

BRENO AZEVEDO DA SILVA

METAMORFOSE DE IMAGENS COM
ALINHAMENTO AUTOMÁTICO DE RECURSOS
FACIAIS E INTERPOLAÇÃO MORFOLÓGICA

Dissertação de Mestrado apresentada ao
Programa de Pós-Graduação em Informática Aplicada
da Pontifícia Universidade Católica do Paraná como
requisito parcial para obtenção do título de
Mestre em Informática

Linha de Pesquisa: Visão Computacional
Orientador: Prof. Dr. Jacques Facon

Curitiba - PR
2016

SILVA, Breno Azevedo da, *Metamorfose de Imagens com Alinhamento Automático de Recursos Faciais e Interpolação Morfológica*. Curitiba, 2016. 95p.

Dissertação de Mestrado – Pontifícia Universidade Católica do Paraná.
Programa de Pós-Graduação em Informática Aplicada.

1. Visão Computacional 2. Metamorfose 3. Morfologia Matemática 4. Envelhecimento Digital. I. Pontifícia Universidade Católica do Paraná. Centro de Ciências Exatas e de Tecnologia. Programa de Pós-Graduação em Informática II-t

Página deixada em branco intencionalmente

Agradecimentos

A minha mãe, com carinho

Não poderia começar estes agradecimentos sem dedicar este trabalho a minha mãe, que sempre prezou pelos meus estudos e nunca mediu esforços para proporcioná-los. Mãe, você foi a definição de um anjo do Criador na minha vida e na dos meus irmãos. Agradeço a Deus ainda, não só pela família e condições privilegiadas de vida e estudo que tive, mas sobretudo por ter me provido força e saúde nos momentos em que levei meu corpo e mente a seus limites extremos para realizar este trabalho. Ao meu pai, primeiro filósofo graduado do meu estado natal, fonte inesgotável de inspiração na sua sabedoria, capacidade de comunicação e criatividade.

A todos os meus professores e orientadores, que dedicaram seu precioso tempo para compartilhar suas experiências e conhecimento, minimizando os tropeços ao longo do árduo caminho do aprendizado.

Ao Prof. Fábio Vinícius Binder, coordenador do BEPiD PUCPR, que me estimulou a fazer o programa de pós graduação e que sempre foi compreensivo quando precisei me ausentar das minhas atividades para me dedicar às demandas da família e do mestrado.

A minha família, em especial à minha querida irmã e madrinha de apresentação Dinara, que mesmo tão longe sempre acompanha com carinho e suporte todos os meus passos. Por último mas de forma alguma menos importante, agradeço à minha esposa Márcia. Serei sempre imensamente grato por não me deixar desistir no meio desta difícil jornada. Por não me permitir optar pelo caminho imediatamente mais vantajoso e confortável mesmo nos momentos mais difíceis, com a exaustiva sobrecarga de trabalho além da necessidade de cuidados e recursos extras com a chegada da minha amada filha Anna Paula.

Sempre foi, e é, por vocês e para vocês. Muito obrigado por acreditarem em mim.

Sumário

Agradecimentos	iv
Sumário	v
Resumo	vii
Abstract	viii
Lista de figuras	ix
Lista de Tabelas	x
Lista de Abreviaturas	xi
Lista de Símbolos	xii
1. Introdução	13
1.1. Desafio – Questão da Pesquisa	15
1.2. Motivação.....	15
1.3. Definição do Problema.....	18
1.4. Objetivos da Pesquisa	19
1.4.1. <i>Objetivo Geral</i>	19
1.4.2. <i>Objetivos Específicos</i>	19
1.5. Motivação.....	20
1.6. Desafios	20
1.7. Proposta	21
1.8. Contribuições	22
2. Fundamentação Teórica	24
2.1. Morfologia Matemática	24
2.1.1. <i>Operações Morfológicas</i>	25
2.1.2. <i>Dilatação</i>	25
2.1.3. <i>Erosão</i>	25
2.1.4. <i>Abertura</i>	26
2.1.5. <i>Fechamento</i>	27
2.2. Detecção de Características	29
2.2.1. <i>Algoritmo SIFT</i>	29
2.2.2. <i>Algoritmo SURF</i>	30
2.3. Triangulação de Delaunay	33
2.4. Modelagem 3D	34
2.5. Topologia Geométrica	36
2.6. Parametrização UV (<i>Parameterization</i>)	36
2.7. Interpolação Gouraud	37
2.8. Nós (<i>Joints</i>) e Aplicação de Pesos.....	38
2.9. Distorção (<i>Warping</i>)	40
2.10. Cálculo de Distorção.....	41
2.10.1. <i>Mapeamento Direto (Forward Mapping)</i>	42
2.10.2. <i>Mapeamento Inverso (Inverse Mapping)</i>	43
2.10.3. <i>Mapeamento Afim (Affine Mapping)</i>	44
2.10.4. <i>Mapeamento Projetivo (Projective Mapping)</i>	44

2.10.5. Mapeamento Por Peças (<i>Piecewise Mapping</i>).....	45
2.11. Considerações	47
3. Estado da Arte	48
3.1. Metamorfose por Campos (<i>Field Morphing</i>)	48
3.2. Deformação Multi-nível de forma livre (MFFD)	51
3.3. Metamorfose de Imagens baseado em Interpolação Morfológica.....	52
3.4. Metamorfose colorida por morfologia matemática	54
3.5. Considerações	58
4. Método Proposto.....	60
4.1. Fluxograma.....	60
4.2. Descrição do Método	62
4.2.1. <i>Visão Geral</i>	62
4.2.2. <i>Deteccção Facial e dos Pontos de Interesse</i>	63
4.2.3. <i>Linha do cabelo, área da testa e queixo</i>	63
4.2.4. <i>Conversão do Espaço de coordenadas 2D para o 3D</i>	64
4.2.5. <i>Métodos de Deformação</i>	65
4.2.6. <i>Alinhamento das Regiões Faciais</i>	67
4.2.7. <i>Deteccção dos recursos da face</i>	67
4.2.8. <i>Deteccção dos FPs (NRFT)</i>	68
4.2.9. <i>Construção da Máscara Facial “ADAMM”</i>	70
4.2.10. <i>Esqueleto de deformação</i>	72
4.2.11. <i>Âncoras de Borda (Border Anchors)</i>	76
4.2.12. <i>Alinhamento dos recursos da face</i>	77
4.2.13. <i>“Baking” geométrico</i>	78
4.2.14. <i>Aplicação de UV padrão</i>	78
4.2.15. <i>Captura das imagens</i>	79
4.2.16. <i>Processamento das Medianas Morfológicas</i>	79
5. Experimentos e Resultados Obtidos	80
5.1. Base de Imagens.....	80
5.2. Tecnologia e Ferramentas.....	80
5.3. Análise de Nitidez.....	81
5.4. Experimentos	82
5.4.1. <i>Protocolo Experimental</i>	82
5.4.2. <i>Experimento 1 – Faces Medianas</i>	83
5.4.1. <i>Experimento 2 – Mãe e Filha</i>	85
5.4.2. <i>Experimento 3 – Hilary Duff</i>	86
5.4.3. <i>Experimento 4 – Reese Witherspoon</i>	87
5.5. Tabela de Resultados	88
6. Conclusão e Trabalhos Futuros	90
7. Referências	92

Resumo

Metamorfose de Imagens (do inglês “image morphing”), é um efeito digital que consiste em produzir uma ou mais imagens que representam a transformação de formas ou aparência entre duas imagens ou dois vídeos. Desde que foi introduzido na primeira linha da indústria do cinema, no início dos anos 90, o objetivo deste efeito tem sido a perfeita ilusão de dois seres ou objetos estarem se transformando um no outro. No modo animado, a metamorfose apresenta uma transição suave e visualmente interessante entre estes dois estados.

Por outro lado, os processos utilizados nos softwares comerciais têm sido extremamente manuais e demorados, e em muitos casos apresentam resultados visuais com baixa nitidez. Existem aplicações que podem ser beneficiadas por resultados mais nítidos do processo de geração de uma imagem individual, indo do entretenimento, com a simulação de filhos ainda por nascer, até a simulação artística forense (Envelhecimento Digital) a partir de imagens de pais e filhos com semelhanças fisionômicas. Os processos atuais de envelhecimento ou progressão digital são inerentemente manuais e lentos, podendo levar cerca de trinta dias para entregar um único resultado. Para atender esta demanda atual por maior qualidade e agilidade no processo, o presente trabalho oferece uma abordagem automática.

A técnica proposta, denominada “ADAMM”, consiste na detecção por visão computacional dos pontos de interesse da face, alinhamento automático entre os segmentos encontrados, e por fim geração da imagem final através da sua Mediana Morfológica. Um programa de computador, o SharpMorpher, foi desenvolvido para demonstrar a eficácia do método. Também foram comparados os resultados do SharpMorpher com os resultados de softwares comerciais que utilizam técnicas tradicionais de morphing. Em todos os testes realizados houve ganho percentual de nitidez, de 66,12% até 625,71%.

Palavras-Chave: Visão Computacional, Metamorfose, Morfologia Matemática, Envelhecimento Digital

Abstract

Image Morphing is a digital visual effect consisting in the generation of one or multiple images depicting the transformation of shapes and appearance between two images or two video sequences. Since its mainstream introduction in the movie industry, at the early 90s, the aim of this effect has been to provide the perfect illusion of a creature or object transforming from one form into another. In the animated sequence version, the morphing presents a smooth and visually interesting transition between these two states.

In the other hand, the workflow required by the specialized commercial software has been extremely time-consuming and laborious for the human operator, often providing results with sub-optimal sharpness. There are applications that can benefit from sharper single-image morphing generation, from entertainment apps that simulate unborn children up to the forensic art simulation (digital aging) field, using images of relatives with physiognomic similarities. Current processes of digital aging are inherently manual and slow, taking around a month to deliver a single result. To cater for the increased demand for quality and speed in the image morphing generation process, the present work presents an automatic approach.

The proposed technique consists in the computer-vision detection of feature points, followed by the automatic alignment of correlated segments and then the final image generation through the Morphological Median operation. A computer program, *SharpMorpher*, was developed to demonstrate the efficacy of the method. *SharpMorpher* results were also compared with results from commercial softwares employing more traditional morphing techniques. In all tests there was a percentage gain in sharpness with the new technique, from 66,12% up to 625,71%.

Keywords: Computer Vision, Morphing, Metamorphosis, Mathematical Morphology, Digital Aging

Lista de figuras

FIGURA 1: PROGRESSÃO DE IDADE DE CRIANÇA DESAPARECIDA	16
FIGURA 2: POLIMORFISMO	17
FIGURA 3: METAMORFORSE POR GRADE DE CONTROLE	18
FIGURA 4: FUSÃO SIMPLES ENTRE DUAS IMAGENS	18
FIGURA 5: DILATAÇÃO E EROÇÃO	26
FIGURA 6: ABERTURA MORFOLÓGICA	27
FIGURA 8: APROXIMAÇÃO DA LAPLACIANA	30
FIGURA 9: APROXIMAÇÃO DE GAUSSIANA	31
FIGURA 10: WAVELET DE HAAR.....	31
FIGURA 11: CONDIÇÃO DE DELAUNAY	33
FIGURA 12: COMPONENTES BÁSICOS DE UMA MALHA 3D	35
FIGURA 13: INTERPOLAÇÃO GOURAUD	38
FIGURA 14: INFLUÊNCIA DE NÓS.....	39
FIGURA 15: DEFORMAÇÃO POR GRADE DE CONTROLE	41
FIGURA 17: MAPEAMENTO INVERSO;	43
FIGURA 18: MAPEAMENTOS PROJETIVO E AFIM	45
FIGURA 19: MAPEAMENTO POR PEÇAS	46
FIGURA 20: ARTEFATOS DO MAPEAMENTO POR PEÇAS.....	46
FIGURA 21: PROCESSO DE METAMORFOSE POR CAMPOS	49
FIGURA 23: MULTILEVEL FREEFORM DEFORMATION	52
FIGURA 24: MEDIANA MORFOLÓGICA BINÁRIA.....	54
FIGURA 25: MEDIANA MORFOLÓGICA DISTORCIDA	57
FIGURA 26: INTERPOLAÇÕES MORFOLÓGICAS SUCESSIVAS	58
FIGURA 27: FLUXOGRAMA DO MÉTODO PROPOSTO.....	61
FIGURA 28: TESTE DE DEFORMAÇÃO DE REGIÃO RETANGULAR	66
FIGURA 29: TESTE DE DEFORMAÇÃO DE ROSTO UTILIZANDO [SCHAEFER2006]	67
FIGURA 30: TESTE COM CLASSIFICADOR HAAR CASCADE	68
FIGURA 31: ALGUNS MODOS DE VARIAÇÃO FACIAL DETECTADOS.....	69
FIGURA 32: DIAGRAMA DE MAPEAMENTO E ALINHAMENTO DAS ROIS	71
FIGURA 33: MÁSCARA POLIGONAL ADAMM.....	72
FIGURA 34: NOMENCLATURA DOS NÓS.....	74
FIGURA 35: HIERARQUIA DE NÓS	75
FIGURA 36: DISTORÇÃO DE LINHAS DE BORDA.....	76
FIGURA 37: APLICAÇÃO DE ÂNCORAS DE BORDA	77
FIGURA 38: PARAMETRIZAÇÃO UV	78
FIGURA 39: FACES MEDIANAS.....	84
FIGURA 40: DETALHE COMPARATIVO	85
FIGURA 41: MODELO DE CRESCIMENTO CRÂNIO-FACIAL	85
FIGURA 42: MORPHING DE MÃE E FILHA	86
FIGURA 43: MORPHING DE HILARY DUFF.....	87
FIGURA 44: COMPARATIVO “GROUND TRUTH”	87
FIGURA 45: MORPH DE REESE WITHERSPOON	88

Lista de Tabelas

Tabela 1 – Resultados dos Experimentos	89
--	----

Lista de Abreviaturas

FP	<i>Feature Point</i> , Marco ou Ponto de Interesse
RoI	<i>Region of Interest</i> , Região de Interesse
RGB	Red-Green-Blue, Vermelho-Verde-Azul
ADAMM	Automatic Detection and Alignment Morphology Morphing, técnica proposta
NRFT	Non-Rigid Face Tracking, método de [Baggio2012]
MM	Morfologia Matemática
VP	Variação Percentual

Lista de Símbolos

ε^g	Erosão utilizando g
δ^g	Dilatação utilizando g
γ^g	Abertura utilizando g
ϕ^g	Fechamento utilizando g
λ	Lâmbda
g	Elemento Estruturante
σ	Gaussiano
lum	Luminância
Min	Mínimo
Max	Máximo

1. Introdução

Antes da Computação Gráfica (CG) se tornar acessível e eficiente, quando a indústria de cinema e televisão necessitava de um efeito de transformação de uma pessoa em outra ou em um monstro, aplicavam-se técnicas rudimentares de maquiagem, corte e fusão de cena. Um exemplo famoso é o do filme “O Médico e o Monstro” (Dr. Jekyll and Mr. Hyde), de 1932.

Em meados da década de 1980, a empresa de efeitos visuais *Industrial Light and Magic* (ILM) desenvolveu o primeiro algoritmo de metamorfose por CG, aplicado nos filmes “Willow” (1988) e “Indiana Jones e a Última Cruzada” (1989). Hoje a metamorfose por CG encontra-se disponível em softwares bem difundidos de efeitos visuais, como o After Effects da empresa Adobe¹. Todavia, o processo de trabalho exigido pelos algoritmos de metamorfose ainda requer bastante preparação e intervenção por parte do usuário.

Originalmente explorados pelo impacto visual, como no videoclip “Black & White” do cantor e compositor Michael Jackson, hoje em dia a metamorfose é aplicada para fins mais sutis e com o resultado mais natural possível. Atualmente, dentre outras aplicações, processadores dedicados dentro de televisores LCD e LED propiciam efeitos convincentes de interpolação de quadros por segundo, chegando a dobrar a taxa de reprodução dos quadros com um número mínimo de artefatos perceptíveis. Nestas aplicações, quadros com diferenças muito grandes entre si simplesmente não são interpolados via metamorfose.

O processo de metamorfose de quadros extremamente distintos é bastante complexo, entretanto, envolvendo a distorção da imagem original para tomar um formato similar ao da imagem de destino (*warping*), ao mesmo tempo em que uma fusão (*fade* ou *cross-dissolve*) é realizada entre as duas. Numa metamorfose entre dois rostos, o processo tradicional envolve marcar pontos-chaves na primeira face – tais quais o contorno do nariz ou os cantos dos olhos – e posicionar marcadores nos pontos correspondentes da imagem da segunda face.

Este trabalho apresenta uma nova técnica para gerar uma imagem de metamorfose facial intermediária a partir de duas imagens frontais de diferentes rostos, através das

¹ Tutorial do processo básico de metamorfose no After Effects. Disponível em: <<http://www surfacedstudio.com/tutorials/after-effects-morph-warp-tutorial>>. Acessado em 11/5/2015

medianas morfológicas de duas imagens em que áreas retangulares correspondentes em cada imagem são alinhadas automaticamente. Estas áreas retangulares são definidas a partir dos extremos dos recursos faciais (*Fps* ou *Feature Points*) de cada região de interesse (RoI ou Region of Interest), por exemplo, região do nariz ou boca.

Em contraste com outras abordagens, esta técnica não requer qualquer tipo de treinamento ou supervisão, podendo ser executado de forma completamente automática.

Para garantir a mais relevante correspondência dos pontos e a menor distorção possível, as *RoIs* das imagens são reposicionadas para o ponto médio entre os seus centros, aproximando-a da outra imagem.

Serão apresentadas métricas absolutas de nitidez nos quadros medianos gerados a partir da técnica proposta, e comparados os resultados com os de aplicativos com características similares.

1.1. Desafio – Questão da Pesquisa

É possível automatizar o processo de metamorfose de imagens de faces humanas, removendo a necessidade de marcação de pontos e ainda aumentar mensuravelmente a taxa de nitidez da imagem-resultado?

1.2. Motivação

A metamorfose de imagens foi considerado um campo de pesquisa esgotado, com praticamente nada novo apresentado desde os últimos trabalhos de Serra e Iwanowski, no princípio dos anos 2000 [Iwano2002].

Todavia, esta presunção deve-se restringir à metamorfose de imagens em movimento, comumente empregada no cinema e publicidade televisiva. Neste tipo de conteúdo a nitidez dos resultados dos quadros intermediários é pouco relevante, já que os resultados são reproduzidos num curto período de tempo, para aumentar o impacto visual. O tempo para montagem destes efeitos é longo, porém pela escassez de demanda é considerado adequado.

Por outro lado, quando o objetivo específico é a geração de um dos quadros intermediários, o aspecto da nitidez é fundamental. Novas aplicações de entretenimento que focam exatamente no quadro intermediário gerado não são bem atendidas pelos processos tradicionais. Podem-se citar neste contexto aplicações de simulação de rostos híbridos entre uma pessoa famosa e outra comum, além de simulação do rosto de um bebê que está para nascer¹.

Além do aspecto de entretenimento geralmente associado à metamorfose de imagens, aplicações de arte forense para a simulação (“envelhecimento digital”) de crianças desaparecidas há anos podem se beneficiar de processos mais rápidos e com melhores taxas totais de nitidez e contraste. O processo artístico manual da atualidade [Taister2000], muitas vezes levando mais de um mês para ser realizado, não comporta a crescente demanda das autoridades e dos próprios familiares.

¹ Website “Make Me Babies”. Acesso por:
<<http://makemebabies.com/#page/1>>



Figura 1: Progressão de idade de criança desaparecida, realizada pelo Serviço de Investigação de Crianças Desaparecidas do Paraná – SICRIDE

Fonte: <http://www.sicride.pr.gov.br/modules/conteudo/conteudo.php?conteudo=8>

Estas simulações digitais, além de serem utilizadas em divulgações da polícia para auxiliar na localização dos menores desaparecidos, são consideradas por especialistas importante fonte de amparo psicológico para os pais e parentes destas crianças.

Hoje existem softwares de geração automática de envelhecimento digital a partir de medições forenses para cada faixa etária, além de outros recursos como aplicação de rugas e manchas. Mesmo o trabalho recente de [Shliz2014] apresenta desfoque devido à filtragem utilizada na etapa de distorção (*warping*), indicando em sua seção de avaliação a necessidade de modificação nas formas e textura das imagens ao invés de recorrer apenas à mescla linear (*linear blend*). Este desfoque ou perda de nitidez, durante os experimentos, foi associado a uma idade percebida inferior à desejada [Burt95]. Além disto, trabalhos como [Burt95] e [Shliz2014] utilizam-se de apenas uma imagem como fonte em seus algoritmos. No caso de bebês por volta de um a dois anos, em que há distorções significativas dos recursos faciais e acúmulo de gordura facial na face [Ram2006], os resultados podem ser aprimorados pela combinação com fotos de pais ou parentes com características fisionômicas similares, que já possuam uma configuração mais definitiva do rosto. Vale notar ainda que existem profissionais da área que consideram artificial o resultado destas técnicas baseadas em uma única imagem¹.

¹ Reportagem sobre envelhecimento digital de crianças desaparecidas, disponível em: <<http://jornalhoje.inf.br/wp/?p=3929>>. Link acessado em 12/10/2015.

Mais um exemplo prático de aplicação relacionado a metamorfose de imagens é a possibilidade de uso na composição de retratos falados que possam auxiliar no trabalho de identificação de suspeitos e investigação policial. [Lee98] descreve o processo de polimorfismo (*Polymorphing*) aplicado a vários segmentos faciais diferentes para a composição de um novo rosto, como pode ser observado na Figura 2.

