



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELECOMUNICAÇÕES
CAMPUS DE FORTALEZA

SAULO ANDERSON FREITAS DE OLIVEIRA

**UMA ABORDAGEM COM BASE EM ALGORITMOS GENÉTICOS PARA O
REDIMENSIONAMENTO DE IMAGENS SENSÍVEL AO CONTEÚDO**

Fortaleza - Ceará
Agosto 2016

Saulo Anderson Freitas de Oliveira

**UMA ABORDAGEM COM BASE EM ALGORITMOS
GENÉTICOS PARA O REDIMENSIONAMENTO DE
IMAGENS SENSÍVEL AO CONTEÚDO**

Dissertação submetida à Coordenação do Programa de Pós-Graduação em Engenharia de Telecomunicações do Instituto Federal do Ceará, como requisito parcial para obtenção do grau de Mestre em Engenharia de Telecomunicações.

Área de concentração: Sistemas de Telecomunicações.

Orientador: Prof. Dr. Ajalmar Rêgo da Rocha Neto.

Fortaleza
2016

Catálogo na fonte: Biblioteca Waldyr Diogo de Siqueira

- O48a Oliveira, Saulo Anderson Freitas de.
 Uma abordagem com base em algoritmos genéticos para o
 redimensionamento de imagens sensível ao conteúdo / Saulo
 Anderson Freitas de Oliveira. - Fortaleza: IFCE, 2016.
 102p.
- Dissertação (Engenharia de Telecomunicações).
 Orientador: Prof. Dr. Ajalmar Rêgo da Rocha Neto.
- 1.PROCESSAMENTO DIGITAL DE IMAGENS. 2.
 REDIMENSIONAMENTO SENSÍVEL AO CONTEÚDO 3.
 ALGORITMOS GENÉTICOS. I. Título.

CDD 621.367



INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO CEARÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE TELECOMUNICAÇÕES

Saulo Anderson Freitas de Oliveira

Esta dissertação foi julgada adequada para a obtenção do título de Mestre em Engenharia de Telecomunicações, sendo aprovada pela Coordenação do Programa de Pós-Graduação em Engenharia de Telecomunicações do Instituto Federal de Educação, Ciência e Tecnologia do Ceará e pela banca examinadora:

Ajalmar Rêgo da Rocha Neto

Orientador: Prof. Dr. Ajalmar Rêgo da Rocha Neto
Instituto Federal do Ceará (IFCE)

Guilherme de Alencar Barreto

Prof. Dr. Guilherme de Alencar Barreto
Universidade Federal do Ceará (UFC)

Francisco Nivando Bezerra

Prof. Dr. Francisco Nivando Bezerra
Instituto Federal do Ceará (IFCE)

Regis Cristiano Pinheiro Marques

Prof. Dr. Regis Cristiano Pinheiro Marques
Instituto Federal do Ceará (IFCE)

Fortaleza, 11 de Agosto de 2016

Este trabalho é dedicado aos meus pais, Antônio Chagas de Oliveira e Cleuma Janete Freitas Silva, pelo exemplo de vida, dedicação aos filhos e por sempre apoiarem minhas escolhas. Dedico também ao meu irmão, Artur Jefferson Freitas de Oliveira, minha querida tia Cleuba Maria Freitas Silva. Por fim, agradeço aos meus colegas do Laboratório de Processamento de Sinais (PDS), Alan Magalhães Braga, Danilo de Souza Avilar, Shara Shami Araújo Alves, Madson Luiz Dantas Dias, Elizângela de Souza Rebouças, David Clifte da Silva Vieira e Lucas Silva de Sousa, pelas discussões que enriqueceram este trabalho.

Agradecimentos

Primeiramente a Deus, pois sem sua permissão nada seria possível;

Aos meus pais por me motivarem a estudar desde cedo, em especial à minha mãe por sempre ser uma fonte eterna de incentivos;

Ao meu orientador, Prof. Dr. Ajalmar Rêgo da Rocha Neto, pela orientação, conhecimentos e entre muitas outras qualidades, pela perseverança em incentivar-me nos momentos cruciais para execução deste trabalho;

Ao Prof. Dr. Francisco Nivando Bezerra pelas suas contribuições;

À minha família que sempre me apoiou durante os estudos;

Aos meus amigos pesquisadores Alan Magalhães Braga, Danilo de Souza Avilar, Shara Shami Araújo Alves, Madson Luiz Dantas Dias, Elizângela de Souza Rebouças, David Clifte da Silva Vieira e Lucas Silva de Sousa por tornarem a convivência no laboratório PDS uma experiência única e enriquecedora #pdssquad;

À CAPES pelo apoio financeiro;

Aos demais professores do IFCE, a Dona Rose do Café, as secretárias da pós-graduação e a todos os funcionários do departamento pela amizade e colaboração durante o decorrer do mestrado;

A todos que de alguma forma contribuíram para a realização deste trabalho.

*“Cedo ou tarde, você vai aprender, assim como eu aprendi,
que existe uma diferença entre CONHECER o caminho e
TRILHAR o caminho.”*

Morpheus (filme Matrix)

Resumo

A diversidade de condições de consumo de imagens apresentam um novo desafio: as imagens devem ser apresentadas de forma ideal em diferentes resoluções e diferentes proporções de tela. Ajustar as imagens a tais mudanças, muitas vezes, não resulta em boas representações da imagem original. Para superar esta limitação, uma classe de técnicas chamada de redimensionamento sensível ao conteúdo almeja alcançar boas representações levando em conta não somente as restrições espaciais, como também a preservação de regiões importantes e a minimização de distorções visuais. O operador *Seam Carving* pertencente a esta classe de técnicas, funciona através da identificação e processamento de caminhos ótimos de pixels conexos que cruzam a imagem de uma extremidade à outra. Processar estes caminhos ótimos resulta em uma imagem com poucas mudanças perceptíveis ao observador. Neste sentido, os caminhos ótimos são obtidos através da minimização de uma função custo que analisa todos os possíveis caminhos de forma exaustiva. É nessa abordagem exaustiva que reside o maior desafio no tocante ao custo computacional do *Seam Carving*. Diversos trabalhos posteriores a sua publicação almejaram mitigar este caráter exaustivo. Tais trabalhos incorporaram soluções subótimas e com isto foi reduzido o número de iterações, porém a abordagem exaustiva foi mantida. Neste contexto, a proposta deste trabalho, denominada *Genetic Seam Carving*, visa uma abordagem sem caráter exaustivo na qual a busca por caminhos ótimos é realizada através de Algoritmos Genéticos, um método de pesquisa meta-heurístico. Inspirados na evolução natural, os Algoritmos Genéticos, são comumente empregados para gerar soluções para problemas de otimização nos quais o espaço de busca é explorado de forma não exaustiva. As simulações computacionais realizadas atestam um desempenho superior em termos de qualidade visual via índices qualitativos, bem como melhor custo computacional para imagens de alta resolução. Além disto, também foi proposto um índice de avaliação para os resultados de redimensionamento.

Palavras-chave: Processamento Digital de Imagens; Redimensionamento Sensível ao Conteúdo em Imagens; Algoritmos Genéticos; Seam Carving; Genetic Seam Carving.

Abstract

The diversity and conditions of image consumption in devices face a new challenge: the images should be displayed optimally at different displays with different resolutions. Adjust images to such changes, often does not result in solutions that are good representations of the original image itself. To overcome this limitation, a class of techniques called content-aware image retargeting aims to achieve good representations taking into account not only the geometric constraints, as well as the preservation of important regions and the minimization of visual distortions. Seam Carving operator, one belonging to this class of techniques, works by identifying and processing optimal paths of neighboring pixels that crosses the image from one end to another. Process these optimal paths results in an image with few noticeable changes to the observer. In that sense, the optimal paths are obtained by minimizing a cost function that analyzes all possible paths exhaustively. It is in this exhaustive approach that lies the greatest challenge in terms of computational cost in Seam Carving. After the Seam Carving publication, several works focused on mitigating this exhaustive component. By incorporating suboptimal solutions, such works reduced the number of iterations, but the exhaustive component was maintained. In this context, the purpose of this dissertation, called Genetic Seam Carving, aims an approach without this exhaustive component in which the search for optimal paths is performed using Genetic Algorithms, a meta-heuristic search method. Inspired by the natural evolution, Genetic Algorithms are commonly used to generate solutions for optimization problems in which the search space is explored not exhaustively. The carried out computer simulations attest superior performance in terms of visual quality evidenced by qualitative metrics, as well as better computational cost for high-resolution images. In addition, it was also proposed an evaluation metric for the retargeting results.

Keywords: Digital Image Processing; Content-aware Image Retargeting; Genetic Algorithms ; Seam Carving; Genetic Seam Carving;

Conteúdo

Lista de Figuras

Lista de Tabelas

Lista de Símbolos

Lista de Abreviações

1	Introdução	18
1.1	Motivação	19
1.2	Objetivo geral	21
1.3	Objetivos específicos	21
1.4	Produção científica	21
1.5	Organização	22
2	Redimensionamento de imagens sensível ao conteúdo	24
2.1	Introdução	24
2.2	Conceitos gerais	25
2.3	Trabalhos relacionados	26
2.4	Seam Carving	31
2.4.1	Busca pelo <i>seam</i> ótimo	32
2.4.2	Função de energia	33
2.4.3	<i>Multi-seam carving</i>	34
2.4.4	Outras aplicações	35

2.5	Avaliação de qualidade em imagens redimensionadas	38
2.5.1	Conceitos gerais	39
2.5.2	Descritores locais de imagens	41
2.5.2.1	<i>Scale invariant feature transform</i>	43
2.5.2.2	<i>Speeded up robust features</i>	44
2.5.3	Métricas de Avaliação	47
2.5.3.1	SIFT Flow	47
2.5.3.2	<i>Image retargeting SSIM</i>	48
2.6	Conclusão	50
3	Algoritmos Genéticos	52
3.1	Introdução	52
3.2	Visão geral	53
3.3	Representação genética	55
3.4	Função de aptidão	56
3.5	População	57
3.6	Operadores genéticos	57
3.6.1	Operador de seleção	57
3.6.2	Operadores de reprodução	60
3.6.3	Cruzamento	60
3.6.4	Operador de mutação	63
3.7	Crerios de convergência	65
3.8	Conclusão	65
4	Genetic Seam Carving (Proposta)	67
4.1	Indivduos ou cromossomos	67
4.2	Operadores genéticos	69

4.3	Função de aptidão	70
4.4	<i>Multi-seam carving</i>	73
4.5	Algoritmo para <i>Genetic Seam Carving</i>	74
4.6	SURF-SSIM (proposta)	75
4.7	Conclusão	77
5	Simulações Computacionais	79
5.1	Metodologia	79
5.2	Resultados	81
5.3	Validação	84
5.3.1	Validação da abordagem genética	85
5.3.2	Análise de qualidade das imagens redimensionadas	88
6	Conclusões e Trabalhos Futuros	93
6.1	Contribuições deste Trabalho	94
6.2	Trabalhos Futuros	94
	Referências	96

Lista de Figuras

1	Redimensionamento de 50% da largura por recorte em (a) e proporcional em (b).	27
2	Redimensionamento por RSVP.	28
3	Redimensionamento por segmentação. Adaptado de Setlur <i>et al.</i> (2005).	29
4	Redimensionamento com base em continuidade.	30
5	Redimensionamento por remendos.	30
6	Esquema de um modelo com base em multi-operadores.	31
7	<i>Seam</i> vertical (a) e um <i>seam</i> horizontal (b) dispostos.	32
8	Mapas de saliência representando imagens no domínio da energia, isto é, em termos de $e(.,.)$. Quanto mais clara a região, mais relevante ela é. Adaptado de Zhang e Sclaroff (2013).	34
9	Imagem original em (a) e sua representante em termos de função de energia $e(.,.)$ em (b).	35
10	Exemplo de ampliação de conteúdo usando redimensionamento proporcional e <i>Seam Carving</i>	35
11	Exemplo de remoção de conteúdo em <i>Seam Carving</i>	36
12	Redimensionamento em vídeo usando <i>Seam Carving</i> , adaptado de Rubinstein, Shamir e Avidan (2008).	38
13	Correspondência entre <i>pixels</i> em imagens redimensionadas.	39
14	Representação de uma pirâmide de imagem com cinco níveis. Adaptado de Moliński (2016).	42
15	Ilustração do processo de criação de oitavas e Diferença de Gaussianas (adaptado a partir de Lei (2011)).	44
16	Ilustração do processo de criação da pilha de imagens em SURF.	45

17	Exemplo de imagem integral.	45
18	Fluxograma do IR-SSIM (adaptado de Fang <i>et al.</i> (2014)).	49
19	Fluxo genérico de GAs.	55
20	Exemplos de diferentes codificações de indivíduos em Algoritmos Genéticos. .	56
21	Método da roleta.	59
22	Cruzamento pelo operador uniponto.	61
23	Cruzamento pelo operador de dois pontos.	62
24	Cruzamento pelo operador multi-ponto.	62
25	Cruzamento pelo operador uniforme.	63
26	Operador de mutação por giro.	64
27	Operador de mutação por permutação.	64
28	Operador de mutação por reversão.	64
29	Transformação de um indivíduo em <i>seam</i> em que o pixel em destaque representa o pivô.	68
30	Operador de cruzamento em GSC.	69
31	Operador de mutação em GSC.	70
32	Cenário com objetos relevantes representados pelas áreas preenchidas em cinza.	71
33	Pontos indicativos de regiões relevantes em S-GSC usando SURF.	71
34	Cenário com um <i>seam</i> inválido.	72
35	Fluxograma geral do <i>Genetic Seam Carving</i>	73
36	Região de cobertura do SURF-SSIM.	75
37	Deteção de pontos de interesse (PoIs).	76
38	Região de interesse em SURF-SSIM.	76
39	Algumas correspondências entre os pontos de interesse (PoIs).	77
40	Influência da não correspondência entre pontos de interesse.	77
41	Resultados de redimensionamento no conjunto RetargetMe.	82

42	Resultados de redimensionamento no conjunto NRID.	83
43	Resultados de redimensionamento no conjunto IRCCyN/IVC. Os resultados foram ampliados para fins comparativos.	84
44	Valores de aptidão dos piores e melhores indivíduos, bem como o valor médio da população para cada geração da última iteração do S-GSC.	85
45	Número de indivíduos mais aptos e o de removidos em cada execução dos GAs.	86
46	<i>Seams</i> azuis representam os mais aptos e os em vermelho, os processados. . . .	87
47	Diagramas de caixa do desempenho dos métodos no conjunto NRID.	89
48	Diagramas de caixa do desempenho dos métodos no conjunto RetargetMe. . . .	90
49	Imagem usada no experimento de medição de tempo em diferentes escalas: (a) 1 ×, (b) 1,5 ×, (c) 2 × e (d) 2,5 × o tamanho original.	91
50	Gráfico de tempo de execução em função da escala da imagem. Adaptado de Oliveira, Rocha Neto e Bezerra (2016).	92

Lista de Tabelas

1	Terminologia para Algoritmos Genéticos.	54
2	Conjunto de parâmetros selecionados para cada métrica nos experimentos. . . .	80
3	Desempenho dos algoritmos em função dos escores obtidos a partir das métricas SIFT Flow, IR-SSIM e SURF-SSIM, Os símbolos μ e σ se referem aos valores melhor, médio, menor e o desvio padrão dos escores, Os valores sublinhados representam os melhores escores em termos de desempenho.	91

Lista de Símbolos

I	Imagem original.
I'	Imagem redimensionada
m	Número de linhas da imagem original.
n	Número de colunas da imagem original.
m'	Número de linhas da imagem redimensionada.
n'	Número de colunas da imagem redimensionada.
s^v	<i>Seam</i> vertical.
s^h	<i>Seam</i> horizontal.
$M(.,.)$	Função de custo de <i>seam</i> .
$e(.,.)$	Função de energia.
s^{v*}	<i>Seam</i> vertical ótimo.
$\arg \min(.)$	Função que retorna o argumento com o menor valor.
$E(s^v)$	Função de energia do <i>seam</i> .
$L_I(x,y,t)$	Imagem I simplificada na escala t .
I_Σ	Imagem integral de I .
σ	Escala referente ao espaço de escala atual.
$\mathcal{H}(\mathbf{p},\sigma)$	Matrix Hessiana do ponto $\mathbf{p}$ na escala σ .
$SSIM_I$	Mapa de similaridade em IR-SSIM.
S_I^{TD}	Mapa de saliência <i>top-down</i> em IR-SSIM.
S_I^{BU}	Mapa de saliência <i>bottom-up</i> em IR-SSIM.
S_I	Mapa de saliência global em IR-SSIM.

Lista de Abreviações

PDI	Processamento Digital de Imagens.
IC	Inteligência Computacional.
GA	Algoritmos Genéticos.
SC	<i>Seam Carving</i> .
IQA	Avaliação de Qualidade de Imagem.
PSNR	<i>Peak to Signal Noise Ratio</i> .
SSIM	<i>Structural SIMilarity</i> .
VIF	<i>Visual Information Fidelity</i> .
EMD	<i>Earth Mover's Distance</i> .
SF	<i>SIFT-Flow</i> .
IR-SSIM	<i>Image Retargeting SSIM</i> .
DoG	Diferença de Gaussianas.
SURF	<i>Speeded Up Robust Features</i> .
SIFT	<i>Scale Invariant Feature Transform</i> .
CR	Operador de Recorte.
MOP	Operador de Multi-operador.
SCL	Operador de Redimensionamento Proporcional.
SM	Operador <i>Shift-maps</i> .
SNS	Operador <i>Scale-and-Stretch</i> .
SV	Operador <i>Streaming Video</i> .
WARP	Operador <i>Nonhomogeneous warping</i> .

INTRODUÇÃO

Inteligência Computacional (IC) é uma área do conhecimento que busca o desenvolvimento de técnicas inteligentes que imitam aspectos dos seres vivos, tais como: aprendizado (aprender com os erros e acertos de forma a, no futuro, agir de maneira mais eficaz), percepção (reconhecer padrões), raciocínio (aplicar regras lógicas a um conjunto de dados para chegar a uma conclusão), evolução (buscar uma melhor solução para um determinado problema através do aprimoramento de soluções) e adaptação (habilidade de aprender a partir de seu ambiente e com isso melhorar seu desempenho). Geralmente estas técnicas inteligentes são inspiradas em comportamentos encontrados na Natureza. Em uma Rede Neural, por exemplo, tem-se como inspiração o neurônio biológico, ao passo que em Algoritmos Genéticos, a evolução biológica é o cerne da abstração, enquanto que na Lógica Fuzzy a inspiração se dá pela ambiguidade e incertezas inerentes ao processamento linguístico.

Atualmente, verifica-se uma disseminação de técnicas de IC combinada ou integrada à área de Processamento Digital de Imagens (PDI) para resolução problemas. Exemplos de aplicações, tais como controle de robôs industriais ou veículos autônomos, detecção de eventos (monitoramento), aplicações médicas, modelagem de objetos ou ambientes, interação homem-máquina, meteorologia, agricultura, percepção e análise de cenas, representação de imagens, dentre outras. O potencial de aplicações existentes é enorme e os exemplos supracitados evidenciam perfeitamente as vantagens de se combinar IC e PDI no dia a dia.

No entanto, projetar uma abordagem capaz de tratar entradas de problemas de PDI para entradas de problemas de IC, nem sempre acontece de forma direta. Grande parte do desafio reside em encontrar uma transformação entre estes domínios de modo que as informações possam ser tratadas por diferentes modelos. Trabalhos como o de Tao, Tian e Liu (2003), Azazi *et al.* (2015) e Zhang *et al.* (2015) são exemplos de como o projeto em si, a transformação entre os domínios, é a maior contribuição.

Em Azazi *et al.* (2015), Paiva, Toledo e Pedrini (2016), Daniel e Anitha (2015) e Zhang *et al.* (2015) são apresentadas aplicações de técnicas de IC em diversas áreas de Visão Computacional, tais como reconhecimento de faces, segmentação de imagens, imageamento

por satélite, diagnóstico médico, dentre outras. Além disso, vale ressaltar que técnicas de IC, tais como Algoritmos Genéticos, Lógica *Fuzzy* e Redes Neurais Artificiais, têm contribuído bastante para a área de PDI, dando um novo fôlego onde antes havia limitações.

1.1 Motivação

Atualmente, uma gama de aplicações multimídia fazem uso de imagens, além de textos, para ilustração de conteúdo. Aplicações multimídia dinâmicas adaptam-se bem às mudanças a elas impostas, readequando bem a parte textual. O mesmo não pode ser dito quanto às imagens. Ajustar as imagens a tais mudanças, diversas vezes, significa cortá-las ou redimensioná-las de forma proporcional o que, em suma, resulta em uma imagem que muitas vezes não pode ser considerada uma boa representação do seu estado original. Quando escolhe-se o recorte, por exemplo, ocorre diretamente o descarte de informação. Já ao adotar o redimensionamento proporcional, é produzida uma imagem com aspecto “espremido” ou “esticado”. Em nenhuma destas abordagens tradicionais de redimensionamento leva em consideração aspectos pertinentes da imagem e acabam resultando em soluções que não agradam ao usuário.

Em contrapartida a essas abordagens, o redimensionamento sensível ao conteúdo (*Content-aware Image Retargeting*) (SHAMIR; SORKINE, 2009; VAQUERO *et al.*, 2010) é uma técnica em que se realiza o redimensionamento e, simultaneamente, preservam-se as regiões de interesse nas imagens. A ideia básica do processo é encontrar regiões de baixo interesse e, em seguida, expandi-las ou reduzi-las, a fim de produzir uma imagem de saída redimensionada com poucas mudanças perceptíveis ao observador.

O operador *Seam Carving* (AVIDAN; SHAMIR, 2007), pertencente à esta classe de técnicas, funciona através da identificação e processamento de caminhos ótimos de *pixels* conexos que cruzam a imagem de uma extremidade à outra, de modo que, ao se processar estes caminhos ótimos, uma imagem com poucas mudanças perceptíveis ao observador é obtida. Na formulação deste operador, os caminhos ótimos são obtidos através da minimização de uma função custo que analisa todos os possíveis caminhos de forma exaustiva. É nesta abordagem exaustiva que reside o maior desafio no tocante ao custo computacional do *Seam Carving*, pois entradas de alta dimensionalidade demandam mais processamento.

Alguns trabalhos posteriores à publicação do *Seam Carving* almejaram mitigar este caráter exaustivo pela incorporação de soluções subótimas, reduzindo assim apenas o número de iterações necessárias para se alcançar o redimensionamento. Esta estratégia foi posteriormente

denominada de *multi-seam carving*. No entanto, as formulações do operador com base *multi-seam carving* são bastante escassas na literatura. Os poucos trabalhos correlatos são descritos a seguir. Em Domingues, Alahi e Vanderghelynst (2010) foi proposto o uso de caminhos mais largos no intuito de se realizar uma estratégia *multi-seam carving*. Motivados pela localização dos caminhos ótimos, os autores tiraram proveito da vizinhança destes caminhos de modo que, ao estender a largura de um caminho ótimo, se obtém soluções subótimas e com isto, menos iterações do operador são realizadas. Conger, Kumar e Radha (2011) propuseram uma extensão da transformada *seamlets* (CONGER; RADHA; KUMAR, 2010) na qual, após a identificação de um caminho ótimo, os pixels da vizinhança deste caminho e os pertencentes a ele são filtrados por uma transformada *wavelet* discreta (Haar) e analisados para posterior remoção com base nos coeficientes obtidos.

Embora esses trabalhos foquem numa estratégia *multi-seam carving* e reduzam o número de iterações necessárias para se alcançar o resultado, o espaço de busca continua a ser explorado por completo a cada iteração, ou seja, o caráter exaustivo ainda é mantido. Explorar o espaço de busca de forma não exaustiva e ainda assim ser capaz de encontrar soluções úteis para problemas é deveras desejado, visto que em problemas como este em que o espaço de busca é demasiado grande, muitas vezes, é impraticável explorá-lo por uma abordagem exaustiva.

Neste contexto, Algoritmos Genéticos (HASHEMI *et al.*, 2010), um método de busca meta-heurístico, podem ser usados para gerar soluções úteis para problemas de otimização ou em problemas em que o espaço a ser explorado é demasiado grande e seria impraticável explorá-lo por uma abordagem exaustiva. Algoritmos Genéticos exploram o espaço de soluções de maneira direcionada utilizando a informação pertinente ao problema alinhada com algumas componentes aleatórias. Diferenciam-se de modelos estritamente aleatórios pois, as componentes aleatórias não são as principais condutoras da exploração.

Uma outra vertente pouco difundida na literatura atual, objeto de estudo e contribuição desta dissertação, consiste na utilização de Algoritmos Genéticos com foco na exploração de caminhos ótimos e subótimos através da minimização da função custo presente no *Seam Carving*. O método proposto em Oliveira, Bezerra e Rocha Neto (2015) (*Genetic Seam Carving*) e aperfeiçoado em Oliveira, Rocha Neto e Bezerra (2016) (*SURF-Based Genetic Seam Carving*) tem como ponto de partida as soluções obtidas por Algoritmos Genéticos, ao passo que as múltiplas soluções obtidas são analisadas para posteriormente serem usadas numa abordagem *multi-seam carving*. Ainda, segundo Oliveira, Rocha Neto e Bezerra (2016) este método supera o *Seam Carving* na sua formulação original em termos de qualidade de resultados e tempo de execução, especialmente quando se utiliza imagens de alta resolução, o que o torna atraente

quando se quer realizar redimensionamento sensível ao conteúdo.

1.2 Objetivo geral

O objetivo deste trabalho é a proposição de uma abordagem para o problema redimensionamento sensível ao conteúdo em imagens usando Algoritmos Genéticos.

1.3 Objetivos específicos

Em um sentido mais específico, esta dissertação tem os seguintes objetivos particulares:

- Realizar um estudo comparativo de métodos que visam o redimensionamento sensível ao conteúdo em imagens.
- Apresentar uma nova função de energia (função guia da pesquisa) com base em descritores de pontos salientes.
- Apresentar um novo método de redimensionamento sensível ao conteúdo com base em Algoritmos Genéticos.
- Avaliar os resultados obtidos a partir da proposta com outros métodos pertencentes ao estado da arte, tais como Recorte, Redimensionamento proporcional, *Seam Carving*, Multi-operador, *Shift-Maps*, *Scale-and-Stretch*, *Streaming Video* e o *Nonhomogeneous Warping*.
- Avaliar o desempenho do método proposto com base em medidas de similaridade e dissimilaridade.

1.4 Produção científica

Os seguintes artigos foram aceitos e publicados no transcorrer deste trabalho, a saber:

Artigos em conferências:

- **Saulo A. F. Oliveira**, Francisco N. Bezerra e Ajalmar R. Rocha Neto. “Genetic Seam Carving: a Genetic Algorithm Approach to Content-Aware Image Retargeting”. *7th Iberian Conference on Pattern Recognition and Image Analysis, IbPRIA*. 2015.

- **Saulo A. F. Oliveira** e Ajalmar R. Rocha Neto. “An Improved Genetic Algorithms-based Seam Carving Method”. *2nd Latin American Congress on Computational Intelligence, LA-CCI*. 2015.
- Shara S. A. Araújo, **Saulo A. F. Oliveira** e Ajalmar R. Rocha Neto. “A Novel Educational Timetabling Solution through Recursive Genetic Algorithms”. *2nd Latin American Congress on Computational Intelligence, LA-CCI*. 2015.
- **Saulo A. F. Oliveira**, Ajalmar R. Rocha Neto e João P. P. Gomes. “Towards Fixation Prediction: a Nonparametric Estimation-Based Approach Through Key-points”. *5th Brazilian Conference on Intelligent System, BRACIS (Aceito)*. 2016.

Cápítulo de livro:

- Shara S. A. Araújo, **Saulo A. F. Oliveira** e Ajalmar R. Rocha Neto. ‘A Recursive Genetic Algorithm-based approach for Educational Timetabling Problems’. *Designing with Computational Intelligence, Springer Book Series Studies in Computational Intelligence (Aceito)*. 2016.

Artigos em periódicos:

- **Saulo A. F. Oliveira** e Ajalmar R. Rocha Neto. “An Improved Genetic Algorithms-based Seam Carving Method”. *International Journal of Innovative Computing & Applications (Aceito)*. 2016.
- **Saulo A. F. Oliveira**, Ajalmar R. Rocha Neto e Francisco N. Bezerra. “A Novel Genetic Algorithms and SURF-Based Approach for Image Retargeting”. *Expert Systems with Applications*. 2016.

1.5 Organização

Este trabalho é formado por seis capítulos que descrevem os diversos conceitos relacionados a métodos de redimensionamento sensível ao conteúdo, avaliação qualitativa de resultados de redimensionamento e Algoritmos Genéticos. Mais especificamente, o Capítulo 2 traz uma descrição do problema de redimensionamento sensível ao conteúdo, salientando alguns dos mais famosos métodos e uma visão geral de como se avaliar qualitativamente resultados de redimensionamento e a descrição dos métodos de avaliação utilizados nos experimentos. Já no

Capítulo 3, os aspectos gerais de Algoritmos Genéticos, fundamental para o desenvolvimento desta dissertação, são apresentados. No Capítulo 4, é apresentado tanto uma proposta de método de redimensionamento com base em Algoritmos Genéticos, denominado *Genetic Seam Carving*, assim como um índice de qualidade para resultados de redimensionamento, denominado *SURF-SSIM*. O Capítulo 5 trata dos experimentos realizados para avaliação do método proposto. E, por fim, o Capítulo 6 traz as considerações finais, comentários e análises dos resultados obtidos nesta dissertação.

