

ANIMAÇÃO DE PERSONAGEM 3D: EXPRESSÃO FACIAL

3D CHARACTERS ANIMATION: FACIAL EXPRESSION
LA ANIMACIÓN DE PERSONAJES 3D: EXPRESIÓN FACIAL

José Carlos Miranda * (jcmira@ippg.pt)

RESUMO

Durante os últimos anos, emergiu um interesse crescente em animação 3D de caracteres e expressões faciais. Este facto está relacionado com a nova participação de actores virtuais em televisão, cinema e indústrias de jogos interactivos, entre outros. Nestes casos, o objectivo principal é criar personagens virtuais que simulam o corpo e todos os detalhes de movimento do rosto, da forma mais realista possível. Neste artigo é efectuada uma revisão dos métodos mais usados na animação 3D. O método base do esqueleto é o mais utilizado para criar os movimentos principais, como o andar, por exemplo; os métodos base de mistura de formas são usados para criar deformações mais subtis. Em áreas que exigem um maior controlo e realismo, como alguns músculos do rosto, é possível adicionar camadas de deformação. Normalmente, curvas de NURBS ou grelhas de FFD são usadas para realçar algumas características do rosto. Os vários métodos têm vantagens e desvantagens. Assim, a melhor solução é combinar as diferentes técnicas. Palavras-chave: Animação de personagem, Objectos de Deformação, Transformação, Animação do esqueleto, Expressão Facial.

ABSTRACT

During the last years a growing interest in 3D animation of characters and facial expressions has emerged. This is related with the new participation of virtual actors in television, cinema and interactive games' industries among others. In these cases, the main goal is to create virtual characters that simulate body and all the facial movement details, in the most realistic possible way. In this paper a study about the most used 3D animation methods is presented. The skeleton based methods are the most used to create the character main movements, like walking, and the blend shapes based methods are used to create more subtle deformations. In regions that require a bigger control and realism, like some

face muscles, it is possible to add deformation layers. Usually, NURBS curves or FFD grids are used to emphasize some facial features. The various methods have advantages and disadvantages, thus the best solution is to combine the different techniques.

Keywords: Character Animation, Deformable Objects, Morphing, Skeletal Animation, Facial Expression.

RESUMEN

Durante los últimos años ha surgido un interés creciente en la animación 3D de caracteres y expresiones faciales ha surgido. Esto está relacionado con la nueva participación de actores virtuales en la televisión, el cine y en industrias de juegos interactivos entre otros. En estos casos, el principal objetivo es de crear caracteres virtuales que simulan cuerpo y todos los detalles de movimiento de rostro, de forma más realista. En este artículo se realiza una revisión de los métodos más utilizados en la animación 3D. El método de base del esqueleto es el más usado para crear los principales movimientos, como andar, y los métodos base de mezcla de formas son utilizados para crear deformaciones más sutiles. En las áreas que requieren un mayor control y realismo, como algunos músculos de rostro, es posible añadir capas de deformación. Normalmente, curvas de NURBS o cuadrículas de FFD son utilizadas para acentuar algunas características de rostro. Los varios métodos tienen ventajas y desventajas, así la mejor solución es de combinar las diferentes técnicas.

Palabras-clave: Animación del personaje, Objetos de Deformación, Transformación, Animación del Esqueleto, Expresión Facial.

* Mestre em Tecnologia Multimédia, Professor da Escola de Tecnologia e Gestão do Instituto Politécnico da Guarda.

1. INTRODUÇÃO

Os primeiros registos de animação datam de 2000 a.C., quando os egípcios pintavam nas paredes das cavernas sequências de lutas e cenas de adoração.

Somente em 1828 a animação pôde ser explicada pelo entendimento do princípio fundamental do olho humano: a persistência da visão. Na animação tradicional, a ilusão do movimento era conseguida através dos vários desenhos criados pelo artista e gravados *frame a frame* numa sequência de imagens.

A animação por computador foi baseada na animação tradicional e começou por utilizar métodos muito simples, como a técnica de *Keyframe*. Um *keyframe* é qualquer *frame* de uma animação onde supostamente ocorre um evento específico importante. Os *frames* localizados entre os *keyframes* são chamados de *inbetweens* e são calculados por interpolação linear. Trata-se de uma técnica antiga, introduzida por Burtnyk e Wein (1971), que produz resultados pouco satisfatórios, como movimentos pouco suaves, descontinuidade na velocidade e distorção nas rotações, como se pode ver na Figura 1.

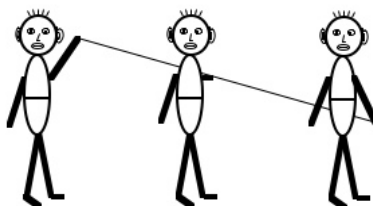


Figura 1 – A interpolação linear produz uma redução indesejada no comprimento do braço

(Magenat-Thalman ; Thalman, 1993)

Kochanek e Bartels (1984) introduziram um método de interpolação mais robusto, denominado de “interpolação *spline*” e que utiliza funções cúbicas. O método permite controlar cada ponto de uma curva através de 3 parâmetros: tensão, continuidade e inclinação, obtendo desta forma movimentos mais suaves. Watt e Watt (1992) apresentam uma perspectiva geral sobre os diferentes tipos de *splines*.

Posteriormente, técnicas de simulação baseadas na cinemática e dinâmica foram importadas da Robótica, ao mesmo

tempo que eram desenvolvidos métodos de *rendering* sofisticados. Este tipo de animação, conhecido como “Animação Procedimental e Simulação”, consiste basicamente na utilização de modelos matemáticos, implementados em linguagens de programação para simulação de forças físicas, como, por exemplo, o caminhar de um personagem, a expressão facial, o movimento de roupas, etc. Algoritmos de “cinemática inversa (IK)” são apresentados em Welman (1993) como necessários para “ligar” um personagem ao meio ambiente, dando-lhe a possibilidade de, através de movimentos minimamente realistas, interagir com esse ambiente; por exemplo, caminhar sobre um chão irregular. Por outro lado, quando se pretende que o personagem 3D agarre ou toque num objecto, já é necessária alguma integridade “dinâmica”. Enquanto a cinemática procura descrever os movimentos sem se preocupar com as forças que os originam, a dinâmica estuda a relação existente entre as forças e o movimento. Arnaldi *et al.* (1989) apresentam vários métodos de animação baseados na dinâmica.

Com o desenvolvimento da tecnologia e o aparecimento de dispositivos de realidade virtual, começaram a emergir sistemas de captura de movimento. “Captura de Movimento (MoCap)” é uma técnica que permite mapear os movimentos de um actor real para um personagem 3D. Os sistemas de captura de movimento evoluíram muito nas últimas décadas com algoritmos capazes de reconhecer os marcadores no corpo de uma pessoa.