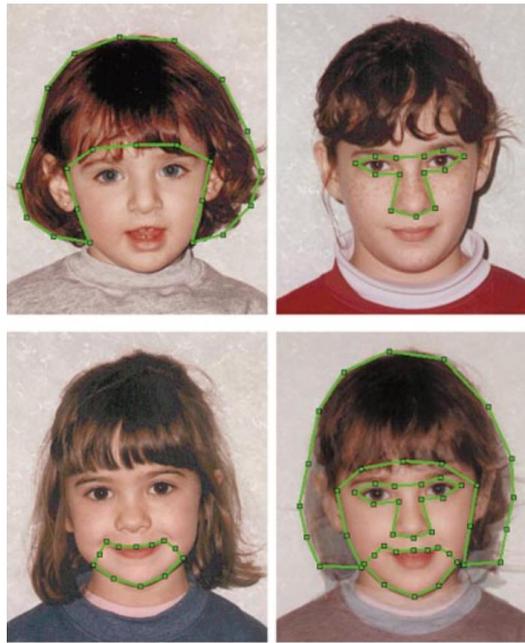


Figura 2: Polimorfismo de seções faciais de várias imagens, marcadas em verde, gerando o rosto final (baixo, direita)

Fonte: [Lee98]

1.3. Definição do Problema



Figura 3: Metamorfose por grade de controle

Fonte: [Wolberg96]

Em um processo de metamorfose de imagens tradicional utilizando a técnica de distorção-e-fusão (*warping-dissolve*), é necessária a adição de marcadores por um operador humano. Um exemplo de resultado do processo é exibido na Figura 3. Apesar do resultado visualmente interessante, pode-se observar a perda de nitidez na interpolação do quadro central, em comparação com os quadros de origem e destino, à esquerda e direita respectivamente. Além disso, o processo de definição dos marcadores é longo e tedioso, visto que precisa ser repetido em ambas as imagens.



Figura 4: Fusão Simples entre duas imagens (Cross fading)

Fonte: [Wolberg96]

Já na Figura 4, gerada automaticamente pelo processo tradicional porém sem uso de marcadores, é visível que apenas uma fusão simples (*cross fade*) foi realizada, com resultados completamente inapropriados ao objetivo, já que não ocorreu sequer um alinhamento dos pontos faciais característicos.

O problema abordado consiste em remover a necessidade de configuração humana ao mesmo tempo em que sejam geradas imagens com nitidez numericamente superior aos resultados das técnicas atualmente empregadas.

1.4. Objetivos da Pesquisa

1.4.1. Objetivo Geral

Desenvolver uma nova técnica de geração de metamorfose de rostos, de forma automática e com maior nitidez numérica [Batten2001] que a das técnicas tradicionais, considerando métodos de avaliação de qualidade ‘*blind*’, ou seja, que dispensam referência ou outras imagens comparativas [XinLi2008].

1.4.2. Objetivos Específicos

- Implementar a operação de Mediana Morfológica entre duas imagens coloridas, conforme [Iwano2002];
- Analisar as técnicas existentes de detecção facial e eleger uma técnica consistente de detectar os pontos de controle das faces em diferentes condições de iluminação;
- Criar um programa capaz de realizar uma nova técnica de metamorfose envolvendo detecção e alinhamento automático de RoIs de face humana e aplicação de Mediana Morfológica, sem provocar distorções ou perda de nitidez nos pixels contidos nas regiões;
- Utilizar métricas exatas e aceitas pela comunidade científica para mensurar os resultados de nitidez das imagens geradas pelo novo método de metamorfose
- Criar uma base de imagens com figuras públicas que possuam similaridades faciais com seus filhos;

1.5. Motivação

O presente trabalho foi motivado originalmente pela perda visível e mensurável de nitidez nos resultados apresentados no mais recente trabalho publicado de Morphing com morfologia matemática [Iwano2002]. Surgiu posteriormente a idéia de que a melhoria na nitidez nos resultados atuais de metamorfose digital de imagens poderia tornar estes resultados adequados para acelerar os trabalhos de envelhecimento digital, notadamente para fins de auxílio à localização de crianças desaparecidas.

1.6. Desafios

Dentre os diversos obstáculos encontrados neste estudo, destacaram-se a detecção dos pontos de interesse (*FPs*) correspondentes em imagens totalmente distintas.

Outro agravante, este bem conhecido na área de visão computacional, encontra-se nas diferentes condições de iluminação no momento em que cada imagem foi registrada, que dificultam a correta identificação dos *FPs*. Problema similar é encontrado em expressões faciais extremas, especialmente sorrisos, assim como com fotos de pessoas de tonalidades de pele muito próximas dos seus pelos faciais.

Um dos maiores desafios consistiu em encontrar um mecanismo de deslocamento de partes da imagem, no caso as próprias *RoIs*, com perfeita manutenção de nitidez dentro destas áreas e mantendo a contiguidade com as demais áreas adjacentes. Várias abordagens testadas resultaram em emendas aparentes nos limites das regiões ou em distorções indesejadas dentro das *RoIs*.

Foi também desafiador o desenvolvimento de uma máscara facial poligonal que mantivesse os principais *FPs* e também maximizasse o espaçamento entre as *RoI*, para evitar artefatos de sobreposição nos casos de rotações e expressões faciais mais exageradas.

1.7. Proposta

A proposta principal deste trabalho de pesquisa é oferecer uma nova técnica de *morphing* com a preparação e alinhamento automático das duas imagens originais, para anteceder o processo de mediana morfológica e proporcionar os objetivos propostos aos artistas forenses de envelhecimento digital. A técnica envolve o uso de uma máscara facial poligonal tridimensional onde as imagens originais são projetadas e podem ter suas ROI reposicionadas, tanto pelo algoritmo quanto pelo operador humano.

Para garantir a preservação das ROIs, utiliza-se da técnica de interpolação de [Gouraud71], inerente à projeção de texturas em objetos tridimensionais – ainda que, neste caso, estes objetos sejam planos bidimensionais. Desde que respeitadas as proporções e tamanhos originais das ROIs, fica garantido ao operador do *software* a preservação exata dos pixels internamente àquelas regiões. As distorções ficam restritas às regiões de mais baixa frequência de detalhes ou menor relevância para a identificação de rostos, como têmporas e bochechas. Mesmo nas áreas distorcidas, a projeção conformal [Haker2000] realizada garante a preservação da nitidez nos *pixels* das bordas das ROI.

Para acelerar o processo de alinhamento de recursos, um processo de visão computacional para detecção de FPs [Baggio2012] é realizado, seguido do alinhamento automático dos FPs das duas imagens originais. É oferecida também ao operador do *software* a capacidade de ajustar o resultado do processo automático de alinhamento, possibilitando o ajuste de proporções faciais para uma determinada faixa etária almejada.

1.8. Contribuições

Este trabalho aliou técnicas de computação gráfica 2D e 3D a técnicas de visão computacional para evoluir de forma mensurável o processo e os resultados da metamorfose de imagens. A técnica desenvolvida inclui montagem e disposição automática do alinhamento dos recursos faciais e automatização do método de composição das imagens medianas, que é aplicado ao final do processo.

A literatura pesquisada não apresenta qualquer solução com o mesmo foco e abordagem. A aplicação de morfologia matemática, especificamente a mediana morfológica, opera sobre toda a imagem indistintamente, associando grupamentos de pixel pela menor distância euclidiana e não por similaridade macroestrutural. As abordagens de mesma finalidade pesquisadas utilizam técnicas de fusão simples e distorção da imagem sobre uma malha de pontos, definidos direta ou indiretamente a partir de um processo manual.

Além da opção pela operação totalmente automática, a técnica proposto neste trabalho oferece melhoria no coeficiente de nitidez dos resultados gerados por um processo de metamorfose específico para rostos humanos. Como produto final do trabalho, foi desenvolvida uma ferramenta computacional interativa e visual capaz de produzir e exibir as imagens de metamorfose utilizando a técnica proposta.

Dois processos de uso foram disponibilizados na ferramenta, com diferentes graus de automação e controle. No processo totalmente automático, os recursos são detectados e alinhados sem intervenção do usuário. Neste caso o software também utiliza configurações pré-definidas da mediana morfológica, que durante os experimentos apresentaram os melhores resultados.

No processo interativo, as RoIs detectadas automaticamente são apresentadas ao operador, superimpostos nas imagens correspondentes. A partir daí o operador pode reposicionar livremente os pontos de controle (FPs), o que atualiza as RoIs interativamente. Uma vez satisfeito com o resultado e finalizando a operação de ajuste, os recursos são alinhados e a mediana morfológica é gerada.

Foi definida também uma estrutura flexível de armazenamento e gerência das regiões faciais, além de um cálculo facilitado de conversão de coordenadas entre os espaços 2D e 3D. A máscara

2. Fundamentação Teórica

Neste capítulo serão apresentados alguns conceitos e operações fundamentais utilizadas no *morphing*, visando estabelecer um vocabulário básico como pré-requisito dos demais capítulos.

2.1. Morfologia Matemática

A Morfologia Matemática (aqui denotada por “MM”) é um campo teórico do Processamento de Imagens que oferece um conjunto de ferramentas úteis para análise de imagem. A MM é uma abordagem de agrupamento dos elementos componentes da imagem (“pixels”) em estruturas que podem ser manipuladas pela teoria dos conjuntos, aplicando técnicas de topologia, funções aleatórias e teoria de treliça (*lattice*). Distancia-se, portanto, de abordagens estritamente aritméticas. A MM foi desenvolvida a partir de idéias de J. Serra e Georges Matheron, da École des Mines de Paris, França, tendo o seu trabalho inicial desenvolvido por Serra (1983). Diversas outras contribuições significativas podem ser encontradas em Serra (1986), Haralick, Sternberg e Zhuang (1987) e Haralick e Shapiro (1992, Cap. 5).

Inicialmente focada em processamento de imagens binárias, a MM foi posteriormente estendida para trabalhar com Imagens em tons de cinza. Em MM, uma imagem é um subconjunto do espaço euclidiano n -dimensional. Por exemplo, o espaço bidimensional (2D), com cada ponto contendo as informações de abscissa e ordenada dos pixels brancos ou pretos na imagem. Analogamente, em imagens de tons de cinza, o espaço utilizado é o tridimensional (3D), onde o terceiro componente indica o valor ou brilho do tom de cinza do pixel na posição definida pelas primeiras duas coordenadas.

As operações básicas da MM podem ser encontradas em diversos pacotes de software para Processamento de Imagens, notadamente o OpenCV, biblioteca de código livre com diversas funções de visão computacional.

2.1.1. Operações Morfológicas

A Morfologia Matemática surgiu em meados dos anos 60, voltada para o tratamento de imagens binárias, sendo posteriormente estendida para imagens em níveis de cinza. Os operadores morfológicos básicos são a dilatação e a erosão, havendo outros operadores compostos baseados na dilatação e erosão. Os operadores mais comuns são descritos a seguir:

2.1.2. Dilatação

A dilatação é uma transformação comutativa, que objetiva “engordar” partículas e preencher pequenas falhas na imagem. Elementos de tonalidade escura de tamanho inferior ao tamanho do elemento estruturante também serão eliminados. A dilatação de níveis de cinza, que possui um valor a cada par de coordenadas x,y , pode ser definida por uma função f aplicando uma outra imagem g , chamada de elemento estruturante [Facon96]. A operação é definida pela Equação 1:

Equação 1

$$\delta^g (f(x)) = \text{Max}\{f(x) + g(x - y): y \in E\}$$

onde Max representa o máximo entre os valores e E representa o domínio do elemento estruturante g .

2.1.3. Erosão

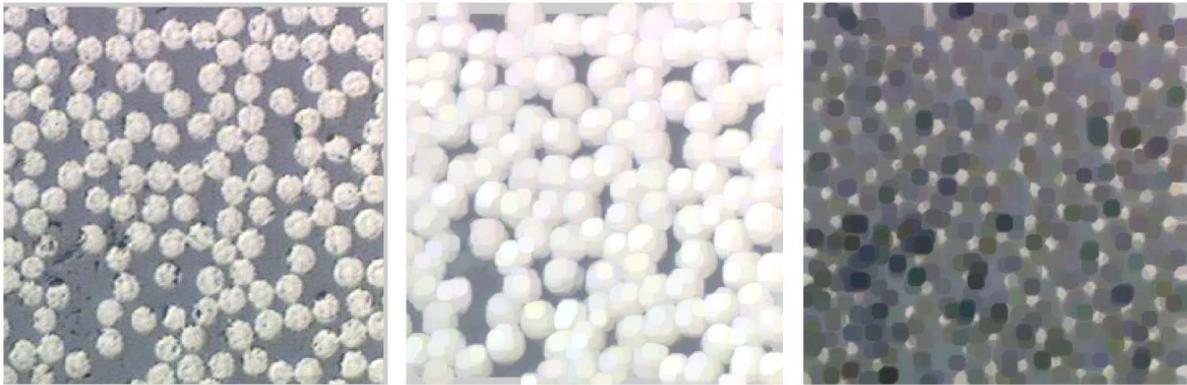
A Erosão é a segunda operação fundamental da MM, junto da dilatação. A erosão de A por B é definida como o conjunto dos deslocamentos z , pelos quais, se B for transladado, o resultado estará contido em A . Em termos de aplicação, a erosão é utilizada para gerar linhas e áreas mais finas em relação à imagem original, diminuindo as partículas e aumentando as áreas de separação ou “buracos” na imagem. Matematicamente, teremos a Equação 2.

Equação 2

$$\varepsilon^g (f(x)) = \text{Min}\{f(x) - g(x - y): y \in E\}$$

Onde *Min* representa o mínimo entre valores e *E* representa o domínio do elemento estruturante [Facon1996]. Na Figura 5 pode ser visto um exemplo do resultado dos operadores de dilatação e erosão sobre a mesma imagem.

Figura 5: Dilatação e Erosão; Resultado de Dilatação (Centro) e de Erosão (Direita) sobre a imagem da esquerda



Fonte: [Facon1996]

2.1.4. Abertura

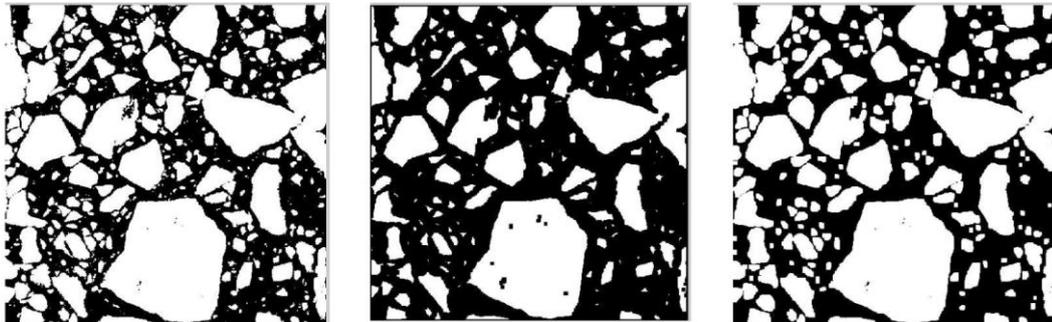
A Abertura é uma operação composta da MM, em que uma operação de erosão é seguida de uma operação de dilatação, utilizando o mesmo elemento estruturante. É utilizada comumente para suavização e remoção de ruído na imagem, sendo capaz de eliminar pequenos relevos e danos na imagem. A operação é descrita matematicamente conforme a Equação 3.

Equação 3

$$\gamma^g(f) = \delta^g[\varepsilon^g(f)]$$

Pode-se visualizar mentalmente o processo de abertura como o deslocamento do elemento estruturante *g*, preso aos limites da de *A* – ou seja, sem nenhum de seus pixels atravessando as bordas de *A*. Todos os pixels de *A* tocados por um dos pixels de *g* durante este processo compõem o resultado da operação de abertura.

Figura 6: Abertura Morfológica; À direita, resultado da operação de Abertura sobre a imagem da esquerda



Fonte: [Facon1996]

2.1.5. Fechamento

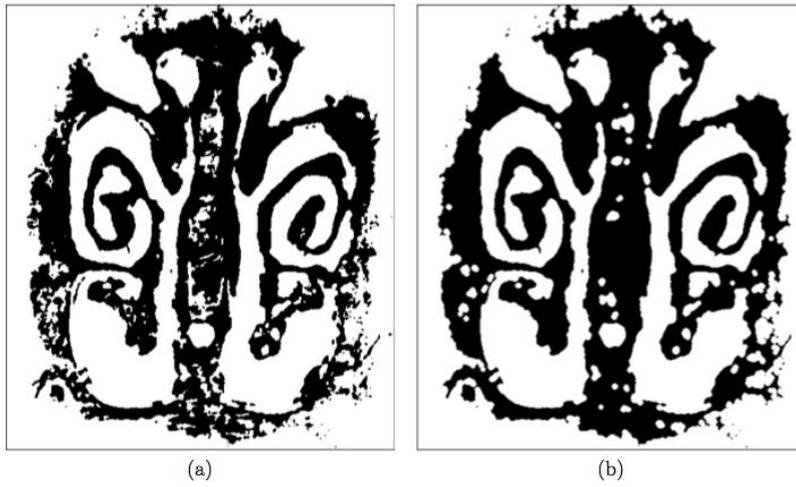
O Fechamento é uma operação composta, assim como a opção de Abertura. É realizada por uma operação de dilatação seguida de uma operação de erosão. O Fechamento também é utilizado para suavização de superfície, mas diferentemente da Abertura o Fechamento funde pequenas fendas, preenchendo áreas vazias. Matematicamente, temos a Equação 4.

Equação 4

$$\phi^g(f) = \varepsilon^g[\delta^g(f)]$$

O processo de Fechamento pode ser visualizado como o deslocamento do elemento estruturante g pela parte exterior dos pixels de A . Apenas os pixels de g durante esta trajetória comporão o resultado da operação. Na Figura 7 pode-se verificar a aplicação do fechamento sobre a imagem original (a), resultando numa imagem de forma similar mas com significativa redução de ruído interno e consolidação dos contornos externos (b).

Figura 7: *Fechamento Morfológico*. (a) Imagem original; (b) Resultado



Fonte: [Facon1996]

2.2. Detecção de Características

O Reconhecimento de Características (*Feature Recognition*, em inglês) é uma das mais importantes aplicações da Visão Computacional. O objetivo é reconhecer estruturas ou pontos muito significativos em uma imagem, para oferecer uma descrição discriminativa destas áreas em relação aos demais pixels da imagem. As aplicações são inúmeras, incluindo classificação de cenas, análise de textura, colagem panorâmica (*Panorama Stitching*), recuperação de imagens e mineração de vídeo. “Características Locais” (*Local Features*) refere-se a padrões de imagem que diferem da sua vizinhança imediata. São portanto de especial interesse para este trabalho, por oferecer um conjunto limitado e claramente localizado de pontos-âncora. Existem diversos algoritmos dos quais o reconhecimento de características se utiliza, alguns dos mais importantes estão listados a seguir.

2.2.1. Algoritmo SIFT

[Lowe99] descreveu um algoritmo para detecção e descrição de características locais em imagens chamado SIFT, acrônimo para “Transformação de Características Invariantes por Escala” (Scale-invariant feature transform).

O efeito de desfoque gaussiano (*gaussian blur*) é frequentemente aplicado para a geração de imagens em diversas escalas, estando disponível na maioria dos programas de edição e tratamento de imagens. Para gerar um processo que não fosse suscetível a variações de escala, Lowe buscou uma forma de aproximar, de forma mais eficiente, a “Laplaciana”, que é um filtro linear capaz de detectar estruturas do tipo bolha (blob) em uma imagem. Abordagens similares à Laplaciana já eram utilizadas em algoritmos de detecção de cantos, como no detector de Harris, porém apresentando um alto custo computacional devido ao cálculo de derivadas dos pixels da imagem.

Localizações-chave no SIFT são portanto definidas como máximos e mínimos do resultado da diferença de Gaussianos (DoG ou *Difference of Gaussians*) aplicadas numa série de imagens suavizadas pela matriz de convolução gaussiana, o que define o “espaço de escala” (*scale space*). A DoG funciona como uma aproximação da derivada da imagem na direção de escala, uma vez que a diferença entre pontos vizinhos em uma dada direção é

uma aproximação da derivada nesta direção. Um exemplo do resultado do processo pode ser visto na Figura 8.



Figura 8: Aproximação da Laplaciana pela diferença de duas imagens filtradas por Gaussiana

O SIFT é capaz de distinguir pontos chave que são invariantes em relação à localização, escala e rotação, além de serem robustos quanto a transformações afins, que incluem distorções de inclinação, e mudanças de iluminação. Por todas estas características o SIFT ganhou amplo reconhecimento para fins de reconhecimento de objetos.

2.2.2. Algoritmo SURF

[Bay2008], no contexto de detecção de faces em tempo real, apresentou um novo processo de detecção de características invariantes-por-escala em uma imagem, denominado SURF (*Speeded Up Robust Features*), conhecido como uma versão acelerada do SIFT. O detector do SURF utiliza uma imagem intermediária chamada “Imagem Integral”, computada a partir da imagem de entrada e utilizada para acelerar os cálculos em qualquer área retangular. A partir da Imagem Integral, apenas quatro operações são necessárias para calcular a área de uma região retangular, independente do seu tamanho. Isto possibilita um cálculo extremamente rápido de um filtro de convolução do tipo caixa, com resultados similares como pode ser visto na Figura 9, além de dispensar o processamento adicional da discretização dos tons de cinza gerados pela gaussiana.

