Motion Capture Facial Baseado em Aquisição Estereoscópica

Rodrigo Assaf[1] José Hiroki Saito[1]

Departamento de Computação – Universidade Federal de São Carlos (UFSCar)
São Carlos, SP – Brasil

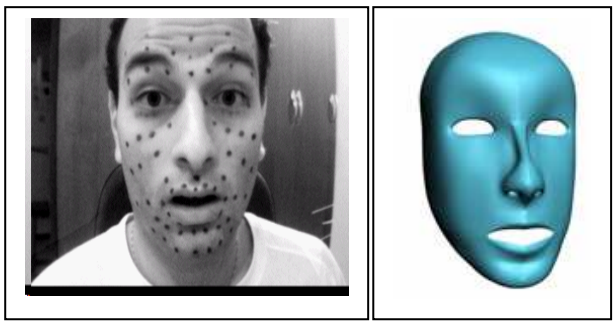

Figura 1: Expressão facial obtida através do processo de *motion capture* baseado em aquisição estereoscópica utilizando duas câmeras.

Abstract

Facial animation has been widely used in many areas of applications, presenting high levels of details and technology. However, its implementation still demands lots of hard work from the producer team. Aiming at the adaptation of a stereoscopic system to a financial low cost facial animation system, this research is based on facial control points on a human face, which is 3D modeled using the box modeling technique. The 3D virtual face is used to apply an ending result of a facial expression, or lip sync, obtained from the methods of stereoscopy, association and interpolation.

Keywords: facial animation, 3D, *motion capture*.

Authors' contact:
rodassaf@terra.com.br,
saito@dc.ufscar.br

Resumo

A animação facial tem sido amplamente utilizada em diversas áreas de aplicação e entretenimento, apresentando altos níveis de detalhamento e avanços tecnológicos. Contudo, sua implementação ainda exige um trabalho árduo da equipe de produção podendo desprender muito tempo nesta etapa. Com o objetivo de adequar um sistema estereoscópico a um sistema de animação facial de baixo custo financeiro, este trabalho baseia-se na demarcação de pontos de controle faciais em uma face humana real, que foi modelada em 3D utilizando a técnica de box modeling, para a realização dos métodos de estereoscopia, associação e interpolação obtendo uma expressão facial e de sincronismo labial.

1. Introdução

A animação facial, bem como outras técnicas de animação em geral, tem sido amplamente usada em diversas áreas de aplicação, tais como: cinema, jogos, desenhos animados, medicina, ambientes virtuais, Internet, entre outras.

Na área de cinema, o desenvolvimento da animação facial tem mostrado um alto nível de detalhamento e avanço tecnológico perante as outras áreas, em razão à alta resolução requerida nos filmes digitais e ao tempo de produção do produto, ou seja, os filmes não são processados em tempo real. Com isso, as animações faciais mais complexas podem passar horas ou talvez meses no estágio de produção final (estágio que consiste em gerar a animação digital final com todos os elementos da animação juntos na cena, incluindo iluminação, reflexos, entre outros efeitos) para depois serem apresentadas ao público.

Em contrapartida, existe uma problemática no processamento em tempo real na arte de animar faces e expressões em jogos eletrônicos, Internet, ambientes virtuais, entre outros meios. Nos jogos eletrônicos tridimensionais, por exemplo, a animação dos personagens ou dos objetos da cena ainda é realizada em baixa quantidade de polígonos (abaixo de 10.000 polígonos, considerando a média de polígonos utilizados em personagens de jogos) deixando, portanto, as animações faciais simplificadas em vista do elevado número de músculos faciais envolvidos (na ordem de 40 músculos) e da qualidade gráfica do jogo em tempo real.

Tradicionalmente, a animação facial por computador segue dois procedimentos básicos: (1) construção de uma malha 3D da face, e (2) animação da malha 3D em alguma forma de controle na qual simule ações faciais.

No procedimento (1), é desejável a construção de uma malha com o máximo de detalhes, como peles

dobráveis, rugas, entre outros, quando for construída em alta quantidade de polígonos (acima de 10.000 polígonos, considerando o nível de detalhe perceptível ao olho humano). Contudo, se a aplicação exigir menos quantidade de polígonos (abaixo de 10.000), o desafio será de fazer uma malha 3D da face convincente. Para isso, outros recursos são utilizados, como textura e iluminação.