Actualmente, existem basicamente quatro tipos: óptico, mecânico, magnético e acústico. O sistema óptico é o mais utilizado uma vez que oferece ao actor uma maior liberdade de movimentos, já que não há a necessidade de cabos ligados ao corpo. O sistema utiliza pequenos discos reflectivos que são fixados ao corpo do actor. A área de actuação é gravada com várias câmaras de vídeo, cada uma equipada com uma fonte de luz, geralmente infra-vermelha, dirigida para iluminar a área de visão da câmara. Os discos reflectem a luz na direcção da câmara e um computador capta e sincroniza a imagem. Os dados das posições x, y, z de cada marcador são aplicados a um sistema de cinemática inversa para animar um esqueleto 3D. Uma desvantagem deste sistema é a facilidade com que o corpo do actor pode obstruir a reflexão dos discos. Esse problema pode ser contornado com a utilização de mais câmaras. Os sistemas de captura de movimento apresentam a grande vantagem de produzir animações bastante realistas.

A criação de novos movimentos através de animações predefinidas denomina-se de “Síntese de Movimento” e implica a existência de uma biblioteca de movimentos previamente definidos e algoritmos para adaptar esses movimentos a outros objectos. Síntese de movimento é, no momento em que se escreve este documento, um importante e activo campo de investigação. Em Liu e Popovic (2002) ou Arikan *et al.* (2003) pode ser encontrada informação fundamental sobre este assunto.

Nos últimos anos tem-se assistido a um crescente interesse na animação de personagens 3D e na expressão facial. Este interesse está relacionado com a crescente participação de actores virtuais na indústria da televisão, cinema, jogos interactivos, comunidades virtuais, e outros. Nestas novas áreas de aplicação, a interactividade e o tempo real são aspectos fundamentais. Embora recentes trabalhos produzam resultados realistas com relativa rapidez de execução, todo o processo de produção existente num *pipeline* de animação requer muita intervenção humana. O grande objectivo na investigação de novas técnicas de animação 3D tem sido a procura de metodologias capazes de automatizar algumas dessas tarefas e que operem em tempo real.

Este documento tenta sintetizar vários assuntos relacionados com a animação de personagens 3D e com a expressão da face humana, e está organizado da seguinte forma. Na secção 2 são feitas algumas considerações fundamentais sobre a animação do esqueleto e apresentados, com algum formalismo, alguns dos métodos mais utilizados na animação de personagens 3D. A animação facial é tratada na secção 3, onde são abordados assuntos relacionados com a parametrização e com o *rigging* facial. Por último, na secção 4, é apresentado um sumário e a conclusão.

2. ANIMAÇÃO DE PERSONAGENS 3D

A animação de personagens envolve tipicamente dois tipos de movimentos. Os movimentos primários, que consistem, basicamente, em definir os movimentos principais do personagem, como, por exemplo, o caminhar ou rodar a cabeça, e os movimentos secundários, relacionados com pequenas subtilezas

que adicionam realismo à animação, como é o caso da contracção de pequenos músculos da face ou a respiração.

2.1. DEFINIÇÃO DE UM ESQUELETO

Para representar personagens animadas, é conveniente armazenar as relações entre cada parte móvel diferente dos elementos geométricos que compõem o personagem. Tomando como exemplo o homem, as partes responsáveis pela deformação e movimento do personagem 3D recebem o nome dos seus correspondentes com a estrutura real do corpo humano. Os ossos são elementos individuais definidos a partir do seu comprimento e vector “direcção” e estabelecem a ligação entre duas articulações. As articulações são os pontos *pivots* de rotação que estabelecem a ligação entre os ossos e podem apresentar até três eixos de rotação. Como se pode ver na Figura 2, o esqueleto é uma estrutura hierárquica formada por ossos e articulações. Além da parte óssea, ainda há que considerar a pele e os músculos, que, além de limitar o movimento do esqueleto, permitem a deformação da malha.

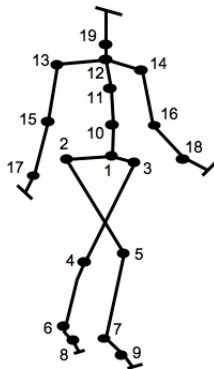


Figura 2: Um esqueleto básico, constituído por ossos (linhas) e articulações (círculos)

O processo de criar um esqueleto e “ligá-lo” a uma determinada personagem é conhecido como *rigging*. Criar o *rig* de um personagem requer tempo, experiência e conhecimento da anatomia do personagem (Henry-Biskup, 1998). Para fixar a malha do personagem ao esqueleto, isto é, definir quais os vértices da malha que são influenciados por determinado osso, podem ser usados algoritmos heurísticos (Mohr e Gleicher, 2003). O método

manual de fazer esta atribuição de vértices aos ossos é conhecido por *weight-painting* e está implementado na maioria dos *softwares* de animação 3D.

2.2. CINEMÁTICA PARA ANIMAÇÃO DO ESQUELETO

Para criar a animação de um esqueleto é necessário definir um conjunto de *keyframes*. Essas posições-chave são geralmente criadas através de técnicas de cinemática directa (FK) e cinemática inversa (IK), tal como descrito em Welman (1993).

A cinemática directa usa a metodologia de hierarquia *top-down*, onde a posição das articulações “pai” interfere nas articulações “filho” (Figura 3).

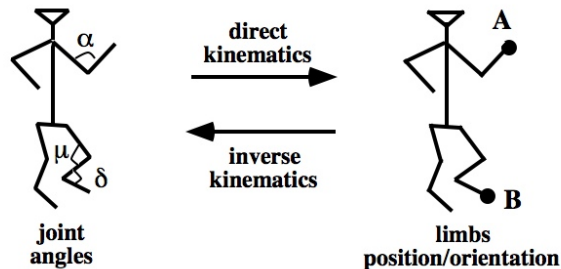


Figura 3: Cinemática directa (FK) e inversa (IK)

(Magnat-Thalmann, Thalmann, 1993)

A cinemática directa é uma boa introdução para a compreensão da estrutura básica de árvore de uma animação, mas torna-se complexa, porque todos os ossos “filhos” ficam dependentes de qualquer animação que se fizer no osso “pai”. Os esqueletos FK são geralmente usados em cenas com pouco movimento.