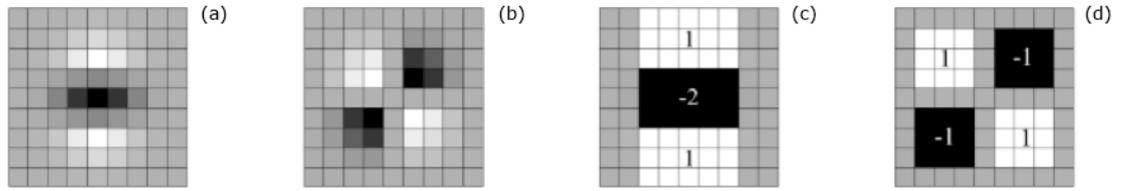
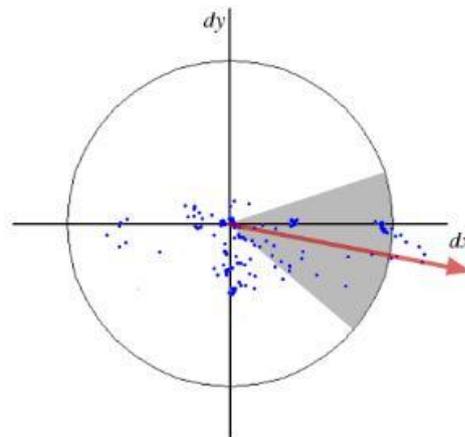


Figura 9: Aproximação de Gaussiana; (a) e (b) Gaussiana nas direções y e xy; (c) e (d) Aproximação de filtro-caixa do SURF nas direções y e xy. Regiões em cinza equivalem a zero.
Fonte: [Bay2008]

Wavelets de Haar são utilizadas para indicar a direção de uma nuvem de pontos, como é possível ver na Figura 10.

Figura 10: Wavelet de Haar, que determina a orientação do SURF. Os pontos azuis são as respostas de Wavelet de Haar para diferentes escalas, a seta vermelha indica a direção atribuída



Fonte: [Bay2008]

O descritor do SURF é baseado na determinante da matriz Hessiana, que é uma matriz quadrada das derivadas parciais de segunda ordem de uma dada função, capaz de descrever a curvatura local de uma função de múltiplas variáveis. A matriz Hessiana foi desenvolvida no século 19 pelo matemático alemão Ludwig Otto Hesse e posteriormente nomeada em sua homenagem.

No caso do SURF esta determinante descreve como as intensidades de pixel são distribuídas dentro da vizinhança de cada ponto de interesse detectado pela Hessiana Rápida (*Fast Hessian*). A matriz Hessiana é aproximada grosseiramente neste processo de *Fast Hessian*, utilizando um conjunto de filtros do tipo caixa, e nenhuma suavização é aplicada

quando da mudança de escala. Isto oferece resultados similares aos das derivadas gaussianas de segunda-ordem e podem ser calculados muito rapidamente. Os resultados em termos de precisão são comparáveis aos obtidos por gaussianas discretizadas, já que existem muitas fontes de ruído na cadeia de processamento.

Por ser um processo invariante quanto a escala e rotação, de desempenho ótimo, além de completamente não-supervisionado, o SURF é uma das técnicas mais utilizadas em Visão Computacional para detecção de características importantes em uma imagem, como no caso de FPs em rostos.

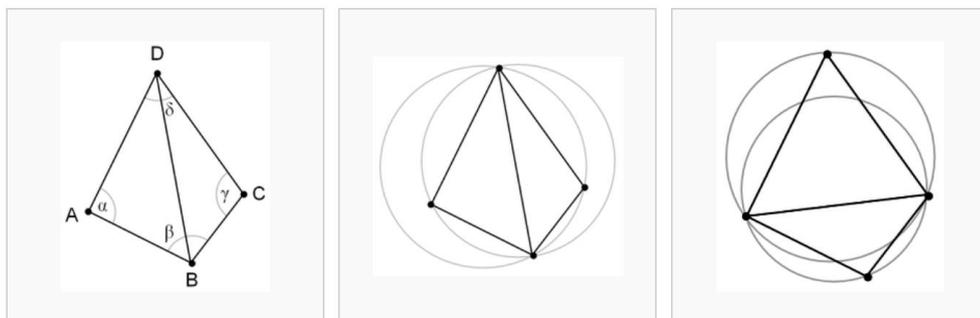
2.3. Triangulação de Delaunay

Em 1934, o francês Boris Delaunay, criou uma técnica revolucionária para a definição de triângulos a partir de um conjunto de pontos em um plano, processo conhecido por “triangulação” [DeLoera2010]. O objetivo era o de minimizar a ocorrência de triângulos com ângulos internos muito pequenos, que ocasionam uma série de problemas práticos nas transformações e projeções gráficas. Para tal, o menor ângulo interno dentre os do triângulo é maximizado.

Além disso, o conjunto de três pontos não pode ser colinear, ou seja, estar na mesma linha, o que não resultaria num triângulo visível. Para quatro ou mais pontos que formam um retângulo, a triangulação sempre possuirá duas configurações possíveis. Ambas ou apenas uma satisfará a “condição Delaunay” que é de que, dados os pontos que compõem um triângulo, o círculo que envolve estes pontos jamais conterá mais de três pontos, e que a soma dos ângulos dos vértices que não estão ligados pela aresta central não ultrapasse 180° .

Quando esta condição não é satisfeita para um conjunto de quatro pontos onde dois triângulos estão definidos, a aresta central que define a separação dos triângulos é rotacionada, de forma que passa a ligar os dois pontos que não possuíam conexão. O processo pode ser visualizado na Figura 11.

Figura 11: *Condição de Delaunay*; (Esquerda) Possui a soma de α e γ maior que 180° ; (meio) Circunferência esquerda contém 4 pontos; (direita) Satisfaz a condição



Fonte: [DeLoera2010]

O procedimento é repetido até todas as circunferências dos triângulos possuírem interiores vazios, ou seja, com três pontos ou menos. Alcançando esta condição, o algoritmo é satisfeito para o conjunto de pontos apresentado.

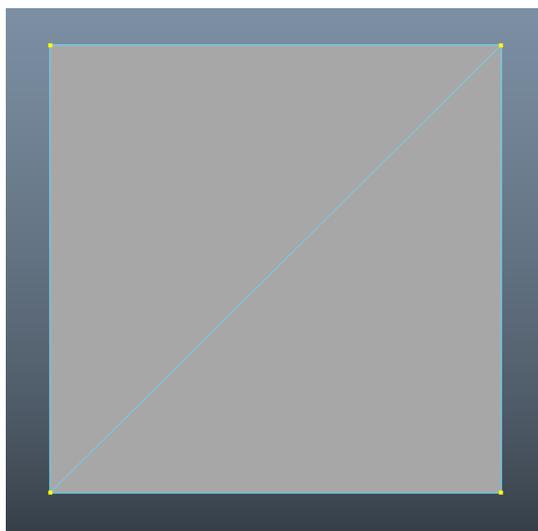
Em termos de implementação, os primeiros algoritmos computacionais para o cálculo de Delaunay basearam-se em seleção aleatória de pontos e posterior otimização. Todavia isto gerava um tempo de processamento $O(n^2)$, ou seja, inconsistente e potencialmente bastante longo dependendo da triangulação inicial. Diversos algoritmos foram propostos para otimizar o processo, destacando-se recentemente o *Sweep Hull (S-Hull)* [Sinclair2010] que gera triangulações sem sobreposição, baseados no pareamento dos triângulos a analisar em relação aos anteriormente processados, resultando em complexidade de processamento $O(n \log(n))$.

2.4. Modelagem 3D

O processamento de imagens opera sobre os *pixels*, menores elementos que compõem uma imagem. No conteúdo destas imagens podem ser representadas linhas e polígonos, que podem ser armazenados e manipulados como uma única entidade.

Da mesma forma, num nível mais alto de abstração, os modelos tridimensionais também são compostos no seu nível mais básico por pontos, sendo a conexão entre dois pontos no espaço chamada de aresta (*edge*). Um conjunto de três arestas compostas por apenas três pontos no espaço define um triângulo, elemento básico de toda representação de malha geométrica tridimensional.

Figura 12: Componentes Básicos de uma malha 3D. Vértices destacados em amarelo, arestas na cor ciano e triângulos em cinza



Fonte: Autor

Triângulos podem também ser definidos no espaço bidimensional, utilizado pelas imagens mapa-de-bits, sendo a única diferença dos triângulos tridimensionais a coordenada de profundidade associada aos seus vértices componentes. Supondo uma câmera no espaço tridimensional que capture imagens com projeção ortográfica (ou seja, sem escorço¹), com sua direção visual ortogonal à de um triângulo qualquer, a imagem resultante conterá toda a informação visual do triângulo, sem distorções. Utilizando uma malha 3D com todos os triângulos coplanares, é trivial realizar um mapeamento planar conformal de uma imagem mapa-de-bitmaps 2D. [Shef2006]

Outra característica importante dos componentes de malha 3D é o do sombreamento (*shading*). O sombreamento utiliza um programa chamado *shader* para definir como a superfície 3D será desenhada na imagem 2D gerada pela câmera virtual posicionada no espaço tridimensional. Os *shaders* necessitam de algum tipo de Parametrização [Gu2003] para saber como mapear pixels das imagens mapa-de-bits, denominadas "texturas", para os vértices dos modelos.

¹ Escorço ou *Foreshortening* é a distorção visual provocada pela percepção da perspectiva, pela qual objetos mais distantes possuem tamanho menor que os mais próximos do observador. [Loomis56]

2.5. Topologia Geométrica

A topologia geométrica de uma malha poligonal consiste na organização e distribuição dos seus vértices e arestas. Os princípios de topologia geométrica são aplicados para a disposição dos triângulos que compõem uma dada malha poligonal a ser representada com a menor quantidade de triângulos e o espaçamento mais amplo e regular possível. Estas características são desejáveis em parte pela necessidade de performance interativa do modelo, mais eficiente quanto menor o conjunto de triângulos sendo manipulado, e por outro lado por garantir a máxima capacidade de deformação possível.

Triângulos mal-formados, definidos por [Chew93] como aqueles que possuem algum ângulo interno inferior a 30° ou superior a 120° , ou compostos por um ou mais pontos irrelevantes à descrição da superfície, não apenas dificultam sua edição numa interface visual, como também propiciam problemas quando texturizados e posteriormente deformados. Estas áreas, se expandidas por um processo qualquer de distorção geométrica, apresentarão um alto grau de artefatos visuais.

Por outro lado, uma quantidade de pontos maior que a necessária em uma região impõe dificuldades para a compressão da área. Exemplificando, para comprimir uma área retangular verticalmente, se temos apenas os quatro pontos dos cantos, necessitaremos manipular apenas os dois pontos compartilhados pela aresta superior ou inferior. Uma quantidade maior de pontos nesta região resultaria numa sobrecarga desnecessária de armazenamento e controle dos pontos adicionais, o que facilitaria a ocorrência de erros de preparação e edição da malha.

2.6. Parametrização UV (*Parameterization*)

A parametrização UV consiste no mapeamento de uma superfície geométrica 3D em um plano 2D [Shef2006]. O processo, chamado de mapeamento de vértices (*vertex map*), possibilita aplicações essenciais, como o mapeamento de imagens mapa-de-bits em objetos tridimensionais.

O mapeamento "conformal" preserva as formas locais, ou constituintes, do objeto mapeado, que são intrínsecas à topologia (distribuição de linhas) e geometria da malha [Gu2003]. Este tipo de mapeamento é independente de resolução do objeto original,

preserva a orientação dos seus elementos internos e não apresenta bordas de descontinuidade, tornando o processo um candidato ideal à transformação rígida de recursos faciais.

Em computação gráfica, o sistema de parametrização utilizado neste processo é chamado de "coordenadas UV", fazendo uma analogia às letras X e Y relacionadas aos eixos do espaço bidimensional.

Uma propriedade interessante do mapeamento de UV, em especial para o presente trabalho, é a interpolação linear aplicada aos pixels da textura [Gouraud71]. Esta parametrização garante continuidade perfeita dos pixels ao longo das arestas dos triângulos adjacentes, sem nenhum tipo de distorção nos pixels mapeados para um triângulo que seja deslocado no espaço, desde que mantida a sua forma original.

2.7. Interpolação Gouraud

O mapeamento de textura conecta pontos no espaço a pixels na textura, pixels estes de coordenadas inteiras. [Gouraud71] definiu uma técnica conformal de mapeamento de textura, como segue: Seja um triângulo composto de três pontos v_1 , v_2 e v_3 com as coordenadas mapeadas $\langle i_1, j_1 \rangle$, $\langle i_2, j_2 \rangle$ e $\langle i_3, j_3 \rangle$. Seja i uma coordenada entre i_1 e i_2 , $i_1 \leq i \leq i_2$, teremos um α definido na Equação 5.

Equação 5

$$\alpha = \frac{i - i_1}{i_2 - i_1}$$

O valor da posição j deste pixel na aresta do triângulo pode ser calculada pela Equação 6.

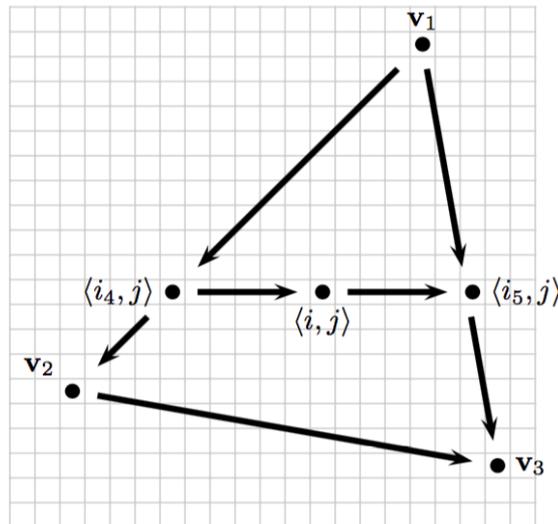
Equação 6

$$j(i) = [(1-\alpha)j_1 + \alpha j_2]$$

Sendo $[x]$ a operação de arredondamento para o número inteiro menor ou igual a x . Utilizando a mesma fórmula, que é basicamente a fórmula da interpolação linear, podemos calcular o valor da cor dos pixels interpolados. Primeiro computam-se as cores dos pixels nos vértices, de acordo com suas coordenadas UV. Em seguida os pixels ao longo das arestas dos triângulos são computados, seguidos da interpolação dos pixels da área interna.

O sentido é sempre de cima para baixo e da direita para a esquerda, conforme as setas na Figura 13.

Figura 13: Interpolação Gouraud, diagrama da técnica



Fonte: [Buss2004]

2.8. Nós (*Joints*) e Aplicação de Pesos

Um Nó (*joint*) é um ponto no espaço que não possui nenhuma geometria associada a si. O nó possui a função de servir de centro e pivô de transformação para um mais vértices que fazem parte de uma malha poligonal, possibilitando deformações locais em objetos tridimensionais.

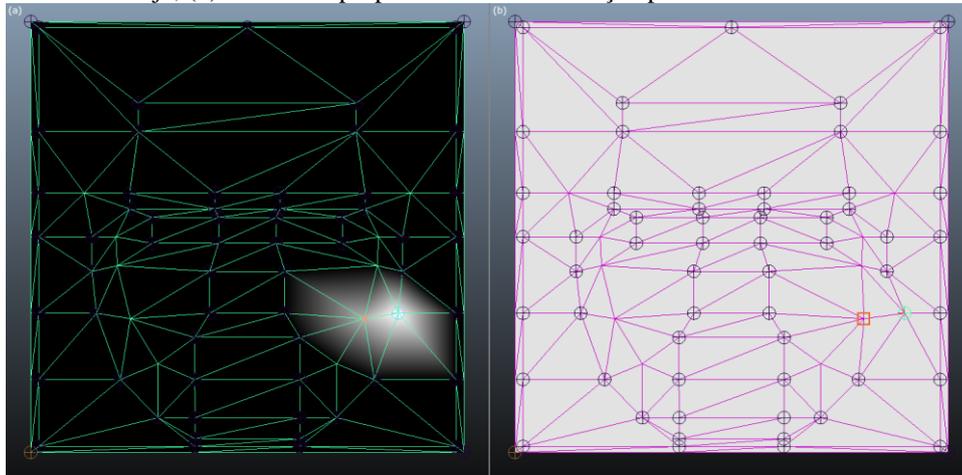
Para controlar como um nó afeta um ou mais vértices da malha, um mapeamento de vértices específico é utilizado, chamado de *Weight Map* (Mapeamento de Pesos) ou *Skinning* (Pelagem). O peso no caso é um valor numérico que corresponde ao percentual de influência da movimentação de um nó sobre a movimentação daquele vértice.

Dois ou mais nós podem dividir suas influências sobre um único vértice, por isso algum algoritmo de normalização destes valores é necessário nas aplicações 3D. Desta forma, o valor máximo de 1, correspondente a 100% da influência de transformação espacial, pode ser dividido entre os nós associados ao vértice. É possível observar no exemplo da Figura 14 como o nó da lateral do rosto exerce influência parcial sobre o vértice adjacente da região da bochecha, destacado em laranja na figura. O valor de 0.5 de influência é denotado

pelo tons acinzentado, sendo a cor branca utilizada para denotar influência máxima do nó e a cor preta exibida onde não há nenhum efeito da influência selecionada.

Utilizando esta técnica, é possível exercer influência total ou parcial de um nó sobre determinados vértices. Isto aumenta a flexibilidade da deformação onde exigido, sem abdicar da correlação 1:1 de posicionamento entre os nós e seus vértices justapostos, nas regiões que não podem ser deformadas.

Figura 14: Influência de Nós; (a) Influência parcial de nó, em ciano, sobre um vértice, em laranja; (b) Resultado proporcional de deformação pelo deslocamento do nó



Fonte: Autor

2.9. Distorção (Warping)

A distorção ou warping é um processo que desloca os pixels de uma imagem para novas posições dentro da própria área da imagem. O processo é dividido em duas etapas, a do deslocamento dos pixels ou áreas de pixels definidos como “fixos” para as suas posições finais, seguido da interpolação dos pixels restantes para posições intermediárias.

A distorção é a primeira etapa num processo tradicional de metamorfose, onde é aplicado tanto na imagem origem quanto na imagem destino. O processo pode ser descrito como um mapeamento 2D-para-2D da seguinte forma:

- O pixel (x,y) é mapeado para a posição (u,v)
- A nova posição é dada pelas funções $u(x,y)$ e $v(x,y)$

Um processo de distorção ideal [Wolberg90] deve oferecer os seguintes resultados:

- Continuidade, sem falhas ou rachaduras
- Paridade um-para-um, sem dobras ou sobreposições
- Suavidade nos contornos gerados
- Flexibilidade de representação

Uma das técnicas de distorção mais utilizadas, segundo Wolberg, é a de recálculo de coordenadas, que é realizado seguindo os seguintes passos:

1. Um conjunto de pontos de controle é definido sobre a imagem original, definido pelo usuário do programa

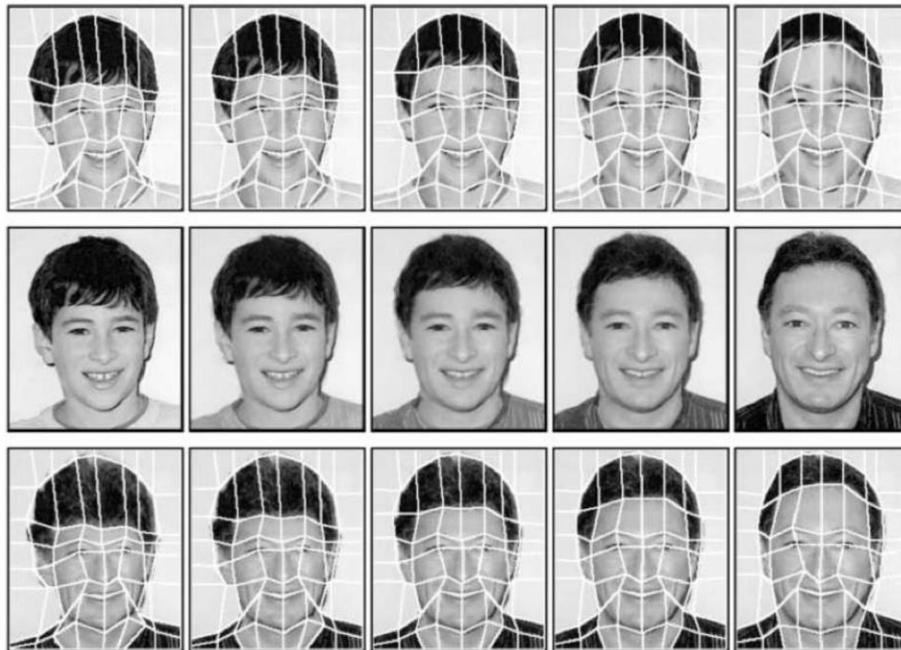
2. Pontos extras são adicionados, manual ou automaticamente, à borda da imagem, para garantir a formação de quadriláteros. Na prática uma grade tendo estes pontos como vértices é disposta sobre a imagem, e este fica definido como o estado neutro ou original.

3. Os pontos de controle são deslocados sobre a área da imagem, deformando os quadriláteros em relação à sua configuração original.

4. Para calcular o resultado da distorção, o valor de cada ponto de controle é copiado da imagem original para a imagem resultado. Os valores dos demais pixels é interpolado bi-linearmente utilizando os valores dos quatro pixels mais próximos e conhecidos na imagem inicial.

Utilizando a interpolação bi-linear, evita-se a aparência serrilhada que resulta de interpolação do tipo vizinho-mais-próximo (*nearest neighbor*) nos quadriláteros que aumentam de tamanho no processo. O resultado do processo de distorção e a interpolação de imagens pode ser observado na Figura 15.

Figura 15: Deformação por grade de controle; (Acima) Deformação da Imagem de Origem para a disposição Final da Grade de Controle; (Abaixo) Deformação da Imagem de Destino para a disposição inicial da Grade de Controle; (Centro) Sequência final de Metamorfose gerada pela Interpolação linear progressiva das duas seqüências de distorção



Fonte: [Wolberg96]

2.10. Cálculo de Distorção

Quando temos uma transformação $T: U \rightarrow U'$, do espaço, é possível interpretá-la de duas formas:

T é considerada como pontos de transformação do espaço U, tomando cada ponto P e movendo-o para uma nova posição $p' = T(p)$ no espaço.