REDIMENSIONAMENTO DE IMAGENS SENSÍVEL AO CONTEÚDO

Neste capítulo são abordados aspectos relacionados ao redimensionamento sensível ao conteúdo; uma visão geral sobre outros métodos aplicados ao *retargeting* junto do método que inspirou este trabalho, denominado *Seam Carving*. Além disso, neste capítulo são abordados métodos relacionados à avaliação de qualidade em imagens redimensionadas. Para tal, três métodos: SIFT Flow, IR-SSIM e SURF-SSIM, foram escolhidos para posteriormente serem utilizados na avaliação dos resultados presentes neste trabalho. A escolha por estes métodos se deu por causa da sua disponibilidade e características. O primeiro, SIFT Flow, é um indicador de dissimilaridade, o IR-SSIM, um indicador de similaridade que se baseia em características mais salientes e detalhes, e, por fim, o SURF-SSIM um indicador de pontos salientes. Quanto às citações, além das apresentadas neste capítulo quando devidas, toma-se como base para escrita os artigos de Wang *et al.* (2004), Avidan e Shamir (2007), Shamir e Avidan (2009), Liu *et al.* (2011) e Liu, Yuen e Torralba (2011), Fang *et al.* (2014).

2.1 Introdução

Atualmente uma gama de aplicações multimídia faz uso de imagens para ilustração de conteúdo. Estas aplicações possuem seus próprios padrões definidos para a publicação e visualização dessas imagens, tais como: tamanho do arquivo, dimensões, formato, compressão, dentre outras. Normalmente uma mesma imagem é utilizada em mais de uma aplicação e nas mais variadas plataformas. Decorre disto que uma imagem pode não estar adequada à plataforma alvo devido às suas características intrínsecas, gerando visualizações inapropriadas.

Restrições impostas pela aplicação fazem com que a imagem seja apresentada de uma forma diferente em vista da geometria do novo espaço. Como somente o aspecto geométrico é levado em consideração, o conteúdo da imagem é muitas vezes negligenciado e sua apresentação se dá de uma forma inadequada e não tão agradável visualmente.

Já no caso de aplicações com leiaute dinâmico¹, o conteúdo textual geralmente consegue se adequar ao novo leiaute enquanto as imagens são muitas vezes prejudicadas por nem sempre conseguirem se ajustar a tais mudanças. Uma abordagem muito comum é o recorte ou aplicação de algum redimensionamento proporcional. Neste caso, o problema reside basicamente no fato de nenhuma dessas abordagens levar em consideração o conteúdo da imagem, resultando em soluções não tão eficientes.

Em contrapartida a estes métodos tradicionais, o redimensionamento sensível ao conteúdo (*content-aware image retargeting*), ou simplesmente *retargeting*, oferece uma abordagem mais sofisticada e eficaz para o problema. A ideia básica deste processo é encontrar regiões de baixo interesse e, em seguida, expandir ou reduzir ao redor das mesmas, a fim de gerar uma imagem com poucas mudanças perceptíveis ao observador.

2.2 Conceitos gerais

Redimensionamento de imagem é uma técnica em que se ajusta imagens de entrada em dimensões arbitrárias com linhas e colunas previamente definidas. O redimensionamento eficaz não deve apenas usar restrições geométricas, mas também considerar o conteúdo da imagem durante o ajuste.

Métodos tradicionais tais como recorte (*cropping*) e redimensionamento proporcional (*scaling*), lidam bem com a questão geométrica, mas falham na preservação do conteúdo. O redimensionamento proporcional introduz distorções quando o aspecto da imagem muda, podendo tornar objetos importantes irreconhecíveis devido à mudança de tamanho. Já o recorte, mesmo utilizando algum critério pré-definido tal como, o centro da imagem ou o canto superior esquerdo pode falhar ao não incluir áreas importantes no resultado.

Diferentemente destes métodos tradicionais, o redimensionamento sensível ao conteúdo baseia-se na preservação de regiões de alto interesse e no processamento efetivo nas regiões de baixo interesse, observando sempre o impacto causado pelas distorções ou artefatos visuais gerados durante o processo.

O problema do redimensionamento sensível ao conteúdo pode ser definido como segue. A partir de uma imagem I de tamanho $m \times n$, deve ser produzida uma nova imagem I' de tamanho $m' \times n'$ que representa a imagem original I . Não há uma definição clara de um indicador em termos qualitativos, de como inferir quão bom uma imagem redimensionada I' está para

¹Aplicações de leiaute dinâmico são aquelas em que o usuário pode redimensionar ou controlar o modo de como o leiaute se ajusta a mudanças.

a sua original I (SHAMIR; AVIDAN, 2009). Isso acontece porque este indicador qualitativo esta totalmente ligado a uma análise subjetiva, vai depender do consumidor. Entretanto, são definidos três objetivos principais para este tipo redimensionamento:

- o conteúdo importante em I deve ser preservado em I' ;
- as relações espaciais importantes em I também devem ser preservadas em I' ;
- I' deve ser livre de artefatos visuais.

A definição de importante é subjetiva, um conceito totalmente dependente da aplicação alvo. De fato, os operadores de redimensionamento sensível ao conteúdo tentam encontrar um equilíbrio entre a preservação de áreas importantes ao mesmo tempo que buscam satisfazer as restrições espaciais. Além do mais, a qualidade do mapa de importância² influencia diretamente no resultado final. Diversos trabalhos definem diferentes formas de indicar o que é ou não importante. Alguns quantificam aspectos visuais em níveis de detalhes, tais como borda, contraste, dentre outros; enquanto outros na preservação de grandes regiões ou objetos mais específicos, tais como pessoas, faces ou animais, dentre outras, para realizar este indicativo.

A seguir são apresentados alguns trabalhos relacionados ao redimensionamento sensível ao conteúdo de imagens.

2.3 Trabalhos relacionados

Diversos métodos para redimensionamento foram propostos, a fim de preservar partes do conteúdo relevante da imagem de entrada. Um trabalho interessante com esta finalidade pode ser encontrado em Vaquero *et al.* (2010). No referido trabalho foi realizado um levantamento de métodos para redimensionamento de forma categórica a fim de dar um panorama do estado da arte. Além da categorização, uma breve discussão sobre a avaliação dos resultados destes métodos é levantada.

As subseções que seguem apresentam categorias que condensam algumas abordagens em um sentido mais amplo de acordo com suas características.

Operadores com base em recorte

A ideia principal das abordagens que se baseiam em recorte como as apresentadas nos trabalhos de Suh *et al.* (2003), Zhang *et al.* (2005), Stentiford (2007), Amrutha *et al.* (2009),

²Um mapa de importância é um caso particular de um mapa de saliência, no qual o valor de cada *pixel* é ponderado pelo objetivo da aplicação.



Figura 1: Redimensionamento de 50% da largura por recorte em (a) e proporcional em (b).

Santella *et al.* (2006) e Store *et al.* (2007) é realizar o redimensionamento através de técnicas que busquem regiões importantes e que recortem uma janela com base no tamanho da imagem de saída. Um exemplo de aplicação de um processo de recorte é apresentado na Figura 1(b). A busca por estas regiões importantes geralmente é automática e é viabilizada por algum mapa de importância. Trabalhos como os de Santella *et al.* (2006) e Store *et al.* (2007) sugerem incluir informação vinda do usuário no processo para facilitar a estimativa das regiões importantes. Embora as técnicas de recorte não restrinjam o tamanho da imagem de saída, permitindo que o tamanho da janela de recorte seja variado, tais técnicas ainda podem ser classificadas como técnicas de redimensionamento sensível ao conteúdo. Basta que ao final da delimitação da janela de recorte seja aplicada um redimensionamento proporcional na janela para que a mesma fique com o tamanho desejado ou, alternativamente, restringir o espaço de busca daquela janela de recorte de modo a alcançar o tamanho desejado.

Operadores com base no redimensionamento proporcional

Uma desvantagem do operador de redimensionamento proporcional (*scaling*) como pode ser visto nos trabalhos de Samadani, Lim e Tretter (2007) e Muñoz, Blu e Unser (2001), é que os detalhes geralmente são perdidos durante o processo de redimensionamento, como mostrado na Figura 1(c). Para superar esta deficiência, são utilizadas técnicas que tentam minimizar essa perda de detalhes com o intuito de obter resultados satisfatórios. Um exemplo bem simples de minimização da perda de detalhes pelo operador de redimensionamento uniforme, é a extração de detalhes da imagem original e reinserção posterior na imagem redimensionada.

Operadores com base em representação visual serial rápida

Para efetivamente visualizar uma imagem em um tamanho completamente diferente,

especialmente em um tamanho menor, a representação visual serial rápida (*Rapid Serial Visual Representation*, RSVP) (BRUIJN; SPENCE, 2000; LIU *et al.*, 2003, 2007; HASAN; KIM, 2009) atua no contexto do tempo e não apenas no contexto espacial. A ideia principal é particionar a imagem em subimagens de alta resolução e exibí-las por um período de tempo. Um exemplo de redimensionamento por RSVP é mostrado na Figura 2, na qual a imagem 2(a) foi particionada em 3 outras subimagens 2(b), 2(c) e 2(d).



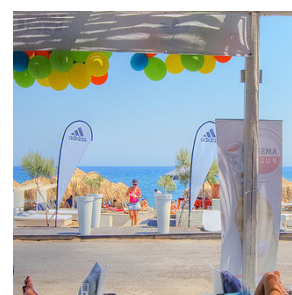
(a) Imagem original.



(b) Subimagem #1.



(c) Subimagem #2.



(d) Subimagem #3.

Figura 2: Redimensionamento por RSVP.

Embora essa abordagem permita que mais informações sejam visualizadas através do aumento do número de imagens exibidas ao longo do tempo, as técnicas que produzem uma única imagem, como se espera, permitem a visualização completa da informação em um único instante de tempo.

Operadores com base em segmentação

Em Setlur *et al.* (2005) é proposto um método baseado em segmentação. Inicialmente, são segmentadas as regiões importantes e em seguida tais espaços gerados são preenchidos com objetivo de incluir um fundo na região alterada. Em seguida, o fundo da imagem é redimensionado para o tamanho desejado e só então os objetos são colocados de volta com algum redimensionamento se necessário. É apresentado na Figura 3 um exemplo de redimensionamento com base em segmentação. O processo de segmentação de regiões importantes é mostrado na Figura 3(b); o processo de inclusão de fundo é mostrado na Figura 3(b); e, por fim, a inserção dos objetos de interesse é mostrada na como pode ser visto

na Figura 3(d).



(a) Imagem original.



(b) Segmentação de regiões salientes.



(c) Novo fundo.



(d) Resultado.

Figura 3: Redimensionamento por segmentação. Adaptado de Setlur *et al.* (2005).

Como este método baseia-se na segmentação precisa de objetos importantes, o mesmo pode gerar alguns artefatos visuais ou distorções. No entanto, ele tem a capacidade de reter as regiões importantes enquanto descarta fundos sem importância para cenas com duas ou mais regiões de interesse esparsas.

Operadores com base em continuidade

Uma outra classe de algoritmos de redimensionamento são os métodos que se baseiam em deformação (*Warping*), comumente chamados de métodos contínuos. Tais métodos realizam uma distorção não-linear para obter a imagem redimensionada, ver Figura 4.

A intensidade da distorção em áreas de alta importância é restringida, enquanto em regiões



Figura 4: Redimensionamento com base em continuidade.

de baixa importância são permitidas distorções mais grosseiras. Desta forma, regiões de alta e baixa importâncias são mantidas na imagem final, o que pode ser útil na preservação de objetos relevantes. Entretanto, dependendo da intensidade da distorção, áreas de baixa importância podem desaparecer, o que resulta na remoção de conteúdo.

Operadores com base em remendos

Os operadores baseados em remendos³ (*patches*) como os descritos nos trabalhos de Cho *et al.* (2008), Simakov *et al.* (2008), Barnes *et al.* (2009) e Pritch, Kav-Venaki e Peleg (2009) buscam através da manipulação de remendos na imagem de entrada, alcançar o redimensionamento. Tal manipulação está associada à minimização de uma função custo que capta as nuances de similaridade entre duas imagens durante o rearranjo dos remendos. Na Figura 5 é ilustrado um exemplo de redimensionamento por remendos. No contexto de redimensionamento, basta fazer uma analogia de encaixar subimagens em um espaço menor que resulte em uma boa representante da original.

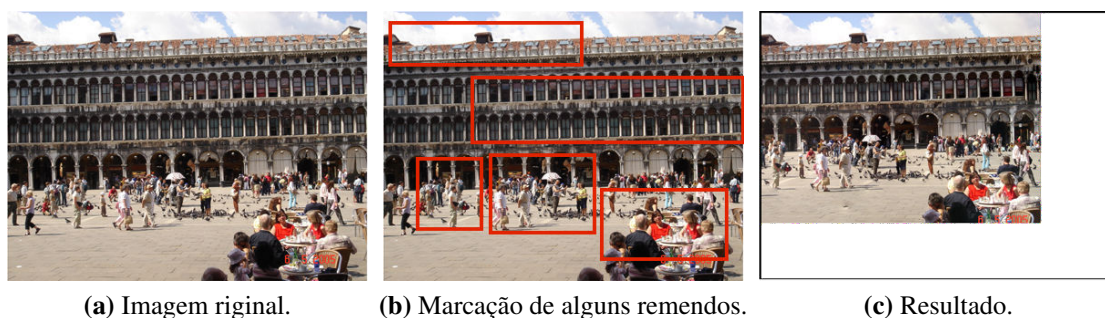


Figura 5: Redimensionamento por remendos.

Abordagem com base em multi-operadores

³Tradução livre do autor.

Por fim, temos os métodos com base em multi-operadores que são aqueles que combinam mais de um operador de redimensionamento. No contexto de multi-operadores, um espaço multi-dimensional de redimensionamento é criado. Espera-se, através da combinação dos operadores de forma ótima, alcançar resultados superiores quando comparados aos provenientes de um operador único. Na Figura 6 é ilustrado o esquema de um modelo que se baseia em multi-operadores no qual a combinação ótima de operadores é definida e guiada pela avaliação da qualidade do redimensionamento.

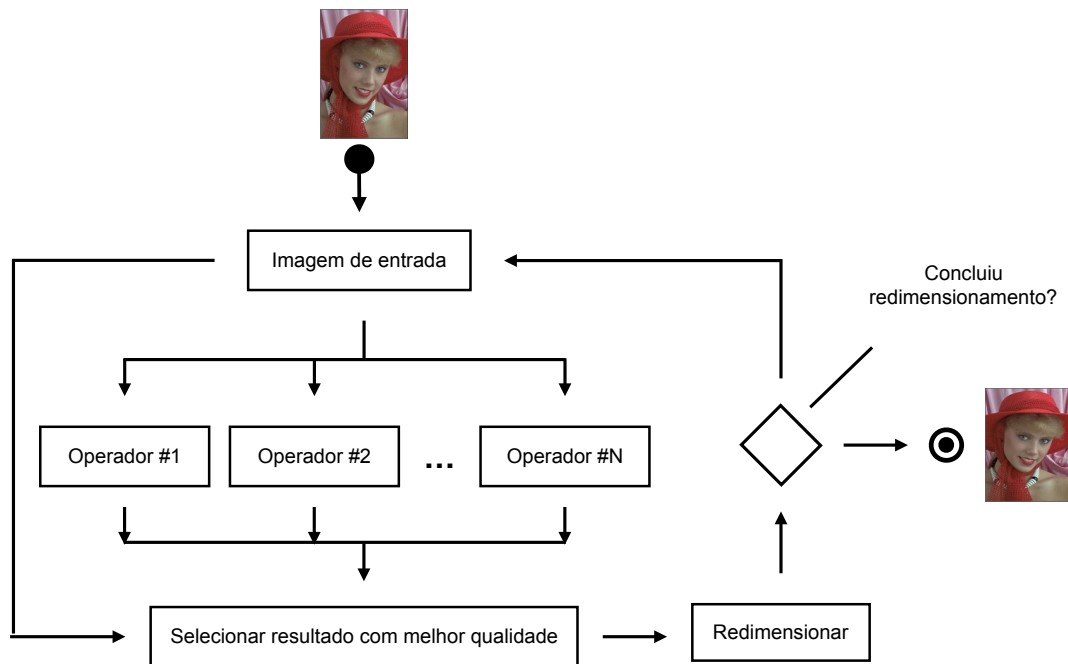


Figura 6: Esquema de um modelo com base em multi-operadores.

A fim de esclarecer o processo de redimensionamento de imagens, é apresentado em detalhes o operador *Seam Carving* que também é utilizado para fins de comparação com a nossa proposta.

2.4 Seam Carving

Ao buscar uma alternativa menos restritiva do que recortes ou remoção de colunas, mas que ainda preserve o conteúdo da imagem, Avidan e Shamir (2007) propuseram o operador *Seam Carving* (SC). O operador funciona com base na exclusão seletiva de *pixels* em forma de caminho que vão de uma extremidade da imagem à sua oposta, doravante denominado *seam*. O *seam* ótimo é definido por meio de uma função de energia que quantifica o grau de importância dos *pixels* da imagem. Em uma imagem I de tamanho $m \times n$, tem-se um *seam* vertical definido

por

$$\mathbf{s}^v = \{i, \alpha(i)\}_{i=1}^m \text{ tal que } \forall i, |\alpha(i) - \alpha(i-1)| \leq 1, \quad (2.1)$$

em que α é um mapeamento $\alpha : [1, \dots, m] \rightarrow [1, \dots, n]$. Similarmente, um *seam* horizontal é definido por

$$\mathbf{s}^h = \{(\beta(i), i)\}_{i=1}^n \text{ tal que } \forall i, |\beta(i) - \beta(i-1)| \leq 1, \quad (2.2)$$

em que β é um mapeamento $\beta : [1, \dots, n] \rightarrow [1, \dots, m]$. Como exemplo, na Figura 7 são apresentados dois *seams* dispostos em uma imagem. Na primeira imagem 7(a) tem-se um *seam* vertical, enquanto na segunda 7(b) tem-se um *seam* horizontal.

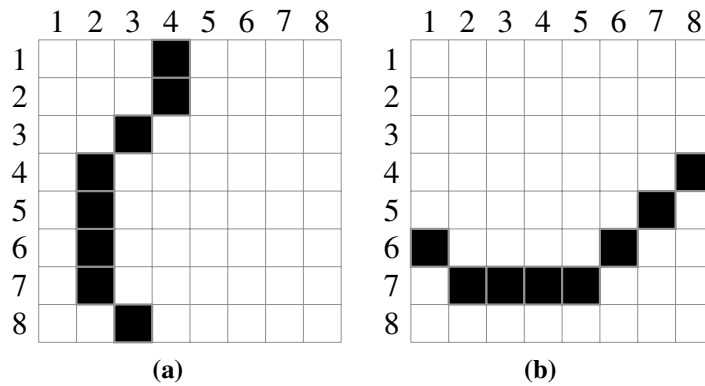


Figura 7: *Seam* vertical (a) e um *seam* horizontal (b) dispostos.

O processamento de *seams*, seja para fins de remoção ou de inserção, introduz somente efeitos locais na imagem já que todos os outros *pixels* são deslocados em um *pixel* para cobrir o espaço onde o seam foi processado de tal maneira que as diferenças são perceptíveis apenas ao redor daquela região. A fim de simplificar a notação para descrição do operador *Seam Carving*, considera-se somente o caso de *seams* verticais uma vez que a extensão para *seams* horizontais é similar. Além disto, enfatiza-se que o sistema de coordenadas tem sua origem no canto superior esquerdo nas descrições seguintes.

2.4.1 Busca pelo *seam* ótimo

A busca inicia-se da primeira linha da imagem ($x = 0$) e recursivamente continua pelos *pixels* vizinhos da próxima linha. O *pixel* vizinho escolhido é aquele que tem o menor custo adicional, sendo este custo definido pela função

$$M(x, y) = e(x, y) + \min(M(x+1, y-1), M(x+1, y), M(x+1, y+1)), \quad (2.3)$$

em que $e(.,.)$ representa a significância associado ao *pixel* localizado nas coordenadas (x,y) da imagem. A recursão se dá até o caso base que é a chegada no máximo valor de x e y , sendo estes valores máximos dados pela altura e largura da imagem, respectivamente. Para recuperar todos os *pixels* que fazem parte do *seam*, é realizado um *backtracking*⁴ na imagem a partir do caminho com o menor custo. Dito isto, o *seam* vertical ótimo $\mathbf{s}^{v*}$ é obtido então por

$$\mathbf{s}^{v*} = \arg \min_{\mathbf{s}^v} (\{M(\mathbf{s}_i^v)\}_{i=1}^n), \text{ tal que } M(\mathbf{s}_i^v) = M(1, i). \quad (2.4)$$

Para fins de otimização, Avidan e Shamir (2007) utilizam técnicas de programação dinâmica para resolver o problema da busca, evitando assim ter que calcular os valores de $M(.)$ mais de uma vez.

2.4.2 Função de energia

Muitos problemas de visão computacional envolvem a atribuição de um rótulo a cada *pixel*, sendo que os rótulos representam alguma categoria ou classe. Problemas de classificação de *pixels* são normalmente representados em termos de minimização de energia, em que a função de energia possui dois termos: um termo penaliza soluções que são inconsistentes com os dados observados, enquanto que o outro termo aplica algum tipo de coerência espacial (SZELISKI *et al.*, 2008).

A função de energia descreve uma medida de informação, sendo esta responsável por atribuir valores de significância a cada *pixel* na imagem. Neste contexto, as palavras energia e significância são praticamente termos intercambiáveis. A função de energia guia o operador SC para que o mesmo consiga encontrar *pixels* de baixa energia, formando assim *seams* também com essa característica de baixa energia. Dada uma função de energia $e(.,.)$ (Eq. (2.3)), o custo de um *seam* vertical $\mathbf{s}^v$ é definido como sendo a soma das energias dos *pixels* que o compõem, a saber:

$$E(\mathbf{s}^v) = \sum_{i=1}^m e(i, \alpha(i)). \quad (2.5)$$

Assim, a busca pelo *seam* vertical ótimo $\mathbf{s}^{v*}$ também pode ser definida como encontrar o *seam* com o menor custo pertencente à imagem $\mathbf{I}$, em que

$$\mathbf{s}^{v*} = \arg \min_{\mathbf{s}^v} (E(\mathbf{s}^v)), \forall \mathbf{s}^v \in \mathbf{I}. \quad (2.6)$$

O operador SC em si não define exatamente uma função de energia, contanto que esta

⁴*Backtracking* é uma estratégia na qual se retorna pelo mesmo caminho percorrido com a finalidade de encontrar a solução ótima.

função lide com a questão de dar maior energia para os objetos importantes e menor energia em caso contrário. Veja na Figura 8 diversos métodos encontrados na literatura com diferentes funções de energia: a saber, segmentação, entropia, métodos baseados em gradiente, mapas de saliência, decomposição via transformada *Wavelet*, descritores de imagem, dentre outros (AVIDAN; SHAMIR, 2007; ITO, 2012; KIM; KIM, 2011; ACHANTA; SUSSTRUNK, 2009; CONGER; KUMAR; RADHA, 2011). Quanto mais clara a região, mais energia ela tem e mais relevante ela é. Já na Figura 9 é ilustrada uma imagem de entrada e suas representantes em termos de função de energia.

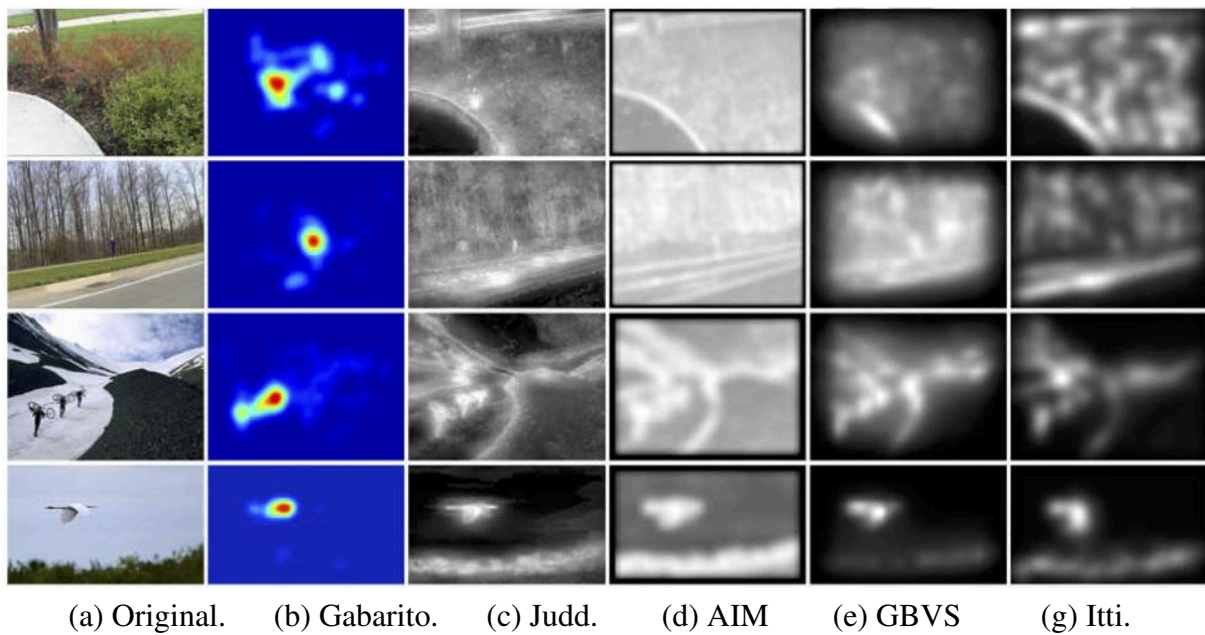


Figura 8: Mapas de saliência representando imagens no domínio da energia, isto é, em termos de $e(.,.)$. Quanto mais clara a região, mais relevante ela é. Adaptado de Zhang e Sclaroff (2013).

Devido à restrições de custo computacional, incorporar métodos complexos de cômputo de energia alinhados com fatores de qualidade se torna inviável (FANG *et al.*, 2014). Neste cenário, o projetista deve escolher entre a qualidade e o tempo de execução.

2.4.3 *Multi-seam carving*

Procurar pelo *seam* ótimo a cada iteração pode ser uma tarefa custosa, especialmente se o custo computacional para se computar a função de energia for elevado. Aproveitar a iteração da busca para encontrar outros *seams* ótimos ou até mesmo subótimos beneficia o algoritmo em termos de tempo de execução. Neste contexto, diz-se que uma abordagem é *multi-seam carving* se ela possuir esse comportamento de indicar mais *seams* durante uma iteração.

Trabalhos como os de Domingues, Alahi e Vanderghenst (2010) e Conger, Kumar e Radha

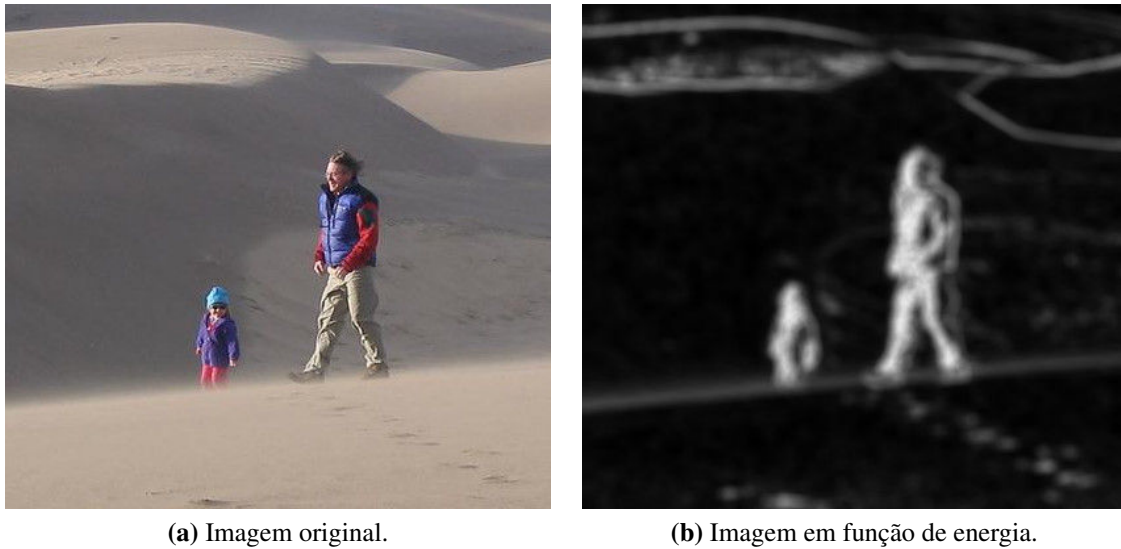


Figura 9: Imagem original em (a) e sua representante em termos de função de energia $e(\cdot, \cdot)$ em (b).

(2011) utilizam estratégias *multi-seam carving* motivadas pelas regiões habituais onde *seams* se encontram na imagem. Geralmente, nestas regiões pode-se processar um número maior de *seams*, alargando ou encolhendo estas regiões, sem provocar distorções visuais relevantes.

2.4.4 Outras aplicações

Além do redimensionamento, o operador SC pode ser utilizado em outras aplicações bastando combinar ou alterar a ordem de execução.

Ampliação de conteúdo



Figura 10: Exemplo de ampliação de conteúdo usando redimensionamento proporcional e *Seam Carving*.

Através da combinação de alguma técnica de ampliação de imagem, o SC pode ser utilizado para amplificar o conteúdo enquanto preserva suas dimensões, como mostrado na Figura 10.

Primeiramente, aplica-se alguma técnica de ampliação de imagem, como o redimensionamento uniforme e só então para preservar o conteúdo da imagem, tanto quanto possível, aplica-se o SC na imagem ampliada para reduzi-la de volta ao seu tamanho original.

Remoção de objetos

A técnica consiste em atribuir valores de baixa energia aos objetos a serem removidos. Tal abordagem implica na categorização de uma região de baixíssimo interesse, na qual *seams* de baixo custo existem. A priorização da exclusão destes *seams* acarreta na remoção dessas áreas e, conseqüentemente, nos objetos ali presentes.

