Sujeto	Sujeto	Sujeto	Sujeto
Carcajada extrema	s01	s02	s01	s02	s06
Carcajada	s04	s04	s06	s01	s02
Risa clamorosa	s05	s05	s06	s01	s01
Sonrisa boca cerrada	s06	s01	s01	s01	s04
Risa sofocada	s05	s03	s07	s07	s06
Sonrisa melancólica	s03	s01	s04	s04	s04
Sonrisa ávida	s02	s02	s07	s02	s02
Sonrisa congraciante	s07	s02	s01	s02	s01
Sonrisa astuta	s04	s04	s01	s01	s04
Sonrisa desleal	s02	s01	s07	s01	s07
Sonrisa embarazosa	s01	s06	s05	s02	s01
Falsa sonrisa	s02	s01	s02	s06	s06
Falsa carcajada	s05	s01	s04	s01	s01

Tabla 11: Sujetos más representativos para cada controlador de las diferentes expresiones de sonrisa

En esta tabla solo aparecen los valores de los controladores de la parte izquierda tanto del *Blink* como del *Twich* y no hay ninguna referencia a las asimetrías que pueden generarse en la parte de la boca y que estarían dirigidas por los controladores *SmileFrown* y *Jaw*. En cualquier caso estos elementos serán calculados de forma análoga al proceso descrito y será el animador que en su momento valorará a qué controlador asignar la información. Es decir, después del proceso de seguimiento y captura de las características faciales y una vez elegidos los más representativos, se dará libertad al animador para que decida qué controladores activar y una vez activados, graduar su intensidad.

Se pueden visualizar los resultados obtenidos de la asignación de las animaciones a los controladores del *rig* en la siguiente URL:

http://dmi.uib.es/~miquel/webdedocencia/Video_Results_1_3.html

En la web mencionada se recogen 13 vídeos uno para cada una de las expresiones descritas por Faigin. En estos vídeos se observa en primer lugar la base de datos con los siete sujetos diferentes ejecutando la expresión, esta secuencia se continua con el seguimiento de las características faciales según el Facetracker de Saragih, el seguimiento se observa sobre un sujeto de la base de datos, que en cada secuencia va variando, en concreto se ha escogido el sujeto más representativo, según el algoritmo FastDWT, respecto al controlador del *SmileFrown* y termina el vídeo con la secuencia renderizada de los datos de animación asignados a tres avatares diferentes: el naturalista y el primitivo, con un *rig* definido según el método de Digital Tutors y el modelo Emily con el *rig* definido por *Faceware Technologies*. Todas las secuencias especificadas se visualizan dos veces y se alargan los cuadros inicial y final para su mejor apreciación. En la imagen 49 se resume el contenido de estos vídeos en ella aparece en la parte izquierda el sujeto más representativo según el controlador *smileFrown* y a la derecha uno de los avatares en el momento

más álgido de la expresión representada.

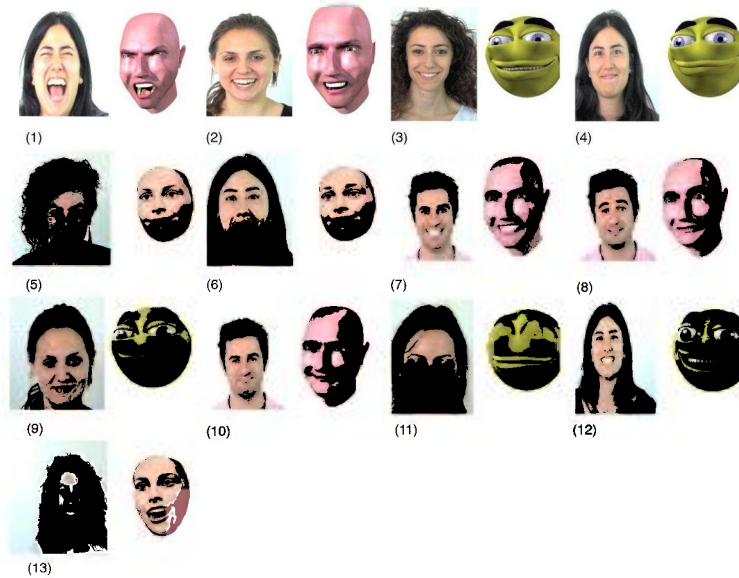


Figura 49: Reproducción de las expresiones de risa más representativas: (1) Carcajada extrema (2) carcajada (3) sonrisa clamorosa (4) sonrisa boca cerrada (5) sonrisa sofocada (6) sonrisa melancólica (7) sonrisa ávida (8) sonrisa congraciante (9) sonrisa astuta (10) sonrisa desleal (11) sonrisa embarazosa (12) falsa sonrisa y (13) falsa carcajada

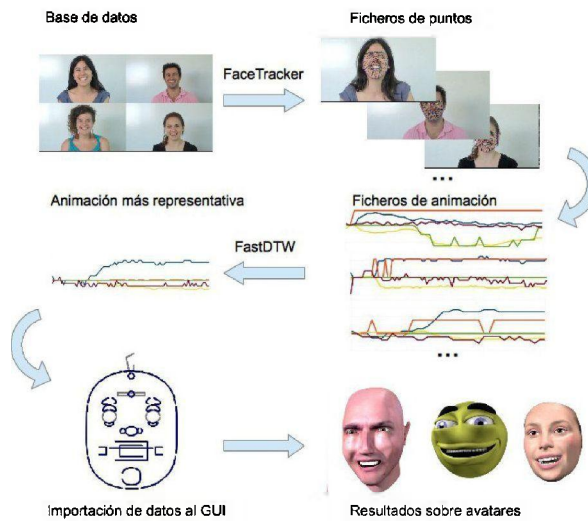


Figura 50: Proceso de generación de la expresión facial a partir de la base de datos de sonrisa

Para la realización de los 13 expresiones se ha ejecutado nuestra aplicación sobre los 91 ficheros de puntos resultantes de la aplicación *FaceTracker* al mismo número de ficheros de vídeo que forma nuestra base de datos de expresiones. Como resultado nuestra aplicación genera la animación de los controladores del *rig*. Estos datos son los que se envían al algoritmo FastDTW que comparándolos dos a dos elige cuales son los mas representativos de entre los diferentes sujetos para cada controlador y para cada expresión. Éstos son los que se importan al *rig* que controla el avatar en un entorno 3D. Un resumen gráfico de este proceso se observa en la figura 50.

7.5 Ajustes generales sobre las curvas de animación

Como se ha especificado con anterioridad, una vez asignada la animación al controlador correspondiente ésta se puede ajustar a criterio del animador, debemos notar aquí que en el caso de la animación correspondiente a las expresiones de risa más representativas que hemos realizado a partir de nuestra base de datos, ha habido una serie de ajustes que se han llevado a término de forma generalizada.

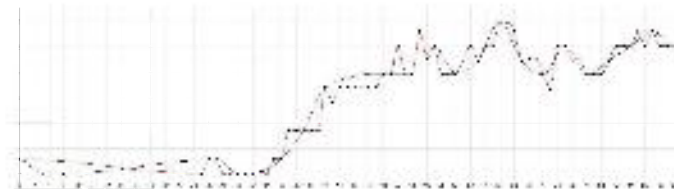


Figura 51: Simplificación de curvas

La **simplificación del número de claves** de las curvas de animación es un proceso habitual en los sistemas de MOCAP. Existen numerosas referencias bibliográficas sobre esta cuestión, como en los trabajos de Xiao et al. [112] y en los de Onder et al. [113]. Así mismo la mayoría de aplicaciones de animación 3D incluyen operaciones que permiten esta simplificación de curvas de forma que animaciones creadas realizando movimientos de forma interactiva que prácticamente tienen una clave de animación en cada cuadro puedan ser simplificadas para su posterior edición y manejabilidad. De esta manera las curvas de animación de los controladores de las cejas, la mandíbula y la sonrisa han sido simplificadas. En la figura 51 se aprecia el resultado del algoritmo de simplificación de curvas de datos densos de

Autodesk Maya sobre la curva correspondiente al SmileFrown de la expresión sonrisa astuta, en rojo, con los datos provenientes del seguimiento facial y en azul, una vez ejecutada la operación de simplificado.

Se pueden tener en consideración ajustes más finos sobre las curvas como la eliminación de oscilaciones que pueden resultar incorrectas visualmente desde el punto de vista del animador. En el gráfico anterior se observa cómo se activa la sonrisa a partir del cuadro 33, se estabiliza en el 50 y aparecen oscilaciones a partir del cuadro 54. Quedaría a criterio del animador actuar sobre los 3 picos de la curva, para eliminarlos, suavizarlos o mantenerlos. De forma similar, y sobre todo en el contexto de la animación de personajes primitivos menos realistas, tiene sentido hablar de escalado de los valores de la curva de animación con el objetivo de exagerar la expresión.

De forma generalizada, durante la generación de las expresiones, se ha observado la necesidad de ajustar las curvas que controlan el movimiento de elevación de las mejillas, etiquetado como la unidad de acción AU6 según el FACS, debido a su relación directa con el estiramiento de la esquina de los labios por el músculo cigomático mayor, etiquetado con la unidad de acción AU12. En el FACS, estas dos acciones aparecen con estas dos etiquetas, como movimientos independientes e ilustrados con las fotografías que se ven en la figura 52.

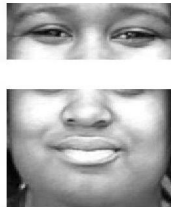


Figura 52: Imágenes que ilustran las acciones AU6 y AU12 en el FACS

En el estudio pormenorizado de la base de datos que hemos construido para el análisis de las 13 expresiones de risa, hemos observado que en 12 de ellas existe una sincronía visual completa entre la AU6 y la AU12, coincidiendo en intensidad y en duración temporal; esto ocurre además en todos los sujetos de la base de datos. Solo en la expresión etiquetada con el nombre de risa sofocada, en donde los sujetos de forma voluntaria bloquean la AU12, se pueden observar cierta independencia entre las dos acciones.

Siguiendo el tutorial para la generación del *rig* de Digital Tutors o bien observando las geometrías de otros *rigs* se observa una relación directa entre estas dos expresiones, es decir la geometría que se usa para la interpolación de la sonrisa presenta los puntos desplazados de la comisura de los labios y a la vez la elevación de la mejilla. Los *rigs* analizados en este estudio, aparte de

esta relación directa entre las acciones AU6 y AU7, contienen otro controlador para realzar el efecto que tiene la elevación de la mejilla en la presión que ejerce sobre el párpado inferior del ojo correspondiente. En el caso de *rig* de Digital Tutors nos referimos al controlador Twitch y en el caso del *rig* de Faceware para el controlador del ojo se tienen tres parámetros que permiten controlar respectivamente el cerrado del párpado superior, la presión que ejerce el párpado inferior y la elevación de las mejillas respecto a su efecto en la zona de los ojos.

Como se ha explicado en punto 4.2 anterior, para el cálculo de la animación del controlador Twitch que simula la AU6 se ha introducido un proceso de filtrado con el objeto de reducir oscilaciones que ocasione un parpadeo no deseado. Este proceso que resulta válido con los tests realizados con la base de datos Kohn - Kanade, no lo es, sin embargo, cuando existe la sincronía a que nos estamos refiriendo. Si tenemos en cuenta que por cuestiones de resolución los datos faciales asignados al controlador SmileFrown que se han extraído con la aplicación de seguimiento son más robustos que los datos filtrados a que nos referimos concluimos que sustituiremos los datos del controlador Twitch por los del SmileFrown para conseguir la sincronía exacta. Si el *rig* se ha definido con la misma escala numérica de valores para los diferentes controladores no será necesario ningún ajuste de escala posterior.

7.6 Estudio particularizado de las curvas de cada expresión

Pasamos a describir a continuación las particularidades que nos hemos encontrado a la hora de asignar los datos de animación para realizar cada una de las 13 expresiones analizadas.

1. Carcajada extrema

Para esta expresión aparecen problemas de sincronía graves a partir de los datos que se extraen de *DTW*, debidos a que el controlador de la mandíbula y el del músculo cigomático mayor se corresponden a los datos generados por el sujeto 1 y 2 respectivamente. En esta expresión se observan, en la mayoría de sujetos, movimientos de inclinación adelante y atrás de la cabeza que dificultan el seguimiento facial correcto. En el caso del sujeto 1, aunque existe un movimiento pronunciado hacia adelante prácticamente no hay cambio en lo que se refiere al plano frontal de la cámara. De forma que para garantizar la sincronía entre mandíbula y movimiento de la boca asignamos al controlador Jaw y al *SmileFrown* los datos extraídos del sujeto 1.

El *Twich* es sustituido por el *SmileFrown* para garantizar la sincronía entre boca y mejillas. Las demás curvas de animación han sido simplificadas limpiando las claves de la captura de movimiento.

2. Carcajada

La animación de la mandíbula ha tenido que ser escalada sobre todo en el caso del personaje primitivo. En esta expresión el controlador de los párpados superiores permanece inactivo.