A modelagem da malha facial pode ser realizada manualmente através de técnicas de modelagem conhecidas como *Box modeling* e *Patch*, entre outras. Ambas as técnicas possuem características distintas, porém geram os mesmos resultados. A técnica de *Box modeling* consiste em começar a modelagem de uma face, a partir de uma forma primitiva básica como um cubo ou uma esfera chegando em sua forma final através de extrusões de arestas e translações dos vértices das faces poligonais da mesma. A técnica *Patch*, visa usar formas geométricas 2D, usando uma imagem referência, formando assim, a forma 3D final através de translação de vértices.

A modelagem da malha facial também poderá ser realizada, usando uma face real como informações iniciais, automaticamente através de métodos estereoscópicos (usando duas ou mais fotos) ou através de *scanners* à laser e de luz estruturada. A grande vantagem destes métodos automáticos é a facilidade quanto à modelagem. Porém, para se obter uma boa precisão é necessário equipamentos de altos custos e algoritmos robustos para resolver problemas como a oclusão, eventuais ruídos, entre outros.

No procedimento (2), o animador deve decidir quais vértices da malha deverão ser movimentados para que se atinja uma animação facial adequada e convincente. Vários paradigmas foram propostos como: modelos parametrizados [PARKE, 1989], modelos com pontos de controle [KURIHARA & ARAI, 1991], modelos musculares com *Kinematic* [WALTERS, 1987], modelos *spline* [NAHAS, 1990], modelos de rastreamento [WILLIAMS, 1990], modelos musculares dinâmicos [PLATT & BARDLER, 1981].

Outros sistemas de animação facial, utilizando diferentes paradigmas, são encontrados no mercado ou no meio acadêmico. Pode-se citar o modelo muscular de Parke [PARKE & WATERS, 1996], onde há o controle individual de certos músculos faciais, o Facade (Universidade de Stanford) e o *CrazyTalk* cuja animação labial é realizada através da sincronização das ondas da voz (*Speech Driven*) e o sistema de animação facial da Universidade de Ohio [SOMASUNDARAM, 2005], cujos pontos de controle são refletores de luzes fixados na face de uma pessoa.

O paradigma utilizado neste trabalho foi o de animação facial baseada em desempenho real. Este envolve o uso de informações derivadas de ações de um modelo humano real para serem usadas em um modelo virtual 3D.

A animação facial pode ser realizada usando mecanismos de captura de expressões reais como o *motion capture*, que consiste na criação de uma representação 3D, a partir de uma performance real humana utilizando emissores (ópticos, mecânicos, magnéticos ou acústicos) que geram sinais captados por sensores estrategicamente distribuídos. Softwares especializados, que podem receber e "entender" os dados captados, geram as coordenadas tridimensionais.

2. Trabalhos Relacionados

2.1 *Motion Capture* Óptico da Universidade de Ohio

O sistema por *motion capture* óptico da Universidade de Ohio é usado para simular os principais músculos faciais. Utiliza-se 85 fontes que serão captadas por 14 câmeras e depois repassadas para um software que manipulará os dados (SOMASUNDARAM, 2005).

Os emissores deste sistema, em geral, são refletores (emissores LED) que são posicionados em qualquer parte do corpo para a captação do movimento. Diversas câmeras (em torno de doze) são posicionadas no ambiente para a captação do movimento e a geração das coordenadas 2D. Após este procedimento, *softwares* irão realizar soluções geométricas (de acordo com a disposição das câmeras de captura) para fornecer as coordenadas 3D.

As taxas de amostragem destes sistemas ópticos podem atingir até 200 quadros por segundo, porém seu custo financeiro é muito elevado (US$ 20,000 a US$ 150,000). Outra característica do sistema óptico, é a liberdade do equipamento, ou seja, o usuário que simulará movimentos reais não ficará coberto por fios e outros equipamentos que prejudiquem seu movimento (SILVA, 1997).

Algumas desvantagens podem ser destacadas como a oclusão, que pode ser solucionado com o uso de várias câmeras. Além disso, ruídos e erros de precisão podem ocorrer se o sistema não estiver bem configurado.

A tabela 1 mostra as principais características dos sistemas ópticos.

Tabela 1: Características dos sistemas ópticos (Fonte: SILVA, 1997).

	Sistemas Ópticos
Resolução espacial	~ 0.0015 a 0.2 % do campo de visão (CV)
Acurácia espacial	~ 0.004 a 0.5 % do CV
Precisão	~ 0.0055 a 0.02 % do CV
Taxa de amostragem	> 200 quadros por segundo
Campo de trabalho	raio de 1 a 4 m
Exemplos de sistemas	ExpertVision HiRES 3-D System, Multitrax Motion Capture System, OPTOTRAK
***Softwares* compatíveis**	ZoeTrax, The Creative Motion Editor, PowerAnimator V7

3. Objetivos

O objetivo geral deste trabalho consiste em animar uma face 3D, previamente obtida por métodos de modelagem, através de uma performance humana real captada por meio de duas câmeras utilizando o procedimento da estereoscopia.