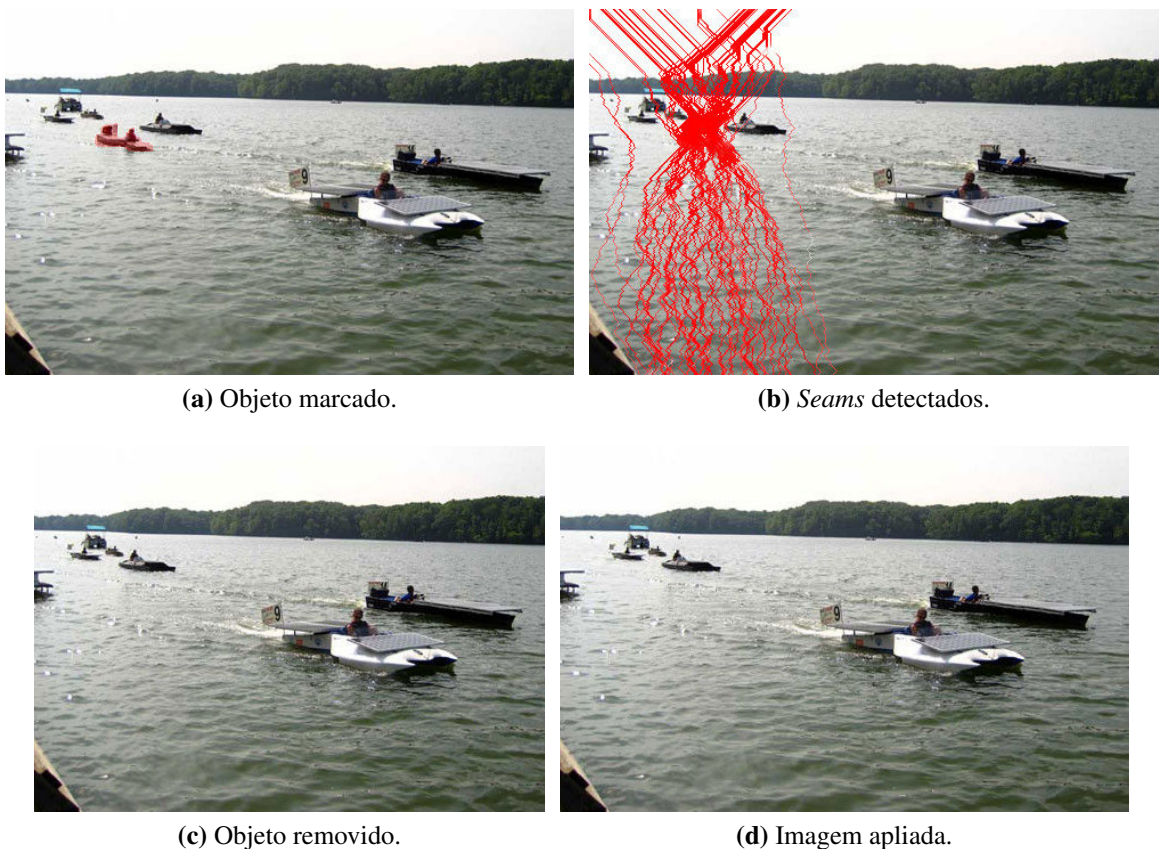


Figura 11: Exemplo de remoção de conteúdo em *Seam Carving*.

Imagens multi-tamanho

A ideia de imagens multi-tamanho está relacionada ao fato de uma mesma imagem possuir tamanhos distintos, podendo esta ser utilizada em diversas aplicações em que diferentes restrições geométricas são impostas. Neste caso, computa-se um determinado número de *seams* e a informação dos mesmos é armazenada junto à imagem. Feito isto, o redimensionamento acontece pela exclusão de *seams* previamente computados. Desta maneira, uma mesma imagem pode ser redimensionada em diversos tamanhos tendo um custo de tempo bem menor quando

comparado com o processo sob demanda (*on demand*).

Redimensionamento em vídeos

Proposto em Rubinstein, Shamir e Avidan (2008), uma nova formulação do SC pode ser aplicada em vídeos a fim de adaptar o conteúdo do mesmo em diversos dispositivos com resoluções distintas. Seu funcionamento é semelhante quando usado em imagens, com o adicional que uma técnica de antecipação de quadros para análise de *seams* na linha do tempo com o intuito de encontrar os *seams* ótimos. Na Figura 12 é ilustrado o comportamento do operador em vídeo.

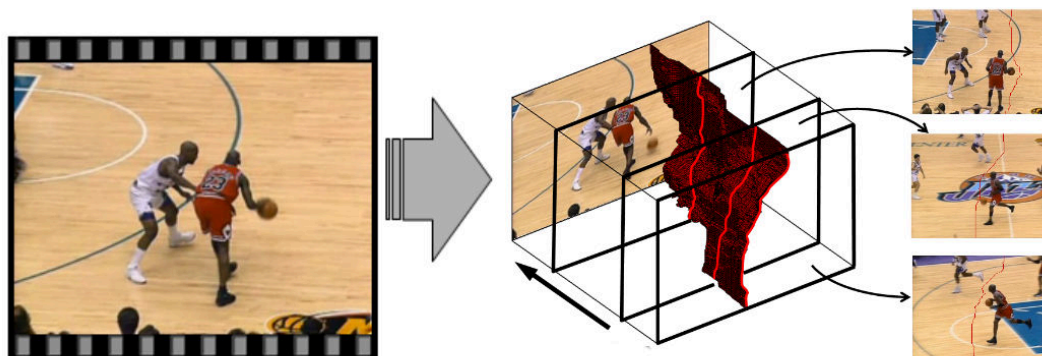


Figura 12: Redimensionamento em vídeo usando *Seam Carving*, adaptado de Rubinstein, Shamir e Avidan (2008).

2.5 Avaliação de qualidade em imagens redimensionadas

Quando se olha para o ambiente ao nosso redor, a atenção é direcionada principalmente para objetos visualmente importantes ou distintos. Tais objetos são referidos como visualmente importantes ou salientes. O sistema visual humano dedica a maior parte de seus recursos de processamento para analisar tais objetos.

O mesmo processo de análise destes objetos salientes deve ser empregado ao se avaliar resultados de redimensionamento. A preservação de tais objetos deve ser passível de medição em termos de qualidade. Como se sabe, a medida que o processo de redimensionamento ocorre, tem-se a remoção/inserção de alguns *pixels* e, dependendo da intensidade do redimensionamento, algumas regiões de interesse são inevitavelmente modificadas.

Métricas populares de avaliação de qualidade de imagem (*Image Quality Assessment*, IQA), tais como *Peak to Signal Noise Ratio* (PSNR), *Structural SIMilarity* (SSIM) de Wang *et al.* (2004) e *Visual Information Fidelity* (VIF) de Sheikh e Bovik (2006) não podem ser aplicados diretamente em resultados de redimensionamento porque requerem que o tamanho das imagens seja o mesmo. Como ressaltado anteriormente, projetar uma métrica capaz de comparar

conteúdos de imagem sob variações de proporção é significativamente mais desafiador, pois exige tanto uma análise semântica, quanto uma análise em termos de correspondência de conteúdo, ver Figura 13.

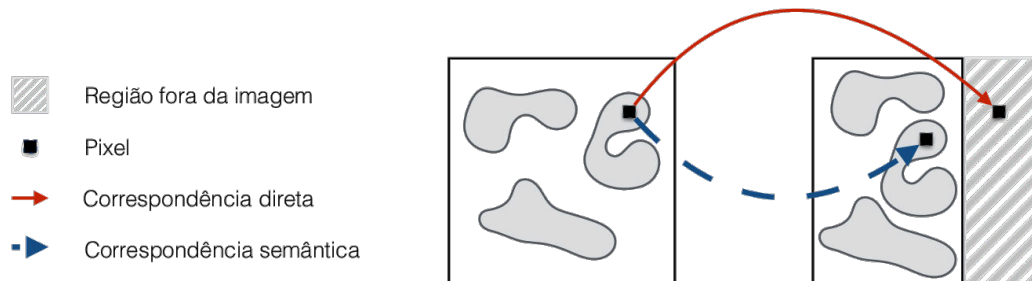


Figura 13: Correspondência entre *pixels* em imagens redimensionadas.

Várias métricas de avaliação de qualidade em imagens com base em distância (SIMAKOV *et al.*, 2008; RUBINSTEIN; SHAMIR; AVIDAN, 2009; LIU; YUEN; TORRALBA, 2011; PELE; WERMAN, 2009; KASUTANI; YAMADA, 2001; MESSING; BEEK; ERRICO, 2001; BARNES *et al.*, 2009) também têm sido amplamente empregadas na avaliação de qualidade em imagens redimensionadas, pois podem prever a percepção humana. Em tais métricas, o custo de transformação da imagem original na redimensionada é representado em termos de distância. Este custo geralmente está relacionado à distorção da imagem redimensionada, isto é, quanto menor esse custo mais elevada é a similaridade entre as imagens.

Por outro lado, um outro grupo de métricas com base em mapas de similaridade (FANG *et al.*, 2014; CASTILLO; JUDD; GUTIERREZ, 2011; CHAMARET *et al.*, 2012) também tem sido empregado de forma bem sucedida na avaliação dos resultados de redimensionamento. Estas métricas baseiam-se na criação de mapas que indicam em cada *pixel* como a informação é preservada após o redimensionamento. A qualidade é dada então pela aplicação de algum critério de similaridade ou medida de distância entre os mapas. Estas métricas, como esperado, não cobrem a totalidade da imagem, mas geralmente regiões importantes previamente mapeadas por algum método que, na grande maioria, detecta e descreve estruturas locais em imagens.

2.5.1 Conceitos gerais

Uma vez que os seres humanos são os consumidores finais das imagens processadas e, portanto, os juízes da qualidade da imagem, faz-se necessário que as métricas sejam próximas aos critérios de avaliação subjetivos.

Um estudo realizado em Rubinstein *et al.* (2010) mostrou que os seres humanos, em geral, concordam uns com os outros sobre a qualidade das imagens redimensionadas. Adicionalmente,

mostrou que algumas métricas são consistentemente mais favoráveis do que outras. As métricas *Earth Mover's Distance* (EMD) (PELE; WERMAN, 2009) e a SIFT-Flow (LIU; YUEN; TORRALBA, 2011) geralmente concordaram melhor com as preferências das pessoas de acordo com os critérios de avaliação do estudo. Isto foi observado através da forte correlação entre as preferências das pessoas, os resultados subjetivos, e os escores obtidos pelas métricas. Além disso, em um outro estudo realizado em Ma *et al.* (2012) foi observado que os seres humanos são muito sensíveis à distorção em rostos e estruturas geométricas, ao passo que eles podem tolerar mais distorções em cenários naturais, especialmente em regiões de textura.

Grande parte dos algoritmos de redimensionamento utilizam-se de técnicas de detecção de saliência para gerar mapas de importância. Nestes mapas, os valores de cada *pixel* variam no intervalo $[0, 1]$ em que 1 significa que o *pixel* é altamente relevante e valores próximos a 0 são atribuídos aos *pixels* sem importância. É através da preservação dos *pixels* de alta importância que se tenta alcançar resultados satisfatórios. Tais mapas são obtidos pela exploração, individual ou combinada, de dois tipos de características de mecanismos de atenção, a saber: a característica guiada por estímulo (*bottom-up*) e a guiada pela expectativa (*top-down*).

Características *bottom-up* geralmente utilizam informação de intensidade, orientação, contraste e cor como um meio para detectar estruturas salientes. São extremamente fáceis de implementar e requerem pouco processamento. Entretanto, em cenários do mundo real, utilizar apenas estas características *bottom-up* não são suficientes para a detecção de regiões importantes. Já as características *top-down* estão ligadas a tarefas mais específicas, tais como reconhecimento de faces, linhas, texto, animais, dentre outras. É sabido que os seres humanos prestam mais atenção em características *top-down* em virtude do valor semântico agregado a elas. Entretanto, a estabilidade e custo computacional associados à obtenção das mesmas é o que impacta o desempenho (FANG *et al.*, 2014).

As três métricas escolhidas para avaliação dos resultados deste trabalho, SIFT Flow (LIU; YUEN; TORRALBA, 2011), *Image Retargeting SSIM* (IR-SSIM) (FANG *et al.*, 2014) e SURF-SSIM (OLIVEIRA; ROCHA NETO; BEZERRA, 2016), utilizam-se de técnicas baseadas em características *bottom-up*, exceto IR-SSIM que usa um detector de faces caracterizando assim uma abordagem mista. Antes de detalhar o funcionamento das métricas, é feita uma breve introdução aos descritores locais, em especial aqueles descritores invariantes à escala, pois os mesmos compõe parte substancial das métricas adotadas neste trabalho.

2.5.2 Descritores locais de imagens

Descritores são vetores de atributos de imagens ou de determinadas áreas da imagem que representam informações sobre as mesmas. Descritores podem ser de duas naturezas, globais e locais, e são amplamente utilizados em aplicações de correspondência entre imagens. Descritores globais informam sobre a imagem de uma maneira global, usando todo o conhecimento dela. Já os locais, usam janelas, pequenas regiões e trazem informação acerca das estruturas contidas nas mesmas. Um tipo de descritor simples e global é a distribuição de intensidades de uma imagem representada pelo seu histograma.

Descritores, de um modo geral, permitem construir uma representação da imagem com informações relacionadas, tais como: orientação, escala, nível de detalhes, entropia e etc. O uso de descritores locais tem sido empregado em diversas aplicações que realizam a correspondência direta entre imagens, uma vez que se torna dispensável etapas de pré ou pós-processamento, tais como, utilizar estruturas de dados complexas ou até mesmo resolver problemas de otimização global (MAIA, 2010).

Como o foco desta seção está na avaliação de qualidade e, mais especificamente, nas métricas, é apresentada uma visão mais geral sobre a utilização destes descritores com foco na detecção de *pixels* invariável à escala, doravante denominados pontos-chaves e a correspondência entre imagens através destes pontos. Isso se faz necessário, uma vez que certos métodos de redimensionamento utilizam alguma variação do redimensionamento proporcional para acomodar o conteúdo. Essa adequação faz com que os objetos fiquem em escalas diferentes, isto é, em outra resolução. Identificar estes objetos com essa variação de escala se faz necessário visto que se busca mapear o quão preservadas estas estruturas estão após o processo de redimensionamento. Sendo assim, o uso de descritores invariáveis à escala torna-se atraente porque os mesmos são capazes de identificar estas estruturas independente da escala na qual elas se encontram.

A invariância à escala é uma propriedade que descreve o quão realçadas estão as características fortes independente da escala na qual elas se encontram em uma imagem. Para tal, uma representação via pirâmide de imagens, árvore quaternária (*Quadtree*) ou imagem integral (*Integral image*), é adotada e então pontos estáveis em escalas diferentes são encontrados através de um processo que consiste em buscar pontos que sejam invariantes à mudanças de escala. No detector SIFT (LOWE, 2004), por exemplo, cada camada dessa pirâmide é subtraída pelo nível conseguinte a fim de se obter imagens nas quais bordas e grandes objetos binários (*binary large objects, blobs*) podem ser facilmente encontrados, como ilustrado na Figura 14.

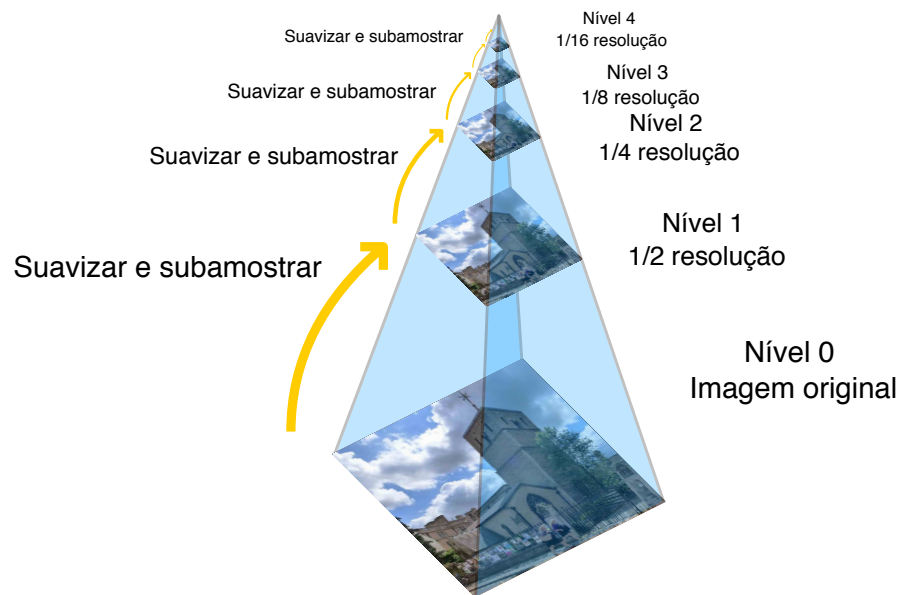


Figura 14: Representação de uma pirâmide de imagem com cinco níveis. Adaptado de Moliński (2016).

Com origem na teoria do espaço de escala (LINDBERG, 2013), a ideia é gerar imagens $L_I(x,y,t)$ derivadas da imagem original $I(x,y)$, tal que $L_I(x,y,0) = I(x,y)$ e t representa o nível de simplificação da imagem. Assim, as estruturas presentes na imagem em escalas mais grosseiras, em que o valor de t é grande, são simplificações das estruturas correspondentes em escalas mais finas com um valor de t pequeno (MAIA, 2010). Além disso, ao se adotar a abordagem de descritores locais, a tarefa de obter correspondências entre imagens é dividida em três etapas:

- **Selecionar pontos-chave:** etapa em que se detectam pontos que caracterizam a imagem, nos quais as informações de orientação e/ou escala de cada *pixel* também são obtidas;
- **Associar descritor local:** etapa na qual cada ponto-chave tem sua vizinhança representada por um vetor de características;
- **Gerar correspondência entre pontos-chaves:** etapa em que se encontra a correspondência entre duas imagens através da correspondência entre os descritores locais que as representam.

Na etapa de seleção de pontos-chave, que é um processo não supervisionado, toda a representação obtida depende basicamente de alguns parâmetros, tais como número de escalas observadas, limiar de correspondência entre pontos-chaves, dentre outros, que em suma carregam pouca informação sobre quais *pixels* são preferíveis para representar as regiões chaves. Na etapa de associação de descritor local, são usados métodos que geralmente observam

a variação de alguma propriedade ao longo de uma transformação. Ao final desta etapa, são associados aos pontos-chaves atributos com base no comportamento destes durante a transformação observada.

A correspondência entre pontos-chaves é dada pelo cômputo da similaridade dos descritores correspondentes. Como os descritores são vetores de atributos, estes podem ser comparados utilizando distância euclidiana, por exemplo. Normalmente, os candidatos a melhor correspondência são pontos próximos, de maneira que o melhor candidato é o ponto que apresenta a menor distância (GONZÁLES, 2010).

2.5.2.1 *Scale invariant feature transform*

O descritor *Scale Invariant Feature Transform* (SIFT) (LOWE, 2004) tem sido amplamente usado em diversas áreas: reconhecimento de objetos, rastreamento, correspondência de imagens panorâmicas, dentre outras. Neste descritor, o processo de detecção de pontos-chaves é dividido em quatro etapas: (i) detecção de extremos; (ii) localização de pontos-chaves; (iii) atribuição de orientação; e (iv) extração e atribuição de descritores aos pontos-chaves.

Na primeira etapa, uma Diferença de Gaussianas (DoG) (ver Figura 15) é criada para identificar os potenciais pontos-chave. A imagem de entrada é, então, repetidamente suavizada⁵ para produzir um conjunto de imagens, denominada oitava. A DoG é calculada como a diferença entre duas escalas de suavização adjacentes na pilha gaussiana (parte esquerda da Figura 15). Para criar a próxima oitava (Figura 15, à direita) cada imagem suavizada será subamostrada por um fator dois e o processo de cômputo das DoG é repetido. Os extremos nas imagens DoG, valores de máximos e mínimos, são detectados comparando o valor do *pixel* aos seus 26 vizinhos em regiões 3×3 nas escalas atual e adjacentes, ver as regiões em roxo na Figura 15. Três parâmetros principais controlam esta fase: o número de oitavas, a suavização inicial para o primeiro nível de cada oitava, definido pelo parâmetro σ e o número de escalas amostradas em cada oitava.

Em seguida, os pontos de interesse candidatos são localizados com precisão de *sub-pixel*⁶. Nesta etapa, pontos potencialmente instáveis⁷ em áreas de baixo contraste são eliminados quando seus valores, considerando os valores de extremos, são menores do que o valor do

⁵O processo de suavização consiste em borrar uma imagem através de um filtro gaussiano. É um efeito amplamente usado para reduzir o ruído da imagem e reduzir o detalhes.

⁶Dado que não há um mapeamento direto entre um mesmo *pixel* em oitavas distintas, a precisão dessa correspondência entre oitavas é aproximada e por isso se diz que os *pixels* são localizados com precisão de *sub-pixel*.

⁷A instabilidade neste contexto se refere ao fato de o ponto ser encontrado somente em algumas escalas, ao passo que os pontos estáveis são encontrados em todas as escalas observadas.

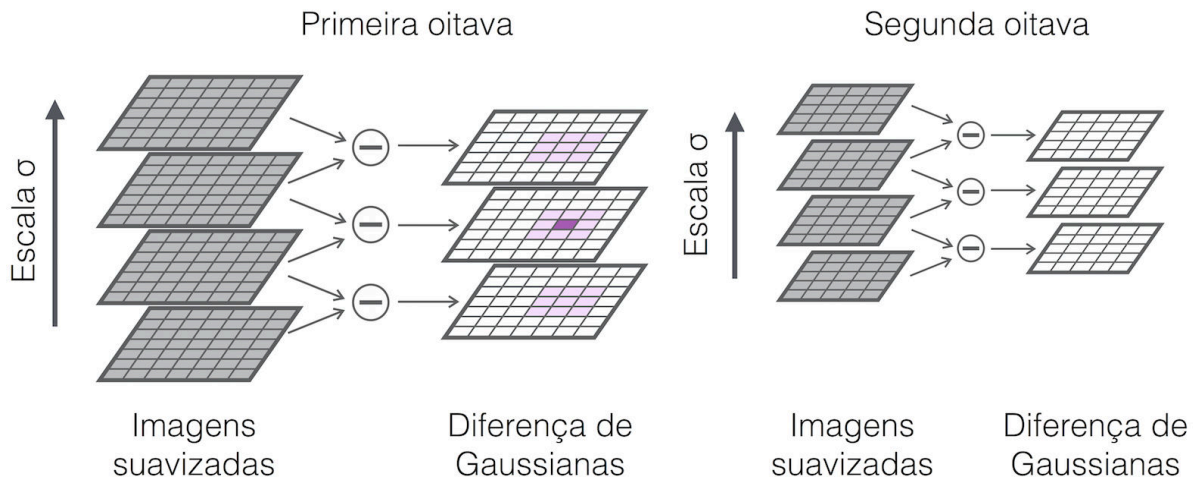


Figura 15: Ilustração do processo de criação de oitavas e Diferença de Gaussianas (adaptado a partir de Lei (2011)).

parâmetro de limiar de contraste. Do mesmo modo, os pontos de interesse instáveis localizados ao longo das bordas também são eliminados quando estes possuem valores menores do que o valor do parâmetro de limiar de borda.

Na terceira fase, uma orientação principal é atribuída a cada ponto chave com base nas direções do gradiente da imagem local. Finalmente, um conjunto de histogramas de 8 entradas baseados na vizinhança 4×4 para cada ponto-chave é obtido. Por fim, na quarta e última etapa, o descritor resultante é um vetor de 128 elementos que é normalizado para tratar as diferenças de iluminação (SIMA; BUCKLEY, 2013). Mais detalhes sobre parametrização e formulação do método podem ser encontrados em Lowe (2004).

2.5.2.2 *Speeded up robust features*

O descritor *Speeded Up Robust Features* (SURF) também tem sido amplamente utilizado em visão computacional e áreas afins como alternativa ao SIFT. SURF também é bem conhecido pelo seu desempenho, por vezes com mais de 10% de melhoria em termos de precisão, superando os outros descritores de maneira sistemática e significativa (BAY; TUYTELAARS; GOOL, 2006).

Diferentemente do SIFT, em que o espaço de escala é implementado através de pirâmides de imagens, em SURF considera-se o espaço de escala como uma pilha de imagens integrais de mesmo tamanho de modo que cada imagem desta pilha representa uma oitava. No entanto, cada imagem da pilha é o resultado da suavização da imagem integral por janelas de tamanhos

distintos. Na Figura 16 é mostrada uma ilustração da criação do espaço de escala em forma de pilhas em SURF.

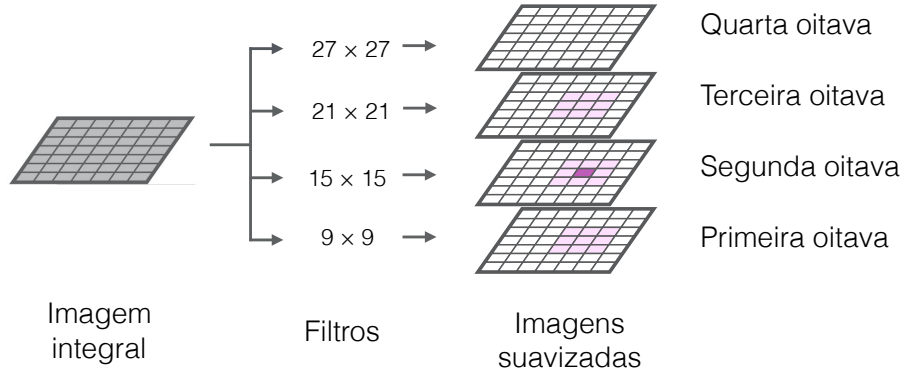


Figura 16: Ilustração do processo de criação da pilha de imagens em SURF.

Em uma imagem integral I_Σ , cada *pixel* é dado pela soma dos valores dos *pixels* à esquerda e acima, isto é,

$$I_\Sigma(x, y) = \sum_{\substack{x' \leq x \\ y' \leq y}} I(x', y'). \quad (2.7)$$

Na Figura 17 é apresentado um exemplo de produção de uma imagem integral I_Σ (Figura 17 (b)) a partir da imagem original I (Figura 17 (a)) de acordo a Eq. (2.7).

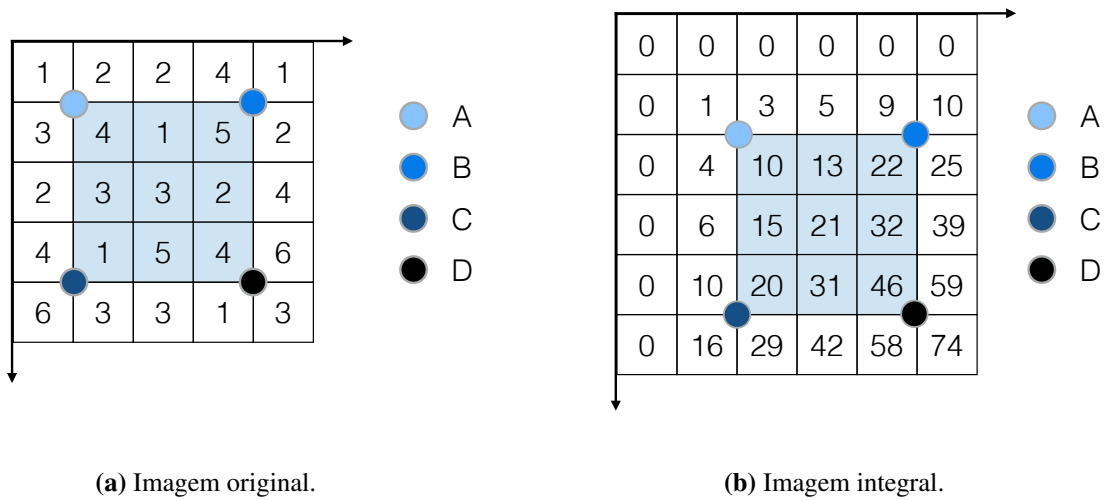


Figura 17: Exemplo de imagem integral.

Uma vez obtida a imagem integral, a tarefa de avaliar a área de qualquer janela requer

apenas o uso de quatro pontos na imagem. Seja a imagem integral I_Σ , a área da região delimitada pelos pontos $A = (x_0, y_0)$, $B = (x_1, y_0)$, $C = (x_0, y_1)$ e $D = (x_1, y_1)$ dispostos na Figura 17 é dada por

$$\sum_{\substack{x_0 < x \leq x_1 \\ y_0 < y \leq y_1}} I(x, y) = I_\Sigma(D) + I_\Sigma(A) - I_\Sigma(B) - I_\Sigma(C). \quad (2.8)$$

Esta propriedade a respeito do cômputo de área independente do tamanho é de fundamental relevância, visto que a operação de suavização requer esta informação de área para produzir as imagens da pilha no espaço de escala. Por exemplo, o cômputo da área da janela delimitada pelos pontos A, B, C e D na Figura 17 é

$$\begin{aligned} \sum_{\substack{x_0 < x \leq x_1 \\ y_0 < y \leq y_1}} I(x, y) &= I_\Sigma(D) + I_\Sigma(A) - I_\Sigma(B) - I_\Sigma(C) \\ 4 + 1 + 5 + 3 + 3 + 2 + 1 + 5 + 4 &= 1 + 46 - 9 - 10 \\ 28 &= 28. \end{aligned} \quad (2.9)$$

É fácil notar pela Eq. (2.9) que dependendo do tamanho da janela, o número de termos usado na imagem original aumenta ao passo que na imagem integral este cômputo é reduzido a apenas 4 termos.

As oitavas em SURF são o resultado da aplicação do filtro com tamanho diferente na mesma imagem integral de entrada, tarefa esta que pode até ser realizada de forma paralela. Isto se dá pelo fato de SURF adotar uma aproximação das derivadas parciais de segunda ordem das funções gaussianas através do determinante de uma matriz para detectar os pontos-chaves usando as imagens integrais. Tal cômputo se dá de forma rápida, independentemente do tamanho da imagem de entrada (BAY; TUYTELAARS; GOOL, 2006).