Se ha sustituido el *Twich* para garantizar la sincronía y se han simplificado las curvas.

3. Risa clamorosa

Para el controlador de las cejas el algoritmo *FastDTW* da como más representativa la animación correspondiente al sujeto 1, estos datos no producen una visualización satisfactoria debido a la expresión llevada a cabo por este sujeto es incorrecta como se aprecia en la secuencia de vídeo de la base de datos. Debido a que en los demás sujetos no es apreciable actividad en esta zona se ha optado por no animar este controlador.

El controlador de la mandíbula ha sido escalado. El controlador de los párpados superiores está inactivo. Se han sustituido los datos del controlador de las mejillas por los de la boca y se han suavizado las curvas eliminando las claves de animación innecesarias.

4. Sonrisa boca cerrada

En esta expresión se sustituyen los datos del controlador de la elevación de las mejillas por los de la elevación de los labios, se suavizan las curvas de animación del cuello y boca. En este caso la zona de las cejas, mandíbula y párpados superiores permanecen mayormente inactivos.

5. Risa sofocada

En el estudio de la base de datos para esta expresión se observa de forma generalizada cómo todos los sujetos ejercen presión en los labios reprimiendo la risa, en lo que en el FACS está codificado como la AU24 *Lip Pressor* o la AU28 *Lip Suck*. Para llevar estas acciones el *rig* de nuestros avatares usa el controlador llamado *M F U MouthRolls*, en la activación de estas acciones va implícita la AU17 que se encarga de la elevación de la barbilla. En el caso del *rig* de *Faceware* existe un controlador específico para la AU17 y para la presión en los labios existen tres controladores diferentes que permiten modular la presión de los labios en su parte izquierda, derecha y central. Tal y como se ha especificado en el punto 4.2 del presente documento nuestra aplicación no recoge datos específicos para este controlador. De forma empírica se han conseguido buenos resultados visuales de la simulación de

esta expresión adaptando los valores obtenidos para el controlador *SmileFrown*. Efectivamente, el movimiento realizado involucra la zona de la boca y en concreto los mismos puntos que son las comisuras y los puntos centrales.

Para esta expresión se observa cómo pueden aparecer valores negativos para el controlador de la comisura de los labios, provocando valores que tradicionalmente se asocian a expresiones de tristeza, en este caso suelen ir seguidos de oscilaciones en la curva de animación típicas del movimiento de represión de la risa.

En este caso la zona de los párpados tanto superiores como inferiores se encuentra mayormente inactiva, con también lo está la mandíbula, acaso en esta última se observan pequeñas oscilaciones fruto del ciclo de la risa.

Como en las otras expresiones se han suavizado las curvas de los controladores con claves de animación provenientes de la captura de datos faciales.

6. *Sonrisa melancólica*

En esta expresión es fundamental el papel que juega la animación de las cejas que suben como lo hacen típicamente en las expresiones asociadas a la emoción de tristeza. Los datos extraídos de los puntos de las cejas a parte de ser asignados al controlador *UpsetSad* que se encarga de realizar tanto la AU1 *Inner Brow Raiser* como la AU4 *Brow Lowerer*, también se asignan al *BrowRaise*. En nuestro *rig* este controlador se divide en su parte derecha e izquierda y lleva a cabo la AU2 *Outer Brow Raiser* a la vez que cierra levemente el ojo contrario al que se aplica. En ambos controladores los valores de la curva han sido suavizados y además se han eliminado las oscilaciones manualmente para mejorar la visualización de la expresión.

Igual que en la expresión descrita anteriormente en ésta también aparece una presión en los labios AU24 en la mayoría de sujetos de la base de datos, de forma similar que en el caso anterior se ha activado el controlador *M_F_U_MouthRolls*.

En esta expresión también se han sustituido los valores del *Twisch* por los del *SmileFrown* y se han suavizado además las curvas de animación del controlador del cuello, de los labios y de la mandíbula, si bien esta última aparece básicamente inactiva.

En el controlador del párpado superior se han eliminado las oscilaciones aunque se ha mantenido un parpadeo que aparece en la secuencia elegida con el *FastDTW*.

7. *Sonrisa ávida*

Para la generación de esta expresión se han detectado dos problemas diferentes. Por una parte en la zona de la boca nos encontramos con que nuestro *rig* tiene dificultades para la reproducción del movimiento consistente en bajar el labio inferior enseñando los dientes inferiores. De hecho esta acción no está codificada en FACS como una acción independiente y se generaría con la AU20 *Lip Strecher* sin llegar a su cenit, más la AU26 *Jaw Drop*. En el caso del *rig* de *Faceware Technologies* si existen dos controladores que permiten bajar la parte derecha e izquierda respectivamente del labio inferior. Asignando a estos controladores los mismos valores que hemos calculado para el *Jaw* se consigue un movimiento sincronizado y correcto.

Por otra parte la problemática del movimiento de las cejas según se desprende de la observación de la base de datos depende bastante de cada individuo: en algunos casos se aprecia movimientos de bajada AU4 *Brow Lowerer*, o de subida AU1 *Inner Brow Raise*, o incluso con ciertas asimetrías debidas a la AU2 *Outer Brow Raise*. En todos ellos se verifica que son movimientos de baja intensidad y en ningún caso se llega al cenit de la expresión. El algoritmo *FastDTW* nos da al sujeto 2 como más representativo para la animación de las cejas y la boca, según su ejecución de la expresión hemos asignado los datos de animación de los puntos al controlador *BrowRaise* simulando la asimetría que se observa en el vídeo.

En esta expresión tanto el cuello como el movimiento de párpados es prácticamente inexistente, también, en esta expresión sustituimos el *Twitch* por el *SmileFrown* y suavizamos todas las curvas.

8. *Sonrisa congraciante*

A diferencia de la expresión anterior, en ésta, la asimetría debida a la elevación de una ceja se puede observar en todos los sujetos de la base de datos. Por tanto se han asignado los datos de animación de los puntos al controlador *BrowRaise*. Se ha ajustado además el controlador del parpadeo eliminando oscilaciones y suavizando la asimetría.

9. *Sonrisa astuta*

En esta expresión básicamente se han asignado a los controladores los datos extraídos de la aplicación de seguimiento sin mayores modificaciones, en todo caso citar que se ha añadido al controlador del cuello una animación de rotación en el eje X coincidente con la animación de los labios. Además, en el caso del *rig* de *Faceware* se ha añadido la animación del *SmileFrown* a la presión de los labios que aumenta el parecido con la secuencia de vídeo del sujeto más representativo.

En la animación de las cejas se ha suavizado la curva y eliminado las oscilaciones erróneas de los datos.

10. *Sonrisa desleal*

En referencia a esta expresión hay que decir que las expresiones de la base de datos difieren bastante entre ellas, incluso la expresión del sujeto 1 presenta dudas que pueda ser interpretada como sonrisa desleal. Aun así el algoritmo *FastDTW* la selecciona como la expresión más representativa para el *SmileFrown*. Así ha sido asignada y después simplificada.

La animación del cuello se han mantenido las oscilaciones que simulan los ciclos de la expresión. La animación de las cejas se ha asignado al controlador *UpsetSad* y al controlador *BrowRaise*, suavizando la curva y eliminando las oscilaciones erróneas.

11. *Sonrisa embarazosa*

En esta expresión, a diferencia de las demás, el papel que juega la animación de los párpados es fundamental. Se asigna la animación de los puntos al controlador *Blink* con su correspondiente suavizado de curva. En el caso de la reproducción de la expresión con nuestro avatar naturalista se observa una presión en los ojos que es debida a la definición de la geometría de ojos cerrados, no a la animación, de hecho esto se confirma al no aparecer este efecto en los demás avatares.

El controlador que dirige el movimiento de las cejas se anima con los datos propios correspondientes a la extracción de características faciales, se le añade un escalado positivo para aumentar el efecto de arqueamiento de las cejas.

12. *Falsa sonrisa*

En esta expresión al igual que lo que ocurría en el caso de la carcajada extrema no coinciden los sujetos elegidos por el *FastDTW* que ejecutan el movimiento de labios y el de mandíbula, por lo que aparecen problemas de sincronía en la animación de la zona de la boca. Estos problemas solo afectan al momento de la activación del movimiento de la mandíbula ya que la apertura de la misma es similar en los dos casos.

Para esta expresión hemos eliminado las oscilaciones de error que aparecen en la animación del controlador de las cejas.

Una característica de las expresiones de sonrisa y carcajada falsas es la menor presión que existe en la elevación de las mejillas y en la zona inferior de los ojos, con el objetivo de simular este efecto se ha asignado al controlador *Twich* sus valores suavizados y con una operación de reducción de escala. Además se han llevado a cabo las operaciones habituales de suavizado de las demás curvas.

13. Falsa carcajada

En esta expresión, igual que en la anterior, también aparece el problema de que el controlador de los labios y el de la mandíbula pertenecen a diferentes sujetos de la base de datos, en concreto a los sujetos 1 y 5 respectivamente, en este caso al igual que en el de la sonrisa desleal aparecen dudas sobre la validez de la expresión llevada a cabo por el sujeto 1 a pesar de que el algoritmo *FastDTW* lo elija como el más representativo. Por tanto se ha optado por asignar a los dos controladores los datos provenientes del sujeto 5.

Igual que en el caso de la expresión falsa anterior, en ésta se han suavizado también los datos del controlador *Twich* y de forma análoga se han realizado las operaciones de suavizado habituales.

Capítulo 8 Validación de las expresiones generadas a partir de la base de datos

En este capítulo se explica la prueba de validación realizada para probar la calidad de las expresiones generadas con la aplicación del método desarrollado. Se presentan y analizan los resultados obtenidos de la prueba.

8.1 Metodología aplicada

Para la validación de las expresiones faciales generadas con nuestro método se ha llevado a cabo un experimento enfocado a evaluar la calidad del movimiento facial y la representatividad de la expresión realizada. Así mismo se han realizado dos versiones del experimento para determinar el grado de influencia en la expresión que puede tener el modelo del avatar que la reproduce. Ambos tests se han realizado mediante la visualización de una serie de trece vídeos contestando a preguntas sencillas después de cada visualización.

Los sujetos de nuestro experimento son ciento cincuenta y cinco estudiantes universitarios de ambos sexos de entre diecinueve y veinte años de edad, pertenecientes a estudios de informática y pedagogía. No se ha considerado el efecto de que la muestra esté sesgada ya que como bien expone Orvalho et al. [40] en su estado del arte: cualquier observador tiene unas expectativas muy altas a la hora de analizar el movimiento facial, y es un experto en la detección de expresiones. La anomalía más pequeña de la forma de la cara, la proporción, la textura de la piel o el movimiento es detectada y clasificada como errónea inmediatamente.

La visualización de los vídeos y la siguiente respuesta a las preguntas del test se ha realizado en una aula equipada con un proyector de vídeo y mesas y sillas dónde se encuentran los sujetos. El experimento se ha realizado en dos tandas, una con setenta y siete y otra con setenta y ocho sujetos. Antes de empezar el experimento se les ha hecho entrega de una hoja en la que en la

primera cara se encuentra una breve explicación del experimento y en la segunda están las tablas que serán usadas para la respuesta. En el apéndice 2 se encuentra una muestra de este enunciado.

8.1.1 Test de validación

Validación de la calidad animación de 13 expresiones de risa y sonrisa.

Para ello se han visualizado 13 vídeos uno para cada expresión. En los vídeos aparece en primer lugar el nombre de la expresión en inglés igual que el que aparece en la hoja que se entrega a cada sujeto junto con un dibujo que representa la expresión, seguidamente desaparece el nombre de la expresión y en su lugar aparece el avatar en la posición neutra (sin expresión), después el avatar ejecuta la expresión y el último cuadro se congela unos segundos con la expresión en su grado máximo; después se repite una segunda vez para su mejor observación. La duración estimada de la prueba es de unos 25 minutos.

Para cada vídeo el sujeto debe contestar dos sencillas preguntas según una escala que va del 0 (muy poco) a 5 (mucho), estas preguntas son: ¿El movimiento te parece correcto? y ¿El movimiento representa la descripción? Se ha de recalcar que las preguntas se refieren al movimiento que realiza el avatar y nunca a la forma y/o textura del avatar. Las preguntas se contestan en una plantilla suministrada a cada sujeto, en la que aparecen las dos preguntas junto con la escala de respuesta, delante de las preguntas está el número que identifica a cada expresión, con el nombre en inglés (igual que en el vídeo) de cada expresión junto con una breve descripción de la expresión. Cuando se visualizan los vídeos la persona que conduce la prueba explica cada expresión leyendo en voz alta estas descripciones y el nombre de las expresiones.