A cinemática inversa utiliza a metodologia de hierarquia inversa, na qual a posição e orientação das articulações e membros “pai” vai ser calculada a partir da posição final da articulação “filho” da hierarquia, ou seja, um osso do pé, na sua posição final, definirá como os outros ossos do joelho e anca deverão rodar e posicionar-se, tornando a animação de um andar muito mais realista (Figura 3).

Comparando-se as duas técnicas, percebe-se que algoritmos IK possibilitam a criação de movimentos mais complexos e realistas; porém, o facto de o movimento ser definido por uma

função pode resultar em movimentos indesejados (Azevedo e Conci, 2003).

2.3. ANIMAÇÃO E DEFORMAÇÃO DE SUPERFÍCIES

Inicialmente, a animação era realizada apenas com objectos rígidos; mais tarde surgiram os objectos articulados e, recentemente, aqueles ditos flexíveis ou deformáveis. A animação de personagens ou objectos deformáveis exige que se possua um modelo geométrico que permita a mudança de forma ao longo do tempo. Em computação gráfica, isso pode ser definido através de métodos numéricos (elementos finitos ou de contorno) ou de modelos físicos de curvas, superfícies, ou mesmo sólidos deformáveis.

As deformações da superfície do personagem justificam-se por um variadíssimo número de razões. De acordo com Ghyme *et al.* (2000), a maior parte dessas deformações podem ser classificadas em seis categorias:

1. *Bending*;
2. *Twisting*;
3. Saliências;
4. Rugas;
5. Face;
6. Cabelo e Roupa;

Bending pode ser visualizada quando se dobra os braços e as pernas. *Twisting* ocorre quando se roda o pescoço ou a cintura. Estas são as deformações mais visíveis que ocorrem num personagem. As rugas são muitas das vezes efeitos secundários dessas deformações. A contracção dos músculos causa saliências e rugas. A deformação facial é provocada pelo movimento dos músculos faciais. É menos visível, mas muito importante nos grandes planos e para expressar emoções. As deformações do cabelo e roupa são necessárias para adicionar realismo extra, mas muito complexas para executar em tempo real.

As malhas poligonais são geralmente deformáveis, movendo os seus vértices. Para tal, são necessárias adequadas técnicas de deformação. A seguir, vão ser apresentadas três técnicas de deformação muito utilizadas na animação de personagens.

2.3.1. *Morphing with Blend Shape Interpolation*

Morphing é uma técnica que transforma uma imagem A numa imagem B através de uma transição suave. Esta técnica pode

ser aplicada a modelos 3D através do algoritmo *blend shape interpolation*, também conhecido como *morph target animation*, *shape blending* ou *multi-target morphing*. Este assunto é amplamente discutido em Lewis *et al.* (2000).

A primeira fase consiste em criar diferentes modelos da mesma personagem. Estes modelos são chamados *blend shapes* e apresentam deformações específicas da personagem (Figura 4).

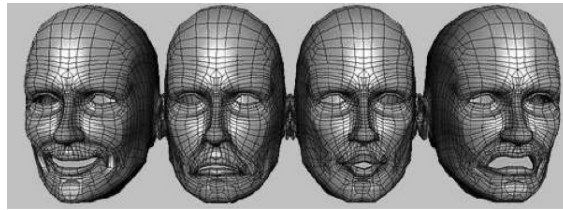


Figura 4: Blend shapes que representam as expressões básicas: feliz, triste, zangado e surpreso

(Orvalho, 2007)

Estes modelos devem apresentar a mesma estrutura (mesmo nº de vértices, mesma conectividade, correspondência de vértices nos diferentes modelos). A deformação do personagem é a combinação linear do número total de *blend shapes*. A equação 1 descreve a forma de calcular os novos vértices:

$$\bar{v} = \sum_{k=0}^{n-1} w_k v_k \quad (1)$$

v_k é o vértice do *blend shape* k , w_k é a influência do *blend shape* k e n é o nº total de *blend shapes*.

A grande vantagem desta técnica reside no facto de as expressões e deformações pretendidas serem definidas na modelação, o que torna possível realizar deformações de grande subtilidade. Pelo contrário, e como já foi referido na secção 1, as interpolações lineares apresentam problemas nas rotações, causando, assim, efeitos indesejados em algumas deformações básicas, como o *bending* e o *twisting* (Figura 5).



Figura 5: Distorção visível no movimento de rotação do braço

(Garstenauer; Kurka, 2006)

Outra desvantagem é o grande consumo de memória, uma vez que é necessário armazenar a posição de todos os vértices dos *blend shapes*.

Esta técnica é muito utilizada na animação facial, para criar o efeito da contracção dos músculos e também é usada para simular o efeito da roupa amarrotada.

2.3.2. *Skeletal Animation with Smooth Skinning*

Skeletal Animation é actualmente a técnica mais popular na animação de personagens 3D. A ideia-base é a de que a superfície do modelo é deformada sobre um esqueleto. O algoritmo mais utilizado é conhecido como *smooth skinning* e citado em várias publicações como Mohr e Gleicher (2003), Lewis *et al.* (2000), Sloan *et al.* (2001), Wang e Phillips (2002), com diferentes designações, tais como *linear blend skinning*, *skeleton-subspace deformation (SSD)*, *matrix deformation* ou *single-weight enveloping*.

A primeira fase consiste em criar o modelo da personagem numa posição neutra, conhecida como *rest pose*, ou *dress pose*, ou *binding pose* (Figura 6), e produzido o seu *rigging*, que envolve a criação de um esqueleto e a “ligação” dos ossos à superfície do modelo.

Cada osso tem o seu próprio sistema de coordenadas. Assim, quando um osso se move, todos os vértices que estão sob a sua influência sofrem o mesmo movimento. No caso de um vértice estar atribuído a ossos diferentes, o deslocamento resultante é um somatório “pesado” dos movimentos de cada osso. A influência de um osso k sobre um vértice v é estimada por um valor w_k ($w_k \in [0, 1]$ e o $\sum w_k = 1$).



Figura 6: Personagem 3D com esqueleto, em posição neutra
(Garstenauer; Kurka, 2006)

Para criar uma nova *pose* do modelo basta definir uma nova posição do esqueleto. Cada vértice v da nova posição é calculado pela equação 2.

$$\bar{v} = \sum_{k=0}^{n-1} w_k M_k M_{Rest,k}^{-1} v \quad (2)$$

$M_{Rest,k}$ é a matriz de transformação do osso k , na posição neutra (*rest pose*), e M_k é a matriz de transformação do osso k , na nova posição.

Inicialmente, $M_{Rest,k}^{-1} v$ transforma o vértice v da posição neutra (*rest pose*) para o sistema de coordenadas do osso k . A seguir, M_k transforma esse vértice para a nova posição do espaço. Isto é feito para todos os ossos que influenciam o vértice v . As posições resultantes sofrem, depois, uma combinação linear tendo em conta os valores w_k .

