

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DEPARTAMENTO DE CIÊNCIA DA COMPUTAÇÃO

Reconstrução de Faces 3D usando Análise de Componentes Principais

– Texto de Qualificação de Doutorado –

Candidato

Jesús P. Mena-Chalco

Orientador

Roberto Marcondes Cesar Junior

Resumo

O processo de reconstrução de modelos faciais 3D (geometria facial) dada uma fotografia 2D (textura facial) é um tópico relevante na área de *Fotografia Digital e Visão Computacional*, e que recentemente vem recebendo especial interesse na comunidade científica. Neste trabalho, apresentamos um sistema de *fotografia facial em 3D* baseada em um banco de dados de expressões faciais compostos de geometria e textura facial. O sistema proposto permite obter uma representação de geometria facial 3D dada uma fotografia 2D. Os dados correspondentes à fotografia 2D sofrem uma série de transformações através de espaços de textura e geometria previamente estimados. Na etapa de treinamento do sistema, pontos característicos faciais das amostras do banco de dados, e análise de componentes principais (PCA), são usados para representar o banco de dados completo, definindo uma base ortonormal de textura e outra de geometria. Na etapa de reconstrução de uma face, a textura facial com seus correspondentes pontos característicos são usados para projetar a face de entrada na base de geometria obtida no treinamento.

Testes experimentais, considerando um banco de dados criado especialmente para a análise de expressões faciais 3D, mostram reconstruções computacionalmente rápidas, mantendo coerência espacial similar à percepção humana, corroborando assim a eficiência e aplicabilidade do sistema proposto.

Palavras-chave: Face 3D, reconstrução facial, análise de componentes principais.

São Paulo, Setembro de 2009

Sumário

Lista de figuras	iii
Lista de tabelas	iv
1 Introdução	1
1.1 Objetivos	2
1.2 Organização do trabalho	2
2 Conjunto de dados de faces 3D	3
2.1 Equipamento	3
2.2 Protocolo de aquisição	4
2.3 Conjunto de dados IMPA-FACE3D	6
3 Reconstrução facial 3D através de espaços PCA	9
3.1 Modelo de face	10
3.2 Modelo matemático proposto	11
3.3 Método de reconstrução facial 3D	12
3.3.1 Treinamento	12
3.3.2 Reconstrução	14
4 Resultados	17
4.1 Reconstrução de amostras presentes no treinamento	17
4.2 Reconstrução de amostras não presentes no treinamento	19
4.3 Reconstrução de amostras com expressões não presentes no treinamento	20
5 Próximos passos	22
5.1 Utilização de um novo modelo de face 3D	22
5.2 Cronograma de atividades	24
6 Conclusões	27
Bibliografia	28

Lista de figuras

2.1	Equipamento considerado na aquisição de dados	5
2.2	Exemplo de captura de dados: textura e geometria	8
3.1	Fluxograma do sistema de reconstrução facial 3D	10
3.2	Exemplo de marcos característicos faciais e triangulação considerada na proposta	10
3.3	Formação do conjunto de treinamento	12
3.4	Espaços de textura e geometria	13
4.1	Expressões faciais do conjunto de dados considerado na reconstrução . . .	18
4.2	Face média obtida na etapa de treinamento	18
4.3	Reconstrução facial 3D de amostras presentes no treinamento	19
4.4	Reconstrução facial 3D de amostras não presentes no treinamento	20
4.5	Reconstrução facial 3D de amostras com expressões faciais não presentes na etapa de treinamento	21
5.1	Novo modelo de face inicial	22
5.2	Exemplificação de subdivisão e projeção do modelo inicial	23
5.3	Exemplo de subdivisão e projeção sucessiva	24

Lista de tabelas

2.1	Códigos de expressão ou posição facial considerados na criação do conjunto de dados.	6
2.2	Características do conjunto de dados IMPA-FACE3D	7
5.1	Informações do esquema de subdivisão e projeção do modelo inicial sobre a superfície facial 3D obtida do escaner.	23
5.2	Cronograma de atividades de trabalho pendente	26

Capítulo 1

Introdução

O processo de reconstrução de modelos faciais 3D é um tópico relevante na área de *Visão Computacional* que recentemente vem recebendo especial interesse na comunidade científica. Este é um exemplo da chamada *fotografia digital* onde métodos de visão computacional e computação gráfica são usados para resolver um problema. Usualmente, modelos com geometria exata são preferidos para reconhecimento de faces, e modelos mais simples são mais adequados em aplicações onde a rapidez do processo é considerado um fato crítico (Elyan e Ugail, 2007), e.g. transmissão de faces ou realidade aumentada.

As imagens de faces têm um papel central nas diferentes aplicações de visão computacional e computação gráfica. Diferentes métodos para detecção, rastreamento e representação de faces 3D foram desenvolvidos para permitir aplicações tais como o reconhecimento de faces (Banz, 2006; Jiang *et al.*, 2005; Onofrio e Tubaro, 2005; Wang e Chua, 2005), análise de expressões faciais (Soyel e Demirel, 2007; Yabui *et al.*, 2003), síntese de faces (Lee e Ranganath, 1999; Zhang e Xu, 2007) e vídeo *puppeteering* (Chai *et al.*, 2003; Hong *et al.*, 2002).

Na síntese de faces, muitos métodos propostos de reconstrução 3D baseiam-se em modelos de face artificiais como os avatares descritos em (Chai *et al.*, 2003). Embora o resultados sejam coerentes com a percepção humana, as reconstruções não são realistas. Por outro lado, Jiang *et al.* (2005) explora um algoritmo de alinhamento e extração de características faciais para reconstrução automática de face 3D. Nessa proposta, o algoritmo aplica análise de componentes principais considerando somente a forma da face para obter coeficientes de forma 3D. Entretanto, uma imagem frontal de uma pessoa com iluminação normal e expressão neutra é requerida, tornando o método não aplicável facilmente.

Apesar de mais de três décadas de pesquisa em fotografia facial 3D (Kanade, 1973; Kittler *et al.*, 2005), ainda existem importantes problemas em aberto como, por exemplo, a criação de métodos de fácil utilização que permitam a rápida e realista reconstrução e

caricaturização de faces 3D. É importante frisar que uma grande dificuldade que existe nessa área deve-se, em parte, à carência de dados reais de faces 3D com diferentes tipos de expressão facial, os quais permitam obter modelos realísticos.

Este trabalho apresenta uma nova proposta para a área de fotografia computacional de faces 3D usando modelos baseados em dados reais. As principais características do nosso trabalho é estar baseado em dados de geometria reais, obtendo faces 3D reconstruídas mais realistas. Atualmente, o sistema está sendo desenvolvido considerando poucas amostras de treinamento e algoritmos de representação padrão de visão e computação gráfica, tendo assim espaço para diferentes melhoras no futuro.

1.1 Objetivos

O objetivo geral do projeto de doutorado é desenvolver uma metodologia para a reconstrução de faces 3D dada uma fotografia 2D. A metodologia proposta está baseada em espaços gerados mediante análises de componentes principais sobre conjuntos de dados de faces 3D reais, de forma que o procedimento seja realizado rapidamente e a reconstrução seja simples e realista.

Os objetivos específicos do projeto de doutorado estão listados a seguir:

- (a) Desenvolver um protocolo de aquisição e criar um banco de dados de expressões faciais 3D;
- (b) Reduzir o tempo de reconstrução de geometria facial mediante modelos simples;
- (c) Definir, implementar e avaliar medidas de exatidão de reconstrução e coerência espacial facial 3D;
- (d) Mostrar limitações da reconstrução de faces 3D dada uma fotografia;
- (e) Desenvolver um método para a caricaturização de faces 3D considerando a metodologia de reconstrução de faces sobre diferentes espaços de caricaturas faciais.

1.2 Organização do trabalho

Este texto está organizado como segue: No Capítulo 2, descrevemos o protocolo adotado para a aquisição de dados, e o conjunto de dados de faces 3D criado especialmente para a análise de expressões faciais. No Capítulo 3 são mostrados uma revisão do esquema matemático proposto, e uma descrição do sistema de reconstrução. Os resultados experimentais obtidos até o momento são mostrados no Capítulo 4. Finalmente, nos Capítulos 5 e 6 deste manuscrito, são descritos brevemente os próximos passos a serem realizados, e as conclusões obtidas até o momento, respectivamente.

Capítulo 2

Conjunto de dados de faces 3D

As expressões faciais são o resultado de uma ou mais emoções que alteram os músculos da face. Esses movimentos mostram o estado emocional da pessoa aos observadores. Devido as faces terem uma quantidade limitada de movimentos, as expressões faciais baseiam-se em pequenas diferenças na proporção e posição relativa das características faciais. Assim, a interpretação ou entendimento requer considerável “sensibilidade”.