Para la primera tanda, o versión del experimento, todos los vídeos son expresiones realizadas por el avatar naturalista de Faceware Thecnologies conocido con el nombre de Emily. Este avatar ha sustituido a nuestro avatar naturalista básicamente por los errores de textura y geométricos expuestos en el punto 6.1.1 anterior. Mientras que en la segunda tanda los vídeos son las mismas expresiones realizadas por nuestro avatar primitivo. Las expresiones de ambos avatares están realizadas con las mismas curvas de animación sintetizadas con la aplicación de nuestro método expuesto en este trabajo.

Los vídeos visualizados en test se pueden consultar en la página:

<http://dmi.uib.es/~miquel/webdedocencia/Recerca.html>

Los de la primera tanda están etiquetados como "Validación A" y los de la segunda como "Validación B".

8.2 Resultados obtenidos

8.2.1 Valoración de la calidad del movimiento y de la expresión

A partir de los datos obtenidos con la prueba podemos decir de forma generalizada que los observadores evalúan de forma positiva los vídeos generados. Exceptuando un resultado debido a un error en el vídeo, todas las valoraciones se encuentran en la mitad superior de la escala. En el caso de los movimientos realizados por el personaje naturalista o realista, los encuestados dan una valoración media de 3.37 puntos sobre 5 a la pregunta de si el movimiento parece correcto y de 3.35 para la pregunta sobre si la expresión representa la expresión descrita, para el caso de los movimientos realizados por el avatar primitivo o con estética *cartoon* los resultados son similares, en concreto, a la evaluación del movimiento se le dan 3.27 puntos sobre 5 y a la representación de la expresión 3.37. En ambos casos y en ambas preguntas el error es el mismo y tiene un valor de +/-1.1.

La discrepancia en la calidad del movimiento está en concordancia con la teoría del "valle misterioso" de Mori[89], según la cual el avatar naturalista estaría mejor valorado que el primitivo debido a su mayor "humanidad", el hecho de presentar el avatar naturalista como una máscara flotando en fondo blanco, sin pelo y sin su cuerpo, le dotan de características sintéticas que lo alejan del "valle misterioso" en donde las características más realistas pero no completamente humanas hubiesen podido influir en una valoración negativa por parte del observador.

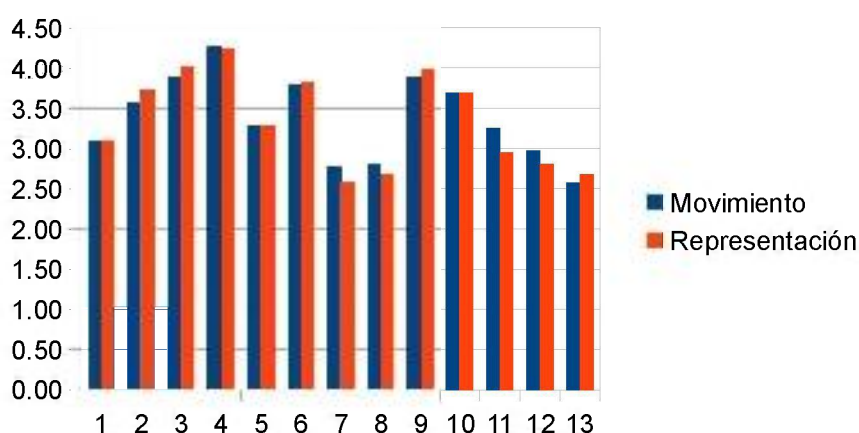


Gráfico 3: Promedio de la valoración de las distintas expresiones realizadas por el avatar naturalista

No todas las expresiones han sido valoradas de la misma manera, en los gráficos 3 y 4, se observan las gráficas correspondientes a los resultados de las 13 expresiones de risa y sonrisa de nuestro estudio. En el eje de las X están los números que referidos a las expresiones y en el Y la puntuación media en una escala que va de 0 a 5. En azul aparece la valoración sobre la calidad del movimiento y en rojo la valoración sobre si ese movimiento representa la expresión deseada. El gráfico 3 es la valoración para el personaje naturalista y el gráfico 4 lo es para el personaje primitivo. Estos gráficos representan los valores de la tabla 12.

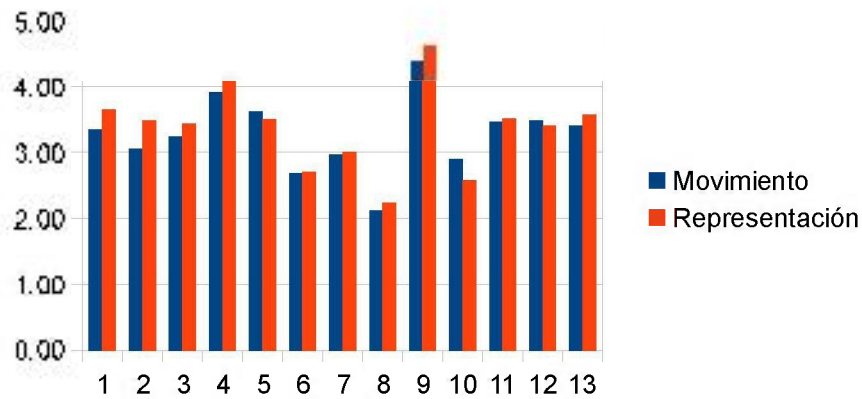


Gráfico 4: Promedio de la valoración de las distintas expresiones realizadas por el avatar primitivo

	Modelo naturalista				Modelo primitivo			
	Mov	Des	Rep	Des	Mov	Des	Rep	Des
1 Carcajada	3.10	1.05	3.10	0.93	3.35	0.87	3.64	0.95
2 Carcajada extrema	3.57	1.08	3.73	1.00	3.05	1.16	3.47	1.26
3 Sonrisa boca abierta	3.90	0.98	4.01	0.92	3.23	1.02	3.44	1.10
4 Sonrisa boca cerrada	4.26	0.88	4.25	0.86	3.91	0.97	4.10	1.00
5 Risa reprimida	3.29	1.29	3.29	1.26	3.62	1.19	3.50	1.35
6 Sonrisa melancólica	3.81	1.12	3.82	1.00	2.68	1.41	2.71	1.50
7 Sonrisa ávida	2.77	1.11	2.57	1.24	2.96	1.00	3.00	1.14
8 Sonrisa congraciante	2.81	1.08	2.68	1.15	2.12	1.40	2.23	1.41
9 Sonrisa astuta	3.90	1.29	3.99	1.29	4.41	0.84	4.64	0.74
10 Sonrisa desleal	3.69	1.32	3.69	1.37	2.90	1.15	2.58	1.23
11 Sonrisa embarazosa	3.25	1.31	2.95	1.39	3.46	1.27	3.51	1.16
12 Falsa sonrisa	2.97	1.29	2.81	1.32	3.47	1.03	3.40	1.13
13 Falsa carcajada	2.57	1.38	2.68	1.46	3.40	1.02	3.56	1.10

Tabla 12: Resultados de la prueba de valoración de las expresiones generadas

De estos datos deducimos que se valora de forma similar tanto la calidad del

movimiento como la fidelidad de representación de las expresiones, y esto es independiente del tipo de avatar que las ejecute.

La expresión mejor valorada es la sonrisa astuta realizada por el modelo primitivo (4.41 movimiento y 4.64 representación), seguida por la sonrisa con la boca cerrada realizada por el modelo naturalista (4.26 movimiento y 4.25 representación) en todos los casos con un error inferior a 0.9. La expresión peor valorada es la sonrisa congraciante reproducida por el modelo primitivo (2.12 movimiento y 2.23 representación) seguida de la sonrisa ávida del modelo naturalista (2.77 movimiento y 2.57 representación).

Se observa un error en el vídeo correspondiente a la sonrisa congraciante del modelo primitivo. En la secuencia animada el personaje no mueve la parte superior de la cara, la animación de cejas y ojos está congelada de manera que solo se ve el efecto de las claves correspondientes a la expresión final, es decir, no realiza el movimiento que va desde la posición neutra a la expresión. Este error justifica la mala valoración de la expresión en los dos aspectos que se evalúan.

Existe una relación entre las expresiones mejor valoradas y las peor valoradas, entre las últimas se encuentran aquellas cuyo significado puede resultar más confuso mientras que en las mejor valoradas están expresiones cuyo significado es menos ambiguo. Así la sonrisa congraciante que en el test ha sido descrita como "insinuante" o "para captar la simpatía", junto con la sonrisa ávida descrita como "ansiosa" o "eufórica" reciben las peores puntuaciones coincidiendo con la ambigüedad que presentan sus descripciones.

Confusión similar a la que ya habíamos detectado en 3.3 cuando habíamos realizado las pruebas de emparejamiento de fotogramas fijos del modelo primitivo con los dibujos de Faigin y con fotogramas de películas. Los estudios de estos errores nos llevan a la conclusión que independientemente de la calidad de la reproducción del movimiento de las expresiones y del realismo del avatar, para la correcta interpretación de las expresiones deben considerarse básicamente dos factores, que quedan fuera de nuestro estudio. Como ya apuntábamos en 3.3 la situación contextual de los agentes es fundamental en la interpretación del mensaje que envían las expresiones faciales, y el otro factor básico es el sonido que acompaña a la secuencia audiovisual, y que como indica Chion[114] constituye el 50% de la misma; concretamente expresiones como la risa falsa o la risa reprimida tienen asociado un sonido característico que las identifica plenamente.

Continuando con el análisis de los resultados, a las expresiones de risa congraciante y ávida les siguen la falsa carcajada y la falsa sonrisa, cuyo descriptor indica que la falsedad de la expresión es "intencionada", es decir el agente "finge" la expresión, en la mayoría de contextos esta expresión sería

continuada con una expresión de desagrado que completaría su significado. Sin ella, estas dos expresiones son necesariamente ambiguas. En la parte central de la valoración de expresiones se encuentran la carcajada, la sonrisa melancólica, la desleal, la carcajada extrema, la sonrisa embarazosa y la risa reprimida todas ellas con valores promedio que van desde el 3.2 al 3.4. Esto incluye la mayoría de las sonrisas a las que Faigin se refiere como sonrisas complejas, siendo las dos mejor puntuadas expresiones fácilmente describibles, la sonrisa embarazosa o con "ojos cerrados", "avergonzada", "humillada" y la risa reprimida con el descriptor "quiere pero no puede". En la parte superior de la tabla aparecen las dos expresiones de la sonrisa estándar con diferente intensidad, estas dos expresiones no tienen más descriptor que el físico: la sonrisa de mayor intensidad enseña los dientes la otra no. Finalmente la expresión mejor valorada es la sonrisa astuta, cuyo descriptor "maliciosa" o "con mala intención" ha resultado suficientemente claro a los encuestados.

En resumen, aún con el error en la secuencia de la expresión comentada anteriormente, los encuestados validan como aceptable el movimiento generado con nuestro método y creen que las expresiones representan suficientemente el objetivo fijado en cada una de ellas.

Capítulo 9 Conclusiones y trabajo futuro

En el desarrollo de esta tesis doctoral se han realizado diferentes aproximaciones que analizaremos por separado a la hora de describir las conclusiones.

1. En el proceso de la caracterización de la risa, hemos comprobado que los estudios que pertenecen al área de la psicología se interesan en temas como la distinción de sonrisas reales frente a las actuadas, las sonrisas referidas a diferentes estados emocionales y significado de las sonrisas. En los tratados de anatomía consultados hemos visto cómo se describen los músculos, su función y el resultado de su acción, pero en ellos no se encuentran ningún tipo de clasificación de sonrisas. En los manuales artísticos hemos encontrado un compendio de las diferentes técnicas de reproducción de expresiones faciales y en la obra de Faigin[16] hemos hallado una clasificación exhaustiva de diferentes expresiones faciales, entre ellas, un capítulo completo dedicado a expresiones realizadas con la intervención del músculo cigomático mayor. Como resultado del estudio pormenorizado y transdisciplinar realizado, hemos tomado esta clasificación como base para reproducción de las expresiones que hemos llevado a cabo. En este aspecto nuestra aportación ha consistido en completar el trabajo de Faigin añadiendo la acción específica de cada músculo que interviene en cada una de las expresiones.
2. En el ámbito de las Técnicas de representación facial 3D, se ha realizado un estudio sobre el estado del arte de la animación facial y de todos los procesos que involucra. Se han relacionado las diferentes técnicas profesionales que usan los artistas actualmente. Basándonos en este estudio, se ha optado por la elección de un sistema de *rigging* para el control de la interpolación geométrica de veintiséis superficies poligonales que se corresponden con la agrupación de acciones musculares capaces de combinarse para reproducir las AUs descritas por Ekman en su FACS[12]. En nuestro trabajo, la principal aportación es la aplicación de esta técnica como método para validar un sistema de animación facial realista.
3. En la síntesis de las trece expresiones risa o sonrisa, se han generado

manualmente las expresiones, tomando las indicaciones de la actividad muscular y las imágenes de referencia de la bibliografía. Para validar estas expresiones se ha diseñado y realizado una breve encuesta en la que se tenía que relacionar las imágenes del personaje virtual con las imágenes estáticas dibujadas por Fagin y fotogramas de películas en los que actores realizan cada una de las expresiones. Estas películas han sido escogidas mediante otra encuesta dónde se han votado las escenas emblemáticas con las expresiones de nuestro interés. Se ha demostrado la dificultad que existe a la hora de identificar expresiones a partir de imágenes estáticas, así como relacionar las mismas expresiones entre diferentes personajes, ya sean reales, dibujados o generados por ordenador.

