

C1

Universidade Estadual de Campinas
Faculdade de Engenharia Elétrica e de Computação

Morphing aplicado ao envelhecimento de imagens faciais

Autor: Greyce Nogueira Schroeder

Orientador: Prof. Dr. Léo Pini Magalhães

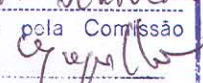
Dissertação de Mestrado apresentada à Faculdade de Engenharia Elétrica e de Computação como parte dos requisitos para obtenção do título de Mestre em Engenharia Elétrica. Área de concentração: **Engenharia de Computação**.

Banca Examinadora

Prof. Dr. Léo Pini Magalhães DCA/FEEC/Unicamp
Prof. Dr. Clésio Luis Tozzi DCA/FEEC/Unicamp
Prof. Dr. Lee Luan Ling DECOM/FEEC/Unicamp
Prof. Dr. Alexandre Xavier Falcão IC/Unicamp

Campinas, SP

20/Abril/2007

Este exemplar corresponde à redação final da tese defendida por:	Greyce Nogueira Schroeder
	e aprovada pela Comissão
Julgada em	20 / 04 / 2007
	
	Orientador

FICHA CATALOGRÁFICA ELABORADA PELA
BIBLIOTECA DA ÁREA DE ENGENHARIA E ARQUITETURA - BAE - UNICAMP

Sch76m Schroeder, Greyce Nogueira
Morphing aplicado ao envelhecimento de imagens faciais
/ Greyce Nogueira Schroeder. – Campinas, SP: [s.n.], 2007.

Orientador: Léo Pini Magalhães.
Tese (mestrado) - Universidade Estadual de Campinas,
Faculdade de Engenharia Elétrica e de Computação.

1. Animação por computador. 2. Computação Gráfica
3. Processamento de imagens. I. Magalhães, Léo Pini. II.
Universidade Estadual de Campinas. Faculdade de
Engenharia Elétrica e de Computação. III. Título

Título em Inglês:	Aging of face image using Morphing
Palavras-chave em Inglês:	Morphing, Warping, Radial Basis Function, Aging Face, Computer Animation, Computer Graphics, Image Processing
Área de concentração:	Engenharia de Computação
Títuloção:	Mestre em Engenharia Elétrica
Banca Examinadora:	Clésio Luis Tozzi, Lee Luan Ling, Alexandre Xavier Falcão
Data da defesa:	20/04/2007
Programa de Pós-Graduação:	Engenharia Elétrica

Resumo

O *morphing* de imagens é uma transformação que mapeia uma imagem em outra, alterando tanto a sua forma quanto as suas intensidades. Esta transformação possui diversas aplicações em imagens médicas e especialmente na indústria de entretenimento. Este trabalho objetiva, através do uso de *morphing*, apresentar um protótipo para a simulação de envelhecimento de faces frontais, utilizando um método chamado Funções de Base Radial (RBF) juntamente com um modelo quantitativo para expressar o processo de envelhecimento no rosto humano. O protótipo trabalha com imagens de pessoas a partir de 20 anos e realiza o envelhecimento até no máximo 70 anos. Além disso, o trabalho apresenta uma revisão bibliográfica sobre as principais técnicas de *morphing* e sobre o que já foi realizado sobre simulações computacionais de envelhecimento facial por imagens.

Palavras-chave: morphing, warping, funções de base radial, envelhecimento de faces.

Abstract

Image morphing is a transformation that maps one image into another, altering both its shape and intensities. These types of transformation have a wide range of applications in medical imaging and, specially, in entertainment industry. This work attempts to present a prototype for aging simulation on frontal face images, using a method of morphing called Radial Basis Functions (RBF) together with a quantitative model for expressing human aging. The prototype works with images from people with 20 years old up and performs the aging up to 70 years. Moreover, the work presents a bibliographical revision on the main techniques of morphing and on the state of the art on computational simulations of image face aging.

Keywords: morphing, warping, radial basis function, aging face.

À minha mãe, minha avó e meu marido

Agradecimentos

Ao meu orientador, Prof. Léo Pini Magalhães, sou grata pela orientação.

Ao meu esposo Ricardo, por me ajudar a realizar este trabalho.

A minha amiga Letícia, pelas sugestões e apoio dados durante a realização deste trabalho.

Aos demais colegas de pós-graduação, pelas críticas e sugestões.

Aos meus amigos da minha terra (RS), que mesmo longe, sempre me apoiaram e estiveram comigo.

As meninas que moraram comigo em Campinas, as quais foram a minha família, minhas companheiras e amigas. Alexandra, Érica, Mariana, Marina, Michelly, Paula, Rita, Thais, obrigada, foi um prazer morar com vocês.

A minha família pelo apoio durante esta jornada.

Ao CNPq, pelo apoio financeiro.

Sumário

Lista de Figuras	viii
Lista de Tabelas	x
1 Introdução	1
1.1 Motivação para o estudo sobre <i>morphing</i>	1
1.2 Motivação para a aplicação em envelhecimento facial	2
1.3 Objetivo	3
1.4 Estrutura da dissertação	3
2 Morphing	4
2.1 Definição de Morphing	4
2.1.1 Especificação de características	5
2.1.2 Deformação	5
2.1.3 Transformação de tonalidades	15
2.1.4 Controle de transição	17
2.2 Aplicações de Morphing	18
2.2.1 View Morphing	18
2.2.2 Polymorph	19
2.3 Considerações Finais	21
3 Funções de Base Radial	22
3.1 Transformação Espacial	22
3.1.1 Interpolação e Aproximação	23
3.1.2 Transformações Locais e Globais	23
3.1.3 Modelos Rígidos e Deformáveis	23
3.2 Descrição das Funções de Base Radial	24
3.2.1 Propriedades das Funções de Base Radial	25
3.2.2 Resolução do Sistema Linear	27
3.2.3 Funções de base	29
3.3 Considerações Finais	30
4 Envelhecimento Facial	32
4.1 Características do envelhecimento facial	33

4.2	Simulação Computacional do envelhecimento	33
4.2.1	Métodos para estimar a idade facial	34
4.2.2	Métodos de textura	35
4.2.3	Métodos que alteram a estrutura da face	38
4.3	Considerações Finais	40
5	Modelagem do Envelhecimento	43
5.1	Regiões faciais sujeitas ao envelhecimento	43
5.2	Curvas de envelhecimento	47
5.2.1	Região Frontal - Altura da Testa	49
5.2.2	Região Frontal - Altura da Sobrancelha	49
5.2.3	Região Orbitária - Altura da Bolsa Palpebral	50
5.2.4	Região Nasal - Altura e largura do nariz	51
5.2.5	Região Orolabial - Altura do lábio superior e espessura labial	53
5.2.6	Região Mentoneana	54
5.3	Considerações Finais	55
6	Implementação e Resultados	56
6.1	Implementação do Protótipo	56
6.1.1	Pré-processamento da Imagem	57
6.1.2	Cálculo das medidas dos segmentos de cada parâmetro	58
6.1.3	Cálculo das novas medidas dos segmentos e pontos finais	60
6.1.4	Cálculo dos coeficientes da função RBF	61
6.1.5	Transformação RBF	64
6.2	Resultados	65
7	Conclusões e Trabalhos Futuros	72
7.1	Trabalhos futuros	73
	Referências bibliográficas	74

Lista de Figuras

1.1	Comercial utilizando <i>morphing</i>	2
2.1	Mesh warping.	6
2.2	Vetores iniciais e finais.	8
2.3	Exemplo do método minimização de energia.	13
2.4	Configuração inicial da treliça sobre a imagem.	14
2.5	Exemplo do método MFFD.	15
2.6	Mistura de cores (Blend).	15
2.7	Interpolação de cores (<i>Cross-Dissolve</i>).	16
2.8	Letra 'F' sofrendo deformações para tornar-se a letra 'A' (esquerda->direita).	17
2.9	Letra 'A' sofrendo deformações para tornar-se a letra 'F' (direita->esquerda).	17
2.10	Resultado final do morphing entre as letras 'F' e 'A'.	17
2.11	Esquema de vistas.	19
2.12	Exemplo do View Morphing.	19
2.13	Exemplo da técnica <i>Polymorph</i> entre 3 imagens.	20
3.1	Gráfico de algumas funções base.	29
3.2	Comparação entre valores do parâmetro σ da função Gaussiana.	31
4.1	Envelhecimento.	32
4.2	Regiões da face usadas na estimação da idade.	35
4.3	Transferência de textura para simular envelhecimento.	37
4.4	Processamento de textura por wavelet.	38
4.5	Exemplos de envelhecimento.	39
4.6	Exemplo dos resultados obtidos por Gandhi para a simulação de envelhecimento.	40
4.7	Exemplos de envelhecimento por BRDF.	41
5.1	Regiões Faciais.	44
5.2	Pontos marcados na imagem facial.	45
5.3	Medidas de um mesmo parâmetro em quatro pacientes.	48
5.4	Curvas de envelhecimento da testa - central e lateral.	50
5.5	Curvas de envelhecimento da sobrancelha - central e lateral.	51
5.6	Curvas de envelhecimento da bolsa palpebral.	52
5.7	Curvas de envelhecimento do nariz.	52
5.8	Curvas de envelhecimento dos lábios.	53

5.9	Curvas de envelhecimento do mento.	54
6.1	Diagrama.	57
6.2	Normalização da face.	58
6.3	Linhas verticais sobre a face.	59
6.4	Comportamento da RBF alterando suas constantes.	63
6.5	Exemplo de imagens do banco de dados FGNET.	65
6.6	Simulação de envelhecimento de 20 para 60 anos ($\sigma = 30, \lambda = 1$).	66
6.7	Simulação de envelhecimento de 20 para 30 anos ($\sigma = 30, \lambda = 1$).	67
6.8	Simulação de envelhecimento de 20 para 40 anos ($\sigma = 30, \lambda = 1$).	68
6.9	Simulação de envelhecimento de 10 em 10 anos ($\sigma = 30, \lambda = 1$).	69
6.10	Simulação de comparação e comportamento da RBF. ($\sigma = 30, \lambda = 1$)	70
6.11	Simulação de envelhecimento de 22 para 42 anos. ($\sigma = 30, \lambda = 1$)	71

Lista de Tabelas

3.1	Comparação entre os diferentes tipos de base.	31
4.1	Métodos para estimar a idade.	41
5.1	Segmentos medidos na face.	46
5.2	Coeficientes dos polinômios da altura da testa (equação 5.5).	49
5.3	Coeficientes dos polinômios da altura da sobrancelha (equação 5.5).	50
5.4	Coeficientes dos polinômios da altura da bolsa palpebral (equação 5.5).	51
5.5	Coeficientes dos polinômios de altura e largura do nariz (equação 5.5).	51
5.6	Coeficientes dos polinômios da espessura e altura dos lábios (equação 5.5).	53
5.7	Coeficientes dos polinômios para a região mentoneana (equação 5.5).	54
6.1	Coeficientes dos polinômios para cada segmento da face.	60
6.2	Pontos calculados - Figura 6.3.	61

Capítulo 1

Introdução

O processamento de imagens lida com a manipulação de imagens, por exemplo, transformações que tomam imagens como entrada produzindo outras imagens como saída. O *warping* e o *morphing* de imagens referem-se a dois tipos fundamentais de transformações de imagens. O *warping* de imagens é uma transformação geométrica genérica que modifica a relação espacial de pontos na imagem, variando desde uma simples translação ou rotação de imagem à deformações no espaço bastante complexas. O *morphing* de imagens é uma transformação completa de uma imagem, onde não apenas a relação espacial entre os pontos é alterada (transformações de *warping*), mas o valor dos pixels na imagem nestes pontos também. Ambos os tipos de transformação têm diversas aplicações e são importantes tópicos de pesquisa.

As aplicações de *morphing* têm usualmente se enquadrado em duas categorias principais: aquelas usadas para gerar uma sequência de transformações e aquelas que criam um objeto intermediário como um fim. Assim, o *morphing* pode ser usado para visualizar transformações, ou como uma ferramenta de modelagem. Neste trabalho, o foco foi utilizar o *morphing* como uma ferramenta de modelagem.

A computação gráfica atualmente encontra aplicações em várias áreas do conhecimento humano. Mais recentemente ela tem adquirido cada vez mais espaço, tornando-se indispensável em diversas áreas, como por exemplo na área médica. As questões relativas ao envelhecimento e à longevidade humana vêm motivando o interesse de diversos especialistas, que buscam a melhoria na qualidade de vida das pessoas. A idéia de representar o envelhecimento de imagens faciais utilizando *morphing* motivou a realização deste trabalho.

1.1 Motivação para o estudo sobre *morphing*

Existem importantes aplicações de *morphing* nos campos de sensoriamento remoto, imagens médicas, visão computacional e entretenimento. Um problema clássico é o casamento de um mosaico de fotografias. Nestes casos, *warping* tem sido usado para deformar, rotacionar e transladar as imagens de forma que haja um encaixe perfeito.

As imagens médicas também se beneficiaram das técnicas de *warping* desenvolvidas inicialmente para aplicações de sensoriamento remoto. Problemas de desalinhamento ocorrem em exames que produzem imagens, devido ao movimento do paciente entre as tomadas ou se exames diferentes pre-

cisam ser combinados. Usualmente, pontos de referência conhecidos servem como indicadores para deformar a imagem de tal forma que as principais características casem aproximadamente, simplificando a interpretação dos resultados.

Um dos principais usos de *morphing* é na indústria de entretenimento, como pode ser visto em diversas produções. O primeiro filme a utilizar *morphing* foi *Willow* em 1988, devido a *Doug Smythe* que fez um *software* de deformação e metamorfose. Houve também um comercial da Esso, no qual um carro se transforma em um tigre, em 1991 (figura 1.1).



Fig. 1.1: Comercial utilizando *morphing*.

1.2 Motivação para a aplicação em envelhecimento facial

A necessidade de ter um sistema computacional para mostrar como as pessoas serão no futuro tem levado muitos pesquisadores a desenvolver e estudar protótipos que podem fazer envelhecimento de faces através de fotografias. As pessoas tem curiosidade nesse assunto, afinal quem não gostaria de se ver a alguns anos à frente e, quem sabe, poder até retardar ou atenuar os efeitos do tempo sobre a pele. Mas, além da curiosidade e estética, algumas das aplicações do envelhecimento computacional citadas em [Gandhi (2004)] são as seguintes:

- Achar fugitivos procurados pela polícia: muitos bandidos não possuem fotos atualizadas, somente de quando eram mais novos, portanto ter um sistema que permita atualizar a aparência física de tais indivíduos permite localizá-los com maior facilidade. [CrimeLab (2006)]
- Crianças desaparecidas: uma aplicação especial é para auxiliar na localização de crianças e pessoas desaparecidas por muitos anos.
- Fazer atualização de fotografias de empregados: muitas empresas mantêm seus empregados por muitos anos, porém sem atualizar suas fotos. Com um sistema de envelhecimento é possível atualizar esses dados sem a necessidade de pedir uma nova fotografia ao empregado.
- Reconhecimento de faces: os sistemas de reconhecimento de faces tem a necessidade de um sistema de envelhecimento para funcionarem corretamente e por muitos anos, sem a necessidade do cadastramento das faces para o reconhecimento.

1.3 Objetivo

Este trabalho tem o objetivo de apresentar um protótipo para a simulação de envelhecimento de imagens faciais frontais. Para isto, foi realizada uma revisão de métodos de *warping* e de técnicas de simulação de envelhecimento de faces. O protótipo utilizou a técnica chamada de Funções de Base Radial (RBF) juntamente com um modelo quantitativo de envelhecimento. O protótipo trabalha com imagens de pessoas a partir de 20 anos e realiza o envelhecimento até no máximo 70 anos.

1.4 Estrutura da dissertação

Capítulo 2 : define os conceitos de *morphing* e *warping* e descreve o processo de *morphing*, mostrando como ele é realizado. As principais técnicas existentes de *morphing* são abordadas e, ainda são ilustradas algumas aplicações de *morphing*.

Capítulo 3 : descreve conceitualmente e matematicamente as funções de base radial, que foi a técnica de *warping* utilizada neste trabalho para realizar o envelhecimento.

Capítulo 4 : aborda algumas características do envelhecimento facial e apresenta uma revisão bibliográfica sobre os principais sistemas de envelhecimento existentes e quais técnicas são utilizadas.

Capítulo 5 : apresenta a modelagem matemática do envelhecimento utilizada para representar o envelhecimento facial.

Capítulo 6 : mostra como foi realizada a implementação utilizando os conceitos envolvidos nos capítulos 3 e 5. Além disso, são apresentados os resultados obtidos.

Capítulo 7 : apresenta as conclusões e trabalhos futuros que podem ser desenvolvidos como continuidade ou para o aprimoramento deste trabalho.

Capítulo 2

Morphing

Neste capítulo serão apresentadas a definição de *morphing*, as fases que o compõem, as principais técnicas de deformação existentes até o momento e por fim, aplicações de *morphing*. O estudo dos principais algoritmos de deformação foi de fundamental importância para decidir qual método se adequava melhor para a aplicação proposta de envelhecimento facial por imagens.

Antes de iniciar o capítulo, cabe salientar que não existe uma tradução de uso corrente para o termo *morphing*, então durante o decorrer desta dissertação, o termo em inglês será usado.

2.1 Definição de Morphing

Morphing é a metamorfose gradual entre duas ou mais imagens diferentes envolvendo deformação espacial e transformação de tonalidades. Assim, uma imagem inicial será gradualmente transformada em uma imagem final. Para que isso aconteça, é necessário antes de mais nada, a definição de pares de características, ou seja, a definição de elementos correspondentes em ambas as imagens, como por exemplo, a região dos olhos e/ou da boca (se as imagens forem de faces). Essas regiões de interesse podem ser marcadas com pontos, retas, curvas, etc. Essa fase é chamada de “especificação de características” e é usada para criar funções de mapeamento que definem uma relação espacial entre todos os pontos em ambas as imagens [Wolberg (1998)]. Essas funções de mapeamento fazem parte do processo de *warping*, ou seja, da fase de deformação e elas são chamadas de *warp functions*. As funções de mapeamento são usadas para interpolar as posições dos elementos correspondentes nas imagens durante esta fase do processo de *morphing*.

Neste trabalho o termo “*warping*” e o termo “deformação” serão usados como sinônimos. O termo “*warp functions*” pode ser expresso por “funções de deformação” ou “funções de mapeamento”. Já foi dito anteriormente que o termo “*morphing*” não será traduzido para o português.

O *morphing* de imagens é composto basicamente de 4 fases: especificação de características, deformação (ou *warping*), transformação de tonalidades e controle de transição. Esta última fase permite criar animações com uma taxa de transição variável, como por exemplo, transformar uma imagem em outra deformando partes diferentes no decorrer do tempo.

2.1.1 Especificação de características

A especificação de características é a fase mais entediante no processo de *morphing*. Esta especificação é realizada com base em uma variedade de primitivas, exigindo atenção especial do animador no posicionamento preciso das mesmas. Estas primitivas podem ser: pontos, segmentos de reta, grades, curvas, etc. Estas primitivas são posicionadas na imagem sobre as características que elas irão representar, por exemplo, regiões da face como boca, nariz e olhos (se a imagem contém uma face).

As primitivas precisam ser especificadas em ambas imagens, a inicial e a final. E, precisa haver uma correspondência entre as características da imagem inicial e final, ou seja, se existe uma primitiva na imagem inicial sobre a boca, então deverá existir a mesma primitiva sobre a boca da imagem final.

Conforme será visto no decorrer deste capítulo, cada técnica de deformação existente apresenta a especificação de características de uma forma diferente, alterando a primitiva a ser usada. Com isso, houve uma evolução na maneira de especificar as características tentando diminuir o incômodo desta fase. Algoritmos foram desenvolvidos para minimizar a interação com o usuário e, entre eles destaca-se, um método chamado *snakes*. Este conceito foi adotado primeiramente em visão computacional como um modelo de contorno ativo [Kass e Terzopoulos (1988)]. *Snakes* ajudam na especificação das características de interesse porque as primitivas somente precisam ser posicionadas perto das mesmas. Então forças empurram as *snakes* para bordas salientes na imagem. Para isso, utilizam funções de minimização de energia que "puxam" a *snake* para uma característica desejada da imagem. *Snakes* tem um melhor resultado quando as características a serem extraídas da imagem tem um valor de gradiente grande.

2.1.2 Deformação

Esta fase é composta pela deformação das imagens através do alinhamento das características especificadas na fase anterior (especificação de características). É através da especificação de elementos que é possível realizar um mapeamento entre as imagens permitindo que a metamorfose ocorra. Segundo [Wolberg (1998)] existe uma relação direta entre a complexidade da especificação de características e a deformação. A medida que a especificação de características se torna mais conveniente para o usuário, a deformação se torna mais precisa, portanto a deformação depende de como é realizada essa especificação.

A deformação final obtida difere em função do método utilizado. Os seguintes métodos serão explicados em seguida:

- Mesh warping [Wolberg (1994)];
- Field morphing [Beier e Neely (1992)];
- Radial basis function [Arad et al. (1994); Litwinowicz e Williams (1994)];
- Thin Plate splines [Lee et al. (1994)];
- Energy minimization [Lee et al. (1996)];
- Multilevel free-form deformation [Lee, Chwa e Shin (1995)].

Mesh Warping

Mesh warping foi o primeiro método de *warping* desenvolvido sendo utilizado pela empresa Industrial Light Magic (ILM) por D. Smythe no filme Willow em 1988. Posteriormente, foi usado em muitos outros filmes subsequentes. Em 1990, *George Wolberg* escreveu um livro descrevendo este método e explicando todo o processo de *Morphing* e *Warping* [Wolberg (1994)].

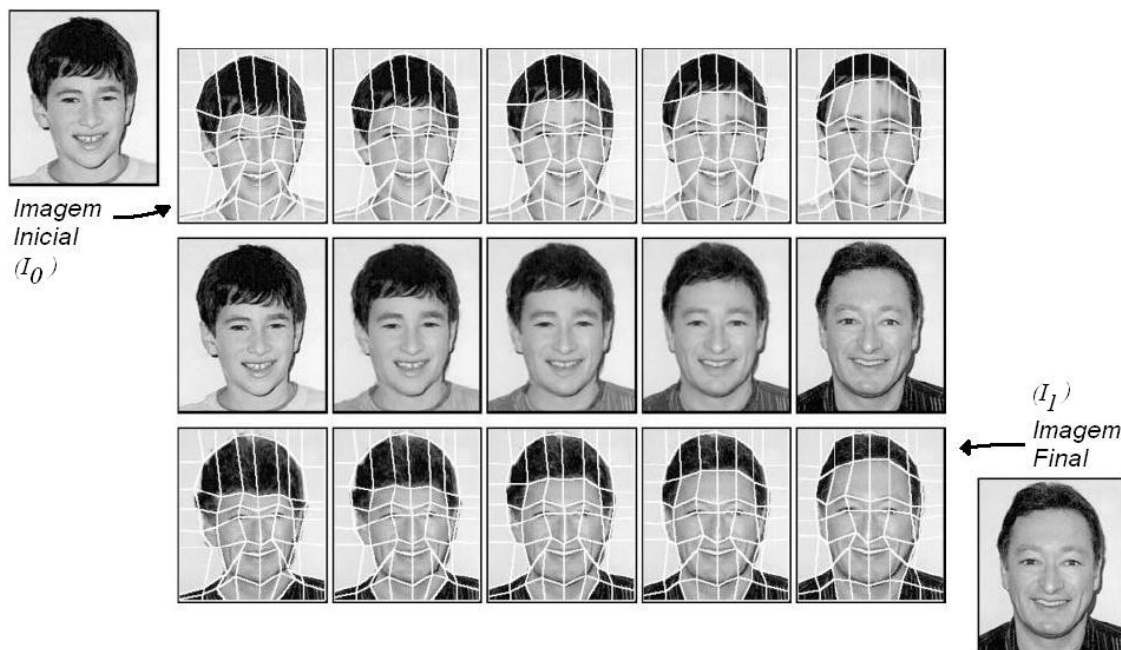


Fig. 2.1: Mesh warping.

Este método de *warping* é baseado em uma transformação entre duas imagens: uma inicial chamada de I_0 e uma imagem final chamada de I_1 . Após a definição das imagens que serão usadas no processo, são identificadas correspondências entre as imagens afim de se estabelecer um critério para a deformação (transformação). Essa correspondência, no caso deste algoritmo é feita por uma grade em ambas as imagens, como ilustrado na figura 2.1. Cada linha da grade das imagens inicial e final representam a mesma característica, conforme a figura. A correspondência de características está associada aos pontos de intersecção das linhas nas grades, portanto, o número de pontos de intersecção (nós) necessita ser o mesmo em ambas imagens. Em seguida, cada ponto de intersecção da grade da imagem inicial é deslocado para a posição correspondente do ponto de intersecção da imagem final e as células das grades são então deformadas de tal forma que acompanhem este deslocamento dos pontos de intersecção. Essa deformação pode ocorrer em várias etapas, gerando vários quadros de imagens, até conseguir ter a mesma forma de I_1 . O inverso também ocorre: a imagem I_1 é deformada até a imagem I_0 , utilizando a mesma função, e também gerando vários quadros afim de que os quadros da primeira seqüência sejam interpolados com os quadros da segunda seqüência, gerando a metamorfose.

A grade associada à imagem inicial I_0 será chamada de M_0 e a da imagem final I_1 será chamada de M_1 na qual as posições dessa estão relacionadas com as posições da grade M_0 . Juntas, M_0 e M_1 são

usadas para definir uma transformação espacial que mapeia todos os pontos de I_0 em I_1 . As grades são construídas para serem topologicamente iguais, por exemplo, nenhuma dobra ou descontinuidade é permitida, conseqüentemente, os pontos de intersecção, ou nós, das grades podem deslocar-se se for necessário, contanto que não causem intersecção.

Todos os quadros intermediários na seqüência de metamorfose são produtos de um processo de quatro passos:

1. Uma interpolação linear precisa ser realizada entre a grade da imagem inicial M_0 e a grade da imagem final M_1 . Essa interpolação linear vai gerar uma nova grade que é chamada de M ;
2. Ocorre uma deformação da imagem inicial I_0 usando como base para essa deformação as grades M_0 e M , onde a grade M representa como será a deformação final da imagem nesta etapa. Esta nova imagem deformada é chamada de I_2 ;
3. Aqui ocorre um processo idêntico ao descrito no item anterior. Entretanto, a deformação é realizada na imagem final I_1 usando como base para essa deformação as grades M_1 e M onde a grade M representa como será a deformação final da imagem nesta etapa. Esta nova imagem deformada é chamada de I_3 ;
4. Por fim, é realizada uma interpolação linear entre as imagens I_2 e I_3 que foram geradas nos itens anteriores, gerando assim a imagem final que será um dos quadros da seqüência.

Para ilustrar esse algoritmo, considera-se a seqüência de imagens na figura 2.1. Os três quadros no centro da figura (linha central), representam a metamorfose entre as duas faces representadas pelas imagens inicial e final (imagens I_0 e I_1). Na linha superior da figura, a grade M_0 é deformada para a grade M_1 , produzindo uma grade intermediária chamada M , para cada quadro. Essas grades são usadas para deformar I_0 aos poucos de forma que se transforme em I_1 . O processo idêntico, porém em ordem reversa, ou seja, a grade M_1 é deformada até M_0 , é mostrado na linha inferior da figura 2.1. O propósito desse procedimento é manter o alinhamento das grades entre as imagens iniciais e finais, I_0 e I_1 , onde ambas imagens deformam-se para um estado intermediário, produzindo pares de imagens I_2 e I_3 , mostradas em cima e em baixo na figura.

Somente depois deste alinhamento ser estabelecido, ou seja, a deformação ser realizada, uma transformação de tonalidades, conforme será descrito na seção 2.1.3 e comumente chamada de *cross-dissolve*, entre os pares sucessivos de imagens I_2 e I_3 , é realizada criando uma nova imagem (quadro) para a seqüência de *morphing*, como mostrado nas imagens na linha central da figura 2.1.

Field Morphing

Esta técnica de *warping* foi proposta por Thaddeus Beier and Shawn Neely [Beier e Neely (1992)] em 1992. Na época, diferenciou-se da técnica da seção anterior (*Mesh Warping*) pela utilização de segmentos de reta, ou vetores, na fase de especificação de características de transformação. Estes segmentos de reta determinam a área de deformação nas imagens. Dependendo do número de vetores, e da localidade, um vetor pode influenciar na área de deformação de outro.

O vetor da figura 2.2, denotado por $\overrightarrow{PQ}$ é posicionado sobre alguma característica da imagem inicial I_0 , assim como o vetor $\overrightarrow{P'Q'}$ é posicionado sobre esta mesma característica da imagem final I_1 .

Por exemplo, como mostrado na figura, o vetor $\overrightarrow{PQ}$ é definido sobre as características de interesse (a figura apresenta 9 possíveis vetores $\overrightarrow{PQ}$). O ponto X indica o pixel a ser analisado e conforme a posição dos vetores iniciais e finais ($\overrightarrow{PQ}$ e $\overrightarrow{P'Q'}$), da posição do pixel sobre a imagem e dos valores de constantes (que serão especificadas mais adiante), o pixel poderá ter sua localização na imagem destino alterada para uma outra localização ou não.