Atualmente, existem extensos bancos de dados de faces em 2D (fotografias) e 3D (em forma de malhas ou imagens de profundidade) e bancos de dados multimodais que além de informação em 2D e 3D incluem outras características biométricas como, por exemplo, a voz. Entretanto, nos bancos de dados atuais de faces 3D não são considerados um conjunto de expressões faciais representativas dos estados emocionais das pessoas. A falta desse tipo de informação, relevante para a análise de expressões faciais de pessoas, nos levou à criação de um conjunto de dados de expressões faciais 3D, denominado IMPA-FACE3D.

Neste capítulo, são descritos os equipamentos usados no processo de criação do conjunto de dados, e o protocolo de aquisição. Finalmente, são apresentadas algumas amostras do conjunto de dados. Parte deste capítulo foi tratado em um relatório técnico (Mena-Chalco *et al.*, 2008a).

2.1 Equipamento

Os dados de faces 3D foram adquiridos usando um escaner 3D (não-intrusivo) KONICA MINOLTA VIVID 910. O escaner está equipado por um sensor de distância laser, e uma câmera digital integrada. As especificações de aquisição do escaner são: $111 \times 84 \times 40mm$ (min) a $1200 \times 903 \times 400mm$ (max) (largura \times profundidade \times altura, respectivamente), e lente com distância focal de 8mm.

A aquisição para cada amostra é composta de dados de textura e geometria registradas. As imagens de textura foram adquiridas com uma resolução de 640×480

pixels (24 bits). A geometria de face 3D associada a cada imagem de textura contém aproximadamente 50000 pontos. O tempo de captura de dados para uma amostra é de aproximadamente 2.5 segundos.

Adicionalmente, dispositivos que permitiram controlar as condições de iluminação foram considerados. Quatro refletores de luz quente (iluminação halógena) equipados com difusores para a eliminação/suavização de sombras foram usados:

- Uma lâmpada de 1000 watts, marca Lowel, modelo Lowel-Omni;
- Duas lâmpadas de 100 watts, marca Dixel Lighting modelo Soft Compact 53-55;
- Uma lâmpada de 600 watts, marca Dixel lighting, modelo Fresnel Vulcano 50-02, com diâmetro de lente de 120mm.

Cabe destacar que a câmera do escaner foi calibrada para a cor de faces e iluminação com luz quente (*white-balance*). O programa Polygon Editing Tool 2.10 da KONICA MINOLTA sob o sistema operacional Windows XP (Intel Pentium 4, CPU 3.20Ghz, com 3GB de memória RAM) foi usado para a aquisição e armazenamento dos dados.

Na Figura 2.1 são mostrados os equipamentos considerados na aquisição de dados.

2.2 Protocolo de aquisição

A criação do conjunto de dados envolve fatores que influenciam as condições de aquisição, tais como iluminação, número de sessões, número de pessoas, imagens por pessoa, e distâncias. Não foram manipulados os dados capturados pelo escaner, como tratamento de buracos, filtro dos dados, fusão de dados. Em um processo posterior à captura serão consideradas operações sobre os dados. A seguir são listadas algumas características da aquisição dos dados:

- Pessoa sempre com olhos abertos;
- Não serão permitidos óculos, bonés nem outros objetos que alterem o aspecto facial;
- Sem restrições de vestimenta, penteado ou corte de cabelo.

A nomenclatura adoptada é sXXX-YY em que XXX representa o número da pessoa, e YY representa o tipo de expressão ou posição. As expressões consideradas estão listadas na Tabela 2.1. Além das expressões básicas, estão sendo consideradas algumas expressões fortemente diferenciadas (códigos 07-11) também com orientação facial frontal. Note-se que para uma futura reconstrução 3D da cabeça foram adquiridos complementarmente dados dos dois perfis das pessoas com expressão neutra (códigos 12-13).

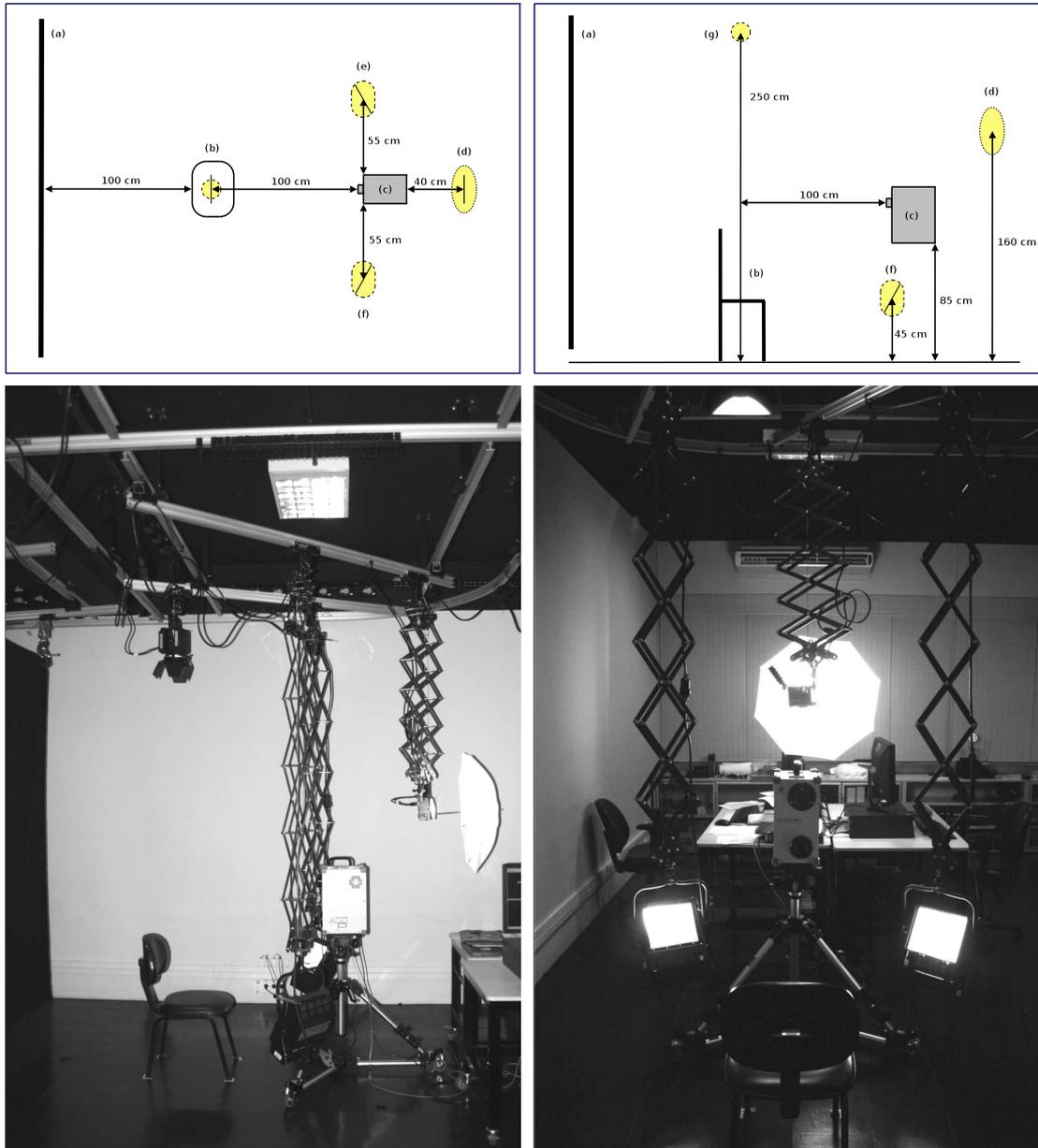


Figura 2.1: Equipamento considerado na aquisição de dados. Acima: Diagrama de configuração (a) Lona preta, (b) cadeira, (c) escaner 3D, (d) reflector halógeno de 1000 watts (guarda-chuva), (e)-(f) reflector halógeno de 100 watts, (g) reflector halógeno de 600 watts. Abaixo: Fotografias do ambiente.

Tabela 2.1: *Códigos de expressão ou posição facial considerados na criação do conjunto de dados.*

Código	Expressão ou posição
00	Neutra e frontal (<i>neutral</i>)
01	Alegria (<i>joy</i>)
02	Tristeza (<i>sadness</i>)
03	Surpresa (<i>surprise</i>)
04	Raiva (<i>anger</i>)
05	Desgosto (<i>disgust</i>)
06	Medo (<i>fear</i>)
07	Boca e olhos abertos
08	Boca e olhos fechados
09	Beijo
10	Boca e olho esquerdo abertos
11	Boca e olho direito abertos
12	Perfil esquerdo com expressão neutra
13	Perfil direito com expressão neutra

2.3 Conjunto de dados IMPA-FACE3D

O conjunto de dados está formado por aquisições de expressões faciais de 38 pessoas das quais 16 foram mulheres (42%) e 22 foram homens (58%). Ao todo, 14 capturas de textura e geometria 3D (imagem de profundidade) para cada pessoa foram obtidas, em uma única sessão de aproximadamente 30 minutos. As aquisições foram realizadas no Estúdio de fotografia computacional e vídeo do Instituto de Matemática Pura e Aplicada no período de Outubro a Novembro de 2008.