4. En la asignación automática de las curvas de animación a los controladores del rig facial, hemos aportado un sistema que a partir de una captura del movimiento facial, convierte estos datos en información para la animación del rig. Se ha realizado este proceso mediante el Face Tacker de Saragih[98] que puede aplicarse a cualquier secuencia de vídeo independientemente de la cámara con que haya sido grabada. Estas curvas pueden ser importadas en un entorno de creación 3D y se ha demostrado que incluso pueden asignarse a estructuras *rig* diferentes a la usada en nuestro caso. La integración de esta información en el entorno 3D es tal que permite su posterior refinado y modificación por parte del animador.
5. En la aplicación del método sobre secuencias cinematográficas, se ha probado el método analizando las secuencias cinematográficas seleccionadas anteriormente. Se ha comprobado de manera visual la correspondencia entre expresiones y se ha analizado la problemática inherente a la recogida de datos de estas secuencias.
6. La aplicación del método sobre bases de datos de expresiones faciales y específicas de risas se ha probado utilizando las bases de datos más reconocidas (Cohn-Kanade, Uva-Nemo y MANHOB). Los resultados son similares en los tres casos y se han reproducido diferentes secuencias de la base de datos obteniendo las expresiones que ejecutan nuestros avatares sintéticos calcando los movimientos de los sujetos. En las secuencias de las tres bases de datos se verifica la ausencia de problemas como el encuadre o iluminación que se encuentran en las secuencias cinematográficas y una clasificación entre sonrisas reales y fingidas. Por el contrario, en todas estas bases de datos públicas no se encuentra ninguna clasificación de otros tipos de risa.
7. Con todos los datos anteriores se ha generado una base de datos con

las trece expresiones de risa o sonrisa. Aportamos la generación de esta base de datos con la interpretación llevada a cabo por diferentes sujetos de las trece expresiones descritas por Faigin. Para una validación exhaustiva del método con nuestra propia base de datos, se ha aplicado el método a todas las expresiones de cada sujeto de la base de datos reproduciendo sus expresiones faciales con los diferentes avatares considerados en esta tesis.

8. Uno de los resultados más significativos ha sido la generación de las curvas más representativas de cada expresión. Se han analizado las curvas de animación de los diferentes controladores de todos los sujetos y se han elegido las más representativas para cada expresión. De manera que nuestro estudio aporta una librería de curvas de animación representativas para animar las expresiones de risa y sonrisa en personajes sintéticos. Mejorando mucho la eficiencia del proceso de animación.
9. Con el objetivo de validar las animaciones más representativas de cada expresión se ha llevado a cabo un test con una valoración de la calidad del movimiento y de la representación de la expresión de interés con resultados favorables.

A partir de nuestro estudio se han abierto diferentes líneas que podrían definir diferentes ampliaciones del trabajo.

1. La generalización para las demás expresiones faciales. El sistema de *rig* permite la generación de las expresiones faciales descritas por el sistema de las AUs de Ekman que además se hallan convenientemente clasificadas en el trabajo de Faigin. Nuestro sistema podría generalizarse creando una librería completa de animaciones predefinidas atendiendo a la nomenclatura usada en la bibliografía, esto daría a los animadores una librería completa de expresiones faciales predefinidas que podría agilizar sus tareas de animación.
2. La integración del sistema de captura. Actualmente el sistema de captura consiste en la ejecución de la aplicación de código abierto Facetracker de Saragih sobre una secuencia de vídeo previamente grabada, o bien con los datos grabados usando la cámara integrada en el ordenador. Los datos procedentes de esta aplicación son tratados y enviados a una aplicación propia que los convierte en valores de curvas de animación que pueden ser importadas y asignadas a los controladores del *rig* en el entorno de la aplicación Maya Autodesk. Este sistema es susceptible de ser integrado en cualquier API 3D, ya sea de código abierto como Blender o mediante la escritura de los *plugins* pertinentes en cualquier otro.

3. La ampliación de la base de datos. La base de datos que hemos desarrollado puede ampliarse grabando y analizando las expresiones de risa y sonrisa de un número mayor de individuos, lo que nos daría mayor robustez a la hora de ofrecer nuestra librería de expresiones predefinidas y por otra parte podríamos llevar a cabo la grabación de las secuencias de las demás expresiones faciales descritas por Faigin, esta base de datos proporcionaría una información muy útil no solo a animadores sino a artistas de todo tipo interesados en la reproducción de características faciales.
4. Creación de una herramienta para la animación procedural de expresiones de risa y sonrisa. A partir de las curvas más representativas para la animación de los controladores faciales se puede desarrollar una herramienta integrada en un API 3D que permita la elección de expresiones que puedan asignarse a diferentes avatares con un *rig* compatible.

Apéndice 1. Tablas con los resultados del algoritmo FastDTW

Las tablas aquí presentadas son el resultado de la aplicación del algoritmo de FastDTW a los datos de los controladores del *rig* facial para las 13 expresiones realizadas ejecutadas por los 7 sujetos de nuestra base de datos, comparados dos a dos. En ellas los *outliers* han sido convenientemente eliminados. Se muestran las tablas correspondientes a los controladores *SmileFrown*, *Jaw*, *UpsetSad - Shocked* y *LeftBlink* usados para la elección de las curvas de animación más representativas de cada expresión.

	s01	s02	s04	s05	s06	s07
s01		59.92	111.14	49.33	8.23	72.69
s02	59.92		28.98	87.96	62.06	18.74
s04	111.14	28.98		78.10	106.85	27.68
s05	49.33	87.96	78.10		50.83	102.50
s06	8.23	62.06	106.85	50.83		64.44
s07	72.69	18.74	27.68	102.50	64.44	
PROMEDIO	60.26	51.53	70.55	73.74	58.48	57.21
DES. EST	37.32	27.80	40.58	23.29	35.26	34.26

Tabla 13: Controlador: *SmileFrown* / Expresión: Carjada extrema

	s01	s03	s04	s05	s06	s07
s01		8.33	6.99	52.02	32.61	47.71
s03	8.33		2.92	48.84	33.69	45.19
s04	6.99	2.92		47.48	35.38	45.25
s05	52.02	48.84	47.48		50.07	52.22
s06	32.61	33.69	35.38	50.07		11.05
s07	47.71	45.19	45.25	52.22	11.05	
PROMEDIO	29.53	27.79	27.61	50.13	32.56	40.28
DES. EST	21.23	21.08	21.22	2.04	13.95	16.59

Tabla 14: Controlador: *SmileFrown* / Expresión: Carcajada

	s01	s02	s04	s05	s06	s07
s01		12.85	27.82	16.68	36.82	11.31
s02	12.85		10.90	8.66	16.17	9.16
s04	27.82	10.90		4.78	4.30	13.42
s05	16.68	8.66	4.78		9.56	12.43
s06	36.82	16.17	4.30	9.56		13.66
s07	11.31	9.16	13.42	12.43	13.66	
PROMEDIO	21.10	11.55	12.24	10.42	16.10	12.00
DES. EST	10.91	3.06	9.55	4.44	12.43	1.84

Tabla 15: Controlador: *SmileFrown* / Expresión: Risa clamorosa

	s01	s02	s03	s04	s05	s06	s07
s01		4.15	22.91	9.23	18.83	15.94	11.67
s02	4.15		24.80	20.96	18.01	17.44	10.58
s03	22.91	24.80		51.81	24.44	4.24	26.63
s04	9.23	20.96	51.81		34.84	32.17	19.37
s05	18.83	18.01	24.44	34.84		23.22	26.14
s06	15.94	17.44	4.24	32.17	23.22		28.21
s07	11.67	10.58	26.63	19.37	26.14	28.21	
PROMEDIO	13.79	15.99	25.80	28.06	24.25	20.20	20.43
DES. EST	6.80	7.46	15.19	14.88	6.08	9.97	7.83

Tabla 16: Controlador: *SmileFrown* / Expresión: Sonrisa boca cerrada

	s01	s02	s03	s04	s05	s06	s07
s01		121.15	45.00	130.36	21.55	155.05	96.29
s02	121.15		44.582	45.37	112.04	40.75	32.51
s03	45.00	44.58		40.79	61.16	60.02	25.41
s04	130.36	45.37	40.79		95.78	73.25	36.97
s05	21.55	112.04	61.16	95.78		133.50	89.85
s06	155.05	40.75	60.02	73.25	133.50		24.40
s07	96.29	32.51	25.41	36.97	89.85	24.40	
PROMEDIO	94.90	66.07	46.16	70.42	85.65	81.16	50.91
DES. EST	51.84	39.51	13.27	37.06	39.54	52.08	33.05

Tabla 17: Controlador: *SmileFrown* / Expresión: Risa sofocada

	s01	s02	s03	s04	s05	s06	s07
s01		21.26	10.55	22.94	50.32	47.57	14.34
s02	21.26		17.41	60.68	88.82	16.71	6.33
s03	10.55	17.41		32.09	55.74	43.84	7.71
s04	22.94	60.68	32.09		24.56	80.31	44.78
s05	50.32	88.82	55.74	24.56		111.30	72.57
s06	47.57	16.71	43.84	80.31	111.30		26.50
s07	14.34	6.33	7.71	44.78	72.57	26.50	
PROMEDIO	27.83	35.20	27.89	44.23	67.22	54.37	28.71
DES. EST	16.99	32.28	19.32	22.63	30.58	35.40	25.83

Tabla 18: Controlador: *SmileFrown* / Expresión: Sonrisa melancólica

	s01	s02	s03	s05	s06	s07
s01		5.49	4.06	5.46	5.46	42.49
s02	5.49		7.13	4.09	4.02	30.74
s03	4.06	7.13		13.50	4.27	37.06
s05	9.67	4.09	13.50		9.78	52.96
s06	5.46	4.02	4.27	9.78		42.46
s07	42.49	30.74	37.06	52.96	42.46	
PROMEDIO	13.43	10.29	13.20	17.16	13.20	41.14
DES. EST	16.38	11.50	13.87	20.35	16.52	8.18

Tabla 19: Controlador: *SmileFrown* / Expresión: Sonrisa Ávida

	s01	s02	s03	s04	s05	s06	s07
s01		29.47	57.03	43.43	59.30	60.36	57.54
s02	29.47		44.72	24.57	38.07	43.15	23.34
s03	57.03	44.72		89.38	8.80	17.70	33.64
s04	43.43	24.57	89.38		77.60	81.99	39.63
s05	59.30	38.07	8.80	77.60		12.07	19.70
s06	60.36	43.15	17.70	81.99	12.07		45.72
s07	57.54	23.34	33.64	39.63	19.70	45.72	
PROMEDIO	51.19	33.89	41.88	59.43	35.92	43.50	36.60
DES. EST	12.30	9.36	29.12	26.83	27.77	26.17	14.15

Tabla 20: Controlador: *SmileFrown* / Expresión: Sonrisa Congraciante

	s01	s02	s03	s04	s05	s06
s01		101.03	11.38	22.49	41.64	5.13
s02	101.03		130.63	24.87	6.76	98.81
s03	11.38	130.63		37.75	56.70	18.82
s04	22.49	24.87	37.75		20.99	22.26
s05	41.64	6.76	56.70	20.99		45.11
s06	5.13	98.81	18.82	22.26	45.11	
PROMEDIO	36.34	72.42	51.06	25.67	34.24	38.03
DES. EST	38.74	53.56	47.84	6.89	20.05	36.89

Tabla 21: Controlador: *SmileFrown* / Expresión: Sonrisa astuta

	s01	s02	s03	s04	s05	s06
s01		13.12	36.14	18.91	48.57	52.77
s02	13.12		47.85	10.09	52.22	62.72
s03	36.14	47.85		27.85	52.42	14.29
s04	18.91	10.09	27.85		56.17	58.62
s05	48.57	52.22	52.42	56.17		51.78
s06	52.77	62.72	14.29	58.62	51.78	
PROMEDIO	33.90	37.20	35.71	34.33	52.23	48.04
DES. EST	17.56	24.01	15.39	21.99	2.70	19.38

Tabla 22: Controlador: *SmileFrown* / Expresión: Sonrisa desleal

	s01	s03	s04	s05	s06	s07
s01		127.40	36.28	50.73	35.03	34.19
s03	127.40		75.16	112.80	26.11	118.50
s04	36.28	75.16		96.80	<u>38.73</u>	94.21
s05	50.73	112.80	96.80		74.13	37.90
s06	35.03	26.11	<u>38.73</u>	74.13		75.37
s07	34.19	118.50	94.21	37.90	75.37	
PROMEDIO	56.72	91.99	68.23	74.47	49.87	72.03
DES. EST	40.08	41.87	29.28	31.10	23.17	36.26