Em posse das imagens filtradas, a tarefa de detectar os pontos-chaves se dá pela análise do determinante da matriz hessiana de cada *pixel* $\mathbf{p} = (x, y)$ da imagem integral I_Σ

$$\mathcal{H}(\mathbf{p}, \sigma) = \begin{bmatrix} L_{xx}(\mathbf{p}, \sigma) & L_{xy}(\mathbf{p}, \sigma) \\ L_{xy}(\mathbf{p}, \sigma) & L_{yy}(\mathbf{p}, \sigma) \end{bmatrix}, \quad (2.10)$$

em que $L_{xx}(\mathbf{p}, \sigma)$ é o resultado da convolução da segunda derivada da função gaussiana $\frac{\partial^2 g(\sigma)}{\partial x^2}$ na escala σ com a imagem integral I_Σ . O mesmo ocorre de forma similar para $L_{xy}(\mathbf{p}, \sigma)$ e

$L_{yy}(\mathbf{p}, \sigma)$. A análise dos candidatos a pontos-chaves se dá pela supressão de não-máximos⁸ numa vizinhança $3 \times 3 \times 3$ (dimensões e escala) dos determinantes. A invariância à orientação é obtida através da análise das respostas da transformada *Wavelet* Haar usando uma janela circular centrada no ponto-chave em questão. Por fim, o descritor final com 64 atributos é computado. Mais detalhes sobre a formulação do método podem ser encontrados em Bay, Tuytelaars e Gool (2006).

2.5.3 Métricas de Avaliação

A seguir, são apresentadas as métricas utilizadas neste trabalho para avaliar a qualidade do processo de redimensionamento das imagens, a saber: SIFT Flow, IR-SSIM e SURF-SSIM.

2.5.3.1 SIFT Flow

A métrica SIFT Flow é uma técnica baseada em fluxo óptico (*optical flow*) que busca através da medição do deslocamento de cada *pixel* ou conjunto de *pixels* vizinhos e correspondência, inferir quão dissimilares as imagens são.

O fluxo óptico estima o rastreamento de movimento de um certo objeto numa sequência de imagens em que cada *pixel* está associado a um único vetor de velocidade (FARIA, 1992). Ele busca encontrar o padrão de movimento aparente dos objetos em uma cena, o qual é computado usando a restrição de brilho. Para tal, o fluxo óptico assume como constante o brilho dos *pixels* em quadros consecutivos correspondentes em uma imagem. Como o redimensionamento não interfere no brilho, não se tem uma variação na intensidade do brilho. Em síntese, o fluxo óptico localiza em cada *pixel* um vetor que indica o deslocamento que ocorreu dentro da imagem por meio do comparativo de matrizes.

Na métrica SIFT Flow, descritores SIFT são extraídos de cada *pixel* para caracterizar estruturas locais de imagem e codificar informação referente ao *pixel*. No entanto, ao invés de realizar a correspondência entre valores RGB, SIFT Flow usa os descritores SIFT no lugar. Ao final, um campo de deslocamento denso é obtido entre as imagens e o impacto visual é dado pela minimização da função custo que segue:

⁸Supressão de não máximos é a tarefa de encontrar todos os máximos locais na imagem e suprimi-los.

$$\begin{aligned}
SIFTflow(\mathbf{I}, \mathbf{I}') = \varphi(\mathbf{w}) = & \overbrace{\sum_{\mathbf{p}} \min(\|\mathbf{I}(\mathbf{p}) - \mathbf{I}'(\mathbf{p} + \mathbf{w}(\mathbf{p}))\|_1, t)}^{\text{termo de dados}} + \\
& \overbrace{\sum_{\mathbf{p}} \eta(|u(\mathbf{p})| + |v(\mathbf{p})|)}^{\text{termo de deslocamento,}} + \\
& \overbrace{\sum_{(\mathbf{p}, \mathbf{q}) \in \varepsilon} \min(\alpha|u(\mathbf{p}) - u(\mathbf{q})|, d) + \min(\alpha|v(\mathbf{p}) - v(\mathbf{q})|, d)}^{\text{termo de suavização}}, \quad (2.11)
\end{aligned}$$

na qual $\mathbf{w}$ é o vetor de deslocamento, u e v são as componentes espaciais do vetor de deslocamento $\mathbf{w}$ em $\mathbf{p} = (x, y)$, isto é, $\mathbf{w} = \{u(\mathbf{p}), v(\mathbf{p})\}$ e ε representa a vizinhança-4 de $\mathbf{p}$ em que $\mathbf{q}$ é um *pixel* $\in \varepsilon$.

A primeira linha da equação, o termo de dados, restringe os descritores SIFT a serem correspondidos com o vetor de fluxo $\mathbf{w}$. O termo de deslocamento, localizado na segunda linha, restringe os vetores de fluxo de forma que os mesmos sejam os menores possíveis quando nenhuma outra informação estiver disponível. O termo de suavização na terceira linha, restringe os vetores de fluxo de *pixels* adjacentes a serem similares. Para levar em consideração os valores de possíveis correspondências atípicas, isto é, correspondências inválidas, são empregadas penalizações com normas L1, tanto nos termos de dados e suavização em que t e d são limiares e η e α são pesos, quanto para os termos de deslocamento e suavização, respectivamente.

Por fim, a métrica mede a distorção do redimensionamento pela distância entre vetores de fluxos correspondentes entre imagens, sendo esta distância uma representação do impacto visual do redimensionamento. Em suma, quanto menor este custo for, menos perceptível deve ser o impacto visual.

2.5.3.2 Image retargeting SSIM

A métrica *Image retargeting SSIM* (IR-SSIM) (FANG *et al.*, 2014) se baseia no cômputo de similaridade através do ponderamento de um mapa de saliência global $\mathbf{I}_{SAL}$ por um mapa de similaridade local $\mathbf{I}_{SIM}$. O mapa de saliência global $\mathbf{I}_{SAL}$ é obtido pela combinação de dois mapas de importância gerados a partir da imagem original $\mathbf{I}$, isto é, $\mathbf{I}_{SAL} = \frac{1}{2}\mathbf{I}_{BU} + \frac{1}{2}\mathbf{I}_{TD}$ em que $\mathbf{I}_{BU}$ é o mapa de saliência gerado com base em características *bottom-up* e $\mathbf{I}_{TD}$ é o mapa de saliência gerado com base em características *top-down*. O mapa de similaridade local $\mathbf{I}_{SIM}$ é obtido através do cômputo da métrica SSIM (WANG *et al.*, 2004) em sub-janelas centradas em correspondências de *pixels* entre as imagens original $\mathbf{I}$ e a redimensionada $\mathbf{I}'$. Para cada *pixel*

$p = \{x, y\} \in \mathbf{I}$, o melhor *pixel* correspondente $p' = \{x', y'\} \in \mathbf{I}'$ é encontrado e suas respectivas vizinhanças formam as sub-janelas W_p e $W_{p'}$ que serão usadas no cômputo da similaridade local de modo que $\mathbf{I}_{SIM}(p) = SSIM(W_p, W_{p'})$. O escore final do IR-SSIM é o ponderamento do mapa de similaridade $\mathbf{I}_{SIM}$ pelo mapa de saliência global $\mathbf{I}_{SAL}$:

$$IR-SSIM(\mathbf{I}, \mathbf{I}') = \frac{1}{w \times h} \sum_{x=1}^w \sum_{y=1}^h \mathbf{I}_{SAL}(x, y) \times \mathbf{I}_{SIM}(x, y), \quad (2.12)$$

em que w e h correspondem à largura e à altura da imagem original $\mathbf{I}$. Neste caso, como cada *pixel* de $\mathbf{I}$ tem uma correspondência em $\mathbf{I}'$, o cômputo de similaridade é o que informa de fato o quão aquele *pixel* e sua vizinhança foram preservados no redimensionamento, independente se a correspondência for fraca, isto é, o melhor *pixel* correspondente de p é na verdade um *pixel* de outra região da imagem. Na Figura 18 é ilustrado o fluxograma da métrica no qual as regiões escuras que indicam correspondências fracas.

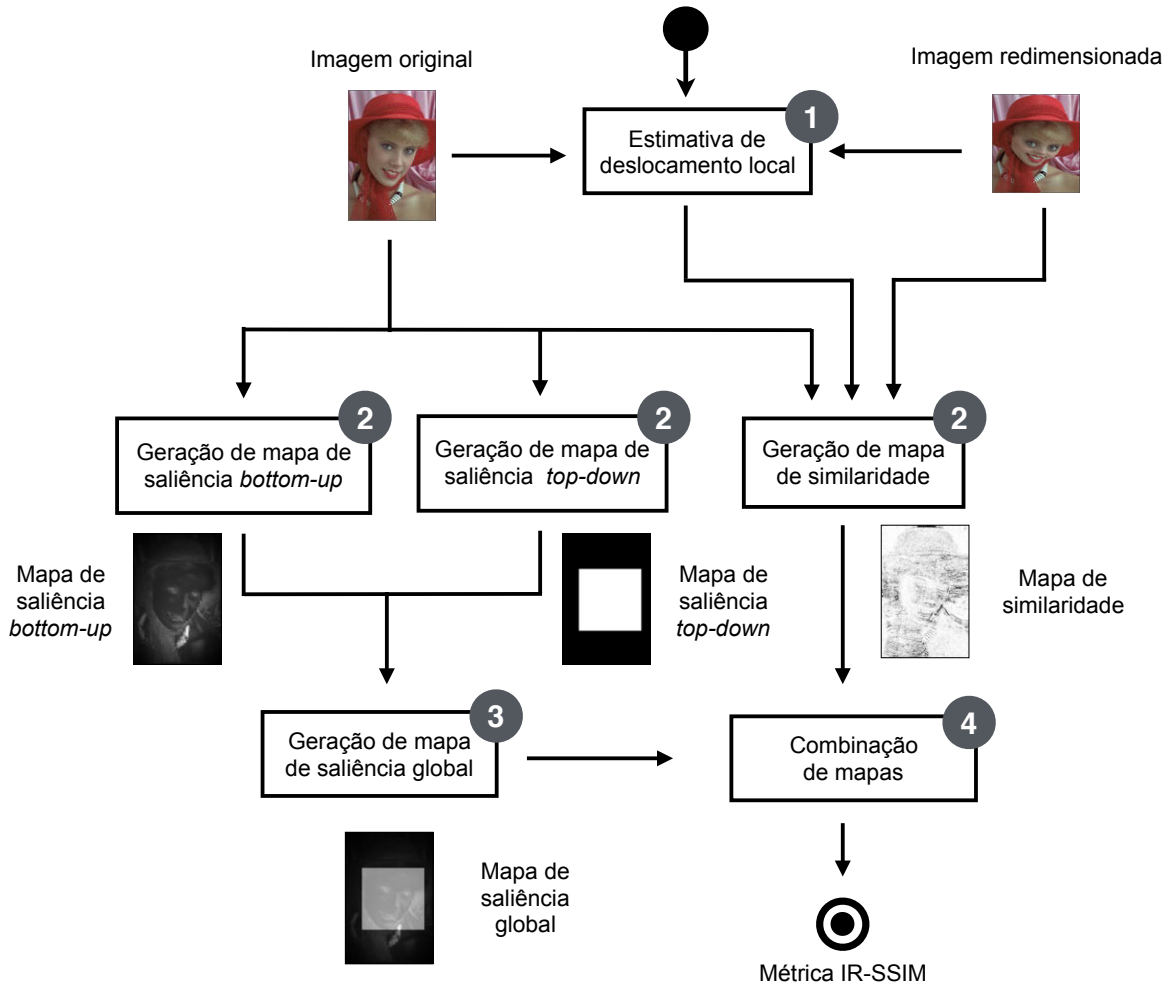


Figura 18: Fluxograma do IR-SSIM (adaptado de Fang *et al.* (2014)).

Sem conhecimento adicional sobre a imagem, é difícil determinar uma lista completa

das características *top-down* para ser incluídas na modelagem de detecção de saliência. Na Figura 18, por exemplo, é dado um enfoque maior aos rostos humanos, uma vez que este fator tem se mostrado atraente em termos de atenção visual. Entretanto, vale ressaltar que outras características *top-down* podem ser empregadas com o intuito de melhorar a avaliação dos resultados. Além disso, a métrica em si pode ser implementada de diversas maneiras, dependendo do algoritmo utilizado para estimativa de deslocamento e correspondência, estimativa de saliência e cômputo de similaridade.

2.6 Conclusão

Redimensionar imagens levando em consideração seu conteúdo é uma tarefa complexa, especialmente quando se almeja uma boa representação do estado original da imagem. Neste contexto, o problema é definido como preservar as regiões mais importantes, enquanto minimiza-se as distorções visuais ao passo que se satisfazem as preferências do usuário sob restrições geométricas.

Dentro do arcabouço de técnicas para realizar a tarefa de redimensionamento sensível ao conteúdo, o operador *Seam Carving* fornece uma abordagem sofisticada que se dá através da exclusão seletiva de *pixels* em forma de caminhos, chamados de *seams*. Apesar de uma abordagem menos restritiva, no tocante ao custo computacional, originalmente a busca por estes caminhos ocorre de forma exaustiva. A fim de atenuar tanto o custo computacional, assim como fornecer uma abordagem *multi-seam*, propõe-se no Capítulo 4, uma nova abordagem de busca pelo uso de Algoritmos Genéticos, uma técnica de Computação Evolutiva.

Além disso, foi visto que projetar uma métrica de qualidade para redimensionamento capaz de comparar o conteúdo da imagem sob variações de proporção é significativamente mais desafiador, uma vez que o problema também exige tanto uma análise semântica quanto uma análise em termos de correspondência de conteúdo. Uma vez que os seres humanos são os consumidores finais das imagens redimensionadas e, portanto, os juízes da qualidade da imagem é necessário que as métricas sejam bem relacionadas com os critérios de avaliação subjetivos.

Pode-se notar também uma forte tendência entre as métricas que usam descritores locais para realizar a correspondência entre imagens e as avaliações subjetivas, uma vez que se torna dispensável a realização de etapas de pré ou pós-processamento, tais como, utilizar estruturas de dados complexas ou até mesmo resolver problemas de otimização global. Nota-se também que estes escores nem sempre concordam entre si. Esta discordância ocorre em função do processo de formulação das métricas que são distintos e seguem caminhos alternativos para medir uma

mesma informação subjetiva.

ALGORITMOS GENÉTICOS

Algoritmos Genéticos pertencem ao grupo de modelos computacionais que usam a evolução natural e a genética como metáfora para resolver problemas. Algoritmos Genéticos podem ser usados para gerar soluções úteis para problemas de otimização ou em problemas onde o espaço de busca é demasiado grande e seria impraticável explorá-lo por uma abordagem exaustiva. O objetivo deste capítulo é descrever a fundamentação teórica necessária sobre Algoritmos Genéticos com o intuito de facilitar o entendimento da proposta para a obtenção de *seams* apresentada no próximo capítulo. Vale ressaltar que grande parte da fundamentação teórica apresentada neste capítulo encontram-se nas obras de Linden (2008) e Sivanandam e Deepa (2008).

3.1 Introdução

A sobrevivência dos mais aptos é um dos pilares da Teoria da Evolução de Charles Darwin, um dos maiores feitos da ciência do século XIX (LINDEN, 2008). A partir da década de 60 do século passado, com o advento dos computadores, os cientistas passaram a aplicar o mesmo princípio para tentar resolver problemas matemáticos difíceis. A ideia é gerar um grande conjunto de possíveis soluções numéricas (no papel do indivíduo), gerar novas soluções a partir das atuais (no papel da reprodução e mutação) e eliminar as demais. Deste modo, após algum processamento (gerações), algumas soluções deveriam estar muito próximas de uma resposta ótima do problema.

Alguns anos depois, em 1975, o pesquisador John Holland publicou um livro, *Adaptation in Natural and Artificial Systems*, no qual usava tais algoritmos, denominados Algoritmos Genéticos (*Genetic Algorithms*, GAs) para simular processos evolutivos, o que impulsionou o seu uso no campo hoje conhecido como Inteligência Computacional. Holland estudou formalmente a evolução das espécies e propôs um modelo heurístico computacional que quando implementado poderia oferecer boas soluções para problemas extremamente difíceis que eram insolúveis computacionalmente até aquele momento (LINDEN, 2008).

Algoritmos Genéticos pertencem ao grupo de modelos computacionais que compreendem a área de Computação Evolutiva (CE). CE usa modelos computacionais dos processos naturais de evolução como uma ferramenta para resolver problemas. Apesar de haver uma grande variedade de modelos computacionais propostos, há quatro paradigmas históricos que serviram como base para grande parte da atividade do campo: Programação Evolutiva (FOGEL; OWENS; WALSH, 1966), Estratégias Evolutivas (RECHENBERG, 1973), Algoritmos Genéticos (HOLLAND, 1992) e Programação Genética (KOZA, 1992, 1994). As diferenças básicas entre os paradigmas estão na natureza dos modelos de representação, os operadores de reprodução e métodos de seleção.

A popularização relacionada a utilização de GAs quando comparados a outros métodos de CE, justifica-se principalmente pela eficácia na solução encontrada, como também pela flexibilidade e simplicidade de implementação (CARDOSO, 1998; SIVAPATHASEKARAN *et al.*, 2010).

3.2 Visão geral

Algoritmos Genéticos são um método de pesquisa meta-heurístico inspirado na evolução natural que utiliza de técnicas como herança, mutação, a seleção natural e reprodução. Esta meta-heurística pode ser usada para gerar soluções úteis para problemas de otimização ou em problemas onde o espaço de busca é demasiado grande e seria impraticável pesquisá-lo por uma abordagem exaustiva.

Devido às características de GAs, é mais fácil solucionar alguns tipos de problemas do que utilizar outros métodos matemáticos que dependem da hipótese de linearidade, diferenciabilidade, continuidade, ou convexidade da função objetivo (OLIVEIRA; BEZERRA; ROCHA NETO, 2015). Além disto, GAs diferenciam-se de modelos estritamente aleatórios por serem uma busca direcionada que utiliza informação pertinente ao problema e não trabalham apenas com exploração aleatória (*random walks*) pelo espaço de soluções. Isto é garantido pelo fato de que GAs usam o valor da função de avaliação como guia na escolha dos indivíduos a serem reproduzidos. Apesar de existirem fatores estocásticos associados ao processo, eles são usados para tentar caminhar na direção correta.

Em GAs, mantém-se um conjunto de indivíduos (população) que evolui de forma semelhante à evolução das espécies, de tal maneira que cada indivíduo é uma representação abstrata de uma solução do problema (SIVANANDAM; DEEPA, 2008). Cada indivíduo da população recebe uma nota (valor de aptidão) indicando a sua qualidade como solução. Com

base nesta aptidão são selecionados os indivíduos de tal maneira que os mais aptos (medido pela função de adaptação) são geralmente mais suscetíveis de serem selecionados.

Os indivíduos selecionados servem de entrada para as etapas de cruzamento e mutação. A etapa de cruzamento tem como objetivo a criação de novos indivíduos a partir de indivíduos previamente selecionados. Tais indivíduos herdam algumas características dos pais envolvidos, de forma similar à reprodução sexuada. Depois da etapa de cruzamento, é realizado um processo de mutação. O objetivo desta etapa de mutação é diversificar a população, alterando o material genético de alguns poucos indivíduos através de um método probabilístico. Ao final deste processo de aplicação de operadores genéticos, tem-se uma geração de novos indivíduos.

Após as etapas de cruzamento e mutação, uma estratégia para fortalecer uma nova geração pode ser aplicada de modo a melhorar a qualidade das soluções existentes. Uma dessas estratégias, o elitismo, visa garantir este fortalecimento da população através do aproveitamento, isto é, da inserção, de uma certa quantidade indivíduos mais aptos da geração anterior para a atual. Tal abordagem permite que GAs não descartem possíveis soluções ótimas a cada geração.

Após a formação de uma geração, os critérios de convergência são verificados. Geralmente procura-se algum indivíduo naquela geração que seja bom o suficiente para ser usado como uma solução para o problema ou se uma determinada condição previamente estabelecida aconteceu. Se o critério de parada não for alcançado, o processo se repete a partir da etapa de seleção até que algum critério de parada seja atendido. Um fluxo genérico em GAs é mostrado na Figura 19.

Uma analogia entre os termos/conceitos utilizados em genética e a teoria da evolução das espécies é apresentada na Tabela 1.

Tabela 1: Terminologia para Algoritmos Genéticos.

Linguagem Natural	Algoritmos Genéticos
Cromossomo	Indivíduo ou Solução
Gene	Característica
Alelo	Valor de um gene
Locus	Posição do gene no indivíduo
Genótipo	Estrutura
Fenótipo	Conjunto de parâmetros

A modelagem adequada da representação genética e seus operadores é crítico, uma vez que o comportamento dos GAs é extremamente dependente deles. Frequentemente, torna-se difícil

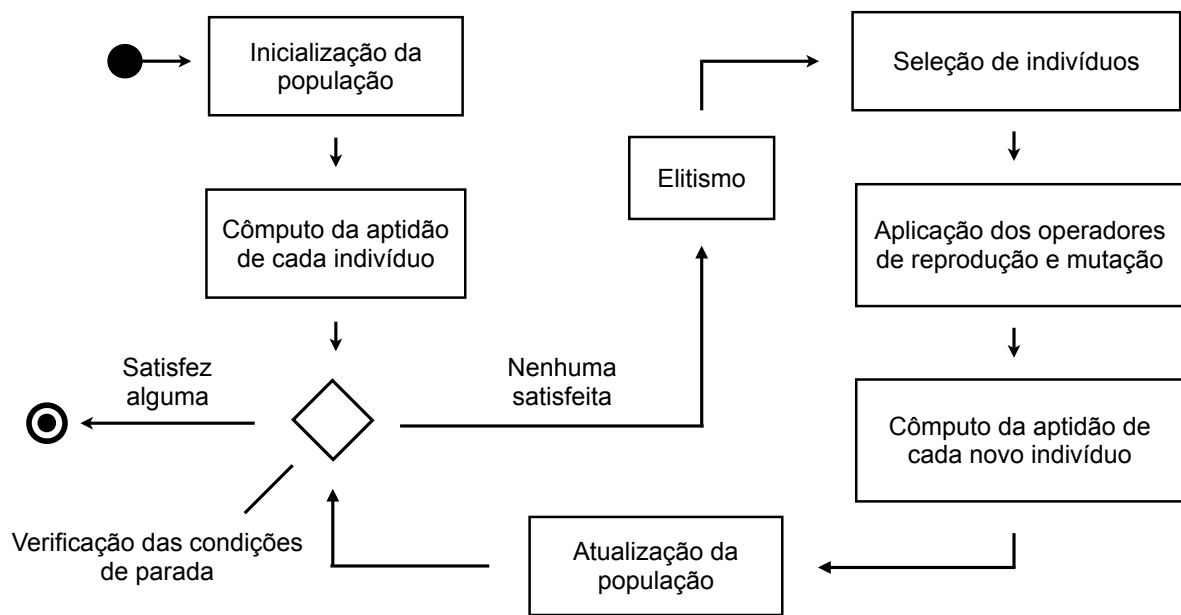


Figura 19: Fluxo genérico de GAs.

encontrar uma representação genética que respeite a estrutura do espaço de busca e operadores genéticos que sejam coerentes com as propriedades do problema (SIVANANDAM; DEEPA, 2008).

Nas subseções seguintes são apresentadas informações sobre representação genética, função de aptidão, operadores de cruzamento, de mutação e de seleção, e por fim condições de parada em GAs.

3.3 Representação genética

A representação genética é o ato de traduzir a informação do problema para uma representação que seja tratável pelo computador. Inicialmente, os indivíduos são representados por cadeias de *bits* (*strings*) ou algum outro conjunto de símbolos de um conjunto pré-definido. A representação de soluções por uma cadeia de *bits* se fundamenta na ideia de que a memória do computador é feita de vetores de *bits*, e que qualquer dado armazenado no computador também pode ser codificado em uma cadeia de bits de tamanho suficiente.

Entretanto, vale ressaltar que a representação genética é completamente arbitrária e deve ser feita de forma simples. Além disso, as restrições de representação a serem impostas devem estar bem mapeadas, assim como as soluções não permitidas devem ser invalidadas ou penalizadas. A representação está diretamente ligada à função de aptidão, uma vez que esta utiliza todo o

conhecimento acerca do problema para atribuir valores de aptidão aos indivíduos.

O termo indivíduo e cromossomo geralmente são intercambiáveis, uma vez que abordagens que utilizem indivíduos com somente um cromossomo são muito comuns. Os indivíduos são formados por cromossomos e estes por genes. Genes representam características da solução e possuem um valor associado, dentre os possíveis no contexto do problema. Alelo é o nome dado ao valor de um gene, enquanto a sua posição no material é tratada por locus.

A codificação da informação em cromossomos é um ponto crucial no projeto de GAs e é, junto com a função de aptidão, o que conecta GAs ao problema a ser resolvido. Se a codificação for feita de forma adequada, incluindo as peculiaridades do problema (restrições sobre o domínio, etc), são evitados testes de viabilidade de cada uma das soluções geradas. Na Figura 20 são ilustrados exemplos de codificação de indivíduos em Algoritmos Genéticos.

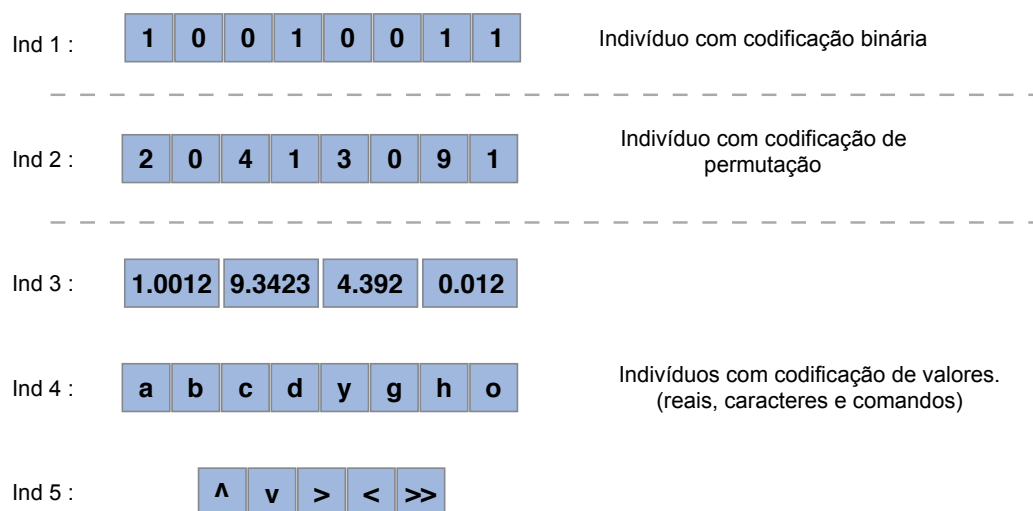


Figura 20: Exemplos de diferentes codificações de indivíduos em Algoritmos Genéticos.

3.4 Função de aptidão

Informa a pontuação (aptidão) de um indivíduo em relação aos demais da população em função do objetivo a ser otimizado. Para o cálculo dessa aptidão, o indivíduo tem de ser primeiro decodificado para então ser entrada da função de aptidão. Essa pontuação em última análise representará a chance deste indivíduo participar do processo reprodutivo nas próximas gerações (WHITLEY, 1994). Além disso, essa pontuação também traz informação sobre quão próximo o indivíduo está do indivíduo ótimo daquela iteração (mais apto). Um ponto muito importante a ser levado em consideração quando se calcula a aptidão de um indivíduo é que a

função de aptidão tenha seu domínio nos reais positivos. Isto se faz necessário, pois alguns critérios de paradas e outros mecanismos que se utilizam da média dos valores de aptidão selecionem indivíduos fracos para continuarem para a próxima geração devido aos valores negativos.

3.5 População

A população é uma coleção de indivíduos e seu tamanho depende da complexidade do problema alvo. A geração da inicialização inicial feita de forma aleatória, mas também pode ser feita com base em alguma heurística que seja capaz de fornecer bons indivíduos iniciais. Assim, a média de aptidão da população pode começar alta, resultando em uma convergência mais rápida em GAs. Entretanto, para isto de fato funcionar, deve-se ter a certeza que o conjunto de indivíduos iniciais é suficientemente representativo, ou seja, que estejam distribuídos pelo espaço de busca. Caso contrário, a população carece de diversidade e com isso o algoritmo irá explorar apenas uma pequena parcela do espaço de busca, podendo nunca encontrar as melhores soluções globais (RAHNAMAYAN; TIZHOOSH; SALAMA, 2007).

O tamanho da população também levanta algumas discussões. Quanto maior o número, mais próximo de uma busca exaustiva os GAs se aproximam. Entretanto, uma população pequena pode não convergir tão facilmente. Na prática, um tamanho de população de por volta de 100 indivíduos é bastante comum. Entretanto, este tamanho pode ser alterado de acordo com o processamento e memória disponíveis a fim de melhorar a qualidade do resultado a ser atingido (SIVANANDAM; DEEPA, 2008).

3.6 Operadores genéticos

A aplicação dos operadores de reprodução é de fundamental importância em GAs (MAN; TANG; KWONG, 2012). Os operadores de cruzamento e mutação, entram no processo evolutivo para atuar como intensificadores e diversificadores da busca, respectivamente, almejando criar indivíduos mais aptos.

3.6.1 Operador de seleção

Na natureza, o processo de seleção garante que os indivíduos mais adaptados tenham maiores chances de sobreviver. O conceito utilizado em GAs é o mesmo. A seleção é

aplicada na população corrente para criar uma população intermediária que irá passar pelos processos de reprodução e mutação (apresentados mais adiante) para formar, então, a próxima geração (GROSKO; GORSKI; DIAS, 2006). Nas subseções que seguem são apresentados alguns métodos de seleção e suas características.