T é considerada uma mudança de coordenadas do espaço U . Para cada ponto $p = (x_1, \dots, x_n)$, T associa novas coordenadas $Y = (y_1, \dots, y_n) = T(x_1, \dots, x_n)$. As coordenadas de p no novo sistema são definidas por $T^{-1}(y_1, \dots, y_n)$

Considerando as duas interpretações possíveis para uma transformação, temos duas formas distintas de realizar suas operações: mapeamento direto e inverso.

Ao tratar uma imagem por inteiro, a escolha de aplicar um mapeamento direto ou inverso é indiferente, já que o mapeamento é feito ponto-a-ponto. Já quando manipulamos seções da imagem, as situações são significativamente distintas.

A discussão generaliza-se facilmente a representações matriciais de objetos gráficos de diferentes dimensões, como é o caso na metamorfose de imagens.

2.10.1. Mapeamento Direto (*Forward Mapping*)

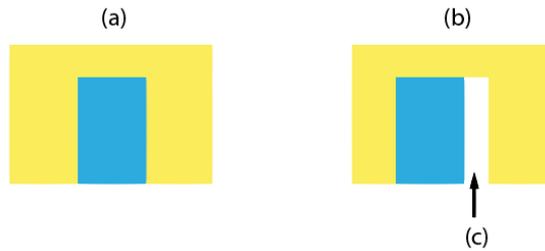
O processo de mapeamento direto varre toda a imagem de origem pixel a pixel, copiando a informação de cor para o pixel correspondente na imagem de destino. A desvantagem deste processo é que alguns pixels da imagem de destino podem não ser associados, por estarmos trabalhando com áreas separadas nas imagens. Este fato gera falhas (buracos) na imagem interpolada.

Por outro lado, já que a imagem de saída também é composta de várias seções, muitas amostras de pontos podem se sobrepor, sendo mapeadas para o mesmo pixel de saída. No geral, mapeamento direto pode exigir computação excessiva na amostragem de pontos localizados em áreas não-importantes da imagem, deixando muitas vezes áreas relevantes sem amostragem.

Cálculos complexos são necessários para tirar uma amostragem apropriada e regularmente distribuída na imagem de origem. Considerar a geometria de pixels como uma área ao invés de um ponto evita a maioria dos problemas citados. O pixel de destino mapeado (p_0) é aproximado por um quadrilátero, que após interseccionado com a grade regular de imagem, pode ser utilizado para determinar as influências de p_i nos pixels da imagem final, como pode ser visto na Figura 16. O mapeamento direto é recomendado para situações em que a imagem de origem precisa ser lida sequencialmente, ou quando não reside integralmente na memória, como no caso de leitura progressiva (*streaming*) via

internet.

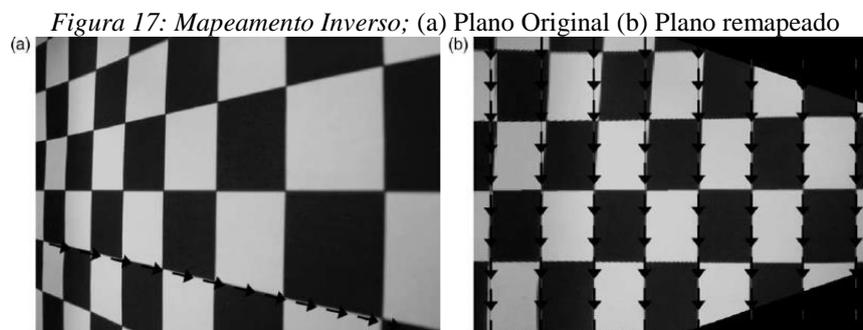
Figura 16: Mapeamento Direto; (a) Imagem de entrada; (b) imagem de saída; (c) falha gerada na Transformação dos pixels azuis



Fonte: Autor

2.10.2. Mapeamento Inverso (*Inverse Mapping*)

Para resolver as limitações do processo de mapeamento direto de imagens, o processo inverso foi proposto. O mapeamento inverso varre a imagem de *destino* pixel a pixel, e copia a informação de cor do pixel correspondente na imagem de origem. Assim como no mapeamento direto, é realizado um arredondamento para obter uma coordenada de pixel com números inteiros, só que agora o arredondamento é feito nas coordenadas da imagem de entrada. Opcionalmente algum processo de interpolação pode ser aplicado, considerando as cores dos pixels vizinhos em uma certa região. Além de ser simples e rápido de computar, a grande vantagem do processo de mapeamento inverso é a garantia de preenchimento de toda a imagem interpolada, sem falhas de continuidade mesmo com grandes distorções, como pode ser visto na Figura 17.



Fonte: [Tan2006]

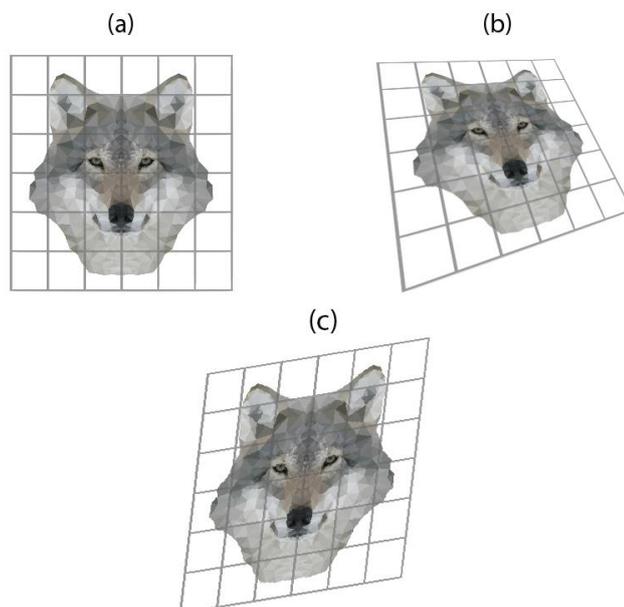
2.10.3. Mapeamento Afim (*Affine Mapping*)

Este tipo de mapeamento é caracterizado por uma matriz de translação, rotação, escala e inclinação (shearing) que é aplicada sobre a imagem. Corresponde a uma projeção ortográfica – ou projeção de plano paralelo – do plano de origem uv para o plano de destino xy . Assim sendo, o mapeamento afim preserva as linhas paralelas da imagem, impossibilitando deformações de perspectiva (escorço ou *foreshortening*) durante o mapeamento. Pontos equidistantes são preservados, tornando a técnica apropriada para mapeamentos de triângulos. Todavia, as transformações não são genéricas o suficiente para mapear quadriláteros, restringindo a aplicabilidade deste tipo de mapeamento.

2.10.4. Mapeamento Projetivo (*Projective Mapping*)

O mapeamento Projetivo, ou transformação de Perspectiva, trabalha com um plano de projeção num dado plano de visualização, conhecido por “perspectiva”. Este mapeamento preserva linhas retas, mas não as linhas paralelas, exceto quando as mesmas são paralelas no plano de projeção, caso contrário as linhas convergirão na direção de um ponto de fuga. Este subproduto, conhecido como escorço, é ideal para a sintetização de imagens realísticas. O mapeamento Projetivo compartilha a qualidade de paridade linha-a-linha do mapeamento afim, inexistente em outros mapeamentos mais genéricos, agregando ainda a capacidade de mapear de forma planar quadrilátero-a-quadrilátero. Um exemplo dos mapeamentos projetivo e afim pode ser visto na Figura 18.

Figura 18: Mapeamentos Projetivo e Afim; (a) imagem original; (b) imagem distorcida por mapeamento projetivo; (c) imagem distorcida por mapeamento afim



Fonte: Autor; Arte do lobo disponibilizada por licença *Creative Commons Zero*¹

2.10.5. Mapeamento Por Peças (*Piecewise Mapping*)

Também conhecido por “Transformação por peças” (*Piecewise Transformation*), o mapeamento por peças surgiu para suprir as limitações dos mapeamentos afim e projetivo, que não conseguem expressar deformações que afetem diferentemente partes distintas da imagem. O processo particiona a imagem de entrada em peças separadas e aplica deformação (*warp*) em cada peça individualmente.

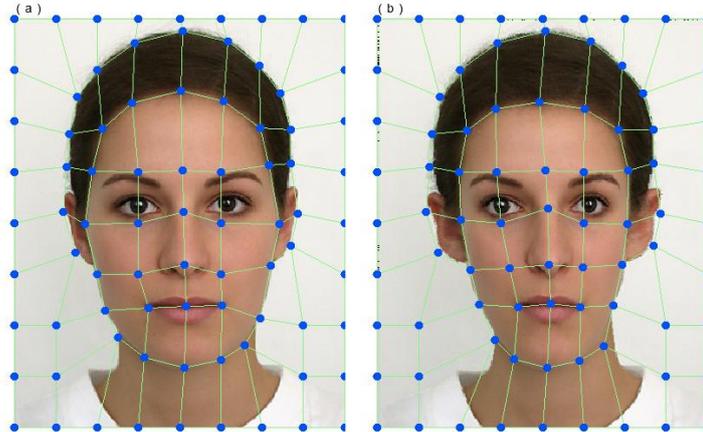
Na sua forma mais simples, o mapeamento afim por peças, os elementos de partição são triângulos. Vértices nas bordas da imagem são adicionados para garantir o suporte à triangulação de toda a imagem.

A triangulação ideal neste processo utiliza o algoritmo de Delaunay, que garante a maximização dos menores ângulos internos dos triângulos. Como resultado prático, este processo de triangulação oferece uma distribuição ótimas de arestas internas, sem triângulos alongados ou achatados, que são indesejáveis pelos artefatos visuais que provocam no processo de distorção.

¹ Imagem em formato vetorial aberto SVG. Disponível em: <https://openclipart.org/image/800px/svg_to_png/261833/Low-Poly-Gray-Wolf.png>. Acessado em 11/11/2016

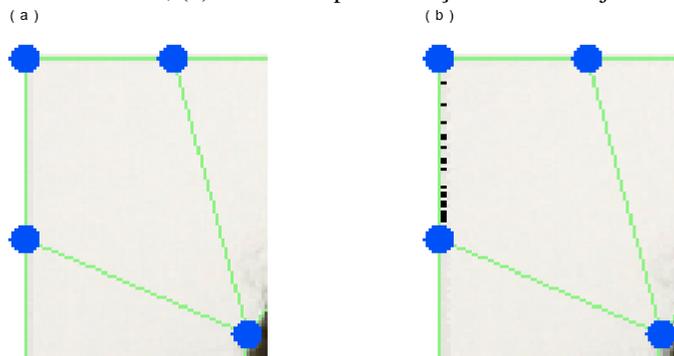
Cabe notar que as técnicas de mapeamento e distorção por peças estudadas apresentam artefatos em áreas que mantiveram seu formato original, por conta de distorções em outras áreas, como pode-se ver em destaque na Figura 20.

Figura 19: Mapeamento por peças; (a) imagem original com áreas poligonais definidas; (b) deformação resultante do deslocamento dos pontos de controle



Fonte: Image Understanding 2012 Course Assignment¹

Figura 20: Artefatos do Mapeamento por peças; (a) detalhe da imagem original em uma área não distorcida; (b) resultado após distorção de áreas adjacentes



Fonte: Image Understanding 2012 Course Assignment¹

¹ Image Understanding 2012 Course Assignment – Brown University. Disponível em: < <http://mesh.brown.edu/engn1610/assignment3.html> >. Acessado em 11/11/2016

2.11. Considerações

Neste capítulo foram apresentados os fundamentos e conceitos teóricos que foram úteis no estudo e desenvolvimento deste trabalho.

As operações morfológicas são necessárias para a implementação do algoritmo de mediana morfológica, utilizadas na etapa final da técnica proposta.

As técnicas de reconhecimento de características genéricas como o SIFT e SURF geram um grande volume de pontos, de difícil tratamento e manipulação, indicando a necessidade do uso de algoritmos específicos para o reconhecimento de recursos faciais. O algoritmo de reconhecimento deve ser capaz de oferecer um conjunto limitado de pontos (FPs) de forma consistente e resistente a variações de iluminação e rotação da cabeça dos indivíduos nas imagens a serem processadas.

Dentre os procedimentos de triangulação existentes o de Delaunay se destaca pela qualidade da malha gerada, entretanto para o sucesso da técnica proposta existe a necessidade de correspondência exata entre não apenas os FPs, mas também entre os triângulos que os interligam. Esta característica não é garantida na Triangulação Delaunay, o que inviabilizaria o processo.

Desde que garantida a detecção consistente dos FPs, e sendo este um conjunto limitado de pontos, a definição explícita e prévia desta malha poligonal, com o auxílio de uma ferramenta de modelagem 3D, apresentou-se como a forma mais simples de garantir a consistência do mapeamento das RoIs nos rostos das duas imagens. Um benefício adicional foi o da possibilidade de polir a topologia geométrica desta “máscara poligonal”, capaz de suportar grandes deformações nas áreas definidas, como frequentemente é encontrado nas expressões faciais humanas.

A parametrização UV possibilita a aplicação de imagens (texturas) nas malhas poligonais 3D, com o uso da parametrização conformal de [Gouraud71] que é um tipo de mapeamento inverso que garante imutabilidade dos pixels nas arestas dos triângulos e evita qualquer tipo de distorção interna em uma região por conta da deformação de uma região adjacente.

Estas características e conclusões possibilitaram definir os recursos e mecanismos necessários para alcançar os objetivos propostos neste trabalho.

3. Estado da Arte

Neste capítulo são apresentados alguns artigos publicados relacionados a este trabalho, junto a uma descrição sucinta das técnicas aplicadas e resultados por eles obtidos. O objetivo desta análise foi a compreensão do processo evolutivo das técnicas de *morphing* e quais aspectos das mesmas podem ser aproveitados no presente trabalho.

3.1. Metamorfose por Campos (*Field Morphing*)

Trabalhando para a Pacific Data Images, Beier e Neely [Beier92] apresentaram uma técnica de aplicação de primitivas bidimensionais de controle, especificamente linhas, sobre a imagem original. Em torno de cada uma destas linhas é gerado um “campos de influência” que atrai os pixels em torno desta de forma proporcional à proximidade da linha. Após a disposição das linhas na imagem de origem é necessário dispor as linhas sobre a imagem destino, de forma a manter a direção e o comprimento das linhas acompanhando os contornos das formas nas duas imagens. O objetivo principal foi o de simplificar a correspondência de pontos de interesse entre as imagens de origem e destino. Ainda assim, este processo é eminentemente supervisionado, o artigo inclusive cita tempos dez vezes mais longos para preparar o processo, por um artista, do que para o cômputo dos quadros resultantes.

Na Figura 21 é possível ver o processo de aplicação da técnica.

Figura 21: Processo de Metamorfose por Campos, linhas de Correspondência

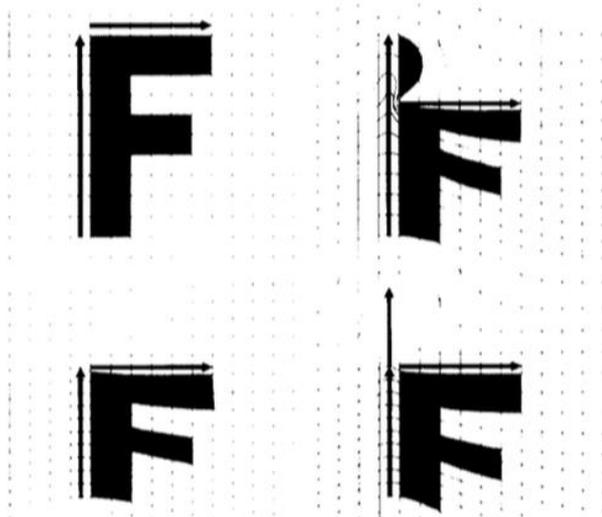


Fonte: [Beier92]

É citada como vantagem sobre a técnica de distorção de malha tradicional [Wolberg90] a expressividade da técnica, exigindo do artista apenas o uso de linhas de contorno importantes nas duas imagens, gerando resultados consistentes e previsíveis. Em contrapartida, a técnica complica a geração da distorção. Isto deve-se ao fato de que todos os pares de linhas devem ser considerados antes que o mapeamento de cada ponto de origem seja conhecido. Este algoritmo global é mais lento do que a distorção por malha, que aplica interpolação bicúbica para determinar o mapeamento de todos os pontos fora da grade da malha.

Um problema mais grave ainda é que deslocamentos inesperados podem ser gerados após serem consideradas as influências de todos os pares de linha sobre cada ponto individualmente. O resultado é a geração de artefatos de gargalo (*bottlenecks*) ou “fantasmas”, que exige um processo de tentativa-e-erro dos artistas na colocação de linhas-guias adicionais para corrigir estas falhas, como pode ser visto na Figura 22.

Figura 22: Artefatos da Metamorfose por Campos



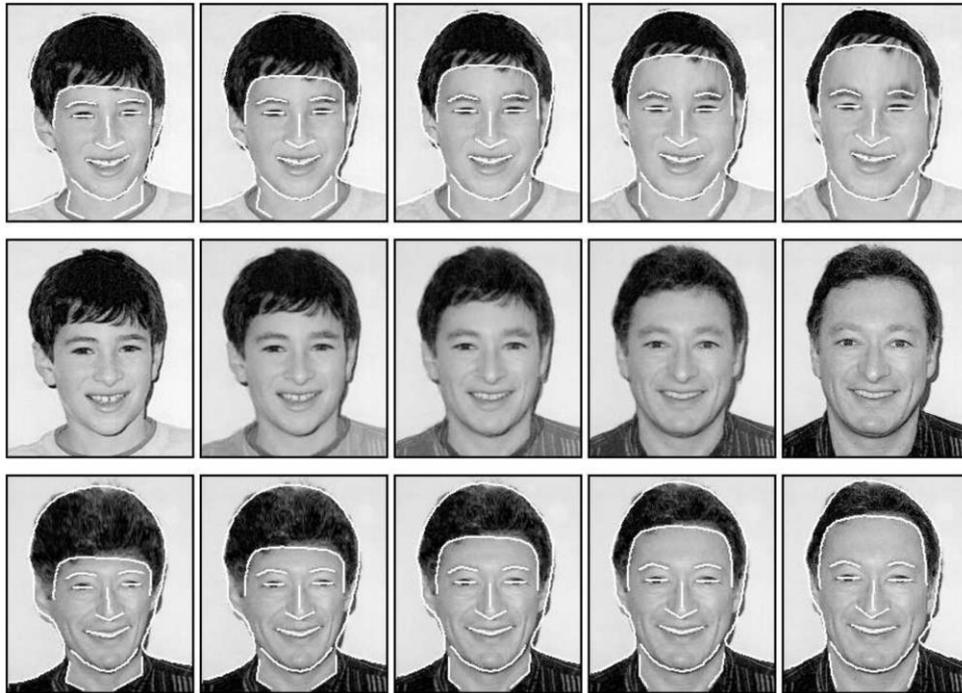
Fonte: [Wolberg90]

3.2. Deformação Multi-nível de forma livre (MFFD)

As técnicas de distorção e interpolação linear não garantem que partes da imagem não se “dobrem”, ou seja, sejam replicadas, por sobre outras áreas da imagem. Esta é a chamada propriedade um-para-um das funções de geração de distorção. Para resolver este problema, uma técnica de minimização de energia foi proposta para derivar funções de distorção um-para-um, em [Lee96]. Um amplo conjunto de primitivas de especificação de elementos é permitido, como pontos, polígonos e curvas, internamente reduzidos a uma coleção de pontos. Estes pontos são então utilizados para gerar uma deformação bidimensional de uma chapa retangular. A partir das restrições posicionais, é aplicada uma técnica de deformação baseada na representação de energia da distorção de cada ponto e na minimização das suas somas, o que possibilita a geração de distorções naturais, por ser baseada em princípios físicos. Todavia, um alto custo computacional resulta da aplicação da técnica.

Já em [Wolberg95] um novo método de geração de deformação foi apresentado, mais simples e de muito mais rápido processamento que o de minimização de energia anterior. O MFFD é uma extensão da deformação de forma-livre (FFD ou *Free-form deformation*), gerando uma hierarquia de treliças (*lattices*) de controle para satisfazer precisamente as restrições posicionais. Para garantir a propriedade um-para-um, foi introduzida uma condição suficiente para que a superfície 2D cúbica B-spline seja um-para-um. O MFFD gera distorções um-para-um e C^2 -contínuas, que resultam em distorções de imagem fluidas. Um exemplo de resultado pode ser visto na Figura 23. Há um aparente ganho em nitidez na imagem intermediária, em comparação com a técnica tradicional de distorção-fusão, mas o foco da abordagem é no processo, oferecendo melhor performance e evitando problemas introduzidos pelas técnicas anteriores.

Figura 23: Multilevel Freeform Deformation (MFFD)



Fonte: [Wolberg96]

[Wolberg95] conclui a sua análise delineando que trabalhos futuros envolveriam a metamorfose automática entre classes de imagens, como as de imagens de rosto, assim como o metamorfose entre sequências de imagens e não apenas entre duas imagens.

3.3. Metamorfose de Imagens baseado em Interpolação Morfológica

[Iwano2002] propôs uma nova abordagem para a metamorfose de imagens coloridas, baseado numa combinação de ferramentas de processamento de imagens através de morfologia matemática e filtragem linear. A metamorfose em si é gerada a partir de interpolação morfológica, aplicando o cálculo da mediana morfológica, e a geração sucessiva de medianas morfológicas geraria a sequência de quadros da transformação de uma imagem de entrada em outra. As ferramentas lineares são propostas para melhorar a qualidade visual da sequência de metamorfose. A principal vantagem desta técnica é a dispensa da disposição de pontos de controle por um operador humano.

A mediana morfológica foi introduzida para imagens no formato binário e mosaico [Meyer96] e posteriormente estendida para imagens em tons de cinza [Beuch98]. O cálculo utiliza os operadores morfológicos de erosão e dilatação.

Seja G a imagem de saída, ε o operador de erosão, δ o operador de dilatação, P um domínio de imagem, F e G as duas imagens a serem processadas. A operação de mediana morfológica, segundo [Beuch98], será definida conforme a Equação 7.