Além do objetivo geral, vale destacar outros objetivos específicos como: (a) determinar métodos para atingir a animação facial a partir de um modelo real humano usando o método da estereoscopia; (b) atingir um grau de resultado e realismo visivelmente de acordo com a performance humana real; (c) animar expressões e *Lip-Sync* (sincronização labial) utilizando a técnica de interpolação; (d) aplicar a performance humana facial em faces de personagens 3D; (e) desenvolver métodos para um sistema de baixo custo financeiro.

4. Estereoscopia

A estereoscopia é fundamentada no sistema de visão humano, que através do uso concomitante de dois olhos consegue perceber a profundidade.

Usando soluções geométricas entre os olhos e o valor da disparidade (nome dado à separação física de pontos semelhantes em duas imagens estereoscópicas), pode ser calculada a profundidade (eixo Z) dos pontos da cena.

A estereoscopia é uma técnica passiva, não sendo necessário o uso de sensores para detecção de profundidade. Seu uso estende-se a diversas áreas como cartografias, reconhecimento de objetos industriais e navegação de robôs portáteis.

4.1 Sistema Estereoscópico Adotado

Conforme a figura 2, a estereoscopia envolve o uso de no mínimo duas imagens separadas e um único ponto **P** (pertencente ao objeto real). A distância entre os centros das duas lentes é chamado de linha base (*baseline*), e o objetivo é encontrar as coordenadas (*X, Y, Z*) do ponto **P** tendo dois pontos de imagens P_1 (x_1, y_1) e P_2 (x_2, y_2).

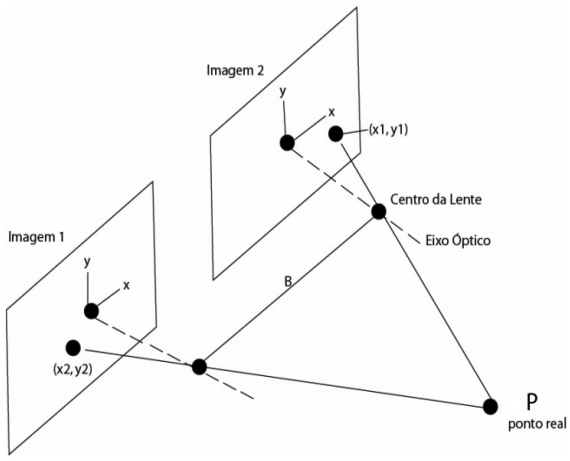

Figura 2. Modelo de processamento de estereoscopia.

Assume-se que as câmeras são iguais e que o sistema de coordenadas de ambas as câmeras estão perfeitamente alinhadas, diferindo apenas na localização de suas origens. Portanto, a coordenada *Z* de **P** é exatamente a mesma para ambos os sistemas de coordenadas das câmeras.

Através de uma relação geométrica do sistema estereoscópico, é conhecido que:

$$X_1 = \frac{x_1}{\lambda}(\lambda - Z_1) \quad (1)$$

$$X_2 = \frac{x_2}{\lambda}(\lambda - Z_2) \quad (2)$$

sendo lambida igual a distância focal das câmeras utilizadas.

Entretanto, por causa da separação entre as câmeras, e visto que a coordenada Z de **P** é semelhante para ambos sistemas de coordenadas das câmeras, segue que,

$$X_2 = X_1 + B \quad (3)$$

$$Z_2 = Z_1 = Z \quad (4)$$

sendo *B* a linha base, substituindo as equações 3 e 4 nas equações 1 e 2 tem-se:

$$X_1 = \frac{x_1}{\lambda}(\lambda - Z) \quad (5)$$

$$X_1 + B = \frac{x_2}{\lambda}(\lambda - Z) \quad (6)$$

Segundo [GONÇALEZ & WINTZ, 1987], A relação entre o eixo Z e os demais parâmetros do sistema é dada por:

$$Z = \lambda - \frac{\lambda B}{x_2 - x_1} \quad (7)$$

4.2 Definição dos Métodos para um Sistema de *Motion Capture* Facial baseado em Estereoscopia

As etapas propostas para a conclusão de uma animação facial 3D, baseada na demarcação de pontos de controle faciais em uma face humana real (paradigma de performance humana real), que é previamente modelada em 3D utilizando a técnica de *box modeling são:*