Na Tabela 2.2 estão listadas as principais características do conjunto de dados criado. Cabe destacar que toda aquisição está armazenada nos seguintes formatos de arquivos: Bitmap, TIFF, TGA, PLY, CMD e VVD. Os dois últimos formatos listados são próprios da KONICA MINOLTA e usados mediante o programa Polygon Editing Tool.

Nas Figura 2.2 são mostrados todas as aquisições de dados pertencentes a uma pessoa considerada no conjunto de dados. As diferentes expressões ou poses correspondem aos 14 tipos descritos na Tabela 2.1.

Tabela 2.2: *Características do conjunto de dados IMPA-FACE3D. O conjunto de dados de faces 3D está constituído de aquisições de 16 mulheres (42%) e 22 homens (58%).*

Código	Sexo	Idade	Código	Sexo	Idade
s001	mulher	26	s020	mulher	22
s002	mulher	23	s021	homem	25
s003	mulher	63	s022	mulher	20
s004	homem	65	s023	homem	27
s005	homem	30	s024	homem	21
s006	mulher	22	s025	homem	33
s007	homem	24	s026	homem	26
s008	homem	29	s027	mulher	44
s009	homem	21	s028	mulher	34
s010	homem	26	s029	mulher	33
s011	mulher	24	s030	homem	31
s012	homem	29	s031	mulher	24
s013	homem	19	s032	mulher	28
s014	mulher	26	s033	homem	27
s015	mulher	29	s034	homem	63
s016	homem	45	s035	mulher	63
s017	homem	54	s036	homem	49
s018	homem	33	s037	mulher	25
s019	homem	23	s038	homem	29

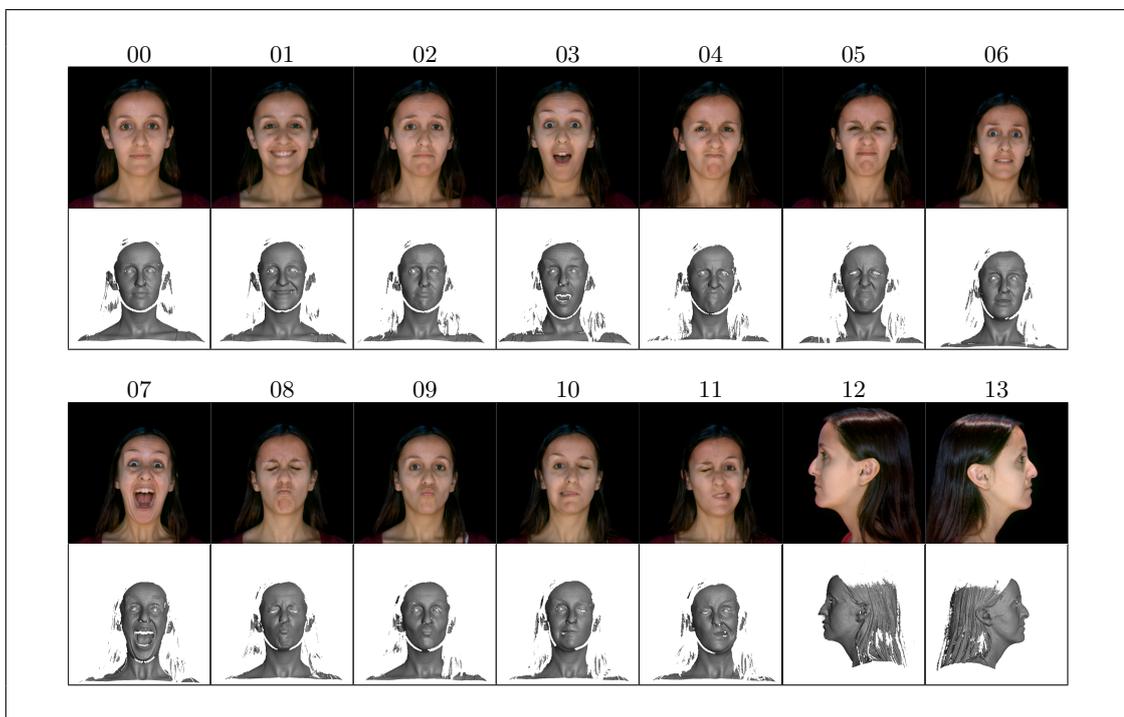


Figura 2.2: Exemplo de captura de dados: Textura e geometria (imagem de profundidade). O código mostrado no topo de cada coluna está associado a um tipo de expressão ou posição facial descrita na Tabela 2.1.

Capítulo 3

Reconstrução facial 3D através de espaços PCA

Baseados no trabalho de [Vlasic et al. \(2005\)](#) e [Macêdo et al. \(2006\)](#), foi criada uma metodologia para a reconstrução facial 3D a partir de imagens 2D coloridas, usando um conjunto pequeno de treinamento composto por textura e geometria. Os dados de treinamento contém textura e geometria registradas.

É interessante notar que estes trabalhos prévios não exploram informação 3D. O conjunto de treinamento é composto por um número pequeno de imagens de profundidade correspondente a sete expressões faciais diferentes de 10 pessoas. Nossa abordagem usa análise de componentes principais (PCA) para representar o modelo de face (textura e geometria separadamente). No sistema, o modelo de face PCA é composto por duas bases ortonormais separadas que representam a textura e a geometria, respectivamente.

Dada como entrada uma imagem de face frontal a ser reconstruída em 3D, são extraídos os marcos ou pontos característicos faciais 2D¹ usando um método de *Active Shape Model* (ASM). O conjunto de marcos é usado para normalizar a textura de entrada. A normalização é obtida mapeando (*warping*) os dados de textura e geometria à forma de face média. A geometria facial 3D é reproduzida projetando a textura normalizada sobre um espaço de geometria obtido na fase de treinamento. A projeção é realizada usando uma base vetorial PCA e uma função de otimização linear para relacionar a informação de textura 2D com a geometria 3D. Finalmente, a reconstrução 3D é obtida mapeando diretamente a textura normalizada sobre a geometria. Na Figura 3.1 é resumida a arquitetura proposta do sistema. A presente abordagem de reconstrução facial 3D foi apresentado no Simpósio Brasileiro de Computação Gráfica e Processamento de Imagens ([Mena-Chalco et al., 2008b](#)), e posteriormente foi publicada uma versão estendida do trabalho na revista *The Visual Computer* ([Mena-Chalco et al., 2009](#)).

¹Na literatura, os pontos característicos faciais também são conhecidos como *landmarks*. Neste manuscrito usaremos o termo *marcos* para referir-nos aos pontos característicos faciais.

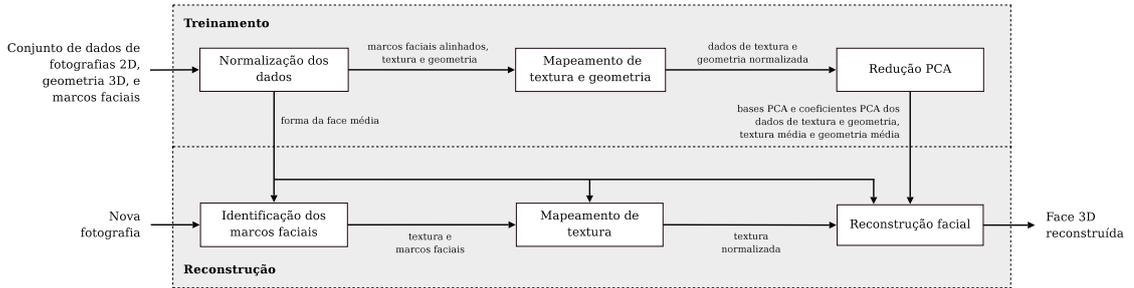


Figura 3.1: Fluxograma do sistema de reconstrução facial 3D (linhas pontilhadas). Cada bloco representa um processo, e cada seta representa o fluxo de informação entre os processos.

3.1 Modelo de face

Nosso estudo está concentrado na reconstrução e futuramente na caricaturização de pessoas de diferente gênero. Assim, inicialmente consideramos faces 3D com 7 expressões diferentes: neutra, alegria, tristeza, surpresa, raiva, desgosto e medo. Os dados de geometria foram pré-processados mediante um filtro *Laplaciano* para preencher os buracos possivelmente existentes na aquisição dos dados. Dessa forma, valores de profundidade são associados a cada pixel das imagens de textura.

Após os dados faciais serem adquiridos e pré-processados, os marcos faciais foram manualmente posicionados sobre as imagens de textura e alinhados a seus correspondentes dados de geometria 3D (visto que os dados de geometria e textura estão registrados). Adoptamos um modelo de face com $K = 77$ marcos e uma triangulação que contém 120 elementos (triângulos). Na Figura 3.2 mostramos um exemplo de marcos faciais e triangulação considerados.