Para ser realizado o cálculo da nova localização do pixel na imagem final, é necessário usar as coordenadas dos pontos X , P , P' , Q e Q' e algumas medidas que podem ser observadas na figura 2.2. Essas medidas são: v , é a distância do ponto X ao ponto que corta o vetor $\overrightarrow{PQ}$ através de uma linha perpendicular a esse vetor, representando a distância na direção perpendicular do vetor ao ponto; u é a distância medida pelo ponto P até onde a linha perpendicular que passa pelo ponto X corta o vetor; é chamada de distância ao longo do vetor. Esses valores são calculados para se obter a nova localização do pixel X , que será chamado de X' na imagem deformada. Eles são calculados da seguinte maneira:

$$u = \frac{(X - P)(Q - P)}{\|Q - P\|^2} \quad (2.1)$$

$$v = \frac{(X - P)(Q - P)^\perp}{\|Q - P\|} \quad (2.2)$$

$$X' = P' + u(Q' - P') + \frac{v(Q' - P')^\perp}{\|Q' - P'\|} \quad (2.3)$$

os pontos denotados por $'$ são os pontos da imagem destino. A notação $\|...\|$ expressa o cálculo da distância euclidiana entre dois pontos e a notação $(...)^{\perp}$ significa o vetor $(-y, x)$ do vetor (x, y) .

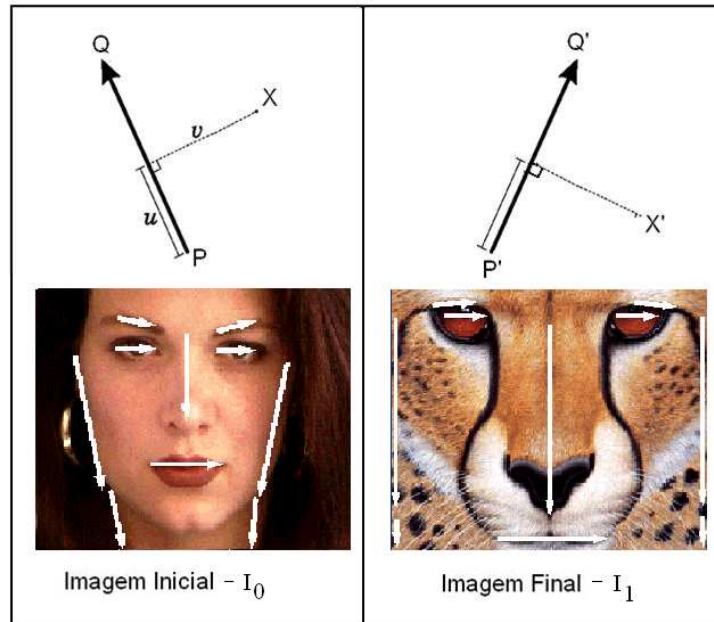


Fig. 2.2: Vetores iniciais e finais.

Pode-se notar que conforme a distância do pixel (ponto X) ao vetor aumenta (v), menos influenciado ele será pelo segmento. Entretanto, essa influência pode ser mudada alterando os valores das seguintes constantes:

- *peso* : tendo mais de um vetor definido, um pixel pode ser influenciado por ambos segmentos, portanto é necessário que exista um peso para cada segmento e, este foi definido como sendo inversamente proporcional à distância do ponto a esse vetor, pela seguinte equação:

$$peso = \left(\frac{tamanho^p}{a + dist} \right)^b \quad (2.4)$$

onde *tamanho* é o módulo do vetor, *dist* é a distância do pixel ao vetor, e a , b e p são as constantes usadas para mudar a influência que cada vetor tem sobre os pixels;

- a : essa constante pode ser vista como a aderência ao vetor, para valores próximos a zero, pontos próximos ao vetor serão mapeados exatamente como o vetor determina; para valores maiores, pontos estão mais soltos, e livres para se movimentar, mesmo que estejam localizados sobre o vetor;
- p : a importância do comprimento do vetor é controlada por p , se ele é zero, o comprimento é ignorado; caso contrário, vetores mais longos são mais importantes;
- b : a interpretação de b pode ser vista como a concentração de força de uma característica, valores altos aumentam a força próxima do vetor, entretanto a força é pequena quando distante e com decaimento bastante rápido; pequenos valores tornam esse decaimento mais lento, reduzindo a localidade da característica.

Nota-se que diferentes valores de constantes podem ser associados a cada característica, apesar de talvez ser mais conveniente usar um valor global para b .

Radial Basis Function e Thin Plate Splines

Radial Basis Function (RBF) ou funções de base radial é um método de deformação que surgiu ao início da década de 90. As principais referências são: Ruprecht e Müller [Ruprecht e Müller (1995)], Ruprecht e Müller [Ruprecht e Müller (1994)], Arad [Arad et al. (1994)] e Litwinowicz e Williams [Litwinowicz e Williams (1994)]. Foi um método que inovou na época, diferindo dos outros existentes em relação a dois aspectos:

- a maneira como são especificados os pares de características entre as imagens utiliza pontos nas imagens iniciais e finais para estabelecer a correspondência de elementos entre ambas imagens;
- utilização de funções de base radial na fase de deformação das imagens.

Uma RBF é uma transformação que faz um mapeamento de uma imagem em outra a partir de um conjunto de pontos. Estes pontos são chamados de pontos de controle e eles descrevem a posição de uma característica na imagem. Cada ponto na imagem inicial tem somente um ponto correspondente na imagem final (deformada), resultando numa sequência de pares de controle. Portanto, são informados pares de pontos de controle $\vec{p}_i, \vec{q}_i$ nas imagens inicial e final, onde $\vec{p}_i = (x_i, u_i)$ é a coordenada

no estado inicial e $\vec{q}_i = (u_i, v_i)$ é a coordenada no estado final, ou seja, o lugar onde a coordenada inicial deverá estar após a transformação. Portanto, uma função de base radial é uma transformação que descreve uma relação espacial entre estes pontos de controle correspondentes, ou seja, mapeia um conjunto de n pontos de controle $\vec{p}_i$, da imagem inicial em pontos correspondentes $\vec{q}_i$ da imagem deformada ($i = 1, \dots, n$). Esse mapeamento é representado por funções definidas como uma combinação de funções de base radialmente simétricas, cada uma centrada em um ponto de controle. Após essas funções serem encontradas, elas são usadas em todo o domínio da imagem para realizar a deformação na imagem inteira.

As funções de base radial apresentam um comportamento facilmente controlável já que não existem restrições nas posições dos pontos de controle sobre as imagens. Esse tipo de função também permite realizar deformações tanto locais como deformações globais. Algumas das funções de base são: *Thin-plate Spline*(TPS), Gaussiana, Multiquadrática(MQ) e *Shifted-LOG*. Essas funções base e uma descrição melhor deste método serão abordados no capítulo 3, tendo em vista que este método foi o utilizado no presente trabalho.

Entretanto, a função *Thin plate splines* merece uma atenção especial, apesar desta função ser utilizada como uma função base para as funções de base radial, ela é destacada como um outro método de *morphing*. *Thin plate splines* são funções que apresentam características interessantes para a aplicação em deformações. Estas funções modelam uma superfície infinita de metal delgado, na qual pontos de carga provocam deformações. A sua formulação assegura as restrições de que a superfície interpolante apresenta mínima energia de deformação, com relação aos pontos de carga impostos, e que seja suave. *Thin plate splines* são formadas através de combinação linear de funções radiais, o que significa que os valores da função são obtidos pela diferença (ou distância) de coordenadas do ponto onde a função deve ser avaliada para as coordenadas de pontos de referência. A medida que se afasta dos pontos de carga, a função apresenta um comportamento quase linear. Assim, embora seja uma função de interpolação global, a influência das deformações locais tendem a diminuir a medida que se afasta do ponto de deformação. Com a função TPS, os efeitos das distorções locais afetam toda a imagem, porém seu efeito diminui a medida que se afasta do ponto de influência da deformação.

Minimização de energia - Energy minimization

O nome deste método vem do fato de utilizar funções de energia para representar as exigências para uma deformação suave, ou seja, a forma desejada do objeto é derivada pela minimização de energias associadas. Os problemas de minimização de energia são então transformados em equações diferenciais parciais, geralmente resolvidas por métodos numéricos. O método proposto por Lee [Lee, Chwa e Shin (1995)] utiliza *Multigrid Relaxation* para resolver tanto as equações diferenciais do sistema linear na derivação das funções de deformações, como para solucionar as funções de controle de transição (ver seção 2.1.4). Este método também não é restritivo quanto ao conjunto de pontos, ou seja, os pontos podem ser localizados em qualquer lugar na imagem.

Para este método, considera-se a imagem como uma placa retangular no plano xy e um conjunto de pontos dispersos no espaço representando o conjunto de suas características. Cada ponto é conectado à placa por uma mola comprimida. O esticamento das molas causa um aumento da energia potencial que se reduz enquanto a placa é deformada para os pontos finais. Quando a placa é deformada, a elasticidade da placa impede uma dobra, resultando em uma forma suave. Esse modelo de placa provê uma interpretação intuitiva do problema de construção de superfícies. O problema

tem como entrada um conjunto de pontos P que são interpolados permitindo uma superfície com continuidade de primeira ordem (C^1). A deformação dessa superfície pode ser representada por uma função $z = f(x, y)$. A interpolação dos pontos é conseguida pela minimização da energia potencial devido às molas,

$$E_I(f) = \beta \sum_{k=1}^p (f(x_k, y_k) - z_k)^2 \quad (2.5)$$

onde (x_k, y_k, z_k) são as coordenadas do conjunto de pontos P , p é o tamanho de P e β é uma constante de mola, real e positiva, controlando a tensão das molas e restringindo a interpolação dos pontos dependendo do seu valor.

A maleabilidade da superfície em deformar-se pode ser aproximada pela minimização da integral da variação de curvatura sobre a superfície,

$$E_C(f) = \int \int_{\Omega} [(\frac{\partial^2 f}{\partial x^2})^2 + 2(\frac{\partial^2 f}{\partial x \partial y})^2 + (\frac{\partial^2 f}{\partial y^2})^2] dx dy \quad (2.6)$$

onde Ω denota o domínio retangular do plano xy em que a superfície f é definida. Como uma dobra causa uma variação de curvatura infinita, a superfície resultante f tem a primeira derivada parcial contínua. A área de deformação pode ser relacionada com uma cola entre a placa e o plano xy . Quando a placa destaca-se do plano xy , a cola ocasiona uma energia que pode ser estimada por

$$E_L(f) = \alpha \int \int_{\Omega} f^2 dx dy \quad (2.7)$$

O parâmetro α corresponde à viscosidade da cola, ele controla a resistência da superfície f em desgrudar-se do plano xy .

Conseqüentemente, a superfície desejada f pode ser obtida pela minimização da seguinte energia funcional, que é a soma dos termos de energia das equações 2.5, 2.6 e 2.7.

$$E(f) = \frac{1}{2} (E_C(f) + E_L(f) + E_I(f)) \quad (2.8)$$

Quando o parâmetro β é suficientemente grande, a superfície construída f estritamente interpola os pontos em P . Um valor alto para α , permite que a superfície f desgrude-se do plano xy somente na vizinhança de pontos do conjunto P .

Para se ter a minimização de energia representada pela equação 2.8, a derivada da soma dos termos na equação deve ser igual a zero.

$$\frac{\delta E}{\delta f} = [\frac{\partial^4 f}{\partial x^4} + 2\frac{\partial^4 f}{\partial x^2 \partial y^2} + \frac{\partial^4 f}{\partial y^4}] + \alpha f + \beta(f(x_k, y_k) - z_k) = 0 \quad (2.9)$$

A equação diferencial parcial 2.9 é conhecida como a equação de Euler-Lagrange. Para se aplicar esta equação em uma imagem o domínio de Ω foi discretizado em uma malha regular $M \times N$, que representa o número de pixels e a superfície f é representada pelos valores nos nós da malha. Cada nó da malha é representado por índices (i, j) e o valor da função f em cada um destes nós é representado por f_{ij} . O conjunto de pontos P é convertido em um conjunto (i, j, z_{ij}) onde z_{ij} representa o valor de f_{ij} . A discretização da equação diferencial apresentada em 2.9 é feita utilizando o método das

diferenças finitas transformando-a em um sistema de equações lineares com $M.N$ equações e $M.N$ variáveis desconhecidas. Para resolvê-la é necessário reescrever o sistema na forma matricial:

$$Af = b \quad (2.10)$$

onde A é uma matriz de dimensão $M.N \times M.N$ que compreende os múltiplos coeficientes multiplicados pelas variáveis; f corresponde à função de deformação e, b é um vetor de dimensão $M.N$ que contém as partes referentes da equação 2.9 com os coeficientes β e α .

Então, como já foi dito anteriormente, tendo duas imagens I_0 e I_1 , busca-se encontrar uma sequência de imagens $I(t)$ tal que $I(0) = I_0$ e $I(1) = I_1$. Neste caso, para derivar uma função de deformação W_0 , um animador precisa especificar um conjunto de pontos S sobre as imagens I_0 e I_1 . Sendo W_0 a função de deformação que especifica uma correspondência de pontos de I_0 para cada ponto em I_1 , W_0 é composta por duas funções X e Y , como a seguir:

$$W_0(x, y) = (X(x, y), Y(x, y)) \quad (2.11)$$

sendo usado um deslocamento x de cada par de pontos em S para derivar a função X . Primeiro define-se X como

$$X(x, y) = x + \Delta X(x, y) \quad (2.12)$$

A função ΔX precisa satisfazer as seguintes restrições

$$\Delta X(x_0, y_0) = \Delta X = x_1 - x_0, \quad \text{onde } p_i = (x_i, y_i), \quad i = 0, 1 \quad \text{e} \quad (p_0, p_1) \in S \quad (2.13)$$

W_0 mapeia o ponto p_1 para cada par de pontos (p_0, p_1) em S . ΔX pode ser obtido pela geração da superfície deformável, onde o conjunto de ΔX de S executa um papel de restrição de interpolação. As funções ΔY e Y podem ser derivadas pelo mesmo método.

Na construção de superfícies com as funções ΔX e ΔY , β grande é geralmente usado quando a função de deformação move os pontos exatamente para o ponto correspondente de cada ponto no conjunto S . O parâmetro α controla o quão longe o movimento de um ponto é refletido na sua vizinhança. Valores elevados de α fazem as funções ΔX e ΔY serem zero sobre o domínio exceto na vizinhança dos pontos interpolados. Este resultado em uma função de deformação distorce uma imagem somente perto do pontos característicos.

Para ilustrar uma simulação utilizando o método de minimização de energia, tem-se o exemplo da figura 2.3 que foi extraído de Lee [Lee et al. (1994)]. Nesta figura, ocorre uma transformação entre um face de um homem (figura 2.3(a)) e a figura de gato (figura 2.3(b)). As imagens 2.3(c) e 2.3(d) ilustram os pontos marcando as características estabelecidas entre as imagens. A imagem 2.3(e) representa a imagem intermediária da transição de uma imagem para outra. A figura 2.3(f) também é uma figura intermediária, entretanto com uma taxa de transição diferente (seção 2.1.4) para cada parte do rosto (olhos, boca e nariz), portanto a figura parece mais com a imagem humana na região da face, ao contrário do contorno do rosto que já adquiriu as características do gato.

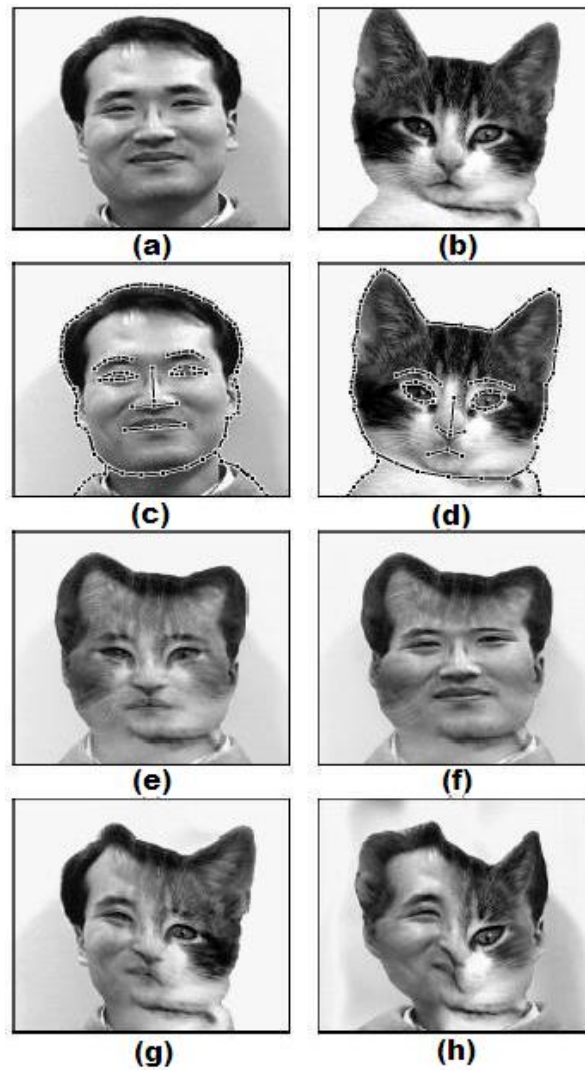


Fig. 2.3: Exemplo do método minimização de energia.

Multilevel free-form deformation-MFFD

Este método origina-se de uma técnica chamada *Free-form deformation* (FFD) proposta inicialmente por Sederberg and Parry em 1986 [Sederberg e Parry (1986)] como uma poderosa ferramenta para deformar objetos tridimensionais. A idéia básica é deformar um objeto pela manipulação de um paralelepípedo que contém este objeto.

A utilização de *Free-form deformation* (FFD) para deformar objetos bidimensionais foi proposta por Lee [Lee et al. (1996)]. A deformação ocorre com a sobreposição de uma treliça Φ (placa quadriculada onde os nós dessa placa podem ser deslocados) sobre a imagem Ω (veja a Figura 2.4). A deformação desta treliça implica na deformação da imagem.

Seja $p = (u, v)$ pontos sobre a imagem de tamanho $m \times n$. Seja Φ a treliça sobre a imagem de tamanho $(m+2) \times (n+2)$ onde $\phi_{ij}^0 = (i, j)$ é o ij ésimo nodo dessa treliça. A deformação da imagem no plano xy pode ser representada pela função $w(p) = (x(p), y(p))$ que é dada pelo deslocamento

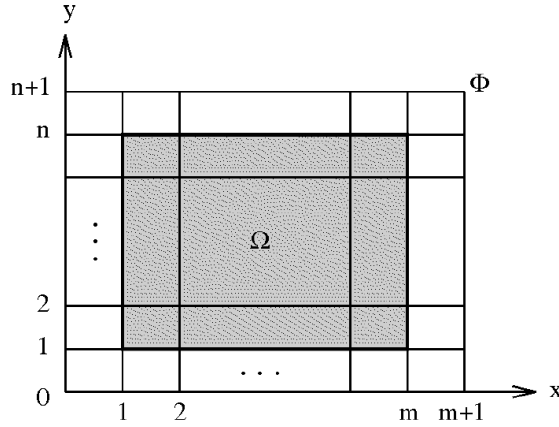


Fig. 2.4: Configuração inicial da treliça sobre a imagem.

dos pontos de controle da treliça (ϕ_{ij}).

A função de deformação $w(p)$ é realizada pelo produto tensor B-spline cúbico bivalente (equação 2.14). Esse produto é utilizado porque permite controle local quando um ponto de controle é deslocado para uma posição específica.

$$w(u, v) = \sum_{k=0}^3 \sum_{l=0}^3 B_k(s) B_l(t) \phi_{(i+k)(j+l)} \quad (2.14)$$

onde $i = \lfloor u \rfloor - 1$, $j = \lfloor v \rfloor - 1$, $s = u - \lfloor u \rfloor$, $t = v - \lfloor v \rfloor$. $B_k(s)$ e $B_l(t)$ são funções uniformes B-spline cúbica estimadas em s e t , respectivamente.

Em MFFD, uma hierarquia de treliças, $\Phi_0, \Phi_1, \dots, \Phi_m$ é usada para derivar uma seqüência de funções de deformação utilizando a manipulação por FFD.

Seja h_k o espaçamento entre pontos de controle na configuração inicial da treliça Φ_k . Quando a imagem Ω é deformada com uma treliça de maior espaçamento entre seus pontos, o resultado é uma deformação suave, entretanto os pontos finais não obedecem a sua posição final desejada. Como o objetivo é chegar na posição desejada, uma nova manipulação FFD é aplicada à deformação anterior onde o espaçamento vai ser dado por $h_k = 2h_{k+1}$ e um erro a cada iteração será computado.

Portanto, em MFFD, a manipulação FFD começa com uma treliça de controle Φ_0 com um grande espaçamento entre seus nodos. As manipulações FFD são computadas com uma treliça de controle Φ_k , e esta interage até que o erro atinja um valor limite. Então, uma treliça mais fina Φ_{k+1} é usada para sucessivas manipulações FFD, contanto que Φ_k não seja mais fina que a treliça da próxima iteração. Esse processo continua enquanto o erro exceder o valor limite e até que se atinja a deformação desejada.

A figura 2.5 mostra um exemplo utilizando o método MFFD. As figuras (a) e (b) ilustram as imagens inicial e final. As figuras (c) e (d) mostram a especificação de características sendo realizada utilizando *snakes* que proporcionam uma captura de características mais precisas em torno dos olhos, boca e etc. A figura (e) é a imagem intermediária do processo ($t = 0.5$) onde a mesma taxa de transição (seção 2.1.3) é aplicada sobre todos os pixels. E a figura (f) mostra a mesma imagem intermediária porém utilizando uma taxa de transição diferente para determinados pixels.

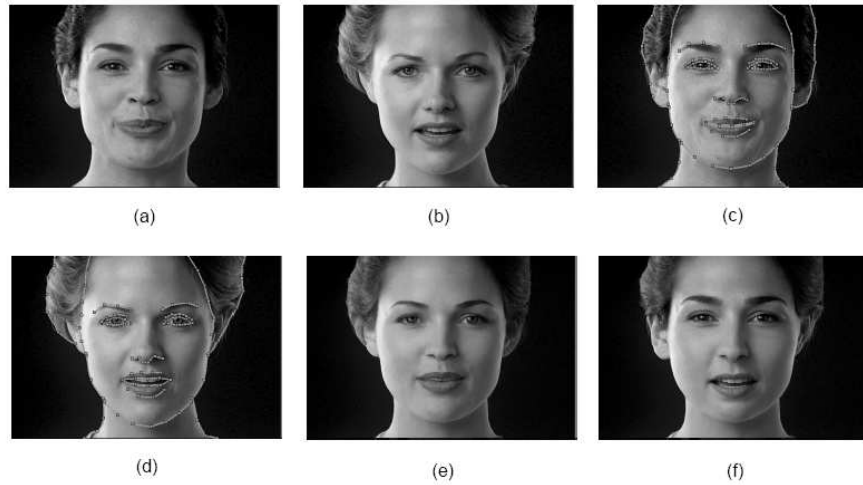


Fig. 2.5: Exemplo do método MFFD.

2.1.3 Transformação de tonalidades

Transformação de tonalidades é um processo onde as tonalidades das imagens geradas ao longo da deformação são interpoladas (figura 2.6). Assim, com a aplicação de uma função de deformação (ou *warping*) o processo de *morphing* cria uma seqüência de quadros, onde uma imagem inicial se transforma em outra final. O número de quadros definirá o tempo da animação. Esse número de quadros é importante, pois é através destes quadros que as imagens serão deformadas gradualmente. No processo de *morphing*, tanto a imagem inicial é deformada para a imagem final, como a imagem final é deformada para a imagem inicial. Portanto, as funções de deformação são usadas tanto na imagem inicial, como na imagem final. Então, após ocorrer a deformação, tem-se imagens intermediárias (quadros) que representam a transformação da imagem inicial na imagem final e vice-versa. Estas imagens intermediárias precisam passar pelo processo de transformação de tonalidades conhecido na literatura como *cross-dissolve*.

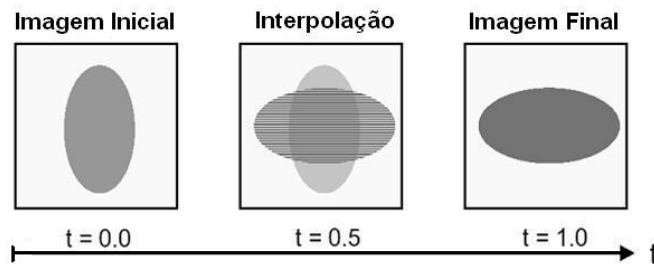


Fig. 2.6: Mistura de cores (Blend).

Matematicamente, tendo duas imagens I_0 e I_1 , o algoritmo *warping* busca as funções de deformação que definem uma seqüência de imagens $I(t)$ tal que $I(0) = I_0$ e que $I(1) = I_1$. Assume-se que o tempo t varia de 0 a 1 enquanto a imagem I_0 se transforma continuamente na imagem final I_1 .

Sendo $W_0(p)$ a função de deformação que especifica uma correspondência de cada ponto p em I_0 para cada ponto em q em I_1 , e quando aplicada a I_1 , $W_0(p)$ tem que gerar uma imagem deformada que se alinhe com I_0 nas posições das características. A função inversa $W_1(q)$ distorce I_0 para I_1 . Para gerar as imagens intermediárias $I(t)$, derivam-se duas funções $W_0(p; t)$ e $W_1(q; t)$ de $W_0(p)$ e $W_1(q)$ por interpolação linear no tempo t . I_0 e I_1 são então deformadas, resultando em imagens intermediárias $I_0(t)$ e $I_1(t)$, respectivamente. Finalmente, $I(t)$ é obtido pela interpolação de tonalidades entre $I_0(t)$ e $I_1(t)$. As equações 2.15 à 2.19 representam este processo.

$$W_0(p; t) = (1 - t) \cdot p + t \cdot W_0(p) \quad (2.15)$$

$$W_1(q; t) = t \cdot q + (1 - t) \cdot W_1(q) \quad (2.16)$$

$$I_0(t) = W_0(p; t) \bullet I_0 \quad (2.17)$$

$$I_1(t) = W_1(q; t) \bullet I_1 \quad (2.18)$$

$$I(t) = (1 - t) \cdot I_0(t) + t \cdot I_1(t) \quad (2.19)$$

onde $W \bullet I$ representa a aplicação da função de deformação W para a imagem I .

Para exemplificar esse processo, imagina-se que o ponto $p = (0, 0)$ da imagem I_0 e o ponto $q = (5, 5)$ da imagem I_1 são os pontos iniciais e finais, respectivamente. Aplicando as equações acima para $t = 0$, tem-se $W_0(p; 0) = p$ e $W_1(q; 0) = W_1(q)$, e sabendo que $W_1(q)$ é uma função de deformação que no ponto q deve ter como resultado o ponto p . Portanto, $W_0(p; 0)$ e $W_1(q; 0)$ são iguais. Após isso, $W_0(p; 0)$ e $W_1(q; 0)$ são aplicadas às imagens I_0 e I_1 , respectivamente. Agora, se $t = 1$, tem-se $W_0(p; 1) = W_0(p)$ e $W_1(q; 1) = q$, sabendo que $W_0(p) = q$. Esse processo deve ser realizado para cada ponto das imagens e t varia de 0 a 1 em um intervalo que depende do número de quadros da sequência.

Para exemplificar melhor o processo de transformação de tonalidade, observe-se a figura 2.7. As imagens (a) e (e) são as imagens iniciais e finais, respectivamente. As imagens (b), (c) e (d) são resultantes da transformação de tonalidades.

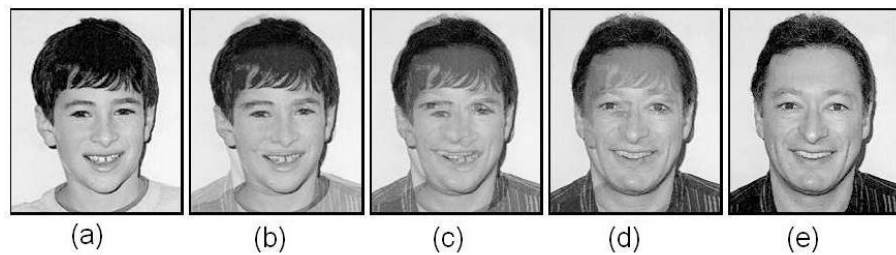


Fig. 2.7: Interpolação de cores (*Cross-Dissolve*).

Esse processo mostra que o *morphing* é um *cross-dissolve* aplicado a uma imagem deformada, ou seja, observando-se a figura 2.10, nota-se que esta figura possui imagens que representam o processo de *morphing* realizado para transformar uma letra 'F' em uma letra 'A'. Cada imagem desta figura foi criada pela transformação de tonalidades entre as imagens da figuras 2.8 e 2.9 na ordem em que aparecem nas figuras, utilizando as equações 2.15-2.19 para definir a influência que as imagens - figuras 2.8 e 2.9 - devem ter sobre a figura final (figura 2.10).