Tabla 23: Controlador: *SmileFrown* / Expresión: Sonrisa embarazosa

	s01	s02	s03	s04	s05	s06	s07
s01		14.93	6.77	4.32	4.26	5.92	14.46
s02	14.93		20.97	10.41	10.92	15.19	15.23
s03	6.77	20.97		16.55	14.89	13.26	13.90
s04	4.32	10.41	16.55		4.46	15.93	10.27
s05	4.26	10.92	14.89	4.46		11.15	17.49
s06	5.92	15.19	13.26	15.93	11.15		34.82
s07	14.46	15.23	13.90	10.27	17.49	34.82	
PROMEDIO	8.44	14.61	14.39	10.32	10.53	16.05	17.70
DES. EST	4.94	3.81	4.64	5.31	5.37	9.88	8.71

Tabla 24: Controlador: *SmileFrown* / Expresión: Falsa sonrisa

	s01	s03	s04	s05	s06	s07
s01		51.94	60.69	34.31	37.98	16.15
s03	51.94		168.52	44.96	87.91	142.78
s04	60.69	168.52		132.39	98.87	61.80
s05	34.31	44.96	132.39		55.13	105.60
s06	37.98	87.91	98.87	55.13		74.68
s07	16.15	142.78	61.80	105.60	74.68	
PROMEDIO	40.21	99.22	104.45	74.48	70.91	80.20
DES. EST	17.14	54.79	46.50	42.37	24.59	47.52

Tabla 25: Controlador: *SmileFrown* / Expresión: Falsa carcajada

	s01	s02	s03	s04	s05	s06	s07
s01		17.45	70.60	66.35	61.11	15.20	15.20
s02	17.45		94.89	73.72	62.44	8.61	41.15
s03	70.60	94.89		48.38	34.69	81.45	116.59
s04	66.35	73.72	48.38		39.19	66.23	111.35
s05	61.11	62.44	34.69	39.19		54.22	94.21
s06	15.20	8.61	81.45	66.23	54.22		40.57
s07	23.58	41.15	116.59	111.35	94.21	40.57	
PROMEDIO	42.38	49.71	74.43	67.54	57.64	44.38	69.85
DES. EST	26.21	33.42	30.06	25.03	21.22	28.61	42.82

Tabla 26: Controlador: *Jaw* / Expresión: Carcajada extrema

	s01	s02	s03	s04	s05	s06	s07
s01		45.88	23.19	24.91	34.50	34.30	41.98
s02	45.88		39.32	36.81	24.55	15.07	40.91
s03	23.19	39.32		3.45	18.07	23.18	9.84
s04	24.91	36.81	3.45		19.95	19.67	6.78
s05	34.50	24.55	18.07	19.95		23.50	27.85
s06	34.30	15.07	23.18	19.67	23.50		19.55
s07	41.98	40.91	9.84	6.78	27.85	19.55	
PROMEDIO	34.13	33.75	19.51	18.60	24.74	22.54	24.48
DES. EST	9.00	11.60	12.43	12.19	5.90	6.52	15.10

Tabla 27: Controlador: *Jaw* / Expresión: Carcajada

	s01	s02	s03	s04	s05	s06	s07
s01		9.19	7.24	6.95	4.60	10.34	3.51
s02	9.19		13.73	8.55	3.45	2.48	8.99
s03	7.24	13.73		19.26	7.69	14.17	3.15
s04	6.95	8.55	19.26		6.01	5.59	15.49
s05	4.60	3.45	7.69	6.01		2.84	4.77
s06	10.34	2.48	14.17	5.59	2.84		10.38
s07	3.51	8.99	3.15	15.49	4.77	10.38	
PROMEDIO	6.97	7.73	10.87	10.31	4.89	7.63	7.71
DES. EST	2.61	4.15	5.87	5.69	1.76	4.72	4.82

Tabla 28: Controlador: *Jaw* / Expresión: Risa clamorosa

	s01	s02	s03	s04	s05	s06	s07
s01		7.14	2.14	6.21	2.43	4.31	3.33
s02	7.14		6.07	2.58	8.58	2.08	3.53
s03	2.14	6.07		1.70	0.87	1.39	1.39
s04	6.21	2.58	1.70		3.75	1.30	2.42
s05	2.43	8.58	0.87	3.75		2.36	1.36
s06	4.31	2.08	1.39	1.30	2.36		1.43
s07	3.33	3.53	1.39	2.42	1.36	1.43	
PROMEDIO	4.26	5.00	2.26	2.99	3.22	2.14	2.24
DES. EST	2.04	2.65	1.91	1.79	2.81	1.14	1.01

Tabla 29: Controlador: *Jaw* / Expresión: Sonrisa boca cerrada

	s01	s02	s04	s05	s06	s07
s01		12.80	0.92	5.98	4.72	8.46
s02	12.80		12.82	3.29	5.69	3.42
s04	0.92	12.82		6.37	4.30	8.16
s05	5.98	3.29	6.37		3.66	1.08
s06	4.72	5.69	4.30	3.66		6.29
s07	8.46	3.42	8.16	1.08	6.29	
PROMEDIO	6.58	7.60	6.52	4.08	4.93	5.48
DES. EST	4.42	4.85	4.44	2.16	1.06	3.18

Tabla 30: Controlador: *Jaw* / Expresión: Risa sofocada

	s01	s02	s03	s04	s05	s06	s07
s01		1.91	1.61	4.41	7.06	1.95	7.80
s02	1.91		4.07	2.08	4.94	0.71	7.10
s03	1.61	4.07		8.26	16.90	4.51	12.68
s04	4.41	2.08	8.26		1.89	3.08	1.45
s05	7.06	4.94	16.90	1.89		6.09	2.79
s06	1.95	0.71	4.51	3.08	6.09		8.16
s07	7.80	7.10	12.68	1.45	2.79	8.16	
PROMEDIO	4.12	3.47	8.00	3.53	6.61	4.08	6.36
DES. EST	2.76	2.35	5.83	2.55	5.40	2.75	4.46

Tabla 31: Controlador: *Jaw* / Expresión: Sonrisa melancólica

	s01	s02	s03	s05	s06	s07
s01		8.16	14.12	5.74	22.80	14.77
s02	8.16		11.05	4.20	16.77	9.40
s03	14.12	11.05		17.64	5.64	2.38
s05	5.74	4.20	17.64		28.36	18.38
s06	22.80	16.77	5.64	28.36		6.54
s07	14.77	9.40	2.38	18.38	6.54	
PROMEDIO	13.12	9.92	10.17	14.86	16.02	10.29
DES. EST	6.64	4.59	6.19	9.99	9.95	6.38

Tabla 32: Controlador: *Jaw* / Expresión: Sonrisa ávida

	s01	s02	s03	s04	s05	s06	s07
s01		6.36	18.36	27.15	13.07	15.15	14.25
s02	6.36		25.23	38.83	12.86	22.48	20.26
s03	18.36	25.23		10.19	7.37	5.34	3.43
s04	27.15	38.83	10.19		23.36	8.02	10.43
s05	13.07	12.86	7.37	23.36		10.66	4.95
s06	15.15	22.48	5.34	8.02	10.66		2.29
s07	14.25	20.26	3.43	10.43	4.95	2.29	
PROMEDIO	15.72	21.00	11.65	19.66	12.05	10.66	9.27
DES. EST	6.85	11.13	8.45	12.23	6.39	7.29	7.05

Tabla 33: Controlador: *Jaw* / Expresión: Sonrisa congraciante

	s01	s03	s04	s05	s06	s07
s01		4.00	2.14	36.64	3.83	8.22
s03	4.00		3.28	41.10	1.20	4.87
s04	2.14	3.28		34.12	3.33	4.44
s05	36.64	41.10	34.12		39.95	28.65
s06	3.83	1.20	3.33	39.95		5.21
s07	8.22	4.87	4.44	28.65	5.21	
PROMEDIO	10.97	10.89	9.46	36.09	10.70	10.28
DES. EST	14.53	16.94	13.81	4.99	16.41	10.38

Tabla 34: Controlador: *Jaw* / Expresión: Sonrisa astuta

	s01	s02	s03	s04	s05	s06	s07
s01		18.77	39.62	3.23	26.97	20.25	11.92
s02	18.77		3.23	24.34	2.49	1.73	9.68
s03	39.62	3.23		45.12	7.16	4.37	19.76
s04	3.23	24.34	45.12		30.79	25.78	17.22
s05	26.97	2.49	7.16	30.79		1.60	12.89
s06	20.25	1.73	4.37	25.78	1.60		9.09
s07	11.92	9.68	19.76	17.22	12.89	9.09	
PROMEDIO	20.13	10.04	19.87	24.41	13.65	10.47	13.43
DES. EST	12.50	9.52	18.47	13.94	12.52	10.24	4.24

Tabla 35: Controlador: *Jaw* / Expresión: Sonrisa desleal

	s01	s03	s04	s05	s06	s07
s01		2.39	4.78	2.77	1.34	2.01
s03	2.39		4.74	3.14	1.30	3.14
s04	4.78	4.74		4.03	6.15	3.19
s05	2.77	3.14	4.03		4.02	2.05
s06	1.34	1.30	6.15	4.02		3.06
s07	2.01	3.14	3.19	2.05	3.06	
PROMEDIO	2.66	2.94	4.58	3.20	3.17	2.69
DES. EST	1.30	1.26	1.09	0.85	2.03	0.60

Tabla 36: Controlador: *Jaw* / Expresión: Sonrisa embarazosa

	s01	s02	s03	s04	s06	s07
s01		21.60	26.29	23.71	18.33	20.75
s02	21.60		6.02	2.55	3.14	3.36
s03	26.29	6.02		3.15	10.24	11.10
s04	23.71	2.55	3.15		7.05	5.40
s06	18.33	3.14	10.24	7.05		1.74
s07	20.75	3.36	11.10	5.40	1.74	
PROMEDIO	22.14	7.33	11.36	8.37	8.10	8.47
DES. EST	3.02	8.09	8.95	8.76	6.62	7.72

Tabla 37: Controlador: *Jaw* / Expresión: Falsa sonrisa

	s01	s02	s03	s04	s05	s07
s01		25.33	28.44	13.87	8.77	10.56
s02	25.33		8.14	24.51	10.27	17.29
s03	28.44	8.14		24.51	29.01	48.76
s04	13.87	24.51	24.51		10.27	17.29
s05	8.77	10.27	29.01	10.27		15.20
s07	10.56	17.29	48.76	17.29	15.20	
PROMEDIO	17.40	17.11	27.77	18.09	14.71	21.82
DES. EST	8.92	7.90	14.48	6.36	8.36	15.31

Tabla 38: Controlador: *Jaw* / Expresión: Falsa carcajada

	s01	s02	s04	s05	s06	s07
s01		12.13	135.02	56.75	16.66	45.88
s02	12.13		174.06	73.02	20.47	100.09
s04	135.02	174.06		122.25	52.58	73.95
s05	56.75	73.02	122.25		52.58	73.95
s06	16.66	20.47	52.58	52.58		58.48
s07	45.88	100.09	73.95	73.95	58.48	
PROMEDIO	53.29	75.96	111.57	75.71	40.15	70.47
DES. EST	49.45	65.89	48.64	27.71	19.90	20.31

Tabla 39: Controlador: *UpsetSad* y *Shocked* / Expresión: Carcajada extrema

	s01	s02	s03	s04	s05	s06	s07
s01		24.63	39.57	59.93	22.21	38.74	41.46
s02	24.63		25.26	33.70	17.08	23.36	23.07
s03	39.57	25.26		44.48	19.15	13.98	29.71
s04	59.93	33.70	44.48		45.90	29.99	26.61
s05	22.21	17.08	19.15	45.90		18.08	34.55
s06	38.74	23.36	13.98	29.99	18.08		26.67
s07	41.46	23.07	29.71	26.61	34.55	26.67	
PROMEDIO	37.76	24.52	28.69	40.10	26.16	25.14	30.34
DES. EST	13.59	5.36	11.73	12.42	11.60	8.81	6.66

Tabla 40: Controlador: *UpsetSad* y *Shocked* / Expresión: Carcajada

	s01	s02	s03	s04	s05	s06	s07
s01		15.76	7.89	15.84	13.24	7.12	12.46
s02	15.76		39.84	45.08	6.70	11.58	26.29
s03	7.89	39.84		14.79	23.63	16.01	15.26
s04	15.84	45.08	14.79		31.95	20.95	22.28
s05	13.24	6.70	23.63	31.95		10.95	17.87
s06	7.12	11.58	16.01	20.95	10.95		9.19
s07	12.46	26.29	15.26	22.28	17.87	9.19	
PROMEDIO	12.05	24.21	19.57	25.15	17.39	12.63	17.22
DES. EST	3.78	15.63	11.12	11.52	9.21	5.04	6.31