Este algoritmo tem uma excelente *performance*, já que utiliza cálculos matemáticos simples; tudo se resume a combinações lineares e multiplicação de matrizes. As deformações também são simples de controlar com esqueletos. Outra vantagem reside no facto de personagens diferentes poderem partilhar o mesmo

esqueleto, o que significa que animações produzidas para um modelo podem ser reutilizadas em diferentes modelos.

Para além das vantagens enunciadas, o algoritmo apresenta alguns problemas nos vértices partilhados por mais do que um osso, como acontece com o cotovelo (Figura 7).

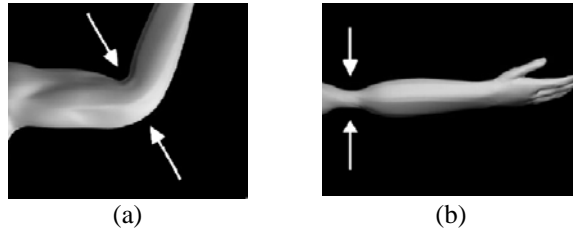


Figura 7: Problema do cotovelo conhecido como (a) “collapsing elbow” e (b) “candy-wrapper”

(Lewis et al., 2000)

Mohr e Gleicher (2003) e Weber (2000) apresentam algumas estratégias para resolver este problema.

Outra desvantagem do *smooth skinning* é que este algoritmo não permite o controlo directo sobre os vértices da superfície, mas sim sobre os ossos, o que prejudica algumas deformações mais complexas, como, por exemplo, a simulação de rugas. Para ultrapassar este problema é necessário adicionar mais ossos na região onde se pretende aplicar a deformação.

Este algoritmo é muito utilizado nas deformações principais dos personagens, como o caminhar, rodar a cabeça ou a cintura, etc. Também pode ser utilizado para simular a contracção de músculos, mas, como já foi referido, é necessário adicionar mais ossos de controlo na região de deformação.

Como se pode ver em Unreal (2007), a próxima geração de motores de jogos como o Unreal Engine 3 vai usar extensivamente o *smooth skinning*, utilizando modelos com diferentes níveis de detalhe que podem ir dos 100 aos 200 ossos.

2.3.3. Free-Form Deformation

Free-Form deformation (FFD) foi inicialmente introduzido por Sederberg e Parry (1986) como uma técnica para deformar geometricamente objectos. Mais tarde, trabalhos apresentados por Coquillart e Jancene (1991) demonstraram que FFD pode ser usado como uma técnica de animação. O princípio desta técnica consiste

em colocar uma grelha 3D de pontos de controlo (*lattice*¹) sobre a superfície do objecto que se pretende deformar. A grelha de controlo é então deformada, e essa deformação é aplicada à superfície do objecto (Figura 8). O algoritmo pode ser usado para deformar superfícies paramétricas ou malhas poligonais.

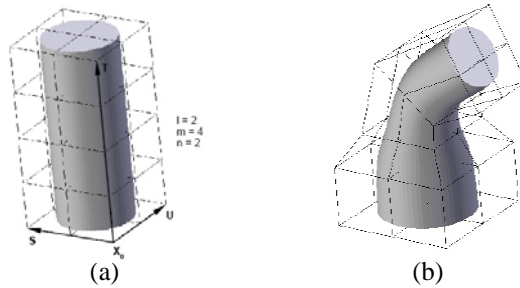


Figura 8: FFD aplicado a uma superfície cilíndrica. (a) posição neutra. (b) deformação

(Garstenauer e Kurka, 2006)

A grelha de controlo é constituída por pontos de controlo P_{ijk} e tem o seu próprio sistema de coordenadas (s, t, u) .

Cada ponto X do espaço cartesiano pode ser mapeado para o espaço de coordenadas (s, t, u) e vice-versa, através das equações 3 e 4.

$$X = X_0 + sS + tT + uU \quad (3)$$

e

$$s = \frac{T \times U \cdot (X - X_0)}{T \times U \cdot S}, \quad t = \frac{S \times U \cdot (X - X_0)}{S \times U \cdot T}, \quad u = \frac{S \times T \cdot (X - X_0)}{S \times T \cdot U} \quad (4)$$

A posição dos pontos de controlo é definida pela equação 5.

$$P_{ijk} = X_0 + \frac{i}{l}S + \frac{j}{m}T + \frac{k}{n}U \quad (5)$$

¹ um objecto *lattice* é um objecto simples que serve para controlar um objecto mais complexo.

Inicialmente, cada ponto da superfície tem de ser mapeado para o sistema de coordenadas (s, t, u) . Posteriormente, a grelha de controlo é deformada deslocando as posições dos seus pontos de controlo P_{ijk} . A posição correspondente X_{ffd} da superfície pode ser obtida através da equação 6.

$$X_{ffd} = \sum_{i=0}^l \sum_{j=0}^m \sum_{k=0}^n B_{l,i}(s) B_{m,j}(t) B_{n,k}(u) P_{ijk} \quad (6)$$

São utilizados polinómios de *Bernstein* para calcular os $B_{l,i}(s)$, $B_{m,j}(t)$ e $B_{n,k}(u)$. Mais informação sobre este assunto pode ser encontrada em Coquillart e Jancene (1991).

O inconveniente desta técnica é o tempo de processamento, uma vez que exige muitos cálculos por ponto de superfície. MacCracken e Joy (1996) apresentam alguns melhoramentos ao clássico FFD.

As grelhas de controlo FFD podem ser colocadas em regiões específicas do personagem, permitindo, assim, a deformação dos músculos do corpo ou da face, a animação de roupas ou do cabelo.

3. ANIMAÇÃO FACIAL

Um dos aspectos mais importantes a considerar na animação de personagens é a “expressão facial”. A face é uma estrutura muito irregular, que varia de personagem para personagem e desempenha um papel fundamental na transmissão de emoções. O problema torna-se mais complexo quando se trata de uma personagem humana e é necessário simular todos os detalhes que compõem a face, tais como os músculos, os ossos ou as rugas. A simulação destes detalhes deve ser o mais realista possível, para captar todas as subtilezas únicas de uma expressão facial, sob pena de se obterem resultados estranhos, conhecidos como *Uncanny Valley* (Mori, 1982).