Fig. 2.8: Letra 'F' sofrendo deformações para tornar-se a letra 'A' (esquerda->direita).



Fig. 2.9: Letra 'A' sofrendo deformações para tornar-se a letra 'F' (direita->esquerda).

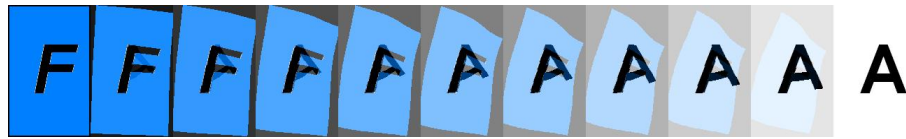


Fig. 2.10: Resultado final do morphing entre as letras 'F' e 'A'.

2.1.4 Controle de transição

A taxa de transição no *morphing* permite que partes da imagem definidas pelo usuário, sejam deformadas em diferentes tempos. A taxa de transição é o tempo t utilizado nas equações 2.15-2.19. Entretanto, para estas equações, a mesma taxa de transição é aplicada a todos os pontos para gerar a imagem intermediária $I(t)$. Conseqüentemente, as características das imagens I_0 e I_1 são refletidas com a mesma taxa sobre a imagem $I(t)$. Cada ponto pode ter uma taxa da transição diferente, produzindo imagens intermediárias mais interessantes.

Uma função de transição é construída para controlar o comportamento de transição na geração das imagens intermediárias. Portanto, uma função de transição especifica uma taxa da transição para cada ponto em uma imagem sobre o tempo. O tempo é que determina a taxa da transição entre as imagens I_0 e I_1 e a taxa de transição é um valor entre 0 e 1. Com uma taxa da transição perto de zero, as imagens intermediárias são mais similares a imagem inicial I_0 . As taxas da transição próximas do valor 1, implicam em imagens intermediárias parecidas com a imagem final I_1 .

Sendo T_0 a função de transição definida para a imagem inicial I_0 , para um determinado tempo t , $T_0(p; t)$ é uma função que determina o quão rápido cada ponto p em I_0 move-se em direção ao ponto correspondente q na imagem final I_1 . $T_0(p; t)$ também determina a contribuição de cor de cada ponto p em I_0 para o ponto correspondente em uma imagem intermediária $I(t)$.

Sendo T_1 a função de transição definida para a imagem final I_1 . Para cada ponto q em I_1 , $T_1(q; t)$ é definida para ter a mesma taxa de transição que $T_0(p; t)$ se q corresponde a p em I_0 . Então, $T_1(q; t)$ pode ser derivado de $T_0(p; t)$ usando a função de deformação W_1 . Isto é, $T_1(q; t) = T_0(W_1(q); t)$.

Portanto, as equações 2.15-2.19 podem ser reescritas para ter uma taxa de transição variável:

$$\begin{aligned}\bar{W}_0(p; t) &= (1 - T_0(p; t)) \cdot p + T_0(p; t) \cdot W_0(p) \\ \bar{W}_1(q; t) &= T_1(q; t) \cdot q + (1 - T_1(q; t)) \cdot W_1(q) \\ \bar{I}_0(p; t) &= \bar{W}_0(p; t) \bullet ((1 - T_0(p; t)) \cdot I_0(p)\end{aligned}$$

$$\begin{aligned}\bar{I}_1(q; t) &= \bar{W}_1(q; t) \bullet (T_1(q; t) \cdot I_1(q)) \\ I(t) &= \bar{I}_0(p; t) + \bar{I}_1(q; t)\end{aligned}$$

onde $W \bullet I$ é a aplicação da função de deformação W para a imagem I . E, $0 \leq T_0, T_1 \leq 1$, e $0 \leq t \leq 1$.

2.2 Aplicações de Morphing

O processo de *morphing* entre imagens além de ser usado para criar animações transformando uma imagem em outra, é usado em outras aplicações como as que serão citadas a seguir:

2.2.1 View Morphing

View morphing é uma técnica criada por Steven Seitz e Charles Dyer [Seitz e Dyer (1996)]. O objetivo do algoritmo de *View Morphing* é, a partir de duas fotos de pontos de vista diferentes de um objeto, obter uma terceira vista arbitrária. A diferença entre *View Morphing* e *Morphing* é que o primeiro preserva a forma tridimensional do objeto ou da cena nas imagens, já o segundo, só mantém a forma 3D se as vistas do objeto na imagem forem paralelas, isso significa que a imagem I_1 foi obtida afastando-se o objeto numa direção paralela ao plano da imagem I_0 , ou que o objeto permaneceu no mesmo lugar e a câmera foi movida numa direção paralela para a obtenção da imagem I_1 . Portanto, a idéia do *View Morphing* é aplicar transformações projetivas em imagens onde as vistas não são paralelas.

O *View Morphing* é composto por três fases: *pré-warping*, *morphing* e *pós-warping*. Durante o *pré-warping*, reprojetam-se as fotos para alterar a direção em que foram obtidas para torná-las paralelas, em seguida aplica-se o *morphing* que gera como saída outra imagem paralela e por último reprojeta-se novamente esse resultado para obter a foto na direção que foi especificada (*pós-warping*). Para aplicar transformações projetivas é necessário ter as matrizes de projeção das imagens, e existe uma variedade de técnicas para obter-se essas matrizes, como por exemplo, sabendo-se os parâmetros internos das câmeras ou conhecendo a posição 3D de alguns pontos das imagens. Os autores utilizaram um algoritmo chamado “algoritmo dos 8 pontos” proposto por Hartley [Hartley (1995)] na qual sabendo a posição 2D de oito pontos correspondentes (ou mais) em ambas imagens, é possível descobrir as matrizes de transformação.

Existem alguns casos onde esta técnica não funciona corretamente, casos onde uma superfície não pode ser visível em ambas imagens, com isso, “dobras” e “buracos” podem aparecer na nova vista.

Na figura 2.11 tem-se uma representação esquemática das transformações que ocorrem no *view morphing*. Nela I_0 e I_1 representam as imagens originais, $\hat{I}_0$ e $\hat{I}_1$ as imagens intermediárias após o *pré-warping*, $\hat{I}_S$ o resultado do *morphing* e, por último, I_S é a imagem final. As imagens $\hat{I}$ são vistas paralelas. Os centros óticos das câmeras estão representados por C_0 , C_1 e C_S .

Um exemplo de *View Morphing* pode ser visualizado na figura 2.12, que ilustra a imagem do quadro pintado por Leonard da Vinci conhecido por Monalisa. A imagem no centro ilustra o resultado de *View Morphing*.

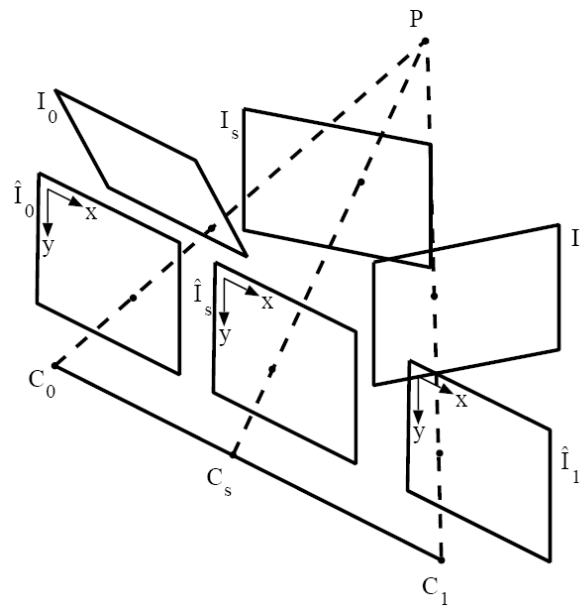


Fig. 2.11: Esquema de vistas.

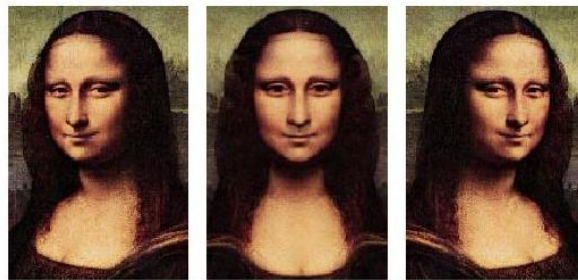


Fig. 2.12: Exemplo do View Morphing.

2.2.2 Polymorph

O processo de *morphing* tradicional de imagens considera somente duas imagens no tempo: a imagem inicial e a imagem final. *Morphing* entre múltiplas imagens é chamado de *polymorphing* e envolve uma série de transformações de uma imagem para outra. Por exemplo, utilizando *polymorphing* pode-se gerar uma imagem facial com características (olhos, nariz e boca) de várias faces de entrada. Rowland and Perrett [Rowland e Perrett (1995)] utilizaram *polymorph* para isso e obtiveram um protótipo de face derivado de diversas faces de adolescentes.

No trabalho de Lee, Wolberg e Shin [Lee, Wolberg e Shin (1998)] foi apresentada uma estrutura geral para *polymorphing* estendendo o conceito de *morphing* entre duas imagens. Cada imagem de entrada foi estipulada como um vértice de um poliedro de tamanho n (n é o número de imagens de entrada). Um número arbitrário de imagens intermediárias de transição pode ser especificado por um ponto no poliedro. As coordenadas baricêntricas desse ponto determinam os pesos usados na mistura das imagens da entrada para gerar imagens intermediárias. Ao considerar somente duas imagens, o poliedro degenera-se em uma linha e os pontos ao longo da linha correspondem às imagens

intermediárias em uma sequência do *morphing*. Este é o caso do *morphing* tradicional. Entretanto, quando se considera mais de duas imagens, as imagens intermediárias encontram-se em algum lugar dentro do poliedro.

No *morphing* entre duas imagens, pode-se criar animações com uma taxa de transição variável, que permite realizar deformações bem mais interessantes no tempo, como por exemplo, transformar uma imagem em outra começando pelo topo da imagem até o fundo desta. No caso de *polymorph*, isso também é possível de realizar. Por exemplo, considerando o *morphing* entre 4 faces, a imagem facial final pode ser gerada com os olhos da face 1, o nariz da face 2, a boca da face 3 e orelhas da face 4.

O *Polymorph* é ideal para aplicações de composição de imagens, ou seja, cria uma imagem composta pela metamorfose de regiões selecionadas em diversas imagens da entrada. As regiões misturam-se respeitando a geometria da deformação. A técnica produz imagens compostas de alta qualidade com menos esforço do que técnicas de composição convencionais de imagens. Por isso, *polymorph* permite imagens mais sofisticadas derivadas de uma mínima interação com o usuário. A técnica de *morphing* utilizada no *Polymorphing* é Multilevel Free-form Deformation, portanto *snakes* fazem as extrações das características nas imagens [Ruprecht e Müller (1993)].

A figura 2.13 mostra um exemplo de *polymorph*. A imagem à esquerda é resultado da mistura (*cross-dissolve*) das 3 imagens de entrada. Observa-se que esta imagem aparece tripla devido à mistura de características desalinhadas. As imagens à direita ilustram o processo para gerar a imagem final I . Essa imagem é resultado da mistura das faces utilizando a proporção de $1/3$ de cada imagem de entrada. Portanto, a imagem resultante I é uma mistura igual do formato do rosto, das posições e cores dos olhos, do nariz e da boca das faces de entrada.

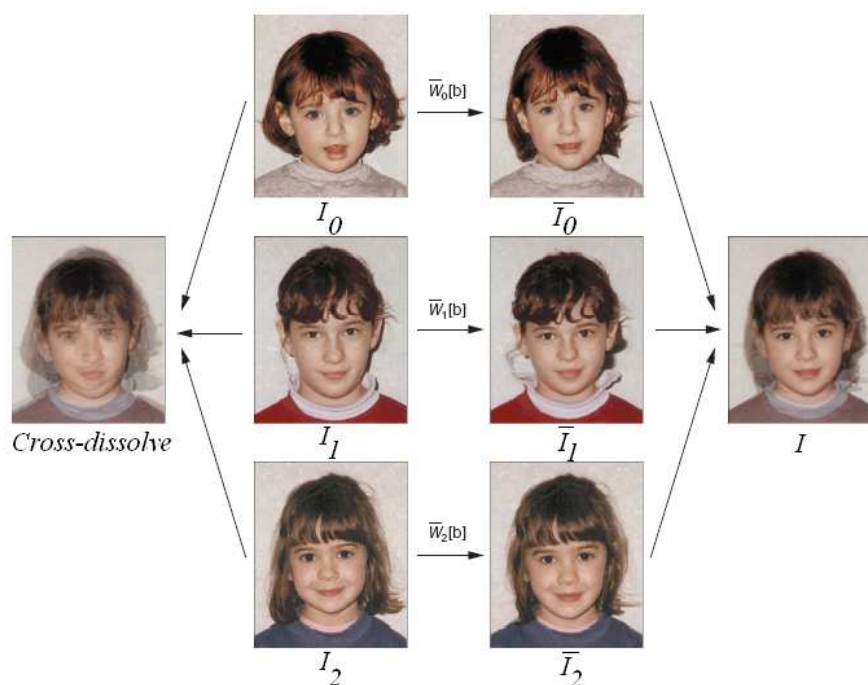


Fig. 2.13: Exemplo da técnica *Polymorph* entre 3 imagens.

2.3 Considerações Finais

Neste capítulo foi explicado como a técnica de *morphing* de imagens transforma uma imagem em outra. Foi visto que segundo [Wolberg (1998)], o *morphing* é dividido em quatro fases:

- Fase de especificação de elementos correspondentes nas imagens: é a fase onde são definidos pares de características em ambas imagens;
- Fase de deformação das imagens: é o processo de deformação nas imagens pelas funções de deformação;
- Fase de transformação de tonalidades: é o processo de interpolação das tonalidades das imagens deformadas;
- Fase de controle de transição: permite criar animações interessantes variando a taxa de transição ao longo da imagem.

Existem cinco métodos principais de deformação (*warping*) sendo estes:

- Mesh warping: utiliza uma grade para especificar os elementos nas imagens;
- Field morphing: utiliza vetores na fase de especificação de elementos, tendo um baixo custo computacional;
- Radial basis function e Thin Plate splines: primeiro a utilizar pontos na especificação de elementos da imagem. Difere por usar funções de base radial;
- Energy minimization: é baseado na minimização de funções de energia associadas a superfície que representa a imagem para obter a deformação desejada;
- Multilevel free-form deformation: contém uma nova manipulação direta da técnica *free-form deformation* (FFD) aplicando uma hierarquia de treliças de controle sobre a imagem para deformá-la.

Foi visto também que *morphing* pode ser utilizado entre diversas imagens, método chamado de *Polymorph*.

E, em *View Morphing*, Seitz [Seitz e Dyer (1996)] utilizou transformações projetivas no *morphing* para preservar a forma tridimensional dos objetos nas imagens, permitindo visualizar metamorfoses mais realistas de imagens.

Capítulo 3

Funções de Base Radial

Este capítulo apresenta uma descrição do método de *morphing* que utiliza funções de base radial para realizar a deformação nas imagens. O método de funções de base radial foi utilizado para resolver o problema de transformação e interpolação das imagens.

A utilização de funções de base radial para o *morphing* de imagens foi proposto pela primeira vez por Nur Arad em [Arad e Reisfeld (1995)]. Este método de *morphing* é conhecido na literatura como “Funções de Base Radial” exatamente porque utiliza estas funções para realizar a deformação de imagens. As funções de base radial já foram usadas em diversas outras aplicações, como por exemplo em: redes neurais, registro de imagens, etc.

Funções de base radial vem do termo em inglês: *Radial Basis Function*, portanto quando for usada a abreviação RBF, quer-se referir a tais funções.

Antes de explicar detalhadamente como funciona uma Função de Base Radial, será feita a introdução de alguns conceitos matemáticos necessários para o entendimento do funcionamento de uma RBF. Serão apresentados os seguintes conceitos: transformação espacial, interpolação e aproximação, modelos deformáveis. Após, será mostrada uma descrição e o funcionamento das RBFs. Este método será usado no presente trabalho para realizar o envelhecimento de imagens faciais humanas.

3.1 Transformação Espacial

Uma transformação espacial no caso de imagens, é uma transformação que faz um mapeamento de uma imagem em outra a partir de um conjunto de pontos da imagem definidos antecipadamente. Estes pontos são chamados de pontos de controle. Um ponto de controle descreve a posição de uma característica na imagem. Cada ponto na imagem fonte tem somente um ponto correspondente na imagem deformada, resultando numa sequência de pares de controle. Portanto, são informados pares de pontos de controle $\{\vec{p}_i, \vec{q}_i\}$ nas imagens inicial e final, onde $\vec{p}_i = (x_i, y_i)$ é a coordenada no estado inicial e $\vec{q}_i = (u_i, v_i)$ é a coordenada no estado final, ou seja, o lugar onde a coordenada inicial deverá estar após a transformação. Uma transformação espacial matematicamente descreve uma relação espacial entre estes pontos de controle correspondentes. Uma transformação espacial em 2-dimensões, denotada por $T(x_i, y_i)$, mapeia um conjunto de n pontos de controle $\vec{p}_i$, da imagem inicial em pontos correspondentes $\vec{q}_i$ da imagem deformada ($i = 1, \dots, n$), sendo essa transformação realizada em todo o domínio da imagem. Uma transformação espacial bidimensional é composta por

duas funções de mapeamento tal que

$$T(\vec{p}_i) = [T_U(\vec{p}_i), T_V(\vec{p}_i)] = q_i \quad (3.1)$$

onde $T_U(\vec{p}_i)$ representa a função que mapeia a primeira dimensão (x), $T_V(\vec{p}_i)$ mapeia a segunda dimensão (y):

$$\begin{aligned} T_U(\vec{p}_i) &= u_i \\ T_V(\vec{p}_i) &= v_i \end{aligned}$$

3.1.1 Interpolação e Aproximação

Dado n pares de pontos de controle em $2D$, $\vec{p}_i = (x_i, y_i)$ e $\vec{q}_i = (u_i, v_i)$, $i = 1, \dots, n$, representando os pontos na imagem inicial e na imagem deformada respectivamente, o problema de interpolação é encontrar duas funções de mapeamento $T_U(\vec{p}_i)$ e $T_V(\vec{p}_i)$ que satisfazem $T_U(\vec{p}_i) = u_i$ e $T_V(\vec{p}_i) = v_i$, para $i = 1, \dots, n$.

O problema de aproximação, por outro lado, requer que $T_U(\vec{p}_i) \approx u_i$ e $T_V(\vec{p}_i) \approx v_i$, ou seja, o resultado da transformação será um valor próximo de u_i e v_i .

Então, a interpolação de dados dispersos é um problema de aproximar uma superfície de forma suave através de uma distribuição de dados dispersa ou não uniforme. A vantagem de usar interpolação de dados dispersos para deformar é que são necessários menos pontos de controle do que a interpolação usando pontos no contorno da imagem.

3.1.2 Transformações Locais e Globais

Transformações espaciais podem ser classificadas pelo seu domínio. O domínio de uma transformação refere-se à região de influência das funções $T_U(\vec{p}_i)$ e $T_V(\vec{p}_i)$ que tanto pode ser local ou global. Uma transformação é considerada global se essas funções afetam todos os pontos da imagem, não importando a sua distância de $\vec{p}_i$ ou, se a remoção de um ponto de controle afeta toda a transformação. Se essas funções não dependem de pontos de controle que estão a mais do que uma certa distância de $\vec{p}_i$, então a transformação é considerada como local.

Transformações localizadas, geralmente produzem uma função com a primeira derivada contínua, denotada por C^1 , isso faz com que a superfície denotada por esta função tenha uma aparência suave [Wirth (2000)]. Transformações globais possuem funções que não têm a primeira derivada contínua (C^0). A notação C^i é usada para denotar uma função que é contínua nas primeiras i derivadas, isso determina o quão suave uma transformação será.

3.1.3 Modelos Rígidos e Deformáveis

Uma transformação espacial pode ser deformável ou rígida. É considerada uma transformação rígida se a distância entre quaisquer pontos consecutivos é preservada. Esse tipo de transformação pode ser decomposta em uma translação, escalonamento e/ou rotação, onde uma translação é um deslocamento constante no espaço. Portanto, uma transformação rígida é essencialmente um subconjunto de uma transformação afim (classe de transformações que incluem: rotação, translação, escala,

reflexão e *shearing*). Uma transformação afim pode ser decomposta em uma transformação linear e uma translação. Em $2D$ uma transformação afim pode ser expressa por um polinômio linear na forma:

$$P_k(\vec{p}) = \alpha_{1k} + \alpha_{2k}x + \alpha_{3k}y \quad k = U, V \quad (3.2)$$

Entretanto, a equação acima só é útil quando se tem 3 pontos no plano, no entanto quer-se uma transformação para mais de 3 pontos no plano.

Uma transformação deformável não mantém relações espaciais entre os pontos de controle, alterando a forma do modelo ou objeto em questão. Exemplos de transformações deformáveis são obtidos com o uso de funções de base radial; entretanto, essas funções podem ter um componente rígido incorporado.

3.2 Descrição das Funções de Base Radial

Imagens são objetos bi-dimensionais e por isso, uma imagem é um domínio finito de um plano com um nível de cinza ou uma tonalidade associada para cada ponto. Uma deformação de uma imagem é uma transformação da localidade dos pixels (mapeamento) no plano da imagem podendo ainda os valores de nível de cinza dos pixels ser alterados de acordo com uma transformação associada as coordenadas.

O objetivo das funções de base radial é realizar a deformação de imagens (planas) que é determinada pela transformação espacial de um número de pontos de controle (n). Estes pontos pré-determinam o mapeamento entre os dois conjuntos de pontos, um associado a imagem inicial e outro associado a imagem final, cada um com n pontos [Arad e Reisfeld (1995)]. O mapeamento deve ser realizado usando interpolação, quando se deseja um mapeamento exato para os pontos definidos na imagem final, ou usando aproximação quando estes pontos não necessitam estar exatamente no local definido pelos pontos finais. Portanto, uma RBF mapeia os pontos de controle de uma imagem nos pontos de controle correspondentes na outra imagem e faz uma interpolação desse mapeamento em pontos intermediários. Uma transformação feita por RBF em 2 dimensões, denotada por $T(\vec{p}_i)$, é composta por duas funções de mapeamentos conforme a equação 3.1.

O pontos de controle definidos nas imagens não são restritos quanto à sua localidade, podendo estar dispersos ou muito próximos uns dos outros. Cada uma das funções de mapeamento na equação 3.1 pode ser decomposta em uma componente global e uma local. Embora as duas componentes sejam distintas, elas podem ser calculadas simultaneamente, resultando em uma única transformação. Portanto, o que se deseja para o mapeamento é uma maneira de interpolar e/ou aproximar dados dispersos. Isso pode ser realizado, definindo uma função de interpolação como uma combinação de funções de base radialmente simétricas, cada uma, centrada num ponto de controle.

O nome função radial sugere que os valores da função são obtidos pela diferença (ou distância) de coordenadas do ponto onde a função deve ser avaliada para as coordenadas dos pontos de controle. Portanto, uma função radial é representada pela função:

$$T(\vec{p}) = \sum_{i=1}^n a_i g(r), \quad (3.3)$$

onde r denota a norma euclidiana em R^d entre dois pontos, sendo para $d = 2$:

$$r = [(x_1 - x_2)^2 + (y_1 - y_2)^2]^{\frac{1}{2}} \quad (3.4)$$

Em outras palavras, isso quer dizer que o valor da função g em um ponto $\vec{p} = (x, y)$ somente depende da norma de $\vec{p}$. É por isso, que uma função radial não tem orientação preferida e, o valor da função é igual para todos os pontos que possuem a mesma distância do ponto de controle que está sendo analisado ($\vec{p}_i$). A equação 3.3 é conhecida como uma pura soma radial que representa o somatório de pesos elásticos ou funções de base não lineares $g(r_i)$, onde r_i representa a norma Euclidiana entre dois pontos. Entretanto, usar funções puramente radiais resulta em uma péssima aproximação para pontos que estão distantes dos pontos de controle $\vec{p}_i$. Para resolver este problema de aproximação de pontos distantes, acrescenta-se a equação 3.2 na equação 3.3:

$$T(\vec{p}) = P_K(\vec{p}) + \sum_{i=1}^n a_i g(r_i) \quad (3.5)$$

A componente $P_K(\vec{p})$ é um polinômio linear, que tem uma influência global no mapeamento porque ela determina uma transformação afim, ou seja, esta componente é capaz de transladar, rotacionar e escalar a imagem. Isso resolve o problema da péssima aproximação para pontos distantes dos pontos de controle. A forma geral de um polinômio linear pode ser dada pela equação 3.2.

Para a equação 3.5 ser resolvida, é preciso conhecer suas variáveis: a função $g(r)$ é conhecida e os tipos de funções estão descritos na seção 3.2.3, mas é preciso descobrir os valores de a_i que são chamados de pesos da função. Esses pesos são encontrados resolvendo um sistema de equações lineares conforme será descrito na seção 3.2.2.

3.2.1 Propriedades das Funções de Base Radial

Como descrito, o objetivo do uso das RBFs é realizar um mapeamento de uma imagem bidimensional em outra imagem bidimensional, ou seja, deseja-se mapeamentos em $\mathbf{R}^2 \rightarrow \mathbf{R}^2$, só que a interpolação de dados dispersos lida com funções do tipo $\mathbf{R}^d \rightarrow \mathbf{R}^1$. Para solucionar este problema, é necessário usar pares de funções $\mathbf{R}^2 \rightarrow \mathbf{R}^1$ para realizar o mapeamento desejado. Portanto, dado dois conjuntos de pontos $\vec{p} = (x_i, y_i)$ e $\vec{q} = (u_i, v_i)$, $i = 1, 2, \dots, n$ que são os pontos de controle, procura-se por uma transformação $T(\vec{p}) = (T_U(\vec{p}), T_V(\vec{p})) : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ com as seguintes propriedades:

- (a) $T(\vec{p})$ possui uma função radial como uma de suas componentes. Isso faz com que o resultado de T para cada ponto de controle $\vec{p}$ seja o mesmo para todos pontos equidistantes de $\vec{p}$;
- (b) $T(x_i, y_i) = (u_i, v_i)$ para todo $i = 1, 2, \dots, n$;
- (c) o número de pontos a ser interpolado é qualquer número maior que 3: se o número de pontos de controle for igual a 3, o mapeamento é afim;
- (d) as componentes de T reproduzem polinômios lineares em $\mathbf{R}^2$. Essa condição garante que T será uma transformação afim sempre que os dados de interpolação admitem tal transformação;
- (e) deve existir uma relação de equilíbrio entre a condição de deformação (a deformação do plano deve ser a menor possível) e a condição de localidade (a interpolação de cada ponto de controle terá um efeito local);

(f) em alguns casos, o mapeamento dos pontos de controle não precisa ser exato. Nesses casos deseja-se encontrar uma relação de equilíbrio entre o erro de interpolação (de cada ponto de controle) e uma condição mínima de deformação indicada na condição anterior (e).

As condições (a),(b),(c) e (d) são satisfeitas pela transformação $T = (T_U, T_V)$, onde T_U e T_V são funções radiais como descritas na equação 3.5, com $K = 1$. Entretanto, a condição (e), precisa de uma formulação matemática que será realizada a seguir. A condição de localidade não será comentada por enquanto, mas a condição de deformação será descrita a seguir.

Dada uma função $T : \mathbf{R}^2 \rightarrow \mathbf{R}^1$ é comum usar a função

$$J(f) = \int \int_{\mathbf{R}^2} [(\frac{\partial^2 f}{\partial x^2})^2 + 2(\frac{\partial^2 f}{\partial x \partial y})^2 + (\frac{\partial^2 f}{\partial y^2})^2] d(x, y)$$

como uma medida de aproximação da energia de curvatura, ou seja da deformação ocorrida, total da superfície $(x, y, f(x, y))$ [Duchon (1979)]. O funcional J fornece uma medida para a suavidade da deformação realizada na imagem, quanto menor for o valor de J , mais suave a deformação será.

Embora o mapeamento seja definido para cada coordenada separadamente ($\mathbf{R}^2 \rightarrow \mathbf{R}^1$), procura-se por uma transformação $T = (T_U(x, y), T_V(x, y))$ ($\mathbf{R}^2 \rightarrow \mathbf{R}^2$) tal que $J(T_U) + J(T_V)$ seja mínimo, onde

$$\begin{aligned} T_U \in f | f(x_i, y_i) &= u_i \quad , \quad i = 1, 2, \dots, n \\ T_V \in f | f(x_i, y_i) &= v_i \quad , \quad i = 1, 2, \dots, n \end{aligned}$$

Essa é uma aproximação para o problema de minimização da condição de deformação induzida pelo mapeamento. Entretanto, como a transformação é realizada por partes, a minimização da expressão $J(T_U) + J(T_V)$ pode ser conseguida por partes também, ou seja, pela minimização em separado de $J(T_U)$ e $J(T_V)$.