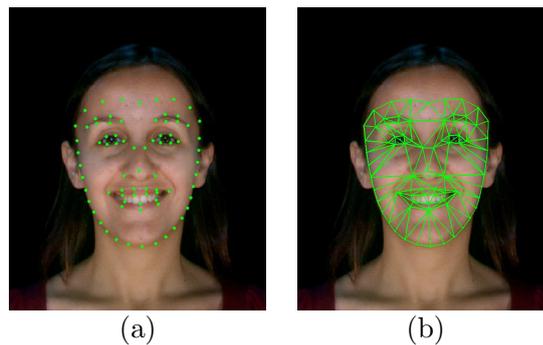


Figura 3.2: Exemplo de (a) marcos característicos faciais; e (b) triangulação considerada na proposta.

3.2 Modelo matemático proposto

Com o intuito de auxiliar na compreensão da abordagem adotada para a reconstrução de faces 3D, apresentamos uma lista de símbolos usados:

- l_i : marcos característicos da i -ésima face;
- \bar{l} : forma de face média;
- x_i^t, x_i^g : i -ésima textura de face e sua correspondente geometria pertencente ao conjunto de dados;
- $\mathbf{x}^t, \mathbf{x}^g$: textura de face de entrada normalizada e a correspondente geometria reconstruída, respectivamente;
- L_i^t, L_i^g : textura e geometria, da i -ésima face, recortada e mapeada a \bar{l} , respectivamente;
- E^t, E^g : bases de textura e geometria;
- α^t, α^g : coeficientes de textura e geometria expressos em termos de E^t, E^g , respectivamente;
- s_x : vetor de pesos de \mathbf{x}^t em termos das amostras de treinamento.

A abordagem adotada baseia-se na aprendizagem de modelos faciais 3D usando textura e geometria de faces de treinamento para algumas expressões faciais. Uma imagem de entrada de face 2D (e.g. textura) dada como entrada é projetada sobre um espaço de textura, obtido na etapa de treinamento. Posteriormente, ela é decomposta em coeficientes correspondentes às amostras de treinamento. Tais coeficientes são usados com a base de geometria para construir um modelo de face 3D da imagem de entrada.

O conjunto de treinamento está composto por tuplas de dados de textura e geometria de diferentes pessoas com diferentes expressões faciais. Um conjunto de marcos $\{l^1, l^2, \dots, l^K\}$ é colocado na imagem de textura e usado para representar a informação de textura de entrada. Assim, a textura facial é representada por L^t composta pela informação limitada pelos marcos faciais. Dado que a textura e a geometria estão registradas, os marcos faciais de textura tem seus correspondentes marcos de geometria, que são utilizados para definir a geometria facial L^g . Este esquema de representação por marcos inicial é ilustrado na Figura 3.3.

A etapa de treinamento consiste em definir bons espaços de representações de textura e geometria baseados em um conjunto de amostras de expressões faciais de pessoas.

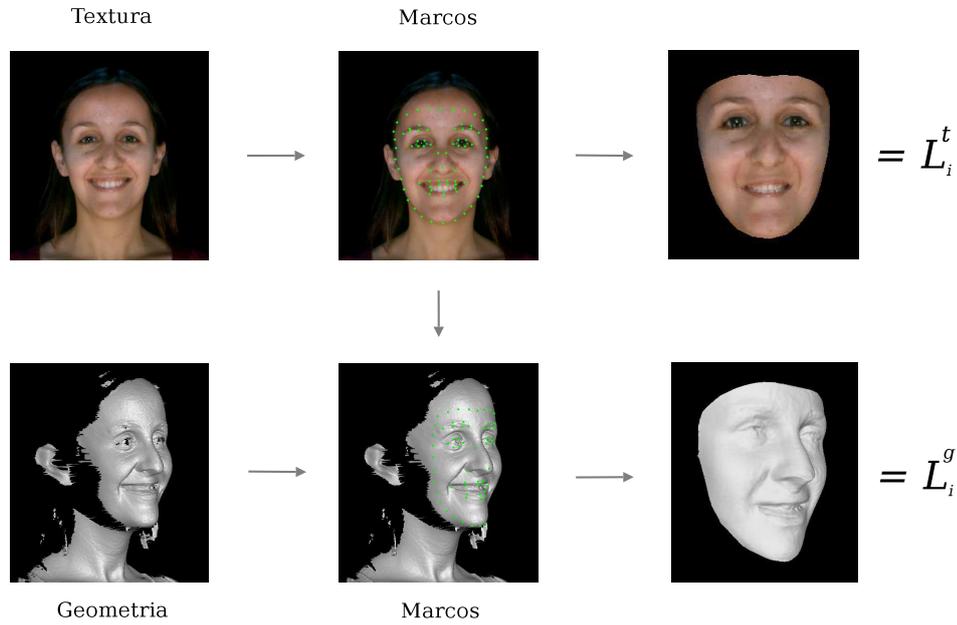


Figura 3.3: Formação do conjunto de treinamento: textura e geometria da i -ésima amostra. Os mesmos marcos faciais são usados na textura e na geometria de cada amostra.

Portanto, textura e geometria são obtidas para diferentes expressões faciais de N pessoas, sendo denotadas como $\{L_1^t, L_2^t, \dots, L_N^t\}$ e $\{L_1^g, L_2^g, \dots, L_N^g\}$, respectivamente. Estes dados ajudam a definir os espaços iniciais de textura e geometria, como mostrado nas Figuras 3.4(a) e (e). A fim de ter uma representação eficiente e otimizada, tanto o espaço de textura quanto o de geometria são transformados mediante PCA (Figuras 3.4(b) e 3.4(d), respectivamente). Cada amostra de treinamento representa um vetor expresso nesses espaços.

O objetivo principal do sistema de fotografia 3D é obter a representação de geometria de uma dada face \mathbf{x} , fornecida como uma imagem de textura. Um conjunto de marcos faciais é automaticamente identificado, e uma textura de entrada normalizada \mathbf{x}^t é extraída da imagem de entrada, sofrendo uma série de transformações através dos espaços de textura e geometria, como ilustrado na Figura 3.4. O resultado final é a geometria reconstruída da imagem de entrada \mathbf{x}^g , i.e. um ponto no espaço de geometria.

3.3 Método de reconstrução facial 3D

3.3.1 Treinamento

O processo de treinamento é composto por três fases. Primeiramente, os marcos característicos faciais de entrada são normalizados mediante um análise *Procrustes* (Dryden e Mardia,

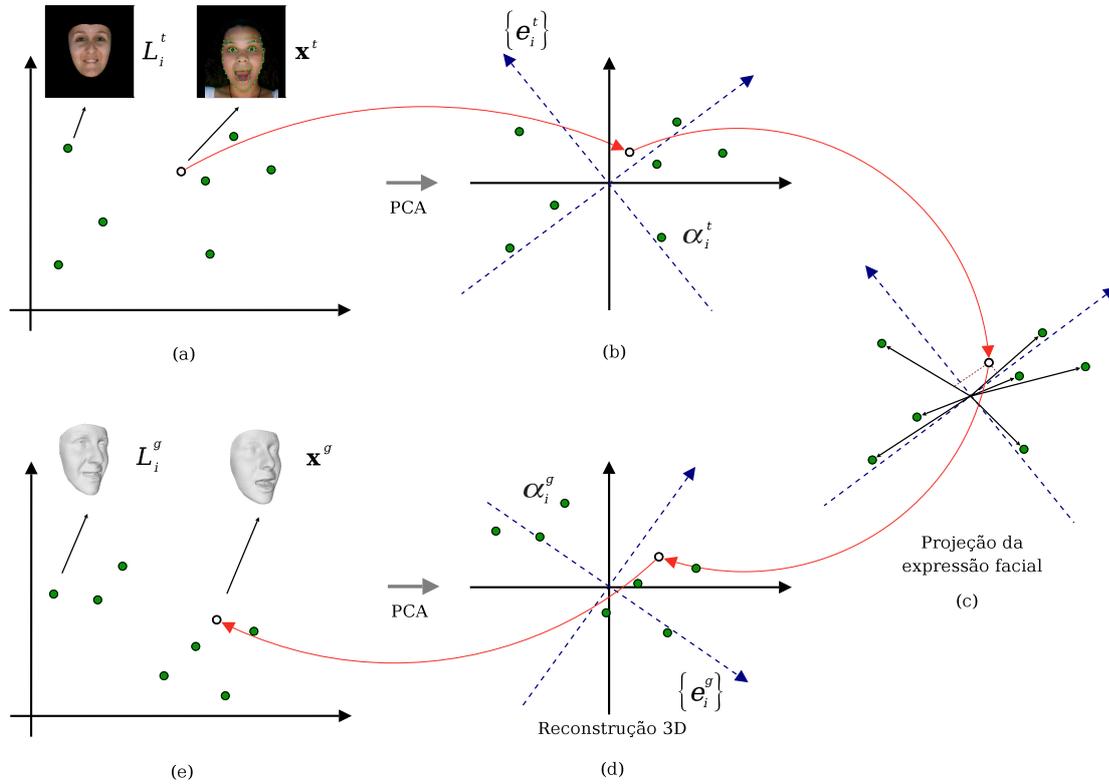


Figura 3.4: Espaços de textura e geometria: \mathbf{x}^t é uma textura de face de entrada que sofre uma série de transformações através dos espaços de textura e geometria, até que sua geometria \mathbf{x}^g seja construída.