Método da roleta

O método da roleta (LIPOWSKI; LIPOWSKA, 2012) é um dos mais tradicionais métodos de seleção de indivíduos em GAs. A ideia básica do método da roleta consiste em aplicar uma busca linear por meio de uma roleta com as divisões ponderados na proporção dos valores de aptidão dos indivíduos. O funcionamento do método da roleta pode ser resumido da seguinte maneira: o valor de aptidão de um indivíduo (aptidão absoluta) é dividido pelo valor de aptidão atual da população, resultando na aptidão relativa do indivíduo. A cada indivíduo é atribuído uma fatia da roleta, no qual o tamanho da fatia é proporcional ao valor de aptidão dos indivíduos (aptidão relativa). A roleta é girada N vezes, onde N é o número de indivíduos da população. Em cada giro, o indivíduo sob o marcador da roleta é selecionado para estar no conjunto dos pais para a próxima geração. Uma visão geral do método pode ser vista na Figura 21 em que a aptidão relativa está sendo representada por um gráfico de pizza, enquanto que a aptidão absoluta se encontra na tabela desta mesma figura. Um algoritmo para o método da roleta é apresentado a seguir:

1. Seja $T = \sum_{i=1}^N fitness(c_i)$ o valor da soma de todas as aptidões da população.
2. Gire a roleta N vezes:
 - (a) Gere um valor aleatório $r \in [0, T]$.
 - (b) Itere sobre a população, somando os valores de aptidão até que a soma seja maior ou igual a r . O indivíduo cuja contribuição na roleta ultrapasse esse limite r é selecionado.

O método da roleta é fácil de implementar, porém é suscetível ao ruído de modo que a taxa de evolução depende da variância da aptidão da população (SIVANANDAM; DEEPA, 2008).

Seleção aleatória

Esta técnica seleciona aleatoriamente um dos pais a partir da população. Em termos de desmembramento de material genético, a seleção aleatória possui um pouco mais de variabilidade que o método da roleta. Um algoritmo para o método da seleção aleatória é apresentado a seguir:

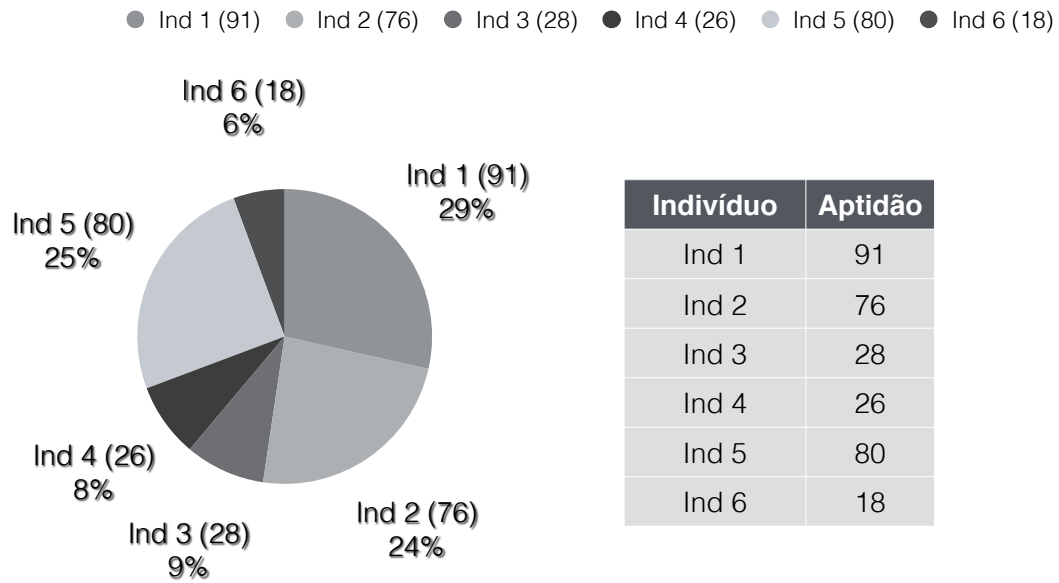


Figura 21: Método da roleta.

1. Seja M o tamanho da população.
2. Execute as instruções abaixo N vezes:
 - (a) Gere um valor aleatório r entre $\in [0, M]$.
 - (b) O indivíduo cuja posição na população é dada por r é selecionado.

Seleção ranqueada

O método da roleta sofre quando os valores de aptidão diferem muito entre si. Se o valor de aptidão relativa do indivíduo mais apto é 90%, sua circunferência também ocupa 90% da roleta de modo que a probabilidade de seleção dos demais indivíduos é baixa. Para evitar este cenário, a seleção ranqueada (SAWYERR; ALI; ADEWUMI, 2011) classifica a população através da atribuição de uma posição usando a aptidão de cada indivíduo em que o pior indivíduo tem aptidão 1 e o melhor tem aptidão N . Isto resulta em convergência lenta, mas impede uma convergência rápida que previne, de certo modo, mínimos locais e se beneficia quando a variação de aptidão na população é baixa, preservando a diversidade e, por esta razão, leva a uma busca bem sucedida. São apresentadas duas possíveis versões do operador de seleção ranqueada, a saber:

1. (a) Ranqueie a população por aptidão.
- (b) Defina um valor de $r \in [0, 1]$

- (c) Selecione um par de indivíduos de forma aleatória, isto é, gere dois números aleatórios, p e $q \in [1, N]$, $p \neq q$.
 - (d) Gere um número aleatório $R \in [0, 1]$.
 - (e) Se $R < r$, use o primeiro indivíduo do par como pai. Caso contrário, então use o segundo indivíduo.
 - (f) Repita isso até selecionar todos os pais.
2. (a) Ranqueie a população por aptidão.
- (b) Selecione um par de indivíduos de forma aleatória.
 - (c) Selecione o indivíduo do par com o maior valor de aptidão para ser o primeiro pai.
 - (d) Repita o processo para encontrar o segundo pai.
 - (e) Repita todos os passos anteriores até selecionar todos os pais.

Seleção por torneio

Em uma estratégia de seleção ideal, o ajuste da pressão de seleção¹ e diversidade da população devem refinar o desempenho da busca em GAs. Diferentemente do método da roleta, a seleção por torneio provê o ajuste da pressão de seleção através de uma competição em forma de torneio entre N_u indivíduos selecionados aleatoriamente. O melhor indivíduo do torneio, e também o vencedor do torneio, é aquele cujo valor de aptidão é o mais alto entre os N_u indivíduos. O vencedor do torneio é então inserido no conjunto dos pais. Este processo é repetido até que se preencha o conjunto de pais com os vencedores caracterizando um conjunto diversificado em termos de aptidão.

Seleção por elitismo

Neste método de seleção são mantidos os melhores indivíduos da geração atual na geração seguinte. Uma variação dela substitui os piores indivíduos da geração atual pelos melhores da anterior. Ambas abordagens visam fortalecer a geração seguinte a fim de melhorar a qualidade das soluções existentes.

3.6.2 Operadores de reprodução

3.6.3 Cruzamento

Cruzamento (*crossing over*) é o processo através do qual ocorre a recombinação do material genético de pares de soluções (pais) com o intuito de criar soluções derivadas (filhas) melhores.

¹ A pressão de seleção é a capacidade de melhorar a seleção quando a variação de aptidão é baixa.

Após o operador de seleção escolher os indivíduos, tem-se um conjunto de possíveis pais no qual um operador de reprodução atua como segue:

1. Seleciona-se de forma aleatória dois indivíduos do conjunto de pais;
2. Cadeias de material genético de ambos os pais são selecionadas de acordo com algum critério pré-definido;
3. Por fim, estas cadeias são trocadas entre os pais e novos filhos são incluídos na população da próxima geração.

Cruzamento uniponto

No operador de cruzamento uniponto, dois pais têm seus materiais genéticos divididos a partir de um ponto escolhido de forma aleatória. As partes do material genético de ambos pais antes desse ponto são trocadas. Se um ponto apropriado for escolhido, filhos melhores podem ser obtidos através das combinações de pais com bons valores de aptidão, caso contrário esse operador pode produzir filhos com baixa qualidade. O funcionamento deste operador é ilustrado na Figura 22.



Figura 22: Cruzamento pelo operador uniponto.

Cruzamento por dois pontos

Diferentemente do operador de ponto único, neste operador dois pontos são escolhidos e os materiais genéticos (blocos) entre os dois pais são trocados entre si. A vantagem imediata desse tipo de operador com mais de um ponto é que a busca no espaço de soluções é mais intensa. Entretanto, blocos de genes são mais prováveis de serem quebrados. É ilustrado na Figura 23 o funcionamento de um operador de dois pontos.

Cruzamento por multi-pontos



Figura 23: Cruzamento pelo operador de dois pontos.

Existem duas formas de realizar o cruzamento com um operador multi-ponto. Uma forma é ter uma quantidade par de pontos de corte e a outra é ímpar. No caso da quantidade par, os pontos são escolhidos aleatoriamente de forma circular e os materiais genéticos são trocados. No caso ímpar, sempre um ponto de corte é fixo no início do materiais genético. É apresentado na Figura 24 o funcionamento de um operador multi-ponto com uma quantidade par de cortes.



Figura 24: Cruzamento pelo operador multi-ponto.

Cruzamento uniforme

No operador de cruzamento uniforme (SEMENKIN; SEMENKINA, 2012), diferentemente do multi-ponto, cada gene nos filhos é copiado diretamente de um pai ou de outro de acordo com uma máscara binária de mesmo comprimento do material genético. Nessa máscara, quando se tem o valor 1 o gene é copiado do primeiro pai e quando há o valor 0 o gene é copiado do segundo pai. Uma nova máscara é aleatoriamente gerada para cada par de pais. Desta maneira os filhos contêm uma mistura de genes de cada par, sendo que o número de pontos de corte não é fixo como mostrado nos operadores anteriormente apresentados, mas em média fica em torno da metade do comprimento do material genético. É apresentado na Figura 25 o funcionamento

de um operador multi-ponto com uma quantidade par de cortes.

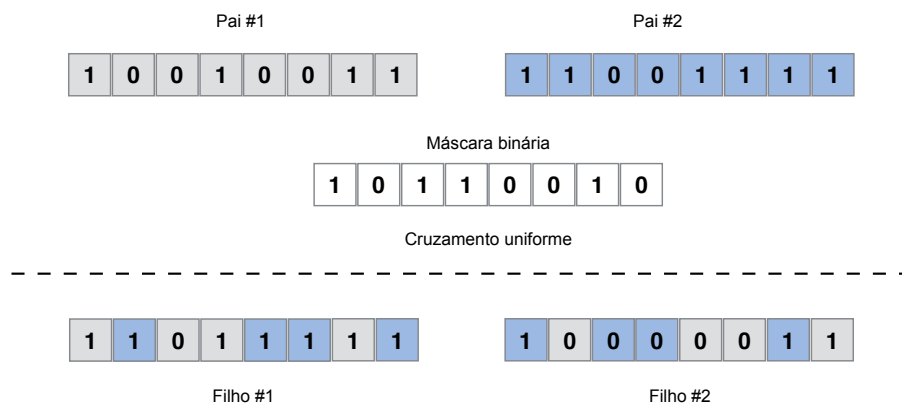


Figura 25: Cruzamento pelo operador uniforme.

3.6.4 Operador de mutação

No processo de evolução das soluções, os indivíduos são sujeitos à mutação. Nesta etapa, o operador de mutação atua sobre uma quantidade muito pequena da população, oferecendo uma componente aleatória a fim de evitar que o algoritmo fique preso em mínimos locais. O operador de mutação é tradicionalmente considerada como um operador de busca simples e geralmente é visto como um operador de plano de fundo que mantém a diversidade genética na população. Se o operador de cruzamento é tido como o fator intensificador da busca em GAs, o de mutação atua então como o fator diversificador permitindo explorar mais regiões do espaço de busca.

Um parâmetro importante da mutação é a probabilidade da mutação que informa com que frequência o material genético deve ser modificado. Se não existir mutação, os filhos são gerados imediatamente após o cruzamento sem qualquer modificação posterior. No entanto, mais partes de um indivíduo são modificadas quando realizada. Se a probabilidade de mutação é 100% todo o indivíduo é modificado, se é 0% então nada é modificado. De um modo geral, a mutação não deve ocorrer muito frequentemente, pois configura uma busca puramente aleatória (SIVANANDAM; DEEPA, 2008).

Mutação por giro

Mutação por giro (*flipping*) envolve modificar os valores de um gene baseado na faixa de valores possíveis, por exemplo, usando uma representação binária, mudar de 0 para 1 ou vice-versa em um indivíduo gerado. Na Figura 26 é mostrado o funcionamento do operador de mutação por giro.



Figura 26: Operador de mutação por giro.

Mutação por permutação

Neste operador posições aleatórias do material genético são escolhidas e seus genes correspondentes a estas posições são permutados (SERPELL; SMITH, 2010). Um exemplo de funcionamento do operador de mutação por permutação é apresentado na Figura 27.

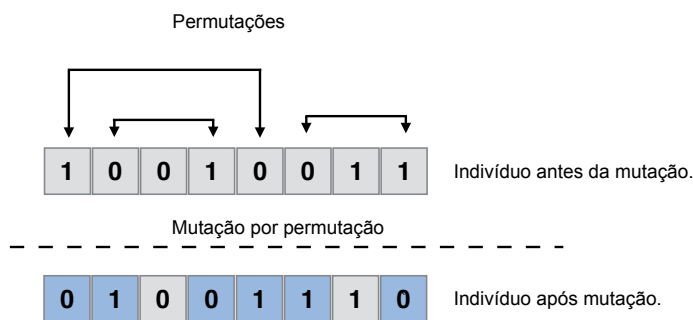


Figura 27: Operador de mutação por permutação.

Mutação por reversão

Neste operador uma posição é escolhida aleatoriamente e os genes seguintes àquela posição sejam invertidos para a geração de um novo filho. Um exemplo de funcionamento do operador de mutação por reversão é apresentado na Figura 28.

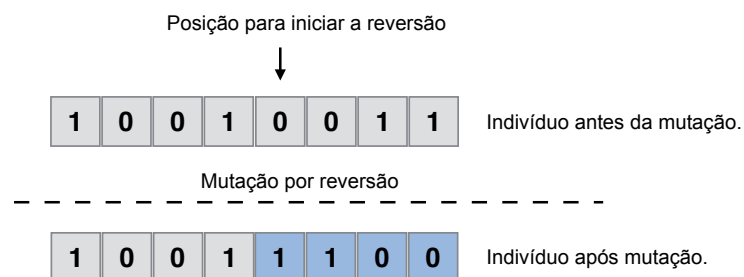


Figura 28: Operador de mutação por reversão.

3.7 Critérios de convergência

Os critérios de convergência (ou término) encerram a busca em GAs. Critérios podem ser combinados para prover um mecanismo capaz de verificar uma possível estagnação do processo evolutivo ou alternativamente, para verificar uma restrição em termos de processamento computacional. Alguns critérios de convergência são apresentados a seguir:

1. **Número máximo de gerações:** o processo de busca é finalizado com base no número de gerações já realizadas.
2. **Tempo máximo de execução:** o processo de busca é finalizado com base no tempo já transcorrido. Lembrando que se o número máximo de gerações for alcançado primeiro, GAs tem sua execução encerrada também.
3. **Sem mudança no melhor valor de aptidão:** a busca é encerrada quando não há mudança no melhor valor de aptidão encontrado durante um determinado número de gerações.
4. **Sem mudança na função objetivo em consecutivas gerações:** o processo de busca é interrompido quando não há uma melhoria na função objetivo para um número definido de gerações consecutivas.
5. **Sem mudança na função objetivo durante um intervalo de tempo:** o processo de busca é finalizado quando não há uma melhoria na função objetivo durante um intervalo de tempo definido.

3.8 Conclusão

GAs são ferramentas de busca inspiradas na Teoria da Evolução. Uma das características mais marcantes dessa ferramenta é o fato de realizar uma busca paralela e direcionada no espaço de soluções durante o processo de evolução. Outra capacidade interessante de GAs é a de conseguir diversidade na população a fim de evitar máximos/mínimos locais.

Devido as suas características, é mais fácil solucionar alguns tipos de problemas usando GAs do que utilizar alguns métodos matemáticos clássicos que dependem da hipótese de linearidade, diferenciabilidade, continuidade ou convexidade da função objetivo. Tal comportamento tornam GAs uma ferramenta bem atraente.

A escolha no projeto de GAs da representação do indivíduo, da função de aptidão e dos operadores genéticos impactam na convergência também. Outros parâmetros, tais como o

tamanho da população e configuração de operadores também têm sua parcela de influência no processo. Logo, a escolha incorreta de tais fatores pode acarretar em soluções não tão satisfatórias ou até mesmo custosas em termos de computação.

GENETIC SEAM CARVING (PROPOSTA)

Neste capítulo, nossa proposta, denominada *Genetic Seam Carving* (GSC) é apresentada. GSC surgiu de um esforço inicial em que se almejava realizar a etapa de busca de *seams* via GAs. Posteriormente, o operador GSC foi aprimorado com base na adição do SURF. Esta proposta aprimorada foi denominada *SURF-Based Genetic Seam Carving* (S-GSC) a qual é explicada mais à frente. Além da proposta do operador redimensionamento, neste trabalho também tem-se a contribuição de uma métrica para avaliar resultados de redimensionamento como as vistas no Capítulo 3. Para que haja uma modelagem bem sucedida, é preciso definir adequadamente o indivíduo, a função de aptidão, bem como utilizar os operadores genéticos padrões ou definir novos quando necessário. A seguir, são apresentados os elementos que compõem o *Genetic Seam Carving*, e em seguida, a formulação da métrica SURF-SSIM.

4.1 Indivíduos ou cromossomos

O indivíduo para o *Genetic Seam Carving* é composto somente por um cromossomo

$$\mathbf{c} = [g_1, \dots, g_{i-1}, g_{i=p}, g_{i+1}, \dots, g_m],$$

em que m é a altura da imagem e, ao mesmo tempo, o comprimento do indivíduo, enquanto p é a posição do gene pivô. Além disso, cada gene é do tipo inteiro e seus valores são limitados como descrito a seguir.

$$g_i \in \begin{cases} \{1, \dots, n\}, & \text{se } i = p; \\ \{-1, 0, 1\}, & \text{caso contrário;} \end{cases}$$

O gene pivô indica no *seam* o ponto no qual a reconstrução (*decoding*) será realizada. Para

realizar essa reconstrução, a função

$$f(i) = \begin{cases} g_i, & \text{se } i = p \text{ (gene pivô);} \\ \sum_{j=p}^i g_j, & \text{se } i > p \text{ (somatório do alelo do pivô até o alelo do gene } g_i); \\ \sum_{j=i}^p g_j, & \text{se } i < p \text{ (somatório do alelo do gene em questão } g_i \text{ até o gene pivô);} \end{cases} \quad (4.1)$$

é utilizada para obter o valor de um gene qualquer em sua coordenada correspondente na imagem. A transformação de um indivíduo em um *seam* é dada pela reconstrução de cada gene do indivíduo em sua respectiva coordenada. A definição de um *seam* vertical é dada agora em termos de genes por

$$\mathbf{s}^v = \{(i, f(i))\}_{i=1}^m,$$

e um *seam* horizontal por

$$\mathbf{s}^h = \{(f(i), i)\}_{i=1}^n.$$

Como pode-se perceber que

$$|f(i) - f(i+1)| \leq 1 \text{ tal que } , \forall i, 1 \leq i \leq m-1,$$

e, assim, esta é uma função de mapeamento adequada, pois permite que os indivíduos representem caminhos 8-conexos.

O processo de transformação de um indivíduo em *seam* é exemplificado na Figura 29.

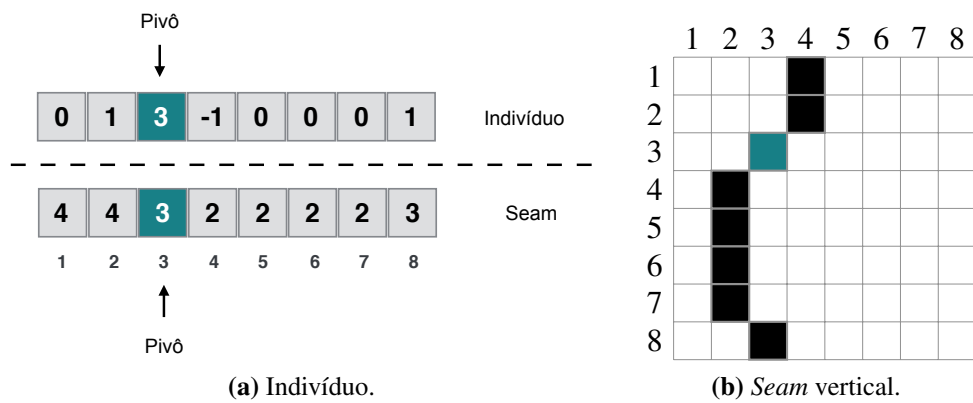


Figura 29: Transformação de um indivíduo em *seam* em que o pixel em destaque representa o pivô.

4.2 Operadores genéticos

Para a modelagem adequada do GSC foram desenvolvidos operadores genéticos específicos para explorar o espaço de busca. Uma modificação do operador de cruzamento de ponto simples foi realizada para lidar com a seguinte restrição: cada indivíduo gerado deve possuir um pivô. Sem esta modificação, é possível gerar indivíduos com dois pivôs ou até mesmo sem pivô.

Operador de seleção

O operador de seleção escolhido em GSC foi o de seleção aleatória. Além disso, de modo a aproveitar indivíduos com altos valores de aptidão de uma geração para a outra, foi adotado o elitismo.

Operador de cruzamento

Inicialmente, o gene pivô é removido do material genético e, posteriormente, com o material composto apenas por genes $\in \{-1, 0, 1\}$ se escolhe um ponto aleatoriamente para dividir o material de ambos os pais. Em seguida, a primeira parte do material genético de cada pai é anexada ao outro para gerar filhos. Por fim, os genes pivôs são colocados de volta na posição original como mostra a Figura 30.

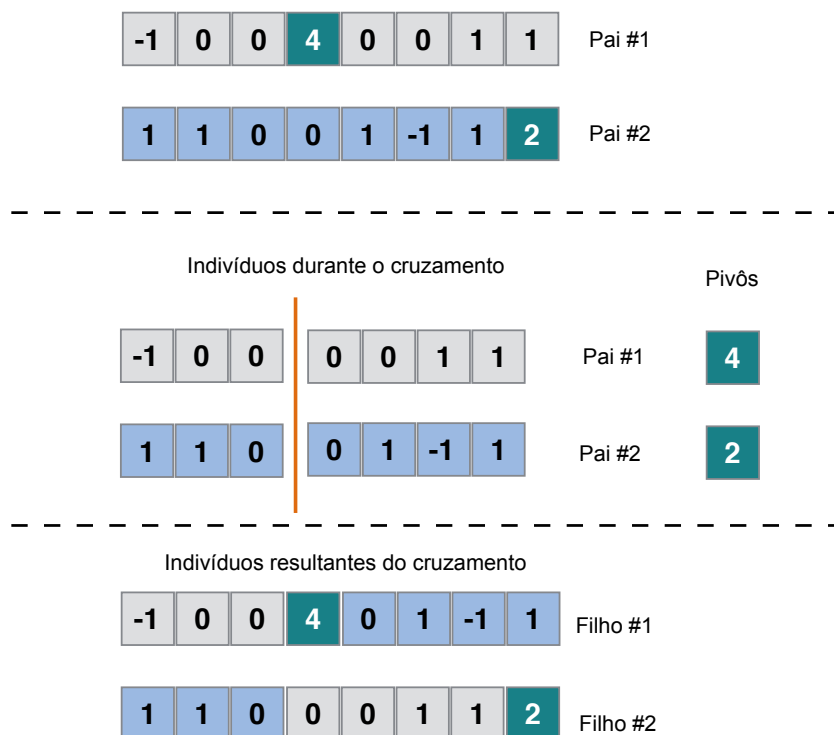


Figura 30: Operador de cruzamento em GSC.

Operador de mutação

O operador de mutação para o GSC também é resultado de uma modificação no operador de mutação uniforme convencional que funciona com base na escolha de um número aleatório de genes para aplicar a mutação. No entanto, para trazer mais diversidade à população, garantiu-se que o gene pivô pertença ao conjunto de genes a serem mutados. Apesar de tal modo trazer benefícios no tocante à diversidade das soluções, foi alcançada mais variabilidade de soluções pela possibilidade de mudança de posição do gene pivô dentro do indivíduo, como mostrado na Figura 31. Nesta figura pode-se notar que não somente os valores dos genes marcados foram alterados mas também a posição do pivô.

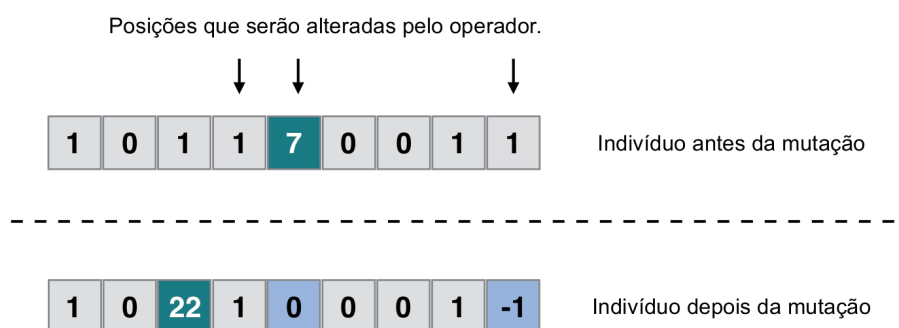


Figura 31: Operador de mutação em GSC.

4.3 Função de aptidão

A ideia principal da função de aptidão para o GSC é atribuir aos *seams* que causam menos degradação, aqueles que estão longe de regiões de interesse, valores baixos de aptidão. Em um cenário ideal de redimensionamento, quanto menos as características dos objetos são modificadas, menor é o impacto visual percebido pelo observador. Por isso, estes *seams* distantes de regiões de interesse são preferíveis, uma vez que eles quando processados minimizam a distorção global.

Um cenário artificial é mostrado na Figura 32, no qual dois *seams* são dispostos em uma imagem. Neste cenário, os objetos relevantes (áreas preenchidas em cinza) são aqueles nos quais qualquer mudança causa um impacto visual que pode ser facilmente percebido. Portanto, neste cenário, remover o *seam* tracejado na Figura 32 ao invés do sólido é preferível porque não há alteração nos objetos relevantes.

Detectar regiões relevantes é uma tarefa bastante importante na função de aptidão. Em GSC,

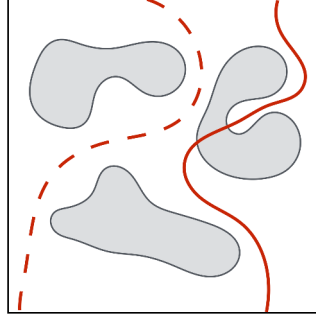
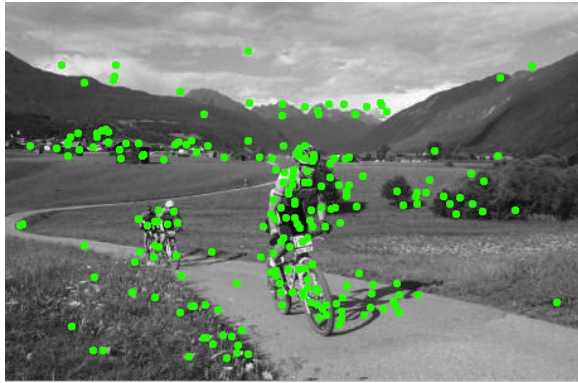


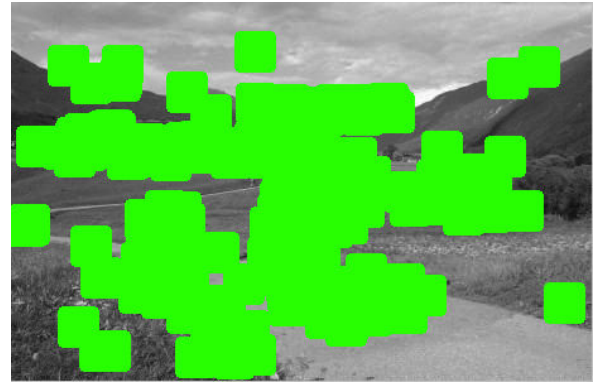
Figura 32: Cenário com objetos relevantes representados pelas áreas preenchidas em cinza.

o detector de pontos-chaves do SURF é usado para dar um indicativo das regiões relevantes. Após a etapa de detecção de pontos-chaves, via janelamento são obtidas as regiões a serem preservadas. Deste modo, a função de aptidão deve atribuir valores aos *seams* a partir da informação de proximidade dos mesmos em relação ao conjunto de pontos-chaves detectados.

Inicialmente, via SURF detecta-se o conjunto de pontos-chaves $\mathbb{P} = \{p_j\}_{j=1}^P$ na imagem original como mostrado na Figura 33(a). Para cada ponto-chave p_j , uma região w_j é obtida através da centralização de uma janela de tamanho W , isto é, $w_j = \{p_i\}_{i=1}^{|w_j|}$, tal que $\|p_i - p_j\| \leq \frac{1}{2}W$, $|\cdot|$ representa a cardinalidade de um conjunto e $\|\cdot\|$ a distância Euclidiana. Essas janelas representam regiões (ver Figura 33(b)) as quais devem ser preservadas ao máximo. Para tal, cada *seam* tem seu valor de de aptidão com base na proximidade destes pontos $\in \mathbb{P}$.



(a) Pontos SURF detectados.



(b) Janelas centradas nos pontos detectados.

Figura 33: Pontos indicativos de regiões relevantes em S-GSC usando SURE.