Retornando às propriedades de transformação vistas anteriormente, itens (a)-(f), mais precisamente na propriedade (f), viu-se que em alguns casos o mapeamento não necessita ser uma interpolação exata de $T(x_i, y_i) = (u_i, v_i)$, $i = 1, 2, \dots, n$, ou seja, pode ser $T(x_i, y_i) \approx (u_i, v_i)$. Para isso, encontra-se duas funções diferenciáveis T_U e T_V tal que para um dado $\lambda > 0$ as funções

$$\tilde{J}(T_U) = \sum_{i=1}^n [u_i - T_U(x_i, y_i)]^2 + \lambda J(T_U) \quad (3.6)$$

e

$$\tilde{J}(T_V) = \sum_{i=1}^n [v_i - T_V(x_i, y_i)]^2 + \lambda J(T_V) \quad (3.7)$$

possuem um valor mínimo [Arad et al. (1994)].

Existem alguns casos especiais das equações 3.6 e 3.7 que são os seguintes:

- $n = 3$. A transformação do plano se reduz a uma transformação afim ($J = 0$). Isso foi visto no item (c).

- $\lambda = 0$. Quanto menor for $\tilde{J}$, mais exata será a aproximação, tornando esta uma interpolação. Isso significa que se deseja um mapeamento exato dos pontos de controle.
- $\lambda \rightarrow \infty$. A minimização de $\tilde{J}$ é dada por um plano, que resulta em $J(T_U) = 0$; ou seja, o termo $\lambda J(T_U)$ tende a zero. Então o primeiro termo da equação que representa o erro volta a ser o fator preponderante na minimização de $J(T_U)$ e determina o plano interpolante.
- $0 < \lambda < \infty$. Deve-se encontrar um valor para o parâmetro λ entre 0 e ∞ para que a transformação ocorra como desejado. Para isso, é importante saber que λ controla a relação entre a transformação afim (quando $\lambda \rightarrow \infty$, isso faz com que fator de curvatura J seja anulado) e o fator radial que permite uma condição mínima de deformação e interpola os dados, como mostrado na propriedade (f). Nessa formulação, o mesmo valor de λ é escolhido para todos os pontos.

Retornando à propriedade (e) da condição de localidade, o funcional J é de natureza global, então uma pequena perturbação de um ponto de controle pode afetar todos pontos na transformação do plano. Em alguns casos, é preferível que todos os pontos de controle tenham um efeito local. Para isso é preciso escolher a função base mais apropriada para cada caso.

3.2.2 Resolução do Sistema Linear

Segundo [Arad et al. (1994)], a solução que minimiza o funcional J e assegura uma superfície de interpolação suave é dada pela utilização das seguintes equações:

$$T_U(x, y) = \alpha_1 + \alpha_2 x + \alpha_3 y + \sum_{i=1}^n a_i g(r_i) \quad (3.8)$$

$$T_V(x, y) = \beta_1 + \beta_2 x + \beta_3 y + \sum_{i=1}^n b_i g(r_i) \quad (3.9)$$

os coeficientes $\alpha_1, \alpha_2, \alpha_3, a_1, a_2, \dots, a_n$ e $\beta_1, \beta_2, \beta_3, b_1, b_2, \dots, b_n$ são soluções procuradas para realizar a minimização do funcional J .

A partir de agora, será mostrada a solução somente para $T_U(x, y)$, entretanto pode-se aplicar a mesma solução para $T_V(x, y)$. Então, para encontrar os coeficientes que permitem a solução do sistema, é necessário resolver o seguinte sistema linear de tamanho $n + 3$:

$$\begin{cases} T_U(x, y) = \alpha_1 + \alpha_2 x + \alpha_3 y + \sum_{i=1}^n a_i g(r_i) \\ \sum_{i=1}^n a_i = 0 \\ \sum_{i=1}^n a_i x = 0 \\ \sum_{i=1}^n a_i y = 0 \end{cases} \quad (3.10)$$

Neste sistema, a primeira equação assegura uma superfície suave e as outras três equações do sistema asseguram que a função de base radial cresce quase linearmente quanto mais distante dos pontos de controle for a avaliação. O sistema pode ser escrito na seguinte forma matricial:

$$\begin{bmatrix} (G + \lambda * I) & P \\ P^T & Z \end{bmatrix} \begin{bmatrix} a \\ \alpha \end{bmatrix} = \begin{bmatrix} F \\ Z \end{bmatrix} \quad (3.11)$$

com Z sendo uma matriz de zeros, I sendo a matriz identidade, e

$$G = g(\|\vec{p}_i - \vec{p}_j\|) \text{ onde } i, j = 1, 2, \dots, n \quad (3.12)$$

portanto,

$$G = \begin{bmatrix} g(r_{11}) + \lambda & g(r_{12}) & \cdots & g(r_{1n}) \\ g(r_{21}) & g(r_{22}) + \lambda & \cdots & g(r_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ g(r_{n1}) & g(r_{n2}) & \cdots & g(r_{nn}) + \lambda \end{bmatrix} \quad (3.13)$$

sendo formada pelas funções bases calculadas para todos os pontos de controle. Nesta matriz, todos os elementos da diagonal principal possuem o valor 1 associado devido ao fato de a distância do ponto de controle em relação a ele mesmo ser zero e utilizando a função gaussiana, temos $e^0 = 1$. Usa-se o parâmetro λ multiplicado pela matriz identidade I acrescentando a G para controlar o mapeamento realizado pela função (equações 3.6 e 3.7).

A função $g(r)$ é chamada de função de base e significa que dado um ponto de controle, os cálculos são feitos com base na distância (raio) desse ponto até os demais da sua vizinhança. Essa função, $g(r)$, é aplicada a cada pixel na imagem para realizar a deformação, e ela pode ser alterada, conforme o comportamento que desejarmos atribuir à deformação. Na próxima seção serão apresentadas as principais funções de base utilizadas.

A matriz que contém os pontos de controle $\vec{p}_i = (x_i, y_i)$ da imagem inicial é:

$$P = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & y_n \end{bmatrix} \quad (3.14)$$

e da imagem final $\vec{q}_i = (u_i, v_i)$:

$$F = \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \\ \vdots & \vdots \\ u_n & v_n \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (3.15)$$

Portanto, o sistema ficará da seguinte forma:

$$\begin{bmatrix} g(r_{11}) + \lambda & g(r_{12}) & \cdots & g(r_{1n}) & 1 & x_1 & y_1 \\ g(r_{21}) & g(r_{22}) + \lambda & \cdots & g(r_{2n}) & 1 & x_2 & y_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ g(r_{n1}) & g(r_{n2}) & \cdots & g(r_{nn}) + \lambda & 1 & x_n & y_n \\ 1 & 1 & \cdots & 1 & 0 & 0 & 0 \\ x_1 & x_2 & \cdots & x_n & 0 & 0 & 0 \\ y_1 & y_2 & \cdots & y_n & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \vdots & \vdots \\ a_n & b_n \\ \alpha_1 & \beta_1 \\ \alpha_2 & \beta_2 \\ \alpha_3 & \beta_3 \end{bmatrix} = \begin{bmatrix} u_1 & v_1 \\ u_2 & v_2 \\ \vdots & \vdots \\ u_n & v_n \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (3.16)$$

Resolvendo o sistema linear, serão obtidos os seguintes coeficientes como solução do sistema:

$$a = (a_1, a_2, \dots, a_n)^T \text{ e } \alpha = (\alpha_1, \alpha_2, \alpha_3)^T \quad (3.17)$$

e

$$b = (b_1, b_2, \dots, b_n)^T \text{ e } \beta = (\beta_1, \beta_2, \beta_3)^T \quad (3.18)$$

Esses coeficientes serão utilizados para calcular o mapeamento de todos os pixels restantes da imagem utilizando a função base escolhida. Uma solução similar existe para T_V .

3.2.3 Funções de base

Como já dito anteriormente, as funções de base radial utilizam uma função, chamada de base a qual recebe como entrada a distância r entre os pontos de controle. Algumas dessas funções crescem à medida em que r cresce, como por exemplo as funções *Thin Plate Spline*, *Multiquadrática* e *Linear*; outras não, como a função *Gaussiana*. A escolha da função mais apropriada é realizada por fatores como a dimensão do problema, as condições de interpolação e as propriedades desejadas para a própria interpolação. O intervalo de influência da função base pode ser controlado ajustando seus parâmetros. As funções aqui descritas podem ser aplicadas a quaisquer dimensões; mas vale salientar que algumas funções são específicas para determinadas dimensões, como algumas para 2D, por exemplo. A figura 3.1 ilustra o gráfico das principais funções bases que serão mencionadas logo a seguir.

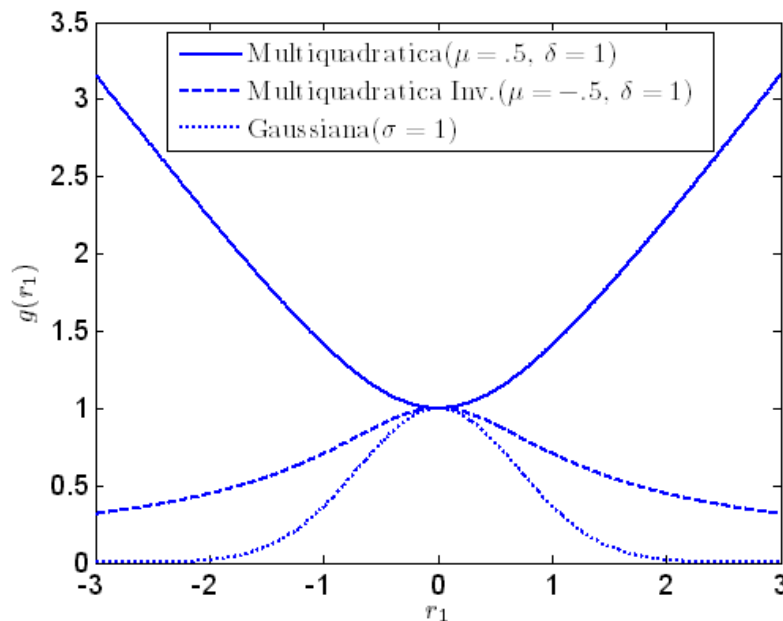


Fig. 3.1: Gráfico de algumas funções base.

Thin Plate Splines (TPS)

A primeira função a ser apresentada é a *Thin Plate Splines* proposta inicialmente por Jean Duchon em 1979 [Duchon (1979)]. Esta função é definida por:

$$g(r) = r^2 \log r^2 \quad (3.19)$$

onde $g(r) = 0$ na origem devido ao fato que a função $\log$ não é definida na origem.

O uso dessa função diminui a localidade da transformação, ou seja, a *Thin Plate Spline* é uma transformação de natureza global. É como se esta função modelasse uma superfície infinita de metal delgado, na qual pontos de controle definidos pelo usuário definissem a deformação. A sua formulação assegura um baixo fator de curvatura, portanto apresenta uma energia mínima de deformação.

Multiquádrica (MQ)

Essa função tem a seguinte forma geral:

$$g(r) = (r_i^2 + \delta)^{+\mu} \quad (3.20)$$

com $\delta > 0$ e $\mu > 0$. Esta função incorpora um parâmetro de localidade, δ , o qual é utilizado para controlar seu comportamento, dando menos peso para a distância entre os pontos de controle e mais para os da vizinhança próxima.

Multiquádrica Inversa

Assume a mesma forma da MQ, porém a condição é que $\mu < 0$. Seu comportamento torna-se similar ao da função Gaussiana, com o parâmetro σ se aproximando de zero para a resposta se tornar mais localizada, e aumentando de valor para aumentar a região de influência dos pontos de controle.

Gaussiana

Essa função é dada por

$$g(r) = e^{(-r_i^2/\sigma)} \quad (3.21)$$

A medida que o valor de σ cresce, a contribuição de pontos mais distantes aumenta, produzindo uma aproximação suavizada e mais global, enquanto a energia total de curvatura é decrementada. Se σ diminui, as contribuições da vizinhança próxima aumentam, produzindo aproximações mais localizadas e, portanto, mais bruscas. A figura 3.2 mostra essa variação no parâmetro σ .

Portanto, as funções gaussianas são de natureza mais local, permitindo que se tenha um controle localizado sobre os pontos de controle apenas alterando o valor do parâmetro σ .

3.3 Considerações Finais

Este capítulo ofereceu uma introdução sobre transformações espaciais que descrevem uma relação espacial entre pontos, ou seja, utilizando n pontos de controle nas imagens iniciais e finais, um

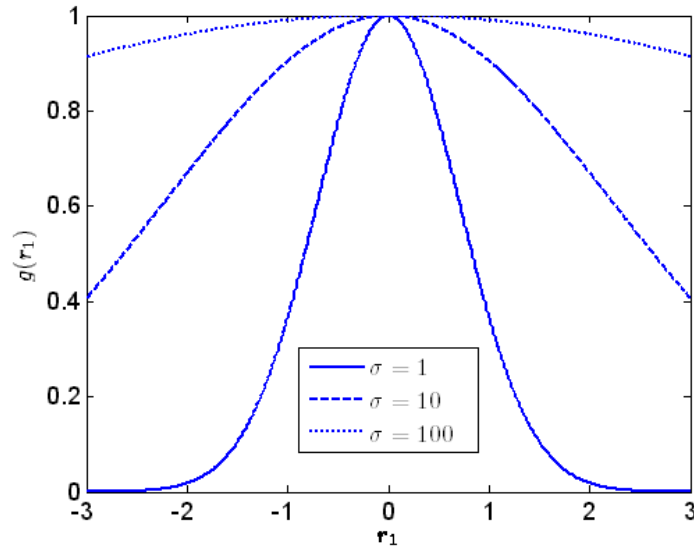


Fig. 3.2: Comparação entre valores do parâmetro σ da função Gaussiana.

mapeamento entre eles é realizado sendo que todos os pixels da imagem inicial são mapeados para a imagem final. O domínio de uma transformação pode ser local, afetando somente a região na imagem que contém o ponto de controle, ou global, a imagem inteira é afetada pelos pontos de controle.

Também foram explicados os conceitos de interpolação e aproximação e, que uma função de base radial realiza transformações entre pontos de controle dispersos utilizando funções de bases radialmente simétricas centradas nos pontos de controle. Algumas funções de base são mostradas na tabela 3.1 a seguir.

<i>Função</i>	$g(r_i)$	<i>Parâmetros</i>	<i>Dimensão</i>
Linear	r_i	-	$d < \infty$
Thin-Plate Spline (TPS)	$r_i^2 \log r_i^2$	-	2
Cubic Spline	$\ r_i\ ^3$	-	$d < \infty$
Multiquádrica (MQ)	$(r_i^2 + \delta)^{+\mu}$	$\delta > 0, 0 < \mu$	$d < \infty$
Gaussiana	$e^{(-r_i^2/\sigma)}$	$\sigma > 0$	$d < \infty$
Shifted-LOG	$\log(r_i^2 + \delta)^{\frac{3}{2}}$	$\delta \geq 1$	$d < \infty$

Tab. 3.1: Comparação entre os diferentes tipos de base.

Capítulo 4

Envelhecimento Facial

Ao longo do processo de envelhecimento humano, pode-se perceber algumas características, que são comuns à maioria das pessoas, como a perda da elasticidade da pele, rugas na face, afinamento dos lábios, cabelos brancos (figura 4.1), perda de cabelos (principalmente nos homens, processo chamado calvície). Entretanto, cada pessoa envelhece de uma maneira diferente e isso faz com que seja difícil obter uma simulação de envelhecimento realista. Muitos fatores como etnia, genética, estilo de vida, alimentação, peso, *stress*, entre outros, influenciam no envelhecimento de cada pessoa [(PITANGUY et al., 1977)]. Por exemplo, é comprovado que pessoas que fumam e são dependentes da nicotina, envelhecem de uma maneira mais rápida do que pessoas que não fumam [Leta (1998)]. Podem ser citados aqui, outros exemplos de pessoas que podem ter seu processo de envelhecimento mais acelerado como pessoas que se expõem muito ao sol, pessoas que vivem estressadas e são infelizes. Todos estes fatores não necessariamente aceleram o envelhecimento, mas pessoas que possuem um estilo de vida saudável e pessoas que praticam exercícios regularmente, podem ter um processo de envelhecimento mais lento.



Fig. 4.1: Envelhecimento.

4.1 Características do envelhecimento facial

Segundo Shirakabe (1988) existem diferenças entre o envelhecimento facial de mulheres orientais e caucasianas devido as variações morfológicas e estruturais entre as duas etnias. Após vários estudos, este autor comprovou que a pele das mulheres orientais é mais grossa e resistente e a face é em geral mais larga e plana. Esta comprovação permite que o envelhecimento das mulheres orientais seja diferenciado do envelhecimento de mulheres caucasianas.

Outro fato que faz com que seja perceptível o efeito da idade nas pessoas é a composição da pele. É comprovado que a composição da pele facial muda com a idade. Este fato pode ser constatado comparando-se as proporções de colágeno e elastina na pele. A pele de um recém-nascido possui uma maior quantidade de colágeno do que elastina, porém esta relação se inverte em pessoas idosas. As rugas se originam devido a uma deficiência de colágeno. As fibras que contribuem para a elasticidade e firmeza da pele tornam-se mais finas e se fragmentam. Esse fenômeno se acelera após exposições prolongadas e repetidas ao sol. Os primeiros sinais de envelhecimento aparecem aproximadamente aos trinta anos, mas geralmente, é um pouco antes dos 40 anos que o processo toma força. As primeiras rugas aparecem freqüentemente acima das bochechas e nas maçãs do rosto.

A aparência harmoniosa da pele é determinada por uma relação balanceada entre todas as estruturas que compõem a face. Com o envelhecimento esta relação se perde devido a alterações dos ossos, músculos e pele, que geram mudanças progressivas no formato facial. Alguns pontos que caracterizam o processo de envelhecimento são: o esqueleto sofre uma redução de volume ocasionado especialmente pelo processo de osteoporose, ocorre o afinamento da camada subcutânea, algumas regiões apresentam um aumento de gordura devido ao processo de degeneração da pele e a massa muscular diminui, perdendo conseqüentemente sua flexibilidade [Psillakis, Rumeley e Camargos (1988)]. O estudo da estrutura muscular facial desperta interesse especialmente para entender e simular animações de expressões faciais e de fala, podendo-se citar os trabalhos de [Ezzat e Poggio (1998)], [Faruque et al. (2001)] e [Bregler, Covell e Slaney (1997)].

4.2 Simulação Computacional do envelhecimento

Simulação Computacional do envelhecimento ou *Facial Age Simulation* é o processo que simula efeitos de envelhecimento em uma imagem facial para refletir a aparência de um indivíduo no futuro, ou no passado. Segundo Gandhi (2004), os métodos utilizados para realizar esse processo podem ser classificados em dois grupos. O primeiro grupo é baseado em métodos de aprendizado, ou seja, características de textura da pele são extraídas de uma ou várias imagens que serão usadas para modelar o sistema de aprendizado das mudanças na face. O segundo grupo é baseado em um modelo utilizando alterações da estrutura e formato do rosto para simular o processo de envelhecimento nas faces.

A partir de agora, será feita uma revisão bibliográfica sobre este assunto mostrando o que já foi feito em envelhecimento facial computacional.

A revisão dos principais trabalhos envolvendo simulação de envelhecimento foi dividida em três abordagens, sendo duas que simulam o envelhecimento e a terceira refere-se a estimativa da idade:

- abordagens que permitem estimar a idade de pessoas a partir de imagens faciais;
- abordagens de envelhecimento por aprendizado de alteração de textura;

- abordagens de envelhecimento por medidas estatísticas ocorridas com o passar do tempo, alterando a estrutura da face.

4.2.1 Métodos para estimar a idade facial

Estimar a idade da face de uma pessoa utilizando imagens envolve um treinamento a partir de um conjunto de imagens de faces em diferentes idades. Para isto, cria-se um modelo que permite estimar a idade em uma imagem que não faça parte das imagens de treinamento que, geralmente fazem parte de um banco de dados. Todavia, para se obter bons resultados, é preciso ter um banco de dados com muitas imagens de rostos em diversas idades.

O primeiro trabalho a ser citado, é o estudo realizado por Hwon e Lobo [Kwon e Lobo (1999)]. Estes autores propuseram um método para classificar imagens de faces em três diferentes grupos: crianças, adolescentes e adultos. O estudo foi baseado na relação geométrica entre regiões do rosto e na análise de rugas na pele. Primeiramente, foram analisadas na geometria da face, características que permitissem diferenciar crianças de adultos. Entre essas características, foram analisadas: a distância entre os olhos; e a distância entre os olhos e o nariz. Para classificar adolescentes foi considerada a ausência de rugas na pele. Por último, uma regra de fusão foi usada para combinar a relação geométrica e a quantidade de rugas para classificar uma face em uma categoria de idade. Porém, segundo os autores, esse método foi testado em um banco de dados com somente 47 imagens de alta resolução contendo crianças, adolescentes e adultos. Apesar de ter sido obtido 100% de sucesso na classificação destes dados, muitas dúvidas são pertinentes nessa pesquisa, entre elas destacam-se: o fato de poucas imagens terem sido testadas, tornando difícil a avaliação; a classificação não foi realizada por idade, ou seja, a classificação foi realizada em grupos (crianças, adolescente e adultos); e, o método requer imagens de no mínimo 256x256 *pixels* para analisar as rugas na face.

Baseado no trabalho de Kwon e Lobo (1999), Horng [Horng, Lee e Chen (2001)] propôs um método mais rápido e um sistema potencialmente mais robusto para a classificação de idade. Esse método é composto pelo mesmo fundamento, mas com diferentes técnicas para analisar a relação geométrica do rosto e a quantidade de rugas. A classificação foi feita usando redes neurais, conseguindo 84.6% de sucesso para 230 imagens experimentais. Entretanto, como o outro método, apenas três grupos de classificação foram utilizados (crianças, adolescentes e adultos).

[Hayashi et al. (2002)] focaram seus estudos na análise de rugas faciais para estimar a idade e o sexo. Eles usaram um banco de dados com 300 indivíduos entre 15 e 64 anos de idade. Primeiro, foram extraídas das imagens as regiões que contém pele na face, após foi realizada uma equalização de histograma para realçar as rugas. Então, tanto rugas pequenas como rugas grandes das faces foram extraídas utilizando a transformada de Hough (DTHT - *Digital Template Hough Transform*). Uma tabela de referência foi usada para estimar a idade, na qual, a classificação era realizada em grupos de 10 em 10 anos de idade. Entretanto, esse experimento não teve muito sucesso na classificação, conseguindo apenas 27% de sucesso na estimativa da idade, e 83% na classificação do sexo. É importante salientar que os autores não mencionaram o tamanho ou a fonte de seu banco de dados para o qual obtiveram esses valores, mas sabe-se que na fase de treinamento foram utilizadas somente faces japonesas. Os autores relataram também a dificuldade em analisar faces entre 20 e 30 anos devido ao fato de que as pessoas, geralmente nessa faixa de idade, ainda não possuem rugas no rosto.

Outro trabalho a ser citado é de Lanitis [Lanitis (2002)] que separou a face em partes distintas e

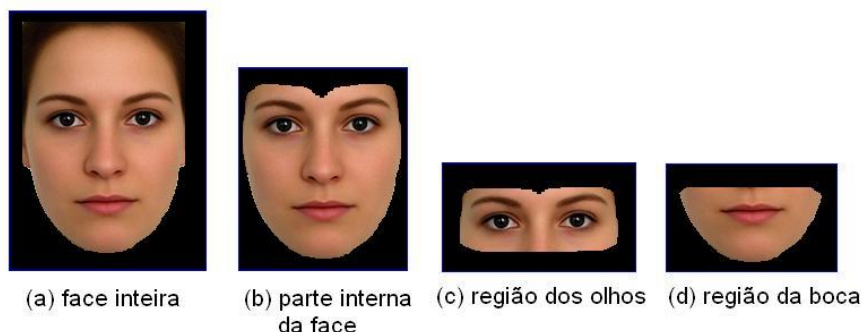


Fig. 4.2: Regiões da face usadas na estimação da idade.

estudou a significância e a relevância de cada parte para estimar a idade. Seu trabalho envolveu a avaliação das seguintes regiões da face: a face inteira (incluindo o cabelo), a parte interna da face (sem cabelo) e, a parte superior da face (região dos olhos) e inferior da face (região da boca). Um exemplo destas regiões pode ser visto na figura 4.2. Os resultados mostraram que a área em torno dos olhos provê um melhor resultado para estimar a idade em faces. O autor afirma que o uso da face inteira (com cabelo) tem um efeito negativo nos resultados. Seu estudo foi limitado a pessoas entre 0 e 35 anos, e em um banco de dados contendo 330 imagens onde somente 80 foram usadas para esse propósito. Devido a esta limitação de idade das pessoas (0-35 anos), observou-se que faces com um número maior de rugas não foram usadas, deixando uma dúvida na habilidade de estimar a idade em pessoas acima de 35 anos.

Vê-se então, das pesquisas citadas, que os trabalhos nessa área são limitados ao tamanho e qualidade do banco de dados. Alguns pesquisadores focam seus trabalhos em um grupo particular, enquanto outros tentam usar uma larga escala de classificação de idade. Porém, não se tem ainda um sistema que tenha uma larga escala de classificação para qualquer idade devido a dificuldade de encontrar grandes bancos de imagens, e estas com uma ótima qualidade de resolução para tal função. E, a partir do último trabalho citado aqui, pode-se constatar que a região em torno dos olhos é a melhor para o propósito de estimar a idade (exceto em crianças) ao contrário de faces com cabelo que impedem uma boa estimativa da idade.

4.2.2 Métodos de textura

Nos métodos de envelhecimento facial por textura, as características (detalhes de textura) do envelhecimento são extraídas de um conjunto de imagens de treinamento que contém as imagens do indivíduo jovem e velho. O envelhecimento de uma nova imagem facial (que não pertence ao conjunto de treinamento) é então realizada aplicando-se estas características que foram aprendidas.

Lanitis, Taylor e Cootes [Lanitis, Taylor e Cootes (2002)] descrevem um trabalho que estima a idade da face em uma imagem e depois envelhece esta imagem para a idade desejada pelo usuário. Eles comentam como os efeitos do envelhecimento na imagem da face podem ser realizados utilizando transformações obtidas pelo aprendizado de textura. Seus estudos apresentam resultados que mostram uma razoável precisão na estimativa de idade para imagens de faces que não pertencem ao banco de imagens e, que portanto, não foram usadas na fase de treinamento. Já para simular os

efeitos de envelhecimento, as imagens das faces são codificadas para depois serem reconstruídas, e imagens de treinamento foram usadas para o sistema aprender uma relação entre a representação da face codificada e a atual idade dos sujeitos. Uma vez que esta relação é estabelecida, foi possível estimar a idade de uma pessoa que não pertence as imagens de treinamento. Essa representação codificada é um modelo facial que tem uma forma tridimensional porém possui uma textura bidimensional. Para a extração da textura foi utilizada uma técnica conhecida como análise de componentes principais (*Principal Component Analysis* - PCA). Esta técnica é usada para extrair características em imagens e é muito utilizada na área de reconhecimento de padrões e reconhecimento de faces [Turk M.A.; Pentland (1991)].

Shan, Liu e Zhang [Shan, Liu e Zhang (2001)] apresentaram um método chamado *Image-Based Surface Detail Transfer (IBSDT)*, onde detalhes geométricos de uma superfície são transferidos para outra superfície. Não é necessário o conhecimento da geometria em 3D para este método, características de textura podem ser extraídas da imagem de um objeto de forma independente das propriedades de reflectância da superfície e esses detalhes podem ser transferidos para modificar a aparência de outros objetos em outras imagens diferentes. Existe um filtro gaussiano que controla a escala de detalhes que será transferida da superfície. Os autores propuseram uma aplicação desse algoritmo para simular o envelhecimento facial. Geometricamente, eles supuseram que a diferença entre a superfície da pele de uma pessoa nova para uma pessoa mais velha é que a pele de uma pessoa mais velha tem mais rugas na face. Então, se as rugas de uma pessoa são transferidas para uma pessoa mais nova, está se parecerá mais velha. Neste trabalho, o alinhamento das imagens é feito por componentes da face como os olhos, o nariz e a boca. Uma função de *warping* é usada para deformar a imagem fonte para a imagem destino. O método proposto, IBSDT, é então aplicado nos pixels da face, ou seja, onde existe pele, enquanto os pixels de regiões como sobrancelhas, pupilas, nariz e boca não são modificadas pela IBSDT. Como já dito antes, um filtro permite limitar a quantidade de detalhes que será transferido para as imagens.

Resultados são mostrados na figura 4.3, onde a imagem (a) é a face original de uma pessoa nova, e a figura (d) é a face original de uma pessoa mais velha. As imagens (b) e (e) ilustram faces simuladas com pequenos detalhes de textura de rugas transferidos, entretanto a figura (b) está envelhecendo e a figura (e) está rejuvenescendo. Já nas imagens (c) e (f) grandes detalhes são transferidos tornando a figura (c) mais envelhecida do que a apresentada na imagem (b) e a figura (f) está mais jovem que na figura (e). A principal vantagem deste método (IBSDT) é que somente uma imagem é necessária para capturar os detalhes e produzir bons resultados, entretanto o maior incômodo é determinar uma taxa de transferência de textura para obter uma progressão de idade realista.