1998; Goodall, 1991), resultando assim em um banco de dados com marcos faciais alinhados \bar{l} em um sistema de coordenadas comum.

A normalização *Procrustes* permite remover componentes de translação, rotação e escala das formas, definidas mediante os marcos, a fim de encontrar o melhor ajuste a todas as formas. Esta análise, alinha iterativamente todas as formas de modo que a soma das distâncias de cada forma à forma média seja minimizada, i.e., esta análise permite minimizar a distância $\sum_{i=1}^N |l_i - \bar{l}|^2$ (Goodall, 1991).

Os marcos faciais podem ser alinhados para as diferentes imagens de entrada devido à homologia entre as representações individuais. Esse fato permite mapear cada imagem de entrada para a forma da face média usando uma triangulação previamente definida. Os dados de geometria são também mapeados para a forma da face média obtida mediante a análise *Procrustes*.

Dois procedimentos PCA foram realizados separadamente tanto para os dados de

geometria L_i^g quanto para os de textura L_i^t . Esta análise permite obter:

- Um modelo de textura média (t_0), uma base ortonormal ($E^t = \{e_i^t\}$) para o espaço de textura facial e os coeficientes ($\{\alpha_i^t\}$) para cada imagem de textura no conjunto de dados de treinamento expressos em relação a $\{e_i^t\}$. Veja um exemplo na Figura 4.2(a);
- Um modelo de geometria média (g_0), uma base ortonormal ($E^g = \{e_i^g\}$) para o espaço de geometria e os coeficientes ($\{\alpha_i^g\}$) para cada dado de geometria 3D no conjunto de treinamento expressos em relação a $\{e_i^g\}$. Veja um exemplo na Figura 4.2(b).

A fim de trabalhar com o mesmo número de componentes principais nos espaços acima descritos, foi usada uma quantidade mínima de componentes que representem a variância total a ser mantida em ambas bases (nível de variância definida *a priori*). Os resultados mostrados neste capítulo foram realizados unicamente considerando componentes principais que mantenham pelo menos 95% do total de variância.

O algoritmo seguinte resume o processo de treinamento:

Face3D-Treinamento ($X^t = \{x_1^t, \dots, x_N^t\}$, $X^g = \{x_1^g, \dots, x_N^g\}$, $marcos = \{l_1, \dots, l_N\}$, *triangulação*)

- 1 $\bar{l} \leftarrow \text{PROCRUSTES-ANALYSIS}(marcos)$
- 2 $L^t \leftarrow \text{WARP}(X^t, marcos, \bar{l}, triangulação)$
- 3 $L^g \leftarrow \text{WARP}(X^g, marcos, \bar{l}, triangulação)$
- 4 $\{t_0, g_0\} \leftarrow \text{MEAN}(L^t, L^g)$
- 5 $\{E^t, E^g\} \leftarrow \text{PRINCIPAL-COMPONENT-ANALYSIS}(L^t, L^g)$

O procedimento WARP permite normalizar as texturas X^t e as geometrias X^g do conjunto de dados, mapeando-os na forma de face média usando uma triangulação definida. O procedimento MEAN calcula o modelo de textura média t_0 , e o modelo de geometria média g_0 , baseados em L^t e L^g , respectivamente.

3.3.2 Reconstrução

A entrada para o sistema é uma fotografia 2D (com uma imagem frontal de face) à qual uma técnica de ASM é aplicada para automaticamente detectar os marcos faciais. Esta técnica permite extrair os marcos faciais usando como entrada a imagem em tons de cinza. ASM foi proposto por [Cootes et al. \(1995\)](#) e permite o alinhamento e a representação de formas contidas nas imagens usando modelos estatísticos obtidos de dados de

treinamento. O modelo de face é representado por um conjunto de marcos manualmente colocados nas imagens de faces de treinamento (não necessariamente aqueles obtidos para o modelo 3D). Os conjuntos de marcos para as imagens de treinamento são alinhados para minimizar a distância entre os pontos correspondentes (i.e. pontos homólogos). Um modelo de distribuição de pontos (PDM, *point distribution model*) é obtido da variância das distâncias entre os diferentes pontos. O PDM é usado para restringir a variação de forma no processo iterativo considerado com ASM.

Os marcos faciais são alinhados à forma de face média obtida no processo de treinamento. A textura é mapeada à forma da face média (similar ao processo realizado na etapa de treinamento). Este processo permite normalizar a textura de uma imagem de entrada.

Seja \mathbf{x}^t a textura mapeada da imagem de entrada \mathbf{x} , e t_0 a textura média normalizada obtida na etapa de treinamento. Os coeficientes de textura α_x^t , são calculados projetando $(\mathbf{x}^t - t_0)$ sobre a respectiva base ortonormal $(\{e_i^t\})$:

$$\alpha_x^t = E^t \cdot (\mathbf{x}^t - t_0) \quad (3.1)$$

em que E^t é a matriz de transformação definida pela base ortonormal para o espaço de textura obtido no processo de treinamento.

Uma vez que os coeficientes de textura α_x^t são obtidos, os coeficientes de textura de todas as imagens consideradas no treinamento α^t são usadas para calcular os coeficientes s_x , definidos como:

$$\alpha^t \cdot s_x = \alpha_x^t \quad (3.2)$$

em que α^t é a matriz definida pelos coeficientes para cada imagem de textura no treinamento. Intuitivamente, s_x representa os pesos dos coeficientes obtidos na projeção de α_x^t sobre α^t (veja a Figura 3.4(c)). É importante destacar que cada amostra representada em α^t está associada a uma expressão facial diferente. Portanto, s_x representa a decomposição de \mathbf{x}^t em termos das diferentes expressões faciais aprendidas pelo sistema (e.g. poderíamos dizer que \mathbf{x}^t esta $a\%$ feliz, $b\%$ com raiva, $c\%$ com expressão neutra, etc.). Os coeficientes de geometria α_x^g de \mathbf{x}^t são então calculados usando os coeficientes de geometria de todas as amostras de geometria consideradas no treinamento α^g :

$$\alpha_x^g = \alpha^g \cdot s_x \quad (3.3)$$

A geometria normalizada \mathbf{x}^g da imagem de face de teste \mathbf{x} é reconstruída por:

$$\mathbf{x}^g = (E^g \cdot \alpha_x^g) + g_0 \quad (3.4)$$

em que E^g é a matriz de transformação definida pela base ortonormal para o espaço de geometria aprendido no treinamento. A fim de reduzir o ruído da geometria facial reconstruída (superfície), uma técnica melhorada para suavização (*smoothing*) de superfícies poligonais proposto por Vollmer *et al.* (1999) foi aplicada. Finalmente, a textura de entrada, normalizada à forma de face média, é diretamente mapeada sobre a geometria 3D resultante. As áreas onde o mapeamento de textura não é atingido são preenchidas mediante interpolação dos pontos 3D adjacentes.

O seguinte algoritmo resume a abordagem proposta:

Face3D-Reconstrução(\mathbf{x})

- 1 $l_x \leftarrow \text{ASM}(\mathbf{x}) \triangleright$ São identificados os marcos da fotografia de entrada \mathbf{x}
- 2 $\mathbf{x}^t \leftarrow \text{WARP}(\mathbf{x}, l_x, \bar{l})$
- 3 $\alpha_x^t \leftarrow E^t \cdot (\mathbf{x}^t - t_0)$
- 4 $s_x \leftarrow \text{PROJECT}(\alpha_x^t, \alpha^t)$
- 5 $\alpha_x^g \leftarrow \alpha^g \cdot s_x$
- 6 $\mathbf{x}^g \leftarrow (E^g \cdot \alpha_x^g) + g_0$
- 7 **return** \mathbf{x}^g

O procedimento WARP normaliza a textura de entrada \mathbf{x} usando os correspondentes marcos faciais l_x e a forma de face média \bar{l} obtida na etapa de treinamento. O procedimento PROJECT permite representar α_x^t em termos de α^t , como indicado na Equação (3.2).

Capítulo 4

Resultados

O sistema foi implementado como descrito nas seções anteriores usando uma versão padrão de MATLAB. Os experimentos apresentados nesta seção foram produzidos usando um computador Intel Core 2 Duo 1.5GHz com 2 Gb de memória RAM sob o sistema operacional GNU/Linux.

O sistema foi testado usando dados reais. Todos os experimentos foram realizados com sete expressões faciais de 10 pessoas (5 mulheres e 5 homens), consistindo de um total de 70 imagens de treinamento (textura e geometria). Na Figura 4.1 são mostradas as imagens usadas na etapa de treinamento. Foi usado um escaner 3D KONICA MINOLTA VIVID 910 para a aquisição de dados como mencionado no Capítulo 2. Em todos os experimentos, para uma dada imagem de textura, uma geometria de face de 26651 vértices foi obtido em no máximo 4 segundos. Oito componentes principais foram requeridos para representar a variância total do banco de dados. A face média do conjunto de dados é mostrada na Figura 4.2.