Equação 7

$$M(F, G) = \sup_{\forall \lambda} \{ \inf [\delta^{(\lambda)}(\inf(F, G)), \varepsilon^{(\lambda)}(\sup(F, G))] \}$$

onde $\lambda = 1, 2, \dots$ são valores inteiros crescentes. Os símbolos '*sup*' e '*inf*' representam imagens de *supremum* e *infimum*, termos definidos pelo autor respectivamente conforme a Equação 8 e a Equação 9.

Equação 8

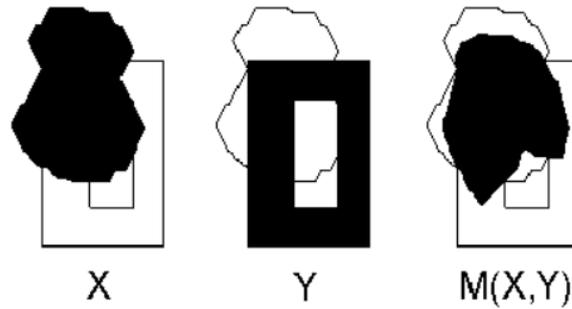
$$G = \sup(F_1, F_2) \Leftrightarrow \forall p \in P \ G(p) = \max\{F_1(p), F_2(p)\}$$

Equação 9

$$G = \inf(F_1, F_2) \Leftrightarrow \forall p \in P \ G(p) = \min\{F_1(p), F_2(p)\}$$

Importante notar que, no caso de imagens binárias, a interseção das duas imagens não pode ser vazia [Beuch98]. Um exemplo de resultado da mediana morfológica binária é apresentado na Figura 24.

Figura 24: Mediana Morfológica Binária $M(X,Y)$, gerada a partir das imagens X e Y



Fonte: [Iwano2002]

3.4. Metamorfose colorida por morfologia matemática

Marcin Iwanowski e Jean Serra têm sido nomes de referência na área de morfologia matemática, dentre os pioneiros na aplicação de medianas morfológicas para produção de metamorfose de imagens. Em [Iwano99] os dois pesquisadores apresentaram uma técnica de metamorfose de formas chamada “Mediana Morfológica Distorcida” (*warped morphological median*). O processo consiste de três etapas: classificação, distorção e geração da mediana morfológica final.

Inicialmente uma classificação dos valores coloridos dos pixels é feita, baseados na importância visual dos canais de cor para a percepção humana. A comparação dos componentes de cor (nos espaços de cor cartesianos) requer a comparação de triplas numéricas, por exemplo os valores de R, G e B, considerando o espaço de cor RGB. A ordenação lexicográfica é uma forma popular de comparação neste tipo de situação, baseada na comparação sucessiva de elementos dos vetores começando com por aqueles de menor índice, pode ser expressa pela Equação 10, Equação 11 e Equação 12.

Equação 10

$$P < Q \Leftrightarrow \exists i: 0 < i \leq n; \forall k: 0 < k < i, p_k = q_k \wedge p_i < q_i$$

Equação 11

$$P = Q \Leftrightarrow \exists i: 0 < i \leq n, p_i = q_i$$

Equação 12

$$P > Q \Leftrightarrow \exists i: 0 < i \leq n; \forall k: 0 < k < i, p_k = q_k \wedge p_i > q_i$$

Sendo $P = [p_1, p_2, \dots, p_n]$ e $Q = [q_1, q_2, \dots, q_n]$ dois vetores no espaço vetorial n -dimensional.

No caso do espaço de cor RGB, o componente R (vermelho) tem maior peso na classificação do que o componente G (verde) que por sua vez tem maior peso que o componente B (azul), o menos importante caso a ordenação tradicional dos elementos de cor seja utilizada. Para corrigir o que consideram uma distorção, os autores propõem o uso de um novo espaço de cor, chamado “espaço de cor comparativo”, que utiliza ordenação lexicográfica. Utilizando combinação linear, a Equação 13 é proposta.

Equação 13

$$[v_1 v_2 v_3]^T = T_1 \cdot [rgb]^T$$

Sendo $V = [v_1 v_2 v_3]^T$ um vetor no espaço vetorial comparativo e $[rgb]^T$ é um vetor no espaço de cor RGB.

Considerando a ordem de importância dos canais de cor para a percepção humana, em que o componente mais associado à luminância, o G, tem o maior peso, seguido do R e finalmente do B. Desta forma uma matriz de transformação para a ordenação GRB é obtida (Equação 14).

Equação 14

$$T_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Para melhor ponderar o peso da luminância a partir dos componentes de cor, é proposta a Equação 15.

Equação 15

$$lum = 0.3r + 0.6g + 0.1b$$

Os experimentos de [Iwano99] demonstraram que esta fórmula produziu os melhores resultados visuais nos operadores morfológicos, quando comparados com os resultados da aplicação da chamada “Norma Manhattan”, que calcula o comprimento do vetor multiplicando igualmente cada componente por 0,33 e somando os resultados.

Na sequência a distorção das imagens de origem de destino é gerada, partindo de um conjunto de pontos de interesse – por exemplo, olhos, boca, nariz – definidos correspondentemente nas duas imagens através de um processo supervisionado. O objetivo é o de controlar e garantir a vinculação destes elementos, já que a mediana morfológica é calculada globalmente, não permitindo associações de elementos distantes, sem interseção e muitas vezes com outras áreas distintas entre eles.

Com o objetivo de alinhar os pontos de interesse das imagens de origem e destino, o método de distorção-fusão bi-linear (*bilinear mesh-warping*) é o escolhido pelo artigo. Os pontos de controle são deslocados nas duas imagens para uma posição mediana, conforme o seguinte algoritmo:

Sejam I,J as imagens iniciais, n o número de pontos de controle, $C_1 = \{p_1, .. , p_n\}$, $C_j = \{q_1, .. , q_n\}$ os conjuntos de pontos de controle em I e J

1. Crie o novo conjunto $C_k = \{r_1, .. , r_n\}$ contendo as posições intermediárias dos pontos de controle de C_1 e C_j : para cada $i = 1, .. , n$ $r_i = 0.5(p_i + q_i)$
2. Desloque I de C_1 para C_k (resultado: I')
3. Desloque J de C_j para C_k (resultado: J')
4. Calcule a mediana morfológica de I' e J'

A geração da mediana morfológica final é feita entre as imagens I' e J' geradas. Os resultados podem ser observados na Figura 25.

Figura 25: Mediana Morfológica Distorcida; (a) imagem de origem I. (b) imagem de destino J. (d) e (f) são (a) e (b), respectivamente, com a grade de pontos de controle sobreposta. (c) imagem mediana de (a) e (b). (e) Mediana Morfológica Distorcida.

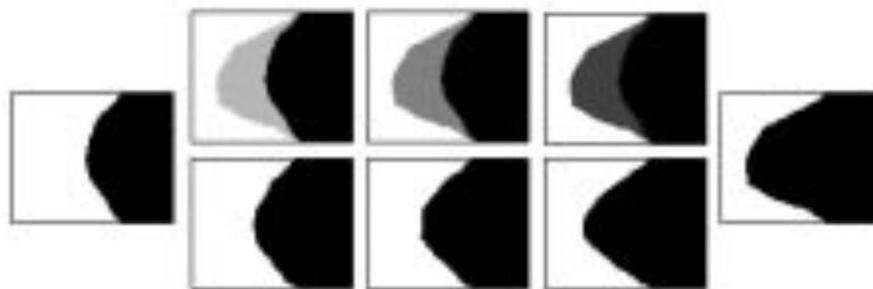


Fonte: [Iwano99]

Para gerar uma sequência de imagens para animação, novas medianas morfológicas são geradas entre o novo quadro intermediário e as imagens originais, com o detalhe de que a posição mediana dos pontos de controle é armazenada e aplicada automaticamente. Devido à geração sucessiva dos quadros intermediários partindo das duas imagens iniciais, o número de imagens em uma sequência final será sempre de uma potência de dois.

Neste artigo os autores propõem a possibilidade do uso de métodos automáticos de detecção de pontos de interesse (*feature points*), mas não se aprofundam a respeito de como o processo seria feito. Chama a atenção que o método de distorção-fusão proposto como passo anterior à geração da mediana morfológica é basicamente uma metamorfose “clássica” que, como discutido, resulta em perda de nitidez visual. Note-se que a mediana morfológica consegue efetivamente transformar contornos mantendo a nitidez, como demonstrado na Figura 26.

Figura 26: Interpolações morfológicas sucessivas, na sequência de baixo. Na sequência de cima, sequência sucessiva de fusões (cross dissolve).



Fonte: [Iwano99]

Entretanto, esta vantagem é, em grande parte, perdida pela interpolação bi-linear executada na etapa de pré-processamento deste método, como é aparente no desfoque do item (e) da Figura 25.

3.5. Considerações

As técnicas de morphing apresentadas por [Beier92], [Wolberg95] e [Lee96] apoiam-se sempre, ao final do processo, na fusão simples para a geração final das imagens. Por sua vez, a fusão possui a característica de gerar “fantasmas” em áreas de diferentes contrastes nas imagens interpoladas, sem efetivamente gerar contornos intermediários, como pode ser visto na Figura 26. Esta limitação impõe forte dependência no pré-processamento da distorção, exigindo grande número de pontos de correspondência e dificultando qualquer tipo de automatização do *morphing* facial.

Dentre os trabalhos apresentados, considerando a finalidade do presente trabalho, distinguem-se os de [Iwano99] e [Iwano2002]. Descrita em [Iwano99], a técnica de Mediana Morfológica Distorcida apoia-se na possibilidade de pré-processamento para alinhar recursos faciais. Todavia, a abordagem adotada para tal pelos autores foi a da distorção por malha, idêntica à de [Wolberg95], que produz os artefatos de desfoque do filtro bilinear nas imagens originais antes do cálculo da mediana morfológica.

Posteriormente [Iwano2002] descreveu uma técnica de processamento de medianas morfológicas coloridas baseadas na importância de percepção visual, aplicável em imagens coloridas. No quebra-cabeças restou apenas desenvolver uma abordagem de alinhamento

destes recursos que não gere os artefatos de perda de nitidez inerentes ao processo de distorção-fusão.

4. Método Proposto

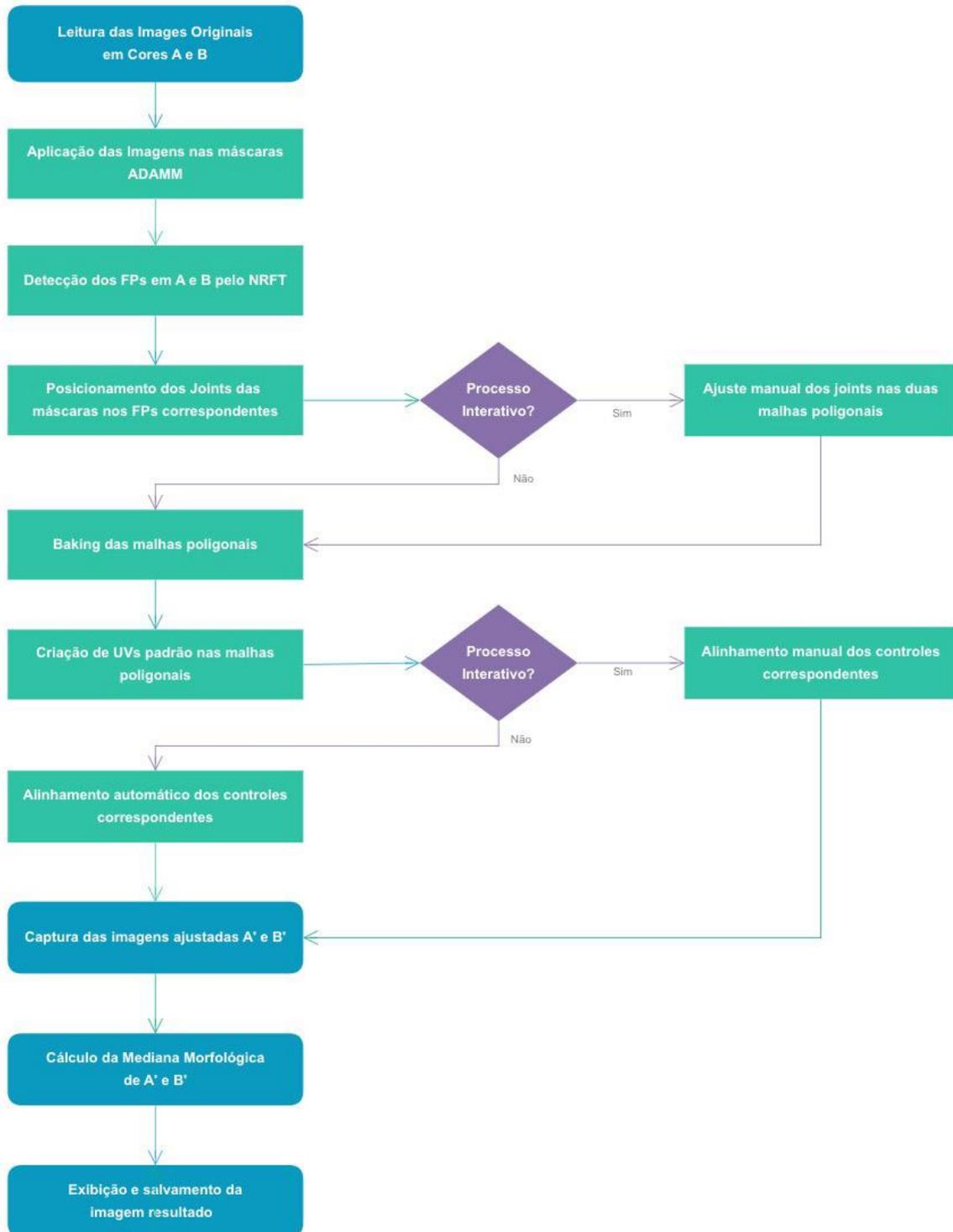
Neste capítulo será apresentado a técnica proposta para a metamorfose facial por detecção e alinhamento automático de características e morfologia matemática (*Automatic Detection and Alignment Morphology Morphing - ADAMM*).

4.1. Fluxograma

Na Figura 27 é apresentado o fluxograma do método proposto, o ADAMM. No modo semi-automático da aplicação desenvolvida, após a etapa de alinhamento das RoIs, o operador humano assume total controle sobre as regiões e nós das malhas, possibilitando o ajuste das proporções da face resultante, assim como eventuais correções dos limites das RoIs detectadas automaticamente pelo método NRFT.

No modo automático o processo é não-supervisionado, todos os parâmetros são definidos para seus valores sugeridos e o processo ocorre sem intervenção do operador.

Figura 27: Fluxograma do Método Proposto



Fonte: Autor

4.2. Descrição do Método

4.2.1. Visão Geral

A mediana morfológica mescla ou combina proporcionalmente pixels entre duas imagens, relacionando-os por sua proximidade e similaridade de luminância ou crominância de acordo com critérios pré-estabelecidos. No caso da metamorfose de imagens de rostos, se os dois não forem pré-alinhados, teremos situações como uma orelha sendo mesclada a um olho na outra imagem, supondo que os dois estejam na mesma posição relativa da imagem.

Para resolver esta questão, [Iwano2002] sugeriu um processo manual de marcação dos pontos de interesse correspondentes em ambas as imagens, seguido de uma distorção (*warp*) bilinear que deslocaria os marcadores para as posições intermediárias entre os pares correspondentes. Além do processo manual de marcação ser demorado e pouco prático, a distorção bilinear acarreta em perda significativa de nitidez em ambas as imagens, o que não pode ser revertido por um processo subsequente de mediana morfológica.

O método aqui proposto objetiva um avanço significativo não apenas no sentido de eliminar a necessidade do processo supervisionado de marcação das RoI (*Regions of Interest* ou Regiões de Interesse), mas também na preservação da nitidez nas RoI das imagens resultante do processo.

Para tanto, as regiões de interesse são movidas porém nunca deformadas de forma não-proporcional ou tendo sua escala aumentada, o que conforme [Ken90] levaria a uma perda de frequência do sinal original e, conseqüentemente, menor nitidez pelo demonstrado em [Batten2001].

As áreas externas às RoI são deformadas pelo método de distorção bilinear, o que gera deformações não-conformais. Todavia, estas regiões alheias às RoIs comumente envolvem cabelo, vestuário e o plano de fundo da imagem, sendo de menor relevância para o reconhecimento facial. Ainda assim, foi aplicada uma parametrização preservadora de ângulos [Shef2006] para minimizar as distorções.

4.2.2. Detecção Facial e dos Pontos de Interesse

Existem diversos métodos de detecção facial e de pontos de interesse ou recursos da face (*face features*). Para o presente trabalho, foram adotados como critérios de avaliação os princípios de [TRIVEDI2009]. O método deve ser preciso, monocular (ou seja, funcional a partir de uma única imagem), autônomo, flexível para identidade e iluminação, independente de resolução e capaz de detectar a rotação da cabeça. A multipessoalidade do método, ou seja, a característica de identificar vários rostos na mesma imagem, é dispensável na aplicação aqui proposta, já que é restrita ao morphing de duas imagens com apenas um rosto em cada. Também pela natureza da aplicação proposta, o método não necessita rodar em tempo real, ou seja, a trinta quadros por segundo ou mais.

Foi inicialmente avaliado o OpenFace, proposto por [Baltrusaitis2016] e baseado no Cambridge Face Tracker (CLM-Framework). Este framework possui ótima precisão e estabilidade nos pontos de interesse, mesmo em condições precárias de iluminação. Como pontos desfavoráveis, durante os testes o OpenFace comumente falhou na detecção do lábio inferior, além de oferecer uma licença de uso um pouco restritiva.

Foi também testado o método NRFT (Non-Rigid Face Tracking), parte do EHCI (*Enhanced human computer interface through webcam image processing library*), conforme apresentado em [Baggio2012]. Durante a avaliação o OpenFace foi mais consistente em aplicações de tempo-real, com menor perda de rastreamento dos pontos (*drifting*) à medida que a cabeça é rotacionada para ângulos extremos. Como o método proposto não tem natureza tempo-real, o método NRFT foi o escolhido.

4.2.3. Linha do cabelo, área da testa e queixo

O algoritmo de detecção de recursos faciais adotado, assim como na maioria dos algoritmos similares, não detecta recursos acima da linha da sobrancelha. Esta limitação inviabilizaria o alinhamento automático das regiões superiores do rosto das pessoas nas fotos. Para resolver este problema, aplica-se uma definição do desenho anatômico, conforme [Loomis56]. seja h a altura da linha horizontal que atravessa as sobrancelhas até a linha

horizontal da base do nariz, h será também a altura da linha horizontal das sobrancelhas até a linha da raiz do cabelo (*hair line*), no topo da testa.

A linha da base da região da testa foi definida como de um terço da altura de h , para garantir um espaçamento adequado entre os retângulos e funcionar com a maior variedade possível de formato de rostos.

Por último, foi definida a metade da altura entre a base da boca e a base do queixo como linha de topo da região do queixo, pelos mesmos motivos delineados em relação à altura da região da frente.

4.2.4. Conversão do Espaço de coordenadas 2D para o 3D

Uma vez mapeados os FPs dos dois rostos que alimentarão o processo de *morphing*, é necessário mapeá-los do espaço bidimensional do OpenCV para posições no espaço tridimensional para onde os nós serão deslocados. Para tanto, a biblioteca OpenCV oferece alguns recursos de conversão de espaço de coordenadas baseados em calibração da câmera virtual 3D da biblioteca. Este processo é projetado para alinhamento de posições detectadas a partir de imagens capturadas por uma câmera real, consequentemente sendo complexo demais para a aplicação do método proposto por este trabalho. Considerando que os nós são registrados por uma câmera virtual ortográfica e os pontos detectados no OpenCV estão num plano bidimensional, é necessário apenas aplicar dois processos sucessivos de interpolação linear. Foi desenvolvido o seguinte cálculo para realizar este mapeamento:

De acordo com [Meijering2002], a fórmula de interpolação linear que obtém um valor t desconhecido, sabidos os mínimos e máximos do seu eixo ortogonal s , quais sejam s_0 e s_1 , e os mínimos e máximos do eixo t , quais sejam t_0 e t_1 , é dado pela Equação 16.

Equação 16

$$t = t_0 + (t_1 - t_0) \frac{(s - s_0)}{s_1 - s_0}$$

Sejam s_0 e s_1 as coordenadas horizontais das extremidades esquerda e direita da imagem no espaço 2D do OpenCV, e um ponto no mesmo sistema de coordenadas, no intervalo (s_0, s_1) e com coordenada horizontal s_n . O valor t no intervalo $(0, 1)$ correspondente a s_n é encontrado pela aplicação da Equação 17.

Equação 17

$$t = \frac{s_n}{s_1 - s_0}$$

Sejam x_1 e x_2 as coordenadas horizontais conhecidas dos nós nas extremidades esquerda e direita da malha 3D. Seja x_n a coordenada horizontal 3D procurada para uma determinada posição no espaço 2D do OpenCV, no intervalo (x_1, x_2) . O valor de x_n é obtido pela Equação 18.

Equação 18

$$x_n = (x_1 + x_2) \cdot t$$

Aplicando a Equação 18 à Equação 16, encontramos a Equação 19.

Equação 19

$$x_n = (x_1 + x_2) \cdot \left(\frac{s_n}{s_1 - s_0}\right)$$

que é a fórmula para o mapeamento da coordenada em um eixo do espaço OpenCV para o espaço tridimensional. A mesma fórmula pode ser aplicada ao eixo y e uma vez encontrados os dois valores temos a coordenada completa convertida para o espaço 3D, registrado por uma câmera virtual ortográfica.