4.2.1. Desempenho do movimento facial humano
4.2.2. Captura de quadros
4.2.3. Seleção de quadros
4.2.4. *Tracking* dos pontos correspondentes
4.2.5. Aquisição volumétrica: Estereoscopia
4.2.6. Modelagem facial 3D
4.2.7. Associação (*Fitting*)
4.2.8. Interpolação
4.2.9. Animação facial

Para a captação dos movimentos faciais do modelo real humano (etapa 4.2.1), faz-se necessário a

demarcação de pontos de controles na face de uma pessoa (Figura 3). Este trabalho considerou 66 pontos distribuídos na face, os quais abrangem todos os principais músculos do movimento facial. É importante ressaltar que os músculos faciais orbitais são simulados na modelagem 3D para facilitar a movimentação dos vértices no espaço tridimensional. A realização e captura dos movimentos faciais do modelo real foram realizadas a uma distância fixa de 40cm (sem mobilidade) da face para com o sistema de câmeras, tendo como relevâncias, o enquadramento total das faces em ambos canais esquerdo e direito, e o problema da oclusão.

Figura 3. Pontos demarcados no rosto humano.

Após a demarcação dos pontos de controle, dá-se o início da etapa 4.2.2, que é a de captura de quadros a partir do sistema de duas câmeras sincronizadas. A saída desta gravação gera imagens ou quadros correspondentes (da câmera da esquerda e da direita) de expressões faciais feitas pelo modelo humano.

A seleção de quadros (etapa 4.2.3), que ocorre logo após a captura, é considerada uma etapa de extrema importância para a animação final, pois é nela que se escolhem as poses-chave (Figura 4). A escolha de poses certas resulta em uma animação com pouca necessidade de cálculos estereoscópicos melhorando seu desempenho. Contudo, para animações mais complexas, é necessário a escolha de um maior número de quadros, suficientes para garantir que a animação gerada corresponda ao movimento real.

A escolha dos quadros é realizada visualmente, através da habilidade técnica do animador, caracterizando um processo artístico.

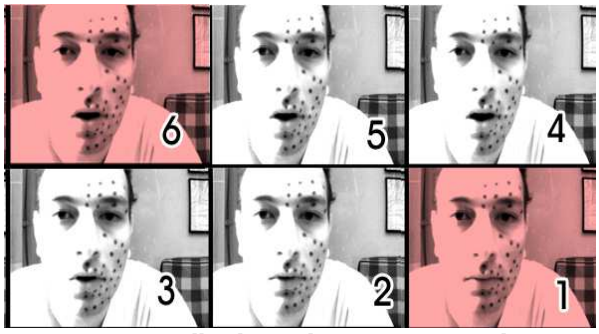

Figura 4. Escolha de quadros para a interpolação.

O *tracking* dos pontos correspondentes (etapa 4.2.4) é o processo de localização destes pontos durante a movimentação humana. Para cada par de quadros gerado pelo sistema estereoscópico, foi realizada a busca por pontos semelhantes de maneira manual através de uma biblioteca gráfica do MatLab, assim como é demonstrado na figura 5.

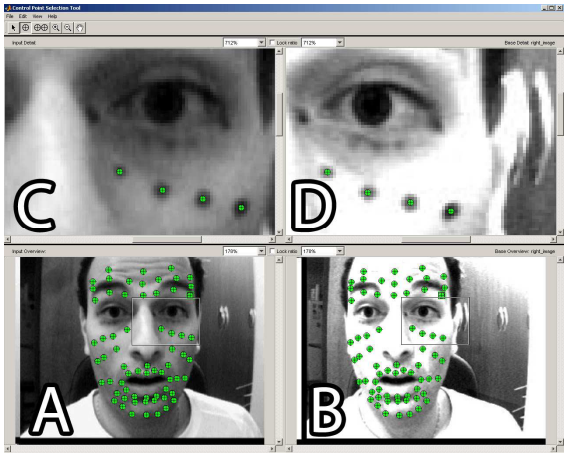

Figura 5. Escolhendo os pontos correspondentes.

Tendo as localizações dos pontos correspondentes (*X* e *Y*) em todas os quadros correspondentes, faz-se necessário a aquisição volumétrica desses, através do método de estereoscopia. O algoritmo estereoscópico recebe dados de entrada como a distância focal das câmeras, a distância física entre elas e a resolução das imagens geradas, para seus cálculos. Após estes dados, o algoritmo faz a leitura do par de imagens correspondentes e em seguida abre este par para a seleção dos pontos correspondentes (*tracking*). Tendo gerado uma matriz contendo os valores X e Y dos pontos de controle do par estereoscópico, o algoritmo realiza a translação da origem do sistema de coordenadas das imagens para o seu centro, ou seja, para a média da largura e altura. Tendo sua origem alterada, os dados X e Y captados no *tracking* são atualizados e assim, é realizado o cálculo estereoscópico e armazenado em uma matriz. contendo as localizações (X, Y e Z) dos pontos demarcados da face humana, de cada par de imagens correspondentes.