Três diferentes tipos de experimentos foram realizados para avaliar a qualidade de reconstrução facial 3D gerada pelo sistema.

4.1 Reconstrução de amostras presentes no treinamento

Este experimento foi realizado para investigar a reconstrução 3D considerando sete diferentes expressões faciais de uma pessoa vista no processo de treinamento. As fotografias 2D mostradas na Figura 4.3(a) foram usadas como entrada para o sistema. Nesse caso, a pessoa pertence ao conjunto de treinamento. A Figura 4.3(b) mostra a correspondente geometria facial reconstruída. A Figura 4.3(c) mostra a geometria facial reconstruída com a imagem de textura mapeada. Como pode ser visto, o sistema foi capaz de reconstruir satisfatoriamente a estrutura de face 3D da pessoa mantendo a coerência espacial facial.

Observe que a boca está ligeiramente aberta nas faces reconstruídas para as ex-



Figura 4.1: Expressões faciais do conjunto de dados considerado na reconstrução: O conjunto completo está composto de 70 fotografias 2D e 70 imagens de profundidade correspondentes a 7 expressões faciais (apresentadas nas linhas) de 10 pessoas (apresentadas nas colunas).

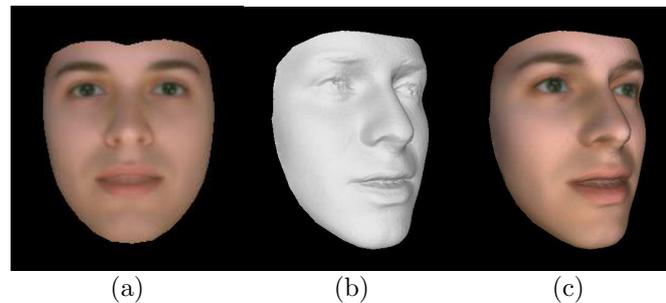


Figura 4.2: Face média obtida na etapa de treinamento: (a) textura, t_0 ; (b) geometria, g_0 ; (c) geometria com a textura facial mapeada.

pressões neutra e de tristeza. Isso deve-se ao fato que os coeficientes de geometria obtidos não foram suficientemente maiores (em módulo) para modificar a boca ligeiramente aberta da geometria média g_0 calculada na etapa de treinamento.



Figura 4.3: *Reconstrução facial 3D de amostras presentes no treinamento: (a) Expressões faciais: neutra, alegria, tristeza, surpresa, raiva, desgosto e medo; (b) Geometria 3D reconstruída; (c) Geometria 3D reconstruída com textura facial mapeada. No sistema, fotografias 2D com distintas expressões faciais de uma pessoa presente no conjunto de dados foram usadas na reconstrução 3D. Note que a geometria reconstruída mostra coerência espacial em relação a cada expressão facial.*

4.2 Reconstrução de amostras não presentes no treinamento

Este experimento foi realizado para reconstruções 3D usando sete expressões faciais de uma pessoa diferente das consideradas na etapa de treinamento. Fotografias de expressões faciais foram obtidas em um ambiente controlado (a informação de textura obtida foi capturada mantendo as mesmas especificações técnicas usadas no treinamento do sistema). Neste caso, os marcos faciais foram identificados usando o método de ASM como descrito no Capítulo 3. Subsequentemente, estes dados foram usados no sistema a fim de reconstruir a geometria 3D a partir de fotografias.

Na Figura 4.4 mostramos as reconstruções faciais 3D para algumas expressões faciais. As reconstruções de geometria calculadas estão mostradas nas Figura 4.4(b). A Figura 4.4(c) mostra a geometria 3D onde a fotografia 2D foi mapeada. Na reconstrução 3D, a coerência espacial facial é também mantida. Note que, para expressão facial de raiva, a geometria obtida apresenta uma face com boca fechada. Isso acontece pois diferentes amostras de raiva no banco de dados foram expressos com a boca fechada, como mostrado na expressão de raiva na Figura 4.3.

Em geral, a coerência espacial é mantida em todas as faces reconstruídas. Resultados

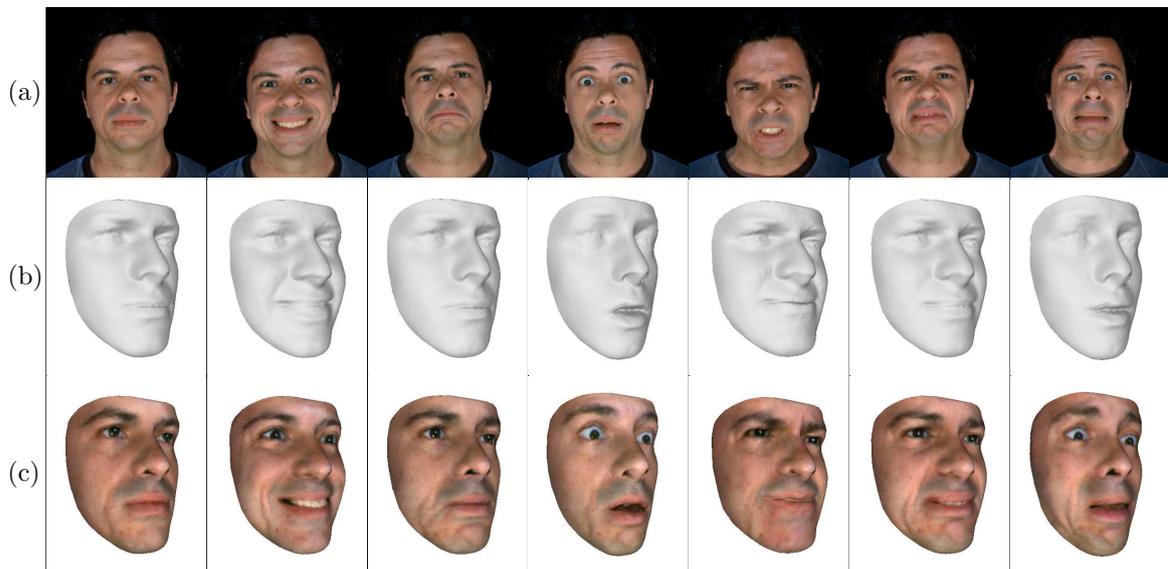


Figura 4.4: Reconstrução facial 3D de amostras não presentes no treinamento: (a) Expressões faciais: neutra, alegria, tristeza, surpresa, raiva, desgosto e medo; (b) Geometria 3D reconstruída; (c) Geometria 3D reconstruída com textura facial mapeada. No sistema, fotografias 2D com distintas expressões faciais de uma pessoa não presente no conjunto de dados foram usadas na reconstrução 3D. Note que a geometria reconstruída mostra coerência espacial em relação a cada expressão facial.

qualitativamente similares foram obtidos em outros experimentos, independentemente de a pessoa estar ou não presente no conjunto de treinamento.

4.3 Reconstrução de amostras com expressões não presentes no treinamento

Neste experimento, consideramos as duas pessoas presentes nos dois primeiros experimentos, mas usando como imagens de entrada expressões faciais que não foram vistas na etapa de treinamento: boca e olhos abertos, e expressão de beijo com olhos abertos. As texturas mostradas na Figura 4.5(a) foram usadas como entradas para o sistema. Logo foram identificados os marcos faciais usando ASM (ambos dados usados no sistema de reconstrução de faces). Observe-se que na Figura 4.5(b), embora a estrutura total seja recuperada, detalhes, como os da boca aberta, não são completamente reconstruídos como se esperava. As expressões obtidas não são realistas quando comparadas com as fotografias de entrada pois o banco de dados não inclui nenhuma amostra com expressões faciais similares. Essa limitação indica que ainda existe um grande espaço para melhorar o sistema de reconstrução.