Em posse do conjunto de pontos-chaves detectados $\mathbb{P}$, para se realizar o cômputo da aptidão de um dado *seam* s^v , deve-se computar primeiramente o vetor de distâncias mínimas D_{s^v} entre o *seam* s^v e os pontos $p_j \in \mathbb{P}$. Dessa forma, o vetor de distâncias mínimas D_{s^v} é dado por

$$D_{s^v} = \{d_j\}_{j=1}^P, d_j = \arg \min_{\|p_j - s^v(i)\|} \left\{ \|p_j - s^v(i)\| \right\}_{i=1}^m, \quad (4.2)$$

em que p_j representa o pixel localizado no centro da j -ésima janela ($\mathbf{w}_j$), P é a quantidade de pontos detectados, isto é, $P = |\mathbb{P}|$ e $\mathbf{s}^v(i) = (i, f(i))$. Com base nisto, considere

$$\lambda(s) = \frac{1}{P} \sum_{j=1}^P \frac{0.5 W - d_j}{0.5 W}, \forall d_j < \frac{W}{2} \in \mathbf{D}_s, \quad (4.3)$$

em que W é o tamanho da janela. Esta função $\lambda(\cdot)$ penaliza com base na proximidade do *seam* em relação ao centro de uma janela. Em suma, quanto mais próximo um *seam* está de um ponto, mais este *seam* será penalizado. Se $\mathbf{D}_{s^v} = \emptyset$, então $\lambda(\cdot) = 0$. Por fim, tem-se a seguinte função de aptidão

$$\text{aptidão}(\mathbf{s}^v) = \begin{cases} \lambda(\mathbf{s}^v), & \text{se } \mathbf{s} - \mathbf{I}' = \emptyset, \\ |\mathbf{s}^v - \mathbf{I}'| + 1, & \text{caso contrário;} \end{cases} \quad (4.4)$$

em que $|\mathbf{X}|$ representa a cardinalidade do conjunto $\mathbf{X}$ e $\mathbf{I}'$ representa a imagem redimensionada.

Um indivíduo pode produzir um *seam* com alguma coordenada fora dos limites da imagem ($\mathbf{s}^v - \mathbf{I}' \neq \emptyset$) durante a etapa de reprodução ou mutação, tal como o *seam* pontilhado da Figura 34. A este tipo de *seam* se atribui um valor alto de aptidão a fim de reduzir a probabilidade de escolha dos mesmos para cruzamento ou mutação em outro momento. De fato, o material genético deste tipo de *seam* é deixado de fora da próxima geração.

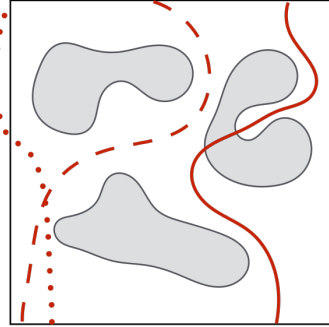


Figura 34: Cenário com um *seam* inválido.

Em geral, as regiões relevantes possuem uma concentração de muitos pontos-chaves. Por causa disto, a maioria dos pontos estão próximos uns dos outros e portanto mais cálculos de distância são necessários. Para garantir uma melhor avaliação de aptidão dos indivíduos, foi aplicada uma técnica de quantização vetorial visando reduzir a quantidade de cálculos.

4.4 Multi-seam carving

A ideia que fundamenta o *multi-seam carving* no GSC reside em utilizar os *seams* mais aptos pertencentes à população final. Assim, são utilizados os indivíduos mais promissores conjuntamente ao invés de usar somente o indivíduo mais adaptado. Entretanto, dentro do conjunto de *seams* promissores S^* podem existir *seams* que se cruzam entre si. Nesta situação, estes *seams* compartilham pelo menos um pixel e, dado que só se pode remover um *pixel* uma única vez, deve-se evitar que este tipo de situação ocorra empregando uma estratégia de descarte de *seams* que se cruzam. Vale a pena ressaltar que após a remoção de um *pixel* compartilhado, aquela coordenada tem sido modificada e não representa mais o *pixel* compartilhado ou pode ser que ela esteja agora indicando uma região inválida.

O algoritmo de descarte de *seams* que se cruzam é simples. Caso os *seams* possuam o mesmo valor de aptidão, são descartados os indivíduos com uma maior quantidade de cruzamentos. Organizados os indivíduos por quantidade de cruzamentos em ordem decrescente, remove-se de S^* aquele com o maior número de cruzamentos. Este processo é repetido até que o número de cruzamentos em S^* seja zero. Um fluxograma do GSC é mostrado na Figura 35.

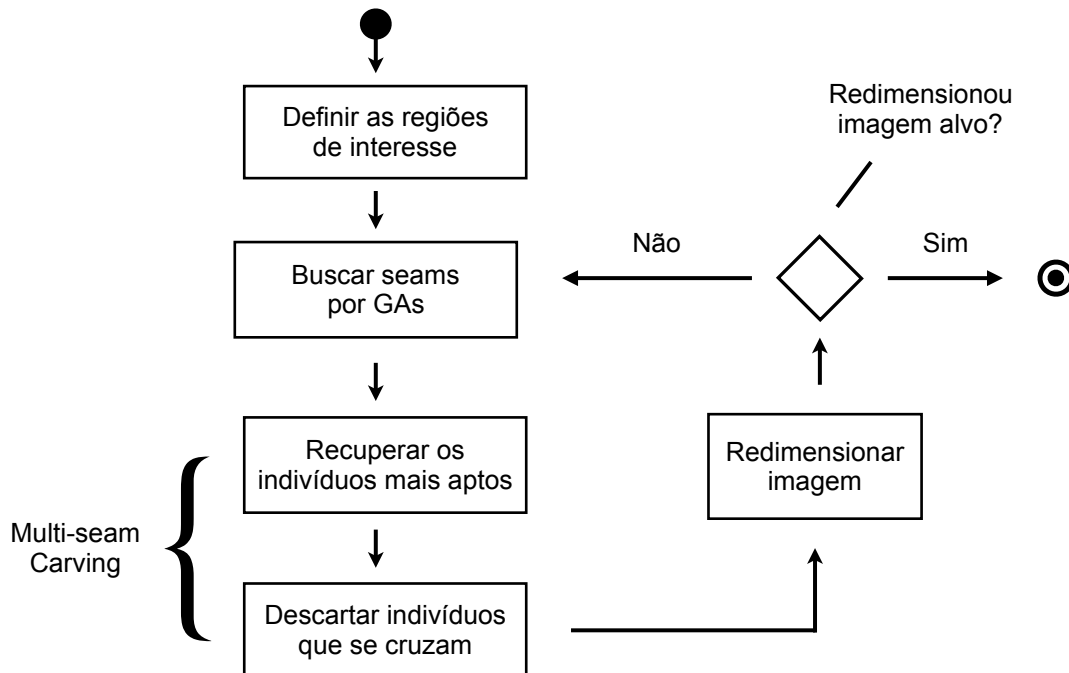


Figura 35: Fluxograma geral do *Genetic Seam Carving*.

4.5 Algoritmo para *Genetic Seam Carving*

O pseudocódigo S-GSC para um exemplo de redução de largura é apresentado a seguir.

GENETIC-SEAM-CARVING(I, m', t_{max})

I : imagem original.
 m' : nova largura.
 t_{max} : número máximo de gerações.

```

1   $I' \leftarrow I$ .
2   $\mathbb{P} \leftarrow \text{DETECTAR-PONTOS-SURF}(I)$ .
3  enquanto LARGURA( $I'$ )  $\leq m'$  :
     $\triangleright$  Algoritmos Genéticos.
4     $Pop[0] \leftarrow \text{INICIAR-POPULAÇÃO}()$ .
5    Computar APTIDÃO( $s^v$ ),  $\forall s^v \in Pop[0]$ .
6    para  $t \leftarrow 1$  até  $t_{max}$  :  $\triangleright$  Condição de parada.
7         $Pais \leftarrow \text{SELECIONAR-PAIS}(Pop[t-1])$ .
8         $Filhos \leftarrow \text{APLICAR-OPERADORES-DE-REPRODUÇÃO}(Pais)$ .
9         $Pop[t] \leftarrow Filhos \cup \text{ELITISMO}(Pop[t-1])$ .
10       Computar APTIDÃO( $s^v, \mathbb{P}$ ),  $\forall s^v \in Pop[t]$ .
     $\triangleright$  Multi-seam a partir daqui.
11     $\mathbb{S} \leftarrow \text{RESGATAR-MAIS-APTOS}(P[t_{max}])$ .
12     $\mathbb{S}^* \leftarrow \text{DESCARTAR-SEAMS-CRUZADOS}(\mathbb{S})$ .
13    para cada seam  $s^v \in \mathbb{S}^*$  :
14         $I' = I' - s^v \triangleright$  Imagem sendo redimensionada.
15        se LARGURA( $I'$ )  $\leq m'$  :
16            retorne  $I'$ .
17 retorne  $I'$ .
```

Como dito, além da proposta do operador redimensionamento, neste trabalho também tem-se a contribuição de uma métrica para avaliar resultados de redimensionamento como as vistas no Capítulo 3.

4.6 SURF-SSIM (proposta)

Inspirado pela função de aptidão, foi proposta também neste trabalho a métrica SURF-SSIM. Nesta métrica, o cômputo da similaridade em imagens redimensionadas é dado pelo escore SSIM (WANG *et al.*, 2004) obtido a partir de correspondências entre pontos-chaves oriundos do detector SURF (BAY; TUYTELAARS; GOOL, 2006). Através do SURF, pontos-chaves são detectados na imagem original (ver Figura 36(a)). Como esperado e, devido as suas características, os pontos detectados são colocados sobre regiões salientes. Tais regiões devem ser preservadas, uma vez que qualquer impacto visual sobre as mesmas pode ser facilmente percebido pelo observador.

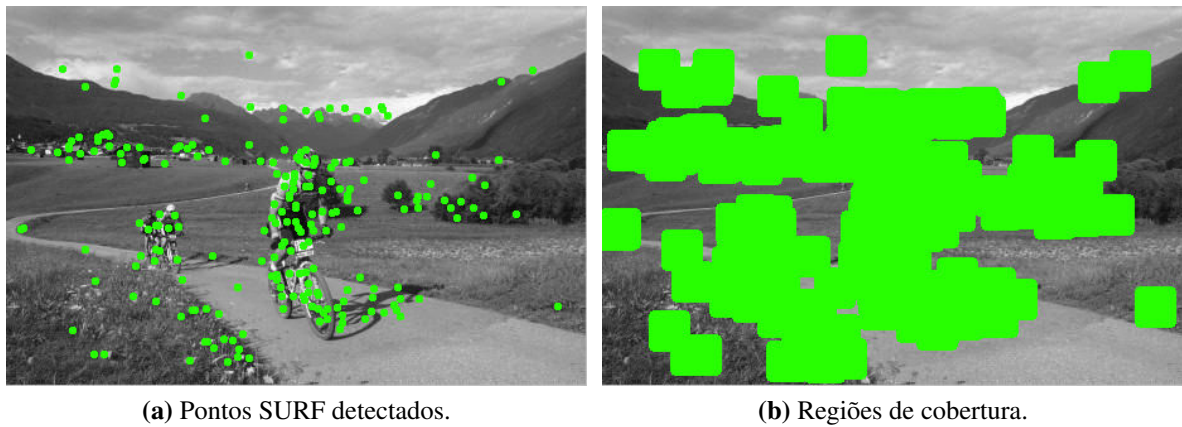


Figura 36: Região de cobertura do SURF-SSIM.

A fim de cobrir a área de interesse, janelas podem ser dispostas centradas nos diversos pontos-chaves obtidos pelo detector SURF, como pode ser visto na Figura 36(b). No entanto, ao se proceder desta maneira, tem-se uma elevada quantidade de janelas sobrepostas. O problema de ter janelas sobrepostas está diretamente ligado ao custo de se avaliar uma mesma região diversas vezes. A fim de melhorar esta condição, o número de pontos neste método é reduzido através da aplicação de um algoritmo de quantização vetorial, tal como o K-médias (AHMAD; DEY, 2007), sobre os pontos detectados.

K-médias é um algoritmo de quantização vetorial que é capaz de obter a partir de um conjunto de dados, vetores representativos chamados de protótipos (ver os pontos duplamente circulos na Figura 37(a)). Depois de aplicar o algoritmo K-médias sobre os pontos, tem-se K protótipos que não são pontos oriundos do detector SURF (ver Figura 37(a)). Para se obter tais pontos, doravante denominados pontos de interesse (*Points of Interest*, PoIs), para cada protótipo encontra-se o ponto real mais próximo (ver novamente os pontos duplamente circulos na Figura 37(b)). Pode-se notar que ao final tem-se um conjunto de pontos oriundos

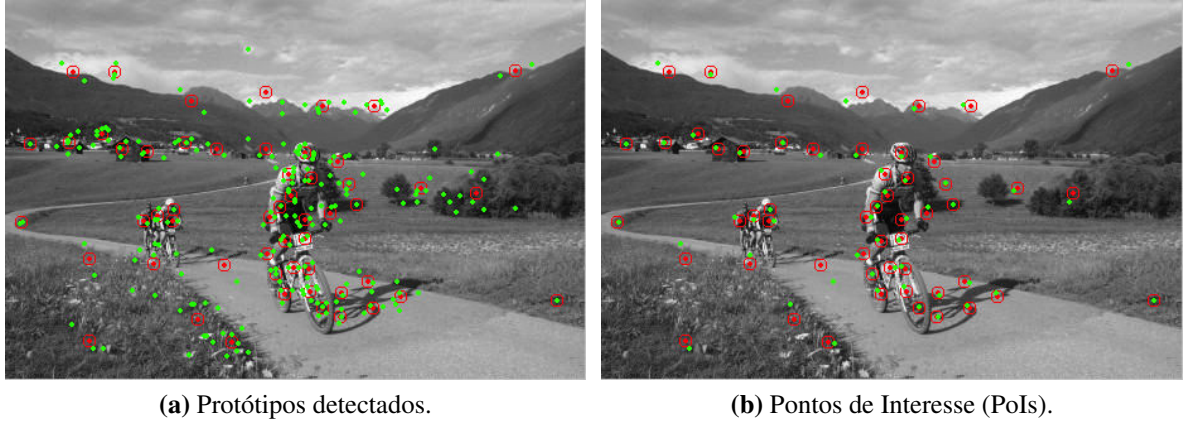


Figura 37: Detecção de pontos de interesse (PoIs).

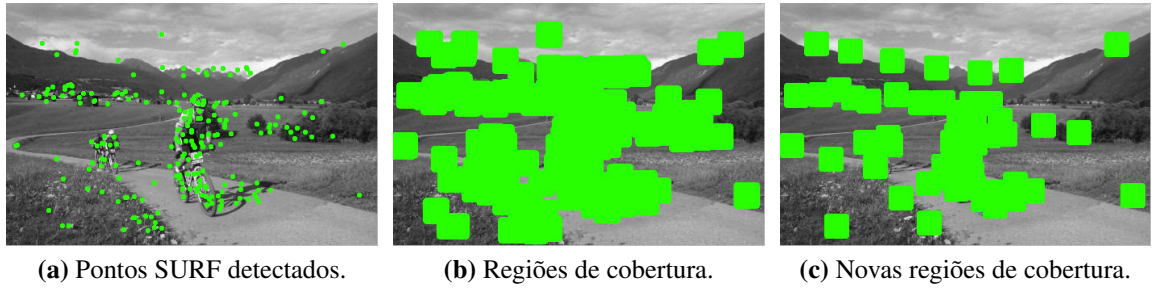


Figura 38: Região de interesse em SURF-SSIM.

do detector SURF e que se apresentam mais próximos dos protótipos. Em seguida, as novas regiões de interesse são obtidas a partir de janelas sobre cada PoI (ver Figura 38(c)).

Detectados os PoIs, é feita a correspondência com os pontos detectados na imagem redimensionada e, consequentemente, a correspondência entre janelas. Feita a correspondência entre janelas, veja Figura 39, o cômputo da similaridade entre janelas é realizado para obter o escore final de SURF-SSIM. Em posse dos escores de similaridades, um ponderamento baseado na relação de correspondências entre as imagens é aplicado como mostrado na Eq. 4.5.

De forma resumida, o SURF-SSIM é dado por:

$$SURF - SSIM(I, I') = \frac{1}{K} \sum_{i=1}^Q SSIM(w_i, w'_i) \text{ com } Q \leq K, \quad (4.5)$$

em que I e I' são as imagens original e redimensionada, respectivamente; w_i e w'_i são os conteúdos da i -ésima janela; K é o número de janelas em I ; e Q é o número de janelas com correspondências entre I e I' . O valor máximo para a $SSIM(.,.)$ é 1, o que significa regiões com maior preservação. No caso de $Q < K$, é alcançada uma condição em que alguns pontos não puderam ter suas correspondências, veja a falta de correspondências ilustrada na Figura 40.

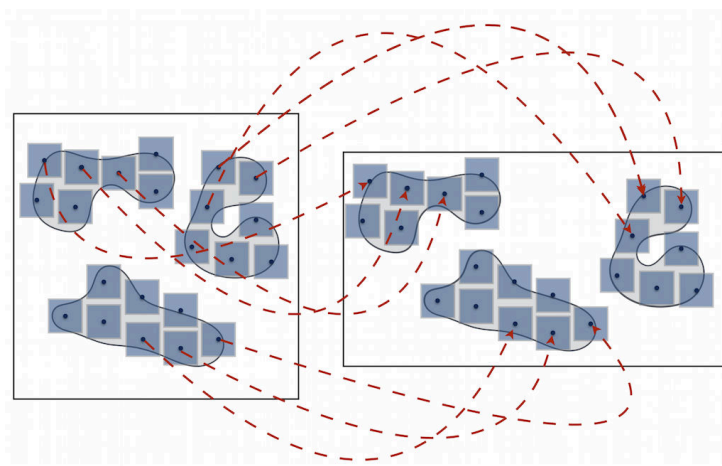


Figura 39: Algumas correspondências entre os pontos de interesse (POIs).

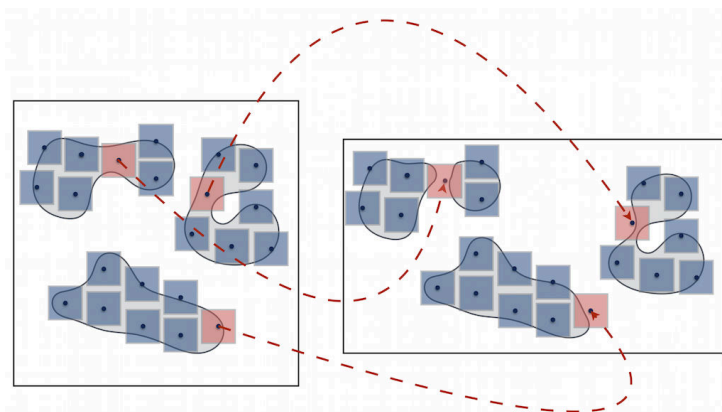


Figura 40: Influência da não correspondência entre pontos de interesse.

Nesta situação indesejada, a distorção nas regiões é forte demais e, assim, deve-se penalizar o valor global de qualidade. Além disso, K é um termo para valores de $\text{SURF-SSIM}(.,.) \in [-1,1]$.

4.7 Conclusão

Como mencionado, o operador *Seam Carving* possui uma componente de busca exaustiva para encontrar os *seams* ótimos. É nesta abordagem exaustiva que reside o maior desafio no tocante ao custo computacional, pois entradas de alta dimensionalidade demandam mais processamento. A fim de mitigar este caráter exaustivo, foi apresentado neste Capítulo uma formulação deste operador com base em Algoritmos Genéticos, denominado *Genetic Seam Carving*.

Para que haja uma modelagem bem sucedida em Algoritmos Genéticos, foram definidos o indivíduo, a função de aptidão, e novos operadores genéticos. Além disto, neste Capítulo, tem-se a proposição de uma métrica de qualidade para resultados de redimensionamento. A

métrica proposta, denominada SURF-SSIM, realiza o cálculo da similaridade em imagens redimensionadas a partir de correspondências entre pontos-chaves oriundos do detector SURF.

SIMULAÇÕES COMPUTACIONAIS

Os resultados obtidos nas simulações computacionais realizadas no transcorrer deste trabalho são apresentados neste capítulo. De início, é apresentada a metodologia aplicada, seguida dos resultados obtidos a partir do S-GSC e outros operadores listados a seguir quando aplicado a duas bases de imagens, a saber: RetargetMe (RUBINSTEIN *et al.*, 2010) e NRID (FANG *et al.*, 2012), as quais são amplamente usadas neste tipo de problema. Adicionalmente, utiliza-se a base IRCCyN/IVC (CALLET; AUTRUSSEAU, 2005) com o intuito de comparar nossa proposta com o SC convencional. Além disso, são apresentadas tanto validações em termos de convergência e aptidão da proposta em Algoritmos Genéticos, quanto em termos de qualidade de imagens utilizando três métricas distintas, a saber: SIFT-flow, IR-SSIM e SURF-SSIM. Ao final do capítulo, são apresentadas algumas considerações acerca do método proposto.

5.1 Metodologia

Para avaliar a nossa proposta foram realizadas simulações computacionais usando imagens de dois conjuntos de imagens específicos para este tipo de problema, a saber RetargetMe e NRID. Um experimento menor foi realizado na base IRCCyN/IVC. Nas simulações envolvendo os conjuntos RetargetMe e NRID, foi realizada uma redução sobre cada imagem em uma única dimensão. Para cada imagem, é definida na especificação do conjunto, o fator de redução que pode ser 25% ou 50%, bem como a dimensão que deve ser reduzida. Já na base IRCCyN/IVC cada imagem foi reduzida em 20% tanto na largura quanto na altura.

O conjunto RetargetMe é composto de 37 imagens. Para cada imagem há um grupo de 8 saídas provenientes da aplicação de diferentes algoritmos de redimensionamento, a saber: Recorte (CR) Multi-operador (MOP) de Rubinstein, Shamir e Avidan (2009), *Seam Carving* (SC) de Avidan e Shamir (2007), *Scaling* (SCL), *Shift-maps* (SM) de Pritch, Kav-Venaki e Peleg (2009), *Scale-and-Stretch* (SNS) de Wang *et al.* (2008), *Streaming Video* (SV) de Krähenbühl *et al.* (2009) e *Nonhomogeneous warping* (WARP) de Wolf, Guttman e Cohen-Or

(2007). Existem também resultados para outros algoritmos de redimensionamento dentro deste conjunto, entretanto, apenas os algoritmos supracitados têm resultados para todas as imagens. O conjunto NRID, é composto de um total de 35 imagens e apresenta resultados para os seguintes algoritmos de redimensionamento: Recorte (CR), *Seam Carving* (SC), *Scaling* (SCL), *Mutli-operador* (MOP), *Shift-maps* (SM) e *Nonhomogeneous warping* (WARP). Já na base IRCCyN/IVC, existe um conjunto de 10 imagens, nas quais foram obtidos resultados usando o *Seam Carving* (SC) na sua formulação convencional.

Além disso, foram selecionadas 3 iterações do S-GSC: a inicial, uma intermediária e a final, para análise da população em 3 momentos distintos (gerações). Para cada momento, foram extraídos da população, os indivíduos mais aptos e os viáveis para ilustrar a colocação dos *seams* durante o processo de evolução. Ressalta-se que para este experimento, foi usada uma imagem que não faz parte do conjunto de imagens avaliados. Adicionalmente, um outro experimento foi realizado utilizando uma imagem em 4 escalas, isto é, a imagem de entrada foi redimensionada de forma proporcional por fator de 1, 1,5, 2 e 2,5 de vezes o seu tamanho, para avaliação do tempo de execução entre o S-GSC e o SC.

Os parâmetros de configuração de GAs escolhidos para o S-GSC foram os mesmos para todos os experimentos realizados. Os parâmetros foram selecionados por tentativa e erro. O tamanho da população foi definido com 120 indivíduos e as taxas de cruzamento e mutação fixadas em 80% e 15%, respectivamente. A fim de completar uma nova população, os 5% restantes foram selecionadas via elitismo a partir da população anterior. No tocante ao critério de parada, o número máximo para todos os experimentos é de 80 gerações.

A respeito da validação em termos de qualidade de imagem foram utilizadas as duas métricas apresentadas no Capítulo 3, a saber, SIFT Flow e IR-SSIM, e a métrica proposta no Capítulo 5, SURF-SSIM. A respeito dos parâmetros usados em SIFT-Flow e SURF-SSIM, todos eles foram fixados nos experimentos e estão apresentados na Tabela 2.

Tabela 2: Conjunto de parâmetros selecionados para cada métrica nos experimentos.

Métrica	Parâmetro	Descrição
SIFT flow	$\eta = 1,275$	Coeficiente de regularização.
	$\alpha = 2 \times 255$	Coeficiente de regularização .
	$d = 40 \times 255$	Limiar de truncamento.
SURF-SSIM	$W = 27$	Tamanho da janela.
	$P = 10\% \text{ de } K$	Número de pontos representativos.

5.2 Resultados

Nas Figuras 41, 42 e 43 são apresentados os resultados de redimensionamento para os conjuntos RetargetMe, NRID e IRCCyN/IVC, respectivamente. Com exceção da primeira linha, cada coluna representa a saída de um algoritmo de redimensionamento de modo a facilitar a comparação visual. Na Figura 41 tem-se que os métodos CR, SM e WARP foram os que mais descartaram elementos visuais em relação aos demais métodos. Além disso, pode-se notar que o resultado do S-GSC foi superior ao do SC, entretanto ambos obtiveram mais distorções que os resultados de MOP, SNS e SV. Estes três últimos métodos aplicaram distorções mais severas em partes menos relevantes, o que acarretou nos melhores resultados, especialmente MOP e SV.

Já na Figura 42, pode-se notar que S-GSC foi o método alcançou resultados mais agradáveis. Os métodos SC, SCL e WARP foram os que mais distorceram e descartaram informação relevante (ver os braços e parte do corpo da mulher que sofreram distorções severas). O método SM, apesar de alcançar resultados satisfatórios na maioria das imagens, não conseguiu produzir um resultado agradável na última imagem. Quando se tem informação relevante bem distribuída pela imagem, o SM não consegue produzir uma boa representante da imagem original. Além disso, destaca-se a qualidade dos resultados produzidos pelo S-GSC sobre os do SC.

Por fim, na Figura 43 pontua-se desde distorções leves, como as das asas na imagem Avion e a garagem na House, até mesmo distorções mais severas como as presentes na face, braços e perna da mesa na imagem Barba e os ombros e moldura na imagem Lena. Como neste experimento somente o SC foi usado para fins comparativos, a comparação pontual entre o método original (SC) e a sua variação (S-GSC) é mais evidente. A qualidade dos resultados produzidos pelo S-GSC sobre os do SC continua atestando a superioridade do S-GSC. Isto se dá pelo fato de SC usar uma função baseada na magnitude do gradiente e a função de energia do S-GSC é baseada em pontos-chaves oriundos do detector SURF. De fato, a função de energia adotada em S-GSC se mostrou mais adequada pelos resultados analisados.

É notável que o método S-GSC produziu resultados competitivos em relação aos demais métodos, com poucas distorções. Além disso, a preservação de conteúdo relevante também se destaca visto que, na grande maioria dos resultados as saídas do SC sofreram distorções severas, enquanto que as produzidas pelo S-GSC ficaram visualmente mais agradáveis.



Figura 41: Resultados de redimensionamento no conjunto RetargetMe.