4.2.5. Métodos de Deformação

Para gerar formas realísticas, as deformações devem ser tão rígidas quanto possível, conforme [Schaefer2006]. O ideal é que os pixels que compõem uma área contígua e sólida da imagem não sejam alterados quando esta área for movida. Este é também um dos objetivos principais deste trabalho, visto que qualquer tipo de deformação interna em uma das áreas de interesse gerará interpolação dos pixels e, conseqüentemente, perda de nitidez.

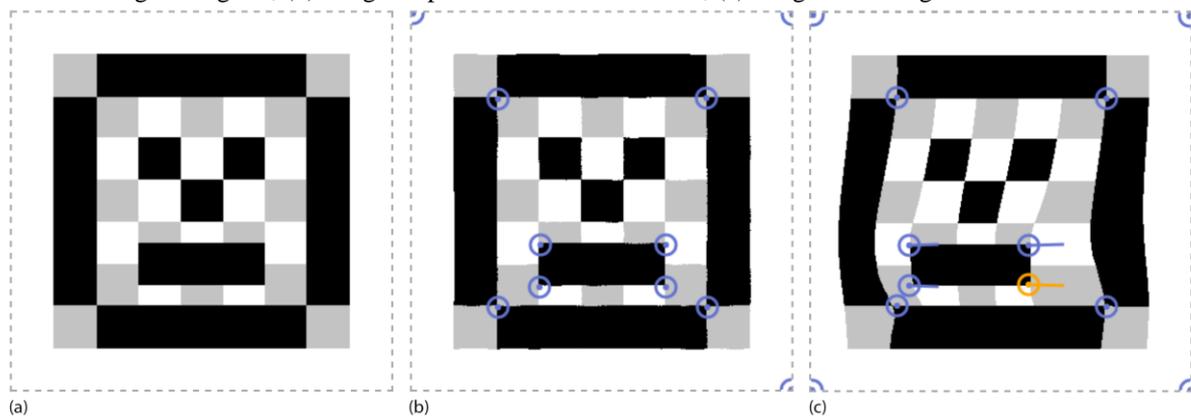
[Schaefer2006] afirma em seu artigo que ao utilizar o método de [Bookstein89] as formas testadas sofreram de escalonamento não-uniforme e inclinações visuais, o que seria indesejável em muitas aplicações.

Ambos os autores supra-citados utilizaram deformação rígida, que abrange apenas transformação de posição, rotação e escala proporcional dos elementos. Ainda assim, não há

presunção de perfeita preservação de áreas internas, como nos métodos de transformação conformal em que a distorção angular é zero [Gu2003].

Durante a avaliação dos métodos de deformação, foram realizados testes em uma aplicação online¹ que implementa o método de mapeamento por peças de [Schaefer2006]. No primeiro teste foram observadas ondulações indesejáveis dentro das áreas transformadas e em quase todas as outras áreas da imagem, mesmo com o uso de pontos-âncora nos limites da imagem, como pode ser visto na Figura 28. Na Figura 29 é possível visualizar a ondulação nos contornos da foto, que deveria ser retilíneo, além de artefatos de distorção e descontinuidade na região do queixo da atriz na foto. Estes resultados comprovaram a natureza não-conformal [Shef2006] da técnica, que até então era considerada para alcançar o objetivo proposto por este trabalho.

Figura 28: Teste de deformação de região retangular utilizando [Schaefer2006]. (a) imagem original; (b) imagem apenas com os marcadores; (c) imagem com região interna deslocada



Fonte: Autor

¹ Image Warp Java demonstration application, disponível em: <http://chenxing.name/fun/imgwarp-js/>, acessado em 14/3/2016

Figura 29: Teste de deformação de rosto utilizando [Schaefer2006]



Fonte: Autor, imagem da base fornecida

4.2.6. Alinhamento das Regiões Faciais

Em muitos casos, as imagens utilizadas pelo *morphing* tem rostos com tamanhos diferentes. Para aproximar a escala dos rostos, a opção adotada foi a de reduzir o tamanho da maior região facial, em uma das imagens, até a largura da menor região facial dentre as duas imagens.

No caso do aumento de imagens mapa-de-bits, as aplicações de edição de imagens aplicam por padrão algum tipo de filtro bilinear, o que resulta em algum prejuízo na nitidez da imagem processada. Como o objetivo do método proposto é o de preservar a nitidez, apenas a redução de tamanho da imagem e de suas regiões internas apresentou-se como alternativa viável.

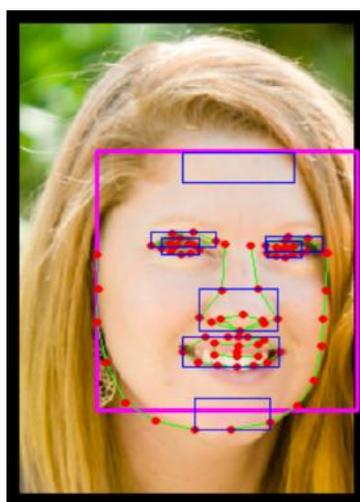
Note-se que atualmente o aplicativo *SharpMorpher* está limitado a trabalhar com imagens de mesma resolução, porém está prevista a possibilidade futura de utilizar imagens de diferentes resoluções.

4.2.7. Detecção dos recursos da face

Para realizar a detecção da região facial, foram testados dois classificadores faciais de OpenCV "*haar cascade frontal face*" e o "*lbp cascade frontal face*". Nos testes realizados com estes classificadores, ambos ofereceram resultados insatisfatórios. Na Figura 30 a área

detectada pelo *Haar Cascade Frontal Face* pode ser visualizada na cor magenta, enquanto as regiões de interesse detectadas pelo algoritmo NRFT estão em azul. É possível observar que parte da região do queixo não foi considerada pelo *Haar Cascade Frontal Face*, enquanto uma parte do cabelo foi capturada, excedendo os limites da face. Interessante notar que a estimativa de [Loomis56] adotada para encontrar a *linha do cabelo* no algoritmo proposto coincidiu exatamente com o que os classificadores ofereceram. Apesar do classificador *Haar Cascade Frontal Face* ser utilizado para a definição da área facial inicial pelo algoritmo NRFT, o NRFT utiliza uma série de outras relações de distância entre os FPs. Por este motivo, o método proposto utiliza a caixa delimitadora (bounding box) envolvendo todos os pontos classificados pelo NRFT para o enquadramento dos nós do *FaceFrame* na região facial.

Figura 30: Teste com classificador Haar Cascade (em Magenta) e marcadores do NRFT



Fonte: Autor; Foto original da St.Louis Post-Dispatch¹

4.2.8. Detecção dos FPs (NRFT)

[Baggio2012] apresenta um método de reconhecimento de recursos faciais, o “Non Rigid Face Tracking” (NRFT).

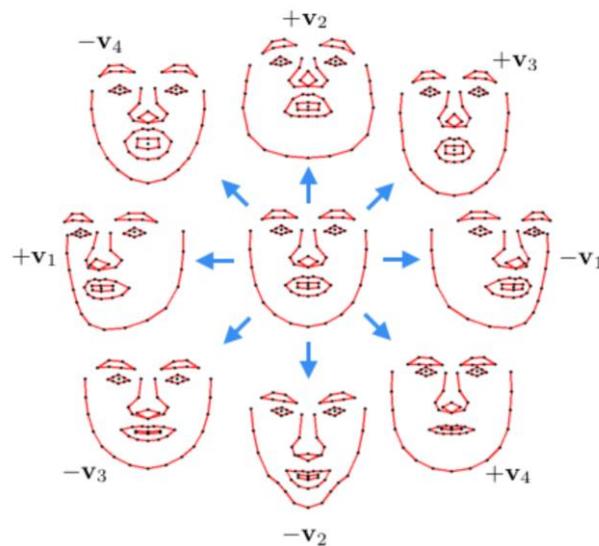
¹ St.Louis Mother-Daughter lookalike contest, disponível em:
<http://www.stltoday.com/lifestyles/relationships-and-special-occasions/parenting/aisha-sultan/erica-losito-and-ellie-weber-of-shrewsbury/article_5da630da-bb01-11e2-81f9-001a4bcf6878.html>.
Acessado em 11/8/2016

O NRFT apresenta uma descrição do modelo geométrico da estrutura facial, chamado de "shape model" (modelo de formas). Este modelo objetiva capturar variações de expressão entre formatos faciais distintos encontrados nas múltiplas etnias humanas.

Segundo os autores, a geometria facial pode ser descrita por dois elementos constituintes, um de transformação global ou rígida, que afeta toda a face, e outro de transformação local ou não-rígida. Os elementos não-rígidos correspondem aos recursos faciais como nariz, boca e olhos. As chamadas deformações locais consistem nas diferenças entre formatos de rosto devido às expressões faciais como as de sorriso e surpresa.

Inicialmente o algoritmo realiza uma remoção dos componentes da movimentação global rígida da face, referentes à escala, translação e rotação no plano da imagem. O algoritmo é capaz de detectar uma série de rotações da cabeça no espaço tridimensional, como pode ser visto na Figura 31.

Figura 31: Alguns modos de Variação Facial detectados pelo NRFT, por transformação local ou global.



Fonte: [Baggio2012]

No segundo estágio, as direções de variância oriundas das transformações rígidas são removidos dos dados, gerando um sub-espço de deformação ortogonal ao sub-espço de transformação rígida. Esta característica possibilita a computação simplificada e extração dos FPs.

O modelo descrito pelo NRFT é composto de pontos, que denominamos FPs (feature points) que têm por inferência linhas que os conectam.

Ainda segundo [Baggio2012] o número de FPs faciais detectados não aumenta a qualidade da detecção facial após cerca de cem FPs, além do aumento linear em complexidade computacional que se segue deste aumento. Por isto o NRFT limitou-se a detectar 76 FPs ou marcos faciais. O método entrega uma lista de vetores bidimensionais no sistema de coordenadas do OpenCV. No SharpMorpher, a anotação destes pontos com seus correspondentes faciais foi feita por uma estrutura de dados específica, chamada *Feature*. Esta classe é capaz de armazenar o índice, o nome do nó na máscara facial e as coordenadas do ponto tanto no sistema de coordenadas do OpenCV quanto no sistema de coordenadas da Unity3D.

Esta estrutura foi desenvolvida junto a um conjunto de métodos auxiliares para conversão entre os sistemas de coordenadas e mapeamento automático dos *Features* lógicos para os nós existentes na cena da Unity. Uma vez implementada, esta estrutura forneceu agilidade e confiabilidade ao processo de manipulação dos nós pelos demais métodos.

Note-se que apesar do método NRFT poder ser aprimorado com o uso de dados de treinamento específicos para um determinado conjunto de imagens, a versão do *SharpMorpher* utilizada para validação deste trabalho utilizou exclusivamente um modelo genérico disponível publicamente¹.

4.2.9. Construção da Máscara Facial “ADAMM”

Existem diversos algoritmos para a geração automática destas malhas a partir de um conjunto arbitrário de pontos, de uma forma comprovavelmente ideal, como [Boi2005] e [Chew93]. Entretanto, considerando que o conjunto de pontos derivados da detecção de FPs é limitado e conhecido, não fazia sentido uma geração procedural de topologia geométrica.

¹ Modelo Genérico de Treinamento de Reconhecimento Facial, disponível em:

<https://github.com/EnoxSoftware/FaceTrackerSample/blob/master/StreamingAssets/tracker_model.json>

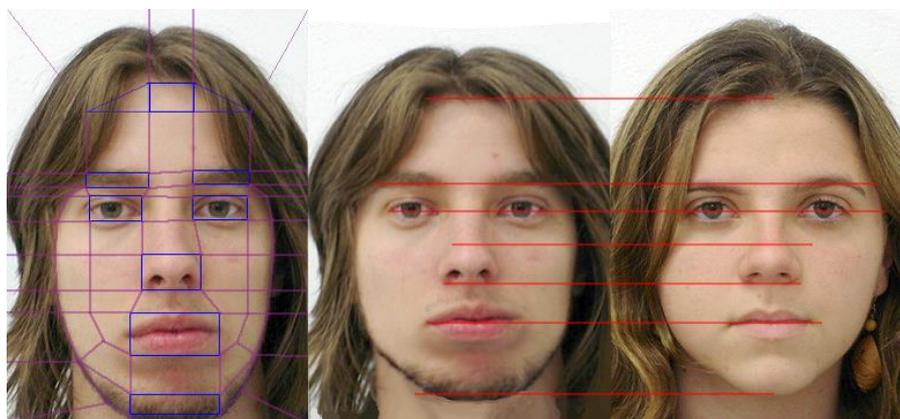
Acessado em 26/03/2016

Foram pesquisados modelos de máscaras poligonais para finalidades de mapeamento similares, como os “Modelos de Aparência Ativa” (*Active Appearance Models*) [Matt2004]. Todavia estes modelos apresentam um eixo central que, apesar de facilitar a triangulação, dificulta o controle e a distorção da malha, conforme os conceitos de topologia geométrica. Portanto, para o método descrito neste trabalho, foi desenvolvida uma máscara facial geométrica a partir do conjunto de FPs entregue pelo NRFT de [Baggio2012].

Um diagrama inicial definindo retângulos delimitadores das regiões de interesse, em azul, e as linhas de conexão entre os outros pontos de interesse, em laranja, pode ser vista na Figura 32. As RoI definidas foram a frente, sobrancelha esquerda, sobrancelha direita, olho direito, olho esquerdo, nariz e boca. Na mesma figura, as principais linhas de correspondência de recursos faciais são apresentadas em vermelho.

No teste apresentado na Figura 32 o algoritmo foi simplificado para apenas mover as RoI do maior rosto para a posição (considerando o centro) das RoI correspondentes no menor rosto. Na implementação final ambas as RoI correspondentes são movidas para a posição média entre estas.

Figura 32: Diagrama de mapeamento e alinhamento das RoIs

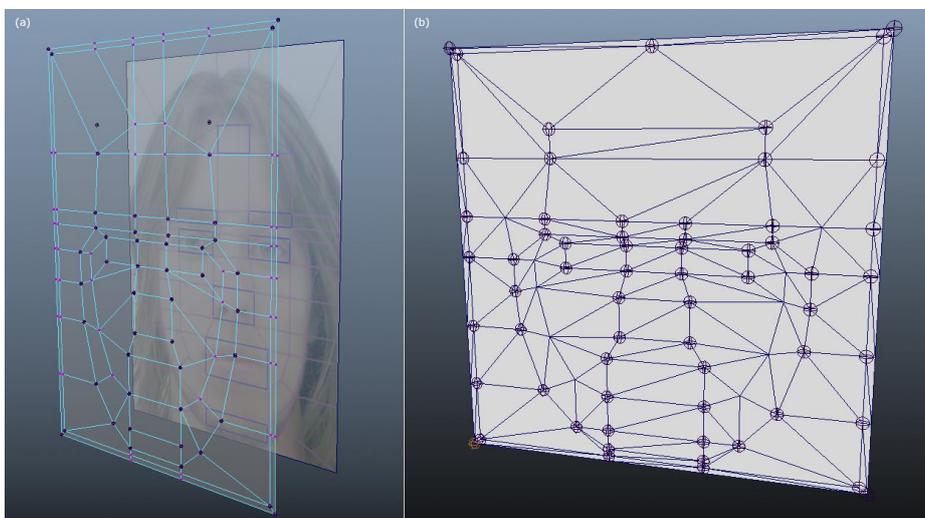


Fonte: Autor; Imagens da base de dados fornecida

Inicialmente a malha foi definida por quadrângulos (*quads*), respeitando as boas práticas topológicas de disposição das fileiras poligonais (*loops*), conforme [Vaughan2011]. As áreas externas à face foram delimitadas por linhas horizontais paralelas, minimizando distorções de ângulo e tornando a projeção de textura nestas áreas conformal, de acordo com [Shef2005].

Por último, a malha poligonal foi triangulada, processo que consiste na inserção de uma aresta interna em cada quadrângulo, resultando numa malha composta exclusivamente por triângulos. Este processo foi executado para garantir um fluxo consistente das arestas internas da malha, visto que polígonos com mais que três lados são subdivididos internamente de forma automática e arbitrariamente durante a importação no motor 3D utilizado (o Unity3D). Na Figura 33, podem ser visualizadas a primeira versão, construída sobre o diagrama de planeamento original, e a versão definitiva triangulada da malha 3D.

Figura 33: Máscara poligonal ADAMM; (a) Primeira versão (b) versão definitiva triangulada



Fonte: Autor

4.2.10. Esqueleto de deformação

Para possibilitar o reposicionamento das RoIs, nós (*joints* ou controles de influência de deformação) foram criados na mesma posição dos vértices da máscara facial. Foram também definidos **grupos de controle**, englobando múltiplos nós de uma ou mais RoIs, assim como dos demais FPs. Graças ao recurso de agrupamento de nós, é oferecido ao operador da aplicação um controle simplificado de reposicionamento de uma RoI inteira, ao invés de sempre necessitar mover os nós individualmente. Isto permite a rápida correção de eventuais falhas na detecção das áreas, além de possibilitar a simulação de proporções faciais típicas de uma determinada faixa etária.

Para organizar as dezenas de nós utilizados, foi aplicado um padrão de nomenclatura padrão de "rigging", que é o processo de criação de marionetes digitais tridimensionais. Os nomes são dados do ponto de vista do "personagem", ou seja, o sufixo "L" (de *Left*, esquerda)

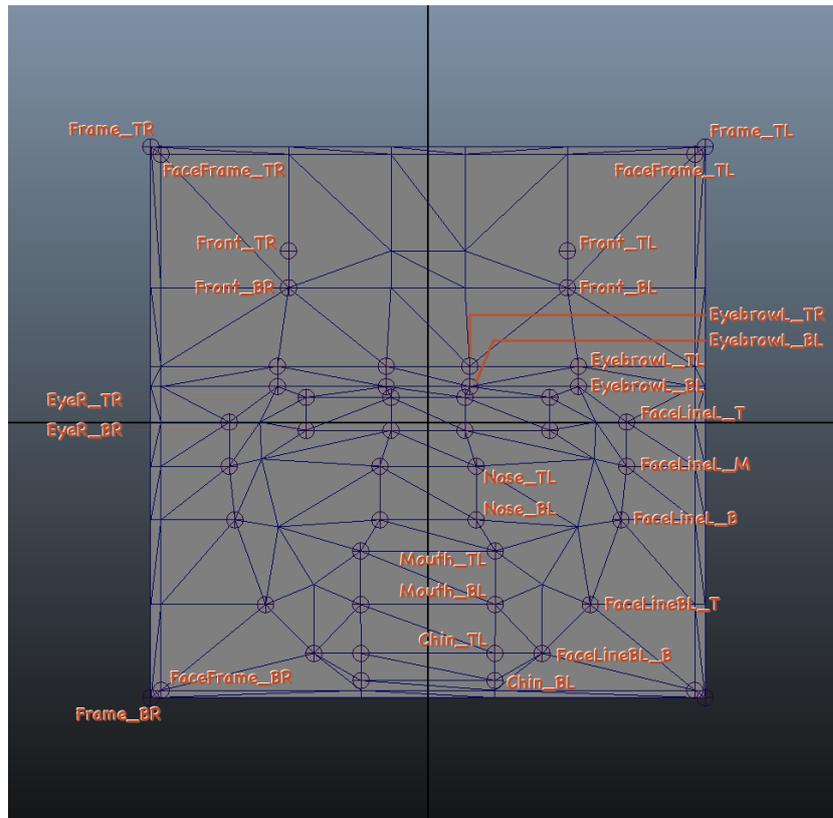
é dado ao lado que, do ponto de vista do observador, é o direito. Os sufixos "T" (*Top*) e "B" (*Bottom*) se aplicam, respectivamente, aos nós que estão no topo e na parte de baixo das suas regiões. Em alguns casos, o sufixo "M" (*Middle*) foi utilizado para nós que estão no meio de uma região.

Para nomear os controles foi utilizado um padrão similar. *Frame* foi o nome dado ao quadro externo, sempre ancorado aos limites da malha e cujos nós correspondem aos limites da imagem original projetada. Estes nós nunca são movidos no processo de alinhamento automático ou de edição interativa da máscara.

O *FaceFrame*, ou quadro facial, corresponde à caixa delimitadora (*bounding box*) que envolve todos os FPs detectados. O controle do *FaceFrame* pode ser redimensionado ou ter os seus nós movidos para alinhar os limites externos do rosto das duas imagens a serem combinadas. *Front* é a região da frente ou testa da pessoa na imagem. *EyeBrow* corresponde à região da Sobrancelha, havendo um controle para cada um dos lados da face. *FaceLine* é o termo que foi adotado para a linha de silhueta ou contorno lateral da face, sendo este grupo composto por três nós da lateral e dois nós da diagonal inferior da cabeça, considerando a sua visão frontal.

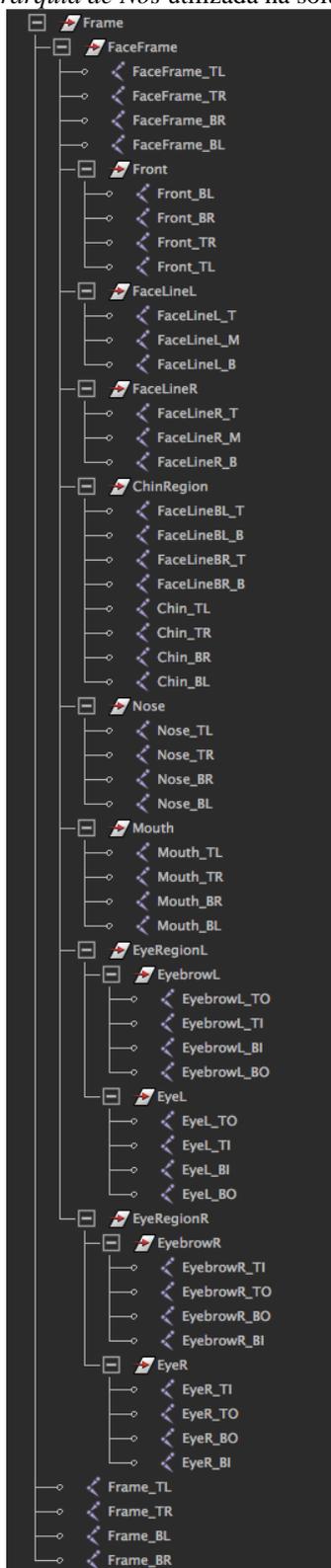
Na Figura 34 o nome de cada nó é exibido ao seu lado ou por uma seta indicativa. Estão omitidos alguns nomes que são espelhados trocando apenas o sufixo indicativo do lado do rosto ("L" por "R"). A hierarquia completa dos nós, conforme disposição na aplicação 3D, define que nós funcionam como grupos de macro-ajuste, transformando relativamente todos os nós subordinados. Esta hierarquia é exibida na Figura 35.