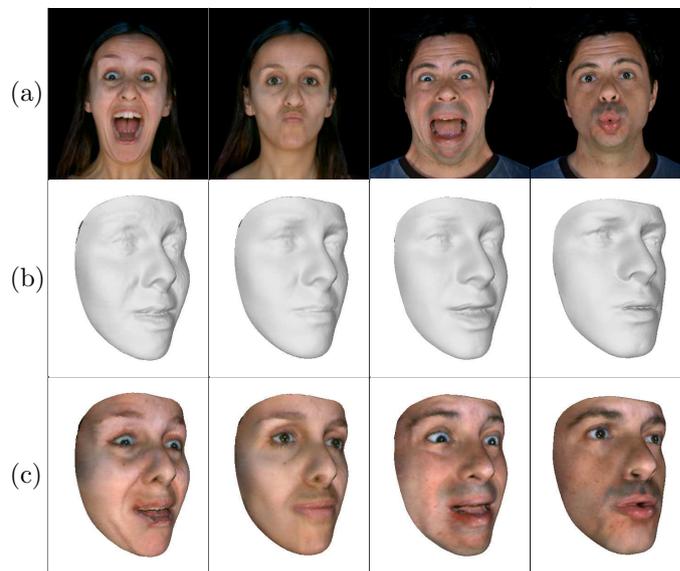


Figura 4.5: *Reconstrução facial 3D de amostras com expressões faciais não presentes na etapa de treinamento: (a) As duas primeiras fotografias correspondem a uma pessoa presente no conjunto de dados e as duas últimas fotografias correspondem a uma pessoa não presente no conjunto de dados; (b) Geometria 3D reconstruída; (c) Geometria 3D reconstruída com textura facial mapeada.*

Capítulo 5

Próximos passos

5.1 Utilização de um novo modelo de face 3D

Para a normalização dos conjunto de dados reais, anterior à etapa de treinamento, esperamos utilizar um processo de refinamento iterativo de um modelo de face que, mediante subdivisões fixas e projeções na geometria de face original, nos permita ter um melhor controle sobre a granularidade da superfície (i.e. quantidade de vértices e triângulos). Nesse sentido, usamos a interpolação linear e o esquema de subdivisão de Loop (Loop, 1987) para refinar ou subdividir um modelo inicial de faces 3D composto de 21 vértices e 30 triângulos, similar ao modelo descrito no trabalho de Golovinskiy *et al.* (2006) (veja a Figura 5.1). Ambas abordagens permitem dividir cada triângulo em quatro triângulos menores. Assim, no limite de um número de subdivisões, espera-se ter um modelo de faces próximo à face obtida mediante o escaner 3D.

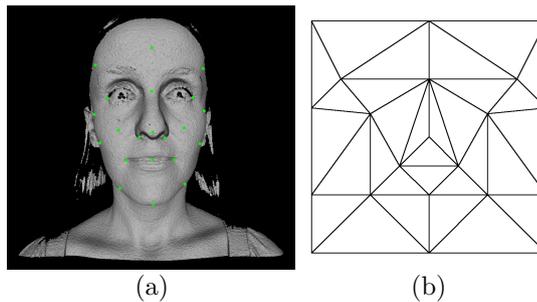


Figura 5.1: Novo modelo de face inicial: (a) As posições dos 21 vértices ou marcos faciais estão representados com pontos verdes; (b) Triangulação composta por 30 unidades. Os marcos e a triangulação foram definidas de forma arbitrária seguindo o trabalho de Golovinskiy *et al.* (2006).

Para uma face 3D em particular, após todo triângulo do modelo ser subdividido em quatro, o valor das posições dos novos vértices constituintes são obtidos projetando-os sobre a face 3D original a fim de obter posições mais próximas à face 3D original. Na

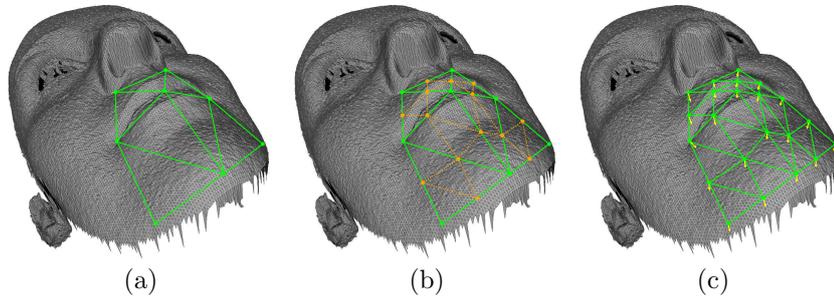


Figura 5.2: Exemplificação de subdivisão e projeção do modelo inicial: (a) Oito marcos e sete triângulos do modelo inicial; (b) Subdivisão triangular em que os novos vértices estão representados na cor laranja; (c) Projeção dos pontos do modelo refinado sobre a face original 3D, considerando os vetores normais (representados na cor amarela).

projeção do novo modelo, os vértices e seus vetores normais são utilizados para calcular a interseção do raio com o triângulo mais próximo da face 3D original obtido do escaner 3D. Na Figura 5.2 é exemplificada a subdivisão triangular e a projeção do modelo na face original para oito vértices da face mostrada na Figura 5.1(a). É importante comentar que os 21 marcos, ou pontos do modelo inicial, nesta etapa, foram seleccionados manualmente para as faces do conjunto de treinamento.

A abordagem de subdivisão e projeção adotada nos permitirá ter um melhor controle sobre a quantidade de triângulos e vértices a serem considerados no modelo de faces. Futuramente, poderemos investigar sobre o modelo de faces ideal tanto para a reconstrução quanto para a caricaturização. Na Tabela 5.1 são listadas informações sobre o número de triângulos e vértices obtidos para cada iteração de sucessivas subdivisões e projeções.

Tabela 5.1: Informações do esquema de subdivisão e projeção do modelo inicial sobre a superfície facial 3D obtida do escaner.

Descrição	Triângulos	Vértices
a. Modelo inicial (malha base)	30	21
b. Subdivisão linear e projeção	120	71
c. Subdivisão linear e projeção	480	261
d. Subdivisão de Loop e projeção	1 920	1 001
e. Subdivisão de Loop e projeção	7 680	3 921
f. Subdivisão de Loop e projeção	30 720	15 521

Veja na Figura 5.3 um exemplo de subdivisão e projeção sucessiva considerando os marcos iniciais da face mostrada na Figura 5.1(a). Observe-se que, somente foram trata-

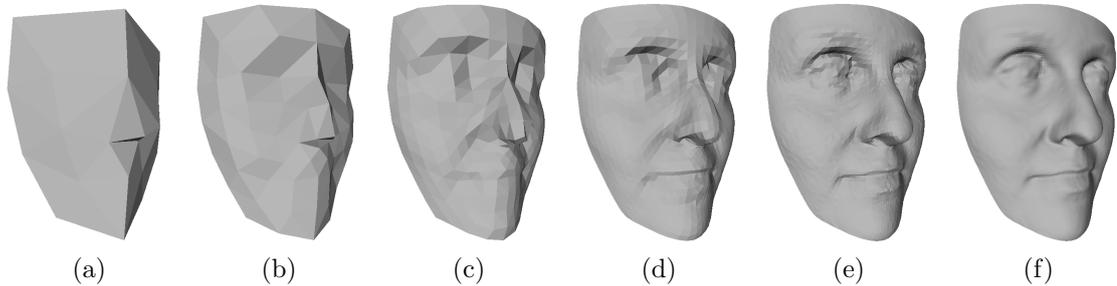


Figura 5.3: Exemplo de subdivisão e projeção sucessiva considerando os marcos iniciais da face mostrada na Figura 5.1(a). A maneira de subdivisão e projeção estão listados na Tabela 5.1.

das informações de geometria e não de textura. Seguindo nossa abordagem, facilmente podemos também considerar informações de textura nas subdivisões e projeções.

5.2 Cronograma de atividades

Na Tabela 5.2 é apresentado um cronograma contendo as principais atividades pendentes no trabalho.

A *atividade 1* refere-se ao pós-processamento dos dados adquiridos usando o scanner 3D. Muitas capturas apresentam alguns problemas como buracos nas zonas onde o sinal do laser não foi recebida (buracos) ou com picos de ruído na superfície (*outliers*). Portanto, esses erros ou limitações na aquisição devem ser, na medida do possível, eliminados de forma automática mediante interpolações (Conde, 2006), a fim de garantir dados com geometria 3D corretos. Observe-se que os dados depurados podem ser de grande utilidade para novos projetos de pesquisa nesta área.

A *atividade 2* refere-se à adoção de um novo modelo facial de 21 marcos faciais (como descrito na Seção 5.1) que permita uma maior flexibilidade no tratamento dos dados. Nesse sentido, para todo o conjunto de dados serão selecionados manualmente os marcos faciais do modelo inicial usando uma interface já desenvolvida neste projeto.

A *atividade 3* refere-se à melhoria, ou utilização de novas funções de mapeamento de textura, e de projeção dos coeficientes de textura (Equação 3.2) considerados no sistema de reconstrução. Por outro lado, a projeção do novo modelo sucessivamente subdividido deverá ter uma implementação com complexidade computacional eficiente. Atualmente a projeção é realizada sem considerar nenhuma técnica de aceleração de *ray tracing*. Para esse fim, pretende-se usar grades ou árvores de hierarquia (Akenine-Möller *et al.*, 2008).

A *atividade 4* refere-se à adoção ou definição de uma métrica de acurácia de reconstrução facial 3D que nos permita determinar numericamente a qualidade da reconstrução e coerência facial obtida. Futuramente, esta medida nos permitirá comparar nossa abor-

dagem de reconstrução com outras similares.

A *atividade 5* refere-se a uma parte importante da pesquisa de doutorado e corresponde à criação de uma série de espaços PCA próprios de relações de exagerações faciais 3D, para a síntese de caricaturas baseadas em amostras reais (mantendo a metodologia apresentada no Capítulo 3). Espera-se criar tantos espaços quanto relações de exageração facial que se deseje modelar.