A etapa 4.2.5, consiste na modelagem de uma face 3D, utilizando a técnica de *box modeling*, a partir da referência do rosto proposto (Figura 6). A face construída é constituída de 734 polígonos (triângulos) em sua modelagem e conta com 66 pontos, ou vértices, de controle para a animação facial. Vale ressaltar que estes vértices de controle são associados às demarcações feitas no rosto humano (etapa 4.2.6), ou seja, todo o movimento realizado por um determinado ponto facial localizado na figura 7.a deve-se ser aplicado ao seu correspondente vértice da face 3D da figura 7.b. Assim, o movimento realizado pelo desempenho humano é associado ao modelo virtual.

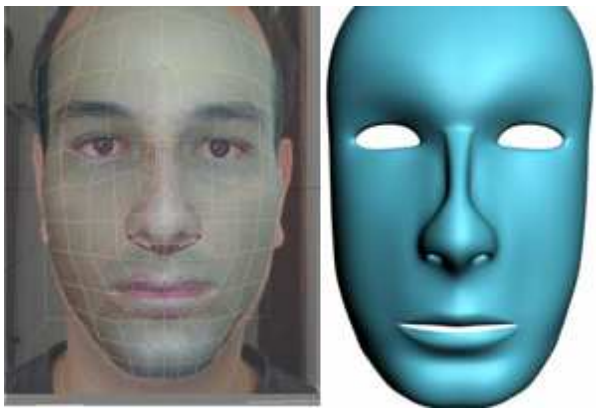

Figura 6. Modelagem 3D da face humana.

(a) **(b)**

Figura 7. (a) Pontos identificados na face real; (b) Mesmos pontos associados nos vértices da face 3D. Total de 66 pontos de controle.

A interpolação, técnica da etapa 4.2.7, é utilizada para a transição entre os quadros ou poses-chave geradas pelas coordenadas do algoritmo de estereoscopia. É nesta etapa em que o animador consegue controlar o tempo de animação a ser usado.

Para a animação final (etapa 4.2.8) propõe-se a utilização da técnica de interpolação entre os quadros-chave, gerados pelas coordenadas do algoritmo de estereoscopia (etapa 7). Esta técnica proporciona o uso de modificadores *morphings* para expressões faciais e para sincronismo labiais.

4.3 Câmeras

Para a gravação de movimentos faciais reais, foram utilizadas duas câmeras do tipo segurança, monocromáticas, com duas saídas RCA de vídeo e som e com uma entrada de força para cada uma. As lentes das câmeras possuem uma distância focal de 3 mm e suas saídas de vídeo possuem uma resolução de 320 x 240 *pixels*.

De acordo com a figura 8, o cabo denominado A, de cor vermelha, é caracterizado pela entrada de energia, o cabo B, de cor branca, é caracterizado pela saída do áudio e por fim, o cabo C, de cor amarela, é caracterizado pela saída de vídeo.

Figura 8. Duas câmeras de segurança monocromáticas.

A captação e digitalização da seqüência de quadros (*frames*) foram realizadas por meio de duas placas gráficas com sinal de entrada de vídeo RCA.

4.4 Sincronização entre as Câmeras

O método estereoscópico avalia dois quadros iguais, tirados no mesmo intervalo de tempo, para a realização de seus cálculos, porém, sem uma sincronização, quadros semelhantes podem ser captados em tempos diferentes. Para evitar este problema, fez-se necessário uma sincronização entre as câmeras, ou seja, para que ambas gerassem vídeo ao mesmo tempo.

Optou-se pelo procedimento de sincronização por *hardware* devido aos fatores de desempenho, pois se esta sincronia fosse realizada via *software*, a execução desta etapa não teria o mesmo resultado.

Através do uso de um osciloscópio, foi realizada a ligação entre os conectores de sincronismo entre uma câmera (*FSI – Field Sync In*) a outra (*FSO – Field Sync Out*). A figura 9 mostra o esquema da ligação de sincronismo entre duas câmeras.