Tabla 41: Controlador: *UpsetSad* y *Shocked* / Expresión: Sonrisa clamorosa

	s01	s02	s03	s04	s05	s06	s07
s01		41.69	19.79	30.16	57.50	10.50	61.54
s02	41.69		63.89	31.15	18.70	62.43	41.96
s03	19.79	63.89		28.89	73.13	18.49	75.58
s04	30.16	31.15	28.89		36.30	32.48	30.73
s05	57.50	18.70	73.13	36.30		72.05	47.85
s06	10.50	62.43	18.49	32.48	72.05		61.24
s07	61.54	41.96	75.58	30.73	47.85	61.24	
PROMEDIO	36.86	43.30	46.63	31.62	50.92	42.86	53.15
DES. EST	20.44	17.59	27.07	2.58	21.19	25.78	16.09

Tabla 42: Controlador: *UpsetSad* y *Shocked* / Expresión: Sonrisa boca cerrada

	s01	s03	s04	s05	s06	s07
s01		134.66	114.17	71.60	110.96	116.89
s03	134.66		26.56	49.18	21.08	32.69
s04	114.17	26.56		39.29	13.79	27.62
s05	71.60	49.18	39.29		32.15	42.86
s06	110.96	21.08	13.79	32.15		22.38
s07	116.89	32.69	27.62	42.86	22.38	
PROMEDIO	109.66	52.83	44.29	47.01	40.07	48.49
DES. EST	23.17	46.94	40.10	15.06	40.16	38.98

Tabla 43: Controlador: *UpsetSad* y *Shocked* / Expresión: Risa sofocada

	s02	s03	s04	s05	s06	s07
s02		27.08	14.65	31.46	22.14	51.41
s03	27.08		24.62	49.59	41.63	27.40
s04	14.65	24.62		43.58	33.52	24.99
s05	31.46	49.59	43.58		17.78	60.34
s06	22.14	41.63	33.52	17.78		56.67
s07	51.41	27.40	24.99	60.34	56.67	
PROMEDIO	29.35	34.06	28.27	40.55	34.35	44.16
DES. EST	13.82	10.96	10.86	16.46	15.62	16.73

Tabla 44: Controlador: *UpsetSad* y *Shocked* / Expresión: Sonrisa melancólica

	s01	s02	s03	s04	s05	s06	s07
s01		46.83	92.75	132.98	58.96	38.16	45.10
s02	46.83		23.02	48.21	84.76	50.60	48.17
s03	92.75	23.02		37.40	123.26	95.36	67.29
s04	132.98	48.21	37.40		162.91	139.54	116.71
s05	58.96	84.76	123.26	162.91		59.68	83.59
s06	38.16	50.60	95.36	139.54	59.68		54.75
s07	45.10	48.17	67.29	116.71	83.59	54.75	
PROMEDIO	69.13	50.26	73.18	106.29	95.53	73.02	69.27
DES. EST	36.80	19.76	37.98	51.48	40.47	37.82	27.20

Tabla 45: Controlador: *UpsetSad* y *Shocked* / Expresión: Sonrisa ávida

	s01	s02	s03	s05	s06	s07
s01		47.51	55.05	38.24	31.37	37.78
s02	47.51		68.82	38.66	64.96	13.68
s03	55.05	68.82		78.50	40.62	98.53
s05	38.24	38.66	78.50		39.16	25.25
s06	31.37	64.96	40.62	39.16		68.28
s07	37.78	13.68	98.53	25.25	68.28	
PROMEDIO	41.99	46.73	68.30	43.96	48.88	48.70
DES. EST	9.30	22.24	22.13	20.17	16.61	34.50

Tabla 46: Controlador: *UpsetSad* y *Shocked* / Expresión: Sonrisa congraciante

	s01	s02	s03	s04	s05	s06	s07
s01		181.35	110.59	58.24	82.02	82.72	29.62
s02	181.35		34.62	116.13	85.30	82.14	249.16
s03	110.59	34.62		54.78	66.36	40.83	157.60
s04	58.24	116.13	54.78		43.41	32.63	68.41
s05	82.02	85.30	66.36	43.41		29.83	90.11
s06	82.72	82.14	40.83	32.63	29.83		117.90
s07	29.62	249.16	157.60	68.41	90.11	117.90	
PROMEDIO	90.76	124.78	77.46	62.27	66.17	64.34	118.80
DES. EST	52.03	77.78	47.61	29.14	24.61	35.42	77.22

Tabla 47: Controlador: *UpsetSad* y *Shocked* / Expresión: Sonrisa astuta

	s01	s02	s03	s04	s05	s06	s07
s01		42.90	69.87	48.03	67.15	37.66	50.27
s02	42.90		57.74	28.83	50.90	43.84	29.97
s03	69.87	57.74		44.79	79.73	53.57	50.39
s04	48.03	28.83	44.79		55.67	38.32	37.05
s05	67.15	50.90	79.73	55.67		82.46	40.74
s06	37.66	43.84	53.57	38.32	82.46		40.11
s07	50.27	29.97	50.39	37.05	40.74	40.11	
PROMEDIO	52.65	42.36	59.35	42.12	62.78	49.33	41.42
DES. EST	13.07	11.39	13.07	9.41	16.56	17.25	7.89

Tabla 48: Controlador: *UpsetSad* y *Shocked* / Expresión: Sonrisa desleal

	s01	s02	s03	s04	s05	s06
s01		69.17	44.02	60.20	35.85	44.86
s02	69.17		107.80	44.30	107.07	39.70
s03	44.02	107.80		80.56	32.57	66.98
s04	60.20	44.30	80.56		74.53	52.85
s05	35.85	107.07	32.57	74.53		53.15
s06	44.86	39.70	66.98	52.85	53.15	
PROMEDIO	50.82	73.61	66.39	62.49	60.63	51.51
DES. EST	13.51	32.85	29.83	15.00	30.85	10.34

Tabla 49: Controlador: *UpsetSad* y *Shocked* / Expresión: Sonrisa embarazosa

	s01	s03	s04	s05	s06	s07
s01		41.09	30.34	30.91	31.02	30.04
s03	41.09		35.25	29.18	21.55	21.73
s04	30.34	35.25		25.70	28.05	42.66
s05	30.91	29.18	25.70		20.78	27.99
s06	31.02	21.55	28.05	20.78		26.42
s07	30.04	21.73	42.66	27.99	26.42	
PROMEDIO	32.68	29.76	32.40	26.91	25.57	29.77
DES. EST	4.72	8.53	6.73	3.91	4.35	7.83

Tabla 50: Controlador: *UpsetSad* y *Shocked* / Expresión: Falsa sonrisa

	s01	s02	s04	s05	s06	s07
s01		31.58	34.02	30.12	35.48	23.83
s02	31.58		34.86	34.40	18.38	36.98
s04	34.02	34.86		32.11	34.64	35.79
s05	30.12	34.40	32.11		35.99	30.68
s06	35.48	18.38	34.64	35.99		38.83
s07	23.83	36.98	35.79	30.68	38.83	
PROMEDIO	31.00	31.24	34.28	32.66	32.66	33.22
DES. EST	4.52	7.44	1.37	2.49	8.14	6.06

Tabla 51: Controlador: *UpsetSad* y *Shocked* / Expresión: Falsa carcajada

	s01	s02	s03	s05	s06	s07
s01		8.50	33.96	14.96	95.55	56.31
s02	8.50		42.33	2.12	139.93	105.28
s03	33.96	42.33		75.30	203.93	72.55
s05	14.96	2.12	75.30		137.81	120.08
s06	95.55	139.93	203.93	137.81		37.49
s07	56.31	105.28	72.55	120.08	37.49	
PROMEDIO	41.86	59.63	85.62	70.05	122.94	78.34
DES. EST	35.32	60.73	68.59	60.77	61.48	34.12

Tabla 52: Controlador: *LeftBlink* / Expresión: Carcajada extrema

	s01	s02	s03	s05	s06	s07
s01		29.61	40.91	30.88	25.65	37.61
s02	29.61		22.38	15.76	7.61	15.76
s03	40.91	22.38		14.84	15.93	14.84
s05	30.88	15.76	14.84		7.60	0.00
s06	25.65	7.61	15.93	7.60		7.60
s07	37.61	15.76	14.84	0.00	7.60	
PROMEDIO	32.93	18.22	21.78	13.82	12.88	15.16
DES. EST	6.20	8.24	11.14	11.46	8.00	14.07

Tabla 53: Controlador: *LeftBlink* / Expresión: Carcajada

	s01	s02	s03	s05	s06	s07
s01		9.65	1.27	14.88	9.65	8.80
s02	9.65		16.84	5.23	0.00	3.11
s03	1.27	16.84		22.07	16.84	16.98
s05	14.88	5.23	22.07		5.23	6.96
s06	9.65	0.00	16.84	5.23		3.11
s07	8.80	3.11	16.98	6.96	3.11	
PROMEDIO	8.85	6.97	14.80	10.88	6.97	7.79
DES. EST	4.87	6.54	7.89	7.42	6.54	5.70

Tabla 54: Controlador: *LeftBlink* / Expresión: Sonrisa clamorosa

	s01	s02	s03	s04	s06	s07
s01		0.00	14.57	11.56	3.81	13.83
s02	0.00		14.57	11.56	3.81	13.83
s03	14.57	14.57		15.53	14.06	14.85
s04	11.56	11.56	15.53		6.72	22.62
s06	3.81	3.81	14.06	6.72		16.94
s07	13.83	13.83	14.85	22.62	16.94	
PROMEDIO	8.75	8.75	14.72	13.60	9.07	16.41
DES. EST	6.49	6.49	0.54	5.93	6.08	3.70

Tabla 55: Controlador: *LeftBlink* / Expresión: Sonrisa boca cerrada

	s01	s03	s04	s05	s06	s07
s01		42.44	30.57	44.09	0.00	22.47
s03	42.44		38.05	25.68	42.44	16.52
s04	30.57	38.05		47.56	30.57	31.95
s05	44.09	25.68	47.56		44.09	34.52
s06	0.00	42.44	30.57	44.09		22.47
s07	22.47	16.52	31.95	34.52	22.47	
PROMEDIO	27.92	33.03	35.74	39.19	27.92	25.59
DES. EST	17.95	11.50	7.29	8.98	17.95	7.45

Tabla 56: Controlador: *LeftBlink* / Expresión: Risa sofocada

	s01	s02	s03	s04	s05	s06	s07
s01		124.22	57.93	40.47	106.21	76.64	83.95
s02	124.22		54.75	64.08	80.60	38.21	28.05
s03	57.93	54.75		49.89	66.01	56.04	73.65
s04	40.47	64.08	49.89		38.88	52.41	32.09
s05	106.21	80.60	66.01	38.88		44.33	25.64
s06	76.64	38.21	56.04	52.41	44.33		36.90
s07	83.95	28.05	73.65	32.09	25.64	36.90	
PROMEDIO	81.57	64.98	59.71	46.30	60.28	50.75	46.71
DES. EST	30.67	34.46	8.62	11.47	29.89	14.77	25.36

Tabla 57: Controlador: *LeftBlink* / Expresión: Sonrisa melancólica

	s01	s02	s03	s04	s06	s07	
s01			0.00	38.56	48.78	23.69	0.00
s02	0.00			38.56	48.78	23.69	0.00
s03	38.56	38.56			69.42	56.28	38.56
s04	48.78	48.78	69.42			36.64	48.78
s06	23.69	23.69	56.28	36.64			23.69
s07	0.00	0.00	38.56	48.78	23.69		
PROMEDIO	22.21	22.21	48.27	50.48	32.80	22.21	
DES. EST	22.15	22.15	14.09	11.82	14.27	22.15	

Tabla 58: Controlador: *LeftBlink* / Expresión: Sonrisa ávida

	s01	s02	s03	s05	s06	s07
s01		23.26	109.57	31.40	13.57	12.11
s02	23.26		166.07	47.38	27.74	22.97
s03	109.57	166.07		91.23	129.69	151.20
s05	31.40	47.38	91.23		26.06	23.52
s06	13.57	27.74	129.69	26.06		7.93
s07	12.11	22.97	151.20	23.52	7.93	
PROMEDIO	37.98	57.48	129.55	43.92	41.00	43.55
DES. EST	40.77	61.52	30.29	28.03	50.28	60.56

Tabla 59: Controlador: *LeftBlink* / Expresión: Sonrisa congraciante

	s02	s03	s04	s06	s07
s02		47.40	8.10	78.59	45.85
s03	47.40		53.51	80.01	29.53
s04	8.10	53.51		109.63	44.45
s06	78.59	80.01	109.63		100.24
s07	45.85	29.53	44.45	100.24	
PROMEDIO	44.99	52.61	53.92	92.12	55.02
DES. EST	28.84	20.91	42.00	15.30	31.04