A animação facial é hoje uma importante área da animação, em crescente desenvolvimento e fortemente influenciada por toda a investigação já iniciada para a animação do resto do corpo. Contudo, comparando a face de um personagem com o resto do

seu corpo, percebe-se, de imediato, diferenças significativas que se devem ter em conta, tais como a necessidade de um maior número de músculos para criar uma expressão facial do que para criar uma posição específica do corpo, ou a cabeça não poder ser animada apenas com uma articulação, como acontece com algumas partes do corpo. Para serem obtidos resultados realistas na animação de uma face, é necessário utilizar um maior número de deformadores para simular as animações produzidas pelos músculos da face, atribuir um maior número de ossos para influenciar regiões específicas da face e implementar um grande número de controladores para manipular toda a estrutura da cabeça e regiões secundárias da face.

A expressão facial humana foi assunto de investigação científica por mais de cem anos. Os esforços em relação à expressão facial por computador começaram há mais de trinta anos, no início da década de setenta do século passado. É interessante perceber que a maioria das técnicas actualmente utilizadas envolve os princípios desenvolvidos várias décadas atrás. A primeira animação facial 3D foi criada por Parke (1972). Desde então, diferentes abordagens sobre o assunto têm sido desenvolvidas e classificadas em duas categorias principais: “manipulação geométrica 3D ou manipulação de imagem 2D”. Não é muito claro perceber a qual categoria um determinado método pertence, uma vez que as fronteiras entre estas duas tecnologias não estão muito bem definidas. Manipulação geométrica normalmente inclui técnicas como *keyframe*, parametrização, modelos baseados na física e outros; manipulação de imagem inclui técnicas de *morphing* e *blend shaping*. Um bom *survey* que apresenta uma classificação para os métodos de animação facial pode ser encontrado em Noh e Neumann (1998). Análise e compreensão facial é uma outra área que influencia as actuais tendências da síntese facial. Zhao *et al.* (2003) apresentam um documento detalhado sobre reconhecimento facial que pode dar uma perspectiva diferente para complementar a investigação actual.

3.1. PARAMETRIZAÇÃO FACIAL

A parametrização facial consiste em definir um conjunto de parâmetros que são usados para controlar os movimentos faciais. Em vez de se manipular directamente a geometria do modelo,

trabalha-se num nível de abstracção mais elevado. Definir uma óptima parametrização é uma tarefa complexa e difícil; a investigação tem mostrado que não existe nenhuma parametrização ideal por ser difícil satisfazer todas as necessidades existentes nesta área da animação. Parke (1974) desenvolveu o primeiro modelo facial 3D parametrizado que permitia obter deformações na face humana através de uma definição *ad-hoc* de parâmetros. Ekman e Friesen (1978) definiram um sistema para descrever e avaliar o comportamento da face humana denominado *Facial Action Coding System* (FACS)². Este sistema tornou-se um *standard* para categorizar as expressões físicas das emoções e é actualmente usado por animadores e psicólogos. Neste sistema, as expressões faciais são codificadas por movimentos faciais básicos, conhecidos como *Action Units* (AUs). Actualmente, existem 66 AUs que representam contracções e relaxações de um ou mais músculos faciais. Na Tabela 1 são apresentados alguns exemplos.

Tabela 1 – Movimentos faciais básicos (AUs)

AU	FACS Name	AU	FACS Name
1	Inner Brow Raiser	12	Lid Corner Puller
2	Outer Bow Raiser	14	Dimpler
4	Brow Lower	15	Lip Corner Depressor
5	Upper Lid Raiser	16	Lower Lip Depressor
6	Cheek Raiser	17	Chin Raiser
7	Lid Tightener	20	Lip Tightener
9	Nose Wrinkler	23	Lip Tightener
10	Upper Lid Raiser	16	Jaw Drop

(Noh e Neumann, 1998)

As expressões faciais são geradas combinando diferentes AUs (Tabela 2). Por exemplo, combinando AU1 (*Inner brow raiser*), AU4 (*Brow Lower*), AU15 (*Lip Corner Depressor*), e AU23 (*Lip Tightener*) é criada a expressão de tristeza.

Definir os parâmetros adequados para controlar e animar os músculos de uma face de forma a obter expressões realistas é uma tarefa complexa mas fundamental. Na secção seguinte é feita uma descrição dos métodos de parametrização baseados em músculos que mais contribuíram para o desenvolvimento da investigação nesta área.

² <http://www.face-and-emotion.com>

Tabela 2 – Expressões faciais básicas

Basic Expressions	Involved Action Units
Surprise	AU1, 2,5, 15, 16, 20, 26
Fear	AU1, 2, 4, 5, 15, 20, 26
Disgust	AU2, 4, 9, 15, 17
Anger	AU2, 4, 7, 9, 10, 20, 26
Happiness	AU1, 6, 12, 14
Sadness	AU1, 4, 15, 23

(Noh; Neumann, 1998)

3.2. MODELOS BASEADOS EM MÚSCULOS

Platt e Badler (1981) desenvolveram o primeiro modelo de animação facial 3D baseado em músculos. O modelo utilizava o conceito de mola elástica para simular as forças geradas pelos músculos e usava codificação FACS.

Waters (1987) definiu três tipos diferentes de músculos que eram ligados à superfície da malha e eram independentes da estrutura dos ossos, o que, para além de simplificar o processo de animação, tornava possível a sua transferência para faces com diferentes topologias. Neste modelo, as expressões faciais eram obtidas através de deformações geométricas simples, controladas por um pequeno número de parâmetros, mas falhava naqueles movimentos que exigiam uma maior subtilidade.

Magenat-Thalmann *et al.* (1988) introduziram um novo conceito denominado *abstract muscle action* (AMA), definido por um conjunto de procedimentos. Um procedimento AMA permite simular uma acção específica do músculo facial e é responsável por um determinado parâmetro da face. Uma expressão facial é considerada como um grupo de parâmetros cujos valores foram obtidos à custa de procedimentos AMA.

Terzopoulos e Waters (1990) desenvolveram o trabalho anterior de Waters com um novo modelo que utiliza técnicas baseadas na física e na anatomia facial, o que tornou possível deformações bastante mais realistas. É curioso perceber que, actualmente, muito modelos baseados na física continuam a seguir os princípios enunciados por Waters.

Lee *et al.* (1995) desenvolveram um modelo de músculos composto por várias camadas que, juntamente com um sistema de

molhas, tornava possível a deformação da superfície da face. Esta abordagem apresentava resultados com grande realismo e fidelidade, mas com um elevado cálculo computacional.

Basu *et al.* (1998) descreveram um método de *tracking* e reconstrução 3D do movimento labial, a partir de um *stream* de vídeo. O modelo tridimensional dos lábios foi criado através de elementos finitos e o método desenvolvido é capaz de, automaticamente, fazer a correspondência entre os dados provenientes do vídeo e os parâmetros correspondentes no modelo 3D.