Após o término deste procedimento, a montagem das câmeras foi realizada, assim como a determinação de uma distância física de 40mm entre suas respectivas lentes ao longo de um plano. Esta distância foi determinada levando-se em consideração a resolução da câmera e a capacidade de enquadramento de uma face inteira em ambas as câmeras, evitando assim o problema da oclusão.

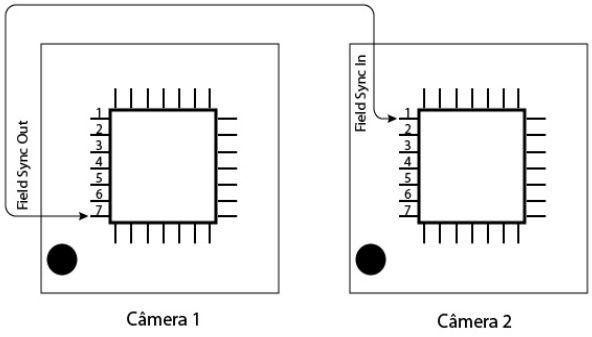

Figura 9. Identificação do FSO e FSI e a ligação entre eles.

5. Resultados

A figura 10 mostra a captação do movimento facial humano pela câmera da esquerda (13 quadros), em ordem cronológica, da esquerda para a direita, e de cima para baixo, da expressão de surpreso.

A figura 11 mostra a captação do movimento facial humano pela câmera da direita (13 quadros), na mesma ordem da figura 9.

A animação do movimento facial "surpreso" foi realizada com 13 quadros, porém, nem todos necessitam estar presentes no cálculo de estereoscopia, uma vez que a interpolação é uma solução mais eficaz para este problema. Contudo, como é observado no caso desta expressão, o vídeo indica que o movimento de sobrancelhas ocorreu antes do movimento de abertura de boca. Este fenômeno não pode ser compreendido na interpolação, caso os quadros-chave escolhidos sejam os extremos (o primeiro e o último). Porém, se mais quadros-chave forem escolhidos indicando este movimento, este problema pode ser solucionado. Vale ressaltar que a interpolação aproxima o movimento realizado entre os quadros-chave. Portanto, para uma maior proximidade do movimento real, deve-se adotar todos os quadros-chave captados na filmagem.

Figura 10. Expressão surpreso captado pela câmera esquerda.

Figura 11. Expressão surpreso captado pela câmera direita.

A figura 12 mostra a seleção de quadros-chave da câmera esquerda (2 quadros extremos) e a figura 13 mostra a seleção de quadros-chave da direita (2 quadros extremos):

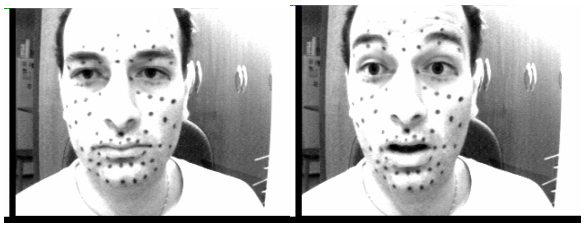

Figura 12. Seleção de quadros-chave da câmera esquerda.

Figura 13. Seleção de quadros-chave da câmera direita.

Após a seleção dos quadros-chave, inicia-se a etapa de estereoscopia. Esta abrange a seleção dos pontos correspondentes previamente marcados na face humana entre os pares de quadros equivalentes. Assim, o primeiro quadro-chave proveniente da câmera esquerda e da direita devem ser as entradas para o cálculo da estereoscopia.

É importante ressaltar que a escolha dos pontos correspondentes entre os quadros foi realizada de forma manual, ou seja, o usuário seleciona os pontos da imagem da esquerda e os associa com os mesmos na imagem da direita um a um.

De acordo com a figura 6, a ordem de escolha dos pontos leva em consideração o processo de *fitting*, ou seja, a indexação dos pontos na face.

A figura 14 mostra a associação no modelo 3D e a interpolação obtendo o resultado final.

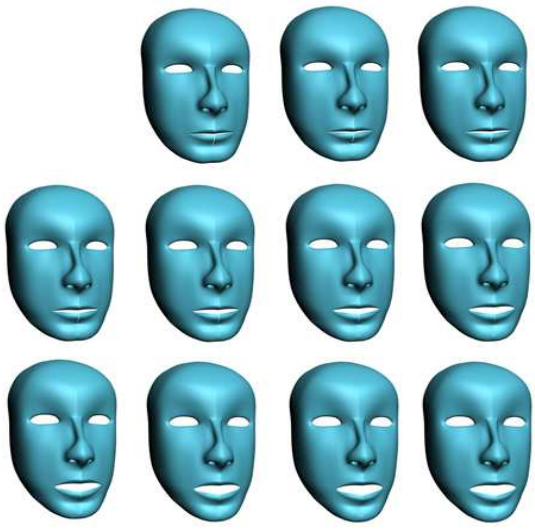

Figura 14. Resultado interpolado numa face 3D.