Tiddeman, Burt e Perret [Tiddeman, Burt e Perrett (2001)] produziram resultados convincentes usando métodos baseados em *wavelet* (tipo de filtro usado para extrair características) para criar protótipos de texturas artificiais mudando a idade das pessoas nas imagens. Estes protótipos faciais são resultados da diferença entre dois conjuntos de imagens (um conjunto de imagens de faces jovens e outro de faces envelhecidas). Os autores criaram um protótipo para cada grupo de idade usando essas diferenças para definir um eixo de transformação através da idade, sexo e etnia estimados. O problema principal encontrado foi a perda de algumas características de texturas (como por exemplo rugas) dos protótipos devido ao processo de mistura de várias imagens para criar o protótipo. Portanto, as rugas não são bem capturadas e a textura facial parece mais suave do que nas imagens reais. A localização de rugas é um processo difícil devido a sua variabilidade de largura, tamanho e sua fina estrutura e, uma restrição de suavidade teve de ser imposta nos métodos de deformação. O método

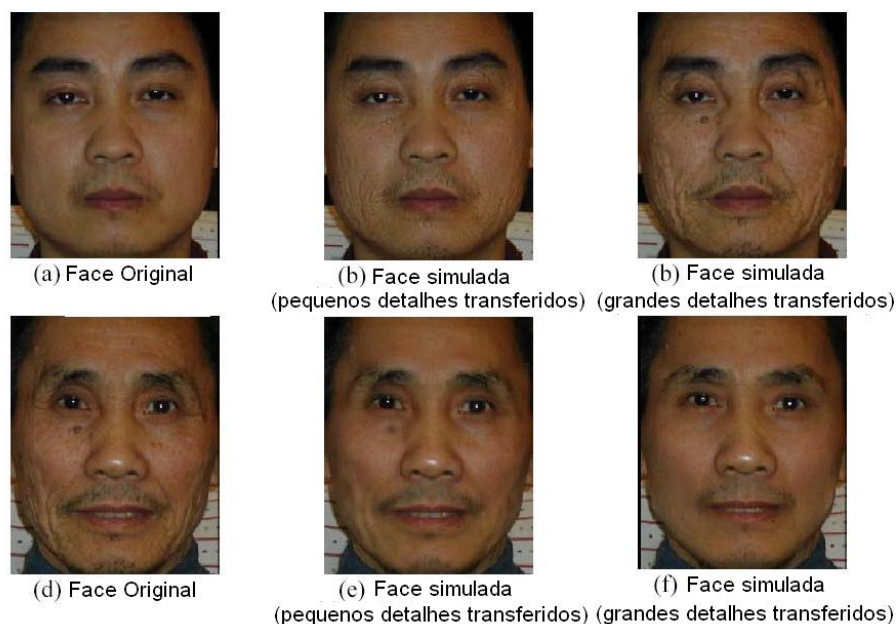


Fig. 4.3: Transferência de textura para simular envelhecimento.

baseado em *wavelet* é usado para melhorar a textura nos protótipos, ele ajusta a amplitude das bordas (mudando a intensidade dos pixels) na imagem de modo que características, como rugas por exemplo, sejam realçadas nos protótipos. Essas características na imagem são capturadas usando o Filtro de Gabor, este filtro permite que as imagens não possuam tantos efeitos de suavização, produzindo excelentes resultados. Experimentos comprovaram que este filtro captura exatamente a informação de envelhecimento perdida no processo de mistura. Transformações de idade que usaram este método de processamento, por *wavelet* provaram ser cinco vezes mais eficazes do que outros métodos. A figura 4.4 ilustra uma aplicação de um protótipo de textura e a figura 4.5 mostra alguns resultados que demonstram o realce da textura e a transformação de idade [Tiddeman, Burt e Perrett (2001)]. Estas imagens mostram o grau de realismo do processo de envelhecimento, podendo-se notar que a identidade da pessoa é mantida enquanto as transformações de idade são aplicadas.

Comparando os métodos por Shan, Liu e Zhang [Shan, Liu e Zhang (2001)] e Tiddeman, Burt e Perret [Tiddeman, Burt e Perrett (2001)], o primeiro apresenta a vantagem de utilizar apenas uma imagem para transferir os detalhes de textura, já o segundo necessita de pelo menos duas imagens para criar o protótipo entretanto, este apresenta melhores resultados visualmente [Gandhi (2004)]. Como já dito anteriormente, a maior dificuldade é a ausência de uma boa base de dados com imagens que possam ser utilizadas para o envelhecimento.

O último trabalho a ser citado foi desenvolvido por Gandhi (2004). Ele criou um sistema que prediz a idade do indivíduo na imagem e depois aplica o processo de envelhecimento ou rejuvenescimento. Para isso, o autor criou seu banco de dados coletando imagens de artistas da internet entre 15 e 99 anos de idade. Essas imagens foram utilizadas no treinamento do *Support Vector Machine (SVM)* que permite estimar a idade das faces nas imagens. A fase de envelhecimento é feita usando ISBDT, entretanto foram utilizadas várias imagens para criar protótipos em diversas idades ao contrário do método de Shan [Shan, Liu e Zhang (2001)] que utilizava somente uma imagem. Portanto, os protóti-

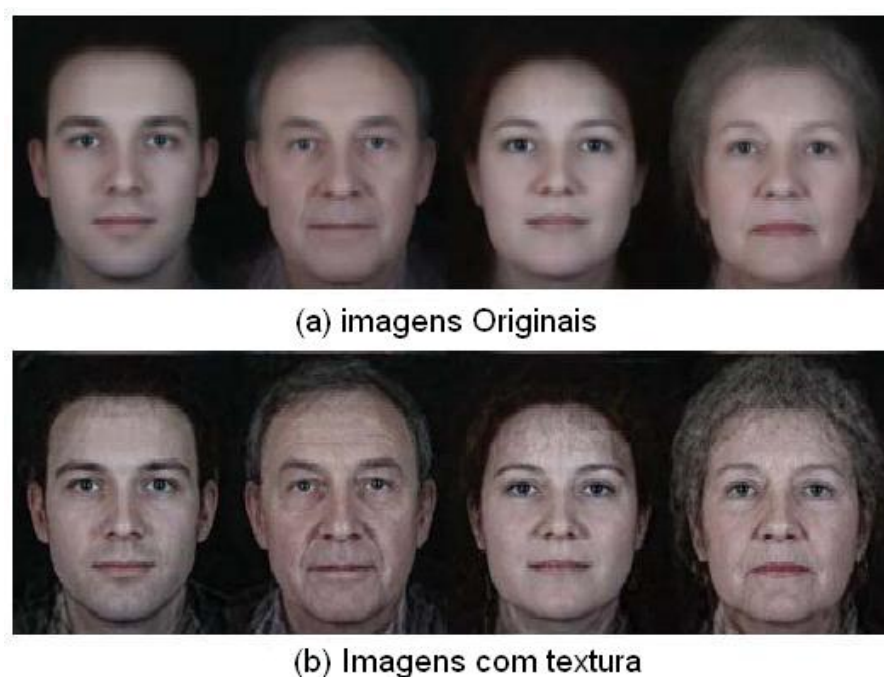


Fig. 4.4: Processamento de textura por wavelet.

pos permitem que rugas sejam transferidas para a imagem inicial a fim de produzir o envelhecimento. Foram criados protótipos de 20 a 80 anos de idade com um intervalo de 5 anos. Um exemplo do seu método pode ser visualizado na figura 4.6, onde o ator Tom Hanks é envelhecido para as idades: 50, 60 e 70 anos de idade.

4.2.3 Métodos que alteram a estrutura da face

Métodos que alteram a estrutura da face oferecem uma melhor compreensão do envelhecimento sob uma perspectiva global. Entretanto para o propósito de envelhecimento, esses métodos podem se tornar mais difíceis de utilização devido a necessidade de medidas precisas. Normalmente, não apresentam mudanças na textura, o que torna o envelhecimento menos realista.

Tentando desenvolver um modelo quantitativo do envelhecimento facial humano a partir de um levantamento estatístico, Leta [Leta (1998)] mediu parâmetros como o tamanho da testa, altura e largura do nariz, espessura dos lábios, etc. Como resultado, obteve um modelo numérico para o envelhecimento facial baseado na variação geométrica. Técnicas de *warping* foram usadas para simular o envelhecimento em mulheres entre 25 e 65 anos de idade, entretanto os resultados visualmente não são muito convincentes, exatamente pela ausência da componente textura. Entretanto, tendo como objetivo cirurgia estética, isto não é um problema.

Com o propósito de simular uma cirurgia facial Koch [Koch et al. (1996)] combinou e estendeu vários métodos de modelamento geométrico, análise de elementos finitos e processamento de imagens para gerar imagens em 3D de uma situação pós cirúrgica. Devido à combinação destas diferentes técnicas, este método mostrou-se muito versátil. Devido a essa versatilidade, esse método pode ser

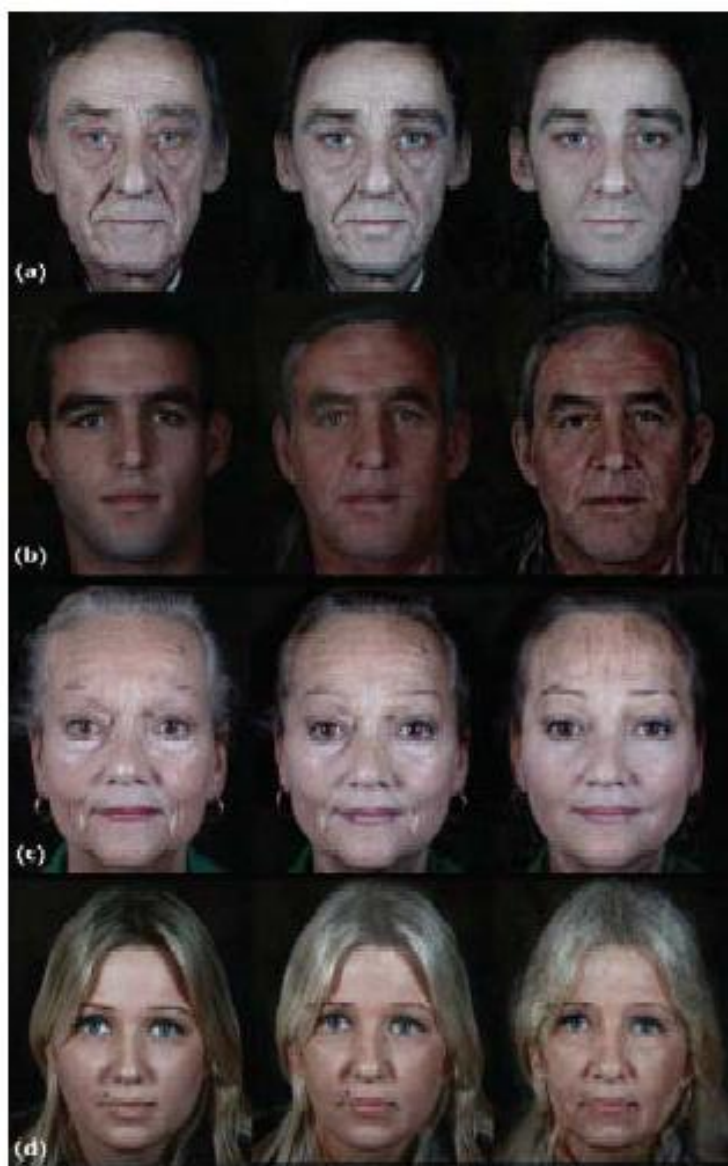


Fig. 4.5: Exemplos de envelhecimento.

empregado em diferentes aplicações incluindo simulação do envelhecimento facial.

Zadoo [Pessa et al. (1999)] fez um estudo sobre alterações no esqueleto e mostrou que a relação do maxilar para a região dos olhos é maior durante a juventude, mas é menor na infância e velhice. Essa relação sugere uma maneira não linear ou multi-modal no processo de envelhecimento facial. As estatísticas médicas dão uma melhor compreensão do processo de envelhecimento de uma perspectiva global, entretanto sua aplicação para a simulação do envelhecimento é difícil devido a falta de precisão nas medidas. Mudanças na estrutura da face são realizadas, mas mudanças na textura são ignoradas, portanto, a simulação não apresenta resultados reais.

Os principais elementos para criar uma animação de envelhecimento são variações na textura da pele e mudanças na geometria do rosto. Hussein [Hussein (2002)] propôs um algoritmo para



Fig. 4.6: Exemplo dos resultados obtidos por Gandhi para a simulação de envelhecimento.

criar simulações faciais de envelhecimento integrando esses elementos. No seu artigo, utilizou duas técnicas: uma para a deformação e outra para a mudança de textura. A deformação baseou-se em conceitos da teoria da antropometria da face (estudo que mede e relaciona o formato e estrutura da face) e, na simulação de rugas utilizou uma função chamada de *Bidirectional Reflectance Distribution function* (BRDF - função que descreve como a luz é distribuída em uma superfície). Portanto, a partir de uma face, a idéia é capturar detalhes da textura da pele e capturar dados antropométricos da face. Então, essas características extraídas são aplicadas em outra face utilizando alguma técnica de deformação afim de gerar uma face com aparência mais velha. A contribuição e vantagem desse artigo comparado com outros é que a técnica proposta é simples de implementar e, somente uma imagem de entrada é necessária, enquanto os outros necessitam de um banco de imagens. A figura 4.7 mostra resultados deste método.

4.3 Considerações Finais

Neste capítulo foram apresentadas características do envelhecimento facial. Foi realizada uma revisão bibliográfica sobre o que foi desenvolvido até o momento. Os sistemas foram classificados em três tipos: sistemas que permitem estimar a idade de pessoas em imagens faciais; sistemas de envelhecimento por aprendizado de alteração de textura; sistemas de envelhecimento realizado por medidas estatísticas ocorridas com o passar do tempo.

A tabela 4.1 resume os vários métodos para estimar a idade em imagens, incluindo detalhes do banco de dados utilizado.

Em sistemas por aprendizado de textura, os principais trabalhos são de:

- Lanitis, 2002 : neste trabalho as faces são codificadas para depois serem reconstruídas, e imagens de treinamento foram usadas para o sistema aprender uma relação entre a representação



Fig. 4.7: Exemplos de envelhecimento por BRDF.

Trabalho	Banco de dados	Dados treinados	Dados testados	Taxa de acerto
Kwon and Lobo	47 imagens classificadas como bebês, jovens e adultos	-	15 imagens	100%
Horng et al.	230 imagens classificadas como bebês, jovens e adultos	-	230 imagens	81.6%
Hayashi	300 imagens entre 15 e 64 anos	-	300 imagens	27%
Lanitis	330 imagens entre 0 e 35 anos	250 imagens	-	erro 3.83 anos

Tab. 4.1: Métodos para estimar a idade.

da face codificada e a atual idade dos sujeitos. Permite o envelhecimento de crianças e adultos até 30 anos de idade.

- Shan et al, 2001: propôs o método chamado *Image Based Surface Detail Transfer* (IBSDT) onde detalhes geométricos de uma superfície são transferidos para outra superfície, permitindo que detalhes de textura, como rugas sejam transferidos para faces jovens.
- Tiddeman et al 2001 : utilizou *wavelet* (tipo de filtro usado para extrair características desejadas) para criar protótipos de texturas artificiais mudando a idade das pessoas nas imagens. Portanto, um protótipo para cada grupo de idade usado foi criado para ser aplicado nas imagens a serem envelhecidas.
- Gandhi, 2004 : criou um sistema para estimar a idade utilizando uma técnica de aprendizado chamada *Support Vector Machine* (SVM). E para a fase de envelhecimento, usou o método de transferência de textura chamado ISBDT.

E, em sistemas que alteram a estrutura da face, foram citados:

- Leta (1998): obteve um modelo numérico para o envelhecimento facial baseado na variação geométrica. Técnicas de *warping* foram usadas para simular o envelhecimento em mulheres entre 25 e 65 anos de idade.
- Koch et al. (1996) : criou um método para análise de faces em uma situação pós cirúrgica. Entretanto, esse método pode ser empregado em diferentes aplicações incluindo simulação do envelhecimento facial.
- Pessa J. et al(1999) : fez um estudo sobre o esqueleto da face e mostrou que a relação do maxilar para o orbital é maior durante a juventude, mas é menor na infância e velhice. Portanto, mudanças na estrutura da face são realizadas, mas mudanças na textura são ignoradas.
- Hussein (2002) : propôs um algoritmo aplicando variações na textura da pele e mudanças na geometria do rosto. Para a simulação de rugas utilizou uma função chamada de *Bidirectional Reflectance Distribution function* (BRDF - função que descreve como a luz é distribuída em uma superfície).

Portanto, analisando todos estes métodos vê-se que ainda não existem simulações que produzam efeitos de envelhecimento realistas. E, existe a necessidade de um bom banco de imagens aplicado ao envelhecimento.

Capítulo 5

Modelagem do Envelhecimento

A partir de agora, será descrito como o envelhecimento foi modelado computacionalmente no presente trabalho. A implementação está baseada em métodos de medidas estatísticas faciais para simular o envelhecimento facial, para isso utilizou-se a metodologia proposta por Leta [Leta (1998)]. Esta metodologia é baseada em métodos estatísticos a partir da análise matemática de medidas faciais em um grupo específico de pessoas, das quais se dispunha de fotografias em idades distintas. A partir dessa análise matemática, obteve-se um padrão para representar e simular o processo de envelhecimento em imagens faciais.

Este método define as regiões faciais onde é perceptível o efeito do avanço da idade e mede as variações destas regiões ao longo do tempo em um grupo de pessoas. A partir das medições que consideram características pessoais, é criado um modelo comum a todo o grupo e, curvas de envelhecimento são geradas para as regiões definidas. A partir destas curvas, é possível definir e calcular o deslocamento dos pontos iniciais, obtendo-se os pontos finais para o processo de deformação (*warping*).

No trabalho [Leta (1998)], para a extração de medidas e definição das curvas de envelhecimento, foi necessário o uso de um banco de dados de imagens faciais. Este banco de dados foi criado a partir de fichas de pacientes que se submeteram à cirurgia plástica na Clínica Ivo Pintanguy ou na Santa Casa de Misericórdia do Rio de Janeiro. Sendo selecionadas 342 fotos de pacientes do sexo feminino, da raça branca, entre 18 e 70 anos. Entretanto, de todas essas fotografias, somente 40 puderam ser utilizadas devido a problemas como: falta de iluminação, má qualidade nas fotos, etc.

A seguir serão apresentadas as regiões da face utilizadas, assim como os pontos definidos sobre o rosto para que pudessem ser definidas distâncias entre estes pontos e a relação entre estas distâncias com o avanço da idade.

5.1 Regiões faciais sujeitas ao envelhecimento

Segundo [Pitanguy et al. (1977)], para avaliar qualitativamente o envelhecimento facial, a face pode ser dividida em oito regiões: frontal, glabellar, orbitária, nasal, lateral, orolabial, mentoneana e cervical (figura 5.1). Destacam-se as seguintes características de envelhecimento cutâneo nestas regiões :

- Frontal (das sobrancelhas até a implantação do cabelo): é a região onde aparecem as primeiras

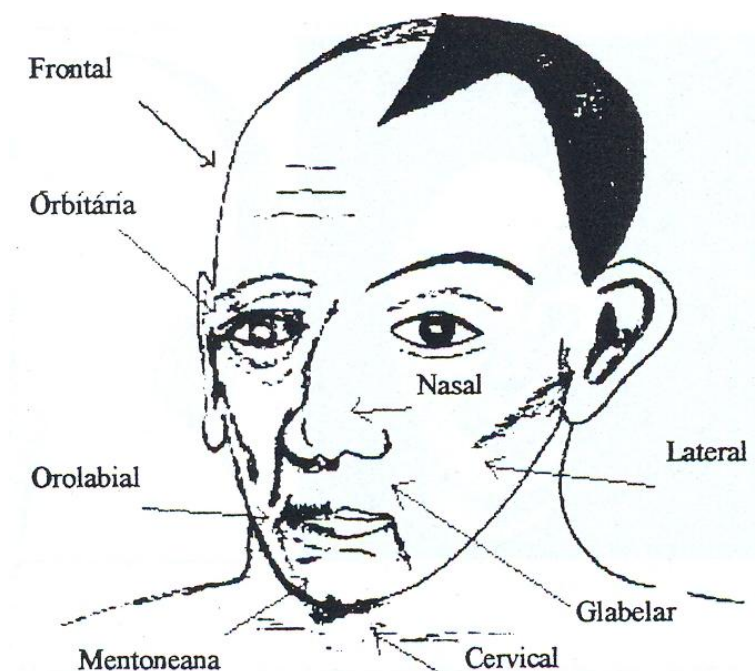


Fig. 5.1: Regiões Faciais.

rugos em torno dos 25 anos;

- Glabellar: é a região onde surgem rugas devido ao movimento dos músculos corrugadores por volta dos 40 anos;
- Orbitária (em torno dos olhos): é considerada um dos pivôs do envelhecimento facial;
- Nasal (definida pelo contorno do nariz): nesta região observa-se o alongamento do nariz com uma rotação de sua porção móvel para baixo;
- Lateral (corresponde à lateral da face): os tecidos desta área deslizam gerando bolsas flácidas;
- Orolabial (é a região em torno dos lábios): observa-se um afinamento dos lábios com o envelhecimento;
- Mentoneana (é definida pelo contorno inferior do rosto): observa-se o chamado mento (queixo) senil devido ao excesso de gordura, que decai em decorrência do relaxamento da pele e da ação da gravidade;
- Cervical (região do pescoço): caracterizada por rugas horizontais.

Com base nestas regiões características do envelhecimento [Pitanguy et al. (1977)], (figura 5.1), a metodologia define 24 pontos na imagem facial como mostrado na figura 5.2. Primeiro são definidos os pontos A e B, no centro das pupilas, depois alguns pontos extras são necessários como pontos nos cantos externos dos olhos e nos cantos externos da boca. A partir da marcação destes pontos, linhas de referência verticais são visualizadas para a marcação do restante dos pontos. Estas linhas de referência, são as que passam pelos seguintes pontos e são definidas como:

- $\overline{AB}$ - essa linha é definida pelo segmento de reta formado pelos pontos A e B . Ela passa pelo centro das pupilas e mede a distância interpupilar;
- $\overline{EW}$ - é definida pelo segmento de reta que passa pelos pontos E e W . É perpendicular a AB passando pelo ponto médio da distância entre pupilas (C).
- $\overline{DU}$ e $\overline{FY}$: perpendiculares ao segmento AB , ligam o mento ao início da implantação do cabelo passando pelas pupilas (pontos A e B);
- $\overline{GT}$ e $\overline{JZ}$: perpendiculares ao segmento AB , passam pelo canto lateral externo do olho unindo a parte superior da sobrancelha à parte inferior da face.

Essas linhas são usadas como base para a marcação do restante dos pontos na imagem pelo usuário. Todos os pontos foram nomeados com as letras do alfabeto e podem ser visualizados na figura 5.2.

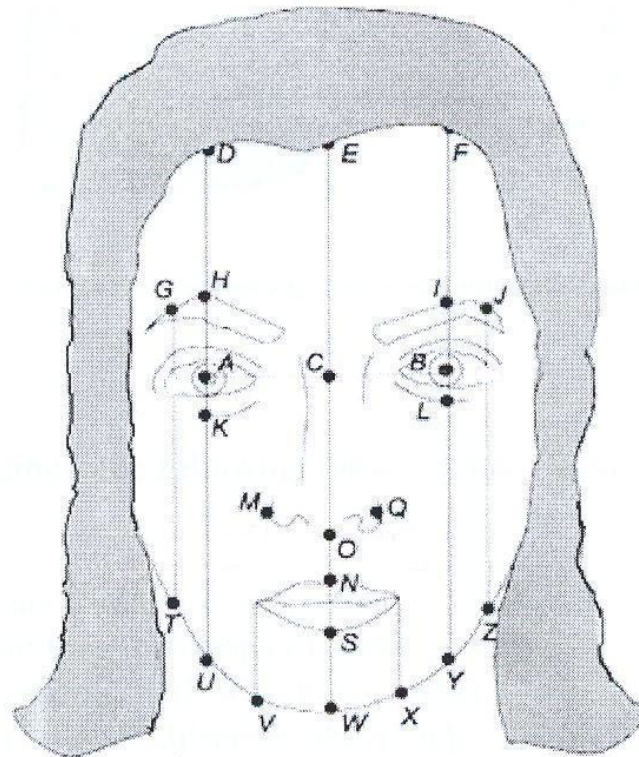


Fig. 5.2: Pontos marcados na imagem facial.

A partir destes pontos definidos sobre a imagem, são calculadas as medidas de segmentos nas imagens frontais de figuras faciais e, estas medidas são usadas para fazer o levantamento estatístico das curvas de envelhecimento. Os segmentos são mostrados na tabela 5.1 e por convenção adota-se a seguinte nomenclatura: a coordenada x do ponto A é chamada de A_x e a coordenada y de A_y .

<i>Nome da distância</i>	<i>Segmento</i>
Distância interpupilar	$B_x - A_x$
Altura da testa - lateral esquerda	$A_x - D_x$
Altura da testa - lateral direita	$B_y - F_y$
Altura da testa - central	$C_y - E_y$
Altura da sobrancelha na direção do canto externo do olho direito	$B_y - J_y$
Altura da sobrancelha na direção do centro do olho direito	$B_y - I_y$
Altura da sobrancelha na direção do centro do olho esquerdo	$A_y - H_y$
Altura da bolsa palpebral - esquerda	$K_y - A_y$
Altura da bolsa palpebral - direita	$L_y - B_y$
Altura do Mento na direção do canto do olho direito	$Z_y - B_y$
Altura do Mento na direção do canto do olho esquerdo	$T_y - A_y$
Altura do Mento na direção do centro do olho direito	$Y_y - B_y$
Altura do Mento na direção do centro do olho esquerdo	$U_y - A_y$
Altura do Mento na direção do canto direito da boca	$X_y - B_y$
Altura do Mento na direção do canto esquerdo da boca	$V_y - A_y$
Altura do Mento central	$(W_y - C_y)$
Altura do nariz	$O_y - C_y$
Largura do nariz	$M_x - Q_x$
Altura do Lábio Superior	$N_y - C_y$
Espessura do Lábio	$S_y - N_y$

Tab. 5.1: Segmentos medidos na face.

5.2 Curvas de envelhecimento

As curvas de envelhecimento foram levantadas conforme as regiões da face apresentadas anteriormente: frontal, orbitária, orolabial, nasal e mentoneana. Partindo-se das medidas da face, deseja-se obter uma curva de envelhecimento genérica para cada parâmetro de envelhecimento. Considerando-se, por exemplo, as medidas de um determinado parâmetro facial de dois pacientes, deseja-se que para ambos o comportamento de envelhecimento siga a mesma regra, ou seja, a mesma curva genérica. Para isso, cada medida de um dos parâmetros de envelhecimento citados anteriormente foi extraída de várias faces de pessoas com idades diferentes. Para exemplificar, considera-se o gráfico ilustrado na figura 5.3(a), i representa o número de uma pessoa, t é a idade dessa pessoa, $Y_i(t)$ é a medida de algum parâmetro da pessoa i na idade t e, o objetivo é construir uma curva genérica u (figura 5.3(b)) para representar cada parâmetro, independente da face.

Então, após todas as medidas serem realizadas de todas as imagens disponíveis, foi usada uma função polinomial para representar a curva genérica u que caracteriza o envelhecimento. Essa função foi definida por uma combinação de polinômios de segunda ordem (equação 5.1). A escolha deste tipo de função deve-se à sua simplicidade e aos bons resultados de ajustagem que foram obtidos [Leta (1998)].

$$p(t) = \begin{cases} a_1 t^2 + b_1 t + c_1, & t \leq t_c \\ a_2 t^2 + b_2 t + c_2, & t > t_c \end{cases} \quad (5.1)$$

onde $(t_c, p(t_c))$ representa um ponto de inflexão.

A função p deve satisfazer às seguintes propriedades no ponto de inflexão:

- (i) ser contínua, e,
- (ii) ter derivada contínua.