Em Byoungwon *et al.* (2001) e Byoungwon e Hyeong-Seok (2001), foi apresentado um sistema para sintetizar expressões faciais através de dados obtidos por captura de movimento (MoCap). Os resultados ainda apresentavam alguma falta de exactidão anatómica.

Um modelo anatomicamente promissor foi descrito por Kahler *et al.* (2001, 2002), que incluía diferentes tipos de músculos e suportava alguns efeitos (*bulging* e *intertwining*). A deformação era simulada pela contracção dos músculos através de um sistema de molhas elásticas.

Em Sy-sem *et al.* (2004) é descrito um sistema de músculos através de curvas NURBS para simular expressões faciais 3D. As curvas NURBS representavam os diferentes músculos, que eram posicionados na face de acordo com o conhecimento da anatomia. As deformações dos músculos eram conseguidas através da manipulação dos diferentes pontos de controlo da curva.

Recentemente, Sifakis *et al.* (2005) desenvolve um dos últimos e mais avançados modelos baseados em músculos. O sistema captura o movimento facial através de marcadores devidamente espalhados pela face e implementa um método de elementos finitos não-lineares para determinar com exactidão a acção de cada músculo. Uma característica interessante desta implementação reside no facto de forças externas, como, por exemplo, a colisão de um objecto, poderem interagir com os músculos e, por consequência, poderem modificar a aparência final da face. O método demonstra o sucesso dos sistemas de captura de movimento nesta área da animação facial.

3.3. STANDARD MPEG-4

Para além de estabelecer métodos para a codificação, transmissão, armazenamento e composição de informação visual e sonora, a especificação do MPEG-4 foi baseada nas abordagens anteriores e tornou-se no primeiro *standard* que estabelece parâmetros e formalismo para a animação facial (Pandzic e Forchheimer, 2002).

Um dos principais conceitos associados à parametrização da animação facial estabelecida pelo MPEG-4 é os denominados “pontos característicos”, ou FPs (*feature points*). Um ponto característico define um ponto de destaque para a definição e animação da face. A Figura 9 apresenta os pontos característicos definidos pelo padrão MPEG-4.

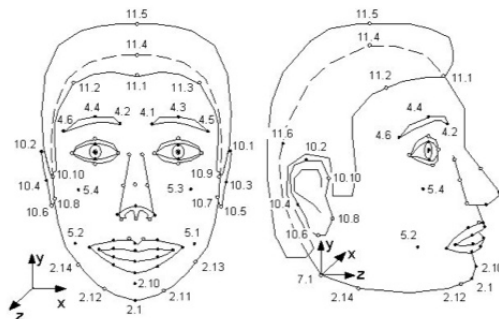


Figura 9: FPs de uma face, definidos no MPEG-4

(Balci, 2005)

A partir de uma cabeça-padrão é possível gerar uma animação através de “parâmetros de animação facial” ou FAPs (*Facial Animation Parameters*). Os FAPs definem os deslocamentos a serem aplicados aos pontos característicos. O MPEG-4 define 2 tipos de FAPs: “Visemas” e “Expressões”. Visema é a representação visual de um fonema e apresenta o posicionamento da boca para a articulação do fonema. As Expressões estão associadas a emoções e condicionadas a determinadas configurações da face. O padrão MPEG-4 define 6 expressões básicas: alegria, tristeza, raiva, medo, rejeição e surpresa.

O número de aplicações que usam algum tipo de parametrização está em constante crescimento e novas tecnologias

estão a emergir. Xface³ é um conjunto de ferramentas *open source* para criar e editar MPEG-4 (Balci, 2005).

3.4. RIGGING FACIAL

Uma das tarefas mais complexas na animação facial é criar um *rig* capaz de adaptar o modelo da face à expressividade de cada personagem. De acordo com Falk *et al.* (2004), *rigging* é o processo de transformar um modelo 3D num personagem preparado para ser animado. O *rigging* tradicional envolve a criação de um esqueleto e a sua "ligação" com a superfície do personagem. Para criar esta ligação entre o esqueleto e a malha é necessário definir qual a influência que os ossos do esqueleto exercem sobre os vértices da malha; normalmente, isto é conseguido através de algoritmos de *smooth skinning*. Num *pipeline* de animação tradicional, podemos ver o *rigging* como uma etapa intermédia entre a modelação e a animação.

Para criar um *rig* facial, normalmente são utilizados métodos de interpolação baseados em *blend shapes* e/ou métodos baseados na animação do esqueleto.

Os *blend shapes* podem definir expressões de toda a região da face ou apenas de pequenas regiões como, por exemplo, o olho ou a boca. Por exemplo, a interpolação entre os *shapes* do olho fechado e do olho aberto cria a animação do personagem a piscar o olho. Criar o *rig* de um modelo facial usando, apenas, *blend shapes* é uma tarefa demorada, uma vez que o artista precisa de modelar um grande número de *shapes* para ter o controlo de toda a região da face. No filme *The Lord of the Rings: The Two Towers*, o *rig* do personagem *Gollum* envolveu a criação de 675 *shapes*. Este processo tem de ser repetido para todos os personagens que se pretende animar, o que consome muito tempo de produção. Orvalho *et al.* (2006) propôs um método para transferir automaticamente os *shapes* entre diferentes personagens, contribuindo, desta forma, para acelerar todo o processo de produção.

Os métodos baseados no esqueleto começam por definir uma estrutura articulada de ossos para toda a região da face. Esta abordagem permite gerar movimentos mais suaves do que o método anterior, mas requer mais tempo de preparação, para se obterem os resultados desejados. Como não se manipulam

³ <http://xface.itc.it/>

directamente os vértices, mas sim os ossos que influenciam esses vértices, a definição do *rig* requer mais tempo de preparação (Ward, 2004).

Na produção de videojogos, este método é muito utilizado em conjunto com técnicas de captura de movimento (Figura 10).



Figura 10: MoCap. Cada sensor colocado na face do actor representa uma marca no modelo 3D

(Copyright 2007 SoftImage)

Na indústria da animação por computador, as grandes preocupações são, naturalmente, os custos e tempos de produção. É fundamental utilizar a melhor técnica de *rigging*. Ambos os métodos baseados em *blend shapes* e “ossos” apresentam vantagens e desvantagens, pelo que, segundo Lewis *et al.* (2000), combinar as duas técnicas é a melhor solução, já que possibilita obter a flexibilidade e suavidade encontradas no método baseado em ossos e a expressividade conseguida com *blend shapes*.