Tabla 60: Controlador: *LeftBlink* / Expresión: Sonrisa astuta

	s01	s02	s03	s04	s05	s06	s07
s01		121.50	147.83	56.56	127.01	53.38	122.48
s02	121.50		81.36	104.49	35.05	153.63	6.63
s03	147.83	81.36		185.85	96.55	65.85	70.69
s04	56.56	104.49	185.85		126.11	110.82	111.12
s05	127.01	35.05	96.55	126.11		158.30	24.12
s06	53.38	153.63	65.85	110.82	158.30		137.43
s07	122.48	6.63	70.69	111.12	24.12	137.43	
PROMEDIO	104.79	83.78	108.02	115.83	94.52	113.23	78.74
DES. EST	39.77	54.89	48.28	41.71	54.07	44.90	54.14

Tabla 61: Controlador: *LeftBlink* / Expresión: Sonrisa desleal

	s01	s02	s03	s04	s05	s06	s07
s01		128.21	124.48	167.29	109.00	111.17	74.50
s02	128.21		42.12	60.10	34.85	34.54	109.29
s03	124.48	42.12		80.57	30.71	28.32	122.37
s04	167.29	60.10	80.57		66.19	76.72	141.06
s05	109.00	34.85	30.71	66.19		3.62	120.96
s06	111.17	34.54	28.32	76.72	3.62		118.78
s07	74.50	109.29	122.37	141.06	120.96	118.78	
PROMEDIO	119.11	68.18	71.43	98.66	60.89	62.19	114.49
DES. EST	30.29	40.69	44.43	44.40	46.52	47.23	22.16

Tabla 62: Controlador: *LeftBlink* / Expresión: Sonrisa embarazosa

	s01	s02	s04	s05	s06	s07
s01		7.90	16.53	16.53	5.09	16.53
s02	7.90		9.04	9.04	6.24	9.04
s04	16.53	9.04		0.00	17.30	0.00
s05	16.53	9.04	0.00		17.30	0.00
s06	5.09	6.24	17.30	17.30		17.30
s07	16.53	9.04	0.00	0.00	17.30	
PROMEDIO	12.52	8.25	8.57	8.57	12.65	8.57
DES. EST	5.59	1.23	8.47	8.47	6.39	8.47

Tabla 63: Controlador: *LeftBlink* / Expresión: Falsa sonrisa

	s01	s02	s03	s04	s05	s07
s01		29.47	28.71	8.08	58.00	28.18
s02	29.47		10.17	9.13	54.26	1.29
s03	28.71	10.17		3.99	26.20	8.88
s04	8.08	9.13	3.99		23.87	7.84
s05	58.00	54.26	26.20	23.87		59.11
s07	28.18	1.29	8.88	7.84	59.11	
PROMEDIO	30.49	20.86	15.59	10.58	44.29	21.06
DES. EST	17.81	21.36	11.11	7.68	17.69	23.52

Tabla 64: Controlador: *LeftBlink* / Expresión: Falsa carcajada

Apéndice 2. Test de validación de resultados

Validación de animación automática de expresiones de risa y sonrisa

Se ha desarrollado un método para facilitar la animación facial de personajes sintéticos creados por ordenador. Este método genera la animación de las expresiones faciales a partir de una base de datos de vídeos, donde diversos actores interpretan las expresiones de nuestro interés. Se pretende ofrecer una herramienta que simplifique la tarea del animador.

Con estos tests se quiere validar los resultados del método aplicados a la animación de expresiones faciales de risa y sonrisa.

Tiempo previsto 25 min.

Test de validación

Validación de la animación de 13 expresiones de risa y sonrisa.

Para ello se visualizarán 13 vídeos uno para cada expresión. En los vídeos aparece en primer lugar el nombre de la expresión en inglés igual que el que aparece en la hoja que se entrega a cada sujeto junto con un dibujo que representa la expresión, seguidamente desaparece el nombre de la expresión y en su lugar aparece el avatar en la posición neutra (sin expresión), después el avatar ejecuta la expresión y el último cuadro se congela unos segundos con la expresión en su grado máximo; después se repite una segunda vez para su mejor observación.

Para cada vídeo el sujeto debe contestar dos sencillas preguntas según una escala que va del 0 (muy poco) a 5 (mucho), estas preguntas son: ¿El movimiento te parece correcto? y ¿El movimiento representa la descripción? Se ha de recalcar que las preguntas se refieren al movimiento que realiza el avatar y nunca a la forma y/o textura del avatar.

Test de validación

Contesta a las preguntas según la escala donde 0 significa poco y 5 mucho.

1	Laughter: Risa de intensidad media	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5
2	Uproarious Laughter: Risa de intensidad máxima	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5
3	Smiling (Open-Mouthed): Sonrisa con la boca abierta menor intensidad que la risa	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5
4	Smiling (Closed-Mouth): Sonrisa con boca cerrada, baja intensidad	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5
5	Stifled Smile: Sonrisa reprimida, quiere pero no puede	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5
6	Melancholy Smile: Sonrisa melancólica, triste, de resignación	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5
7	Eager Smile: Sonrisa ávida, ansiosa, eufórica	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5
8	Ingratiating Smile: Sonrisa insinuante, congradiente, para captar la simpatía	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5
9	Sly Smile: Sonrisa maliciosa, astuta, mala intención	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5
10	Debauched Smile: Sonrisa desleal, perversa, corrupta	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5
11	Closed-Eye (Abashed) Smile: Sonrisa avergonzada, humillada, desconcertada	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5
12	False Smile: Falsa sonrisa intencionada	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5
13	False Laughter: Falsa risa intencionada	¿El movimiento te parece correcto?	0 1 2 3 4 5
		¿El movimiento representa la descripción?	0 1 2 3 4 5

Referencias

- [1] K. M. B. Bridges, "Emotional Development in Early Infancy," *Child Development*, vol. 3, no. 4, p. 341, Blackwell Publishing on behalf of the Society for Research in Child Development, 1932.
- [2] U. Hess, Martin Beaupré, and N. Cheung, "Who to whom and why -- cultural differences and similarities in the function of smiles," in *The smile: Forms, functions, and consequences*, 2002, p. -216.
- [3] W. Ruch and P. Ekman, *The Expressive Pattern of Laughter*. 2001.
- [4] G.-B. Duchenne, *The Mechanism of Human facial Expression*. 1990.
- [5] C. Darwin, *The Expressions of the Emotions in Man and Animals*. 1998.
- [6] R. R. Provine, "Laughing, Tickling, and the Evolution of Speech and Self," *Current Directions in Psychological Science* 2004 13: 215, vol. 13, no. 215, p. 218, 2004.
- [7] C. R. Gruner, *The game of humor: a comprehensive theory of why we laugh*. 2000.
- [8] J. Morreall, *Taking laughter seriously*. 1983.
- [9] H. Rothänger, G. Hauser, A. C. Capellini, and A. Guidotti, "Analysis of Laughter and Speech Sounds in Italian and German Students," *Naturwissenschaften*, vol. 85, no. 8, pp. 402, Springer-Verlag, 1998.
- [10] S. J. Haggbloom, R. Warnick, J. E. Warnick, and V. K. Jones, "The 100 Most Eminent Psychologists of the Twentieth Century," *Review of General Psychology*, no. 6, p. 152, 2002.
- [11] P. Ekman and W. V. Friesen, "The Repertoire of Nonverbal Behavior: Categories, Origins, Usage, and Coding," *Semiotica*, vol. 1, no. 1, p. 98, 1969.
- [12] P. Ekman and W. V. Friesen, "Facial Action Coding System: A Technique for the Measurement of Facial Movement," *Consulting Psychologists Press*, 1978.
- [13] J. A. Russell and J. M. F. Dols, *The psychology of facial expression*.

- 1997.
- [14] J. G. Iglesias, “Revelaciones anatómicas de Leonardo De Vinci.” [Online]. Available: <http://www.latinartmuseum.com/anatomica.htm>.
 - [15] G. B. Bridgman, *Bridgman’s Complete Guide to Drawing From Life: Over 1,000 Illustrations*. 1992.
 - [16] G. Faigin, *The artist’s complete guide to facial expression*. Watson-Guptill, 1990.
 - [17] B. Hogarth, *Drawing The Human Head*. 1965.
 - [18] W. Spalteholz, E. P. Tortella, and S. V. Pedrals, *Atlas de anatomia humana*. Editorial Labor, 1963.
 - [19] J. Noh and U. Neumann, “A Survey of Facial Modeling and Animation Techniques,” 1998.
 - [20] M. Radovan and L. Pretorius, “Facial animation in a nutshell: past, present and future,” in *Proceedings of the 2006 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries*, 2006, pp. 71–79.
 - [21] F. I. Parke, “Computer generated animation of faces,” in *Proceedings of the ACM annual conference 1*, 1972, vol. Siggraph 1, p. 457.
 - [22] F. I. Parke and K. Waters, *Computer Facial Animation*. 1996.
 - [23] J. Ahlberg, “CANDIDE-3 An Updated Parametrised Face,” 2001.
 - [24] S. M. Platt and N. I. Badler, “Animating facial expressions,” in *ACM SIGGRAPH computer graphics*, 1981, vol. 15, pp. 245–252.
 - [25] K. Waters, “A muscle model for animation three-dimensional facial expression,” *ACM SIGGRAPH Computer Graphics*, vol. 21, no. 4, p. 24, 1987.
 - [26] T. W. Sederberg and S. R. Parry, “Free-form deformation of solid geometric models,” *ACM SIGGRAPH Computer Graphics*, vol. 20, no. 4, p. 160, 1986.
 - [27] P. Kalra, A. Mangili, N. M. Thalmann, and D. Thalmann, “3D Interactive Free Form Deformations for Facial Expressions,” in *Compugraphics ’91 Proceedings*, 1991, p. 141.
 - [28] S. Pasquariello and C. Pelachaud, “Greta: A Simple Facial Animation Engine,” in *6th On-Line World Conference on Soft Computing in Industrial Applications, Session on Soft Computing for Intelligent 3D Agents*, 2001.

- [29] T. D. Bui, "Creating Emotions And Facial Expressions For Embodied Agents." 2004.
- [30] T. Beier and S. Neely, "Feature-Based Image Metamorphosis," in *Proceedings of SIGGRAPH 1992*, 1992, p. 42.
- [31] F. Pighin, J. Hecker, D. Lischinski, R. Szeliski, and D. H. Salesin, "Synthesizing Realistic Facial Expressions from Photographs," in *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, 1998, p. 84.
- [32] M. Oka, K. Tsutsui, and A. Ohba, "Real-Time Manipulation of Texture-Mapped Surfaces," in *ACM Computer Graphics*, 1987, vol. 21, p. 188.
- [33] R. Kalra and N. Magnenat-Thalmann, "Modeling of Vascular Expressions in Facial Animation," in *Computer Animation '94*, 1994, p. 58.
- [34] C. Bregler, M. Covell, and M. Slaney, "Video Rewrite: Driving Visual Speech with Audio," in *Proceedings of SIGGRAPH 1997*, 1997, p. 397.
- [35] T. Ezzat, G. Geiger, and T. Poggio, "Trainable Videorealistic Speech Animation," in *In Proceedings of SIGGRAPH 2002*, 2002, p. 397.
- [36] E. Cosatto and H. P. Graf, "Photo-Realistic Talking-Heads from Image Samples," in *Transactions on Multimedia*, 2000, vol. 2, p. 163.
- [37] C. Kouadio, P. Poulin, and P. Lachapelle, "Real-Time Facial Animation based upon a Bank of 3D Facial Expressions," in *Proceedings of Computer Animation '98 Conference*, 1998.
- [38] P. K. Doenges, T. K. Capin, F. Lavagetto, J. Ostermann, I. S. Pandzic, and E. D. Petajan, "MPEG-4: Audio/video and synthetic graphics/audio for mixed media," in *Signal Processing: Image Communication*, 1997, vol. 5, p. 463.
- [39] E. Sifakis, A. Selle, A. Robinson-Mosher, and R. Fedkiw, "Simulating Speech with a Physics-Based Facial Muscle Model," *Eurographics/ACM SIGGRAPH Symposium on Computer Animation*, 2006.
- [40] V. Orvalho, P. Bastos, F. Parke, B. Oliveira, and X. Alvarez, "A Facial Rigging Survey," 2012, pp. 183–204.
- [41] C. Maraffi, *Maya Character Creation: Modeling and Animation Controls*. New Riders, 2004.
- [42] J. Gorden, *LightWave 3D 8 Cartoon Character Creation: Modeling & Texturing*. Wordware Publishing, Inc., 2004.
- [43] K. Ritchie, J. Callery, and K. Biri, *The Art of Rigging*. CG Toolkit, 2005.