Conseqüentemente tem-se no ponto de inflexão:

$$\begin{aligned} a_1 t_c^2 + b_1 t_c + c_1 &= a_2 t_c^2 + b_2 t_c + c_2 \\ 2a_1 t_c + b_1 &= 2a_2 t_c + b_2 \end{aligned} \quad (5.2)$$

Além disso, deseja-se que:

$$\lim_{t \rightarrow t_c^+} p''(t) = - \lim_{t \rightarrow t_c^-} p''(t) \quad (5.3)$$

Assim:

$$\begin{aligned} a &= a_1 = -a_2 \\ b &= b_1 = b_2 - 4a t_c \\ c &= c_1 = 2a t_c^2 + c_2 \end{aligned} \quad (5.4)$$

A partir das equações 5.1 e 5.4, é possível escrever:

$$p(t) = \begin{cases} a_1 t^2 + b_1 t + c_1, & t \leq t_c \\ -a_2 t^2 + (b_2 + 4a_2 t_c)t + c_2 - 2a_2 t_c^2, & t > t_c \end{cases} \quad (5.5)$$

Esta é a forma geral dos polinômios usados para representar as curvas de envelhecimento associadas às diferentes regiões estudadas da face. O ponto de inflexão t_c será a idade de inflexão e, para

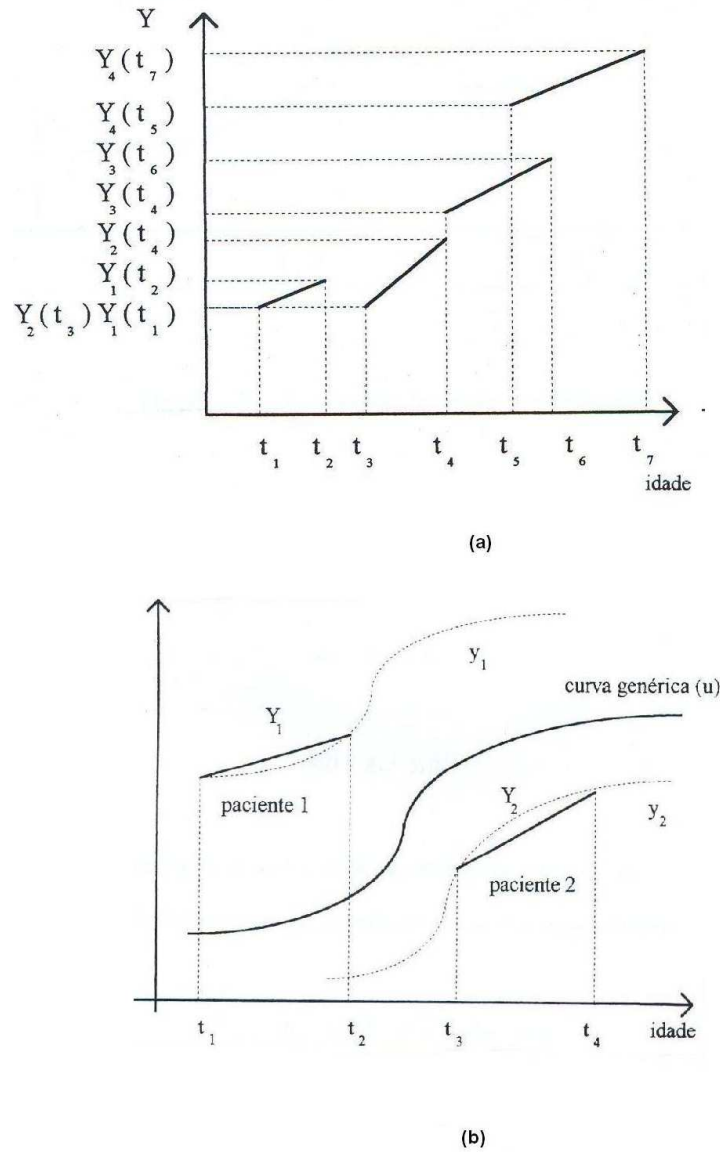


Fig. 5.3: Medidas de um mesmo parâmetro em quatro pacientes.

a composição dos polinômios, foram realizados testes considerando diversas idades de inflexão (20, 30, 45 e 50 anos). Para a idade de inflexão 20 é definido um único polinômio, tendo em vista que esta idade foi adotada como limite inferior para o ajuste das curvas. Para os demais casos ajustam-se dois polinômios com concavidades diferentes, isto é, um com $a \geq 0$ até a idade de inflexão e outro com $a < 0$ a partir desta idade.

Para maiores detalhes, consulte o trabalho [Leta (1998)], de onde foram extraídas estas equações e todo o desenvolvimento das curvas. Portanto, não fizeram parte deste trabalho a extração das medidas e desenvolvimento destas curvas, elas foram utilizadas como uma aplicação para o *morphing* de imagens faciais.

Nos próximos itens, para cada parâmetro de envelhecimento, serão apresentadas as medidas obti-

das, os coeficientes dos polinômios geradores das curvas finais e, as curvas [Leta (1998)].

5.2.1 Região Frontal - Altura da Testa

A região frontal é delimitada pela linha de implantação do cabelo e pelas sobrancelhas. As curvas da testa foram levantadas a partir da medição dos comprimentos entre a posição vertical das pupilas e a implantação pilosa.

Para a testa, existem duas curvas diferentes: uma para a região central, na direção do centro da face formada pelos pontos C_y e E_y e, outra para as regiões laterais, na direção das pupilas esquerda (pontos A_y e D_y) e direita (pontos B_y e F_y). As duas curvas são apresentadas na figura 5.4, a idade de inflexão usada é de 45 anos e os coeficientes encontrados para a equação 5.5 são apresentados na tabela 5.2.

Coeficiente	Lateral	Central
a_1	$0.5585 * 10^{-4}$	$0.0866 * 10^{-4}$
b_1	$-0.1108 * 10^{-2}$	$0.1605 * 10^{-2}$
c_1	41.0478	0.9658
a_2	$-0.5585 * 10^{-4}$	$0.0866 * 10^{-4}$
b_2	$0.8945 * 10^{-2}$	$0.31638 * 10^{-2}$
c_2	0.8216	0.9307

Tab. 5.2: Coeficientes dos polinômios da altura da testa (equação 5.5).

Conforme pode ser observado na figura 5.4 ocorre um contínuo aumento da testa com o envelhecimento, sendo que na parte lateral da testa ocorre um maior crescimento, ou seja, podem ser percebidas as chamadas “entradas” laterais.

5.2.2 Região Frontal - Altura da Sobrancelha

As curvas da sobrancelha foram obtidas a partir da medição dos comprimentos verticais entre o nível das pupilas e os pontos na parte superior da sobrancelha, nas seguintes direções:

- Lateral: canto do olho esquerdo ($A_y - G_y$) e direito ($B_y - J_y$).
- Central: centro do olho esquerdo ($A_y - H_y$) e direito ($B_y - I_y$).

As curvas para esta região possuem como ponto de inflexão a idade de 20 anos, portanto apenas uma curva para a região lateral e apenas uma curva para a região central são necessárias, sendo os coeficientes destas apresentados na tabela 5.3.

Analisando a figura 5.5 pode-se perceber a redução da altura da sobrancelha, com uma pequena diferença de caimento entre a parte lateral e central. Em 50 anos ocorre uma redução de cerca de 20% na altura da sobrancelha.

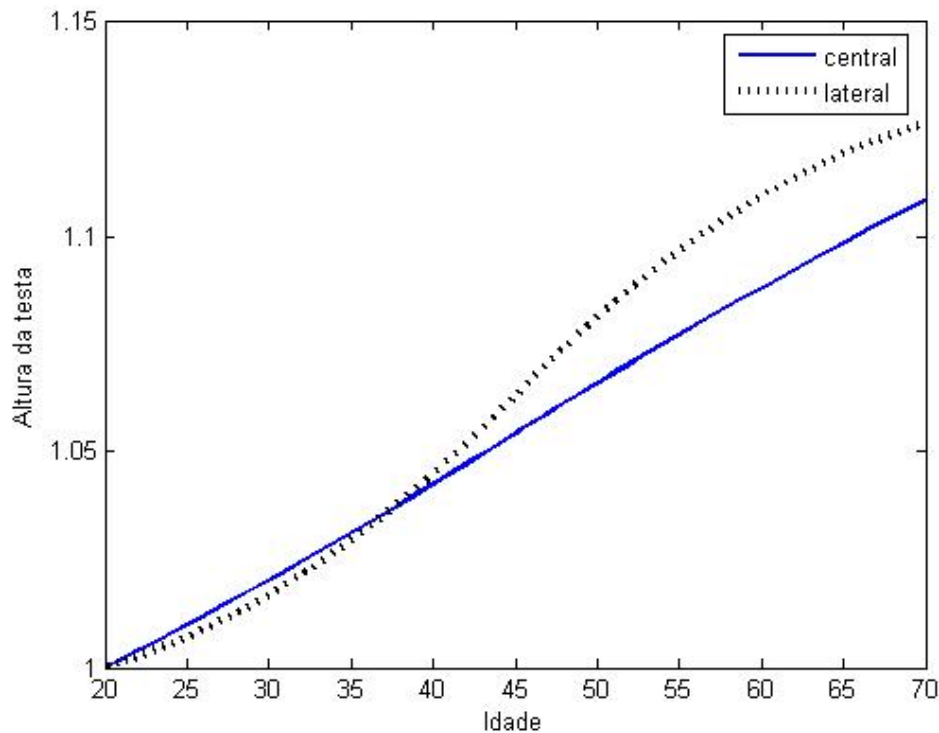


Fig. 5.4: Curvas de envelhecimento da testa - central e lateral.

Coeficiente	Lateral	Central
a_1	$0.4354 * 10^{-4}$	$0.6231 * 10^{-4}$
b_1	$-0.3122 * 10^{-2}$	$-0.2981 * 10^{-2}$
c_1	1.0244	0.9711

Tab. 5.3: Coeficientes dos polinômios da altura da sobrancelha (equação 5.5).

5.2.3 Região Orbitária - Altura da Bolsa Palpebral

A Região Orbitária caracteriza-se pelas curvas da sobrancelha (que delimitam também a região frontal) e da bolsa palpebral. A curva correspondente à bolsa palpebral é obtida a partir da medição do comprimento entre o nível das pupilas e o ponto inferior da bolsa na direção da pupila, sendo para o olho esquerdo ($K_y - A_y$) e, para o olho direito ($L_y - B_y$).

As curvas para esta região possuem como ponto de inflexão a idade de 20 anos, portanto apenas uma curva é necessária, sendo esta definida pelos seguintes coeficientes da tabela 5.4.

Conforme a figura 5.6, a bolsa palpebral aumenta consideravelmente com o envelhecimento. Percebe-se a necessidade de avaliar outros pontos, que permitam uma definição melhor do seu contorno, para obter-se resultados mais realistas. Acredita-se que a partir de uma determinada idade o crescimento da bolsa palpebral deva se estabilizar por limitações anatômicas, apesar da curva obtida não representar tal fato.

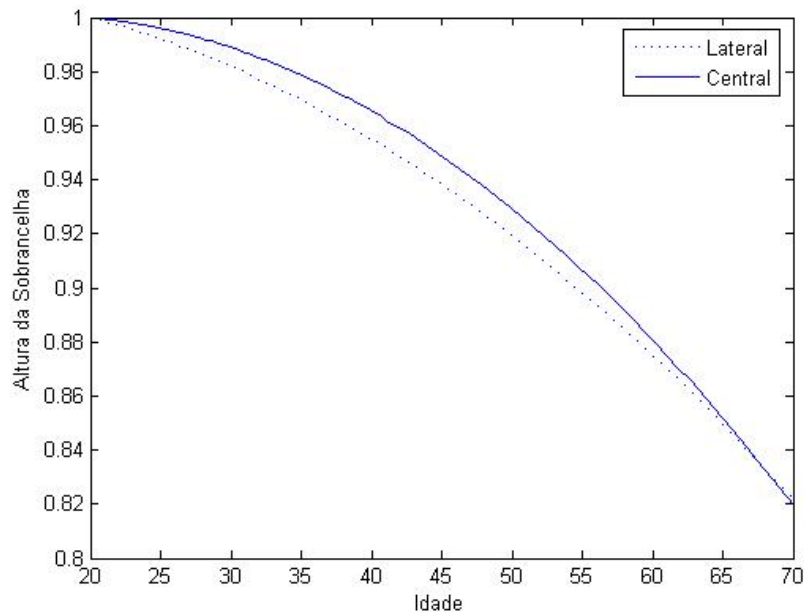


Fig. 5.5: Curvas de envelhecimento da sobrancelha - central e lateral.

Coeficiente	Valor
a_1	$-2.6894 * 10^{-4}$
b_1	$-1.3922 * 10^{-2}$
c_1	0.7026

Tab. 5.4: Coeficientes dos polinômios da altura da bolsa palpebral (equação 5.5).

5.2.4 Região Nasal - Altura e largura do nariz

A região nasal caracteriza-se pelos segmentos que delineiam o nariz, e são definidas por duas curvas de envelhecimento, a primeira referente à altura ou comprimento do nariz ($O_y - C_y$) e a segunda referente à largura do nariz ($Q_x - M_x$). A curva da altura possui uma idade de inflexão de 20 anos, entretanto a curva da largura do nariz tem uma idade de inflexão de 45 anos.

Coeficiente	Espessura	Altura
a_1	$0.0540 * 10^{-4}$	$1.3997 * 10^{-4}$
b_1	$0.7004 * 10^{-2}$	$-0.6810 * 10^{-2}$
c_1	0.8387	1.0658
a_2	-	$-1.3997 * 10^{-4}$
b_2	-	$-0.6558 * 10^{-2}$
c_2	-	0.4989

Tab. 5.5: Coeficientes dos polinômios de altura e largura do nariz (equação 5.5).

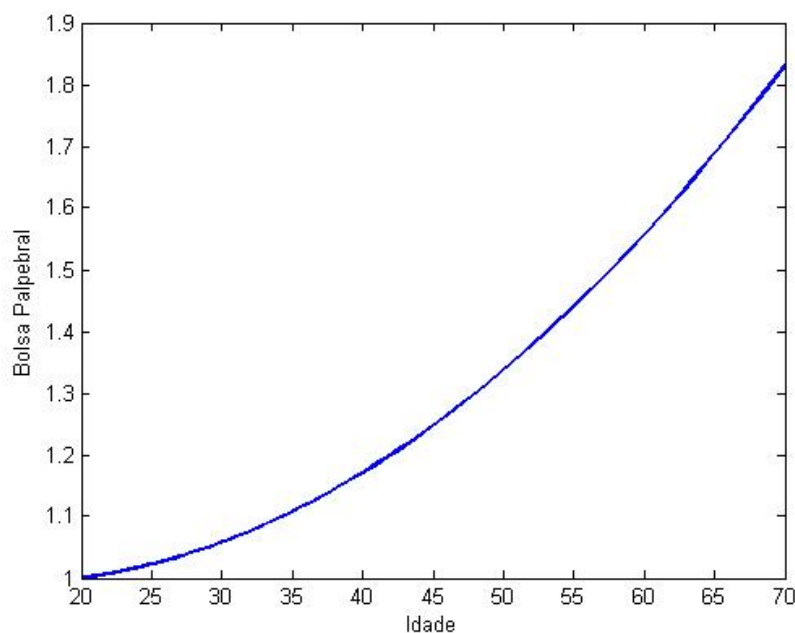


Fig. 5.6: Curvas de envelhecimento da bolsa palpebral.

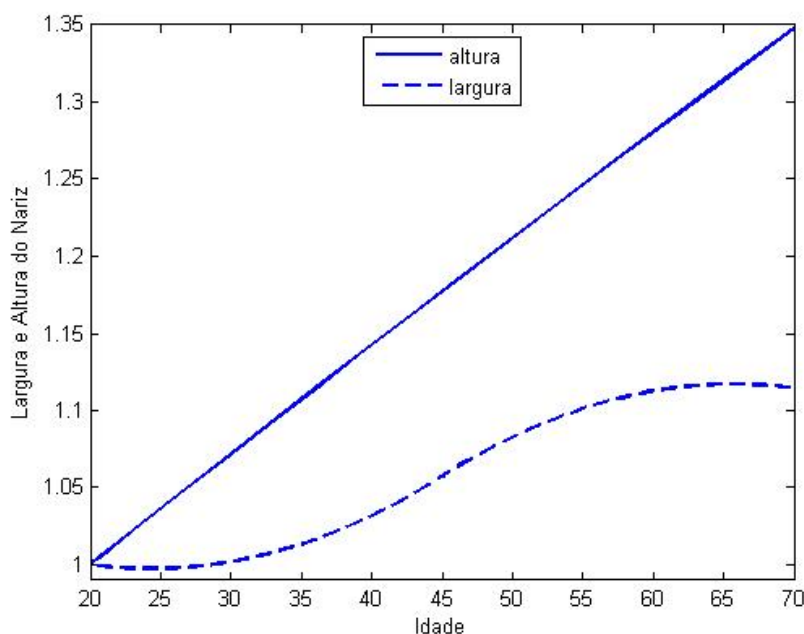


Fig. 5.7: Curvas de envelhecimento do nariz.

Na figura 5.7 pode-se observar um considerável aumento do comprimento do nariz e um gradativo aumento de sua largura, que se torna mais acentuado na faixa de 40 a 50 anos. Observa-se em fotografias faciais de perfil, que a ponta do nariz com o passar do tempo cai girando levemente, este

movimento pode ser uma das razões para o aumento da largura do nariz.

5.2.5 Região Orolabial - Altura do lábio superior e espessura labial

A região orolabial caracteriza-se por segmentos que delineiam os lábios e as rugas naso-genianas. A curva correspondente à altura do lábio superior refere-se à medição do comprimento entre o nível dos olhos e o ponto superior do lábio ($C_y - N_y$) e a espessura labial é dada por ($N_y - S_y$) na mesma direção. Tanto a curva da espessura quanto a curva de altura, possuem uma idade de inflexão de 20 anos.

Coeficiente	Espessura	Altura
a_1	$-0.0262 * 10^{-4}$	$-0.6791 * 10^{-4}$
b_1	$-0.7288 * 10^{-2}$	$0.4747 * 10^{-2}$
c_1	1.0236	0.9073

Tab. 5.6: Coeficientes dos polinômios da espessura e altura dos lábios (equação 5.5).

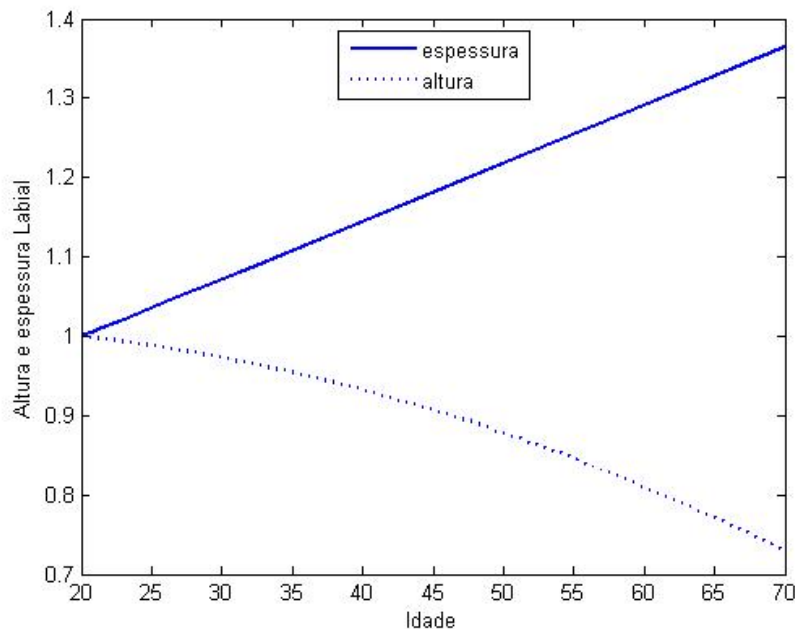


Fig. 5.8: Curvas de envelhecimento dos lábios.

Ocorre um razoável afinamento labial com o envelhecimento. A altura do lábio superior aumenta colaborando com a redução da espessura labial. A medida da espessura labial sofre influência da maquiagem e é muito sensível a pequenas variações de expressão, por esta razão observa-se um erro maior no ajuste de sua curva. Seria interessante medir mais de um ponto no contorno do lábio, especialmente na parte lateral, pois observa-se um leve caimento da boca nesta região e em alguns casos a formação da ruga oromental.

5.2.6 Região Mentoneana

A região mentoneana corresponde a todo o contorno da parte inferior da face, incluindo o mento e a bolsa lateral da face. Sete alturas foram medidas para a obtenção das curvas de envelhecimento desta região, considerando a direção vertical dos seguintes pontos:

- Canto externo do olho - esquerdo ($C_y - T_y$) e direito ($C_y - Z_y$);
- Centro do olho - esquerdo ($C_y - U_y$) e direito ($C_y - Y_y$);
- Canto da boca - esquerdo ($C_y - V_y$) e direito ($C_y - X_y$);
- Centro da face ($C_y - W_y$).

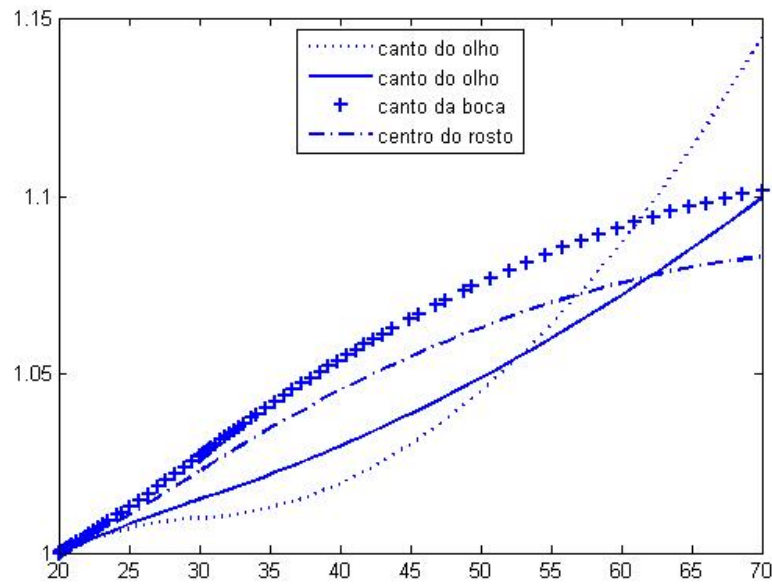


Fig. 5.9: Curvas de envelhecimento do mento.

Coeficiente	Canto olho	Centro olho	Canto boca	Centro face
a_1	$-0.7992 * 10^{-4}$	$-0.2055 * 10^{-4}$	$0.2842 * 10^{-4}$	$0.2566 * 10^{-4}$
b_1	$0.4962 * 10^{-2}$	$0.2524 * 10^{-2}$	$0.1291 * 10^{-2}$	$0.0994 * 10^{-2}$
c_1	40.9391	0.9560	0.9483	0.9814
a_2	$0.7992 * 10^{-4}$	$0.2055 * 10^{-4}$	$-0.2842 * 10^{-4}$	$-0.2566 * 10^{-4}$
b_2	$-0.4604 * 10^{-2}$	$0.0058 * 10^{-2}$	$0.4701 * 10^{-2}$	$0.4073 * 10^{-2}$
c_2	1.0825	0.9929	0.8971	0.9352

Tab. 5.7: Coeficientes dos polinômios para a região mentoneana (equação 5.5).

Foi utilizada a idade de inflexão de 30 anos para o ajuste de todas as curvas da região mentoneana e os coeficientes de ajuste das curvas podem ser visualizados na tabela 5.7.

A partir das curvas obtidas na figura 5.9 observa-se uma das características decorrentes do envelhecimento, o chamado "queixo de bruxa", ou seja ocorre um maior caimento das bolsas laterais e do mento (direção do centro da face) e um aumento menor na direção do canto da boca.

5.3 Considerações Finais

Foi visto neste capítulo como foram obtidas as curvas de envelhecimento e o comportamento destas através dos gráficos. As regiões analisadas e as principais características observadas são resumidas da seguinte maneira:

- Região Frontal - Altura da testa: ocorre um contínuo aumento da testa com o envelhecimento, sendo que na parte lateral da testa ocorre um maior crescimento.
- Região Frontal - Altura da sobrancelha: ocorre a redução da altura da sobrancelha, com uma pequena diferença de caimento entre a parte lateral e central.
- Região Orbitária - Altura da bolsa palpebral: a bolsa palpebral aumenta consideravelmente com o envelhecimento.
- Região Nasal - Altura e largura do nariz: pode-se observar um considerável aumento do comprimento do nariz e um gradativo aumento de sua largura, que se torna mais acentuado na faixa de 40 a 50 anos.
- Região Orolabial - Altura do lábio superior e espessura labial: ocorre um razoável afinamento labial com o envelhecimento.
- Região Mentoneana: ocorre um maior caimento das bolsas laterais e do mento (direção do centro da face) e um aumento menor na direção do canto da boca.

Capítulo 6

Implementação e Resultados

O propósito deste trabalho foi desenvolver um protótipo que permitisse envelhecer imagens de faces utilizando *morphing*. Para isso, primeiramente foi realizado um estudo sobre *morphing* (capítulo 2), pesquisando como funcionam as principais técnicas existentes. Após isso, foi feito um estudo sobre o que existe na área de envelhecimento de imagens faciais (capítulo 4). Tendo esses dois conceitos bem definidos (*morphing* e envelhecimento), uma proposta de envelhecimento possível de implementar foi escolhida e, iniciou-se uma pesquisa para escolher qual método de *morphing* era melhor para essa proposta. Resolveu-se usar o trabalho proposto por Fabiana Leta [Leta (1998)], de onde foram extraídas curvas de envelhecimento aplicadas neste trabalho (capítulo 5). Em seguida foi escolhido para síntese de envelhecimento, um método que utilizasse pontos como primitiva de especificação de características e então, foi adotado o método de *morphing* que utiliza funções de base radial (capítulo 3). Foi visto que a maioria dos sistemas de envelhecimento necessitam de uma boa base de dados, e foram pesquisados bancos de imagens para a aplicação de envelhecimento, mas somente um banco de imagens foi encontrado na internet [FGNET (2006)], entretanto as imagens nesse banco de imagens não possuem um padrão, são imagens de diferentes tamanhos onde as pessoas foram fotografadas em ângulos diferentes e isso impossibilitou o uso desse banco para extrair características que permitissem montar um sistema de envelhecimento.

A partir de agora será explicada e detalhada a implementação realizada e os resultados obtidos.

6.1 Implementação do Protótipo

O protótipo implementado foi baseado nos conceitos vistos nos capítulos anteriores, unindo os conceitos dos capítulos 3 e 5. Para uma melhor implementação, o protótipo desenvolvido foi dividido em diversas partes, e na figura 6.1 pode-se ver uma representação dessas partes e do funcionamento do protótipo por inteiro.

Portanto, o funcionamento do protótipo é da seguinte maneira: o usuário escolhe a imagem desejada e fornece as coordenadas das pupilas na imagem; uma normalização da imagem é realizada para ela ter um tamanho padronizado para este trabalho (300x250). Após, os 24 pontos vistos no capítulo 5 serem informados é feito o cálculo das medidas de cada parâmetro. A partir da idade inicial e final e com essas medidas, são calculadas as novas medidas para a idade de envelhecimento determinada (idade final). Então, os coeficientes para a função de base radial são calculados e fornecidos

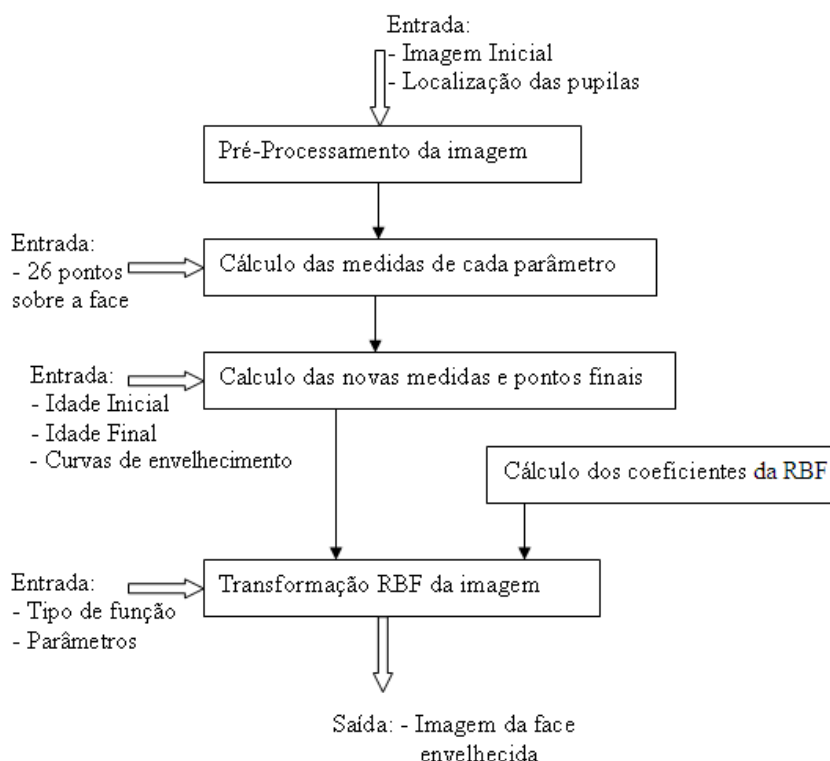


Fig. 6.1: Diagrama.

para o subsistema que fará a transformação da imagem inicial para a imagem final (envelhecimento), para isso, parâmetros pertinentes às funções de base radial, como por exemplo, o tipo de função e a constante λ , precisam ser determinados pelo usuário.

A seguir, em cada subseção deste capítulo, serão detalhados os blocos apresentados no diagrama disposto na figura 6.1 que compõem o protótipo de envelhecimento deste trabalho e que foram implementados no MATLAB, versão 7.0.

6.1.1 Pré-processamento da Imagem

O pré-processamento pode ser feito de duas formas: sem causar incômodo ao usuário, ou necessitando que o usuário informe alguns dados. Neste trabalho, o usuário precisa informar a posição das pupilas na imagem.