Nas regiões da face onde nem os ossos nem os *shapes* produzem os resultados desejados, é possível adicionar novos *layers* de deformação. Estes novos objectos adicionam mais controlo e conferem um maior realismo à animação. Um exemplo deste tipo de objectos de deformação são as curvas NURBS, que podem ser adicionadas em determinadas regiões para enfatizar algumas características da face, como, por exemplo, rugas, músculos, etc.

Resumindo, um *rig* típico pode incluir os seguintes atributos:

esqueleto: estrutura hierárquica composta por ossos e articulações;

objectos de deformação: podem ser curvas e superfícies NURBS, malhas poligonais, objectos *lattice* e outros. Estes objectos, ligados ao esqueleto, adicionam controlo e realismo às animações e podem representar a geometria de determinados músculos da face;

skinning: o processo de “ligar” os ossos do esqueleto aos objectos de deformação e aos restantes vértices da malha

poligonal. A tarefa mais importante durante esta fase é a definição do mapa de influências que é exercida sobre os vértices durante a deformação;

restrições: devem ser definidas restrições de posição, rotação e escala, de modo a evitar movimentos estranhos ou impossíveis;

shapes: utilizados para definir diferentes *poses* ou expressões faciais. A interpolação de diferentes *shapes* resulta na animação do modelo facial ou, apenas, de determinadas regiões da face.

Muitas vezes, durante o processo de animação as deformações obtidas na expressão facial não são exactamente as esperadas, o que implica voltar atrás e redefinir, de novo, o *rig*. Trata-se, assim, de um processo iterativo que requer muito tempo de produção, experiência e conhecimento da anatomia do personagem (Henry-Biskup, 1998). Um artista experiente pode demorar cerca de 4 semanas para criar o *rig* de um personagem. Automatizar algumas tarefas deste processo é fundamental para acelerar toda a produção da animação. Recentemente, Orvalho e Susin (2007) apresentaram um método para transferir automaticamente o *rig* entre diferentes modelos faciais, independentemente da topologia que os modelos apresentam (Figura 11).



Figura 11: Transferência do rig e animações entre modelos de diferentes proporções
(Orvalho, 2007)

O método desenvolvido difere de outros trabalhos anteriores que usavam técnicas de *morphing* (Blanz e Vetter, 1999) e técnicas de *retargeting*⁴ (Noh e Neumann, 2001), uma vez que estes apenas

⁴ Síntese de movimento através da reutilização de dados

transferiam animações entre personagens, enquanto o novo método transfere, para além das animações, todos os atributos definidos no *rig*. De acordo com os autores, transferir o *rig* dá ao artista uma completa liberdade para manipular posteriormente os personagens, permitindo, assim, a criação de novas animações sem estarem limitados àquelas geradas inicialmente.

4. SUMÁRIO E CONCLUSÃO

Para criar a animação de um personagem 3D é necessário definir, em primeiro lugar, um esqueleto para esse personagem. Um esqueleto é uma estrutura hierárquica composta por ossos e articulações. O processo de criar um esqueleto e “fixá-lo” a um determinado modelo é conhecido como *rigging*, e o algoritmo mais utilizado para fazer esta atribuição entre os ossos do esqueleto e os vértices da malha denomina-se *smooth skinning*. A ideia base é a de que a superfície do modelo (*skin*) é deformada sobre o esqueleto. A animação do esqueleto utiliza técnicas de cinemática e dinâmica. A estrutura da cinemática directa é a mais básica e os esqueletos FK são geralmente usados em cenas com pouco movimento. Os algoritmos IK possibilitam a criação de movimentos mais complexos e realistas e estão implementados na maioria dos *softwares* de modelação e animação 3D. Em determinadas situações, como é o caso das colisões e algumas interações do personagem com o ambiente, já são necessários algoritmos que implementem alguma integridade dinâmica.

Os métodos baseados no esqueleto são os mais utilizados para criar os movimentos principais de um personagem, como é o caso do caminhar, rodar a cabeça ou a cintura, levantar os braços ou as pernas. Métodos baseados em *blend shapes* não são muito utilizados em situações que exijam movimentos de rotação, como é o caso deste tipo de movimentos primários. Os *blend shapes* são mais usados para criar deformações mais subtis, como é o caso de algumas contracções musculares. Ambos os métodos apresentam vantagens e desvantagens, pelo que combinar as duas técnicas é a melhor solução, já que possibilita obter a simplicidade e suavidade encontrada nos métodos baseados no esqueleto e a expressividade conseguida com os *blend shapes*. Em regiões onde

nem os “ossos” nem os *shapes* apresentam bons resultados são usados *layers* adicionais de deformação. São normalmente utilizadas curvas NURBS ou grelhas FFD para simular estas animações mais exigentes de determinadas partes do corpo ou da face.

A parametrização é fundamental para controlar os movimentos dos músculos da face e obter expressões realistas. MPEG-4 foi o primeiro *standard* a definir parâmetros e formalismo para a animação facial e foi baseado no sistema de codificação FACS.

A fase intermédia entre a modelação de um personagem e a sua animação, conhecida como *rigging*, implica a definição de um conjunto de atributos. Este processo exige muito tempo de produção e tem de ser repetido todas as vezes que se pretende animar um novo personagem. O recente trabalho de Orvalho (2007) é dirigido para a área concreta da animação facial e vem, em certa medida, alterar o paradigma existente num *pipeline* de animação. Os autores desenvolveram um método automático para transferir, entre diferentes modelos faciais, todos os atributos definidos no *rig*, incluindo as animações, contribuindo desta forma para uma drástica redução no tempo de produção.

REFERÊNCIAS

- Arikan, O.; Forsyth, D. A.; O'Brien, J. F. (2003) *Motion synthesis from annotations*. In: San Diego, CA, United States.
- Arnaldi, B. *et al.* (1989); “Animation control with dynamics”; *State-of-the-Art in Computer Animation; Proceedings of Computer Animation '89*, 113-23.
- Azevedo, E.; Conci, A. (2003); *Computação gráfica teoria e prática*. Rio de Janeiro: Elsevier.
- Balci, K. (2005); *Xface: Open source toolkit for creating 3D faces of an embodied conversational agent*. In: Lecture Notes in Computer Science.
- Basu, S.; Oliver, N.; Pentland, A. (1998); *3D modeling and tracking of human lip motions*. In: Bombay, India.
- Blanz, V.; Vetter, T. (1999); A morphable model for the synthesis of 3D faces. *Computer Graphics Proceedings. SIGGRAPH 99*, 187-94.
- Burtnyk, N.; Wein, M. (1971); Computer-Generated KeyFrame Animation; 80:3; 149-153.
- Byoungwon, C.; Hanook, L.; Hyeong-Seok, K. (2001); Performance-driven muscle-based facial animation. *Journal of Visualization and Computer Animation*. 12:2; 67-79.
- Byoungwon, C.; Hyeong-Seok, K. (2001); *Analysis and synthesis of facial expressions*