A *atividade 6* refere-se à avaliação numérica do método de reconstrução e caricaturização facial 3D. Pode-se dar também especial interesse a uma validação subjetiva ou informal. Finalmente as *atividades 7 e 8*, referem-se à escrita formal, criação das transparências, e defesa de tese.

Tabela 5.2: *Cronograma de atividades de trabalho pendente. Os meses dos anos correspondentes estão representados com números de 2 algarismos.*

Atividade	2009		2010					
	09-10	11-12	01-02	03-04	05-06	07-08	09-10	
1. Depuração do conjunto de dados adquirido	x							
2. Utilização de um novo modelo de faces no conjunto de dados	x	x						
3. Melhoria dos procedimentos de mapeamento e projeção		x	x					
4. Definição e implementação de medidas de exatidão de reconstrução 3D			x	x				
5. Criação de espaços de relações de exageração facial 3D (caricaturização)				x	x	x		
6. Avaliação dos métodos de reconstrução e caricaturização 3D						x		
7. Escrita da tese					x	x	x	
8. Defesa da tese							x	

Capítulo 6

Conclusões

A presente proposta de trabalho de doutorado descreve uma metodologia para a reconstrução de geometria 3D de faces humanas a partir de imagens coloridas monoculares 2D. Esta tarefa de reconstrução 3D é realizada através de uma abordagem de “sínteses mediante análises” em que modelos para textura facial humana, e espaços de geometria são construídos desde dados adquiridos do mundo real.

O modelo matemático está baseado na formação de espaços adequados de textura e geometria a partir de um conjunto de dados de treinamento e de transformações entre tais espaços. A reconstrução de faces 3D dada uma imagem de textura é realizada por transformações do dado de entrada através desses espaços até que um modelo de geometria seja criado. Os resultados obtidos atualmente mostram que nossa abordagem pode ser aplicada para reconstrução de faces 3D de pessoas presentes ou não no conjunto de treinamento.

Para os testes experimentais foram utilizados modelos de faces humanas reais com diferentes expressões faciais. Foi definido um protocolo de aquisição de dados, de textura e geometria, considerando um escaner 3D em um ambiente de iluminação controlado. Ao todo, foram 38 pessoas as consideradas no conjunto de dados, capturando 266 (38x7) amostras com expressões faciais básicas, e 190 (38x5) amostras com expressões fortemente diferenciadas às anteriores.

Uma limitação da reconstrução facial 3D a se considerar é a identificação automática dos marcos ou pontos característicos faciais de fotografias 2D dadas como entrada. Para amostras com expressões faciais diferentes às consideradas no conjunto de treinamento do ASM, alguns marcos são posicionados de forma incorreta, sendo assim o sistema sensível aos dados de entrada. Acreditamos que uma identificação correta dos marcos nas fotografias 2D irá garantir uma boa reconstrução facial 3D. Assim, a consideração de um melhor método para tal identificação irá ao encontro de nosso objetivo de reconstrução facial 3D.

Bibliografia

- Akenine-Möller et al. (2008)** T. Akenine-Möller, E. Haines, e N. Hoffman. *Real-Time Rendering*. A K Peters Ltd., Massachusetts, third edition. Citado na p. [24](#)
- Blanz (2006)** V. Blanz. Face recognition based on a 3D morphable model. In *International Conference on Automatic Face and Gesture Recognition*, pages 617–624. Citado na p. [1](#)
- Chai et al. (2003)** J-X. Chai, J. Xiao, e J. Hodgins. Vision-based control of 3D facial animation. In D. Breen e M. Lin, editors, *Eurographics/SIGGRAPH Symposium on Computer Animation*, pages 193–206, San Diego, California. Eurographics Association. Citado na p. [1](#)
- Conde (2006)** C. Conde. *Biometría: Reconocimiento facial mediante fusión 2D y 3D*. Dykinson SL, Madrid. Citado na p. [24](#)
- Cootes et al. (1995)** T. F. Cootes, C. J. Taylor, D. H. Cooper, e J. Graham. Active shape models: Their training and application. *Computer Vision and Image Understanding*, 61(1):38–59. Citado na p. [14](#)
- Dryden e Mardia (1998)** I. L. Dryden e K. V. Mardia, editors. *Statistical Shape Analysis*. John Wiley and Sons, Chichester. Citado na p. [12](#)
- Elyan e Ugail (2007)** E. Elyan e H. Ugail. Reconstruction of 3D human facial images using partial differential equations. *Journal of Computers*, 2(8):1–8. Citado na p. [1](#)
- Golovinskiy et al. (2006)** A. Golovinskiy, W. Matusik, H. Pfister, S. Rusinkiewicz, e T. Funkhouser. A statistical model for synthesis of detailed facial geometry. *ACM Transactions on Graphics*, 25(3):1025–1034. Citado na p. [22](#)
- Goodall (1991)** C. Goodall. Procrustes methods in the statistical analysis of shape. *Journal of the Royal Statistical Society. Series B*, 53(2):285–339. Citado na p. [13](#)
- Hong et al. (2002)** P. Hong, Z. Wen, T. S. Huang, e H. Y. Shum. Real-time speech-driven 3D face animation. In *3D Data Processing Visualization and Transmission*, pages 713–716. Citado na p. [1](#)
- Jiang et al. (2005)** D. L. Jiang, Y. X. Hu, S. C. Yan, L. Zhang, H. J. Zhang, e W. Gao. Efficient 3D reconstruction for face recognition. *Pattern Recognition*, 38(6):787–798. Citado na p. [1](#)

- Kanade (1973)** T. Kanade. Picture processing system by computer complex and recognition of human faces. In *Doctoral dissertation, Kyoto University*. Citado na p. 1
- Kittler et al. (2005)** J. V. Kittler, A. Hilton, M. Hamouz, e J. Illingworth. 3D assisted face recognition: A survey of 3D imaging, modelling and recognition approaches. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 3, pages 114–114. Citado na p. 1
- Lee e Ranganath (1999)** M. W. Lee e S. Ranganath. 3D deformable face model for pose determination and face synthesis. In *International Conference on Image Analysis and Processing*, pages 260–265. Citado na p. 1
- Loop (1987)** C. Loop. Smooth subdivision surfaces based on triangles. Master’s thesis, University of Utah, Department of Mathematics. Citado na p. 22
- Macêdo et al. (2006)** I. Macêdo, E. V. Brazil, e L. Velho. Expression transfer between photographs through multilinear AAM’s. In *SIBGRAPI*, pages 239–246. IEEE Computer Society. Citado na p. 9
- Mena-Chalco et al. (2008a)** J.P. Mena-Chalco, R.M. Cesar-Jr, e L. Velho. Banco de dados de faces 3D: IMPA-FACE3D. Technical report, Instituto de Matemática Pura e Aplicada - IMPA - VISGRAF Laboratory, Rio de Janeiro, RJ, Brazil. TR01. Citado na p. 3
- Mena-Chalco et al. (2008b)** J.P. Mena-Chalco, I. Macêdo, L. Velho, e R.M. Cesar-Jr. PCA-based 3D face photography. In *Brazilian Symposium on Computer Graphics and Image Processing*, pages 313–320. IEEE Computer Society. Citado na p. 9
- Mena-Chalco et al. (2009)** J.P. Mena-Chalco, I. Macêdo, L. Velho, e R.M. Cesar-Jr. 3D face computational photography using PCA spaces. *The Visual Computer*. doi: <http://dx.doi.org/10.1007/s00371-009-0373-x>. (Aceito para publicação). Citado na p. 9
- Onofrio e Tubaro (2005)** D. Onofrio e S. Tubaro. A model based energy minimization method for 3D face reconstruction. In *ICME*, pages 1274–1277. IEEE. Citado na p. 1
- Soyel e Demirel (2007)** H. Soyel e H. Demirel. Facial expression recognition using 3D facial feature distances. In *International Conference on Image Analysis and Recognition*, pages 831–838. Citado na p. 1
- Vlasic et al. (2005)** D. Vlasic, M. Brand, H. Pfister, e J. Popović. Face transfer with multilinear models. *ACM Transactions on Graphics*, 24(3):426–433. Citado na p. 9
- Vollmer et al. (1999)** J. Vollmer, R. Mencl, e H. Muller. Improved Laplacian smoothing of noisy surface meshes. *Computer Graphics Forum*, 18(3):131–138. Citado na p. 16
- Wang e Chua (2005)** Y. J. Wang e C. S. Chua. Face recognition from 2D and 3D images using 3D Gabor filters. *Image and Vision Computing*, 23(11):1018–1028. Citado na p. 1

Yabui et al.(2003) T. Yabui, Y. Kenmochi, e K. Kotani. Facial expression analysis from 3D range images; comparison with the analysis from 2D images and their integration. In *International Conference on Image Processing*, pages 879–882. Citado na p. 1

Zhang e Xu (2007) Y. Zhang e S. Xu. Data-driven feature-based 3D face synthesis. In *3-D Digital Imaging and Modeling*, pages 39–46. Citado na p. 1