Os principais objetivos do pré-processamento são: localizar e isolar a região da face, evitando interferência do fundo; normalizar o tamanho da imagem da face, o que faz com que o sistema não seja sensível à distância da câmera para o usuário; corrigir problemas de iluminação; eliminar, através do processamento da imagem, objetos como óculos e jóias, etc. Entretanto, para este trabalho o principal objetivo foi estabelecer um padrão de tamanho e normalizar a imagem de entrada para que o rosto não ficasse inclinado, e as curvas de envelhecimento propostas no capítulo 5 pudessem ser aplicadas sobre a imagem.

A localização e normalização da face foram realizadas baseando-se na posição dos olhos e na

distância intra-pupilar que não se altera durante o envelhecimento em adultos [Leta et al. (1996)]. As posições dos olhos são determinadas manualmente, para isso, os pontos A e B, ou seja, os pontos no centro da pupila direita e da pupila esquerda são os primeiros pontos a serem definidos.

Portanto, para o protótipo desenvolvido, a imagem é redimensionada para ter 300x250 pixels, onde os olhos da pessoa na imagem estarão localizados cerca de 100 pixels abaixo do topo da imagem, conforme pode ser visto na figura 6.2. Os números mostrados na figura representam a distância entre os pontos, utilizando pixels como medida de comprimento.

Desta forma, dada as posições dos olhos, o pré-processamento da imagem é feito através dos seguintes passos:

1. Primeiramente, a imagem é rotacionada para que os dois olhos fiquem na mesma linha horizontal;
2. A imagem é redimensionada para que a distância entre os olhos seja exatamente de 100 pixels;
3. A imagem é recortada para que a distância vertical entre a linha dos olhos e o topo da imagem seja de 100 pixels e a imagem resultante tenha o tamanho de 300×250 pixels;

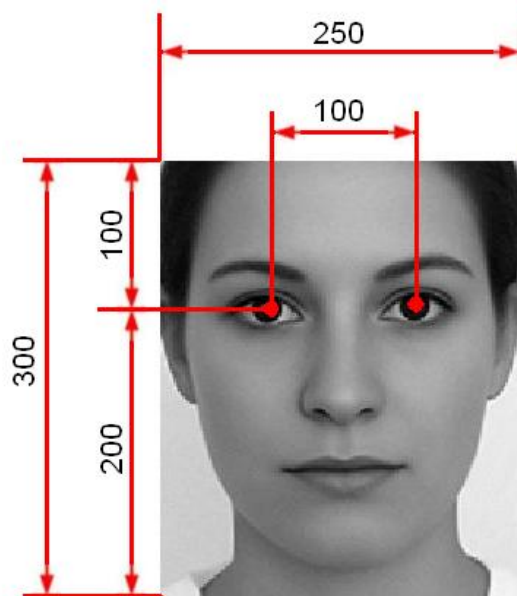


Fig. 6.2: Normalização da face.

6.1.2 Cálculo das medidas dos segmentos de cada parâmetro

Após a normalização descrita na seção anterior, 4 pontos de referência na imagem precisam ser especificados pelo usuário, esses pontos correspondem aos cantos externos dos olhos e cantos externos da boca (pontos 1, 2, 3 e 4 sobre a figura 6.3(a)). Esses pontos servem de referência para desenhar as linhas verticais ilustradas que passam por esses pontos. Essas linhas servem como referência para

que o usuário possa demarcar os 24 pontos definidos no capítulo 5 e ilustrados na figura 6.3(b). Além das 4 linhas ilustradas na figura, existem duas linhas que passam pelo centro das pupilas e uma linha que corta a imagem ao meio que possuem o mesmo propósito.

Portanto, observando a figura 6.3 a linha que passa pelo ponto 1, serve para marcar os pontos G e T; a linha que passa pelo ponto 2 serve para marcar o ponto V; a linha que passa pelo ponto 3 serve para marcar o ponto X; a linha que passa pelo ponto 4 serve para marcar os pontos J e Z; a linha que passa pela pupila direita serve para marcar os pontos D, H, A, K e U; a linha que passa pela pupila esquerda serve para marcar os pontos F, I, B, L e Y; e a linha que corta a face ao meio, marca os pontos E, C, O, N, S e W. Os pontos M e Q são marcados sobre as extremidades das narinas direita e esquerda.

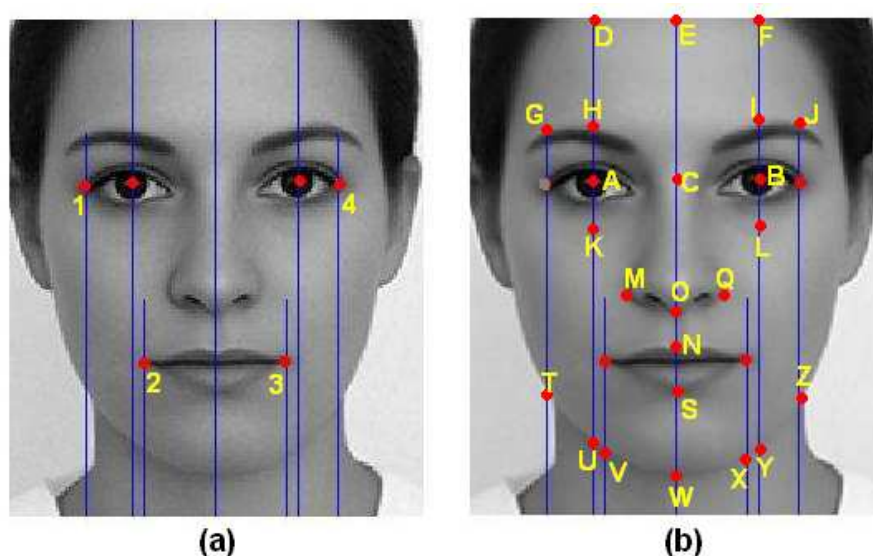


Fig. 6.3: Linhas verticais sobre a face.

Então, tendo esses 24 pontos definidos sobre a imagem, eles representam o conjunto de pontos de controle inicial P utilizados na equação 3.14 do capítulo 3. Esse conjunto de pontos serve para modelar os segmentos na face que irão representar parâmetros para avaliar e quantificar o envelhecimento facial. Esses parâmetros são mostrados na tabela 5.1 do capítulo 5. O cálculo do tamanho de cada segmento é representado pela distância existente entre os pontos que representam esse segmento, conforme foi visto na tabela 5.1. Tendo o tamanho de cada segmento, é possível aplicar as curvas de envelhecimento propostas no capítulo 5 para aumentar ou diminuir cada segmento conforme a idade que se quer representar o envelhecimento.

Por exemplo, considera-se a medida realizada para a altura do nariz. O segmento que representa a altura do nariz é composto pela distância dos pontos 'C' e 'O' na imagem. Na figura 6.3, os pontos 'C' e 'O' possuem as coordenadas (125,99) e (125,184), respectivamente. Portanto, o tamanho desse segmento é de 85 pixels, calculado pela distância euclidiana entre esses pontos. Assim como para a medida do nariz, outras 19 medidas de segmentos são analisadas.

6.1.3 Cálculo das novas medidas dos segmentos e pontos finais

As novas medidas dos segmentos de cada parâmetro são calculadas a partir das curvas de envelhecimento representadas por polinômios de segundo grau conforme visto no capítulo 5. Observando-se a figura 6.1 nota-se que os parâmetros de entrada para este módulo do protótipo são: a idade inicial e final; as medidas dos segmentos calculadas na seção anterior; e as curvas de envelhecimento. Então os segmentos são modificados podendo ter seu tamanho alterado conforme as curvas propostas. Os polinômios que representam as curvas fornecem a proporção que deve ser adicionada ou retirada do segmento. Esses polinômios foram calculados com os coeficientes descritos na tabela 6.1 e a fórmula matemática deles está na equação 6.1 já apresentada no capítulo 5 - equação 5.5, onde t_c é a idade de inflexão.

$$p(t) = \begin{cases} a_1 t^2 + b_1 t + c_1, & t \leq t_c \\ -a_2 t^2 + (b_2 + 4a_2 t_c)t + c_2 - 2a_2 t_c^2, & t > t_c \end{cases} \quad (6.1)$$

Nome	$a_1(10)^{-4}$	$b_1(10)^{-2}$	c_1	$a_2(10)^{-4}$	$b_2(10)^{-2}$	c_2	t_c
Altura da testa lateral	0.5585	-0.1108	1.0478	-0.5585	0.8945	0.8216	45
Altura da testa central	0.0866	0.1605	0.9658	0.0866	0.31638	0.9307	45
Altura da Sobrancelha Lateral	0.4354	-0.3122	1.0244	-	-	-	20
Altura da Sobrancelha Central	0.6231	-0.2981	0.9711	-	-	-	20
Altura da Bolsa Palpebral	-2.6894	-1.3922	0.7026	-	-	-	20
Altura do nariz	0.0540	0.7004	0.8337	-	-	-	20
Largura do nariz	1.3997	-0.6810	1.0658	-1.3997	-0.6558	0.4989	45
Espessura do lábio	-0.0262	-0.7288	1.0236	-	-	-	20
Altura do lábio	-0.6791	0.4747	0.9073	-	-	-	20
Mentoneana-Canto olho	-0.7992	0.4962	0.9391	0.7992	-0.4604	1.0825	30
Mentoneana-Centro olho	-0.2055	0.2524	0.9560	0.2055	0.0058	0.9929	30
Mentoneana-Canto boca	0.2842	0.1291	0.9483	0.2055	0.0058	0.8971	30
Mentoneana-Centro face	0.2566	0.0994	0.9814	-0.2566	0.4073	0.9352	30

Tab. 6.1: Coeficientes dos polinômios para cada segmento da face.

Assim, com o cálculo dos novos segmentos obteve-se o conjunto F de pontos finais de controle. Dependendo de cada segmento, o ponto inicial e/ou ponto final podem ser deslocados na imagem tendo suas coordenadas alteradas. Por exemplo, continuando o exemplo da seção anterior, a medida obtida para a altura do nariz foi de 85 pixels na vertical, se a idade informada para o envelhecimento for de 60 anos, a idade inicial for igual a 20 anos, pelas curvas de envelhecimento para o nariz, a nova distância será 107 pixels, ou seja, o ponto 'O' deverá ser deslocado 22 pixels na vertical de onde ele se encontra. Então, assim como para o nariz, todas as outras medidas tem o seu valor recalculado. Na tabela 6.2 são mostrados todos os pontos do exemplo da figura 6.3 recalculados.

Ponto	Inicial	Final	Calculado
A	(75,101)	(75,124)	(75,116)
B	(175,99)	(175,122)	(175,114)
C	(125,99)	(125,99)	(125,99)
D	(75,2)	(75,-8)	(75,-3)
E	(125,2)	(125,-6)	(125,-2)
F	(175, 2)	(175,-8)	(175,-3)
G	(45,69)	(45,77)	(45,76)
H	(75,69)	(75,77)	(75,77)
I	(175,67)	(175,175)	(175,175)
J	(203,68)	(203,76)	(203,75)
K	(74,126)	(74,139)	(74,138)
L	(175,126)	(175,139)	(175,137)
M	(94, 173)	(92, 173)	(93,178)
N	(125,205)	(125,225)	(125,221)
O	(125,184)	(125,206)	(125,199)
Q	(152,171)	(157,171)	(155,177)
S	(125,233)	(125,240)	(125,243)
T	(45,232)	(45,243)	(45,243)
U	(75,266)	(75,280)	(75,279)
V	(84,270)	(84,284)	(84,284)
W	(125,284)	(125,300)	(125,298)
X	(166,272)	(166,286)	(166,286)
Y	(175,264)	(175,278)	(175,277)
Z	(203,235)	(203,246)	(203,246)

Tab. 6.2: Pontos calculados - Figura 6.3.

6.1.4 Cálculo dos coeficientes da função RBF

Os coeficientes apresentados aqui são aqueles mencionados no capítulo 3, seção 3.2.2. São os coeficientes do método que utiliza as funções de base radial e são usados para realizar a transformação que deforma a imagem inicial. O cálculo desses coeficientes começa com a inicialização da matriz G apresentada no capítulo 3 e reescrita na equação 6.2.

$$G = \begin{bmatrix} g(r_{11}) & g(r_{12}) & \cdots & g(r_{1n}) \\ g(r_{21}) & g(r_{22}) & \cdots & g(r_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ g(r_{n1}) & g(r_{n2}) & \cdots & g(r_{nn}) \end{bmatrix} \quad (6.2)$$

Cada componente $g(r_{ij})$ dessa matriz é calculado utilizando a distância euclidiana entre os pontos de controle iniciais i e j ($i, j = 1, \dots, n$) e, aplicando essa distância na função base escolhida $g(r)$. A função base (seção 3.2.3) pode ser: gaussiana, multiquádrica, *shifted-log*, *thin plate spline*, etc.

A matriz G tem a dimensão $n \times n$ onde n é o número de pontos de controle que no caso é igual a 24. Pode-se notar que o (i, j) -ésimo elemento de G é $g(r_{ij})$ que mede o efeito do j -ésimo ponto de controle na transformação do i -ésimo ponto de controle.

Então, tendo os conjuntos de pontos de controle inicial P e final F , e a matriz G definida, pode-se resolver o sistema linear abaixo, que terá como resultado os coeficientes desejados $\begin{bmatrix} a \\ \alpha \end{bmatrix}$ com dimensão $(N + 3)$.

$$\begin{bmatrix} (G + \lambda * I) & P \\ P^T & 0 \end{bmatrix} \begin{bmatrix} a \\ \alpha \end{bmatrix} = \begin{bmatrix} F \\ 0 \end{bmatrix} \quad (6.3)$$

Neste sistema, na matriz G , todos os elementos da diagonal principal possuem o valor 1 devido ao fato de $e^0 = 1$ (gaussiana). O valor 0 é por expressar o cálculo da distância do ponto de controle em relação a ele mesmo. Usa-se o parâmetro λ multiplicado pela matriz identidade I acrescentado à G , conforme 3.2, para controlar o mapeamento realizado pela função. Quando

- $\lambda \rightarrow 0$: ocorre uma interpolação, ou seja, cada ponto é mapeado exatamente para a posição que foi estipulada.
- $0 < \lambda < \infty$: ocorre uma aproximação, ou seja, o ponto calculado está próximo do ponto de controle final.
- $\lambda \rightarrow \infty$: ocorre uma transformação afim, ou seja, é como se a parte radial da função não existisse e o interpolante fosse um plano.

Outro parâmetro que permite controlar o mapeamento é a constante σ da função gaussiana. Esta função foi escolhida por permitir um ajuste desejável do seu comportamento e, por permitir transformações locais ao contrário da função *Thin Plate Spline* que possui um comportamento global e não tem nenhum parâmetro que permita ajustar seu comportamento. Quanto maior for esse parâmetro, maior será o tamanho das gaussianas e maior será a área de influência dessas, reduzindo a localidade da transformação.

Para ilustrar o comportamento de tal parâmetro junto com a constante λ , observa-se a figura 6.4. Nesta figura, simplesmente usou-se o conjunto de pontos de controle inicial adicionando um deslocamento de 30 pixels para o lado direito e para baixo como conjunto de pontos de controle final. A figura (a) é a imagem original e as outras figuras são imagens deformadas. Observa-se nas figuras que conforme σ aumenta, a área de influência da transformação local aumenta também e, quando λ é igual a zero, o mapeamento é realizado conforme indica os pontos de controle finais. As figuras (b), (c) e (d) possuem pontos fixos nos cantos da imagem, enquanto a figura (e) não possui, o que fez com que ocorresse uma translação da imagem devido a todos os pontos terem sido deslocados por uma mesma distância.

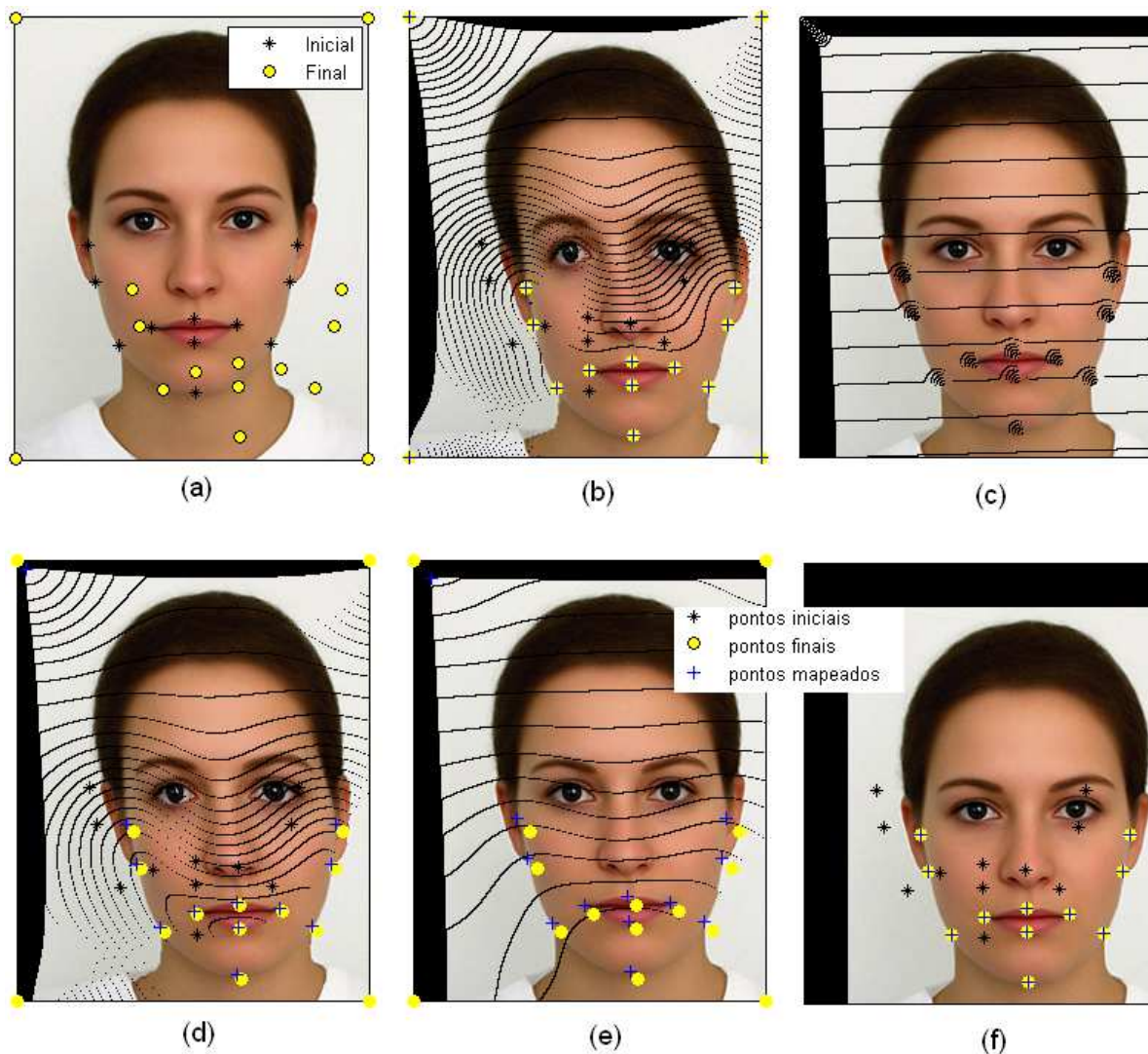


Fig. 6.4: Comportamento da RBF alterando suas constantes.

Os pontos finais foram definidos sendo os pontos iniciais deslocados 30 pixels para a direita. (a) a imagem original com a localização dos pontos de controle; (b) $\lambda = 0$ e $\sigma = 45$; (c) $\lambda = 0$ e $\sigma = 4$; (d) $\lambda = 1$ e $\sigma = 45$; (e) $\lambda = 1$ e $\sigma = 45$.

6.1.5 Transformação RBF

A transformação RBF tem o propósito de realizar um mapeamento entre a imagem inicial e imagem final utilizando os pontos de controle. Esse mapeamento permite que ocorram deformações na imagem inicial para que a face seja envelhecida. Após os coeficientes serem calculados, eles são utilizados nessa transformação para realizar o mapeamento de cada pixel da imagem inicial conforme especificado pelos pontos de controle finais, que foram definidos pelas curvas de envelhecimento, produzindo a imagem envelhecida como deseja-se. Sabendo que a transformação $T(\vec{p}_i)$ é realizada em cada dimensão por vez; que p_i é o i -ésimo ponto de controle inicial; u_i e v_i são as coordenadas x e y , respectivamente, do i -ésimo ponto de controle final, então:

$$T(\vec{p}_i) = [T_U(\vec{p}_i), T_V(\vec{p}_i)] \quad (6.4)$$

onde

$$T_U(\vec{p}_i) = u_i$$

$$T_V(\vec{p}_i) = v_i$$

Portanto, essa transformação permitirá que o mapeamento seja realizado. Cada pixel da imagem terá (ou não) suas coordenadas alteradas na imagem podendo ser mapeado fora das dimensões da imagem. Então, sendo (x, y) as coordenadas do pixel na imagem; $p_i = (x_i, y_i)$ é um ponto de controle inicial; $\alpha_1, \alpha_2, \alpha_3, a_i, \beta_1, \beta_2, \beta_3, b_i$ os coeficientes calculados pelo sistema linear, a transformação de cada dimensão é realizada pelas seguintes equações:

$$T_U(x, y) = \alpha_1 + \alpha_2 x + \alpha_3 y + \sum_{i=1}^N a_i g(r_i) \quad (6.5)$$

$$T_V(x, y) = \beta_1 + \beta_2 x + \beta_3 y + \sum_{i=1}^N b_i g(r_i) \quad (6.6)$$

onde $g(r_i)$ é a função base que neste caso, é representada pela função gaussiana aplicada sobre a distância euclidiana entre o pixel (x, y) e os pontos de controle iniciais:

$$r_i = \|\vec{x} - \vec{p}_i\| = \sqrt{\vec{x} - \vec{p}_i} = \sqrt{(x - x_i)^2 + (y - y_i)^2}$$

$T_U(x, y)$ representa a coordenada u e $T_V(x, y)$ representa a coordenada v da nova imagem, portanto, o pixel (x, y) agora será transferido na nova imagem para a coordenada (T_U, T_V) ou (u, v) .

Ao final do procedimento descrito muitos pixels da imagem inicial não são mapeados na imagem final, aparecendo pontos pretos nesta. Isso acontece porque alguns pixels da imagem inicial não são mapeados na imagem final. Para resolver esse problema pode ser realizada uma interpolação adicional na imagem final. A interpolação de imagens consiste em adicionar novos pontos à ela, baseado nos pontos existentes, aumentando a quantidade total de pontos. Se em uma imagem, tem-se um ponto verde tonalidade 100, e outro verde tonalidade 20. A interpolação, neste caso, simplesmente iria calcular a média e incluir entre os dois um ponto verde tonalidade 60. Com isso, tem-se a imagem final envelhecida. Um procedimento mais preciso seria a utilização de um mapeamento inverso com interpolação linear, ou seja, buscando-se para cada pixel da imagem final um correspondente na imagem inicial, o que no presente caso não alteraria substancialmente a aparência final.

6.2 Resultados

Nesta seção serão apresentados os resultados obtidos implementando as funções de base radial no processo de deformação da imagem para obter um envelhecimento desta.

Antes de iniciar a simulação de envelhecimento, era preciso ter imagens de trabalho, então, foram buscadas na internet imagens de faces onde a pessoa encontrava-se de frente para a fotografia. A imagem da face de uma mulher que foi utilizada em algumas imagens no decorrer desta dissertação e que foi utilizada nas seguintes simulações, é uma imagem criada por um trabalho realizado na Alemanha que corresponde a um protótipo de beleza. Esse trabalho pode ser visto em [Beautycheck (2006)]. Imagem utilizada nos casos de estudo 1,2,3 e 4.

Buscando-se na internet bancos de dados de pessoas envelhecidas, somente um resultado foi encontrado [FGNET (2006)]. Esse banco de dados é composto de fotografias de pessoas em diversas idades, desde bebê até idoso, entretanto são fotos onde não existe um padrão, cada foto foi tirada de um ângulo diferente, com as pessoas em poses diferentes. Isso dificulta a análise dos resultados obtidos, pois nem sempre todos os pontos podem ser localizados na face, além de fatores como franja na testa e resolução da imagem que atrapalham. Esse banco de dados é composto por 1002 imagens de 82 pessoas, entre homens e mulheres. O número de imagens por pessoa varia entre 6 e 18 imagens, e a idade varia de 0 a 69 anos. Existem tanto imagens coloridas como em níveis de cinza e a resolução em média das fotos é de 400×500 pixels. Elas estão no formato JPEG e um exemplo de imagens desse banco de dados pode ser visto na figura 6.5, é um exemplo de fotos de uma pessoa em diversas idades, fotos essas, recolhidas do banco de dados. A ordem de progresso de idade na figura é: 03, 15, 18, 30, 35, 38, 41, 45, 47 e 54 anos. Este banco de dados foi utilizado nos casos de estudo 5 e 6.



Fig. 6.5: Exemplo de imagens do banco de dados FGNET.

Como esta seção tem a intenção de mostrar os testes realizados com a implementação do método de metamorfose por funções de base radial explorando as funções de envelhecimento descritas anteriormente, a partir de agora serão mostradas figuras simulando o processo de envelhecimento. Cabe salientar que o limite de idade para o envelhecimento foi fixado entre 20 e 70 anos e, as figuras testadas estão em preto e branco devido ao fato de que somente a localidade geométrica dos pontos será analisada.

Caso de estudo 1

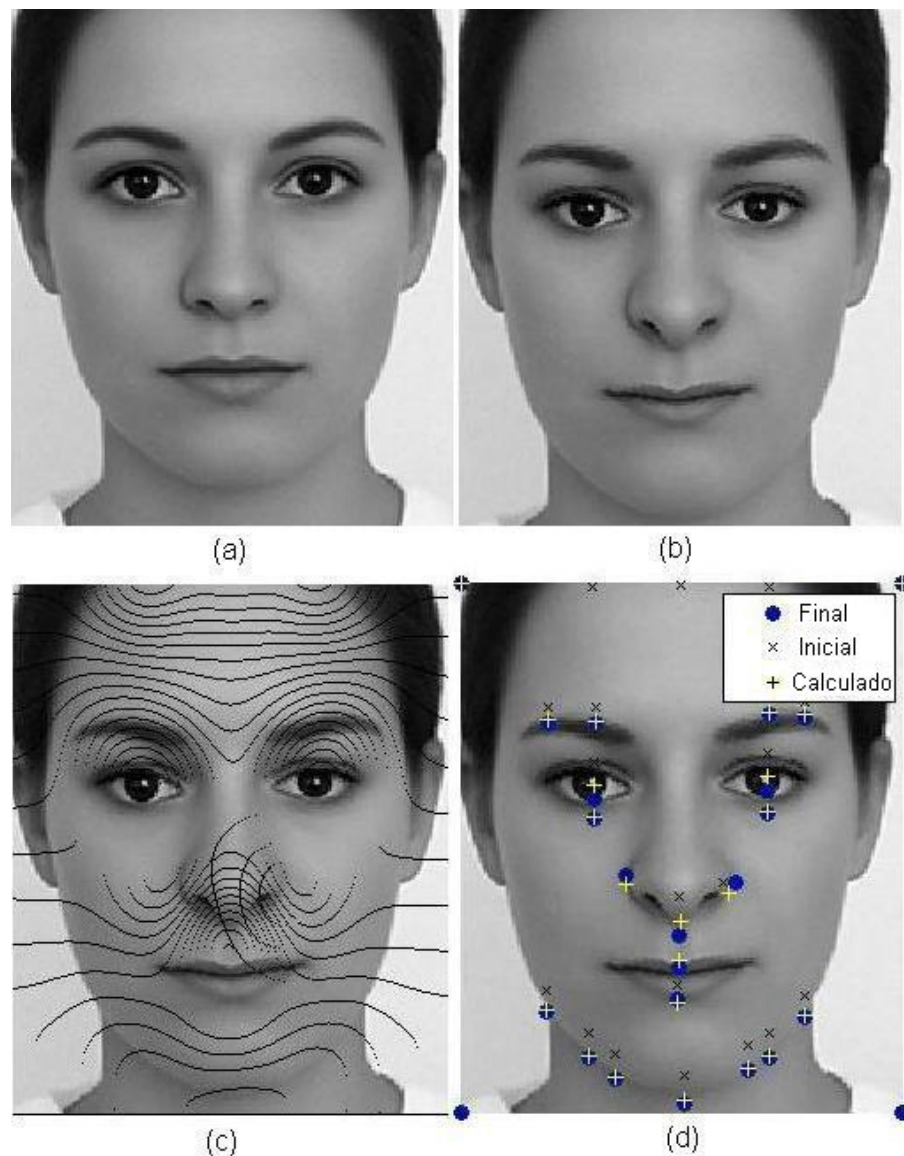


Fig. 6.6: Simulação de envelhecimento de 20 para 60 anos ($\sigma = 30, \lambda = 1$).