Figura 34: Nomenclatura dos Nós sobre a máscara facial



Fonte: Autor

Figura 35: Hierarquia de Nós utilizada na solução proposta

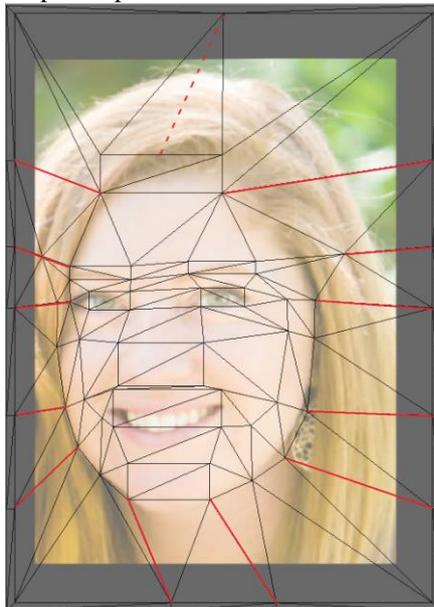


Fonte: Autor

4.2.11. Âncoras de Borda (*Border Anchors*)

Para manter a forma retangular nos polígonos da área exterior à face, não basta ajustar os nós que deformam a malha para as posições-chave detectadas. Isto porque o deslocamento destes nós na direções vertical e horizontal em muitos casos vão gerar quadriláteros com ângulos diferentes de noventa graus, como pode ser visto na Figura 36.

Figura 36: *Distorção de Linhas de borda*; as linhas diagonais, em vermelho, foram geradas pelo reposicionamento dos nós da máscara facial

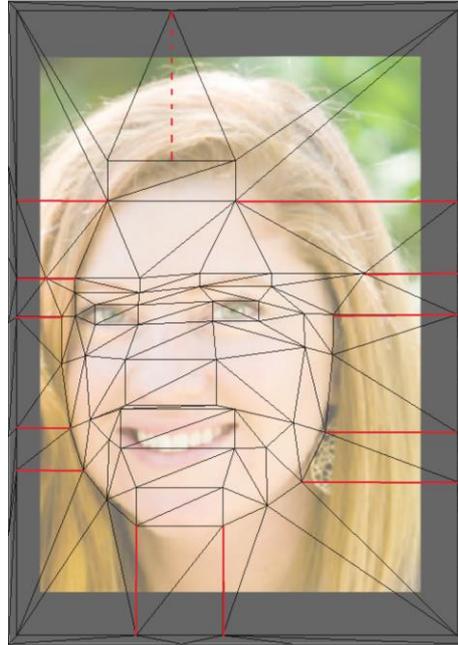


Fonte: Autor; Foto original da St.Louis Post-Dispatch¹

Para evitar estas conformações, que provocam distorções angulares ou *shear* [Shef2006], os nós adicionais nas bordas da área facial ("FaceFrame") devem alinhar-se horizontal ou verticalmente para acompanhar a posição do seu nó facial correspondente. Para este fim, um componente adicional para Unity foi desenvolvido em C#, com a função de copiar a coordenada apenas horizontal ou vertical de outro nó para a posição do nó em que o componente é adicionado. Após uma rápida montagem de cena pareando os nós através deste componente, o resultado visto na Figura 37 foi obtido.

¹ St.Louis Mother-Daughter lookalike contest, disponível em:
<http://www.stltoday.com/lifestyles/relationships-and-special-occasions/parenting/aisha-sultan/erica-losito-and-ellie-weber-of-shrewsbury/article_5da630da-bb01-11e2-81f9-001a4bcf6878.html>.
Acessado em 11/8/2016

Figura 37: Aplicação de âncoras de borda



Fonte: Autor; Foto original da St.Louis Post-Dispatch¹

4.2.12. Alinhamento dos recursos da face

Antes da aplicação da mediana morfológica, é realizado o alinhamento das regiões de interesse (RoI) referentes aos principais recursos faciais. Este passo faz-se necessário para garantir a eficiência dos resultados da mediana morfológica, cuja operação baseia-se na proximidade de pixels similares.

Após a detecção dos FPs, os controles são posicionados no centro das regiões correspondentes. O processo de alinhamento dos controles das RoI pode ocorrer manualmente, caso em que o operador posiciona-os manualmente para encontrar a posição intermediária que mais se adeque ao resultado pretendido. No modo automático, os controles correspondentes nas duas imagens são movidos para a posição intermediária entre eles.

No programa também é possível reposicionar os nós individuais e controladores de RoI sem deformar a imagem respectiva, num modo de operação chamado “*slide*”. Assim é possível corrigir qualquer imperfeição na detecção dos FPs pelo algoritmo do NRFT.

¹ St.Louis Mother-Daughter lookalike contest, disponível em:
<http://www.stltoday.com/lifestyles/relationships-and-special-occasions/parenting/aisha-sultan/erica-losito-and-ellie-weber-of-shrewsbury/article_5da630da-bb01-11e2-81f9-001a4bcf6878.html>.
Acessado em 11/8/2016

4.2.13. “Baking” geométrico

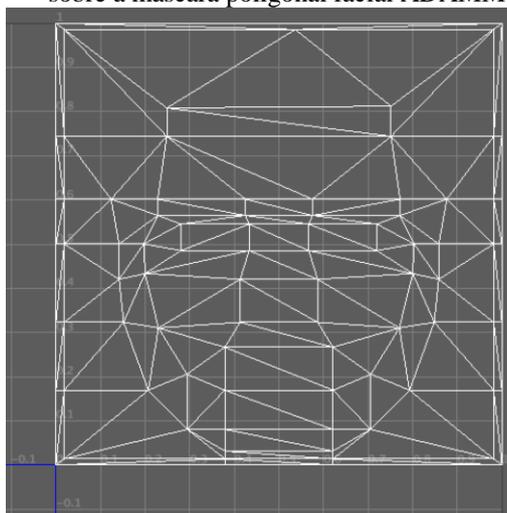
O processo chamado de “queima” ou *baking*, corresponde a tirar uma fotografia instantânea do estado atual da geometria e gerar a partir daí uma nova malha geométrica, independente. Desta forma é possível obter uma versão da malha em que a matriz de transformação de cada vértice corresponda à sua coordenada no mundo. Esta etapa é imprescindível para o método apresentado, já que a geração de UVs realizada na etapa subsequente utiliza a matriz de coordenadas dos vértices para calcular a parametrização.

4.2.14. Aplicação de UV padrão

Uma vez criada uma cópia da malha com a nova configuração dos vértices, é calculada a parametrização UV automática (*Auto UVs*). Este processo faz a correspondência de todos os vértices da malha 3D processada de forma a ocupar todo o espaço de coordenadas UV, que vai de -1 a 1 em U e de -1 a 1 em V.

Note-se que a máscara poligonal 3D é reduzida na dimensão horizontal, para enquadrar à proporção da imagem, mas isso não afeta as coordenadas UVs que sempre permanecem ocupando todo o espaço disponível. O resultado da parametrização pode ser visto na Figura 38.

Figura 38: Parametrização UV automática gerada sobre a máscara poligonal facial ADAMM



Fonte: Autor

4.2.15. Captura das imagens

Uma vez alinhados os recursos faciais, as máscaras poligonais são capturadas por câmeras virtuais no espaço 3D. Estas câmeras são ortográficas e tem seu campo visual definido para capturar exatamente os limites da malha 3D que observam. A resolução das câmeras virtuais também é definida para a mesma resolução das imagens de entrada. Existem três câmeras virtuais configuradas desta forma na cena, uma para as máscaras de cada imagem original (denominadas “A” e “B”) e uma para a máscara da imagem de resultado (de nome “Result”).

Para evitar desfoque das imagens, é sempre definido o filtro de *nearest neighbor* (vizinho mais próximo) nas texturas originais e processadas, antes de aplicá-las nas malhas geométricas, o que garante a correspondência 1:1 dos pixels de origem e dos pixels capturados pelas câmeras 3D.

As imagens capturadas pelas câmeras A e B, que aqui chamaremos de A’ e B’, são arquivadas em disco para a etapa seguinte do método.

4.2.16. Processamento das Medianas Morfológicas

Por fim, a mediana morfológica é calculada sobre a área total das imagens A’ e B’, conforme algoritmo publicado em [Iwano99], e armazenada na imagem de resultado, a partir da câmera virtual “Result”. Além de armazenada em disco, a imagem resultado também é exibida na máscara de resultado, para oferecer *feedback* imediato ao operador.

Os parâmetros do cálculo de mediana morfológica sugeridos de forma automática pelo *software* são o $\Lambda = 1$ e o Tamanho de Elemento = 6. Estes parâmetros podem ser ajustados pelo operador se assim desejado. Os experimentos realizados com vários valores nestes parâmetros indicaram que tamanhos do elemento morfológico maiores resultam em imagens com maior nitidez porém com menor fusão entre os pixels das imagens originais. Neste caso, o aspecto das áreas resultantes passa a ser similar a de várias micro-regiões copiadas e coladas entre as imagens originais, o que implica em maior nitidez do resultado.

Quanto ao Λ , ou número de repetições do processo, este parâmetro mostrou-se mais eficiente com valores menores, próximos a 1. Valores mais altos resultaram em menor fusão entre os pixels, o que é desejável em alguns casos.

5. Experimentos e Resultados Obtidos

5.1. Base de Imagens

É incomum artigos sobre metamorfose de imagens apresentarem referência às bases de imagem utilizadas, ainda que algumas imagens apresentadas sejam clássicas e de domínio público, como a foto de Lena Söderberg.

Para este trabalho, uma pequena base própria foi montada e disponibilizada online¹, utilizando fotos de personalidades públicas e seus filhos com similaridades fisionômicas. Além disso, foram adicionadas duas imagens da base de imagens públicas de alta-resolução “FEI Face Database”, do Laboratório de Processamento de Imagens do Centro Universitário FEI de São Paulo. Diferentemente das imagens de personalidades obtidas pela internet, todas as imagens da FEI Face Database foram obtidas em uma situação homogênea de iluminação e fundo neutro branco.

As imagens das personalidades públicas foram escolhidas dentre as inúmeras disponíveis online, sendo o critério primário de seleção a similaridade de orientação da cabeça. Este critério foi adotado dada a natureza bidimensional da técnica proposta, inadequada para pares de imagens com grandes diferenças de rotação da cabeça. Como critério secundário, a similaridade e uniformidade de iluminação dentre os pares de imagens foi buscada, por minimizar artefatos gerados por sombras presentes em apenas uma das fotos.

A resolução das imagens varia de 134 a 320 pixels horizontais, e de 118 a 400 pixels verticais. Como formato final das imagens da base disponibilizada foi escolhido o PNG, pela sua otimização, compatibilidade e preservação da qualidade original, sem perdas por compressão.

5.2. Tecnologia e Ferramentas

O programa *SharpMorpher* foi desenvolvido na plataforma Windows, utilizando o ambiente de desenvolvimento Visual Studio 2013 CE, que apresenta poderoso instrumental de depuração e manutenção. A linguagem de programação adotada foi a C# (“C-Sharp”), devido a suas modernas características de programação e extensibilidade.

¹ Base de imagens para Morphing, disponível em:
<<http://www.fluidplaystudios.com/adamm/>>

Como plataforma de interface com o usuário e motor 3D, foi adotada a *Unity3D*¹, pelas suas características de flexibilidade, capacidade de publicação multi-plataforma e custo gratuito para fins não comerciais. Foi utilizado dentro deste ambiente o plugin *OpenCV for Unity*², uma versão re-codificada do *Java OpenCV* para C#. Além de performance superior ao de outras soluções como o *EmguCV*³, o *OpenCV for Unity* é totalmente compatível com a publicação do aplicativo para as plataformas iOS e Android.

Utilizando a mesma plataforma e linguagem propostas, já havia sido desenvolvido, durante o processo de pesquisa inicial, programa de computador com a implementação da metamorfose por mediana morfológica [Iwano2002]. Os resultados então obtidos foram idênticos aos apresentados pelo artigo original, o que já cobriu boa parte do desenvolvimento do processo proposto e minimizou incertezas.

Como ferramenta de criação e edição das malhas poligonais, assim como de configuração do esqueleto e mapeamento de deformação pelos nós, foi utilizada a versão acadêmica 2016 do Autodesk Maya⁴.

Ademais, a ferramenta resultante deste trabalho, o *SharpMorpher*, será disponibilizada para a comunidade científica para fins de estudo, validação e pesquisa.

5.3. Análise de Nitidez

A modelagem numérica e exata do resultado da percepção humana, subjetiva por natureza, tem sido um grande desafio para as ciências exatas. Todavia, bons resultados tem sido obtidos em trabalhos como o de [Tiz2005], com o emprego de técnicas de Morfologia Matemática.

[Batten2001] realizou um extenso trabalho comparativo entre as técnicas de medição de nitidez, chegando à conclusão de que a Medida de variância seria a melhor técnica para a finalidade. Esta medida é calculada pelo somatório dos quadrados das diferenças entre cada

¹ Unity3D Editor, disponível em:

<<http://unity3d.com/get-unity/download>>. Acessado em 1/3/2016

² OpenCV for Unity, documentação disponível em:

<<http://enoxsoftware.github.io/OpenCVForUnity/doc/html/>>. Acessado em 19/11/2015

³ EmguCV OpenCV Wrapper, disponível em:

<<https://sourceforge.net/projects/emgucv/files/emgucv/2.4.0/>>. Acessado em 25/10/2015

⁴ Autodesk Maya, versão acadêmica gratuita disponível em:

<<http://www.autodesk.com/education/free-software/maya>>. Acessado em 8/8/2015

pixel e a intensidade média da imagem como um todo. Imagens focadas tem uma maior variação de intensidade que imagens desfocadas e borradas. O autor destaca como pontos fortes da técnica a simplicidade e eficiência da implementação, além da resistência a ruídos.

Foi utilizado o canal de luminância das imagens avaliadas, dado que, segundo [XinLi2008], o canal de luminância (G) é menos serrilhado que os outros dois canais.

Este algoritmo foi aplicado na rodada de experimentos entre a ferramenta proposta, o *SharpMorpher* e duas ferramentas com capacidade de detecção automática de recursos faciais, o FantaMorph e o Morph Thing. O FantaMorph é comercial, a sua versão gratuita de avaliação foi utilizada¹. O Morph Thing pode ser utilizado gratuitamente pela sua página na Internet². Cabe notar que o Morph Thing, diferentemente do FantaMorph, não opera de forma totalmente automática, pedindo ao usuário que marque os limites faciais ao submeter uma nova imagem.

5.4. Experimentos

Nesta seção serão descritos o protocolo experimental, os experimentos escolhidos e os resultados obtidos. Ao final é apresentada uma planilha comparativa dos resultados utilizando o SharpMorpher em relação aos de aplicativos comerciais existentes que aplicam técnicas tradicionais de *morphing* [Wolberg98].

5.4.1. Protocolo Experimental

Inicialmente vamos classificar a pesquisa. Esta é uma pesquisa aplicada ou tecnológica, que busca oferecer resultados práticos e imediatos na evolução da área de interesse, a de auxílio à confecção de imagens forenses de crianças desaparecidas. A natureza da pesquisa é experimental, processando as imagens de amostra nos softwares e aplicando técnicas objetivas e numéricas de medição de nitidez [Batten2001] dos resultados, sendo portanto a nitidez a variável ou parâmetro de estudo.

Foi realizado um estudo controlado randomizado, sendo a forma de abordagem quantitativa (numérica) e de complexidade analítica, ocorrendo a busca de validação dos

¹ FantaMorph, versão de avaliação disponível em:
<<http://www.fantamorph.com/download.html>>. Acessado em 17/11/2015

² Morph Thing, aplicativo online disponível em:
<<http://www.morphthing.com/>>. Acessado em 4/4/2016

questionamentos da pesquisa através de análise de causa-efeito dentre a técnica aplicada e variações nos valores numéricos resultantes.

Quanto aos procedimentos técnicos, o presente trabalho aplicou a pesquisa bibliográfica e a pesquisa laboratorial. Por fim, em termos do desenvolvimento em relação ao tempo, a pesquisa foi aplicada em um curto período de tempo, sendo portanto classificada como uma pesquisa seccional.

5.4.2. Experimento 1 – Faces Medianas

Para início dos experimentos, decidiu-se por definir um “ground truth”, ou valor de referência para os experimentos. Para isso, foram escolhidas as imagens chamadas “FACES Medianas”, conforme definição da empresa alemã que as gerou, a Beautycheck¹. Estas imagens foram construídas a partir do morphing de 32 fotos de homens e 64 fotos de mulher, seguindo a teoria de [Langlois90]. Note-se que para garantir a nitidez durante este processo, foram manualmente posicionados 250 pontos de correspondência em cada imagem, totalizando mais de 75.000 posicionamentos individuais.

As duas imagens de face mediana, uma de homem e a outra de mulher, possuem fundo neutro, pose frontal do rosto, olhos alinhados e resolução de 320x400 pixels. Estas características são desejáveis por evitarem problemas comuns aos processos da geração automática do morphing de rostos, como falhas de detecção de RoIs por deficiência de iluminação ou por poses laterais ou com angulações diferentes. Objetivou-se com isso avaliar a nitidez dos processos de morphing da forma mais pura possível. Os resultados obtidos nas aplicações testadas podem ser vistos na Figura 39.

A

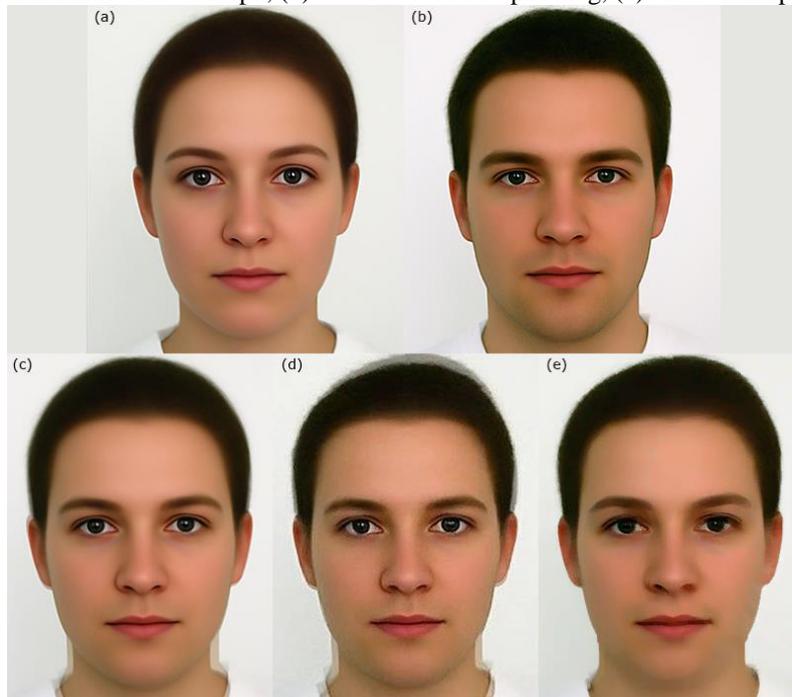
Fonte: Autor; Imagens originais da BeautyCheck

Figura 40 mostra em destaque como o método tradicional, aplicado nas imagens (c) e (d) da Figura 39, falha na fusão dos pixels da região do pescoço. O FantaMorph solucionou o morph do topo da cabeça sem o artefato de “imagem fantasma” (c), que pode ser visto no resultado do Morph Thing (e).

¹ Website da Beauty Check, durchschnittsgesichter (Average Faces). Disponível em: <<http://www.beautycheck.de/cmsms/index.php/durchschnittsgesichter>>. Acessado em 10/8/2016

Em contrapartida, o algoritmo ADAMM combinou as duas regiões citadas sem dificuldade e com nitidez de contornos visivelmente alta (e), o que foi comprovado pela avaliação numérica que se seguiu. O resultado do SharpMorpher apresentou nitidez de 256,01 pontos, 66,12% superior ao do Morph Thing (154,11 pontos) e 236,6% superior ao do FantaMorph (76,05 pontos).