6. Conclusões

O desenvolvimento do método de estereoscopia, na área de animação facial, contribui para uma automatização e facilidade de implementação na mesma, permitindo um resultado de uma animação facial realista.

A precisão do eixo Z resultou em uma variação particular para cada expressão realizada. Expressões de menor intensidade geraram pequenas variações (na ordem de 2 pixels) no eixo Z, enquanto que expressões de maior intensidade resultaram em maiores variações dos valores de Z (na ordem de 8 pixels).

Para cada expressão facial realizada neste trabalho, foram usados não mais do que três quadros para se atingir uma animação realística. Contudo, o número de quadros a ser usado para a realização de uma animação facial, pode variar para cada caso, de acordo com a necessidade do animador.

Como há esta interação com o sistema de animação facial, a taxa de amostragem deste método de captura não chegaria a mais de 30 quadros por segundo, devido à limitação das câmeras utilizadas. A grande vantagem desta característica, é que o sistema não produz informações exageradas e inúteis, facilitando a tarefa de retirar estes quadros, em excesso, produzidos na animação antes dos cálculos de estereoscopia.

O método proposto de estereoscopia mostrou-se eficaz (de acordo com os resultados visualmente obtidos da figura 14) para a obtenção de informações volumétricas a partir de imagens, mesmo utilizando câmeras de baixa resolução (320 x 240 *pixels*). Também é importante ressaltar que pequenas variações do eixo Z dão um grau de realismo não alcançado pela utilização de apenas uma câmera devido a pequenas variações musculares que ocorrem na face humana.

Objetivando um sistema de baixo custo financeiro e de experimentação, as câmeras utilizadas no projeto, custaram em torno de R$ 40,00. Contudo, a precisão deste método pode alcançar melhores níveis com o uso de câmeras de maior qualidade.

A tabela 2 mostra uma comparação entre o sistema de animação facial por *motion capture* óptico da Universidade de Ohio (SOMASUNDARAM, 2005) com o sistema proposto.

Tabela 2: Comparação entre o sistema óptico de *motion capture* da Universidade de Ohio (SOMASUNDARAM, 2005) com o sistema proposto.

	Motion Capture de Ohio	**Sistema Proposto**
Caract. principais dos sistemas	Conjunto de câmeras, fazendo o *tracking* 2D dos refletores para posterior conversão para 3D	2 Câmeras que capturam os pontos marcados na face humana.
Dados obtidos	Posição *XYZ* dos refletores	Posição *XYZ* dos pontos marcados no rosto
Taxa de amostragem	200 quadros por seg. [60]	Até 30 quadros por seg.
Proc. de dados	*Tracking* via software dos conjuntos de posições 2D para posterior geração de coordenadas 3D, não sendo um sistema de tempo real.	Idem
Prinp. Vantagens	Alta taxa de amostragem (até 200 quadros por segundo) , número de refletores ilimitado, liberdade dos atores durante o processo	Baixo custo financeiro, nro. alto de marcadores sem atrapalhar o movimento e geração de quadros-chave apenas
Principais desvantagens	Oclusão de refletores, necessita *tracking* via *software* para geração de coordenadas 3D, alto custo do equipamento e excesso de quadros produzidos	Baixa taxa de amostragem (até 30 quadros por segundo), necessita *traking* via *software* para geração de coordenadas 3D, oclusão dos pontos marcados no corpo
Preço médio	U$ 20.000,00 a	R$ 1000,00

equip.	150.000,00 (SILVA,1997)	(valor dos materiais utilizados)
Interatividade com o sistema	Após a captura de informações volumétricas gerando excesso de dados a serem excluídos na fase de produção.	Antes dos cálculos das informações volumétricas, produzindo apenas quadros-chave.