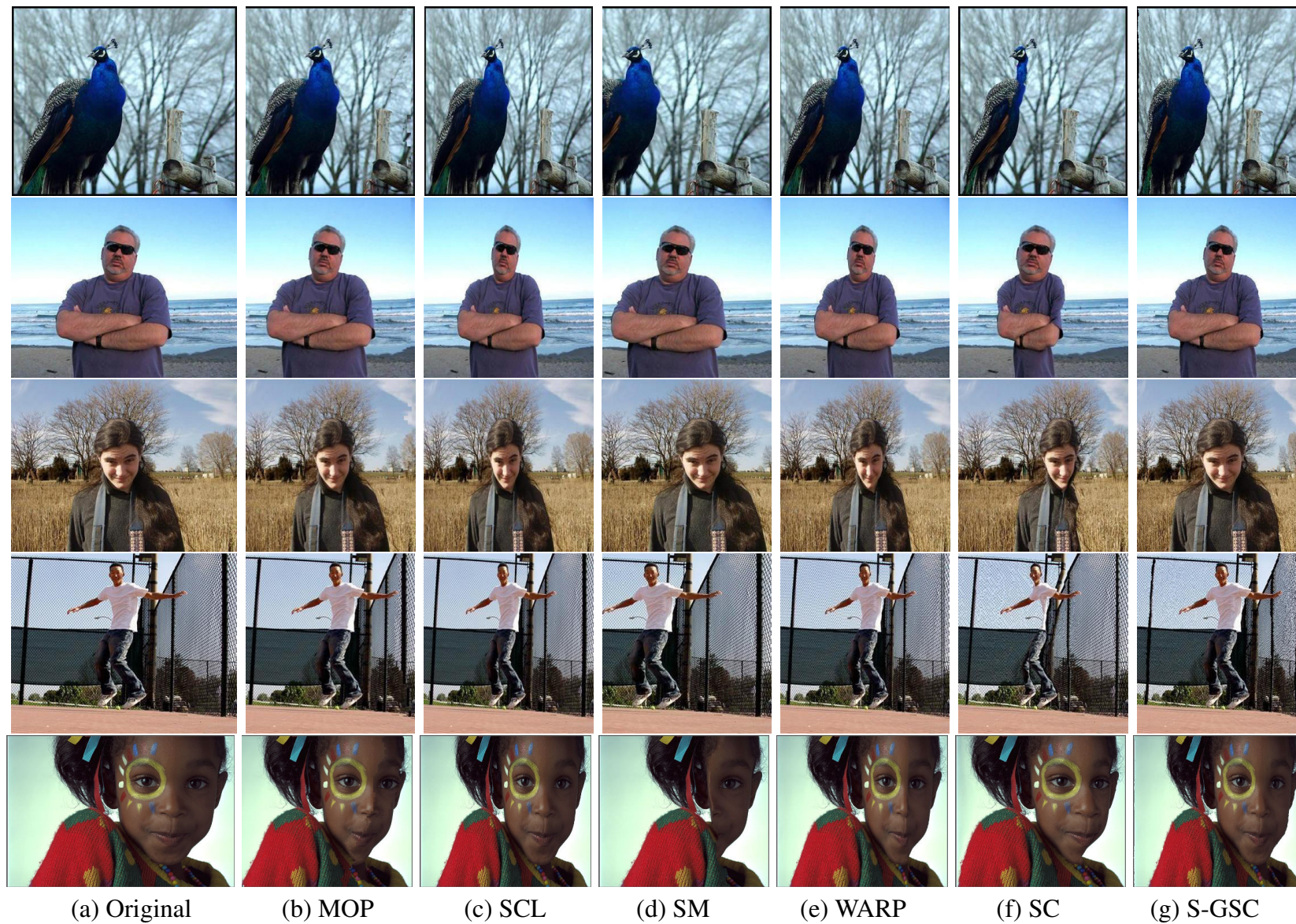
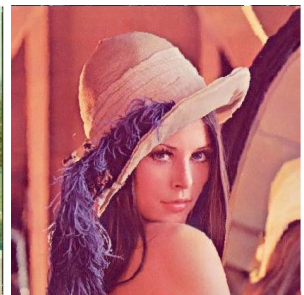
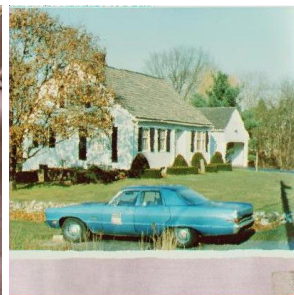
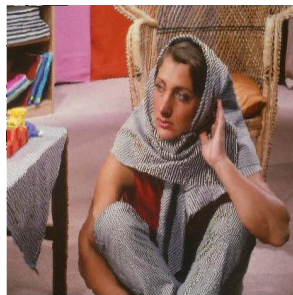


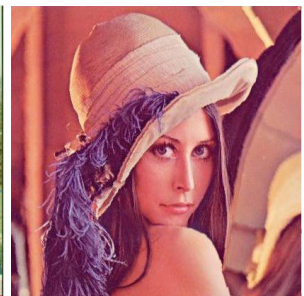
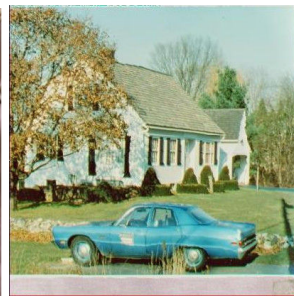
Figura 42: Resultados de redimensionamento no conjunto NRID.



Imagens originais



Resultado do S-GSC



Resultado do SC

(a) Avion

(b) Barbara

(c) House

(d) Lena

Figura 43: Resultados de redimensionamento no conjunto IRCCyN/IVC. Os resultados foram ampliados para fins comparativos.

5.3 Validação

A validação da abordagem proposta foi realizada de duas maneiras: pela validação em termos de GAs e pela abordagem em termos de qualidade de imagem, ou seja, as métricas de redimensionamento. Na validação da abordagem genética, os resultados são avaliados em termos de convergência do algoritmo e valores de aptidão. Enquanto na validação em termos qualitativos de imagem, são apresentados resultados em termos de escores obtidos pelas

métricas de qualidade de imagem para os resultados de redimensionamento.

5.3.1 Validação da abordagem genética

Para avaliar a abordagem genética do método proposto, analisamos a convergência e o comportamento dos valores da função de aptidão da população no decorrer do processo de evolução. Assim, são apresentados na Figura 44 os valores para o pior e melhor indivíduos, bem como o valor para a aptidão média de todos os indivíduos em uma determinada geração para a última execução do laço mais externo apresentado no Algoritmo S-GSC. Devido a variabilidade que os operadores genéticos fornecem ao nosso modelo, a população teve uma ligeira oscilação da aptidão média durante o processo. No entanto, o melhor valor de aptidão foi alcançada logo na primeira geração e mantido via elitismo.

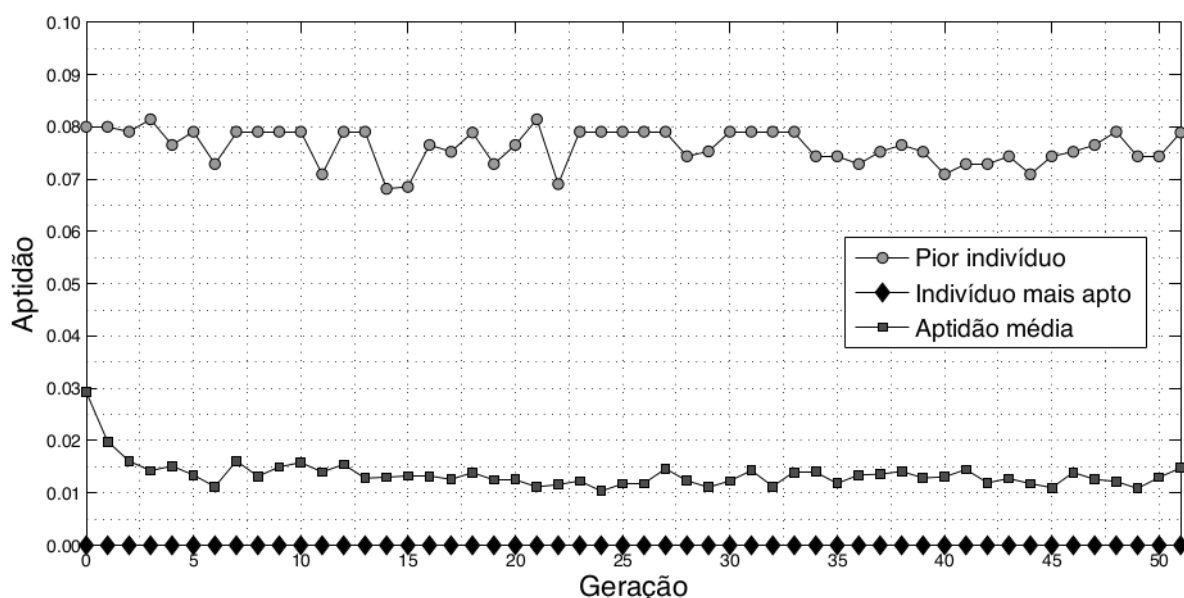


Figura 44: Valores de aptidão dos piores e melhores indivíduos, bem como o valor médio da população para cada geração da última iteração do S-GSC.

Além disso, o número de indivíduos mais aptos e o de *seams* removidos em cada execução do S-GSC é apresentado na Figura 45. S-GSC encontra com sucesso uma boa quantidade de indivíduos para remoção em cada execução. Ressalta-se novamente que os indivíduos mais aptos são aqueles usados na estratégia *multi-seam* em cada iteração do S-GSC.

Como dito anteriormente, foram selecionadas 3 iterações do Algoritmo S-GSC: a inicial, uma intermediária e a final, para análise da população em 3 momentos distintos (gerações). Para cada momento, foram extraídos da população, os indivíduos mais aptos e os viáveis para ilustrar como eles, em forma de *seams*, estão posicionados na imagem. Esta análise tem o

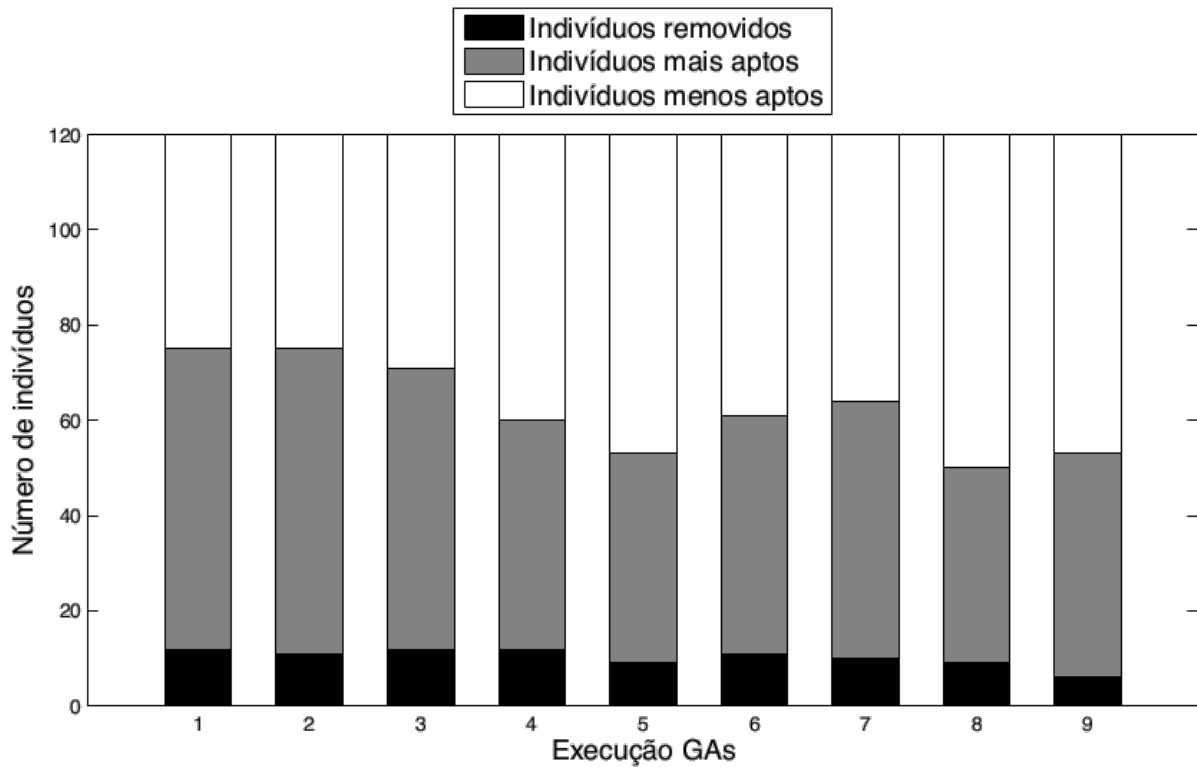
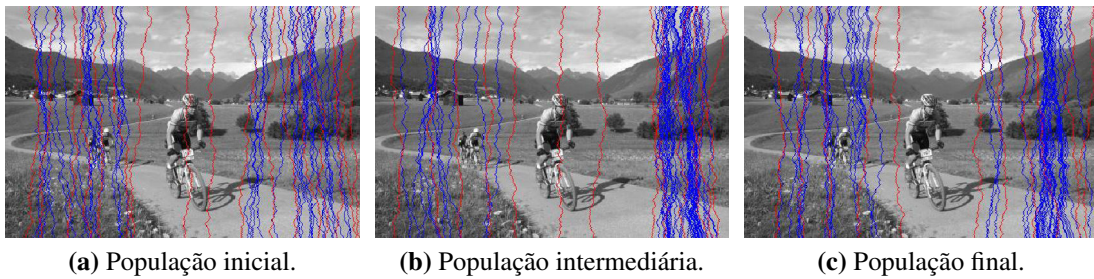


Figura 45: Número de indivíduos mais aptos e o de removidos em cada execução dos GAs.

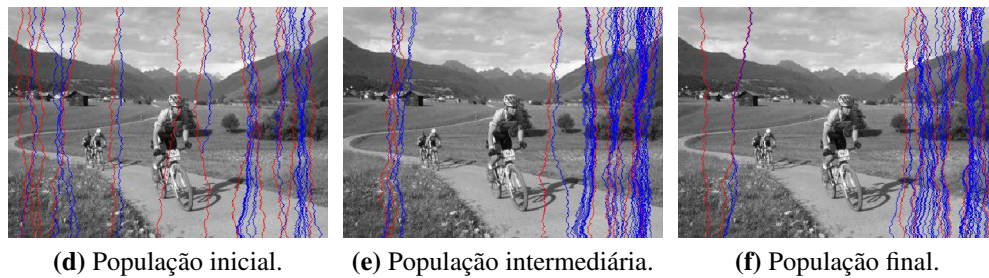
intuito de mostrar como está a colocação dos *seams* durante o processo de evolução. Ressalta-se que para este experimento, foi usada uma imagem que não faz parte do conjunto de imagens avaliados. Os resultados podem ser vistos em Figura 46. Ao se verificar as imagens contidas na Figura 46, percebe-se o agrupamento de *seams* em certas regiões de baixa importância ao passo que se evita o cruzamentos com regiões de mais alta importância.

À primeira vista, parece não tão relevante combinar indivíduos com diferentes pivôs, devido a informação genética de cada um tentar encontrar um *seam* ideal a partir de diferentes regiões na imagem. Além disso, pode parecer muito improvável que tal combinação melhore a solução de uma geração para outra, mas ela o faz. Em uma primeira tentativa, foi projetado um indivíduo com pivô fixo, mas como o indivíduos são gerados aleatoriamente, os pivôs podem ser colocados em regiões de interesse e como não havia possibilidade de mover os pivôs, os mesmos permaneciam inalterados. Nesta condição, indivíduos que possuíam pivôs em regiões de interesse não dispunham de nenhum mecanismo capaz de modificar a coordenada do pivô.

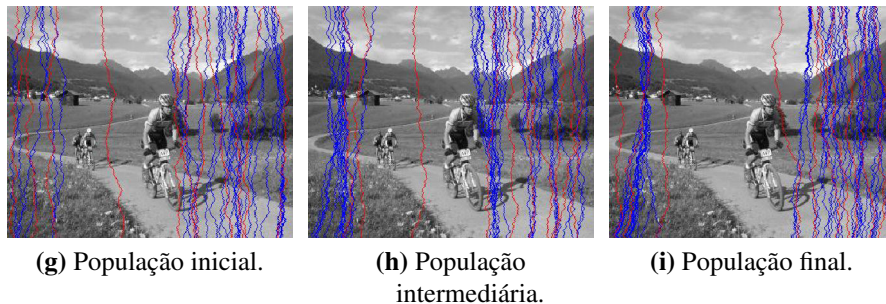
Por um lado, a abordagem que se baseia em pivô fixo é boa em termos de simplicidade de modelagem, mas, por outro, surge uma desvantagem que limita os pivôs a estarem na mesma linha (para um *seam* vertical) ou coluna (para um *seam* horizontal) da imagem. Assim, a falta de flexibilidade para mover o pivô para qualquer posição na imagem durante a evolução faz-se necessária para atingir melhores resultados. Além disso, a abordagem que se baseia em pivô fixo



Primeira execução com a imagem de tamanho 460×300 .



5ª execução com a imagem de tamanho 418×300 .



10ª execução com a imagem de tamanho 374×300 .

Figura 46: *Seams* azuis representam os mais aptos e os em vermelho, os processados.

alcança um *seam* por iteração por execução, tornando-se quase impossível o uso em aplicações em tempo real.

Devido à variabilidade que os operadores genéticos fornecem ao nosso modelo, é mostrado na Figura 44 que a população tem uma ligeira oscilação da aptidão média durante o processo de evolução. Optou-se, entretanto, por continuar o processo de evolução até que o monitor de estagnação¹ perceba que a aptidão média não tem nenhuma mudança significativa (máximo local) durante um certo número de iterações. Tal abordagem é relevante para que melhores *seams* sejam obtidos. Apesar do grande número de indivíduos que estão em conformidade

¹Monitor de estagnação é uma ferramenta dentro do projeto de GAs que ao verificar um estado de estagnação, convergência em um máximo local, encerra a execução de GAs.

com as restrições não serem utilizados como solução, nosso modelo atinge com sucesso uma quantidade razoável de indivíduos mais próximos do ótimo em cada iteração.

5.3.2 Análise de qualidade das imagens redimensionadas

Ao avaliar a Figura 47, percebe-se que o método S-GSC se comporta de forma semelhante em relação aos escores obtidos quando comparado com os demais métodos. Especificamente para o escore SIFT Flow, S-GSC obteve um desempenho superior. Outro fato que pode-se notar é que alguns escores máximos para o IR-SSIM de outros algoritmos de redimensionamento são valores atípicos (*outliers*), isto é, valores diferentes dos demais valores do conjunto. Ao se fazer uma análise pontual, nota-se que o método S-GSC se encontra entre os 3 melhores. No tocante aos escores SURF-SSIM, o método S-GSC alcançou resultados razoáveis e ficou entre os 4 melhores.

Já na Figura 48, o método S-GSC mais uma vez obteve um desempenho superior quando comparado aos demais métodos para o escore SIFT Flow. Além disso, o desvio padrão do método S-GSC para este escore foi o menor dentre todos os métodos. Para o escore IR-SSIM, pode-se também notar uma certa similaridade entre os valores de máximos e médios do S-GSC para o IR-SSIM, entretanto, o valor de desvio padrão foi o menor. Além disso, o valor de mínimo reforça que até o pior resultado de redimensionamento é, em termos de qualidade, melhor que os resultados obtidos para os demais métodos. Para o escore SURF-SSIM, S-GSC continuou alcançando resultados satisfatórios.

Na Tabela 3, são apresentados os valores dos escores SIFT Flow, IR-SSIM e SURF-SSIM por método nas bases RetargetMe e NRID. Ressalta-se mais uma vez que os escores nem sempre concordam entre si porque as métricas buscam, à sua maneira, modos distintos de inferir o quão bom o resultado do redimensionamento foi para sua imagem original. É fácil notar que S-GSC alcançou resultados satisfatórios para os conjuntos avaliados, confirmando sua eficiência e estabilidade. Além disso, na maior parte dos resultados o método S-GSC obteve o menor desvio padrão. Outro fato interessante, que aconteceu no conjunto RetargetMe, foi que, até nos mais baixos escores alcançados pelo S-GSC, ainda assim o S-GSC se manteve superior aos demais.

Apesar do método MOP ter conseguido bons resultados, o que corrobora com a ideia de que combinar diversos operadores impacta positivamente no resultado final, uma vez que o método S-GSC possui uma componente estocástica, o mesmo alcança resultados competitivos. Esperava-se que, com a preservação de pontos SURF, os resultados de S-GSC na métrica SURF-SSIM fossem tendenciosos já que ambos possuem parte da formulação idêntica.

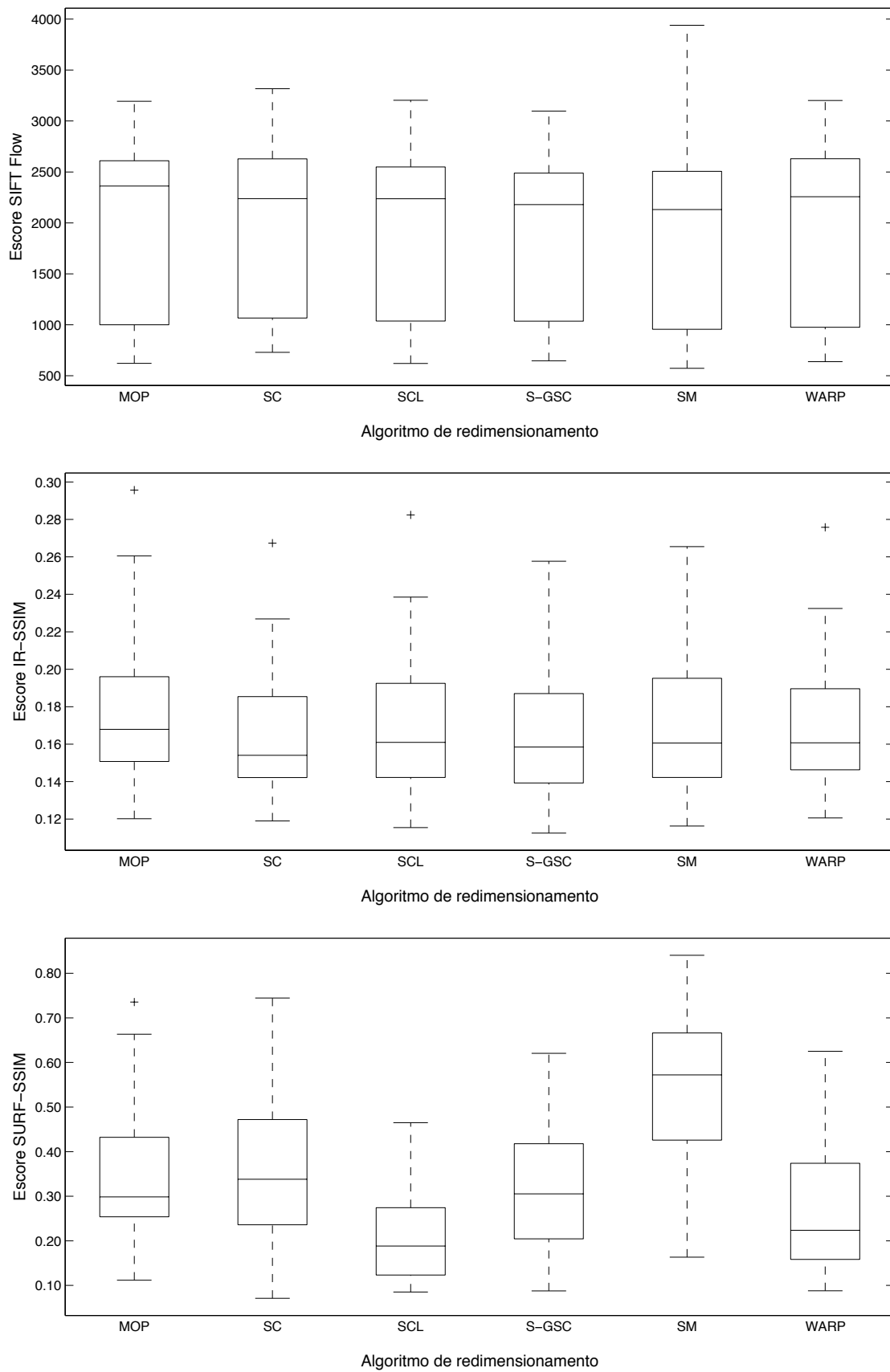


Figura 47: Diagramas de caixa do desempenho dos métodos no conjunto NRID.

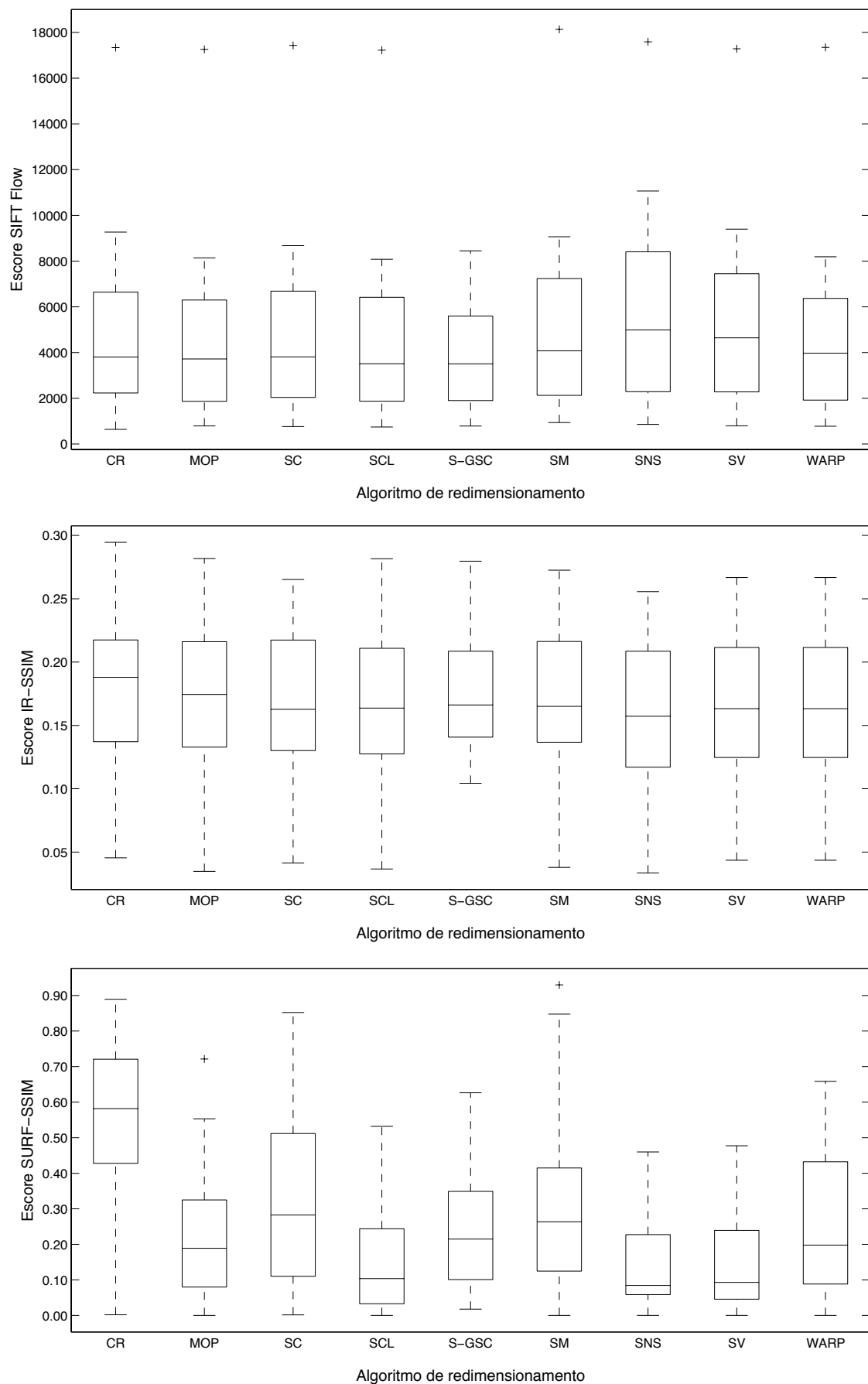


Figura 48: Diagramas de caixa do desempenho dos métodos no conjunto RetargetMe.

Tabela 3: Desempenho dos algoritmos em função dos escores obtidos a partir das métricas SIFT Flow, IR-SSIM e SURF-SSIM, Os símbolos μ e σ se referem aos valores melhor, médio, menor e o desvio padrão dos escores, Os valores sublinhados representam os melhores escores em termos de desempenho.

Base	Métrica		CR	MOP	SC	SCL	S-GSC	SM	SNS	SV	WARP
NRID	SIFT Flow	μ	×	1984,18	2024,61	1949,30	1908,99	1930,10	×	×	1951,65
		σ	×	827,47	833,38	773,52	766,77	910,65	×	×	803,48
	IR-SSIM	μ	×	0,1760	0,1636	0,1664	0,1638	0,1697	×	×	0,1685
		σ	×	0,0383	0,0336	0,0363	0,0332	0,0373	×	×	0,0346
	SURF-SSIM	μ	×	0,3596	0,3718	0,2006	0,3116	0,5416	×	×	0,2675
		σ	×	0,1567	0,1953	0,0934	0,1286	0,1870	×	×	0,1417
RetargetMe	SIFT-Flow	μ	4549,37	4265,82	4483,02	4249,75	3858,45	4725,71	5258,23	4889,42	4331,81
		σ	3329,85	3134,82	3257,13	3134,49	2260,70	3400,63	3569,18	3344,37	3141,84
	IR-SSIM	μ	0,1779	0,1761	0,1715	0,1695	0,1772	0,1682	0,1626	0,1690	0,1690
		σ	0,0568	0,0567	0,0547	0,0555	0,0460	0,0585	0,0600	0,0569	0,0569
	SURF-SSIM	μ	0,5628	0,2227	0,3307	0,1470	0,2353	0,3053	0,1412	0,1459	0,2526
		σ	0,2142	0,1774	0,2473	0,1334	0,1683	0,2323	0,1164	0,1277	0,1985
IVC	SIFT Flow	μ	×	×	3781,14	×	3723,02	×	×	×	×
		σ	×	×	2577,11	×	2755,14	×	×	×	×
	IR-SSIM	μ	×	×	0,1599	×	0,1571	×	×	×	×
		σ	×	×	0,0355	×	0,0366	×	×	×	×
	SURF-SSIM	μ	×	×	0,2813	×	0,3670	×	×	×	×
		σ	×	×	0,1036	×	0,0366	×	×	×	×

Entretanto, como visto na Tabela 3, outros métodos de fato obtiveram um melhor desempenho no que se refere a preservação de pontos salientes.

Outra característica interessante do S-GSC é que por não se tratar de uma abordagem determinística, a imagem redimensionada resultante alcançada em diferentes execuções não é a mesma. Entretanto, como visto nos experimentos em Oliveira, Bezerra e Rocha Neto (2015), verificou-se que o desvio padrão para a qualidade de 10 imagens é algo em torno 1×10^{-3} , indicando que as imagens são similares.

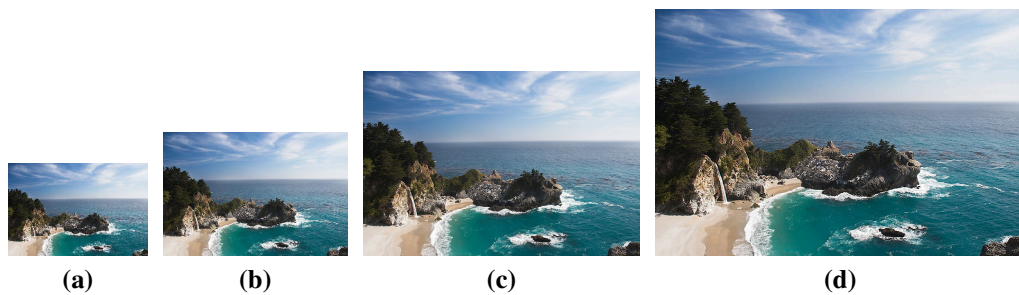


Figura 49: Imagem usada no experimento de medição de tempo em diferentes escalas: (a) $1 \times$, (b) $1,5 \times$, (c) $2 \times$ e (d) $2,5 \times$ o tamanho original.