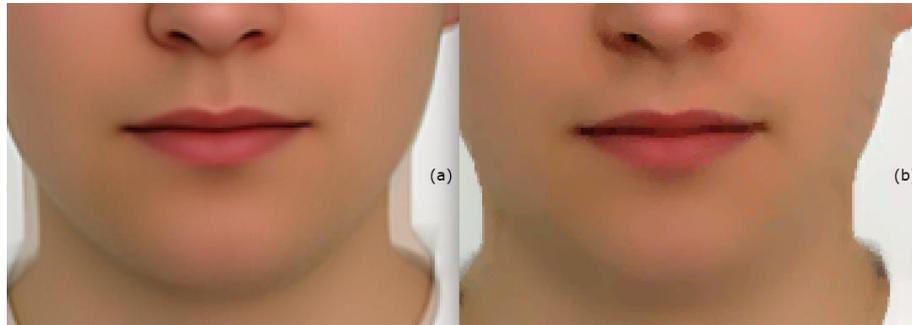
Figura 39: Faces Medianas; (a) Face Mediana de mulher, (b) Face Mediana de Homem, (c) Resultado do FantaMorph, (d) Resultado do MorphThing, (e) Método Proposto



Fonte: Autor; Imagens originais da BeautyCheck¹

¹ Website da Beauty Check, durchschnittsgesichter (Average Faces). Disponível em: <<http://www.beautycheck.de/cmsms/index.php/durchschnittsgesichter>>. Acessado em 10/8/2016

Figura 40: Detalhe comparativo dos resultados do FantaMorph (a) e do ADAMM (b)

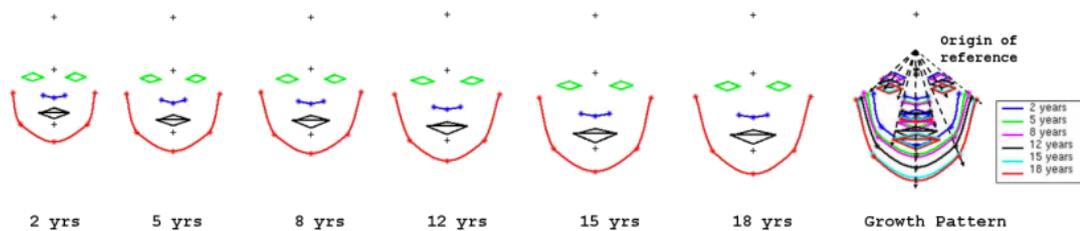


Fonte: Autor; Imagens originais da BeautyCheck¹

5.4.1. Experimento 2 – Mãe e Filha

Este foi o primeiro experimento realizado envolvendo uma mãe e sua filha com traços fisionômicos similares. Foi também a primeira vez em que a capacidade de alterar as proporções faciais interativamente foi colocada em prática para simular digitalmente o processo de envelhecimento do rosto. Utilizaram-se as proporções do modelo de crescimento crânio-facial definido por [RAM2006], cujo gráfico pode ser visto na Figura 41.

Figura 41: Modelo de Crescimento Crânio-Facial



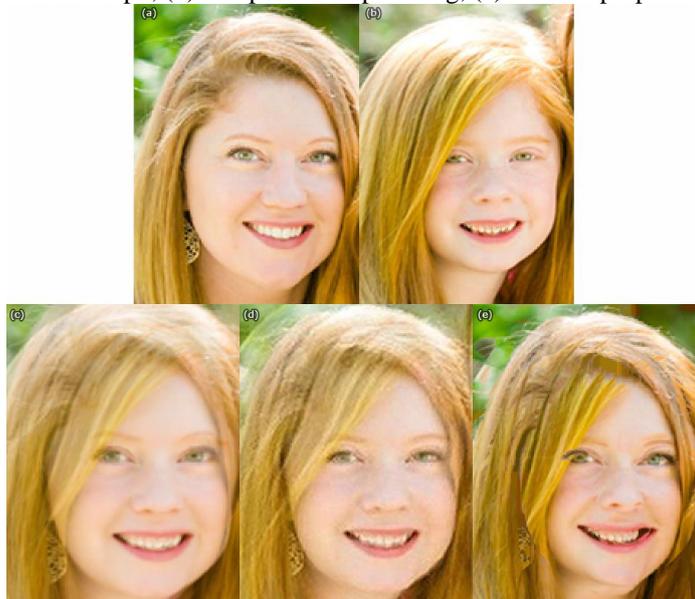
Fonte: [RAM2006]

Os resultados das três aplicações do experimento estão na Figura 42. O resultado do *SharpMorpher* (e) apresentou nitidez algumas vezes superior à das imagens geradas pelo *FantaMorph* (c) e o *Morph Thing* (d), como pode ser visto na Tabela 1. Além disto, apresentou a proporção facial esperada, dentro do estudo da antropometria, para a idade de

¹ Website da Beauty Check, durchschnittsgesichter (Average Faces). Disponível em: <<http://www.beautycheck.de/cmsms/index.php/durchschnittsgesichter>>. Acessado em 10/8/2016

15 anos que foi simulada. A resolução das imagens neste teste foi de 180x256 pixels.

Figura 42: *Morphing de Mãe e Filha*; (a) Foto da mãe; (b) Foto da filha; (c) Morph do FantaMorph; (d) Morph do Morph Thing; (e) Método proposto



Fonte: Autor; Foto original da St.Louis Post-Dispatch¹

5.4.2. Experimento 3 – Hilary Duff

Para o morphing entre as faces da atriz e cantora norte-americana Hilary Duff e seu filho, Luca, foram mais uma vez aplicados alguns ajustes de proporção facial durante o uso do *ADAMM*. Após a detecção automática das regiões, estes ajustes foram feitos exclusivamente nos macro-controles da máscara facial correspondentes a nariz, boca, queixo e laterais dos rostos, objetivando chegar a proporção padrão crânio-facial correta para a idade simulada. Ressalta-se que este não é o enfoque do presente trabalho, tendo sido as proporções faciais de [Ramanathan2006] aplicadas apenas como forma de validar esta possibilidade de uso.

Os resultados podem ser vistos na Figura 43. Para fins de comparação, fotos mais recentes do filho da atriz foram colocadas lado a lado à simulação digital gerada pelo método proposto, na Figura 44. A resolução das imagens utilizadas neste experimento foi de 267x376 pixels.

¹ St.Louis Mother-Daughter lookalike contest, disponível em: http://www.stltoday.com/lifestyles/relationships-and-special-occasions/parenting/aisha-sultan/erica-losito-and-ellie-weber-of-shrewsbury/article_5da630da-bb01-11e2-81f9-001a4bcf6878.html. Acessado em 11/8/2016

Figura 43: Morphing de Hilary Duff (a) e seu filho (b); (c) Morph do FantaMorph; (d) Morph do Morph Thing; (e) Método Proposto



Fonte: Autor; Imagens da base fornecida

Figura 44: Comparativo “Ground Truth” da imagem gerada pelo método proposto (a) e duas fotos mais recentes do Luca (b) e (c), anos depois da foto da Figura 43



Fonte: Autor; Imagens da base fornecida

5.4.3. Experimento 4 – Reese Witherspoon

A atriz e produtora estadunidense Reese Witherspoon possui similaridade fisionômica arbitrariamente alta com seus três filhos. Por este motivo, ela foi a escolhida para o último experimento deste trabalho, junto a seu filho mais novo, Tennessee James. Neste experimento uma resolução de imagens mais baixa que a dos experimentos anteriores

foi utilizada, 134x118 pixels, para avaliar um possível impacto deste parâmetro no resultado de nitidez. A magnitude do ganho de nitidez do ADAMM não foi muito diferente da obtida no Experimento 2, ainda assim foi o maior percentual de ganho de nitidez dentre todos os experimentos: +625,71% em relação ao *FantaMorph* e +214,6% em relação ao *Morph Thing*. Os resultados estão na Figura 45.

Figura 45: *Morph de Reese Witherspoon*; (a) foto da atriz; (b) foto do filho; (c) Morph do FantaMorph; (d) Morph do Morph Thing; (e) Método Proposto



Fonte: Autor; Imagens da base fornecida

5.5. Tabela de Resultados

Na tabela a seguir, seguem os resultados dos experimentos realizados. Foi mensurada a nitidez da técnica proposta e a dos aplicativos concorrentes selecionados, que utilizam a técnica de distorção-fusão tradicional [Wolberg90].

Foram empregadas as mesmas imagens em todas as aplicações, utilizando seus modos automáticos. Tanto o MorphThing quanto o FantaMorph apresentaram falhas nas extremidades de algumas imagens, o que não se verificou com o SharpMorpher. Para não distorcer os resultados, uma vez geradas as imagens, as mesmas foram centralizadas e tiveram as áreas externas à área da cabeça removidas, no programa de edição de imagens

*Adobe Photoshop*¹. Este processo também foi necessário para remover a marca d'água aplicada pelo *Morph Thing* no canto inferior direito das suas imagens.

Nas colunas da direita da Tabela 1, é mostrada a variação percentual coeficiente de nitidez, dos resultados do *SharpMorpher* em relação aos dos outros dois aplicativos comparados. A Variação Percentual é a razão entre o aumento de um valor e o seu valor original. Matematicamente, a Variação Percentual (VP) é definida pela Equação 20.

Equação 20

$$VP = \frac{V_1 - V_0}{V_0}$$

Na última linha da Tabela 1 é exibida a média das colunas para chegar a um valor médio de variação percentual.

Tabela 1: Resultados dos Experimentos de Avaliação de Nitidez

Imagem	(a) Nitidez / FantaMorph	(b) Nitidez / MorphThing	(c) Nitidez / SharpMorpher	Variação % de (c) em relação a (a)	Variação % de (c) em relação a (b)
Referência	76,0529	154,1156	256,0197	+236,63%	+66,12%
Imagem 1 (m&d)	142,3706	351,1793	969,382	+580,88%	+176,03%
Imagem 2 (Hilary)	282,4852	815,0072	1143,042	+304,63%	+40,24%
Imagem 3 (Reese)	481,1482	1.109,893	3.491,786	+625,71%	+214,60%
Média	245,5142	607,5488	1.404,0467	+471,88%	+131,10%

Fonte: Autor

¹ Programa de edição de imagens Adobe Photoshop. Disponível em: <<http://www.adobe.com/br/products/photoshop.html>>. Acessado em 20/7/2016

6. Conclusão e Trabalhos Futuros

A técnica de distorção-e-fusão (*Warp-dissolve*) tem sido utilizada com bons resultados e pouco ou nada nesta área tem sido desenvolvido desde o último trabalho de Jean Serra e Marcin Iwanowski a respeito, no começo dos anos 2000. O tema tem sido considerado "resolvido" desde então, tanto pelo mercado e a comunidade científica, porém esta afirmação só é verdadeira para certos tipos de aplicação.

Se considerarmos apenas as necessidades de gerar transições que serão exibidas em alta velocidade em algum filme ou programa de TV, os métodos tradicionais são bastante maduros e suficientes. Esta foi a característica da demanda original por avanços nas técnicas e usabilidade relacionadas, e é natural que uma vez atendidos esses anseios a pesquisa e o desenvolvimento na área tenham cessado.

Todavia, uma vez que busquemos apenas o quadro intermediário do morphing, a imagem mediana exata entre os dois extremos, o *warping* linear seguido pela fusão das imagens gera perda de nitidez e “fantasmas”, resultados de partes semi-transparentes das imagens, que deixam muito a desejar. O aumento da resolução e densidade visual das telas que utilizamos no dia-a-dia só serve para evidenciar estes problemas, modificando a nossa percepção do que antes eram resultados interessantes e aceitáveis para a de imagens borradas e artificiais.

Novas aplicações como as citadas na introdução fomentam soluções mais específicas e ágeis, por isso grande parte deste trabalho foi orientado exatamente para a automação e aceleração do processo de geração das imagens de morphing.

O principal artefato produzido por este trabalho, o software SharpMorpher, quando comparado a dois softwares modernos dedicados a morphing que aplicam a técnica de distorção-fusão tradicional [Wolberg90], apresentou ganhos percentuais de nitidez [Batten2001] que variaram de 66,12% a 625,71%. Não havia sentido em comparar os resultados da técnica apresentada com os de outras estudadas, como a de [Iwano2002], pelo simples fato destas técnicas não apresentarem qualquer tipo de alinhamento automático de regiões.

Dentre as contribuições mais significativas aqui apresentadas, além do processo adaptativo ADAMM, que evita distorções ou perda de nitidez, está a concepção da máscara

facial poligonal desenvolvida para o processo. Esta máscara poligonal obedece a todos os critérios de qualidade de geometria topológica, tendo sido ajustada em seus detalhes para funcionar com as mais diversas proporções faciais.

Além da detecção e do alinhamento automático dos recursos faciais, a ferramenta de software desenvolvida propicia um modo interativo de ajuste das faces. Desta forma o operador pode corrigir eventuais distorções, simular diferentes proporções faciais correspondentes a cada faixa etária e até mesmo gerar proporções caricatas se assim desejar.

Para as aplicações de arte forense, uma das maiores motivações deste trabalho, pôde ser oferecido um significativo ganho de produtividade e nitidez nas imagens geradas nos processos automático e semi-automático, em comparação com o método clássico de *morphing*. Evidentemente, assim como já ocorre atualmente, uma combinação de técnicas deverá continuar a ser utilizada na produção destas imagens. Destaca-se o objetivo alcançado de oferecer mais um aliado ao leque de recursos à disposição dos profissionais de arte forense e envelhecimento digital.

Como possível trabalho futuro de pesquisa, derivado deste trabalho, inclui-se a busca por melhorias no algoritmo de mediana morfológica, especialmente no sentido de selecionar e alinhar regiões com cores e topologia de imagem similares entre as duas imagens. Tal trabalho poderia propiciar a redução dos artefatos inerentes ao algoritmo atual, tornando-o mais “inteligente”. Outra possibilidade é a de projetar as imagens dos rostos em geometria facial com volume 3D real, ao invés da malha poligonal 2D utilizada. Uma abordagem volumétrica 3D propiciaria resultados mais realistas para o *morphing* entre cabeças com diferentes rotações, como é comum acontecer com fotos em condições não controladas.

7. Referências

[BALTRUSAITIS2016] TADAS BALTRUSAITIS, PETER ROBINSON, LOUIS-PHILIPPE MORENCY, “OPENFACE: AN OPEN SOURCE FACIAL BEHAVIOR ANALYSIS TOOLKIT”, 2016

[BAY2008] BAY, H. E TUYTELAARS, T. AND VAN GOOL, L. “SURF: SPEEDED UP ROBUST FEATURES”, 9TH EUROPEAN CONFERENCE ON COMPUTER VISION, 2008

[BAGGIO2012] DANIEL LÉLIS BAGGIO, DAVID MILLÁN ESCRIVA, NAUREEN MAHMOOD, ROY SHILKROT, SHERVIN EMAMI, KHVEDCHENIA LEVGEN, JASON SARAGIH, “MASTERING OPENCV WITH PRACTICAL COMPUTER VISION PROJECTS”, 2012

[BATTEN2001] C.F. BATTEN, D.M. HOLBURN, B.C. BRETON, N.H.M. CALDWELL, “SHARPNESS SEARCH ALGORITHMS FOR AUTOMATIC FOCUSING IN THE SCANNING ELECTRON MICROSCOPE”, DEPARTMENT OF ENGINEERING - CAMBRIDGE UNIVERSITY, 2001

[BEUCH98] SERGE BEUCHER "INTERPOLATION OF SETS, OF PARTITIONS AND OF FUNCTIONS" IN: H.HEIMANS AND J.ROEDINK, EDITORS, MATHEMATICAL MORPHOLOGY AND ITS APPLICATIONS TO IMAGE AND SIGNAL PROCESSING. KLUWER, 1998

[BUSS2004] SAMUEL R. BUSS, “3D COMPUTER GRAPHICS: A MATHEMATICAL INTRODUCTION WITH OPENGL”, CAMBRIDGE UNIVERSITY, 2004

[BOI2005] BOISSONNAT J.-D., OUDOT S., “PROVABLY GOOD SAMPLING AND MESHING OF SURFACES”, 2005

[DELOERA2010] JESÚS A. DE LOERA, JÖRG RAMBAU, FRANCISCO SANTOS, “TRIANGULATIONS, STRUCTURES FOR ALGORITHMS AND APPLICATIONS. ALGORITHMS AND COMPUTATION IN MATHEMATICS”, 2010

[GU2003] MIAO JIN, YALIN WANG, XIANFENG GU, SHING-TUNG YAU, “GLOBAL CONFORMAL SURFACE PARAMETERIZATION”, ACM SYMPOSIUM ON GEOMETRY PROCESSING, PP. 127–137, 2003

[IGARASHI2005] IGARASHI T., MOSCOVICH T., HUGHES, J. F., “AS- RIGID-AS-POSSIBLE SHAPE MANIPULATION”, ACM TRANS. GRAPH. 24, 3, 1134–1141, 2005

[IWANO99] IWANOWSKI M., SERRA J.: “MORPHOLOGICAL INTERPOLATION AND COLOR IMAGES”, PROCEEDINGS OF 10TH INTERNATIONAL CONFERENCE ON IMAGE ANALYSIS AND PROCESSING, SEPT. 1999

[IWANO2002] IWANOWSKI M., “IMAGE MORPHING BASED ON MORPHOLOGICAL INTERPOLATION COMBINED WITH LINEAR FILTERING”, JOURNAL OF WSCG., VOL. 10, 2002

[IWANO2011] M. IWANOWSKI, IMAGE METHAMORPHOSIS BASED ON UNIVERSAL MORPHOLOGICAL INTERPOLATOR PRZEGLAD ELEKTROTECHNICZNY, 87, PP.234-237, ISSN 033-2097, 2011

[LEE98] LEE S, WOLBERG G, SHIN SY, “POLYMORPH: MORPHING AMONG MULTIPLE IMAGES”, IEEE COMPUT GRAPH APPL 18:58 – 71, 1998

[LOOMIS56] ANDREW LOOMIS, “DRAWING THE HEAD AND HANDS”, 1956

[LOWE99] DAVID G. LOWE, "OBJECT RECOGNITION FROM LOCAL SCALE-INVARIANT FEATURES," INTERNATIONAL CONFERENCE ON COMPUTER VISION, CORFU, GREECE PP. 1150-1157, SEPTEMBER 1999

[MATT2004] IAIN MATTHEWS, SIMON BAKER, “ACTIVE APPEARANCE MODELS REVISITED”, CARNEGIE MELLON UNIVERSITY, 2004

[MEIJERING2002] MEIJERING, ERIK, “A CHRONOLOGY OF INTERPOLATION: FROM ANCIENT ASTRONOMY TO MODERN SIGNAL AND IMAGE PROCESSING”, PROCEEDINGS OF THE IEEE 90, 2002

[PECH2000] J.L. PECH-PACHECO, G. CRISTÓBAL, J. CHAMORRO-MARTÍNEZ, J. FERNÁNDEZ-VALDIVIA, “DIATOM AUTOFOCUSING IN BRIGHTFIELD MICROSCOPY: A COMPARATIVE STUDY”, 2000

[RAM2006] N. RAMANATHAN AND R. CHELLAPPA, “MODELING AGE PROGRESSION IN YOUNG FACES” IN PROCEEDINGS OF IEEE COMPUTER SOCIETY CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2006, PP. 387-394.

[SABIHAH2013] SABIHAH BINTI SHAMSUDDIN, SITI NORASYIKIN BINTI MOHD RAZALI, NUR ADILA BINTI SALEHUDDIN, “REVIEW OF IMAGE MORPHING TECHNIQUES”, SEM1-2012-2013

- [SCHAEFER2006] SCHAEFER S., MCPHAIL T. AND WARREN J., “IMAGE DEFORMATION USING MOVING LEAST SQUARES”, PÁGINAS 533-540, ACM SIGGRAPH 2006
- [SHEF2006] ALLA SHEFFER, EMIL PRAUN, KENNETH ROSE, MESH PARAMETERIZATION METHODS AND THEIR APPLICATIONS”, FOUNDATIONS AND TRENDS IN COMPUTER GRAPHICS AND VISION, VOL. 2, NO. 2, PP 105-171, 2006
- [SHLIZ2014] I. KEMELMACHER-SHLIZERMAN, S. SUWAJANAKORN, S. M. SEITZ, “ILLUMINATION-AWARE AGE PROGRESSION”, 2014
- [SINCLAIR2010] D. A. SINCLAIR, “S-HULL: A FAST RADIAL SWEEP-HULL ROUTINE FOR DELAUNAY TRIANGULATION”, JULY 2010
- [TAISTER2000] M.A. TAISTER, S.D. HOLLIDAY, AND H. BORMAN, “COMMENTS ON FACIAL AGING IN LAW ENFORCEMENT INVESTIGATION”, FORENSIC SCIENCE COMMUNICATIONS, 2000
- [TAN2006] SOVIRA TAN, JASON DALE, ANDREW ANDERSON, ALAN JOHNSTON, “INVERSE PERSPECTIVE MAPPING AND OPTIC FLOW: A CALIBRATION METHOD AND A QUANTITATIVE ANALYSIS”, IMAGE AND VISION COMPUTING, VOL. 24 ISSUE 2, PP. 153-165, 2006
- [TAYLOR2000] KAREN T. TAYLOR, “FORENSIC ART AND ILLUSTRATION”, 2000
- [TIZ2005] S. RAHNAMAYAN, H.R. TIZHOOSH, M.M.A. SALAMA, “RECOGNITION OF SUBJECTIVE OBJECTS BASED ON ONE GOLD SAMPLE”, PROCEEDINGS OF THE 5TH WSEAS INT. CONF. ON SIGNAL, SPEECH AND IMAGE PROCESSING, CORFU, GREECE, 17-19 AGO, PP309-314, 2005
- [TRIVEDI2009] ERIK MURPHY-CHUTORIAN, MOHAN MANUBHAI TRIVEDI, “HEAD POSE ESTIMATION IN COMPUTER VISION: A SURVEY”, 2009
- [VAUGHAN2011] WILLIAM VAUGHAN, “DIGITAL MODELING”, NEW RIDERS, 2011
- [WOLBERG98] G. WOLBERG, "IMAGE MORPHING: A SURVEY", THE VISUAL COMPUTER, PAGES 360-372, 1998
- [WU2011] JONATHAN WU, “FACE RECOGNITION JAMMER USING IMAGE MORPHING”, TECHNICAL REPORT, BOSTON UNIVERSITY, 2011 [HAKER2000] S. HAKER, S. ANGENENT, A.

TANNENBAUM, R. KIKINIS, G. SAPIRO, M. HALLE, “CONFORMAL SURFACE
PARAMETERIZATION FOR TEXTURE MAPPING”, IEEE TVCG, 6(2):181–189, 2000

[XINLI2008] XIN LI, BAHADIR GUNTURK, LEI ZHANG, “IMAGE DEMOSAICING: A
SYSTEMATIC SURVEY”, THE HONG KONG POLYTECHNIC UNIVERSITY, 2008