O sistema estereoscópico pode ser caracterizado pelas contribuições de:

- Re-uso e portabilidade: a captura da seqüência da animação facial é dada na forma de dados, ou seja, o uso destes para um outro personagem ou uma outra cena é possível ao carregar o arquivo de dados.
- Escolha de quadros chave: a produção de quadros chave na animação se dá previamente ao cálculo estereoscópico. Esta característica faz com que o sistema produza apenas os quadros escolhidos, economizando cálculos na fase do *tracking* e na aquisição tridimensional. O método de escolha dos quadros chave é visual fazendo com que este processo se torne artístico.
- Criação de *Targets Morphing*: as expressões faciais podem ser capturadas e armazenadas em formas de *Target Morphings*, ou seja, uma animação pode ser repetida, ou modificada, diversas vezes.
- Sistema de baixo custo financeiro

6.1 Limitações

- Alta quantidade de dados: pode ocorrer se não houver uma seleção de quadros-chave, tornando a animação mais próxima da performance humana, porém, mais carregada para se editar.
- Baixa qualidade das câmeras: O uso de câmeras de baixo custo financeiro resultou em imagens com diferentes e poucos contrastes, prejudicando assim a utilização de um algoritmo de busca.
- Movimento da cabeça: o sistema de câmeras fixas pode gerar problemas devido à movimentação natural da cabeça.
- Automatização do *Fitting*: para um melhor desempenho, faz-se necessário a automatização da etapa de *fitting*.
- Precisão do algoritmo de *tracking*: prejudica o desempenho do sistema, pois quando o mesmo não consegue achar a posição de um marcador, ele entra em modo manual pedindo ao usuário a seleção do ponto.

6.2 Trabalhos Futuros

- Melhoria no algoritmo: melhoria e automatização do algoritmo de *tracking*, dando maior precisão e implementação do algoritmo de *fitting* para um maior desempenho do sistema.
- Portabilidade dos dados: os dados dos quadros-chave provenientes da etapa da estereoscopia devem comunicar-se com o sistema, podendo ser carregado em outras faces 3D, manipulados e salvos com uma extensão proprietária.
- Melhoria na precisão da estereoscopia: para a melhoria da precisão do sistema, faz-se necessário a aquisição de câmeras de melhor qualidade e em maior número destas.

Bibliografia

PARKE, F. PARAMETERIZED MODELS FOR FACIAL ANIMATION REVISITED. EM: SIGGRAPH FACIAL ANIMATION TUTORIAL NOTES, P 43–56. ACM SIGGRAPH, 1989.

KURIHARA, T.; ARAI, K. A TRANSFORMATION METHOD FOR MODELING AND ANIMATION OF THE HUMAN FACE FROM PHOTOGRAPHS. EM: STATE OF THE ART IN COMPUTER ANIMATION, SPRINGER-VERLAG, P. 45–57, 1991.

WATERS, K. A MUSCLE MODEL FOR ANIMATING THREE-DIMENSIONAL FACIAL EXPRESSION. COMPUTER GRAPHICS, V. 22, N. 4, P 17–24, 1987.

YAU, J.; DUFFY, N. 3-D FACIAL ANIMATION USING IMAGE SAMPLES. EM: NEW TRENDS IN COMPUTER GRAPHICS, SPRINGER-VERLAG, P 64–73, 1988.

NAHAS, M. ET. AL. FACIAL IMAGE SYNTHESIS USING SKIN TEXTURE RECORDING. VISUAL COMPUTER, V. 6, N. 6, P. 337–343, 1990.

WILLIAMS, L. PERFORMANCE-DRIVEN FACIAL ANIMATION. EM: SIGGRAPH 24, ACM COMPUTER GRAPHICS, P. 235–242, 1990.

PLATT, S.; BADLER, N. ANIMATING FACIAL EXPRESSION. COMPUTER GRAPHICS, V. 15, N. 3, P. 245–252, 1981.

LI, Z. ET. AL. SPACETIME FACES: HIGH RESOLUTION CAPTURE FOR MODELING AND ANIMATION. UNIVERSITY OF WASHINGTON, 2005.

SOMASUNDARAM, A. REALISTIC FACIAL ANIMATION DURING SPEECH. CSE-ACCAD, THE OHIO STATE UNIVERSITY, 2005. DISPONÍVEL EM: HTTP://ACCAD.OSU.EDU/RESEARCH/MOCAP/PAPERS/ ARUN_ACCADSUMMARY.PDF. DATA DE ACESSO: 07/07/2005.]

SILVA, F. MOTION CAPTURE: INTRODUÇÃO À TECNOLOGIA. LABORATÓRIO DE COMPUTAÇÃO GRÁFICA, LCG. COPPE, UNIVERSIDADE FEDERAL DO RIO DE JANEIRO, UFRJ, 1997.

GONZALEZ, R. C.; WINTZ, P. A. DIGITAL IMAGE PROCESSING. ADDISON – WESLEY, 1987, 2ª ED.

PARKE F. I.; WATERS, K. COMPUTER FACIAL ANIMATION. MASSACHUSETTS: A K PETERS, 1996.