- with hand-generated muscle actuation basis; In: Seoul, South Korea.
- Coquillart, S.; Jancene, P.; (1991); Animated Free-Form Deformation - An Interactive Animation Technique; *Siggraph 91 Conference Proceedings*. 25; 23- 26.
- Ekman, P.; Friesen, W. (1978); Facial Action Coding System; *Consulting Psychologist Press, Palo Alto, CA, USA.*
- Falk, R., *et al.* (2004); Art-directed technology: anatomy of a Shrek2 sequence *ACM SIGGRAPH 2004 Course Notes; Los Angeles, CA.*
- Garstenauer, M.; Kurka, G. (2006); Character Animation in Real-Time. *Magisterarbeit zur Erlangung des akademischen Grades Diplomingenieur im Magisterstudium Informatik. Angefertigt am Institut für Graphische und Parallele Datenverarbeitung; Linz.*
- Ghyme, S.; Kim, K.; Kim, H. (2000); A simplified deformation for real-time 3D character animation *Proceedings of the ACM symposium on Virtual reality software and technology 2000*; 175-182.
- Henry-Biskup, S. (1998); Anatomically Correct Character Modeling. *Game Developer Magazine.*
- Kahler, K.; Haber, J.; Seidel, H. (2001); *Geometry-based muscle modeling for facial animation.* In: Ottawa, Ont..
- Kahler, K., *et al.* (2002); *Head shop: Generating animated head models with anatomical structure.* In: San Antonio, TX, United States.
- Kochanek, D. H. U.; Bartels, R. H. (1984); Interpolating Splines With Local Tension, Continuity, And Bias Control; *Computer Graphics (ACM)*. 18:3; 33-41.
- Lee, Y.; Terzopoulos, D.; Waters, K. (1995); Realistic modeling for facial animation; *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics*; 55-62.
- Lewis, J. P.; Cordner, M.; Fong, N. (2000); *Pose space deformation: A unified approach to shape interpolation and skeleton-driven deformation*; In: New Orleans, LA.
- Liu, C. K.; Popovic, Z. (2002) *Synthesis of complex dynamic character motion from simple animations*; In: United States.
- MacCracken, R.; Joy, K. I. (1996); *Free-form deformations with lattices of arbitrary topology*; In: New Orleans, LA, USA.
- Magenat-Thalmann, N.; Primeau, E.; Thalmann, D. (1988); Abstract muscle action procedures for human face animation; *Visual Computer*. 3:5; 290-7.
- Magenat-Thalmann, N.; Thalmann, D. (1993); Computer Animation. *MIRALab, Centre Universitaire d'Informatique; University of Geneva; Switzerland*, 30.
- Mohr, A.; Gleicher, M. (2003); *Building efficient, accurate character skins from examples.* In: San Diego; CA, United States.
- Mori, M. (1982); The Buddha in the Robot; *Tuttle Publishing*, 192.
- Noh, J.; Neumann, U. (1998); A survey of facial modeling and animation techniques; *University of Southern California Technical Report.*
- Noh, J. Y.; Neumann, U. (2001); Expression cloning; *Proceedings of the ACM SIGGRAPH Conference on Computer Graphics*; 277-288.
- Orvalho V. C. T. (2006); Reusable Facial Rigging and Animation: Create Once, Use Many. PhD Thesis, Faculties of The Universitat Politècnica de Catalunya, Barcelona, June.
- Orvalho, V. C. T.; Zacur, E.; Susin, A. (2006); *Transferring a labeled generic rig to animate face models*; In: Mallorca, Spain.
- Orvalho, V. C.; Susin, A. (2007); Fast and reusable facial rigging and animation *ACM SIGGRAPH 2007 sketches, San Diego, California.*
- Pandzic, I. S.; Forchheimer, R. (2002); The Origins of the MPEG-4 Facial Animation Standard; *MPEG-4 Facial Animation, The Standard, Implementation and Application.* John Wiley & Sons Ltd.

- Parke, F. I. (1974); *A parametric model for human faces*. Univ. Utah, Salt Lake City, UT, USA.
- Parke, F. I. (1972); Computer generated animation of faces; *In Association For Computing Machinery, Proceedings Of The Acm Annual Conference, Boston, August 1972. P. 451-457. 6 Illus. 14 Ref. Work Supported In Part By Department Of Defense Contract F30602-70-c-0300. See Isa 73-479/y. 1972.*
- Platt, S. M.; Badler, N. I. (1981); Animating Facial Expressions; 15:3; 245-252.
- Sederberg, T. W.; Parry, S. R. (1986); Free-Form Deformation Of Solid Geometric Models; *Computer Graphics (ACM)*; 20:4; 151-160.
- Sifakis, E.; Neverov, I.; Fedkiw, R. (2005); Automatic determination of facial muscle activations from sparse motion capture marker data; *ACM Transactions on Graphics*; 24:3; 417-25.
- Sloan, P. P. J.; Rose I. C. F.; Cohen, M. F. (2001); *Shape by example*. In: Research Triangle Park, NC.
- Sy-sen, T.; Liew, A. W. C.; Hong, Y. (2004); *Lip-sync in human face animation based on video analysis and spline models*; In: Brisbane, Qld., Australia.
- Terzopoulos, D.; Waters, K. (1990); Physically-based facial modelling, analysis, and animation; *Journal of Visualization and Computer Animation*; 1:2; 73-80; Unreal, T. - Unreal Engine 3. Dezembro 2007; <<http://www.unrealtechnology.com/html/technology/ue30.shtml>>
- Wang, X. C.; Phillips, C. (1987); *Multi-weight enveloping: Least-squares approximation techniques for skin animation*; In: San Antonio, TX, United States.
- Waters, K. (1987); Muscle Model For Animating Three- Dimensional Facial Expression; *Computer Graphics(ACM)*; 21:4; 17-24.
- Watt, A.; Watt, M. (1992); *Advanced Animation and Rendering Techniques Theory and Practice 0008*. Wokingham; Addison-Wesley Publishing Company.
- Weber, J. (2000); Run-Time Skin Deformation; *Game Developers Conference Proceedings (GDC 2000)*; 703-721.
- Welman, C. P. (1993); Inverse kinematics and geometric constraints for articulated figure manipulation; *M.Sc.*
- Zhao, W., *et al.* (2003); Face recognition: A literature survey; *ACM Computing Surveys*; 35:4; 399-458.