- [44] E. Miller, P. Thuriot, and J. Unay, *Maya Techniques: Hyper-Real Creature Creation*, 2006th ed. Sybex, 2006.
- [45] M. McKinley, *The Game Animator's Guide to Maya*, Pap/Cdr. John Wiley & Sons, 2006.
- [46] R. Bredow, D. Schaub, D. Kramer, M. Hausman, D. Dimian, and R. S. Duguid, "Surf's up: the making of an animated documentary," in *ACM SIGGRAPH 2007 courses*, New York, NY, USA, 2007, pp. 1–123.
- [47] E. Chang and O. Jenkins, "Sketching Articulation and Pose for Facial Animation," in *Data-Driven 3D Facial Animation*, Z. Deng and U. Neumann, Eds. Springer London, 2007, pp. 145–161.
- [48] L. Pardew, *Character Emotion in 2D and 3D Animation*, 1st ed. Boston, MA, United States: Course Technology Press, 2007.
- [49] R. O'Neill, *Digital Character Development: Theory and Practice*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2008.
- [50] S. Villagrassa and A. Susín, "FACe! 3D facial Animation System based on FACS," F. S. O. Rodriguez, Ed. 2009, p. 209, IV Iberoamerican Symposium in Computer Graphics – SIACG 2009.
- [51] O. Alexander, M. Rogers, W. Lambeth, M. Chiang, and P. Debevec, "Creating a Photoreal Digital Actor: The Digital Emily Project," in *Visual Media Production, 2009. CVMP '09. Conference for*, 2009, pp. 176–187.
- [52] J. Osipa, *Stop Staring: Facial Modeling and Animation Done Right*. John Wiley & Sons, 2010.
- [53] A. Arghinenti, "Animation workflow in KILLZONE3™: a fast facial retargeting system for game characters," in *ACM SIGGRAPH 2011 Talks*, New York, NY, USA, 2011, pp. 37:1–37:1.
- [54] T. McLaughlin, "Taxonomy of digital creatures: defining character development techniques based upon scope of use," in *ACM SIGGRAPH 2006 Courses*, New York, NY, USA, 2006.
- [55] I. S. Pandzic and R. Forchheimer, *MPEG-4 facial animation: the standard, implementation and applications*. Wiley. com, 2003.
- [56] J. Lasseter, "Principles of traditional animation applied to 3D computer animation," in *ACM Siggraph Computer Graphics*, 1987, vol. 21, pp. 35–44.
- [57] A. Menache, *Understanding motion capture for computer animation*. Elsevier, 2011.

- [58] A. Arya and S. DiPaola, "Multispace behavioral model for face-based affective social agents," *Journal on Image and Video Processing*, vol. 2007, no. 1, pp. 4–4, 2007.
- [59] D. Arellano, J. Varona, and F. J. Perales, "Generation and visualization of emotional states in virtual characters," *Computer Animation and Virtual Worlds*, vol. 19, no. 3-4, pp. 259–270, 2008.
- [60] R. Bredow, D. Schaub, D. Kramer, M. Hausman, D. Dimian, and R. S. Duguid, "Surf's up: the making of an animated documentary," in *ACM SIGGRAPH 2007 courses*, 2007, pp. 1–123.
- [61] Digital-Tutors, "Facial Rigging in Maya," 2009. [Online]. Available: <http://www.digitaltutors.com>. [Accessed: 15-May-2009].
- [62] O. Alexander, M. Rogers, W. Lambeth, M. Chiang, and P. Debevec, "The Digital Emily project: photoreal facial modeling and animation," in *ACM SIGGRAPH 2009 Courses*, New York, NY, USA, 2009, pp. 12:1–12:15.
- [63] D. Komorowski, V. Melapudi, D. Mortillaro, and G. S. Lee, "A hybrid approach to facial rigging," in *ACM SIGGRAPH ASIA 2010 Sketches*, 2010, p. 42.
- [64] P. M. Plantec, *Virtual humans: a build-it-yourself kit, complete with software and step-by-step instructions*. New York: AMACOM, American Management Association, 2004.
- [65] J. C. Miranda, X. Alvarez, J. Orvalho, D. Gutierrez, A. A. de Sousa, and V. Orvalho, "Sketch Express: Facial Expressions Made Easy.," in *SBM*, 2011, pp. 87–94.
- [66] P. Bergeron and P. Lachapelle, "Controlling facial expressions and body movements in the computer generated animated short 'Tony de Peltrie'}," 1985.
- [67] P. Jackson, *The Lord of the Rings: The Two Towers*. 2002.
- [68] V. C. T. Orvalho, "Reusable facial Rigging And Animation: Create Once, Use Many." 2007.
- [69] H. Li, T. Weise, and M. Pauly, "Example-based facial rigging," *ACM Transactions on Graphics (TOG)*, vol. 29, no. 4, p. 32, 2010.
- [70] J. Noh, D. Fidaleo, and U. Neumann, "Animated deformations with radial basis functions," in *Proceedings of the ACM symposium on Virtual reality software and technology*, 2000, pp. 166–174.
- [71] L. Innes, "The Secret To Drawing Characters Consistently & Expressively." [Online]. Available:

<http://www.paperwingspodcast.com/2011/06/the-secret-to-drawin-expressive-consistent-characters/>. [Accessed: 18-Sep-2012].

- [72] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001*, 2001, vol. 1, pp. I-511 – I-518 vol.1.
- [73] J. Varona, C. Manresa-Yee, and F. J. Perales, "Hands-free vision-based interface for computer accessibility," *Journal of Network and Computer Applications*, vol. 31, no. 4, pp. 357–374, Nov. 2008.
- [74] M. Castrillón, O. Déniz, D. Hernández, and J. Lorenzo, "A comparison of face and facial feature detectors based on the Viola–Jones general object detection framework," *Machine Vision and Applications*, vol. 22, no. 3, pp. 481–494, 2011.
- [75] E. Cerezo, I. Hupont, C. Manresa-Yee, J. Varona, S. Baldassarri, F. Perales, and F. Seron, "Real-Time Facial Expression Recognition for Natural Interaction," in *Pattern Recognition and Image Analysis*, vol. 4478, J. Martí, J. Benedí, A. Mendonça, and J. Serrat, Eds. Springer Berlin / Heidelberg, 2007, pp. 40–47.
- [76] F. Pighin and J. P. Lewis, "Facial motion retargeting," in *ACM SIGGRAPH 2006 Courses*, New York, NY, USA, 2006.
- [77] H. Huang, J. Chai, X. Tong, and H.-T. Wu, "Leveraging motion capture and 3D scanning for high-fidelity facial performance acquisition," in *ACM Transactions on Graphics (TOG)*, 2011, vol. 30, p. 74.
- [78] I. Patras and M. Pantic, "Particle filtering with factorized likelihoods for tracking facial features," in *Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference on*, 2004, pp. 97–102.
- [79] M. Pantic and I. Patras, "Detecting facial actions and their temporal segments in nearly frontal-view face image sequences," in *Systems, Man and Cybernetics, 2005 IEEE International Conference on*, 2005, vol. 4, pp. 3358 – 3363 Vol. 4.
- [80] J. M. Saragih, S. Lucey, and J. F. Cohn, "Face alignment through subspace constrained mean-shifts," *IEEE Internet Computing*, pp. 1034–1041, 2009.
- [81] "Face Tracking." [Online]. Available: <http://msdn.microsoft.com/en-us/library/jj130970.aspx>.
- [82] T. Weise, S. Bouaziz, H. Li, and M. Pauly, "Realtime performance-based facial animation.," *ACM Trans. Graph.*, vol. 30, no. 4, p. 77, 2011.

- [83] K. Kawakami, K. Takai-Kawakami, M. Tomonaga, J. Suzuki, T. Kusaka, and T. Okai, "Origins of smile and laughter: a preliminary study," *Early Human Development*, vol. 82, no. 1, p. 6, 2006.
- [84] P. Ekman and W. V. Friesen, "Felt, False, and Miserable Smiles," *Journal of Nonverbal Behavior*, vol. 6, no. 4, p. 252, 1982.
- [85] P. Ekman, W. V. Friesen, and M. O'Sullivan, "Smiles When Lying," *Journal of Personality and Social Psychology*, vol. 54, no. 3, p. 420, 1988.
- [86] E. E. Nwokah, H.-C. Hsu, P. Davies, and A. Fogel, "The Integration of Laughter and Speech in Vocal Communication: A Dynamic System Perspective," *Journal of Speech, language and Hearing Research*, vol. 42, p. 894, 1999.
- [87] J.-A. Bachorowski and M. J. Smoski, "The acoustic features of human laughter," *Acoustical Society of America Digital Library*, vol. 110, no. 3, p. 1597, 2001.
- [88] D. S. Messinger, A. Fogel, and K. L. Dickson, "All Smiles Are Positive, But Some Smiles Are More Positive Than Others," *Developmental Psychology*, vol. 37, no. 5, p. 653, 2001.
- [89] M. Mori, "The uncanny valley," *Energy*, vol. 7, no. 4, pp. 33–35, 1970.
- [90] J. Unay and R. Grossman, "Hyper-real advanced facial blendshape techniques and tools for production," in *ACM SIGGRAPH 2005 Master Class*, 2005.
- [91] E. Catmull and J. Clark, "Computer Aided Design," vol. 6, 1978, p. Elsevier Science.
- [92] "About Smiley," 2010. [Online]. Available: <http://smileycompany.com/smiley-original/about>.
- [93] E. Lubitsch, *Ninotchka*. 1940.
- [94] A. Minghella, *The Talented Mr. Ripley*. 1999.
- [95] R. Aldrich, *What Ever Happened to Baby Jane?* 1962.
- [96] H. Hathaway, *Kiss of Death*. 1947.
- [97] M. Forman, *Amadeus*. 1984.
- [98] J. Saragih, S. Lucey, and J. Cohn, "Deformable Model Fitting by Regularized Landmark Mean-Shift," *International Journal of Computer Vision*, vol. 91, no. 2, pp. 200–215, 2011.
- [99] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I.

- Matthews, “The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression,” in *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2010, pp. 94–101.
- [100] T. Palamar, *Mastering Autodesk Maya 2014*. 2013.
- [101] “Boris with Facial GUI - Free Character Rigs Downloads for Maya.” [Online]. Available: <http://www.creativecrash.com/maya/downloads/character-rigs/c/boris-with-facial-gui>. [Accessed: 20-Nov-2013].
- [102] “Faceware Technologies.” [Online]. Available: <http://www.facewaretech.com/>. [Accessed: 28-Sep-2012].
- [103] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, “Web-based database for facial expression analysis,” in *Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on*, 2005, p. 5 pp.
- [104] “BBC - Science & Nature - Human Body and Mind - Spot The Fake Smile.” [Online]. Available: <http://www.bbc.co.uk/science/humanbody/mind/surveys/smiles/>. [Accessed: 23-Nov-2012].
- [105] M. E. Hoque, D. J. McDuff, and R. W. Picard, “Exploring Temporal Patterns in Classifying Frustrated and Delighted Smiles,” *Affective Computing, IEEE Transactions on*, vol. 3, no. 3, pp. 323–334, 2012.
- [106] H. Dibeklioglu, A. Salah, and T. Gevers, “Are You Really Smiling at Me? Spontaneous versus Posed Enjoyment Smiles,” in *Computer Vision – ECCV 2012*, vol. 7574, A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, Eds. Springer Berlin Heidelberg, 2012, pp. 525–538.
- [107] S. Petridis, B. Martinez, and M. Pantic, “The MAHNOB Laughter Database,” *Image and Vision Computing Journal*, vol. 31, no. 2, pp. 186–202, Feb. 2013.
- [108] “ELAN.” [Online]. Available: <http://tla.mpi.nl/tools/tla-tools/elan/>. [Accessed: 27-Jan-2014].
- [109] S. Salvador and P. Chan, “Toward accurate dynamic time warping in linear time and space,” *Intelligent Data Analysis*, vol. 11, no. 5, pp. 561–580, Jan. 2007.
- [110] V. Hodge and J. Austin, “A Survey of Outlier Detection Methodologies,” *Artificial Intelligence Review*, vol. 22, no. 2, pp. 85–126, Oct. 2004.
- [111] J. Laurikkala, M. Juhola, E. Kentala, N. Lavrac, S. Miksch, and B. Kavsek, “Informal identification of outliers in medical data,” in

- Proceedings of the 5th International Workshop on Intelligent Data Analysis in Medicine and Pharmacology*, 2000, pp. 20–24.
- [112] J. Xiao, Y. Zhuang, T. Yang, and F. Wu, “An Efficient Keyframe Extraction from Motion Capture Data,” in *Advances in Computer Graphics*, vol. 4035, T. Nishita, Q. Peng, and H.-P. Seidel, Eds. Springer Berlin Heidelberg, 2006, pp. 494–501.
- [113] O. Onder, U. Gudukbay, B. Ozguc, T. Erdem, C. Erdem, and M. Ozkan, “Keyframe Reduction Techniques for Motion Capture Data,” presented at the 3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2008, 2008, pp. 293–296.
- [114] M. Chion, *La Audiovisión Introducción a un análisis conjunto de la imagen y el sonido*. Barcelona, 1993.