A figura 6.6 mostra um envelhecimento de uma face com 20 anos para uma idade de 60 anos com as constantes $\sigma = 30$ e $\lambda = 1$. Na 6.6(b) tem-se a imagem envelhecida, somente alterando a geometria da face. Nota-se um aumento considerável no nariz, um afinamento nos lábios, na região dos olhos ocorre um aumento da bolsa palpebral. A imagem (c) ilustra a imagem envelhecida sem a interpolação adicional (seção 6.1.5) e a imagem (d) mostra os pontos de controle iniciais, os pontos de controle finais estabelecidos pelas curvas e os pontos mapeados pelas funções de base radial, conforme a legenda específica.

Caso de estudo 2

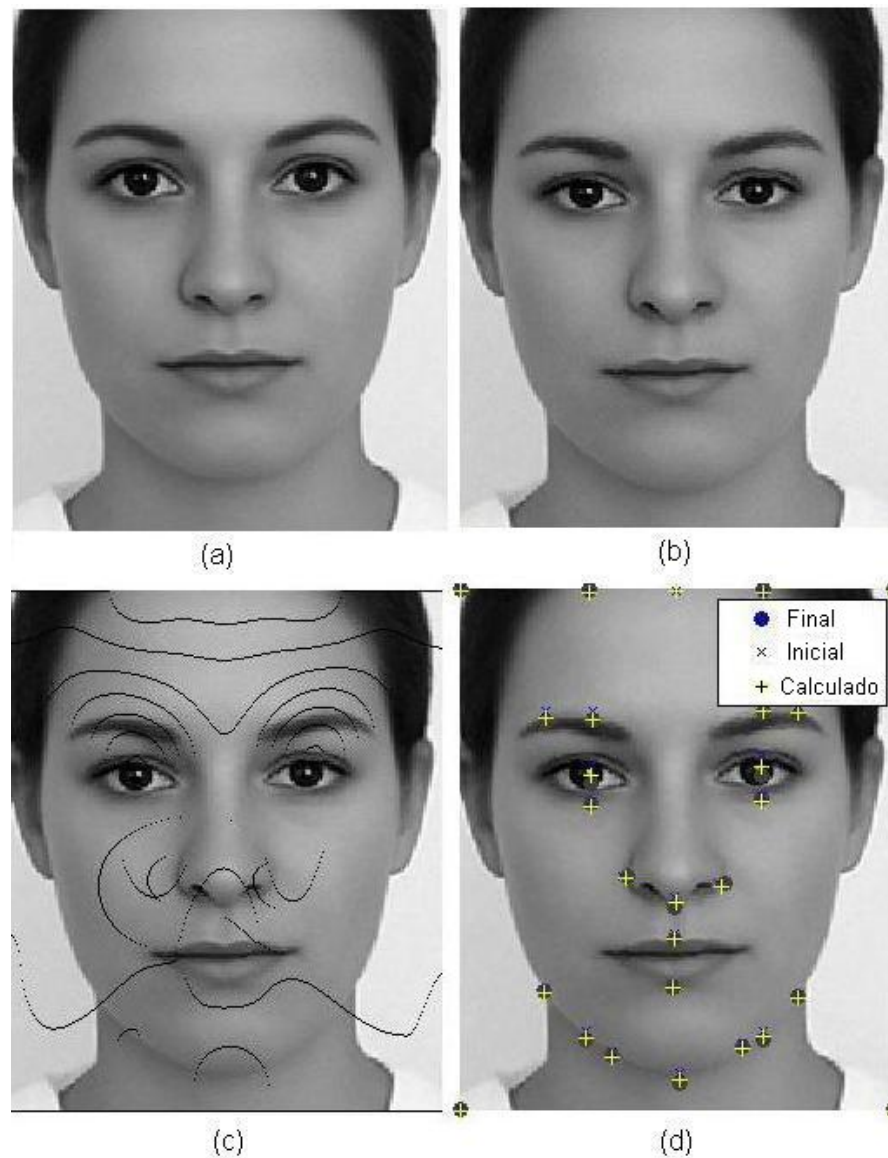


Fig. 6.7: Simulação de envelhecimento de 20 para 30 anos ($\sigma = 30, \lambda = 1$).

A figura 6.7 possui os mesmos parâmetros da figura 6.6 entretanto o envelhecimento foi feito em um intervalo somente de 10 anos, ou seja, a face original (a) tem 20 anos e a face envelhecida (b) tem 30 anos. As constantes utilizadas foram as mesmas usadas anteriormente: $\sigma = 30$ e $\lambda = 1$. A imagem (c) ilustra a face envelhecida sem a interpolação adicional e, (d) é a face envelhecida ilustrando os pontos iniciais, pontos finais e pontos mapeados.

Caso de estudo 3

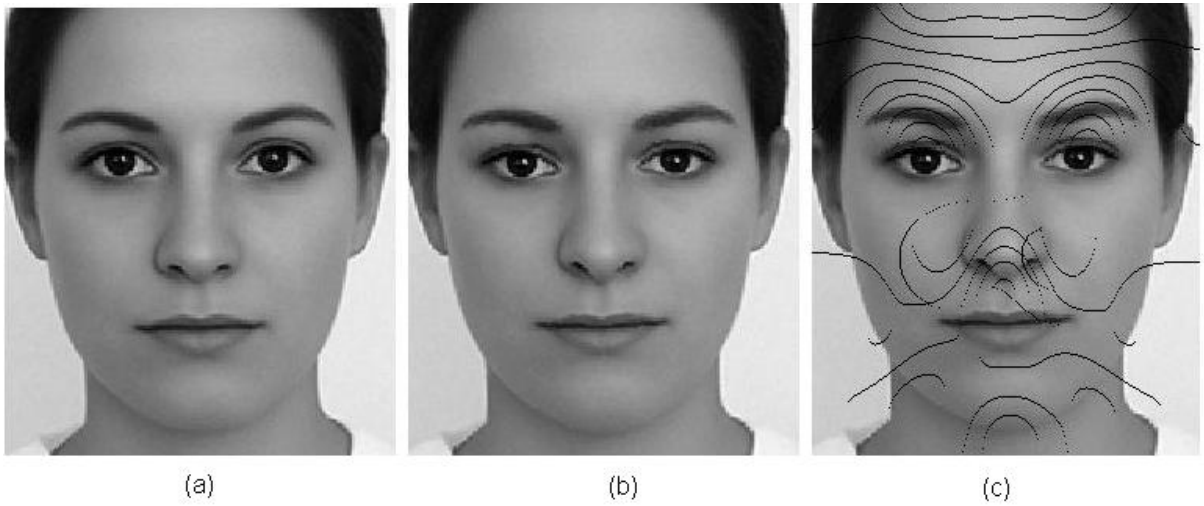


Fig. 6.8: Simulação de envelhecimento de 20 para 40 anos ($\sigma = 30, \lambda = 1$).

A figura 6.8 mostra um envelhecimento para (a) de 20 anos para uma idade de 40 anos com os parâmetros $\sigma = 30, \lambda = 1$. A imagem (b) é a face envelhecida com interpolação e a imagem (c) é a face envelhecida sem interpolação.

Caso de estudo 4

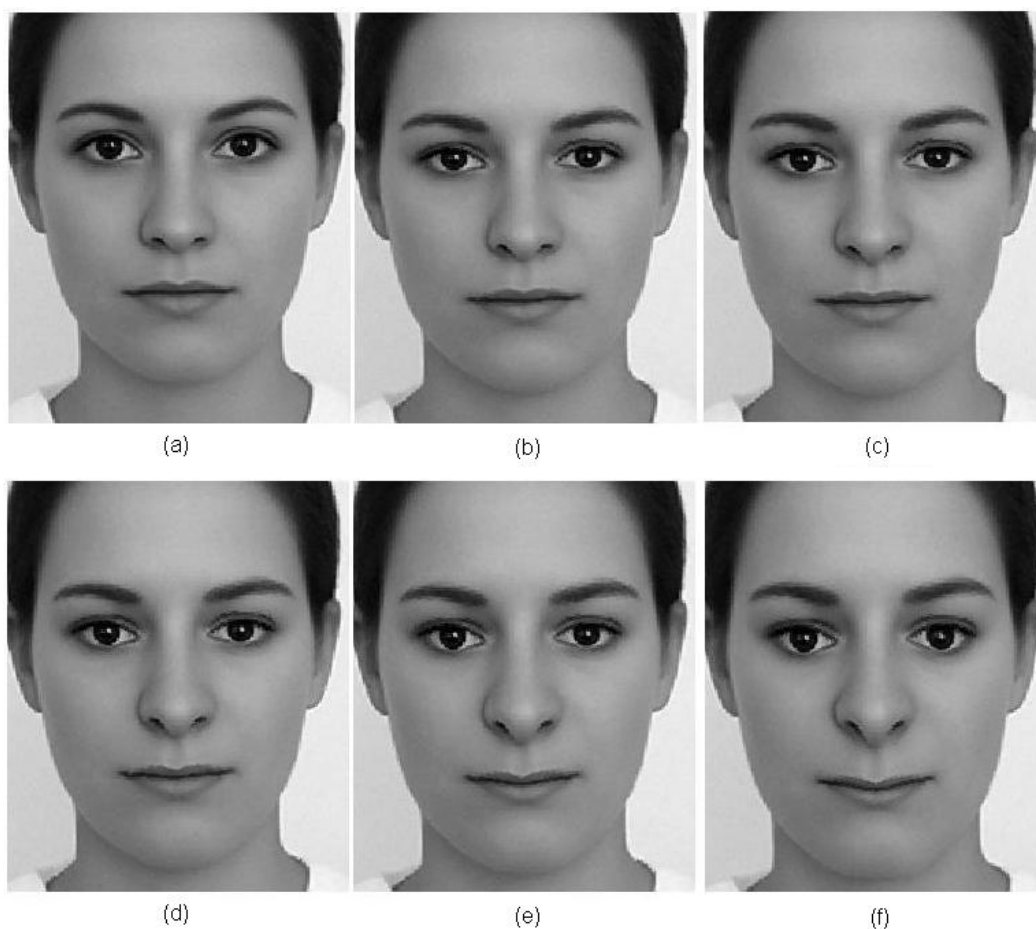


Fig. 6.9: Simulação de envelhecimento de 10 em 10 anos ($\sigma = 30, \lambda = 1$).

A figura 6.9 mostra a face sendo envelhecida de 10 em 10 anos. Convém salientar que a face (a) possui 20 anos e é a imagem original, as outras são faces simuladas, a face (b) tem 30 anos, a face (c) 40 anos, a face (d) 50 anos, a face (e) 60 anos e a face (f) 70 anos. Os valores das constantes usadas neste envelhecimento foram: $\sigma = 30$ e $\lambda = 1$.

Caso de estudo 5

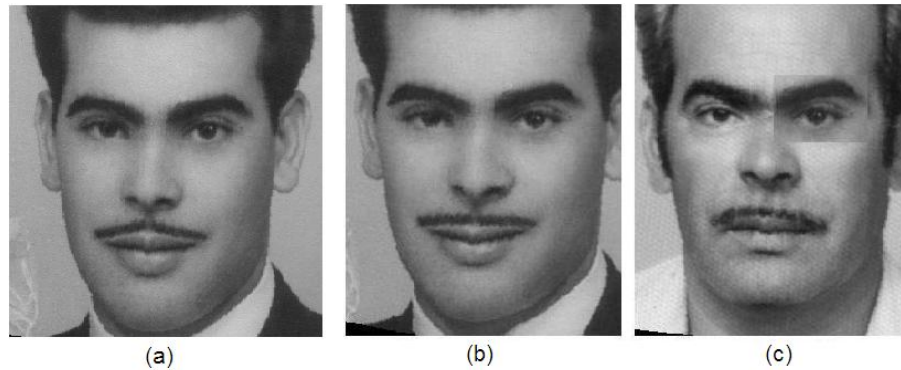


Fig. 6.10: Simulação de comparação e comportamento da RBF. ($\sigma = 30, \lambda = 1$)

A figura 6.10 ilustra três imagens da mesma pessoa, porém em idades diferentes (24, simulação de 42 anos e 42 anos, respectivamente) retiradas do banco de imagens FGNET (2006). A imagem (a) mostra o rosto de 24 anos, a imagem (b) mostra o rosto de 24 anos envelhecido para a idade de 42 anos, (b) ilustra a foto real da pessoa com 42 anos, porém com a sobreposição do olho esquerdo retirada da figura (b) afim de comparação com o envelhecimento.

Caso de estudo 6

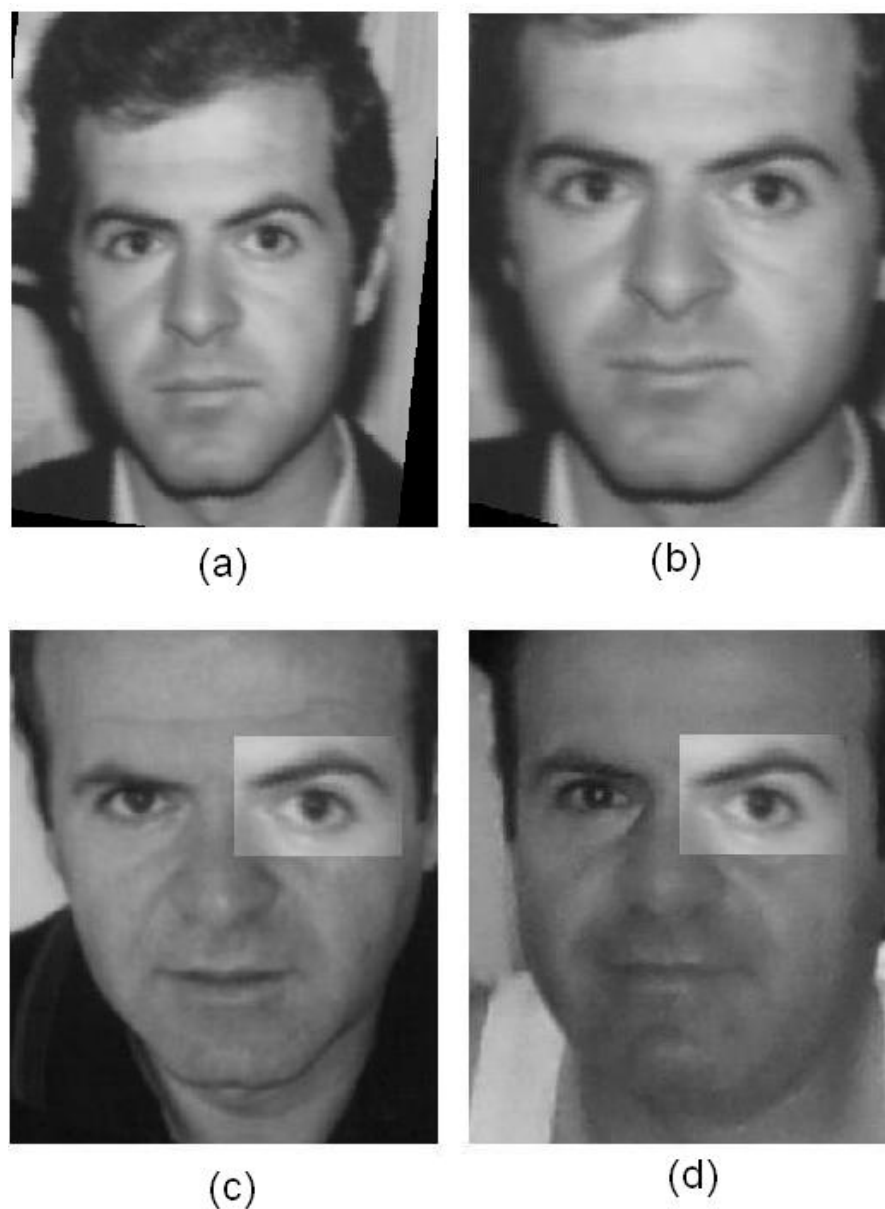


Fig. 6.11: Simulação de envelhecimento de 22 para 42 anos. ($\sigma = 30, \lambda = 1$)

A figura 6.11 ilustra quatro imagens da mesma pessoa, retiradas do banco de imagens FGNET (2006). A figura (a) possui 22 anos, (c) possui 42 e (d) 45 anos (estas 3 figuras são fotos reais). A imagem (b) mostra o rosto (a) mapeado para a idade de 42 anos; (c) e (d) são fotos reais do envelhecimento com a sobreposição do olho esquerdo retirado de (b) para comparação.

Capítulo 7

Conclusões e Trabalhos Futuros

Representar e simular o processo de envelhecimento é um desafio. Portanto, a presente dissertação teve por desafio criar um protótipo que permitisse simular o envelhecimento facial de imagens frontais.

Devido ao caráter multidisciplinar do trabalho, para criar o protótipo proposto, foram envolvidos conceitos nas áreas de computação, processamento de imagens, medicina e matemática.

O trabalho iniciou com um estudo sobre *morphing* pesquisando como funcionam as principais técnicas existentes. A conceituação de transformações de *morphing* apresentada no capítulo 2 foi fundamental na compreensão do processo, necessária para qualquer tentativa de implementação. A importância desta conceituação manifesta-se mais evidentemente na escolha da técnica de *warping* por funções de base radial utilizadas para a criar o protótipo de envelhecimento de imagens faciais.

As funções de base radial mostraram sua adequação para a aplicação proposta devido ao fato de utilizarem pontos para especificar características nas imagens, e estes pontos não apresentarem restrições na sua localidade na imagem. Devido a isso, as funções de base radial apresentaram um comportamento facilmente controlável já que não existem restrições nas posições dos pontos de controle sobre as imagens. Esses tipos de funções também permitiram realizar deformações tanto locais como deformações globais. Além disso, uma vantagem muito interessante desses tipos de funções é que realizam deformações suavizadas.

Além das funções de base radial, o método de *morphing* chamado *Field Morphing* (seção 2.1.2) foi implementado para a aplicação proposta, mas não foram obtidos resultados convincentes. Acredita-se que estes resultados não satisfatórios resultam do fato dessa técnica utilizar segmentos de reta que possuem a restrição de não poder se cruzar, caso contrário, as imagens finais apresentam efeitos como ‘dobras’ na imagem.

O capítulo 4 permitiu avaliar o que já foi realizado na área de modelagem de envelhecimento de imagens faciais e a partir disso, extraiu-se uma metodologia para ser aplicada no presente trabalho. Certamente com a utilização do *morphing*, somente métodos que alteram a geometria da face poderiam ser utilizados. Entretanto, métodos de envelhecimento que não alteram também a textura da pele não produzem um efeito tão realista.

Foi discutido também, que para se ter bons sistemas de envelhecimento é necessário ter bancos de dados com imagens tiradas com um bom ângulo de captura e que possuam uma resolução adequada. Infelizmente, até a presente data, bancos ideais para essa proposta não existem.

Os resultados obtidos e mostrados no capítulo 6 mostraram melhor resultado em algumas regiões,

ou seja, é visível o envelhecimento das regiões do nariz e da boca, e menos realista nos olhos. Assim, analisando a face como um todo, nota-se que muitos aprimoramentos podem ser feitos, por exemplo aumentando o número de pontos na face e agregando-se textura no processo de simulação do envelhecimento.

O protótipo implementado em MATLAB, versão 7.0, foi um instrumento muito útil de estudo, permitindo uma primeira análise dos resultados da modelagem proposta em [Leta (1998)] aplicada à *morphing* por RBF. Uma próxima atividade seria a realização de um *software* autônomo com a mesma finalidade e com uma interface voltada ao usuário final.

Assim, o enfoque desta dissertação contemplando o uso de *morphing* por RBF com os dados gerados pela modelagem do envelhecimento facial [Leta (1998)] e simulação gráfica via MATLAB, representam um primeiro passo de pesquisa unindo estes conceitos. No item 7.1 destacam-se algumas sugestões de trabalhos futuros.

7.1 Trabalhos futuros

Extração automática das características da face: A definição das características faciais foi realizada manualmente. Foram marcados pontos e linhas nas imagens digitalizadas, conforme descrito no capítulo 5. Entretanto, a marcação dos pontos é um trabalho incômodo e impreciso. Métodos que permitissem a extração automática destes levariam menos tempo e ofereceriam uma maior precisão na localidade do ponto a ser marcado.

Implementação de textura: Simulação de envelhecimento facial somente alterando a geometria da face não apresenta resultados visuais finais realistas, embora sejam úteis em diversas áreas, como por exemplo, planejamento cirúrgico. A textura aplicada a imagem representando rugas e manchas na pele, cabelos brancos, pode promover um realismo maior à essa simulação do envelhecimento facial.

Implementação do sistema autônomo: realização de um *software* autônomo com uma interface voltada ao usuário final.

Bibliografia

ARAD, N. et al. Image warping by radial basis functions: applications to facial expressions. *CVGIP: Graph. Models Image Process.*, Academic Press, Inc., Orlando, FL, USA, v. 56, n. 2, p. 161–172, 1994. ISSN 1049-9652.

ARAD, N.; REISFELD, D. Image warping using few anchor points and radial functions. *Computer Graphics Forum*, v. 14, n. 1, p. 35–46, 1995.

BEAUTYCHECK. *Beautycheck*. 2006. Disponível em: <<http://www.beautycheck.de>>.

BEIER, T.; NEELY, S. Feature-based image metamorphosis. In: *SIGGRAPH '92: Proceedings of the 19th annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM Press, 1992. v. 26, n. 2, p. 35–42. ISSN 0097-8930.

BREGLER, C.; COVELL, M.; SLANEY, M. Video rewrite: driving visual speech with audio. In: *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM Press/Addison-Wesley Publishing Co., 1997. p. 353–360. ISBN 0-89791-896-7.

CRIMELAB. *Criminal Minds and Methods Age Progression*. 2006. Disponível em: <<http://www.crimelibrary.com>>.

DUCHON, J. Splines minimizing rotation-invariant semi-norms in sobolev spaces. Birkhauser, Basel, 1979.

EZZAT, T.; POGGIO, T. Miketalk: A talking facial display based on morphing visemes. IEEE Computer Society, Los Alamitos, CA, USA, p. 96, 1998. ISSN 1087-4844.

FARUQUIE, T. A. et al. Audio driven facial animation for audio-visual reality. IEEE Computer Society, Los Alamitos, CA, USA, p. 209, 2001.

FGNET. *Face Gesture Recognition Research Network Aging Database*. 2006. Disponível em: <<http://www.fgnet.rsunit.com/>>.

GANDHI, M. R. Ph.D. Thesis, *A Method for Automatic Synthesis of Aged Human Facial Images*. McGill University, Montreal, Canadá: [s.n.], August 2004.

HARTLEY, R. I. In defence of the 8-point algorithm. In: *ICCV '95: Proceedings of the Fifth International Conference on Computer Vision*. Washington, DC, USA: IEEE Computer Society, 1995. p. 1064. ISBN 0-8186-7042-8.

- HAYASHI, J. et al. Age and gender estimation from facial image processing. *Proceedings of the 41st SICE Annual Conference*, v. 01, p. 13–18, 2002.
- HORNG, W.; LEE, C.; CHEN, C. Classification of age groups based on facial features. *Tamkang Journal of Science and Engineering*, v. 4, n. 3, p. 183–192, 2001.
- HUSSEIN, H. K. Towards realistic facial modeling and re-rendering of human skin aging animation. IEEE Computer Society, Los Alamitos, CA, USA, p. 205, 2002.
- KASS, A. W. M.; TERZOPOULOS. Snakes: active contour models. *Int. Journal Computer Vision*, v. 1, n. 4, p. 321–331, 1988.
- KOCH, R. M. et al. Simulating facial surgery using finite element models. In: *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM, 1996. p. 421–428. ISBN 0-89791-746-4.
- KWON, Y. H.; LOBO, N. da V. Age classification from facial images. *Comput. Vis. Image Underst.*, Elsevier Science Inc., New York, NY, USA, v. 74, n. 1, p. 1–21, 1999. ISSN 1077-3142.
- LANITIS, A. On the significance of different facial parts for automatic age estimation. *14th International Conference on Digital Signal Processing*, v. 2, p. 1027–1030, 2002.
- LANITIS, A.; TAYLOR, C.; COOTES, T. Toward automatic simulation of aging effects on face images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Los Alamitos, CA, USA, v. 24, n. 4, p. 442–455, 2002. ISSN 0162-8828.
- LEE, S. et al. Image metamorphosis with scattered feature constraints. *IEEE Transactions on Visualization and Computer Graphics*, v. 2, n. 4, p. 337–354, 1996.
- LEE, S.; WOLBERG, G.; SHIN, S. Y. Polymorph: Morphing among multiple images. *IEEE Comput. Graph. Appl.*, IEEE Computer Society Press, Los Alamitos, CA, USA, v. 18, n. 1, p. 58–71, 1998. ISSN 0272-1716.
- LEE, S.-Y. et al. Image morphing using deformable surfaces. *Proceedings of Computer Animation 94*, p. 31–39, 1994.
- LEE, S.-Y. et al. Image morphing using deformation techniques. *The Journal of Visualization and Computer Animation*, v. 7, n. 1, p. 3–24, 1996.
- LEE, S.-Y.; CHWA, K.-Y.; SHIN, S. Y. Image metamorphosis using snakes and free-form deformations. *Computer Graphics*, v. 29, p. 439–448, 1995.
- LETA, F. R. Ph.D. Thesis, *Modelagem Matemática e Simulação Gráfica do Envelhecimento Facial*. PUC Rio, Rio de Janeiro, Brasil: [s.n.], 1998.
- LETA, F. R. et al. Manipulating facial appearance through age parameters. *Anais do IX SIBGRAP*, 1996.

- LITWINOWICZ, P.; WILLIAMS, L. Animating images with drawings. In: *SIGGRAPH '94: Proceedings of the 21st annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM Press, 1994. p. 409–412. ISBN 0-89791-667-0.
- PESSA, J. et al. Concertina effect and facial aging: nonlinear aspects of youthfulness and skeletal remodeling, and why, perhaps, infants have jowls. *Plastic and Reconstructive Surgery*, p. 635–644, February 1999.
- PITANGUY, I. et al. Anatomia do envelhecimento da face. *Bulletin of Plastic Surgery No 40*, IEEE Computer Society, Los Alamitos, CA, USA, v. 67, p. 385–390, 1977.
- PSILLAKIS, Y.; RUMELEY, T.; CAMARGOS, A. Subperiosteal approach as an improved concept for correction of the aging face. *Plastic and Reconstructive Surgery*, v. 82, p. 383–394, 1988.
- ROWLAND, D. A.; PERRETT, D. I. Manipulating facial appearance through shape and color. *IEEE Computer Graphics and Applications*, v. 15, n. 5, p. 70–76, 1995.
- RUPRECHT, D.; MÜLLER, H. Free form deformation with scattered data interpolation methods. In: FARIN, G.; HAGEN, H.; NOLTEMEIER, H. (Ed.). *Geometric Modelling (Computing Suppl. 8)*. Wien: Springer Verlag, 1993. p. 267–281. Disponível em: <citeseer.ist.psu.edu/ruptrecht93free.html>.
- RUPRECHT, D.; MÜLLER, H. Deformed cross-dissolves for image interpolation in scientific visualization. *The Journal of Visualization and Computer Animation*, v. 5, n. 3, p. 167–181, – 1994.
- RUPRECHT, D.; MÜLLER, H. Image warping with scattered data interpolation. *IEEE Comput. Graph. Appl.*, IEEE Computer Society Press, Los Alamitos, CA, USA, v. 15, n. 2, p. 37–43, 1995. ISSN 0272-1716.
- SEDERBERG, T. W.; PARRY, S. R. Free-form deformation of solid geometric models. In: *SIGGRAPH'86*. New York, NY, USA: ACM Press, 1986. v. 20, n. 4, p. 151–160. ISSN 0097-8930.
- SEITZ, S. M.; DYER, C. R. View morphing. In: *SIGGRAPH '96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*. New York, NY, USA: ACM, 1996. p. 21–30. ISBN 0-89791-746-4.
- SHAN, Y.; LIU, Z.; ZHANG, Z. Image-based surface detail transfer. In: *Proceedings of Computer Vision and Pattern Recognition*. Los Alamitos, CA, USA: [s.n.], 2001. v. 02, p. 794. ISSN 1063-6919.
- SHIRAKABE, Y. The oriental aging face: An evaluation of a decade of experience with the triangular smas flap technique in facelifting. *Aesthetic Plastic Surgery*, v. 12, p. 25–32, September 1988.
- TIDDEMAN, B.; BURT, M.; PERRETT, D. Prototyping and transforming facial textures for perception research. *IEEE Comput. Graph. Appl.*, IEEE Computer Society Press, Los Alamitos, CA, USA, v. 21, n. 5, p. 42–50, 2001. ISSN 0272-1716.

TURK M.A.; PENTLAND, A. Face recognition using eigenfaces. In: *Proceedings of Computer Vision and Pattern Recognition*. [S.l.: s.n.], 1991. p. 586–591.

WIRTH, M. Ph.D. Thesis, *A Nonrigid Approach to Medical Image Registration: Matching Images of the Breast*. RMIT University: [s.n.], 2000.

WOLBERG, G. *Digital Image Warping*. Los Alamitos, CA, USA: IEEE Computer Society Press, 1994. ISBN 0818689447.

WOLBERG, G. Image morphing: a survey. *The Visual Computer*, v. 14, n. 8/9, p. 360–372, 1998.