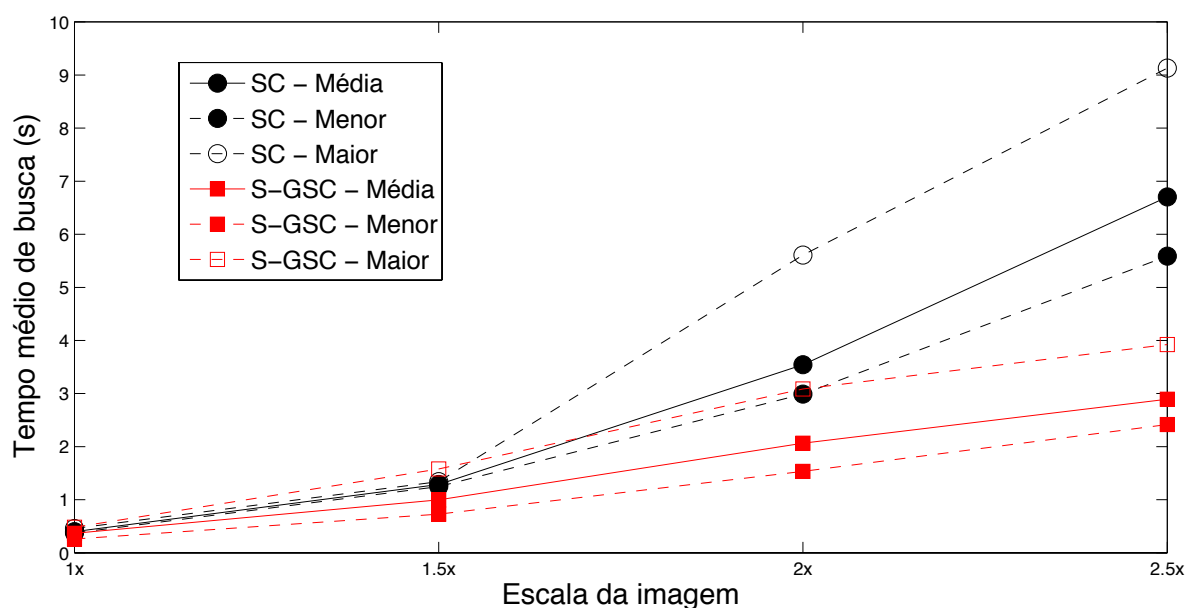


Figura 50: Gráfico de tempo de execução em função da escala da imagem. Adaptado de Oliveira, Rocha Neto e Bezerra (2016).

Acerca dos custos computacionais, um experimento foi realizado utilizando uma imagem em 4 escalas, isto é, a imagem de entrada foi redimensionada uniformemente por fator de 1, 1,5, 2 e 2,5 de vezes o seu tamanho, como mostrado na Figura 49. Depois disto, realizou-se uma redução de 20% da imagem (largura e altura) de cada escala usando os algoritmos SC e S-GSC em 10 execuções.

Neste experimento, nota-se que S-GSC demanda mais processamento e tempo de execução em imagens pequenas (OLIVEIRA; ROCHA NETO; BEZERRA, 2016). No entanto, S-GSC consegue melhor desempenho em termos de custo computacional à medida que o tamanho da imagem aumenta. Isto se dá pelo fato de que a busca realizada em GAs é restringida por parâmetros como o número de gerações ou um monitor de estagnação em alguns casos que não dependem do tamanho da imagem. A Figura 50 mostra a relação entre o tamanho da imagem (via sua escala) pelo tempo de execução referente à busca de *seams*.

Apesar da modelagem do problema via AG ser parte fundamental do modelo, a escolha da função de energia (assim como a de aptidão) com base em pontos salientes é o fator de maior influência no desempenho e qualidade alcançados. Como em SC a função de energia é definida com base na magnitude do gradiente, ao passo que em S-GSC tem-se a função de energia com base em pontos SURF, é perceptível que os pontos SURF definiram regiões de interesse mais apropriadas do que as regiões obtidas pela magnitude do gradiente. Tal fato pode ser verificado pela qualidade dos resultados em que se percebe facilmente a preservação de conteúdo e uma diminuição das distorções.

CONCLUSÕES E TRABALHOS FUTUROS

Neste trabalho, uma nova formulação para o problema de redimensionamento sensível ao conteúdo foi proposta com base em Algoritmos Genéticos. A proposta inclui uma nova representação genética, novos operadores genéticos e uma função de aptidão. Nossa proposta, denominada de *SURF-Based Genetic Seam Carving* (S-GSC), leva em consideração pontos-chaves do operador SURF para obter os caminhos de pixels (*seams*) que vão de uma extremidade à sua oposta na imagem, bem como valores de aptidão com base na proximidade a estes pontos. Tais pontos são colocados em importantes regiões de interesse, nas quais qualquer impacto visual pode ser facilmente percebido pelo observador. Assim, supomos que, processar *seams* que se encontram distantes destes representantes, resulta em uma boa imagem redimensionada.

Além disso, foi proposto uma métrica de qualidade que avalia imagens redimensionadas que se baseia, principalmente, no cômputo de similaridade estrutural da região ao redor de pontos-chaves que possuem correspondência. Na métrica proposta, similar ao que ocorre no S-GSC, um algoritmo de quantização vetorial também é aplicado com o objetivo de reduzir o cômputo de similaridade sobre regiões já avaliadas.

Para avaliar a nossa proposta, foram realizadas simulações computacionais usando imagens de dois únicos conjuntos de imagens específicos para este tipo de problema, a saber RetargetMe e NRID, e um experimento menor foi realizado na base IRCCyN/IVC. Os resultados experimentais indicam que o método S-GSC provê resultados competitivos quando comparado aos outros métodos. Isto foi confirmado pelos escores obtidos em três métricas. Ademais, uma presente limitação do S-GSC é que o mesmo nem sempre encontra o *seam* ótimo como no *Seam Carving*. Entretanto, *seams* subótimos ainda são encontrados, os quais se apresentam adequados como solução do problema.

Acerca dos custos computacionais, S-GSC demanda mais processamento e tempo de execução em imagens pequenas, no entanto, à medida que o tamanho da imagem aumenta, S-GSC consegue melhores resultados. Isto se dá pelo fato de que a busca realizada em Algoritmos Genéticos é restringida por parâmetros como o número de gerações ou um monitor

de estagnação em alguns casos.

6.1 Contribuições deste Trabalho

As contribuições apresentadas nesta dissertação estão relacionadas aos temas: redimensionamento sensível ao conteúdo em imagens, métricas de avaliação para resultados de redimensionamento e Algoritmos Genéticos. Tanto o Capítulo 2 quanto o Capítulo 3 fornecem ao leitor uma compreensão sobre o problema do redimensionamento sensível ao conteúdo em imagens e as particularidades a respeito da avaliação de métodos que se propõem a resolvê-lo. No Capítulo 4 temos a apresentação mais detalhada sobre Algoritmos Genéticos, encerrando as técnicas que compõem este trabalho.

No Capítulo 5 é apresentado a proposta, o *Genetic Seam Carving* (GSC), e sua variação, o *SURF-based Genetic Seam Carving* (S-GSC), que é utilizada para comparação com outras abordagens. De modo geral, o GSC busca caminhos de pixels que atravessam a imagem em regiões de baixo interesse e que uma vez que estes caminhos, ou *seams*, são processados, o impacto percebido pelo observador é minimizado quando comparado com outras técnicas. Além disto, neste mesmo capítulo, são propostos uma nova representação genética, novos operadores genéticos e uma nova função baseada em proximidade de pontos-chaves.

Em relação ao desempenho, S-GSC foi avaliado em função de três métricas e os escores obtidos foram apresentados no Capítulo 6, evidenciando que a proposta dessa dissertação está de acordo com os objetivos inicialmente definidos.

Durante o decorrer deste trabalho, quatro artigos ligados ao tema foram aceitos. Dois deles em congressos internacionais (OLIVEIRA; BEZERRA; ROCHA NETO, 2015; OLIVEIRA; ROCHA NETO, 2015) e os outros dois em periódicos (OLIVEIRA; ROCHA NETO; BEZERRA, 2016; OLIVEIRA; ROCHA NETO, 2016). Os primeiros trabalhos de conferência tratam da proposta do *Genetic Seam Carving* na sua formulação original, enquanto que os de periódico tratam de versões *multi-seam carving* com e sem a função de energia com base em pontos-chaves.

6.2 Trabalhos Futuros

Como trabalhos futuros, pode-se sugerir uma nova função de aptidão que internamente leve em consideração características do tipo *bottom-up* e *top-down*, como proposto na métrica IR-SSIM ou até mesmo outros descritores de imagens. Além disso, uma versão multiobjetivo

que seja capaz de gerar soluções mais esparsas, isto é, sem cruzamento entre si e distribuídas na imagem, substituindo a etapa de análise dos cruzamentos de *seams* na abordagem *multi-seam*.

Outra sugestão consiste na inicialização da população baseada em alguma heurística. É conhecido que os custos computacionais em Algoritmos Genéticos podem ser reduzidos quando se adota uma população bem diversificada. Adicionalmente, em relação aos parâmetros de quantização e o tamanho de janela, alcançar uma maneira automática de ajuste também seria interessante.

Por fim, pretende-se explorar abordagens diferentes, a fim de elevar a qualidade dos indivíduos em termos de representação genética e os operadores genéticos no que compete a obter resultados melhores.

Referências

- ACHANTA, R.; SUSSTRUNK, S. Saliency detection for content-aware image resizing. In: IEEE. **Image Processing (ICIP), 2009 16th IEEE International Conference on**. Cairo, Egito, 2009. p. 1005–1008.
- AHMAD, A.; DEY, L. A k-mean clustering algorithm for mixed numeric and categorical data. **Data & Knowledge Engineering**, Elsevier, v. 63, n. 2, p. 503–527, 2007.
- AMRUTHA, I.; SHYLAJA, S.; NATARAJAN, S.; MURTHY, K. A smart automatic thumbnail cropping based on attention driven regions of interest extraction. In: ACM. **Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human**. Seoul, Coreia do Sul, 2009. p. 957–962.
- AVIDAN, S.; SHAMIR, A. Seam carving for content-aware image resizing. In: ACM. **ACM Transactions on graphics (TOG)**. San Diego, 2007. v. 26, n. 3, p. 10.
- AZAZI, A.; LUTFI, S. L.; VENKAT, I.; FERNÁNDEZ-MARTÍNEZ, F. Towards a robust affect recognition: Automatic facial expression recognition in 3d faces. **Expert Systems with Applications**, v. 42, n. 6, p. 3056 – 3066, 2015. ISSN 0957-4174. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417414006666>>.
- BARNES, C.; SHECHTMAN, E.; FINKELSTEIN, A.; GOLDMAN, D. Patchmatch: A randomized correspondence algorithm for structural image editing. **ACM Transactions on Graphics-TOG**, v. 28, n. 3, p. 24, 2009.
- BAY, H.; TUYTELAARS, T.; GOOL, L. V. Surf: Speeded up robust features. In: **Computer Vision–ECCV 2006**. Graz, Áustria: Springer, 2006. p. 404–417.
- BRUIJN, O. de; SPENCE, R. Rapid serial visual presentation: a space-time trade-off in information presentation. In: ACM. **Proceedings of the working conference on Advanced visual interfaces**. Palermo, Itália, 2000. p. 189–192.
- CALLET, P. L.; AUTRUSSEAU, F. **Subjective quality assessment IRCCyN/IVC database**. 2005. [Http://www.irccyn.ec-nantes.fr/ivcdb/](http://www.irccyn.ec-nantes.fr/ivcdb/). Acesso em: 27-Março-2016.
- CARDOSO, F. A. **Algoritmos Genéticos para Interpolação Não-Linear de Imagem e Decodificação de Códigos Lineares**. Mestrado em Engenharia Elétrica — Universidade Estadual de Campinas, 1998.
- CASTILLO, S.; JUDD, T.; GUTIERREZ, D. Using eye-tracking to assess different image retargeting methods. In: ACM. **Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization**. Anaheim, Estados Unidos, 2011. p. 7–14.

- CHAMARET, C.; MEUR, O. L.; GUILLOT, P.; CHEVET, J.-C. How to measure the relevance of a retargeting approach? In: **Trends and Topics in Computer Vision**. Creta, Grécia: Springer, 2012. p. 156–168.
- CHO, T. S.; BUTMAN, M.; AVIDAN, S.; FREEMAN, W. T. The patch transform and its applications to image editing. In: IEEE. **Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on**. Anchorage, Estados Unidos, 2008. p. 1–8.
- CONGER, D. D.; KUMAR, M.; RADHA, H. Multi-seam carving via seamlets. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **IS&T/SPIE Electronic Imaging**. San Francisco, Estados Unidos, 2011. p. 78700H–78700H.
- CONGER, D. D.; RADHA, H.; KUMAR, M. Seamlets: content-aware nonlinear wavelet transform. In: IEEE. **Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on**. Dallas, Estados Unidos, 2010. p. 1450–1453.
- DANIEL, E.; ANITHA, J. Optimum green plane masking for the contrast enhancement of retinal images using enhanced genetic algorithm. **Optik - International Journal for Light and Electron Optics**, v. 126, n. 18, p. 1726 – 1730, 2015. ISSN 0030-4026. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0030402615003277>>.
- DOMINGUES, D.; ALAHLI, A.; VANDERGHEYNST, P. Stream carving: an adaptive seam carving algorithm. In: IEEE. **Image Processing (ICIP), 2010 17th IEEE International Conference on**. Hong Kong, China, 2010. p. 901–904.
- FANG, Y.; CHEN, Z.; LIN, W.; LIN, C.-W. Saliency detection in the compressed domain for adaptive image retargeting. **Image Processing, IEEE Transactions on**, IEEE, v. 21, n. 9, p. 3888–3901, 2012.
- FANG, Y.; ZENG, K.; WANG, Z.; LIN, W.; FANG, Z.; LIN, C.-W. Objective quality assessment for image retargeting based on structural similarity. **IEEE J. Emerg. Sel. Topics Circuits Syst.**, v. 4, n. 1, p. 95–105, 2014.
- FARIA, A. Fluxo ótico. **Universidade Federal de Minas Gerais, ICEx-DCC-Visão Computacional**, 1992.
- FOGEL, L.; OWENS, A.; WALSH, M. **Artificial Intelligence Through Simulated Evolution**. John Wiley & Sons, 1966. Disponível em: <<https://books.google.com.br/books?id=QMLaAAAAMAAJ>>.
- GONZÁLES, G. L. G. **Aplicação da Técnica SIFT para Determinação de Campos de Deformações de Materiais**. Mestrado em Engenharia Mecatrônica — PUC-Rio, 2010.
- GROSKO, A. P.; GORSKI, J. R.; DIAS, J. d. S. Algoritmo genético: revisão histórica e exemplificação. In: SOCIEDADE BRASILEIRA DE INFORMÁTICA NA SAÚDE. **X Congresso Brasileiro de Informática na Saúde**. Florianópolis, Brasil, 2006.
- HASAN, M. A.; KIM, C. An automatic image browsing technique for small display users. In: IEEE. **Advanced Communication Technology, 2009. ICACT 2009. 11th International Conference on**. Dublin, Irlanda, 2009. v. 3, p. 2044–2049.

- HASHEMI, S.; KIANI, S.; NOROOZI, N.; MOGHADDAM, M. E. An image contrast enhancement method based on genetic algorithm. **Pattern Recognition Letters**, v. 31, n. 13, p. 1816 – 1824, 2010. ISSN 0167-8655. Meta-heuristic Intelligence Based Image Processing. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167865509003365>>.
- HOLLAND, J. H. **Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence**. Cambridge, Estados Unidos: MIT Press, 1992. ISBN 0262082136.
- ITO, I. Gradient-based global features and its application to image retargeting. In: IEEE. **Signal & Information Processing Association Annual Summit and Conference (APSIPA ASC), 2012 Asia-Pacific**. Tóquio, Japão, 2012. p. 1–4.
- KASUTANI, E.; YAMADA, a. The MPEG-7 color layout descriptor: a compact image feature description for high-speed image/video segment retrieval. In: **Proceedings 2001 International Conference on Image Processing (Cat. No.01CH37205)**. Salonica, Grécia: [s.n.], 2001. v. 1, p. 674–677. ISBN 0-7803-6725-1. ISSN 13652893.
- KIM, W.; KIM, C. A novel image importance model for content-aware image resizing. In: IEEE. **Image Processing (ICIP), 2011 18th IEEE International Conference on**. Bruxelas, Bélgica, 2011. p. 2469–2472.
- KOZA, J. R. **Genetic Programming: On the Programming of Computers by Means of Natural Selection**. Cambridge, Estados Unidos: MIT Press, 1992. ISBN 0-262-11170-5.
- KOZA, J. R. **Genetic Programming II: Automatic Discovery of Reusable Programs**. Cambridge, Estados Unidos: MIT Press, 1994. ISBN 0-262-11189-6.
- KRÄHENBÜHL, P.; LANG, M.; HORNUNG, A.; GROSS, M. A system for retargeting of streaming video. In: ACM. **ACM Transactions on Graphics (TOG)**. Hong Kong, China, 2009. v. 28, n. 5, p. 126.
- LEI, R. **Distinctive Image Features from Scale-Invariant Keypoints**. 2011. Disponível em: <<http://cops2point0.com/2016/02/ammai-02-distinctive-image-features-from-scale-invariant/>>. Acesso em: 27-Março-2016.
- LINDEBERG, T. **Scale-space theory in computer vision**. Norwell, Estados Unidos: Springer Science & Business Media, 2013.
- LINDEN, R. **Algoritmos genéticos (2a edição)**. Curitiba, Brasil: Brasport, 2008.
- LIPOWSKI, A.; LIPOWSKA, D. Roulette-wheel selection via stochastic acceptance. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 391, n. 6, p. 2193–2196, 2012.
- LIU, C.; YUEN, J.; TORRALBA, A. Sift flow: Dense correspondence across scenes and its applications. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 33, n. 5, p. 978–994, 2011.
- LIU, H.; JIANG, S.; HUANG, Q.; XU, C.; GAO, W. Region-based visual attention analysis with its application in image browsing on small displays. In: ACM. **Proceedings of the 15th international conference on Multimedia**. Augsburg, Alemanha, 2007. p. 305–308.

- LIU, H.; XIE, X.; MA, W.-Y.; ZHANG, H.-J. Automatic browsing of large pictures on mobile devices. In: **ACM. Proceedings of the eleventh ACM international conference on Multimedia**. Dublin, Irlanda, 2003. p. 148–155.
- LIU, Y.-J.; LUO, X.; XUAN, Y.-M.; CHEN, W.-F.; FU, X.-L. Image retargeting quality assessment. **Computer Graphics Forum**, Blackwell Publishing Ltd, v. 30, n. 2, p. 583–592, 2011. ISSN 1467-8659. Disponível em: <<http://dx.doi.org/10.1111/j.1467-8659.2011.01881.x>>.
- LOWE, D. G. Distinctive image features from scale-invariant keypoints. **International journal of computer vision**, Springer, v. 60, n. 2, p. 91–110, 2004.
- MA, L.; LIN, W.; DENG, C.; NGAN, K. N. Image retargeting quality assessment: a study of subjective scores and objective metrics. **Selected Topics in Signal Processing, IEEE Journal of**, IEEE, v. 6, n. 6, p. 626–639, 2012.
- MAIA, J. **Detecção e Reconhecimento de Objetos usando Descritores Locais**. Doutorado em Ciência da Computação — Universidade Federal do Ceará, 2010.
- MAN, K.-F.; TANG, K. S.; KWONG, S. **Genetic algorithms: concepts and designs**. [S.l.]: Springer Science & Business Media, 2012.
- MESSING, D. S.; BEEK, P. van; ERRICO, J. H. The mpeg-7 colour structure descriptor: Image description using colour and local spatial information. In: IEEE. **Image Processing, 2001. Proceedings. 2001 International Conference on**. Salonica, Grécia, 2001. v. 1, p. 670–673.
- MOLÍŃSKI, S. **Scilab: Gaussian Pyramids**. 2016. Disponível em: <<http://imgsimon.blogspot.com.br/2016/06/scilab-gaussian-pyramids.html>>. Acesso em: 01-Agosto-2016.
- MUÑOZ, A.; BLU, T.; UNSER, M. Least-squares image resizing using finite differences. **Image Processing, IEEE Transactions on**, IEEE, v. 10, n. 9, p. 1365–1378, 2001.
- OLIVEIRA, S. A.; BEZERRA, F. N.; ROCHA NETO, A. R. Genetic seam carving: A genetic algorithm approach for content-aware image retargeting. In: **Pattern Recognition and Image Analysis**. Santiago de Compostela, Espanha: Springer, 2015. p. 700–707.
- OLIVEIRA, S. A.; ROCHA NETO, A. R.; BEZERRA, F. N. A novel genetic algorithms and surf-based approach for image retargeting. **Expert Systems with Applications**, Pergamon, v. 44, p. 332–343, 2016.
- OLIVEIRA, S. A. F.; ROCHA NETO, A. R. An improved genetic algorithms-based seam carving method. In: IEEE. **2015 Latin America Congress on Computational Intelligence (LA-CCI)**. Curitiba, Brasil, 2015. p. 1–6.
- OLIVEIRA, S. A. F.; ROCHA NETO, A. R. An improved genetic algorithms-based seam carving method. **International Journal of Innovative Computing & Applications**, Inderscience, n. 1, p. 1–7, 2016.

- PAIVA, J. L. de; TOLEDO, C. F.; PEDRINI, H. An approach based on hybrid genetic algorithm applied to image denoising problem. **Applied Soft Computing**, v. 46, p. 778 – 791, 2016. ISSN 1568-4946. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1568494615005852>>.
- PELE, O.; WERMAN, M. Fast and robust earth mover's distances. In: IEEE. **Computer vision, 2009 IEEE 12th international conference on**. Quioto, Japão, 2009. p. 460–467.
- PRITCH, Y.; KAV-VENAKI, E.; PELEG, S. Shift-map image editing. In: IEEE. **Computer Vision, 2009 IEEE 12th International Conference on**. Quioto, Japão, 2009. p. 151–158.
- RAHNAMAYAN, S.; TIZHOOSH, H. R.; SALAMA, M. M. A novel population initialization method for accelerating evolutionary algorithms. **Computers & Mathematics with Applications**, Elsevier, v. 53, n. 10, p. 1605–1614, 2007.
- RECHENBERG, I. Evolution strategy: Optimization of technical systems by means of biological evolution. **Fromman-Holzboog, Stuttgart**, v. 104, 1973.
- RUBINSTEIN, M.; GUTIERREZ, D.; SORKINE, O.; SHAMIR, A. A comparative study of image retargeting. **ACM Transactions on Graphics (Proc. SIGGRAPH Asia)**, v. 29, n. 6, p. 160:1–160:10, 2010.
- RUBINSTEIN, M.; SHAMIR, A.; AVIDAN, S. Improved seam carving for video retargeting. In: ACM. **ACM transactions on graphics (TOG)**. Los Angeles, Estados Unidos, 2008. v. 27, n. 3, p. 16.
- RUBINSTEIN, M.; SHAMIR, A.; AVIDAN, S. Multi-operator media retargeting. In: ACM. **ACM Transactions on Graphics (TOG)**. Nova Orleans, Estados Unidos, 2009. v. 28, n. 3, p. 23.
- SAMADANI, R.; LIM, S. H.; TRETTER, D. Representative image thumbnails for good browsing. In: IEEE. **Image Processing, 2007. ICIP 2007. IEEE International Conference on**. San Antonio, Estados Unidos, 2007. v. 2, p. II–193.
- SANTELLA, A.; AGRAWALA, M.; DECARLO, D.; SALESIN, D.; COHEN, M. Gaze-based interaction for semi-automatic photo cropping. In: ACM. **Proceedings of the SIGCHI conference on Human Factors in computing systems**. Montreal, Canadá, 2006. p. 771–780.
- SAWYERR, B. A.; ALI, M. M.; ADEWUMI, A. O. A comparative study of some real-coded genetic algorithms for unconstrained global optimization. **Optimization Methods and Software**, Taylor & Francis, v. 26, n. 6, p. 945–970, 2011.
- SEMENKIN, E.; SEMENKINA, M. Self-configuring genetic algorithm with modified uniform crossover operator. In: SPRINGER. **International Conference in Swarm Intelligence**. [S.l.], 2012. p. 414–421.
- SERPELL, M.; SMITH, J. E. Self-adaptation of mutation operator and probability for permutation representations in genetic algorithms. **Evolutionary Computation**, MIT Press, v. 18, n. 3, p. 491–514, 2010.
- SETLUR, V.; TAKAGI, S.; RASKAR, R.; GLEICHER, M.; GOOCH, B. Automatic image retargeting. In: ACM. **Proceedings of the 4th international conference on Mobile and ubiquitous multimedia**. Christchurch, Nova Zelândia, 2005. p. 59–68.

- SHAMIR, A.; AVIDAN, S. Seam carving for media retargeting. **Commun. ACM**, ACM, Nova York, Estados Unidos, v. 52, n. 1, p. 77–85, jan. 2009. ISSN 0001-0782. Disponível em: <<http://doi.acm.org/10.1145/1435417.1435437>>.
- SHAMIR, A.; SORKINE, O. Visual media retargeting. In: ACM. **ACM SIGGRAPH ASIA 2009 Courses**. Yokohama, Japão, 2009. p. 11.
- SHEIKH, H. R.; BOVIK, A. C. Image information and visual quality. **Image Processing, IEEE Transactions on**, IEEE, v. 15, n. 2, p. 430–444, 2006.
- SIMA, A. A.; BUCKLEY, S. J. Optimizing sift for matching of short wave infrared and visible wavelength images. **Remote Sensing**, Multidisciplinary Digital Publishing Institute, v. 5, n. 5, p. 2037–2056, 2013.
- SIMAKOV, D.; CASPI, Y.; SHECHTMAN, E.; IRANI, M. Summarizing visual data using bidirectional similarity. In: IEEE. **Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on**. Anchorage, Estados Unidos, 2008. p. 1–8.
- SIVANANDAM, S.; DEEPA, S. **Genetic Algorithms**. Heidelberg, Alemanha: Springer Berlin Heidelberg, 2008. 15–37 p. ISBN 978-3-540-73190-0. Disponível em: <http://dx.doi.org/10.1007/978-3-540-73190-0_2>.
- SIVAPATHASEKARAN, C.; MUKHERJEE, S.; RAY, A.; GUPTA, A.; SEN, R. Artificial neural network modeling and genetic algorithm based medium optimization for the improved production of marine biosurfactant. **Bioresource technology**, Elsevier, v. 101, n. 8, p. 2884–2887, 2010.
- STENTIFORD, F. Attention based auto image cropping. In: CITESEER. **The 5th International Conference on Computer Vision Systems, Bielefeld**. Bielefeld, Alemanha, 2007.
- STORE, H.; SIDE, O. L.; PAGE, H. M.; HCI, S. **Photocropr: A first step towards computer-supported automatic generation of photographically interesting cropping suggestions**. 2007.
- SUH, B.; LING, H.; BEDERSON, B. B.; JACOBS, D. W. Automatic thumbnail cropping and its effectiveness. In: ACM. **Proceedings of the 16th annual ACM symposium on User interface software and technology**. Vancouver, Canadá, 2003. p. 95–104.
- SZELISKI, R.; ZABIH, R.; SCHARSTEIN, D.; VEKSLER, O.; KOLMOGOROV, V.; AGARWALA, A.; TAPPEN, M.; ROTHER, C. A comparative study of energy minimization methods for markov random fields with smoothness-based priors. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, IEEE, v. 30, n. 6, p. 1068–1080, 2008.
- TAO, W.-B.; TIAN, J.-W.; LIU, J. Image segmentation by three-level thresholding based on maximum fuzzy entropy and genetic algorithm. **Pattern Recognition Letters**, v. 24, n. 16, p. 3069 – 3078, 2003. ISSN 0167-8655. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0167865503001661>>.
- VAQUERO, D.; TURK, M.; PULLI, K.; TICO, M.; GELFAND, N. A survey of image retargeting techniques. In: INTERNATIONAL SOCIETY FOR OPTICS AND PHOTONICS. **SPIE Optical Engineering+ Applications**. San Diego, Estados Unidos, 2010. p. 779814–779814.

- WANG, Y.-S.; TAI, C.-L.; SORKINE, O.; LEE, T.-Y. Optimized scale-and-stretch for image resizing. In: ACM. **ACM Transactions on Graphics (TOG)**. Singapura, Singapura, 2008. v. 27, n. 5, p. 118.
- WANG, Z.; BOVIK, A. C.; SHEIKH, H. R.; SIMONCELLI, E. P. Image quality assessment: from error visibility to structural similarity. **Image Processing, IEEE Transactions on**, IEEE, v. 13, n. 4, p. 600–612, 2004.
- WHITLEY, D. A genetic algorithm tutorial. **Statistics and computing**, Springer, v. 4, n. 2, p. 65–85, 1994.
- WOLF, L.; GUTTMANN, M.; COHEN-OR, D. Non-homogeneous content-driven video-retargeting. In: IEEE. **2007 IEEE 11th International Conference on Computer Vision**. Rio de Janeiro, Brasil, 2007. p. 1–6.
- ZHANG, J.; SCLAROFF, S. Saliency detection: A boolean map approach. In: IEEE. **Proceedings of the IEEE International Conference on Computer Vision**. Sydney, Austrália, 2013. p. 153–160. ISSN 1550-5499.
- ZHANG, M.; ZHANG, L.; SUN, Y.; FENG, L.; MA, W. Auto cropping for digital photographs. In: IEEE. **Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on**. Amsterdã, Países Baixos, 2005. p. 4–pp.
- ZHANG, X.; LI, X.; FENG, Y.; ZHAO, H.; LIU, Z. Image fusion with internal generative mechanism. **Expert Systems with Applications**, v. 42, n. 5, p. 2382 – 2391, 2015. ISSN 0957-4174. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417414006824>>.